FIRST-ORDER SPLITTING ALGORITHMS FOR NONCONVEX MATRIX OPTIMIZATION PROBLEMS

LEI YANG

Ph.D

The Hong Kong Polytechnic University

2017

The Hong Kong Polytechnic University

Department of Applied Mathematics

First-order Splitting Algorithms for Nonconvex Matrix Optimization Problems

Lei Yang

A thesis submitted in partial fulfilment of the requirements

for the degree of Doctor of Philosophy

June 2017

Certificate of Originality

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

__________________________(Signature)

_________Lei YANG_________(Name of student)

Dedicated to my parents.

Abstract

In this thesis, we consider two classes of matrix optimization problems. One is the *matrix decomposition problem* (MDP), which aims to decompose a given data matrix into the sum of two matrices with different desirable structures. The other one is the *matrix factorization problem* (MFP), which aims to factorize a given data matrix into the product of two small factor matrices with different desirable structures. These two classes of problems cover many existing widely-studied models with many applications in areas such as machine learning and imaging sciences. To solve MDP and MFP, which are possibly nonconvex, nonsmooth and even non-Lipschitz, we develop two efficient first-order splitting algorithms for them, respectively.

We first consider MDP. Specifically, we adapt the alternating direction method of multipliers (ADMM) with a general dual step-size to solve a reformulation of the original problem that contains three blocks of variables, and analyze its convergence. We show that for any dual step-size less than the golden ratio, there exists a computable threshold such that if the penalty parameter is chosen above such a threshold and the sequence thus generated by our ADMM is bounded, then the cluster point of the sequence gives a stationary point of the nonconvex optimization problem. We achieve this via a potential function specifically constructed for our ADMM. Moreover, we establish the global convergence of the whole sequence if, in addition, this special potential function is a Kurdyka-Łojasiewicz function. Furthermore, we present a simple strategy for initializing the ADMM to guarantee boundedness of

the sequence. Finally, we perform numerical experiments comparing our ADMM with the proximal alternating linearized minimization proposed in [8] for the background/foreground extraction problem on real datasets. The numerical results show that our ADMM with a nontrivial dual step-size is efficient.

We then consider MFP. To solve it, we develop a non-monotone alternating updating method based on a potential function. Our method essentially updates two blocks of variables in turn by inexactly minimizing this potential function, and updates another auxiliary block of variables using an explicit formula. The special structure of our potential function allows us to take advantage of efficient computational strategies for non-negative matrix factorization to perform the alternating minimization over two blocks of variables. A suitable line search criterion is also incorporated to improve the numerical performance of our method. Under some mild conditions, we show that our line search criterion is well defined, and establish that the sequence generated is bounded and any cluster point of the sequence is a stationary point of our problem. Moreover, we discuss the convergence rate for the function value if, in addition, the objective is a Kurdyka-Łojasiewicz function. Finally, we conduct some numerical experiments using real datasets to compare our method with some existing efficient methods for non-negative matrix factorization and matrix completion. The numerical results show that our method can outperform these methods for these specific applications.

Publications Arising from the Thesis

- L. Yang, T. K. Pong and X. Chen. ADMM for a class of nonconvex and nonsmooth problems with applications to background/foreground extraction. *SIAM Journal on Imaging Science*, 10(1): 74 - 110, 2017.
- L. Yang, T. K. Pong and X. Chen. A non-monotone alternating updating method for a class of matrix factorization problems. *Submitted*, 2017.

Acknowledgements

First and foremost, I would like to thank my chief supervisor Professor Xiaojun Chen for her kind help and patient guidance throughout my PhD period. Her wealth of knowledge, wisdom in life and generous supports have made working with her a precious experience and treasure for me, not only in my PhD study, but also in my future life. She is undoubtedly my role model in my academic career. Moreover, special gratitude to her for supporting me to visit University of California, Davis. I would also like to thank Dr. An Li, husband of Professor Xiaojun Chen. He is a wise and generous person with a great sense of humor. We all like to talk with him. He always treats us delicious food and good wine during holidays and other important days. Those all colored my PhD life.

My deepest thanks also go to my co-supervisor Professor Ting Kei Pong. In the past three years, he spent a lot of time discussing with me on my research. I really learn and benefit a lot from him. His rich knowledge, clarity of thought, enthusiasm and energy have influenced me profoundly in my research study. In particular, I am thankful to him for recommending me to Professor Michael P. Friedlander and Professor Nicolas Gillis. I feel very fortunate to have had him as my co-supervisor.

My sincere acknowledgements go to Professor Michael P. Friedlander for giving me the opportunity to visit his research group at University of California, Davis. His wisdom, passion and caring attitude have impressed me a lot. During my visit, he spent plenty of time discussing with me and teaching me how to do and how to think when carrying on a new research project. I really benefit a lot from him. I am also thankful to his PhD students Dr. Gabriel Goh and Mr. Will Wright for their kind helps and discussions when I was in Davis.

I am thankful to Professor Zhaosong Lu, Professor Nicolas Gillis and Professor Robert Womersley for their helpful discussions and valuable suggestions on my research when they visited PolyU. I thank Professor Zhenghai Huang, my master supervisor, for his special care and kind recommendation that enabled me to come to PolyU. I would also like to thank Professor Xiaoqi Yang for acting as the chair of my defense committee and thank Professor Kim Chuan Toh and Professor Anthony Man-Cho So for kindly accepting to be my external examiners.

I am very grateful to joint Professor Xiaojun Chen's group and I am very thankful to all my academic brothers and sisters: Professor Congpei An, Professor Wei Bian, Professor Xin Liu, Professor Yafeng Liu, Professor Hailin Sun, Professor Shulin Wu, Professor Zaikun Zhang, Dr. Yanfang Zhang, Dr. Yang Zhou, Dr. Hong Wang, Ms. Lili Pan, Miss. Tianxiang Liu, Miss. Qiyu Wang, Mr. Andy Yat Ming Cheung, Mr. Bo Wen, Mr. Yun Shi, Mr. Guidong Liu, Mr. Jie Jiang and Mr. Rui Zhang for their constant and substantial support. I have benefited a lot from them during the past weekly seminars. Special thanks go to Dr. Yang Zhou and Dr. Hong Wang for their helps when I come to PolyU.

I would also like to thank all the friends I have met in Hong Kong and Davis for their friendships. In particular, I would like to especically acknowledge Mr. Jin Yang, my best friend, for his sincere friendship and remarkably kind help and constant support over the years.

My deepest gratitude and love, of course, belong to my mother Jiafang Li and my father Benquan Yang for their unconditional love and support through my life. Last but not least, I am also greatly indebted to my girlfriend Xiaoqiong Zhuang for her company, understanding and support throughout my PhD study.

Contents

List of Figures

List of Tables

List of Notations

$\mathbb{R}$	The set of real numbers
$\mathbb{R}^m$ ($\mathbb{R}^m_+$)	The set of m-dimensional real vectors (with nonnegative entries)
$\mathbb{R}^{m\times n}$ ($\mathbb{R}^{m\times n}_+$)	The set of $m \times n$ real matrices (with nonnegative entries)
$\\|\boldsymbol{x}\\|$	The Euclidean norm of the vector $\boldsymbol{x}$
$\\|X\\|_0$	The ℓ_0-quasi-norm of the matrix X
$\\|X\\|_1$	The ℓ_1-norm of the matrix X
$\\|X\\|_p$	The ℓ_p-quasi-norm ($0 < p < 1$) of the matrix X
$\\|X\\|$	The spectral norm (maximum singular value) of the matrix X
$\\|X\\|_*$	The nuclear norm of the matrix X
$\\|X\\|_F$	The Frobenius norm of the matrix X
$\\|X\\|_{\mathcal{S}_p}$	The Schatten-p (quasi-)norm ($0 < p < \infty$) of the matrix X
$\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$	The linear maps
$\mathcal{A}^*$, $\mathcal{B}^*$, $\mathcal{C}^*$	The adjoint linear maps of $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$

Chapter 1

Introduction

Nowadays, many optimization problems arising in various areas, e.g., signal processing, image processing and machine learning, are nonconvex, nonsmooth and even non-Lipschitz. Moreover, these problems are usually presented in large scale. Thus, the fast solution methods for these problems are urgently in demand. On the other hand, the first-order splitting methods have played a significant role for solving convex problems and inspired many efficient numerical algorithms in different applications. Hence, it is conceivable that the performance of using the first-order splitting methods for nonconvex, nonsmooth and non-Lipschitz problems can be promising. In fact, some encouraging empirical performances have been verified for different applications in the literature. However, the theoretic analysis is still limited. The goal of this thesis is to develop efficient first-order splitting algorithms for solving some existing or new nonconvex, nonsmooth and non-Lipschitz problems capturing concrete applications in various areas, and analyze their convergences. Specifically, in this thesis, we consider the following two classes of matrix optimization problems:

- **Matrix Decomposition Problem (MDP)**: in this problem, we focus on decomposing a given data matrix $D \in \mathbb{R}^{m\times n}$ into two components $L \in \mathbb{R}^{m\times n}$ and $S \in \mathbb{R}^{m\times n}$ with different desirable structures such that $D \approx L + S$.

- **Matrix Factorization Problem (MFP)**: in this problem, we focus on factorizing a given data matrix $M \in \mathbb{R}^{m\times n}$ into two factors $X \in \mathbb{R}^{m\times r}$ and $Y \in \mathbb{R}^{n\times r}$ with different desirable structures such that $M \approx XY^\top$, where $r \leqslant \min\{m, n\}$.

These two classes of matrix optimization problems arise in many applications from areas such as machine learning and imaging sciences. We will discuss more details for their models in Section 1.1 and Section 1.2, respectively, and develop two efficient first-order splitting algorithms for solving them in Chapter 3 and Chapter 4, respectively. To better distinguish these two classes of problems, throughout this thesis, for MDP, we use $L \in \mathbb{R}^{m\times n}$ and $S \in \mathbb{R}^{m\times n}$ as the decision variables, and use $D \in \mathbb{R}^{m\times n}$ as the given matrix; while, for MFP, we use $X \in \mathbb{R}^{m\times r}$ and $Y \in \mathbb{R}^{n\times r}$ as the decision variables, and use $M \in \mathbb{R}^{m\times n}$ as the given matrix. With a slight abuse of notation, we use $\mathcal{F}(L,S)$ in (1.1) and $\mathcal{F}(X,Y)$ in (1.4) to denote the objectives in the models for MDP and MFP, respectively.

1.1 Matrix decomposition problem

The MDP considered in this thesis can be modeled as

$$\min_{L,S}\ \mathcal{F}(L,S) := \Psi(L) + \Phi(S) + \frac{1}{2}\left\|D - \mathcal{A}\left[\mathcal{B}(L) + \mathcal{C}(S)\right]\right\|_F^2, \tag{1.1}$$

where $\Psi, \Phi : \mathbb{R}^{m\times n} \to \mathbb{R}_+ \cup \{\infty\}$ are proper closed nonnegative functions, and Ψ is *convex*, while Φ is possibly *nonconvex*, *nonsmooth* and *non-Lipschitz*; $\mathcal{A}, \mathcal{B}, \mathcal{C} : \mathbb{R}^{m\times n} \to \mathbb{R}^{m\times n}$ are linear maps and $\mathcal{B}$, $\mathcal{C}$ are *injective*. In particular, $\Psi(L)$ and $\Phi(S)$ in (1.1) can be different regularizers used for inducing the desired structures. For instance, $\Psi(L)$ can be used for inducing low rank in L. One possible choice is $\Psi(L) = \|L\|_*$ (see Chapter 2 for notation and definitions). Alternatively, one may consider $\Psi(L) = \delta_\Omega(L)$, where Ω is a compact convex set such as $\Omega = \{L \in \mathbb{R}^{m\times n} : \|L\|_\infty \leqslant l,\ L_{:1} = L_{:2} = \cdots = L_{:n}\}$ with $l > 0$, or $\Omega = \{L \in \mathbb{R}^{m\times n} : \|L\|_* \leqslant r'\}$ with $r' > 0$; the former choice restricts L to have rank at most 1 and makes (1.1) nuclear-norm-free (see [52, 54]). On the other hand, $\Phi(S)$ can be used for inducing sparsity. In the literature, $\Phi(S)$ is typically separable, i.e., taking the form

$$\Phi(S) = \mu \sum_{i=1}^{m} \sum_{j=1}^{n} \phi(s_{ij}), \tag{1.2}$$

where ϕ is a nonnegative continuous function with $\phi(0) = 0$ and $\mu > 0$ is a regularization parameter. Some concrete examples of ϕ include:

1. bridge penalty [41, 44]: $\phi(t) = |t|^p$ for $0 < p \leqslant 1$;
2. fraction penalty [28]: $\phi(t) = \alpha|t|/(1 + \alpha|t|)$ for $\alpha > 0$;
3. logistic penalty [64]: $\phi(t) = \log(1 + \alpha|t|)$ for $\alpha > 0$;
4. smoothly clipped absolute deviation [24]: $\phi(t) = \displaystyle\int_0^{|t|} \min(1, (\alpha - s/\mu)_+/(\alpha - 1))\,\mathrm{d}s$ for $\alpha > 2$;
5. minimax concave penalty [93]: $\phi(t) = \int_0^{|t|}(1 - s/(\alpha\mu))_+\,\mathrm{d}s$ for $\alpha > 0$;
6. hard thresholding penalty function [25]: $\phi(t) = \mu - (\mu - |t|)_+^2/\mu$.

The bridge penalty and the logistic penalty have also been considered in [20]. Finally, the linear map $\mathcal{A}$ can be suitably chosen to model different scenarios. For example, $\mathcal{A}$ can be chosen to be the identity map for extracting L and S from a noisy data D, and the blurring map for extracting L and S from a blurred data D. The linear map $\mathcal{B}$ can be the identity map or some "dictionary" that spans the data space (see, for example, [57]), and $\mathcal{C}$ can be chosen to be the identity map or the inverse of certain sparsifying transform (see, for example, [65]). More examples of (1.1) can be found in [11, 12, 13, 20, 67, 89].

One representative application that is frequently modeled by (1.1) via a suitable choice of Φ, Ψ, $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ is the background/foreground extraction problem, which is an important problem in video processing; see [9, 10] for recent surveys. In this problem, one attempts to separate the relatively static information called "background" and the moving objects called "foreground" in a video. The problem can be modeled by (1.1), and such models are typically referred to as robust principal component analysis (RPCA)-based models. In these models, each image is stacked as a column of a data matrix D, the relatively static background is then modeled as a low rank matrix, while the moving foreground is modeled as sparse outliers. The data matrix D is then decomposed (approximately) as the sum of a low rank matrix $L \in \mathbb{R}^{m\times n}$ modeling the background and a sparse matrix $S \in \mathbb{R}^{m\times n}$ modeling the foreground. Various approximations are then used to induce low rank and sparsity, resulting in different RPCA-based models, most of which take the form of (1.1). One example is to set Ψ to be the nuclear norm of L, i.e., the sum of singular values of L, to promote low rank in L and Φ to be the ℓ_1 norm of S to promote sparsity in S, as in [13]. Besides convex regularizers, nonconvex models have also been widely studied recently and their performances are promising; see [20, 83] for background/foreground extraction and [6, 18, 35, 63, 64, 97] for other problems in image processing. There are also nuclear-norm-free models that do not require the singular value decomposition of the matrix variable L when solving them, making the model more practical especially when the size of matrix is large. For instance, in [54], the authors set Φ to be the ℓ_1 norm of S and Ψ to be the indicator function of $\Omega = \{L \in \mathbb{R}^{m\times n} : L_{:1} = L_{:2} = \cdots = L_{:n}\}$. A similar approach was also adopted in [52] with promising performances. Clearly, for nuclear-norm-free models, one can also take Φ to be some nonconvex sparsity inducing regularizers, resulting in a special case of (1.1) that has not been explicitly considered in the literature before; we will consider these models in our numerical experiments in Section 3.3. The above discussion shows that problem (1.1) is flexible enough to cover a wide range of RPCA-based models for background/foreground extraction.

Problem (1.1) is nonconvex in general. Thus, in this thesis, we will focus on finding a *stationary point* of the objective $\mathcal{F}$ in (1.1). As we will show later in Section 3.1, model (1.1) can be reformulated into an optimization problem with three blocks of variables. This kind of problems containing several blocks of variables has been widely studied in the literature; see, for example, [54, 62, 67]. Hence, it is natural to adapt the algorithm used there, namely, the alternating direction method of multipliers (ADMM), for solving (1.1). Classically, the ADMM can be applied to solving problems of the following form that contains 2 blocks of variables:

$$\min_{x_1,x_2} \{\, f_1(x_1) + f_2(x_2) : \mathcal{A}_1(x_1) + \mathcal{A}_2(x_2) = b \,\}, \tag{1.3}$$

where f_1, f_2 are proper closed convex functions and $\mathcal{A}_1$, $\mathcal{A}_2$ are linear maps. The iterative scheme of ADMM is

$$\begin{cases} x_1^{k+1} \in \operatorname*{Argmin}_{x_1} \left\{ \mathcal{L}_\beta(x_1, x_2^k, z^k) \right\}, \\ x_2^{k+1} \in \operatorname*{Argmin}_{x_2} \left\{ \mathcal{L}_\beta(x_1^{k+1}, x_2, z^k) \right\}, \\ z^{k+1} = z^k - \tau\beta(\mathcal{A}_1(x_1^{k+1}) + \mathcal{A}_2(x_2^{k+1}) - b), \end{cases}$$

where $\tau \in (0, \frac{\sqrt{5}+1}{2})$ is the dual step-size and $\mathcal{L}_\beta$ is the augmented Lagrangian function for (1.3) defined as

$$\begin{aligned} \mathcal{L}_\beta(x_1, x_2, z) \; := \; & f_1(x_1) + f_2(x_2) - \langle z, \mathcal{A}_1(x_1) + \mathcal{A}_2(x_2) - b \rangle \\ & + \frac{\beta}{2}\|\mathcal{A}_1(x_1) + \mathcal{A}_2(x_2) - b\|^2 \end{aligned}$$

with $\beta > 0$ being the penalty parameter. Under some mild conditions, the sequence $\{(x_1^k, x_2^k)\}$ generated by the above ADMM can be shown to converge to an optimal solution of (1.3); see for example, [5, 23, 27, 33]. However, the ADMM used in [54, 62, 67] does not have a convergence guarantee; indeed, it is shown recently in [16] that the ADMM, when applied to a convex optimization problem with 3 blocks of variables, can be divergent in general. This motivates the study of many provably convergent variants of the ADMM for convex problems with more than 2 blocks of variables; see, for example, [37, 38, 55, 56]. Recently, Hong et al. [40] established the convergence of the multi-block ADMM with $\tau = 1$ for certain types of nonconvex problems whose objective is a sum of a possibly *nonconvex* Lipschitz differentiable function and a bunch of convex nonsmooth functions when the penalty parameter is chosen above a computable threshold. The problem they considered covers (1.1) when Φ is convex, or smooth and possibly nonconvex. Later, Wang et al. [83] considered a more general type of nonconvex problems that contains (1.1) as a special case and allows some nonconvex nonsmooth functions in the objective. To solve this type of problems, they considered a variant of the ADMM whose subproblems are simplified by adding a Bregman proximal term. However, their results cannot be applied to the direct adaptation of the ADMM for solving (1.1). In view of the above, we will develop the three-block ADMM with a non-trivial dual step-size for (1.1) and analyze its convergence in Chapter 3.

1.2 Matrix factorization problem

The MFP considered in this thesis can be modeled as

$$\min_{X,Y} \ \mathcal{F}(X,Y) := \Psi(X) + \Phi(Y) + \frac{1}{2}\left\|\mathcal{A}(XY^\top) - \boldsymbol{b}\right\|^2, \tag{1.4}$$

where $X \in \mathbb{R}^{m\times r}$ and $Y \in \mathbb{R}^{n\times r}$ are decision variables with $r \leqslant \min\{m, n\}$, the functions $\Psi : \mathbb{R}^{m\times r} \to \mathbb{R} \cup \{\infty\}$ and $\Phi : \mathbb{R}^{n\times r} \to \mathbb{R} \cup \{\infty\}$ are proper closed but possibly *nonconvex*, *nonsmooth* and *non-Lipschitz*, $\boldsymbol{b} \in \mathbb{R}^q$ is a given vector and $\mathcal{A} : \mathbb{R}^{m\times n} \to \mathbb{R}^q$ is a linear map with $q \leqslant mn$ and $\mathcal{A}\mathcal{A}^* = \mathcal{I}$[1]. Model (1.4) covers many existing widely-studied models in many application areas such as machine

[1] We here make a blanket assumption on $\mathcal{A}$ that $\mathcal{A}\mathcal{A}^* = \mathcal{I}$ for matrix factorization problem

learning [81] and imaging sciences [96]. Similarly, $\Psi(X)$ and $\Phi(Y)$ in (1.4) can be various regularizers for inducing the desired structure. We take $\Psi(X)$ for example. Ψ can be the Frobenius norm of X for improving the stability of the estimation of factor X. When the factor X is required to be sparse, Ψ can be chosen as a sparsity-inducing norm, which is typically separable, i.e, taking the form $\Psi(X) = \nu \sum_{i=1}^{m} \sum_{j=1}^{r} \psi(x_{ij})$ with different forms of the penalty ψ (the concrete examples of ψ can be found in Section 1.1). When the factor X is required to have low rank, one can consider $\Psi(X) = \nu\|X\|_*$ with $\nu > 0$ or $\Psi(X) = \delta_{\mathcal{X}}(X)$, where $\mathcal{X} = \{X \in \mathbb{R}^{m\times r} : \|X\|_* \leqslant r'\}$ with $0 < r' \leqslant r$. Moreover, one can also impose nonnegativity on X by considering $\Psi(X) = \delta_{\mathcal{X}}(X)$, where $\mathcal{X} = \{X \in \mathbb{R}^{m\times r} : X \geqslant 0\}$. For the linear map $\mathcal{A}$, it can be suitably chosen to model different scenarios. For instance, $\mathcal{A}$ can be the identity map for the factorization of a given matrix, or the sampling map for recovering a matrix from some observations as in the matrix completion problem [14, 15, 68, 69]. More examples of (1.4) can be found in recent surveys [81, 96].

One representative application of (1.4) is the non-negative matrix factorization (NMF) problem, where $\Psi(X)$ and $\Phi(Y)$ are chosen as the indicator functions for $\mathcal{X} = \{X \in \mathbb{R}^{m\times r} : X \geqslant 0\}$ and $\mathcal{Y} = \{Y \in \mathbb{R}^{n\times r} : Y \geqslant 0\}$, respectively, and $\mathcal{A}$ is the identity map. NMF was first introduced by Paatero and Tapper [66], and then popularized by Lee and Seung [46]. It now has been widely used in data mining applications due to its ability to provide interpretable decompositions of data. The basic task of NMF is to find two nonnegative matrices $X \in \mathbb{R}_+^{m\times r}$ and $Y \in \mathbb{R}_+^{n\times r}$ such that $M \approx XY^\top$ for a given nonnegative data matrix $M \in \mathbb{R}_+^{m\times n}$. One can also impose some desired structures (e.g., sparsity) on X or Y by using some suitable regularizers (e.g., ℓ_1-norm). Thus, this type of problems can be modeled by (1.4) via suitable choices of Ψ, Φ and $\mathcal{A}$. We refer readers to [4, 29, 30, 31, 47, 86] for more information on NMF and its variants. Another important application of (1.4) is the matrix completion (MC) problem, which aims to recover an unknown low rank matrix from a sample of its entries. This problem also arises in various applications, e.g., collaborative filtering [70, 77], sensor-network localization [7] and system identification [59]. One popular class of methods for MC is based on nuclear-norm minimization [14, 15, 68, 69], or, more generally, Schatten-p ($0 < p \leqslant 1$) (quasi−)norm minimization [45, 60, 61, 94]. However, this class of methods bears the computational cost of singular value decompositions or eigenvalue decompositions of huge ($m \times n$) matrices, which can become costly for large-scale problems. Recently, an alternative class of methods based on low–rank matrix factorization has attracted more and more attentions in the literature; see, for example, [43, 73, 74, 75, 79, 87]. In these methods, the $m \times n$ matrix variable is replaced by the product of two or more small-sized matrices. Strategies such as alternating minimization are then adopted so that the resulting subproblems only involve small-sized matrix variables and hence can be solved efficiently. In particular, it has been showed in [73, 74, 75] that the Schatten-p ($0 < p \leqslant 1$) (quasi-)norm of any matrix is equivalent to minimizing the weighted sum of Schatten-p_1 quasi-norm and Schatten-p_2 quasi-norm of its two factor matrices for suitable p_1, $p_2 > 0$. Thus, Schatten-p ($0 < p \leqslant 1$) (quasi-)norm minimization for MC can be modeled as (1.4) with suitable choices of Ψ, Φ and $\mathcal{A}$. Promising numerical results of this method have also been reported in [73, 74].

Problem (1.4) is in general nonconvex (even when Ψ, Φ are convex) and NP-hard[2]. Therefore, in this thesis, we also focus on finding a *stationary point* of the objective $\mathcal{F}$ in (1.4). Note that $\mathcal{F}$ here also involves two blocks of variables. This kind of structure has been widely studied in the literature; see, for example, [3, 8, 22, 39, 42, 80, 91, 92, 95]. As we mentioned in Section 1.1, one popular class of methods for tackling this kind of problems is the alternating direction method of multipliers (ADMM) (see [22, 92, 95]), in which each iteration consists of an alter-

[2] Problem (1.4) is NP-hard because it contains NMF as a special case, which is NP-hard in general [82].

nating minimization of an augmented Lagrangian function that involves X, Y and some auxiliary variables, followed by updates of the associated multipliers. However, the conditions presented in [22, 92, 95] that guarantee convergence of the ADMM are too restrictive. In fact, their condition requires that the successive change of the dual variable goes to zero, which is uncheckable in practice. Moreover, updating the auxiliary variables and the multipliers can be expensive for large-scale problems. Another class of methods for (1.4) is the alternating-minimization-based (or block-coordinate-descent-type) methods (see [3, 8, 19, 32, 53, 58, 80, 91]), which alternately (exactly or inexactly) minimizes $\mathcal{F}(X, Y)$ over each block of variables and converges under some mild conditions. When $\mathcal{A}$ is not the identity map, the majorization technique can be used to simplify the subproblems. Some representative algorithms of this class are proximal alternating linearized minimization (PALM) [8], hierarchical alternating least squares (HALS) (for NMF only; see [19, 32, 53, 58]) and block coordinate descent (BCD) [91]. Comparing with ADMM, it was reported in [91] that BCD outperforms ADMM in both CPU time and solution quality for NMF. While methods such as PALM, HALS and BCD are currently the state-of-the-art algorithms for solving problems of the form (1.4), in Chapter 4 of this thesis, we develop a new iterative method for (1.4), which, according to our numerical experiments in Section 4.4, outperforms HALS and BCD for NMF, and PALM for MC.

1.3 Contributions of the thesis

The contributions of this thesis can be divided into two parts.

- **The first part is for MDP.** Following the studies in [40, 83] on convergence of nonconvex ADMM and its variant, and the recent studies in [1, 49, 84], we manage to analyze the convergence of the ADMM applied to solving the possibly nonconvex problem (1.1). In addition, we would like to point out that all the nonconvex ADMM mentioned in Section 1.1 only have a dual step-size of $\tau = 1$. While it is known that the classical ADMM converges for any $\tau \in (0, \frac{\sqrt{5}+1}{2})$ for convex problems, and that empirically $\tau \approx \frac{\sqrt{5}+1}{2}$ works best (see, for example, [26, 27, 33, 55]), to our knowledge, the algorithm with a dual step-size $\tau \neq 1$ has never been studied in the nonconvex scenarios. Thus, we also study the ADMM with a general dual step-size, which will allow more flexibilities in the design of algorithms. The specifical contributions in this part are presented as follows.

- We show that for any positive dual step-size τ less than the golden ratio, the cluster point of the sequence generated by our ADMM gives a stationary point of (1.1) if the penalty parameter is chosen above a computable threshold depending on τ, whenever the sequence is bounded. We achieve this via a potential function specifically constructed for our ADMM. To the best of our knowledge, this is the first convergence result for the ADMM in the nonconvex scenario with a possibly nontrivial dual step-size ($\tau \neq 1$). This result is also new for the convex scenario for the directly extended multi-block ADMM.
- We establish global convergence of the whole sequence generated by the ADMM under the additional assumption that the special potential function is a Kurdyka-Łojasiewicz function.
- Furthermore, we discuss an initialization strategy to guarantee the boundedness of the sequence generated by the ADMM.
- We conduct some numerical experiments to evaluate the performance of our ADMM by using different nonconvex regularizers on real datasets in Section 3.3. Our computational results illustrate the efficiency of our ADMM with a nontrivial dual step-size.

- **The second part is for MFP.** In this part, we develop a new iterative method for (1.4). Our method is based on the following potential function (specifically constructed for the objective $\mathcal{F}$ in (1.4)):

$$\Theta_{\alpha,\beta}(X,Y,Z) := \Psi(X) + \Phi(Y) + \frac{\alpha}{2}\|XY^\top - Z\|_F^2 + \frac{\beta}{2}\left\|\mathcal{A}(Z) - \boldsymbol{b}\right\|^2, \quad (1.5)$$

  where α and β are real numbers. Instead of alternately (exactly or inexactly) minimizing $\mathcal{F}(X,Y)$ or the augmented Lagrangian function, our method alternately updates X and Y by inexactly minimizing $\Theta_{\alpha,\beta}(X,Y,Z)$ over X and Y, and then updates Z by an *explicit formula*. Note that the coupled variables $XY^\top$ is now separated from $\mathcal{A}$ in our potential function. Thus, one can readily take advantage of efficient computational strategies for NMF, such as those used in HALS (see the "hierarchical-prox" updating strategy in Section 4.2), for inexactly minimizing $\Theta_{\alpha,\beta}(X,Y,Z)$ over X or Y. Furthermore, our method can be implemented for NMF and MC without explicitly forming the huge $(m \times n)$ matrix Z (see (4.46) and (4.49)) in each iteration. This significantly reduces the computational cost per iteration. Finally, a suitable non-monotone line search criterion, which is motivated by recent studies on non-monotone algorithms (see, for example, [17, 34, 88]), is also incorporated to improve the numerical performance. The specifical contributions in this part are presented as follows.

  - We propose a potential function $\Theta_{\alpha,\beta}$ and study its properties. If $\mathcal{A}\mathcal{A}^* = \mathcal{I}$ and α, β are chosen such that $\alpha\mathcal{I} + \beta\mathcal{A}^*\mathcal{A} \succ 0$ and $\frac{1}{\alpha} + \frac{1}{\beta} = 1$, then the problem $\min\limits_{X,Y,Z} \{\Theta_{\alpha,\beta}(X,Y,Z)\}$ is equivalent to (1.4) (see Theorem 4.1). Furthermore, under the weaker conditions that $\mathcal{A}\mathcal{A}^* = \mathcal{I}$ and $\frac{1}{\alpha} + \frac{1}{\beta} = 1$, we can show that (i) a stationary point of $\Theta_{\alpha,\beta}$ gives a stationary point of $\mathcal{F}$; (ii) a stationary point of $\mathcal{F}$ can be used to construct a stationary point of $\Theta_{\alpha,\beta}$ (see Theorem 4.2). Thus, one can find a stationary point of $\mathcal{F}$ by finding a stationary point of $\Theta_{\alpha,\beta}$.

- We develop a non-monotone alternating updating method based on this potential function to find a stationary point of $\Theta_{\alpha,\beta}$, and hence of $\mathcal{F}$. The convergence analysis of this method is presented. We show that our non-monotone line search criterion is well defined and any cluster point of the sequence generated by our method is a stationary point of $\mathcal{F}$ under some mild conditions. Moreover, we discuss the convergence rate for the function value if, in addition, the objective is a Kurdyka-Łojasiewicz function.

- Finally, we conduct some numerical experiments to evaluate the performance of our method for NMF and MC on real datasets in Section 4.4. Our computational results illustrate the efficiency of our method.

1.4 Organization of the thesis

This thesis is organized as follows.

- In Chapter 1, we briefly introduce the backgrounds and models for two classes of matrix optimization problems considered in this thesis and summarize the main contributions of this thesis.

- In Chapter 2, we present some basic notions and some well-known theoretical results needed in this thesis.

- In Chapter 3, we develop the three-block ADMM with a non-trivial dual step-size for MDP (1.1) and analyze its convergence. Then, we conduct numerical experiments to evaluate the performance of our ADMM.

- In Chapter 4, we derive a non-monotone alternating updating method for MFP (1.4) and analyze its convergence. The numerical experiments for nonnegative matrix factorization and matrix completion are also conducted to evaluate the performance of our method.

- In Chapter 5, we summarize our main results in this thesis and give some possible further works.

Chapter 2

Notation and Preliminaries

In this chapter, we summarize some basic notation used throughout this thesis, and recall some well-known theoretical results in the literature.

2.1 Basic notation

In this thesis, we use $\mathbb{R}^m$ to denote the set of all m-dimensional vectors and use $\mathbb{R}^{m\times n}$ to denote the set of all $m\times n$ matrices. For a vector $\boldsymbol{x}\in\mathbb{R}^m$, we let x_i denote its i-th entry, $\|\boldsymbol{x}\|$ denote the Euclidean norm of $\boldsymbol{x}$ and $\mathrm{Diag}(\boldsymbol{x})$ denote a diagonal matrix whose i-th diagonal element is x_i. For a matrix $X\in\mathbb{R}^{m\times n}$, x_{ij} denotes its ij-th entry and $\boldsymbol{x}_j$ denotes its j-th column. The ℓ_0-"norm" denoted by $\|X\|_0$ is defined as the number of nonzero entries in X and the ℓ_∞-norm denoted by $\|X\|_\infty$ is defined as the largest entry in magnitude in X. The ℓ_1-norm and ℓ_p-quasi-norm $(0<p<1)$ of X are defined as $\|X\|_1 := \sum_{i=1}^m\sum_{j=1}^n |x_{ij}|$ and $\|X\|_p := \left(\sum_{i=1}^m\sum_{j=1}^n |x_{ij}|^p\right)^{\frac{1}{p}}$, respectively. Moreover, the Schatten-p (quasi-)norm $(0<p<\infty)$ of X is defined as $\|X\|_{S_p} = \left(\sum_{i=1}^{\min(m,n)} \varsigma_i^p(X)\right)^{\frac{1}{p}}$, where $\varsigma_i(X)$ is the i-th singular value of X. For $p=2$, the Schatten-2 norm reduces to the Frobenius norm $\|X\|_F$, and for $p=1$, the Schatten-1 norm reduces to the nuclear norm $\|X\|_*$. The spectral norm of X is denoted by $\|X\|$, which is the largest singular values of X. Additionally, for two matrices X and Y

of the same size, we denote their trace inner product by $\langle X, Y\rangle := \sum_{i=1}^{m}\sum_{j=1}^{n} x_{ij}y_{ij}$. We also use $X \leqslant Y$ (resp. $X \geqslant Y$) to denote $x_{ij} \leqslant y_{ij}$ (resp. $x_{ij} \geqslant y_{ij}$) for all (i, j).

For the linear map $\mathcal{A} : \mathbb{R}^{m\times n} \to \mathbb{R}^{m\times n}$, $\mathcal{A}^*$ denotes the adjoint linear map and $\|\mathcal{A}\|$ denotes the induced operator norm of $\mathcal{A}$, i.e., $\|\mathcal{A}\| = \sup\{\|\mathcal{A}(X)\| : \|X\|_F \leqslant 1\}$. Particularly, for notational simplicity, we use $\lambda_{\max}$ (resp., $\lambda_{\min}$) to denote the largest (resp., smallest) eigenvalue of the linear map $\mathcal{A}^*\mathcal{A}$ in Chapter 3. For a linear map $\mathcal{T} : \mathbb{R}^{m\times n} \to \mathbb{R}^{m\times n}$, we write $\mathcal{T} \succ 0$ if $\mathcal{T}$ is positive definite. Finally, the identity map is denoted by $\mathcal{I}$.

For an extended-real-valued function $f : \mathbb{R}^{m\times n} \to [-\infty, \infty]$, we say that it is *proper* if $f(X) > -\infty$ for all $X \in \mathbb{R}^{m\times n}$ and its domain $\mathrm{dom} f := \{X \in \mathbb{R}^{m\times n} : f(X) < \infty\}$ is nonempty. Such a function is *lower semicontinuous* at a point $X \in \mathbb{R}^{m\times n}$ if $f(X) \leqslant \liminf_{Y\to X} f(Y)$. We say that f is lower semicontinuous or closed on $\mathbb{R}^{m\times n}$ if f is lower semicontinuous at every $X \in \mathbb{R}^{m\times n}$. For a proper function f, we use the notation $Y \xrightarrow{f} X$ to denote $Y \to X$ and $f(Y) \to f(X)$. The basic *(limiting) subdifferential* [71, Definition 8.3] of f at $X \in \mathrm{dom} f$ used in this thesis is

$$\partial f(X) := \left\{D \in \mathbb{R}^{m\times n} : \exists\, X^k \xrightarrow{f} X \text{ and } D^k \to D \text{ with } D^k \in \widehat{\partial} f(X^k) \text{ for all } k\right\},$$

where $\widehat{\partial} f(U)$ denotes the Fréchet subdifferential of f at $U \in \mathrm{dom} f$, which is the set of all $D \in \mathbb{R}^{m\times n}$ satisfying

$$\liminf_{Y\neq U,\, Y\to U} \frac{f(Y) - f(U) - \langle D, Y - U\rangle}{\|Y - U\|_F} \geqslant 0.$$

From the above definition, we can easily observe that

$$\left\{D \in \mathbb{R}^{m\times n} : \exists\, X^k \xrightarrow{f} X,\ D^k \to D,\ D^k \in \partial f(X^k)\right\} \subseteq \partial f(X). \tag{2.1}$$

We also recall that when f is continuously differentiable or convex, the above subdifferential coincides with the classical concept of derivative or convex subdifferential of

f; see, for example, [71, Exercise 8.8] and [71, Proposition 8.12]. Moreover, from the generalized Fermat's rule [71, Theorem 10.1], we know that if $X \in \mathbb{R}^{m\times n}$ is a local minimizer of f, then $0 \in \partial f(X)$. In this thesis, we say that X^* is *stationary point* of f if $0 \in \partial f(X^*)$. Additionally, for a function f with several groups of variables, we write $\partial_X f$ (resp., $\nabla_X f$) for the subdifferential (resp., derivative) of f with respect to the group of variables X.

For a nonempty closed set $\mathcal{X} \subseteq \mathbb{R}^{m\times n}$, its indicator function $\delta_{\mathcal{X}}$ is defined by

$$\delta_{\mathcal{X}}(X) = \begin{cases} 0 & \text{if } X \in \mathcal{X}, \\ +\infty & \text{otherwise.} \end{cases}$$

The normal cone of $\mathcal{X}$ at the point $X \in \mathcal{X}$ is given by $\mathcal{N}_{\mathcal{X}}(X) = \partial\delta_{\mathcal{X}}(X)$. We also use $\mathrm{dist}(X, \mathcal{X})$ to denote the distance from X to $\mathcal{X}$, i.e., $\mathrm{dist}(X, \mathcal{X}) := \inf_{Y\in\mathcal{X}} \|X - Y\|_F$, and $\mathcal{P}_{\mathcal{X}}(X)$ to denote the closest point to X in $\mathcal{X}$.

For a proper closed function $g : \mathbb{R}^m \to (-\infty, \infty]$, the proximal mapping $\mathrm{Prox}_g : \mathbb{R}^m \to \mathbb{R}^m$ of g is defined by

$$\mathrm{Prox}_g(\boldsymbol{z}) := \operatorname*{Argmin}_{\boldsymbol{x}\in\mathbb{R}^m} \left\{ g(\boldsymbol{x}) + \frac{1}{2}\|\boldsymbol{x} - \boldsymbol{z}\|^2 \right\}.$$

For any $\nu > 0$, the matrix shrinkage operator $\mathcal{S}_\nu : \mathbb{R}^{m\times n} \to \mathbb{R}^{m\times n}$ is defined by

$$\mathcal{S}_\nu(X) := U\mathrm{Diag}(\bar{\boldsymbol{s}})V^\top \text{ with } \bar{s}_i = \begin{cases} s_i - \nu, & \text{if } s_i - \nu > 0, \\ 0, & \text{otherwise,} \end{cases}$$

where $U \in \mathbb{R}^{m\times t}$, $\boldsymbol{s} \in \mathbb{R}^t_+$ and $V \in \mathbb{R}^{n\times t}$ are given by the singular value decomposition of X, i.e, $X = U\mathrm{Diag}(\boldsymbol{s})V^\top$.

2.2 Kurdyka-Łojasiewicz property

We next recall the Kurdyka-Łojasiewicz (KL) property, which plays an important role in our global convergence analysis in Section 3.2 and convergence rate in Section 4.3.

For notational simplicity, we use Ξ_η ($\eta > 0$) to denote the class of concave functions $\varphi : [0, \eta) \to \mathbb{R}_+$ satisfying: (1) $\varphi(0) = 0$; (2) φ is continuously differentiable on $(0, \eta)$ and continuous at 0; (3) $\varphi'(x) > 0$ for all $x \in (0, \eta)$. Then, the KL property can be described as follows.

Definition 2.1 (**KL property and KL function**). *Let f be a proper lower semi-continuous function.*

(i) For $\widetilde{X} \in \mathrm{dom}\,\partial f := \{X \in \mathbb{R}^{m\times n} : \partial f(X) \neq \varnothing\}$, if there exist an $\eta \in (0, +\infty]$, a neighborhood V of $\widetilde{X}$ and a function $\varphi \in \Xi_\eta$ such that for all $X \in V \cap \{X \in \mathbb{R}^{m\times n} : f(\widetilde{X}) < f(X) < f(\widetilde{X}) + \eta\}$, it holds that

$$\varphi'(f(X) - f(\widetilde{X}))\mathrm{dist}(0,\, \partial f(X)) \geqslant 1,$$

*then f is said to have the **Kurdyka-Łojasiewicz (KL)** property at $\widetilde{X}$.*

(ii) If f satisfies the KL property at each point of $\mathrm{dom}\,\partial f$, then f is called a KL function.

We refer the interested readers to [2] and references therein for examples of KL functions. Based on the above definition, we can further describe the KL exponent, which is defined [2, 3] as follows.

Definition 2.2. *Suppose that f is a proper closed function satisfying the KL property at $\widetilde{X} \in \mathrm{dom}\,\partial f := \{X \in \mathbb{R}^{m\times n} : \partial f(X) \neq \varnothing\}$ with $\varphi(s) = \tilde{a}s^{1-\vartheta}$ for some $\tilde{a} > 0$ and $\vartheta \in [0, 1)$, i.e., there exist $a, \varepsilon, \eta > 0$ such that*

$$\mathrm{dist}(0,\, \partial f(X)) \geqslant a \left(f(X) - f(\widetilde{X})\right)^{\vartheta}$$

whenever $X \in \mathrm{dom}\,\partial f$, $\|X - \widetilde{X}\|_F \leqslant \varepsilon$ and $f(\widetilde{X}) < f(X) < f(\widetilde{X}) + \eta$. Then, f is said to have the KL property at $\tilde{X}$ with an exponent ϑ. If f is a KL function and *has the same exponent* ϑ *at any* $\widetilde{X} \in \operatorname{dom} \partial f$, *then* f *is said to be a KL function with an exponent* ϑ.

Finally, we recall the following uniformized KL property, which was established in [8, Lemma 6].

