Généralisation de la pénalisation L1 dans les estimations sparses

Assoweh Mohamed Ibrahim

Généralisation de la pénalisation ℓ_1 dans les estimations sparses

Thèse de doctorat de
l'Université Bourgogne Franche-Comté

École doctorale CARNOT-PASTEUR

présentée et soutenue publiquement le 13 juillet 2021
en vue de l'obtention du grade de

Docteur de l'Université Bourgogne Franche-Comté

(mention Mathématiques Appliquées et Applications des Mathématiques)

par

Assoweh Mohamed Ibrahim

Composition du jury :

Président :	M. Julien JACQUES	Université de Lyon 2
Rapporteurs :	M. Pascal BONDON	CNRS et Université Paris-Saclay
	M. Julien JACQUES	Université de Lyon 2
Examinateurs :	M. Landy RABEHAISAINA	Université Bourgogne Franche-Comté
	M. Brahim TAMADAZTE	CNRS et Sorbonne Université
Directeur de thèse :	M. Stéphane CHRETIEN	Université de Lyon 2

Remerciements

Cette thèse n'aurait jamais été écrite sans la contribution de nombreuses personnes. A toutes ces personnes, je souhaite exprimer ma sincère gratitude.

Tout d'abord, je voudrais remercier mon conseiller, Stéphane CHRETIEN, pour ses conseils et son aide constante durant toutes ces années. C'est un énorme et immense plaisir d'avoir travaillé avec lui, car en plus de son soutien scientifique sans faille, il a toujours été là pour me soutenir et me conseiller dans l'élaboration de cette thèse.

Un grand merci à Brahim TAMADAZTE pour nous avoir suggéré les applications à l'OCT de nos travaux. Cette collaboration très stimulante nous a motivés tout au long de ce travail.

J'adresse tous mes remerciements à Pascal BONDON et à Julien JACQUES de l'honneur qu'ils m'ont fait d'accepter d'évaluer cette thèse et de l'intérêt qu'ils ont porté à mon travail.

Je remercie également Landy RABEHAISANA d'avoir également accepté de participer à la soutenance de cette thèse.

Je tiens aussi à remercier Davit VARRON d'avoir été membre du comité de suivi de la thèse.

Je remercie chaleureusement tous les membres du Laboratoire de Mathématiques de Besançon pour la saine atmosphère d'intéraction et de convivialité tout au long de ces années.

Mes plus profonds remerciements vont à mes parents, ils ont su me donner toutes les chances pour réussir. Qu'ils trouvent, dans la réalisation de ce travail, l'aboutissement de leurs efforts ainsi que l'expression de ma plus affectueuse gratitude.

Je remercie avec amour ma chère épouse pour son soutien quotidien indéfectible et son enthousiasme contagieux. Notre couple a grandi en même temps que mon projet scientifique, le premier servant de socle solide à l'épanouissement du second.

Ces remerciements ne peuvent s'achever sans une pensée à mes frères, mes soeurs et à l'ensemble de ma famille et de mes amis pour leur soutien dans mes efforts.

Je dédie ce travail à mes parents, à mes filles Hawa & Hikmat et à mon épouse Hana.

Résumé

De nombreuses applications en engineering, sociologie, neurosciences, biologie, etc nécessitent l'utilisation d'outils mathématiques pour la résolution de problèmes inverses mal conditionnés. Pour cette raison, l'étude des problèmes inverses est l'objet d'une activité très intense depuis de nombreuses décennies.

Une petite révolution a eu lieu au milieu des années 2000, lors qu'est apparue la théorie du Compressed Sensing, proposée par Candès, Donoho, Romberg et Tao, une approche qui a abondamment irrigué par la suite les diverses branches des mathématiques appliquées où avait été démontré l'existence de représentations parcimonieuses a priori des solutions recherchées, comme en traitement d'images, en équations aux dérivées partielles, en génétique, en analyse des réseaux, etc. Les méthodes de résolution sont souvent regroupées sous l'étiquette "régression pénalisée", avec une pénalisation de type "norme ℓ_1 ", "norme ℓ_2" ou des combinaisons. Récemment, la notion de sparsité a évolué et de nombreux problèmes de reconstruction, par exemple matriciels, sont maintenant étudiés sous l'angle de la parcimonie spectrale, où l'objectif est de reconstruire des matrices de rang faible. Pour cela, l'approche la plus puissance s'appuie sur la régression pénalisée par la norme nucléaire, dont la résolution s'appuie sur la programmation semi-définie, et des méthodes de factorisation de type Bürer-Monteiro afin de préserver le passage à l'échelle algorithmique.

L'objet du présent travail est dans un premier temps de proposer une étude de la régression pénalisée par la norme ℓ_1, appelée méthode du LASSO, pour la reconstruction de vecteurs admettant une représentation parcimonieuses, dans le cas où la matrice de design ne satisfait pas aux contraintes usuellement posées dans ce type d'approche (incohérence, inversibilité restreinte, etc). Nous obtenons en particulier des résultats nouveaux sur l'erreur de prédiction pour le LASSO dans le cas où la matrice des colonnes de la matrice de design sont des réalisations d'un mélange Gaussien vectoriel, dont les centres de classes forment une matrice satisfaisant les conditions d'incohérence usuelles. Nos contributions sont théoriques et s'appuient sur des résultats récents en probabilité.

