Analysis on Non-negative Factorizations and Applications

Yat Tin Chow * Kazufumi Ito† Jun Zou*

Abstract

In this work we apply non-negative matrix factorizations (NMF) to some imaging and inverse problems. We propose a sparse low-rank approximation of big data and images in terms of tensor products, and investigate its effectiveness in terms of the number of tensor products to be used in the approximation. A multi-resolution analysis (MRA) framework is presented using a sparse low-rank approximation. We propose a primal-dual active set semi-smooth Newton method for the non-negative factorization. Numerical results are given to demonstrate the effectiveness of the proposed method to capture features in images and structures of inverse problems under no *a-priori* assumption on the data structure, as well as to provide a sparse low-rank representation of the data.

Mathematics Subject Classification (MSC2000): 15A23, 65F22, 65F30, 65F50, 78M25.

Keywords: Non-negative matrix factorization, Clustering, Feature extraction, Multi-resolution analysis, Inverse problems

1 Introduction to Non-negative Factorizations

Non-negative factorization (NMF) has emerged since the last decade, attempting to tackle k-clustering problems and structural analysis of big data. It is very effective in extraction of principle components, features and similarities inside a large set of data or image. NMF was first studied by Paatero [19], and suggested by Lee and Seung [14, 15] into the realm of machine learning and data mining. Since then, the now-known concept of NMF as K-means clustering for principle component analysis has been widely studied theoretically and numerically in many works, e.g. [1, 4, 5, 6, 9, 17, 19, 21]. The concept of tri-factorization was also suggested as a concurrent column and row clustering (a.k.a. co-clustering) of data in, e.g. [7] . Usually, in order to achieve feature extraction as well as to reduce memory complexity, sparsity is imposed to NMF using l_0 or l_1 regularization. In order to actualize these various concepts of NMF, mature toolboxes have been well-developed to provide different choices of regularizers and constraints, e.g. the non-negative matrix factorization toolbox in MATLAB developed by Li and Ngom [18]. Another convex model for NMF is suggested by Ocher *et. al.* [8], where the convex $l_{1,\infty}$ norm is used as the regularizer to enforce row sparsity. Applying this convex model to hyper-spectral end-members selections, the authors succeeded in providing abundance maps of end-members representing different structures inside an image, e.g. roofs, trees, grass, soil and road.

Generally speaking, a NMF of a general matrix $Y \in \mathbb{R}^{N\times M}$ is of the form

$$Y \approx AP \tag{1.1}$$

where $A \in \mathbb{R}^{N\times k}$, $P \in \mathbb{R}^{k\times M}$ and $P \geq 0$ entry-wise. Generally, the rank k is much lower than the rank of X, i.e. $k << \min(N, M)$. The matrix P is regarded as a basis of the information contained in matrix Y. We may further requires P to be nearly orthornormal, i.e. $PP^T \approx I$. In this case, it is similar to a partition of unity in the underlying space considered and the vectors in P are similar to some indicator functions. The matrix A is an assignment matrix, which gives some special weighting to the corresponding vectors in

*Department of Mathematics, The Chinese University of Hong Kong, Shatin, Hong Kong. (ytchow@math.cuhk.edu.hk, zou@math.cuhk.edu.hk).

†Department of Mathematics and Center for Research in Scientific Computation, North Carolina State University, Raleigh, North Carolina (kito@unity.ncsu.edu).

P. It is our aim to obtain a sparse matrix A which has a very small number of non-zero entries. Therefore, A can be interpreted as some sparse assignments of linear combinations of basis vectors in P. If the matrix $Y \geq 0$ entry-wise, we may further requires $A \geq 0$ entry-wise. This constraint is however infeasible if Y is not nonnegative entry-wise and we are bound to relax and drop the non-negativity condition for A if this is the case. The sparsity constraint imposed on A makes the information extraction process to be concise and precise. Moreover, the sorting of vectors in A in descending order of its magnitude indicates to us the most important bases of P. From this sorting process, the physically important components can be effectively extracted. Using the standard l_1 regularization to impose sparsity and almost orthogonality, the problem of NNF for a non-negative matrix Y can be reformulated as the following minimization problem:

$$\min_{A\geq 0, P\geq 0} ||Y - AP||_{F,2}^2 + \alpha||A||_{F,1} + \gamma||PP^T - I||_{F,1} \tag{1.2}$$

over nonnegative matrices $A \in \mathbb{R}^{N\times k}$ and $P \in \mathbb{R}^{k\times M}$, where $||X||_{F,2} := \sqrt{\sum_{i,j} |X_{ij}|^2}$ is the Frobenius norm, $||X||_{F,1} := \sum_{i,j} |X_{ij}|$ and α, γ are some regularization parameters.

A classical and natural approach for matrix factorization is the singular value decomposition (SVD), which helps to obtain the best low-rank approximation of the matrix in l_2 sense. Classical approaches usually use SVD decomposition to extract the most important components of the matrix Y based on the magnitude of their corresponding singular values. The factorization of SVD is of the form

$$Y = U\Sigma V^T \tag{1.3}$$

where we can interpret the matrix U, V as basis of information, Σ as a weighting representing the importance of the corresponding basis vectors in U and V. However, SVD is unstructured, and therefore we turn to NMF to obtain a structural decomposition of the matrix. Combining the non-negativity constraints and the SVD gives rise to the idea of non-negative matrix tri-factorization, which was first suggested in [7]. In this work, we suggest and investigate the following version of non-negative matrix tri-factorization for non-negative matrix Y formulated using l_1 regularisation:

$$\min_{U\geq 0, \Sigma\geq 0, V\geq 0} ||Y - U\Sigma V^T||_{F,2}^2 + \alpha||\Sigma||_{F,1} + \gamma||UU^T - I||_{F,1} + \gamma||VV^T - I||_{F,1}. \tag{1.4}$$

We can interpret the matrices $U \in M_{M\times p}$, $V \in M_{p\times N}$ as basis of information, $\Sigma \in M_{p\times p}$ as a generalized singular matrix. In here, we would like to remark that the matrix Σ is not required to be diagonal in our setting, but is only required to be sparse.

In our work, we suggest to apply the aforementioned model of non-negative matrix tri-factorization to big data and large images because of its low memory complexity when the rank p is small, even if the original data and images do not attain any sparse structure. This is true considering the fact that the factorization is a low rank sparse approximation of the matrix in term of the tensor products of column and row vectors of U and V. The fact that p is small requires the storage of only a few columns and rows in the matrices U and V, and therefore greatly reduces memory complexity. The sparsity of Σ is also very important for the reduction of memory complexity because we only need to store the respective columns and rows of the matrices U and V, i.e. u_i and v_j where the corresponding entry in the singular matrix Σ, i.e. σ_{ij} is significant. These reasons suggest us to apply the above NMF model to big data and imaging. To effectively implement the NMF, we utilize the well-known primal-dual active set semi-smooth Newton method [11] for the optimization process. It is more advantageous than the classical methods suggested in [5] or [7] since Newton-type methods are known to converge faster than these classical methods. Using the result of NMF from the Newton method, we propose a dissection of the image into layers by its order of importance. Moreover, we propose a multi-resolution analysis (MRA) framework of the images based on the NMF, which induces a sparse representation and extraction of features of different scales. Numerical results has shown acceptable resolution of images can be achieved with low memory complexity using this model of tri-factorization without any *a-priori* assumption of their structures, such as sparsity and specific patterns.

This paper is organized as follows. In section 2 the general mathematical framework of non-negative matrix tri-factorization using l_1 regularization is clearly stated, and an optimal choice of the dimension of

generalized singular matrix is investigated. An MRA framework using NMF is introduced in section 3 and a primal-dual active set semi-smooth Newton method for NMF is presented in section 4. Applications of our framework to imaging and inverse problems are provided in section 5, providing numerical evidence for the success of sparse low-rank representation of the data.

2 Mathematical formulation of non-negative matrix tri-factorization using l_1 regularization

In this section we clearly state the matrix tri-factorization that we would like to consider. In the subsequent discussion, we shall often denote $M_{M\times N}$ as the set of $M \times N$ matrices and $M^+_{M\times p} \subset M_{M\times N}$ as those with positive entries. Given a matrix $Y \in M^+_{M\times N}$, we define the following functional $\mathcal{J}^{\alpha,\gamma}_p$ for a fixed set of parameters p, α, γ:

$$\begin{aligned} \mathcal{J}^{\alpha,\gamma}_p &: M^+_{M\times p} \times M^+_{p\times p} \times M^+_{p\times N} \to \mathbb{R} \\ \mathcal{J}^{\alpha,\gamma}_p(U,\Sigma,V) &:= ||Y - U\Sigma V^T||^2_{F,2} + \gamma||\Sigma||_{F,1} + \alpha||UU^T - I||_{F,1} + \alpha||VV^T - I||_{F,1}. \end{aligned} \tag{2.1}$$

Consider a minimizer of the functional, denoted as $[\tilde{U}_p, \tilde{\Sigma}_p, \tilde{V}_p] \in \arg\min \mathcal{J}^{\alpha,\gamma}_p$. With the above definition at hand, we then define the following operator $\mathcal{I}^{\alpha,\gamma}_p$:

$$\begin{aligned} \mathcal{I}^{\alpha,\gamma}_p : M^+_{M\times N} &\to M^+_{M\times N} \\ \mathcal{I}^{\alpha,\gamma}_p(Y) &:= \tilde{U}_p\tilde{\Sigma}_p\tilde{V}_p = \sum_{i,j} \sigma_{ij}(\tilde{u}_p)_i \otimes (\tilde{v}_p)_j \end{aligned} \tag{2.2}$$

where $(\tilde{u}_p)_i, (\tilde{v}_p)_j$ denote the respect column and row vectors of $\tilde{U}_p$ and $\tilde{V}_p$ respectively and σ_{ij} is the (i,j)-th entry of the matrix Σ_p. This non-negative matrix tri-factorization can be regarded as a non-negative version of the SVD, and the matrix Σ_p is called the generalized singular matrix, which is not constrained to be diagonal.

It is obvious yet worth-noting that with a small p, the memory complexity of storing the matrix triple $[\tilde{U}_p, \tilde{\Sigma}_p, \tilde{V}_p]$ is minimal, since the memory complexity is of the order $(M+N)p + p^2$ even if $\tilde{\Sigma}_p$ is not sparse. If $\tilde{\Sigma}_p$ is furthermore a sparse matrix, the memory complexity can be much further reduced, as we only need to store the vectors $(\tilde{u}_p)_i$ and $(\tilde{v}_p)_j$ when σ_{ij} is non-zero. In fact, for a generic matrix Y, if p can be chosen to be small and yet $||Y - \mathcal{I}^{\alpha,\gamma}_p(Y)||^2_{F,2}$ can still be maintained to be a small quantity, then $[\tilde{U}_p, \tilde{\Sigma}_p, \tilde{V}_p]$ may serve as our desired sparse low-rank approximation of Y. However, it is obvious that the smaller the value of p is, the worse the approximation of Y by $\mathcal{I}^{\alpha,\gamma}_p(Y)$ will be. With a smaller p, the error $||Y - \mathcal{I}^{\alpha,\gamma}_p(Y)||^2_F$ and also $\mathcal{J}^{\alpha,\gamma}_p(\tilde{U}_p, \tilde{\Sigma}_p, \tilde{V}_p)$, will be larger. Therefore, in practice, it is an interesting question to ask how we should choose the number p as N, M grow large.

2.1 An Optimal choice of p

In what follows, we aim to find an optimal choice of p with respect to N, M by means of a probabilistic argument. We first obtain a lower bound in terms of p, N, M, δ of the probability that there exists a triple $[U,\Sigma,V]$ such that $\mathcal{J}^{\alpha,\gamma}_p(U,\Sigma,V) < \delta$. From this lower bound, we suggest an optimal choice of p to maximize this probability. The value $\mathcal{J}^{\alpha,\gamma}_p(U_p,\Sigma_p,V_p)$ reflects the derivations of matrices U_p, V_p from being orthogonal, the sparsity of Σ_p and the error of the approximation of Y by $\mathcal{I}^{\alpha,\gamma}_p(Y)$. In particular, if for some $[U,\Sigma,V]$, we have $\mathcal{J}^{\alpha,\gamma}_p(U,\Sigma,V) < \delta$, then

$$||Y - \mathcal{I}^{\alpha,\gamma}_p(Y)||^2_{F,2} \le \mathcal{J}^{\alpha,\gamma}_p(U_p,\Sigma_p,V_p) \le \mathcal{J}^{\alpha,\gamma}_p(U,\Sigma,V) < \delta\,.$$

We begin by showing the following lemmas concerning a set of i.i.d. random vectors. Consider a set of i.i.d random vectors $\{X_i\}_{i=1}^N \in [0,1]^d$, where the probability distribution $d\mathbb{P}_X = f dx$ with dx denoting

the standard Lebesgue measure and $0 < C_1 < f < C_2 < \infty$. Then it is immediate to see that the random variables $\{\omega_i := X_i/||X_i||_2\}_{i=1}^N \in \mathbb{S}^{d-1}$ has a probability density $d\,\mathbb{P}_\omega = g d\omega$ where $d\omega$ is the standard surface measure and $\frac{C_1}{||\omega||_\infty} \leq g \leq \frac{C_2}{||\omega||_\infty}$ for some other constants $0 < C_1, C_2 < \infty$. From this, we can derive the following important results for our subsequent analysis.

Lemma 2.1. *Consider a set of i.i.d random vectors $\{X_i\}_{i=1}^N \in [0,1]^d$, where the probability distribution $d\,\mathbb{P}_X = f dx$ with dx denoting the standard Lebesgue measure and $0 < C_1 < f < C_2 < \infty$. Then the probability of the vectors $\omega_i := X_i/||X_i||_2$ being able to be approximated by p points $\{P_i\}_{i=1}^p \in \mathbb{S}^{d-1} \bigcup [0,1]^d$ within an error of $\varepsilon > 0$ for a small ε can be estimated as follows:*

$$p^N (C_3\varepsilon)^{(d-1)N} \leq \mathbb{P}\left(\exists \{P_i\}_{i=1}^p \ s.t. \ \{\omega_i\}_{i=1}^N \subset \bigcup_{1\leq i\leq P} B_\varepsilon(P_i)\right) \leq p^N (C_4\varepsilon)^{(d-1)N} \tag{2.3}$$

for some $0 < C_3, C_4 < \infty$.

Proof. Using the fact that for small $\varepsilon > 0$, $C\varepsilon < \sin\varepsilon < \varepsilon$ for some $C > 0$, we can actually observe from the assumption of the i.i.d. random vectors and the polynomial theorem that

$$\begin{aligned}
& \mathbb{P}\left(\exists \{P_i\}_{i=1}^p \text{ s.t. } \{\omega_i\}_{i=1}^N \subset \bigcup_{1\leq i\leq P} B_\varepsilon(P_i)\right) \\
= & \sum_{\sum_{i=1}^p N_i = N} \frac{N!}{\prod_i N_i!} \frac{1}{|\mathbb{S}^{d-1} \bigcap [0,1]^d|} \prod_i \int_{\mathbb{S}^{d-1}\bigcap[0,1]^d} \mathbb{P}(||\omega_i - K||_2 < \varepsilon)^{N_i} dK \\
\geq & \sum_{\sum_{i=1}^p N_i = N} \frac{N!}{\prod_i N_i!} (C_3\varepsilon)^{(d-1)\sum_i N_i} \\
\geq & \ p^N (C_3\varepsilon)^{(d-1)N}
\end{aligned}$$

for some $C_3 > 0$. The other side of the inequality is similar. □

Lemma 2.2. *Consider a set of i.i.d random vectors $\{P_i\}_{i=1}^p \in [0,1]^d$, where the probability distribution $d\,\mathbb{P}_\omega = f d\omega$ with $d\omega$ denoting the standard surface measure and $0 < C_1 < f < C_2 < \infty$. Then the probability of the set of vectors P_i being almost mutually orthogonal within an error of ε for a small $\varepsilon > 0$ can be estimated as follows:*

$$p!\, d\, (C_3\varepsilon)^{\frac{(p)(p-1)}{2}+(d-1)} \leq \mathbb{P}\left(|\langle P_i, P_j\rangle - \delta_{ij}| < \varepsilon \, \forall i,j\right) \leq p!\, d\, (C_4 3\varepsilon)^{\frac{(p)(p-1)}{2}+(d-1)} \tag{2.4}$$

for some $0 < C_3, C_4 < \infty$ if $p \leq d$, and is zero if otherwise.

