Unmixing Biological Fluorescence Image Data with Sparse and Low-Rank Poisson Regression

Ruogu Wang[1], Alex A. Lemus[2, 3], Colin M. Henneberry[2, 3], Yiming Ying[1], Yunlong Feng[1, *, $], Alex M. Valm[2, 3, *, $]

1 Department of Mathematics and Statistics, University at Albany, SUNY, Albany, NY 12222, USA
2 Department of Biology, University at Albany, SUNY, Albany, NY 12222, USA
3 RNA Institute, University at Albany, SUNY, Albany, NY 12222, USA

*** To whom correspondence should be addressed**
$ The authors wish it to be known that, in their opinion, the last two authors should be regarded as Joint Last Authors.

Abstract

Multispectral biological fluorescence microscopy has enabled the identification of multiple targets in complex samples. The accuracy in the unmixing result degrades (1) as the number of fluorophores used in any experiment increases and (2) as the signal-to-noise ratio in the recorded images decreases. Further, the availability of prior knowledge regarding the expected spatial distributions of fluorophores in images of labeled cells provides an opportunity to improve the accuracy of fluorophore identification and abundance. We propose a regularized sparse and low-rank Poisson unmixing approach (SL-PRU) to deconvolve spectral images labeled with highly overlapping fluorophores which are recorded in low signal-to-noise regimes. Firstly, SL-PRU implements multi-penalty terms when pursuing sparseness and spatial correlation of the resulting abundances in small neighborhoods simultaneously. Secondly, SL-PRU makes use of Poisson regression for unmixing instead of least squares regression to better estimate photon abundance. Thirdly, we propose a method to tune the SL-PRU parameters involved in the unmixing procedure in the absence of knowledge of the ground truth abundance information in a recorded image. By validating on simulated and real-world images, we show that our proposed method leads to improved accuracy in unmixing fluorophores with highly overlapping spectra.

1 Introduction and motivation

Many biological systems are composed of multiple interacting subcomponents, any of which may be labeled with fluorescent reporters to map their spatial location within cells and tissues. While some progress has been achieved in developing fluorescent dyes with narrow emission spectra, e.g., semiconductor nanocrystals or quantum dots, most widely used organic fluorophores, including fluorescent proteins, have broad excitation and emission spectra [1, 18]. The inherent wide emission spectra of different fluorophores used in a single experiment lead to unavoidable overlap in spectral emission profiles and cross-talk between the recorded channels when samples are imaged with conventional bandpass filters. To overcome this limitation, fluorescence spectral imaging instrumentation and image analysis tools have been developed and applied to biological imaging. Spectral imaging microscopes collect fluorescence intensity information at every pixel in an image to construct a 3-dimensional data cube with spatial and spectral information from the sample as shown in Figure 1.

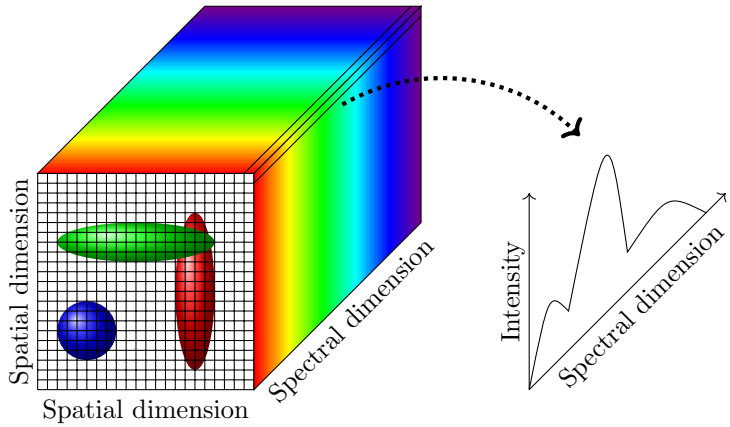

Figure 1: A hyperspectral data cube and spectral intensity information at each pixel with a generalized biological image, in which most foreground pixels record fluorescent signal from only one cell but some pixels record overlapped signals from two or more cells.

Among existing biological spectral imaging analysis methods, spectral unmixing, which aims at extracting the spectral signature of each fluorophore from recorded images and gaining knowledge of each fluorophore's abundance in every pixel, has been widely utilized. In particular, linear unmixing approaches, especially utilizing the least squares framework, have been widely applied since these approaches make no underlying assumptions about the image data except that the signal recorded in the same pixel from multiple fluorophores adds linearly, i.e., no a priori knowledge about the sample is required. Linear unmixing separates each pixel linearly into the spectral signatures of the contributing fluorophores, called *endmembers*, and their contributions, called *abundances*. Due to physical considerations, both endmembers and abundances should satisfy the nonnegativity constraint. Given a spectral image matrix $\mathbf{Y} \in \mathbb{R}_+^{C\times N}$ with C channels and N pixels, denoting $\mathbf{M}^\star \in \mathbb{R}_+^{C\times R}$ as the endmember matrix of the associated R fluorophores and $\mathbf{A}^\star \in \mathbb{R}_+^{R\times N}$ the corresponding abundance matrix, such a linear relationship can be expressed as

$$\mathbf{Y} = \mathbf{M}^\star \mathbf{A}^\star + \mathcal{E},$$

where $\mathcal{E} \in \mathbb{R}^{C\times N}$ denotes an unknown noise matrix.

Least squares (linear) unmixing procedures are implemented by decomposing a spectral image matrix into an endmember matrix and an abundance matrix simultaneously while minimizing the data fidelity error measured by the squared error criterion, e.g., the sum of the squared residuals. Mathematically, it can be formulated as follows

$$\min_{\mathbf{M}\succcurlyeq\mathbf{0},\mathbf{A}\succcurlyeq\mathbf{0}} \|\mathbf{Y} - \mathbf{M}\mathbf{A}\|_F^2,$$

where $\|\cdot\|_F$ denotes the Frobenius norm of a matrix, and $\mathbf{M} \succcurlyeq \mathbf{0}$ and $\mathbf{A} \succcurlyeq \mathbf{0}$ denote element-wise nonnegativity of the endmember matrix $\mathbf{M}$ and the abundance matrix $\mathbf{A}$. In biological spectral imaging, it is frequently the case that reference spectral images that consist of only one fluorophore—and so only one endmember—in each image are available, which is also the scenario considered in this study. In such a case, least squares unmixing procedures can be carried out in a two-step way: first, each reference spectral image matrix is decomposed into the outer product of two vectors, i.e., the endmember and its abundances; then, with the extracted endmember information from the first step,

the abundance matrix is estimated for the mixed image. Such a two-step approach is also considered in this study.

While least squares unmixing has been extensively studied and widely used in the spectral unmixing literature, its limitations are also well-recognized in the community. For instance, due to the involvement of matrix multiplication, least squares unmixing may not admit a unique solution, which leads to imprecise estimates of the abundances. In the literature, this problem is frequently addressed by imposing a certain penalty on A, which leads to

$$\min_{\mathbf{A} \succcurlyeq \mathbf{0}} \|\mathbf{Y} - \mathbf{M}\mathbf{A}\|_F^2 + \lambda \Omega(\mathbf{A}),$$

where $\lambda > 0$ is a tuning parameter, $\Omega(\mathbf{A})$ denotes the penalty imposed on $\mathbf{A}$, and $\mathbf{M}$ is assumed to be known from reference images. The penalty term $\lambda\Omega(\mathbf{A})$ controls the complexity of the matrix space within which $\mathbf{A}$ is searched and thus may lead to a unique solution. Another disadvantage of the least squares unmixing approach lies in the use of the least squares error criterion, which originates from the maximum likelihood estimation when assuming Gaussian noise. In fact, the presence of purely Gaussian noise is rarely the case in biological spectral images. This is because the spectrum at each pixel of a biological fluorescence image corresponds to the photon counts recorded at every spectral channel and the uncertainties in these measurements can be better approximated by Poisson distributions rather than normal distributions (the so-called, photon shot noise); see e.g., [5]. Following this observation, Poisson regression approaches to spectral unmixing have been proposed in the literature; see e.g., [12, 18]. To date, existing applications of Poisson regression approaches to biological spectral unmixing comprise an active area of research, and the strengths and weaknesses of such approaches, when being compared with well-established least squares approaches, are yet to be fully elucidated.

