Obviously, (2) is still a nonnegative least squares (NNLS) problem. The objective function in (2) can be equivalently rewritten as

\[ \min_{A^{(0)},\ldots,A^{(N)}} \frac{1}{2} \|X - [A^{(1)}, \ldots, A^{(N)}]\|_F^2 \]

\[ + \sum_{n=1}^{N} \frac{\alpha_n}{2} \|A^{(n)} - A^{(n-1)}\|_F^2 \]

s.t. \( A^{(n)} \geq 0 \) for \( n = 1, \ldots, N \).

The NCP using proximal algorithm is

\[ A_k^{(n)} = \arg \min_{A \succeq 0} \left\{ \frac{1}{2} \|X_{(n)} - A^{(n)} (B_k^{(n)})^T\|_F^2 \right\} \]

\[ + \frac{\alpha_n}{2} \|A^{(n)} - A^{(n-1)}\|_F^2 \],

where \( B_k^{(n)} = (A_k^{(N)} \otimes \cdots \otimes A_k^{(n+1)} \otimes \cdots \otimes A_k^{(1)}) \).

### Mathematical model

#### Nonnegative Least Squares

The objective function in (2) can be equivalently rewritten as

\[ J_2 = \frac{1}{2} \left\| \frac{X_{(k)}}{\sqrt{\alpha_k}} (A_k^{(n)})^T - \frac{B_k^{(n)}}{\sqrt{\alpha_k}} (A^{(n)})^T \right\|_F^2. \]

Obviously, (2) is still a nonnegative least squares (NNLS) problem. Therefore, we employ the block principal pivoting (BPP) method [2] to solve the subproblem in (2).

### Experimental & Results

#### Data description:

- We synthesized a fourth-order nonnegative tensor by \( X_{\text{syn}} = [S^{(4)}, A^{(2)}, A^{(4)}] \in \mathbb{R}^{100 \times 100 \times 100 \times 100}, \) in which \( S^{(4)} \in \mathbb{R}^{100 \times 7} \) is the signal matrix, and \( A^{(2)}, A^{(4)} \in \mathbb{R}^{100 \times 7}, A^{(4)} \in \mathbb{R}^{5 \times 7} \) are random matrices in uniform distribution.
- We utilized a set of real preprocessed fourth-order event-related potential (ERP) tensor data \( X_{\text{ERP}} \in \mathbb{R}^{7 \times 7 \times 60 \times 142} \) (channel \( \times \) frequency \( \times \) time \( \times \) subject-group = 9 \( \times \) 71 \( \times \) 60 \( \times \) 42).

#### Experimental setting:

- The factor matrices were initialized using random numbers by command \( \text{max}(0, \text{randn}(L, R)) \).
- The number of components are 7 and 40 for the synthetic and real data.
- The values of 1, 1e-2, 1e-4, 1e-6, and 1e-8 were tested for \( \alpha_n \).

### Algorithm

```
Algorithm 1: PROX-BPP with different \( \alpha_n \)
Input: \( A^{(n)}, n = 1, \ldots, N \)
Output: \( A_n^{(n)}, n = 1, \ldots, N \)
1. Initialize \( A^{(n)} \in \mathbb{R}^{L \times R}, n = 1, \ldots, N \), using random numbers;
2. repeat
3. for \( n = 1 \) to \( N \) do
4. Compute MTTKRP \( X_{\text{ERP}} B_n^{(n)} \) and \( (B_n^{(n)})^T B_n^{(n)} \);
5. \( B_n^{(n)} B_n^{(n)} = B_n^{(n)} B_n^{(n)} - \alpha_n J_n^{(n)} \)
6. \( X_n B_n^{(n)} = X_n B_n^{(n)} + \alpha_n A_n^{(n)} \)
7. Update factor \( A_n^{(n)} \) based on (2) using BPP:
8. \( A_n^{(n)} = \arg \min J_2 (A^{(n)}) \)
9. end
10. until some termination criterion is reached;
11. return \( A_n^{(n)}, n = 1, \ldots, N \).
```

### Remarks

- The choice of parameter \( \alpha_n \) for PROX-BPP is said to be related to the noise level in the data. Surprisingly, our method is very robust with different \( \alpha_n \) values. We suggest to select 1e-2 \( \leq \alpha_n \leq 1e-4 \).
- Our proposed NCP method using proximal algorithm can be combined with many other NNLS algorithms.

### References