Tensor decomposition of hyperspectral images to study autofluorescence in age-related macular degeneration

Neel Dey [a],[*], Sungmin Hong [a], Thomas Ach [b], Yiannis Koutalos [c], Christine A. Curcio [d], R. Theodore Smith [e], Guido Gerig [a]

[a] *Department of Computer Science and Engineering, NYU Tandon School of Engineering, NY, USA*
[b] *Department of Ophthalmology, University Hospital Würzburg, Würzburg, Germany*
[c] *Department of Ophthalmology, Medical University of South Carolina, SC, USA*
[d] *Department of Ophthalmology and Visual Sciences, School of Medicine, University of Alabama at Birmingham, AL, USA*
[e] *Department of Ophthalmology, Icahn School of Medicine, Mount Sinai, NY, USA*

A B S T R A C T

Autofluorescence is the emission of light by naturally occurring tissue components on the absorption of incident light. Autofluorescence within the eye is associated with several disorders, such as Age-related Macular Degeneration (AMD) which is a leading cause of central vision loss. Its pathogenesis is incompletely understood, but endogenous fluorophores in retinal tissue might play a role. Hyperspectral fluorescence microscopy of ex-vivo retinal tissue can be used to determine the fluorescence emission spectra of these fluorophores. Comparisons of spectra in healthy and diseased tissues can provide important insights into the pathogenesis of AMD. However, the spectrum from each pixel of the hyperspectral image is a superposition of spectra from multiple overlapping tissue components. As spectra cannot be negative, there is a need for a non-negative blind source separation model to isolate individual spectra. We propose a tensor formulation by leveraging multiple excitation wavelengths to excite the tissue sample. Arranging images from different excitation wavelengths as a tensor, a non-negative tensor decomposition can be performed to recover a provably unique low-rank model with factors representing emission and excitation spectra of these materials and corresponding abundance maps of autofluorescent substances in the tissue sample. We iteratively impute missing values common in fluorescence measurements using Expectation-Maximization and use L_2 regularization to reduce ill-posedness. Further, we present a framework for performing group hypothesis testing on hyperspectral images, finding significant differences in spectra between AMD and control groups in the peripheral macula. In the absence of ground truth, i.e. molecular identification of fluorophores, we provide a rigorous validation of chosen methods on both synthetic and real images where fluorescence spectra are known. These methodologies can be applied to the study of other pathologies presenting autofluorescence that can be captured by hyperspectral imaging.

© 2019 Elsevier B.V. All rights reserved.

1. Introduction

Tensor decompositions are a form of multi-linear algebra widely used in unsupervised machine learning (Anandkumar et al., 2014), signal processing (Cichocki et al., 2015), chemometrics (Smilde et al., 2005), psychometrics (Carroll and Chang, 1970), and have recently found applications in medical image analysis (Bharath et al., 2017; Schultz and Seidel, 2008; Cong et al., 2015; Li et al., 2017). While linear methods like Principal Component Analysis (Hotelling, 1933), Independent Component Analysis (Comon, 1994) and Non-negative Matrix Factorization (Paatero and Tapper, 1994; Lee and Seung, 1999) are common tools for unsupervised data analysis, they are limited when dealing with higher dimensional data. It is common to concatenate or average high dimensional data in order to apply these methods which may ignore covariance information across different dimensions and impede the derivation of accurate and/or interpretable solutions.

Tensor decompositions perform joint low-rank multi-linear approximations of higher dimensional data to extract distinct features for each dimension by accounting for covariance across all dimensions. Moreover, under mild conditions, they give unique solutions without strong constraints (Kruskal, 1989) often improving reliability and interpretability over linear methods. This makes tensor decompositions well suited to data analysis applications where multiple related datasets exist and interpretable features in each

dimension are expected. However, computing the decomposition and determining the number of features without ground truth are NP-hard problems (Hillar and Lim, 2013) that require careful modeling and validation.

As our driving application, microscopic images of human donor retina tissue samples were captured by hyperspectral cameras for a study of Age-related Macular Degeneration (AMD). AMD is a leading cause of legal blindness in the elderly (Bourne et al., 2013), with the projected worldwide prevalence of AMD reaching 288 million people in 2040 (Wong et al., 2014). Currently, there is no cure for the most common form of AMD and its pathogenesis is not well understood. Autofluorescence imaging of the ocular fundus is an important diagnostic tool for ophthalmologists who care for patients with AMD. Our long-term objective in microscopic analysis of human donor eye tissue is to understand the biology underlying hyperspectral autofluorescence imaging for in vivo early diagnosis.

The fluorescence spectrum of a fluorescent compound (i.e., a fluorophore) may provide information about both its molecular identity and cellular environment. In the case of AMD, analysis of retinal tissue autofluorescence holds particular significance, as endogenous fluorophores may have important pathogenic roles (Sparrow and Boulton, 2005). The minute amount of these fluorophores are a major challenge to their identification, yet this important information can, however, be obtained by determining their fluorescence spectra with hyperspectral imaging, in concert with molecular assays in parallel (e.g., mass spectrometry).

Hyperspectral cameras can finely capture the intensity spectrum of a scene over multiple discrete wavelengths. Using hyperspectral fluorescence microscopy, multiple co-registered image stacks illuminated with varying excitation wavelength can be acquired. In this way, the spectra of retinal tissue fluorophores at different excitation wavelengths can be determined. Fluorescence is additive and the signal from each pixel of a hyperspectral image is mixed in nature, with each pixel signal being a superposition of multiple layers of fluorescent tissue. This highlights the need for a blind source separation of multi-excitation fluorescence images, which is a natural extension of tensor methodology. The emission spectrum of a fluorophore typically varies with excitation wavelength in intensity but not in shape (Lakowicz, 2006a). This introduces another dimension of variation contained in the magnitude of the emission spectrum across excitation wavelengths called the *excitation* spectrum of a fluorophore. The emission and excitation spectra of a fluorophore can thus both provide important information about its identity. However, a direct application of tensor decomposition is inapplicable due to the nature of fluorescence measurements.

In this paper, we present a regularized tensor decomposition and data imputation method to extract the spectral and spatial characteristics of autofluorescent human eye tissue using hyperspectral images captured with varying excitation light sources. In the current absence of ground truth of retinal fluorophore spectra, we thoroughly validate our methods against both synthetic and real hyperspectral images with known fluorophore spectra and known tensor rank. To illustrate how the resulting tensor components can be utilized for distinguishing between groups, we present a hypothesis testing framework that considers the resulting emission spectra as functional data.

1.1. Related work

Decompositions of hyperspectral images are widely used in geospatial image analysis (Keshava and Mustard, 2002; Bioucas-Dias et al., 2012; Dobigeon et al., 2016). They possess known spectra of earth-based substances useful for validation, with strong justifications for additivity constraints, pure pixel assumptions, and

physical models of reflection. However, these methodologies do not easily translate to hyperspectral fluorescence microscopy. In the context of blind source separation, Independent Component Analysis has been explored for feature extraction in hyperspectral images, with limited success due to statistical independence assumptions (Nascimento and Dias, 2005). A popular alternative is Nonnegative Matrix Factorization (NMF), in which where a data matrix is decomposed into a set of non-negative basis vectors and coefficients. This non-negativity leads to better interpretation for some kinds of data analysis.

NMF applied to hyperspectral images for spectral unmixing (Sajda et al., 2003; Dobigeon et al., 2016), results in basis vectors and coefficients maps interpreted as spectral signatures and spatial distributions of spectral sources (also known as abundance images), respectively. With hyperspectral fluorescence microscopy of tissues, NMF has been used to extract fluorescence emission spectra (Gobinet et al., 2004; Neher et al., 2009; Montcuquet et al., 2011) including applications to the human retina (Johri et al., 2014). Due to rotational ambiguity, NMF based approaches provide non-unique solutions and may be non-reproducible due to local minima and initialization sensitivity. To address these issues, multi-set concatenation methodologies have been developed in retinal hyperspectral fluorescence microscopy in a series of papers (Smith et al., 2014; Ami et al., 2016; Tong et al., 2016) in which a tissue sample is repeatedly illuminated and imaged using several different excitation wavelengths. The corresponding hyperspectral image stacks are then concatenated and unmixed using NMF. However, multi-set concatenation approaches ignore covariances between the image stacks and can lead to overfitting. As NMF-based approaches have a diverse set of possible solutions, hypothesis testing for differences between groups of interest becomes challenging. Moreover, as they cannot directly account for differing excitation wavelengths, they cannot obtain excitation spectra.

Closely related work to our analysis lies in the field of chemometrics. A fluorescence unmixing in chemometrics typically involves measuring a fluorescence emission signal from a mixture of different fluorophores in a test tube, measured across changing excitation wavelengths, from multiple test tubes containing the same compounds in different ratios (Smilde et al., 2005). Analogously, our paper performs a similar experiment by considering every pixel of a hyperspectral image as a virtual test tube in a fluorescence experiment. However, in chemometric fluorescence unmixing, there is no notion of spatial distribution and chemometricians typically measure excitations and emissions at very fine spectral resolutions, which is not possible with fluorescence microscopy at this time. Finally, Neher et al. (2009) consider an extension of an NMF-based approach to non-negative tensor decomposition for a fluorescence microscopy case study. However, the derived updates are not guaranteed to converge (Lin, 2007) or maintain non-negativity, instead heuristically setting negative values to small positive values during the optimization.

1.2. Proposed approach

For our driving application, a tissue sample is illuminated with different excitation wavelengths and a hyperspectral image is captured for each excitation wavelength. As described above, multiple excitation wavelength hyperspectral fluorescence microscopy of a tissue sample has intuitive extensions to tensor approaches. Computing a non-negative tensor decomposition of multi-excitation hyperspectral images leads to recovery of the emission spectra and the spatial abundances of fluorophores within the tissue sample and to our knowledge, the first extraction of the excitation spectra of retinal fluorophores from hyperspectral images. We adopt a flexible alternating optimization approach towards computing the

Brightfield Image   Manual Segmentation Label   Fluorescence Image
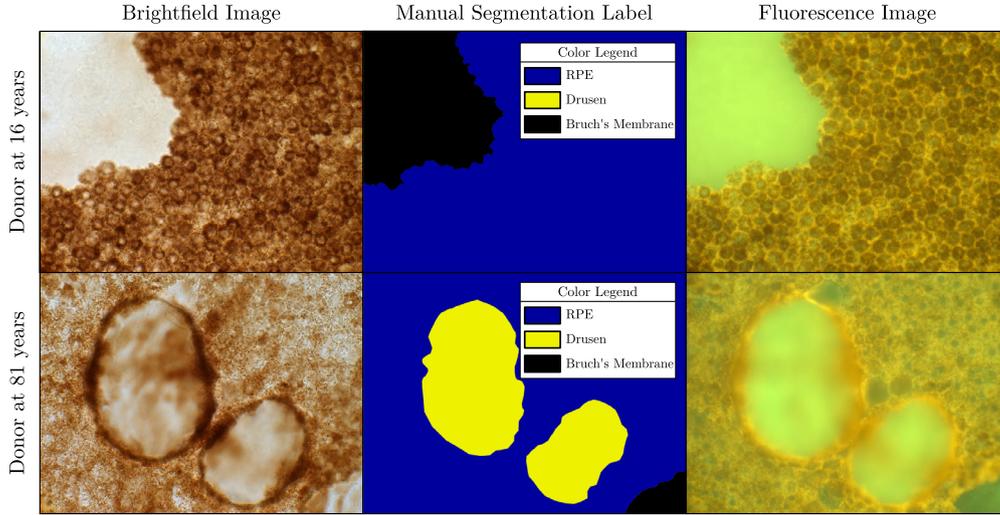
Fig. 1. A comparison of tissue samples from a young adult with a normal retina (top row) and an older adult with advanced AMD (bottom row). Shown are microscopic views of retinal tissue in the same orientation as it would be seen in clinical fundus autofluorescence imaging but at higher magnification. Tissue from the 16-year-old male (top row) shows the hexagonal cells of the RPE layer and in the upper left corner, Bruch's membrane bared by mechanical disruption of the RPE layer. Tissue from the 81-year-old male (bottom row) has two drusen (see description of drusen in Fig. 3). **Left column:** Images of ex-vivo RPE tissue at $40\times$ magnification. The distinct hexagonal structure of RPE cells in the top row is severely disrupted by age and disease in the bottom row, with the RPE cell layer disrupted by two large oval-shaped drusen deposits. **Middle Column:** Regions corresponding to different tissues labeled manually. The Bruch's membrane lies directly under the RPE and is present in the entire background of both samples. **Right Column:** RGB false color composites of the fluorescence emission of tissues shown in the left column excited with light at 436 nm wavelength. Drusen are hyperautofluorescent relative to their surroundings at this excitation wavelength as pixel brightness corresponds to fluorescence intensity. (Color figure online.)

non-negative tensor decomposition, allowing for various cost functions, constraints, and regularity parameters (Huang et al., 2016).