Least squares approaches to spectral unmixing are implemented in a pixel-wise fashion, which requires no assumption regarding fluorophore distribution in the image. In biological spectral images, neighboring pixels are frequently similar in their fluorophore identities and abundances. This spatial correlation can be reflected as linear dependence between the corresponding abundances. As a result, the rank of the matrix formed by these abundance vectors may be limited, especially if we consider neighboring pixels in a small region. In addition, for a real biological spectral image, it is frequently the case that the recorded signal from any given pixel may only comprise one or a small number of endmembers, though, for the whole image, the number of involved endmembers may be far larger. However, least squares unmixing tends to treat all pixels and endmembers equally without taking such sparseness information into account, which may lead to imprecise abundance estimation when the number of contributing endmembers in any pixel is much smaller than that of the candidate pool across the entire image.

Here we address these gaps in knowledge and further take advantage of the prior information available regarding spectral images of fluorescently labeled cells by exploring a Poisson regression approach and by simultaneously seeking endmember-wise sparseness and low-rankness when estimating the abundances in a localized pattern. We propose in this paper a regularized sparse and low-rank Poisson regression approach (SL-PRU) to accurately estimate abundances in multiplex labeled images of cells. The proposed approach takes into account the non-Gaussian nature of the noise in the data and neighboring information, which are physically meaningful considerations in biological spectral imaging. To implement the low-rankness assumption of neighboring pixels, we make use of a sliding window technique, which allows us to consider the spectral signatures of adjacent pixels lying in the window. To make SL-PRU computationally tractable and to pursue further robustness, we consider convex relaxations of the penalty terms on the sparsity and the rank of the abundance matrix.

It is noticed from our empirical studies that unregularized Poisson regression may lead to satisfying endmember extraction results. We thus suspend the penalty terms when implementing SL-PRU to

extract endmembers in the first step from images with known fluorophore identities. In the second step, where SL-PRU is applied to unmix real biological images, we propose a constructive approach for tuning the parameters involved in the estimation without resorting to the unknown abundance matrix. We validate the proposed method on simulated spectral images and on real images of a microbial biofilm. The experimental results show that our proposed method can outperform existing approaches in biological imaging both quantitatively and qualitatively.

2 Materials and methods

2.1 Sample preparation

E. coli K12 (ATCC 10798) cells were grown to the mid-log phase in Luria-Bertani LB Broth (Difco Laboratories, Inc.). E. coli cultures and dental plaque smears were fixed in 2% paraformaldehyde (EMS Diasum) for 1.5 hours at room temperature, then stored in 50% ethanol for 24 hours before FISH labeling. E. coli cells were labeled with the general bacteria probe, EUB338 (GCTGCCTCC-CGTAGGAGT) conjugated to a fluorescent dye at the 5' end (Thermofisher). Plaque smear samples were obtained through self-flossing from healthy volunteers after giving informed consent. The use of human subjects for this study was approved by the University at Albany Institutional Review Board (IRB). Plaque samples were labeled with previously validated taxon-specific FISH probes (See Supplementary Table 1) and acquired as multi-plane z-stack images.

2.2 Imaging and pre-processing

Images were acquired on Zeiss LSM 710 or LSM 880 confocal microscopes with 32 anode spectral detectors. Images were acquired with 488, 561, and 633 nm laser excitation and collected on the 32-anode spectral detector with 9.8 nm width spectral resolution in each channel. E. coli images were acquired as a single plane with a 63x 1.4 NA objective. Plaque smear images and reference E. coli images were acquired as multi-plane Z-stack images with a 20x 0.8 NA objective.

2.3 SL-PRU: The proposed unmixing approach

Since the signal from multiple fluorophores adds linearly in a pixel, we can express the true photon counts as the sum of the endmembers weighted by their abundances. Considering that the recorded photon counts follow the Poisson distribution, a biological spectral image $\mathbf{Y} \in \mathbb{R}_+^{C\times N}$ with C channels and N pixels can be written as [13, 14, 16]

$$\mathbf{Y} = \text{Pois}(\mathbf{MA}),$$

where $\mathbb{R}_+^k$ denotes a nonnegative orthant in a k-dimensional Euclidean space, $\text{Pois}(\cdot)$ is element-wise Poisson probability distribution, $\mathbf{M} \in \mathbb{R}_+^{C\times R}$ represents the endmember matrix which consists of reference spectra of R fluorophores used for labeling, and $\mathbf{A} \in \mathbb{R}_+^{R\times N}$ denotes the abundance matrix.

To extract the endmember $\mathbf{m} \in \mathbb{R}_+^C$ from a reference image $\mathbf{Y_m} \in \mathbb{R}_+^{C\times N}$, we maximize the likelihood of observing $\mathbf{Y_m} \in \mathbb{R}_+^{C\times N}$ given $\mathbf{m}$ and the corresponding abundances $\mathbf{a} \in \mathbb{R}_+^N$:

$$\mathcal{L}(\mathbf{Y_m}; \mathbf{m}, \mathbf{a}) = \Pi_{c=1}^C \Pi_{n=1}^N \text{Pois}(y_{cn}) = \Pi_{c=1}^C \Pi_{n=1}^N \text{Pois}(m_c a_n),$$

where m_c is the intensity of the endmember in channel c, a_n is the corresponding abundance at pixel n, y_{cn} is the observed intensity in channel c at pixel n of the reference image. To maximize

$\mathcal{L}(\mathbf{Y_m}; \mathbf{m}, \mathbf{a})$ is to minimize the negative log of , $\mathcal{L}(\mathbf{Y_m}; \mathbf{m}, \mathbf{a})$:

$$\begin{aligned}
&-\log \mathcal{L}(\mathbf{Y_m}; \mathbf{m}, \mathbf{a}) \\
=&-\log \left(\Pi_{c=1}^{C} \Pi_{n=1}^{N} (m_c a_n)^{y_{cn}} \exp(-m_c a_n)\right) \\
=&\sum_{c=1}^{C} \sum_{n=1}^{N} \left[m_c a_n - y_{cn} \log(m_c a_n)\right],
\end{aligned}$$

where log denotes the natural logarithm. The endmember $\mathbf{m}$ can be obtained from a reference image $\mathbf{Y_m}$ through the maximization likelihood estimation:

$$\min_{\mathbf{m} \in \mathbb{R}_+^C, \mathbf{a} \in \mathbb{R}_+^N} \sum_{c=1}^{C} \sum_{n=1}^{N} \left[m_c a_n - y_{cn} \log(m_c a_n)\right],$$

which, in a compact form, is equivalent to Poisson Nonnegative Matrix Factorization (PNMF):

$$\min_{\mathbf{m} \in \mathbb{R}_+^C, \mathbf{a} \in \mathbb{R}_+^N} \mathbf{1}_C^\top \left[\mathbf{m}\mathbf{a}^\top - \mathbf{Y_m} \circ \log\left(\mathbf{m}\mathbf{a}^\top\right)\right] \mathbf{1}_N, \tag{1}$$

where $\mathbf{1}_C$ and $\mathbf{1}_N$ are vectors of length C and N whose entries are all 1 and $\circ$ denotes element-wise multiplication.

With endmembers extracted through reference images via (1), the abundance estimation of a spectral image that shares the same morphologies can be viewed as a nonnegative Poisson regression problem. As both endmember-wise sparsity and spatial correlation properties rely on the homogeneity of a small area, we utilize a sliding 3×3 window that contains the spectra of the target pixel and spectra of its surrounding pixels for unmixing as shown in Figure 2 [6]. In each window, we assume the contribution of only a few fluorophores which can be reflected as a limited number of endmembers with nonzero abundances. In the literature, the sparsity constraint is usually carried out via an ℓ_1 norm regularization that reduces the number of nonzero entries [7, 15]. To impose endmember-wise sparsity on the abundance matrix $\mathbf{A}$ in a window, we apply the constraint among the rows of $\mathbf{A}$ which turns out to be the $\ell_{2,1}$ norm $\|\mathbf{A}\|_{2,1} = \sum_{r=1}^{R} \|\mathbf{a}_r\|_2$ where $\mathbf{a}_r$ denotes the r-th row of $\mathbf{A}$ and $\|\cdot\|_2$ denotes the ℓ_2 norm [9]. The other property, spatial similarity of neighboring pixels, has also been widely exploited for hyperspectral unmixing [6, 8, 17]. In this work, the spatial correlation is incorporated by imposing the low-rankness constraint on $\mathbf{A}$. As a convex surrogate of the matrix rank, the nuclear norm $\|\mathbf{A}\|_*$ defined as the sum of its singular values, $\{\sigma_i(\mathbf{A})\}_{i=1}^{\text{rank}(\mathbf{A})}$, is used.