Proposition 2.1 (Uniformized KL property). *Suppose that* f *is a proper lower semicontinuous function and* Γ *is a compact set. If* $f \equiv f^*$ *on* Γ *for some constant* f^* *and satisfies the KL property at each point of* Γ, *then there exist* $\varepsilon > 0$, $\eta > 0$ *and* $\varphi \in \Xi_\eta$ *such that*

$$\varphi'(f(X) - f^*) \operatorname{dist}(0, \partial f(X)) \geqslant 1$$

for all $X \in \{X \in \mathbb{R}^{m\times n} : \operatorname{dist}(X, \Gamma) < \varepsilon\} \cap \{X \in \mathbb{R}^{m\times n} : f^* < f(X) < f^* + \eta\}$.

Chapter 3

ADMM for Matrix Decomposition Problem

In this chapter, we consider applying ADMM with a general dual step-size for solving the matrix decomposition problem (1.1), i.e.,

$$\min_{L,S}\ \mathcal{F}(L,S) := \Psi(L) + \Phi(S) + \frac{1}{2}\left\| D - \mathcal{A}\left[\mathcal{B}(L) + \mathcal{C}(S)\right]\right\|_F^2 .$$

Note that this problem is possibly nonconvex, nonsmooth and non-Lipschitz. Thus, we focus on finding a stationary point of the objective $\mathcal{F}$, i.e., finding a point $(L^*,\, S^*)$ such that $0 \in \partial\mathcal{F}(L^*, S^*)$. Moreover, from [71, Exercise 8.8] and [71, Proposition 10.5], we see that

$$\partial\mathcal{F}(L,\, S) = \begin{pmatrix} \partial\Psi(L) + \mathcal{B}^*\mathcal{A}^*\left(\mathcal{A}(\mathcal{B}(L) + \mathcal{C}(S)) - D\right) \\ \partial\Phi(S) + \mathcal{C}^*\mathcal{A}^*\left(\mathcal{A}(\mathcal{B}(L) + \mathcal{C}(S)) - D\right) \end{pmatrix}.$$

Then, a stationary point (L^*, S^*) of $\mathcal{F}$ must satisfy the following first-order necessary conditions:

$$\begin{cases} 0 \in \partial\Psi(L^*) + \mathcal{B}^*\mathcal{A}^*\left(\mathcal{A}(\mathcal{B}(L^*) + \mathcal{C}(S^*)) - D\right), \\ 0 \in \partial\Phi(S^*) + \mathcal{C}^*\mathcal{A}^*\left(\mathcal{A}(\mathcal{B}(L^*) + \mathcal{C}(S^*)) - D\right). \end{cases} \tag{3.1}$$

Later, we will show that for any dual step-size less than the golden ratio $\frac{\sqrt{5}+1}{2}$, there exists a computable threshold $\bar{\beta}$ such that if the penalty parameter β is chosen above

$\bar{\beta}$ and the sequence thus generated by our ADMM is bounded, then the cluster point of the sequence gives a stationary point of $\mathcal{F}$, i.e., a point satisfying (3.1). We achieve this via a potential function specifically constructed for our ADMM. Moreover, we establish the global convergence of the whole sequence if, in addition, this special potential function is a Kurdyka-Łojasiewicz function. Furthermore, we present a simple strategy for initializing the algorithm to guarantee boundedness of the sequence. Finally, we perform numerical experiments comparing our ADMM with the proximal alternating linearized minimization (PALM) proposed in [8] on the background/foreground extraction problem with real data. The numerical results show that our ADMM with a nontrivial dual step-size is efficient.

The rest of this chapter is organized as follows. We first present the ADMM for (1.1) in Section 3.1. We then present the convergence analysis in Section 3.2. Some numerical results are reported in Section 3.3.

3.1 Alternating direction method of multipliers

In this section, we present an ADMM for (1.1). To this end, we first rewrite (1.1) as the following equivalent form:

$$\begin{aligned} \min_{L,S,Z} \quad & \Psi(L) + \Phi(S) + \frac{1}{2}\|D - \mathcal{A}(Z)\|_F^2 \\ \text{s.t.} \quad & \mathcal{B}(L) + \mathcal{C}(S) = Z. \end{aligned} \tag{3.2}$$

We then introduce the augmented Lagrangian function of (3.2) as follows:

$$\begin{aligned} \mathcal{L}_\beta(L, S, Z, \Lambda) = \Psi(L) + \Phi(S) + \frac{1}{2}\|D - \mathcal{A}(Z)\|_F^2 \\ - \langle \Lambda,\ \mathcal{B}(L) + \mathcal{C}(S) - Z\rangle + \frac{\beta}{2}\|\mathcal{B}(L) + \mathcal{C}(S) - Z\|_F^2, \end{aligned}$$

where $\Lambda \in \mathbb{R}^{m\times n}$ is the Lagrangian multiplier and $\beta > 0$ is the penalty parameter. Now, we are ready to present the complete ADMM for (3.2) (equivalently (1.1)) in

Algorithm 1.

Algorithm 1 ADMM for solving (3.2)

Input: Initial point (S^0, Z^0, Λ^0), dual step-size parameter $\tau > 0$, penalty parameter $\beta > 0$. Set $k = 0$.

while a termination criterion is not met, **do**

Step 1. Set

$$\begin{cases} L^{k+1} \in \operatorname*{Argmin}_{L} \ \mathcal{L}_\beta(L, S^k, Z^k, \Lambda^k), & (3.3a) \\ S^{k+1} \in \operatorname*{Argmin}_{S} \ \mathcal{L}_\beta(L^{k+1}, S, Z^k, \Lambda^k), & (3.3b) \\ Z^{k+1} = \operatorname*{argmin}_{Z} \ \mathcal{L}_\beta(L^{k+1}, S^{k+1}, Z, \Lambda^k), & (3.3c) \\ \Lambda^{k+1} = \Lambda^k - \tau\beta(\mathcal{B}(L^{k+1}) + \mathcal{C}(S^{k+1}) - Z^{k+1}). \end{cases}$$

Step 2. Set $k := k + 1$.

end while

Output: $(L^k,\ S^k)$

Comparing with the ADMM considered in [40], the above algorithm has an extra dual step-size parameter $\tau > 0$ in the Λ-update. Such a dual step-size was introduced in [27, 33] for the classical ADMM (i.e., for convex problems with two separate blocks of variables), and was further studied in [26, 55, 78, 90] for other variants of the ADMM. Numerically, it was also demonstrated in [78] that a larger dual step-size ($\tau \approx \frac{\sqrt{5}+1}{2}$) results in faster convergence for the convex problems they consider. Thus, we adapt this dual step-size τ in our algorithm above. Surprisingly, in our numerical experiments, a parameter choice of $\tau \approx \frac{\sqrt{5}+1}{2}$ leads to the worst performance for our nonconvex problems.

When $\tau = 1$, the above algorithm is a special case of the general algorithm studied in [40] when Ψ and Φ are smooth functions, or convex nonsmooth functions. The algorithm is shown to converge when β is chosen above a computable threshold. However, their convergence result cannot be directly applied when $\tau \neq 1$ or when Φ

is nonsmooth and nonconvex. Nevertheless, following their analysis and the related studies [49, 84, 83], the above algorithm can be shown to be convergent under suitable assumptions. We will present the convergence analysis in Section 3.2.

Before ending this section, we further discuss the three subproblems in Algorithm 1. First, notice that the L-update and S-update are given by

$$\begin{cases} L^{k+1} \in \underset{L}{\operatorname{Argmin}} \ \left\{ \Psi(L) + \dfrac{\beta}{2} \|\mathcal{B}(L) + \mathcal{C}(S^k) - Z^k - \dfrac{1}{\beta}\Lambda^k\|_F^2 \right\}, \\ S^{k+1} \in \underset{S}{\operatorname{Argmin}} \ \left\{ \Phi(S) + \dfrac{\beta}{2} \|\mathcal{B}(L^{k+1}) + \mathcal{C}(S) - Z^k - \dfrac{1}{\beta}\Lambda^k\|_F^2 \right\}. \end{cases}$$

In general, these two subproblems are not easy to solve. However, when Ψ and Φ are chosen to be some common regularizers used in the literature, for example, $\Psi(L) = \|L\|_*$ and $\Phi(S) = \|S\|_1$, then these subproblems can be solved efficiently via the proximal gradient method. Additionally, when $\Psi(L) = \delta_\Omega(L)$ with Ω being a closed convex set and $\mathcal{B} = \mathcal{I}$, the L-update can be given explicitly by

$$L^{k+1} = \mathcal{P}_\Omega \left(-\mathcal{C}(S^k) + Z^k + \frac{1}{\beta}\Lambda^k \right),$$

which can be computed efficiently if Ω is simple, for example, when $\Omega = \{L \in \mathbb{R}^{m\times n} \mid \|L\|_\infty \leqslant l,\ L_{:1} = L_{:2} = \cdots = L_{:n}\}$ for some $l > 0$. For the S-update, when Φ is given by (1.2) with ϕ being one of the penalty functions presented in the introduction and $\mathcal{C} = \mathcal{I}$, it can be solved efficiently via a simple root-finding procedure. Finally, from the optimality conditions of (3.3c), the Z^{k+1} can be obtained by solving the following linear system

$$\mathcal{A}^*\mathcal{A}(Z) + \beta Z = \mathcal{A}^*(D) - \Lambda^k + \beta \left(\mathcal{B}(L^{k+1}) + \mathcal{C}(S^{k+1}) \right),$$

whose complexity would depend on the choice of $\mathcal{A}$ in our model (1.1). For example, when $\mathcal{A}$ is just the identity map, the Z^{k+1} is given explicitly by

$$Z^{k+1} = \frac{1}{1+\beta} \left[D - \Lambda^k + \beta \left(\mathcal{B}(L^{k+1}) + \mathcal{C}(S^{k+1}) \right) \right].$$

3.2 Convergence analysis of ADMM

In this section, we discuss the convergence of Algorithm 1 for $0 < \tau < \frac{1+\sqrt{5}}{2}$. We first present the first-order optimality conditions for the subproblems in Algorithm 1 as follows, which will be used repeatedly in our convergence analysis below.

$$\begin{cases} 0 \in \partial\Psi(L^{k+1}) - \mathcal{B}^*(\Lambda^k) + \beta\mathcal{B}^*\left(\mathcal{B}(L^{k+1}) + \mathcal{C}(S^k) - Z^k\right), & \text{(3.4a)} \\ 0 \in \partial\Phi(S^{k+1}) - \mathcal{C}^*(\Lambda^k) + \beta\mathcal{C}^*\left(\mathcal{B}(L^{k+1}) + \mathcal{C}(S^{k+1}) - Z^k\right), & \text{(3.4b)} \\ 0 = \mathcal{A}^*(\mathcal{A}(Z^{k+1}) - D) + \Lambda^k - \beta(\mathcal{B}(L^{k+1}) + \mathcal{C}(S^{k+1}) - Z^{k+1}), & \text{(3.4c)} \\ \Lambda^{k+1} - \Lambda^k = -\tau\beta\left(\mathcal{B}(L^{k+1}) + \mathcal{C}(S^{k+1}) - Z^{k+1}\right). & \text{(3.4d)} \end{cases}$$

Our convergence analysis is largely based on the following potential function:

$$\Theta_{\tau,\beta}(L, S, Z, \Lambda) = \mathcal{L}_\beta(L, S, Z, \Lambda) + \theta(\tau)\beta\|\mathcal{B}(L) + \mathcal{C}(S) - Z\|_F^2,$$

where

$$\theta(\tau) := \max\left\{1 - \tau,\ \frac{(\tau-1)\tau^2}{1+\tau-\tau^2}\right\}, \quad \text{for} \quad 0 < \tau < \frac{1+\sqrt{5}}{2}. \tag{3.5}$$

Note that $\theta(\cdot)$ is a convex and nonnegative function on $\left(0,\ \frac{1+\sqrt{5}}{2}\right)$. Thus, for any (L, S, Z, Λ), we have $\Theta_{\tau,\beta}(L, S, Z, \Lambda) \geqslant \mathcal{L}_\beta(L, S, Z, \Lambda)$ for $0 < \tau < \frac{1+\sqrt{5}}{2}$, and the equality holds when $\tau = 1$ (so that $\theta(\tau) = 0$).

Our convergence analysis also relies on the following assumption.

Assumption 3.1. *Ψ, Φ, $\mathcal{B}$, $\mathcal{C}$, β and τ are chosen such that*

(a1) $\mathcal{B}^*\mathcal{B} \succeq \sigma\mathcal{I}$ for some $\sigma > 0$ and $\mathcal{C}^*\mathcal{C} \succeq \sigma'\mathcal{I}$ for some $\sigma' > 0$;

(a2) Ψ is continuous on its domain;

(a3) the first iterate $(L^1, S^1, Z^1, \Lambda^1)$ satisfies

$$\Theta_{\tau,\beta}(L^1, S^1, Z^1, \Lambda^1) < h_0 := \liminf_{\|L\|_F + \|S\|_F \to \infty} \Psi(L) + \Phi(S).$$

Remark 3.1 (Note on Assumption 3.1). *(i) Since $\mathcal{B}$ and $\mathcal{C}$ in (1.1) are injective, (a1) holds trivially; (ii) (a2) holds for many common regularizers (for example, the nuclear norm) or the indicator function of a set; (iii) (a3) places conditions on the first iterate of the algorithm. It is not hard to observe that this assumption holds trivially if both Ψ and Φ are coercive, i.e., if $\liminf_{\|L\|_F+\|S\|_F\to\infty} \Psi(L)+\Phi(S)=\infty$. We will discuss more sufficient conditions for this assumption after our convergence results, i.e., after Theorem 3.2.*

We now start our convergence analysis by proving the following preparatory lemma, which states that the potential function is decreasing along the sequence generated from Algorithm 1 if the penalty parameter β is chosen above a computable threshold.

Lemma 3.1. *Suppose that $0<\tau<\frac{1+\sqrt{5}}{2}$ and $\{(L^k,S^k,Z^k,\Lambda^k)\}$ is a sequence generated by Algorithm 1. If (a1) in Assumption 3.1 holds, then for $k\geqslant 1$, we have*

$$\begin{aligned}&\Theta_{\tau,\beta}(L^{k+1},S^{k+1},Z^{k+1},\Lambda^{k+1})-\Theta_{\tau,\beta}(L^k,S^k,Z^k,\Lambda^k)\\&\leqslant\left(\max\left\{\tfrac{1}{\tau},\tfrac{\tau^2}{1+\tau-\tau^2}\right\}\cdot\tfrac{\lambda_{\max}^2}{\beta}-\tfrac{\lambda_{\min}+\beta}{2}\right)\|Z^{k+1}-Z^k\|_F^2-\tfrac{\sigma\beta}{2}\|L^{k+1}-L^k\|_F^2,\end{aligned}\tag{3.6}$$

where $\lambda_{\max}$ (resp., $\lambda_{\min}$) denotes the largest (resp., smallest) eigenvalue of the linear map $\mathcal{A}^*\mathcal{A}$. Moreover, if $\beta\geqslant-\frac{\lambda_{\min}}{2}+\frac{1}{2}\sqrt{\lambda_{\min}^2+\max\left\{\frac{1}{\tau},\frac{\tau^2}{1+\tau-\tau^2}\right\}\cdot 8\lambda_{\max}^2}$, then the sequence $\{\Theta_{\tau,\beta}(L^k,S^k,Z^k,\Lambda^k)\}_{k=1}^{\infty}$ is decreasing.

Proof. We start our proof by noticing that

$$\begin{aligned}&\Theta_{\tau,\beta}(L^{k+1},S^{k+1},Z^{k+1},\Lambda^{k+1})-\Theta_{\tau,\beta}(L^{k+1},S^{k+1},Z^{k+1},\Lambda^k)\\&=-\langle\Lambda^{k+1}-\Lambda^k,L^{k+1}+S^{k+1}-Z^{k+1}\rangle=\frac{1}{\tau\beta}\|\Lambda^{k+1}-\Lambda^k\|_F^2,\end{aligned}\tag{3.7}$$

where the last equality follows from (3.4d). We next derive an upper bound of

$\|\Lambda^{k+1} - \Lambda^k\|_F^2$. To proceed, we first note from (3.4c) that

$$\begin{aligned} 0 &= \mathcal{A}^*(\mathcal{A}(Z^{k+1}) - D) + \Lambda^k - \beta(\mathcal{B}(L^{k+1}) + \mathcal{C}(S^{k+1}) - Z^{k+1}) \\ &= \mathcal{A}^*(\mathcal{A}(Z^{k+1}) - D) + \Lambda^k + \frac{1}{\tau}(\Lambda^{k+1} - \Lambda^k) \\ \Longrightarrow \quad \Lambda^{k+1} &= \tau\mathcal{A}^*(D - \mathcal{A}(Z^{k+1})) + (1-\tau)\Lambda^k, \end{aligned}$$

where the second equality follows from (3.4d). Hence, for $k \geqslant 1$,

$$\begin{aligned} &\Lambda^{k+1} - \Lambda^k \\ = \; &[\tau\mathcal{A}^*(D - \mathcal{A}(Z^{k+1})) + (1-\tau)\Lambda^k] - [\tau\mathcal{A}^*(D - \mathcal{A}(Z^k)) + (1-\tau)\Lambda^{k-1}] \\ = \; &\tau\mathcal{A}^*\mathcal{A}(Z^k - Z^{k+1}) + (1-\tau)(\Lambda^k - \Lambda^{k-1}). \end{aligned} \tag{3.8}$$

We now consider two separate cases: $0 < \tau \leqslant 1$ and $1 < \tau < \frac{1+\sqrt{5}}{2}$.

- For $0 < \tau \leqslant 1$, it follows from the convexity of $\|\cdot\|_F^2$ that

$$\begin{aligned} \|\Lambda^{k+1} - \Lambda^k\|_F^2 &= \left\|\tau\mathcal{A}^*\mathcal{A}(Z^k - Z^{k+1}) + (1-\tau)(\Lambda^k - \Lambda^{k-1})\right\|_F^2 \\ &\leqslant \tau\lambda_{\max}^2\|Z^{k+1} - Z^k\|_F^2 + (1-\tau)\|\Lambda^k - \Lambda^{k-1}\|_F^2. \end{aligned}$$

  We further add $-(1-\tau)\left\|\Lambda^{k+1} - \Lambda^k\right\|_F^2$ to both sides of the above inequality and simplify the resulting inequality to get

$$\begin{aligned} &\|\Lambda^{k+1} - \Lambda^k\|_F^2 \\ \leqslant \; &\lambda_{\max}^2\|Z^{k+1} - Z^k\|_F^2 + \tfrac{1-\tau}{\tau}\left(\|\Lambda^k - \Lambda^{k-1}\|_F^2 - \|\Lambda^{k+1} - \Lambda^k\|_F^2\right) \\ = \; &(1-\tau)\tau\beta^2\left(\|\mathcal{B}(L^k) + \mathcal{C}(S^k) - Z^k\|_F^2 - \|\mathcal{B}(L^{k+1}) + \mathcal{C}(S^{k+1}) - Z^{k+1}\|_F^2\right) \\ &+ \lambda_{\max}^2\|Z^{k+1} - Z^k\|_F^2. \end{aligned} \tag{3.9}$$

  where the last equality follows from (3.4d).

- For $1 < \tau < \frac{1+\sqrt{5}}{2}$, dividing τ from both sides of (3.8), we have

$$\begin{aligned}\frac{1}{\tau}\left(\Lambda^{k+1}-\Lambda^{k}\right) &= \mathcal{A}^*\mathcal{A}\left(Z^k - Z^{k+1}\right) + \left(\frac{1}{\tau}-1\right)\left(\Lambda^k - \Lambda^{k-1}\right) \\ &= \frac{1}{\tau}\cdot\tau\mathcal{A}^*\mathcal{A}\left(Z^k - Z^{k+1}\right) + \left(1-\frac{1}{\tau}\right)\left(\Lambda^{k-1}-\Lambda^k\right).\end{aligned}$$

  This together with $0 < \frac{1}{\tau} < 1$ and the convexity of $\|\cdot\|_F^2$, implies that

$$\begin{aligned}\left\|\tfrac{1}{\tau}\left(\Lambda^{k+1}-\Lambda^k\right)\right\|_F^2 &\leqslant \tfrac{1}{\tau}\|\tau\mathcal{A}^*\mathcal{A}\left(Z^k-Z^{k+1}\right)\|_F^2 + \left(1-\tfrac{1}{\tau}\right)\|\Lambda^{k-1}-\Lambda^k\|_F^2 \\ &\leqslant \tau\lambda_{\max}^2\|Z^{k+1}-Z^k\|_F^2 + \left(1-\tfrac{1}{\tau}\right)\|\Lambda^k-\Lambda^{k-1}\|_F^2 \\ \Longrightarrow \quad \|\Lambda^{k+1}-\Lambda^k\|_F^2 &\leqslant \tau^3\lambda_{\max}^2\|Z^{k+1}-Z^k\|_F^2 + \left(\tau^2-\tau\right)\|\Lambda^k-\Lambda^{k-1}\|_F^2.\end{aligned}$$

  Then, adding $-\left(\tau^2-\tau\right)\left\|\Lambda^{k+1}-\Lambda^k\right\|_F^2$ to both sides of the above inequality, simplifying the resulting inequality and using the fact that $1+\tau-\tau^2 > 0$ for $1 < \tau < \frac{1+\sqrt{5}}{2}$, we see that

$$\begin{aligned}&\|\Lambda^{k+1}-\Lambda^k\|_F^2 \\ \leqslant\ & \frac{\tau^3\lambda_{\max}^2}{1+\tau-\tau^2}\|Z^{k+1}-Z^k\|_F^2 + \frac{\tau^2-\tau}{1+\tau-\tau^2}\left(\|\Lambda^k-\Lambda^{k-1}\|_F^2 - \|\Lambda^{k+1}-\Lambda^k\|_F^2\right) \\ =\ & \frac{\tau^3\lambda_{\max}^2}{1+\tau-\tau^2}\|Z^{k+1}-Z^k\|_F^2 + \frac{(\tau-1)\tau^3\beta^2}{1+\tau-\tau^2}\big(\|\mathcal{B}(L^k)+\mathcal{C}(S^k)-Z^k\|_F^2 \\ & - \|\mathcal{B}(L^{k+1})+\mathcal{C}(S^{k+1})-Z^{k+1}\|_F^2\big),\end{aligned} \tag{3.10}$$

  where the equality follows from (3.4d).

Thus, for $0 < \tau < \frac{1+\sqrt{5}}{2}$, combining (3.9), (3.10) and recalling the definition of $\theta(\tau)$ in (3.5), we have

$$\begin{aligned}\frac{1}{\tau\beta}\|\Lambda^{k+1}-\Lambda^k\|_F^2 \ &\leqslant\ \max\left\{\frac{1}{\tau}, \frac{\tau^2}{1+\tau-\tau^2}\right\}\cdot\frac{\lambda_{\max}^2}{\beta}\|Z^{k+1}-Z^k\|_F^2 \\ &+ \theta(\tau)\beta\left(\|\mathcal{B}(L^k)+\mathcal{C}(S^k)-Z^k\|_F^2 - \|\mathcal{B}(L^{k+1})+\mathcal{C}(S^{k+1})-Z^{k+1}\|_F^2\right).\end{aligned} \tag{3.11}$$

Next, note that the function $Z \mapsto \mathcal{L}_\beta(L^{k+1}, S^{k+1}, Z, \Lambda^k)$ is strongly convex with modulus at least $\lambda_{\min} + \beta$. Using this fact and the definition of Z^{k+1} as a minimizer in (3.3c), we see that

$$\begin{aligned}
&\Theta_{\tau,\beta}(L^{k+1}, S^{k+1}, Z^{k+1}, \Lambda^k) - \Theta_{\tau,\beta}(L^{k+1}, S^{k+1}, Z^k, \Lambda^k)\\
=\;& \mathcal{L}_\beta(L^{k+1}, S^{k+1}, Z^{k+1}, \Lambda^k) - \mathcal{L}_\beta(L^{k+1}, S^{k+1}, Z^k, \Lambda^k)\\
&+ \theta(\tau)\beta \left(\|\mathcal{B}(L^{k+1}) + \mathcal{C}(S^{k+1}) - Z^{k+1}\|_F^2 - \|\mathcal{B}(L^{k+1}) + \mathcal{C}(S^{k+1}) - Z^k\|_F^2\right)\\
\leqslant\;& \theta(\tau)\beta \left(\|\mathcal{B}(L^{k+1}) + \mathcal{C}(S^{k+1}) - Z^{k+1}\|_F^2 - \|\mathcal{B}(L^{k+1}) + \mathcal{C}(S^{k+1}) - Z^k\|_F^2\right)\\
&- \tfrac{\lambda_{\min}+\beta}{2}\|Z^{k+1} - Z^k\|_F^2.
\end{aligned} \tag{3.12}$$

Moreover, using the fact that S^{k+1} is a minimizer in (3.3b), we have

$$\begin{aligned}
&\Theta_{\tau,\beta}(L^{k+1}, S^{k+1}, Z^k, \Lambda^k) - \Theta_{\tau,\beta}(L^{k+1}, S^k, Z^k, \Lambda^k)\\
=\;& \mathcal{L}_\beta(L^{k+1}, S^{k+1}, Z^k, \Lambda^k) - \mathcal{L}_\beta(L^{k+1}, S^k, Z^k, \Lambda^k)\\
&+ \theta(\tau)\beta \left(\|\mathcal{B}(L^{k+1}) + \mathcal{C}(S^{k+1}) - Z^k\|_F^2 - \|\mathcal{B}(L^{k+1}) + \mathcal{C}(S^k) - Z^k\|_F^2\right)\\
\leqslant\;& \theta(\tau)\beta \left(\|\mathcal{B}(L^{k+1}) + \mathcal{C}(S^{k+1}) - Z^k\|_F^2 - \|\mathcal{B}(L^{k+1}) + \mathcal{C}(S^k) - Z^k\|_F^2\right).
\end{aligned} \tag{3.13}$$

Finally, note that $L \mapsto \mathcal{L}_\beta(L, S^k, Z^k, \Lambda^k)$ is strongly convex with modulus at least $\sigma\beta$ from (a1) in Assumption 3.1. From this, we can similarly obtain

$$\begin{aligned}
&\Theta_{\tau,\beta}(L^{k+1}, S^k, Z^k, \Lambda^k) - \Theta_{\tau,\beta}(L^k, S^k, Z^k, \Lambda^k)\\
\leqslant\;& \theta(\tau)\beta \left(\|\mathcal{B}(L^{k+1}) + \mathcal{C}(S^k) - Z^k\|_F^2 - \|\mathcal{B}(L^k) + \mathcal{C}(S^k) - Z^k\|_F^2\right)\\
&- \frac{\sigma\beta}{2}\|L^{k+1} - L^k\|_F^2.
\end{aligned} \tag{3.14}$$

Thus, summing (3.7), (3.11), (3.12), (3.13) and (3.14), we obtain (3.6).

Now, suppose in addition that $\beta \geqslant -\frac{\lambda_{\min}}{2} + \frac{1}{2}\sqrt{\lambda_{\min}^2 + \max\left\{\frac{1}{\tau}, \frac{\tau^2}{1+\tau-\tau^2}\right\} \cdot 8\lambda_{\max}^2}$. Then, it is easy to check that

$$\max\left\{\frac{1}{\tau},\ \frac{\tau^2}{1+\tau-\tau^2}\right\} \cdot \frac{\lambda_{\max}^2}{\beta} - \frac{\lambda_{\min}+\beta}{2} \leqslant 0.$$

Hence we see from (3.6) that

$$\Theta_{\tau,\beta}(L^{k+1}, S^{k+1}, Z^{k+1}, \Lambda^{k+1}) - \Theta_{\tau,\beta}(L^k, S^k, Z^k, \Lambda^k) \leqslant 0,$$

which means that $\{\Theta_{\tau,\beta}(L^k, S^k, Z^k, \Lambda^k)\}_{k=1}^{\infty}$ is decreasing. This completes the proof.

□

We next show that the sequence generated by Algorithm 1 is bounded if β is chosen above a computable threshold, under (a1) and (a3) in Assumption 3.1. For notational simplicity, from now on, we let

$$\bar{\beta} := \max\left\{\max\{1/\tau, \tau\} \cdot \lambda_{\max}, -\tfrac{\lambda_{\min}}{2} + \tfrac{1}{2}\sqrt{\lambda_{\min}^2 + \max\left\{\tfrac{1}{\tau}, \tfrac{\tau^2}{1+\tau-\tau^2}\right\} \cdot 8\lambda_{\max}^2}\right\}. \tag{3.15}$$

Proposition 3.1 (**Boundedness of sequence generated by ADMM**)**.** *Suppose that $0 < \tau < \frac{1+\sqrt{5}}{2}$ and $\beta > \bar{\beta}$. If (a1) and (a3) in Assumption 3.1 hold, then a sequence $\{(L^k, S^k, Z^k, \Lambda^k)\}_{k=1}^{\infty}$ generated by Algorithm 1 is bounded.*

Proof. With our choice of β and (a1) in Assumption 3.1, we see immediately from Lemma 3.1 that the sequence $\{\Theta_{\tau,\beta}(L^k, S^k, Z^k, \Lambda^k)\}_{k=1}^{\infty}$ is decreasing. This together with (a3) in Assumption 3.1 shows that, for $k \geqslant 1$,

$$\begin{aligned}
&h_0 \;>\; \Theta_{\tau,\beta}(L^1, S^1, Z^1, \Lambda^1) \;\geqslant\; \Theta_{\tau,\beta}(L^k, S^k, Z^k, \Lambda^k) \\
&= \Psi(L^k) + \Phi(S^k) + \frac{1}{2}\|D - \mathcal{A}(Z^k)\|_F^2 - \langle \Lambda^k, \mathcal{B}(L^k) + \mathcal{C}(S^k) - Z^k\rangle \\
&\quad + (1 + 2\theta(\tau))\,\frac{\beta}{2}\|\mathcal{B}(L^k) + \mathcal{C}(S^k) - Z^k\|_F^2 \\
&= \Psi(L^k) + \Phi(S^k) + \frac{1}{2}\|D - \mathcal{A}(Z^k)\|_F^2 + \frac{\beta}{2}\|\mathcal{B}(L^k) + \mathcal{C}(S^k) - Z^k - \frac{1}{\beta}\Lambda^k\|_F^2 \\
&\quad - \frac{1}{2\beta}\|\Lambda^k\|_F^2 + \theta(\tau)\beta\|\mathcal{B}(L^k) + \mathcal{C}(S^k) - Z^k\|_F^2,
\end{aligned} \tag{3.16}$$

where the last equality is obtained by completing the square. We next derive an upper bound for $\left\|\Lambda^k\right\|_F^2$. We start by substituting (3.4d) into (3.4c) and rearranging terms to obtain

$$\begin{aligned} 0 &= \mathcal{A}^*(\mathcal{A}(Z^k) - D) + \Lambda^{k-1} + \frac{1}{\tau}(\Lambda^k - \Lambda^{k-1}) \\ \Longrightarrow \quad -\tau\Lambda^k &= \tau\mathcal{A}^*(\mathcal{A}(Z^k) - D) + (1-\tau)\,(\Lambda^k - \Lambda^{k-1}). \end{aligned}$$

We now consider two different cases:

- For $0 < \tau \leqslant 1$, it follows from the convexity of $\|\cdot\|_F^2$ and (3.17) that

$$\begin{aligned} \|-\tau\Lambda^k\|_F^2 &\leqslant \tau\|\mathcal{A}^*(\mathcal{A}(Z^k) - D)\|_F^2 + (1-\tau)\|\Lambda^k - \Lambda^{k-1}\|_F^2 \\ &\leqslant \tau\lambda_{\max}\|\mathcal{A}(Z^k) - D\|_F^2 + (1-\tau)\|\Lambda^k - \Lambda^{k-1}\|_F^2 \\ &= \tau\lambda_{\max}\|\mathcal{A}(Z^k) - D\|_F^2 + (1-\tau)\tau^2\beta^2\|\mathcal{B}(L^k) + \mathcal{C}(S^k) - Z^k\|_F^2, \end{aligned}$$

  where the equality follows from (3.4d). Then, we have

$$\|\Lambda^k\|_F^2 \leqslant \frac{\lambda_{\max}}{\tau}\|\mathcal{A}(Z^k) - D\|_F^2 + (1-\tau)\beta^2\|\mathcal{B}(L^k) + \mathcal{C}(S^k) - Z^k\|_F^2. \tag{3.17}$$

- For $1 < \tau < \frac{1+\sqrt{5}}{2}$, by dividing $-\tau$ from both sides of (3.4c), we obtain

$$\Lambda^k = \frac{1}{\tau}\,\tau\mathcal{A}^*(D - \mathcal{A}(Z^k)) + \left(1 - \frac{1}{\tau}\right)(\Lambda^k - \Lambda^{k-1}).$$

  Then, since $0 < \frac{1}{\tau} < 1$, using the convexity of $\|\cdot\|_F^2$ and (3.4d), we have

$$\begin{aligned} \|\Lambda^k\|_F^2 &\leqslant \frac{1}{\tau}\|\tau\mathcal{A}^*(D - \mathcal{A}(Z^k))\|_F^2 + \left(1 - \frac{1}{\tau}\right)\|\Lambda^k - \Lambda^{k-1}\|_F^2 \\ &\leqslant \tau\lambda_{\max}\|D - \mathcal{A}(Z^k)\|_F^2 + (\tau-1)\tau\beta^2\|\mathcal{B}(L^k) + \mathcal{C}(S^k) - Z^k\|_F^2. \end{aligned} \tag{3.18}$$

Thus, combining (3.17) and (3.18), we have

$$\begin{aligned} \|\Lambda^k\|_F^2 &\leqslant \max\{1/\tau, \tau\}\cdot\lambda_{\max}\|D - \mathcal{A}(Z^k)\|_F^2 \\ &\quad + \max\{1-\tau, (\tau-1)\tau\}\beta^2\|\mathcal{B}(L^k) + \mathcal{C}(S^k) - Z^k\|_F^2 \\ \Longrightarrow \quad -\frac{1}{2\beta}\|\Lambda^k\|_F^2 &\geqslant -\tfrac{\max\{1/\tau,\tau\}\lambda_{\max}}{2\beta}\|D - \mathcal{A}(Z^k)\|_F^2 \\ &\quad - \tfrac{\max\{1-\tau,(\tau-1)\tau\}\beta}{2}\|\mathcal{B}(L^k) + \mathcal{C}(S^k) - Z^k\|_F^2. \end{aligned}$$

Substituting (3.19) into (3.16), we have

$$
\begin{aligned}
h_0 > \Theta_{\tau,\beta}(L^k, S^k, Z^k, \Lambda^k) &\geqslant \Psi(L^k) + \Phi(S^k) \\
&+ \frac{1}{2}\left(1 - \max\{1/\tau, \tau\} \cdot \frac{\lambda_{\max}}{\beta}\right) \|D - \mathcal{A}(Z^k)\|_F^2 \\
&+ \frac{\beta}{2}\|\mathcal{B}(L^k) + \mathcal{C}(S^k) - Z^k - \frac{1}{\beta}\Lambda^k\|_F^2 \\
&+ \left[2\theta(\tau) - \max\{1-\tau, (\tau-1)\tau\}\right] \cdot \frac{\beta}{2}\|\mathcal{B}(L^k) + \mathcal{C}(S^k) - Z^k\|_F^2.
\end{aligned}
\tag{3.19}
$$

With (3.19) established, we are now ready to prove the boundedness of the sequence. We start with the observation that for $0 < \tau < \frac{1+\sqrt{5}}{2}$ and $\beta > \overline{\beta}$, we always have

$$
1 - \max\{1/\tau, \tau\} \cdot \frac{\lambda_{\max}}{\beta} > 0 \tag{3.20}
$$

and

$$
2\theta(\tau) - \max\{1-\tau,\, (\tau-1)\tau\} = \begin{cases} 1 - \tau > 0, & \text{for } 0 < \tau < 1, \\ 0, & \text{for } \tau = 1, \\ \frac{\tau(\tau-1)(\tau^2+\tau-1)}{1+\tau-\tau^2} > 0, & \text{for } 1 < \tau < \frac{1+\sqrt{5}}{2}, \end{cases} \tag{3.21}
$$

where $\theta(\tau)$ is defined in (3.5). Then we consider two cases:

- For $\tau \in (0,1) \cup \left(1, \frac{1+\sqrt{5}}{2}\right)$, it follows from (3.19), (3.20), (3.21), and the non-negativity of Ψ and Φ that $\{\|D - \mathcal{A}(Z^k)\|_F\}$, $\{\|\mathcal{B}(L^k) + \mathcal{C}(S^k) - Z^k - \frac{1}{\beta}\Lambda^k\|_F\}$ and $\{\|\mathcal{B}(L^k) + \mathcal{C}(S^k) - Z^k\|_F\}$ are bounded; and moreover,

  $$\Psi(L^k) + \Phi(S^k) < h_0.$$

  The boundedness of $\{L^k\}$ and $\{S^k\}$ follows immediately from this last relation. Furthermore, $\{\Lambda^k\}$ is bounded since

  $$\|\Lambda^k\|_F \;\leqslant\; \beta\|\mathcal{B}(L^k) + \mathcal{C}(S^k) - Z^k - \frac{1}{\beta}\Lambda^k\|_F + \beta\|\mathcal{B}(L^k) + \mathcal{C}(S^k) - Z^k\|_F.$$

Finally, we obtain the boundedness of $\{Z^k\}$ from

$$\begin{aligned}\|Z^k\|_F \;\leqslant\; & \|\mathcal{B}(L^k) + \mathcal{C}(S^k) - Z^k - \frac{1}{\beta}\Lambda^k\|_F + \|\mathcal{B}(L^k)\|_F \\ & + \|\mathcal{C}(S^k)\|_F + \frac{1}{\beta}\|\Lambda^k\|_F.\end{aligned} \tag{3.22}$$

- For $\tau = 1$, it follows from (3.19), (3.20), (3.21), and the nonnegativity of Ψ and Φ that $\{\|D - \mathcal{A}(Z^k)\|_F\}$ and $\{\|\mathcal{B}(L^k) + \mathcal{C}(S^k) - Z^k - \frac{1}{\beta}\Lambda^k\|_F\}$ are bounded; and moreover $\Psi(L^k) + \Phi(S^k) < h_0$, from which we see immediately that $\{L^k\}$ and $\{S^k\}$ are bounded. The boundedness of $\{\Lambda^k\}$ now follows from (3.17) with $\tau = 1$, i.e., $\Lambda^k = \mathcal{A}^*(D - \mathcal{A}(Z^k))$. The boundedness of $\{Z^k\}$ again follows from (3.22).

This completes the proof. □

We are now ready to prove our first global convergence result for Algorithm 1, which also characterizes the cluster point of the sequence generated.

Theorem 3.1 (Global subsequential convergence). *Suppose that* $0 < \tau < \frac{1+\sqrt{5}}{2}$ *and* $\beta > \bar{\beta}$. *If Assumption 3.1 holds, then*

(i) $\lim_{k\to\infty} \|L^{k+1} - L^k\|_F + \|S^{k+1} - S^k\|_F + \|Z^{k+1} - Z^k\|_F + \|\Lambda^{k+1} - \Lambda^k\|_F = 0$;

(ii) *for any cluster point* $(L^*, S^*, Z^*, \Lambda^*)$ *of a sequence* $\{(L^k, S^k, Z^k, \Lambda^k)\}$ *generated by Algorithm 1,* (L^*, S^*) *is a stationary point of* $\mathcal{F}$.

Proof. The boundedness of the sequence $\{(L^k, S^k, Z^k, \Lambda^k)\}$ follows immediately from Proposition 3.1 and thus a cluster point exists. We now prove statement (i).

Suppose that $(L^*, S^*, Z^*, \Lambda^*)$ is a cluster point of the sequence $\{(L^k, S^k, Z^k, \Lambda^k)\}$ and let $\{(L^{k_i}, S^{k_i}, Z^{k_i}, \Lambda^{k_i})\}$ be a convergent subsequence such that

$$\lim_{i\to\infty} (L^{k_i}, S^{k_i}, Z^{k_i}, \Lambda^{k_i}) = (L^*, S^*, Z^*, \Lambda^*).$$

By summing (3.6) from $k = 1$ to $k = k_i - 1$, we have

$$\begin{aligned}
&\Theta_{\tau,\beta}(L^{k_i}, S^{k_i}, Z^{k_i}, \Lambda^{k_i}) - \Theta_{\tau,\beta}(L^1, S^1, Z^1, \Lambda^1) \\
&\leqslant -C \sum_{k=1}^{k_i-1} \|Z^{k+1} - Z^k\|_F^2 - \frac{\sigma\beta}{2} \sum_{k=1}^{k_i-1} \|L^{k+1} - L^k\|_F^2,
\end{aligned} \tag{3.23}$$

where $C := \frac{\lambda_{\min}+\beta}{2} - \max\left\{\frac{1}{\tau}, \frac{\tau^2}{1+\tau-\tau^2}\right\} \cdot \frac{\lambda_{\max}^2}{\beta} > 0$ (since $\beta > \bar{\beta}$). Passing to the limit in (3.23) and rearranging terms in the resulting relation, we obtain

$$\begin{aligned}
&C \sum_{k=1}^{\infty} \|Z^{k+1} - Z^k\|_F^2 + \frac{\sigma\beta}{2} \sum_{k=1}^{\infty} \|L^{k+1} - L^k\|_F^2 \\
&\leqslant \Theta_{\tau,\beta}(L^1, S^1, Z^1, \Lambda^1) - \Theta_{\tau,\beta}(L^*, S^*, Z^*, \Lambda^*) < \infty,
\end{aligned}$$

where the last inequality follows from the properness of Ψ and Φ. This together with $C > 0$ and $\sigma > 0$ implies that

$$\sum_{k=1}^{\infty} \|Z^{k+1} - Z^k\|_F^2 < \infty \quad \text{and} \quad \sum_{k=1}^{\infty} \|L^{k+1} - L^k\|_F^2 < \infty.$$

Hence, we have

$$Z^{k+1} - Z^k \to 0, \; L^{k+1} - L^k \to 0. \tag{3.24}$$

Next, by summing both sides of (3.11) from $k = 1$ to $k = k_i$ and passing to the limit, we have

$$\begin{aligned}
&\sum_{k=1}^{\infty} \|\Lambda^{k+1} - \Lambda^k\|_F^2 \leqslant \max\left\{\frac{1}{\tau}, \frac{\tau^2}{1+\tau-\tau^2}\right\} \cdot \tau\lambda_{\max}^2 \sum_{k=1}^{\infty} \|Z^{k+1} - Z^k\|_F^2 \\
&\quad + \theta(\tau)\tau\beta^2 \left(\|\mathcal{B}(L^1) + \mathcal{C}(S^1) - Z^1\|_F^2 - \liminf_{k\to\infty} \|\mathcal{B}(L^{k+1}) + \mathcal{C}(S^{k+1}) - Z^{k+1}\|_F^2\right),
\end{aligned}$$

from which we conclude that

$$\Lambda^{k+1} - \Lambda^k \to 0. \tag{3.25}$$

Finally, we have $S^{k+1} - S^k \to 0$ from (3.24), (3.25), (3.4d) and (a1) in Assumption 3.1. This proves statement (i).