Dans un deuxième temps, nous considérons les problèmes de reconstruction tensorielle. Nous proposons en particulier une extension des méthodes classiques basées sur la parcimonie spectrale à des problèmes de reconstruction tensorielle, motivés par des applications en tomographie à cohérence optique. Pour cela, nous utilisons le cadre tensoriel introduit

par Kilmer et ses collaborateurs ces dernières années. Pour ces problèmes, nous présentons des résultats nouveaux de nature: - statistique sur la qualité de la reconstruction par minimisation du risque empirique pénalisé par la norme nucléaire. - géométriques démontrant que l'approche par factorisation conduit à la minimisation d'une fonctionnelle non-convexe, mais dont tous les minimiseurs locaux sont aussi globaux.

Nos résultats théoriques sont illustrés par des expériences sur des données réelles en Tomographie à Cohérence Optique.

Mots-clés : Modèle de mélange, LASSO, reconstruction d'un tenseur, tubal SVD, norme nucléaire, norme de Frobenius, Tomographie par cohérence optique.

Abstract

Many applications in engineering, sociology, neuroscience, biology, etc. require the use of sophisticated mathematical tools for the solution of ill-conditioned inverse problems. For this reason, the study of inverse problems has become a subject of an extensive research activity for many decades.

In the middle of the 2000s, the theory of Compressed Sensing, proposed by Candès, Donoho, Romberg and Tao, marked the emergence of a new paradigm which has abundantly irrigated the various branches of applied mathematics where existence of an a priori sparse representations of the sought-after solutions is known, such as in image processing, partial differential equations, genetics, network analysis, etc. The solution methods are often grouped under the label "penalised regression", with a penalisation of the type "ℓ_1 norm", "ℓ_2 norm" or combinations of them. Recently, the notion of sparsity was applied to spectra of matrices, and the topic of low rank reconstruction has emerged as a handy tool in many modern machine learning problems. In many cases, the most powerful approach to spectrally sparse reconstruction problems is penalised regression penalised by the nuclear norm. The solution to such problems is often obtained using semidefinite programming, which can be efficiently solved via factorisation methods of the Bürer-Monteiro type.

The aim of the present work is first to propose a study of l1-penalized regression, called the LASSO method, for the reconstruction of sparse vectors in the case where the design matrix does not satisfy the constraints usually posed in the literature such as incoherence, restricted isometry, etc. In particular, we obtain new results on the prediction error for LASSO in the case where the matrix of columns of the design matrix are realisations of a vector Gaussian mixture, whose class centres form a matrix satisfying the usual incoherence conditions. Our contributions are mainly theoretical and makes use of recent results in probability.

Next, we propose an extension of classical methods based on spectral sparsity to tensor reconstruction problems, motivated by applications in optical coherence tomography. For this purpose, we use the "tubal" tensor framework introduced by Kilmer and co-workers in recent years. For these problems, we present new - statistical results on the quality of the reconstruction by minimising the empirical risk penalised by the nuclear norm. - geometric results showing that the factorization approach leads to the minimization of

a non-convex functional, but whose local minimizers are global as well. Our theoretical results are illustrated by experiments on real data in Optical Coherence Tomography.

Keywords: Mixture model, LASSO, tensor comletion, tubal SVD, nuclear norm, Frobenuis norm, Optical Coherence Tomography.

Table des matières

Liste des figures

Liste des tableaux

Chapter 1

Introduction

De nos jours, de grandes quantités de données sont collectées et exploitées dans presque tous les domaines de la science, de la médecine, du commerce et de l'industrie. En médecine, par exemple, les scientifiques médicaux étudient le génome afin de mettre en oeuvre les meilleurs traitements ou pour connaître les causes de leur maladie. Un autre domaine très en vogue est celui du e-commerce où l'on étudie les évaluations des clients dans le but de leur recommander de nouveaux produits. Les réseaux sociaux, de leur côté, exploitent intensivement les informations collectées sur les personnes connectées et leurs liens afin de comprendre les comportements, détecter des groupes de clients potentiels et vendre ces informations à de grandes entreprises. Les applications pratiques seraient longues à énumérer tant elle sont variées et potentiellement illimitées.

Un des grands enjeux de l'exploration des données accessibles est d'en extraire l'information vraiment essentielle. Par exemple, il est très probable que parmi les 30 000 gènes du corps humain, seulement quelques uns sont directement impliqués dans le processus qui conduit au développement d'un cancer. Concernant le commerce en ligne, les grandes entreprises et les réseaux sociaux sont convaincus que les évaluations d'un client sur peut-être 50 ou 100 films différents suffisent caractériser précisément son profil et ses goûts. Cette hypothèse de l'existence d'un information très simple cachée derrière des fluctations possiblement très fortes est de nos jours rendue pratique dans l'émergence de la notion de parcimonie. Au sens large, un modèle statistique parcimonieux est un modèle dans lequel seul un nombre relativement faible de paramètres joue un rôle explicatif ou prédictif important.

Dans ce mémoire, nous étudierons deux types de parcimonie associées à deux modèles différents.