With a slight abuse of notation, in what follows, we denote $\mathbf{Y} \in \mathbb{R}_+^{C \times N}$ as a spectral window. The abundance matrix $\mathbf{A}$ can be estimated through the following sparse low-rank Poisson regression:

$$\min_{\mathbf{A} \in \mathbb{R}_+^{R \times N}} \mathbf{1}_C^\top \left[\mathbf{M}\mathbf{A} - \mathbf{Y} \circ \log\left(\mathbf{M}\mathbf{A}\right)\right] \mathbf{1}_N + \lambda_1 \|\mathbf{A}\|_* + \lambda_2 \|\mathbf{A}\|_{2,1}, \tag{2}$$

where λ_1 and λ_2 are nonnegative parameters that balance the fidelity term, the penalty term on sparsity, and the penalty term on low-rankness.

As convex relaxations of the ℓ_0 norm, i.e., the number of nonzero entries of a vector, and the rank function of a matrix, the ℓ_1 norm and the nuclear norm suffer from the influence of the magnitude [4, 11]. We thus use weighted formulations of nuclear norm and $\ell_{2,1}$ norm to democratically penalize the nonzero entries:

$$\|\mathbf{A}\|_{\mathbf{w}_p,*} = \sum_{i=1}^{\text{rank}(\mathbf{A})} w_{p,i} \sigma_i(\mathbf{A}),$$

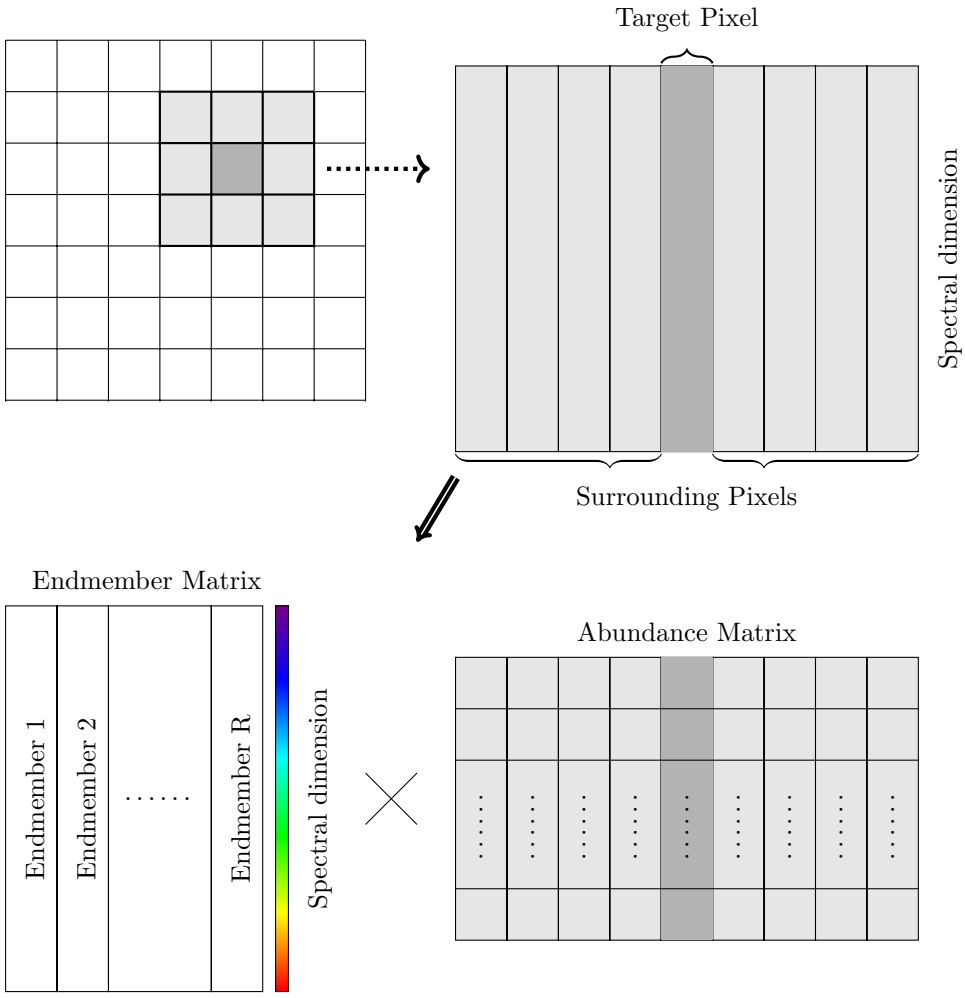

Figure 2: A 3×3 window unmixed into the product of its endmember and abundance matrices

and

$$\|\mathbf{W}_q\mathbf{A}\|_{2,1} = \sum_{r=1}^{R} w_{q,r}\|\mathbf{a}_r\|_2,$$

where

$$\mathbf{w}_p = \begin{bmatrix} w_{p,1} \\ w_{p,2} \\ \vdots \\ w_{p,\mathrm{rank}(\mathbf{A})} \end{bmatrix}, \mathbf{W}_q = \begin{bmatrix} w_{q,1} & 0 & \dots & 0 \\ 0 & w_{q,2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & w_{q,R} \end{bmatrix},$$

and $\{w_{p,i}\}_{i=1}^{\mathrm{rank}(\mathbf{A})}$, $\{w_{q,r}\}_{r=1}^{R}$ are nonnegative weighting coefficients that will be determined later. As a result, we arrive at the following variant of the above regularized sparse and low-rank Poisson regression method:

$$\begin{aligned} \min_{\mathbf{A}\in\mathbb{R}_+^{R\times N}} & \mathbf{1}_C^\top \left[\mathbf{M}\mathbf{A} - \mathbf{Y}\circ\log\left(\mathbf{M}\mathbf{A}\right)\right]\mathbf{1}_N + \lambda_1\|\mathbf{A}\|_{\mathbf{w}_p,*} \\ & + \lambda_2\|\mathbf{W}_q\mathbf{A}\|_{2,1}, \end{aligned} \tag{3}$$

an algorithm for solving which is detailed below.

2.4 Algorithms

Endmember extraction method (1) can be solved by *multiplicative update algorithm* [10] which is a diagonally rescaled version of gradient descent. Denoting any variable $\mathbf{X}$ at the t-th, iteration as $\mathbf{X}^{(t)}$ and the maximum norm as $\|\cdot\|_\infty$, the pseudocode of the multiplicative update algorithm is provided in Algorithm 1. The abundance vector $\mathbf{a}$ and the endmember vector $\mathbf{m}$ are initialized with random values that follow the standard uniform distribution $U(0,1)$. At each iteration, the endmember vector is standardized by dividing by its maximum for uniqueness. The algorithm stops when the relative change of the standardized endmember $\widetilde{\mathbf{m}}$ between the $(t-1)$-th and the t-th iterations given by

$$\frac{\|\widetilde{\mathbf{m}}^{(t)} - \widetilde{\mathbf{m}}^{(t-1)}\|_2^2}{\|\widetilde{\mathbf{m}}^{(t-1)}\|_2^2}$$

is less than a small threshold value.