Further, multi-excitation fluorescence measurements are prone to missing values being present across the dataset, which we iteratively impute during the tensor decomposition (Tomasi and Bro, 2005). To reduce the ill-posedness of retrieving both factor matrices and imputing missing values simultaneously, we employ L_2 regularization. Additionally, while the uniqueness of tensor decompositions makes them attractive for data mining applications, determining the number of components for the decomposition and avoiding overfitting is not straightforward. We thus present thorough validation of our results using metrics from chemometrics on synthetic and real fluorescence experiments where the ground truth is known.

To demonstrate the utility of tensor decomposition-derived emission spectra for our application, we present a novel perspective on performing group testing on spectral images by considering the spectra as discrete realizations of functions of wavelength. This allows the use of two-sample non-parametric functional hypothesis tests to detect group differences. As a preliminary result, we find a statistically significant difference in emission spectra in a small dataset of AMD and control group donors, that merits further validation with a larger sample size. While chemometrics has recently considered the usage of functional representations of spectra (Saeys et al., 2008; Möller et al., 2016), to our knowledge, this is the first use of functional hypothesis tests to distinguish between groups of hyperspectral images.

2. Methods

2.1. Tensor algebra

2.1.1. Background

A tensor represents a multidimensional array, i.e. data with n indices for each entry forms an n-dimensional tensor. We borrow notation from Kolda and Bader (2009) with vectors, matrices and tensors being denoted as $\mathbf{x}$, $\mathbf{X}$, and $\mathcal{X}$, respectively. Vectors within tensors are referred to as *fibers*, i.e. a mode-i fiber is a vector in the tensor formed by fixing every index in the tensor except the ith index and letting the data vary along i. Re-arranging a tensor into a matrix is called matricizing. To matricize a tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \ldots \times I_n}$ in mode-i into $\mathbf{X}_{(i)}$, we take the mode-i fibers of $\mathcal{X}$ and concatenate them as columns of $\mathbf{X}_{(i)} \in \mathbb{R}^{I_i \times (I_1 \ldots I_{i-1} I_{i+1} \ldots I_n)}$ (Kolda, 2006).

If an n-dimensional tensor can be expressed as the outer product of one set of n vectors, it is a rank-1 tensor $\mathcal{X} = \mathbf{x}^{(1)} \circ \mathbf{x}^{(2)} \circ \ldots \circ \mathbf{x}^{(n)}$. The overall rank of a tensor is defined as the minimum number of rank-1 tensors that need to be summed in order to achieve perfect reconstruction (Hitchcock, 1927). A three-dimensional tensor $\mathcal{X}$ of size $I \times J \times K$ has rank-R if R is the smallest number of sets of vectors $\mathbf{a}_{1 \times i}$, $\mathbf{b}_{1 \times j}$ and $\mathbf{c}_{1 \times k}$ that reconstruct the tensor as,

$$x_{ijk} = \sum_{r=1}^{R} a_{ir} b_{jr} c_{kr} \text{ where } i \in \mathbb{R}^I; \ j \in \mathbb{R}^J; \ k \in \mathbb{R}^K.$$
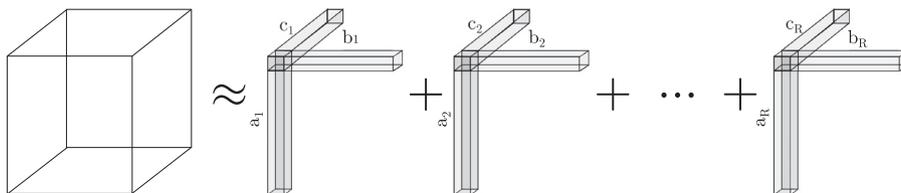
Fig. 2. Visualizing the rank-R Canonical Polyadic Decomposition/Parallel Factorization of a three-dimensional tensor. The summation over the outer product of the set of fibers, $\mathbf{a}_{i \times r}$, $\mathbf{b}_{i \times r}$, $\mathbf{c}_{i \times r}$ from 1 to R reconstructs a tensor of rank-R.

In terms of factor matrices, we can write the tensor decomposition of $\mathcal{X}$ as $\mathcal{X} = \sum_{p=1}^{R} \lambda_p \, \mathbf{a}_p \circ \mathbf{b}_p \circ \mathbf{c}_p$ where λ_p is a scalar absorbing all of the weights, and $\circ$ is the vector outer product. The tensor $\mathcal{X}$ can be matricized in 3 different modes as, $\mathbf{X}_{(1)} = \lambda \mathbf{A}(\mathbf{C} \odot \mathbf{B})^T$, $\mathbf{X}_{(2)} = \lambda \mathbf{B}(\mathbf{C} \odot \mathbf{A})^T$, and $\mathbf{X}_{(3)} = \lambda \mathbf{C}(\mathbf{A} \odot \mathbf{B})^T$ where λ is the vector containing all λ_p's, $\odot$ is the matrix Khatri-Rao product and $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are factor matrices containing all $\mathbf{a}_p$'s, $\mathbf{b}_p$'s and $\mathbf{c}_p$'s, respectively. [[**A, B, C**]] represents the tensor reconstructed by the factor matrices.

2.1.2. Non-negative tensor decomposition

To recover the individual rank-1 tensors, we use parallel factor analysis (PARAFAC) (Harshman, 1970), also referred to as Canonical Polyadic Decomposition (CPD). For the rest of this paper, unconstrained CPD/PARAFAC is referred to as CP and non-negative CPD/PARAFAC is referred to as NNCP. With real-world noisy data, we often solve for the factors iteratively in an alternating optimization (AO) manner.

We require all factors to be non-negative and wish to introduce regularization on the factor matrices to reduce ill-posedness. To this end, we adopt and extend the Alternating Optimization-Alternating Direction Method of Multipliers (AO-ADMM) method for constrained tensor factorization (Huang et al., 2016) for our problem. AO-ADMM is a block coordinate update-style algorithm. It alternates between solving the tensor factorization in each matricized form in its outer loop and in the inner loop solves each individual constrained optimization problem in an ADMM fashion (Boyd et al., 2011) which allows for the usage of multiple constraints. As the updates made in the constrained sub-problems are element-wise, we implement the algorithm on graphics processing units (GPUs) for a $2 \times$ improvement in speed on the specific problems considered in this paper.

When the tensor has missing values, there are several approaches in order to fit a CP model (Kroonenberg, 2007), most common of which are weighted regression with the weights being contained in an indicator tensor $\mathcal{W}$ with entries of 0 and 1 for missing and not-missing, i.e, fitting a model only to the available data. AO-ADMM can perform data imputation in its inner iterations, but in its original formulation cannot handle the case where entire fibers of the tensor are systematically missing (as in our application). To solve this, we implement an Expectation-Maximization method to impute the missing values at each iteration from the current estimate of the CP model (Kroonenberg, 2007; Smilde et al., 2005, Section 6.9). We choose the EM model for its straightforward integration into the outer loop of the AO-ADMM procedure used. Further, weighted regression and EM imputation have been proven to be equivalent in CP models (Kiers, 1997).

The entire procedure (including rank estimation in Section 2.2) is illustrated in Algorithm 1 . Regularization is written as $r(\cdot)$. Smoothness-promoting L_2 regularization is used here, although AO-ADMM can accommodate other regularizers as well. Consider the tensor $\mathcal{X}$ with missing values that we want to decompose, and the indicator tensor $\mathcal{W}$ from earlier. In the middle loop of Algorithm 1, the missing values of $\mathcal{X}$ are iteratively updated as $\mathcal{X} \rightarrow \mathcal{W} * \mathcal{X} + (1 - \mathcal{W}) * \hat{\mathcal{X}}$ where $\hat{\mathcal{X}}$ is the current estimate of the NNCP model and $*$ is the Hadamard (element-wise) product. As the NNCP model is initialized randomly in this paper, the initial estimates of the missing data are random values as well. In Section 4.2.1 we find that the iterative imputation method works well in recovering the ground truth factor matrices in our data.

2.2. Model selection

Despite the attractive properties of CP, determining the number of factors (rank) is an NP-hard problem (Håstad, 1990). Since

---

Algorithm 1: NNCP with unknown rank estimation and data imputation.

Input: $\mathcal{X} \in \mathbb{R}^{I \times J \times K}$.
Output: $\mathbf{A}, \mathbf{B}, \mathbf{C}$ and λ.

1 Create indicator tensor $\mathcal{W} \in \mathbb{R}^{I \times J \times K}$ with $w_{ijk} = 0$ if x_{ijk} is missing and $w_{ijk} = 1$ otherwise.
2 Initialize rank of decomposition as $g = 1$.
3 **repeat**
4   Initialize $\lambda \in \mathbb{R}^g$ and Factor Matrices $\mathbf{A} \in \mathbb{R}^{I \times g}, \mathbf{B} \in \mathbb{R}^{J \times g}, \mathbf{C} \in \mathbb{R}^{K \times g}$.
5   Impute initial estimate of missing values in $\mathcal{X}$ as $\mathcal{X} \rightarrow \mathcal{W} * \mathcal{X} + (1 - \mathcal{W}) * [[\mathbf{A}, \mathbf{B}, \mathbf{C}]]$ (all element-wise).
6   Define $\hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{C}}$ as $\hat{\mathbf{A}} = \mathbf{A} \times \mathrm{diag}(\lambda), \hat{\mathbf{B}} = \mathbf{B} \times \mathrm{diag}(\lambda), \hat{\mathbf{C}} = \mathbf{C} \times \mathrm{diag}(\lambda)$.
7   **repeat**
8   Fix B and C, solve $\|\mathbf{X}_{(1)} - \hat{\mathbf{A}}(\mathbf{C} \odot \mathbf{B})^T\|_F^2 + r(\hat{\mathbf{A}})$ for $\hat{\mathbf{A}}$ s.t. $\hat{\mathbf{A}} \geq 0$.
9         $\mathcal{X} \rightarrow \mathcal{W} * \mathcal{X} + (1 - \mathcal{W}) * [[\mathbf{A}, \mathbf{B}, \mathbf{C}]]$.
10   Fix A and C, solve $\|\mathbf{X}_{(2)} - \hat{\mathbf{B}}(\mathbf{C} \odot \mathbf{A})^T\|_F^2 + r(\hat{\mathbf{B}})$ for $\hat{\mathbf{B}}$ s.t. $\hat{\mathbf{B}} \geq 0$.
11         $\mathcal{X} \rightarrow \mathcal{W} * \mathcal{X} + (1 - \mathcal{W}) * [[\mathbf{A}, \mathbf{B}, \mathbf{C}]]$.
12   Fix A and B, solve $\|\mathbf{X}_{(3)} - \hat{\mathbf{C}}(\mathbf{B} \odot \mathbf{A})^T\|_F^2 + r(\hat{\mathbf{C}})$ for $\hat{\mathbf{C}}$ s.t. $\hat{\mathbf{C}} \geq 0$.
13         $\mathcal{X} \rightarrow \mathcal{W} * \mathcal{X} + (1 - \mathcal{W}) * [[\mathbf{A}, \mathbf{B}, \mathbf{C}]]$.
14   **until** *convergence.*
15   **return** $\lambda, \mathbf{A}, \mathbf{B}, \mathbf{C}$.
16   Compute CORCONDIA between $\mathcal{X}$ and $[[\mathbf{A}, \mathbf{B}, \mathbf{C}]]$.
17     $g = g + 1$.
18 **until** *CORCONDIA decrease exceeds threshold.*
19 **return** $\lambda, \mathbf{A}, \mathbf{B}, \mathbf{C}$.