Proof. From direct counting, and the fact that $||P_i - P_j||^2 = 2 - 2\langle P_i, P_j\rangle$ together with half angle formula, we have for $p \leq d$,

$$\begin{aligned}
& \mathbb{P}\left(\langle |P_i, P_j\rangle - \delta_{ij}| < \varepsilon \, \forall i,j\right) \\
\geq & \ p!\, d\, (C_3\varepsilon)^{d-1} \prod_{1\leq i\leq p} (C_3\varepsilon)^i |(\mathbb{B}_1^i \times \mathbb{B}_1^{n-i}) \bigcap [0,1]^d| \\
\geq & \ p!\, d\, (C_3\varepsilon)^{\frac{(p)(p-1)}{2}+(d-1)}
\end{aligned}$$

for some $C_3 > 0$. The other side of the inequality is similar. The probability of this being zero when $p > d$ follows directly from dimensionality argument. □

Lemma 2.3. *Consider a set of i.i.d random vectors $\{X_i\}_{i=1}^N \in [0,1]^d$, where the probability distribution $d\mathbb{P}_X = f dx$ with dx denoting the standard Lebesgue measure and $0 < C_1 < f < C_2 < \infty$. Then the*

probability of the event $E_{p,\varepsilon}$ *representing* $\exists\{P_i\}_{i=1}^p$ *s.t.* $\{\omega_i\}_{i=1}^N \subset \bigcup_{1\le i\le P} B_\varepsilon(P_i)$ *and* $|\langle P_i, P_j\rangle - \delta_{ij}| < \varepsilon\, \forall i,j$ *for a small* $\varepsilon > 0$ *can be estimated as follows:*

$$\left(p^N-(p-1)^N\right) pl!\, d\,(C_3\varepsilon)^{\frac{p(p-1)}{2}+(d-1)(N+1)} \le \mathbb{P}\left(E_{p,\varepsilon}\backslash E_{p-1,\varepsilon}\right) \le \left(p^N-(p-1)^N\right) pl!\, d\,(C_4\varepsilon)^{\frac{p(p-1)}{2}+(d-1)(N+1)}$$

for some $0 < C_3, C_4 < \infty$ *if* $p \le N$*, and therefore*

$$\sum_{l=1}^{p}\left(l^N-(l-1)^N\right) l!\, d\,(C_3\varepsilon)^{\frac{l(l-1)}{2}+(d-1)(N+1)} \le \mathbb{P}\left(E_{p,\varepsilon}\right) \le \sum_{l=1}^{p}\left(l^N-(l-1)^N\right) l!\, d\,(C_4\varepsilon)^{\frac{l(l-1)}{2}+(d-1)(N+1)}.$$

We also have the following lower bound estimate

$$\mathbb{P}\left(E_{p,\varepsilon}\right) \ge dp^N (C_3\varepsilon)^{(d-1)(N+1)+\frac{(p)(p-1)}{2}}. \tag{2.5}$$

Proof. The following inequality follows directly from the argument of the above two lemmas

$$\begin{aligned} & \sum_{\sum_{i=1}^p N_i=N\,,\,N_i>0} \frac{N!}{\prod_i N_i!} p!\, d\,(C_3\varepsilon)^{\frac{p(p-1)}{2}+(d-1)(N+1)} \\ \le\ & \mathbb{P}\left(E_{p,\varepsilon}\backslash E_{p-1,\varepsilon}\right) \\ \le\ & \sum_{\sum_{i=1}^p N_i=N\,,\,N_i>0} \frac{N!}{\prod_i N_i!} p!\, d\,(C_4\varepsilon)^{\frac{p(p-1)}{2}+(d-1)(N+1)}. \end{aligned}$$

Now since the last term can be simplified as follows:

$$\begin{aligned} & p!\, d\,(C_3\varepsilon)^{\frac{p(p-1)}{2}+(d-1)(N+1)} \sum_{\sum_{i=1}^p N_i=N\,,\,N_i>0} \frac{N!}{\prod_i N_i!} \\ =\ & p!\, d\,(C_3\varepsilon)^{\frac{p(p-1)}{2}+(d-1)(N+1)} \left(\sum_{\sum_{i=1}^p N_i=N} \frac{N!}{\prod_i N_i!} - \sum_{\sum_{i=1}^{p-1} N_i=N} \frac{N!}{\prod_i N_i!} \right) \\ =\ & \left(p^N-(p-1)^N\right) p!\, d\,(C_3\varepsilon)^{\frac{p(p-1)}{2}+(d-1)(N+1)}, \end{aligned}$$

we directly have

$$\sum_{l=1}^{p}\left(l^N-(l-1)^N\right) l!\, d\,(C_3\varepsilon)^{\frac{l(l-1)}{2}+(d-1)(N+1)} \le \mathbb{P}\left(E_{p,\varepsilon}\right) \le \sum_{l=1}^{p}\left(l^N-(l-1)^N\right) l!\, d\,(C_4\varepsilon)^{\frac{l(l-1)}{2}+(d-1)(N+1)}.$$

The last inequality comes readily from

$$\mathbb{P}\left(E_{p,\varepsilon}\right) \ge \sum_{l=1}^{p}\left(l^N-(l-1)^N\right)\, d\,(C_3\varepsilon)^{\frac{p(p-1)}{2}+(d-1)(N+1)} = dp^N (C_3\varepsilon)^{(d-1)(N+1)+\frac{(p)(p-1)}{2}}.$$

□

Now we consider a general image or large data $Y = \sum_{i,j} Y_{ij}\, e_i \otimes e_j$ comprised of non-negative entries. Without loss of generality, we may always assume $\max_{i,j}|Y_{ij}| = 1$. Write $Y_i := \sum_j Y_{ij}\, e_j$, and $\omega_i = Y_i/||Y_i||_2$, then $Y = \sum_i ||Y_i||_2\, e_i \otimes \omega_i$. If it is possible that there exists a set of $\{P_i\}_{i=1}^p$ such that $\{\omega_i\}_{i=1}^N \subset \bigcup_{1\le i\le P} B_\varepsilon(P_i)$ and $|\langle P_i, P_j\rangle - \delta_{ij}| < \varepsilon\ \forall i,j$, then we can write $\{\omega_{k_j}\}_{j=1}^{K_j} \in B_\varepsilon(P_j)$ for some K_j with $1 \le j \le p$ and then intuitively, we have

$$I = \sum_i ||Y_i||_2 e_i \otimes \omega_i \approx \sum_{j=1}^{p}\sum_{k_j=1}^{K_j} ||Y_{k_j}||_2\, e_{k_j} \otimes P_j\,.$$

Writing $Q_j := (\sum_{k_j=1}^{K_j} ||Y_{k_j}||_2 e_{k_j})/\sqrt{\sum_{k_j=1}^{K_j} ||Y_{k_j}||_2}$ and denoting $\sigma_{ij} = \delta_{ij}\sqrt{\sum_{k_j=1}^{K_j} ||Y_{k_j}||_2}$, then we have

$$I \approx \sum_i \sigma_{ij}\, Q_i \otimes P_j$$

where $|\langle P_i, P_j\rangle - \delta_{ij}| < \varepsilon\ \forall i,j$ and $|\langle Q_i, Q_j\rangle - \delta_{ij}| = 0\ \forall i,j$. Writing $\Sigma = (\sigma_{ij})$, $P = (P_i)^T$ $Q = (Q_j)$, then we actually have directly that

$$||I - \sum_i \sigma_{ij}\, Q_i \otimes P_j||_{F_2} \leq \sum_{j=1}^{p}\sum_{k_j=1}^{K_j} ||Y_{k_j}||_2 |\omega_{k_j} - P_j| \leq ||I||_{F,2}\varepsilon \leq NM\varepsilon\,,$$

as well as

$$\begin{aligned}\mathcal{J}_p^{\alpha,\gamma}(\tilde{U}_p, \tilde{\Sigma}_p, \tilde{V}_p) \leq \mathcal{J}_p^{\alpha,\gamma}(Q,\Sigma,P) &\leq ||I||_{F,2}\varepsilon + \gamma\sum_j \sqrt{\sum_{k_j=1}^{K_j} ||Y_{k_j}||_2} + \alpha p(p-1)\varepsilon \\ &\leq NM\varepsilon + NM\gamma + \alpha\, p(p-1)\varepsilon\,.\end{aligned}$$

The measure of the event $E_{p,\varepsilon}$ such that the above happens has an estimate of

$$\begin{aligned}\mathbb{P}\,(E_{p,\varepsilon}) &\geq \sum_{l=1}^{p} \left(l^N - (l-1)^N\right) l!\, M\, (C_3\varepsilon)^{\frac{l(l-1)}{2}+(M-1)(N+1)} \\ &\geq M\, p^N (C_3\varepsilon)^{(M-1)(N+1)+\frac{(p)(p-1)}{2}}\,.\end{aligned}$$

Similarly, switching the columns and rows of the image, we may follow the above argument and analysis to conclude the same with N, M switched. Combining to the two statements, we have

$$\begin{aligned}&\mathbb{P}\left(\mathcal{J}_p^{\alpha,\gamma}(\tilde{U}_p, \tilde{\Sigma}_p, \tilde{V}_p) < NM\varepsilon + NM\gamma + \alpha\, p(p-1)\varepsilon\right) \\ \geq\ &\sum_{l=1}^{p}\left(l^{\max(N,M)} - (l-1)^{\max(N,M)}\right) l!\, \min(N,M)\,(C_3\varepsilon)^{\mu(N,M,l)} \\ \geq\ &\min(N,M) p^{\max(N,M)} (C_3\varepsilon)^{\mu(N,M,p)}\,.\end{aligned}$$

where the function $\mu(\,\cdot\,,\cdot\,,\cdot\,)$ is defined for all $N, M, l \in \mathbb{N}$ as

$$\mu(N,M,l) := \frac{l(l-1)}{2} + MN - |N-M| - 1\,. \tag{2.6}$$

If we further choose the parameter $\gamma \leq (K-1)\varepsilon$ for some $K > 1$, then we have directly that

Lemma 2.4. *For any small $\varepsilon > 0$ and for all $N, M \in \mathbb{N}$, then we have*

$$\begin{aligned}&\mathbb{P}\left(\mathcal{J}_p^{\alpha,\gamma}(\tilde{U}_p, \tilde{\Sigma}_p, \tilde{V}_p) < \left(KNM + \min(N,M)^2\right)\varepsilon\right) \\ \geq\ &\sum_{l=1}^{p}\left(l^{\max(N,M)} - (l-1)^{\max(N,M)}\right) l!\, \min(N,M)\,(C_3\varepsilon)^{\mu(N,M,l)} \qquad (2.7) \\ \geq\ &\min(N,M) p^{\max(N,M)} (C_3\varepsilon)^{\mu(N,M,p)} \qquad (2.8)\end{aligned}$$

where the function $\mu(\,\cdot\,,\cdot\,,\cdot\,)$ is defined as in (2.6) *and γ is such that $\gamma \leq (K-1)\varepsilon$ for some $K > 1$.*

Before we derive a sharp bound of an optimal choice for p, let us consider a rough lower bound introduced in the last inequality (2.8). Now if we consider the function

$$F(p) := \min(N,M)p^{\max(N,M)}\,(C_3\varepsilon)^{\mu(N,M,p)}$$

for $p \geq 1$, then we have

$$\begin{aligned} F'(p) \quad &= \quad \frac{F(p)}{p}\left(\max(N,M) + \frac{|\log(C_3\varepsilon)|}{16} - |\log(C_3\varepsilon)|(p-\frac{3}{4})^2\right) \\ &\begin{Bmatrix} > \\ = \\ < \end{Bmatrix} \quad 0 \Leftrightarrow p \begin{Bmatrix} < \\ = \\ > \end{Bmatrix} \frac{3}{4} + \sqrt{\frac{1}{16} + \frac{\max(M,N)}{|\log(C_3\varepsilon)|}}\,. \end{aligned}$$

Therefore we can propose a primitive optimal choice of p to maximize the lower bound of the possibility $\mathbb{P}\left(\mathcal{J}_p^{\alpha,\gamma}([\tilde{U}_p,\tilde{\Sigma}_p,\tilde{V}_p]) < \left(KNM+\min(M,N)^2\right)\varepsilon\right)$, i.e. to choose

$$p = \sqrt{\frac{\max(M,N)}{|\log(C_3\varepsilon)|}} \tag{2.9}$$

for large N, M. Following some basic substitutions, we obtain the following theorem.

Theorem 2.5. *For any small $\delta > 0$, we have*

$$\mathbb{P}\left(\min_p \mathcal{J}_p^{\alpha,\gamma}(\tilde{U}_p,\tilde{\Sigma}_p,\tilde{V}_p) < \delta\right) \geq \min(N,M)\,p_{N,M,\delta}^{\max(N,M)}\,(C_3\varepsilon)^{\mu(N,M,p_{N,M,\delta})} \tag{2.10}$$

whenever $\gamma \leq (K-1)\,\varepsilon$, where $\varepsilon := \delta\left(KNM+\min(M,N)^2\right)^{-1}$ for some $K > 1$, the function $\mu(\,\cdot\,,\,\cdot\,,\,\cdot\,)$ is defined as in (2.6)*, and $p_{N,M,\delta}$ stands for the following constant*

$$p_{N,M,\delta} := \sqrt{\frac{\max(M,N)}{|\log(C_3\varepsilon)|}} = \sqrt{\frac{\max(M,N)}{\log(KNM+\min(M,N)^2) - |\log\delta| - \log C_3}}\,. \tag{2.11}$$

When $M = N$, it is then obvious that the above optimal choice of p for a fixed $\delta > 0$ is of the form

$$p = p_{N,N,\delta} = \sqrt{\frac{N}{2\log N - |\log\delta| - \log C_3 + \log(K+1)}} \sim \sqrt{\frac{N}{2\log N}} \tag{2.12}$$

as N goes to infinity. The last asymptotic actually gives a precise approximation and

$$\sqrt{\frac{N}{2\log N}} \leq p_{N,N,\delta} \leq \sqrt{\frac{N}{\log N}} \tag{2.13}$$

if N is large enough such that $N > C_3\delta^{-1}$. Hence (2.12) serves as an optimal choice of p for large N. Furthermore, with this choice of p, the memory complexity is asymptotically $\sqrt{\frac{2N^3}{\log N}}$ as N goes to infinity.