We now turn to the discussion of the proposed algorithm for solving SL-PRU (3). Inspired by the work in [6], an alternating direction method of multipliers (ADMM) technique [3] is adopted in our study by first letting all elements of $\mathbf{w}_p$ be equal to ensure the convexity of the low-rankness regularization term in SL-PRU (3). Similar to the work in [6], we introduce auxiliary variables $\mathbf{V}_1 \in \mathbb{R}^{C\times N}$, $\mathbf{V}_2, \mathbf{V}_3, \mathbf{V}_4 \in \mathbb{R}^{R\times N}$ and reformulate SL-PRU (3) as follows

$$\begin{aligned} \min_{\mathbf{U},\mathbf{V}_1,\mathbf{V}_2,\mathbf{V}_3,\mathbf{V}_4} \quad & \mathbf{1}_C^\top \left[\mathbf{V}_1 - \mathbf{Y}\circ\log\left(\mathbf{V}_1\right)\right]\mathbf{1}_N + \lambda_1\|\mathbf{V}_2\|_{\mathbf{w}_p,*} \\ & + \lambda_2\|\mathbf{W}_q\mathbf{V}_3\|_{2,1} + \mathcal{I}_{\mathbb{R}_+}(\mathbf{V}_4), \\ \text{s.t.} \quad & \mathbf{V}_1 = \mathbf{M}\mathbf{U}, \mathbf{V}_2 = \mathbf{U}, \mathbf{V}_3 = \mathbf{U}, \mathbf{V}_4 = \mathbf{U}, \end{aligned} \tag{4}$$

where $\mathcal{I}_{\mathbb{R}_+}(\cdot)$ is the indicator function which is zero if all the entries are nonnegative and infinity otherwise. Then we obtain the following augmented Lagrangian function for the optimization problem

Input: A reference image $\mathbf{Y_m} \in \mathbb{R}_+^{C\times N}$;
Output: The standardized endmember vector $\widetilde{\mathbf{m}}$;
Initialization: $\mathbf{a}^{(0)}$, $\mathbf{m}^{(0)}$;
repeat

$$a_n^{(t)} \leftarrow \widetilde{a}_n^{(t-1)} \cdot \left(\frac{\sum_c y_{cn}\widetilde{m}_c^{(t-1)}}{\widetilde{m}_c^{(t-1)}\widetilde{a}_n^{(t-1)}}\right) \Big/ \left(\sum_c \widetilde{m}_c^{(t-1)}\right);$$

$$m_c^{(t)} \leftarrow \widetilde{m}_c^{(t-1)} \cdot \left(\frac{\sum_n y_{cn}a_n^{(t)}}{\widetilde{m}_c^{(t-1)}a_n^{(t)}}\right) \Big/ \left(\sum_n a_n^{(t)}\right);$$

$$\widetilde{m}_c^{(t)} \leftarrow m_c^{(t)} / \|\mathbf{m}^{(t)}\|_\infty;$$

$$\widetilde{a}_n^{(t)} \leftarrow a_n^{(t)} \cdot \|\mathbf{m}^{(t)}\|_\infty;$$

until *the stopping criterion is satisfied*;

Algorithm 1: Multiplicative update algorithm for PNMF (1)

(4):

$$\begin{aligned}
&\mathcal{L}_1(\mathbf{U},\mathbf{V}_1,\mathbf{V}_2,\mathbf{V}_3,\mathbf{V}_4,\mathbf{D}_1,\mathbf{D}_2,\mathbf{D}_3,\mathbf{D}_4)\\
=&\,\mathbf{1}_C^\top\left[\mathbf{V}_1-\mathbf{Y}\circ\log\left(\mathbf{V}_1\right)\right]\mathbf{1}_N+\lambda_1\|\mathbf{V}_2\|_{\mathbf{w}_p,*}+\lambda_2\|\mathbf{W}_q\mathbf{V}_3\|_{2,1}\\
&+\mathcal{I}_{\mathbb{R}_+}(\mathbf{V}_4)+\mathrm{tr}(\mathbf{D}_1^\top(\mathbf{V}_1-\mathbf{M}\mathbf{U}))+\mathrm{tr}(\mathbf{D}_2^\top(\mathbf{V}_2-\mathbf{U}))\\
&+\mathrm{tr}(\mathbf{D}_3^\top(\mathbf{V}_3-\mathbf{U}))+\mathrm{tr}(\mathbf{D}_4^\top(\mathbf{V}_4-\mathbf{U}))+\frac{\mu}{2}\left(\|\mathbf{M}\mathbf{U}-\mathbf{V}_1\|_F^2\right.\\
&\left.+\|\mathbf{U}-\mathbf{V}_2\|_F^2+\|\mathbf{U}-\mathbf{V}_3\|_F^2+\|\mathbf{U}-\mathbf{V}_4\|_F^2\right),
\end{aligned} \tag{5}$$

where $\mathbf{D}_1 \in \mathbb{R}^{C\times N}$, $\mathbf{D}_2,\mathbf{D}_3,\mathbf{D}_4 \in \mathbb{R}^{R\times N}$ denote the Lagrange multipliers, $\mathrm{tr}(\cdot)$ denotes the trace of a matrix, $\mu > 0$ is a Lagrange multiplier regularization parameter, and $\|\cdot\|_F$ denotes the Frobenius norm. Denoting the identity matrix of size $k\times k$ as $\mathbf{I}_k$ and scaled Lagrange multipliers $\mathbf{D}_i' = \mathbf{D}_i/\mu, i = 1,2,3,4$, the augmented Lagrangian function $\mathcal{L}_1$ can be rewritten as

$$\begin{aligned}
\mathcal{L}_2(\mathbf{U},\mathbf{V},\mathbf{D}) =&\,\mathbf{1}_C^\top\left[\mathbf{V}_1-\mathbf{Y}\circ\log\left(\mathbf{V}_1\right)\right]\mathbf{1}_N+\lambda_1\|\mathbf{V}_2\|_{\mathbf{w}_p,*}\\
&+\lambda_2\|\mathbf{W}_q\mathbf{V}_3\|_{2,1}+\mathcal{I}_{\mathbb{R}_+}(\mathbf{V}_4)\\
&+\frac{\mu}{2}\|\mathbf{G}\mathbf{U}+\mathbf{B}\mathbf{V}-\mathbf{D}\|_F^2,
\end{aligned} \tag{6}$$

where

$$\mathbf{V}=\begin{bmatrix}\mathbf{V}_1\\\mathbf{V}_2\\\mathbf{V}_3\\\mathbf{V}_4\end{bmatrix},\mathbf{D}=\begin{bmatrix}\mathbf{D}_1'\\\mathbf{D}_2'\\\mathbf{D}_3'\\\mathbf{D}_4'\end{bmatrix},\mathbf{G}=\begin{bmatrix}\mathbf{M}\\\mathbf{I}_R\\\mathbf{I}_R\\\mathbf{I}_R\end{bmatrix},\mathbf{B}=-\mathbf{I}_{C+3R}.$$

The proposed ADMM-type algorithm for solving SL-PRU sequentially optimizes (5) or (6) with respect to each variable while the other variables remain as the latest values. Optimizing $\mathcal{L}_2$ w.r.t. $\mathbf{U}$ gives

$$\begin{aligned}
\mathbf{U}^{(t)} =&\arg\min_{\mathbf{U}}\mathcal{L}_2\left(\mathbf{U},\mathbf{V}^{(t-1)},\mathbf{D}^{(t-1)}\right)\\
=&\left(\mathbf{M}^\top\mathbf{M}+3\mathbf{I}_R\right)^{-1}\left[\mathbf{M}^\top\left(\mathbf{V}_1^{(t-1)}+\mathbf{D}_1^{'(t-1)}\right)+\mathbf{V}_2^{(t-1)}\right.\\
&\left.+\mathbf{D}_2^{'(t-1)}+\mathbf{V}_3^{(t-1)}+\mathbf{D}_3^{'(t-1)}+\mathbf{V}_4^{(t-1)}+\mathbf{D}_4^{'(t-1)}\right].
\end{aligned} \tag{7}$$

Letting $\mathbf{V}_1^{\prime(t)} = \mathbf{M}\mathbf{U}^{(t)} - \mathbf{D}_1^{'(t-1)} - 1/\mu$, the optimization of $\mathcal{L}_2$ w.r.t. $\mathbf{V}_1$ gives

$$\begin{aligned}\mathbf{V}_1^{(t)} =& \arg\min_{\mathbf{V}_1} \mathcal{L}_2 \left(\mathbf{U}^{(t)}, \begin{bmatrix} \mathbf{V}_1 \\ \mathbf{V}_2^{(t-1)} \\ \mathbf{V}_3^{(t-1)} \\ \mathbf{V}_4^{(t-1)} \end{bmatrix}, \mathbf{D}^{(t-1)} \right) \\ =& \left(\mathbf{V}_1^{\prime(t)} + \sqrt{\mathbf{V}_1^{\prime(t)} \circ \mathbf{V}_1^{\prime(t)} + 4\mathbf{Y}/\mu} \right) \Big/ 2,\end{aligned} \tag{8}$$

where $\sqrt{\cdot}$ denotes the element-wise square root.

