Reconstructing 3D Spatial Structures of RNA-Tomography Transcriptomes by Tensor Factorization

Tianci Song[1], Quoc Nguyen[1], and Rui Kuang[1]

Department of Computer Science and Engineering,
University of Minnesota Twin Cities, Minneapolis, 55414, MN, USA
kuang@umn.edu

Abstract. Cells are structurally organized in complex tissues. Mapping the organization of the cells requires reliable 3D molecular profiling. Technologies based on microdissections of tissues provide deep RNA sequencing of orthographically arranged cryosections of a tissue such that the full 3D spatial structure can be reconstructed from deeply sequenced transcriptomes in three views projected like a 3D tomography. In this work, we introduce a method, TFacTomo, to learn a Tensor Factorization for RNA Tomography data from cryosections for constructing 3D spatially resolved gene expressions. TFacTomo combines tensor factorization with a loss function supervised by collapsing the tensor entries to match the gene expressions and tissue mask in each cryosection. To incorporate spatial and functional prior, a regularization by a product graph of the PPI network and spatial graphs is also introduced. In the experiments, TFacTomo is applied to two simulation datasets constructed from a 3D ST mouse brain dataset and a 3D ST human heart dataset, and two RNA Tomography datasets, one from zebrafish embryo and the other from mouse olfactory mucosa. In the simulation, TFacTomo significantly outperforms the benchmark method IPF in the accuracy of predicting the ground-truth gene expressions in the ST datasets based on the projected 1D spatial gene expressions of three orthographic views. In the two RNA Tomography datasets, TFacTomo detects 3D spatial expressions of several marker genes that are consistent with the developmental or functional regions in comparison to staining images. Visual comparisons also suggest that TFacTomo detects more focused regions than the state-of-the-art method.

Keywords: Spatial Genomics · RNA Tomography Sequencing · Tensor Factorization

1 Introduction

Cells in multicellular organisms are organized to form complex 3D tissue structures, which create distinct microenvironments to drive essential events such as cell movement, cell communication and interactions, and spatially varying gene

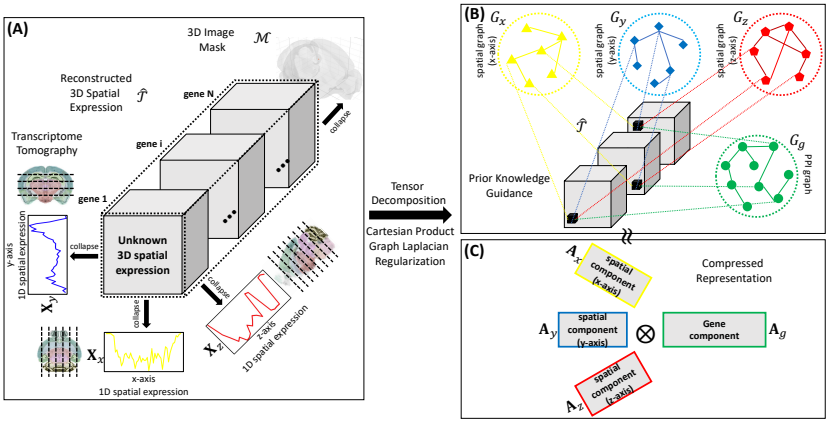

Fig. 1. Overview of TFacTomo. **(A)** Given 1D spatial gene expression of each gene along three orthogonal spatial axes ($\mathbf{X}_x, \mathbf{X}_y, \mathbf{X}_z$), TFacTomo reconstructs 3D spatial gene expressions in a 4-way tensor $\mathcal{T}$ to minimize the fitting error to the gene expression along different spatial axes and the 3D mask. **(B)** TFacTomo integrates the prior knowledge in spatial graphs (G_x, G_y, G_z) and PPI network G_g to inform 3D spatial gene expression reconstruction. **(C)** TFacTomo learns a compressed representation by CP decomposition to estimate $\hat{\mathcal{T}}$ with four factor matrices ($\mathbf{A}_x$,$\mathbf{A}_y$,$\mathbf{A}_z$,$\mathbf{A}_g$).

expressions. To reveal the spatial organization of the cells and their functions in the tissues, spatial transcriptomic technologies based on in-situ hybridization (ISH) or in-situ capturing (ISC) have been widely used to profile gene expression with retained spatial localization information in the tissue [11]. Concurrently, technologies based on microdissection adopt a different strategy, which isolates slices or regions of tissues for separate RNA sequencing such that the location of the cells is informed by the organization of the arranged slices or regions [10]. Transcriptome tomography sequencing (Tomo-seq) [5] is a notable method in this category inspired by image tomography. In Tomo-seq, replicates of a 3D tissue are dissected into successive slices along three orthogonal views (axes) and then the RNAs in the cells in each slice are collected for deep bulk RNA sequencing. Finally, the 3D structures of the gene expressions can be reconstructed by the gene expressions measured from the slices stacked in each of the three axes using a computational method. Tomo-seq collects more reliable transcriptome-wide gene expression measures with highly sensitive bulk sequencing while ISC-based methods suffer from low-RNA capture rate and ISH-based methods suffer from lower detection sensitivity in transcriptome-wide scale profiling. More importantly, since both ISH and ISC inherently profile RNAs in a 2D space, a challenging step is required to align stacked slices by manual registration of annotated regions, tissue histological images, or spatial coordinates for 3D construction. In contrast, Tomo-seq performs RNA sequencing of slices dissected in three views from replicates for a real 3D measuring of the gene expressions for full 3D reconstruction. Despite these advantages, one significant limitation of Tomo-seq and other microdissection-based methods is that the re-

