**Automatic Model Determination for Quaternion NMF**

GIANCARLO SANCHEZ¹, ERIK SKAU², AND BOIAN ALEXANDROV ²

¹Florida International University, Miami, FL 33199 USA (e-mail: gsanc018@fiu.edu)
²Los Alamos National Laboratory, Los Alamos, NM 87545

This research was funded by the U.S. Department of Energy National Nuclear Security Administration under Contract No. 89233218CNA000001 and supported by Los Alamos National Laboratory: Institutional Computing Program, National Security Education Center, and LDRD Grant 20190020DR.

**ABSTRACT** Nonnegative Matrix Factorization (NMF) is a well-known method for Blind Source Separation (BSS). Recently, BSS for polarized signals in spectropolarimetric data, containing both polarization and spectral information, was introduced. This information was encoded in 4-dimensional Stokes vectors represented by quaternion numbers. In the proposed Quaternion NMF (QNMF), the common challenge of determining the (usually) unknown number of quaternion signals remained unaddressed. Estimating the number of signals (aka model determination) is important, since an underestimation of this number results in poor source separation and omission of signals, while overestimation leads to extraction of noisy signals without physical meaning. Here, we introduce a method for determining the number of polarized signals in spectropolarimetric data, named QNMF:k. QNMF:k integrates: (a) Quaternion Alternating Direction Method of Multipliers (QADMM) implemented for QNMF, (b) random resampling of the initial quaternion data, and (c) custom clustering of sets of QADMM solutions with same number of sources, k, needed to estimate the stability of the solutions. The appropriate latent dimension is determined based on the stability of the solutions. We demonstrate that, without any prior information, QNMF:k accurately extracts the correct number of signals used to generate synthetic quaternion datasets and a benchmark spectropolarimetric data.

**INDEX TERMS** Blind Source Separation, Model Determination, Polarization, Quaternion NMF, Spectropolarimetric Imaging

**INTRODUCTION**

**P**olarization is a property of transverse waves that describes oscillations perpendicular to the direction of propagation of the wave. Modern sensors allow us to capture both the spectral signatures and the polarization information of different wave types, such as, electromagnetic waves [1], seismic waves [2], acoustic waves [3], and many others. The polarized wave includes two real-valued signals \((u(t), v(t))\) which can be represented as a complex function \(\phi(t) = u(t) + iv(t)\) [4]. The complex function, \(\phi(t)\), describes an elliptical trajectory and the corresponding wave field is called elliptically polarized. When the direction of the arrival of such wave field is known, a standard description of the polarization ellipse can be done by the Stokes parameters, \(S_0, S_1, S_2,\) and \(S_3\), which are measurable quantities [5].

Recently, it has become clear that obtaining both spectral and polarization information often leads to new results and understanding in various applications [6]–[10]. This new experimental technique, that involves recording the Stokes vector \((S_0, S_1, S_2, S_3)\) for every pixel of the image to obtain such information, is called spectropolarimetric imaging.

In practice, spectropolarimetric imaging information usually contains mixtures of multiple polarized sources which remain latent (not directly observed). To infer these original sources solely from data, one needs the power of model-free unsupervised blind source separation (BSS) techniques [11]. BSS separates the source signals from their observed mixtures with no additional information about the original signals [12], [13]. BSS includes factorization methods, such as, Independent Component Analysis [14] and Non-negative Matrix Factorization (NMF) [15], which decompose the observational matrix (containing mixture of the original signals) into a product of two low-rank matrices, called factors. One of these factor matrices contains the original sources and the other – their activities. A unique advantage of NMF is that the extracted latent sources are explainable because they are parts of the original data [16]. The identification of the (usually unknown in practice) number of original sources is...
an important step in BSS. There are various heuristics used to find this important number. The Automatic Relevance Determination (ARD) method, originally introduced for neural networks by MacKay [17] and then adapted for probabilistic interpretation of NMF by Fevotte and Tan [18] is one of the popular Bayesian techniques for determination of the latent dimensionality.