Denoting the singular value decomposition of $\mathbf{U}^{(t)} - \mathbf{D}_2^{'(t-1)}$ as $\mathbf{S}_l \mathbf{\Sigma}^{(t)} \mathbf{S}_r^\top$, the optimization w.r.t. $\mathbf{V}_2$ gives

$$\begin{aligned}\mathbf{V}_2^{(t)} =& \arg\min_{\mathbf{V}_2} \mathcal{L}_2 \left(\mathbf{U}^{(t)}, \begin{bmatrix} \mathbf{V}_1^{(t)} \\ \mathbf{V}_2 \\ \mathbf{V}_3^{(t-1)} \\ \mathbf{V}_4^{(t-1)} \end{bmatrix}, \mathbf{D}^{(t-1)} \right) \\ =& \mathbf{S}_l [\mathrm{sign}(\mathbf{\Sigma}^{(t)}) \circ \max\{\mathbf{0}, \mathbf{\Sigma}^{(t)} - \lambda_1 \mathrm{diag}(\mathbf{w}_p)/\mu\}] \mathbf{S}_r^\top,\end{aligned} \tag{9}$$

where $\mathrm{sign}(\cdot)$ is the element-wise sign function, $\max\{\cdot,\cdot\}$ denotes the element-wise max function, and $\mathrm{diag}(\cdot)$ creates a matrix with diagonal elements equal to the vector elements.

Denoting $\mathbf{V}_{3,r}^{(t)}$ as the r-th row of $\mathbf{V}_3^{(t)}$ and $\mathbf{x}_r^{(t)}$ as the r-th row of $\mathbf{U}^{(t)} - \mathbf{D}_3^{'(t-1)}$ where $r = 1, \ldots, R$, each row of $\mathbf{V}_3$ is updated sequentially as

$$\begin{aligned}\mathbf{V}_{3,r}^{(t)} =& \arg\min_{\mathbf{V}_{3,r}} \mathcal{L}_2 \left(\mathbf{U}^{(t)}, \begin{bmatrix} \mathbf{V}_1^{(t)} \\ \mathbf{V}_2^{(t)} \\ \mathbf{V}_{3,1}^{(t)} \\ \vdots \\ \mathbf{V}_{3,r-1}^{(t)} \\ \mathbf{V}_{3,r} \\ \mathbf{V}_{3,r+1}^{(t-1)} \\ \vdots \\ \mathbf{V}_{3,R}^{(t-1)} \\ \mathbf{V}_4^{(t-1)} \end{bmatrix}, \mathbf{D}^{(t-1)} \right) \\ =& \frac{\mathbf{x}_r^{(t)} \max\{\|\mathbf{x}_r^{(t)}\|_2 - \lambda_2 w_{q,r}/\mu, 0\}}{\max\{\|\mathbf{x}_r^{(t)}\|_2 - \lambda_2 w_{q,r}/\mu, 0\} + \lambda_2 w_{q,r}/\mu}.\end{aligned} \tag{10}$$

The optimization w.r.t. $\mathbf{V}_4$ gives

$$\begin{aligned}\mathbf{V}_4^{(t)} =& \arg\min_{\mathbf{V}_4} \mathcal{L}_2 \left(\mathbf{U}^{(t)}, \begin{bmatrix} \mathbf{V}_1^{(t)} \\ \mathbf{V}_2^{(t)} \\ \mathbf{V}_3^{(t)} \\ \mathbf{V}_4 \end{bmatrix}, \mathbf{D}^{(t-1)} \right) \\ =& \max\left\{ \mathbf{U}^{(t)} - \mathbf{D}_4^{'(t-1)}, \mathbf{0} \right\}.\end{aligned} \tag{11}$$

The scaled Lagrange multipliers, $\mathbf{D}_1'$, $\mathbf{D}_2'$, $\mathbf{D}_3'$, and $\mathbf{D}_4'$, are updated as follows

$$\begin{aligned}\mathbf{D}_1'^{(t)} =& \mathbf{D}_1'^{(t-1)} - \mathbf{M}\mathbf{U}^{(t)} + \mathbf{V}_1^{(t)},\\ \mathbf{D}_i'^{(t)} =& \mathbf{D}_i'^{(t-1)} - \mathbf{U}^{(t)} + \mathbf{V}_i^{(t)}, \quad i = 2,3,4.\end{aligned}$$

The stopping criteria adopted in the algorithm are based on the primal and dual residuals [3] $\mathbf{r}_p$ and $\mathbf{r}_d$ given by

$$\mathbf{r}_p = \mathbf{G}\mathbf{U}^{(t)} + \mathbf{B}\mathbf{V}^{(t)}, \quad \mathbf{r}_d = \mu\mathbf{G}^\top\mathbf{B}\left(\mathbf{V}^{(t)} - \mathbf{V}^{(t-1)}\right),$$

that go to 0, respectively, as $t \to \infty$. The algorithm terminates whenever any of the ℓ_2 norm of $\mathbf{r}_p$ or $\mathbf{r}_d$ is less than a small threshold value or some number of iterations is reached.

To enhance the performance of the algorithm, as done in [6], we also update the weights $\mathbf{w}_p$ and $\mathbf{W}_q$ based on $\mathbf{A}^{(t)}$ at each iteration t as follows

$$w_{p,i}^{(t)} = \frac{1}{\sigma_i(\mathbf{A}^{(t)}) + \varepsilon}, \quad w_{q,r} = \frac{1}{\|\mathbf{a}_r^{(t)}\|_2 + \varepsilon},$$

where $\varepsilon > 0$ is assigned a small value to avoid singularities.

The pseudo-code of the proposed algorithm for solving SL-PRU (3) is presented in Algorithm 2. Detailed explanations on how the abundance matrix $\mathbf{U}$ and the auxiliary variables $\mathbf{V}_1$, $\mathbf{V}_2$, $\mathbf{V}_3$, $\mathbf{V}_4$ in Algorithm 2 are updated are deferred to the Detailed Explanations in the Supplemental Document.

Input: The data matrix $\mathbf{Y} \in \mathbb{R}_+^{C\times N}$, and the endmember matrix $\mathbf{M} \in \mathbb{R}_+^{C\times R}$;
Output: The abundance matrix $\mathbf{U}$;
Initialization: $\mathbf{U}^0, \mathbf{V}_i^0, \mathbf{D}_i^0, \quad i = 1,2,3,4$;
repeat

$\mathbf{U}^{(t)} \leftarrow \left(\mathbf{M}^\top\mathbf{M} + 3\mathbf{I}_R\right)^{-1}\left[\mathbf{M}^\top\left(\mathbf{V}_1^{(t-1)} + \mathbf{D}_1^{'(t-1)}\right) + \mathbf{V}_2^{(t-1)} + \mathbf{D}_2^{'(t-1)} + \mathbf{V}_3^{(t-1)} + \mathbf{D}_3^{'(t-1)} + \mathbf{V}_4^{(t-1)} + \mathbf{D}_4^{'(t-1)}\right]$;

$\mathbf{V}_1^{(t)} \leftarrow \left(\mathbf{V}_1^{(t)} + \sqrt{\mathbf{V}_1^{(t)} \circ \mathbf{V}_1^{(t)} + 4\mathbf{Y}/\mu}\right)\Big/2$;

$\mathbf{V}_2^{(t)} \leftarrow \mathbf{S}_l[\mathrm{sign}(\mathbf{\Sigma}) \circ \max\{\mathbf{0}, \mathbf{\Sigma} - \lambda_1\mathrm{diag}(\mathbf{w}_p)/\mu\}]\mathbf{S}_r^\top$;

$\mathbf{V}_{3,r}^{(t)} \leftarrow (\mathbf{x}_r^{(t)}\max\{\|\mathbf{x}_r^{(t)}\|_2 - \lambda_2 w_{q,r}/\mu, 0\})/(\max\{\|\mathbf{x}_r^{(t)}\|_2 - \lambda_2 w_{q,r}/\mu, 0\} + \lambda_2 w_{q,r}/\mu)$, $r = 1,\ldots,R$;

$\mathbf{V}_4^{(t)} \leftarrow \max\left\{\mathbf{U}^{(t)} - \mathbf{D}_4'^{(t-1)}, \mathbf{0}\right\}$;

$\mathbf{D}_1'^{(t)} \leftarrow \mathbf{D}_1'^{(t-1)} - \mathbf{M}\mathbf{U}^{(t)} + \mathbf{V}_1^{(t)}$;

$\mathbf{D}_i'^{(t)} \leftarrow \mathbf{D}_i'^{(t-1)} - \mathbf{U}^{(t)} + \mathbf{V}_i^{(t)}, \quad i = 2,3,4$;

until *the stopping criteria are satisfied*;

Algorithm 2: The proposed ADMM-type algorithm for SL-PRU (3)

3 Experiments and results

The performance of our proposed approach, SL-PRU, is compared with the commonly used linear unmixing methods: nonnegative least squares approach (NLS) [19], sparse NLS (S-NLS) [2,15], and sparse and low-rank NLS (SL-NLS) [6].