However, we note that the optimal choice of p obtained above is only based on a rough lower bound (2.8). In what follows, we peruse a sharper bound by using (2.7). Since (2.7) always increases with respect to p, we get an optimal choice of p by controlling the increment of (2.7) with respect to p. In order to do so, we investigate the ratio of the terms

$$a_l := \left(l^{\max(N,M)} - (l-1)^{\max(N,M)}\right) l!\,\min(N,M)\,(C_3\varepsilon)^{\mu(N,M,l)}$$

explicitly given as follows:

$$\frac{a_{l+1}}{a_l} = \frac{(l+1)^{\max(N,M)} - l^{\max(N,M)}}{l^{\max(N,M)} - (l-1)^{\max(N,M)}} l e^{-|\log(C_3\varepsilon)|(l+1)}\,.$$

From the l'Hospital rule, we can directly see that for a fixed pair of N, M, the above term $a_{l+1}/a_l \to 0$ as $l \to \infty$. Therefore, given a small $\eta < 1$, there is always a $\hat{p}_{N,M,\eta,\varepsilon}$ such that $a_{l+1}/a_l \le \eta$ whenever $l > \hat{p}_{N,M,\eta,\varepsilon}$. Then for all $p > \hat{p}_{N,M,\eta,\varepsilon}$ we have

$$\begin{aligned}
&\mathbb{P}\left(\mathcal{J}_p^{\alpha,\gamma}(\tilde{U}_p, \tilde{\Sigma}_p, \tilde{V}_p) < \left(KNM + \min(N,M)^2\right)\varepsilon\right) \\
\ge &\sum_{l=1}^{\hat{p}_{N,M,\eta,\varepsilon}-1} \left(l^{\max(N,M)} - (l-1)^{\max(N,M)}\right) l!\, \min(N,M)\, (C_3\varepsilon)^{\mu(N,M,l)} \\
&+ \frac{1}{1-\eta}\left((\hat{p}_{N,M,\eta,\varepsilon})^{\max(N,M)} - (\hat{p}_{N,M,\eta,\varepsilon}-1)^{\max(N,M)}\right)(\hat{p}_{N,M,\eta,\varepsilon})!\, \min(N,M)\, (C_3\varepsilon)^{\mu(N,M,\hat{p}_{N,M,\eta,\varepsilon})}
\end{aligned}$$

whenever $\gamma \le (K-1)\,\varepsilon$, and that the increment of p from $\hat{p}_{N,M,\eta,\varepsilon}$ onward brings insignificant increment to (2.7). Now we aim to find an explicit $\hat{p}_{N,M,\eta,\varepsilon}$ in terms of N, M, thus obtaining an optimal choice of p. From Holder's inequality, we readily have

$$a_{p+1}/a_p = \frac{\sum_{i=0}^{\max(N,M)-1}(1+1/p)^i}{\sum_{i=0}^{\max(N,M)-1}(1-1/p)^i} p\, e^{-|\log(C_3\varepsilon)|(p+1)} \le \frac{p(p+1)^{\max(N,M)-1}}{(p-1)^{\max(N,M)-1}} p\, e^{-|\log(C_3\varepsilon)|(p+1)}. \quad (2.14)$$

Now if we consider the function

$$G(N_0, p) := \frac{p(p+1)^{N_0-1}}{(p-1)^{N_0-1}} p\, e^{-|\log(C_3\varepsilon)|(p+1)}$$

for $p \ge 2$ and $N_0 \ge 2$, then we have

$$\begin{aligned}
\frac{\partial}{\partial p} G(N_0,p) \quad &= \quad G(N_0,p)\left(\frac{1}{p} + \frac{N_0-1}{p+1} - \frac{N_0-1}{p-1} - |\log(C_3\varepsilon)|\right) \\
\begin{Bmatrix} > \\ = \\ < \end{Bmatrix} \quad &0 \Leftrightarrow p \begin{Bmatrix} < \\ = \\ > \end{Bmatrix} p_0(N_0).
\end{aligned}$$

where $p_0(N_0)$ is the unique real zero of $-|\log(C_3\varepsilon)|p^3 + p^2 + (|\log(C_3\varepsilon)| - 2N_0 + 2)p - 1 = 0$ which can be found explicitly by the Cardano's formula or the Lagrange's method. Fixing N_0, we get that $G(N_0, p_0(N_0))$ is the global maximum of $G(N_0, p)$ on $(2, \infty)$, and from $p_0(N_0)$ onward, the function is decreasing. Together with the fact that $G(N_0, 2) = 4(5/3)^{N_0-1}(C_3\varepsilon)^4$, we have that $G(N_0,\cdot)^{-1} : (0, 4(C_3\varepsilon)^4) \to (p_0, \infty)$ is a well-defined smooth function and is monotone by inverse function theorem, and that the implicity function $g : (1,\infty) \to (1,\infty)$ defined by $G(N_0, g(N_0)) = \eta$ is well-defined and smooth by implicit function theorem as $g(N_0) = [G(N_0,\cdot)]^{-1}(\eta)$. Moreover

$$\begin{aligned}
g' = -\frac{\frac{\partial N_0}{\partial p}(N_0, g(N_0))}{\frac{\partial G}{\partial p}(N_0, g(N_0))} \quad &= \quad -\log\left(\frac{g+1}{g-1}\right)\left(\frac{1}{g} + \frac{N_0-1}{g+1} - \frac{N_0-1}{g-1} - |\log(C_3\varepsilon)|\right)^{-1} \\
&= \quad \log\left(\frac{g+1}{g-1}\right)\frac{g(g+1)(g-1)}{|\log(C_3\varepsilon)|g^3 - g^2 - (|\log(C_3\varepsilon)| - 2N_0 + 2)g + 1}
\end{aligned}$$

Now since $g(N_0) > p_0(N_0) + \hat{\delta} > 1$ for some $\hat{\delta} > 0$ by our choice of domain, we have $|\log(C_3\varepsilon)|p^3 - p^2 - (|\log(C_3\varepsilon)| - 2N_0 + 2)p + 1 > \hat{C} > 0$ for some $\hat{C}$, and $0 < g'(N_0) < \infty$ for all N_0 as well as $g'(N_0) \to \infty$ as $N_0 \to \infty$. Moreover putting these inequalities back into the expression of g', we have $g'(N_0) \to 0$ as $N_0 \to \infty$, and that g satisfies the following differential inequality for large N_0,

$$g' \le \log\left(\frac{g+1}{g-1}\right)\frac{2}{|\log(C_3\varepsilon)|} \le \frac{4}{(g-1)|\log(C_3\varepsilon)|}$$

Now using the Gronwall-Bellman-Bihari's inequality, we directly infer that

$$g \le H^{-1}(H(a(\eta)) + N_0) \quad (2.15)$$

for some constant $a(\eta)$ depending only on η, where the function H is defined as

$$H(s) := \frac{|\log(C_3\varepsilon)|}{4}\int (s-1)ds = \frac{|\log(C_3\varepsilon)|(s-1)^2}{8} + K_0(\eta) \tag{2.16}$$

for some $K_0(\eta)$. Therefore the following inequality holds for g:

$$g \leq \sqrt{\frac{K_1(\eta)N_0 - K_2(\eta)}{|\log(C_3\varepsilon)|}} + K_3(\eta)\,.$$

for some $K_1(\eta), K_2(\eta), K_3(\eta)$. Choosing $\hat{p}_{N,M,\eta,\varepsilon}$ such that

$$\hat{p}_{N,M,\eta,\varepsilon} = K_\eta\sqrt{\frac{\max(N,M)}{|\log(C_3\varepsilon)|}} = K_\eta p_{N,M,\delta} \tag{2.17}$$

for some K_η depending only on η, where $p_{N,M,\delta}$ is defined as in (2.11), then for all

$$p > \hat{p}_{N,M,\eta,\varepsilon} \geq g\left(\max(N,M)\right) = \left[G\left(\max(N,M),\cdot\right)\right]^{-1}(\eta)\,,$$

we have

$$\frac{p(p+1)^{\max(N,M)-1}}{(p-1)^{\max(N,M)-1}}p\,e^{-|\log(C_3\varepsilon)|(p+1)} < \eta\,.$$

Therefore the growth of the probability $\mathbb{P}\left(\mathcal{J}_p^{\alpha,\gamma}(\tilde{U}_p,\tilde{\Sigma}_p,\tilde{V}_p) < \left(KNM + \min(N,M)^2\right)\varepsilon\right)$ with respect to p becomes insignificant from $\hat{p}_{N,M,\eta,\varepsilon}$ onward. This gives another optimal choice for p. Amazingly enough, we notice that

$$\hat{p}_{N,M,\eta,\varepsilon} \sim p_{N,M,\delta}\,.$$

i.e. the two choices of p are of the same order. This gives the following theorem.

Theorem 2.6. *For any small $\delta > 0$ and, the following bound for the probability holds*

$$\mathbb{P}\left(\mathcal{J}_p^{\alpha,\gamma}(\tilde{U}_p,\tilde{\Sigma}_p,\tilde{V}_p) < \delta\right) \geq \sum_{l=1}^{p}\left(l^{\max(N,M)} - (l-1)^{\max(N,M)}\right) l!\,\min(N,M)\,(C_3\varepsilon)^{\mu(N,M,l)} \tag{2.18}$$

whenever $\gamma \leq (K-1)\,\varepsilon$, where $\varepsilon := \delta\left(KNM + \min(M,N)^2\right)^{-1}$ for some $K > 1$ and the function $\mu(\,\cdot\,,\cdot\,,\cdot\,)$ is defined as in (2.6)*. For a given small constant η, the growth of value of the last summation with respect to p can be controlled by η when $p > K_\eta\, p_{N,M,\delta}$ for some K_η depending on only η where $p_{N,M,\delta}$ is defined as* (2.11)*. Furthermore, whenever the above event happens, $||Y - \mathcal{I}_p^{\alpha,\gamma}(Y)||_{F,2}^2 < \delta$.*

Therefore, in the particular case when $M = N$, the following asymptotic order for p

$$p \sim \sqrt{\frac{N}{\log N}} \tag{2.19}$$

is genuinely an optimal choice of p, and that they are equivalent up to a multiplicative constant whenever $N > C_3\delta^{-1}$. Following this optimal choice of p, the memory complexity grows in the order $\sqrt{\frac{N^3}{\log N}}$ as N goes to infinity.

2.2 Importance of magnitudes of entries in generalized singular matrix

In this subsection, we discuss a further reduction of memory complexity by truncating the generalized singular matrix $\tilde{\Sigma}_p = (\sigma_{ij})$. We aim to remove the less important components $(\tilde{u}_p)_i \otimes (\tilde{v}_p)_j$ in (2.2) while keeping the most important ones so that it still serves as a good approximation of the original matrix Y.

In order to attain this aim, we rearrange σ_{ij} from the largest value to the smallest value as $\sigma_{i_1 j_1} \geq \sigma_{i_2 j_2} \geq .. \geq \sigma_{i_{p^2} j_{p^2}}$. We then denote $\tilde{\sigma}_l = \sigma_{i_l j_l} e_l \otimes e_l$, and write $\tilde{\Sigma}_{p,\tilde{p}} = \sum_{l=1}^{\tilde{p}} \tilde{\sigma}_l$ as the truncated generalized singular matrix for all $\tilde{p} \leq p^2$. The sequence $\{\tilde{\sigma}_l\}_{l=1}^{p^2}$ represents the components of $\tilde{\Sigma}_p$ in descending order by its importance of magnitude. With the above definition at hand, we then define the following operator $\mathcal{I}_{p,\tilde{p}}^{\alpha,\gamma}$:

$$\begin{aligned} \mathcal{I}_{p,\tilde{p}}^{\alpha,\gamma} : M_{M\times N}^{+} &\rightarrow M_{M\times N}^{+} \\ \mathcal{I}_{p,\tilde{p}}^{\alpha,\gamma}(Y) &:= \tilde{U}_p \tilde{\Sigma}_{p,\tilde{p}} \tilde{V}_p = \sum_{l=1}^{\tilde{p}} \sigma_{i_l j_l} (\tilde{u}_p)_{i_l} \otimes (\tilde{v}_p)_{j_l} \end{aligned} \tag{2.20}$$

where $[\tilde{U}_p, \Sigma_p, \tilde{V}_p]$ is a chosen minimizer of the functional (2.1) and $\tilde{\Sigma}_{p,\tilde{p}}$ is the truncated generalized singular matrix up to $\tilde{p}$ as stated above.

The approximation $Y \approx \mathcal{I}_{p,\tilde{p}}^{\alpha,\gamma}(Y) = \tilde{U}_p \tilde{\Sigma}_{p,\tilde{p}} \tilde{V}_p$ is a truncation of the approximation (2.2) of Y up to $\tilde{p}$. This truncated approximation removes the less important components but keep the more important parts. With this approach, we only need to save the vectors $(\tilde{u}_p)_{i_l}$ and $(\tilde{v}_p)_{j_l}$ for $1 \leq l \leq \tilde{p}$. This further reduces the memory complexity and serves as our desired sparse low-rank approximation of Y.

In what follows, we give a brief analysis for the aforementioned truncated approximation of Y. Indeed, from the pigeon-hole principle, we directly obtain that

$$\begin{aligned} \mathcal{J}_p^{\alpha,\gamma}(\tilde{U}_p, \tilde{\Sigma}_{p,\tilde{p}}, \tilde{V}_p) &< \mathcal{J}_p^{\alpha,\gamma}(\tilde{U}_p, \tilde{\Sigma}_p, \tilde{V}_p) + C||I||_1 \sum_{i=0}^{p^2-\tilde{p}} \frac{1}{p^2 - i} \\ &< \mathcal{J}_p^{\alpha,\gamma}(\tilde{U}_p, \tilde{\Sigma}_p, \tilde{V}_p) + C||I||_1 \int_{\frac{\tilde{p}}{p^2}}^{1} 1/x dx \\ &< \mathcal{J}_p^{\alpha,\gamma}(\tilde{U}_p, \tilde{\Sigma}_p, \tilde{V}_p) + CNM \log\left(\frac{p^2}{\tilde{p}}\right) \\ &< \left((K + LC)NM + \min(N, M)^2\right)\varepsilon \end{aligned}$$

whenever $\mathcal{J}_p^{\alpha,\gamma}(\tilde{U}_p, \tilde{\Sigma}_p, \tilde{V}_p) < \left((KNM + \min(N, M)^2\right)\varepsilon$ and $\tilde{p} > e^{-L\varepsilon} p^2$. Combining this with Theorem 2.5, the following corollary follows directly.

Corollary 2.7. *For any small $\delta > 0$, we have*

$$\mathbb{P}\left(\min_p \mathcal{J}_p^{\alpha,\gamma}(\tilde{U}_p, \tilde{\Sigma}_{p,\tilde{p}}, \tilde{V}_p) < \delta\right) \geq \min(N, M)\, p_{N,M,\delta}^{\max(N,M)}\, (C_3\varepsilon)^{\mu(N,M,p_{N,M,\delta})}\ .$$

whenever $\gamma \leq (K-1)\,\varepsilon$, where $\varepsilon := \delta((K + CL)NM + \min(M, N)^2)^{-1}$, $p_{N,M,\delta}$ is stated as (2.11) and $\tilde{p} > e^{-L\varepsilon} p^2$ for some C, K, L. Moreover, whenever the above event happens, $||Y - \mathcal{I}_{p,\tilde{p}}^{\alpha,\gamma}(Y)||_{F,2}^2 < \delta$.

3 Multi-resolution analysis (MRA) of non-negative tri-factorization

In this section, we introduce an MRA framework based on the aforementioned tri-factorization. We notice that, for a matrix Y, especially for those representing an image, there are features of different scales in Y which usually represent different objects in the image. We aim at extracting these features of different scales and represent them in a sparse low-rank approximation in terms of tensor products. By combining the concepts of MRA and NMF, we obtain a framework which helps to get a sparse

representation of the features of multiple scales, ranging from the coarsest scale to the finest scale in the image Y.

For the sake of exposition, we shall introduce the following several operators which are very useful in the subsequent discussion. We first define an interpolation operator $\iota_s : M_{M\times N} \to M_{\frac{M}{r^s}\times\frac{N}{r^s}}$ as the following averaging operator:

$$\begin{aligned} \iota_s : M_{M\times N} &\to M_{\frac{M}{r^s}\times\frac{N}{r^s}} \\ \iota_s(Y) &:= \sum_{1\le i\le M/r^s, 1\le j\le N/r^s} \frac{1}{r^{2s}} \sum_{k,l\in Q_{I_i,J_j}} Y_{kl} e_i \otimes e_j \end{aligned} \tag{3.1}$$

where Q_{I_i,J_j} contains the entries $iM/r^s \le k < (i+1)M/r^s, jM/r^s \le l < (j+1)M/r^s$. Then we define

$$\mathcal{I}^{\alpha,\gamma}_{s,p} : M^+_{M\times N} \to M^+_{M\times N}$$

as follows,

$$\mathcal{I}^{\alpha,\gamma}_{s,p} := \iota_s^T \circ \mathcal{I}^{\alpha,\gamma}_p \circ \iota_s \,. \tag{3.2}$$

The approximation $\mathcal{I}^{\alpha,\gamma}_{s,p}(Y)$ represents the approximation of the $(s_{\max}-s)$-th layer of the image Y by NMF where $s_{\max} \le [\log(\min(N,M))/\log(r)]$. Similarly, we define

$$\mathcal{I}^{\alpha,\gamma}_{s,p,\tilde{p}} : M^+_{M\times N} \to M^+_{M\times N}$$

as

$$\mathcal{I}^{\alpha,\gamma}_{s,p,\tilde{p}} := \iota_s^T \circ \mathcal{I}^{\alpha,\gamma}_{p,\tilde{p}} \circ \iota_s \,, \tag{3.3}$$

which serves as a truncated approximation of the $(s_{\max}-s)$-th layer of Y.

Now we are ready to analyse this MRA framework. In fact, it is easy to see by combining the arguments in previous sections and the Poincare inequality that

$$\begin{aligned} ||Y - \iota_s^T \circ \mathcal{I}^{\alpha,\gamma}_{p,\tilde{p}} \circ \iota_s(Y)||^2_{F,2} &\le r^{2s}||\iota_s(Y) - \mathcal{I}^{\alpha,\gamma}_{p,\tilde{p}} \circ \iota_s(Y)||^2_{F,2} + \sum_{I,J} ||\nabla_\delta Y_{IJ}||^2_{F,2} \\ &\le r^{2s}\mathcal{J}^{\alpha,\gamma}_p(\tilde{U}_p, \tilde{\Sigma}_{p,\tilde{p}}, \tilde{V}_p) + \sum_{I,J} ||\nabla_\delta Y_{IJ}||^2_{F,2} \\ &\le r^{2s}\mathcal{J}^{\alpha,\gamma}_p(\tilde{U}_p, \tilde{\Sigma}_{p,}, \tilde{V}_p) + r^{2s}\left(Cr^{-2s}NM\log\left(\frac{p^2}{\tilde{p}}\right)\right) + \sum_{I,J} ||\nabla_\delta Y_{IJ}||^2_{F,2} \end{aligned}$$

where ∇_δ is the difference gradient operator defined as $(\nabla_\delta Y)_{i,j} = (Y_{i+1,j} - Y_{i,j}, Y_{i,j+1} - Y_{i,j})$, the matrix Y_{IJ} are the (I,J)-th block of the Y matrices, $[\tilde{U}_p, \tilde{\Sigma}_p, \tilde{V}_p]$ is an argument minimum of (2.1) with Y replaced by $\iota_s(Y)$ and $\tilde{\Sigma}_{p,\tilde{p}}$ is the truncation of $\tilde{\Sigma}_p$ up to $\tilde{p}$ as stated in the previous section.