Recently, Quaternion Non-Negative Matrix Factorization (QNMF) was introduced as a new BSS method for analyzing of wide-band signals and frequency-dependent polarization fields present in spectropolarimetric data [19]. QNMF is a generalization of NMF that is able to work with polarized signals. QNMF relies on quaternion numbers to encode polarization information and the physical constraints of Stokes parameters. However, to obtain meaningful and explainable factors, the introduced QNMF needs prior information of the number of the latent polarized (quaternion) sources in the system - information that is usually unavailable.

In this paper, we introduce a method for estimating the true number, \( k \), of original quaternion sources, whose mixture generates the analyzed data. We refer to this method as QNMF\( k \). QNMF\( k \) is a generalization of a recent method for automatic model determination, called NMF\( k \), for quaternion matrices. When applied to a large number of synthetic datasets with predetermined latent dimensions, NMF\( k \) has demonstrated a superior performance in comparison to other heuristics [20]. QNMF\( k \) determines the latent dimension based on the cluster stability of QNMF solutions of a set of random quaternion matrices with mean value the original observational data. Thus, QNMF\( k \) utilizes a specific quaternion bootstrap and determines the unknown number of latent sources based on stability of QNMF solutions (with different number of quaternion sources, \( k \)) without prior assumptions. We demonstrate that QNMF\( k \) accurately determines the number of latent sources and factors in several synthetically generated polarized signals as well as in a previous benchmark quaternion data with predetermined number of polarized sources.

**BACKGROUND**

**POLARIZATION AND STOKES PARAMETERS**

Stokes parameters provide a characterization of polarized waves by means of four intensity parameters \( S_0, S_1, S_2, S_3 \). Here, \( S_0 \) is the total intensity of the wave while \( S_1, S_2, S_3 \) describe the polarization. The degree of polarization \( P \in [0, 1] \) is given by \( P = \frac{\sqrt{S_1^2 + S_2^2 + S_3^2}}{S_0} \) where \( P = 1 \) corresponds to a completely polarized wave, \( P = 0 \) to completely unpolarized wave, and anything in between is deemed partially polarized. Stokes parameters can be represented by a 4-dimensional vector \( S = (S_0, S_1, S_2, S_3) \in \mathbb{R}^4 \) known as the Stokes vector. Physical properties of waves lead to algebraic constraints on the space of Stokes vectors. The first constraint \( i \) \( S_0 \geq 0 \) indicates that there cannot be a negative intensity wave, while the second constraint \( ii \) \( S_0^2 \geq S_1^2 + S_2^2 + S_3^2 \), imposes that the total wave intensity is greater than or equal to the polarized components magnitude.

**QUATERNIONS**

Instead of considering Stokes vectors as elements in \( \mathbb{R}^4 \), every Stokes vector can be represented algebraically as a single quaternion number in \( \mathbb{H} \). This number completely describes the polarization state and encodes the physical constraints of the parameters. This encoding of the data allows us to represent spectropolarimetric data as an \( M \times N \) matrix with quaternion elements. Formally, quaternion numbers are defined as the set, \( \mathbb{H} = \{ a + bi + cj + dk \mid a, b, c, d \in \mathbb{R} \} \) where \( i^2 = j^2 = k^2 = ijk = -1 \). For any \( q \in \mathbb{H} \), the real part of \( q \) is denoted by \( \text{Re}(q) = a \) while the imaginary part (or vector part) is \( \text{Im}(q) = b \, i + c \, j + d \, k \). Quaternions have component-wise addition and component-wise scalar multiplication. The multiplicative group structure is known as the Hamilton product and is given by the distributive law and multiplication of basis elements \( \{ 1, i, j, k \} \). The Hamilton product is associative but not commutative. Similarly to the complex numbers, the conjugate of \( q \) is defined as the number \( q^{-1} = a - bi - cj - dk \). The norm of a quaternion \( q \) is given by \( \|q\|^2 = \sqrt{a^2 + b^2 + c^2 + d^2} \). Equipping \( \mathbb{H} \) with the metric \( d(p, q) = \|p - q\| \) makes \( \mathbb{H} \) a formal metric space.