---

prior knowledge of the number of spectral sources in autofluorescent tissue is not available, an appropriate validation metric is required to verify the number of components. In chemometrics, the core consistency diagnostic (CORCONDIA) metric is often used to decide on an appropriate rank (Bro and Kiers, 2003) by exploiting the relationship between CP and the Tucker3 model (Tucker, 1964).

Briefly, the Tucker3 model decomposes a tensor $\mathcal{X} \in \mathbb{R}^{I \times J \times K}$ into a smaller 'core' tensor $\mathcal{G}$ multiplied by a matrix along each mode as shown below,

$$\mathcal{X} = \sum_{p=1}^{P} \sum_{q=1}^{Q} \sum_{r=1}^{R} g_{pqr} \mathbf{a}_p \circ \mathbf{b}_q \circ \mathbf{c}_r,$$

where $\mathbf{A} \in \mathbb{R}^{I \times P}$, $\mathbf{B} \in \mathbb{R}^{J \times Q}$, $\mathbf{C} \in \mathbb{R}^{K \times R}$ are factor matrices along each mode, and $\mathcal{G} \in \mathbb{R}^{P \times Q \times R}$ is the core tensor whose entries show the level of interaction between different factors. Several different CP models with increasing numbers of components may be appropriate and the model with the highest rank which is still appropriate is considered to be the most reasonable choice. For the CP model to be valid, the core tensor $\mathcal{G}$ should be an identity tensor when the Tucker3 model is fit with CP components.

Let $\mathcal{T}$ be an $F \times F \times F$ identity tensor. To assess the similarity between $\mathcal{G}$ and $\mathcal{T}$, CORCONDIA is defined as,

$$\mathrm{CORCONDIA} = 100 \left(1 - \frac{\sum_{d=1}^{F} \sum_{e=1}^{F} \sum_{f=1}^{F} (g_{def} - t_{def})^2}{F} \right),$$

where g_{def} and t_{def} are the *def*-th elements of $\mathcal{G}$ and $\mathcal{T}$, respectively. CORCONDIA close to 100% implies an appropriate model, with closeness to 100% understood relative to the changes compared with models of lower rank. Core consistency decreases gradually with increasing rank, until the maximum rank is exceeded

and the decrease becomes dramatic. It is at this inflection point that the rank should be chosen.

2.3. Group testing with spectral images

As shown in Section 4, NNCP can disentangle mixed emission and excitation spectra. However, the multiple resulting spectra specific to each donor can be difficult to interpret, especially across groups (AMD and non-AMD in this study). This highlights a need for a principled framework for detecting group differences in the results we obtain from NNCP. Parametric multivariate tests can be applied, but they would assume normality, treat adjacent channels as individual features, and would not take into account their sequential nature. Further, the number of features would be larger than the sample size, precluding their straightforward application. As the resulting emission spectra are discrete representations of continuous curves indexed by wavelength, they can be considered infinite-dimensional functional data (Ramsay and Silverman, 2007). Functional data analysis has a rich history in fiber tract-based statistics in Diffusion Tensor Imaging (Geng et al., 2012; Zhu et al., 2011). Recently, Pomann et al. (2016) proposed a two-sample non-parametric test for fiber tract data which we adopt for our problem and briefly describe below.

After NNCP, each donor has a set of emission spectra $X_i \in \mathbb{R}^m$ where m is the number of image channels and $i = 1, 2, \ldots, k$ for a rank-k decomposition with individual vector corresponding to individual fluorophores. The null hypothesis is that the curves from the two groups are equal in distribution,

$$H_0 : X^1(\cdot) \overset{d}{=} X^2(\cdot),$$

where the superscript denotes group membership. The spectra corresponding to the same fluorophore across m donors are concatenated into a matrix $Y_k \in \mathbb{R}^{m \times n}$. Repeating this across k fluorophores creates k matrices of size $m \times n$. Computing the functional principal components analysis (Hall et al., 2006) of each matrix Y_k retrieves a set of functional basis vectors and projections ξ, with the number of bases U set by the fraction of variance explained (0.95 in this study). The null hypothesis above then becomes that the U projections of the curves from the two groups on the eigenbasis are equal in distribution,

$$H_0^U : \{\xi_u^1\}_{u=1}^U \overset{d}{=} \{\xi_u^2\}_{u=1}^U,$$

where 1 and 2 in the superscript of ξ indicate group membership. Each of the U hypotheses is tested by the two-sample non-parametric univariate Anderson-Darling test (Scholz and Stephens, 1987), which is similar to but more powerful than the Kolmogorov-Smirnov test (Stephens, 1974). Since there are U comparisons corresponding to U hypotheses, we perform a Bonferroni multiple comparisons adjustment. Combining all of the above, the original null hypothesis is rejected if

$$\min_{1 \le u \le U} p_u \le (\alpha/U),$$

where p_u is the u-th p-value returned by the multiple Anderson-Darling tests and α is the chosen level of significance.

3. Materials

3.1. Data acquisition

Hyperspectral autofluorescence microscopy was performed ex-vivo on human donor retinal tissue. Twenty μm thick flatmounts of Retinal Pigment Epithelium (RPE) cells attached to the underlying Bruch's Membrane (BrM) close to the macula are imaged at $40\times$ magnification using an upright microscope with a hyperspectral camera (Nuance, Perkin-Elmer). The dataset includes ten donor
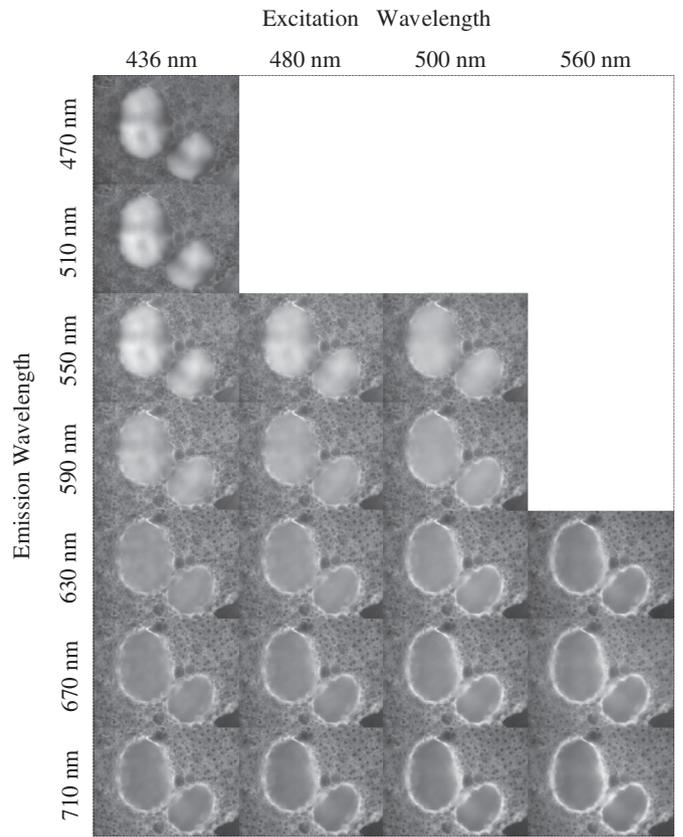
Fig. 3. Normalized hyperspectral images on a grid of emission vs. excitation wavelengths from a donor with AMD. Each column shows the sample illuminated with a different excitation wavelength, with the rows indicating the corresponding emission. The blank spaces in the figure are unobservable emission channels. 7 out of 26 emission channels are displayed due to space constraints. Images show two drusen, the characteristic lipid-rich extracellular dome-shaped deposit of AMD, surrounded by retinal pigment epithelium (RPE), a layer of support cells for the light-sensing photoreceptors. In the lower right corner of each panel the RPE is absent, revealing Bruch's membrane, an extracellular matrix and vessel wall that underlies the RPE. At shorter wavelength emissions, the druse contents and Bruch's membrane are visible. At higher emission wavelengths, the RPE layer is not visible on the apex of the druse dome (which comes up out of the plane of the tissue). Instead the RPE is thickened around the drusen bases, forming two haloes.

samples with non-neovascular dry AMD (mean age 86.20 years with 3.46 years standard deviation) and five age-matched control donors (mean age 84.40 years with 2.19 years standard deviation). Four filter cubes corresponding to 436, 480, 500, and 560 nm excitation wavelengths were used to filter the excitation light source from the microscope. The hyperspectral camera was used to capture autofluorescence emission spectra from these tissues evenly at 10 nm intervals from 420 nm to 720 nm for each of the excitation wavelengths, leading to 31 channels per excitation wavelength. For further details on acquiring this dataset, see Tong et al. (2016). See Gao and Smith (2015) for image acquisition methodology and Ach et al. (2014) for details on tissue preparation.

3.2. Data pre-processing

The hyperspectral images are pre-processed to remove variation due to intrinsic camera properties, experimental design, and acquisition parameters. The image stacks are first recalibrated by the efficiency of the corresponding emission filter cube and then by the quantum efficiency of the camera detector. Additionally, the fluorescence emission spectrum is positioned along the wavelength axis at longer wavelengths than the excitation wavelengths. Therefore, while the camera captures 31 channels of emission, only

channels with wavelengths longer than the excitation wavelength can be measured, making the shorter wavelength channels missing values in the statistical sense. Further, emission filters which filter light from the sample before it reaches the detector attenuate several channels longer than the excitation wavelength. Therefore, the channels of the image stack at wavelengths shorter than the cutoff of the emission filter used are also modeled as missing values.

The intensity value of a pixel in a hyperspectral image is a total photon count. To have well-exposed images, the exposure time of the camera has to vary with the intensity of tissue autofluorescence. The shorter the excitation wavelength, the shorter the exposure time needed to collect an adequate number of photons. The longer the excitation wavelength, the longer the exposure time. For example, in our study, an image stack corresponding to a 436 nm excitation wavelength may be exposed for 30–50 ms, whereas a 560 nm excitation wavelength image stack may be exposed for 200–300 ms. Therefore, the image stacks are scaled by their exposure time in milliseconds with the intensity of a pixel representing photons detected per millisecond.