We next prove statement (ii). From the lower semicontinuity of $\Theta_{\tau,\beta}$ (since Ψ and Φ are lower semicontinuous), we have

$$\begin{aligned}\liminf_{i\to\infty}\ \Theta_{\tau,\beta}(L^{k_i+1}, S^{k_i+1}, Z^{k_i}, \Lambda^{k_i}) \geqslant\ &\Psi(L^*) + \Phi(S^*) + \frac{1}{2}\|D - \mathcal{A}(Z^*)\|_F^2 \\ &- \langle \Lambda^*,\ \mathcal{B}(L^*) + \mathcal{C}(S^*) - Z^*\rangle + (1 + 2\theta(\tau))\frac{\beta}{2}\|\mathcal{B}(L^*) + \mathcal{C}(S^*) - Z^*\|_F^2.\end{aligned} \tag{3.26}$$

On the other hand, from the definition of S^{k_i+1} as a minimizer in (3.3b), we have

$$\begin{aligned}&\Theta_{\tau,\beta}(L^{k_i+1}, S^{k_i+1}, Z^{k_i}, \Lambda^{k_i}) \leqslant \Theta_{\tau,\beta}(L^{k_i+1}, S^*, Z^{k_i}, \Lambda^{k_i}) \\ &\quad + \theta(\tau)\beta\left(\|\mathcal{B}(L^{k_i+1}) + \mathcal{C}(S^{k_i+1}) - Z^{k_i}\|_F^2 - \|\mathcal{B}(L^{k_i+1}) + \mathcal{C}(S^*) - Z^{k_i}\|_F^2\right).\end{aligned}$$

Taking limit in above equality, and invoking statement (i) and (a2) in Assumption 3.1, we see that

$$\begin{aligned}\limsup_{i\to\infty}\ \Theta_{\tau,\beta}(L^{k_i+1}, S^{k_i+1}, Z^{k_i}, \Lambda^{k_i}) \leqslant\ &\Psi(L^*) + \Phi(S^*) + \frac{1}{2}\|D - \mathcal{A}(Z^*)\|_F^2 \\ &- \langle \Lambda^*,\ \mathcal{B}(L^*) + \mathcal{C}(S^*) - Z^*\rangle + (1 + 2\theta(\tau))\frac{\beta}{2}\|\mathcal{B}(L^*) + \mathcal{C}(S^*) - Z^*\|_F^2.\end{aligned} \tag{3.27}$$

Then, combining (3.26) and (3.27), we see that

$$\begin{aligned}\lim_{i\to\infty}\ \Theta_{\tau,\beta}(L^{k_i+1}, S^{k_i+1}, Z^{k_i}, \Lambda^{k_i}) =\ &\Psi(L^*) + \Phi(S^*) + \frac{1}{2}\|D - \mathcal{A}(Z^*)\|_F^2 \\ &- \langle \Lambda^*,\ \mathcal{B}(L^*) + \mathcal{C}(S^*) - Z^*\rangle + (1 + 2\theta(\tau))\frac{\beta}{2}\|\mathcal{B}(L^*) + \mathcal{C}(S^*) - Z^*\|_F^2,\end{aligned}$$

which, together with (a2) in Assumption 3.1, $L^{k+1} - L^k \to 0$, $S^{k+1} - S^k \to 0$ and the definition of $\Theta_{\tau,\beta}$, implies that

$$\lim_{i\to\infty} \Phi(S^{k_i+1}) = \Phi(S^*). \tag{3.28}$$

Thus, passing to the limit in (3.4a)-(3.4d) along $\{(L^{k_i}, S^{k_i}, Z^{k_i}, \Lambda^{k_i})\}$ and invoking statement (i), (3.28) and (2.1), we see that

$$\begin{cases} 0 \in \partial\Psi(L^*) - \mathcal{B}^*(\Lambda^*) + \beta\mathcal{B}^*\left(\mathcal{B}(L^*) + \mathcal{C}(S^*) - Z^*\right), \\ 0 \in \partial\Phi(S^*) - \mathcal{C}^*(\Lambda^*) + \beta\mathcal{C}^*\left(\mathcal{B}(L^*) + \mathcal{C}(S^*) - Z^*\right), \\ 0 = \mathcal{A}^*(\mathcal{A}(Z^*) - D) + \Lambda^* - \beta(\mathcal{B}(L^*) + \mathcal{C}(S^*) - Z^*), \\ \mathcal{B}(L^*) + \mathcal{C}(S^*) = Z^*. \end{cases} \tag{3.29}$$

Rearranging terms in (3.29), it is not hard to obtain

$$\begin{cases} 0 \in \partial\Psi(L^*) + \mathcal{B}^*\mathcal{A}^*\left(\mathcal{A}(\mathcal{B}(L^*) + \mathcal{C}(S^*)) - D\right), \\ 0 \in \partial\Phi(S^*) + \mathcal{C}^*\mathcal{A}^*\left(\mathcal{A}(\mathcal{B}(L^*) + \mathcal{C}(S^*)) - D\right). \end{cases}$$

This shows that (L^*, S^*) is a stationary point of $\mathcal{F}$. This completes the proof. □

Remark 3.2 (**Comments on the computable threshold**). *From the above discussions, we establish under Assumption 3.1 the convergence of the ADMM with* $0 < \tau < \frac{1+\sqrt{5}}{2}$ *when the penalty parameter* β *is chosen above a computable threshold* $\bar{\beta}$ *which depends on* τ*. The existence of this kind of threshold is also obtained in the recent studies [1, 40, 49, 83, 84] on the nonconvex ADMM and its variants with* $\tau = 1$*. In Fig. 3.1, we plot* $\bar{\beta}$ *against* τ *with* $\mathcal{A}$ *being the identity map (hence,* $\lambda_{\max} = \lambda_{\min} = 1$*). It is not hard to see from Fig. 3.1 that for a given penalty parameter* $\beta > 1$*, we can always choose a dual step-size* τ *from an interval containing 1 so that the corresponding ADMM is convergent.*

Remark 3.3 (**Practical computation consideration on penalty parameter**). *In computation, for a* $0 < \tau < \frac{1+\sqrt{5}}{2}$*, the* $\bar{\beta}$ *in* (3.15) *may be too large and hence fixing a* β *close to it can lead to slow convergence. As in [50, 78], one could possibly accelerate the algorithm by initializing the algorithm with a small* β *(less than* $\bar{\beta}$*) and then increasing the* β *by a constant ratio until* $\beta > \bar{\beta}$ *if the sequence generated*

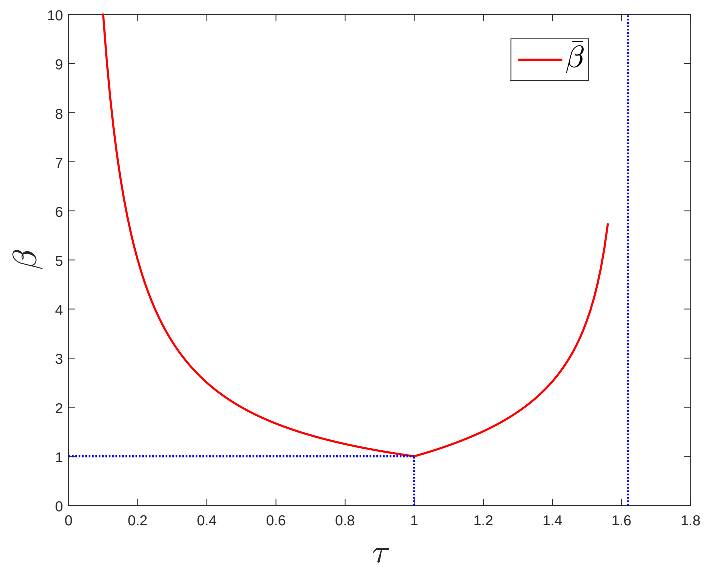


Figure 3.1: The computable threshold $\bar{\beta}$ for $0 < \tau < \frac{1+\sqrt{5}}{2}$.

becomes unbounded or the successive change does not vanish sufficiently fast. Clearly, after at most finitely many increases, the penalty parameter β gets above the threshold $\bar{\beta}$ and the convergence of the resulting algorithm is guaranteed by Theorem 3.1 under Assumption 3.1. On the other hand, if β is never increased, this means that the successive change goes to zero and the sequence is bounded. Then it is routine to show that any cluster point is a stationary point if Φ is continuous in its domain.

Under the additional assumption that the potential function $\Theta_{\tau,\beta}$ is a KL function, we show in the next theorem that the whole sequence generated by Algorithm 1 is convergent if β is greater than a computable threshold, again under Assumption 3.1. Our proof makes use of the uniformized KL property; see Proposition 2.1. This technique was previously used in [8] to prove the convergence of the proximal alternating linearized minimization algorithm for nonconvex and nonsmooth problems, and later in [83, 84] to prove the global convergence of the Bregman ADMM with $\tau = 1$. Our analysis, though follows a similar line of arguments as in [83, 84], is much more intricate. This is because when $\tau \neq 1$, the successive change in the dual variable cannot be controlled solely by the successive changes in the primal variables.

Theorem 3.2 (**Global convergence of the whole sequence**). *Let $0 < \tau < \frac{1+\sqrt{5}}{2}$ and $\beta > \bar{\beta}$. Suppose in addition that Assumption 3.1 holds and the potential function $\Theta_{\tau,\beta}(\cdot)$ is a KL function. Then, the sequence $\{(L^k, S^k, Z^k, \Lambda^k)\}_{k=1}^{\infty}$ generated by Algorithm 1 has a cluster point $(L^*, S^*, Z^*, \Lambda^*)$. Moreover, $\{(L^k, S^k, Z^k, \Lambda^k)\}_{k=1}^{\infty}$ converges to $(L^*, S^*, Z^*, \Lambda^*)$ and (L^*, S^*) is a stationary point of (1.1).*

Proof. In view of Theorem 3.1, we only need to show that the sequence is convergent. We start by noting from (3.19), (3.20) and (3.21) that $\{\Theta_{\tau,\beta}(L^k, S^k, Z^k, \Lambda^k)\}_{k=1}^{\infty}$ is bounded below. Since this sequence is also decreasing from Theorem 3.1, we conclude that $\lim_{k\to\infty} \Theta_{\tau,\beta}(L^k, S^k, Z^k, \Lambda^k) =: \theta^*$ exists. In the following, we will consider two cases.

Case 1) Suppose first that $\Theta_{\tau,\beta}(L^N, S^N, Z^N, \Lambda^N) = \theta^*$ for some $N \geqslant 1$. Since $\{\Theta_{\tau,\beta}(L^k, S^k, Z^k, \Lambda^k)\}_{k=1}^{\infty}$ is decreasing, we must have $\Theta_{\tau,\beta}(L^k, S^k, Z^k, \Lambda^k) = \theta^*$ for all $k \geqslant N$. Then, it follows from (3.6) that $L^{N+t} = L^N$ and $Z^{N+t} = Z^N$ for all $t \geqslant 0$. Hence, $\{L^k\}$ and $\{Z^k\}$ converge finitely. Moreover, from (3.8), we have

$$\|\Lambda^{k+1} - \Lambda^k\|_F = |1-\tau| \cdot \|\Lambda^k - \Lambda^{k-1}\|_F = \cdots = |1-\tau|^{k+1-N} \cdot \|\Lambda^N - \Lambda^{N-1}\|_F$$

for all $k \geqslant N$. Since $0 < \tau < \frac{1+\sqrt{5}}{2}$, we have $0 < 1 - |1-\tau| \leqslant 1$ and hence we see further that

$$\sum_{k=N}^{\infty} \|\Lambda^{k+1} - \Lambda^k\|_F \leqslant \frac{1}{1-|1-\tau|} \|\Lambda^N - \Lambda^{N-1}\|_F < \infty, \tag{3.30}$$

which implies the convergence of $\{\Lambda^k\}$. Additionally, for all $k \geqslant N$, we have

$$\begin{aligned}
\|S^{k+1} - S^k\|_F &\leqslant \frac{1}{\sqrt{\sigma'}} \|\mathcal{C}(S^{k+1}) - \mathcal{C}(S^k)\|_F \\
&= \frac{1}{\sqrt{\sigma'}} \left\| \frac{1}{\tau\beta}(\Lambda^k - \Lambda^{k+1}) - \frac{1}{\tau\beta}(\Lambda^{k-1} - \Lambda^k) \right\|_F \\
&\leqslant \frac{1}{\tau\beta\sqrt{\sigma'}} \|\Lambda^{k+1} - \Lambda^k\|_F + \frac{1}{\tau\beta\sqrt{\sigma'}} \|\Lambda^k - \Lambda^{k-1}\|_F,
\end{aligned}$$

where the first inequality follows from (a1) in Assumption 3.1 and the equality follows from (3.4d). This together with (3.30), implies that $\sum_{k=N}^{\infty} \|S^{k+1} - S^k\|_F < \infty$. Thus, $\{S^k\}$ is also convergent. Consequently, we see that $\{(L^k, S^k, Z^k, \Lambda^k)\}_{k=1}^{\infty}$ is a convergent sequence in this case.

Case 2) From now on, we consider the case where $\Theta_{\tau,\beta}(L^k, S^k, Z^k, \Lambda^k) > \theta^*$ for all $k \geqslant 1$. In this case, we will divide the proof into three steps: **1.** we first prove that $\Theta_{\tau,\beta}$ is constant on the set of cluster points of the sequence $\{(L^k, S^k, Z^k, \Lambda^k)\}_{k=1}^{\infty}$ and then apply the uniformized KL property; **2.** we bound the distance from 0 to $\partial\Theta_{\tau,\beta}(L^k, S^k, Z^k, \Lambda^k)$; **3.** we show that the sequence $\{(L^k, S^k, Z^k, \Lambda^k)\}_{k=1}^{\infty}$ is a Cauchy sequence and hence is convergent. The complete proof is presented as follows.

*Step **1.*** We recall from Proposition 3.1 that the sequence $\{(L^k, S^k, Z^k, \Lambda^k)\}_{k=1}^{\infty}$ generated by Algorithm 1 is bounded and hence must have at least one cluster point. Let Γ denote the set of cluster points of $\{(L^k, S^k, Z^k, \Lambda^k)\}_{k=1}^{\infty}$. We will show that $\Theta_{\tau,\beta}$ is constant on Γ.

To this end, take any $(L^*, S^*, Z^*, \Lambda^*) \in \Gamma$ and consider a convergent subsequence $\{(L^{k_i}, S^{k_i}, Z^{k_i}, \Lambda^{k_i})\}$ with $\lim_{i\to\infty}(L^{k_i}, S^{k_i}, Z^{k_i}, \Lambda^{k_i}) = (L^*, S^*, Z^*, \Lambda^*)$. Then from the lower semicontinuity of $\Theta_{\tau,\beta}$ (since Ψ and Φ are lower semicontinuous) and the definition of θ^*, we have

$$\theta^* = \lim_{i\to\infty} \Theta_{\tau,\beta}(L^{k_i}, S^{k_i}, Z^{k_i}, \Lambda^{k_i}) \geqslant \Theta_{\tau,\beta}(L^*, S^*, Z^*, \Lambda^*). \tag{3.31}$$

On the other hand, notice from the definition of S^{k+1} as a minimizer in (3.3b) that

$$\begin{aligned}
&\Theta_{\tau,\beta}(L^{k_i}, S^{k_i}, Z^{k_i-1}, \Lambda^{k_i-1}) - \Theta_{\tau,\beta}(L^{k_i}, S^{*}, Z^{k_i-1}, \Lambda^{k_i-1}) \\
&= \mathcal{L}_\beta(L^{k_i}, S^{k_i}, Z^{k_i-1}, \Lambda^{k_i-1}) - \mathcal{L}_\beta(L^{k_i}, S^{*}, Z^{k_i-1}, \Lambda^{k_i-1}) \\
&\quad + \theta(\tau)\beta\left(\|\mathcal{B}(L^{k_i}) + \mathcal{C}(S^{k_i}) - Z^{k_i-1}\|_F^2 - \|\mathcal{B}(L^{k_i}) + \mathcal{C}(S^{*}) - Z^{k_i-1}\|_F^2\right) \\
&\leqslant \theta(\tau)\beta\left(\|\mathcal{B}(L^{k_i}) + \mathcal{C}(S^{k_i}) - Z^{k_i-1}\|_F^2 - \|\mathcal{B}(L^{k_i}) + \mathcal{C}(S^{*}) - Z^{k_i-1}\|_F^2\right).
\end{aligned}$$

This together with Theorem 3.1(i), the continuity of $\Theta_{\tau,\beta}$ with respect to L (from (a2) in Assumption 3.1), Z and Λ; and the definition of θ^* implies that

$$\theta^* = \lim_{i\to\infty} \Theta_{\tau,\beta}(L^{k_i}, S^{k_i}, Z^{k_i}, \Lambda^{k_i}) \leqslant \Theta_{\tau,\beta}(L^*, S^*, Z^*, \Lambda^*). \tag{3.32}$$

Combining (3.31) and (3.32), we conclude that $\Theta_{\tau,\beta}(L^*, S^*, Z^*, \Lambda^*) = \theta^*$. Since $(L^*, S^*, Z^*, \Lambda^*) \in \Gamma$ is arbitrary, we conclude further that the potential function $\Theta_{\tau,\beta}$ is constant on Γ.

The fact that $\Theta_{\tau,\beta} \equiv \theta^*$ on Γ together with our assumption that $\Theta_{\tau,\beta}(\cdot)$ is a KL function and Proposition 2.1 implies that there exist $\varepsilon > 0$, $\eta > 0$ and $\varphi \in \Xi_\eta$, such that

$$\varphi'\left(\Theta_{\tau,\beta}(L, S, Z, \Lambda) - \theta^*\right) \operatorname{dist}\left(0,\ \partial\Theta_{\tau,\beta}(L, S, Z, \Lambda)\right) \geqslant 1$$

for all (L, S, Z, Λ) satisfying $\operatorname{dist}((L, S, Z, \Lambda), \Gamma) < \varepsilon$ and $\theta^* < \Theta_{\tau,\beta}(L, S, Z, \Lambda) < \theta^* + \eta$. On the other hand, since $\lim_{k\to\infty} \operatorname{dist}((L^k, S^k, Z^k, \Lambda^k), \Gamma) = 0$ by the definition of Γ, and $\Theta_{\tau,\beta}(L^k, S^k,\ Z^k,\ \Lambda^k) \to \theta^*$, then for such ε and η, there exists $k_1 \geqslant 3$ such that $\operatorname{dist}((L^k, S^k, Z^k, \Lambda^k), \Gamma) < \varepsilon$ and $\theta^* < \Theta_{\tau,\beta}(L^k, S^k, Z^k, \Lambda^k) < \theta^* + \eta$ for all $k \geqslant k_1$. Thus, for $k \geqslant k_1$, we have

$$\varphi'\left(\Theta_{\tau,\beta}(L^k, S^k, Z^k, \Lambda^k) - \theta^*\right) \operatorname{dist}\left(0,\ \partial\Theta_{\tau,\beta}(L^k, S^k, Z^k, \Lambda^k)\right) \geqslant 1. \tag{3.33}$$

*Step **2**.* We next consider the subdifferential $\partial\Theta_{\tau,\beta}(L^k, S^k, Z^k, \Lambda^k)$. Looking at the partial subdifferential with respect to L, we have

$$
\begin{aligned}
&\partial_L \Theta_{\tau,\beta}(L^k, S^k, Z^k, \Lambda^k) \\
&= \partial\Psi(L^k) - \mathcal{B}^*(\Lambda^k) + (1+2\theta(\tau))\beta\mathcal{B}^*(\mathcal{B}(L^k) + \mathcal{C}(S^k) - Z^k) \\
&= \partial\Psi(L^k) - \mathcal{B}^*(\Lambda^{k-1}) + \beta\mathcal{B}^*(\mathcal{B}(L^k) + \mathcal{C}(S^{k-1}) - Z^{k-1}) \\
&\quad + 2\theta(\tau)\beta\mathcal{B}^*(\mathcal{B}(L^k) + \mathcal{C}(S^k) - Z^k) - \mathcal{B}^*(\Lambda^k - \Lambda^{k-1}) \\
&\quad + \beta\mathcal{B}^*(\mathcal{C}(S^k) - Z^k - \mathcal{C}(S^{k-1}) + Z^{k-1}) \\
&\ni 2\theta(\tau)\beta\mathcal{B}^*(\mathcal{B}(L^k) + \mathcal{C}(S^k) - Z^k) - \mathcal{B}^*(\Lambda^k - \Lambda^{k-1}) \\
&\quad + \beta\mathcal{B}^*(\mathcal{C}(S^k) - Z^k - \mathcal{C}(S^{k-1}) + Z^{k-1}) \\
&\overset{\text{(i)}}{=} -\Big(1 + \tfrac{2\theta(\tau)}{\tau}\Big)\mathcal{B}^*(\Lambda^k - \Lambda^{k-1}) + \beta\mathcal{B}^*[(\mathcal{C}(S^k) - Z^k) - (\mathcal{C}(S^{k-1}) - Z^{k-1})] \\
&\overset{\text{(ii)}}{=} -\Big(1 + \tfrac{2\theta(\tau)}{\tau}\Big)\mathcal{B}^*(\Lambda^k - \Lambda^{k-1}) + \beta\mathcal{B}^*\big[\big(-\mathcal{B}(L^k) - \tfrac{\Lambda^k - \Lambda^{k-1}}{\tau\beta}\big) \\
&\quad - \big(-\mathcal{B}(L^{k-1}) - \tfrac{\Lambda^{k-1} - \Lambda^{k-2}}{\tau\beta}\big)\big] \\
&= -\Big(1 + \tfrac{2\theta(\tau)+1}{\tau}\Big)\mathcal{B}^*(\Lambda^k - \Lambda^{k-1}) + \tfrac{1}{\tau}\mathcal{B}^*(\Lambda^{k-1} - \Lambda^{k-2}) - \beta\mathcal{B}^*\mathcal{B}(L^k - L^{k-1}),
\end{aligned}
$$

where the inclusion follows from (3.4a), and the equalities (i) and (ii) follow from (3.4d). Similarly,

$$
\begin{aligned}
&\partial_S \Theta_{\tau,\beta}(L^k, S^k, Z^k, \Lambda^k) \\
&= \partial\Phi(S^k) - \mathcal{C}^*(\Lambda^k) + (1+2\theta(\tau))\beta\mathcal{C}^*(\mathcal{B}(L^k) + \mathcal{C}(S^k) - Z^k) \\
&= \partial\Phi(S^k) - \mathcal{C}^*(\Lambda^{k-1}) + \beta\mathcal{C}^*(\mathcal{B}(L^k) + \mathcal{C}(S^k) - Z^{k-1}) \\
&\quad + 2\theta(\tau)\beta\mathcal{C}^*(\mathcal{B}(L^k) + \mathcal{C}(S^k) - Z^k) - \mathcal{C}^*(\Lambda^k - \Lambda^{k-1}) - \beta\mathcal{C}^*(Z^k - Z^{k-1}) \\
&\ni 2\theta(\tau)\beta\mathcal{C}^*(\mathcal{B}(L^k) + \mathcal{C}(S^k) - Z^k) - \mathcal{C}^*(\Lambda^k - \Lambda^{k-1}) - \beta\mathcal{C}^*(Z^k - Z^{k-1}) \\
&= -\left(1 + \frac{2\theta(\tau)}{\tau}\right)\mathcal{C}^*(\Lambda^k - \Lambda^{k-1}) - \beta\mathcal{C}^*(Z^k - Z^{k-1}),
\end{aligned}
$$

where the inclusion follows from (3.4b) and the last equality follows from (3.4d).

Moreover,

$$
\begin{aligned}
\nabla_Z \Theta_{\tau,\beta}(L^k, S^k, Z^k, \Lambda^k) &= \mathcal{A}^*(\mathcal{A}(Z^k) - D) + \Lambda^k - \beta(\mathcal{B}(L^k) + \mathcal{C}(S^k) - Z^k) \\
&\quad - 2\theta(\tau)\beta(\mathcal{B}(L^k) + \mathcal{C}(S^k) - Z^k) \\
&= \mathcal{A}^*(\mathcal{A}(Z^k) - D) + \Lambda^{k-1} - \beta(\mathcal{B}(L^k) + \mathcal{C}(S^k) - Z^k) \\
&\quad - 2\theta(\tau)\beta(\mathcal{B}(L^k) + \mathcal{C}(S^k) - Z^k) + (\Lambda^k - \Lambda^{k-1}) \\
&= -2\theta(\tau)\beta(\mathcal{B}(L^k) + \mathcal{C}(S^k) - Z^k) + (\Lambda^k - \Lambda^{k-1}) \\
&= \left(1 + \frac{2\theta(\tau)}{\tau}\right)(\Lambda^k - \Lambda^{k-1}),
\end{aligned}
$$

where the third equality follows from (3.4c) and the last equality follows from (3.4d). Finally,

$$
\nabla_\lambda \Theta_{\tau,\beta}(L^k, S^k, Z^k, \Lambda^k) = -(\mathcal{B}(L^k) + \mathcal{C}(S^k) - Z^k) = \frac{1}{\tau\beta}(\Lambda^k - \Lambda^{k-1}),
$$

where the last equality follows from (3.4d). Thus, from the above relations, there exists $a > 0$ so that

$$
\begin{aligned}
&\operatorname{dist}\left(0,\ \partial\Theta_{\tau,\beta}\left(L^k, S^k, Z^k, \Lambda^k\right)\right) \\
&\leqslant a\left(\|L^k - L^{k-1}\|_F + \|Z^k - Z^{k-1}\|_F + \|\Lambda^k - \Lambda^{k-1}\|_F + \|\Lambda^{k-1} - \Lambda^{k-2}\|_F\right).
\end{aligned}
\tag{3.34}
$$

*Step **3.*** We now prove the convergence of the sequence by combining (3.34) with (3.33). For notational simplicity, define

$$
\Delta^k := \varphi\left(\Theta_{\tau,\beta}(L^k, S^k, Z^k, \Lambda^k) - \theta^*\right) - \varphi\left(\Theta_{\tau,\beta}(L^{k+1}, S^{k+1}, Z^{k+1}, \Lambda^{k+1}) - \theta^*\right).
$$

Since $\Theta_{\tau,\beta}$ is decreasing and φ is monotonic, it is easy to see $\Delta^k \geqslant 0$ for $k \geqslant 1$. Then we have for all $k \geqslant k_1$ that

$$
\begin{aligned}
& a\left(\|L^k - L^{k-1}\|_F + \|Z^k - Z^{k-1}\|_F + \|\Lambda^k - \Lambda^{k-1}\|_F + \|\Lambda^{k-1} - \Lambda^{k-2}\|_F\right) \cdot \Delta^k \\
\geqslant\ & \operatorname{dist}(0,\ \partial\Theta_{\tau,\beta}(L^k, S^k, Z^k, \Lambda^k)) \cdot \Delta^k \\
\geqslant\ & \operatorname{dist}(0,\ \partial\Theta_{\tau,\beta}(L^k, S^k, Z^k, \Lambda^k)) \cdot \varphi'\left(\Theta_{\tau,\beta}(L^k, S^k, Z^k, \Lambda^k) - \theta^*\right) \\
& \cdot \left[\Theta_{\tau,\beta}(L^k, S^k, Z^k, \Lambda^k) - \Theta_{\tau,\beta}(L^{k+1}, S^{k+1}, Z^{k+1}, \Lambda^{k+1})\right] \\
\geqslant\ & \Theta_{\tau,\beta}(L^k, S^k, Z^k, \Lambda^k) - \Theta_{\tau,\beta}(L^{k+1}, S^{k+1}, Z^{k+1}, \Lambda^{k+1}) \\
\geqslant\ & b_1\|L^{k+1} - L^k\|_F^2 + b_2\|Z^{k+1} - Z^k\|_F^2 \\
\geqslant\ & \frac{1}{2}\min\{b_1, b_2\} \cdot \left[\|L^{k+1} - L^k\|_F + \|Z^{k+1} - Z^k\|_F\right]^2,
\end{aligned}
\tag{3.35}
$$

where the first inequality follows from (3.34), the second inequality follows from the concavity of φ, the third inequality follows from (3.33), the fourth inequality follows from (3.6) with $b_1 := \frac{\sigma\beta}{2}$ and $b_2 := \frac{\lambda_{\min}+\beta}{2} - \max\left\{\frac{1}{\tau}, \frac{\tau^2}{1+\tau-\tau^2}\right\} \cdot \frac{\lambda_{\max}^2}{\beta}$.

Dividing both sides of (3.35) by $c := \frac{1}{2}\min\{b_1, b_2\}$, taking the square root and using the inequality $\sqrt{uv} \leqslant \frac{u+v}{2}$ for $u, v \geqslant 0$ to further upper bound the left hand side of the resulting inequality, we obtain that

$$
\begin{aligned}
& \tfrac{1}{2\gamma}\left(\|L^k - L^{k-1}\|_F + \|Z^k - Z^{k-1}\|_F + \|\Lambda^k - \Lambda^{k-1}\|_F + \|\Lambda^{k-1} - \Lambda^{k-2}\|_F\right) + \tfrac{\gamma a}{2c}\Delta^k \\
& \geqslant \|L^{k+1} - L^k\|_F + \|Z^{k+1} - Z^k\|_F,
\end{aligned}
\tag{3.36}
$$

where γ is an arbitrary positive constant. On the other hand, it follows from (3.8) that

$$
\begin{aligned}
\|\Lambda^k - \Lambda^{k-1}\|_F &= \|\tau\mathcal{A}^*\mathcal{A}(Z^{k-1} - Z^k) + (1-\tau)(\Lambda^{k-1} - \Lambda^{k-2})\|_F \\
&\leqslant \tau\lambda_{\max}\|Z^k - Z^{k-1}\|_F + |1-\tau| \cdot \|\Lambda^{k-1} - \Lambda^{k-2}\|_F.
\end{aligned}
$$

Adding $-|1-\tau| \cdot \|\Lambda^k - \Lambda^{k-1}\|_F$ to both sides of the above inequality and simplifying the resulting inequality, we obtain that

$$
\begin{aligned}
&\|\Lambda^k - \Lambda^{k-1}\|_F \\
\leqslant\ & \tfrac{\tau\lambda_{\max}}{1-|1-\tau|}\|Z^k - Z^{k-1}\|_F + \tfrac{|1-\tau|}{1-|1-\tau|}\left(\|\Lambda^{k-1} - \Lambda^{k-2}\|_F - \|\Lambda^k - \Lambda^{k-1}\|_F\right) \\
=\ & d_1\|Z^k - Z^{k-1}\|_F + d_2\left(\|\Lambda^{k-1} - \Lambda^{k-2}\|_F - \|\Lambda^k - \Lambda^{k-1}\|_F\right),
\end{aligned}
\tag{3.37}
$$

where we write $d_1 := \frac{\tau\lambda_{\max}}{1-|1-\tau|}$ and $d_2 := \frac{|1-\tau|}{1-|1-\tau|}$ for notational simplicity. Similarly,

$$
\|\Lambda^{k-1}-\Lambda^{k-2}\|_F \leqslant d_1\|Z^{k-1}-Z^{k-2}\|_F + d_2\left(\|\Lambda^{k-2}-\Lambda^{k-3}\|_F - \|\Lambda^{k-1}-\Lambda^{k-2}\|_F\right). \tag{3.38}
$$

Then substituting (3.37) and (3.38) into (3.36) and rearranging terms, we have

$$
\begin{aligned}
&\left(1 - \tfrac{1}{2\gamma}\right)\|L^{k+1} - L^k\|_F + \left(1 - \tfrac{1}{2\gamma} - \tfrac{d_1}{\gamma}\right)\|Z^{k+1} - Z^k\|_F \\
\leqslant\ & \tfrac{1}{2\gamma}\left(\|L^k - L^{k-1}\|_F - \|L^{k+1} - L^k\|_F\right) \\
&+ \left(\tfrac{1}{2\gamma} + \tfrac{d_1}{\gamma}\right)\left(\|Z^k - Z^{k-1}\|_F - \|Z^{k+1} - Z^k\|_F\right) \\
&+ \tfrac{d_1}{2\gamma}\left(\|Z^{k-1} - Z^{k-2}\|_F - \|Z^k - Z^{k-1}\|_F\right) \\
&+ \tfrac{d_2}{2\gamma}\left(\|\Lambda^{k-1} - \Lambda^{k-2}\|_F - \|\Lambda^k - \Lambda^{k-1}\|_F\right) \\
&+ \tfrac{d_2}{2\gamma}\left(\|\Lambda^{k-2} - \Lambda^{k-3}\|_F - \|\Lambda^{k-1} - \Lambda^{k-2}\|_F\right) + \tfrac{\gamma a}{2c}\Delta^k.
\end{aligned}
\tag{3.39}
$$

Thus, summing (3.39) from $k = k_1$ to ∞, we have

$$
\begin{aligned}
&\left(1 - \tfrac{1}{2\gamma}\right)\textstyle\sum_{k=k_1}^{\infty}\|L^{k+1} - L^k\|_F + \left(1 - \tfrac{1}{2\gamma} - \tfrac{d_1}{\gamma}\right)\textstyle\sum_{k=k_1}^{\infty}\|Z^{k+1} - Z^k\|_F \\
\leqslant\ & \tfrac{1}{2\gamma}\|L^{k_1} - L^{k_1-1}\|_F + \left(\tfrac{1}{2\gamma} + \tfrac{d_1}{\gamma}\right)\|Z^{k_1} - Z^{k_1-1}\|_F + \tfrac{d_1}{2\gamma}\|Z^{k_1-1} - Z^{k_1-2}\|_F \\
&+ \tfrac{d_2}{2\gamma}\|\Lambda^{k_1-1} - \Lambda^{k_1-2}\|_F + \tfrac{d_2}{2\gamma}\|\Lambda^{k_1-2} - \Lambda^{k_1-3}\|_F \\
&+ \tfrac{a\gamma}{2c}\,\varphi\left(\Theta_{\tau,\beta}(L^{k_1}, S^{k_1}, Z^{k_1}, \Lambda^{k_1}) - \theta^*\right) < \infty.
\end{aligned}
$$

Recall that γ introduced in (3.36) is an arbitrary positive constant. Taking $\gamma > \frac{1+2d_1}{2}$

and hence $1 - \frac{1}{2\gamma} > 1 - \frac{1}{2\gamma} - \frac{d_1}{\gamma} > 0$, we have from the above inequality that

$$\sum_{k=k_1}^{\infty} \|L^{k+1} - L^k\|_F < \infty \quad \text{and} \quad \sum_{k=k_1}^{\infty} \|Z^{k+1} - Z^k\|_F < \infty.$$

Hence $\{L^k\}$ and $\{Z^k\}$ are convergent. Additionally, summing (3.37) from $k = k_1$ to ∞, we have

$$\sum_{k=k_1}^{\infty} \|\Lambda^k - \Lambda^{k-1}\|_F \leqslant d_1 \sum_{k=k_1}^{\infty} \|Z^k - Z^{k-1}\|_F + d_2\|\Lambda^{k_1-1} - \Lambda^{k_1-2}\|_F < \infty,$$

which implies that $\{\Lambda^k\}$ is convergent. Finally, from (3.4d) and (a1) in Assumption 3.1, we see that $\{S^k\}$ is also convergent. Consequently, we conclude that $\{(L^k, S^k, Z^k, \Lambda^k)\}_{k=1}^{\infty}$ is a convergent sequence. This completes the proof. □

Our convergence analysis relies on Assumption 3.1. While (a3) in Assumption 3.1 appears restrictive since it makes assumptions on the first iterate of Algorithm 1, we show below that this assumption would hold upon a suitable choice of initialization. Specifically, if we initialize at $(L^0, S^0, Z^0, \Lambda^0)$ satisfying

$$\begin{cases} \Theta_{\tau,\beta}(L^1, S^1, Z^1, \Lambda^1) \leqslant \Theta_{\tau,\beta}(L^0, S^0, Z^0, \Lambda^0), & \text{(3.40a)} \\ \Theta_{\tau,\beta}(L^0, S^0, Z^0, \Lambda^0) < h_0, & \text{(3.40b)} \end{cases}$$

then it is easy to check that (a3) in Assumption 3.1 holds. In the next proposition, we demonstrate that (3.40a) can always be satisfied with a suitable initialization. After this, we will propose a specific way to initialize Algorithm 1 for a wide range of problems so that both (3.40a) and (3.40b) are satisfied.

Proposition 3.2. *Suppose that* $0 < \tau < \frac{1+\sqrt{5}}{2}$ *and* $\beta > \bar{\beta}$. *If the initialization* $(L^0, S^0, Z^0, \Lambda^0)$ *is chosen as* $(L^0, S^0) \in \operatorname{dom}\Psi \times \operatorname{dom}\Phi$ *and*

$$\Lambda^0 = \mathcal{A}^*(D - \mathcal{A}(Z^0)), \tag{3.41}$$

then we have

$$\Theta_{\tau,\beta}(L^1, S^1, Z^1, \Lambda^1) \leqslant \Theta_{\tau,\beta}(L^0, S^0, Z^0, \Lambda^0).$$

Proof. First, from (3.4c), we have

$$\begin{aligned} & 0 = \mathcal{A}^*(\mathcal{A}(Z^1) - D) + \Lambda^0 - \beta(\mathcal{B}(L^1) + \mathcal{C}(S^1) - Z^1) \\ \Longrightarrow\ & \mathcal{B}(L^1) + \mathcal{C}(S^1) - Z^1 = \frac{1}{\beta}\Lambda^0 + \frac{1}{\beta}\mathcal{A}^*(\mathcal{A}(Z^1) - D) = \frac{1}{\beta}\mathcal{A}^*\mathcal{A}(Z^1 - Z^0), \end{aligned} \tag{3.42}$$

where the last equality follows from (3.41). Then,

$$\begin{aligned} & \Theta_{\tau,\beta}(L^1, S^1, Z^1, \Lambda^1) - \Theta_{\tau,\beta}(L^1, S^1, Z^1, \Lambda^0) \\ = \ & -\langle \Lambda^1 - \Lambda^0, \mathcal{B}(L^1) + \mathcal{C}(S^1) - Z^1 \rangle = \tau\beta\|\mathcal{B}(L^1) + \mathcal{C}(S^1) - Z^1\|_F^2 \\ = \ & (\tau + \theta(\tau))\,\beta\|\mathcal{B}(L^1) + \mathcal{C}(S^1) - Z^1\|_F^2 - \theta(\tau)\beta\|\mathcal{B}(L^1) + \mathcal{C}(S^1) - Z^1\|_F^2 \\ = \ & (\tau + \theta(\tau))\,\beta\left\|\tfrac{1}{\beta}\mathcal{A}^*\mathcal{A}(Z^1 - Z^0)\right\|_F^2 - \theta(\tau)\beta\|\mathcal{B}(L^1) + \mathcal{C}(S^1) - Z^1\|_F^2 \\ \leqslant \ & (\tau + \theta(\tau))\,\frac{\lambda_{\max}^2}{\beta}\|Z^1 - Z^0\|_F^2 - \theta(\tau)\beta\|\mathcal{B}(L^1) + \mathcal{C}(S^1) - Z^1\|_F^2, \end{aligned} \tag{3.43}$$

where the second equality follows from (3.4d) and the fourth equality follows from (3.42). Additionally, using the same arguments as in the proof of Lemma 3.1 leading to (3.12), (3.13) and (3.14), it is easy to see that

$$\begin{aligned} & \Theta_{\tau,\beta}(L^1, S^1, Z^1, \Lambda^0) - \Theta_{\tau,\beta}(L^1, S^1, Z^0, \Lambda^0) \leqslant -\tfrac{\lambda_{\min}+\beta}{2}\|Z^1 - Z^0\|_F^2 \\ & \qquad + \theta(\tau)\beta\left(\|\mathcal{B}(L^1) + \mathcal{C}(S^1) - Z^1\|_F^2 - \|\mathcal{B}(L^1) + \mathcal{C}(S^1) - Z^0\|_F^2\right), \end{aligned} \tag{3.44}$$

$$\begin{aligned} & \Theta_{\tau,\beta}(L^1, S^1, Z^0, \Lambda^0) - \Theta_{\tau,\beta}(L^1, S^0, Z^0, \Lambda^0) \\ & \qquad \leqslant \theta(\tau)\beta\left(\|\mathcal{B}(L^1) + \mathcal{C}(S^1) - Z^0\|_F^2 - \|\mathcal{B}(L^1) + \mathcal{C}(S^0) - Z^0\|_F^2\right), \end{aligned} \tag{3.45}$$

$$\begin{aligned} & \Theta_{\tau,\beta}(L^1, S^0, Z^0, \Lambda^0) - \Theta_{\tau,\beta}(L^0, S^0, Z^0, \Lambda^0) \\ & \qquad \leqslant \theta(\tau)\beta\left(\|\mathcal{B}(L^1) + \mathcal{C}(S^0) - Z^0\|_F^2 - \|\mathcal{B}(L^0) + \mathcal{C}(S^0) - Z^0\|_F^2\right). \end{aligned} \tag{3.46}$$

Summing (3.43), (3.44), (3.45) and (3.46), we obtain

$$\begin{aligned}&\Theta_{\tau,\beta}(L^1, S^1, Z^1, \Lambda^1) - \Theta_{\tau,\beta}(L^0, S^0, Z^0, \Lambda^0) \\ &\leqslant \left((\tau + \theta(\tau))\,\tfrac{\lambda_{\max}^2}{\beta} - \tfrac{\lambda_{\min}+\beta}{2}\right)\|Z^1 - Z^0\|_F^2 - \theta(\tau)\beta\|\mathcal{B}(L^0) + \mathcal{C}(S^0) - Z^0\|_F^2.\end{aligned} \tag{3.47}$$

We now consider two cases:

- For $0 < \tau \leqslant 1$, it is easy to see $\theta(\tau) = 1 - \tau$ and

  $$\beta > \max\left\{\tfrac{\lambda_{\max}}{\tau},\quad -\tfrac{\lambda_{\min}}{2} + \tfrac{1}{2}\sqrt{\lambda_{\min}^2 + \tfrac{8}{\tau}\lambda_{\max}^2}\right\}.$$

  Then, we have

  $$(\tau + \theta(\tau))\tfrac{\lambda_{\max}^2}{\beta} - \tfrac{\lambda_{\min}+\beta}{2} = \tfrac{\lambda_{\max}^2}{\beta} - \tfrac{\lambda_{\min}+\beta}{2} \leqslant \tfrac{\lambda_{\max}^2}{\tau\beta} - \tfrac{\lambda_{\min}+\beta}{2} < 0.$$

- For $1 < \tau < \frac{1+\sqrt{5}}{2}$, it is easy to see $\theta(\tau) = \frac{(\tau-1)\tau^2}{1+\tau-\tau^2}$ and

  $$\beta > \max\left\{\tau\lambda_{\max},\quad -\tfrac{\lambda_{\min}}{2} + \tfrac{1}{2}\sqrt{\lambda_{\min}^2 + \tfrac{8\tau^2}{1+\tau-\tau^2}\lambda_{\max}^2}\right\}.$$

  Then, we have

  $$(\tau + \theta(\tau))\tfrac{\lambda_{\max}^2}{\beta} - \tfrac{\lambda_{\min}+\beta}{2} = \tfrac{\tau\lambda_{\max}^2}{(1+\tau-\tau^2)\beta} - \tfrac{\lambda_{\min}+\beta}{2} < \tfrac{\tau^2\lambda_{\max}^2}{(1+\tau-\tau^2)\beta} - \tfrac{\lambda_{\min}+\beta}{2} < 0.$$

Thus, combining the above with (3.47) and $\theta(\tau) \geqslant 0$, we conclude that

$$\Theta_{\tau,\beta}(L^1, S^1, Z^1, \Lambda^1) \leqslant \Theta_{\tau,\beta}(L^0, S^0, Z^0, \Lambda^0).$$

This completes the proof. □

From Proposition 3.2, we see that if the initialization $(L^0, S^0, Z^0, \Lambda^0)$ is chosen to satisfy the conditions in Proposition 3.2, then (3.40a) holds. Based on this, we can now present one specific way to initialize Algorithm 1 so that both (3.40a) and

(3.40b) are satisfied for a class of problems, whose objective functions $\Psi(L)$ and $\Phi(S)$ take forms $\delta_\Omega(L)$ and (1.2), respectively; here, Ω is a compact convex set.