Given a Stokes vector \( (S_0, S_1, S_2, S_3) \in \mathbb{R}^4 \) we identify it with the quaternion \( q = S_0 + S_1i + S_2j + S_3k \). To be admissible as a Stokes vector, the parameters must satisfy constraints \( (i) \) and \( (ii) \) imposed by the data as stated in the section 2.1. This dictates that we work with quaternions \( q = a + bi + cj + dk \) which satisfy the non-negativity constraints \( S \) given by \( a > 0 \) and \( b^2 + c^2 + d^2 \leq a^2 \). Therefore, we will be interested in a ‘non-negative’ subset \( \mathbb{H}_S \subset \mathbb{H} \) given by \( \mathbb{H}_S = \{ q \mid \text{Re}(q) > 0 \} \) and \( |\text{Im}(q)|^2 \leq |\text{Re}(q)|^2 \). The identification of Stokes vectors to this constrained set is crucial for our studies as it formally generalizes the notion of ‘non-negativity’ to quaternion numbers. \( \mathbb{H}_S \) is isomorphic to the set of \( 2 \times 2 \) non-negative (or Positive Semi-Definite) Hermitian matrices \( \mathbb{K}^+ \) given by the map

\[
q \leftrightarrow M_q = \frac{1}{2} \begin{pmatrix} a + c & d - bi \\ d + bi & a - c \end{pmatrix}
\]

For an arbitrary quaternion \( q \in \mathbb{H} \), we can project it to \( \mathbb{H}_S \) by considering its associated 2x2 complex matrix defined above and projecting to \( \mathbb{K}^+ \). The projection is given by

\[
\Pi_{\mathbb{H}_S}(M_q) = \sum_{i=0}^{2} \max(0, \eta_i) v_i v_i^\dagger = \begin{pmatrix} p_1 & p_2 \\ p_2 & p_3 \end{pmatrix}
\]

where \( \eta_i \) and \( v_i \) are the \( i^{th} \) eigenvalue and eigenvectors. Then, our projection is given by

\[
p_1 + p_3 + 2\text{Im}(p_2) + j(p_1 - p_3) + 2k\text{Re}(p_2)
\]

Flamant et al. [19] introduced a QNMF algorithm, which is viable when the underlying number of polarized sources, \( k \), is known beforehand. The original paper relies on a
Input Quaternion Matrix: $X \in \mathbb{H}^{M \times N}$

Generate Ensemble of Resampled Quaternion Matrices
$
\{X_r | X_r = \Pi_{HS}(X + \delta_r)\}
$

Explore $k$ topics by applying QNMF to resampled matrices
$$
\begin{aligned}
\arg \min_{W_q, H_q} & \frac{1}{2} \|X_q - W_q H_q\|^2_F \\
\text{subject to} & \quad W_q \in \mathbb{H}^{M \times K}_S, \quad H_q \in \mathbb{R}^{K \times N}_+, \\
& \quad \text{for } 1 \leq q \leq r.
\end{aligned}
$$

Cluster features into $k$ clusters

Check Cluster Stability

increase $k$

FIGURE 1: Flowchart depicting the QNMF/k scheme.

alternating least squares method to approximate the factorized matrices. Alternatively, one can employ a quaternion alternating direction method of multipliers (QADMM) [22]. From our testing, we found that QADMM provides faster and more consistent results. This is valuable since our QNMF/k algorithm, which determines $k$, requires computing multiple factorizations with different values of $k$.

