In this supplement file, we prove the presented results. Note that Theorem 2.3 and Corollary 2.1 follow directly from Theorem 2.1 by basic algebra, so the proof are omitted. In this paper, $C$ is a generic constant that may vary from occurrence to occurrence.

A Proof of Theorem 2.1

The key is to prove the following lemma, which the conditions are similar to those of Theorem 2.1, except for that we replace condition (2.9) in Theorem 2.1 by a slightly stronger condition:

$$\frac{(y_0, y_i)}{\|y_i\|} \geq \sqrt{1 - 1/K}, \quad \text{for all } 1 \leq i \leq n. \quad (A.1)$$

Lemma A.1 Fix $K \geq 2$, $n \geq K$, and $0 \leq m \leq K/2$. Consider the NMF problem as in (1.1), where $J_{K,m}$ and $Y$ are as in (2.7). Suppose there is a vector $y_0 \in S_{K,m}$ such that

$$\frac{(y_0, y_i)}{\|y_i\|} \geq \sqrt{1 - 1/K}, \quad \text{for all } 1 \leq i \leq n.$$

There is a $K \times K$ orthogonal matrix $Q$ such that simultaneously

- For all $1 \leq i \leq n$, $Qy_i$ falls in the first orthant of $\mathbb{R}^K$.
- $QJ_{K,m}Q'$ is non-negative.
- The NMF problem for $\Omega$ is solvable by writing $YJ_{K,m}Y' = ZPZ'$ with $Z = YQ'$ and $P = QJ_{K,m}Q'$.

Lemma A.1 is proved below.

We now explain why Theorem 2.1 follows once Lemma A.1 is proved. In Theorem 2.1 $\Omega = YJ_{K,m}Y'$, and $Y = [y_1, y_2, \ldots, y_n]$. By the assumption of Theorem 2.1

$$\frac{|(y_0, y_i)|}{\|y_i\|} \geq \sqrt{1 - 1/K}, \quad 1 \leq i \leq n.$$

Therefore, for each $1 \leq i \leq n$, there is a number $s_i \in \{-1, 1\}$ such that

$$\frac{(y_0, s_iy_i)}{\|y_i\|} \geq \sqrt{1 - 1/K}.$$

We have two cases.

- Case 1. $s_i = -1$ for all $1 \leq i \leq n$. 

• Case 2. \( s_i = 1 \) for at least one \( 1 \leq i \leq n \).

In case 1, let \( \tilde{Y} = -Y \). Theorem 2.1 follows directly if we write \( \Omega = \tilde{Y}J_{K,m}\tilde{Y} \) and then apply Lemma A.1. In case 2, the key is to show that

\[
\text{if } s_i = 1 \text{ for some } 1 \leq i \leq n, \text{ then } s_i = 1 \text{ for all } 1 \leq i \leq n. \tag{A.2}
\]

Note that once (A.2) is proved, Theorem 2.1 follows by directly applying Lemma A.1 in this case.

It remains to show (A.2). Let \( \tilde{Y} = [\tilde{y}_1, \tilde{y}_2, \ldots, \tilde{y}_n]' \) where \( \tilde{y}_i = s_iy_i \) and let

\[
\tilde{\Omega} = \tilde{Y}J_{K,m}\tilde{Y}'.
\]

It is seen that \( \tilde{\Omega} \) satisfies all conditions of Lemma A.1 so there is a \( K \times K \) orthogonal matrix \( Q \) such that \( QJ_{K,m}Q' \) is non-negative, and that \( Q\tilde{y}_i \) is non-negative for all \( 1 \leq i \leq n \).

Now, let

\[
S = \{ 1 \leq i \leq n : s_i = 1 \}.
\]

If (A.2) is not true, \( S^c \) is non-empty. Since \( \Omega \) is irreducible, we can find an \( i \in S \) and \( j \notin S \) such that \( \Omega(i,j) > 0 \). Note that

\[
\Omega(i,j) = y_i'J_{K,m}y_j = s_is_jy_i'y_j = s_is_j(Q\tilde{y}_i)'[QJ_{K,m}Q'](Q\tilde{y}_j).
\]

Therefore, on one hand, \( \Omega(i,j) > 0 \). On the other hand, since \( QJ_{K,m}Q' \), \( Q\tilde{y}_i \), and \( Q\tilde{y}_j \) are all non-negative, \( (Q\tilde{y}_i)'[QJ_{K,m}Q'](Q\tilde{y}_j) \geq 0 \). Note however \( s_is_j = -1 \). Comparing these with (A.3) gives a contradiction. The contradiction shows that \( S^c \) is empty and completes the proof.

### A.1 Proof of Lemma A.1

For \( 1 \leq k \leq K/2 \), let \( h_k \) be the vector where

\[
h_k(i) = \begin{cases} 1/\sqrt{2}, & i = 2k - 1, \\ -1/\sqrt{2}, & i = 2k, \end{cases}
\]

and let

\[
Q^{(m)} = [h_1, h_2, \ldots, h_m].
\]

Note that in the special case of \( m = 0 \), \( Q^{(m)} \) is empty.