- Dans le premier chapitre, nous étudierons le modèle linéaire multiple classique, mais dans le cas où il y a plus de covariables que d'observations, un cas a priori infaisable sans plus d'hypothèses sous-jacentes. Une façon de résoudre le problème est de supposer comme en analyse génétique que le vecteur de régression est suffisamment parcimonieux, i.e. le nombre de composantes non nulles du vecteur de régression est plus petit que le nombre d'observations. Entrons dans les détails formels. Nous

observons n observations d'une variable de résultat y_i et p variables prédictives associées (ou caractéristiques) $x_i = (x_{i1}, \dots x_{ip})^T$. L'objectif est de prédire le résultat à partir des prédicteurs, à la fois pour une prédiction réelle avec des données futures et aussi pour découvrir quels prédicteurs jouent un rôle important. Un modèle de régression linéaire suppose que

$$y_i = \beta_0 + \sum_{j=1}^{p} x_{ij}\beta_j + e_i$$

où β_0 et $\beta = (\beta_1, \beta_2, \dots \beta_p)$ sont des paramètres inconnus et e_i est un terme d'erreur. La méthode des moindres carrés fournit des estimations des paramètres par la minimisation de la fonction objectif des moindres carrés

$$\underset{\beta_0,\beta}{\text{minimize}} \sum_{i=1}^{n} \left(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j \right)^2$$

En particulier, si $p > n$, les estimations des moindres carrés ne sont pas uniques. Il existe un ensemble infini de solutions, ce qui rend la fonction objectif égale à zéro, ce qui nous amène à une situation d'overfitting.

Il est donc nécessaire de contraindre ou de régulariser le critère d'estimation. Dans la régression LASSO, encore appelée ℓ_1 -régularisée, nous estimons les paramètres en résolvant le problème suivant

$$\underset{\beta_0,\beta}{\text{minimize}} \sum_{i=1}^{N} \left(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j \right)^2 \text{ subject to } \|\beta\|_1 \leqslant \tau$$

où $\|\beta\|_1 = \sum_{j=1}^{p} |\beta_j|$ est la norme ℓ_1 de β, et τ est un paramètre spécifié par l'utilisateur. Nous pouvons considérer τ comme un budget sur la norme totale de ℓ_1 du vecteur de paramètres, et le LASSO trouve le meilleur ajustement dans ce budget.

Malheureusement, les hypothèses usuellement requises pour démontrer le bon fonctionnement de la méthode du LASSO sont souvent très contraignantes. Ainsi, RIP ou l'incohérence impliquent en particulier que les colonnes de la matrice de design soient très peu colinéaires, ce qui est souvent mis en défaut par les données pratiques. Pour pallier ce problème, nous introduisons un nouveau modèle de mélange Gaussien pour les colonnes de la matrice de design et nous montrons que pour ce modèles, de bonnes propriétés en prédictions sont encore possibles.

- Dans le deuxième chapitre, nous étudions les problèmes du type suivant : étant donné des données sous la forme d'un tenseur X, trouvez un tenseur $\widehat{X}$ qui approxime X dans un sens approprié. Ce problème est abordé d'une manière générale comme celui d'acquérir une compréhension de la matrice X à travers une approximation $\widehat{X}$ qui a une structure simple. Plusieurs scénarios sont possibles dans le processus d'acquision des mesures sur le tenseur X. Dans le problème qui nous intéresse spécifiquement, des

mesures sont prises partiellement sur le tenseur à connaître, et le problème résultant est une extention du problème de Matrix Completion, devenu célèbre après l'annonce du prix Netflix en 2009, au cas tensoriel.

L' approche générale consiste à considérer des estimateurs basés sur des problèmes d'optimisation de la forme

$$\widehat{X} = \arg \min_{M \in \mathbb{R}^{n_1 \times n_2 \times n_3}} \|\mathcal{A}(X) + \varepsilon - \mathcal{A}(M)\|_F^2 \text{ sujet à } \Phi(M) \leqslant c \tag{1.1}$$

où $\|\cdot\|_F^2$ est la norme de Frobenius (au carré) d'un tenseur (définie comme la somme des carrés par élément), et $\Phi(\cdot)$ est une fonction de contrainte qui encourage $\widehat{X}$ à être clairsemé dans un certain sens général. Une fonction de pénalisation assez classique est la norme nucléaire, qui n'est rien d'autre que la norme ℓ_1 du spectre pour les matrices. Dans la cas tensoriel, la SVD n'est pas unique et peut être définie de plusieurs manières différentes. Nous étudions le problème dans le cas des tenseurs tubaux, définis récemment par Kilmer et ses collaborateurs. Une de nos contributions est alors d'analyser une norme nucléaire naturelle pour les tenseurs et de l'utiliser pour donner une borne précise sur la qualité de la reconstruction de X par $\hat{X}$, en appliquant des techniques intialement développées par Koltchinski, Lounici et Tsybakov.

- Dans la dernière partie de ce travail, nous nous concentrons sur l'analyse d'une approche par régularisation de rang faible, c'est à dire que nous nous concentrons sur une approche où la norme nucléaire étudiée dans le chapitre précédent pour la reconstrution de tenseurs tubaux de rang faible, peut être remplacée par une contrainte de rang faible. Contrairement à la partie précédente où les questions étaient de nature purement statistique, nous regardons maintenant le problème sous l'angle optimisation et prouvons dans l'esprit de travaux récent de Rong Ge et ses collaborateurs, que la factorisation de Bürer Monteiro est telle que tous les minimiseurs locaux de problème (1.1) sont en fait des optimiseurs globaux. Ce résultat justifie donc les techniques de factorisations pour les problèmes de reconstructions où les dimensions sont trop grandes et passer en dimension faible permet un passage à l'échelle significatif. Nous illustrons l'efficacité de l'approche sur un problème d'OCT comme pour le chapitre précédent.