METHODS

**QUATERNION NMF**

Quaternion Non-negative Matrix Factorization (QNMF) finds the original polarized sources from the observed data matrix, $X$, where every entry of $X$ is a superposition/mixture of these original polarized sources. This means that if we have $M$ wavelengths and $N$ sensors, for some wavelength $\lambda$ and sensor $u$, the polarized signal $x(\lambda, u)$ is assumed to be linear combination of $K < \min\{4M, N\}$ distinct basis sources $w_1, \ldots, w_K \in \mathbb{H}_S$ and non-negative activations $h_1, \ldots, h_K \in \mathbb{R}$, i.e.,

$$
x(\lambda, u) = \sum_{k=1}^{K} w_k(\lambda) h_k(u).
$$

QNMF approximates the data matrix $X \in \mathbb{H}^{M \times N}_S$ as a product of two low-rank factor matrices $W \in \mathbb{H}^{M \times K}_S$ and $H \in \mathbb{R}^{K \times N}_+$, such that $X \approx WH$. Here, the columns of matrix $W$ are the latent original polarized sources, represented as quaternion vectors, while the rows of matrix $H$ are activations observed at each sensor. QNMF seeks a numerical solution to the optimization,

$$
\begin{aligned}
\arg \min_{W, H} & \quad \frac{1}{2} \|X - WH\|^2_F \\
\text{subject to} & \quad W \in \mathbb{H}^{M \times K}_S, \quad H \in \mathbb{R}^{K \times N}_+.
\end{aligned}
$$

**QUATERNION ADMM UPDATES**

ADMM is an optimization procedure used to solve convex problems by breaking the problem into easier subproblems [23]. QNMF quaternion problem can be written in a form suitable to ADMM by splitting the variables $W$ and $H$ into primary, $W_p$ and $H_p$, and auxiliary, $W_a$ and $H_a$, variables and moving the constraint to the objective function. With this, Equation (1) becomes,

$$
\begin{aligned}
\arg \min_{W_p, W_a, H_p, H_a} & \quad \frac{1}{2} \|X - W_p H_p\|^2_F + \chi_{\mathbb{H}_S}(W_a) + \chi_+(H_a) \\
\text{subject to} & \quad W_p = W_a, \quad H_p = H_a,
\end{aligned}
$$

where $\chi_{\mathbb{H}_S}$ is the characteristic function indicating if the matrix adheres to the Stokes constraint, and $\chi_+$ is the non-negative characteristic function. Since our problem is convex in $W$ and $H$ independently, we alternate optimizing over $W_p, W_a, W_d$ and $H_p, H_a, H_d$ for the $W$ and $H$ updates respectively, where $W_d$ and $H_d$ are the dual are the dual variables for the equality constraints. The resulting QADMM algorithm in scaled form to solve for $W$ iterates the update rules described in Algorithm 1.

**Algorithm 1 $W_{\text{Update}}$**

**Input:** $X, W_a, W_d, H, \rho_W$  
**Output:** $W_a, W_d$  

$W_p \leftarrow (XH^T + \rho_W W_a - \rho_W W_d)(HH^T + \rho_W I)^{-1}$  
$W_a \leftarrow \Pi_{\mathbb{H}_S}(W_p + W_d)$  
$W_d \leftarrow W_d + W_p - W_a$

where $\rho_W$ is the convergence parameter. Similarly the scaled form of the QADMM algorithm for $H$ consists of the updates shown in Algorithm 2.

**Algorithm 2 $H_{\text{Update}}$**

**Input:** $X, W, H_a, H_d, \rho_H$  
**Output:** $H_a, H_d$  

$H_p \leftarrow (W^*W + \rho_H I)^{-1}(\text{Re}(W^*X) + \rho_H H_a - \rho_H H_d)$  
$H_a \leftarrow \Pi_{\mathbb{R}^+_S}(H_p + H_d)$  
$H_d \leftarrow H_d + H_p - H_a$

It is standard practice to use the auxiliary variables $W_a$ and
$H_a$ as the fixed variables in the alternating optimizations to ensure that the constraints are met. Additionally, faster overall convergence is achieved by interleaving the QADMM updates for $W$ and $H$ rather than solving each $W_k$ and $H$ subproblems to convergence as depicted in Algorithm 3.