We need the following lemma. As before, let \( e_0 \) be the \( K \)-dimensional vector \( K^{-1/2}(1, \ldots, 1)' \).

**Lemma A.2** Any \( K \)-dimensional unit-norm vector \( x \) with \( (x, e_0) \geq \sqrt{1 - 1/K} \) is non-negative.

**Proof of Lemma A.2** Without loss of generality, we assume the first \( k \) entries of \( x \) are non-negative, and the remaining \( (K - k) \) entries are strictly negative. All we need to show is \( k = K \). If \( k < K \), then by definition and Cauchy-Schwartz inequality,

\[
(x, e_0) < K^{-1/2} \sum_{i=1}^{k} x_i \leq K^{-1/2} \sqrt{k \sum_{k=1}^{k} x_i^2} < \sqrt{k/K},
\]

where we have used

\[
\sum_{k=1}^{k} x_i^2 < \|x\|^2 = 1.
\]

This contradicts with \( (x, e_0) \geq \sqrt{1 - 1/K} \), and so the claim follows.

We also need the following lemma.

**Lemma A.3** Fix \( K \geq 2 \), \( 0 \leq m \leq K/2 \), and a unit-norm vector \( y_0 \in S_m \). There exists a \( K \times (K - m) \) matrix \( B \) such that

\[
B'B = I_{K-m}, \quad B'Q^{(m)} = 0, \quad \text{and} \quad [B, Q^{(m)}]'e_0 = y_0.
\]
Proof of Lemma A.3} The proof for the case of \( m = 0 \) is trivial, so we only consider the case of \( 1 \leq m \leq K/2 \). It is seen that \( \{h_1, h_2, \ldots, h_m\} \) expands a \( m \)-dimensional sub-space of \( \mathbb{R}^K \). By basic algebra, we can always expand \( \{h_1, h_2, \ldots, h_m\} \) to a full set of orthogonal basis vectors of \( \mathbb{R}^K \), denoted by
\[
\{q_1, q_2, \ldots, q_{K-m}, h_1, h_2, \ldots, h_m\}.
\]
Note that since \( e_0 \perp \text{span}\{h_1, h_2, \ldots, h_m\} \) (i.e., the linear spanned by \( h_1, \ldots, h_m \)), there is a \((K - m)\) vector \( \delta_1 \) such that
\[
e_0 = \sum_{i=1}^{K-m} \delta_i q_i = Q_0 \delta,
\]
where \( Q_0 = [q_1, q_2, \ldots, q_{K-m}] \).
Since \( Q_0 Q^{(m)} = 0 \) and \( Q_0' Q_0 = I_{K-m} \),
\[
\|\delta\| = \|Q_0 \delta\| = \|e_0\| = 1.
\]
At the same time, since \( y_0 \in S_m \), by definition, there is a unit-norm \((K - m)\) dimensional vector \( a \) such that
\[
y_0 = (a, 0, \ldots, 0)',
\]
and there is a \((K - m) \times (K - m)\) orthogonal matrix \( U \) such that
\[
U' \delta = a.
\]
Let \( B = Q_0 U \).
We have
\[
\begin{align*}
B' B &= U' Q_0 Q_0 U = I_{K-m}, \\
B' Q^{(m)} &= U' Q_0 Q^{(m)} = 0, \\
[B, Q^{(m)}]' e_0 &= \left[ U' Q_0 Q_0 \delta \right] = \left[ \begin{array}{c} a \\ 0 \end{array} \right] = y_0.
\end{align*}
\]
This proves Lemma A.3.

We now come back to prove Lemma A.1. Let \( B \) by any matrix in Lemma A.3, and let \( Q = [B, Q^{(m)}] \).
By Lemma A.3, \( Q'e_0 = y_0 \), and so \( Qy_0 = e_0 \).
At the same time, for all \( 1 \leq i \leq n \),
\[
(y_i', y_0) = (Q y_i, Q y_0) = (Q y_i, e_0), \quad \|y_i\| = \|Q y_i\|.
\]
Combining these with Lemma A.1,
\[
\begin{align*}
\frac{(Q y_i, e_0)}{\|Q y_i\|} &= \frac{(y_i, y_0)}{\|y_i\|} \geq \sqrt{1 - 1/K}.
\end{align*}
\]
It follows from Lemma A.2 that
\[
\text{for all } 1 \leq i \leq n, Q y_i \text{ is a non-negative vector. (A.4)}
\]
At the same time, since \([B, Q^{(m)}] \) is an orthogonal matrix, \( B B' + Q^{(m)}(Q^{(m)})' = I_K \), and so
\[
Q J_{K,m} Q' = B B' - Q^{(m)}(Q^{(m)})' = I_K - 2Q^{(m)}(Q^{(m)})'.
\]
By direct calculations, \( 2Q^{(m)}(Q^{(m)})' \) is a \( K \times K \) block-wise diagonal matrix, where the first \( m \) diagonal blocks are
\[
\begin{bmatrix}
1 & -1 \\
-1 & 1
\end{bmatrix};
\]
note that except for these \( m \) diagonal blocks, the matrix is 0 elsewhere. Therefore, \( Q J_{K,m} Q' = I_K - 2Q^{(m)}(Q^{(m)})' \) and is non-negative. (A.5)
Combining this with (A.4)-(A.5) gives Lemma A.1.