Therefore if we can choose $[\tilde{U}_p, \tilde{\Sigma}_p, \tilde{V}_p]$ such that $\mathcal{J}^{\alpha,\gamma}_p(\tilde{U}_p, \tilde{\Sigma}_p, \tilde{V}_p) < r^{-2s}\left((KNM + \min(N,M)^2\right)\varepsilon$ and $\tilde{p} > e^{-L\varepsilon}p^2$, then

$$||Y - \mathcal{I}^{\alpha,\gamma}_{s,p,\tilde{p}}(Y)||^2_{F,2} \le \left((K+CL)NM + \min(N,M)^2\right)\varepsilon + \sum_{I,J} ||\nabla_\delta Y_{IJ}||^2_{F,2} \,.$$

From the discussions made in the previous section, The probability of the above event, denoted as $E_{p,\tilde{p},\delta}$, is bounded below by

$$\mathbb{P}(E_{p,\tilde{p},\delta}) \ge r^{-s}\min(N,M)\, p^{r^{-s}\max(N,M)}_{r^{-s}N,r^{-s}M,\delta}\,(C_3\varepsilon)^{\mu(r^{-s}N,r^{-s}M,p_{r^{-s}N,r^{-s}M,\delta})}$$

whenever $\gamma \le (K-1)\,\varepsilon$, where $p_{r^{-s}N,r^{-s}M,\delta}$ is defined as in (2.12).

Now, in general, we have no hope that either $||\nabla_\delta Y||_{F,2}^2$ or $\sum_{I,J}||\nabla_\delta Y_{IJ}||_{F,2}^2$ can be controlled, since we did not impose any regularity conditions for Y in general. However, if we further assume that Y has some regularity, for instance $\sum_{I,J}||\nabla_\delta Y_{IJ}||_{F,2}^2 < K_0 MN\varepsilon$, then

$$||Y - \mathcal{I}_{s,p,\tilde{p}}^{\alpha,C\varepsilon}(Y)||_{F,2}^2 \quad \leq \quad \left((K + CL + K_0)NM + \min(N,M)^2\right)\varepsilon\,.$$

Therefore we have the following theorem combining all the previous arguments and theorems.

Theorem 3.1. *For any small $\delta > 0$, let the event $E_{p,\tilde{p},\delta}$ be such that the following inequality holds:*

$$||Y - \mathcal{I}_{s,p,\tilde{p}}^{\alpha,\gamma}(Y)||_{F,2}^2 \leq \left((K + CL)NM + \min(N,M)^2\right)\varepsilon + \sum_{I,J}||\nabla_\delta Y_{IJ}||_{F,2}^2$$

where $\varepsilon := -r^{2s}\delta\left((K + CL + K_0)NM + \min(N,M)^2\right)^{-1}$ for some K_0, K, C, L. If all the $\tilde{p}$ corresponding to p is chosen such that $\tilde{p} > e^{L\varepsilon}p^2$, then we have for any s,

$$\mathbb{P}\left(\bigcup_p E_{p,\tilde{p},\delta}\right) \geq r^{-s}\min(N,M)\, p_{r^{-s}N,r^{-s}M,\delta}^{r^{-s}\max(N,M)}\,(C\varepsilon)^{\mu\left(r^{-s}N,r^{-s}M,p_{r^{-s}N,r^{-s}M,\delta}\right)} \tag{3.4}$$

whenever $\gamma \leq (K-1)\,\varepsilon$ where the function $\mu(\cdot,\cdot,\cdot)$ is defined as in (2.6) and $p_{r^{-s}N,r^{-s}M,\delta}$ is defined as in (2.11). Moreover, the following bound for the probability holds for all s and $p < r^{-s}\min(N,M)$, that whenever $\tilde{p} > e^{L\varepsilon}p^2$ and $\gamma \leq (K-1)\,\varepsilon$,

$$\mathbb{P}\left(E_{p,\tilde{p},\delta}\right) \geq \sum_{l=1}^{p}\left(l^{r^{-s}\max(N,M)} - (l-1)^{r^{-s}\max(N,M)}\right) l!\, r^{-s}\min(N,M)\;(C_3\varepsilon)^{\mu(r^{-s}N,r^{-s}M,l)}\,. \tag{3.5}$$

For a given small constant η, the growth of value of the last summation with respect to p can be controlled by η when $p > K_\eta\, p_{r^{-s}N,r^{-s}M,\delta}$ for some K_η depending on only η. Furthermore, when the above event $E_{p,\tilde{p},\delta}$ happens as well as the inequality $\sum_{I,J}||\nabla_\delta Y_{IJ}||_{F,2}^2 < K_0 MN\varepsilon$ holds, we have

$$||Y - \mathcal{I}_{s,p,\tilde{p}}^{\alpha,\gamma}(Y)||_{F,2}^2 \leq \delta\,. \tag{3.6}$$

Now that given a threshold δ, when $M = N$, if Y has the regularity such that $||\nabla_\delta Y||_{F,2}^2 < \tilde{K}\delta$ for some $\tilde{K} < 1$, then the above theorem infers that the lower bound of the probability of $||Y - \mathcal{I}_{s,p,\tilde{p}}^{\alpha,\gamma}(Y)||_{F,2}^2 < \delta$ is higher than that of $||Y - \mathcal{I}_{p,\tilde{p}}^{\alpha,\gamma}(Y)||_{F,2}^2 < \delta$ if $\tilde{p}$ is smartly chosen. Furthermore, for each s, the optimal choice of p has the same order as $p_{r^{-s}N,r^{-s}M,\delta}$, which is asymptotically

$$p \sim r^{-s/2}\sqrt{\frac{N}{\log N - 2s\log r}}\,, \tag{3.7}$$

with the memory complexity of $\mathcal{I}_{s,p,\tilde{p}}^{\alpha,\gamma}$ growing in the order $r^{-3s/2}\sqrt{\frac{N^3}{\log N - 2s\log r}}$ as N goes to infinity. This tells us that, by increasing s, the probability of fine approximation by NMF is increased as well as the memory complexity is decreased. Moreover, from numerical experiments, we can observe that for larger values of s, the resulted approximation $\mathcal{I}_{s,p,\tilde{p}}^{\alpha,\gamma}$ captures the coarser features of Y, with features finer and finer as s decreases.

4 Semi-smooth Newton method for non-negative factorizations

In this section, we describe the numerical algorithm to obtain the NMF of the image or big data Y as previously mentioned. We aim to utilize an efficient and cost-effective algorithm for this numerical process.

From numerical experiments, we notice that instead of finding the optimal $[\tilde{U}_p, \Sigma_p, \tilde{V}_p]$ inside the functional (2.1), it is enough to perform the following alternative NMF to obtain an approximation of $\mathcal{I}_p^{\alpha,\gamma}(Y)$ as follows:

$$Y \approx AV^T\,, A^T \approx \Sigma^T U^T\,,\ \text{then combine to get } Y \approx U\Sigma V^T\,. \tag{4.1}$$

where in each of the NMF, we minimize the functional (1.2) via a primal-dual active set semi-smooth Newton method [11, 12] which would be derived and stated below. The primal-dual active set semi-smooth Newton method is more advantageous than the classical methods suggested in [5] or [7] since it converges faster than the classical methods. We notice also that although this process does not obtain the optimal $[\tilde{U}_p, \Sigma_p, \tilde{V}_p]$ of the functional (2.1), this gives us a approximation of $\mathcal{I}_p^{\alpha,\gamma}(Y)$, which is good enough as confirmed by numerical experiments. It is also convenient to use the alternative NMF in practice, since the necessary optimality condition coming from a linearization of the functional that is used in the semi-smooth Newton method is more convenient to be evaluated numerically as one-sided matrix multiplications, than as a two-sided matrix multiplication which will be encountered when one aims to minimize (2.1) directly. Therefore numerically speaking, alternative factorization is more convenient to implement and less computationally expensive.

4.1 Primal-dual active set semi-smooth Newton method for NMF

Before we discuss the alternative NMF for a approximation of $\mathcal{I}_p^{\alpha,\gamma}(Y)$, we first discuss the optimization algorithm for the important non-convex problem (1.2). The primal-dual active set semi-smooth Newton method are proposed and derived in [11, 12] to solve either convex or non-convex non-smooth optimization problem effectively by combining the ideas of active sets and a Newton type update. In this section, we briefly introduce this method for solving the optimization of the non-smooth non-convex problem (1.2):

$$\min_{A\geq 0, P\geq 0} J(A,P) := ||Y - AP||_{F,2}^2 + \alpha||A||_{F,1} + \gamma||PP^T - I||_{F,1}. \tag{4.2}$$

4.1.1 Complementary Conditions

Before stepping into the optimization process, we shall first recall two complementary conditions for the characterization of some constraints conditions from [12], which is crucial for the development of the algorithm in the subsequent analysis. For this purpose, We first recall that for the function $|\cdot| : \mathbb{R} \to \mathbb{R}$, the sub-differential $\partial|\cdot|$ is the set-valued signum function defined as

$$\partial|\cdot|(x) = \begin{cases} 1 & \text{if } x > 0\,, \\ [-1,1] & \text{if } x = 0\,, \\ -1 & \text{if } x < 0\,. \end{cases} \tag{4.3}$$

We also recall the following celebrated complementarity condition [12] which characterizes the set-value sub-differential $\partial|\cdot|$ as follows:

$$\lambda = \frac{\lambda + cx}{\max(1, |\lambda + cx|)} \tag{4.4}$$

for any given $c > 0$. It is handy to obtain the following lemma by directly comparison of point-wise values:

Theorem 4.1.

$$\lambda = \frac{\lambda + cx}{\max(1, |\lambda + cx|)} \Leftrightarrow \lambda \in \partial|\cdot|(x) \tag{4.5}$$

where ∂ is regarded as the sub-differential operator.

Proof. Suppose $\lambda = \frac{\lambda+cx}{\max(1,|\lambda+cx|)}$. If $|\lambda + cx| \leq 1$, then $\lambda = \lambda + cx$, which gives $x = 0$, hence $|\lambda| \leq 1$ and $\lambda \in \partial|\cdot|(x)$. Else if $|\lambda + cx| > 1$, then $\lambda = \frac{\lambda+cx}{|\lambda+cx|} = \pm 1$. Assume that $\lambda = 1$, then we have $|1 + cx| > 1$, which directly gives $x > 0$, and therefore $\lambda \in \partial|\cdot|(x)$. The case for $\lambda = -1$ is the similar.

On the other hand, suppose that $\lambda \in \partial|\cdot|(x)$. If $x = 0$, then $|\lambda| \leq 1$ and therefore $\lambda = \frac{\lambda}{\max(1,|\lambda|)} = \frac{\lambda+cx}{\max(1,|\lambda+cx|)}$. Else if $x > 0$, then $\lambda = 1$ and $|\lambda + cx| > 1$, therefore $\frac{\lambda+cx}{\max(1,|\lambda+cx|)} = \frac{\lambda+cx}{|\lambda+cx|} = \lambda$. The case for $x < 0$ is similar. □

Note that in the above complementary condition, the choice of c is arbitrary. However, for a practical implementation of an algorithm using the complementary condition, c will be chosen as a fixed constant which acts as a parameter for a stable implementation.

Now for any matrix A, note that $||A||_{F,1} = \sum_{i,j} |A_{i,j}|$, we then have the set-valued sub-differential function $\partial||\cdot||_{F,1}(A)$ as follows:

$$(\,\partial||\cdot||_{F,1}[A]\,)_{i,j} = \begin{cases} 1 & \text{if } A_{i,j} > 0\,, \\ [-1,1] & \text{if } A_{i,j} = 0\,, \\ -1 & \text{if } A_{i,j} < 0\,. \end{cases} \tag{4.6}$$

Using the aforementioned celebrated complementarity condition for a dual variable λ, we can then put

$$\lambda_{i,j} = \frac{\lambda_{i,j} + cA_{i,j}}{\max(1, |\lambda_{i,j} + cA_{i,j}|)} \Leftrightarrow \lambda \in \partial||\cdot||_{F,1}(A)\,. \tag{4.7}$$

or in short $\lambda = \frac{\lambda+cA}{\max(1,|\lambda+cA|)}$, where the division, the maximum and the absolute values are taken pointwise.

The second complementary condition we would like to introduce is to characterise an inequality constrain $x \geq 0$, which is introduced in [12]. We sketch the argument from [12] to arrive at our desired complementary condition. Consider the following constrained optimization for a functional $F : \mathbb{R}^N \to \mathbb{R}$.

$$\min F(x) \quad \text{subject to} \quad x \geq 0\,. \tag{4.8}$$

We consider the following equivalent formulation of the variation problem with the same necessary optimality condition with the help of the introduction of a dummy variable z and the augmented Lagrangian formulation with a dual variable μ and constant c:

$$\min F(x) + \langle \mu, x - z\rangle + \frac{c}{2}||x - z||_2^2 \quad \text{subject to} \quad x = z \text{ and } z \geq 0\,. \tag{4.9}$$

Minimizing over $z \geq 0$, with the convexity of this functional in z, we obtain the following entry-wise necessary and sufficient condition for z with the help of the linearization $\langle \mu, \cdot\rangle + c\langle x - z, \cdot\rangle$ of the functional in z :

$$\begin{cases} \mu_i + c(x_i - z_i) < 0 & \text{if } z_i = 0\,, \\ \mu_i + c(x_i - z_i) = 0 & \text{if } z_i > 0\,, \end{cases} \tag{4.10}$$

or equivalently

$$\begin{cases} 0 > \mu_i + cx_i & \text{if } z_i = 0\,, \\ z_i = \frac{\mu_i + cx_i}{c} & \text{if } z_i > 0\,. \end{cases} \tag{4.11}$$

which gives the following unique minimizer for the variable z

$$z_i = \max\left(0, \frac{\mu_i + cx_i}{c}\right)\,. \tag{4.12}$$

or in compact form $z = \max(0, \frac{\mu+cx}{c})$ where the operations are understood point-wisely. Now putting this into the expression $\langle \mu, x-z\rangle + \frac{c}{2}||x-z||_2^2$, we have by direct computation

$$\begin{aligned} & \langle \mu, x-z\rangle + \frac{c}{2}||x-z||_2^2 \\ = \quad & \frac{1}{c}\langle \mu, \min(cx, -\mu)\rangle + \frac{1}{2c}||\min(cx,-\mu)||_2^2 \\ = \quad & \frac{1}{2c}\left(||\min(cx,-\mu)+\mu||_2^2 - ||\mu||_2^2\right) \\ = \quad & \frac{1}{2c}\left(||\min(\mu+cx,0)||_2^2 - ||\mu||_2^2\right) . \end{aligned} \tag{4.13}$$

Substituting this expression back to the augmented lagrangian, we obtain the following equivalent minimization problem

$$\min F(x) + \frac{1}{2c}\left(||\min(\mu+cx,0)||_2^2 - ||\mu||_2^2\right) \tag{4.14}$$

The necessary optimality condition of the above problem is therefore given by the following set-valued equation

$$\begin{cases} 0 \in \partial F(x) + \min(\mu+cx,0)\partial\min(\cdot,0)\left(\min(\mu+cx,0)\right) \\ 0 \in -\mu + \min(\mu+cx,0)\partial\min(\cdot,0)\left(\min(\mu+cx,0)\right) . \end{cases} \tag{4.15}$$

Equivalently, by point-wise comparison, if $\mu + cx \leq 0$, then $\min(\mu+cx,0)\partial\min(\cdot,0)\left(\min(\mu+cx,0)\right) = 0$ as a set value function, then from the above necessary optimality condition, we have $\mu = 0$ and $\min(\mu+cx,0) = 0$, and therefore $\mu = \min(\mu+cx,0)$; whereas, if $\mu + cx > 0$, then we have that the following is satisfies $\min(\mu+cx,0)\partial\min(\cdot,0)\left(\min(\mu+cx,0)\right) = \min(\mu+cx,0)$; which gives, from the necessary optimality condition, that $\mu = \min(\mu+cx,0)$. Therefore, after considering both of the cases, we arrive at the following equivalent optimality condition for μ, namely $\mu = \min(\mu+cx,0)$. Therefore, as in [12], we arrive at the following necessary optimality condition of the above problem.