The initialization we consider is:

$$L^0 = \mathcal{P}_\Omega(\kappa D),\ S^0 = 0,\ Z^0 = \mathcal{B}(L^0),\ \Lambda^0 = \mathcal{A}^*\left(D - \mathcal{A}(Z^0)\right), \tag{3.48}$$

where κ is a scaling parameter. One can easily check that this initialization satisfies (3.41). Moreover,

$$\Theta_{\tau,\beta}(L^0, S^0, Z^0, \Lambda^0) = \frac{1}{2}\left\|D - \mathcal{A}\left(Z^0\right)\right\|_F^2 = \frac{1}{2}\left\|D - \mathcal{A}\left(\mathcal{B}(\mathcal{P}_\Omega(\kappa D))\right)\right\|_F^2.$$

Thus, the condition (3.40b) is equivalent to

$$\frac{1}{2}\left\|D - \mathcal{A}\left(\mathcal{B}(\mathcal{P}_\Omega(\kappa D))\right)\right\|_F^2 < \liminf_{\|L\|_F + \|S\|_F \to \infty} \Psi(L) + \Phi(S) = \liminf_{\|S\|_F \to \infty} \Phi(S). \tag{3.49}$$

We further discuss this inequality for some concrete examples of Φ presented in the introduction.

Example 3.1. *Suppose that Φ is coercive. Then $\liminf_{\|S\|_F \to \infty} \Phi(S) = \infty$ and hence (3.49) holds trivially for any choice of κ.*

Example 3.2. *Suppose that $\Phi(S) = \mu \sum_{i=1}^m \sum_{j=1}^n \frac{\alpha|s_{ij}|}{1+\alpha|s_{ij}|}$ for $\alpha > 0$. Then $\liminf_{\|S\|_F \to \infty} \Phi(S) = \mu$. Hence, (3.49) holds if the parameter κ can be chosen so that $\frac{1}{2}\left\|D - \mathcal{A}\left(\mathcal{B}(\mathcal{P}_\Omega(\kappa D))\right)\right\|_F^2 < \mu$.*

Example 3.3. *Suppose that $\Phi(S) = \mu \sum_{i=1}^m \sum_{j=1}^n \int_0^{|s_{ij}|} \min(1, (\alpha - t/\mu)_+/(\alpha - 1))\ \mathrm{d}t$ for $\alpha > 2$. Then $\liminf_{\|S\|_F \to \infty} \Phi(S) = \frac{1}{2}(\alpha+1)\mu^2$. Hence, (3.49) holds if κ can be chosen so that $\frac{1}{2}\left\|D - \mathcal{A}\left(\mathcal{B}(\mathcal{P}_\Omega(\kappa D))\right)\right\|_F^2 < \frac{1}{2}(\alpha+1)\mu^2$.*

Example 3.4. *Suppose that $\Phi(S) = \mu \sum_{i=1}^m \sum_{j=1}^n \int_0^{|s_{ij}|} (1 - t/(\alpha\mu))_+\ \mathrm{d}t$ for $\alpha > 0$. Then, $\liminf_{\|S\|_F \to \infty} \Phi(S) = \frac{1}{2}\alpha\mu^2$. Hence, (3.49) holds if κ can be chosen so that $\frac{1}{2}\left\|D - \mathcal{A}\left(\mathcal{B}(\mathcal{P}_\Omega(\kappa D))\right)\right\|_F^2 < \frac{1}{2}\alpha\mu^2$.*

Example 3.5. *Suppose that* $\Phi(S) = \mu \sum_{i=1}^{m} \sum_{j=1}^{n} \mu - (\mu - |s_{ij}|)_+^2/\mu$. *Then it is not hard to show that* $\liminf\limits_{\|S\|_F \to \infty} \Phi(S) = \mu^2$. *Hence, (3.49) holds if* κ *can be chosen so that* $\frac{1}{2}\|D - \mathcal{A}(\mathcal{B}(\mathcal{P}_\Omega(\kappa D)))\|_F^2 < \mu^2$.

3.3 Numerical experiments

In this section, we conduct numerical experiments to show the performances of our algorithm. All experiments are run in MATLAB R2015b on a 64-bit PC with an Intel Core i7-4790 CPU (3.60 GHz) and 32 GB of RAM equipped with Windows 10 OS.

3.3.1 Implementation details

Testing model We consider the problem of extracting background/foreground from a given video under different scenarios. Specifically, we consider:

$$\begin{array}{ll} \min\limits_{L,S} & \Phi(S) + \frac{1}{2}\|D - \mathcal{A}(L+S)\|_F^2 \\ \text{s.t.} & L \in \Omega, \end{array} \tag{3.50}$$

where $\Omega = \{L \in \mathbb{R}^{m\times n} \mid \|L\|_\infty \leqslant 1,\ L_{:1} = L_{:2} = \cdots = L_{:n}\}$ and $\mathcal{A}$ is a linear map. This model corresponds to (1.1) with $\Psi(L) = \delta_\Omega(L)$ and $\mathcal{B} = \mathcal{C} = \mathcal{I}$. We compare the performances of the ADMM with different choices of τ, as well as the proximal alternating linearized minimization (PALM) proposed in [8], on solving (3.50). For ease of future reference, we recall that the PALM for solving (3.50) is given by

$$\begin{cases} L^{k+1} = \mathcal{P}_\Omega\left(L^k - \dfrac{1}{c_k}\mathcal{A}^*(\mathcal{A}(L^k + S^k) - D)\right), \\ S^{k+1} \in \operatorname*{Argmin}\limits_{S}\left\{\Phi(S) + \dfrac{d_k}{2}\left\|S - S^k + \dfrac{1}{d_k}\mathcal{A}^*(\mathcal{A}(L^{k+1} + S^k) - D)\right\|_F^2\right\}, \end{cases}$$

where c_k and d_k are positive numbers.

In our experiments, we consider the following three choices of the sparse regularizer $\Phi(S)$:

- bridge regularizer: $\Phi(S) = \mu\|S\|_p^p$ for $0 < p \leqslant 1$;
- fraction regularizer: $\Phi(S) = \mu\sum_{i=1}^{m}\sum_{j=1}^{n}\frac{\alpha|s_{ij}|}{1+\alpha|s_{ij}|}$ for $\alpha > 0$;
- logistic regularizer: $\Phi(S) = \mu\sum_{i=1}^{m}\sum_{j=1}^{n}\log(1+\alpha|s_{ij}|)$ for $\alpha > 0$;

and two choices of the linear map $\mathcal{A}$:

- $\mathcal{A}(L+S) := L+S$: in this case, model (3.50) can be applied to extracting background/foreground from a surveillance video with noise.
- $\mathcal{A}(L+S) := H(L+S)$ with $H \in \mathbb{R}^{m\times m}$ being the matrix representation of a regular blurring operator (the blurring is assumed to occur frame-wise): in this case, model (3.50) can be applied to extracting background/foreground from a blurred and noisy surveillance video.

Testing videos We choose four real videos, “Hall”, “Bootstrap”, “Fountain” and “ShoppingMall”, from the dataset I2R[1] provided by Li et al. [51]. The details of these videos are as follows:

- **Hall** video contains 200 144×176 frames (from airport2001 to airport2200);
- **Bootstrap** video contains 200 120×160 frames (from b01801 to b02000);
- **Fountain** video contains 200 128×160 frames (from Fountain1301 to Fountain1500);
- **ShoppingMall** video contains 200 256×320 frames (from ShoppingMall1501 to ShoppingMall1700).

[1] This dataset is available in `http://perception.i2r.a-star.edu.sg/bk_model/bk_index.html`. The authors also provide 20 ground-truth images of foregrounds for each video in this dataset.

We show one frame of each testing video under two different scenarios (noisy and noisy blurred), and their ground-truth images of foregrounds in Fig. 3.2. Additionally, all pixel values of the testing videos are re-scaled into $[0, 1]$ in our numerical experiments.

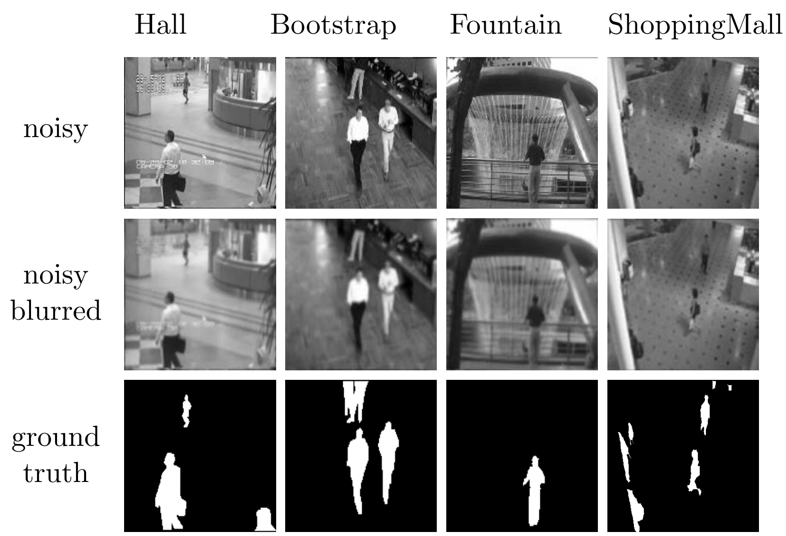


Figure 3.2: One frame (from left to right: airport2180, b01842, Fountain1440 and ShoppingMall1535) of each testing video under different scenarios (the first two rows) and the ground-truth image of foreground of each testing video (the last row).

Parameters setting For the ADMM, we use the following heuristics[2] to update β: we initialize $n_s = 0$ and $\beta = 0.6\bar{\beta}$, where $\bar{\beta}$ is given in (3.15). In the k-th iteration, we compute

$$\begin{aligned} fnorm^k &= \|L^k\|_F + \|Z^k\|_F, \\ succ\_chg^k &= \|L^k - L^{k-1}\|_F + \|Z^k - Z^{k-1}\|_F. \end{aligned}$$

Then, we increase n_s by 1 if $succ\_chg^k > 0.99 \cdot succ\_chg^{k-1}$. Obviously, n_s is non-decreasing in this procedure. We then update β as 1.1β whenever $\beta \leqslant 1.01\bar{\beta}$ and

[2] Note from Theorem 3.1(i) that the successive change of each variable goes to zero as $k \to \infty$. Thus, intuitively, it is more favorable to see a decrease in the successive change as k increases. This heuristic is designed based on this intuition.

the sequence satisfies either $n_s \geqslant 0.3k$ or $fnorm^k > 10^{10}$. On the other hand, for PALM, we set $c_k = d_k = \frac{\lambda_{\max}}{0.99}$.

We initialize our algorithm and the PALM at the point specified in (3.48) with $\kappa = 1$. Moreover, we terminate our ADMM by the following two-stage criterion[3]: in each iteration, we check if

$$\frac{\|L^k - L^{k-1}\|_F + \|Z^k - Z^{k-1}\|_F}{\|L^k\|_F + \|Z^k\|_F + 1} < \mathrm{Tol}_{A,1}$$

for some $\mathrm{Tol}_{A,1} > 0$; if it holds, then we further check if

$$\frac{\|S^k - S^{k-1}\|_F + \|\Lambda^k - \Lambda^{k-1}\|_F}{\|S^k\|_F + \|\Lambda^k\|_F + 1} < \mathrm{Tol}_{A,2}$$

for some $\mathrm{Tol}_{A,2} > 0$. We terminate the algorithm if this latter condition is also satisfied. For the PALM, we terminate it when

$$\frac{\|L^k - L^{k-1}\|_F + \|S^k - S^{k-1}\|_F}{\|L^k\|_F + \|S^k\|_F + 1} < \mathrm{Tol}_P$$

for some $\mathrm{Tol}_P > 0$. The specific values of $\mathrm{Tol}_{A,1}$, $\mathrm{Tol}_{A,2}$ and Tol_P are given in the following experiments.

3.3.2 Comparisons between ADMM with different τ and PALM

In this subsection, we use the performance profile to evaluate the performances of the ADMM with different τ and the PALM for extraction under different scenarios. The performance profile is proposed by Dolan and Moré [21] as a tool for evaluating

[3] We use this two-stage criterion rather than computing the relative errors of all four variables (L, S, Z, Λ) in each iteration of our algorithm because computing matrix Frobenius norms can be expensive, especially for large scale problems. This strategy will help reduce the cost per iteration. We examine $\|L^k - L^{k-1}\|_F$ and $\|Z^k - Z^{k-1}\|_F$ in the first stage because these quantities being small intuitively implies that $\|S^k - S^{k-1}\|_F$ and $\|\Lambda^k - \Lambda^{k-1}\|_F$ are small; see the proof of Theorem 3.1, particularly (3.24), (3.25) and the discussions that follow.

and comparing the performance of a collection of solvers $\mathcal{K}$ on a set of test problems $\mathcal{J}$.

To describe this method, we assume that we have K solvers and J problems, and we use the iteration number as a performance measure. Then, for each problem j and solver k, we set

$$\mathrm{iter}_{j,k} = \text{the iteration number required to solve problem } j \text{ by solver } k.$$

and compute the performance ratio

$$r_{j,k} = \frac{\mathrm{iter}_{j,k}}{\min\{\mathrm{iter}_{j,k} : k \in \mathcal{K}\}}. \tag{3.51}$$

The performance profile of iteration numbers is then defined as the distribution function for the performance ratio, i.e.,

$$\rho_k(\nu) = \frac{1}{J} \sharp\{j \in \mathcal{J} : r_{j,k} \leqslant \nu\}$$

for $\nu \geqslant 1$. Similarly, the performance profile of function values is obtained by using $\mathrm{fval}_{j,k}$ in place of $\mathrm{iter}_{j,k}$ in (3.51), where $\mathrm{fval}_{j,k}$ denotes the function value at the solution given by solver k for solving problem j. Generally speaking, for solver $k \in \mathcal{K}$, the higher $\rho_k(\nu)$ indicates a better performance within the factor ν.

In our experiments, we evaluate the following solvers: the ADMM with $\tau = 0.8$, the ADMM with $\tau = 1$, the ADMM with $\tau = 1.6$ and the PALM.

For $\mathcal{A}(L + S) = L + S$, our test problems are described in Table 3.1, where we use the four real videos introduced above as our input data in (3.50), with 3 choices of sparse regularizers, 10 choices of μ, and 6 choices of p and α. Thus, we have 4 solvers and a total of 720 test problems, with 240 test problems for each sparse regularizer. Moreover, we set $\mathrm{Tol}_{A,1} = 10^{-4}$, $\mathrm{Tol}_{A,2} = 5 \times 10^{-3}$ and $\mathrm{Tol}_P = 10^{-4}$. Fig. 3.3 shows the performance profiles of iteration numbers and function values for different regularizers under this scenario.

Table 3.1: Problem setting for $\mathcal{A}(L+S) = L+S$

data	μ	regularizers
4 real videos	5e-1, 1e-1, 5e-2, 1e-2, 5e-3 1e-3, 5e-4, 1e-4, 5e-5, 1e-5	bridge: $p = 0.2, 0.4, 0.5, 0.6, 0.8, 1$ fraction/logistic: $\alpha = 0.01, 0.1, 1, 2, 5, 10$

For $\mathcal{A}(L+S) = H(L+S)$, our test problems are described in Table 3.2, where we use 2 choices of p and α. Thus, we have 4 solvers and a total of 240 test problems, with 80 test problems for each sparse regularizer. In our experiments, we use the method described in [36] to generate the blurring matrix H, which can be represented as a Kronecker product $H = H_r \otimes H_c$ under the periodic boundary condition. The matlab codes[4] that generate H_r and H_c are shown below, where "frame_size" is the size of each frame:

```
[P, center] = psfGauss(frame_size, 1);
[Hr, Hc] = kronDecomp(P, center, 'periodic');
```

Moreover, we set $\mathrm{Tol}_{A,1} = 5 \times 10^{-3}$, $\mathrm{Tol}_{A,2} = 10^{-2}$ and $\mathrm{Tol}_P = 3 \times 10^{-3}$. Fig. 3.4 shows the performance profiles under this scenario.

Table 3.2: Problem setting for $\mathcal{A}(L+S) = H(L+S)$

data	μ	regularizers
4 real videos	5e-1, 1e-1, 5e-2, 1e-2, 5e-3 1e-3, 5e-4, 1e-4, 5e-5, 1e-5	bridge: $p = 0.5, 1$ fraction/logistic: $\alpha = 1, 2$

It is not hard to see from Fig. 3.3 and Fig. 3.4 that the performance profiles of iteration numbers for the ADMM with $\tau = 0.8$ and $\tau = 1$ usually lie above those for the PALM; and their performance profiles of function values are almost the same. This shows that the ADMM with $\tau = 0.8$ or $\tau = 1$ takes less iterations

[4] The codes are available at http://www.imm.dtu.dk/~pcha/HNO/ as a supplement to the book [36].

for solving all the test problems while giving comparable function values. For bridge regularizer in the case where $\mathcal{A}(L+S) = L+S$ (see Fig. 3.3(a)) and in the case where $\mathcal{A}(L+S) = H(L+S)$ (see Fig. 3.4(a)), we can see that the ADMM with $\tau = 0.8$ sightly outperforms the ADMM with $\tau = 1$ in terms of the number of iterations. For other regularizers, their performances are comparable. Additionally, for the ADMM with $\tau = 1.6$, we can see from Fig. 3.3 and Fig. 3.4 that it always terminates with the worst function value, although it is always fastest in the case where $\mathcal{A}(L+S) = H(L+S)$ (see Fig. 3.4).

To better visualize the performance of the algorithms in terms of function values, we also plot $\mathrm{RelErr}^k := |\mathcal{F}(L^k, S^k) - \mathcal{F}_{\min}|/\mathcal{F}_{\min}$ against the number of iterations for each algorithm, where $\mathcal{F}(L^k, S^k)$ denotes the objective value obtained by each algorithm at (L^k, S^k) and $\mathcal{F}_{\min}$ denotes the minimum of the objective values obtained from all algorithms. We only consider the ADMM with $\tau = 0.8$, the ADMM with $\tau = 1$ and the PALM, and terminate them only after *at least* 500 iterations *and* the termination criteria are satisfied with $\mathrm{Tol}_{A,1} = 10^{-5}$, $\mathrm{Tol}_{A,2} = 5 \times 10^{-4}$ and $\mathrm{Tol}_P = 10^{-5}$. For brevity, we focus on the scenario $\mathcal{A}(L+S) = L+S$ and use the "Hall" video. The results are presented in Fig. 3.5, from which we can see that the ADMM with $\tau = 1$ or $\tau = 0.8$ performs better than PALM for those particular instances.

3.3.3 Simulation results

In this subsection, we present some simulation results for the background/foreground extraction problem. In order to evaluate the performance in background/foreground extraction, we compare the support of the recovered foreground S^* with the support of the ground-truth $\tilde{S}$ by computing the following measurement:

$$\text{F-measure} := 2 \times \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}},$$

where precision and recall are defined as

$$\text{precision} := \frac{\text{TP}}{\text{TP} + \text{FP}}, \quad \text{recall} := \frac{\text{TP}}{\text{TP} + \text{FN}},$$

in which,

- TP stands for true positives: the number of true foreground pixels that are recovered;
- FP stands for false positives: the number of background pixels that are misdetected as foreground;
- FN stands for false negatives: the number of true foreground pixels that are missed.

The support of the recovered foreground S^* is obtained by thresholding S^* entry-wise with a threshold value (we use 1e-3 in our numerical experiments). We would like to point out that F-measure varies between 0 and 1 according to the similarity of the support of S^* and $\tilde{S}$. The higher the F-measure value, the better the recovery accuracy of the support of $\tilde{S}$. The F-measure approaches the maximum value 1 if the supports of S^* and $\tilde{S}$ are the same, which means the foreground is recovered completely.

In our experiments below, we choose $\tau = 0.8$ for the ADMM. We also use the aforementioned four real videos as input with 3 choices of sparse regularizers and 2 choices of p and α. For each fixed p and α, we experiment with different regularization parameters μ (5e-1, 1e-1, 5e-2, 1e-2, 5e-3, 1e-3, 5e-4, 1e-4, 5e-5, 1e-5) and present only the μ corresponding to the maximal F-measure.[5]

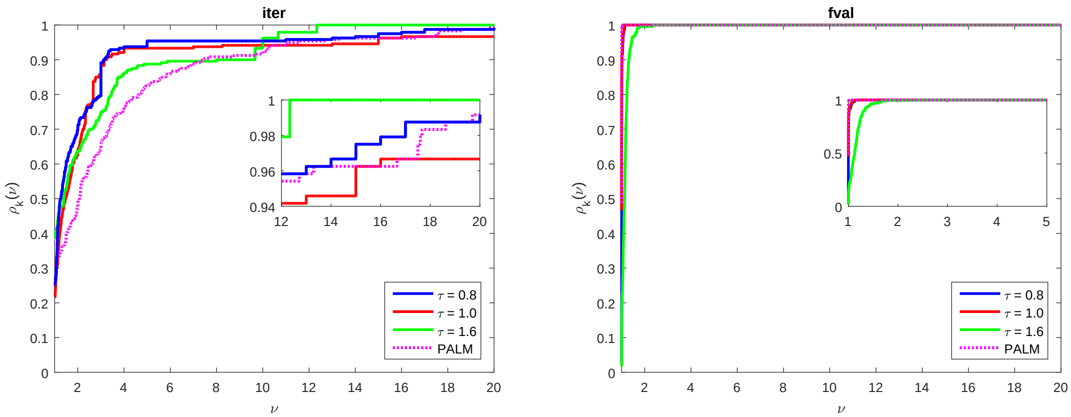
Extraction from noisy surveillance videos In this case, $\mathcal{A}(L + S) = L + S$, $\lambda_{\max} = \lambda_{\min} = 1$ and we set $\text{Tol}_{A,1} = 10^{-4}$, $\text{Tol}_{A,2} = 5 \times 10^{-3}$ and $\text{Tol}_P = 10^{-4}$. The

[5] If the F-measures are the same, we pick the μ that corresponds to the minimal number of iterations.

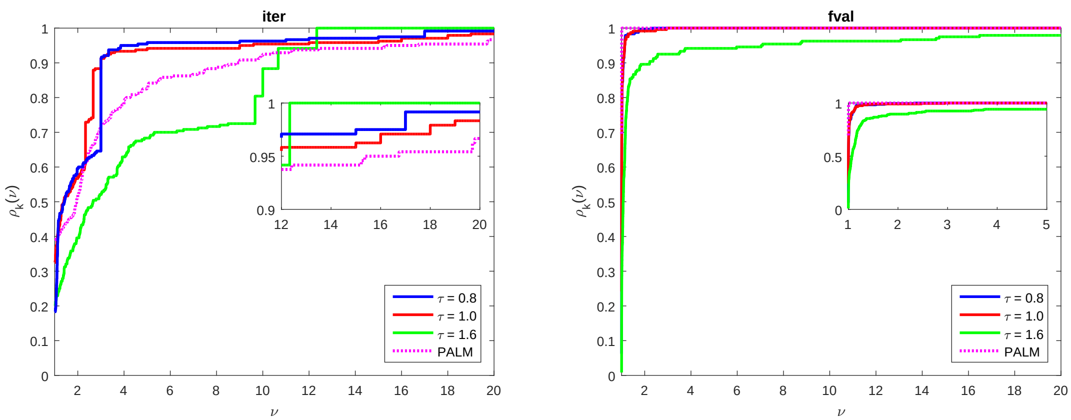
computational results are reported in Table 3.3, where we report p and α, the optimal μ, the number of iterations, the CPU time (seconds) and F-measure. We also show the extracted backgrounds and foregrounds given by the ADMM in Fig. 3.6.

Extraction from noisy and blurred surveillance videos In this case, $\mathcal{A}(L+S) = H(L+S)$, $\lambda_{\max} = \lambda_{\max}(H^*H)$, $\lambda_{\min} = \lambda_{\min}(H^*H)$ and we set $\mathrm{Tol}_{A,1} = 5\times 10^{-3}$, $\mathrm{Tol}_{A,2} = 10^{-2}$ and $\mathrm{Tol}_P = 3\times 10^{-3}$. The blurring matrix H is generated by the same method introduced in Subsection 3.3.2. One frame of each corrupted video is shown in the second row in Fig. 3.2. We report the computational results in Table 3.4 and show the extracted backgrounds and foregrounds by the ADMM in Fig. 3.7.

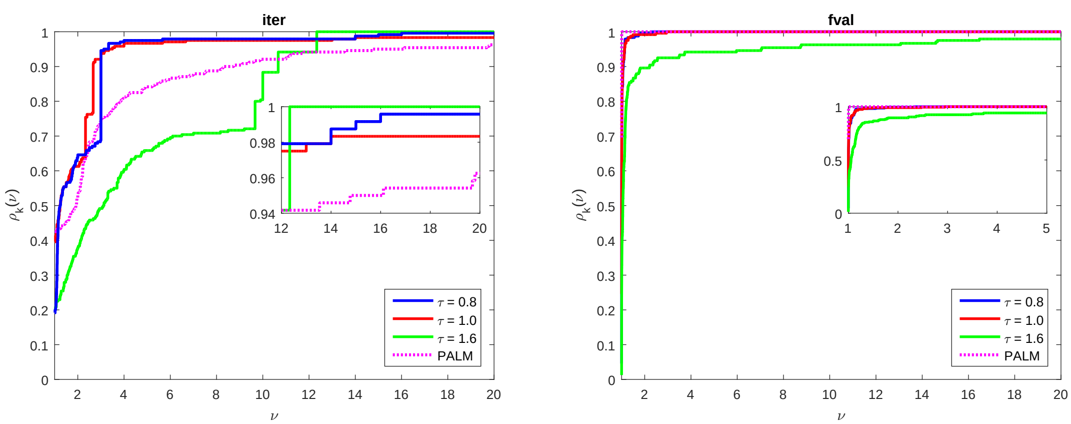
Summary From the results above, it can be seen that the ADMM with $\tau = 0.8$ performs better in the sense that it takes less CPU time for solving most test problems while returning comparable F-measures. The performances of our ADMM for extraction are also promising from Fig. 3.6 and Fig. 3.7.



(a) bridge regularizer

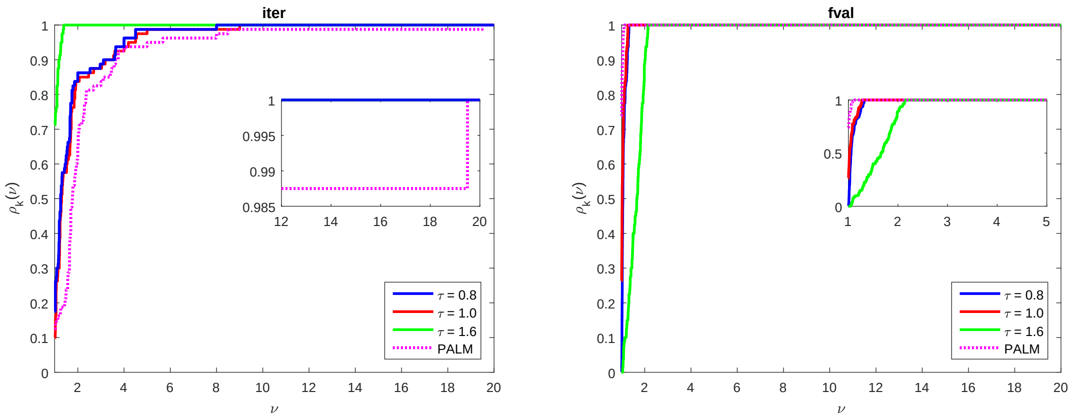


(b) fraction regularizer

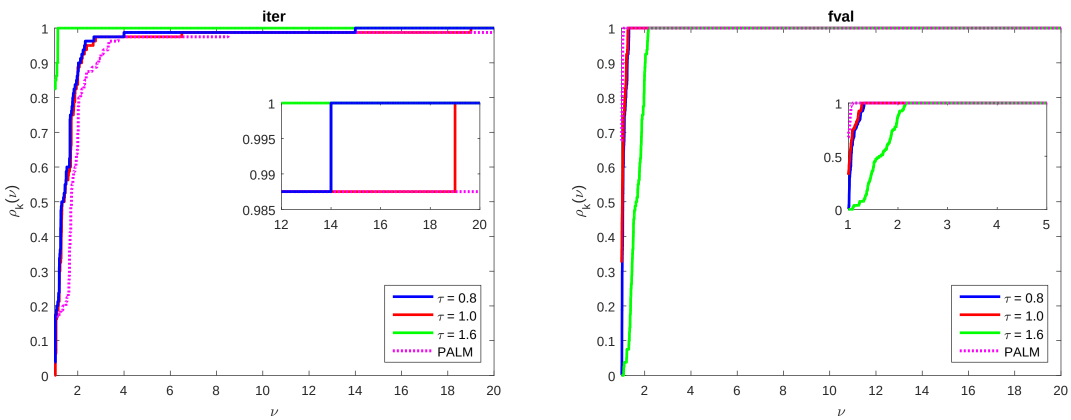


(c) logistic regularizer

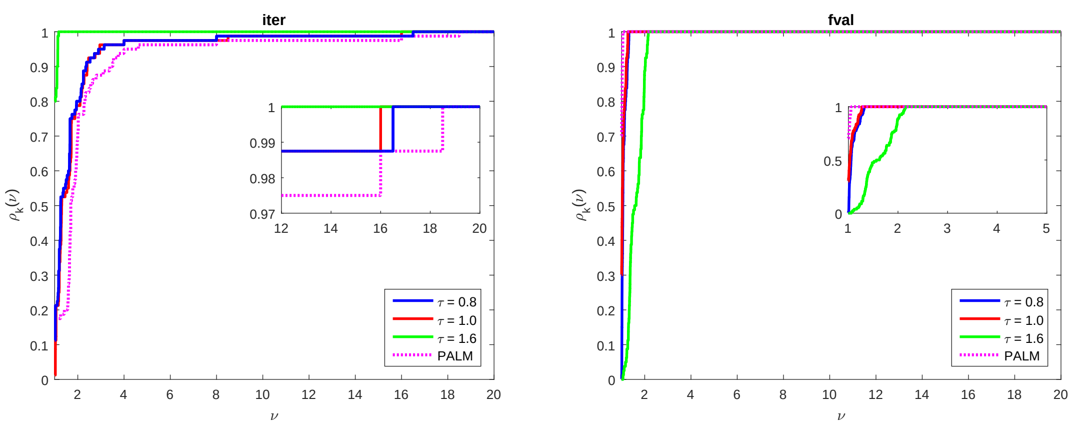
Figure 3.3: Performance profiles of iteration numbers (denoted by "**iter**" on the left) and function values (denoted by "**fval**" on the right) for each sparse regularizer with $\mathcal{A}(L + S) = L + S$. The blown-up subfigures are used to highlight the differences in a specific range of ν.



(a) bridge regularizer

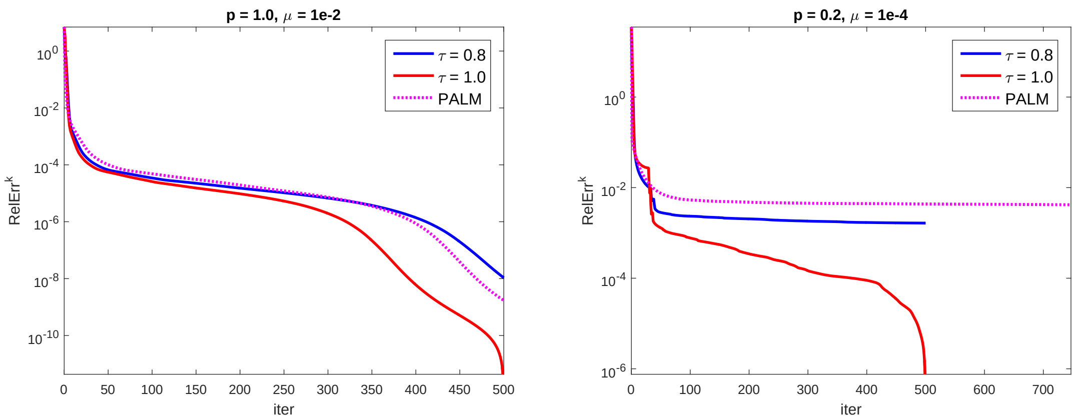


(b) fraction regularizer

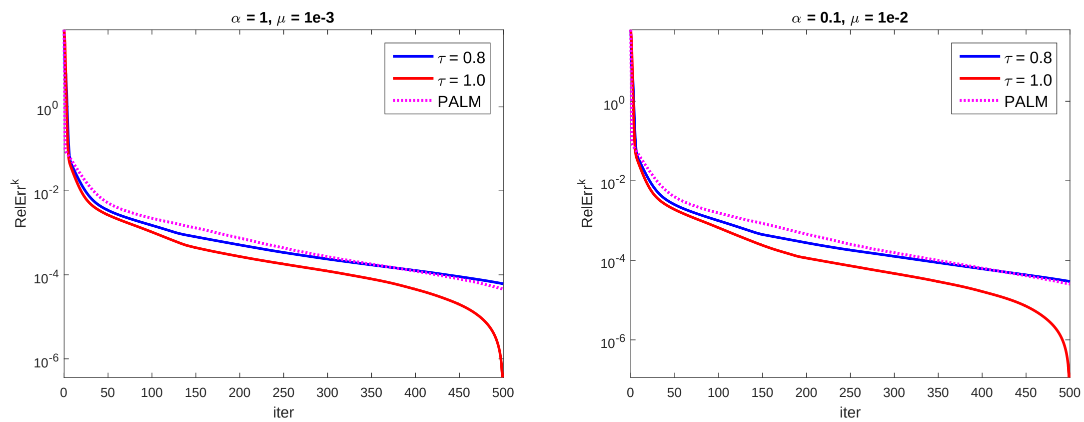


(c) logistic regularizer

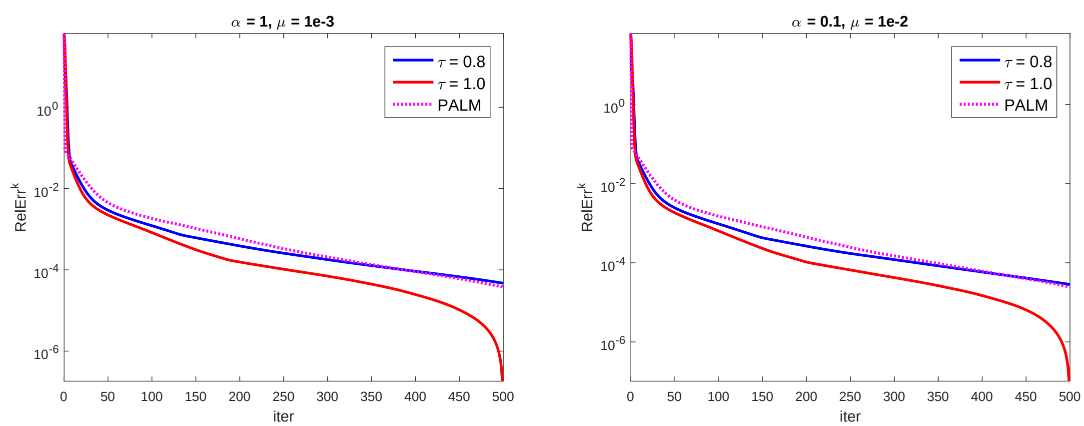
Figure 3.4: Performance profiles of iteration numbers (denoted by "**iter**" on the left) and function values (denoted by "**fval**" on the right) for each sparse regularizer with $\mathcal{A}(L+S) = H(L+S)$. The blown-up subfigures are used to highlight the differences in a specific range of ν.



(a) bridge regularizer



(b) fraction regularizer



(c) logistic regularizer

Figure 3.5: The RelErr^k vs the number of iterations for each sparse regularizer

Table 3.3: Numerical results for extraction for case $\mathcal{A}(L+S) = L+S$

Data	regularizer		ADMM μ	iter	time	F-measure	fval	PALM μ	iter	time	F-measure	fval
Hall	bri. p	1.0	5e-02	10	**2.07**	0.7562	1869.66	5e-02	19	2.58	0.7560	1869.62
		0.5	1e-02	32	8.42	0.7634	1149.09	1e-02	36	**6.40**	0.7624	1149.12
	fra. α	1.0	5e-02	23	**5.55**	0.7578	1595.60	5e-02	33	5.56	0.7578	1595.60
		2.0	5e-02	12	2.79	0.7368	2106.98	5e-02	15	**2.36**	0.7368	2106.98
	log. α	1.0	5e-02	12	**11.34**	0.7566	1721.64	5e-02	39	36.11	0.7576	1721.36
		2.0	5e-02	12	**9.61**	0.7368	2426.38	5e-02	16	14.57	0.7368	2426.39
Bootstrap	bri. p	1.0	1e-01	14	2.13	0.8180	8832.69	1e-01	19	**1.82**	0.8180	8832.69
		0.5	5e-02	23	5.27	0.8206	7960.33	5e-02	22	**3.70**	0.8209	7959.47
	fra. α	1.0	1e-01	15	**3.36**	0.8163	7155.41	1e-01	20	3.44	0.8165	7155.41
		2.0	1e-01	14	3.15	0.8264	9709.11	1e-01	16	**2.49**	0.8261	9709.11
	log. α	1.0	1e-01	16	**11.37**	0.8195	7941.13	1e-01	22	14.56	0.8195	7941.13
		2.0	1e-01	12	5.06	0.8363	11473.90	1e-01	10	**3.70**	0.8363	11473.90
Fountain	bri. p	1.0	1e-01	9	1.38	0.7749	2446.89	1e-01	7	**0.65**	0.7749	2446.88
		0.5	5e-02	13	2.23	0.7000	2511.64	5e-02	11	**1.23**	0.6922	2511.32
	fra. α	1.0	1e-01	9	1.66	0.7717	2284.07	1e-01	8	**1.03**	0.7717	2284.07
		2.0	5e-02	10	1.86	0.7717	2151.58	5e-02	9	**1.16**	0.7717	2151.58
	log. α	1.0	1e-01	9	6.81	0.7738	2365.23	1e-01	7	**4.86**	0.7738	2365.23
		2.0	5e-02	9	6.37	0.7717	2294.74	5e-02	8	**5.55**	0.7717	2294.74
ShoppingMall	bri. p	1.0	1e-01	10	6.10	0.7046	16920.08	1e-01	13	**4.96**	0.7043	16920.09
		0.5	1e-02	39	**41.28**	0.7087	6234.82	1e-02	79	70.33	0.7078	6235.19
	fra. α	1.0	1e-01	12	**9.94**	0.7055	15220.65	1e-01	18	11.31	0.7055	15220.65
		2.0	5e-02	15	**12.93**	0.7061	13797.45	5e-02	26	16.00	0.7061	13797.42
	log. α	1.0	1e-01	11	**33.15**	0.7055	16070.07	1e-01	16	44.65	0.7055	16070.08
		2.0	5e-02	12	**27.84**	0.7057	15326.69	5e-02	18	49.90	0.7057	15326.69

Table 3.4: Numerical results for extraction for case $\mathcal{A}(L+S) = H(L+S)$

			ADMM					PALM				
Data	regularizer		μ	iter	time	F-measure	fval	μ	iter	time	F-measure	fval
Hall	bri. p	1.0	5e-02	24	**9.18**	0.6801	6616.61	5e-02	36	15.52	0.6626	6412.11
		0.5	1e-02	44	**19.25**	0.6358	5439.40	1e-02	45	21.11	0.6357	5448.62
	fra. α	1.0	5e-02	57	25.50	0.5265	5692.80	5e-02	49	**22.39**	0.5616	5811.12
		2.0	1e-02	66	33.11	0.5381	4779.83	1e-02	61	**32.21**	0.5445	4852.03
	log. α	1.0	5e-02	42	**48.90**	0.5970	6066.31	5e-02	44	54.91	0.6033	6071.10
		2.0	1e-02	54	67.45	0.5188	5121.25	1e-02	52	**64.82**	0.5211	5167.56
Bootstrap	bri. p	1.0	1e-01	22	**5.97**	0.7651	11098.68	1e-01	50	15.13	0.7364	10866.62
		0.5	1e-02	56	**24.10**	0.6694	4046.93	1e-02	86	37.02	0.6589	3864.54
	fra. α	1.0	5e-02	66	**24.98**	0.5705	5427.10	5e-02	99	36.25	0.5270	5175.45
		2.0	5e-02	73	**25.45**	0.5267	6070.65	5e-02	111	38.21	0.4676	5776.34
	log. α	1.0	5e-02	47	**44.05**	0.5651	6480.76	5e-02	87	81.41	0.5666	6172.91
		2.0	5e-02	73	**68.53**	0.4891	7670.60	5e-02	115	109.46	0.4179	7329.97
Fountain	bri. p	1.0	5e-02	28	**7.83**	0.7229	6167.61	5e-02	64	20.55	0.6970	5688.45
		0.5	1e-02	51	**17.03**	0.6881	5271.96	1e-02	78	27.25	0.6606	5054.11
	fra. α	1.0	5e-02	64	**22.24**	0.5155	5384.44	5e-02	84	29.27	0.5000	5259.77
		2.0	1e-02	62	**24.29**	0.4482	4784.08	1e-02	87	34.54	0.4341	4596.22
	log. α	1.0	5e-02	50	**47.58**	0.6095	5676.79	5e-02	77	74.04	0.5760	5451.61
		2.0	1e-02	53	**50.74**	0.4438	5042.31	1e-02	80	77.10	0.4525	4819.92
ShoppingMall	bri. p	1.0	5e-02	22	**31.99**	0.6431	14000.88	5e-02	25	39.06	0.6411	13957.07
		0.5	5e-03	54	107.27	0.6271	5590.31	5e-03	40	**82.12**	0.6328	5878.83
	fra. α	1.0	1e-02	33	**65.56**	0.5045	5267.76	5e-02	60	108.54	0.5106	10698.09
		2.0	1e-02	51	94.69	0.5810	6137.38	1e-02	38	**74.95**	0.5935	6450.37
	log. α	1.0	5e-02	58	237.06	0.5453	12096.71	5e-02	39	**158.97**	0.5967	12514.39
		2.0	1e-02	38	154.29	0.5856	7182.66	1e-02	33	**135.12**	0.5913	7332.29

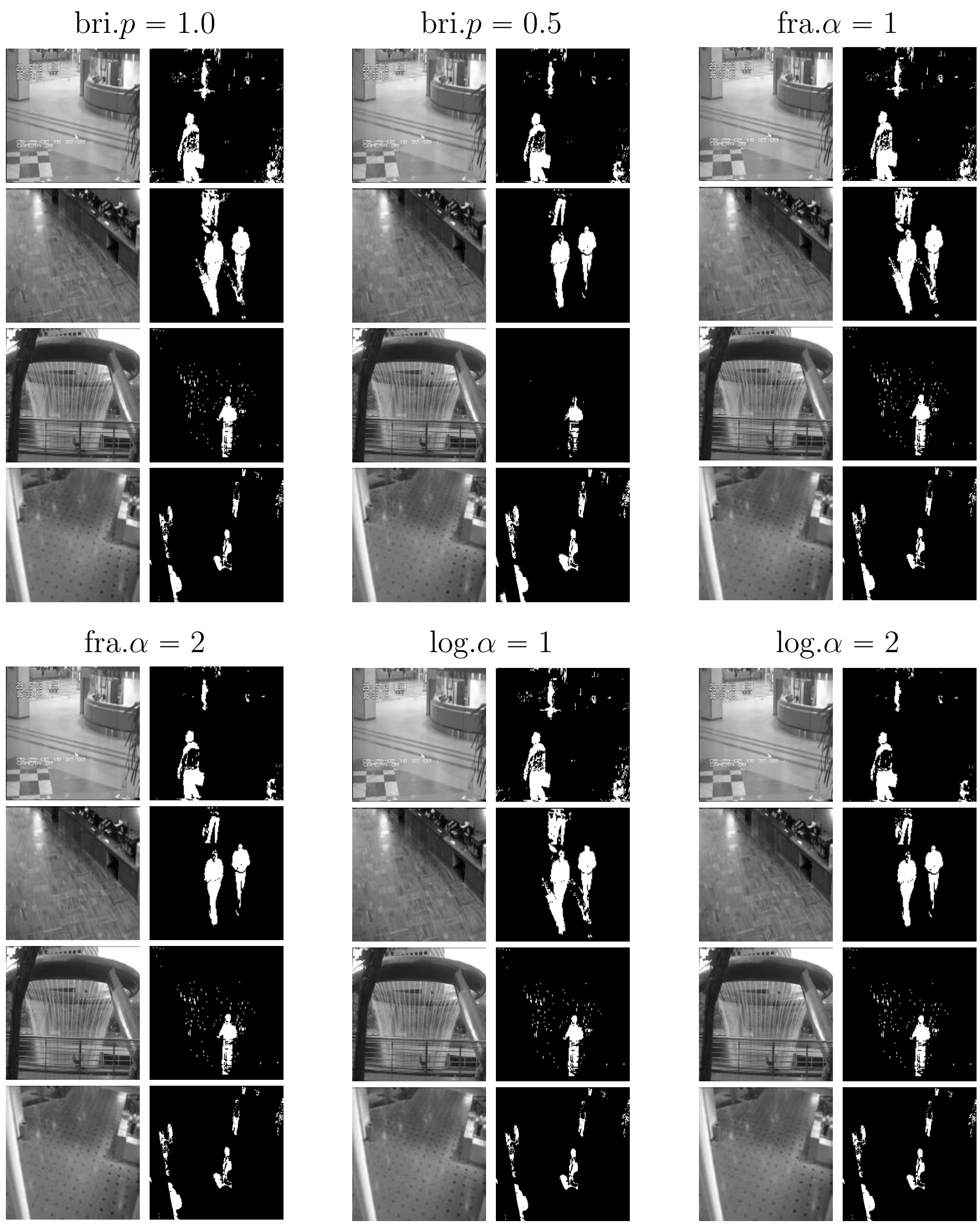


Figure 3.6: Extracted backgrounds and foregrounds given by the ADMM for noisy surveillance videos.