3
B Proof of Theorem 2.2

By Theorem 2.1 we only need to check when \( K = 2 \),

- we must have \( m \leq K/2 \).
- the condition (2.9) always holds without any extra conditions.

Since \( \Omega \) has at least one positive eigenvalues, so we can only have \( m = 0 \) or \( m = 1 \). Therefore, we must have \( 0 \leq m \leq K/2 \). This checks the first bullet point above.

We now consider the second bullet point. In detail, let \( e_1 = (1, 0)' \) as before. Let \( 0 \leq \theta_i < 2\pi \) be the angle from \( e_1 \) to \( y_i \) counterclockwise, and let \( \theta_{\text{min}} \) and \( \theta_{\text{max}} \) be the smallest and largest values of all \( \theta_i \). Recall that we can only have \( m = 0 \) or \( m = 1 \). Also, we assume the following in the lemma.

- When \( m = 0 \), \( y_0 \) is the unit vector where the angle from \( e_1 \) to \( y_0 \) is \( (\theta_{\text{max}} + \theta_{\text{min}})/2 \), counterclockwise.
- When \( m = 1 \), \( y_0 = (1, 0) \).

We now consider the case of \( m = 0 \) and \( m = 1 \) separately.

Consider the case of \( m = 0 \) first. Note that all \( y_i \) fall in the sector (with the apex at 0) bounded by two rays, where the angle from \( e_1 \) to the two rays (counterclockwise) are \( \theta_{\text{min}} \) and \( \theta_{\text{max}} \), respectively. Recall that \( \Omega = Y J_{K,m} Y' \). When \( m = 0 \), \( J_{K,m} = I_K \) and \( \Omega = Y Y' \). Therefore, for \( 1 \leq i,j \leq n \),
\[
(y_i, y_j) = \Omega(i,j) \geq 0.
\]
This says the angle of the sector is no bigger than \( \pi/2 \):
\[
\theta_{\text{max}} - \theta_{\text{min}} \leq \pi/2.
\]
If we take \( y_0 \) as above (so the angle between \( e_1 \) and \( y_0 \) is \( (\theta_{\text{max}} + \theta_{\text{min}})/2 \)), it follows that for each \( 1 \leq i \leq n \), the angle between \( y_0 \) and \( y_i \) is no bigger than \( \pi/4 \). Therefore,
\[
\frac{|(y_i, y_0)|}{\|y_i\|} \geq \sqrt{1/2},
\]
and (2.9) holds.

Consider the case of \( m = 1 \). In this case,
\[
J_{K,m} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.
\]
For each \( 1 \leq i \leq n \), write \( y_i = (a_i, b_i)' \). Since \( \Omega = Y J_{K,m} Y' \) and \( \Omega(i,i) > 0 \),
\[
0 \leq \Omega(i,i) = y'_i J_{K,m} y_i = a_i^2 - b_i^2,
\]
and so
\[
|b_i| \leq |a_i|.
\]
If we take \( y_0 = e_1 \) as above, then for all \( 1 \leq i \leq n \),
\[
\frac{|(y_i, y_0)|}{\|y_i\|} = \frac{|a_i|}{\sqrt{a_i^2 + b_i^2}} \geq \sqrt{1/2},
\]
and the claim follows directly.

C Proof of Theorem 2.4

It is sufficient to justify

- \( y_0 = e_1 \) (and so especially \( y_0 \in S_m \)).
- condition (2.10) is equivalent to
\[
\sum_{k=1}^{K-1} |\lambda_{k+1}| \cdot r_i^2(k) \leq \lambda_1/(K-1),
\]
where \( r_i(k) = \xi_{k+1}(i)/\xi_1(i), 1 \leq k \leq K - 1, 1 \leq i \leq n \).
We now justify both bullet points. Consider the first bullet point first. Note that in this case, 
\[ \Omega = YJ_{K,m}Y', \quad Y = [\xi_1, \xi_2, \ldots, \xi_K]D^{1/2}, \]
and 
\[ y^{(w)} = Y'w, \]
with 
\[ w = c\xi_1, \]
where \( c = 1/\|\xi_1\|_1 \), and 
\[ D = \text{diag}(|\lambda_1|, |\lambda_2|, \ldots, |\lambda_K|). \]
By Perron’s theorem \( \Pi \), \( \lambda_1 > 0 \) and all entries of \( \xi_1 \) are positive. Also, note that \( \xi_1 \perp \xi_k \) for all \( 2 \leq k \leq K \). It follows 
\[ y^{(w)} = cD^{1/2}[\xi_1, \xi_2, \ldots, \xi_K]'\xi_1 = c\sqrt{\lambda_1}e_1, \]
so 
\[ y_0 = y^{(w)}/\|y^{(w)}\|_1 = e_1. \]
This justifies the first bullet point.
Consider the second bullet point. By definition, 
\[ \beta^{(w)}_i = \Omega w = c[\xi_1, \xi_2, \ldots, \xi_K]D^{1/2}J_{K,m}D^{1/2}[\xi_1, \xi_2, \ldots, \xi_K]'\xi_1 = c\lambda_1\xi_1, \]
It follows 
\[ w'\tilde{\Omega}w = w'\beta^{(w)} = c^2\lambda_1\xi_1^2 = c^2 \lambda_1. \]
At the same time, by definition, 
\[ \tilde{\Omega} = YY' = \sum_{k=1}^{K} |\lambda_k|\xi_k\xi_k', \]
so 
\[ \tilde{\Omega}(i, i) = \sum_{k=1}^{K} |\lambda_k|\xi_k^2(i). \]
Combining these, 
\[ \frac{|\beta^{(w)}_i|}{\sqrt{(w'\tilde{\Omega}w)\tilde{\Omega}(i, i)}} = \frac{c\lambda_1\xi_1(i)}{\sqrt{c^2\lambda_1 \sum_{k=1}^{K} |\lambda_k|\xi_k^2(i)}} = \frac{1}{\sqrt{1 + \sum_{k=1}^{K-1} (|\lambda_{k+1}|/\lambda_1)[\xi_{k+1}(i)/\xi_1(i)]^2}}. \]
Therefore, condition \( 2.10 \) of Theorem \( 2.3 \) reduces to 
\[ \frac{1}{\sqrt{1 + \sum_{k=1}^{K-1} (|\lambda_{k+1}|/\lambda_1)[\xi_{k+1}(i)/\xi_1(i)]^2}} \geq \sqrt{1 - 1/K}, \]
which is equivalent to \( C.6 \). This completes the proof.