Theorem 4.2. *The necessary optimality condition for the minimization problem* (4.8) *is*

$$0 \in \partial F(x) + \mu \quad \textit{and} \quad \mu = \min(\mu+cx,0) . \tag{4.16}$$

where ∂ is regarded as the sub-differential operator.

The condition $\mu = \min(\mu+cx,0)$ for the dual variable μ is regarded as a complementary condition in [12] which serves as a characterization of the constraint $x \geq 0$. This complementary condition may also be regarded as a project of the solution to the convex set as the epigraph defined by the constraint.

4.1.2 Necessary optimality condition for optimization of (4.2)

Combining the above two theorems from the last subsection, we arrive at the necessary optimality condition for optimization of (4.2), which we intend to solve numerically using semi-smooth Newton method. In fact, directly applying Theorem 4.2 and calculating the sub-differentials, we have the following necessary optimality condition for (4.2) using the primal-dual variables (A, P, μ_A, μ_P) for a given c_1:

$$\begin{cases} 0 & \in \partial_A J(A,P) + \mu_A = 2APP^T - 2YP^T + \mu_A + \alpha\partial||\cdot||_{F,1}(A) \\ \mu_A & = \min(\mu_A + c_1 A, 0) . \\ 0 & \in \partial_P J(A,P) + \mu_P = -2A^TY + 2A^TAP + \mu_P + \gamma\partial||\cdot||_{F,1}(PP^T - I)\left(P + T\circ P^T \circ T\right) \\ \mu_P & = \min(\mu_P + c_1 P, 0) . \end{cases} \tag{4.17}$$

where $T : M_{M\times N} \to M_{N\times M}$, $A \to A^T$ is the transpose operator. Now, applying Theorem 4.1 to the above system and introducing several more variables R, L, we arrive at the following conclusion.

Theorem 4.3. *The following system of equations with the primal-dual variables* $(A, P, R, L, \mu_A, \lambda_A, \mu_P, \lambda_L)$ *provides the necessary optimality condition for* (4.2) *for a given* c_1, c_2*:*

$$\begin{cases} 0 &= 2APP^T - 2YP^T + \mu_A + \alpha\lambda_A \\ \lambda_A &= \frac{\lambda_A + c_2 A}{\max(1, |\lambda_A + c_2 A|)} \\ \mu_A &= \min(\mu_A + c_1 A, 0) \\ 0 &= -2A^T Y + 2A^T AP + \mu_P + \gamma\lambda_L R \\ L &= PP^T - I \\ R &= P \circ T + T \circ P^T \circ T \\ \lambda_L &= \frac{\lambda_L + c_2 L}{\max(1, |\lambda_L + c_2 L|)} \\ \mu_P &= \min(\mu_P + c_1 P, 0)\,. \end{cases} \tag{4.18}$$

where $T : M_{M\times N} \to M_{N\times M}$ *is the transpose operator.*

This is the starting point of the algorithm for solving the variational problem (4.2).

4.1.3 Semi-smooth Newton strategy

From the previous subsection, we derived a necessary optimality condition for solving the optimization problem (4.2). In what follows, we shall develop a semi-smooth Newton method for solving such a system, which can be immediately shown to be Newton differentiable [12]. To further develop our algorithm, we separate the variables $(A, P, R, L, \mu_A, \lambda_A, \mu_P, \lambda_L)$ into three set, i.e. (A, μ_A, λ_A), (P, μ_P) and (L, R, λ_L), and perform a directional search. This is computationally advantageous to separate the variables and solve independently by considering the variables from the other sides as constants thanks to the following reasons. On one hand, the separate systems are easier for us to perform active set method which greatly reduces computational cost, and on the other hand, each system consists of much fewer non-linear terms, and is therefore more stable when performing semi-smooth Newton method. With these motivation, we therefore separate (4.18) into three separate set of equations and consider the following three systems:

(1) For a fixed P, solve for (A, μ_A, λ_A) such that

$$\begin{cases} 0 &= 2APP^T - 2YP^T + \mu_A + \alpha\lambda_A \\ \lambda_A &= \frac{\lambda_A + c_2 A}{\max(1, |\lambda_A + c_2 A|)} \\ \mu_A &= \min(\mu_A + c_1 A, 0)\,. \end{cases} \tag{4.19}$$

(2) For a fixed A, L, R, λ_L, solve for (P, μ_P) such that

$$\begin{cases} 0 &= -2A^T Y + 2A^T AP + \mu_P + \gamma\lambda_L R \\ \mu_P &= \min(\mu_P + c_1 P, 0)\,. \end{cases} \tag{4.20}$$

(3) For a fixed P, solve for (L, R, λ_L) such that

$$\begin{cases} L &= PP^T - I \\ R &= P \circ T + T \circ P^T \circ T \\ \lambda_L &= \frac{\lambda_L + c_2 L}{\max(1, |\lambda_P + c_2 L|)}\,. \end{cases} \tag{4.21}$$

Now if we define the following active sets and inactive sets:

$$\begin{aligned} \mathcal{A}_{A,1} &= \{(i,j) : (\mu_A)_{i,j} + c_1 A_{i,j} > 0\} & \mathcal{I}_{A,1} &= \{(i,j) : (\mu_A)_{i,j} + c_1 A_{i,j} \le 0\} \\ \mathcal{A}_{A,2} &= \{(i,j) : |(\lambda_A)_{i,j} + c_2 A_{i,j}| \le 1\} & \mathcal{I}_{A,2} &= \{(i,j) : |(\lambda_A)_{i,j} + c_2 A_{i,j}| > 1\} \\ \mathcal{A}_P &= \{(i,j) : (\mu_P)_{i,j} + c_1 P_{i,j} > 0\} & \mathcal{I}_P &= \{(i,j) : (\mu_P)_{i,j} + c_1 P_{i,j} \le 0\} \\ \mathcal{A}_L &= \{(i,j) : |(\lambda_L)_{i,j} + c_2 L_{i,j}| \le 1\} & \mathcal{I}_L &= \{(i,j) : |(\lambda_L)_{i,j} + c_2 L_{i,j}| > 1\}\,, \end{aligned}$$

we can further reduce the system into the following much simpler form thanks to the direct substitution and point-wise comparison of the complementary conditions:
(1) For a fixed P, on $\mathcal{A}_{A,1}\bigcup\mathcal{A}_{A,2}$, we have $A=0$, whereas on $\mathcal{I}_{A,1}\bigcap\mathcal{I}_{A,2}$, (A,λ_A) satisfies

$$\begin{cases} 0 &= 2APP^T - 2YP^T + \alpha\lambda_A \\ 0 &= \lambda_A|\lambda_A + c_2A| - (\lambda_A + c_2A)\,. \end{cases} \tag{4.22}$$

(2) For a fixed A, L, R, λ_L, on $\mathcal{A}_P$, we have $P=0$, whereas on $\mathcal{I}_P$, P satisfies

$$0 = -2A^TY + 2A^TAP + \gamma\lambda_L R\,. \tag{4.23}$$

(3) For a fixed P, on $\mathcal{A}_L$, we have $L=0$, whereas on $\mathcal{I}_L$, (L,R,λ_L) satisfies

$$\begin{cases} L &= PP^T - I \\ R &= P\circ T + T\circ P^T\circ T \\ 0 &= \lambda_L|\lambda_L + c_2L| - (\lambda_L + c_2L)\,. \end{cases} \tag{4.24}$$

For the non-linear constraint on λ_A in the first system as well as that on λ_L in the third system, we propose a semi-smooth Newton step update as in [11] to solve the corresponding equations. One might suggest the Uzawa explicit method [12] instead, however this method is only conditionally stable and the convergence is slower than a semi-smooth update, and therefore we prefer the Newton step update. In here, we only sketch a derivation of the method following [11]. We first consider the first system. In fact, assume that (A,λ_A) are perturbed to (A^h,λ_A^h) such that the increment is of order $O(h)$ and satisfies $0=\lambda_A|\lambda_A+c_2A|-(\lambda_A+c_2A)$, we actually have

$$(\lambda_A^h-\lambda_A)|\lambda_A+c_2A| + \lambda_A\left(|\lambda_A+c_2A| + \frac{\lambda_A+c_2A}{|\lambda_A+c_2A|}(\lambda_A^h+c_2A^h-\lambda_A-c_2A)\right) - (\lambda_A^h+c_2A^h) + (\lambda_A+c_2A) = O(h^2)$$

which gives the following Newton update from (A,λ_A) to (A^+,λ_A^+):

$$\lambda_A^+|\lambda_A+c_2A| + \lambda_A\frac{\lambda_A+c_2A}{|\lambda_A+c_2A|}\left(\lambda_A^+ + c_2A^+\right) = \lambda_A|\lambda_A+c_2A| + (\lambda_A^+ + c_2A^+)$$

We then suggest the following damping term and regularization following [12] for the update

$$\lambda_A^+|\lambda_A+c_2A| + \frac{\lambda_A+c_2A}{|\lambda_A+c_2A|}\left(\lambda_A^+ + c_2A^+\right)\frac{\theta\lambda_A}{\max(1,|\lambda_A|)} = |\lambda_A+c_2A|\frac{\theta\lambda_A}{\max(1,|\lambda_A|)} + (\lambda_A^+ + c_2A^+)$$

where θ is automatically chosen to achieve stability and the regularizer $\lambda_A/\max(1,|\lambda_A|)$ is set to automatically constrain λ_A to be in $[-1,1]$. Following [11], we set $\theta\,|\lambda_A+c_2A|/\left(|\lambda_A+c_2A| - 1 + \theta\frac{\lambda_A(\lambda_A+c_2A)}{\max(1,|\lambda_A|)|\lambda_A+c_2A|}\right) = 1$ which gives $\theta\le 1$ to simplify the iteration and get the following update after direct substitution,

$$0 = \lambda_A^+ - c_2\frac{1-a_Ab_A}{d_A-1}A^+ + a_A$$

where $a_A = \frac{\lambda_A}{\max(1,|\lambda_A|)}$, $b_A = \frac{\lambda_A+c_2A}{|\lambda_A+c_2A|}$ and $d_A = |\lambda_A+c_2A|$, which is used as the semi-smooth update for the first system

$$\begin{cases} 0 &= 2A^+PP^T - 2YP^T + \alpha\lambda_A^+ \\ 0 &= \lambda_A^+ - \frac{c_2}{d_A-1}\left(I - a_Ab_A^T\right)A^+ + a_A\,. \end{cases} \tag{4.25}$$

We can linearize the constraint for the variable λ_L in the third system with this same method.

For the other two variables (L,R) in the third system, we can solve for them by direct substitutions; however, it is not easy. Although the equation $R = P\circ T + T\circ P^T\circ T$ is linear, this is computational

costly to evaluate as the transpose operator T is involved. We therefore derive a semi-smooth update for R from L and P instead of a direct substitution. Assume (L, R, P) are perturbed to (L^h, R^h, P^h) such that the increment is of order $O(h)$ and satisfies $L = PP^T - I$, we then actually get that

$$(L^h - L) = R^h(P^h - P) + O(h^2)\,,$$

and hence we get the following update after linearization,

$$R^+ \quad = (L^+ - L)[P^+ - P]^{-1}\,.$$

Combining the above update with the aforementioned strategy for λ_L, we obtain the following semi-smooth Newton update from (L, R, P) to (L^+, R^+, P^+) for the third system:

$$\begin{cases} L^+ & = P^+(P^+)^T - I \\ R^+ & = (L^+ - L)[P^+ - P]^{-1} \\ \lambda_L^+ & = \frac{c_2}{d_L - 1}\left(I - a_L b_L^T\right) L^+ - a_L \end{cases} \tag{4.26}$$

where $a_L = \frac{\lambda_L}{\max(1, |\lambda_L|)}$, $b_L = \frac{\lambda_L + c_2 L}{|\lambda_L + c_2 L|}$ and $d_L = |\lambda_L + c_2 L|$.

Using the numerical values of the variables in the previous iteration, we can predict the locations of the active sets, and therefore greatly reduce the computational cost of the algorithm.

4.1.4 The algorithm

Combining all the techniques and results from the previous subsections, we are ready to propose the following primal-dual active set semi-smooth Newton method for solving the system (4.18) to tackle the minimization problem (4.2).

Algorithm 1

Step 1 Fix two parameters c_1, c_2. Initialize $(A^0, P^0, \mu_A^0, \lambda_A^0, \mu_P^0, \lambda_P^0)$.

Step 2 For $k = 0, ..., K$,

Step 2.1 Compute $\mu_A^{(k)}$ as

$$\mu_A^{(k)} := -2A^{(k)}P^{(k)}{P^{(k)}}^T + 2Y(P^{(k)})^T - \alpha\lambda_A^{(k)}\,.$$

Step 2.2 Set the active sets $\mathcal{A}_{A,i}^k$ and inactive sets $\mathcal{I}_{A,i}^k$ for $i = 1, 2$ respectively as follows:

$$\begin{aligned} \mathcal{A}_{A,1}^{(k)} &= \{(i,j) : (\mu_A)_{i,j}^{(k)} + c_1 A_{i,j}^{(k)} > 0\} & \mathcal{I}_{A,1}^{(k)} &= \{(i,j) : (\mu_A)_{i,j}^{(k)} + c_1 A_{i,j}^{(k)} \le 0\} \\ \mathcal{A}_{A,2}^{(k)} &= \{(i,j) : |(\lambda_A)_{i,j}^{(k)} + c_2 A_{i,j}^{(k)}| \le 1\} & \mathcal{I}_{A,2}^{(k)} &= \{(i,j) : |(\lambda_A)_{i,j}^{(k)} + c_2 A_{i,j}^{(k)}| > 1\}\,. \end{aligned}$$

Step 2.3 Compute $a_A^{(k)}, b_A^{(k)}, d_A^{(k)}$ as

$$a_A^{(k)} := \frac{\lambda_A^{(k)}}{\max(1, |\lambda_A^{(k)}|)}\,, \quad b_A^{(k)} := \frac{\lambda_A^{(k)} + c_2 A^{(k)}}{|\lambda_A^{(k)} + c_2 A^{(k)}|} \quad \text{and} \quad d_A^{(k)} := |\lambda_A^{(k)} + c_2 A^{(k)}|\,.$$

Step 2.4 On $\mathcal{A}_{A,1}^{(k)} \bigcup \mathcal{A}_{A,2}^{(k)}$, set $A^{(k+1)} := 0$. On $\mathcal{I}_{A,1}^{(k)} \bigcap \mathcal{I}_{A,2}^{(k)}$, solve $(A^{(k+1)}, \lambda_A^{(k+1)})$ from the system

$$\begin{cases} 0 & = 2A^{(k+1)}P^{(k)}{P^{(k)}}^T - 2Y(P^{(k)})^T + \alpha\lambda_A^{(k+1)} \\ 0 & = \lambda_A^{(k+1)} - \frac{c_2}{d_A^{(k)} - 1}\left(I - a_A^{(k)}[b_A^{(k)}]^T\right) A^{(k+1)} + a_A^{(k)}\,. \end{cases}$$

Step 2.5 Compute $\mu_P^{(k)}$ as

$$\mu_P^{(k)} := 2(A^{(k+1)})^T Y - 2(A^{(k+1)})^T A^{(k+1)} P^{(k)} - \gamma \lambda_L^{(k)} R^{(k)} .$$

Step 2.6 Set the active sets $\mathcal{A}_P^{(k)}$ and inactive sets $\mathcal{I}_P^{(k)}$ respectively as follows:

$$\mathcal{A}_P^{(k)} = \{(i,j) \,:\, (\mu_P)_{i,j}^{(k)} + c_1 P_{i,j}^{(k)} > 0\} \qquad \mathcal{I}_P^{(k)} = \{(i,j) \,:\, (\mu_P)_{i,j}^{(k)} + c_1 P_{i,j}^{(k)} \le 0\} .$$

Step 2.7 On $\mathcal{A}_P^{(k)}$, set $P^{(k+1)} := 0$. On $\mathcal{I}_P^{(k)}$, solve $P^{(k+1)}$ from

$$0 \;\; = -2(A^{(k+1)})^T Y + 2(A^{(k+1)})^T A^{(k+1)} P^{(k+1)} + \gamma \lambda_L^{(k)} R^{(k)} .$$

Step 2.8 Set the active sets $\mathcal{A}_L^{(k)}$ and inactive sets $\mathcal{I}_L^{(k)}$ respectively as follows:

$$\mathcal{A}_L^{(k)} = \{(i,j) \,:\, |(\lambda_L)_{i,j}^{(k)} + c_2 L_{i,j}^{(k)}| \le 1\} \qquad \mathcal{I}_L^{(k)} = \{(i,j) \,:\, |(\lambda_L)_{i,j}^{(k)} + c_2 L_{i,j}^{(k)}| > 1\} .$$

Step 2.9 Compute $a_L^{(k)}, b_L^{(k)}, d_L^{(k)}$ as

$$a_L^{(k)} := \frac{\lambda_L^{(k)}}{\max(1, |\lambda_L^{(k)}|)} , \quad b_L^{(k)} := \frac{\lambda_L^{(k)} + c_2 L^{(k)}}{|\lambda_L^{(k)} + c_2 L^{(k)}|} \quad \text{and} \quad d_L^{(k)} := |\lambda_L^{(k)} + c_2 L^{(k)}| .$$

Step 2.10 On $\mathcal{A}_L^{(k)}$, set $L^{(k+1)} = 0$. On $\mathcal{I}_L^{(k)}$, evaluate $(L^{(k+1)}, R^{(k+1)}, \lambda_L^{(k+1)})$ as

$$\begin{cases} L^{(k+1)} & = P^{(k+1)}(P^{(k+1)})^T - I \\ R^{(k+1)} & = \left(L^{(k+1)} - L^{(k)}\right)[P^{(k+1)} - P^{(k)}]^{-1} \\ \lambda_L^{(k+1)} & = \frac{c_2}{d_L^{(k)} - 1}\left(I - a_L^{(k)}[b_L^{(k)}]^T\right) L^{(k+1)} - a_L^{(k)} . \end{cases}$$

Step 2.11 Check the stopping criterion.