Les trois chapitres suivants contiennent nos résultats sur les trois directions que nous venons de présenter et consiste en des articles dont le premier est à soumettre, le deuxième est publié et le troisième est en révision.

Nous allons maintenant présenter plus en détail les notions l'estimation pénalisée dans le cadre vectoriel (avec le LASSO) et matriciel/tensoriel (pour les problèmes de complétion).

1.1 Le LASSO et ses généralisations

Soit le modèle de régression linéaire multiple

$$Y = X\theta + \varepsilon \tag{1.2}$$

où $Y = (Y_1, \ldots, Y_n)^\top$, $X \in \mathbb{R}^{n \times d}$ est une matrice de données, $\theta \in \mathbb{R}^d$ est le vecteur des paramètres inconnus et $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)^\top$ est le vecteur d'erreur gaussien, indépendant et identiquement distribué, $\mathcal{N}(0, \sigma^2 I)$.

L'estimateur LASSO $\hat{\theta}$ de θ est défini par

$$\hat{\theta} \in \underset{\theta \in \mathbb{R}^d}{\operatorname{Argmin}} L_\lambda(\theta) = \frac{1}{n}\|y - X\theta\|_2^2 + \lambda\|\theta\|_1. \tag{1.3}$$

où $\lambda \geqslant 0$ est un paramètre de régularisation.

D'autre part, l'estimateur LASSO vérifie la condition de dualité suivante

$$\|X^\top \left(y - X\hat{\theta}\right)\|_\infty \leqslant \lambda.$$

Etant donnée que $\hat{\theta}$ minimise la fonction $L_\lambda(\theta)$, alors par définition, $\forall \theta \in \mathbb{R}^d$, on a:

$$L_\lambda\left(\hat{\theta}\right) \leqslant L_\lambda(\theta).$$

Cela implique le théorème suivant qui correspond à contrôler l'erreur quadratique moyenne dans le cas où aucune hypothèse est requise pour la matrice X.

Theorem 1.1.1 *Supposons que le modèle linéaire (1.2) est vérifié lorsque $\varepsilon \sim subG_n(\sigma^2)$. D'autre part, supposons que les colonnes de X sont normalisées de tel sorte que $\max_j \|X_j\|_2 \leqslant \sqrt{n}$. Alors, l'estimateur LASSO $\hat{\theta}$ avec le paramètre de régularisation*

$$2\lambda = 2\sigma\sqrt{\frac{2\log(2d)}{n}} + 2\sigma\sqrt{\frac{2\log(1/\delta)}{n}}$$

satisfait

$$\frac{1}{n}\left\|X\hat{\theta} - X\theta\right\|_2^2 \leqslant 4\|\theta\|_1 \sigma\sqrt{\frac{2\log(2d)}{n}} + 4\|\theta\|_1 \sigma\sqrt{\frac{2\log(1/\delta)}{n}}$$

avec une probabilité d'au moins $1 - \delta$.

Le taux dans le théorème précédent est de l'ordre $\sqrt{\log(2d)/n}$ (taux faible). Par la suite, le théorème suivant montre que le taux rapide est atteint par l'estimateur LASSO et l'erreur quadratique moyenne lorsque la matrice X satifait la condition d'incohérence.

Definition 1.1.2 *(Hypothèse INC(s)) On dit que la matrice X a s incohérence pour un entier $s > 0$ si*

$$\|\frac{X^\top X}{n} - I_d\|_\infty \leqslant \frac{1}{32s}.$$

Theorem 1.1.3 *Pour $n \geqslant 2$ fixé. Supposons que le modèle linéaire est vérifié pour $\varepsilon \sim subG(\sigma^2)$. De plus, supposons que $\|\beta\|_0 \leqslant s$ et que X satisfait l'hypothèse de cohérence INC(s). Alors l'estimateur LASSO $\hat{\beta}$ avec le paramètre de régularisation défini par*

$$2\lambda = 8\sigma\sqrt{\frac{2\log(2d)}{n}} + 8\sigma\sqrt{\frac{2\log(1/\delta)}{n}}$$

satisfait

$$\frac{1}{n}\left\|X\hat{\theta} - X\theta\right\|_2^2 \lesssim s\sigma^2\frac{\log(2d/\delta)}{n}$$

et

$$\left\|\hat{\theta} - \theta\right\|_2^2 \lesssim s\sigma^2\frac{\log(2d/\delta)}{n}.$$

1.2 La complétion matricielle et tensorielle

Après l'expansion des travaux de Compressed Sensing dans les problèmes inverses, les problèmes de complétion de matrice ou de tenseur est devenu l'objet d'une activité de recherche intensive. Les problèmes de complétion matricielle ont des applications dans le filtrage collaboratif, l'apprentissage automatique, les réseaux de capteurs, le traitement du signal, etc. L'objectif est de compléter une matrice de dimension potentiellement importante basée sur un petit (et potentiellement bruité) sous-ensemble de ses entrées (Srebro et al. en 2005, Candès et Plan en 2010). Une application populaire consiste à construire des systèmes de recommandation automatiques, où les lignes correspondent à des utilisateurs, les colonnes à des items et les entrées peuvent être des notes. L'objectif est alors de prévoir les préférences d'utilisateur à partir d'un sous-ensemble des entrées.