\section{Proof of Theorem \( 2.5 \)}

It is sufficient to show 
- \( a_m = (K - 1) \) when \( m = K - 1 \). 
- \( a_m = 1 \) when \( m \leq K/2 \). 
- The NMF problem is solvable for \( \Omega \) if \( r_i'r_iD_0r_i \leq 1/[a_m(K - 1)] \) for all \( 1 \leq i \leq n \).
Consider the first bullet point. By the definition of $Q$, for all $Q \in \mathcal{Q}$, the first column of $Q$ is $K^{-1/2}e_1$. Therefore, when $m = (K - 1)$, for any such $Q$, we have

$$Q = [e_0, Q^{(m)}].$$

By basic algebra,

$$2Q^{(m)}(Q^{(m)})' - I = 2(I_K - e_0e_0') - I_K = I - (2/K)1_K1_K',$$

where the maximum entry of the matrix is $(K - 2)/K$. By the definition, in this case, $a_m = 1 + K \cdot [(K - 2)/K] = (K - 1)$. This proves the first bullet point.

Consider the second bullet point. The goal is to show $a_m = 1$ for all $0 \leq m \leq K/2$. The case of $m = 0$ is trivial, so we suppose $m \geq 1$. First, we show

$$a_m \leq 1, \quad \text{if } 1 \leq m \leq K/2. \quad \text{(D.7)}$$

For $1 \leq k \leq K/2$, let $h_k$ be the vector where

$$h_k(i) = \begin{cases} 
1/\sqrt{2}, & i = 2k - 1, \\
-1/\sqrt{2}, & i = 2k.
\end{cases}$$

Construct $Q$ such that

- the first column is $e_0$,
- the last $m$ columns are $h_1, h_2, \ldots, h_m$, respectively.

For such a $Q$, by basic algebra, we have

- $Q \in \mathcal{Q}$,
- for the matrix $2Q^{(m)}(Q^{(m)})' - I_K$, none of the entries is positive.

The second item is true because, by construction, $2Q^{(m)}(Q^{(m)})'$ is a blockwise diagonal matrix, where all nonzero entries appear in $2 \times 2$ diagonal blocks of

$$\begin{bmatrix} 1 & -1 \\
-1 & 1 \end{bmatrix}.$$

This can be either checked by direct calculations or quoted from the proof of Lemma A.1. Combining these with the definition, $a_m \leq 1$, and (D.7) follows.

Next, we show

$$a_m \geq 1, \quad \text{if } 1 \leq m \leq K/2. \quad \text{(D.8)}$$

We claim that in $m$ dimensional space, the maximum number of unit-norm vectors $r_i$ where the pairwise inner product are all (strictly) negative is no greater than $m + 1$. We prove this by mathematical induction. Note that this holds trivially when $m = 1$. Now, suppose this holds for $m \leq m_0$, we show the claim continues to hold for $m = m_0 + 1$. If the claim is not true for $m = m_0 + 1$, then there are $m_0 + 3$ different vectors $r_1, r_2, \ldots, r_{m_0+3}$ of $(m_0 + 1)$ dimension where the pairwise inner product is strictly negative. Since all the pairwise inner products remain unchanged if we rotate these vectors by the same orthogonal matrix, we assume $r_{m_0+3} = (1, 0, \ldots, 0)'$ without loss of generality. Write

$$r_k = (a_k, s_k)', \quad 1 \leq i \leq m_0 + 3.$$ 

where $a_k$ is the first entry of $r_k$ and $s_k$ is $m_0$-dimensional sub-vector of $r_k$. Since for all $1 \leq k \leq m_0 + 2$,

$$(r_k, r_{m_0+3}) < 0,$$

we must have

$$a_k < 0, \quad \text{for all } 1 \leq k \leq m_0 + 2.$$ 

Therefore, for all $1 \leq k, \ell \leq K$ and $k \neq \ell$,

$$(s_k, s_\ell) = (r_k, r_\ell) - a_k a_\ell < 0.$$
We now consider the last bullet point. Recall that $\lambda$ where since $e$ where $e$ where $e$ where $e$.