Algorithm 3 QADMM

Input: $X, k, \rho_W, \rho_H$

Output: $W_a, W_d, H_a, H_d$

Initialize $W_a, W_d, H_a, H_d$

while not converged do

$H_a, H_d \leftarrow H\_update(X, W_a, H_a, H_d, \rho_H)$

$W_a, W_d \leftarrow W\_update(X, W_a, W_d, H_a, \rho_W)$

end while

RESULTS

QNMF WITH AUTOMATIC MODEL DETERMINATION

The QNMF algorithm is adequate for extraction of latent polarized sources when their true number $k$ is known, however this information is not typically known in practice. To determine this (usually unknown) number, we introduce here, QNMF$k$, a method that integrates: (a) QADMM minimization, (b) random resampling of the initial data, and (c) custom clustering. For the determination of the unknown number of latent features we follow the ideas of a recent NMF algorithm [24]–[26]. Algorithm 4 details this procedure and a flowchart of QNMF$k$ can be seen in Figure 1.

Algorithm 4 QNMF$k$

Input: $X$, $k$, $\rho_W, \rho_H$

Output: $k$

for $k = 1, \ldots, \min(4M, N) - 1$ do

for $r = 1, \ldots, R$ do

$X_r \leftarrow I_{4g}(X + \delta_r)$

$W_r, H_r \leftarrow QADMM(X_r, k, \rho_w, \rho_H)$

end for

$C_k \leftarrow Cluster(k, W_1, \ldots, W_R)$

$s_k \leftarrow SilhouetteStatistics(C_k)$

end for

Based on the original data, QNMF$k$ creates an ensemble of randomly perturbed quaternion matrices on which it performs QADMM minimization. While QNMF$k$ can explore all possible numbers of quaternion sources, $k = 1, \ldots, \min(4M, N) - 1$, in practice this range of investigated $k$ values can be tighter. For each $k$, a set of $R \sim 100$ minimizations is computed with random initial conditions. Thus, each QADMM minimization works on a random quaternion matrix: $X_r = I_{4g}(X + \delta_r)$, where each $\delta_r$ is a small uniformly distributed quaternion random perturbation centered at 0. After we construct the set of $R$ solutions (with the explored $k$), the $k$ columns of all quaternion feature solutions, $\{W_1, \ldots, W_R\}$, are subjected to custom clustering into $k$ clusters. This custom clustering is similar to k-means in that it iteratively assigns vectors to centroids but it differs in that it holds that each one of the $k$ clusters must contain exactly one column from each $W_r$ matrix (for this $k$) [20]. These conditions need to be enforced since each QADMM minimization contributes one solution $W_r$ with the same number of columns, $k$, and accordingly it has to supply exactly one element to each of the $k$ clusters. In the clustering, the similarity between the vectors is measured with cosine similarity, defined as $\cos(\theta) = \frac{u^H v}{||u|| ||v||}$ for two vectors $u$ and $v$. After each iteration of clustering, the centroids of each cluster are computed as the mean of assigned vectors and the process is iterated until convergence. Finally, to determine the number of latent features, QNFM$k$ calculates the stability of the clusters for each $k$, via Silhouette statistics, $s_k$ [27]. Silhouette statistics quantifies (a) the cohesion and (b) the separability of clusters. The Silhouettes measure the average similarity between elements belonging to the same cluster (cohesion) and compare it to the similarity of the other clusters (separability). The average Silhouette value, $s_k$, ranges from -1 to +1, where +1 indicates that the vector matches to its assigned cluster well and poorly matches the neighboring clusters, while -1 shows that the vector does not belong to the cluster it was assigned by the clustering. The number of latent quaternion sources, $k$, is determined as the maximum number, $k$, of stable clusters. The centroids of the clusters for the determined latent dimension can be used to reconstruct the initial quaternion matrix, $X$.