construction of the 3D spatial gene expressions is a significant computational challenge. The current solution for 3D reconstruction with Tomo-seq is the iterative proportional fitting (IPF) algorithm [4], which is a relatively naive heuristic optimization method for minimizing fitting error for each individual gene independently. Accordingly, IPF is prone to noise in RNA sequencing data and might fail to capture complex spatial patterns [15].

In this paper, we introduce tensor factorization for tomography (TFacTomo) to learn a tensor factorization representation of spatially resolved 3D gene expressions from RNA tomography data of cryosections. TFacTomo reconstructs 3D spatial transcriptomics based on the 1D bulk RNA sequencing data obtained for the consecutive slices in each orthogonal spatial axis of tomography along the tissue as outlined in Fig. 1. The 1D gene expression data represent the expression of each gene in each consecutive slice in one orthogonal axis. Then, the 1D data in three views is combined to reconstruct the 3D spatial transcriptomes in a 4-way tensor with x-y-z spatial modes and gene mode as shown in Fig. 1A. TFacTomo combines tensor factorization with a loss function supervised by collapsing the tensor entries to match the gene expressions in each cryosection and the 3D tissue mask. TFacTomo outputs a CP decomposition of the reconstructed tensor aligned with the spatial axes (Fig. 1C). TFacTomo also incorporates spatial relations among the slices along different axes and gene functional relations among genes with regularization by a product of the PPI network and spatial graphs to utilize the critical prior information in modeling (Fig. 1B). The source code of TFacTomo is available at https://github.com/kuanglab/TFacTomo.

2 Method

2.1 Preliminaries

Here, we only present the minimal necessary definitions for introducing TFacTomo. A more complete review is given in supplementary section S1.

Canonical Polyadic Decomposition (CPD). A M-way tensor $\mathcal{T}$ can be approximated in a compressed rank-R CPD representation as follows,

$$\mathcal{T} \approx [\![\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_M]\!] = \sum_{r=1}^{n_r} \mathbf{a}_1^r \circledcirc \mathbf{a}_2^r \circledcirc \ldots \circledcirc \mathbf{a}_M^r, \tag{1}$$

where $[\![\cdot]\!]$ denotes the Kruskal operator, $\circledcirc$ denotes vector outer product, n_r is the rank of the decomposition, $\mathbf{A}_i \in \mathbb{R}^{n_i \times n_r}$ represents the factor matrix along mode-i, and $\mathbf{a}_i^r$ is a vector in the r-th column of $\mathbf{A}_m$.

Matricization Property. A tensor represented by CP decomposition can be matricized along i-th mode as $\mathcal{T}_{(i)} = \mathbf{A}_i(\mathbf{A}_1 \odot \ldots \odot \mathbf{A}_{i-1} \odot \mathbf{A}_{i+1} \odot \ldots \odot \mathbf{A}_M)^T$, where $\odot$ denotes Khatri-Rao product, and $\mathcal{T}_{(i)} \in \mathbb{R}^{n_i \times \prod_{j \neq i} n_j}$. As a special case of matricization, a tensor in the form of CP decomposition can be vectorized as $\text{vec}(\mathcal{T}) = (\mathbf{A}_1 \odot \ldots \odot \mathbf{A}_j \odot \ldots \odot \mathbf{A}_M)\mathbf{1}^T$, where $\text{vec}(\cdot)$ represents the vectorization

function and $\mathbf{1}$ is an all-ones vector of size n_r.

Collapsing Property. Collapsing operation collapses a tensor into a lower-order one with given modes by summing over entries along the remaining modes. Collapsing given any two modes on the CPD of a tensor can be written as $\text{collapse}(\llbracket \mathbf{A}_1, \ldots, \mathbf{A}_M \rrbracket, i, j) = \llbracket \mathbf{a}_1, \ldots, \mathbf{a}_{i-1}, \mathbf{A}_i, \mathbf{a}_{i+1}, \ldots, \mathbf{a}_{j-1}, \mathbf{A}_j, \mathbf{a}_{j+1}, \ldots, \mathbf{a}_M \rrbracket$, where collapse$(\cdot)$ represents the collapsing function and $\mathbf{a}_i^T = \mathbf{1}_i^T \mathbf{A}_i$ denotes the row summation of $\mathbf{A}_i$. Collapsing operation can similarly be defined on the tensor of CP decomposition form given more than two modes.