We carry out three different sets of experiments in which different types of spectral image data are used. In the first set of experiments, we unmix reference images of E. coli cells, in which every cell in the image contains a single endmember of knonw identity. Therefore, the performance of the different approaches in unmixing this data set can be compared by evaluating the proportion of the correct endmember assigned to each foreground pixel in the images while assuming all endmembers are also involved in each image. In our second set of experiments, we consider simulated spectral images, which are generated by using simulated abundance matrices and uncorrelated and correlated endmembers and by adding Poisson noise with different signal-to-noise ratios (SNRs). The root mean square errors (RMSEs) of the unmixing solutions from the different methods are evaluated for comparison. As a third set of experiments, we also evaluate the effectiveness of our proposed approach on real biological spectral images of labeled microbial biofilms and compare the results with NLS.

For all three experiments, the endmember spectra used for abundance estimation are extracted through PNMF (1) from known reference images. The tuning parameters λ_1 and λ_2 in (3) are chosen from the set $\{0, 10^{-3}, 10^{-2}, 10^{-1}, 1, 10\}$, and μ in Algorithm 2 is fixed to 0.01.

3.1 Reference images: endmember extraction and unmixing

We first extract endmembers using PNMF (1) on thirteen reference images of labeled E. coli cells and apply SL-PRU to unmix them to estimate abundances. The standardized fluorometer measured spectra of thirteen endmembers against wavelength and the standardized endmember matrices extracted by the arithmetic mean method [15] and PNMF are displayed in Supplementary Figure 1 for comparison. See Supplementary Table 2 for the full names of all fluorophores and their abbreviations used here.

Given an estimated abundance vector $\mathbf{a} = (a_1, a_2, \cdots, a_R)^\top$, where a_r, $r = 1, \cdots, R$, denotes the estimated abundance of the r-th endmember of a pixel from a reference image being unmixed. Then, the proportion of the involved endmember in this pixel is defined as

$$\mathrm{Prop}_{r_0}(\mathbf{a}) = \frac{a_{r_0}}{\sum_{r=1}^{R} a_r},$$

where a_{r_0} denotes the abundance of this endmember in the pixel. For a reference image, the average proportion (of the involved endmember) is defined as the average of such proportions across all pixels in this reference image. Note that for each reference image, only one endmember is involved. Therefore, the closer the average proportion is to 1, the better the estimated abundance vector.

With each tuning parameter for S-NLS or each pair of tuning parameters for SL-NLS or SL-PRU, we can obtain thirteen average proportions for all thirteen reference images. We choose the minimum of the thirteen average proportions for comparisons, since this represents the most conservative measure of unmixing accuracy. The tuning parameter values that reach the highest minimum of the thirteen average proportions obtained through each method are regarded as the optimal values. The average proportions of thirteen reference images obtained through each method: NLS, S-NLS, SL-NLS, and our proposed model SL-PRU with the optimal parameters are reported in Table 3.1. As shown in Table 3.1, our proposed method provides the highest average proportions among all the methods for all thirteen reference images.

3.2 Unmixing simulated spectral images

Using the endmembers extracted from the same reference images as in the first set of experiments above, we simulate spectral images of size 3×3 pixels with two uncorrelated endmembers AF514 and

	AF488	AF514	TET	AF532	AF546	AF555	RRX	AF568	AF594	AF647	AF660	AF680	AF700
NLS	93.1%	60.2%	70.6%	43.6%	79.0%	44.5%	68.8%	61.9%	77.0%	93.9%	64.5%	54.2%	74.4%
S-NLS	97.3%	86.9%	90.0%	62.6%	92.8%	69.7%	82.9%	78.3%	84.9%	98.1%	79.6%	77.3%	81.1%
SL-NLS	98.7%	96.4%	93.2%	72.1%	96.8%	76.8%	91.2%	89.2%	91.3%	99.0%	86.7%	85.4%	87.7%
SL-PRU	**99.5%**	**99.5%**	**94.5%**	**93.6%**	**98.6%**	**84.1%**	**97.8%**	**95.0%**	**93.7%**	**99.2%**	**95.7%**	**94.4%**	**89.4%**

Table 1: Optimal average proportions of endmembers estimated by NLS, NLS with sparsity constraint (S-NLS), NLS with sparsity and low-rank constraints (SL-NLS), and Poisson regression with sparsity and low-rank constraints (SL-PRU).

RRX and two correlated ones, AF555 and RRX, respectively. The root mean square error criterion is used to compare the performance of different methods, where

$$\mathrm{RMSE}_{\mathbf{A}}(\hat{\mathbf{A}}) = \sqrt{\frac{\|\mathbf{A} - \hat{\mathbf{A}}\|_F^2}{NR}},$$

where $\mathbf{A}$ is the simulated abundance matrix that serves as the underlying truth and $\hat{\mathbf{A}}$ the estimated abundance matrix.

3.2.1 Simulation 1: Uncorrelated endmembers

The abundances of AF514 and RRX are generated from the uniform distribution $U[0, 1]$ and the abundances of other endmembers are set to zero. Then the simulated spectral image is the multiplication of the generated abundance matrix and the endmember matrix. Poisson noise with nine SNRs ranging from 2 to 10 is added to the simulated data, and 1000 realizations of such simulated images are generated for each SNR.

With the optimal parameter(s) being selected for the cases with different SNRs, the averaged RMSEs in $1,000$ repetitions of all the unmixing methods under comparison are plotted in Figure 3. In particular, when SNR is set to 5, the optimal average RMSE is obtained from SL-PRU with $\lambda_1 = 0.1$ and $\lambda_2 = 1$. Moreover, the average of the estimated abundance matrices for this scenario is plotted in the top row of Figure 5.

From the reported experimental results in Figure 3 and the top row of Figure 5, it is observed that our proposed method can outperform other unmixing approaches under comparison here. For instance, as shown in Figure 3, our proposed method gives the lowest RMSEs with different SNRs. And as illustrated in Figure 5, the estimated average abundance matrix is more close to the true abundance matrix.

3.2.2 Simulation 2: Correlated endmembers

We next apply SL-PRU to unmix simulated spectral images with correlated endmembers AF555 and RRX under the same experimental setup as in Simulation I. Similarly, the averaged RMSEs from $1,000$ repetitions of all the unmixing methods are plotted in Figure 4, and when SNR is set to 5, the average of the estimated abundance matrices for this scenario is plotted in the bottom row of Figure 5.

As shown in Figure 4 and the bottom row of Figure 5, SL-PRU also outperforms other unmixing approaches under comparison here. It is noted that the correlation between endmembers does have an impact on the performance of SL-PRU as well as other unmixing approaches, which coincides

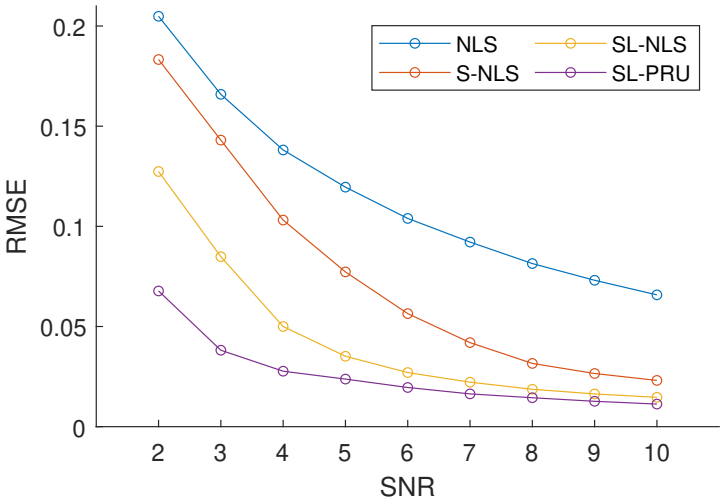

Figure 3: Averages of RMSEs of the abundance matrices from NLS, S-NLS, SL-NLS, and SL-PRU with simulated data of AF514 and RRX with Poisson noise and SNR of 2 to 10.