Le problème est difficile à traiter si la matrice à récupérer n'a pas de structure particulière. Une découverte importante dans les littératures est que lorsque la matrice à récupérer a un rang faible, alors elle peut être récupérée sur la base de quelques observations seulement. Dans l'application aux systèmes de recommandation, l'hypothèse de rang faible est judicieuse car il est communément admis que seul un petit nombre de facteurs contribue aux préférences de l'utilisateur. La structure de rang faible implique donc une sorte de collaboration entre les différents utilisateurs/items.

On observe généralement une version bruitée Y_{ij} de certaines entrées $(i,j) \in \Omega$ où $\Omega \subset \{1,\dots,m\} \times \{1,\dots,n\}$. Pour $(i,j) \in \Omega$

$$Y_{ij} = \operatorname{tr}\left(X_{ij}^\top Z_{ij}^*\right) + \varepsilon_{ij} \tag{1.4}$$

où X est une matrice $m \times n$ aléatoire déterministe, Z^* est une matrice $m \times n$ inconnu et $\varepsilon_{ij} \sim \mathcal{N}(0,\sigma^2)$ indépendant et identiquement distribué. La complétion de matrice de rang faible peut être effectuée par la résolution du problème d'optimisation suivant

$$\hat{Z} = \underset{Z}{\operatorname{Argmin}} \frac{1}{2\sigma^2}\sum_{(i,j)\in\Omega}\left(Y_{ij} - \operatorname{tr}\left(X_{ij}^\top Z_{ij}^*\right)\right)^2 + \lambda\ \operatorname{rang}(Z) \tag{1.5}$$

où λ est un paramètre de paramétrisation. Pour un sous-ensemble quelconque Ω, le problème d'optimisation (1.6) est très complexe et beaucoup d'auteurs ont préconisé l'utilisation d'une relaxation convexe de (1.6) (Fazel en 2002, Candès et Plan en 2010), donnant le problème d'optimisation convexe suivant

$$\hat{Z} = \underset{Z}{\text{Argmin}} \frac{1}{2\sigma^2} \sum_{(i,j)\in\Omega} \left(Y_{ij} - \text{tr}\left(X_{ij}^\top Z_{ij}^*\right)\right)^2 + \lambda \, \|Z\|_\otimes \tag{1.6}$$

où $\|Z\|_\otimes$ est la norme nucléaire de Z, soit la somme des valeurs singulières de Z.

Le théorème suivant joue un rôle important en contrôlant l'erreur quadratique moyenne

Theorem 1.2.1 *Soit $\mathcal{A} \subset \mathbb{R}^{m\times n}$ un ensemble quelconque de matrices. Soit la matrice aléatoire*

$$M = \frac{1}{n} \sum_{(i,j)\in\Omega} (Y_{ij}X_{ij} - \mathbb{E}[Y_{ij}X_{ij}])$$

Si $\lambda \geqslant 2\|M\|_\infty$, alors

$$\|\hat{Z} - Z^*\|_2^2 \leqslant \inf_{Z\in\mathcal{A}} \left[\|Z - Z^*\|_2^2 + 2\lambda\|Z\|_\otimes\right] \tag{1.7}$$

La complétion tensorielle est une généralisation naturelle du problème classique de complétion des matrices (Candès et Recht, 2009). De même, elle a un large gamme d'applications, notamment dans les systèmes de recommandation, le traitement des signaux et des images, l'analyse des données en ingénierie et en sciences, l'analyse harmonique, etc. Cependant, contrairement à la complétion de la matrice, il existe un large fossé entre la théorie et la pratique. D'abord, le rang est NP-difficile à calculer en toute généralité (Landsberg, 2012). Deuxièmement, il existe de nombreuses décompositions de valeurs singulières différentes. La décomposition de Tucker (Landsberg, 2012) étend la décomposition utile de la somme de rang un aux tenseurs, mais elle est NP-dure à calculer en général. Une autre SVD a été proposée dans la référence (Gross, 2011) avec l'avantage principal que les vecteurs singuliers généralisés forment une famille orthogonale. Une autre approche très intéressante, appelée t-SVD, a été proposée récemment par (Kilmer, 2013) pour les tenseurs 3D, avec des applications à la reconnaissance des visages et au débrouillage des images, à la tomographie informatisée et à la complétion de données. Dans ce cadre, la décomposition ressemble beaucoup à la SVD en utilisant un tenseur diagonal au lieu d'une matrice diagonale comme pour la SVD 2D standard. Un autre avantage intéressant de la t-SVD est que la norme nucléaire est bien définie et que son sous-différentiel est facile à calculer.

Chapter 2

Mixture model for designs in high dimensional regression and the LASSO

2.1 Introduction

The goal of the present paper is the study of the high dimensional regression problem $y = X\beta + z$, where $X \in \mathbb{R}^{n\times p}$, with $p \gg n$ and $z \sim \mathcal{N}(0, \sigma^2 I)$. This problem has been the subject of an extensive research activity. This high dimensional setting, where more variables are involved than observations, occurs in many different applications such as image processing and denoising, gene expression analysis, time series (filtering) [22], [26], graphical models [25], biochemistry [1], etc. One very popular approach is the Least Angle Shrinkage and Selection Operator (LASSO) introduced in [30] for the purpose of variable selection. The LASSO estimator is given as a solution*, for $\lambda > 0$, of

$$\hat{\beta} = \underset{b\in\mathbb{R}^p}{\text{Argmin}} \ \frac{1}{2}\|y - Xb\|_2^2 + \lambda \ \|b\|_1. \tag{2.1}$$

The main advantage of the LASSO over more traditional penalized likelihood optimization procedures such as BIC, AIC, etc, is that a solution can be obtained in polynomial time by solving a convex optimisation problem. Very efficient scalable algorithms are available, based on Nesterov's method [2], the Alternating Direction of Method of Multipliers [28], the Frank-Wolfe algorithm [21] or online versions of them [23]. Extensions to more general statistical estimation problems is proposed in e.g. [32], [15]. Extensions to Lagrange duality based sequential ℓ_1-penalised schemes have been proposed in [14].