2.2 Loss Function

The overview of TFacTomo is illustrated in Fig. 1. The inputs of TFacTomo are the tomography transcriptomics data along different spatial axes as gene expression matrices $\mathbf{X}_x \in \mathbf{R}_+^{n_g \times n_x}, \mathbf{X}_y \in \mathbf{R}_+^{n_g \times n_y}, \mathbf{X}_z \in \mathbf{R}_+^{n_g \times n_z}$, and a 3D image mask as a 3-way tensor $\mathcal{M} \in \{0,1\}^{n_x \times n_y \times n_z}$ as shown in Fig. 1A. The 1D spatial chain relations along three orthogonal axes and functional relations among genes are given in four knowledge graphs $\{G_i, i = g, x, y, z\}$ (Fig. 1C). TFacTomo reconstructs 3D spatial gene expression in a 4-way tensor $\hat{\mathcal{T}} \in \mathbb{R}_+^{n_g \times n_x \times n_y \times n_z}$ represented by CP decomposition while being guided by prior knowledge encoded in the spatial and functional graphs (Fig. 1B). We next formulate the learning problem and define each term in the loss function of TFacTomo.

Reconstruction Loss. We first define reconstruction loss $\mathcal{J}_1$, which ensures the collapse of the reconstructed 3D spatial gene expression tensor $\hat{\mathcal{T}}$ along the gene and a given spatial axes should be identical to the 1D gene expression matrix $\mathbf{X}_i$, and the loss term can be written as

$$\mathcal{J}_1 = \frac{1}{2} \sum_{i=\{x,y,z\}} \left\| \mathbf{X}_i - \text{collapse}(\hat{\mathcal{T}}, g, i) \right\|_F^2 . \tag{2}$$

Here, 4-way tensor $\hat{\mathcal{T}}$ is collapsed (summed) over the other two spatial modes along each of x-axis, y-axis and z-axis, exactly modeling the tomography projection onto the cryosections in each view.

3D Masking Loss. To prevent the reconstructed 3D spatial gene expression from overspreading to regions outside of the tissue, a loss term $\mathcal{J}_2$ is introduced to fit the given 3D mask $\mathcal{M}$ that outlines the tissue shape as follows,

$$\mathcal{J}_2 = \frac{1}{2} \left\| \mathcal{M} \circledast \text{collapse}(\hat{\mathcal{T}}, x, y, z) \right\|_F^2 \tag{3}$$

where $\circledast$ denotes Hadamard product. Here, 4-way tensor $\hat{\mathcal{T}}$ is collapsed (summed) over all the genes in each spatial location (spot), and the loss penalizes the non-zero sums in the spots outside of the tissue.

Product Graph Laplacian Regularization. To leverage both spatial relations in 3D space and functional relations among genes for reconstruction, a Laplacian regularization of the Cartesian product of gene and spatial graphs $\{G_i, i = g, x, y, z\}$

in $\mathcal{J}_3$ is used to smooth the entries in $\mathcal{T}$ over the manifolds in the Cartesian product graph G_p, ensuring that the expressions of adjacent nodes in this high-order graph share similar values, defined as follows,

$$\mathcal{J}_3 = \frac{1}{2}\text{vec}(\hat{\mathcal{T}})^T \mathbf{L}_p \text{vec}(\hat{\mathcal{T}}), \tag{4}$$

Here $\mathbf{L}_p = \mathbf{L}_g \oplus \mathbf{L}_x \oplus \mathbf{L}_y \oplus \mathbf{L}_z$ is the graph Laplacian of G_p, where $\mathbf{L}_i, \forall i = g, x, y, z$, is the Laplacian graph of G_i, and $\oplus$ denotes Kronecker sum [8].

Total Loss function. Lastly, the total loss function $\mathcal{J}$ to learn a rank-R CP decomposition $\hat{\mathcal{T}} = [\![\mathbf{A}_g, \mathbf{A}_x, \mathbf{A}_y, \mathbf{A}_z]\!]$ is the sum of $\mathcal{J}_1$, $\mathcal{J}_2$ and $\mathcal{J}_3$ as follows,

$$\begin{aligned} &\min_{\{\mathbf{A}_g, \mathbf{A}_x, \mathbf{A}_y, \mathbf{A}_z\}} \mathcal{J} = \mathcal{J}_1 + \alpha \mathcal{J}_2 + \beta \mathcal{J}_3 \\ &\mathcal{J}_1 = \frac{1}{2} \sum_{i=\{x,y,z\}} \|\mathbf{X}_i - \text{collapse}([\![\mathbf{A}_g, \mathbf{A}_x, \mathbf{A}_y, \mathbf{A}_z]\!], g, i)\|_F^2 \\ &\mathcal{J}_2 = \frac{1}{2} \|\mathcal{M} \circledast \text{collapse}([\![\mathbf{A}_g, \mathbf{A}_x, \mathbf{A}_y, \mathbf{A}_z]\!], x, y, z)\|_F^2 \\ &\mathcal{J}_3 = \frac{1}{2}\text{vec}([\![\mathbf{A}_g, \mathbf{A}_x, \mathbf{A}_y, \mathbf{A}_z]\!])^T \mathbf{L}_p \text{vec}([\![\mathbf{A}_g, \mathbf{A}_x, \mathbf{A}_y, \mathbf{A}_z]\!]), \end{aligned} \tag{5}$$

where α and β are hyperparameters weighting the loss terms in the loss function.

Table 1. Auxiliary variables.