At the same time, note that $D$. Also, recall that $Q$ such that $Q$ such that $Q$ such that $Q$. Therefore, we need to show is that, there is an orthogonal matrix $Q$ such that $Q$ such that $Q$ such that $Q$ such that $Q$.

Let $Q_{\tilde{J},m}$ be the $K \times K$ diagonal matrix

$$
\text{diag}(a_m, 1, \ldots, 1, -1, \ldots, -1).
$$

That is, the first diagonal entry is $a_m$, the next $(K - m - 1)$ diagonal entries are $-1$, and all other diagonal entries are $-1$. It follows from spectral decomposition that

$$
\Omega = \Xi \cdot \text{diag}(\lambda_1, \ldots, \lambda_K) \cdot \Xi' = \tilde{Y} Q_{\tilde{J},m} \tilde{Y}'.
$$

Therefore, we need to show is that, there is an orthogonal matrix $Q$ such that both matrix $\tilde{Y} Q'$ and $Q \tilde{J}_{K,m} Q'$ are non-negative.

To show the claim, note that by definition, there is a matrix $Q \in Q$ such that

$$
\max_{1 \leq i, j \leq K} \{ H(i, j) : H = 2Q^{(m)}(Q^{(m)})' - I_k \} = (a_m - 1)/K,
$$

Recall that by definition, the first column of any matrix $Q \in Q$ is $e_0$. It follows

$$
Q \tilde{J}_{K,m} Q' = (a_m - 1)e_0 e_0' + I_K - 2Q^{(m)}(Q^{(m)})',
$$

where since $e_0 = K^{-1} 1_K$, the right hand side is

$$
\frac{(a_m - 1)}{K} 1_K 1_K' - [2Q^{(m)}(Q^{(m)})' - I_K].
$$

Combining (D.12), (D.14) shows that the matrix $Q \tilde{J}_{K,m} Q'$ is non-negative.

At the same time, note that

$$
Q e_1 = e_0,
$$

where $e_1 = (1, 0, \ldots, 0)'$. Recall that $r_i \in \mathbb{R}^{K-1}$ and

$$
r_i(k) = \xi_{k+1}(i)/\xi_1(i), \quad 1 \leq k \leq K - 1, 1 \leq i \leq n.
$$

Also, recall that $D_0 = \text{diag}(|\lambda_2|, \ldots, |\lambda_K|)$. By (D.10),

$$
\tilde{y}_i \propto [1, r_i'] D^{1/2} = [\sqrt{\lambda_1/a_m}, r_i D_0^{1/2}].
$$

By basic algebra, for any $1 \leq i \leq n$,

$$
\frac{|\langle \tilde{y}_i, e_1 \rangle|}{\|\tilde{y}_i\|} = \frac{1}{\sqrt{1 + (a_m/\lambda_1) r_i^2 D_0 r_i^2}}.
$$
where by the condition of Theorem 2.5

\[ a_m r_i' D_0 r_i \leq \lambda_1/(K - 1). \]  

(D.15)

It follows that

\[ \frac{|(\tilde{y}_i, e_1)}{\|\tilde{y}_i\|} \geq \sqrt{1 - 1/K}. \]

Combining this with \( Qe_1 = e_0 \),

\[ \frac{|(Q\tilde{y}_i, e_0)}{\|Q\tilde{y}_i\|} = \frac{|(Q\tilde{y}_i, e_1)|}{\|Q\tilde{y}_i\|} = \frac{|(\tilde{y}_i, e_1)|}{\|\tilde{y}_i\|} \geq \sqrt{1 - 1/K}. \]

By Lemma 2.4, \( Q\tilde{y}_i \) falls in the first orthant. Since this holds for all \( 1 \leq i \leq n \), \( Y'Q' \) is non-negative.

This finishes the proof of Theorem 2.4.