One of the most surprising and important discoveries is that, under appropriate assumptions on the design matrix X, and for at least most regression vectors β, the support of β can be recovered exactly when its size is up to the order of $n/\log(p)$ and the nonzero components are sufficiently large; see [3], [6], [10], [34] for instance. Moreover, under similar assumptions, the prediction error can be controlled adaptively as a function of the sparsity of β and the noise variance; see for instance [10].

*Conditions for uniqueness of the minimizer in this last expression are discussed in [20], [27] and [19]

A great amount of work has been devoted to finding error bounds on $X\beta$ [10], [13], [35], etc. Oracle inequalities for this problem are divided into two different classes depending on the so called "regime": the first class describes the slow rate regime and does not require any particular assumption on X, while the second class describes the fast rate regime which does require some structural assumptions on X.

For simplicity, we will assume throughout this paper that the columns of X have unit ℓ_2-norm.

Historically, the first is the Restricted Isometry Property [11] [9], which requires that

$$(1-\delta)\|\beta_S\|_2^2 \leq \quad \|X_S\beta_S\|_2^2 \quad \leq (1+\delta)\|\beta_S\|_2^2, \tag{2.2}$$

for $S \subset \{1,\ldots,n\}$ with $|S| = s'$ and all $\beta \in \mathbb{R}^p$. This property is satisfied by high probability for most random matrices with i.i.d. entrees with variance $1/n$ [†] such as Gaussian or Rademacher variables and for $s' \leq C_{rip}\ n/\log(p)$, where the constant C_{rip} depends on the distribution of the individual entrees. RIP has been extensively used in signal processing after the emergence of the so-called Compressed Sensing paradigm [8].

A second assumption which is often considered is the Incoherence Condition, which requires that

$$\mu(X) \quad = \quad \max_{j\neq j'=1\ldots,p} |\langle X_j, X_{j'}\rangle|$$

is small, e.g. $\mu(X) \leq C_\mu/\log(p)$ as in [10], which is garanteed for random matrices with i.i.d. gaussian entrees with variance $1/n$ in the range $n \geq C_{ic}\log(p)^3$.

The main advantage of the Incoherence Condition over the Restricted Isometry Property is that it can be checked in $p(p-1)/2$ operations, whereas RIP is NP-hard to verify. The main relationship between the Incoherence Condition and RIP is that under the Incoherence Condition, (2.2) holds, not for all, but for most supports $S \subset \{1,\ldots,n\}$ with cardinal s', where $s' \leq C_s\ p/(\|X\|\log(p))$, for some constant C_s controlling the proportion of such supports.

The objective of the present paper is to extend the analysis based on the Incoherence Condition to more general situations where X may have a lot of very colinear columns. The main idea is to assume that the columns are drawn from a mixture model of K clusters, and that the set of centers forms a (usually) much smaller matrix for which is it quite realistic to impose the Incoherence Condition.

2.2 Main results on the LASSO

In this section, we summarise the main results on the LASSO.

[†] the $1/n$ assumption on the variance and standard concentration bounds imply that the resulting random matrix has almost normalized columns

2.2.1 Background

We will study the linear regression model

$$y = X\beta + z \tag{2.3}$$

where $y \in \mathbb{R}^n$ is the data, $X \in \mathbb{R}^{n\times p}$ is the matrix of explanatory variables, $\beta \in \mathbb{R}^p$ is the parameter of interest and z is a centered gaussian i.i.d noise vector $\mathcal{N}(0, \sigma^2 I)$.

The LASSO estimator of β is defined by any $\hat{\beta}$ such that

$$\hat{\beta} = \underset{b\in\mathbb{R}^p}{\text{Argmin}} \frac{1}{2}\|y - Xb\|_2^2 + \lambda\|b\|_1. \tag{2.4}$$

Lemma 2.2.1 *The LASSO estimator obeys*

$$\|X^t\left(y - X\hat{\beta}\right)\|_\infty \leqslant \lambda. \tag{2.5}$$

2.2.2 Statistical viewpoint

The LASSO estimator has been the subject of intense research in the recent years in the statistics community. Several results have been obtained about the mean squared error. The first result below is about the case where no specific assumption is required about X.

Theorem 2.2.2 *Assume that the linear model (2.3) holds where $z \sim subG(\sigma^2)$. Moreover, assume that the columns of X are normalized in such a way that $\max_j \|X_j\|_2 \leqslant \sqrt{n}$. Then, the Lasso estimator $\hat{\theta}$ with regularization parameter*

$$2\lambda = 2\sigma\sqrt{\frac{2\log(2p)}{n}} + 2\sigma\sqrt{\frac{2\log(1/\delta)}{n}}$$

satisfies

$$\frac{1}{n}\left\|X\hat{\beta} - X\beta\right\|_2^2 \leqslant 4\|\beta\|_1\sigma\sqrt{\frac{2\log(2p)}{n}} + 4\|\beta\|_1\sigma\sqrt{\frac{2\log(1/\delta)}{n}}$$

with probability at least $1-\delta$.