l

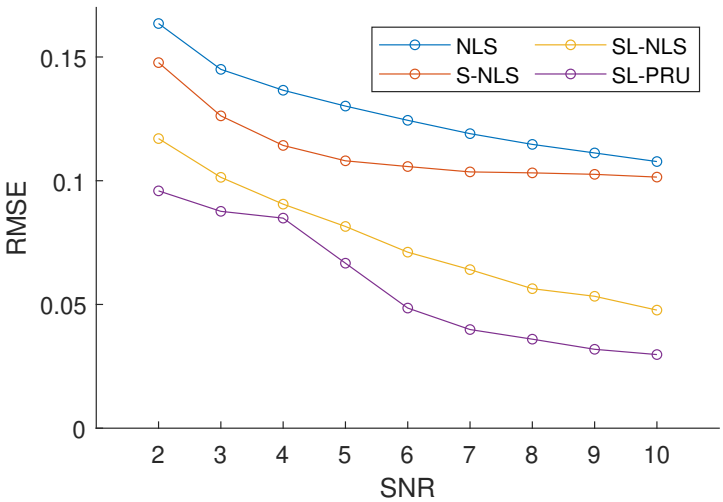

Figure 4: Averages of RMSEs of the abundance matrices estimated by NLS, S-NLS, SL-NLS, and SL-PRU from simulated data of AF555 and RRX with Poisson noise and SNR of 2 to 10.

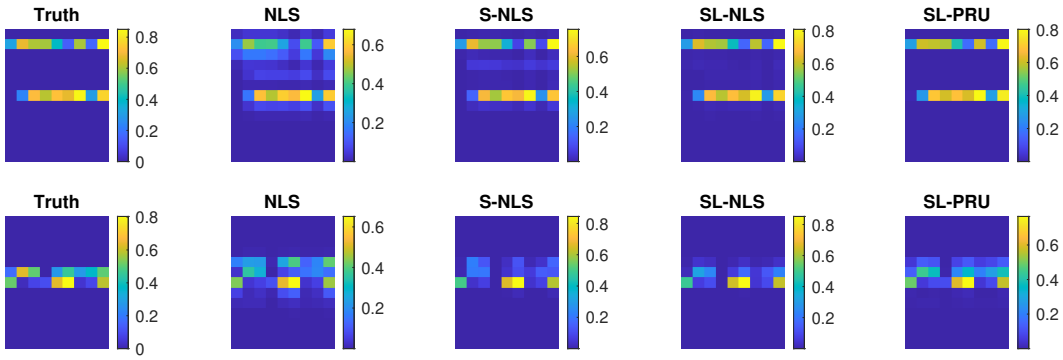

Figure 5: Graphical representation of simulated image pixels that contain two colocalized endmembers, either AF514 and RRX (highly uncorrelated endmembers, (Top row)) or AF555 and RRX (highly correlated endmembers (Bottom row)). In each row, the "Truth" matrix represents the ground truth starting simulation. Subsequent matrices represent the results of estimated abundances obtained from each of the unmixing methods that we considered. For each matrix, the 13 rows represent the 13 different endmembers used in the simulation and each column represents an independent pixel with varying intensity, scaled from 0-1. The color represents the mean abundance measure for each fluorophore from 1000 simulations of the same ground truth model after applying Poisson noise with SNR = 5 to each pixel.

with our intuitive understanding of linear unmixing problems. However, the experiments carried out in our study also suggest that SL-PRU may outperform other unmixing approaches in unmixing spectral images with correlated endmembers, which could be an appealing feature of SL-PRU in some practical scenarios.

3.3 Unmixing real mixed biological images

We next apply SL-PRU to unmix a real biological image with known fluorophore labels, but unknown spatial distributions and abundances. A dental plaque smear hedgehog structure was obtained from a healthy volunteer via dental flossing and labeled in a FISH experiment with taxon-specific probes for 8 different genera or families of bacteria, with each probe conjugated to a different fluorescent reporter. Reference spectra were obtained from images of separate populations of E. coli cells labeled with FISH probes conjugated to the eight fluorophores used in the plaque smear experiment and imaged under identical acquisition settings. Reference spectra were extracted using our Poisson endmember extraction procedure. The values for sparseness and low-rank tuning parameters were selected using a heuristic approach. Visual inspection of the unmixed plaque smear images was performed over the range of tuning parameters described above and values that maximized expected cell morphologies and minimized salt and pepper noise in the image were chosen for the final unmixing result (Figure 6 A-F and Supplementary Movie File 1). To evaluate the performance of our SL-PRU on this real image, we performed a comparative, quantitative cellular morphological analysis against the unmixing result we obtained for the same image using the commercial microscope vendor's linear unmixing algorithm. One plaque smear image set was used for quantitative comparison. The full z-stack image was unmixed in Zeiss Zen software, with endmember reference spectra extracted from single E. coli cells in images acquired with the same settings as the plaque smear image. The same spectral image data set was then unmixed using our SL-PRU approach. One central z-plane from both images

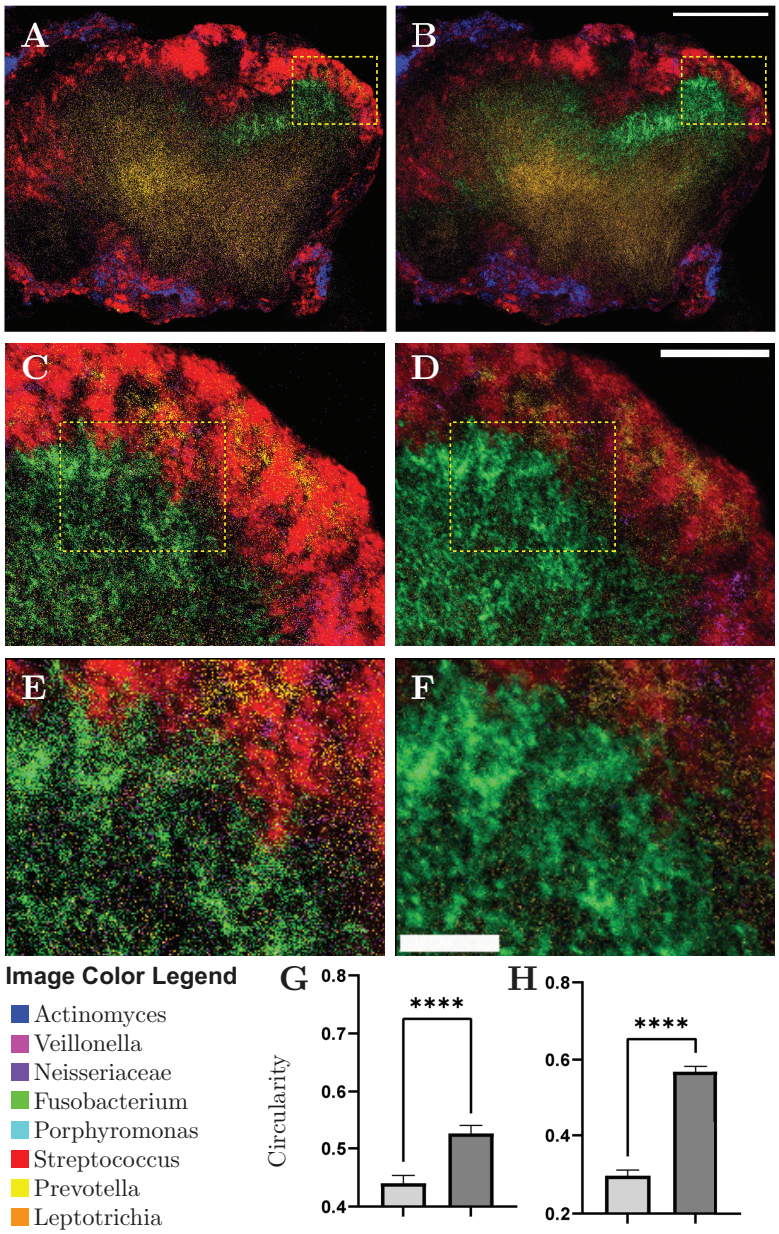

Figure 6: Qualitative and quantitative comparison of least squares and SL-PRU unmixing on a real biological sample. A-F: Full field of view multi-spectral image of a dental plaque smear hedgehog structure after (A,C,E) least squares unmixing and (B,D,F) SL-PRU unmixing. Dashed boxes in (A)and (B) and in (C) and (D) indicate zoom area in C-D, and E-F respectively. Scale bars equal 100 µm (B), 25 µm, (D), 10 µm (F). G-H: Quantitative comparison of mean circularity measurement per cell for two coccoid-shaped cells in the plaque structure: Streptococcus (G) and Veillonella (H). Light grey bars = results from least squares unmixing, dark grey bars = results from SL-PU. Error bars represent 95% confidence intervals. **** = p ¡ 0.001 Welch's t-test.