Variable	Definition
$\Phi_j^{(-i)}$	$\circledast_{k \neq i,j}(\mathbf{a}_k \mathbf{a}_k^T) \circledast (\mathbf{A}_j^T \mathbf{A}_j)$
$\Phi_{j,k}^{(-i)}$	$\circledast_{l \neq i,j,k}(\mathbf{a}_k \mathbf{a}_k^T) \circledast (\mathbf{A}_j^T \mathbf{A}_j) \circledast (\mathbf{A}_k^T \mathbf{A}_k)$
$\Theta_j^{(-i)}$	$\mathbf{X}_{(i,j)}(\odot_{k \neq i,j} \mathbf{a}_k^T \odot \mathbf{A}_j)$
$\Theta_{j,k}^{(-i)}$	$\mathbf{1}_i \text{vec}(\mathbf{X}_{(j,k)})^T (\odot_{l \neq i,j,k} \mathbf{a}_l^T \odot \mathbf{A}_j \odot \mathbf{A}_j)$
$\Upsilon^{(-i)}$	$\odot_{j \neq i} \mathbf{A}_j$
$\Upsilon_j^{(-i)}$	$\odot_{k \neq i,j} \mathbf{A}_k \odot \mathbf{a}_j^T$
$\Psi_j^{(-i)}$	$(\mathbf{A}_j^T \mathbf{L}_j \mathbf{A}_j) \circledast_{k \neq i,j} (\mathbf{A}_k^T \mathbf{A}_k)$

2.3 Multiplicative Updating

To minimize the loss function in Eq. 5, we propose an iterative optimization method based on multiplicative updating rules [2, 7, 9], which ensures a stationary solution of $\hat{\mathcal{T}}$ by alternatively updating each $\{\mathbf{A}_i, i = g, x, y, z\}$ with the derivatives of $\mathcal{J}$ with respect to the $\mathbf{A}_i$. For simplicity, we decompose the loss function into multiple terms and introduce their derivatives with respect to $\mathbf{A}_g$ and $\mathbf{A}_i, \forall i = x, y, z$ individually. To facilitate the calculation of the derivatives, we define the auxiliary variables in Table 1.

With the definitions in Eq. 1, the partial derivatives of $\mathcal{J}_1$ with respect to $\mathbf{A}_g$ and $\mathbf{A}_i, \forall i = x, y, z$ can be represented as follows:

$$\begin{aligned} \frac{\partial \mathcal{J}_1}{\partial \mathbf{A}_g} &= \mathbf{A}_g \sum_{i=\{x,y,z\}} \Phi_i^{(-g)} - \sum_{i=\{x,y,z\}} \Theta_i^{(-g)} \\ \frac{\partial \mathcal{J}_1}{\partial \mathbf{A}_i} &= \mathbf{A}_i \Phi_g^{(-i)} + \mathbf{1}_i \mathbf{1}_i^T \mathbf{A}_i (\sum_{j \neq i} \Phi_{j,g}^{(-i)}) - \Theta_g^{(-i)} - \sum_{j \neq i} \Theta_{j,g}^{(-i)}. \end{aligned} \tag{6}$$

Similarly, the partial derivatives of $\mathcal{J}_2$ with respect to $\mathbf{A}_g$ and $\mathbf{A}_i, \forall i = x, y, z$ can be represented as follows:

$$\begin{aligned} \frac{\partial \mathcal{J}_2}{\partial \mathbf{A}_g} &= \mathbf{1}_g(\mathrm{vec}(\mathcal{M}) \circledast \mathrm{vec}([\![\mathbf{a}_g, \mathbf{A}_x, \mathbf{A}_y, \mathbf{A}_z]\!]))^T \Upsilon^{(-g)} \\ \frac{\partial \mathcal{J}_2}{\partial \mathbf{A}_i} &= (\mathbf{M}_{(i)} \circledast [\![\mathbf{a}_g, \mathbf{A}_x, \mathbf{A}_y, \mathbf{A}_z]\!])\Upsilon_g^{(-i)}, \end{aligned} \tag{7}$$

where $\mathbf{M}_{(i)} \in \{0,1\}^{n_i \times (\prod_{j \neq i} n_j)}$ denotes the mode-i matricization of tensor $\mathcal{M}$.

Following the derivation in [9], the partial derivatives of $\mathcal{J}_3$ with respect to $\mathbf{A}_i, \forall i = g, x, y, z$ can be represented as follows:

$$\begin{aligned} \frac{\partial \mathcal{J}_3}{\partial \mathbf{A}_i} &= (\mathbf{L}_i \mathbf{A}_i)(\circledast_{j \neq i} \mathbf{A}_j^T \mathbf{A}_j) + \mathbf{A}_i(\sum_{j \neq i} \Psi_j^{(-i)}) \\ &= (\mathbf{D}_i \mathbf{A}_i)(\circledast_{j \neq i} \mathbf{A}_j^T \mathbf{A}_j) + \mathbf{A}_i(\sum_{j \neq i} \Psi_j^{(-i)}) - (\mathbf{W}_i \mathbf{A}_i)(\circledast_{j \neq i} \mathbf{A}_j^T \mathbf{A}_j), \end{aligned} \tag{8}$$

where $\mathbf{W}_i$ and $\mathbf{D}_i$ are the adjacency matrix and degree matrix of the graph G_i, and $\mathbf{L}_i = \mathbf{D}_i - \mathbf{W}_i$.