3.3. Data arrangement and interpretation

A hyperspectral image forms a three-dimensional cube of data with two spatial dimensions and one spectral dimension. A single hyperspectral image can be converted to a matrix via mode-3 matricization. The columns of this matrix correspond to the emission spectra of the pixels in the hyperspectral image. To create a tensor, we take hyperspectral images of the same tissue sample under varying excitation wavelengths. Hyperspectral images are captured at four excitation wavelengths and undergo mode-3 matricization and are inserted into a tensor as a frontal slice arranged in ascending order of excitation wavelength. In this arrangement, the magnitude scalings of the emission corresponding to a pixel, i.e. the excitation spectrum of a pixel is captured in a mode-3 fiber.

We capture 31 channels of image size 1040×1392 pixels over the visible spectrum, forming a $1040 \times 1392 \times 31$ hyperspectral image. After mode-3 matricization, an image channel becomes a 1,447,680-length vector and the hyperspectral image forms a matrix of size $31 \times 1,447,680$. After capturing four hyperspectral images for each tissue sample, we insert the corresponding matrices as the frontal slices of a tensor of size $31 \times 1,447,680 \times 4$. The first five emission channels are discarded as there are no observed values in them with any of the excitation wavelengths. A series of hyperspectral images across excitation wavelength channels now forms a three-dimensional tensor $\mathcal{X} \in \mathbb{R}^{I \times J \times K}$, where the number of emission wavelength channels $I = 26$, the number of pixels $J = 1,447,680$, and the number of excitation wavelengths $K = 4$. $\mathcal{X}$ contains 24.04% missing values because of the cut-off of the emission filter as explained in Section 3.2. Low-rank decomposition using NNCP on this tensor provides mode-1 factors representing the emission spectra of fluorophores and mode-3 factors representing their excitation spectra, while mode-2 factors represent their spatial abundances.

3.4. Brief overview of AMD and fluorescence microscopy

In this section, we introduce ophthalmological and chemical concepts required to interpret the results shown in Section 4. Over its lifespan, a single RPE cell ingests hundreds of thousands of photoreceptor outer segment tips and these segment tips ultimately give rise to fluorescent lysosome-related intracellular organelles called Lipofuscin (LF) and Melanolipofuscin (ML) (Kennedy et al., 1995). LF/ML accumulation within the RPE cells is strongly correlated to aging within the eye (Sparrow and Boulton, 2005). RPE autofluorescence is largely due to LF/ML accumulation. The RPE layer lies upon the weakly autofluorescent Bruch's membrane

(BrM). LF/ML can be thought of as RPE inclusions and drusen/BrM can be thought of as Sub-RPE structures. To study AMD mechanisms, finding age or disease-related changes in the fluorescence spectra of RPE and adjacent tissues is of importance.

The term fluorophore generally refers to a single fluorescent compound. However, as spectral ground truth is not available in this study, a spectrum recovered could be the spectrum of a single fluorophore or a family of multiple fluorescent compounds which are closely related in a spectral and spatial sense. Therefore, we use the term fluorophore interchangeably between a single fluorophore and a family of closely related fluorophores. Fluorescence signals are strongly sensitive to fluorophore environment (Lakowicz, 2006b). Therefore, how the tissue samples were prepared, the retinal location that they were harvested from, and age can all affect the spectral and quantitative characteristics of the data.

In fluorescence spectroscopy, the same fluorescence emission spectrum is generally observed irrespective of the excitation wavelength (Kasha, 1950) and the intensities of fluorescence emission spectra are directly proportional to the concentration of the respective fluorescing compounds in the test tube. In our formulation, we retrieve a weight vector λ which contains the weights of the individual components and can be thought of as an analogue for the amount of the respective fluorophores in the donor sample.

4. Experiments and results

4.1. Tissue experiments

In this section, quantitative and qualitative results are presented with the data acquired, pre-processed and arranged as per Sections 3.1–3.3 respectively. An appropriate rank of factorization was determined, NNCP was applied, and the two groups of AMD and non-AMD tissues underwent statistical testing to determine differences in emission spectra. The experiments were implemented in MATLAB R2018a using a Linux workstation with an Intel i7-5930K CPU and 32GB RAM and a server virtual machine with an NVIDIA Tesla P100 for GPU accelerated tensor decomposition. Research toolboxes heavily utilized included the Tensor Toolbox v2.6 (Bader and Kolda, 2015), the N-Way Toolbox (Andersson and Bro, 2000), and the PACE toolbox for functional data analysis (Yao et al., 2003). All initializations for NNCP procedures were drawn uniformly at random with random number generator states set based on system time. All displayed spectra are normalized to unit L_2 norm, unless stated otherwise.

4.1.1. Rank selection

The difficulty in choosing an appropriate rank was detailed in Section 2.2. Performing a core consistency analysis of the NNCP model on this dataset finds that two components appropriately model the data as shown in Fig. 4 and ensures that the factors are not too closely correlated with each other. The CORCONDIA results shown in Fig. 4 were obtained on spatially downsampled images due to computation time. This approximation was verified to give similar CORCONDIA results to three test images analyzed at full resolution. For consensus, Bayesian rank estimation techniques for NNCP (Mørup and Hansen, 2009) and non-subjective versions of CORCONDIA which account for noisy data and reconstruction error (Liu et al., 2013b) determine the same rank as well.

4.1.2. NNCP decomposition

Rank-2 NNCP was performed on both groups with L_2 regularization applied on emission and excitation spectra. No regularization was applied on abundance images as they may not have smooth or sparse characteristics, although this can be applied if required. The algorithm was allowed to perform a maximum of 30,000 outer
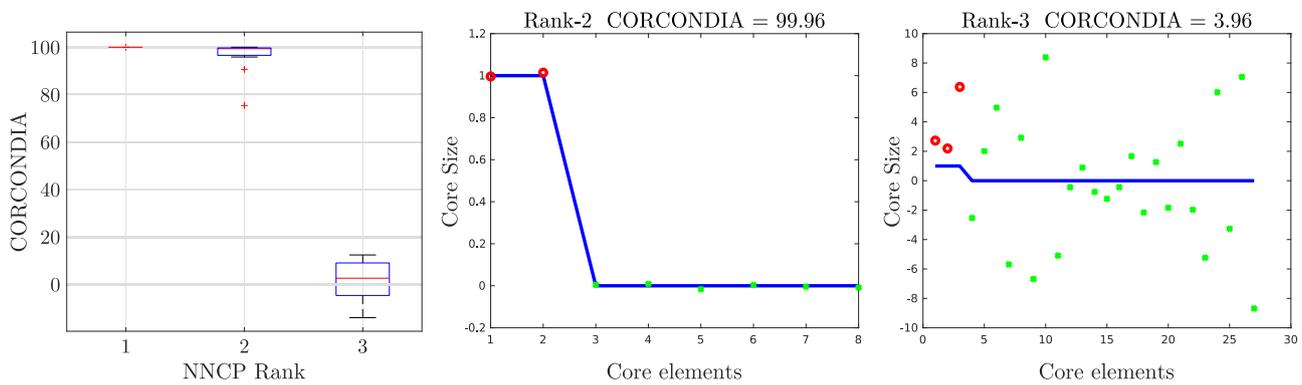
Fig. 4. Left: Boxplot of CORCONDIA values on the entire dataset for NNCP using rank-1 to rank-3 factorizations. The sharp inflection after rank-2 suggests that NNCP is over-fitting the data and that two components are appropriate. **Center and Right:** CORCONDIA plots of rank-2 and rank-3 NNCP models applied to a specific image. The red dots represent the values along the super-diagonal of the core tensor which should be close to one and the green dots represent the off-diagonal values which should be as close to zero as possible. The blue line represents ideal values. Increasing rank from two to three shows abrupt undesirable changes in core values. (For interpretation of the references to color in this figure, the reader is referred to the online version of this article.)

iterations with a convergence tolerance of 10^{-6}, each initially performing 30 inner ADMM iterations per sub-problem with a tolerance of 10^{-3}, eventually increasing to 40 inner ADMM iterations with tolerance 10^{-4} after 4000 outer iterations. Qualitatively, the algorithm was found to give similar results with fewer outer iterations (4,000).

The decomposition finds two major components, which can be attributed to RPE and Sub-RPE structures via their spatial distributions. Fig. 5 shows results for four out of fifteen donor samples. The brightfield and fluorescence images show strong heterogeneity of tissue morphology, whereas the spectra recovered are stable across donors. Subtle changes in the emission spectra between donor samples are observed. For example, the sub-RPE emission spectrum (red curve) in Column 3 of Fig. 5 shows a noticeable blue-shift (peak shifting towards the left) in comparison to other donor samples. Large variations are noticed in the fluorescence intensity, which is indicative of the amount of fluorophores being different across the tissue samples, as well as tissue preparation and experimental conditions. The excitation spectra are found to be largely consistent between different donors. However, there are only four points available (corresponding to four excitation wavelengths) and there may be variation in between these points.

4.1.3. Group testing

On retrieving the emission spectra for all fifteen donors, they are assigned to their respective groups (AMD or non-AMD) as per Section 3.1 and made to carry the weights of the NNCP decomposition. Functional hypothesis testing is performed using the methodology presented in Section 2.3. The functional bases are determined with functional principal components analysis with a Gaussian kernel, with most other parameters chosen via cross-validation in PACE with random effects computed using numerical integration (Yao et al., 2003). The number of bases was chosen based on the fraction of variance explained, chosen to be 0.95 in this study, finding two bases to be appropriate. Qualitatively, a higher fraction was found to return noise-like bases. A threshold for significant differences between groups was chosen to be a 5% chance of a Type-I error. A two-sample univariate non-parametric Anderson Darling test is applied to each set of projections. As two comparisons are made corresponding to two sets of bases and projections, a Bonferroni correction is applied which reduces the threshold for significance to $p = .025$.

The bases are shown in Fig. 6. We find statistically significant differences for both RPE ($p = 0.0233 < 0.025$) and Sub-RPE ($p = 0.0237 < 0.025$) emission spectra in the near periphery (just outside the macula). It is important to contextualize these results and

state limitations. First, the sample size ($n = 15$) is small and outliers can cause Type I errors. Second, donors within the AMD group are at various stages of the disease which could affect these results. However, it should be highlighted that non-parametric tests have less statistical power than parametric tests, and especially so with small sample sizes. Further, the Bonferroni correction is a highly conservative adjustment. Additional validation will be performed in the future with a larger sample size and in tissues with better preserved central macula, where differences in sub-RPE structures between cases and controls are expected to be higher (Curcio et al., 2013).

4.2. Validation

Here we test the ability of NNCP to distinguish between fluorophores spectrally and spatially with known ground truth available for validation. Section 4.2.1 presents fluorescence reference standards captured under the same imaging framework. Additionally, it tests the applicability of additional constraints often employed in geospatial image analysis. Section 4.2.2 presents synthetic data and tests the advantages of the multi-linear approach over NMF-style approaches.