The next theorem states that when X satisfies some incoherence-type assumption, more can be obtained for the LASSO estimator and the mean squared error decreases faster.

Assumption Incoherence IN(s)

We say that the design matrix X has incoherence s for some integer $s > 0$ if

$$\|\frac{X^\top X}{n} - I_d\|_\infty \leqslant \frac{1}{32s}.$$

Theorem 2.2.3 *Fix $n \geqslant 2$. Assume that the linear model holds where $z \sim \mathrm{subG}_n(\sigma^2)$. Moreover, assume that $\|\beta\|_0 \leqslant s$ and that X satisfies assumption* $\mathrm{INC}(s)$. *Then the Lasso estimator $\hat{\beta}$ with regularization parameter defined by*

$$2\lambda = 8\sigma\sqrt{\frac{\log(2d)}{n}} + 8\sigma\sqrt{\frac{\log(1/\delta)}{n}}$$

satisfies

$$\frac{1}{n}\left\|X\hat{\beta} - X\beta\right\|_2^2 \lesssim s\sigma^2\frac{\log(2d/\delta)}{n}$$

and

$$\left\|\hat{\beta} - \beta\right\|_2^2 \lesssim s\sigma^2\frac{\log(2d/\delta)}{n}$$

with probability at least $1-\delta$.

In this paper, our goal is to extend this last result to the case where the design matrix has potentially many almost co-linear columns, using a mixture model as a generating model for the columns.

2.3 Our mixture model and a sketch of our main result

2.3.1 The mixture model

In order to relax the Incoherence Condition, one needs a model for the design matrix X allowing for a certain amount of almost parallel columns while keeping some of the algebraic structure in the same spirit as in (2.2) for at least most supports indexing a subset of relevant covariates. In what follows, we study such a model, where the columns can be considered as drawn from a finite mixture of K Gaussian distributions.

An important parameter for the theoretical analysis is a separation index for the centers in the mixture model. This separation index we chose to study in this work is simply the coherence of the matrix of centers which is much smaller than the original design matrix X.

2.3.1.1 Detailed presentation

Let K be the number of clusters of our model. Consider a matrix $\mathfrak{C}$ in $\mathbb{R}^{n\times K}$. The columns $\mathfrak{C}_k$, $k = 1, \ldots, K$ of the matrix $\mathfrak{C}$ are the "centers" of each cluster.

In our model,

- the matrix X is obtained as follows.
  - Choose n_k, $k = 1, \ldots, K$.

– Let $X_o \in \mathbb{R}^{n\times p}$ be a random matrix with independent columns such that the n_1 columns follow the distribution ϕ_1, the n_2 next columns follow the distribution ϕ_2, etc, where

$$\phi_k(x) \;=\; \frac{1}{(2\pi\mathfrak{s}^2)^{\frac{n}{2}}}\exp\left(-\frac{\|x-\mathfrak{C}_k\|_2^2}{2\mathfrak{s}^2}\right).$$

For each $j \in \{1,\ldots,p\}$, $k_j \in \{1,\ldots,K\}$ will denote the index of the Gaussian component from which columns j was drawn, and $\mathcal{J}_k$ will denote the set of indices of the columns drawn from the k^{th} Gaussian component. For any $S \subset \{1,\ldots,p\}$, k_S will denote the subset of $\{1,\ldots,K\}$ indexing the centers of the distributions from which the columns of X_S were drawn.

– The matrix X is obtained by a random permutation of the columns of X_o and column-wise ℓ_2-normalization.

- the support of β will drawn at random as follows.

– We will assume that the support T of the true regression vector β is drawn in such a way that T has the uniform distribution on the subsets of $\{1,\ldots,K\}$ with cardinality equal to s.

2.3.1.2 Best approximation of the class centers and projection of β

For any index set $S \subset \{1,\ldots,p\}$, let $\mathcal{K}_S$ denote the list (with possible repetitions)

$$\mathcal{K}_S \;=\; \{k_j \mid j \in S\}.$$

For each $j \in \{1,\ldots,p\}$, the deviation of column $X_{o,j}$ from center $\mathfrak{C}_{k_j}$ will be denoted by ε_j:

$$\varepsilon_j \;=\; X_{o,j} - \mathfrak{C}_{k_j}. \sim \mathcal{N}\left(0, \mathfrak{s}^2 I\right).$$

The matrix E is defined as

$$E \;=\; (\varepsilon_{i,j})_{i\in\{1,\ldots,n\},\ j\in\{1,\ldots,p\}}.$$

For each $k \in \{1,\ldots,K\}$, let j_k^* be the best approximation of the center $\mathfrak{C}_k$ from the set of columns X_j, $j \in \mathcal{J}_k$, i.e.

$$j_k^* \;=\; \underset{j\in\mathcal{J}_k}{\operatorname{Argmin}} \|X_j - \mathfrak{C}_k\|_2.$$

Moreover, set

$$T^* \;=\; \{j_k^* \mid k = 1,\ldots,K\}.$$

Notice that in particular, $|T^*| = s^*$.

Let β^* be the vector defined as

$$\mathfrak{C}_{\mathcal{K}_{T^*}}\beta^*_{T^*} \;=\; \mathfrak{C}_{\mathcal{K}_T}\beta_T. \tag{2.6}$$

A simple expression of β^* can be obtained by taking

$$\beta^*_{j^*} \;=\; \sum_{j\in J_{k_{j^*}}\cap T} \beta_j \tag{2.7}$$

for all $j^* \in T^*$. Moreover, this expression is unique whenever X_{T^*} has rank equal to s^*[‡].