After combining the derivatives in Eq. 6, 7, 8, the derivative of $\mathcal{J}$ with respect to $\mathbf{A}_i, \forall i = g, x, y, z$ can be formalized as follows:

$$\begin{aligned} \frac{\partial \mathcal{J}}{\partial \mathbf{A}_i} &= \frac{\partial \mathcal{J}_1}{\partial \mathbf{A}_i} + \frac{\partial \mathcal{J}_2}{\partial \mathbf{A}_i} + \frac{\partial \mathcal{J}_3}{\partial \mathbf{A}_i} \\ &= \left[\frac{\partial \mathcal{J}_1}{\partial \mathbf{A}_i}\right]^+ - \left[\frac{\partial \mathcal{J}_1}{\partial \mathbf{A}_i}\right]^- + \alpha \left[\frac{\partial \mathcal{J}_2}{\partial \mathbf{A}_i}\right]^+ + \beta\Big(\left[\frac{\partial \mathcal{J}_3}{\partial \mathbf{A}_i}\right]^+ - \left[\frac{\partial \mathcal{J}_3}{\partial \mathbf{A}_i}\right]^-\Big), \end{aligned} \tag{9}$$

where $\left[\frac{\partial \mathcal{J}_k}{\partial \mathbf{A}_i}\right]^+$ and $\left[\frac{\partial \mathcal{J}_k}{\partial \mathbf{A}_i}\right]^-$ are the non-negative components in the partial derivative of $\mathcal{J}_k, \forall k = 1, 2, 3$ with respect to $\mathbf{A}_i, \forall i = g, x, y, z$.

The loss function $\mathcal{J}$ will be monotonically decreased by alternatively updating each factor matrix $\mathbf{A}_i, \forall i = g, x, y, z$ using the following multiplicative updating rule in each iteration until convergence.

$$[\mathbf{A}_i]_{ab} \leftarrow [\mathbf{A}_i]_{ab} \left(\frac{\left[\frac{\partial \mathcal{J}_1}{\partial \mathbf{A}_i}\right]_{ab}^- + \left[\frac{\partial \mathcal{J}_3}{\partial \mathbf{A}_i}\right]_{ab}^-}{\left[\frac{\partial \mathcal{J}_1}{\partial \mathbf{A}_i}\right]_{ab}^+ + \left[\frac{\partial \mathcal{J}_2}{\partial \mathbf{A}_i}\right]_{ab}^+ + \left[\frac{\partial \mathcal{J}_3}{\partial \mathbf{A}_i}\right]_{ab}^+} \right). \tag{10}$$

2.4 Time and Space Complexity

Let $n_r = R$ and $|\cdot|$ denote the number of non-zero entries in either a matrix or a tensor, the time complexity to update $\mathbf{A}_i, \forall i = g, x, y, z$ in each iteration is $\mathcal{O}(\sum_{j \neq i} n_r |X_j| + n_r |\mathcal{M}| + \sum_j (n_r^2 n_j + n_r n_j^2))$, which is derived based on the time complexity of matricized tensor times Khatri-Rao product (MTTKRP) [3, 18]. Since in RNA tomography data, the number of genes is larger in several

Table 2. Reconstruction performance in simulations. The best scores are bold.

Dataset	Method	Evaluation Metrics			
		MSE ↓	*MAE* ↓	*spot-wise* R^2 ↑	*gene-wise* R^2 ↑
Human Heart ($14,726\times 20\times 20\times 9$)	IPF	0.160 ± 0.081	0.190 ± 0.061	0.285 ± 0.205	0.061 ± 0.132
	TFacTomo w/o graph reg	0.176 ± 0.163	0.217 ± 0.192	0.269 ± 0.230	0.015 ± 0.245
	TFacTomo w/ graph reg	$\mathbf{0.147 \pm 0.093}$	$\mathbf{0.156 \pm 0.055}$	$\mathbf{0.386 \pm 0.181}$	$\mathbf{0.139 \pm 0.224}$
Mouse Brain ($10,239\times 30\times 30\times 30$)	IPF	0.426 ± 0.159	0.389 ± 0.081	0.262 ± 0.245	0.066 ± 0.293
	TFacTomo w/o graph reg	0.288 ± 0.196	0.321 ± 0.104	0.517 ± 0.289	0.096 ± 0.193
	TFacTomo w/ graph reg	$\mathbf{0.245 \pm 0.172}$	$\mathbf{0.298 \pm 0.097}$	$\mathbf{0.600 \pm 0.174}$	$\mathbf{0.171 \pm 0.119}$

magnitudes, i.e. $n_i << n_g, \forall i = x, y, z, r$, the time complexity is likely to be upper bounded by either $\mathcal{O}(n_r n_g^2)$ or $\mathcal{O}(n_r n_x n_y n_z)$, in neither of which the complexity in the size of the full 4-ways of the tensor is needed. The space required for the inputs RNA tomography data, spatial and functional graphs, as well as 3D image mask is $\mathcal{O}(\sum_{i\neq g} |\mathbf{X}_i| + |\mathcal{M}| + \sum_i (|\mathbf{W}| + n_g n_r))$. In practice, TFacTomo converges within 10 minutes and requires up to 10GB of memory for computation on the real datasets tested in this study.