A natural choice of the stopping criterion is based on checking the change of the active sets: if the active sets for two consecutive iterations are the same, we may stop the iteration [11, 12]. We can see that, as A, P, L becomes more and more sparse, the system of linear equations solved in the above iteration drip drastically and therefore the inversion of the linear system involved will be less computational costly as k increases.

Finally, for effective implementation of the algorithm, the following remarks are in force:

1. With the enforcement of the constraint $A \ge 0$ by the dual variable μ_A, if the initial guess $A^{(0)}$ is chosen to be non-negative, then we automatically have $A^{(k)} \ge 0$ for all k. The algorithm can therefore be simplified by automatically setting the dual variable $\lambda_A^{(k)}$ to be a constant $\lambda_A^{(k)} = 1$ and drop the the active/inactive sets $\mathcal{A}_{A,2}^{(k)}$ and $\mathcal{I}_{A,2}^{(k)}$.

2. In order to further simplify the algorithm, we may normalize the row vectors of P after **Step 2.7** so that PP^T has unit diagonal entries. If this step is added, then $L^{(k)} \ge 0$ for all k. $\lambda_L^{(k)}$ can be automatically set as a constant $\lambda_L^{(k)} = 1$ and $\mathcal{A}_L^{(k)}$ and $\mathcal{I}_L^{(k)}$ can be dropped.

3. In the development of the algorithm above, we assume $Y \ge 0$ entry-wise, and therefore the constraint $A \ge 0$ entry-wise is legitimate to be enforced. This non-negativity condition for A is however infeasible and shall be dropped if Y is not nonnegative entry-wise. In this case, nonetheless, we can still utilize the above algorithm for a non-negative factorization with the following minor modification: drop the dual variable μ_A and the active/inactive sets $\mathcal{A}_{A,1}^{(k)}$ and $\mathcal{I}_{A,1}^{(k)}$.

4.2 Numerical algorithm for NMF of an image

With the availability of **Algorithm 1** to minimize the functional (4.2), we are ready to propose the following algorithm to approximate $\mathcal{I}_p^{\alpha,\gamma}(Y)$ in (2.2) and $\mathcal{I}_{p,\tilde{p}}^{\alpha,\gamma}(Y)$ in (2.20) for the NMF of an image Y which is convenient in application:

Algorithm 2

Step 1 Fix four parameters α, γ, p, $\tilde{p}$.

Step 2 Minimize

$$\min_{A\geq 0, V\geq 0} ||Y - AV^T||_{F,2}^2 + \alpha||A||_{F,1} + \gamma||V^TV - I||_{F,1}.$$

using **Algorithm 1** to obtain a minimizer $[A_0, V_0]$.

Step 3 Minimize

$$\min_{\Sigma\geq 0, U\geq 0} ||A_0^T - \Sigma^T U^T||_{F,2}^2 + \alpha||\Sigma||_{F,1} + \gamma||U^TU - I||_{F,1}.$$

using **Algorithm 1** to obtain a minimizer $[\Sigma_0, U_0]$.

Step 4 Compute $\mathcal{I}_p^{\alpha,\gamma}(Y)$ from $[U_0, \Sigma_0, V_0]$ as

$$\mathcal{I}_p^{\alpha,\gamma}(Y) := U_0\Sigma_0V_0^T\,.$$

Step 5 Sort the entries of Σ_0 from the largest to the smallest, enumerated as $\sigma_{i_1j_1} \geq \sigma_{i_2j_2} \geq .. \geq \sigma_{i_{p^2}j_{p^2}}$.

Step 6 Compute $\tilde{\sigma}_l := \sigma_{i_lj_l} e_l \otimes e_l$ and evaluate

$$\Sigma_{0,\tilde{p}} := \sum_{l=1}^{\tilde{p}} \tilde{\sigma}_l.$$

Step 7 Evaluate $\mathcal{I}_{p,\tilde{p}}^{\alpha,\gamma}$ as

$$\mathcal{I}_{p,\tilde{p}}^{\alpha,\gamma}(Y) := U_0\Sigma_{0,\tilde{p}}V_0^T\,.$$

4.3 Numerical algorithm for MRA using NMF

Finally we state the algorithm for the MRA framework based on the results from NMF.

Algorithm 3

Step 1 Fix a scaling parameter r and s_{max} such that $s_{max} < \log N/\log r$.

Step 2 Set two parameters α, γ and two arrays of parameters $[p(1), ..., p(s_{max})]$, $[\tilde{p}(1), ..., \tilde{p}(s_{max})]$.

Step 3 For $s = 1, ..., s_{max}$,

Step 3.1 Compute $\iota_s(Y)$ as in (3.1).

Step 3.2 Calculate $\mathcal{I}_{p(s),\tilde{p}(s)}^{\alpha,\gamma}[\iota_s(Y)]$ using **Algorithm 2**.

Step 3.3 Calculate $\mathcal{I}_{s,p(s),\tilde{p}(s)}^{\alpha,\gamma}(Y)$ as $\mathcal{I}_{s,p(s),\tilde{p}(s)}^{\alpha,\gamma}(Y) := \iota_s^T\mathcal{I}_{p(s),\tilde{p}(s)}^{\alpha,\gamma}\iota_s(I)$.

5 Applications to photo images, EIT and DOT images from DSMs

In this section we shall present of some of the applications of both the NMF and the MRA framework of NMF suggested in the previous sections. We investigate two applications, the first being an MRA for photo images using NMF, and the second being an NMF over the images from an inverse problem algorithm for a broad class of coefficient determination inverse problem. In the first application, we aim at capturing features of different scales in the image and obtain a sparse low-rank representation of them; while in the second application, we aim to identify the major components in the image, which represents some structures in the coefficients to be determined in the corresponding inverse problems, and remove noise from the reconstruction. We observe, from the numerical experiments of both of the applications, that a much fewer number of p than that being suggested by (2.19) and (3.7) are necessary for good approximation of the data involved. This is very desirable for memory reduction and feature identifications.

5.1 Applications to photo images

We perform an MRA over several grey-scaled photo images Y using NMF. In all of the following examples, we utilize **Algorithm 3** for the MRA to approximate $\mathcal{I}^{\alpha,\gamma}_{s,p(s),\tilde{p}(s)}(Y)$, in which **Algorithm 2** is used to calculate $\mathcal{I}^{\alpha,\gamma}_{p(s),\tilde{p}(s)}[\iota_s(Y)]$ and **Algorithm 1** is used to minimize (4.2) .

Example 1. In this example, we set Y as the grey image presented in Figure 1. We use the following set of parameters in **Algorithm 3** as

$$r = 2, \quad s_{max} = [\log(\min(N, M))/\log(r) - 3], \quad \alpha = 0.2, \quad \gamma = 0.02,$$

as well as the two arrays of parameters

$$p(s) = [20\log(s_{max} - s + e)], \quad \tilde{p}(s) = [50\log(s_{max} - s + e)^2(s_{max} - s + 1)]$$

where $1 \le s \le s_{max}$ and $[\cdot]$ denotes the floor function. The parameters c_1, c_2 in **Algorithm 1** are both set to be 1. The resulting images are shown in Figure 2. The memory compression ratios for the $(s_{max} - s)$-th layer, defined as the quotient between the memory size of $\mathcal{I}^{\alpha,\gamma}_{s,p(s),\tilde{p}(s)}(Y)$ and that of Y, are shown as follows:

$s_{max} - s$	:	1	2	3	4	5	6	Total
compression ratio	:	0.0017	0.0046	0.0108	0.0230	0.0495	0.1015	0.1910

We can see from Figure 2 that details of different scales are captured in different layers, while a reasonably low compression ratio is attained.

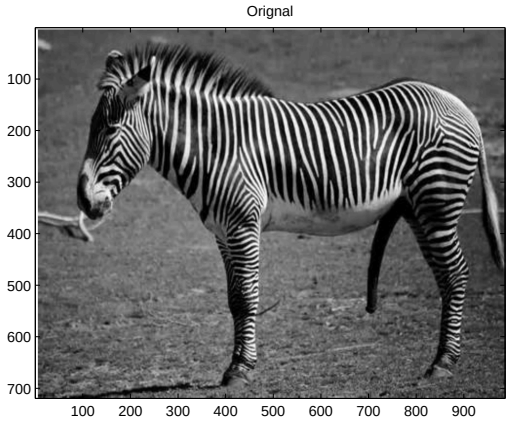


Figure 1: Original image in Example 1

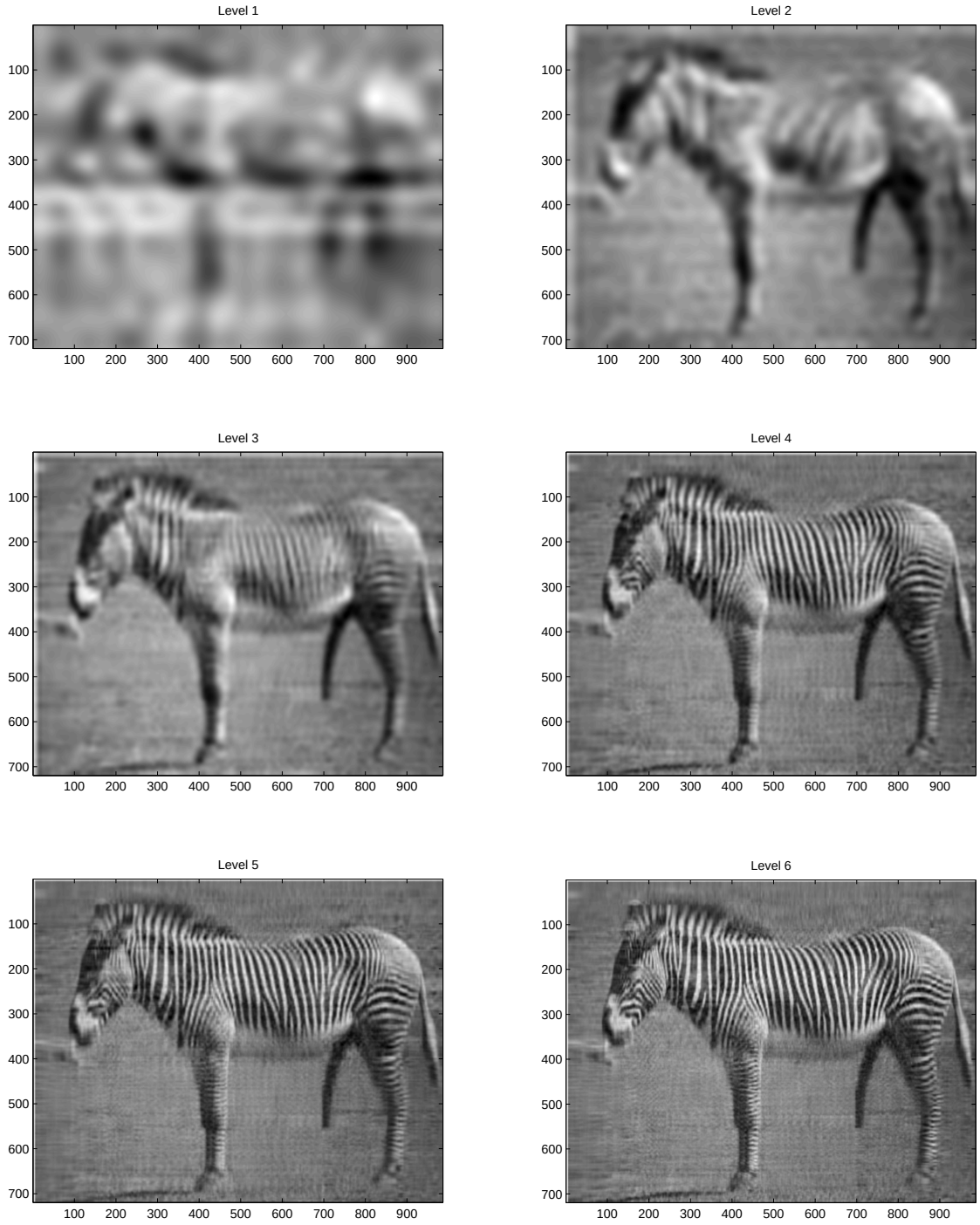


Figure 2: MRA for the image in Example 1 using NMF

Example 2. In this example, we set Y as the image presented in Figure 3. The parameters are the

same as in the previous example. The resulting images are shown in Figure 4. The memory compression ratios for the $(s_{max} - s)$-th layer are shown as follows:

$s_{max} - s$	:	1	2	3	4	Total
compression ratio	:	0.0053	0.0146	0.0343	0.0733	0.2861

We can see from Figure 4 finer and finer details are present as the layer number increases, while a reasonably low compression ratio is attained.