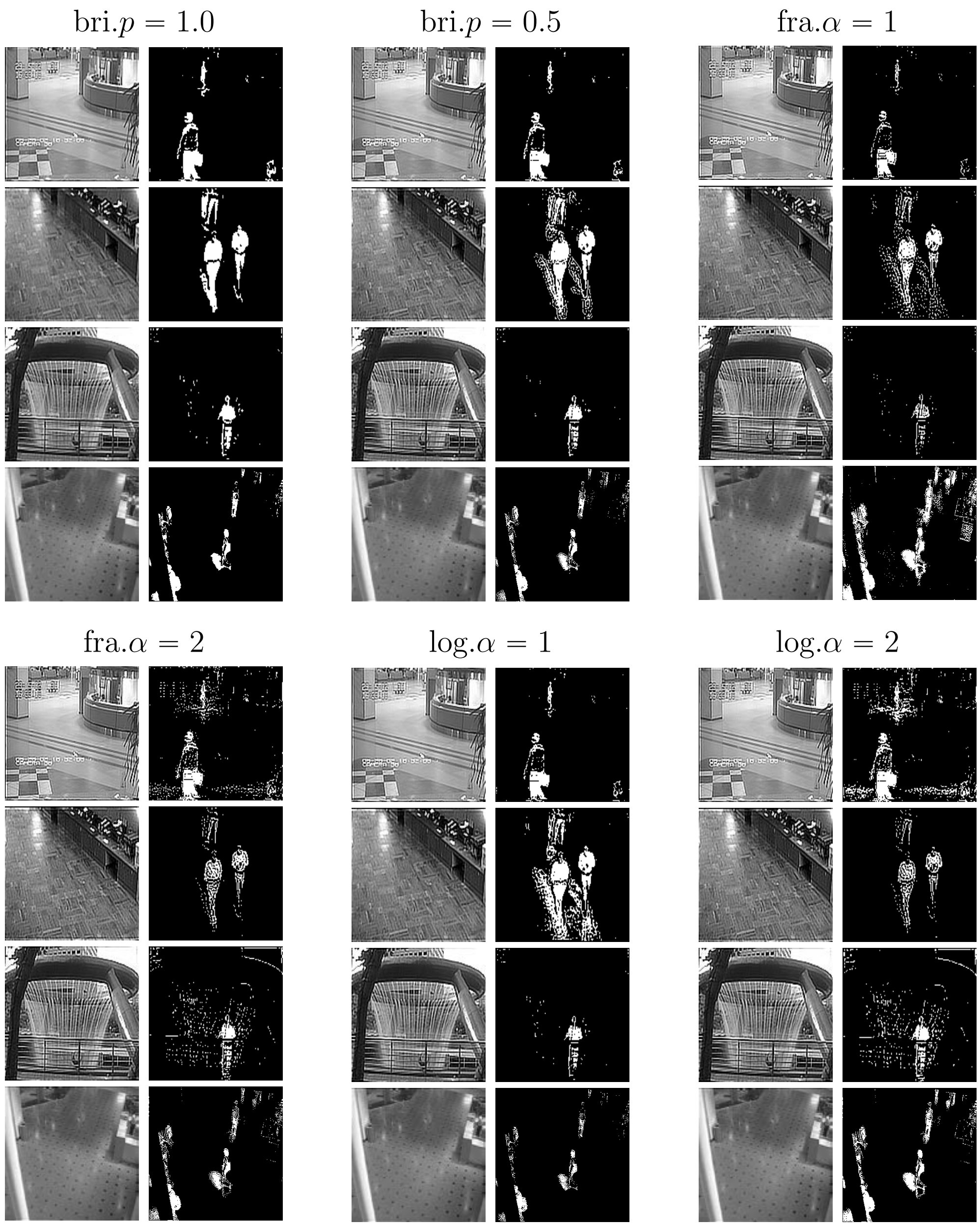


Figure 3.7: Extracted backgrounds and foregrounds given by the ADMM for noisy and blurred surveillance videos.

Chapter 4

NAUM for Matrix Factorization Problem

In this chapter, we consider the matrix factorization problem (1.4), i.e.,

$$\min_{X,Y} \ \mathcal{F}(X,Y) := \Psi(X) + \Phi(Y) + \frac{1}{2}\left\|\mathcal{A}(XY^\top) - \boldsymbol{b}\right\|^2.$$

To solve this possibly nonconvex, nonsmooth and non-Lipschitz problem, we introduce a potential function $\Theta_{\alpha,\beta}$ in (1.5), which is specifically constructed for the objective $\mathcal{F}$ in (1.4). For ease of future reference, we recall that $\Theta_{\alpha,\beta}$ is given by

$$\Theta_{\alpha,\beta}(X,Y,Z) := \Psi(X) + \Phi(Y) + \frac{\alpha}{2}\|XY^\top - Z\|_F^2 + \frac{\beta}{2}\left\|\mathcal{A}(Z) - \boldsymbol{b}\right\|^2,$$

where α and β are real numbers. The relation between $\mathcal{F}$ and $\Theta_{\alpha,\beta}$ is discussed in Section 4.1. We then, in Section 4.2, develop a non-monotone alternating updating method (NAUM) with a suitable line search criterion based on this potential function. In Section 4.3, under some mild conditions, we show the well-definedness of our line search criterion, and establish that the sequence generated by NAUM is bounded and any cluster point of the sequence gives a stationary point of $\mathcal{F}$. Moreover, we discuss the convergence rate for the function value if, in addition, the objective is a Kurdyka-Łojasiewicz function. Finally, in Section 4.4, we conduct numerical experiments to compare our method with some existing efficient methods for non-negative matrix factorization and matrix completion on real datasets. The numerical results show that our method can outperform these methods for these specific applications.

Before proceeding, we discuss the first-order necessary conditions for (1.4). First, from [71, Exercise 8.8] and [71, Proposition 10.5], we see that

$$\partial\mathcal{F}(X, Y) = \begin{pmatrix} \partial\Psi(X) + \mathcal{A}^*\left(\mathcal{A}(XY^\top) - \boldsymbol{b}\right)Y \\ \partial\Phi(Y) + \left(\mathcal{A}^*\left(\mathcal{A}(XY^\top) - \boldsymbol{b}\right)\right)^\top X \end{pmatrix}.$$

Then, it follows from the generalized Fermat's rule [71, Theorem 10.1] that any local minimizer $(\bar{X}, \bar{Y})$ of (1.4) satisfies $0 \in \partial\mathcal{F}(\bar{X}, \bar{Y})$, i.e.,

$$\begin{cases} 0 \in \partial\Psi(\bar{X}) + \mathcal{A}^*(\mathcal{A}(\bar{X}\bar{Y}^\top) - \boldsymbol{b})\bar{Y}, \\ 0 \in \partial\Phi(\bar{Y}) + (\mathcal{A}^*(\mathcal{A}(\bar{X}\bar{Y}^\top) - \boldsymbol{b}))^\top \bar{X}, \end{cases} \tag{4.1}$$

which implies that $(\bar{X}, \bar{Y})$ is a stationary point of $\mathcal{F}$. In this chapter, we again focus on finding a stationary point (X^*, Y^*) of $\mathcal{F}$ in (1.4), i.e., (X^*, Y^*) satisfies (4.1) in place of $(\bar{X}, \bar{Y})$.

Additionally, we present the following two propositions, which will be useful for developing our NAUM.

Proposition 4.1. *Suppose that $\mathcal{A}\mathcal{A}^* = \mathcal{I}$ and $\alpha(\alpha+\beta) \neq 0$. Then, $\alpha\mathcal{I} + \beta\mathcal{A}^*\mathcal{A}$ is invertible and its inverse is given by $\frac{1}{\alpha}\mathcal{I} - \frac{\beta}{\alpha(\alpha+\beta)}\mathcal{A}^*\mathcal{A}$.*

Proof. It is easy to check that $\frac{1}{\alpha}\mathcal{I} - \frac{\beta}{\alpha(\alpha+\beta)}\mathcal{A}^*\mathcal{A}$ is well defined since $\alpha(\alpha+\beta) \neq 0$, and that

$$\left(\alpha\mathcal{I} + \beta\mathcal{A}^*\mathcal{A}\right)\left(\frac{1}{\alpha}\mathcal{I} - \frac{\beta}{\alpha(\alpha+\beta)}\mathcal{A}^*\mathcal{A}\right) = \mathcal{I}.$$

This completes the proof. $\square$

Proposition 4.2. *Let $\psi : \mathbb{R}^m \to (-\infty, \infty]$ and $\phi : \mathbb{R}^n \to (-\infty, \infty]$ be proper closed functions. Given $P, Q \in \mathbb{R}^{m\times n}$ and $\boldsymbol{a} \in \mathbb{R}^n$, $\boldsymbol{b} \in \mathbb{R}^m$ with $\|\boldsymbol{a}\| \neq 0$, $\|\boldsymbol{b}\| \neq 0$, the following statements hold.*

(i) *The problem* $\min_{\boldsymbol{x}\in\mathbb{R}^m} \left\{\psi(\boldsymbol{x}) + \frac{1}{2}\|\boldsymbol{x}\boldsymbol{a}^\top - P\|_F^2\right\}$ *is equivalent to the problem*

$$\min_{\boldsymbol{x}\in\mathbb{R}^m} \left\{\psi(\boldsymbol{x}) + \frac{\|\boldsymbol{a}\|^2}{2}\left\|\boldsymbol{x} - \frac{P\boldsymbol{a}}{\|\boldsymbol{a}\|^2}\right\|^2\right\};$$

(ii) *The problem* $\min_{\boldsymbol{y}\in\mathbb{R}^n} \left\{\phi(\boldsymbol{y}) + \frac{1}{2}\|\boldsymbol{b}\boldsymbol{y}^\top - Q\|_F^2\right\}$ *is equivalent to the problem*

$$\min_{\boldsymbol{y}\in\mathbb{R}^n} \left\{\phi(\boldsymbol{y}) + \frac{\|\boldsymbol{b}\|^2}{2}\left\|\boldsymbol{y} - \frac{Q^\top\boldsymbol{b}}{\|\boldsymbol{b}\|^2}\right\|^2\right\}.$$

Proof. Statement (i) can be easily proved by noticing that

$$\begin{aligned}\|\boldsymbol{x}\boldsymbol{a}^\top - P\|_F^2 = \|\boldsymbol{x}\boldsymbol{a}^\top\|_F^2 - 2\langle \boldsymbol{x}\boldsymbol{a}^\top, P\rangle + \|P\|_F^2 &= \|\boldsymbol{a}\|^2\|\boldsymbol{x}\|^2 - 2\langle \boldsymbol{x}, P\boldsymbol{a}\rangle + \|P\|_F^2 \\ &= \|\boldsymbol{a}\|^2\left\|\boldsymbol{x} - P\boldsymbol{a}/\|\boldsymbol{a}\|^2\right\|^2 - \|P\boldsymbol{a}\|^2/\|\boldsymbol{a}\|^2 + \|P\|_F^2.\end{aligned}$$

Then, statement (ii) can be easily proved by using statement (i) and $\|\boldsymbol{b}\boldsymbol{y}^\top - Q\|_F^2 = \|\boldsymbol{y}\boldsymbol{b}^\top - Q^\top\|_F^2$. $\square$

4.1 The potential function

In this section, we analyze the relation between $\mathcal{F}$ in (1.4) and its potential function $\Theta_{\alpha,\beta}$ in (1.5). Intuitively, $\Theta_{\alpha,\beta}$ originates from $\mathcal{F}$ by separating the coupled variables $XY^\top$ from the linear mapping $\mathcal{A}$ via introducing an auxiliary variable Z and penalizing $XY^\top = Z$. We will see later that the stationary point of $\mathcal{F}$ can be characterized by the stationary point of $\Theta_{\alpha,\beta}$. Before proceeding, we prove the following technical lemma.

Lemma 4.1. *Suppose that* $\mathcal{A}\mathcal{A}^* = \mathcal{I}$ *and* $\frac{1}{\alpha} + \frac{1}{\beta} = 1$. *Then, for any* (X, Y, Z) *satisfying*

$$Z = \left(\mathcal{I} - \tfrac{\beta}{\alpha+\beta}\mathcal{A}^*\mathcal{A}\right)\left(XY^\top\right) + \tfrac{\beta}{\alpha+\beta}\mathcal{A}^*(\boldsymbol{b}), \tag{4.2}$$

we have $\mathcal{F}(X,Y) = \Theta_{\alpha,\beta}(X,Y,Z)$.

Proof. First, from (4.2), we have

$$\begin{aligned} XY^\top - Z &= \tfrac{\beta}{\alpha+\beta}\mathcal{A}^*(\mathcal{A}(XY^\top) - \boldsymbol{b}), \qquad (4.3) \\ \mathcal{A}(Z) - \boldsymbol{b} &= \mathcal{A}\left(XY^\top - \tfrac{\beta}{\alpha+\beta}\mathcal{A}^*\mathcal{A}(XY^\top) + \tfrac{\beta}{\alpha+\beta}\mathcal{A}^*(\boldsymbol{b})\right) - \boldsymbol{b} \\ &= \mathcal{A}(XY^\top) - \tfrac{\beta}{\alpha+\beta}\mathcal{A}\mathcal{A}^*\mathcal{A}(XY^\top) + \tfrac{\beta}{\alpha+\beta}\mathcal{A}\mathcal{A}^*(\boldsymbol{b}) - \boldsymbol{b} \\ &= \tfrac{\alpha}{\alpha+\beta}\left(\mathcal{A}(XY^\top) - \boldsymbol{b}\right), \qquad (4.4) \end{aligned}$$

where the last equality follows from $\mathcal{A}\mathcal{A}^* = \mathcal{I}$. Then, we see that

$$\begin{aligned} &\frac{\alpha}{2}\|XY^\top - Z\|_F^2 + \frac{\beta}{2}\left\|\mathcal{A}(Z) - \boldsymbol{b}\right\|^2 \\ &= \frac{\alpha}{2}\left\|\tfrac{\beta}{\alpha+\beta}\mathcal{A}^*(\mathcal{A}(XY^\top) - \boldsymbol{b})\right\|_F^2 + \frac{\beta}{2}\left\|\tfrac{\alpha}{\alpha+\beta}\left(\mathcal{A}(XY^\top) - \boldsymbol{b}\right)\right\|^2 \\ &= \tfrac{\alpha\beta^2}{(\alpha+\beta)^2}\cdot\frac{1}{2}\left\|\mathcal{A}^*(\mathcal{A}(XY^\top) - \boldsymbol{b})\right\|_F^2 + \tfrac{\alpha^2\beta}{(\alpha+\beta)^2}\cdot\frac{1}{2}\left\|\mathcal{A}(XY^\top) - \boldsymbol{b}\right\|^2 \\ &= \tfrac{\alpha\beta^2}{(\alpha+\beta)^2}\cdot\frac{1}{2}\left\|\mathcal{A}(XY^\top) - \boldsymbol{b}\right\|^2 + \tfrac{\alpha^2\beta}{(\alpha+\beta)^2}\cdot\frac{1}{2}\left\|\mathcal{A}(XY^\top) - \boldsymbol{b}\right\|^2 \\ &= \tfrac{\alpha\beta}{\alpha+\beta}\cdot\frac{1}{2}\left\|\mathcal{A}(XY^\top) - \boldsymbol{b}\right\|^2, \end{aligned}$$

where the first equality follows from (4.3) and (4.4); and the third equality follows from $\mathcal{A}\mathcal{A}^* = \mathcal{I}$. This, together with $\frac{1}{\alpha} + \frac{1}{\beta} = 1$ and the definitions of $\mathcal{F}$ and $\Theta_{\alpha,\beta}$ completes the proof. □

Based on the above lemma, we now establish the following property of $\Theta_{\alpha,\beta}$.

Theorem 4.1. *Suppose that $\mathcal{A}\mathcal{A}^* = \mathcal{I}$. If α and β are chosen such that $\alpha\mathcal{I}+\beta\mathcal{A}^*\mathcal{A} \succ 0$ and $\frac{1}{\alpha} + \frac{1}{\beta} = 1$, then the problem $\min\limits_{X,Y,Z}\{\Theta_{\alpha,\beta}(X,Y,Z)\}$ is equivalent to (1.4).*

Proof. First, it is easy to see from $\alpha\mathcal{I} + \beta\mathcal{A}^*\mathcal{A} \succ 0$ that the function $Z \longmapsto \Theta_{\alpha,\beta}(X,Y,Z)$ is strongly convex. Thus, for any fixed X and Y, the optimal solution Z^* to the problem $\min\limits_{Z}\{\Theta_{\alpha,\beta}(X,Y,Z)\}$ exists and is unique, and can be obtained explicitly. Indeed, from the optimality condition, we have

$$\alpha(Z^* - XY^\top) + \beta\mathcal{A}^*(\mathcal{A}(Z^*) - \boldsymbol{b}) = 0.$$

Then, since $\alpha\mathcal{I} + \beta\mathcal{A}^*\mathcal{A}$ is invertible (as $\alpha\mathcal{I} + \beta\mathcal{A}^*\mathcal{A} \succ 0$), we see that

$$\begin{aligned}
Z^* &= (\alpha\mathcal{I} + \beta\mathcal{A}^*\mathcal{A})^{-1}\left[\alpha XY^\top + \beta\mathcal{A}^*(\boldsymbol{b})\right] \\
&= \left[\tfrac{1}{\alpha}\mathcal{I} - \tfrac{\beta}{\alpha(\alpha+\beta)}\mathcal{A}^*\mathcal{A}\right]\left[\alpha XY^\top + \beta\mathcal{A}^*(\boldsymbol{b})\right] \\
&= \left(\mathcal{I} - \tfrac{\beta}{\alpha+\beta}\mathcal{A}^*\mathcal{A}\right)(XY^\top) + \left[\tfrac{\beta}{\alpha}\mathcal{A}^*(\boldsymbol{b}) - \tfrac{\beta^2}{\alpha(\alpha+\beta)}\mathcal{A}^*\mathcal{A}\mathcal{A}^*(\boldsymbol{b})\right] \\
&= \left(\mathcal{I} - \tfrac{\beta}{\alpha+\beta}\mathcal{A}^*\mathcal{A}\right)(XY^\top) + \left[\tfrac{\beta}{\alpha} - \tfrac{\beta^2}{\alpha(\alpha+\beta)}\right]\mathcal{A}^*(\boldsymbol{b}) \\
&= \left(\mathcal{I} - \tfrac{\beta}{\alpha+\beta}\mathcal{A}^*\mathcal{A}\right)(XY^\top) + \tfrac{\beta}{\alpha+\beta}\mathcal{A}^*(\boldsymbol{b}),
\end{aligned}$$

where the second equality follows from Proposition 4.1 and the fourth equality follows from $\mathcal{A}\mathcal{A}^* = \mathcal{I}$. This, together with Lemma 4.1, implies that $\mathcal{F}(X,Y) = \Theta_{\alpha,\beta}(X,Y,Z^*)$. Then, we have that

$$\begin{aligned}
\min_{X,Y,Z}\left\{\Theta_{\alpha,\beta}(X,Y,Z)\right\} &= \min_{X,Y}\left\{\min_{Z}\left\{\Theta_{\alpha,\beta}(X,Y,Z)\right\}\right\} \\
&= \min_{X,Y}\left\{\Theta_{\alpha,\beta}(X,Y,Z^*)\right\} \\
&= \min_{X,Y}\left\{\mathcal{F}(X,Y)\right\}.
\end{aligned}$$

This completes the proof. □

Remark 4.1 (Comments on Theorem 4.1). *From the proof of Lemma 4.1, we see that if Φ and Ψ are the indicator functions of some nonempty closed sets, then $\mathcal{F}(X,Y) = \left(\frac{1}{\alpha} + \frac{1}{\beta}\right)\Theta_{\alpha,\beta}(X,Y,Z)$ holds with the special choice of Z in (4.2) whenever $\mathcal{A}\mathcal{A}^* = \mathcal{I}$ and $\frac{1}{\alpha} + \frac{1}{\beta} > 0$. Thus, the result in Theorem 4.1 remains valid whenever $\mathcal{A}\mathcal{A}^* = \mathcal{I}$ and α, β are chosen such that $\alpha\mathcal{I} + \beta\mathcal{A}^*\mathcal{A} \succ 0$ and $\frac{1}{\alpha} + \frac{1}{\beta} > 0$.*

It can be seen from Theorem 4.1 that (1.4) is equivalent to minimizing $\Theta_{\alpha,\beta}$ with some suitable choices of α and β. On the other hand, we can also characterize the relation between the stationary points of $\mathcal{F}$ and $\Theta_{\alpha,\beta}$ under weaker conditions on α and β.

Theorem 4.2. *Suppose that $\mathcal{A}\mathcal{A}^* = \mathcal{I}$ and α, β are chosen such that $\frac{1}{\alpha} + \frac{1}{\beta} = 1$. Then, the following statements hold.*

(i) *If (X^*, Y^*, Z^*) is a stationary point of $\Theta_{\alpha,\beta}$, then (X^*, Y^*) is a stationary point of $\mathcal{F}$;*

(ii) *If (X^*, Y^*) is a stationary point of $\mathcal{F}$, then (X^*, Y^*, Z^*) is a stationary point of $\Theta_{\alpha,\beta}$, where Z^* is given by*

$$Z^* = \left(\mathcal{I} - \tfrac{\beta}{\alpha+\beta}\mathcal{A}^*\mathcal{A}\right)\left(X^*(Y^*)^\top\right) + \tfrac{\beta}{\alpha+\beta}\mathcal{A}^*(\boldsymbol{b}). \tag{4.5}$$

Proof. First, if (X^*, Y^*, Z^*) is a stationary point of $\Theta_{\alpha,\beta}$, then we have $0 \in \partial\Theta_{\alpha,\beta}(X^*, Y^*, Z^*)$, i.e.,

$$\begin{cases} 0 \in \partial\Psi(X^*) + \alpha(X^*(Y^*)^\top - Z^*)Y^*, & \text{(4.6a)} \\ 0 \in \partial\Phi(Y^*) + \alpha(X^*(Y^*)^\top - Z^*)^\top X^*, & \text{(4.6b)} \\ 0 = \alpha(Z^* - X^*(Y^*)^\top) + \beta\mathcal{A}^*(\mathcal{A}(Z^*) - \boldsymbol{b}). & \text{(4.6c)} \end{cases}$$

Since $\frac{1}{\alpha} + \frac{1}{\beta} = 1$, we have $\alpha(\alpha + \beta) \neq 0$ and hence $\alpha\mathcal{I} + \beta\mathcal{A}^*\mathcal{A}$ is invertible from Lemma 4.1. Then, using the same arguments in the proof of Theorem 4.1, we see from (4.6c) that (X^*, Y^*, Z^*) satisfies (4.5). Moreover, using (4.5) and the same arguments in (4.3) and (4.4), we have

$$X^*(Y^*)^\top - Z^* = \tfrac{\beta}{\alpha+\beta}\mathcal{A}^*(\mathcal{A}(X^*(Y^*)^\top) - \boldsymbol{b}), \tag{4.7}$$

$$\mathcal{A}(Z^*) - \boldsymbol{b} = \tfrac{\alpha}{\alpha+\beta}\left(\mathcal{A}(X^*(Y^*)^\top) - \boldsymbol{b}\right). \tag{4.8}$$

Thus, substituting (4.7) into (4.6a) and (4.6b), we see that

$$\begin{cases} 0 \in \partial\Psi(X^*) + \frac{\alpha\beta}{\alpha+\beta}\mathcal{A}^*(\mathcal{A}(X^*(Y^*)^\top) - \boldsymbol{b})Y^*, \\ 0 \in \partial\Phi(Y^*) + \frac{\alpha\beta}{\alpha+\beta}\left(\mathcal{A}^*(\mathcal{A}(X^*(Y^*)^\top) - \boldsymbol{b})\right)^\top X^*. \end{cases} \tag{4.9}$$

This together with $\frac{1}{\alpha} + \frac{1}{\beta} = 1$ implies (X^*, Y^*) is a stationary point of $\mathcal{F}$. This proves statement (i).

We now prove statement (ii). First, if (X^*, Y^*) is a stationary point of $\mathcal{F}$, then invoking $\frac{1}{\alpha} + \frac{1}{\beta} = 1$ and (4.1), we have (4.9). Next, we consider (X^*, Y^*, Z^*) with Z^* given by (4.5). Then, (X^*, Y^*, Z^*) satisfies (4.7) and (4.8). Thus, substituting (4.7) into (4.9), we obtain (4.6a) and (4.6b). Moreover, we have from (4.7) and (4.8) that

$$\begin{aligned} &\alpha(Z^* - X^*(Y^*)^\top) + \beta\mathcal{A}^*(\mathcal{A}(Z^*) - \boldsymbol{b}) \\ = &-\tfrac{\alpha\beta}{\alpha+\beta}\mathcal{A}^*\left((\mathcal{A}(X^*(Y^*)^\top) - \boldsymbol{b}\right) + \beta\mathcal{A}^*\left(\tfrac{\alpha}{\alpha+\beta}\left(\mathcal{A}(X^*(Y^*)^\top) - \boldsymbol{b}\right)\right) = 0. \end{aligned} \tag{4.10}$$

This together with (4.6a) and (4.6b) implies that (X^*, Y^*, Z^*) is a stationary point of $\Theta_{\alpha,\beta}$. This proves statement (ii). $\square$

Remark 4.2 (**Comments on Theorem 4.2**). *From the proof in Theorem 4.2, one can see that if $\partial\Psi$ and $\partial\Phi$ are cones, Theorem 4.2 remains valid under the weaker conditions that $\mathcal{A}\mathcal{A}^* = \mathcal{I}$ and $\frac{1}{\alpha} + \frac{1}{\beta} > 0$.*

From Theorem 4.2, we see that a stationary point of $\mathcal{F}$ can be obtained from a stationary point of $\Theta_{\alpha,\beta}$ with a suitable choice of α and β, i.e., $\frac{1}{\alpha} + \frac{1}{\beta} = 1$. Since the linear map $\mathcal{A}$ is no longer associated with the coupled variables $XY^\top$ in $\Theta_{\alpha,\beta}$, finding a stationary point of $\Theta_{\alpha,\beta}$ is conceivably easier. Thus, one can consider finding a stationary point of $\Theta_{\alpha,\beta}$ in order to find a stationary point of $\mathcal{F}$. Note that some existing alternating-minimization-based methods (see, for example, [3, 91]) can be used to find a stationary point of $\Theta_{\alpha,\beta}$, and hence of $\mathcal{F}$, under the conditions that

$\mathcal{A}\mathcal{A}^* = \mathcal{I}$ and α, β are chosen so that $\alpha\mathcal{I} + \beta\mathcal{A}^*\mathcal{A} \succ 0$ and $\frac{1}{\alpha} + \frac{1}{\beta} = 1$. These conditions further imply that $\alpha > 1$ and $\beta = \frac{\alpha}{\alpha-1} > 1$. However, as we will see from our numerical results in Section 4.4, finding a stationary point of $\Theta_{\alpha,\beta}$ with $\alpha > 1$ can be slow. In view of this, in the next section, we develop a new non-monotone alternating updating method for finding a stationary of $\Theta_{\alpha,\beta}$ (and hence of $\mathcal{F}$) under the weaker conditions that $\mathcal{A}\mathcal{A}^* = \mathcal{I}$ and $\frac{1}{\alpha} + \frac{1}{\beta} = 1$. This allows more flexibilities in choosing α and β.

4.2 Non-monotone alternating updating method

In this section, we consider a non-monotone alternating updating method (NAUM) for finding a stationary point of $\Theta_{\alpha,\beta}$ with $\frac{1}{\alpha} + \frac{1}{\beta} = 1$. Compared to the existing alternating-minimization-based methods [3, 91] applied to $\Theta_{\alpha,\beta}$, which update X, Y, Z by alternately solving subproblems related to $\Theta_{\alpha,\beta}$, NAUM updates Z by an *explicit formula* (see (4.15)) and updates X, Y by solving subproblems related to $\Theta_{\alpha,\beta}$ in a Gauss-Seidel manner. Before presenting the complete algorithm, we first comment on the updates of X and Y.

Let (X^k, Y^k) denote the value of (X, Y) after the $(k-1)$-st iteration, and let (U, V) denote the candidate for (X^{k+1}, Y^{k+1}) at the k-th iteration (we will set (X^{k+1}, Y^{k+1}) to be (U, V) if a line search criterion is satisfied; more details can be found in Algorithm 2). For notational simplicity, we also define

$$\mathcal{H}_\alpha(X, Y, Z) := \frac{\alpha}{2}\|XY^\top - Z\|_F^2$$

for any (X, Y, Z). Then, at the k-th iteration, we first compute Z^k by (4.15) and, in the line search loop, we compute U in one of the following 3 ways for a given $\mu_k > 0$:

- **Proximal**

$$U \in \operatorname*{Argmin}_X \; \Psi(X) + \mathcal{H}_\alpha(X, Y^k, Z^k) + \frac{\mu_k}{2}\|X - X^k\|_F^2. \tag{4.11a}$$

- **Prox-linear**

$$U \in \underset{X}{\text{Argmin}} \;\; \Psi(X) + \langle \nabla_X \mathcal{H}_\alpha(X^k, Y^k, Z^k),\, X - X^k \rangle + \frac{\mu_k}{2}\|X - X^k\|_F^2. \quad (4.11\text{b})$$

- **Hierarchical-prox** If Ψ is column-wise separable, i.e., $\Psi(X) = \sum_{i=1}^r \psi_i(\boldsymbol{x}_i)$ for $X = [\boldsymbol{x}_1, \cdots, \boldsymbol{x}_r] \in \mathbb{R}^{m\times r}$, we can update U column-by-column. Specifically, for $i = 1, 2, \cdots, r$, compute

$$\boldsymbol{u}_i \in \underset{\boldsymbol{x}_i}{\text{Argmin}} \;\; \psi_i(\boldsymbol{x}_i) + \mathcal{H}_\alpha(\boldsymbol{u}_{j<i}, \boldsymbol{x}_i, \boldsymbol{x}_{j>i}^k, Y^k, Z^k) + \frac{\mu_k}{2}\|\boldsymbol{x}_i - \boldsymbol{x}_i^k\|^2. \quad (4.11\text{c})$$

After computing U, we compute V in one of the following 3 ways for a given $\sigma_k > 0$:

- **Proximal**

$$V \in \underset{Y}{\text{Argmin}} \;\; \Phi(Y) + \mathcal{H}_\alpha(U, Y, Z^k) + \frac{\sigma_k}{2}\|Y - Y^k\|_F^2. \quad (4.12\text{a})$$

- **Prox-linear**

$$V \in \underset{Y}{\text{Argmin}} \;\; \Phi(Y) + \langle \nabla_Y \mathcal{H}_\alpha(U, Y^k, Z^k),\, Y - Y^k \rangle + \frac{\sigma_k}{2}\|Y - Y^k\|_F^2. \quad (4.12\text{b})$$

- **Hierarchical-prox** If Φ is column-wise separable, i.e., $\Phi(Y) = \sum_{i=1}^r \phi_i(\boldsymbol{y}_i)$ for $Y = [\boldsymbol{y}_1, \cdots, \boldsymbol{y}_r] \in \mathbb{R}^{n\times r}$, we can update V column-by-column. Specifically, for $i = 1, 2, \cdots, r$, compute

$$\boldsymbol{v}_i \in \underset{\boldsymbol{y}_i}{\text{Argmin}} \;\; \phi_i(\boldsymbol{y}_i) + \mathcal{H}_\alpha(U, \boldsymbol{v}_{j<i}, \boldsymbol{y}_i, \boldsymbol{y}_{j>i}^k, Z^k) + \frac{\sigma_k}{2}\|\boldsymbol{y}_i - \boldsymbol{y}_i^k\|^2. \quad (4.12\text{c})$$

For notational simplicity, we further let

$$\rho := \left\| \mathcal{I} - \tfrac{\beta}{\alpha+\beta}\mathcal{A}^*\mathcal{A} \right\|^2 \quad (4.13)$$

and let $\gamma \geqslant 0$ be a nonnegative number satisfying

$$(\alpha + \gamma)\mathcal{I} + \beta\mathcal{A}^*\mathcal{A} \succeq 0. \quad (4.14)$$

Now, we are ready to present NAUM as Algorithm 2.

Remark 4.3 (**Comments on "hierarchical-prox"**). *The hierarchical-prox updating scheme requires the column-wise separability of Ψ or Φ. This is satisfied for many common regularizers, for example, $\|\cdot\|_F^2$, $\|\cdot\|_1$, $\|\cdot\|_p^p$ $(0 < p < 1)$, and the indicator function of a column-wise separable constraint.*

Remark 4.4 (**Comments on ρ and γ**). *Since $\mathcal{A}\mathcal{A}^* = \mathcal{I}$, we see that the eigenvalues of $\mathcal{A}^*\mathcal{A}$ are either 0 or 1. Then, the eigenvalues of $\mathcal{I} - \frac{\beta}{\alpha+\beta}\mathcal{A}^*\mathcal{A}$ must be either 1 or $\frac{\alpha}{\alpha+\beta}$, and hence $\rho = \max\{1, \alpha^2/(\alpha+\beta)^2\}$. Similarly, the eigenvalues of $-(\alpha\mathcal{I} + \beta\mathcal{A}^*\mathcal{A})$ are either $-\alpha$ or $-(\alpha+\beta)$. Then, (4.14) is satisfied whenever $\gamma \geqslant \max\{0, -\alpha, -(\alpha+\beta)\}$.*

In Algorithm 2, the update for Z^k is given explicitly. This is motivated by the condition on Z at a stationary point of $\Theta_{\alpha,\beta}$; see (4.6c). In fact, following the same arguments in (4.10), we see that (4.6c) always holds at (X^k, Y^k, Z^k) with Z^k given in (4.15) when $\mathcal{A}\mathcal{A}^* = \mathcal{I}$ and $\frac{1}{\alpha} + \frac{1}{\beta} = 1$. If, in addition, $\alpha\mathcal{I} + \beta\mathcal{A}^*\mathcal{A} \succ 0$ holds, one can show that Z^k is actually the optimal solution to the problem $\min\limits_{Z} \left\{\Theta_{\alpha,\beta}(X^k, Y^k, Z)\right\}$. In this case, our NAUM with $N = 0$ in (4.16) can be considered as an alternating-minimization-based method (see, for example, [3, 91]) applied to the problem $\min\limits_{X,Y,Z} \{\Theta_{\alpha,\beta}(X, Y, Z)\}$.

Our NAUM also allows U and V to be updated in three different ways. Thus, one can choose suitable updating schemes to fit different applications. In particular, if Ψ or Φ are column-wise separable, taking advantage of the structure of $\Theta_{\alpha,\beta}$ and the fact that $XY^\top$ can be written as $\sum_{i=1}^r \boldsymbol{x}_i\boldsymbol{y}_i^\top$ with $X = [\boldsymbol{x}_1, \cdots, \boldsymbol{x}_r] \in \mathbb{R}^{m\times r}$ and $Y = [\boldsymbol{y}_1, \cdots, \boldsymbol{y}_r] \in \mathbb{R}^{n\times r}$, one can update X or Y column-wise even when $\mathcal{A} \neq \mathcal{I}$. The motivation for updating X (or Y) column-wise rather than updating the whole X (or Y) is that the resulting subproblems (4.11c) (or (4.12c)) can be reduced to the computation of the proximal mapping of ψ_i (or ϕ_i), which is easy for many commonly

Algorithm 2 NAUM for finding a stationary point of $\mathcal{F}$

Input: (X^0, Y^0), α and β such that $\frac{1}{\alpha} + \frac{1}{\beta} = 1$, ρ as in (4.13), $\gamma \geqslant 0$ satisfying (4.14), $\tau > 1$, $c > 0$, $\mu^{\min} > 0$, $\sigma^{\max} > \sigma^{\min} > 0$, and an integer $N \geqslant 0$. Set $k = 0$.

while a termination criterion is not met, **do**

Step 1. Compute Z^k by

$$Z^k = \left(\mathcal{I} - \frac{\beta}{\alpha+\beta}\mathcal{A}^*\mathcal{A}\right)\left(X^k(Y^k)^\top\right) + \frac{\beta}{\alpha+\beta}\mathcal{A}^*(\boldsymbol{b}). \tag{4.15}$$

Step 2. Choose $\mu_k^0 \geqslant \mu^{\min}$ and $\sigma_k^0 \in [\sigma^{\min}, \sigma^{\max}]$ arbitrarily. Set $\tilde{\mu}_k = \mu_k^0$, $\sigma_k = \sigma_k^0$ and $\mu_k^{\max} = (\alpha + 2\gamma\rho)\|Y^k\|^2 + c$.

(2a) Set $\mu_k \leftarrow \min\{\tilde{\mu}_k, \mu_k^{\max}\}$. Compute U by either (4.11a), (4.11b) or (4.11c).

(2b) Compute V by either (4.12a), (4.12b) or (4.12c).

(2c) If

$$\mathcal{F}(U,V) - \max_{[k-N]_+ \leqslant i \leqslant k} \mathcal{F}(X^i, Y^i) \leqslant -\frac{c}{2}\left(\|U - X^k\|_F^2 + \|V - Y^k\|_F^2\right), \tag{4.16}$$

then go to **Step 3**.

(2d) If $\mu_k = \mu_k^{\max}$, set $\sigma_k^{\max} = (\alpha + 2\gamma\rho)\|U\|^2 + c$, $\sigma_k \leftarrow \min\{\tau\sigma_k, \sigma_k^{\max}\}$ and then, go to step **(2b)**; otherwise, set $\tilde{\mu}_k \leftarrow \tau\mu_k$ and $\sigma_k \leftarrow \tau\sigma_k$ and then, go to step **(2a)**.

Step 3. Set $X^{k+1} \leftarrow U$, $Y^{k+1} \leftarrow V$, $\bar{\mu}_k \leftarrow \mu_k$, $\bar{\sigma}_k \leftarrow \sigma_k$, $k \leftarrow k+1$ and go to **Step 1**.

end while

Output: (X^k, Y^k)

---

used ψ_i (or ϕ_i). Indeed, from (4.11c) and (4.12c), $\boldsymbol{u}_i$ and $\boldsymbol{v}_i$ are given by

$$\begin{cases} \boldsymbol{u}_i \in \operatorname*{Argmin}_{\boldsymbol{x}_i} \left\{\psi_i(\boldsymbol{x}_i) + \dfrac{\alpha}{2}\left\|\boldsymbol{x}_i(\boldsymbol{y}_i^k)^\top - P_i^k\right\|_F^2 + \dfrac{\mu_k}{2}\|\boldsymbol{x}_i - \boldsymbol{x}_i^k\|^2\right\}, \\ \boldsymbol{v}_i \in \operatorname*{Argmin}_{\boldsymbol{y}_i} \left\{\phi_i(\boldsymbol{y}_i) + \dfrac{\alpha}{2}\left\|\boldsymbol{u}_i\boldsymbol{y}_i^\top - Q_i^k\right\|_F^2 + \dfrac{\sigma_k}{2}\|\boldsymbol{y}_i - \boldsymbol{y}_i^k\|^2\right\}, \end{cases} \tag{4.17}$$

where P_i^k and Q_i^k are defined by

$$\begin{aligned} P_i^k &:= Z^k - \textstyle\sum_{j=1}^{i-1}\boldsymbol{u}_j(\boldsymbol{y}_j^k)^\top - \sum_{j=i+1}^{r}\boldsymbol{x}_j^k(\boldsymbol{y}_j^k)^\top, \\ Q_i^k &:= Z^k - \textstyle\sum_{j=1}^{i-1}\boldsymbol{u}_j\boldsymbol{v}_j^\top - \sum_{j=i+1}^{r}\boldsymbol{u}_j(\boldsymbol{y}_j^k)^\top. \end{aligned} \tag{4.18}$$

Then, from Proposition 4.2, we can reformulate the subproblems in (4.17) and obtain the corresponding solutions by computing the proximal mappings of ψ_i and ϕ_i, which can be computed efficiently when ψ_i and ϕ_i are some common regularizers used in the literature. In particular, when $\psi_i(\cdot)$ and $\phi_i(\cdot)$ are $\|\cdot\|_1$, $\|\cdot\|_2^2$ or the indicator function of the box constraint, these subproblems have closed-form solutions. This updating strategy has also been used for nonnegative matrix factorization; see, for example, [19, 53, 58]. However, the methods used in [19, 53, 58] can only be applied for some specific problems with $\mathcal{A} = \mathcal{I}$, while NAUM can be applied for more general problems with $\mathcal{A}\mathcal{A}^* = \mathcal{I}$.

Additionally, NAUM adapts a non-monotone line search criterion (see Step 2 in Algorithm 2) to improve the numerical performance. This is motivated by recent studies on non-monotone algorithms with promising performances; see, for example, [17, 34, 88]. However, different from the non-monotone line search criteria used there, NAUM only includes (U, V) in the line search loop and checks the stopping criterion (4.16) after updating a pair of (U, V), rather than checking (4.16) immediately once U or V is updated. Thus, we do not need to compute the function value after updating each block of variable. This may reduce the cost of the line search and make NAUM more practical, especially when computing the function value is relatively expensive.