4.2.1. Real data: fluorescent beads

Fluorescent beads (also known as microspheres) are used to calibrate fluorescence microscopes as they have known emission and excitation spectra provided by the manufacturers. They are localized to specific regions on the imaged slide, which is ideal for testing both spectral and spatial capabilities of the NNCP framework. Four hyperspectral images of fluorescent microspheres (Thermo Fisher Scientific) were captured using excitation wavelengths of 420, 450, 480, and 500 nm using the same imaging system used to acquire the tissue images. To account for instrumental variation, emission and excitation spectra were physically measured at those wavelengths independently prior to unmixing and are treated as ground truth shown in Fig. 7A. The two fluorophores in the image are highly similar in both emission spectra shape and peak wavelength. However, bead 2 is significantly dimmer than bead 1. Additionally, the blank regions of the image have an intrinsic noise spectrum induced by the physics of the camera, referred to here as the background spectrum. These details combined create a challenging blind source separation problem.

The unmixing results for beads are shown in Fig. 7B1. There is a very close correspondence between the NNCP recovered emission and excitation spectra and the measured ground truth. Spatially, NNCP is able to localize the fluorophores correctly.
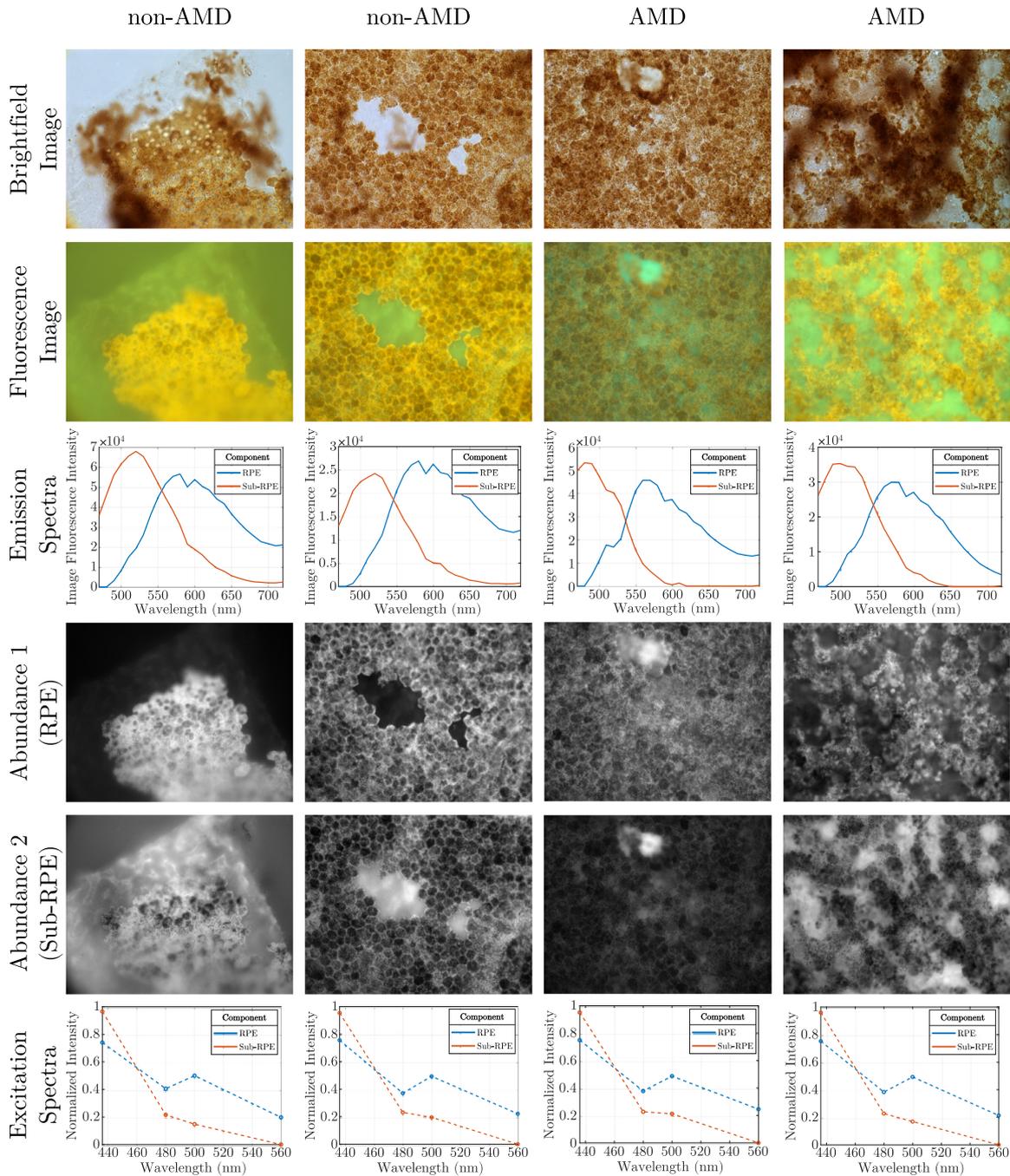
Fig. 5. A collage of four tissue samples from donors with AMD and controls, and their corresponding NNCP decomposition. Readers are encouraged to zoom in for tissue details. In Row 1, by transmitted light, the RPE layer appears as a continuous sheet of hexagonal cells with brown pigment. In the non-AMD samples, the absence of RPE represents mechanical damage that reveals Bruch's membrane. In the AMD samples are small drusen (third column) and ill-defined deposits overlaying incompletely removed pigmented tissue (out of focus brown material; fourth column). In Row 2, the false color microscopic images were taken at an excitation wavelength of 436 nm. The yellow-brown hexagonal cells are the RPE layer. Green regions are underlying Sub-RPE tissues. In Row 3, the emission spectra are multiplied by the corresponding weights in the decomposition to give a relative fluorescence intensity value of that signal across the entire image. The excitation spectra are linearly interpolated between the four available points for visualization purposes. In Rows 4 and 5, the abundance images give the spatial localization of RPE and Sub-RPE components, respectively. For example, in column 1 it can be clearly seen that the RPE layer overlays Sub-RPE tissue. In columns 2 and 4, sub-RPE autofluorescence is less visible in RPE regions. In column 3 a bright green druse is hyper autofluorescent compared to other sub-RPE regions. (For interpretation of the references to color in this figure, the reader is referred to the online version of this article.)

Interestingly, in Abundance 3 NNCP localizes the background underneath bead 2 as well indicating that due to its low fluorescence, the spectrum of bead 2 and the background are comparable in intensity. A minor limitation is noticed in our choice of the CORCONDIA metric which is 58.7040 for the rank-3 decomposition here. CORCONDIA values in this range are acceptable in practice (Smilde et al., 2005), However, it should ideally be

closer to 100, which highlights that the CORCONDIA heuristic is conservative for very correlated fluorophores such as bead 1 and bead 2.

In contrast to typical approaches in geospatial spectral unmixing, an additive sum to one constraint on the abundance images is not imposed in this work as the physics of multi-layered fluorescence in tissue are unclear, and ground truth does not
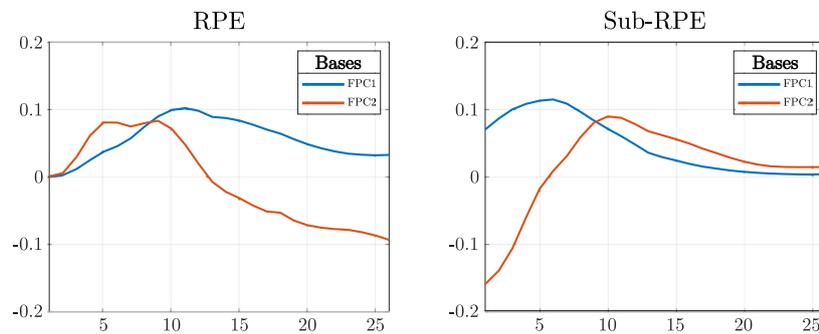
Fig. 6. Eigenbases/functional principal components (FPCs) of the NNCP derived emission spectra corresponding to RPE (left) and Sub-RPE (right) components from the entire dataset. Group testing is performed on the projections of the emission spectra onto these bases. (Color figure online.)

currently exist for these tissue spectra. Thus, reinforcing this constraint might cause approximation errors. To explore this, we perform a preliminary analysis with an additional sum to one constraint on the abundance images. Simultaneous non-negativity and sum to one constraints can be efficiently implemented in the AO-ADMM framework by replacing the abundance non-negative least squares sub-problem with a simplex projection introduced in Duchi et al. (2008), Wang and Carreira-Perpinán (2013).

The sum to one results are shown in Fig. 7B2. Qualitatively, we find that using the sum to one constraint gives the abundance images stronger interpretability, with the background and beads correctly localized. However, significant distortions are present in the bead 2 emission spectra shown in the red highlighted region. As we are most interested in the emission characteristics of fluorophores, limiting constraints to non-negativity only in this kind of data may be appropriate. Quantitatively, the CORCONDIA measure of trilinearity reduces from 58.7040 to 57.5242 with the addition of the sum to one constraint as well.

4.2.2. Synthetic data

To verify whether the correct factors and rank are returned by NNCP, we can generate synthetic data with the required additive multi-excitation nature of the hyperspectral images of autofluorescent tissue. Also, to evaluate the advantages of using multiple excitation wavelengths to recover emission spectra in an NNCP formulation instead of decomposing each image stack individually, we compare the results of NMF and NNCP. In order to let NMF and NNCP work with the same data, no data is assumed to be cutoff by the simulated excitation, i.e. there is no missing data.

To generate the tensor, 5 excitation wavelengths and 31 emission channels of size 500×500 are chosen. To simulate emission spectra, three Gaussian probability density functions are used to represent emission spectra of three different phantom substances. A hyperspectral image is then constructed with a spatial overlap of the three substances, with one substance placed on the left, one on the right and one present throughout the entire image causing each pixel to be a mixture of two to three substances. To simulate concentration differences of substances within each region, the signals are scaled spatially. The images are corrupted using both Poisson and Gaussian noise typical to fluorescence microscopy (Waters, 2009). Five hyperspectral images are then constructed by using the excitation spectra shown in Fig. 8 to scale the spectra of each component, with some components becoming brighter and some becoming dimmer with higher excitation wavelength. The five synthetic hyperspectral images are arranged as per Section 3.3, i.e. into a tensor of size $31 \times 250,000 \times 5$.

The results of NMF and NNCP are shown in Fig. 9. The NNCP recovered emission and excitation spectra are in strong correspondence with ground truth, with abundance images showing clear spatial separation. We vary the rank of the NNCP decom-

position of synthetic images and achieve CORCONDIA values of $100, 99.9763, 99.5701$, and -1.4732 for rank 1 to 4. From these values, rank-3 is appropriate, which is consistent with ground truth and thus supports the usage of CORCONDIA to decide on a real-world rank. The NMF results were obtained using an alternating minimization strategy (Kim and Park, 2008) and are sub-optimal compared to NNCP in both emission spectral recovery and the separability and noise-robustness of the abundance images. Additionally, by definition, NMF does not return excitation spectra for the three substances. The results from the original multiplicative algorithm for NMF (Lee and Seung, 1999) are highly noisy.

5. Discussion

5.1. Improvements over previous work

Tensor decompositions are unique under mild conditions, therefore the NNCP approach guarantees stable recovery of fluorophore signatures in a joint decomposition. In contrast, previous related work uses NMF-style approaches extensively which are in general non-unique. This leads to NMF approaches having several feasible solutions, which may make them unreliable for performing group hypothesis testing. Within the framework described, we notice subtle shifts in emission spectra across groups, which may be relevant to AMD biochemistry. Without a tensor approach, previous work mentioned in Section 1.1 utilized multi-excitation wavelength image stacks by concatenating the image stacks. Alternatively, they could also perform independent NMF analysis for each image stack. However, both of these approaches do not take the relationships between the image stacks into account when retrieving spectral factors and can lead to overfitting.