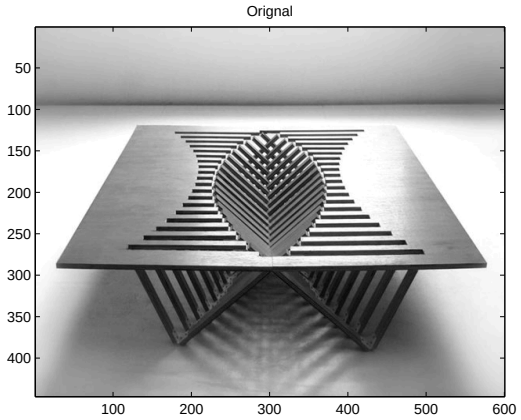
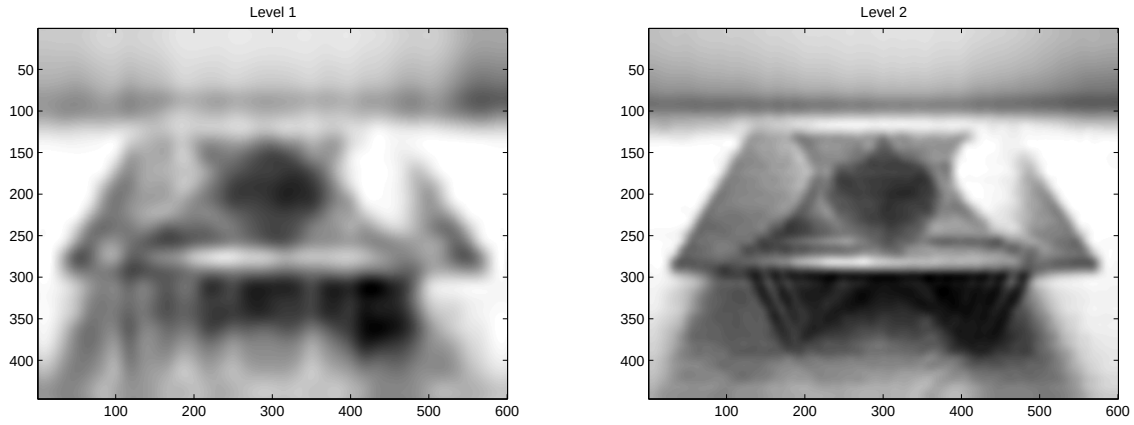


Figure 3: Original image in Example 2



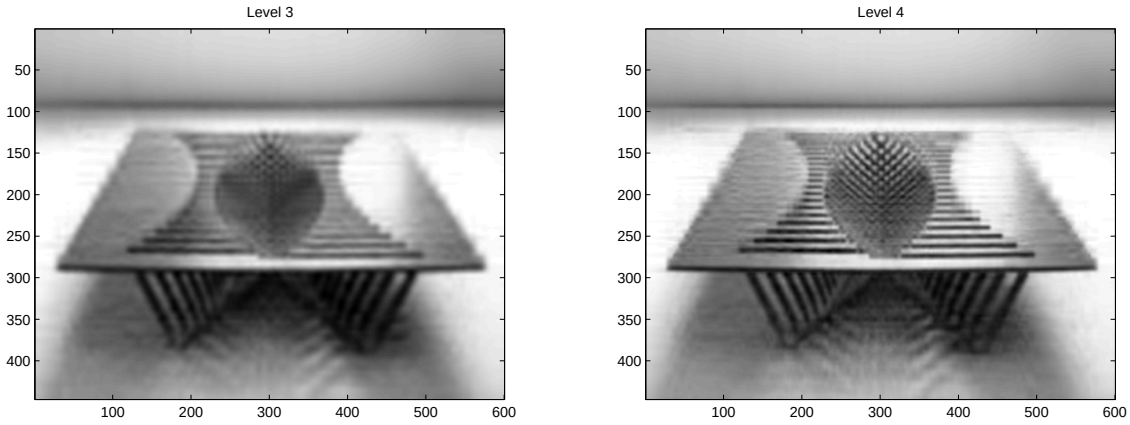


Figure 4: MRA for the image in Example 2 using NMF

Example 3. In this example, Y is set as the image in Figure 5. A same set of parameters is used, and the resulting images are shown in Figure 6. The memory compression ratios for the $(s_{max} - s)$-th layer are shown as follows:

$s_{max} - s$	:	1	2	3	4	5	6	Total
compression ratio	:	0.0020	0.0055	0.0130	0.0275	0.0590	0.1206	0.2274

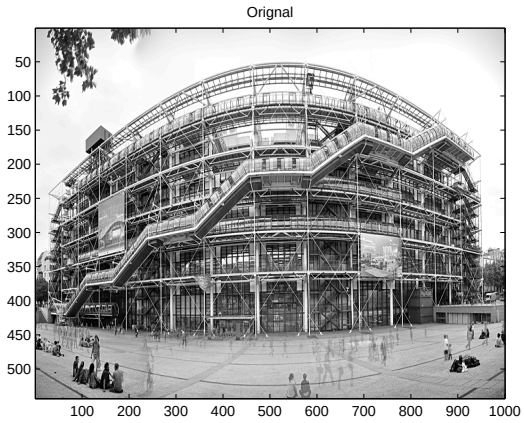


Figure 5: Original image in Example 3

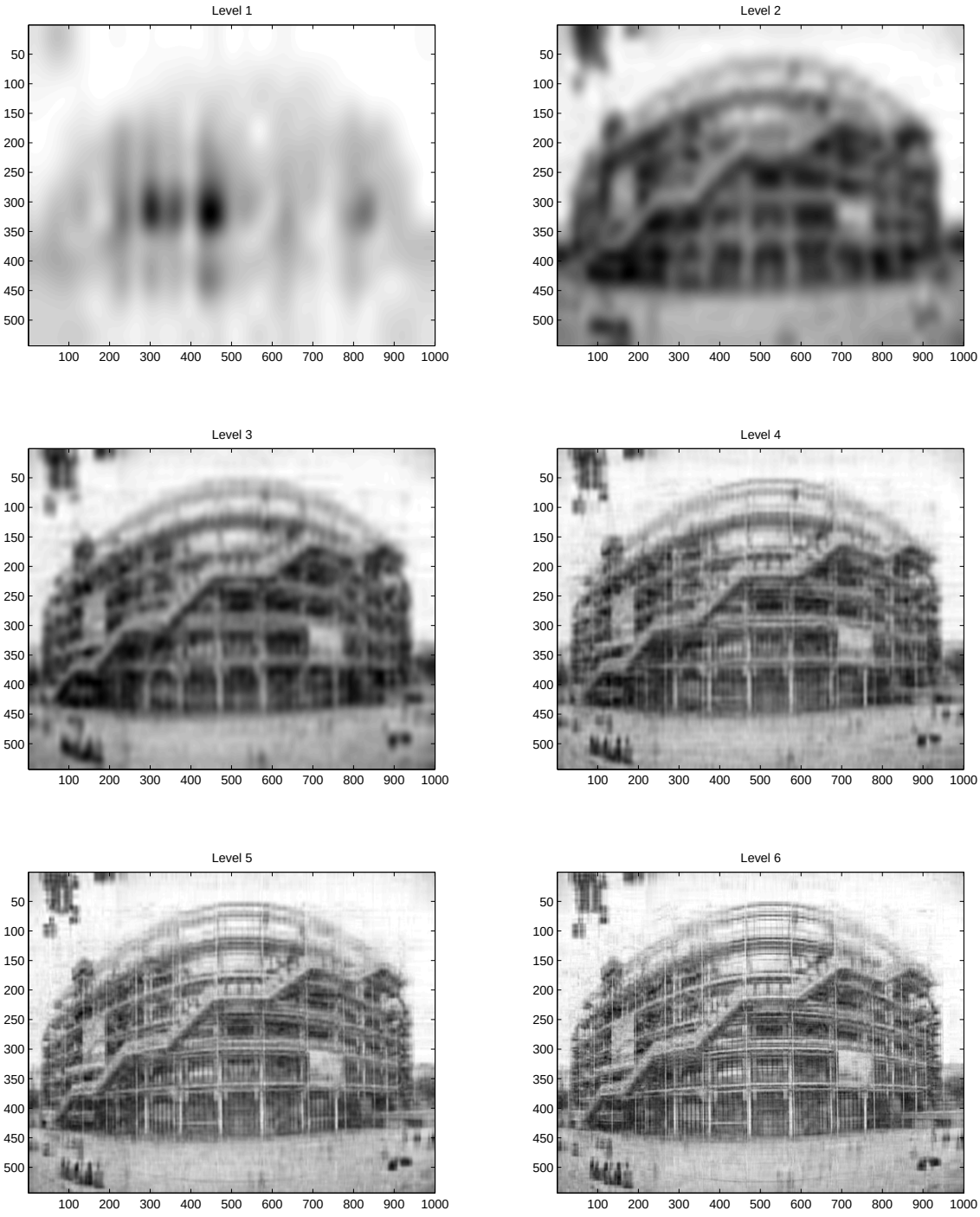


Figure 6: MRA for the image in Example 3 using NMF

Example 4. In this example, we use the same set of parameters except that we set $s_{max} = [\log(\min(N, M))/\log(r) - 4]$ instead. Y is set as the image in Figure 7, and the resulting images are

shown in Figure 8. The memory compression ratios for the $(s_{max} - s)$-th layer are shown as follows:

$s_{max} - s$	:	1	2	3	4	5	Total
compression ratio	:	0.0034	0.0090	0.0212	0.0458	0.1003	0.1796

Finer details are present as the layer number increases, and a reasonably low compression ratio is maintained. Meanwhile, the Chinese characters are already readable from the 4-th layer, which requires only 4% of the original memory.

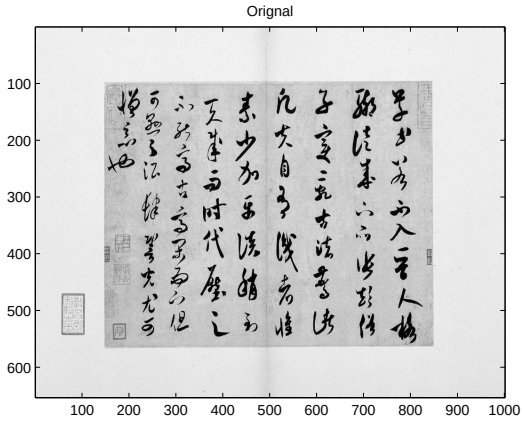
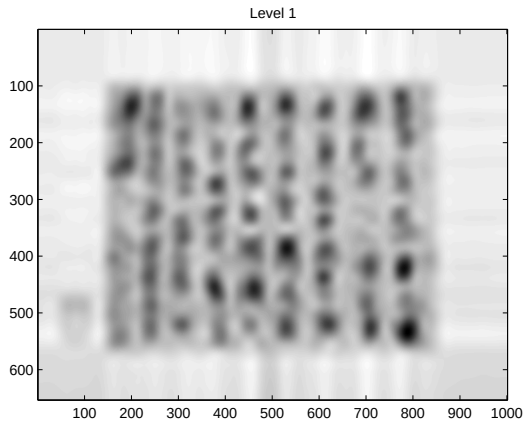
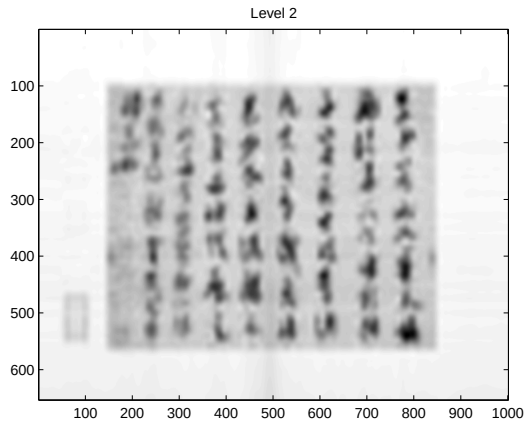


Figure 7: Original image in Example 4





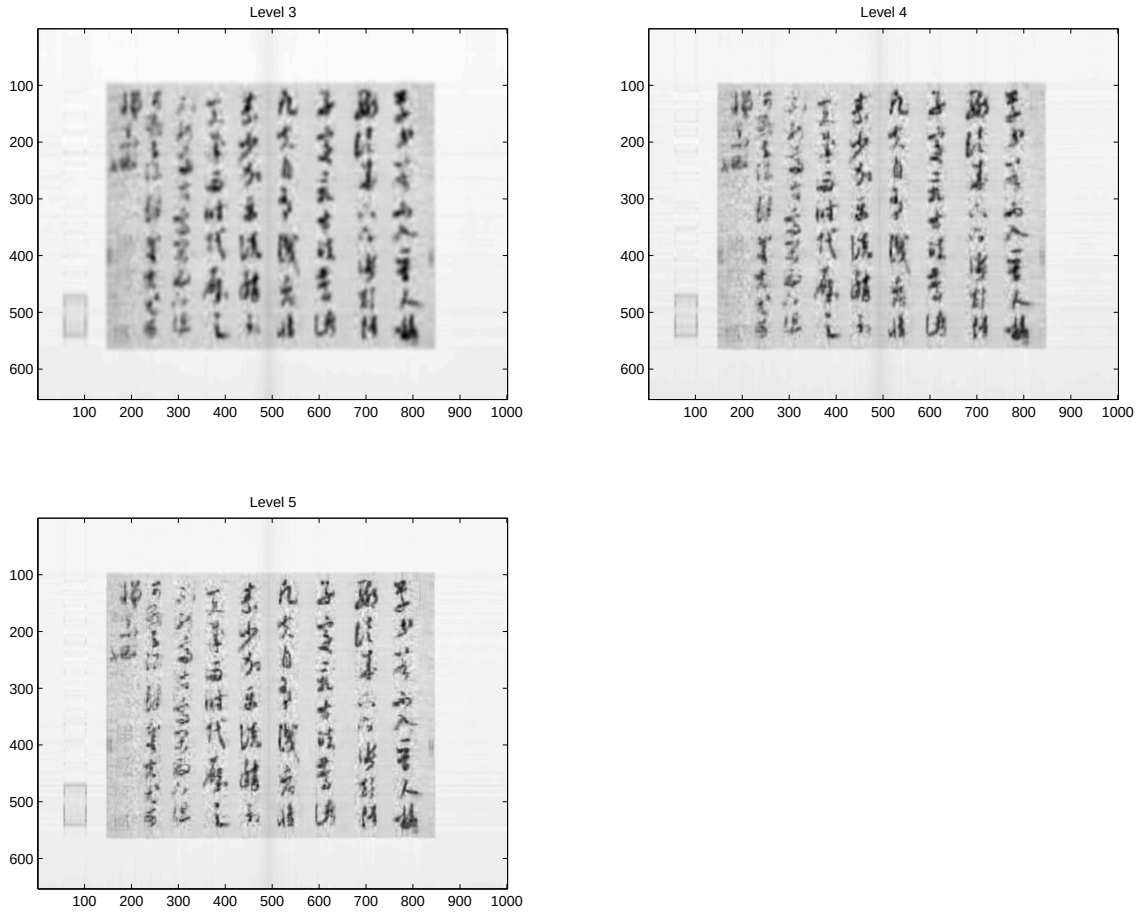


Figure 8: MRA for the image in Example 4 using NMF

5.2 Images from the DSMs

In this subsection, we shall present the application of NMF to the images obtained from an inverse problem algorithm, namely the direct sampling methods (DSMs). The DSMs are a family of simple and efficient inversion methods which aim at providing a good estimate of the locations of inhomogeneities inside a homogeneous background representing various physical media from a single or a small number of boundary data in both full and limited aperture cases. They were first introduced and studied in [16] [20] using far-field data and in [13] using near-field data for locating inhomogeneities in inverse acoustic medium scattering, and was later extended to various other coefficient determination inverse problems, such as the electrical impendence tomography (EIT) [3], the diffusive optical tomography (DOT) [2] and the electromagnetic inverse scattering problem [10]. In each of the aforementioned tomography, a family of probing functions is introduced and an indicator function is defined as a dual product between the observed data and the probing function under an appropriate choice of Sobolev scale. The evaluation of the index function is very cost-effective and the images obtained from the index functions are proven to be effective in locating small abnormalities. For further information of DSMs, the readers may consult [2, 3, 10, 13, 16, 20].

In what follows, we shall apply the NMF to the DSM images from two tomography, namely the DOT and EIT. DOT is a popular non-invasive imaging technique that measures the optical properties of a

medium and creates images which show the distribution of absorption coefficient inside the body. It is very useful for medical imaging, e.g. breast cancer imaging, brain functional imaging, stroke detection, muscle functional studies, photodynamic therapy, and radiation therapy monitoring; see ref. in [2]. In our subsequent discussion, we consider the numerical experiments of the DOT using DSM as in the Section 6 in [2], and the same numerical setting described therein. The medium coefficient inside all the inhomogeneous inclusions are set as $\mu = 50$. The images generated from the scattered potential using the DSM algorithm described in that work are then put into **Algorithm 2** for NMF, with parameters set to be $\alpha = 0.2, \gamma = 0.02, p = 5, \tilde{p} = 3$ and $c_1 = c_2 = 1$ in all the following examples.

Example 5. In this example, we consider the case of two circular inclusions of radius 0.065, which are respectively centered at $(-0.5, 0.25)$ and $(0.25, 0.15)$; see Figure 9 (top). The squared reconstructed images from the indices $\widetilde{I}^2$ described in [2] is presented in Figure 9 (second). The three images $\sigma_{i_l j_l} (\tilde{u}_p)_{i_l} \otimes (\tilde{v}_p)_{j_l}$, for $l = 1, 2, 3$ after NMF obtained in **Algorithm 2** are shown in Figure 9 (third to fifth). The squared image of the final approximation to $\mathcal{I}^{\alpha,\gamma}_{p,\tilde{p}}$ after normalization is given in Figure 9 (last). From the figure, we can see that with an appropriate cutoff, e.g. a 50% cutoff, both the sizes and locations of inhomogeneities obtained from the image are reasonable accurate.

Example 6. This example tests a medium with 4 circular inclusions of radius 0.065 with their corresponding positions: $(-0.5, 0.3)$, $(-0.3, -0.1)$, $(0.3, 0.1)$ and $(0.5, 0.3)$; see Figure 10 (top). Figure 10 (second) shows the squared reconstructed images from the indices $\widetilde{I}^2$ described in [2]. Components $\sigma_{i_l j_l} (\tilde{u}_p)_{i_l} \otimes (\tilde{v}_p)_{j_l}$, for $l = 1, 2, 3$ after NMF are shown in Figure 10 (third to fifth). Figure 10 (last) gives the squared image of the final approximation to $\mathcal{I}^{\alpha,\gamma}_{p,\tilde{p}}$ after normalization. The first two components decomposed from NMF represent the inhomogeneous inclusions inside the original medium, although the combined image shows a dominance by the bottom two inclusions.