3 Experiments

TFacTomo is applied to two simulation datasets constructed from a 3D ST mouse brain dataset (40 slices) [12] and a 3D ST human heart dataset (9 slices) [1], and two RNA tomography datasets, one from zebrafish embryo ($(50, 49, 56)$ slices) [5] and the other from mouse olfactory mucosa ($(56, 54, 60)$ slices) [16]. TFacTomo is benchmarked against the state-of-the-art method, IPF algorithm [4]. Note that while there exist other reconstruction methods such as probabilistic graphical model Tomographer [15], they are only applicable to 2D reconstruction and require sophisticated tissue sections along the same axis and cuts at different angles into parallel strips, and thus not comparable.

3.1 Simulation with 3D ST Data

To establish a benchmark with ground-truth spatially resolved 3D gene expressions for evaluation, we constructed RNA tomography simulation data from two 3D ST spatial transcriptomics datasets. The first ST dataset contains a stack of 9 adjacent slices of 2D ST spatial transcriptomics data from a developing human heart at 6.5 post-conception weeks (PCW) sectioned along the ventrodorsal axis. The second dataset contains a stack of 40 slices sectioned along the anteroposterior axis of an adult mouse brain. In both datasets, each ST slice contains 33×35 spots, and the 3D data were manually registered based on their associated H&E staining images, batch-corrected across slices, and normalized in log-space [1,12].

We first binned the spots in the stack of 2D expression data into a 3D tensor to reconcile the coordinate shift across the slices. The dimensions after binning are $14,726 \times 20 \times 20 \times 9$ for the human heart and $10,239 \times 30 \times 30 \times 30$ for the mouse brain, after removing non-expressed genes and background regions for both tissues. The 3D mask was derived for the no-expression spots by collapsing gene expressions along the gene axis in the tensor. A spatial chain graph was constructed

for each spatial axis and the protein-protein interaction (PPI) networks for *homo sapiens* and *mus musculus* were obtained from BioGRID [13] using mainly only physical interactions. To tune the hyperparameters α and β, we applied the grid search to find the best combination that minimizes the reconstruction loss $\mathcal{J}_1$ while restricting the over-expression outside tissue mask to be less than a very small threshold such as 1e$-$4 by imposing a large enough α such that a small $\mathcal{J}_2$ loss is assured. We determined the rank using $\min(\prod_i k_i, 500), \forall i = x, y, z$, where k_i denotes the number of principal components explaining more than 95% variance of the simulated tomography transcriptomics data $\mathbf{X}_i, \forall i = x, y, z$. The complete results for hyperparameter tuning are shown in Fig. S1 and S2.

The reconstruction performance on the simulated data is evaluated by mean squared error (MSE) and mean absolute error (MAE) between the reconstructed expression tensor and the ground-truth ST expressions. We also measured both spot-wise and gene-wise R^2 to assess how well the reconstruction recovers spot and gene patterns. As shown in Table 2, TFacTomo consistently outperformed IPF in reconstruction with the lowest MSE and MAE and the highest spot-wise and gene-wise R^2 on both simulated datasets. The scatter plots for comparison in Fig. S3 and S4 show the difference is highly significant. Note that to further investigate the contribution of spatial and functional information in the reconstruction, we performed an ablation study by removing the corresponding graph Laplacian regularization in the total loss function. We observed that TFacTomo outperformed its counterpart without graph Laplacian regularization, which suggests both spatial and functional relations are important in informing reconstruction. Interestingly, in the simulated human heart data, TFacTomo without graph regularization performed worse than IPF, suggesting that regularization is often necessary and required for high-order modeling used in TFacTomo.

3.2 3D Transcriptome Reconstruction for Zebrafish Embryo

Next, we applied IPF and TFacTomo to reconstruct 3D expressions using the RNA tomography data from a zebrafish embryo at the shield stage [5]. This dataset contains $(50, 49, 56)$ slices along three spatial axes, respectively. For a more comprehensive representation, the PPI network of *Danio rerio* was curated from STRING [19] with the interactions of confidence scores at least 0.8 retained. Keeping 6,153 genes present in both the PPI and expression profiles resulted in a tensor of size $6,153 \times 50 \times 49 \times 56$ to be reconstructed. We followed the same procedure in the original study to normalize each 1D spatial expression data based on embryonic geometry derived from the image mask along the corresponding spatial axis and then log-transformed the normalized data. We used the same strategy in the simulation experiment for hyperparameter tuning. Parameters $\alpha = 1$, $\beta = 1$, and rank$= 500$ were chosen for this dataset based on the hyperparameter tuning results shown in Fig. S5.