Further, the tensor representation gives us the excitation spectra of these fluorophores as well, which is not possible with matrix-based approaches. This excitation information can be used clinically to determine an appropriate wavelength of light used for illuminating the retina in clinical imaging. Finally, to our knowledge, this is the first work to perform functional group statistical testing on hyperspectral images. This novel perspective allows the testing of clinically relevant hypotheses, such as whether autofluorescence properties change with disease.

5.2. Limitations and future work

The excitation wavelengths used to generate the dataset are unequally spaced on the wavelength axis, however, we implicitly assume equal separation. We could use tensor completion (Liu et al., 2013a; Zhang et al., 2014) from the sparse samples available and re-sample evenly. However, tensor completion requires the data to be missing randomly which is not the case here and cannot be reliably applied. In this study, we assume that there are no
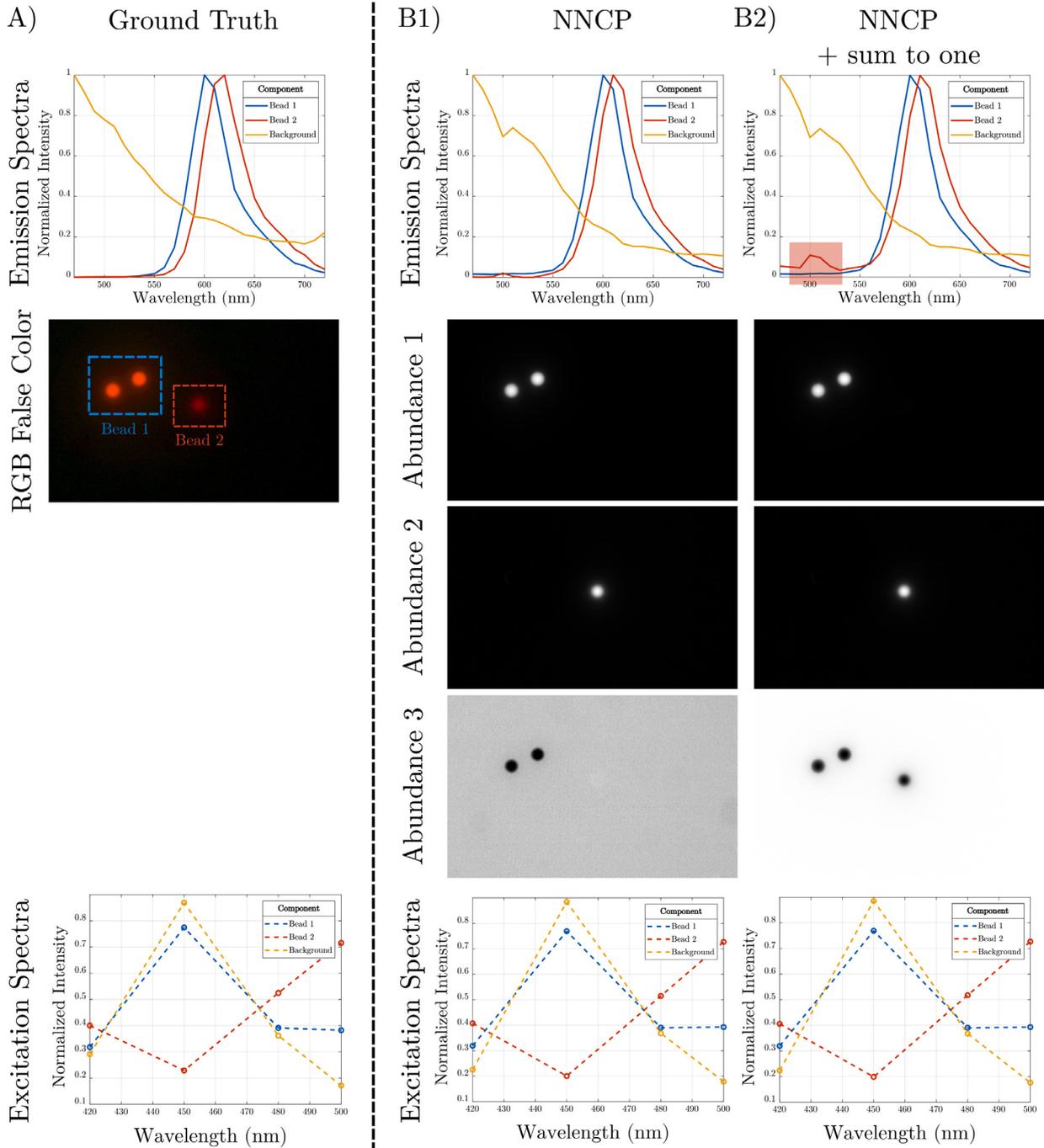
Fig. 7. NNCP analysis of fluorescent beads with known spectral properties on the same microscope used for tissue images. **Column A.** Ground truth emission and excitation spectra for the three beads shown in the false-color image. Here, emission spectra are normalized with respect to the maximum intensity. The two beads shown in the blue bounding box are the same fluorophore, the bead in the red bounding box is minutely shifted spectrally and dimmer in intensity. The dark regions possess a spectrum referred to here as the background. **Column B.** Results of NNCP decomposition both with (**B2**) and without (**B1**) a sum to one constraint. Both approaches adequately recover spectral and spatial factors corresponding to the ground truth. Enforcing the sum to one constraint causes distortions (highlighted in red bounding box) in the recovered emission spectra. (For interpretation of the references to color in this figure, the reader is referred to the online version of this article.)

non-linearities and the data are well described by a multi-linear model. Non-linear models have been extensively explored in geospatial spectral unmixing (Heylen et al., 2014). However, the underlying physics of reflectance imaging in remote sensing is different from that of fluorescence microscopy and many assumptions do not translate. Ikoma et al. (2014) create NMF and NNCP model that handles attenuation-based nonlinearities in hyperspectral images. However, their method requires known reference spectra which are currently unavailable in our application. Future investi-

gations should incorporate the physics of fluorescence microscopy into a blind non-multilinear NNCP model or be purely data-driven.

Future work will also explore incorporating domain knowledge into the decomposition. For example, semi-supervised NMF methods can incorporate domain knowledge into the matrix factorization by leveraging a small set of labeled observations (Lee et al., 2009). For this dataset, we find two components to be appropriate based on data-driven analysis. However, a semi-supervised NNCP approach may allow the user to define rank in the case
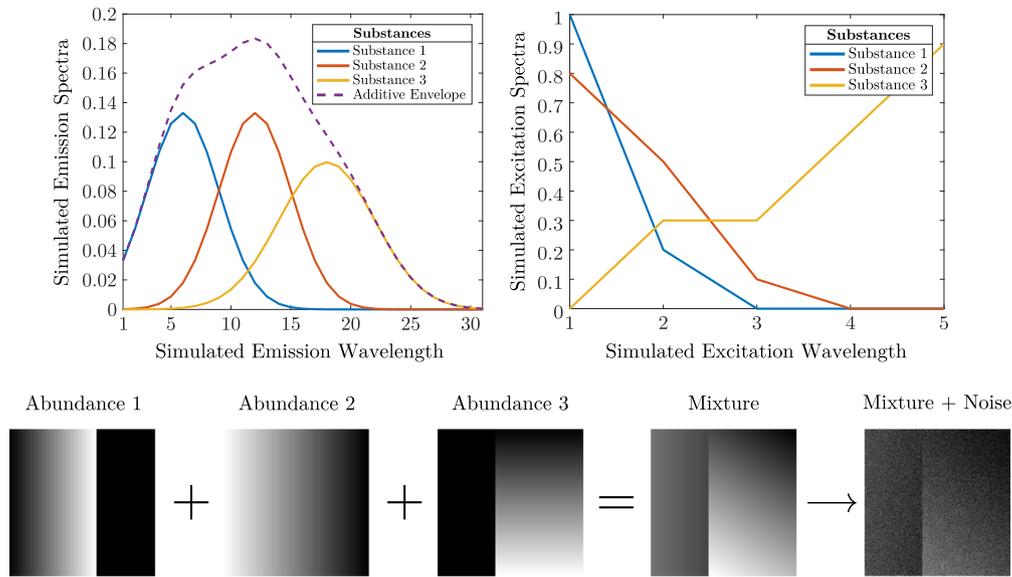
Fig. 8. Synthetic spectra and their spatial distributions in a hyperspectral image used to generate five hyperspectral images. Three Gaussian probability density functions of mean 6, 18 and 24 with sigma 3, 3 and 4 were chosen to simulate emission spectra, while excitation spectra were chosen in order to simulate phantom fluorescent substances either becoming more or less intensely fluorescent with increasing excitation wavelength. The normalized 16th emission channel of the second hyperspectral image is displayed as the abundance images. All five hyperspectral images are corrupted with both Poisson and Gaussian noise. (Color figure online.)
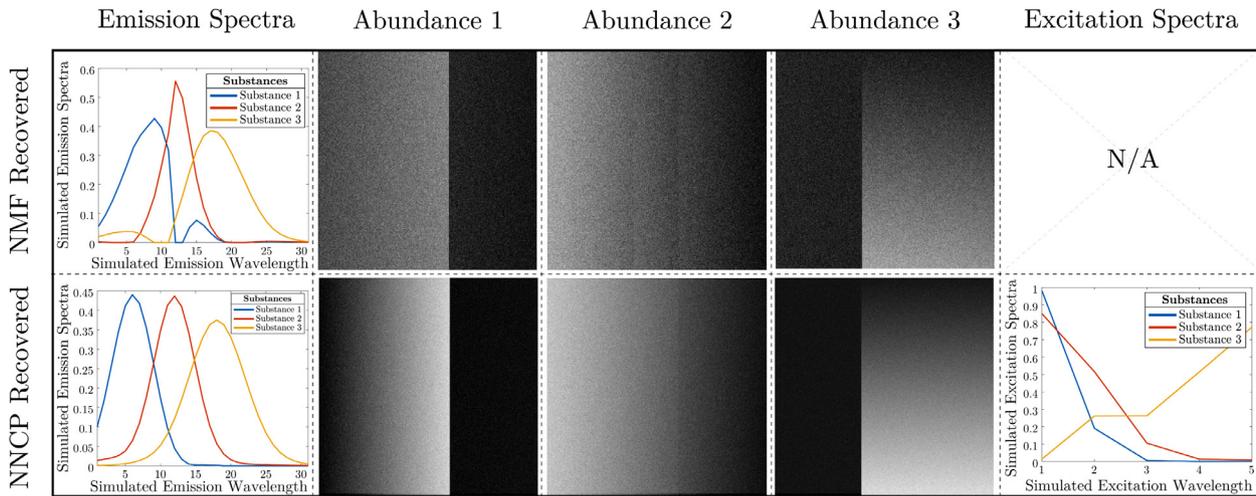
Fig. 9. Abundance images, emission and excitation spectra recovered using NMF and NNCP. NMF returned emission spectra suffer from poor reproducibility and are not able to adequately recover the ground truth. NNCP achieves a closer approximation to the ground truth and also returns the excitation spectra. (Color figure online.)

where strong domain knowledge is available and guide the decomposition.