EXPERIMENTS

Here, we demonstrate and validate QNMF$k$ on datasets with predetermined patterns and numbers of sources. The application to synthetic data allows the validation of both the latent dimension, as well as correct feature extraction. First, we show QNMF$k$ applied to synthetically constructed data where the polarization orientations contain the six primary polarization orientations. Additionally, we examine QNMF$k$ applied to several randomly generated datasets, as well as to spectropolarimetric data generated in [19]. In all experiments a grid search was conducted to determine the convergence parameters. The convergence parameters for $\rho_W = \rho_H$ that minimized the objective function after 1000 iterations were selected independently for each latent dimension.

Synthetic Constructed Data

To display and test our methodology, we first apply it to a synthetic matrix indexed by four wavelength channels and one hundred sensors, $X \in \mathbb{H}^{4 \times 100}$. Each sensor’s observed signal was generated as a linear combination of six constructed latent signals. The six latent signals were 100% polarized and correspond to vertical, horizontal, 45 degree, -45 degree, clockwise, and counterclockwise polarization schemes in pairs of wavelength channels. The polarization ellipses of the six generative vectors are depicted in Figure 2a. Random exponential combinations of these six signals were added together to simulate the observed signal of each of the 100 sensors in the matrix. We applied QNMF$k$ to the resulting matrix to investigate latent dimensions $k \in [1, 9]$. This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/
Figure 2b depicts the silhouette and relative error selection criteria metrics. The results clearly indicate stable clustering at the correct latent dimension, \( k = 6 \), as the silhouette score is approximately 1, while at \( k = 7 \) the silhouette score drops significantly. To verify that the extracted features are correct, Figure 2c depicts the cosine similarities between the recovered features and the generative features, after the recovered features were permuted to correspond with the generative features. The cosine similarities close to 1 on the diagonal indicate that the extracted features are functionally identical to the generative features. The cosine similarities close to zero show the orthogonal nature between the pairs of signals in the synthetic construction.

Figure 3: Panel (a): Convergence curves for parameters \( \rho = \rho_W = \rho_H \). Panels (b), (c), (d), (e), and (f) show: Silhouette scores, \( S \), in red, the relative errors, \( R \), in blue, and the identified latent dimension, dashed black, for our synthetic random data.

**Synthetic Random Data**

To test QNMF\( k \) on less structured data we constructed random datasets with predetermined number of sources. This was accomplished by independently synthesizing quaternion feature matrices \( W \in \mathbb{H}^{m \times k} \) and activation matrix \( H \in \mathbb{R}^{k \times n} \). Each element of \( W \) was sampled as a complex \( 2 \times 2 \) matrix from the complex Wishart distribution \( \mathcal{CW}(I, 2, 2) \) \[28\]. This distribution, which is a generalized multidimensional gamma distribution, was chosen as it has a support of positive semidefinite matrices with a reasonable eigenvalue distribution, and therefore the sampled elements adhere to the physicality constraints of the Stokes parameters. Each element of \( H \) was sampled from the uniform distribution \( \mathcal{U}(0, 1) \) to ensure non-negativity. Figure 1, panel a, depicts the convergence curves of the QADMM implementation for parameter \( \rho \). Figure 1, panels b, c, d, e, and f depict the average Silhouette, \( S \), and relative error, \( R \), plots for 5 different latent dimension constructed with \( M = 50 \) and \( N = 100 \). We see that at the right latent dimensions, the Silhouette is close to 1, and with a small relative error, thus selecting the correct number of predetermined latent sources in each of the synthetic datasets.
Spectropolarimetric Data

Finally, we tested our method on the spectropolarimetric data generated in Ref. [19]. This data is contained in a wavelength by spatial index matrix and is a linear combination of three fully polarized sources. From Figure 4a we see that our method correctly identifies that there are three polarized sources. Figure 4b shows comparison of the ground truth activations with QNMF\(k\) extracted activations. Finally, the basis latent quaternion vectors, we found, also coincide with these used to construct the data, showing Pearson correlations > 0.9999 with the ground truth sources.