4.3 Convergence analysis of NAUM

In this section, we discuss the convergence properties of Algorithm 2. First, we present the first-order optimality conditions for the three different updating schemes in (2a) of Algorithm 2 as follows:

- **Proximal**

$$0 \in \partial\Psi(U) + \alpha\left(U(Y^k)^\top - Z^k\right)Y^k + \mu_k(U - X^k), \tag{4.19a}$$

- **Prox-linear**

$$0 \in \partial\Psi(U) + \alpha \left(X^k(Y^k)^\top - Z^k\right) Y^k + \mu_k(U - X^k), \tag{4.19b}$$

- **Hierarchical-prox** For $i = 1, 2, \cdots, r$,

$$0 \in \partial\psi_i(\boldsymbol{u}_i) + \alpha \left(\sum_{j=1}^{i} \boldsymbol{u}_j(\boldsymbol{y}_j^k)^\top + \sum_{j=i+1}^{r} \boldsymbol{x}_j^k(\boldsymbol{y}_j^k)^\top - Z^k\right) \boldsymbol{y}_i^k + \mu_k(\boldsymbol{u}_i - \boldsymbol{x}_i^k). \tag{4.19c}$$

Similarly, the first-order optimality conditions for the three different updating schemes in (2b) of Algorithm 2 are

- **Proximal**

$$0 \in \partial\Phi(V) + \alpha \left(UV^\top - Z^k\right)^\top U + \sigma_k(V - Y^k), \tag{4.20a}$$

- **Prox-linear**

$$0 \in \partial\Phi(V) + \alpha \left(U(Y^k)^\top - Z^k\right)^\top U + \sigma_k(V - Y^k), \tag{4.20b}$$

- **Hierarchical-prox** For $i = 1, 2, \cdots, r$,

$$0 \in \partial\phi_i(\boldsymbol{v}_i) + \alpha \left(\sum_{j=1}^{i} \boldsymbol{u}_j\boldsymbol{v}_j^\top + \sum_{j=i+1}^{r} \boldsymbol{u}_j(\boldsymbol{y}_j^k)^\top - Z^k\right)^\top \boldsymbol{u}_i + \sigma_k(\boldsymbol{v}_i - \boldsymbol{y}_i^k). \tag{4.20c}$$

We also need to make the following assumptions.

Assumption 4.1.

(a1) Ψ, Φ *are proper, closed, level-bounded functions and continuous on their domains;*

(a2) $\mathcal{A}\mathcal{A}^* = \mathcal{I}$*;*

(a3) $\frac{1}{\alpha} + \frac{1}{\beta} = 1.$

Remark 4.5 (**Note on Assumption 4.1**). *(i) From* (a1)*, it is easy to see from [71, Theorem 1.9] that* $\inf \Psi$ *and* $\inf \Phi$ *are finite, i.e.,* Ψ *and* Φ *are bounded from below. In particular, the iterates* (4.11a)*,* (4.11b)*,* (4.11c)*,* (4.12a)*,* (4.12b) *and* (4.12c) *are well defined; (ii) The continuity assumption in* (a1) *holds for many common regularizers, for example,* ℓ_1*-norm, nuclear norm and the indicator function of a nonempty closed set; (iii)* (a2) *is satisfied for some commonly used linear maps, for example, the identity map and the sampling map.*

We start our convergence analysis by proving the following auxiliary lemma.

Lemma 4.2 (**Sufficient descent of $\mathcal{F}$**). *Suppose that Assumption 4.1 holds. Let* $\{(X^k, Y^k)\}$ *be the sequence generated by Algorithm 2 and* (U, V) *be the candidate generated by steps (2a) and (2b) at the k-th iteration. Then, for any integer* $k \geqslant 0$*, we have*

$$\begin{aligned}&\mathcal{F}(U,V) - \mathcal{F}(X^k,Y^k)\\ &\leqslant -\frac{\mu_k - (\alpha+2\gamma\rho)\|Y^k\|^2}{2}\|U - X^k\|_F^2 - \frac{\sigma_k - (\alpha+2\gamma\rho)\|U\|^2}{2}\|V - Y^k\|_F^2.\end{aligned} \tag{4.21}$$

Proof. First, from Lemma 4.1 and (4.15), we see that $\mathcal{F}(X^k, Y^k) = \Theta_{\alpha,\beta}(X^k, Y^k, Z^k)$. For any (U, V), let

$$W = \left(\mathcal{I} - \tfrac{\beta}{\alpha+\beta}\mathcal{A}^*\mathcal{A}\right)\left(UV^\top\right) + \tfrac{\beta}{\alpha+\beta}\mathcal{A}^*(\boldsymbol{b}). \tag{4.22}$$

Then, from Lemma 4.1, we have $\mathcal{F}(U, V) = \Theta_{\alpha,\beta}(U, V, W)$. Thus, we only need to consider the difference $\Theta_{\alpha,\beta}(U, V, W) - \Theta_{\alpha,\beta}(X^k, Y^k, Z^k)$.

We start by noting that

$$\begin{aligned}\mathcal{A}^*\mathcal{A}(W) &= \left(\mathcal{A}^*\mathcal{A} - \tfrac{\beta}{\alpha+\beta}\mathcal{A}^*\left(\mathcal{A}\mathcal{A}^*\right)\mathcal{A}\right)\left(UV^\top\right) + \tfrac{\beta}{\alpha+\beta}\mathcal{A}^*\left(\mathcal{A}\mathcal{A}^*\right)(\boldsymbol{b})\\ &= \tfrac{\alpha}{\alpha+\beta}\mathcal{A}^*\mathcal{A}\left(UV^\top\right) + \tfrac{\beta}{\alpha+\beta}\mathcal{A}^*(\boldsymbol{b}),\end{aligned} \tag{4.23}$$

where the last equality follows from (a2) in Assumption 4.1. Then, we obtain that

$$
\begin{aligned}
&\nabla_Z \Theta_{\alpha,\beta}(U,V,W) \\
&= \alpha(W - UV^\top) + \beta\mathcal{A}^*\mathcal{A}(W) - \beta\mathcal{A}^*(\boldsymbol{b}) \\
&= \alpha\left[-\tfrac{\beta}{\alpha+\beta}\mathcal{A}^*\mathcal{A}(UV^\top) + \tfrac{\beta}{\alpha+\beta}\mathcal{A}^*(\boldsymbol{b})\right] + \beta\left[\tfrac{\alpha}{\alpha+\beta}\mathcal{A}^*\mathcal{A}\left(UV^\top\right) + \tfrac{\beta}{\alpha+\beta}\mathcal{A}^*(\boldsymbol{b})\right] - \beta\mathcal{A}^*(\boldsymbol{b}) \\
&= 0,
\end{aligned}
$$

where the second equality follows from (4.22) and (4.23). Moreover, since γ is chosen such that $(\alpha+\gamma)\mathcal{I} + \beta\mathcal{A}^*\mathcal{A} \succeq 0$ (see (4.14)), we see that, for any $k \geqslant 0$, the function $Z \longmapsto \Theta_{\alpha,\beta}(U,V,Z) + \frac{\gamma}{2}\|Z - Z^k\|_F^2$ is convex and hence

$$
\begin{aligned}
&\Theta_{\alpha,\beta}(U,V,Z^k) + \underbrace{\frac{\gamma}{2}\|Z^k - Z^k\|_F^2}_{=0} \\
&\geqslant \Theta_{\alpha,\beta}(U,V,W) + \frac{\gamma}{2}\|W - Z^k\|_F^2 + \langle \underbrace{\nabla_Z \Theta_{\alpha,\beta}(U,V,W)}_{=0} + \gamma(W - Z^k),\, Z^k - W\rangle,
\end{aligned}
$$

which implies that

$$
\Theta_{\alpha,\beta}(U,V,W) - \Theta_{\alpha,\beta}(U,V,Z^k) \leqslant \frac{\gamma}{2}\|W - Z^k\|_F^2. \tag{4.24}
$$

Then, substituting (4.15) and (4.22) into (4.24), we obtain

$$
\begin{aligned}
&\Theta_{\alpha,\beta}(U,V,W) - \Theta_{\alpha,\beta}(U,V,Z^k) \\
&\leqslant \frac{\gamma}{2}\left\|\left(\mathcal{I} - \tfrac{\beta}{\alpha+\beta}\mathcal{A}^*\mathcal{A}\right)\left(UV^\top - X^k(Y^k)^\top\right)\right\|_F^2 \\
&\leqslant \frac{\gamma}{2}\left\|\mathcal{I} - \tfrac{\beta}{\alpha+\beta}\mathcal{A}^*\mathcal{A}\right\|^2 \cdot \|UV^\top - X^k(Y^k)^\top\|_F^2 \\
&= \frac{\gamma\rho}{2}\|U(V-Y^k)^\top + (U-X^k)(Y^k)^\top\|_F^2 \\
&\leqslant \frac{\gamma\rho}{2}\Big(\|U(V-Y^k)^\top\|_F + \|(U-X^k)(Y^k)^\top\|_F\Big)^2 \\
&\leqslant \frac{\gamma\rho}{2}\Big(\|U\|\|V-Y^k\|_F + \|Y^k\|\|U-X^k\|_F\Big)^2 \\
&\leqslant \gamma\rho\Big(\|U\|^2\|V-Y^k\|_F^2 + \|Y^k\|^2\|U-X^k\|_F^2\Big),
\end{aligned} \tag{4.25}
$$

where the equality follows from the definition of ρ in (4.13), the second last inequality follows from the relation $\|AB\|_F \leqslant \|A\|\|B\|_F$ and the last inequality follows from the relation $\|a+b\|^2 \leqslant 2\|a\|^2 + 2\|b\|^2$.

Next, we show that

$$\Theta_{\alpha,\beta}(U,V,Z^k) - \Theta_{\alpha,\beta}(U,Y^k,Z^k) \leqslant \frac{\alpha\|U\|^2 - \sigma_k}{2}\|V-Y^k\|_F^2. \tag{4.26}$$

To this end, we consider the following three cases.

- Proximal: in this case, we have

$$
\begin{aligned}
&\Theta_{\alpha,\beta}(U,V,Z^k) - \Theta_{\alpha,\beta}(U,Y^k,Z^k) \\
&= \Phi(V) + \mathcal{H}_\alpha(U,V,Z^k) - \Phi(Y^k) - \mathcal{H}_\alpha(U,Y^k,Z^k) \\
&= \left[\Phi(V) + \mathcal{H}_\alpha(U,V,Z^k) + \frac{\sigma_k}{2}\|V-Y^k\|_F^2\right] - \left[\Phi(Y^k) + \mathcal{H}_\alpha(U,Y^k,Z^k)\right] \\
&\quad - \frac{\sigma_k}{2}\|V-Y^k\|_F^2 \\
&\leqslant -\frac{\sigma_k}{2}\|V-Y^k\|_F^2,
\end{aligned}
$$

where the inequality follows from the definition of V as a minimizer of (4.12a). This implies (4.26).

- Prox-linear: in this case, we have

$$
\begin{aligned}
&\Theta_{\alpha,\beta}(U, V, Z^k) - \Theta_{\alpha,\beta}(U, Y^k, Z^k) \\
&= \Phi(V) + \mathcal{H}_\alpha(U, V, Z^k) - \Phi(Y^k) - \mathcal{H}_\alpha(U, Y^k, Z^k) \\
&\leqslant \Phi(V) + \mathcal{H}_\alpha(U, Y^k, Z^k) + \langle \nabla_Y \mathcal{H}_\alpha(U, Y^k, Z^k),\, V - Y^k \rangle + \frac{\alpha\|U\|^2}{2}\|V - Y^k\|_F^2 \\
&\quad - \Phi(Y^k) - \mathcal{H}_\alpha(U, Y^k, Z^k) \\
&= \Phi(V) + \langle \nabla_Y \mathcal{H}_\alpha(U, Y^k, Z^k),\, V - Y^k \rangle + \frac{\sigma_k}{2}\|V - Y^k\|_F^2 - \Phi(Y^k) \\
&\quad + \frac{\alpha\|U\|^2 - \sigma_k}{2}\|V - Y^k\|_F^2 \\
&\leqslant \frac{\alpha\|U\|^2 - \sigma_k}{2}\|V - Y^k\|_F^2,
\end{aligned}
$$

where the first inequality follows from the fact that $Y \mapsto \nabla_Y \mathcal{H}_\alpha(X, Y, Z)$ is Lipschitz with modulus $\alpha\|X\|^2$ and the last inequality follows from the definition of V as a minimizer of (4.12b).

- Hierarchical-prox: in this case, for any $1 \leqslant i \leqslant r$, we have

$$
\begin{aligned}
&\Theta_{\alpha,\beta}(U, \boldsymbol{v}_{j<i}, \boldsymbol{v}_i, \boldsymbol{y}_{j>i}^k, Z^k) - \Theta_{\alpha,\beta}(U, \boldsymbol{v}_{j<i}, \boldsymbol{y}_i^k, \boldsymbol{y}_{j>i}^k, Z^k) \\
&= \phi_i(\boldsymbol{v}_i) + \mathcal{H}_\alpha(U, \boldsymbol{v}_{j<i}, \boldsymbol{v}_i, \boldsymbol{y}_{j>i}^k, Z^k) - \phi_i(\boldsymbol{y}_i^k) - \mathcal{H}_\alpha(U, \boldsymbol{v}_{j<i}, \boldsymbol{y}_i^k, \boldsymbol{y}_{j>i}^k, Z^k) \\
&= \Big[\phi_i(\boldsymbol{v}_i) + \mathcal{H}_\alpha(U, \boldsymbol{v}_{j<i}, \boldsymbol{v}_i, \boldsymbol{y}_{j>i}^k, Z^k) + \frac{\sigma_k}{2}\|\boldsymbol{v}_i - \boldsymbol{y}_i^k\|^2\Big] - \frac{\sigma_k}{2}\|\boldsymbol{v}_i - \boldsymbol{y}_i^k\|^2 \\
&\quad - \big[\phi_i(\boldsymbol{y}_i^k) + \mathcal{H}_\alpha(U, \boldsymbol{v}_{j<i}, \boldsymbol{y}_i^k, \boldsymbol{y}_{j>i}^k, Z^k)\big] \\
&\leqslant -\frac{\sigma_k}{2}\|\boldsymbol{v}_i - \boldsymbol{y}_i^k\|^2,
\end{aligned}
$$

where the inequality follows from the definition of $\boldsymbol{v}_i$ as a minimizer of (4.12c).

Then, summing the above relation from $i = r$ to $i = 1$ and simplifying the resulting inequality, we obtain (4.26).

Similarly, we can show that

$$\Theta_{\alpha,\beta}(U, Y^k, Z^k) - \Theta_{\alpha,\beta}(X^k, Y^k, Z^k) \leqslant \frac{\alpha\|Y^k\|^2 - \mu_k}{2}\|U - X^k\|_F^2 \tag{4.27}$$

by considering the following three cases.

- Proximal: in this case, we have

$$\begin{aligned}
&\Theta_{\alpha,\beta}(U, Y^k, Z^k) - \Theta_{\alpha,\beta}(X^k, Y^k, Z^k) \\
&= \Psi(U) + \mathcal{H}_\alpha(U, Y^k, Z^k) - \Psi(X^k) - \mathcal{H}_\alpha(X^k, Y^k, Z^k) \\
&= \left[\Psi(U) + \mathcal{H}_\alpha(U, Y^k, Z^k) + \frac{\mu_k}{2}\|U - X^k\|_F^2\right] - \left[\Psi(X^k) + \mathcal{H}_\alpha(X^k, Y^k, Z^k)\right] \\
&\quad - \frac{\mu_k}{2}\|U - X^k\|_F^2 \\
&\leqslant -\frac{\mu_k}{2}\|U - X^k\|_F^2,
\end{aligned}$$

  where the inequality follows from the definition of U as a minimizer of (4.11a). This implies (4.27).

- Prox-linear: in this case, we have

$$
\begin{aligned}
&\Theta_{\alpha,\beta}(U, Y^k, Z^k) - \Theta_{\alpha,\beta}(X^k, Y^k, Z^k) \\
&= \Psi(U) + \mathcal{H}_\alpha(U, Y^k, Z^k) - \Psi(X^k) - \mathcal{H}_\alpha(X^k, Y^k, Z^k) \\
&\leqslant \Psi(U) + \mathcal{H}_\alpha(X^k, Y^k, Z^k) + \langle \nabla_X \mathcal{H}_\alpha(X^k, Y^k, Z^k), U - X^k \rangle + \frac{\alpha \|Y^k\|^2}{2} \|U - X^k\|_F^2 \\
&\quad - \Psi(X^k) - \mathcal{H}_\alpha(X^k, Y^k, Z^k) \\
&= \Psi(U) + \langle \nabla_X \mathcal{H}_\alpha(X^k, Y^k, Z^k), U - X^k \rangle + \frac{\mu_k}{2} \|U - X^k\|_F^2 - \Psi(X^k) \\
&\quad + \frac{\alpha \|Y^k\|^2 - \mu_k}{2} \|U - X^k\|_F^2 \\
&\leqslant \frac{\alpha \|Y^k\|^2 - \mu_k}{2} \|U - X^k\|_F^2,
\end{aligned}
$$

where the first inequality follows from the fact that $\nabla_X \mathcal{H}_\alpha(X, Y, Z)$ is Lipschitz with modulus $\alpha\|Y\|^2$ and the last inequality follows from the definition of U as a minimizer of (4.11b).

- Hierarchical-prox: in this case, for any $1 \leqslant i \leqslant r$, we have

$$
\begin{aligned}
&\Theta_{\alpha,\beta}(\boldsymbol{u}_{j<i}, \boldsymbol{u}_i, \boldsymbol{x}_{j>i}^k, Y^k, Z^k) - \Theta_{\alpha,\beta}(\boldsymbol{u}_{j<i}, \boldsymbol{x}_i^k, \boldsymbol{x}_{j>i}^k, Y^k, Z^k) \\
&= \psi_i(\boldsymbol{u}_i) + \mathcal{H}_\alpha(\boldsymbol{u}_{j<i}, \boldsymbol{u}_i, \boldsymbol{x}_{j>i}^k, Y^k, Z^k) - \psi_i(\boldsymbol{x}_i^k) - \mathcal{H}_\alpha(\boldsymbol{u}_{j<i}, \boldsymbol{x}_i^k, \boldsymbol{x}_{j>i}^k, Y^k, Z^k) \\
&= \left[\psi_i(\boldsymbol{u}_i) + \mathcal{H}_\alpha(\boldsymbol{u}_{j<i}, \boldsymbol{u}_i, \boldsymbol{x}_{j>i}^k, Y^k, Z^k) + \frac{\mu_k}{2}\|\boldsymbol{u}_i - \boldsymbol{x}_i^k\|^2\right] - \frac{\mu_k}{2}\|\boldsymbol{u}_i - \boldsymbol{x}_i^k\|^2 \\
&\quad - \left[\psi_i(\boldsymbol{x}_i^k) + \mathcal{H}_\alpha(\boldsymbol{u}_{j<i}, \boldsymbol{x}_i^k, \boldsymbol{x}_{j>i}^k, Y^k, Z^k)\right] \\
&\leqslant -\frac{\mu_k}{2}\|\boldsymbol{u}_i - \boldsymbol{x}_i^k\|^2,
\end{aligned}
$$

where the inequality follows from the definition of $\boldsymbol{u}_i$ as a minimizer of (4.11c). Then, summing the above relation from $i = r$ to $i = 1$ and simplifying the resulting inequality, we obtain (4.27).

Now, summing (4.25), (4.26) and (4.27), and using $\mathcal{F}(U,V) = \Theta_{\alpha,\beta}(U,V,W)$ and $\mathcal{F}(X^k,Y^k) = \Theta_{\alpha,\beta}(X^k,\,Y^k,Z^k)$, we obtain (4.21). This completes the proof. $\square$

From the above lemma, we see that the sufficient descent of $\mathcal{F}(X,Y)$ can be guaranteed as long as μ_k and σ_k are sufficiently large. Thus, based on this lemma, we can show in the following proposition that our line search criterion (4.16) in Algorithm 2 is well defined.

Proposition 4.3 (**Well-definedness of the line search criterion**). *Suppose that Assumption 4.1 holds. Let (X^k,Y^k) be the sequence generated by Algorithm 2. Then, for each $k \geqslant 0$, the line search criterion (4.16) is satisfied after finitely many inner iterations.*

Proof. We prove this proposition by contradiction. Assume that there exists a $k \geqslant 0$ such that the line search criterion (4.16) cannot be satisfied after finitely many inner iterations. Note from (2a) and (2d) in Step 2 of Algorithm 2 that $\mu_k \leqslant \mu_k^{\max} = (\alpha+2\gamma\rho)\|Y^k\|^2 + c$ and hence $\mu_k = \mu_k^{\max}$ must be satisfied after finitely many inner iterations. Let n_k denote the number of inner iterations when $\mu_k = \mu_k^{\max}$ is satisfied for the *first* time. If $\mu_k^0 \geqslant \mu_k^{\max}$, then $n_k = 1$; otherwise, we have

$$\mu^{\min}\tau^{n_k-2} \leqslant \mu_k^0\tau^{n_k-2} < \mu_k^{\max},$$

which implies that

$$n_k \leqslant \left\lfloor \frac{\log(\mu_k^{\max}) - \log(\mu^{\min})}{\log\tau} + 2 \right\rfloor. \tag{4.28}$$

Then, from (2d) in Step 2 of Algorithm 2, we have $U \equiv U_{\mu_k^{\max}}$ and $\sigma_k^{\max} = (\alpha + 2\gamma\rho)\|U_{\mu_k^{\max}}\|^2 + c$ after at most n_k+1 inner iterations, where $U_{\mu_k^{\max}}$ is computed by (4.11a), (4.11b) or (4.11c) with $\mu_k = \mu_k^{\max}$. Moreover, we see that $\sigma_k = \sigma_k^{\max}$ must be satisfied after finitely many inner iterations. Similarly, let $\hat{n}_k$ denote the number of inner iterations when $\sigma_k = \sigma_k^{\max}$ is satisfied for the *first* time. If $\sigma_k^0 > \sigma_k^{\max}$, then $\hat{n}_k = n_k$; if $\sigma_k^0 = \sigma_k^{\max}$, then $\hat{n}_k = 0$; otherwise, we have

$$\sigma^{\min}\tau^{\hat{n}_k-1} \leqslant \sigma_k^0\tau^{\hat{n}_k-1} < \sigma_k^{\max},$$

which implies that

$$\hat{n}_k \leqslant \left\lfloor \frac{\log(\sigma_k^{\max}) - \log(\sigma^{\min})}{\log\tau} + 1 \right\rfloor.$$

Thus, after at most $\max\{n_k,\, \hat{n}_k\}+1$ inner iterations, we must have $V \equiv V_{\sigma_k^{\max}}$, where $V_{\sigma_k^{\max}}$ is computed by (4.12a), (4.12b) or (4.12c) with $\sigma_k = \sigma_k^{\max}$. Therefore, after at most $\max\{n_k,\, \hat{n}_k\} + 1$ inner iterations, we have

$$\begin{aligned}
&\mathcal{F}(U_{\mu_k^{\max}}, V_{\sigma_k^{\max}}) - \mathcal{F}(X^k, Y^k) \\
&\leqslant -\frac{\mu_k^{\max} - (\alpha+2\gamma\rho)\|Y^k\|^2}{2}\cdot\|U_{\mu_k^{\max}} - X^k\|_F^2 - \frac{\sigma_k^{\max} - (\alpha+2\gamma\rho)\|U_{\mu_k^{\max}}\|^2}{2}\cdot\|V_{\sigma_k^{\max}} - Y^k\|_F^2 \\
&= -\frac{c}{2}\left(\|U_{\mu_k^{\max}} - X^k\|_F^2 + \|V_{\sigma_k^{\max}} - Y^k\|_F^2\right),
\end{aligned}$$

where the inequality follows from (4.21) and the equality follows from $\mu_k^{\max} = (\alpha + 2\gamma\rho)\|Y^k\|^2 + c$ and $\sigma_k^{\max} = (\alpha+2\gamma\rho)\|U_{\mu_k^{\max}}\|^2 + c$. This together with

$$\mathcal{F}(X^k, Y^k) \leqslant \max_{[k-N]_+ \leqslant i \leqslant k} \mathcal{F}(X^i, Y^i)$$

implies that (4.16) must be satisfied after at most $\max\{n_k,\, \hat{n}_k\} + 1$ inner iterations, which leads to a contradiction. $\square$

Now, we are ready to prove our main convergence result, which characterizes a cluster point of the sequence generated by Algorithm 2. Our proof of statement (ii) in the following theorem is similar to that of [88, Lemma 4]. However, the arguments involved are more intricate since we have two blocks of variables in our line search loop.

Theorem 4.3. *Suppose that Assumption 4.1 holds. Let $\{(X^k, Y^k)\}$ be the sequence generated by Algorithm 2. Then,*

(i) ***(boundedness of sequence)*** *$\{(X^k, Y^k)\}$, $\{\bar{\mu}_k\}$ and $\{\bar{\sigma}_k\}$ are bounded;*

(ii) ***(diminishing successive changes)*** *$\lim_{k\to\infty} \|X^{k+1} - X^k\|_F + \|Y^{k+1} - Y^k\|_F = 0$;*

(iii) ***(global subsequential convergence)*** *any cluster point (X^*, Y^*) of $\{(X^k, Y^k)\}$ is a stationary point of $\mathcal{F}$.*

Proof. *Proof of Statement (i).* We first show that

$$\mathcal{F}(X^k, Y^k) \leqslant \mathcal{F}(X^0, Y^0) \tag{4.29}$$

for all $k \geqslant 1$. We will prove it by induction. Indeed, for $k = 1$, it follows from Proposition 4.3 that

$$\mathcal{F}(X^1, Y^1) - \mathcal{F}(X^0, Y^0) \leqslant -\frac{c}{2}\left(\|X^1 - X^0\|_F^2 + \|Y^1 - Y^0\|_F^2\right) \leqslant 0$$

is satisfied after finitely many inner iterations. Hence, (4.29) holds for $k = 1$. We now suppose that (4.29) holds for all $k \leqslant K$ for some integer $K \geqslant 1$. Then, we only need to show that (4.29) also holds for $k = K+1$. For $k = K+1$, we have

$$\begin{aligned} \mathcal{F}(X^{K+1}, Y^{K+1}) - \mathcal{F}(X^0, Y^0) &\leqslant \mathcal{F}(X^{K+1}, Y^{K+1}) - \max_{[K-N]_+ \leqslant i \leqslant K} \mathcal{F}(X^k, Y^k) \\ &\leqslant -\frac{c}{2}\left(\|X^{K+1} - X^K\|_F^2 + \|Y^{K+1} - Y^K\|_F^2\right) \leqslant 0, \end{aligned}$$

where the first inequality follows from the induction hypothesis and the second inequality follows from (4.16). Hence, (4.29) holds for $k = K+1$. This completes the induction. Then, from (4.29), we have that for any $k \geqslant 0$,

$$\mathcal{F}(X^0, Y^0) \geqslant \mathcal{F}(X^k, Y^k) = \Psi(X^k) + \Phi(Y^k) + \frac{1}{2}\left\|\mathcal{A}(X^k(Y^k)^\top) - \boldsymbol{b}\right\|^2,$$

which, together with (a1) in Assumption 4.1, implies that the sequences $\{X^k\}$, $\{Y^k\}$ and $\{\|\mathcal{A}(X^k(Y^k)^\top)-\boldsymbol{b}\|\}$ are bounded. Moreover, from Step 2 and Step 3 in Algorithm 2, it is easy to see $\bar{\mu}_k \leqslant \mu_k^{\max} = (\alpha + 2\gamma\rho)\|Y^k\|^2 + c$ for all k. Since $\{Y^k\}$ is bounded, the sequences $\{\mu_k^{\max}\}$ and $\{\bar{\mu}_k\}$ are bounded. Next, we prove the boundedness of $\{\bar{\sigma}_k\}$. Indeed, at the k-th iteration, there are three possibilities:

- $\bar{\mu}_k < \mu_k^{\max}$: In this case, we have $\bar{\sigma}_k \leqslant \sigma_k^0 \tau^{\tilde{n}_k} \leqslant \sigma^{\max}\tau^{\tilde{n}_k}$, where $\tilde{n}_k$ denotes the number of inner iterations for the line search at the k-th iteration and $\tilde{n}_k \leqslant \max\left\{1, \left\lfloor \frac{\log(\mu_k^{\max})-\log(\mu^{\min})}{\log\tau} + 2\right\rfloor\right\}$ (see (4.28) and the discussions preceding it).

- $\bar{\mu}_k = \mu_k^{\max}$ and $\bar{\sigma}_k > \sigma_k^{\max}$: In this case, we have $\bar{\sigma}_k \leqslant \sigma_k^0 \tau^{\tilde{n}_k} \leqslant \sigma^{\max}\tau^{\tilde{n}_k}$, where $\tilde{n}_k \leqslant \max\left\{1, \left\lfloor \frac{\log(\mu_k^{\max})-\log(\mu^{\min})}{\log\tau} + 2\right\rfloor\right\}$.

- Otherwise, we have $\bar{\sigma}_k \leqslant \sigma_k^{\max} = (\alpha + 2\gamma\rho)\|X^{k+1}\|^2 + c$.

Note that $\{\tilde{n}_k\}$ is bounded as $\{\mu_k^{\max}\}$ is bounded. Thus, $\{\bar{\sigma}_k\}$ is bounded as the sequences $\{X^k\}$ and $\{\tilde{n}_k\}$ are bounded. This proves statement (i).

Proof of Statement (ii). For notational simplicity, from now on, we let $\Delta_{X^k} := X^{k+1} - X^k$, $\Delta_{Y^k} := Y^{k+1} - Y^k$, $\Delta_{Z^k} := Z^{k+1} - Z^k$ and

$$\ell(k) = \arg\max_i \{\, \mathcal{F}(X^i, Y^i) \,:\, i = [k-N]_+, \cdots, k \,\}. \tag{4.30}$$

Then, the line search criterion (4.16) can be rewritten as

$$\mathcal{F}(X^{k+1}, Y^{k+1}) - \mathcal{F}(X^{\ell(k)}, Y^{\ell(k)}) \leqslant -\frac{c}{2}\left(\|\Delta_{X^k}\|_F^2 + \|\Delta_{Y^k}\|_F^2\right) \leqslant 0. \tag{4.31}$$

Observe that

$$\begin{aligned}
\mathcal{F}(X^{\ell(k+1)}, Y^{\ell(k+1)}) &= \max_{[k+1-N]_+ \leqslant i \leqslant k+1} \mathcal{F}(X^i, Y^i) \\
&= \max\left\{\mathcal{F}(X^{k+1}, Y^{k+1}), \max_{[k+1-N]_+ \leqslant i \leqslant k} \mathcal{F}(X^i, Y^i)\right\} \\
&\overset{\text{(i)}}{\leqslant} \max\left\{\mathcal{F}(X^{\ell(k)}, Y^{\ell(k)}), \max_{[k+1-N]_+ \leqslant i \leqslant k} \mathcal{F}(X^i, Y^i)\right\} \\
&\leqslant \max\left\{\mathcal{F}(X^{\ell(k)}, Y^{\ell(k)}), \max_{[k-N]_+ \leqslant i \leqslant k} \mathcal{F}(X^i, Y^i)\right\} \\
&\overset{\text{(ii)}}{=} \max\left\{\mathcal{F}(X^{\ell(k)}, Y^{\ell(k)}), \mathcal{F}(X^{\ell(k)}, Y^{\ell(k)})\right\} \\
&= \mathcal{F}(X^{\ell(k)}, Y^{\ell(k)}),
\end{aligned}$$

where (i) follows from (4.31) and (ii) follows from (4.30). Therefore, the sequence $\{\mathcal{F}(X^{\ell(k)}, Y^{\ell(k)})\}$ is non-increasing. Since $\mathcal{F}(X^{\ell(k)}, Y^{\ell(k)})$ is also bounded from below (due to (a1) in Assumption 4.1), we conclude that there exists a number $\widetilde{\mathcal{F}}$ such that

$$\lim_{k\to\infty} \mathcal{F}(X^{\ell(k)}, Y^{\ell(k)}) = \widetilde{\mathcal{F}}. \tag{4.32}$$

We next prove by induction that for all $j \geqslant 1$,

$$\begin{cases}
\lim_{k\to\infty} \Delta_{X^{\ell(k)-j}} = \lim_{k\to\infty} \Delta_{Y^{\ell(k)-j}} = 0, & \text{(4.33a)} \\
\lim_{k\to\infty} \mathcal{F}(X^{\ell(k)-j}, Y^{\ell(k)-j}) = \widetilde{\mathcal{F}}. & \text{(4.33b)}
\end{cases}$$

We first prove (4.33a) and (4.33b) for $j = 1$. Applying (4.31) with k replaced by $\ell(k) - 1$, we obtain

$$\mathcal{F}(X^{\ell(k)}, Y^{\ell(k)}) - \mathcal{F}(X^{\ell(\ell(k)-1)}, Y^{\ell(\ell(k)-1)}) \leqslant -\frac{c}{2}\left(\|\Delta_{X^{\ell(k)-1}}\|_F^2 + \|\Delta_{Y^{\ell(k)-1}}\|_F^2\right).$$

Thus, from this and (4.32), we have

$$\lim_{k\to\infty} \frac{c}{2}\left(\|\Delta_{X^{\ell(k)-1}}\|_F^2 + \|\Delta_{Y^{\ell(k)-1}}\|_F^2\right) = 0,$$

which implies that

$$\lim_{k\to\infty} \Delta_{X^{\ell(k)-1}} = \lim_{k\to\infty} \Delta_{Y^{\ell(k)-1}} = 0. \tag{4.34}$$

Then, from (4.32) and (4.34), we have

$$\begin{aligned} \widetilde{\mathcal{F}} &= \lim_{k\to\infty} \mathcal{F}(X^{\ell(k)}, Y^{\ell(k)}) \\ &= \lim_{k\to\infty} \mathcal{F}(X^{\ell(k)-1} + \Delta_{X^{\ell(k)-1}}, Y^{\ell(k)-1} + \Delta_{Y^{\ell(k)-1}}) \\ &= \lim_{k\to\infty} \mathcal{F}(X^{\ell(k)-1}, Y^{\ell(k)-1}), \end{aligned}$$

where the last equality follows because $\{(X^k, Y^k)\}$ is bounded and $\mathcal{F}$ is uniformly continuous on any compact subset of $\mathrm{dom}\mathcal{F}$ under (a1) in Assumption 4.1. Thus, (4.33a) and (4.33b) hold for $j = 1$.

We next suppose that (4.33a) and (4.33b) hold for $j = J$ for some $J \geqslant 1$. It remains to show that they also hold for $j = J + 1$. Indeed, from (4.31) with k replaced by $\ell(k) - J - 1$ (here, without loss of generality, we assume that k is large enough such that $\ell(k) - J - 1$ is nonnegative), we have

$$\begin{aligned} &\mathcal{F}(X^{\ell(k)-J}, Y^{\ell(k)-J}) - \mathcal{F}(X^{\ell(\ell(k)-J-1)}, Y^{\ell(\ell(k)-J-1)}) \\ \leqslant &-\frac{c}{2}\left(\|\Delta_{X^{\ell(k)-J-1}}\|_F^2 + \|\Delta_{Y^{\ell(k)-J-1}}\|_F^2\right), \end{aligned}$$

which implies that

$$\begin{aligned} &\|\Delta_{X^{\ell(k)-J-1}}\|_F^2 + \|\Delta_{Y^{\ell(k)-J-1}}\|_F^2 \\ \leqslant &\frac{2}{c}\left(\mathcal{F}(X^{\ell(\ell(k)-J-1)}, Y^{\ell(\ell(k)-J-1)}) - \mathcal{F}(X^{\ell(k)-J}, Y^{\ell(k)-J})\right). \end{aligned}$$

This together with (4.32) and the induction hypothesis implies that

$$\lim_{k\to\infty} \Delta_{X^{\ell(k)-(J+1)}} = \lim_{k\to\infty} \Delta_{Y^{\ell(k)-(J+1)}} = 0.$$

Thus, (4.33a) holds for $j = J+1$. From this, we further have

$$\begin{aligned}
&\lim_{k\to\infty} \mathcal{F}(X^{\ell(k)-(J+1)},\, Y^{\ell(k)-(J+1)}) \\
&= \lim_{k\to\infty} \mathcal{F}(X^{\ell(k)-J} - \Delta_{X^{\ell(k)-(J+1)}},\, Y^{\ell(k)-J} - \Delta_{Y^{\ell(k)-(J+1)}}) \\
&= \lim_{k\to\infty} \mathcal{F}(X^{\ell(k)-J},\, Y^{\ell(k)-J}) = \widetilde{\mathcal{F}},
\end{aligned}$$

where the second equality follows because $\{(X^k, Y^k)\}$ is bounded and $\mathcal{F}$ is uniformly continuous on any compact subset of $\mathrm{dom}\mathcal{F}$ under (a1) in Assumption 4.1. Hence, (4.33b) also holds for $j = J+1$. This completes the induction.

We are now ready to prove the main result in this statement. Indeed, from (4.30), we can see $k - N \leqslant \ell(k) \leqslant k$ (without loss of generality, we assume that k is large enough such that $k \geqslant N$). Thus, for any k, we must have $k - N - 1 = \ell(k) - j_k$ for $1 \leqslant j_k \leqslant N+1$. Then, we have

$$\begin{aligned}
\|\Delta_{X^{k-N-1}}\|_F &= \|\Delta_{X^{\ell(k)-j_k}}\|_F \leqslant \max_{1\leqslant j\leqslant N+1} \|\Delta_{X^{\ell(k)-j}}\|_F, \\
\|\Delta_{Y^{k-N-1}}\|_F &= \|\Delta_{Y^{\ell(k)-j_k}}\|_F \leqslant \max_{1\leqslant j\leqslant N+1} \|\Delta_{Y^{\ell(k)-j}}\|_F.
\end{aligned}$$

This together with (4.33a) implies that

$$\begin{aligned}
\lim_{k\to\infty} \Delta_{X^k} &= \lim_{k\to\infty} \Delta_{X^{k-N-1}} = 0, \\
\lim_{k\to\infty} \Delta_{Y^k} &= \lim_{k\to\infty} \Delta_{Y^{k-N-1}} = 0.
\end{aligned}$$

This proves the statement (ii).

Proof of Statement (iii). First, since $\{(X^k, Y^k)\}$ is bounded from statement (i), there exists at least one cluster point. Suppose that (X^*, Y^*) is a cluster point of $\{(X^k, Y^k)\}$ and let $\{(X^{k_i}, Y^{k_i})\}$ be a convergent subsequence such that $\lim_{i\to\infty}(X^{k_i}, Y^{k_i}) = (X^*, Y^*)$. Then, it is easy to see from (4.15) that $\lim_{i\to\infty} Z^{k_i} = Z^*$, where Z^* is given by (4.5). Thus, it can be shown as in (4.10) that

$$\alpha(Z^* - X^*(Y^*)^\top) + \beta\mathcal{A}^*(\mathcal{A}(Z^*) - \boldsymbol{b}) = 0. \tag{4.35}$$

We next show that

$$\begin{cases} 0 \in \partial \Psi(X^*) + \alpha(X^*(Y^*)^\top - Z^*)Y^*, & (4.36a) \\ 0 \in \partial \Phi(Y^*) + \alpha(X^*(Y^*)^\top - Z^*)^\top X^*. & (4.36b) \end{cases}$$

We start by showing (4.36a) in the following cases:

- Proximal & Prox-linear: in these two cases, passing to the limit along $\{(X^{k_i}, Y^{k_i})\}$ in (4.19a) or (4.19b) with X^{k_i+1} in place of U and $\bar{\mu}_{k_i}$ in place of μ_k, and invoking (a1) in Assumption 4.1, statements (i), (ii) and (2.1), we obtain (4.36a).

- Hierarchical-prox: in this case, passing to the limit along $\{(X^{k_i}, Y^{k_i})\}$ in (4.19c) with X^{k_i+1} in place of U and $\bar{\mu}_{k_i}$ in place of μ_k, and invoking (a1) in Assumption 4.1, statements (i), (ii) and (2.1), we have

$$0 \in \partial \psi_i(\boldsymbol{x}_i^*) + \alpha(X^*(Y^*)^\top - Z^*)\boldsymbol{y}_i^*$$

 for any $i = 1, 2, \cdots, r$. Then, rearranging the above relations, we obtain (4.36a).

Similarly, we can obtain (4.36b). Thus, combining (4.35), (4.36a) and (4.36b), we see that (X^*, Y^*, Z^*) is a stationary point of $\Theta_{\alpha,\beta}$, which further implies (X^*, Y^*) is a stationary point of $\mathcal{F}$ from Theorem 4.2. This proves statement (iii). $\square$

Remark 4.6 (**Comment on (a3) in Assumption 4.1**). *If Φ and Ψ are the indicator functions of some nonempty closed sets, the results in Theorem 4.3 remain valid under the weaker condition on α and β that $\frac{1}{\alpha} + \frac{1}{\beta} > 0$ with a slight modification in (4.16) of Algorithm 2. Indeed, when Φ and Ψ are the indicator functions, one can see from Remark 4.1 and the proofs of Lemma 4.2 and Proposition 4.3 that if $\frac{1}{\alpha} + \frac{1}{\beta} > 0$, then*

$$\mathcal{F}(U, V) - \mathcal{F}(X^k, Y^k) = \left(\tfrac{1}{\alpha} + \tfrac{1}{\beta}\right)\left(\Theta_{\alpha,\beta}(U, V, W) - \Theta_{\alpha,\beta}(X^k, Y^k, Z^k)\right)$$

$$\leqslant -\left(\tfrac{1}{\alpha} + \tfrac{1}{\beta}\right)\left(\tfrac{\mu_k - (\alpha + 2\gamma\rho)\|Y^k\|^2}{2} \cdot \|U - X^k\|_F^2 + \tfrac{\sigma_k - (\alpha + 2\gamma\rho)\|U\|^2}{2} \cdot \|V - Y^k\|_F^2\right),$$

and the line search criterion is well defined with c replaced by $\left(\frac{1}{\alpha}+\frac{1}{\beta}\right)c$. *Moreover, recalling [71, Exercise 8.14], we see that* $\partial\Psi$ *and* $\partial\Phi$ *are normal cones. Thus, following Remark 4.2 and the similar augments in Theorem 4.3, we can obtain the same results when* $\frac{1}{\alpha}+\frac{1}{\beta}>0$ *with c replaced by* $\left(\frac{1}{\alpha}+\frac{1}{\beta}\right)c$ *in* (4.16) *of Algorithm 2.*

Remark 4.7 (**Comments on updating** $\mu_k^{\max}$ **and** $\sigma_k^{\max}$). *In Algorithm 2, we need to evaluate* $\mu_k^{\max}=(\alpha+2\gamma\rho)\|Y^k\|^2+c$ *and* $\sigma_k^{\max}=(\alpha+2\gamma\rho)\|U\|^2+c$ *in each iteration. However, computing the spectral norms of* Y^k *and* U *might be costly, especially when* r *is large. Hence, in our experiments, instead of computing* $\|Y^k\|^2$ *and* $\|U\|^2$*, we compute* $\|Y^k\|_F^2$ *and* $\|U\|_F^2$*, and update* $\mu_k^{\max}$ *and* $\sigma_k^{\max}$ *by* $\mu_k^{\max}=(\alpha+2\gamma\rho)\|Y^k\|_F^2+c$ *and* $\sigma_k^{\max}=(\alpha+2\gamma\rho)\|U\|_F^2+c$ *instead. Since* $\|Y^k\|\leqslant\|Y^k\|_F$ *and* $\|U\|\leqslant\|U\|_F$*, it follows from* (4.21) *that*

$$\begin{aligned}&\mathcal{F}(U,V)-\mathcal{F}(X^k,Y^k)\\ \leqslant&-\frac{\mu_k-(\alpha+2\gamma\rho)\|Y^k\|_F^2}{2}\cdot\|U-X^k\|_F^2-\frac{\sigma_k-(\alpha+2\gamma\rho)\|U\|_F^2}{2}\cdot\|V-Y^k\|_F^2.\end{aligned}$$

Then, one can show that Proposition 4.3 and Theorem 4.3 remain valid. Additionally, the quantities $\|U\|_F^2$ *and* $\|Y^k\|_F^2$ *can also be reused in the computation of the successive changes* $\|U-X^k\|_F^2$ *and* $\|V-Y^k\|_F^2$ *to reduce the cost of line search.*

Next, under the additional assumption that the objective function $\mathcal{F}$ in (1.4) is a KL function with an exponent ϑ, we can discuss the local convergence rate of Algorithm 2 with respect to the function value. To this end, we first give the following supporting lemma.