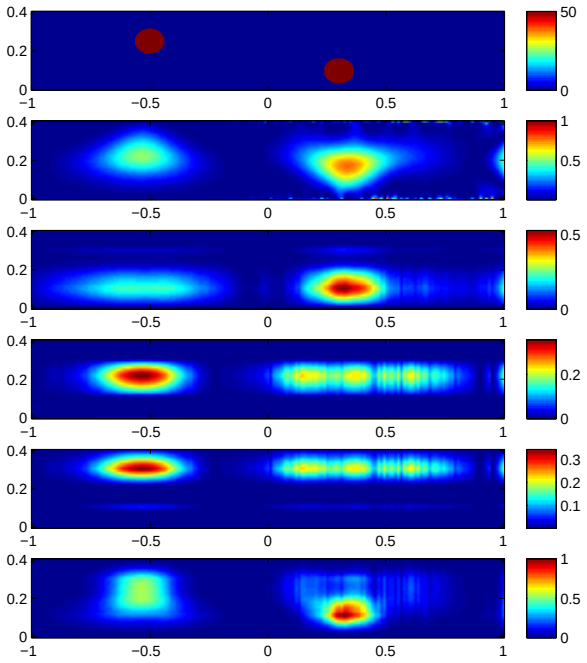


Figure 9: NMF decomposition of the DSM images from DOT in Example 5

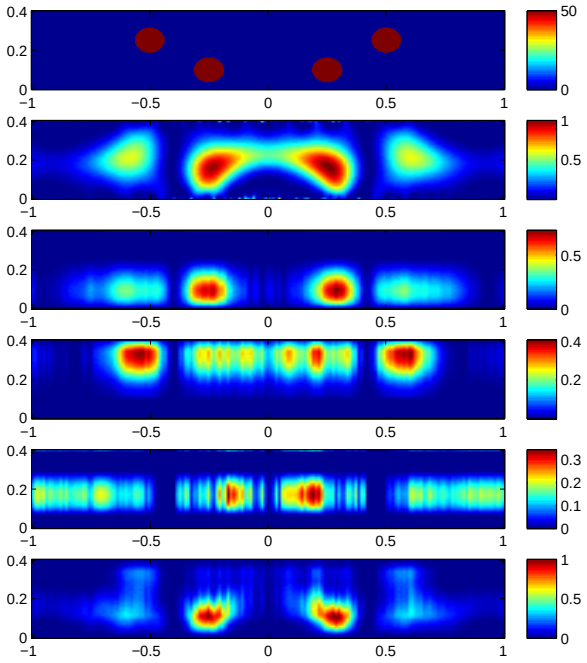


Figure 10: NMF decomposition of the DSM images from DOT in Example 6

Next, we shall apply the NMF to the DSM images from EIT, which is an effective noninvasive evaluation method that creates images of the electrical conductivity of an inhomogeneous medium by applying currents at a number of electrodes on the boundary and measuring the corresponding voltages. It has found applications in many areas, such as oil and geophysical prospection, medical imaging, physiological measurement, early diagnosis of breast cancer, monitoring of pulmonary functions and detection of leaks from buried pipes, etc; see ref. in [3]. In what follows, we consider the same numerical setting as in the numerical experiments of EIT for a circular domain using DSM described in the Section 6 in [3]. The physical coefficient of the inhomogeneous inclusions are all set to be $\sigma = 5$. The images generated from the scattered potential field using the DSM algorithm are then put into **Algorithm 2** for NMF, with parameters set to be $\alpha = 0.2, \gamma = 0.02, p = 5, \tilde{p} = 3$ and $c_1 = c_2 = 1$ in all the following examples.

Example 7. We now investigate an example with 2 inclusions of size 0.1×0.1 respectively at the positions $(-0.44, 0.36)$ and $(0.36, -0.44)$; see Figure 11 (a). The squared reconstructed images from the indices $\widetilde{I}^2$ after normalization as described in [3] is presented in Figure 11 (b). The components $\sigma_{i_l j_l} (\tilde{u}_p)_{i_l} \otimes (\tilde{v}_p)_{j_l}$, for $l = 1, 2, 3$ obtained from NMF using **Algorithm 2** over the image $\widetilde{I}$ are shown in Figure 11 (c-e). The squared image of the approximation to $\mathcal{I}^{\alpha,\gamma}_{p,\tilde{p}}$ after normalization is in Figure 11 (f). The components of inhomogeneous inclusions sitting inside the original medium are decomposed into different components from the NMF.

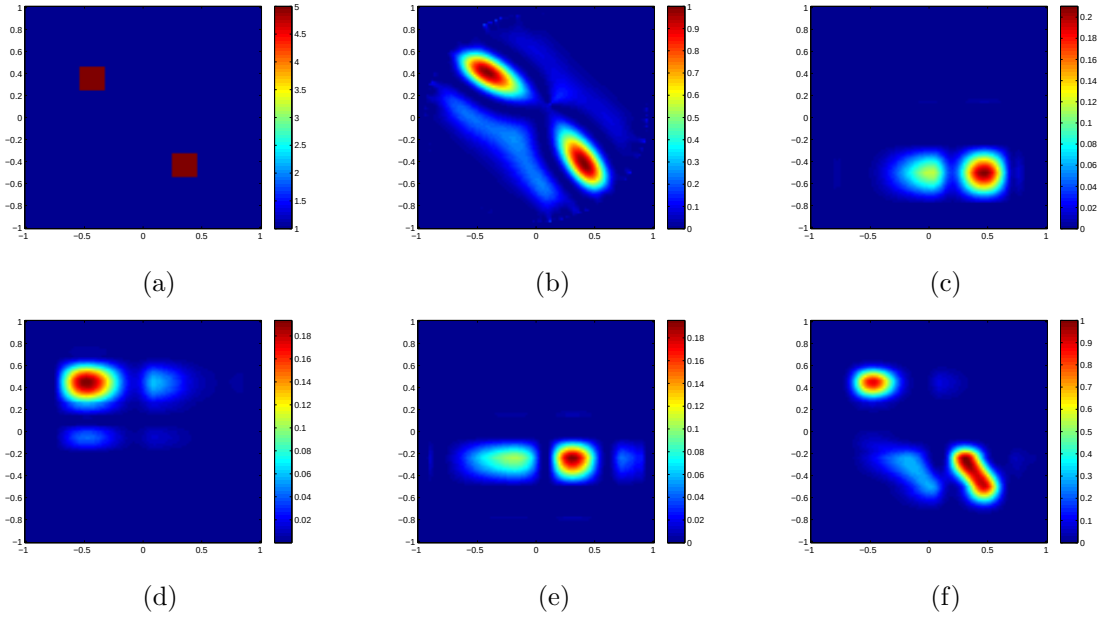


Figure 11: NMF decomposition of the DSM images from EIT in Example 7

Example 8. In this example, we consider the case of 4 inclusions with same size as in Example 7 sitting inside the sampling region, which are placed at positions of $(0.36, 0.36)$, $(0.36, -0.44)$, $(-0.44, 0.36)$ and $(-0.44, -0.44)$; see Figure 12 (a). The squared reconstructed images from the indices $\widetilde{I}^2$ after normalization is shown in Figure 12 (b). Figure 12 (c-e) presents the images of $\sigma_{i_l j_l} (\tilde{u}_p)_{i_l} \otimes (\tilde{v}_p)_{j_l}$, for $l = 1, 2, 3$ after NMF over the image $\widetilde{I}$. The squared image of the approximation to $\mathcal{I}^{\alpha,\gamma}_{p,\tilde{p}}$ after normalization is in Figure 12 (f).

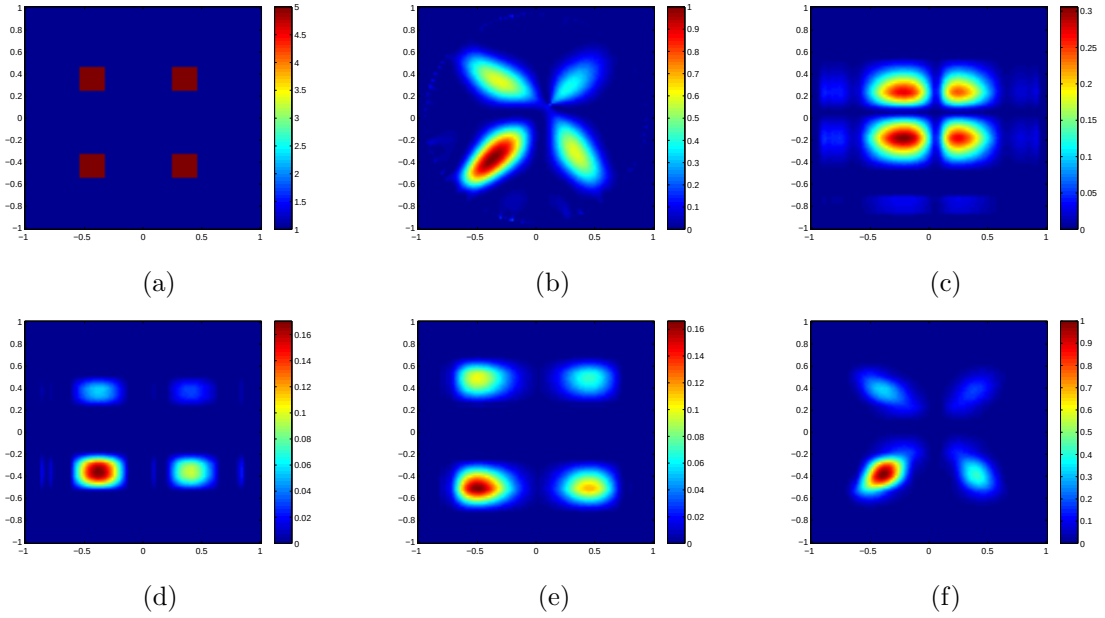


Figure 12: NMF decomposition of the DSM images from EIT in Example 8

Example 9. In this example, 2 inclusions of the same size as in Example 7 are introduced in the homogeneous background, and they are respectively placed at the positions $(-0.36, 0.36)$ and $(0.36, 0.36)$ inside the domain; see Figure 13 (a). The squared reconstructed images from the indices $\widetilde{I}^2$ after nor-

malization is given in Figure 13 (b). The images of $\sigma_{i_l j_l} (\tilde{u}_p)_{i_l} \otimes (\tilde{v}_p)_{j_l}$, for $l = 1, 2, 3$ after NMF over the image $\widetilde{I}$ are shown in Figure 13 (c-e). Figure 13 (f) presents the squared image of the approximation to $\mathcal{I}_{p,\tilde{p}}^{\alpha,\gamma}$ after normalization. From the figure, both the sizes and locations of inhomogeneities can be reasonably obtained from the NMF image after the introduction of a appropriate cutoff.

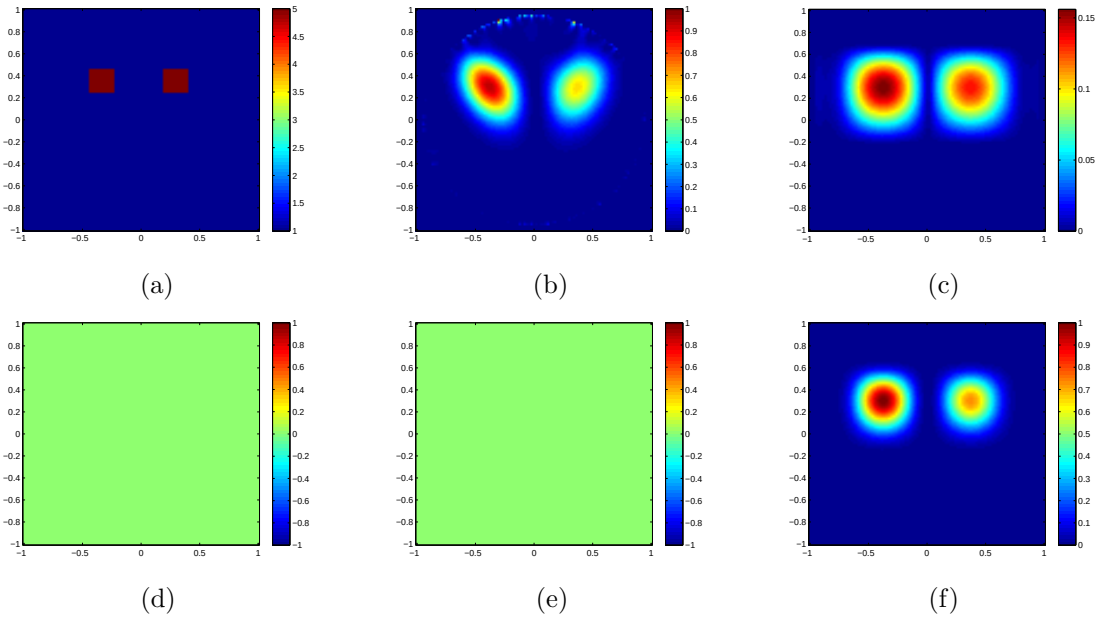


Figure 13: Image of Example 3(EIT) using Non-negative factorization

References

[1] J.-P. Brunet, P. Tamayo, T.R. Golub, and J.P. Mesirov, *Metagenes and molecular pattern discovery using matrix factorization*, Proc. Natl Academy of Sciences USA 102(12), (2004), pp. 4164-4169.

[2] Yat Tin Chow, Kazufumi Ito, Keji Liu, Jun Zou, *Direct sampling method for diffusive optical tomography*, preprint, arXiv:1410.1275.

[3] Yat Tin Chow, Kazufumi Ito, Jun Zou, *A direct sampling method for electrical impedance tomography*, Inverse Problems 30 (2014), 095003.

[4] M. Cooper and J. Foote, *Summarizing video using non-negative similarity matrix factorization*, Proc. IEEE Workshop on Multimedia Signal Processing, (2002), pp. 25-28.

[5] C. Ding and X. He. *K-means clustering via principal component analysis*. Proc. of Int'l Conf. Machine Learning, 2004, pp 225-232.

[6] C. Ding, X. He, and H.D. Simon, *On the equivalence of nonnegative matrix factorization and spectral clustering*, Proc. SIAM Data Mining Conf, (2005).

[7] C. Ding, T. Li, W. Peng, H. Park, *Orthogonal nonnegative matrix tri-factorizations for clustering*, Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining, ACM, 2006, pp. 126-135.

[8] E. Esser, M. Muller, S. Osher, G. Sapiro, and J. Xin, *A convex model for nonnegative matrix factorization and dimensionality reduction on physical space*, IEEE Trans. on Image Processing 21 (7), (2012), pp. 3239-3252.

[9] P. O. Hoyer, *Non-negative matrix factorization with sparseness constraints*, J. Machine Learning Research 5,(2004), pp.1457-1469.

[10] K. Ito, B. Jin and J. Zou, *A direct sampling method for inverse electromagnetic medium scattering*, Inv. Prob., 29 (2013) 095018 (19pp).

[11] K. Ito, B. Jin and J. Zou, *A two-stage method for inverse medium scattering*, J. Comput. Phys. 237 (2013), pp. 211-223.

[12] K. Ito and K. Kunisch, Lagrange Multiplier Approach to Variational Problems and Applications, SIAM, Philadelphia, PA, (2008).

[13] K. Ito, B. Jin and J. Zou, A direct sampling method to inverse medium scattering problem, Inverse Problems, 28(2) (2012), 025003.

[14] D.D. Lee and H.S. Seung, *Learning the parts of objects by non-negative matrix factorization*, Nature 401, (1999), pp. 788-791.

[15] D.D. Lee and H.S. Seung, *Algorithms for non-negatvie matrix factorization*, Advances in Neural Information Processing Systems, 13 (2001), pp. 556-562.

[16] J. Li and J. Zou, A direct sampling method for inverse scattering using far-field data, Inverse Problems and Imaging, 7 (2013), pp. 757-775.

[17] S.Z. Li, X. Hou, H. Zhang, and Q. Cheng, *Learning spatially localized, parts-based representation*, Proc. of IEEE Computer Vision and Pattern Recognition, (2001), pp. 207-212.

[18] Y. Li and A. Ngom, *The non-negative matrix factorization toolbox for biological data mining*, BMC Source Code for Biology and Medicine, 8 (2013), pp. 10-25.

[19] P. Paatero and U. Tapper, *Positive matrix factorization: Anon-negative factor model with optimal utilization of error estimates of data values*, Environmetrics 5, (1994), pp. 111-126.

[20] R. Potthast, A study on orthogonality sampling, Inverse Problems, 26 (2010), 074015 (17pp).

[21] W. Xu, X. Liu, and Y. Gong, *Document clustering based on non-negative matrix factorization*, Proc. ACM conf. Research and development in IR (SIRGIR), (2003), pp.267-273.