CONCLUSION

In this work, we introduced and demonstrated the effectiveness of a method for determining the number of latent quaternion sources based on their stability. QNMF\(k\) utilizes a QADMM version of previously introduced QNMF minimization, integrated with custom quaternion bootstrapping and clustering algorithm to statistically compare factorization models with different number of latent quaternion sources. We showed that QNMF\(k\) correctly identifies the number of latent quaternion sources in multiple synthetic data sets. To further validate the use of quaternion based models, additional research on quaternion domain optimization as well as experiments on real world spectropolarimetric datasets is required. Current results, however, indicate a promising direction for quaternion based modeling in signal processing.

FIGURE 4: Analysis of spectropolarimetric data generated in [19].

REFERENCES

[1] C. Brosseau, Fundamentals of polarized light: a statistical optics approach. Wiley-Interscience, 1998.
[2] J. Samson, “Matrix and stokes vector representations of detectors for polarized waveforms: theory, with some applications to teleseismic waves,” Geophysical Journal International, vol. 51, no. 3, pp. 583–603, 1977.
[3] A. Bacigalupo and M. Lepidi, “Acoustic wave polarization and energy flow in periodic beam lattice materials,” International Journal of Solids and Structures, vol. 147, pp. 183–203, 2018.
[4] J. M. Lilly and S. C. Olhede, “Bivariate instantaneous frequency and bandwidth,” IEEE Transactions on Signal Processing, vol. 58, no. 2, pp. 591–603, 2009.
[5] G. G. Stokes, “Parameters characterizing electromagnetic wave polarization,” Trans. Cambridge Philos. Soc., vol. 9, p. 399, 1852.
[6] S. Sattar, P.-J. Lapray, A. Foulonneau, and L. Bigué, “Review of spectral and polarization imaging systems,” in Unconventional Optical Imaging II, vol. 11351. International Society for Optics and Photonics, 2020, p. 113511Q.
[7] J. C. Smit, J. H. Rietjens, G. van Harten, A. Di Noia, W. Lauwren, B. E. Rheingans, D. J. Diner, B. Cairns, A. Wasilewski, K. D. Knobelspiesse et al., “Spex airborne spectropolarimeter calibration and performance,” Applied optics, vol. 58, no. 21, pp. 5695–5719, 2019.
[8] M. Rahman, I. H. Cairns, and P. I. McAuley, “Spectropolarimetric imaging of metric type iii solar radio bursts,” SoPh, vol. 295, no. 3, p. 51, 2020.
[9] A. Varin, J. Dellingner, J. Rehbinder, C. Draman, M. Torzynski, C. Heinrich, and J. Zallat, “Ultra-stable spectropolarimeter for dermatology,” in Advanced Biomedical and Clinical Diagnostic and Surgical Guidance Systems XVIII, vol. 11229. International Society for Optics and Photonics, 2020, p. 112290Q.
[10] R. Foster, D. Gray, J. Bowles, D. Korwan, I. Slutsker, M. Sorokin, M. Roche, A. Smith, and L. Pezzaniti, “Mantis: an all-sky visible-to-near-infrared hyper-angular spectropolarimeter,” Applied optics, vol. 59, no. 20, pp. 5896–5909, 2020.
[11] A. Belouchnani, K. Abed-Meraim, J.-F. Cardoso, and E. Moulines, “A blind source separation technique using second-order statistics,” IEEE Transactions on signal processing, vol. 45, no. 2, pp. 434–444, 1997.
[12] Y. Xie, K. Xie, and S. Xie, “Underdetermined blind separation of source using lp-norm diversity measures,” Neurocomputing, vol. 411, pp. 259–267, 2020.
[13] ———, “Underdetermined convolutive blind separation of sources integrating tensor factorization and expectation maximization,” Digital Signal Processing, vol. 87, pp. 145–154, 2019.
[14] S.-i. Amari, A. Cichocki, and H. H. Yang, “A new learning algorithm for blind signal separation,” in Advances in neural information processing systems, 1996, pp. 757–763.
[15] P. Paatero and U. Tapper, “Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values,” Environmetrics, vol. 5, no. 2, pp. 111–126, 1994.
[16] D. D. Lee and H. S. Seung, “Learning the parts of objects by non-negative matrix factorization,” *Nature*, vol. 401, no. 6755, pp. 788–791, 1999.