### E  Proof of Lemma 3.1 and Lemma 3.2

Recall that \( D = \text{diag}(|\lambda_1|, |\lambda_2|, \ldots, |\lambda_K|) \) and \( D_0 = \text{diag}(|\lambda_2|, \ldots, |\lambda_K|) \). Similarly, define

\[ H = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_K), \quad \text{and} \quad H_0 = \text{diag}(\lambda_2, \lambda_3, \ldots, \lambda_K). \]

It is sufficient to show

- If \( \lambda_1(G) \leq c_0 \lambda_K(G) \) for a constant \( c_0 > 0 \), then as \( n \to \infty \), the two conditions of \( \max_{2 \leq k \leq K} \{ |\lambda_k/\lambda_1| \} \to 0 \) and \( \max_{2 \leq k \leq K} \{ |\lambda_k(P)/\lambda_1(P)| \} \to 0 \) are equivalent.
- If \( \lambda_1(G) \leq c_0 \lambda_K(G) \), then for all \( 1 \leq i \leq n \), \( \|r_i\| \leq CM(\Omega) \).
- \( B = \text{diag}(b_1)[1_K, V] \), \( P = BH' \), \( P(k,k) = b_1^2(k)/[\lambda_1 + \lambda_k H_0 v_k] \), \( 1 \leq k \leq K \), and \( b_1 \) is an eigenvector of \( PG \) with \( \lambda_1 \) being the corresponding eigenvalue.
- If \( (Y, P) \) are non-negative, then first, \( PG \) is an irreducible non-negative matrix and \( b_1 \) is the Perron vector (so all entries of \( b_1 \) are strictly positive). Second, all rows of \( r_i \) lives with a simplex with \( v_1, v_2, \ldots, v_K \) being the vertices. Last, if \( \lambda_1(G) \leq c_0 \lambda_K(G) \), then \( M(\Omega) \leq \max_{1 \leq k \leq K} \{ \|b_1\|/\|b_1(k)\| \} \).

Consider the first bullet point. It is sufficient to show

- if \( \max_{2 \leq k \leq K} \{ |\lambda_k/\lambda_1| \} \to 0 \) then \( \max_{2 \leq k \leq K} \{ |\lambda_k(P)/\lambda_1(P)| \} \to 0 \).
- if \( \max_{2 \leq k \leq K} \{ |\lambda_k(P)/\lambda_1(P)| \} \to 0 \) then \( \max_{2 \leq k \leq K} \{ |\lambda_k/\lambda_1| \} \to 0 \).

Consider the first item. Recall that \( B = G^{-1/2}Q \) for a \( K \times K \) orthogonal matrix \( Q \), and \( G \) is a positive definite matrix satisfying

\[ \lambda_1(G) \leq C \lambda_K(G). \]  

(E.16)

Recall that \( B = [b_1, b_2, \ldots, b_K] \). Write \( Q = [q_1, q_2, \ldots, q_K] \) and let

\[ G = UDU' \]

be the spectral decomposition of \( G \), where \( U \in \mathbb{R}^{K, K} \) is orthogonal and

\[ D = \text{diag}(\lambda_1(G), \lambda_2(G), \ldots, \lambda_K(G)). \]

It follows

\[ B = G^{-1/2}Q = UD^{-1/2}U'Q, \quad b_1 = UD^{-1/2}U'q_1. \]

Therefore, by (E.16),

\[ \|B\| = \|G^{-1/2}Q\| = \|G^{-1/2}\| \leq 1/\sqrt{\lambda_K(G)}, \]  

(E.17)

and for \( 1 \leq k \leq K \),

\[ \|b_k\|^2 = \|UD^{-1/2}U'q_k\|^2 = q_k^2 UD^{-1}U'q_k \cong (1/\lambda_K(G))\|U'q_k\| = 1/\lambda_K(G). \]  

(E.18)

Now, using the first bullet point, we write

\[ P = BH' = B \cdot \text{diag}(\lambda_1, \ldots, \lambda_K) \cdot B = P_0 + P_1, \]
where

\[ P_0 = B \cdot \text{diag}(\lambda_1, 0, \ldots, 0) \cdot B', \quad P_1 = B \cdot \text{diag}(0, \lambda_2, \ldots, \lambda_K) \cdot B'. \]

Note that

\[ P_0 = \lambda_1 b_1 b_1', \]

where \( \lambda_1 > 0 \). Therefore,

\[ \lambda_k(P_0) = \begin{cases} \lambda_1 ||b_1||^2, & k = 1, \\ 0, & 2 \leq k \leq K. \end{cases} \quad (E.19) \]

By Weyl’s theorem \( \Pi \), for any \( 1 \leq k \leq K \),

\[ |\lambda_k(P) - \lambda_k(P_0)| \leq \|P_1\|, \]

where by (E.18),

\[ \|P_1\| \leq \|B\| \cdot \max_{\{2 \leq k \leq K\}} \{|\lambda_k|\}. \]

Therefore,

\[ |\lambda_k(P) - \lambda_k(P_0)| \leq (1/\lambda_K(G)) \cdot \max_{\{2 \leq k \leq K\}} \{|\lambda_k|\}. \quad (E.20) \]

Combining (E.19)–(E.20) with (E.18),

\[ \lambda_1(P) \geq \lambda_1 ||b_1||^2 - (1/\lambda_K(G)) \cdot \max_{\{2 \leq k \leq K\}} \{|\lambda_k|\} \geq (1/\lambda_K(G))(C\lambda_1 - \max_{\{2 \leq k \leq K\}} \{|\lambda_k|\}), \]

and

\[ |\lambda_k(P)| \leq (1/\lambda_K(G)) \cdot \max_{\{2 \leq k \leq K\}} \{|\lambda_k|\}, \quad 2 \leq k \leq K. \]

Therefore, if \( (1/\lambda_1) \max_{\{2 \leq k \leq K\}} \{|\lambda_k|\} \to 0 \),

then \( \lambda_1(P) > 0 \) and

\[ (1/\lambda_1(P)) \max_{\{2 \leq k \leq K\}} \{|\lambda_k(P)|\} \to 0. \]

This proves the first item.