was extracted and used for quantitative comparison. Unmixed images were imported into ImageJ. The Streptococcus and Veillonella channels were segmented using an intensity threshold determined algorithmically using the same algorithm, either "Otsu" or "Triangle" for the Streptococcus channels and Veillonella channels respectively. Circularity analysis was performed on each segmented image, defined as

$$\text{Circ} = 4\pi \frac{\text{area}}{\text{perimeter}^2}.$$

Streptococcus and Veillonella cells in the sample have characteristic, near-perfect spherical shapes. With this a priori information about these two labeled cells, we measured the circularity of these two cell populations in one central plane of the plaque smear image and found that SL-PRU improved the circularity measure by 16% for Streptococcus and 18.4% for Veillonella (Figure 6 G-H). Lastly, we performed a line scan analysis on the same region of interest in the Streptococcus channel in both unmixed images and found that the SL-PRU approach generated an image with less noise and more identifiable cell boundaries than the commercial vendor least squares unmixing (Figure S1).

4 Discussion

Multispectral imaging has allowed the visualization and quantification of large numbers of targets simultaneously within biological specimens. The linear model assumption behind spectral unmixing holds for most biological images acquired from specimens that are thin and therefore have minimal scattering effects and in which fluorescent tags label components, e.g., cells or macromolecules that are well separated in space, beyond the minimal Förster Resonance Energy Transfer distance. The least squares approach to unmixing is essentially a multi-output regression problem, which assumes a Gaussian noise distribution in the recorded signals; while it is known that the dominant noise source in fluorescence images follows a Poisson distribution. To better accommodate such a Poisson distribution, and to improve fidelity in fluorophore abundance estimation in the unmixing process by using prior information about our labeled samples, we have developed in this paper a sparse and low-rank Poisson (SL-PRU) approach to multispectral unmixing. Specifically, we considered a two-step approach where we first extract endmember information through reference images using unregularized Poisson regression and then learn the abundance information via the proposed regularized Poisson approach. In the case of multiplex labeled cells such as the microbial biofilm samples used here, we assume that while many dozens of fluorophores might be present in the sample, for any individual pixel, the number of fluorophores present approaches one. We demonstrated the effectiveness of the proposed approach through experimental results on model images and real biological samples reported above.

We believe that our SL-PRU unmixing approach will be generally applicable to a wide variety of image datasets of multi-plex, fluorescently labeled cells. As microscope detector technologies improve, and the need for extracting information from ever lower numbers of photons increases, the dominant noise source in fluorescence images of cells is expected to shift ever more toward the Poisson-distributed, physically unavoidable photon shot noise. Even as the number of labeled targets increases, the physical constraints of these targets, whether they be cells or macromolecules, limits their simultaneous occurrence in pixels in digitally recorded images, although we recognize that the finite resolution of the light microscope and the labeling of target molecules that are sufficiently small and diffusible together put limitations on the sparsity assumption in some situations. In implementing our low-rank penalty term, we used a sliding window approach and restricted our neighborhood size to 3 x 3 pixels. In general, the size of the sliding window should be dictated by

the prior information available about the sample, e.g., the relative size of the labeled targets vs pixel dimensions in the image.

Finally, the effectiveness of the proposed sparse and low-rank Poisson approach has justified the importance of the low-rankness constraint in abundance estimation, which indicates the similarity of abundances of neighboring pixels. Note that such similarities, which agree with our intuitive understanding of spectral images, can be used to account for the spatial information among neighboring pixels. Recall that spectral images are three-way tensors and we unfold these tensor data into matrices before unmixing them. Therefore, further improved results may be obtained if the spatial information of the original tensor spectral data is taken into account in unmixing.

5 Conclusion

The existence of photon shot noise in biological fluorescence spectral images motivates us to address the unmixing problem through a Poisson approach instead of NLS. We also incorporate the spatial information by imposing sparsity and low-rankness constraints in a localized pattern. The unmixing results from both simulated data and real-world biological images demonstrate that our proposed approach SL-PRU can identify endmembers and estimate the corresponding abundances with increased accuracy over existing unmixing approaches.

Acknowledgement

This work was supported by the National Science Foundation [DMS-2111080 to Y.F., DMS-2110836, IIS-2103450, and IIS-2110546 to Y.Y.]; the Simons Foundation [#572064 to Y.F.]; and the National Institutes of Health [R01DE030927 to A.M.V.].

References

1. Barroso. Quantum dots in cell biology. *Journal of Histochemistry and Cytochemistry*, 59(3):237–251, 2011.

2. J. M. Bioucas-Dias and M. A. Figueiredo. Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing. In *2010 2nd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, pages 1–4. IEEE, 2010.

3. S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, et al. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine Learning*, 3(1):1–122, 2011.

4. E. J. Candes, M. B. Wakin, and S. P. Boyd. Enhancing sparsity by reweighted ℓ_1 minimization. *Journal of Fourier Analysis and Applications*, 14(5):877–905, 2008.

5. P. Coates. Photomultiplier noise statistics. *Journal of Physics D: Applied Physics*, 5(5):915, 1972.

6. P. V. Giampouras, K. E. Themelis, A. A. Rontogiannis, and K. D. Koutroumbas. Simultaneously sparse and low-rank abundance matrix estimation for hyperspectral image unmixing. *IEEE Transactions on Geoscience and Remote Sensing*, 54(8):4775–4789, 2016.

7. M.-D. Iordache, J. M. Bioucas-Dias, and A. Plaza. Sparse unmixing of hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, 49(6):2014–2039, 2011.

8. M.-D. Iordache, J. M. Bioucas-Dias, and A. Plaza. Total variation spatial regularization for sparse hyperspectral unmixing. *IEEE Transactions on Geoscience and Remote Sensing*, 50(11):4484–4502, 2012.

9. M.-D. Iordache, J. M. Bioucas-Dias, and A. Plaza. Collaborative sparse regression for hyperspectral unmixing. *IEEE Transactions on Geoscience and Remote sensing*, 52(1):341–354, 2013.

10. D. Lee and H. S. Seung. Algorithms for non-negative matrix factorization. *Advances in Neural Information processing Systems*, 13, 2000.

11. C. Lu, J. Tang, S. Yan, and Z. Lin. Generalized nonconvex nonsmooth low-rank minimization. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4130–4137, 2014.

12. R. Neher and E. Neher. Optimizing imaging parameters for the separation of multiple labels in a fluorescence image. *Journal of Microscopy*, 213(1):46–62, 2004.

13. R. A. Neher, M. Mitkovski, F. Kirchhoff, E. Neher, F. J. Theis, and A. Zeug. Blind source separation techniques for the decomposition of multiply labeled fluorescence images. *Biophysical Journal*, 96(9):3791–3800, 2009.

14. D. Novo, G. Grégori, and B. Rajwa. Generalized unmixing model for multispectral flow cytometry utilizing nonsquare compensation matrices. *Cytometry Part A*, 83(5):508–520, 2013.

15. B. J. Rossetti, S. A. Wilbert, J. L. Mark Welch, G. G. Borisy, and J. G. Nagy. Semi-blind sparse affine spectral unmixing of autofluorescence-contaminated micrographs. *Bioinformatics*, 36(3):910–917, 2020.

16. J. Xu, J. Bobin, A. de Vismes Ott, and C. Bobin. Sparse spectral unmixing for activity estimation in γ-ray spectrometry applied to environmental measurements. *Applied Radiation and Isotopes*, 156:108903, 2020.

17. X. Zhang, Y. Sun, J. Zhang, P. Wu, and L. Jiao. Hyperspectral unmixing via deep convolutional neural networks. *IEEE Geoscience and Remote Sensing Letters*, 15(11):1755–1759, 2018.

18. T. Zimmermann. Spectral imaging and linear unmixing in light microscopy. *Microscopy Techniques*, pages 245–265, 2005.

19. Zimmermann, Timo. Spectral imaging and linear unmixing in light microscopy. *Advances in Biochemical Engineering/Biotechnology*, 95:245–265, 2005.