Since there is no 3D spatial transcriptomics data for the same zebrafish embryo at the shield stage available, we evaluated the reconstruction performance by visualizing the expression of individual marker genes and comparing their 2D projections with existing ISH images from the same perspective [5]. ISH images

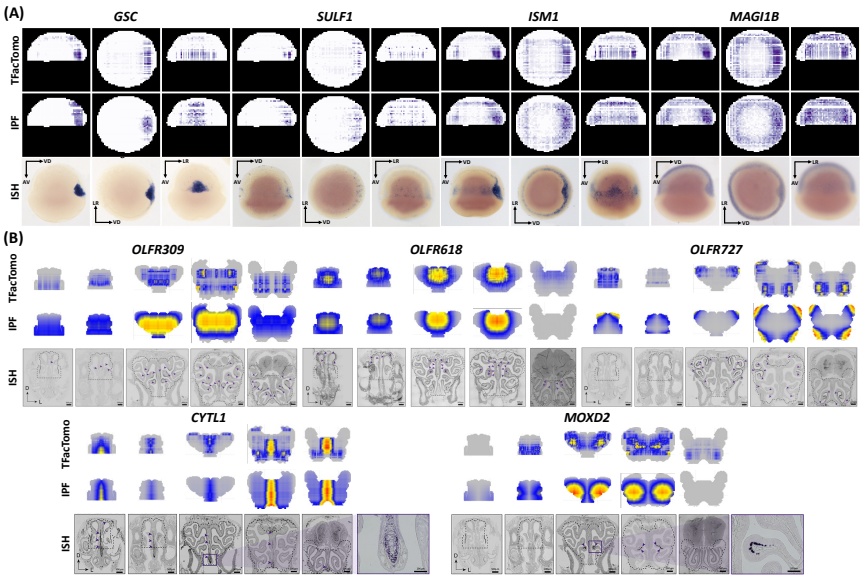

Fig. 2. Comparison of 3D-reconstructions and ISH images of marker genes. **(A)** 2D projection of the reconstructed expression for zebrafish embryo at shield stage along frontal, anteroposterior, and ventrodorsal axes of four genes (*GSC*, *SULF1*, *ISM1*, and *MAGI1B*). Shaded blue areas in the ISH images denote expression regions. **(B)** Consecutive 2D slices of the reconstructed expression patterns for mouse olfactory mucosa along the anteroposterior axis on five genes (*OLFR309*, *OLFR618*, *OLFR727*, *CYTL1* and *MOXD2*). In the ISH images, the dotted curves mark the boundary of the tissue and the small blue arrows mark the expression regions.

are available for four genes, *GSC*, *SULF1*, *ISM1*, and *MAGI1B*, and their reconstruction are shown in Fig. 2A. All these four genes are expressed as a patch or different gradient in the organizer on the dorsal side at the early development stages of zebrafish. *GSC* is an important gene in developing mesoderm. While the 3D reconstruction of both TFacTomo and IPF in some projections, such as along the frontal and anteroposterior axes, highlight the correct expressed regions in the ISH images, TFacTomo reconstructions align better to the ISH images without overspreading expression. Moreover, it is obvious that the projections of 3D reconstructions from TFacTomo along the ventrodorsal axis demonstrated significantly higher agreement with their ISH images compared to IPF.

3.3 3D Transcriptome Reconstruction for Mouse Olfactory Mucosa

Lastly, we experimented with another large RNA tomography dataset from a mouse olfactory mucosa [16]. The dataset contains $(56, 54, 60)$ slices in the three views. We obtained a *mus musculus* PPI network from BioGRID [13] and retained only 9,245 genes present in both the PPI network and expression profiles. Additional nine known marker genes in the original study were also included, totaling $9,254$ genes. The RNA tomography data was initially fitted with polynomial fitting and normalized using the procedure described in [5]. Parameters

$\alpha = 100$, $\beta = 1$, and rank$= 500$ were chosen for this dataset based on the hyperparameter tuning results shown in Fig. S6.

We then visualized the reconstructions from TFacTomo ($9,254 \times 56 \times 54 \times 60$) and IPF for comparisons on five marker genes with ISH images available [16]. The five marker genes (*Olfr309*, *Olfr618*, *Olfr727*, *Cytl1*, *Moxd2*) are shown in Fig. 2B. *Olfr309*, *Olfr727*, and *Olfr618* are three olfactory receptors and genetic markers of different mature olfactory sensory neuron subtypes. *Cytl1* is a cytokine-like protein playing roles in osteogenesis, chondrogenesis, and bone and cartilage homeostasis [17]. *Moxd2* is a mono-oxygenase dopamine hydroxylase-like protein functioning in olfaction [6]. We visualized slices along the anteroposterior axis and selected the slices that best matched the shape of the ISH images (see Fig. S7 and S8 for all the reconstructed slices). These visualizations show that TFacTomo is more precise than IPF for reconstruction. It is most clear when comparing a reconstructed slice of *Olfr618* along the anteroposterior axis with its last ISH image, where TFacTomo accurately reconstructs expression whereas IPF fails to reconstruct any. The fourth image of *Olfr727* shows another example of this, where IPF exaggerates the expression in the bottom two corners and TFacTomo more accurately shows the two concentrated regions of expression. We see a similar example in the fifth image of *Olfr727*, where IPF even incorrectly infers expression in the upper region whereas TFacTomo correctly infers that there is only expression in the bottom region of that slice. However, there are some instances among the visualized genes in which both TFacTomo and IPF incorrectly represent spatial gene expression. For example, both TFacTomo and IPF struggle to accurately reconstruct the expression of *Olfr309*. Note that since *Olfr618*'s expression level is only around 1% of the other genes, we set $\alpha = 1\text{e}{-3}$ to run TFacTomo separately such that the extremely low expressions can be retained in the reconstruction. In summary, through a visual inspection of the reconstructions, we see that TFacTomo can find more focused expression regions than IPF, likely due to IPF's tendency to overspread the regions on such data.