The proof of the second item is similar, so we keep it short. Let \( P = UDU' \) be the spectral decomposition of \( P \), where \( U \) is a \( K \times K \) orthogonal matrix, and

\[ D = \text{diag}(\lambda_1(P), \ldots, \lambda_2(P)). \]

Combining these with \( P = BHB' \) and \( B = G^{-1/2}Q \),

\[ H = B^{-1}U'DU'(B^{-1})' = (Q'G^{1/2}U')D(UG^{1/2}Q). \]

Recall that

\[ H = \text{diag}(\lambda_1, \ldots, \lambda_K). \]

It follows

\[ \text{diag}(\lambda_1, \ldots, \lambda_K) = (Q'G^{1/2}U') \cdot \text{diag}(\lambda_1(P), \ldots, \lambda_K(P)) \cdot (UG^{1/2}Q). \]

The remaining part of the proof is similar, so we skip it. This completes the proof of the first bullet point.

Consider the second bullet point. Recall that

\[ \Xi = [\xi_1, \ldots, \xi_K] = YB, \quad r_i(k) = \xi_{k+1}(i)/\xi_1(k), 1 \leq k \leq K - 1, 1 \leq i \leq n, \]

and that \( Y = [y_1, y_2, \ldots, y_n]' \) and \( B = [b_1, b_2, \ldots, b_n] \). We have

\[ r_i(k) = y_i'b_{k+1}/y_i'b_1. \]

By definition, \( y_i'b_1 = ||y_i|| ||b_1|| \cos(\alpha_i) \), so

\[ |r_i(k)| \leq \frac{||y_i|| ||b_{k+1}||}{||y_i|| ||b_1|| |\cos(\alpha_i)|} = \frac{1}{||b_{k+1}||} \cdot \frac{||b_{k+1}||}{||b_1||}. \]
By (E.18),
\[
\frac{\|b_{k+1}\|}{\|b_1\|} \asymp 1.
\]

Therefore,
\[
|r_1(k)| \leq C/|\cos(\alpha_i)| \leq CM(\Omega).
\]
Since $K$ is finite, the claim follows. This proves the second bullet point.

Consider the third bullet point. The first item follows directly by definition of $V$. For the second item, recall that
\[
\Omega = \Xi H\Xi' = YPY', \quad \text{and} \quad \Xi = YB.
\]
Combining these gives
\[
YBB'Y' = YPY',
\]
and the claim follows since $Y$ is full-rank by our assumption.

For the third item, we combine the first two items, and it follows that
\[
P = BHB' = \text{diag}(b_1)[1_K, V]H[1_K, V]'\text{diag}(b_1) = \text{diag}(b_1)[\lambda_1 1_K 1_K' + VH0V']\text{diag}(b_1).
\]
Recall that
\[
V = [v_1, v_2, \ldots, v_K]'.
\]
For any $1 \leq k \leq K$, comparing the $k$-th diagonal entry of the two matrices, $P$ and $\lambda_1 b_1 + VH0V'$, it follows that
\[
P(k, k) = b_1^2(k)[\lambda_1 + v_k' H0v_k].
\]
This proves the third item.

Consider the last item. By basic algebra, the set of all nonzero eigenvalues of the matrix $YPY'$ are the same as the set of all non-zero eigenvalues of $PY'Y$ or $PG$. Since $PG$ is non-singular, the eigenvalues are $\lambda_1, \ldots, \lambda_K$. By $\Xi = [\xi_1, \ldots, \xi_K] = YB$ and $B = [b_1, b_2, \ldots, b_K]$, we have
\[
\xi_k = Yb_k.
\]
Since $\xi_1$ is the eigenvector corresponding to $\lambda_k$, so on one hand,
\[
\lambda_k \xi_k = \lambda_k Yb_k,
\]
and on the other hand,
\[
\lambda_k \xi_k = \Omega \xi_1 = YPY' \xi_k = YPY' Yb_k = YPGb_k.
\]
Combining these and noting that $Y$ is full rank,
\[
PGb_k = \lambda_k b_k.
\]
Therefore, $b_1, b_2, \ldots, b_K$ are singular vectors corresponding to $\lambda_1, \ldots, \lambda_K$. This proves the last item and completes the proof of the second bullet point.

Consider the last bullet point. Consider the first item first. Since $Y$ and $P$ are non-negative, $PG$ is non-negative. Also, the matrix $PG$ is also irreducible since $\Omega$ is irreducible. Since $b_1$ is the Perron vector of $PG$, all of entries are strictly positive.