Lemma 4.3. *Suppose that Assumption 4.1 holds. Let* $\{(X^k,Y^k)\}$ *be the sequence generated by Algorithm 2. Then, there exists* $d_1>0$ *such that*

$$\operatorname{dist}\left(0,\,\partial\mathcal{F}(X^k,Y^k)\right)\leqslant d_1\left(\|X^k-X^{k-1}\|_F+\|Y^k-Y^{k-1}\|_F\right). \tag{4.37}$$

Proof. We first let $\{Z^k\}$ be the sequence generated by (4.15). Thus, it is easy to see

$$\mathcal{A}^*(\mathcal{A}(X^k(Y^k)^\top) - \boldsymbol{b}) = \frac{\alpha+\beta}{\beta}\left(X^k(Y^k)^\top - Z^k\right) = \alpha\left(X^k(Y^k)^\top - Z^k\right), \quad (4.38)$$

where the last equality follows from $\frac{1}{\alpha} + \frac{1}{\beta} = 1$. Then, we consider the subdifferential $\partial\mathcal{F}(X^k, Y^k)$.

For the partial subdifferential with respect to X or $\boldsymbol{x}_i$, we consider the following three cases:

- Proximal: in this case, we have

$$\begin{aligned}
&\partial_X \mathcal{F}(X^k, Y^k) \\
&= \partial\Psi(X^k) + \mathcal{A}^*(\mathcal{A}(X^k(Y^k)^\top) - \boldsymbol{b})Y^k \\
&= \partial\Psi(X^k) + \alpha\left(X^k(Y^k)^\top - Z^k\right)Y^k \\
&= \partial\Psi(X^k) + \alpha\left(X^k(Y^{k-1})^\top - Z^{k-1}\right)Y^{k-1} + \bar{\mu}_{k-1}(X^k - X^{k-1}) - \bar{\mu}_{k-1}(X^k - X^{k-1}) \\
&\quad + \alpha(X^k(Y^k)^\top Y^k - X^k(Y^{k-1})^\top Y^{k-1}) - \alpha(Z^kY^k - Z^{k-1}Y^{k-1}) \\
&\ni -\bar{\mu}_{k-1}(X^k - X^{k-1}) + \alpha X^k(Y^k - Y^{k-1})^\top Y^k + \alpha X^k(Y^{k-1})^\top(Y^k - Y^{k-1}) \\
&\quad - \alpha(Z^k - Z^{k-1})Y^k - \alpha Z^{k-1}(Y^k - Y^{k-1}),
\end{aligned}$$

  where the second equality follows from (4.38) and the inclusion follows from (4.19a).

- Prox-linear: in this case, we have

$$
\begin{aligned}
&\partial_X \mathcal{F}(X^k, Y^k) \\
&= \partial\Psi(X^k) + \mathcal{A}^*(\mathcal{A}(X^k(Y^k)^\top) - \boldsymbol{b})Y^k \\
&= \partial\Psi(X^k) + \alpha\left(X^k(Y^k)^\top - Z^k\right)Y^k \\
&= \partial\Psi(X^k) + \alpha\left(X^{k-1}(Y^{k-1})^\top - Z^{k-1}\right)Y^{k-1} + \bar{\mu}_{k-1}(X^k - X^{k-1}) - \bar{\mu}_{k-1}(X^k - X^{k-1}) \\
&\quad + \alpha(X^k(Y^k)^\top Y^k - X^{k-1}(Y^{k-1})^\top Y^{k-1}) - \alpha(Z^kY^k - Z^{k-1}Y^{k-1}) \\
&\ni -\bar{\mu}_{k-1}(X^k - X^{k-1}) + \alpha X^k(Y^k - Y^{k-1})^\top Y^k + \alpha X^k(Y^{k-1})^\top(Y^k - Y^{k-1}) \\
&\quad + (X^k - X^{k-1})(Y^{k-1})^\top Y^{k-1} - \alpha(Z^k - Z^{k-1})Y^k - \alpha Z^{k-1}(Y^k - Y^{k-1}),
\end{aligned}
$$

  where the second equality follows from (4.38) and the inclusion follows from (4.19b).

- Hierarchial-prox: in this case, for any $i = 1, 2, \cdots, r$, we have

$$
\begin{aligned}
&\partial_{\boldsymbol{x}_i} \mathcal{F}(X^k, Y^k) \\
&= \partial\psi_i(\boldsymbol{x}_i^k) + \mathcal{A}^*(\mathcal{A}(X^k(Y^k)^\top) - \boldsymbol{b})\boldsymbol{y}_i^k \\
&= \partial\psi_i(\boldsymbol{x}_i^k) + \alpha\left(X^k(Y^k)^\top - Z^k\right)\boldsymbol{y}_i^k \\
&= \partial\psi_i(\boldsymbol{x}_i^k) + \alpha\left(\textstyle\sum_{j=1}^{i}\boldsymbol{x}_j^k(\boldsymbol{y}_j^{k-1})^\top + \sum_{j=i+1}^{r}\boldsymbol{x}_j^{k-1}(\boldsymbol{y}_j^{k-1})^\top - Z^{k-1}\right)\boldsymbol{y}_i^{k-1} + \bar{\mu}_{k-1}\left(\boldsymbol{x}_i^k - \boldsymbol{x}_i^{k-1}\right) \\
&\quad + \alpha\textstyle\sum_{j=1}^{i}\left(\boldsymbol{x}_j^{k-1} - \boldsymbol{x}_j^k\right)(\boldsymbol{y}_j^{k-1})^\top\boldsymbol{y}_i^{k-1} + \alpha\left(X^k(Y^k)^\top\boldsymbol{y}_i^k - X^{k-1}(Y^{k-1})^\top\boldsymbol{y}_i^{k-1}\right) \\
&\quad - \alpha\left(Z^k\boldsymbol{y}_i^k - Z^{k-1}\boldsymbol{y}_i^{k-1}\right) - \bar{\mu}_{k-1}\left(\boldsymbol{x}_i^k - \boldsymbol{x}_i^{k-1}\right) \\
&\ni \alpha\textstyle\sum_{j=1}^{i}\left(\boldsymbol{x}_j^{k-1} - \boldsymbol{x}_j^k\right)(\boldsymbol{y}_j^{k-1})^\top\boldsymbol{y}_i^{k-1} + \alpha\left(X^k - X^{k-1}\right)(Y^k)^\top\boldsymbol{y}_i^k \\
&\quad + \alpha X^{k-1}\left(Y^k - Y^{k-1}\right)^\top\boldsymbol{y}_i^k + \alpha X^{k-1}(Y^{k-1})^\top\left(\boldsymbol{y}_i^k - \boldsymbol{y}_i^{k-1}\right) \\
&\quad - \alpha\left(Z^k - Z^{k-1}\right)\boldsymbol{y}_i^k - \alpha Z^{k-1}\left(\boldsymbol{y}_i^k - \boldsymbol{y}_i^{k-1}\right) - \bar{\mu}_{k-1}\left(\boldsymbol{x}_i^k - \boldsymbol{x}_i^{k-1}\right),
\end{aligned}
$$

  where the second equality follows from (4.38) and the inclusion follows from (4.19c).

Similarly, for the partial subdifferential with respect to Y or $\boldsymbol{y}_i$, we also consider the following three cases:

- Proximal: in this case, we have

$$\begin{aligned}\partial_Y \mathcal{F}(X^k, Y^k) &= \partial\Phi(Y^k) + \left(\mathcal{A}^*(\mathcal{A}(X^k(Y^k)^\top) - \boldsymbol{b})\right)^\top X^k \\ &= \partial\Phi(Y^k) + \alpha\left(X^k(Y^k)^\top - Z^k\right)^\top X^k \\ &= \partial\Phi(Y^k) + \alpha\left(X^k(Y^k)^\top - Z^{k-1}\right)^\top X^k + \bar{\sigma}_{k-1}(Y^k - Y^{k-1}) \\ &\quad - \bar{\sigma}_{k-1}(Y^k - Y^{k-1}) - \alpha\left(Z^k - Z^{k-1}\right)^\top X^k \\ &\ni -\bar{\sigma}_{k-1}(Y^k - Y^{k-1}) - \alpha\left(Z^k - Z^{k-1}\right)^\top X^k,\end{aligned}$$

  where the second equality follows from (4.38) and the inclusion follows from (4.20a).

- Prox-linear: in this case, we have

$$\begin{aligned}&\partial_Y \mathcal{F}(X^k, Y^k) \\ &= \partial\Phi(Y^k) + \left(\mathcal{A}^*(\mathcal{A}(X^k(Y^k)^\top) - \boldsymbol{b})\right)^\top X^k \\ &= \partial\Phi(Y^k) + \alpha\left(X^k(Y^k)^\top - Z^k\right)^\top X^k \\ &= \partial\Phi(Y^k) + \alpha\left(X^k(Y^{k-1})^\top - Z^{k-1}\right)^\top X^k + \bar{\sigma}_{k-1}(Y^k - Y^{k-1}) \\ &\quad - \bar{\sigma}_{k-1}(Y^k - Y^{k-1}) + \alpha(Y^k - Y^{k-1})(X^k)^\top X^k - \alpha\left(Z^k - Z^{k-1}\right)^\top X^k \\ &\ni -\bar{\sigma}_{k-1}(Y^k - Y^{k-1}) + \alpha(Y^k - Y^{k-1})(X^k)^\top X^k - \alpha\left(Z^k - Z^{k-1}\right)^\top X^k,\end{aligned}$$

  where the second equality follows from (4.38) and the inclusion follows from (4.20b).

- Hierarchial-prox: in this case, for any $i = 1, 2, \cdots, r$, we have

$$
\begin{aligned}
&\partial_{\boldsymbol{y}_i} \Theta_{\alpha,\beta}(X^k, Y^k, Z^k) \\
&= \partial \phi_i(\boldsymbol{y}_i^k) + \left(\mathcal{A}^*(\mathcal{A}(X^k(Y^k)^\top) - \boldsymbol{b})\right)^\top \boldsymbol{x}_i^k \\
&= \partial \phi_i(\boldsymbol{y}_i^k) + \alpha \left(X^k(Y^k)^\top - Z^k\right)^\top \boldsymbol{x}_i^k \\
&= \partial \phi_i(\boldsymbol{y}_i^k) + \alpha \left(\textstyle\sum_{j=1}^{i} \boldsymbol{x}_j^k(\boldsymbol{y}_j^k)^\top + \sum_{j=i+1}^{r} \boldsymbol{x}_j^k(\boldsymbol{y}_j^{k-1})^\top - Z^{k-1}\right)^\top \boldsymbol{x}_i^k + \bar{\sigma}_{k-1}\left(\boldsymbol{y}_i^k - \boldsymbol{y}_i^{k-1}\right) \\
&\quad + \alpha \textstyle\sum_{j=1}^{i} \left(\boldsymbol{y}_j^{k-1} - \boldsymbol{y}_j^k\right)(\boldsymbol{x}_j^k)^\top \boldsymbol{x}_i^k + \alpha\left(Y^k - Y^{k-1}\right)(X^k)^\top \boldsymbol{x}_i^k \\
&\quad - \alpha\left(Z^k - Z^{k-1}\right)^\top \boldsymbol{x}_i^k - \bar{\sigma}_{k-1}\left(\boldsymbol{y}_i^k - \boldsymbol{y}_i^{k-1}\right) \\
&\ni \alpha \textstyle\sum_{j=1}^{i} \left(\boldsymbol{y}_j^{k-1} - \boldsymbol{y}_j^k\right)(\boldsymbol{x}_j^k)^\top \boldsymbol{x}_i^k + \alpha\left(Y^k - Y^{k-1}\right)(X^k)^\top \boldsymbol{x}_i^k \\
&\quad - \alpha\left(Z^k - Z^{k-1}\right)^\top \boldsymbol{x}_i^k - \bar{\sigma}_{k-1}\left(\boldsymbol{y}_i^k - \boldsymbol{y}_i^{k-1}\right),
\end{aligned}
$$

  where the second equality follows from (4.38) and the inclusion follows from (4.20c).

Note that, for any i, we have $\|\boldsymbol{x}_i^k - \boldsymbol{x}_i^{k-1}\| \leqslant \|X^k - X^{k-1}\|_F$ and $\|\boldsymbol{y}_i^k - \boldsymbol{y}_i^{k-1}\| \leqslant \|Y^k - Y^{k-1}\|_F$, and moreover,

$$
\begin{aligned}
\|Z^{k+1} - Z^k\|_F &\leqslant \left\|\left(\mathcal{I} - \tfrac{\beta}{\alpha+\beta}\mathcal{A}^*\mathcal{A}\right)\left(X^{k+1}(Y^{k+1})^\top - X^k(Y^k)^\top\right)\right\|_F \\
&\leqslant \left\|\mathcal{I} - \tfrac{\beta}{\alpha+\beta}\mathcal{A}^*\mathcal{A}\right\| \cdot \|X^{k+1}(Y^{k+1})^\top - X^k(Y^k)^\top\|_F \\
&= \sqrt{\rho}\left\|X^{k+1}\left(Y^{k+1} - Y^k\right)^\top + \left(X^{k+1} - X^k\right)(Y^k)^\top\right\|_F \\
&\leqslant \sqrt{\rho}\left\|X^{k+1}\left(Y^{k+1} - Y^k\right)^\top\right\|_F + \sqrt{\rho}\left\|\left(X^{k+1} - X^k\right)(Y^k)^\top\right\|_F \\
&\leqslant \sqrt{\rho}\|X^{k+1}\|\|Y^{k+1} - Y^k\|_F + \sqrt{\rho}\|Y^k\|\|X^{k+1} - X^k\|_F,
\end{aligned}
$$

where the equality follows from the definition of ρ (see (4.13)). Thus, from the above relations and the boundednesses of $\{(X^k, Y^k)\}$, $\{\bar{\mu}_k\}$ and $\{\bar{\sigma}_k\}$ (see statement (i) in Theorem 4.3), we can obtain (4.37). This completes the proof. $\square$

Then, we discuss the convergence rate of our NAUM in the following theorem.

Theorem 4.4. *Suppose that Assumption 4.1 holds and $\mathcal{F}$ in (1.4) is a KL function with an exponent ϑ, i.e., $\mathcal{F}$ satisfies KL property with $\varphi(s) = \tilde{a}s^{1-\vartheta}$ for some $\tilde{a} > 0$ and $\vartheta \in [0, 1)$. Let $\{(X^k, Y^k)\}_{k=0}^{\infty}$ be the sequence generated by Algorithm 2. Then, there exists a constant $\mathcal{F}^*$ such that the following statements hold.*

(i) *If $\vartheta = 0$, then $\mathcal{F}(X^k, Y^k) \leqslant \mathcal{F}^*$ for all large k;*

(ii) *If $\vartheta \in \left(0, \frac{1}{2}\right]$, then there exist $\varrho \in (0, 1)$ and $\zeta > 0$ such that*

$$\mathcal{F}(X^k, Y^k) - \mathcal{F}^* \leqslant \zeta \varrho^k$$

for all large k;

(iii) *If $\vartheta \in \left(\frac{1}{2}, 1\right)$, then there exists $\zeta > 0$ such that*

$$\mathcal{F}(X^k, Y^k) - \mathcal{F}^* \leqslant \zeta\, k^{-\frac{1}{2\vartheta - 1}}$$

for all large k.

Proof. We start by defining an index sequence $\{\xi(t)\}_{t=0}^{\infty}$ as follows:

$$\xi(t) = \ell((N + 1)t), \quad t = 0, 1, 2, \cdots,$$

where $\ell(k)$ is defined in (4.30). Since $k - N \leqslant \ell(k) \leqslant k$ for any $k \geqslant N$, we then have $\xi(t) = \ell((N + 1)t) \geqslant (N + 1)t - N = (N + 1)(t - 1) + 1 \geqslant \ell((N + 1)(t - 1)) + 1 > \xi(t - 1)$ for any $t \geqslant 1$. Thus, $\{\xi(t)\}_{t=0}^{\infty}$ is increasing. Note that $\{\mathcal{F}(X^{\xi(t)}, Y^{\xi(t)})\}_{t=0}^{\infty}$ is a subsequence of $\{\mathcal{F}(X^{\ell(k)}, Y^{\ell(k)})\}_{k=0}^{\infty}$. Then, this sequence is non-increasing and bounded from below, and there exists a number $\mathcal{F}^*$ such that

$$\lim_{t \to \infty} \mathcal{F}(X^{\xi(t)}, Y^{\xi(t)}) =: \mathcal{F}^*.$$

Moreover, it follows from (4.31) with k replaced by $\xi(t)-1$ that

$$\begin{aligned}
&\mathcal{F}(X^{\xi(t)}, Y^{\xi(t)}) \\
&\leqslant \mathcal{F}(X^{\ell(\xi(t)-1)}, Y^{\ell(\xi(t)-1)}) - \frac{c}{2}\left(\|\Delta_{X^{\xi(t)-1}}\|_F^2 + \|\Delta_{Y^{\xi(t)-1}}\|_F^2\right) \\
&\leqslant \mathcal{F}(X^{\ell((N+1)(t-1))}, Y^{\ell((N+1)(t-1))}) - \frac{c}{2}\left(\|\Delta_{X^{\xi(t)-1}}\|_F^2 + \|\Delta_{Y^{\xi(t)-1}}\|_F^2\right) \\
&= \mathcal{F}(X^{\xi(t-1)}, Y^{\xi(t-1)}) - \frac{c}{2}\left(\|\Delta_{X^{\xi(t)-1}}\|_F^2 + \|\Delta_{Y^{\xi(t)-1}}\|_F^2\right),
\end{aligned} \tag{4.39}$$

where the second inequality follows from the facts that $\{\mathcal{F}(X^{\ell(k)}, Y^{\ell(k)})\}_{k=0}^{\infty}$ is non-increasing and $\xi(t)-1 = \ell((N+1)t)-1 \geqslant (N+1)(t-1)$. We next consider two cases.

Case 1: First, we suppose that $\mathcal{F}(X^{\xi(T)}, Y^{\xi(T)}) = \mathcal{F}^*$ for some $T \geqslant 0$. Since the sequence $\{\mathcal{F}(X^{\xi(t)}, Y^{\xi(t)})\}_{t=0}^{\infty}$ is non-increasing, we must have $\mathcal{F}(X^{\xi(t)}, Y^{\xi(t)}) = \mathcal{F}^*$ for all $t \geqslant T$. Then, for all $k \in [(N+1)t-N,\, (N+1)t]$ with any $t \geqslant T$, we have

$$\mathcal{F}(X^k, Y^k) \leqslant \mathcal{F}(X^{\xi(t)}, Y^{\xi(t)}) = \mathcal{F}^*.$$

This proves three statements.

Case 2: From now on, we consider the case where $\mathcal{F}(X^{\xi(t)}, Y^{\xi(t)}) > \mathcal{F}^*$ for all $t \geqslant 0$. From statement (i) in Theorem 4.3, we see that $\{(X^{\xi(t)}, Y^{\xi(t)})\}_{t=0}^{\infty}$ is bounded and hence must have at least one cluster point. Let Γ denote the set of cluster points of $\{(X^{\xi(t)}, Y^{\xi(t)})\}_{t=0}^{\infty}$. Then, from the facts that Ψ, Φ are continuous on their domains ((a1) in Assumption 4.1) and the definition of $\mathcal{F}^*$, it is not hard to see $\mathcal{F}(X^*, Y^*) \equiv \mathcal{F}^*$ for any $(X^*, Y^*) \in \Gamma$. This together with our assumption that $\mathcal{F}$ is a KL function with $\varphi(s) = as^{1-\vartheta}$ and Proposition 2.1 (uniformized KL property) implies that there exist $\varepsilon > 0$ and $\eta > 0$, such that

$$\varphi'\left(\mathcal{F}(X,Y) - \mathcal{F}^*\right) \operatorname{dist}\left(0,\, \partial\mathcal{F}(X,Y)\right) \geqslant 1$$

for all (X,Y) satisfying $\operatorname{dist}((X,Y),\Gamma) < \varepsilon$ and $\mathcal{F}^* < \mathcal{F}(X,Y) < \mathcal{F}^* + \eta$. On the other hand, since $\lim\limits_{t\to\infty} \operatorname{dist}((X^{\xi(t)}, Y^{\xi(t)}),\, \Gamma) = 0$ by the definition of Γ, and

$\mathcal{F}(X^{\xi(t)}, Y^{\xi(t)}) \to \mathcal{F}^*$, then for such ε and η, there exists an integer $T_0 \geqslant 0$ such that $\mathrm{dist}((X^{\xi(t)}, Y^{\xi(t)}), \Gamma) < \varepsilon$ and $\mathcal{F}^* < \mathcal{F}(X^{\xi(t)}, Y^{\xi(t)}) < \mathcal{F}^* + \eta$ for all $t \geqslant T_0$. Thus, for $t \geqslant T_0$, we have

$$\varphi'\left(\mathcal{F}(X^{\xi(t)}, Y^{\xi(t)}) - \mathcal{F}^*\right) \mathrm{dist}\left(0,\ \partial\mathcal{F}(X^{\xi(t)}, Y^{\xi(t)})\right) \geqslant 1. \tag{4.40}$$

For notational simplicity, let $\Delta_{\mathcal{F}}^{\xi(t)} := \mathcal{F}(X^{\xi(t)}, Y^{\xi(t)}) - \mathcal{F}^*$. Since $\{\mathcal{F}(X^{\xi(t)}, Y^{\xi(t)})\}$ is non-increasing, it is easy to see that $\Delta_{\mathcal{F}}^{\xi(t)}$ is non-increasing, $\Delta_{\mathcal{F}}^{\xi(t)} > 0$ for $t \geqslant 0$ and $\Delta_{\mathcal{F}}^{\xi(t)} \to 0$. Then, we have that, for all $t \geqslant T_0$,

$$\begin{aligned}
1 &\leqslant \varphi'\left(\Delta_{\mathcal{F}}^{\xi(t)}\right) \mathrm{dist}\left(0,\ \partial\mathcal{F}(X^{\xi(t)}, Y^{\xi(t)})\right) \\
&\leqslant \frac{a}{1-\vartheta} \cdot \left(\Delta_{\mathcal{F}}^{\xi(t)}\right)^{-\vartheta} \cdot d_1 \left(\|\Delta_{X^{\xi(t)-1}}\|_F + \|\Delta_{Y^{\xi(t)-1}}\|_F\right) \\
&\leqslant \frac{ad_1}{1-\vartheta} \cdot \left(\Delta_{\mathcal{F}}^{\xi(t)}\right)^{-\vartheta} \cdot \sqrt{2\left(\|\Delta_{X^{\xi(t)-1}}\|_F^2 + \|\Delta_{Y^{\xi(t)-1}}\|_F^2\right)} \\
&\leqslant \frac{ad_1}{1-\vartheta} \cdot \left(\Delta_{\mathcal{F}}^{\xi(t)}\right)^{-\vartheta} \cdot \sqrt{\frac{4}{c}\left(\mathcal{F}(X^{\xi(t-1)}, Y^{\xi(t-1)}) - \mathcal{F}(X^{\xi(t)}, Y^{\xi(t)})\right)} \\
&= d_2 \left(\Delta_{\mathcal{F}}^{\xi(t)}\right)^{-\vartheta} \cdot \sqrt{\Delta_{\mathcal{F}}^{\xi(t-1)} - \Delta_{\mathcal{F}}^{\xi(t)}},
\end{aligned} \tag{4.41}$$

where $d_2 := \frac{2ad_1}{(1-\vartheta)\sqrt{c}} > 0$, the first inequality follows from (4.40), the second inequality follows from (4.37), the third inequality follows from $p+q \leqslant \sqrt{2(p^2+q^2)}$ for $p, q \geqslant 0$ and the last inequality follows from (4.39). In the following, we consider three cases.

(i): $\vartheta = 0$. In this case, we see from (4.41) that $\Delta_{\mathcal{F}}^{\xi(t-1)} - \Delta_{\mathcal{F}}^{\xi(t)} \geqslant \frac{1}{d_2^2}$ for all $t \geqslant T_0$, which contradicts $\Delta_{\mathcal{F}}^{\xi(t)} \to 0$. Thus, **Case 2** cannot happen.

(ii): $0 < \vartheta \leqslant \frac{1}{2}$. In this case, we have $0 < 2\vartheta \leqslant 1$. Since $\Delta_{\mathcal{F}}^{\xi(t)} \to 0$, there exists $T_2 \geqslant 0$ such that $\Delta_{\mathcal{F}}^{\xi(t)} \leqslant 1$ for all $t \geqslant \widetilde{T} := \max\{T_0, T_2\}$. Then, for all $t \geqslant \widetilde{T}$, we see from (4.41) that

$$\Delta_{\mathcal{F}}^{\xi(t)} \leqslant \left(\Delta_{\mathcal{F}}^{\xi(t)}\right)^{2\vartheta} \leqslant d_2^2\left(\Delta_{\mathcal{F}}^{\xi(t-1)} - \Delta_{\mathcal{F}}^{\xi(t)}\right),$$

which implies that

$$\Delta_{\mathcal{F}}^{\xi(t)} \leqslant \tilde{\varrho}\, \Delta_{\mathcal{F}}^{\xi(t-1)} \leqslant \cdots \leqslant \tilde{\varrho}^{t-\widetilde{T}+1}\, \Delta_{\mathcal{F}}^{\xi(\widetilde{T}-1)},$$

where $\tilde{\varrho} := \frac{d_2^2}{1+d_2^2} < 1$. Then, for all $k \in [(N+1)t - N,\, (N+1)t]$ with any $t \geqslant \widetilde{T}$, we have

$$\mathcal{F}(X^k, Y^k) - \mathcal{F}^* \leqslant \Delta_{\mathcal{F}}^{\xi(t)} \leqslant \tilde{\varrho}^{t-\widetilde{T}+1}\, \Delta_{\mathcal{F}}^{\xi(\widetilde{T}-1)} \leqslant \tilde{\varrho}^{\frac{k}{N+1}-\widetilde{T}+1}\, \Delta_{\mathcal{F}}^{\xi(\widetilde{T}-1)} = \zeta \varrho^k,$$

where $\zeta := \tilde{\varrho}^{-\widetilde{T}+1}\, \Delta_{\mathcal{F}}^{\xi(\widetilde{T}-1)}$ and $\varrho := \tilde{\varrho}^{\frac{1}{N+1}} < 1$, the last inequality follows from $t \geqslant \frac{k}{N+1}$. This proves statement (ii).

(iii): $\frac{1}{2} < \vartheta < 1$. We define $f(s) := s^{-2\vartheta}$ for $s \in (0, \infty)$. It is easy to see that f is non-increasing. Then, for any $t \geqslant T_0$, we further consider the following two cases.

- If $f(\Delta_{\mathcal{F}}^{\xi(t)}) \leqslant 2f(\Delta_{\mathcal{F}}^{\xi(t-1)})$, it follows from (4.41) that

$$\begin{aligned} 1/d_2^2 &\leqslant (\Delta_{\mathcal{F}}^{\xi(t)})^{-2\vartheta} \cdot (\Delta_{\mathcal{F}}^{\xi(t-1)} - \Delta_{\mathcal{F}}^{\xi(t)}) = f(\Delta_{\mathcal{F}}^{\xi(t)}) \cdot (\Delta_{\mathcal{F}}^{\xi(t-1)} - \Delta_{\mathcal{F}}^{\xi(t)}) \\ &\leqslant 2f(\Delta_{\mathcal{F}}^{\xi(t-1)}) \cdot (\Delta_{\mathcal{F}}^{\xi(t-1)} - \Delta_{\mathcal{F}}^{\xi(t)}) \leqslant 2\int_{\Delta_{\mathcal{F}}^{\xi(t)}}^{\Delta_{\mathcal{F}}^{\xi(t-1)}} f(s)\, \mathrm{d}s \\ &= \frac{2(\Delta_{\mathcal{F}}^{\xi(t-1)})^{1-2\vartheta} - 2(\Delta_{\mathcal{F}}^{\xi(t)})^{1-2\vartheta}}{1-2\vartheta} = \frac{2(\Delta_{\mathcal{F}}^{\xi(t)})^{1-2\vartheta} - 2(\Delta_{\mathcal{F}}^{\xi(t-1)})^{1-2\vartheta}}{2\vartheta-1}, \end{aligned}$$

  which, together with $2\vartheta - 1 > 0$, implies that

$$(\Delta_{\mathcal{F}}^{\xi(t)})^{1-2\vartheta} - (\Delta_{\mathcal{F}}^{\xi(t-1)})^{1-2\vartheta} \geqslant (2\vartheta-1)/(2d_2^2). \tag{4.42}$$

- If $f(\Delta_{\mathcal{F}}^{\xi(t)}) \geqslant 2f(\Delta_{\mathcal{F}}^{\xi(t-1)})$, it is not hard to see that $(\Delta_{\mathcal{F}}^{\xi(t)})^{1-2\vartheta} \geqslant 2^{\frac{2\vartheta-1}{2\vartheta}}(\Delta_{\mathcal{F}}^{\xi(t-1)})^{1-2\vartheta}$. Then, we have

$$\begin{aligned} (\Delta_{\mathcal{F}}^{\xi(t)})^{1-2\vartheta} - (\Delta_{\mathcal{F}}^{\xi(t-1)})^{1-2\vartheta} &\geqslant (2^{\frac{2\vartheta-1}{2\vartheta}} - 1)(\Delta_{\mathcal{F}}^{\xi(t-1)})^{1-2\vartheta} \\ &\geqslant (2^{\frac{2\vartheta-1}{2\vartheta}} - 1)(\Delta_{\mathcal{F}}^{\xi(T_0-1)})^{1-2\vartheta}, \end{aligned} \tag{4.43}$$

where the last inequality follows from the facts that $\Delta_{\mathcal{F}}^{\xi(t)}$ is non-increasing and $1-2\vartheta<0$.

Thus, combining (4.42) and (4.43), we obtain

$$(\Delta_{\mathcal{F}}^{\xi(t)})^{1-2\vartheta}-(\Delta_{\mathcal{F}}^{\xi(t-1)})^{1-2\vartheta}\geqslant d_3:=\min\left\{(2\vartheta-1)/(2d_2^2),\,(2^{\frac{2\vartheta-1}{2\vartheta}}-1)(\Delta_{\mathcal{F}}^{\xi(T_0-1)})^{1-2\vartheta}\right\}.$$

Then, we have

$$\begin{aligned}(\Delta_{\mathcal{F}}^{\xi(t)})^{1-2\vartheta}&\geqslant(\Delta_{\mathcal{F}}^{\xi(t)})^{1-2\vartheta}-(\Delta_{\mathcal{F}}^{\xi(T_0)})^{1-2\vartheta}=\textstyle\sum_{j=T_0+1}^{t}\left((\Delta_{\mathcal{F}}^{\xi(j)})^{1-2\vartheta}-(\Delta_{\mathcal{F}}^{\xi(j-1)})^{1-2\vartheta}\right)\\&\geqslant(t-T_0)d_3\geqslant\frac{d_3}{2}\,t,\end{aligned}$$

where the last inequality holds for $t\geqslant 2T_0$. This implies that $\Delta_{\mathcal{F}}^{\xi(t)}\leqslant d_4\,t^{-\frac{1}{2\vartheta-1}}$, where $d_4:=\left(\frac{d_3}{2}\right)^{-\frac{1}{2\vartheta-1}}$. Then, for all $k\in[(N+1)t-N,\,(N+1)t]$ with any $t\geqslant 2T_0$, we have

$$\mathcal{F}(X^k,Y^k)-\mathcal{F}^*\leqslant\Delta_{\mathcal{F}}^{\xi(t)}\leqslant d_4\,t^{-\frac{1}{2\vartheta-1}}\leqslant\zeta\,k^{-\frac{1}{2\vartheta-1}},$$

where $\zeta:=(N+1)^{\frac{1}{2\vartheta-1}}d_4$ and the last inequality follows from $t\geqslant\frac{k}{N+1}$. This proves statement (iii). $\square$

4.4 Numerical experiments

In this section, we conduct numerical experiments to test our algorithm for NMF and MC on real datasets. All experiments are run in MATLAB R2015b on a 64-bit PC with an Intel Core i7-4790 CPU (3.60 GHz) and 32 GB of RAM equipped with Windows 10 OS.

4.4.1 Non-negative matrix factorization

As we mentioned in the introduction, NMF is a very important problem in many applications. The basic model for NMF is

$$\min_{X,Y} \ \frac{1}{2}\left\|XY^\top - M\right\|_F^2 \quad \text{s.t.} \quad X \geqslant 0, \quad Y \geqslant 0, \tag{4.44}$$

where $X \in \mathbb{R}^{m\times r}$ and $Y \in \mathbb{R}^{n\times r}$ are decision variables. Note that the feasible set of (4.44) is unbounded. We hence focus on the following model:

$$\min_{X,Y} \ \frac{1}{2}\left\|XY^\top - M\right\|_F^2 \quad \text{s.t.} \quad 0 \leqslant X \leqslant X^{\max}, \ 0 \leqslant Y \leqslant Y^{\max}, \tag{4.45}$$

where $X^{\max} \geqslant 0$ and $Y^{\max} \geqslant 0$ are upper bound matrices. One can show that, when $X^{\max}_{ij}$ and $Y^{\max}_{ij}$ are sufficiently large, solving (4.45) gives a solution of (4.44). In our experiments, we set $X^{\max}_{ij} = 10^{16}$ and $Y^{\max}_{ij} = 10^{16}$ for all (i,j). Now, we see that (4.45) corresponds to (1.4) with $\Psi(X) = \delta_{\mathcal{X}}(X)$, $\Phi(Y) = \delta_{\mathcal{Y}}(Y)$ and $\mathcal{A} = \mathcal{I}$, where $\mathcal{X} = \{X \in \mathbb{R}^{m\times r} : 0 \leqslant X \leqslant X^{\max}\}$ and $\mathcal{Y} = \{Y \in \mathbb{R}^{n\times r} : 0 \leqslant Y \leqslant Y^{\max}\}$. We apply NAUM to solving (4.45), and use (4.11c) and (4.12c) to update U and V. The specific updates of Z^k, $\boldsymbol{u}_i$ and $\boldsymbol{v}_i$ are

$$\begin{cases} Z^k = \dfrac{\alpha}{\alpha+\beta} X^k (Y^k)^\top + \dfrac{\beta}{\alpha+\beta} M, \\[2ex] \boldsymbol{u}_i = \max\left\{0, \ \min\left\{\boldsymbol{x}_i^{\max}, \ \dfrac{\alpha P_i^k \boldsymbol{y}_i^k + \mu_k \boldsymbol{x}_i^k}{\alpha\|\boldsymbol{y}_i^k\|^2 + \mu_k}\right\}\right\}, \quad i = 1, 2\cdots, r, \\[2ex] \boldsymbol{v}_i = \max\left\{0, \ \min\left\{\boldsymbol{y}_i^{\max}, \ \dfrac{\alpha (Q_i^k)^\top \boldsymbol{u}_i + \sigma_k \boldsymbol{y}_i^k}{\alpha\|\boldsymbol{u}_i\|^2 + \sigma_k}\right\}\right\}, \quad i = 1, 2\cdots, r, \end{cases}$$

where P_i^k and Q_i^k are defined in (4.18). Note that here it is not necessary to update Z^k explicitly. Indeed, we can directly compute $P_i^k \boldsymbol{y}_i^k$ and $(Q_i^k)^\top \boldsymbol{u}_i$ by substituting

Z^k as below:

$$\begin{cases} P_i^k \boldsymbol{y}_i^k = \frac{\alpha}{\alpha+\beta} X^k (Y^k)^\top \boldsymbol{y}_i^k + \frac{\beta}{\alpha+\beta} M \boldsymbol{y}_i^k - \sum_{j=1}^{i-1} \boldsymbol{u}_j (\boldsymbol{y}_j^k)^\top \boldsymbol{y}_i^k - \sum_{j=i+1}^{r} \boldsymbol{x}_j^k (\boldsymbol{y}_j^k)^\top \boldsymbol{y}_i^k, \\ (Q_i^k)^\top \boldsymbol{u}_i = \frac{\alpha}{\alpha+\beta} Y^k (X^k)^\top \boldsymbol{u}_i + \frac{\beta}{\alpha+\beta} M^\top \boldsymbol{u}_i - \sum_{j=1}^{i-1} \boldsymbol{v}_j \boldsymbol{u}_j^\top \boldsymbol{u}_i - \sum_{j=i+1}^{r} \boldsymbol{y}_j^k \boldsymbol{u}_j^\top \boldsymbol{u}_i. \end{cases} \tag{4.46}$$

When computing $X^k(Y^k)^\top \boldsymbol{y}_i^k$ and $Y^k(X^k)^\top \boldsymbol{u}_i$ in the above, we first compute $(Y^k)^\top \boldsymbol{y}_i^k$ and $(X^k)^\top \boldsymbol{u}_i$ to avoid forming the huge $(m \times n)$ matrix $X^k(Y^k)^\top$. This technique is also used in many popular algorithms for NMF to reduce the computational cost (see, for example, [4, 29, 30, 47, 86]).

We will compare NAUM with some recent algorithms[1] for NMF: the hierarchical alternating least squares (HALS) method[2] (see, for example, [19, 29, 30, 32, 53, 58]) and the block coordinate descent method for NMF (BCD-NMF[3]) (see Algorithm 2 in Section 3.2 in [91]). It shall be mentioned that there are two classical methods for NMF: the multiplicative updating (MU) method [47] and the alternating least squares (ALS) method [4], which can be called simply by the Matlab function **nnmf**[4] with options **mult** and **als**, respectively. However, as observed in [29, 30, 32, 91, 95] and our experiments, the performances of **mult** and **als** are not competitive. Additionally, it has been reported in [91] that BCD-NMF outperforms ADMM in both CPU time and solution quality. Therefore, we do not include MU, ALS and

[1] Most existing algorithms are directly developed for (4.44). However, they need the assumption that the sequence generated is bounded in their convergence analysis. Although this assumption is uncheckable and may fail, these algorithms always work well in practice. Thus, we directly use these algorithms in our comparisons, rather than modifying them for (4.45).

[2] HALS for (4.44) is given by

$$\begin{cases} \boldsymbol{x}_i^{k+1} = \max\left\{0, \dfrac{M\boldsymbol{y}_i^k - \sum_{j=1}^{i-1} \boldsymbol{x}_j^{k+1} (\boldsymbol{y}_j^k)^\top \boldsymbol{y}_i^k - \sum_{j=i+1}^{r} \boldsymbol{x}_j^k (\boldsymbol{y}_j^k)^\top \boldsymbol{y}_i^k}{\|\boldsymbol{y}_i^k\|^2}\right\}, \quad i = 1, \cdots, r, \\ \boldsymbol{y}_i^{k+1} = \max\left\{0, \dfrac{M^\top \boldsymbol{x}_i^{k+1} - \sum_{j=1}^{i-1} \boldsymbol{y}_j^{k+1} (\boldsymbol{x}_j^{k+1})^\top \boldsymbol{x}_i^{k+1} - \sum_{j=i+1}^{r} \boldsymbol{y}_j^k (\boldsymbol{x}_j^{k+1})^\top \boldsymbol{x}_i^{k+1}}{\|\boldsymbol{x}_i^{k+1}\|^2}\right\}, \quad i = 1, \cdots, r. \end{cases}$$

[3] The Matlab codes for BCD-NMF is available at `http://www.math.ucla.edu/~wotaoyin/papers/bcu/nmf/index.html`.

[4] The **nnmf** function is a part of the Statistics and Machine Learning Toolbox for Matlab.

ADMM in our comparisons.

To evaluate the performances of different algorithms, we follow [32] to use an evolution of the objective function value. To define this evolution, we first define

$$e(k) := \frac{\mathcal{F}^k - \mathcal{F}_{\min}}{\mathcal{F}^0 - \mathcal{F}_{\min}},$$

where $\mathcal{F}^k$ denotes the objective function value obtained by an algorithm at (X^k, Y^k) and $\mathcal{F}_{\min}$ denotes the minimum of the objective function values obtained among *all* algorithms across *all* initializations. We also use $\mathcal{T}(k)$ to denote the total computational time after completing the k-th iteration of an algorithm. Thus, $\mathcal{T}(0) = 0$ and $\mathcal{T}(k)$ is non-decreasing with respect to k. Then, the evolution of the objective function value obtained from a particular algorithm with respect to time t is defined as

$$E(t) := \min \{e(k) : k \in \{i : \mathcal{T}(i) \leqslant t\}\}.$$

One can see that $0 \leqslant E(t) \leqslant 1$ (since $0 \leqslant e(k) \leqslant 1$ for all k) and $E(t)$ is non-increasing with respect to t. The $E(t)$ can be considered as a normalized measure of the reduction of the objective function value with respect to time. For a given matrix M and a positive integer r, one can take the average of $E(t)$ over several independent trials with different initializations, and plot the average $E(t)$ within time t for a given algorithm.

The experiments are conducted on face datasets (dense matrices) and text datasets (sparse matrices). For face datasets, we use CBCL[5], ORL[6] [72] and the extended Yale Face Database B (e-YaleB)[7] [48] for our test. CBCL contains 2429 images of

[5] Available in http://cbcl.mit.edu/cbcl/software-datasets/FaceData2.html.

[6] Available in http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html.

[7] Available in http://vision.ucsd.edu/~iskwak/ExtYaleDatabase/ExtYaleB.html. Both original images and cropped images are provided in this dataset. The cropped images except 18 corrupted ones are used in our experiments.

faces with 19×19 pixels, ORL contains 400 images of faces with 112×92 pixels, and e-YaleB contains 2414 images of faces with 168×192 pixels. In our experiments, for each face dataset, each image is vectorized and stacked as a column of a data matrix M of size $m \times n$. For text datasets, we use three datasets from the CLUTO toolkit[8]. The specific values of m and n for each dataset and the values of r used for our tests are summarized in Table 4.1.

Table 4.1: Real data sets

Face Datasets (dense matrices)					Text Datasets (sparse matrices)				
Data	Pixels	m	n	r	Data	Sparsity	m	n	r
CBCL	19×19	361	2429	30, 60	classic	99.92%	7094	41681	10, 20
ORL	112×92	10304	400	30, 60	sports	99.14%	8580	14870	10, 20
e-YaleB	168×192	32256	2414	30, 60	ohscal	99.47%	11162	11465	10, 20

In our experiments, we initialize all the algorithms at the same random initial point (X^0, Y^0), generated by the following Matlab commands:

```
X0 = max(0,randn(m,r)); X0 = X0/norm(X0,'fro')*sqrt(norm(M,'fro'));
Y0 = max(0,randn(n,r)); Y0 = Y0/norm(Y0,'fro')*sqrt(norm(M,'fro'));
```

and set the maximum running time to $\mathrm{T}^{\max}$ for all algorithms. The specific values of $\mathrm{T}^{\max}$ are given in Fig. 4.1 and Fig. 4.2. Additionally, we use the default settings for BCD-NMF. For NAUM, we set $\mu^{\min} = \bar{\mu}_{-1} = 1$, $\sigma^{\min} = \bar{\sigma}_{-1} = 1$, $\sigma^{\max} = 10^6$, $\tau = 4$, $c = 10^{-4}$, $N = 3$, and we choose $\mu_k^0 = \max\left\{0.1\bar{\mu}_{k-1},\, \mu^{\min}\right\}$ and $\sigma_k^0 = \min\left\{\max\left\{0.1\bar{\sigma}_{k-1},\, \sigma^{\min}\right\},\, \sigma^{\max}\right\}$ for any $k \geqslant 0$. Moreover, we set $\beta = \frac{\alpha}{\alpha-1}$, $\gamma = \max\{0, -\alpha, -(\alpha+\beta)\}$ and $\rho = \max\{1, \alpha^2/(\alpha+\beta)^2\}$ for some α. We then test NAUM[9] with $\alpha = 0.6, 0.8, 1.1, 2$, compare them to HALS and BCD-NMF, and plot the average $E(t)$ for each algorithm within time $\mathrm{T}^{\max}$.