Controlling for environmental variation in the image acquisition is of importance as a change in environment affects the NNCP spectral unmixing. Additionally, some natural fluorophores have emission spectra that vary with excitation wavelength not just in magnitude, but also in shape (Lakowicz, 2006a). The proposed unmixing and statistical framework needs to be applied to a larger sample size. This would allow the quantification of spectral changes and intensity changes and would allow us to build statistical models based on other possible covariates such as stage of disease, gender, and age. Ultimately, autofluorescence spectral signatures should be utilized concurrently with other methodologies such as imaging mass spectrometry (Norris and Caprioli, 2013; Crouch et al., 2015), resonance Raman tissue imaging (Bhosale et al., 2009), and fluorescence lifetime microscopy (Suhling et al., 2015) to form a consensus result for precise molecular identification.

5.3. Conclusion

In this paper, we present a novel framework for the analysis of high-dimensional hyperspectral imaging of tissue autofluorescence including the arrangement of images into a low rank tensor, determining the number of autofluorescent substances and recovering their autofluorescence spectra and spatial distribution in the presence of missing data and to our knowledge, the first application of functional hypothesis tests to hyperspectral images. This includes the novel recovery of excitation spectra from hyperspectral fluorescence microscopy images. From a clinical perspective, AMD is notoriously difficult to detect early and treat in time. The extraction of fluorescence spectra that are most associated with its progression can lead to future study of the biochemistry of the disease possibly leading to the development of new drugs and therapies. Better understanding of autofluorescence sources will also improve the interpretation of clinical fundus autofluorescence imaging, which is commonly used in the diagnosis and management of AMD.

We simultaneously estimate spectral and spatial characteristics of autofluorescent tissue within the eye. Further, we present a statistical framework for group testing between AMD and control donors which accounts for the functional nature of spectra. However, the methodology developed here is general and can be applied to the spectral imaging of any pathology. Previous studies have used methods like NMF for source separation extensively. For example, NMF has been used to recover spectral and spatial characteristics of structures of interest in microscopy images of aorta tissue (Galeano et al., 2015), breast cancer cells (Pu et al., 2011), and the colon (Roxbury et al., 2015). Additionally, in contrast to our current goals, autofluorescence is a large source of noise for fluorescent dye-based microscopy and its removal is highly desired. NMF-based procedures have been proposed for the detection and subsequent removal of cellular autofluorescence to enhance fluorescence imaging (Woolfe et al., 2011). Members of the medical image analysis community relying on linear non-negative methods for spectral analysis may benefit by taking the high-dimensional multi-linear non-negative tensor decomposition approach proposed in this study.

Declarations of Competing of Interest

Acknowledgments

Appendix A. List of Abbreviations

AMD Age-related Macular Degeneration
AF Autofluorescence
BrM Bruch's Membrane
CP CPD/PARAFAC
CPD Canonical Polyadic Decomposition
CORCONDIA Core Consistency Diagnostic
LF/ML Lipofuscin/Melanolipofuscin
PARAFAC Parallel Factor Analysis
NMF Non-negative Matrix Factorization
NNCP Non-negative CP
RPE Retinal Pigment Epithelium
AO Alternating Optimization
ADMM Alternating Direction Method of Multipliers

References

Ach, T., Huisingh, C., McGwin, G., Messinger, J.D., Zhang, T., Bentley, M.J., Gutierrez, D.B., Ablonczy, Z., Smith, R.T., Sloan, K.R., et al., 2014. Quantitative autofluorescence and cell density maps of the human retinal pigment epithelium autofluorescence and cell density. Invest. Ophthalmol. Vis. Sci. 55 (8), 4832–4841.

Ami, T.B., Tong, Y., Bhuiyan, A., Huisingh, C., Ablonczy, Z., Ach, T., Curcio, C.A., Smith, R.T., 2016. Spatial and spectral characterization of human retinal pigment epithelium fluorophore families by ex vivo hyperspectral autofluorescence imaging. Transl.Vis.Sci.Technol. 5 (3). 5–5.

Anandkumar, A., Ge, R., Hsu, D., Kakade, S.M., Telgarsky, M., 2014. Tensor decompositions for learning latent variable models. J. Mach. Learn. Res. 15, 2773–2832.

Andersson, C.A., Bro, R., 2000. The N-way toolbox for MATLAB. Chemom. Intell. Lab. Syst. 52 (1), 1–4.

Bader, B.W., Kolda, T.G., 2015. Matlab Tensor Toolbox Version 2.6. Available online

Bharath, H., Sima, D., Sauwen, N., Himmelreich, U., De Lathauwer, L., Van Huffel, S., 2017. Nonnegative canonical polyadic decomposition for tissue-type differentiation in gliomas. IEEE J. Biomed. Health Inform. 21 (4), 1124–1132.

Bhosale, P., Li, B., Sharifzadeh, M., Gellermann, W., Frederick, J.M., Tsuchida, K., Bernstein, P.S., 2009. Purification and partial characterization of a lutein-binding protein from human retina. Biochemistry 48 (22), 4798–4807.

Bioucas-Dias, J.M., Plaza, A., Dobigeon, N., Parente, M., Du, Q., Gader, P., Chanussot, J., 2012. Hyperspectral unmixing overview: geometrical, statistical, and sparse regression-based approaches. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 5 (2), 354–379.

Bourne, R.R.A., Stevens, G.A., White, R.A., Smith, J.L., Flaxman, S.R., Price, H., Jonas, J.B., Keeffe, J., Leasher, J., Naidoo, K., Pesudovs, K., Resnikoff, S., Taylor, H.R., 2013. Causes of vision loss worldwide, 1990–2010: a systematic analysis. Lancet Global Health 1 (6), e339–e349.

Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J., et al., 2011. Distributed optimization and statistical learning via the alternating direction method of multipliers. Found. Trends® Mach.Learn. 3 (1), 1–122.

Bro, R., Kiers, H.A.L., 2003. A new efficient method for determining the number of components in PARAFAC models. J. Chemom. 17 (5), 274–286. doi:10.1002/cem.801.

Carroll, J.D., Chang, J.-J., 1970. Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition. Psychometrika 35 (3), 283–319.

Cichocki, A., Mandic, D., De Lathauwer, L., Zhou, G., Zhao, Q., Caiafa, C., Phan, H.A., 2015. Tensor decompositions for signal processing applications: from two-way to multiway component analysis. IEEE Signal Process Mag 32 (2), 145–163.

Comon, P., 1994. Independent component analysis, a new concept? Signal Processing 36 (3), 287–314.

Cong, F., Lin, Q.-H., Kuang, L.-D., Gong, X.-F., Astikainen, P., Ristaniemi, T., 2015. Tensor decomposition of EEG signals: a brief review. J. Neurosci. Methods 248, 59–69.

Crouch, R.K., Koutalos, Y., Kono, M., Schey, K., Ablonczy, Z., 2015. A2e and lipofuscin. In: Hejtmancik, J.F., Nickerson, J.M. (Eds.), Molecular Biology of Eye Disease. In: Progress in Molecular Biology and Translational Science, 134. Academic Press, pp. 449–463.

Curcio, C.A., Messinger, J.D., Sloan, K.R., McGwin, G., Medeiros, N.E., Spaide, R.F., 2013. Subretinal drusenoid deposits in non-neovascular age-related macular degeneration: morphology, prevalence, topography, and biogenesis model. Retina (Philadelphia, Pa.) 33 (2).

Dobigeon, N., Altmann, Y., Brun, N., Moussaoui, S., 2016. Linear and nonlinear unmixing in hyperspectral imaging. Data Handl. Sci. Technol. 30, 185–224.

Duchi, J., Shalev-Shwartz, S., Singer, Y., Chandra, T., 2008. Efficient projections onto the l1-ball for learning in high dimensions. In: Proceedings of the 25th International Conference on Machine Learning. ACM, pp. 272–279.

Galeano, J., Perez, S., Montoya, Y., Botina, D., Garzón, J., 2015. Blind source separation of ex-vivo aorta tissue multispectral images. Biomed. Opt. Express. 6 (5), 1589–1598.

Gao, L., Smith, R.T., 2015. Optical hyperspectral imaging in microscopy and spectroscopy âĂŞ a review of data acquisition. J. Biophotonics 8 (6), 441–456.

Geng, X., Gouttard, S., Sharma, A., Gu, H., Styner, M., Lin, W., Gerig, G., Gilmore, J.H., 2012. Quantitative tract-based white matter development from birth to age 2 years. Neuroimage 61 (3), 542–557.

Gobinet, C., Perrin, E., Huez, R., 2004. Application of non-negative matrix factorization to fluorescence spectroscopy. In: Signal Processing Conference, 2004 12th European. IEEE, pp. 1095–1098.

Hall, P., Müller, H.-G., Wang, J.-L., 2006. Properties of principal component methods for functional and longitudinal data analysis. Ann.Stat. 1493–1517.

Harshman, R., 1970. Foundations of the PARAFAC procedure: models and conditions for an" explanatory" multi-modal factor analysis.

Håstad, J., 1990. Tensor rank is NP-complete. J. Algorithm. 11 (4), 644–654.

Heylen, R., Parente, M., Gader, P., 2014. A review of nonlinear hyperspectral unmixing methods. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 7 (6), 1844–1868.

Hillar, C.J., Lim, L.-H., 2013. Most tensor problems are NP-hard. J. ACM (JACM) 60 (6), 45.

Hitchcock, F.L., 1927. The expression of a tensor or a polyadic as a sum of products. Stud. Appl. Math. 6 (1–4), 164–189.

Hotelling, H., 1933. Analysis of a complex of statistical variables into principal components.. J Educ Psychol. 24 (6), 417.

Huang, K., Sidiropoulos, N.D., Liavas, A.P., 2016. A flexible and efficient algorithmic framework for constrained matrix and tensor factorization. IEEE Trans. Signal Process. 64 (19), 5052–5065.

Ikoma, H., Heshmat, B., Wetzstein, G., Raskar, R., 2014. Attenuation-corrected fluorescence spectra unmixing for spectroscopy and microscopy. Opt. Express 22 (16), 19469–19483.

Johri, A., Post, R., Ganti, B., Fawzi, A.A., Sajda, P., Ach, T., Curcio, C.A., Smith, T., 2014. Mathematical modeling of retinal pigment epithelium (RPE) autofluorescence (AF) with gaussian mixture models and non-negative matrix factorization (NMF). Invest. Ophthalmol. Vis. Sci. 55 (13). 3827–3827

Kasha, M., 1950. Characterization of electronic transitions in complex molecules. Discuss. Faraday Soc. 9, 14–19.

Kennedy, C.J., Rakoczy, P.E., Constable, I.J., 1995. Lipofuscin of the retinal pigment epithelium: a review. Eye 9 (6), 763–771.

Keshava, N., Mustard, J.F., 2002. Spectral unmixing. IEEE Signal Process Mag. 19 (1), 44–57.

Kiers, H.A.L., 1997. Weighted least squares fitting using ordinary least squares algorithms. Psychometrika 62 (2), 251–266. doi:10.1007/BF02295279.

Kim, H., Park, H., 2008. Nonnegative matrix factorization based on alternating nonnegativity constrained least squares and active set method. SIAM J. Matrix Anal. Appl. 30 (2), 713–730.

Kolda, T., Bader, B., 2009. Tensor decompositions and applications. SIAM Rev. 51 (3), 455–500.

Kolda, T.G., 2006. Multilinear Operators for Higher-order Decompositions.

Kroonenberg, P.M., 2007. Missing Data in Multiway Analysis. John Wiley & Sons, Inc., pp. 143–174.