Consider the second item. Fix $1 \leq i \leq n$. Since $\Omega$ is irreducible, $y_i \neq 0$. Introduce a weight vector $w_i \in \mathbb{R}K$ by $w_i(k) = \frac{y_i(k)b_1(k)}{y_i'b_1}$. Consider the second item. Fix $1 \leq i \leq n$. Since $\Omega$ is irreducible, $y_i \neq 0$. Introduce a weight vector $w_i \in \mathbb{R}K$ by $w_i(k) = \frac{y_i(k)b_1(k)}{y_i'b_1}$. Consider the second item. Fix $1 \leq i \leq n$. Since $\Omega$ is irreducible, $y_i \neq 0$. Introduce a weight vector $w_i \in \mathbb{R}K$ by $w_i(k) = \frac{y_i(k)b_1(k)}{y_i'b_1}$. Recall that $V(j, k) = b_{k+1}(j)/b_1(j)$, $1 \leq k \leq K - 1, 1 \leq j \leq K$. We have
\[
r_i(k) = \frac{y_i'b_{k+1}}{y_i'b_1} = \sum_{j=1}^{K} w_i(j)(b_{k+1}(j)/b_1(j)) = \sum_{j=1}^{K} w_i(j)V(j, k).
\]
Therefore,
\[
r_i = w_i' V,
\]
and $r_i$ is a convex linear combination of the $K$ rows of $V$. This says $r_i$ fall within the simplex with $v_1, v_2, \ldots, v_K$ being the vertices. This proves the second item.
Consider the last item. Note that for any \( 1 \leq i \leq n, y_i \neq 0 \). Otherwise, if \( y_i = 0 \) for some \( i \), then \( \Omega(i, j) = \Omega(j, i) \) for all \( j \), and \( \Omega \) is reducible. Now,

\[
\cos(\alpha_i) = \frac{(y_i, b_1)}{\|y_i\|_1 b_1}
\]

Without loss of generality, assume \( \|y_i\|_1 = 1 \) since \( y_i \) is non-negative vector. It follows

\[
\|y_i\|_1 = \sqrt{\sum_{k=1}^{K} y_i^2(k)} \leq \sqrt{\|y_i\|_1^2} = 1.
\]

Therefore,

\[
|\cos(\alpha_i)| \geq \min_{1 \leq k \leq K} \{b_1(k)\|y_i\|_1/\|y_i\|_1\|b_1\|} \geq \min_{1 \leq k \leq K} \{b_1(k)/\|b_1\|\}.
\]

This proves the last item and completes the proof of the last bullet point.

F Comments on how to find \((Q, y_0)\) in Theorem 2.1 numerically

Given \( \Omega = YJ_{K,m}Y' \) as in Theorem 2.1, an interesting question is how to construct a \( Q \) numerically when condition (2.9) of Theorem 2.1 holds for some \( y_0 \). The proof of Theorem 2.1 relies on a specific construction of \( Q \) (see the proof of Theorem 2.1 especially Lemma A.3 for details) as follows.

- Let \( Q_1 \) be the \( K \times K \) orthogonal matrix with the form of

\[
Q_1 = \begin{bmatrix} Q_0 & \mathbf{0} \\ \mathbf{0} & I_m \end{bmatrix}, \quad \text{the first row of } Q_0 \text{ is } (y_0(1), \ldots, y_0(K-m)).
\]

- Let \( Q_2 \) be the \( K \times K \) orthogonal matrix with the form of

\[
Q_1 = [e_0, q_1, \ldots, q_{K-m-1}, h_1, h_2, \ldots, h_m], \quad e_0 = (1/\sqrt{K})(1, 1, \ldots, 1)',
\]

where for \( k = 1, 2, \ldots, m,

\[
h_k(i) = \begin{cases} 1/\sqrt{2}, & i = 2k - 1, \\ -1/\sqrt{2}, & i = 2k, \\ 0, & \text{otherwise}. \end{cases}
\]

- Let \( Q = Q_2Q_1 \).

Therefore, numerically, all remains is to decide the remaining rows of \( Q_0 \) and the remaining columns of \( Q_2 \), both can be solved quickly by basic algebra, since \( K \) is usually small.

A related question is how to check whether there is a vector \( y_0 \in S_{K,m} \) such that condition (2.9) of Theorem 2.1 holds. Recalling that \( Y = [y_1, y_2, \ldots, y_n]' \), possible candidate are \( y_i^*, 1 \leq i \leq n \), where

\[
y_i^*(k) = \begin{cases} y_i(k), & 1 \leq k \leq K - m, \\ 0, & k > K - m. \end{cases}
\]

The \( y_i^* \) constructed this way belong to \( S_{K,m} \). Now, without loss of generality, assume \( y_i^* \neq 0 \) for all \( 1 \leq i \leq n \). For each \( 1 \leq s \leq n \), let

\[
U_s = \max_{1 \leq i \leq n} \{(y_i^*, y_i)/\|y_i\|_1\|y_i^*\|\}.
\]

If \( \max_{1 \leq s \leq n} \{U_s\} > 1 - 1/\sqrt{K} \), then condition (2.9) does not hold and it is unclear if the NMF problem is solvable for \( \Omega \). If \( \max_{1 \leq s \leq n} \{U_s\} \geq 1 - 1/\sqrt{K} \), then condition (2.9) holds with \( y_0 = y_k^* \), and the NMF problem is solvable.

References