[8] CLUTO: A clustering toolkit that is available in `http://glaros.dtc.umn.edu/gkhome/cluto/cluto/download`.

[9] We observed from our experiments for NMF that NAUM with $\alpha < 0.6$ is not competitive. Thus, we do not choose $\alpha < 0.6$ in our comparisons.

Fig. 4.1 and Fig. 4.2 show the average $E(t)$ of 30 independent trials for NMF on face datasets and text datasets, respectively. From the results, we can see that NAUM with $\alpha = 0.6$ performs best in most cases, and NAUM with $\alpha < 1$ always performs better than NAUM with $\alpha > 1$. This shows that choosing α and β under the weaker condition $\frac{1}{\alpha} + \frac{1}{\beta} = 1$ (hence α can be small than 1) can improve the numerical performance of NAUM.

4.4.2 Matrix completion

We next apply our NAUM to a recent model for MC:

$$\min_{X,Y} \ \frac{\eta}{2}\|X\|_* + \frac{\eta}{2}\|Y\|_* + \frac{1}{2}\left\|\mathcal{P}_\Omega(XY^\top - M)\right\|_F^2, \tag{4.48}$$

where $\eta > 0$ is a penalty parameter, Ω is the index set of the known entries of M, and $\mathcal{P}_\Omega(Z)$ keeps the entries of Z in Ω and sets the remaining ones to zero. This model was first considered in [73, 74] and was shown to be equivalent to Schatten-$\frac{1}{2}$ quasi-norm minimization. Encouraging numerical performance of this model has also been reported in [73, 74]. Note that (4.48) corresponds to (1.4) with $\Psi(X) = \frac{\eta}{2}\|X\|_*$, $\Phi(Y) = \frac{\eta}{2}\|Y\|_*$ and $\mathcal{A} = \mathcal{P}_\Omega$. Thus, we can apply NAUM with (4.11b) and (4.12b) to solving (4.48). The updates of Z^k, U and V are

$$\begin{cases} Z^k = X^k(Y^k)^\top + \dfrac{\beta}{\alpha+\beta}\mathcal{P}_\Omega\left(M - X^k(Y^k)^\top\right), \\ U = \mathcal{S}_{\eta/(2\mu_k)}\left(X^k - \dfrac{\alpha}{\mu_k}(X^k(Y^k)^\top - Z^k)Y^k\right), \\ V = \mathcal{S}_{\eta/(2\sigma_k)}\left(Y^k - \dfrac{\alpha}{\sigma_k}(U(Y^k)^\top - Z^k)^\top U\right). \end{cases}$$

Substituting Z^k into U and V and using $\frac{1}{\alpha}+\frac{1}{\beta}=1$ gives

$$\begin{cases} U = \mathcal{S}_{\eta/(2\mu_k)}\left(X^k - \dfrac{1}{\mu_k}\left[\mathcal{P}_\Omega(X^k(Y^k)^\top - M)\right]Y^k\right), \\ V = \mathcal{S}_{\eta/(2\sigma_k)}\left(Y^k - \dfrac{\alpha}{\sigma_k}Y^k(U-X^k)^\top U - \dfrac{1}{\sigma_k}\left[\mathcal{P}_\Omega(X^k(Y^k)^\top - M)\right]^\top U\right). \end{cases} \tag{4.49}$$

Thus, similar to NAUM for NMF, we do not need to update Z^k explicitly for MC.

We compare NAUM with the proximal alternating linearized minimization (PALM), which was proposed in [8] and was used to solve (4.48) in [73, 74]. For ease of future reference, we recall that the PALM for solving (4.48) is given by

$$\begin{cases} X^{k+1} = \mathcal{S}_{\frac{\eta}{2\|Y^k\|^2}}\left(X^k - \dfrac{1}{\|Y^k\|^2}\left[\mathcal{P}_\Omega(X^k(Y^k)^\top - M)\right]Y^k\right), \\ Y^{k+1} = \mathcal{S}_{\frac{\eta}{2\|X^{k+1}\|^2}}\left(Y^k - \dfrac{1}{\|X^{k+1}\|^2}\left[\mathcal{P}_\Omega(X^{k+1}(Y^k)^\top - M)\right]^\top X^{k+1}\right). \end{cases}$$

For NAUM, we use the same parameter settings as in Section 4.4.1, but choose $\alpha = 0.4, 0.6, 1.1$. All the algorithms are initialized at the same random initialization (X^0, Y^0) generated by the following Matlab commands:

```
X0 = randn(m, r); Y0 = randn(n, r);
```

and terminated if one of the following stopping criteria is satisfied:

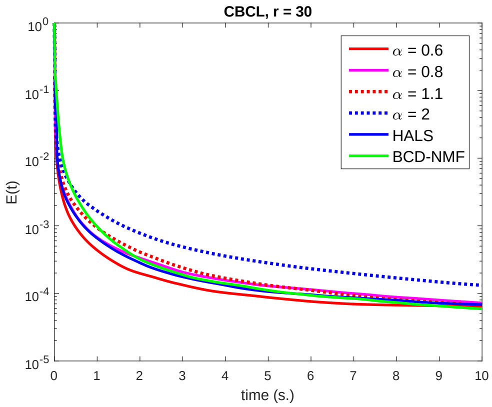
- $\frac{|\mathcal{F}^k - \mathcal{F}^{k-1}|}{\mathcal{F}^k+1} \leqslant 10^{-4}$ holds for 3 consecutive iterations;
- $\frac{\|X^k-X^{k-1}\|_F+\|Y^k-Y^{k-1}\|_F}{\|X^k\|_F+\|Y^k\|_F+1} \leqslant 10^{-4}$ holds;
- the running time is more than 300 seconds,

where $\mathcal{F}^k := \frac{\eta}{2}\|X^k\|_* + \frac{\eta}{2}\|Y^k\|_* + \frac{1}{2}\left\|\mathcal{P}_\Omega(X^k(Y^k)^\top - M)\right\|_F^2$ denotes the objective function value obtained by each algorithm at (X^k, Y^k).

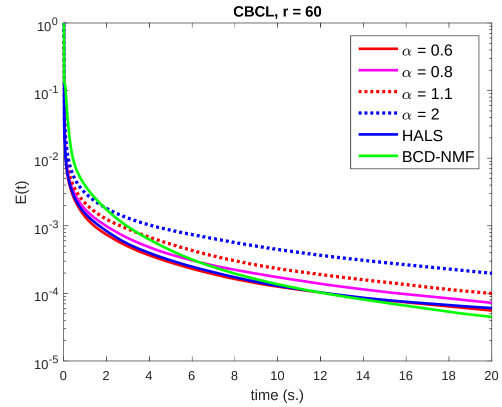
Table 4.2 presents the numerical results of different algorithms for different problems, where two face datasets (CBCL and ORL) are used as our test matrices M and a subset Ω of entries is sampled uniformly at random. In the table, sr denotes the sampling ratio, i.e., a subset Ω of (rounded) $mn * sr$ entries is sampled; r denotes the rank used for test; "iter" denotes the number of iterations; "Normalized fval"[10] denotes the normalized function value $\frac{\mathcal{F}(X^*,Y^*)-\mathcal{F}_{\min}}{\mathcal{F}_{\max}-\mathcal{F}_{\min}}$, where (X^*, Y^*) is obtained by each algorithm, $\mathcal{F}(X^*, Y^*)$ is the function value at (X^*, Y^*) for each algorithm and $\mathcal{F}_{\max}$ (resp. $\mathcal{F}_{\min}$) denotes the maximum (resp. minimum) of the terminating function values obtained from *all* algorithms in *a* trial (one random initialization and Ω); "RecErr" denotes the recovery error $\frac{\|X^*(Y^*)^\top - M\|_F}{\|M\|_F}$. All the results presented are the average of 10 independent trials. Additionally, in each case, the smallest normalized function value, the least CPU time and the smallest recovery error are in bold.

From Table 4.2, we can see that NAUM with $\alpha = 0.4$ gives the smallest function values and the smallest recovery error within least CPU time in most cases. Moreover, NAUM with $\alpha = 0.6$ also performs better than NAUM with $\alpha = 1.1$ and PALM with respect to the function value and the recovery error in most cases. This again shows that a flexible choice of α and β can lead to better numerical performances and the choice of $\alpha = 0.4$ performs best for MC from our experiments.

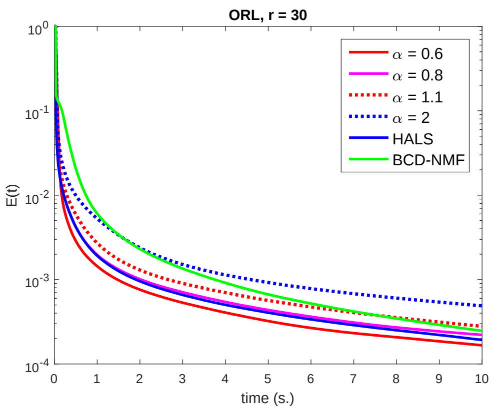
[10] This measure is used to distinguish the function values obtained by different algorithms and it varies between 0 and 1. In a trial, for each algorithm, the smaller the "Normalized fval" value, the better the quality of solution obtained by this algorithm. "0" means that this algorithm attains the smallest function value among all algorithms in this trial.



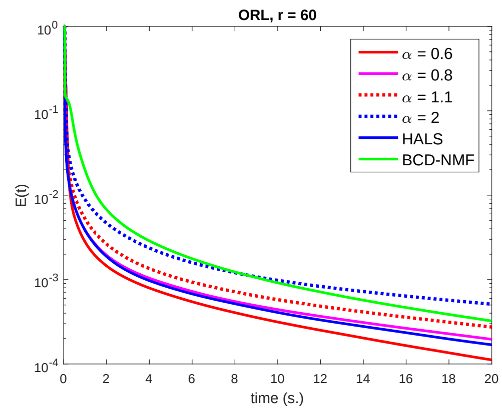
(a) $T^{max} = 10$



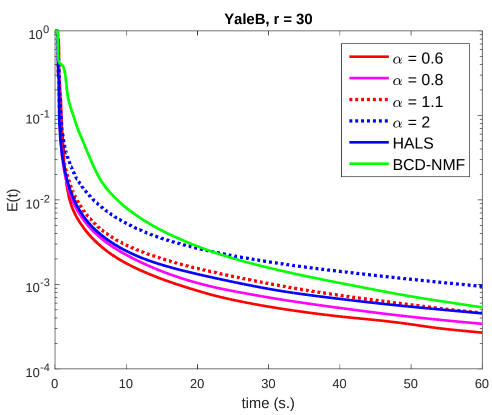
(b) $T^{max} = 20$



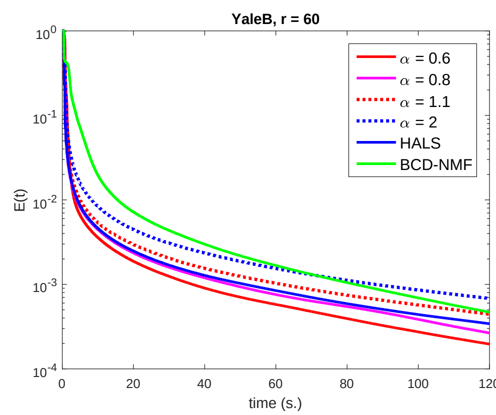
(c) $T^{max} = 10$



(d) $T^{max} = 20$

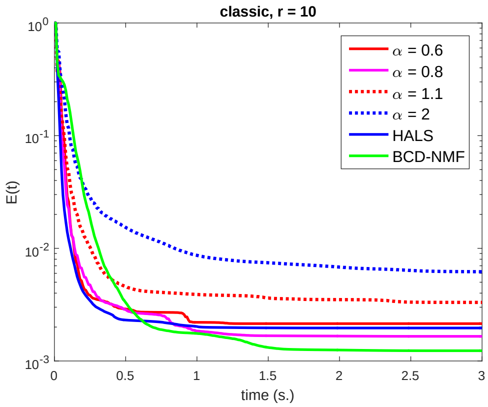


(e) $T^{max} = 60$

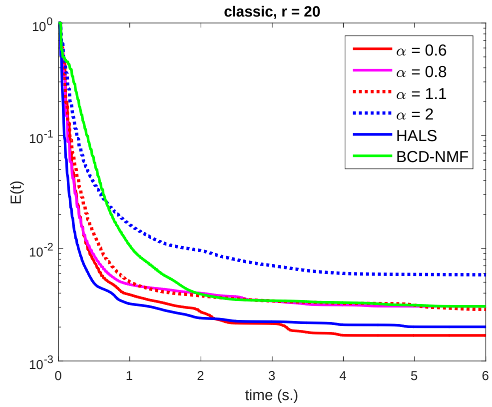


(f) $T^{max} = 120$

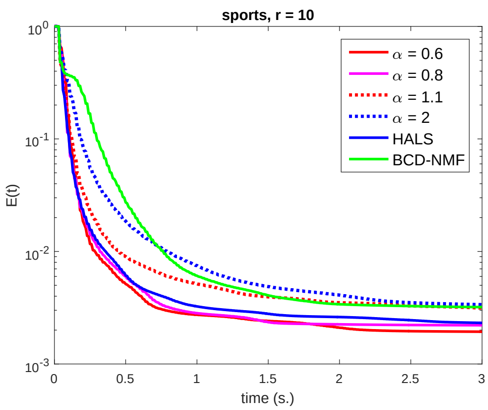
Figure 4.1: Average $E(t)$ of 30 independent trials for NMF on face datasets.



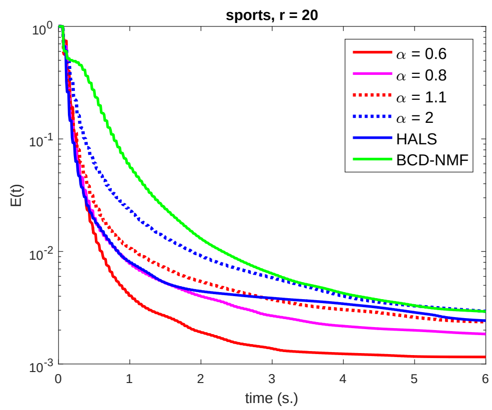
(a) $\mathrm{T}^{\max} = 3$



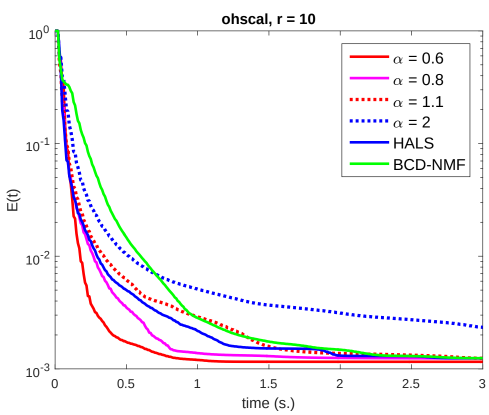
(b) $\mathrm{T}^{\max} = 6$



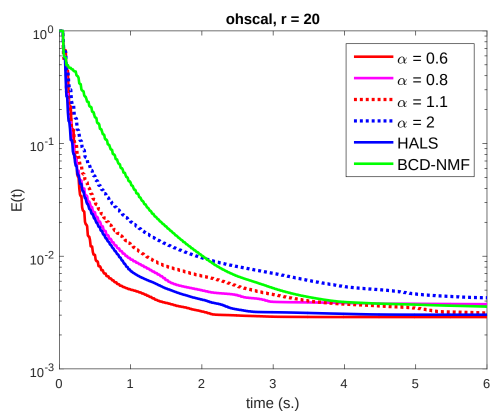
(c) $\mathrm{T}^{\max} = 3$



(d) $\mathrm{T}^{\max} = 6$



(e) $\mathrm{T}^{\max} = 3$



(f) $\mathrm{T}^{\max} = 6$

Figure 4.2: Average $E(t)$ of 30 independent trials for NMF on text datasets.

Table 4.2: Numerical results for MC on face datasets

η	data	sr	r	$\alpha = 0.4$	$\alpha = 0.6$	$\alpha = 1.1$	PALM	$\alpha = 0.4$	$\alpha = 0.6$	$\alpha = 1.1$	PALM
				iter				Normalized fval			
5	CBCL	0.5	30	780	1189	3320	3306	1.13e-01	**7.50e-02**	4.52e-01	1
		0.5	60	921	1218	3850	4654	**3.24e-02**	5.10e-02	3.85e-01	1
		0.2	30	1174	2366	4767	3573	**8.01e-03**	2.21e-01	6.87e-01	9.60e-01
		0.2	60	1577	1919	5360	5037	**1.03e-02**	8.95e-02	8.08e-01	8.86e-01
	ORL	0.5	30	1218	1243	1241	1468	**0**	2.94e-01	5.06e-01	1
		0.5	60	1049	1051	1051	1327	**0**	1	4.00e-01	7.73e-01
		0.2	30	2074	325	385	2691	**2.59e-03**	7.01e-01	1	1.31e-01
		0.2	60	1551	1551	356	2222	**0**	3.82e-01	1	2.12e-01
10	CBCL	0.5	30	457	654	1793	1935	**2.20e-02**	1.29e-01	3.60e-01	9.81e-01
		0.5	60	514	594	1950	2559	2.65e-01	**1.15e-01**	3.79e-01	8.71e-01
		0.2	30	627	1313	2513	2116	**1.91e-02**	3.75e-02	8.35e-01	7.79e-01
		0.2	60	866	1095	2713	2889	**2.07e-02**	2.89e-02	9.22e-01	4.86e-01
	ORL	0.5	30	1003	1186	1192	1402	**3.30e-02**	1.47e-01	4.30e-01	1
		0.5	60	975	1009	1012	1276	**0**	8.58e-01	6.11e-01	9.99e-01
		0.2	30	1409	364	411	2646	**0**	7.16e-01	1	8.10e-02
		0.2	60	1241	1504	376	2185	**4.05e-06**	3.97e-02	1	2.21e-01
				CPU time				RecErr			
5	CBCL	0.5	30	**35.56**	54.14	151.23	119.05	**1.05e-01**	**1.05e-01**	1.06e-01	1.08e-01
		0.5	60	**57.66**	76.09	240.19	206.47	**8.81e-02**	9.02e-02	9.04e-02	8.99e-02
		0.2	30	**34.04**	68.57	137.97	75.56	**1.37e-01**	**1.37e-01**	1.38e-01	1.43e-01
		0.2	60	**72.01**	87.82	245.21	147.08	**1.34e-01**	1.35e-01	1.35e-01	1.36e-01
	ORL	0.5	30	**294.20**	300	300	300	**1.72e-01**	1.84e-01	2.01e-01	2.12e-01
		0.5	60	300	300	300	300	**1.66e-01**	2.11e-01	2.05e-01	2.11e-01
		0.2	30	300	**47.35**	55.86	300	**2.08e-01**	3.04e-01	3.81e-01	2.24e-01
		0.2	60	300	300	**69.21**	300	**2.16e-01**	2.35e-01	3.49e-01	2.61e-01
10	CBCL	0.5	30	**21.01**	30.12	82.45	70.32	**1.16e-01**	1.19e-01	1.18e-01	1.17e-01
		0.5	60	**32.40**	37.38	122.51	113.80	**1.09e-01**	1.11e-01	1.14e-01	1.11e-01
		0.2	30	**18.15**	38.01	72.84	44.62	**1.60e-01**	1.61e-01	1.62e-01	**1.60e-01**
		0.2	60	**39.13**	49.37	123.74	83.52	1.57e-01	1.57e-01	1.58e-01	**1.56e-01**
	ORL	0.5	30	**252.15**	300	300	300	**1.71e-01**	1.77e-01	1.95e-01	2.08e-01
		0.5	60	**289.57**	300	300	300	**1.53e-01**	2.01e-01	2.03e-01	2.09e-01
		0.2	30	207.22	**53.08**	60.54	300	**1.95e-01**	3.06e-01	3.83e-01	2.14e-01
		0.2	60	243.45	295.60	**74.09**	300	**1.87e-01**	1.95e-01	3.60e-01	2.36e-01

Chapter 5

Concluding Remarks

In this chapter, we summarize our main results in this thesis and give some possible directions for the further research.

5.1 Summary

In this thesis, we have considered two classes of matrix optimization problems, which arise in many applications, and developed two efficient algorithms to solve them, respectively.

We first study the *matrix decomposition problem* (MDP), which aims to decompose a given data matrix $D \in \mathbb{R}^{m\times n}$ into two components $L \in \mathbb{R}^{m\times n}$ and $S \in \mathbb{R}^{m\times n}$ with different desirable structures such that $D \approx L + S$, and adapt the ADMM with a general dual step-size τ, which can be chosen in $(0, \frac{1+\sqrt{5}}{2})$, to solve it. Theoretically, we establish that any cluster point of the sequence generated by our ADMM with a non-trivial dual step-size gives a stationary point under some assumptions; we also give simple sufficient conditions for these assumptions. Under an additional assumption that a potential function is a Kurdyka-Łojasiewicz function, we can further establish the global convergence of the whole sequence generated by our ADMM. Numerically, we conduct some experiments for background/foreground extraction and show the efficiency of our algorithm.

We next study the *matrix factorization problem* (MFP), which aims to factorize a given data matrix $M \in \mathbb{R}^{m\times n}$ into two factors $X \in \mathbb{R}^{m\times r}$ and $Y \in \mathbb{R}^{n\times r}$ with different desirable structures such that $M \approx XY^\top$, where $r \leqslant \min\{m, n\}$. To sovle MFP, we introduce a specially constructed potential function $\Theta_{\alpha,\beta}$ defined in (1.5) which contains one auxiliary block of variables. We then develop a non-monotone alternating updating method with a suitable line search criterion based on this potential function. Unlike other existing methods such as those based on alternating minimization, our method essentially updates the two blocks of variables alternately by solving subproblems related to $\Theta_{\alpha,\beta}$ and then updates the auxiliary block of variables by an explicit formula (see (4.15)). Using the special structure of $\Theta_{\alpha,\beta}$, we demonstrate how some efficient computational strategies for NMF can be used to solve the associated subproblems in our method. Moreover, under some mild conditions, we establish that the sequence generated by our method is bounded and any cluster point of the sequence gives a stationary point of our problem. We also discuss the convergence rate for the function value under an additional assumption that the objective is a Kurdyka-Łojasiewicz function. Finally, we conduct some numerical experiments for NMF and MC on real datasets to illustrate the efficiency of our method.

5.2 Future research

Studying MDP and MFP in this thesis is just a start for us to study the first-order splitting methods for the nonconvex, nonsmooth and non-Lipschitz problems. We believe that this topic can be interesting and promising, although it is still challenging. Some possible future works closely related to this thesis are presented as follows.

- Our ADMM may not be beneficial when $\mathcal{B}$ or $\mathcal{C}$ has no special structure, because the corresponding subproblems of ADMM may not have closed-form solutions. Nonetheless, as in [49, 83, 84], it may be possible to add "proximal terms" to simplify the subproblems of our ADMM and lead to some more efficient variants of ADMM. In addition, in view of the recent work [85], it may also be possible to study the convergence of our ADMM for some specially structured nonconvex Ψ.

- The special potential function given in (1.5) for our NAUM is used to separate the coupled variables $XY^\top$ from the linear map $\mathcal{A}$. This technique may be possibly used for other kinds of problems. Indeed, for our MDP (see (1.1)), it is possible to consider the following potential function:

$$\tilde{\Theta}_{\alpha,\beta}(L,S,Z) := \Psi(L) + \Phi(S) + \frac{\alpha}{2}\|\mathcal{B}(L) + \mathcal{C}(S) - Z\|_F^2 + \frac{\beta}{2}\|D - \mathcal{A}(Z)\|_F^2,$$

  where α and β are real numbers. It is conceivable that $\tilde{\Theta}_{\alpha,\beta}$ is closely related to (1.1). Thus, we may derive a new efficient method for (1.1) based on $\tilde{\Theta}_{\alpha,\beta}$ under some conditions.

- Our NAUM may be restricted to the condition $\mathcal{A}\mathcal{A}^* = \mathcal{I}$, which is mainly used to construct our potential function. Thus, for the more general liner map $\mathcal{A}$ in (1.4), we may consider using our ADMM discussed in Chapter 3 to solve (1.4). As we mentioned in Section 1.2, the known conditions that guarantee convergence of the ADMM presented in [22, 92, 95] for NMF are too restrictive. Therefore, it may be interesting to study the ADMM with convergence guarantee for solving problems of the form (1.4).

- Although we consider MDP and MFP separately in this thesis, these two classes of problems have some connections. Note that the matrix L considered in (1.1) is always supposed to have low rank. Thus, it is possible to consider (1.1) incorporated with low-rank matrix factorization, i.e., consider the following problem:

$$\min_{X,Y,S} \ \Psi_1(X) + \Psi_2(Y) + \Phi(S) + \frac{1}{2}\left\| D - \mathcal{A}\left[\mathcal{B}(XY^\top) + \mathcal{C}(S)\right]\right\|_F^2,$$

where $X \in \mathbb{R}^{m\times r}$, $Y \in \mathbb{R}^{n\times r}$ and $S \in \mathbb{R}^{m\times n}$ are decision variables with $r \leqslant \min\{m, n\}$; $\Psi_1 : \mathbb{R}^{m\times r} \to \mathbb{R}\cup\{\infty\}$, $\Psi_2 : \mathbb{R}^{n\times r} \to \mathbb{R}\cup\{\infty\}$ and $\Phi : \mathbb{R}^{m\times n} \to \mathbb{R}\cup\{\infty\}$ are proper closed functions; $\mathcal{A}, \mathcal{B}, \mathcal{C} : \mathbb{R}^{m\times n} \to \mathbb{R}^{m\times n}$ are linear maps. This model also arises in applications such as robust principal component analysis [76]. Thus, it is also a possible future research direction.

Bibliography

[1] B.P.W. Ames and M. Hong. Alternating direction method of multipliers for penalized zero-variance discriminant analysis. *Computational Optimization and Applications*, 64(3):725–754, 2014.

[2] H. Attouch, J. Bolte, P. Redont, and A. Soubeyran. Proximal alternating minimization and projection methods for nonconvex problems: An approach based on the Kurdyka-Łojasiewicz inequality. *Mathematics of Operations Research*, 35(2):438–457, 2010.

[3] H. Attouch, J. Bolte, and B.F. Svaiter. Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward–backward splitting, and regularized Gauss–Seidel methods. *Mathematical Programming*, 137(1):91–129, 2013.

[4] M.W. Berry, M. Browne, A.N. Langville, V.P. Pauca, and R.J. Plemmons. Algorithms and applications for approximate nonnegative matrix factorization. *Computational Statistics & Data Analysis*, 52(1):155–173, 2007.

[5] D. Bertsekas and J.N. Tsitsiklis. *Parallel and Distributed Computation: Numerical Methods*. Prentice Hall, 1989.

[6] W. Bian and X. Chen. Linearly constrained non-Lipschitz optimization for image restoration. *SIAM Journal on Imaging Sciences*, 8(4):2294–2322, 2015.

[7] P. Biswas, T.C. Liang, K.C. Toh, Y. Ye, and T.C. Wang. Semidefinite programming approaches for sensor network localization with noisy distance measurements. *IEEE Transactions on Automation Science and Engineering*, 3(4):360–371, 2006.

[8] J. Bolte, S. Sabach, and M. Teboublle. Proximal alternating linearized minimization for nonconvex and nonsmooth problems. *Mathematical Programming*, 146(1–2):459–494, 2014.

[9] T. Bouwmans. Recent advanced statistical background modelling for foreground detection: A systematic survey. *Recent Patents on Computer Science*, 4(3):147–176, 2011.

[10] T. Bouwmans. Traditional and recent aproaches in background modeling for foreground detection: An overview. *Computer Science Review*, 11–12:31–66, 2014.

[11] T. Bouwmans, A. Sobral, S. Javed, S.K. Jung, and E.-H. Zahzah. Decomposition into low–rank plus additive matrices for background/foreground separation: A review for a comparative evaluation with a large-scale dataset. *Computer Science Review*, 23:1–71, 2017.

[12] T. Bouwmans and E.H. Zahzah. Robust PCA via principal component pursuit: A review for a comparative evaluation in video surveillance. *Computer Vision and Image Understanding*, 122:22–34, 2014.

[13] E.J. Candès, X. Li, Y. Ma, and J. Wright. Robust principal component analysis? *Journal of the ACM*, 58(3):Article No. 11, 2011.

[14] E.J. Candès and B. Recht. Exact matrix completion via convex optimization. *Foundations of Computational Mathematics*, 9:717–772, 2009.

[15] E.J. Candès and T. Tao. The power of convex relaxation: Near-optimal matrix completion. *IEEE Transactions on Information Theory*, 56(5):2053–2080, 2010.

[16] C. Chen, B. He, Y. Ye, and X. Yuan. The direct extension of ADMM for multi-block convex minimization problems is not necessarily convergent. *Mathematical Programming*, 155(1):57–79, 2016.

[17] X. Chen, Z. Lu, and T.K. Pong. Penalty methods for a class of non-Lipschitz optimization problems. *SIAM Journal on Optimization*, 26(3):1465–1492, 2016.

[18] X. Chen and W. Zhou. Smoothing nonlinear conjugate gradient method for image restoration using nonsmooth nonconvex minimization. *SIAM Journal on Imaging Sciences*, 3(4):765–790, 2010.

[19] A. Cichocki, R. Zdunek, and S.I. Amari. Hierarchical ALS algorithms for non-negative matrix and 3D tensor factorization. In *International Conference on Independent Component Analysis and Signal Separation*, pages 169–176. Springer, 2007.

[20] Y. Deng, Q. Dai, R. Liu, Z. Zhang, and S. Hu. Low-rank structure learning via nonconvex heuristic recovery. *IEEE Transactions on Neural Networks and Learning Systems*, 24(3):383–396, 2013.

[21] E.D. Dolan and J.J. Moré. Benchmarking optimization software with performance profiles. *Mathematical Programming*, 91(2):201–213, 2012.

[22] S.S. Du, Y. Liu, B. Chen, and L. Li. Maxios: Large scale nonnegative matrix factorization for collaborative filtering. In *Proceedings of the NIPS 2014 Workshop on Distributed Matrix Computations*, 2014.

[23] J. Eckstein and D. Bertsekas. On the Douglas–Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Mathematical Programming*, 55(1):293–318, 1992.

[24] J. Fan. Comments on "wavelets in statistics: A review" by A. Antoniadis. *Journal of the Italian Statistical Society*, 6(2):131–138, 1997.

[25] J. Fan and R. Li. Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American Statistical Association*, 96(456):1348–1360, 2001.

[26] M. Fazel, T.K. Pong, D. Sun, and P. Tseng. Hankel matrix rank minimization with applications to system identification and realization. *SIAM Journal on Matrix Analysis and Applications*, 34(3):946–977, 2013.

[27] D. Gabay and B. Mercier. A dual algorithm for the solution of nonlinear variational problems via finite element approximations. *Computers & Mathematics with Applications*, 2(1):17–40, 1976.

[28] D. Geman and G. Reynolds. Constrained restoration and the recovery of discontinuities. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(3):367–383, 1992.

[29] N. Gillis. *Nonnegative Matrix Factorization: Complexity, Algorithms and Applications*. PhD thesis, Université catholique de Louvain, 2011.

[30] N. Gillis. The why and how of nonnegative matrix factorization. In *Regularization, Optimization, Kernels, and Support Vector Machines*, pages 257–291. Chapman and Hall/CRC, 2014.

[31] N. Gillis. Introduction to nonnegative matrix factorization. *SIAG/OPT Views and News*, 25(1):7–16, 2017.

[32] N. Gillis and F. Glineur. Accelerated multiplicative updates and hierarchical ALS algorithms for nonnegative matrix factorization. *Neural Computation*, 24(4):1085–1105, 2012.

[33] R. Glowinski and A. Marroco. Sur l'approximation, par éléments finis d'ordre un, et la résolution, par pénalisation-dualité, d'une classe de problèmes de Dirichlet non linéares. *Revue Francaise d'Automatique, Informatique, Recherche Opérationelle*, 9(R-2):41–76, 1975.

[34] P. Gong, C. Zhang, Z. Lu, J.Z. Huang, and J. Ye. A general iterative shinkage and thresholding algorithm for non-convex regularized optimization problems. In *Proceedings of the International Conference on Machine Learning*, volume 28, pages 37–45, 2013.

[35] S. Gu, L. Zhang, W. Zuo, and X. Feng. Weighted nuclear norm minimization with application to image denoising. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2862–2869, 2014.

[36] P.C. Hansen, J.G. Nagy, and D.P. O'leary. *Deblurring Images: Matrices, Spectra, and Filtering*. SIAM, Philadelphia, 2006.

[37] B. He, M. Tao, and X. Yuan. Alternating direction method with Gaussian back substitution for separable convex programming. *SIAM Journal on Optimization*, 22(2):313–340, 2012.

[38] B. He and X. Yuan. Linearized alternating direction method with gaussian back substitution for separable convex programming. *Numerical Algebra, Control and Optimization*, 3(2):247–260, 2013.

[39] M. Hong. Decomposing linearly constrained nonconvex problems by a proximal primal dual approach: Algorithms, convergence and applications. *arXiv preprint arXiv:1604.00543*, 2016.

[40] M. Hong, Z.-Q. Luo, and M. Razaviyayn. Convergence analysis of alternating direction method of multipliers for a family of nonconvex problems. *SIAM Journal on Optimization*, 26(1):337–364, 2016.

[41] J. Huang, J.L. Horowitz, and S. Ma. Asymptotic properties of bridge estimators in sparse high-dimensional regression models. *The Annals of Statistics*, 36(2):587–613, 2008.

[42] B. Jiang, T. Lin, S. Ma, and S. Zhang. Structured nonconvex and nonsmooth optimization: Algorithms and iteration complexity analysis. *arXiv preprint arXiv: 1605.02408*, 2016.

[43] R.H. Keshavan, A. Montanari, and S. Oh. Matrix completion from a few entries. *IEEE Transactions on Information Theory*, 56(6):2980–2998, 2010.

[44] K. Knight and W. Fu. Asymptotics for lasso-type estimators. *The Annals of Statistics*, 28(5):1356–1378, 2000.

[45] M.-J. Lai, Y. Xu, and W. Yin. Improved iteratively rewighted least squares for unconstrained smoothed ℓ_p minimization. *SIAM Journal on Numerical Analysis*, 51(2):927–957, 2013.

[46] D. Lee and H. Seung. Learning the parts of objects by nonnegative matrix factorization. *Nature*, 401:788–791, 1999.

[47] D. Lee and H. Seung. Algorithms for non-negative matrix factorization. In *Proceedings of NIPS*, pages 556–562, 2001.

[48] K.C. Lee, J. Ho, and D. Kriegman. Acquiring linear subspaces for face recognition under variable lighting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(5):684–698, 2005.

[49] G. Li and T.K. Pong. Global convergence of splitting methods for nonconvex composite optimization. *SIAM Journal on Optimization*, 25(4):2434–2460, 2015.

[50] G. Li and T.K. Pong. Douglas-Rachford splitting for nonconvex optimization with application to nonconvex feasibility problems. *Mathematical Programming*, 159(1):371–401, 2016.

[51] L. Li, W. Huang, I.Y.-H. Gu, and Q. Tian. Statistical modeling of complex backgrounds for foreground object detection. *IEEE Transactions on Image Processing*, 13(11):1459–1472, 2004.

[52] L. Li, P. Wang, Q. Hu, and S. Cai. Efficient background modeling based on sparse representation and outlier iterative removal. *IEEE Transactions on Circuits and Systems for Video Technology*, 26(2):278–289, 2014.

[53] L. Li and Y.J. Zhang. Fastnmf: Highly efficient monotonic fixed-point nonnegative matrix factorization algorithm with good applicability. *Journal of Electronic Imaging*, 18(3):033004, 2009.

[54] X. Li, M.K. Ng, and X. Yuan. Median filtering–based methods for static background extraction from surveillance video. *Numerical Linear Algebra with Applications*, 22(5):845–865, 2015.

[55] M. Lin, D. Sun, and K.-C. Toh. A convergent 3-block semi-proximal ADMM for convex minimization problems with one strongly convex block. *Asia-Pacific Journal of Operational Research*, 32(3):1550024(19p), 2015.

[56] Z. Lin, R. Liu, and H. Li. Linearized alternating direction method with parallel splitting and adaptive penalty for separable convex programs in machine learning. *Machine Learning*, 99(2):287–325, 2015.

[57] G. Liu, Z. Lin, S. Yan, J. Sun, Y. Yu, and Y. Ma. Robust recovery of subspace structures by low-rank representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(1):171–184, 2013.

[58] J. Liu, J. Liu, P. Wonka, and J. Ye. Sparse non-negative tensor factorization using columnwise coordinate descent. *Pattern Recognition*, 45(1):649–656, 2012.

[59] Z. Liu and L. Vandenberghe. Interior-point method for nuclear norm approximation with application to system identification. *SIAM Journal on Matrix Analysis and Applications*, 31(3):1235–1256, 2009.

[60] Z. Lu and Y. Zhang. Schatten-p quasi-norm regularized matrix optimization via iterative reweighted singular value minimization. *arXiv preprint arXiv: 1401.0869*, 2014.

[61] K. Mohan and M. Fazel. Iterative reweighted algorithms for matrix rank minimization. *Journal of Machine Learning Research*, 13:3441–3473, 2012.

[62] K. Mohan, P. London, M. Fazel, D. Witten, and S.-I. Lee. Node-based learning of multiple Gaussian graphical models. *Journal of Machine Learning Research*, 15(1):445–488, 2014.

[63] M. Nikolova. Energy minimization methods. In *Handbook of Mathematical Methods in Imaging*, pages 139–185. Springer, 2011.

[64] M. Nikolova, M.K. Ng, S. Zhang, and W.-K. Ching. Efficient reconstruction of piecewise constant images using nonsmooth nonconvex minimization. *SIAM Journal on Imaging Sciences*, 1(1):2–25, 2008.

[65] R. Otazo, E.J. Candès, and D.K. Sodickson. Low–rank plus sparse matrix decomposition for accelerated dynamic MRI with separation of background and dynamic components. *Magnetic Resonance in Medicine*, 73(3):1125–1136, 2015.

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

[67] Y. Peng, A. Ganesh, J. Wright, W. Xu, and Y. Ma. Robust alignment by sparse and low-rank decomposition for linearly correlated images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(11):2233–2246, 2012.

[68] B. Recht. A simpler approach to matrix completion. *Journal of Machine Learning Research*, 12:3413–3430, 2011.

[69] B. Recht, M. Fazel, and P.A. Parrilo. Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM Review*, 52(3):471–501, 2010.

[70] J.D.M. Rennie and N. Srebro. Fast maximum margin matrix factorization for collaborative prediction. In *Proceedings of the 22nd international conference on Machine learning*, pages 713–719. ACM, 2005.

[71] R.T. Rockafellar and R.J-B. Wets. *Variational Analysis*. Springer, 1998.

[72] F. Samaria and A. Harter. Parameterisation of a stochastic model for human face identification. In *Proceedings of the Second IEEE Workshop on Applications of Computer Vision*, pages 138–142. IEEE, 1994.

[73] F. Shang, Y. Liu, and J. Cheng. Scalable algorithms for tractable Schatten quasi-norm minimization. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence*, pages 2016–2022, 2016.

[74] F. Shang, Y. Liu, and J. Cheng. Tractable and scalable Schatten quasi–norm approximations for rank minimization. In *Proceedings of the 19th International Conference on Artificial Intelligence and Statistics*, pages 620–629, 2016.

[75] F. Shang, Y. Liu, and J. Cheng. Unified scalable equivalent formulations for Schatten quasi-norms. *arXiv preprint arXiv: 1606.00668*, 2016.

[76] Y. Shen, Z. Wen, and Y. Zhang. Augmented Lagrangian alternating direction method for matrix separation based on low–rank factorization. *Optimization Methods and Software*, 29(2):239–263, 2014.

[77] N. Srebro. *Learning with Matrix Factorizations*. PhD thesis, Massachusetts Institute of Technology, 2004.

[78] D. Sun, K.-C. Toh, and L. Yang. A convergent 3–block semi-proximal alternating direction method of multipliers for conic programming with 4–type constraints. *SIAM Journal on Optimization*, 25(2):882–915, 2015.

[79] R. Sun and Z.-Q. Luo. Guaranteed matrix completion via non-convex factorization. *IEEE Transactions on Information Theory*, 62(11):6535–6579, 2016.

[80] P. Tseng and S. Yun. A coordinate gradient descent method for nonsmooth separable minimization. *Mathematical Programming*, 117(1):387–423, 2009.

[81] M. Udell, C. Horn, R. Zadeh, and S. Boyd. Generalized low rank models. *Foundations and Trends in Machine Learning*, 9(1):1–118, 2016.

[82] S.A. Vavasis. On the complexity of nonnegative matrix factorization. *SIAM Journal on Optimization*, 20(3):1364–1377, 2009.

[83] F. Wang, W. Cao, and Z. Xu. Convergence of multi-block Bregman ADMM for nonconvex composite problems. *arXiv preprint arXiv:1505.03063*, 2015.

[84] F. Wang, Z. Xu, and H.-K. Xu. Convergence of Bregman alternating direction method with multipliers for nonconvex composite problems. *arXiv preprint arXiv:1410.8625*, 2014.

[85] Y. Wang, W. Yin, and J. Zeng. Global convergence of ADMM in nonconvex nonsmooth optimization. *arXiv preprint arXiv:1511.06324*, 2015.

[86] Y.X. Wang and Y.J. Zhang. Nonnegative matrix factorization: A comprehensive review. *IEEE Transactions on Knowledge and Data Engineering*, 25(6):1336–1353, 2013.

[87] Z. Wen, W. Yin, and Y. Zhang. Solving a low–rank factorization model for matrix completion by a non-linear successive over-relaxation algorithm. *Mathematical Programming Computation*, 4(4):333–361, 2012.

[88] S.J. Wright, R. Nowak, and M.A.T. Figueiredo. Sparse reconstruction by separable approximation. *IEEE Transactions on Signal Processing*, 57(7):2479–2493, 2009.

[89] H. Xu, C. Caramanis, and S. Sanghavi. Robust PCA via outlier pursuit. *IEEE Transactions on Information Theory*, 58(5):3047–3064, 2012.

[90] M.H. Xu and T. Wu. A class of linearized proximal alternating direction methods. *Journal of Optimization Theory and Applications*, 151(2):321–337, 2011.

[91] Y. Xu and W. Yin. A block coordinate descent method for regularized multiconvex optimization with applications to nonnegative tensor factorization and completion. *SIAM Journal on Imaging Sciences*, 6(3):1758–1789, 2013.

[92] Y. Xu, W. Yin, Z. Wen, and Y. Zhang. An alternating direction algorithm for matrix completion with nonnegative factors. *Frontiers of Mathematics in China*, 7(2):365–384, 2012.

[93] C.-H. Zhang. Nearly unbiased variable selection under minimax concave penalty. *The Annals of Statistics*, 38(2):894–942, 2010.

[94] M. Zhang, Z.-H. Huang, and Y. Zhang. Restricted p-isometry properties of nonconvex matrix recovery. *IEEE Transactions on Information Theory*, 59(7):4316–4323, 2013.

[95] Y. Zhang. An alternating direction algorithm for nonnegative matrix factorization. *Preprint*, 2010.

[96] X. Zhou, C. Yang, H. Zhao, and W. Yu. Low-rank modeling and its applications in image analysis. *ACM Computing Survey*, 47(2):36, 2015.

[97] W. Zuo, D. Ren, S. Gu, L. Lin, and L. Zhang. Discriminative learning of iteration-wise priors for blind deconvolution. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3232–3240, 2015.