4 Discussions

Recently, microdissection-based methods such as variations of Tomo-seq, e.g. Geo-seq [14] and STRP-seq [15] are gaining more attention. These newer technologies use adjacent slices as replicates such that no identical biological replicate is necessary for tomography. In addition, more flexible tomography views can be adopted beyond the three orthogonal views [15]. These new developments greatly improved the applicability of microdissection-based methods. Based on our analysis and experiments, we concluded that TFacTomo is a better option than IPF. First, TFacTomo utilizes crucial spatial information and gene functional relations to better solve the difficult reconstruction problem. Second, TFacTomo is a multitask model for whole transcriptome reconstruction while IPF's single-gene construction is very prone to noise. Finally, TFacTomo is a high-order method that better characterizes the 3D organization of the information.

Acknowledgments. We sincerely thank Dr. Mayra Luisa and Dr. Antonio Scialdone for providing the data and scripts necessary to perform the mouse olfactory mucosa experiments and the ISH images for the evaluation of the results.

References

1. Asp, M., Giacomello, S., Larsson, L., et al.: A spatiotemporal organ-wide gene expression and cell atlas of the developing human heart. Cell **179**(7), 1647–1660 (2019)
2. Cai, D., He, X., Wu, X., Han, J.: Non-negative matrix factorization on manifold. In: 2008 eighth IEEE international conference on data mining. pp. 63–72. IEEE (2008)
3. Choi, J.H., Vishwanathan, S.: Dfacto: Distributed factorization of tensors. Advances in neural information processing systems **27** (2014)
4. Fienberg, S.E.: An iterative procedure for estimation in contingency tables. The Annals of Mathematical Statistics **41**(3), 907–917 (1970)
5. Junker, J.P., Noel, E.S., Guryev, V., et al.: Genome-wide rna tomography in the zebrafish embryo. Cell **159**(3), 662–675 (2014)
6. Kim, D.S., Wang, Y., Oh, H.J., et al.: Frequent loss and alteration of the moxd2 gene in catarrhines and whales: a possible connection with the evolution of olfaction. PLoS One **9**(8), e104085 (2014)
7. Lee, D., Seung, H.S.: Algorithms for non-negative matrix factorization. Advances in neural information processing systems **13** (2000)
8. Li, Z., Petegrosso, R., Smith, S., et al.: Scalable label propagation for multi-relational learning on the tensor product of graphs. IEEE Transactions on Knowledge and Data Engineering **34**(12), 5964–5978 (2021)
9. Li, Z., Zhang, W., Huang, R.S., Kuang, R.: Learning a low-rank tensor of pharmacogenomic multi-relations from biomedical networks. In: 2019 IEEE International Conference on Data Mining (ICDM). pp. 409–418. IEEE (2019)
10. Li, Z., Peng, G.: Spatial transcriptomics: new dimension of understanding biological complexity. Biophysics Reports **8**(3), 119 (2022)
11. Moses, L., Pachter, L.: Museum of spatial transcriptomics. Nature methods **19**(5), 534–546 (2022)
12. Ortiz, C., Navarro, J.F., Jurek, A., et al.: Molecular atlas of the adult mouse brain. Science advances **6**(26), eabb3446 (2020)
13. Oughtred, R., Rust, J., Chang, C., et al.: The biogrid database: A comprehensive biomedical resource of curated protein, genetic, and chemical interactions. Protein Science **30**(1), 187–200 (2021)
14. Peng, G., Suo, S., Chen, J., et al.: Spatial transcriptome for the molecular annotation of lineage fates and cell identity in mid-gastrula mouse embryo. Developmental cell **36**(6), 681–697 (2016)
15. Schede, H.H., Schneider, C.G., Stergiadou, J., et al.: Spatial tissue profiling by imaging-free molecular tomography. Nature Biotechnology **39**(8), 968–977 (2021)
16. Segura, M.L.R.T., Abou Moussa, E., Garabello, E., et al.: A 3d transcriptomics atlas of the mouse nose sheds light on the anatomical logic of smell. Cell Reports **38**(12) (2022)
17. Shin, Y., Won, Y., Yang, J.I., Chun, J.S.: Cytl1 regulates bone homeostasis in mice by modulating osteogenesis of mesenchymal stem cells and osteoclastogenesis of bone marrow-derived macrophages. Cell Death & Disease **10**(2), 47 (2019)
18. Smith, S., Ravindran, N., Sidiropoulos, N.D., Karypis, G.: Splatt: Efficient and parallel sparse tensor-matrix multiplication. In: 2015 IEEE International Parallel and Distributed Processing Symposium. pp. 61–70. IEEE (2015)
19. Szklarczyk, D., Kirsch, R., Koutrouli, M., et al.: The string database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic acids research **51**(D1), D638–D646 (2023)