Kruskal, J., 1989. Rank, Decomposition, and Uniqueness for 3-Way and N-Way Arrays. Elsevier Science doi:10.1002/bimj.4710330518.

Lakowicz, J.R., 2006. Introduction to fluorescence. In: Principles of Fluorescence Spectroscopy. Springer, pp. 1–26.

Lakowicz, J.R., 2006. Solvent and Environmental Effects. Springer, pp. 205–235.

Lee, D.D., Seung, H.S., 1999. Learning the parts of objects by non-negative matrix factorization. Nature 401 (6755), 788–791.

Lee, H., Yoo, J., Choi, S., 2009. Semi-supervised nonnegative matrix factorization. IEEE Signal Process. Lett. 17 (1), 4–7.

Li, R., Wu, P., Yakushev, I., Wang, J., Ziegler, S.I., Förster, S., Huang, S.-C., Schwaiger, M., Navab, N., Zuo, C., Shi, K., 2017. Pattern visualization and recognition using tensor factorization for early differential diagnosis of Parkinsonism. In: Medical Image Computing and Computer-Assisted Intervention âĹŠ MICCAI 2017: 20th International Conference, Quebec City, QC, Canada, September 11–13, 2017, Proceedings, Part III. Springer International Publishing, pp. 125–133.

Lin, C.-J., 2007. On the convergence of multiplicative update algorithms for nonnegative matrix factorization. IEEE Trans. Neural Networks 18 (6), 1589–1596.

Liu, J., Musialski, P., Wonka, P., Ye, J., 2013. Tensor completion for estimating missing values in visual data. IEEE Trans. Pattern Anal. Mach. Intell. 35 (1), 208–220.

Liu, K., So, H.-C., da Costa, J.P.C., Huang, L., 2013. Core consistency diagnostic aided by reconstruction error for accurate enumeration of the number of components in PARAFAC models. In: Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on. IEEE, pp. 6635–6639.

Möller, A., Tutz, G., Gertheiss, J., 2016. Random forests for functional covariates. J. Chemom. 30 (12), 715–725.

Montcuquet, A.-S., Herve, L., Navarro, F., Dinten, J.-M., Mars, J.I., 2011. In vivo fluorescence spectra unmixing and autofluorescence removal by sparse nonnegative matrix factorization. IEEE Trans. Biomed. Eng. 58 (9), 2554–2565.

Mørup, M., Hansen, L.K., 2009. Automatic relevance determination for multi-way models. J. Chemom. 23 (7–8), 352–363.

Nascimento, J.M., Dias, J.M., 2005. Does independent component analysis play a role in unmixing hyperspectral data? IEEE Trans. Geosci. Remote Sens. 43 (1), 175–187.

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

Norris, J.L., Caprioli, R.M., 2013. Analysis of tissue specimens by matrix-assisted laser desorption/ionization imaging mass spectrometry in biological and clinical research. Chem. Rev. 113 (4), 2309–2342.

Paatero, P., Tapper, U., 1994. Positive matrix factorization: a non-negative factor model with optimal utilization of error estimates of data values. Environmetrics 5 (2), 111–126.

Pomann, G.-M., Staicu, A.-M., Ghosh, S., 2016. A two-sample distribution-free test for functional data with application to a diffusion tensor imaging study of multiple sclerosis. J. R. Stat. Soc.: Ser. C (Appl. Stat.) 65 (3), 395–414.

Pu, Y., Tang, G.C., Wang, W.B., Savage, H.E., Schantz, S.P., Alfano, R.R., 2011. Native fluorescence spectroscopic evaluation of chemotherapeutic effects on malignant cells using nonnegative matrix factorization analysis. Technol. Cancer Res. Treat. 10 (2), 113–120.

Ramsay, J.O., Silverman, B.W., 2007. Applied Functional Data Analysis: Methods and Case Studies. Springer.

Roxbury, D., Jena, P.V., Williams, R.M., Enyedi, B., Niethammer, P., Marcet, S., Verhaegen, M., Blais-Ouellette, S., Heller, D.A., 2015. Hyperspectral microscopy of near-infrared fluorescence enables 17-chirality carbon nanotube imaging. Sci. Rep. 5.

Saeys, W., De Ketelaere, B., Darius, P., 2008. Potential applications of functional data analysis in chemometrics. J. Chemom. 22 (5), 335–344.

Sajda, P., Du, S., Parra, L.C., 2003. Recovery of constituent spectra using non-negative matrix factorization. In: Wavelets: Applications in Signal and Image Processing X, 5207. International Society for Optics and Photonics, pp. 321–332.

Scholz, F.W., Stephens, M.A., 1987. K-sample Anderson–Darling tests. J. Am. Stat. Assoc. 82 (399), 918–924.

Schultz, T., Seidel, H.-P., 2008. Estimating crossing fibers: a tensor decomposition approach. IEEE Trans. Vis. Comput. Graph 14 (6).

Smilde, A., Bro, R., Geladi, P., 2005. Multi-Way Analysis: Applications in the Chemical Sciences. John Wiley & Sons.

Smith, R.T., Post, R., Johri, A., Lee, M.D., Ablonczy, Z., Curcio, C.A., Ach, T., Sajda, P., 2014. Simultaneous decomposition of multiple hyperspectral data sets: signal recovery of unknown fluorophores in the retinal pigment epithelium. Biomed. Opt. Express 5 (12), 4171–4185.

Sparrow, J.R., Boulton, M., 2005. RPE lipofuscin and its role in retinal pathobiology. Exp. Eye Res. 80 (5), 595–606.

Stephens, M.A., 1974. EDF statistics for goodness of fit and some comparisons. J. Am. Stat. Assoc. 69 (347), 730–737.

Suhling, K., Hirvonen, L.M., Levitt, J.A., Chung, P.-H., Tregidgo, C., Marois, A.L., Rusakov, D.A., Zheng, K., Ameer-Beg, S., Poland, S., Coelho, S., Henderson, R., Krstajic, N., 2015. Fluorescence lifetime imaging (flim): basic concepts and some recent developments. Med. Photonics 27, 3–40. Fluorescence lifetime measurements in the biomedical sciences

Tomasi, G., Bro, R., 2005. PARAFAC and missing values. Chemom. Intell. Lab. Syst. 75 (2), 163–180.

Tong, Y., Ami, T.B., Hong, S., Heintzmann, R., Gerig, G., Ablonczy, Z., Curcio, C.A., Ach, T., Smith, R.T., 2016. Hyperspectral autofluorescence imaging of drusen and retinal pigment epithelium in donor eyes with age-related macular degeneration. Retina 36, S127–S136.

Tucker, L.R., 1964. The extension of factor analysis to three-dimensional matrices. Contrib.Math.Psychol. 110119.

Wang, W., Carreira-Perpinán, M. A., 2013. Projection onto the probability simplex: an efficient algorithm with a simple proof, and an application. arXiv preprint arXiv:1309.1541.

Waters, J.C., 2009. Accuracy and precision in quantitative fluorescence microscopy. J. Cell Biol. 185 (7), 1135–1148. doi:10.1083/jcb.200903097.

Wong, W.L., Su, X., Li, X., Cheung, C.M.G., Klein, R., Cheng, C.-Y., Wong, T.Y., 2014. Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: a systematic review and meta-analysis. Lancet Glob. Health 2 (2), e106–e116.

Woolfe, F., Gerdes, M., Bello, M., Tao, X., Can, A., 2011. Autofluorescence removal by non-negative matrix factorization. IEEE Trans. Image Process. 20 (4), 1085–1093.

Yao, F., Müller, H.-G., Clifford, A.J., Dueker, S.R., Follett, J., Lin, Y., Buchholz, B.A., Vogel, J.S., 2003. Shrinkage estimation for functional principal component scores with application to the population kinetics of plasma folate. Biometrics 59 (3), 676–685.

Zhang, Z., Ely, G., Aeron, S., Hao, N., Kilmer, M., 2014. Novel methods for multilinear data completion and de-noising based on tensor-SVD. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3842–3849.

Zhu, H., Kong, L., Li, R., Styner, M., Gerig, G., Lin, W., Gilmore, J.H., 2011. Fadtts: functional analysis of diffusion tensor tract statistics. NeuroImage 56 (3), 1412–1425.

Neel Dey graduated with an M.S. in Electrical Engineering from New York University. He is currently a Ph.D. candidate in Computer Science at New York University, with research interests pertaining to statistical learning theory and applications in the context of medical image analysis.

Sungmin Hong is a Ph.D. candidate in Computer Science at New York University. He has B.S. and M.S. in Electrical Engineering and Computer Science from Seoul National University, Republic of Korea. His research interests include longitudinal shape analysis, manifold statistics, and deformation analysis for medical image analysis.

Thomas Ach, M.D., M.Sc., is a German board-certified retina specialist currently employed at the Department of Ophthalmology, University Hospital Würzburg, Germany. After residency, he spent >25 months as a research fellow in Dr. Christine Curcio's lab at the University of Alabama at Birmingham, USA. During this time basic science of histologic and microscopic data regarding retinal pigment epithelium cell number and autofluorescence was gained, both in normal aging and disease (i.e., age-related macular degeneration). He has published >40 papers in the field of basic science and clinical retina and has been awarded several prizes for his research.

Yiannis Koutalos has a degree in Physics from the University of Athens, Greece, a Ph.D. in Biophysics from the University of Illinois at Urbana-Champaign, and postdoctoral studies at Johns Hopkins University School of Medicine. He is Professor of Ophthalmology and Neurosciences at the Medical University of South Carolina. His research is on the metabolism of vitamin A in the retina with particular emphasis in the use of fluorescence imaging techniques.

Christine A. Curcio, Ph.D., is the White-McKee Endowed Professor of Ophthalmology and Vision Sciences at the University of Alabama at Birmingham. Trained in neuroanatomy, she has collaborated with eye banks for 33 years in producing seminal observations of human retinal neurobiology, retinal aging and pathology, including description of major aspects of age-related macular degeneration and validation of clinical imaging techniques (optical coherence tomography, adaptive optics scanning laser ophthalmoscopy, and fundus autofluorescence). Techniques include electron microscopy, morphometry, digital pathology, histochemistry, lipid

profiling; collaborative studies have involved epidemiology, visual function testing, and transcriptomics.

R. Theodore Smith, M.D., Ph.D., is Professor of Ophthalmology and Neuroscience in the Department of Ophthalmology, Icahn School of Medicine at Mount Sinai, and Director, Biomolecular Retinal Imaging, at the New York Eye and Ear Infirmary of Mount Sinai. He holds a B.A. in Mathematics from Rice University, a Ph.D. in Pure Mathematics from the University of Warwick, UK, and an M.D. from the Albert Einstein College of Medicine. He has held faculty positions in Mathematics at MIT and Columbia University, and in Ophthalmology at Columbia, NYU and Mount Sinai medical centers.

Guido Gerig is Institute Professor in the Computer Science and Engineering Department of NYU's Tandon School of Engineering. He holds M.S, Ph.D. and Venia Legendi degrees from ETH Zurich. He is a Fellow of AIMBE and MICCAI and a senior member of IEEE. His research supports a number of clinical imaging research studies with novel, innovative image analysis methodologies for segmentation, registration, atlas building, shape analysis, and image statistics. He was previously USTAR Professor of Computer Science at the University of Utah (2007–2015), Taylor Grandy Professor of Computer Science and Psychiatry at UNC Chapel Hill (1998–2007), and Assistant Professor at ETH Zurich (1993–1998).