[17] D. J. MacKay and R. M. Neal, “Automatic relevance determination for neural networks,” in *Technical Report in preparation*. Cambridge University, 1994.

[18] C. Févotte and A. T. Cemgil, “Nonnegative matrix factorizations as probabilistic inference in composite models,” in *2009 17th European Signal Processing Conference*. IEEE, 2009, pp. 1913–1917.

[19] J. Flamant, S. Miron, and D. Brie, “Quaternion non-negative matrix factorization: Definition, uniqueness, and algorithm,” *IEEE Transactions on Signal Processing*, vol. 68, pp. 1870–1883, 2020.

[20] B. T. Nebgen, R. Vangara, M. A. Hombrados-Herrera, S. Kuksova, and B. S. Alexandrov, “A neural network for determination of latent dimensionality in non-negative matrix factorization,” *Machine Learning: Science and Technology*, vol. 2, no. 2, p. 025012, 2021.

[21] W. R. Hamilton, “Xr. on quaternions; or on a new system of imaginaries in algebra,” *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 33, no. 219, pp. 58–60, 1848.

[22] J. Flamant, S. Miron, and D. Brie, “A general framework for constrained convex quaternion optimization,” *arXiv preprint arXiv:2102.02763*, 2021.

[23] S. Boyd, N. Parikh, and E. Chu, “Distributed optimization and statistical learning via the alternating direction method of multipliers,” in *Foundations and Trends in Machine Learning*, vol. 3, p. 1–122, 2010.

[24] L. B. Alexandrov, S. Nik-Zainal, D. C. Wedge, S. A. Aparicio, S. Behjati, A. V. Biankin, G. R. Bignell, N. Bolli, A. Borg, A.-L. Berresen-Dale et al., “Signatures of mutational processes in human cancer,” *Nature*, vol. 500, no. 7463, pp. 415–421, 2013.

[25] B. S. Alexandrov, L. B. Alexandrov, F. L. Iliev, V. G. Stanev, and V. V. Vesselinov, “Source identification by non-negative matrix factorization combined with semi-supervised clustering,” Sep. 15 2020, uS Patent 10,776,718.

[26] R. Vangara, M. Bhattarai, E. Skau, G. Chennupati, H. Djidjev, T. Tierney, J. P. Smith, V. G. Stanev, and B. S. Alexandrov, “Finding the number of latent topics with semantic non-negative matrix factorization,” *IEEE Access*, 2021.

[27] P. J. Rousseeuw, “Silhouettes: a graphical aid to the interpretation and validation of cluster analysis,” *Journal of computational and applied mathematics*, vol. 20, pp. 53–65, 1987.

[28] J. Wishart, “The generalised product moment distribution in samples from a normal multivariate population,” *Biometrika*, pp. 32–52, 1928.

GIANCARLO SANCHEZ is currently pursuing his PhD in Applied Mathematics at Florida International University. He was an intern at Los Alamos National Laboratory for the Applied Machine Learning Summer Research program. He is currently pursuing research in machine learning algorithms for approximation of high dimensional PDEs.

ERIK SKAU received his PhD in Applied Mathematics from North Carolina State University in 2017. His research expertise includes optimization techniques for matrix and tensor decompositions. Erik is a scientist in the Information Sciences Group as Los Alamos National Laboratory.

BOIAN ALEXANDROV is a scientist at the Theoretical Division in Los Alamos National Laboratory. He has MS in Theoretical Physics, a PhD in Nuclear Engineering and second PhD in Computational Biophysics. Alexandrov is specialized in Big Data analytics, non-negative matrix and tensor factorization, unsupervised learning, and latent features extraction.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/