2.3.2 Main result

The following theorem shows a bound on the prediction error which is a function of the sparsity s^*, the number n of observations, the number of columns p.

Theorem 2.3.1 *(Sketch) Let $\lambda = 2\sigma\sqrt{2\alpha\ \log(p)}$. Assume that X is drawn from the Gaussian mixture model of Section 2.3.1. Then, for p sufficiently large, with probability at least $1 - C_\alpha(\rho^\alpha + p^{-\alpha})$, we have*

$$\frac{1}{2}\|X\left(\hat{\beta}-\beta\right)\|_2^2 \le \frac{3}{2}\lambda s^* \frac{1}{1-r^*_{\alpha,n,\rho}(r)}\left(\frac{3}{2}\ \lambda + \sqrt{3}\|X\left(\beta^*-\beta\right)\|_2\right) + \frac{1}{2}\|X\left(\beta^*-\beta\right)\|_2^2.$$

with

$$r^*_{\alpha,n,\rho}(r) = r\left(\frac{1}{2}+0.1\ C_{\alpha,n,\rho}\right)\left(2+\frac{1}{2}r+0.1\ r\ C_{\alpha,n,\rho}\right) \tag{2.8}$$

where $C_{\alpha,n,\rho} = \sqrt{\alpha} + \sqrt{\frac{\log(n)}{\log(\rho^{-1})}}$.

2.4 A general result and its proof

Some parts of the proof closely follow the key arguments in the proof of [10, Theorem 1.2], although many details of the needed adaptation are nontrivial. Our Theorem 2.4.3 below contains the most general statement of our work.

2.4.1 A more general result

We will require a set of assumptions that are described below.

[‡]In Section 4.2, we will show that X_{T^*} is indeed non-singular with high probability under appropriate assumptions on T

2.4.1.1 Assumptions

In the sequel $\alpha \geq 1$ and r will denote a constant in $(0, 1/2)$. The constants ϑ_* et ν will be specified in Assumptions 2.4.2 below. The constants C_μ, C_{spar} et C_{col} will be used in the Assumptions below:

$$C_\mu \ = r/(1+\alpha),\ C_{\text{spar}} = r^2/((1+\alpha)e^2),\ C_{col} = \ \frac{1}{2}\left(\frac{\sqrt{2}}{\sqrt{(1-r)(1+\alpha)}} - (1+r)\right).$$

Let C_χ denote a positive constant such that

$$\mathbb{P}\left(\frac{\|G\|_2^2}{\mathfrak{s}^2} \leq u^2\right) \ \leq \ C_\chi \left(\frac{u^2}{n}\right)^n$$

where G is a n-dimensional centered and unit-variance i.i.d. gaussian vector.

We will make the following assumptions.

Assumptions 2.4.1 *The matrix $\mathfrak{C}$ has a small coherence, i.e. $\mu(\mathfrak{C})$ should satisfying*

$$\mu(\mathfrak{C}) \leq \frac{C_\mu}{\log(K)} \tag{2.9}$$

for some positive constant C_μ.

Assumptions 2.4.2 *The clusters must containt sufficiently many points, i.e. there exists a positive real constant ϑ^* and a positive integer ν such that*

$$\min_{j^* \in T^*} |\mathcal{J}_{k_{j^*}}| \geq \vartheta_* \log(p)^\nu. \tag{2.10}$$

Assumptions 2.4.3 *The proxy β^* must be sufficiently sparse, i.e.*

$$s^* \ \leq \ K_o \, \frac{K}{\log K} \, \frac{C_{\text{spar}}}{\|\mathfrak{C}\|^2}$$

for some positive constant C_{spar} and $K_0 \leq \rho^{-1}$ for some $\rho \in (0, 1)$.

Assumptions 2.4.4 *The number of columns of $\mathfrak{C}$ satisfying*

$$K \leqslant C_{\text{K}} \log(p)$$

for some positive constant C_{K}.

Assumptions 2.4.5 *One must have sufficiently many observations, i.e.*

$$n \ \geq \ \frac{\alpha+1}{c} \log(p) \tag{2.11}$$

for some positive constant c.

Remark 2.4.1 *The number of observations is both controlled by Assumption 2.4.5 and Assumption 2.4.1 on the coherence of* $\mathfrak{C}$*. For instance, if* $\mathfrak{C}$ *comes from a Gaussian i.i.d. random matrix, the coherence will be of the order* $\sqrt{\log(K)/n}$ *as discussed in [10, Section 1.1] and by Assumption 2.4.1,* n *should be at least of the order* $\log(K)^3/C_\mu^2$*. Notice that this is still less than if* X *itself had to satisfy the coherence bound, which would imply that* n *be of the order* $\log(p)^3$*. This demontrates the advantage of using our Gaussian Mixture framework over the standard framework based on incoherence on* X*.*

Assumptions 2.4.6 *The variance inside the clusters must be sufficiently small, so that the clusters are well separated. More precisely, we will require that*

$$\mathfrak{s} \quad \leq \min\left\{\frac{\alpha}{2\sqrt{n}}; \frac{C_{\mathfrak{s},n,p}}{\sqrt{\log(\rho^{-1})}\left(\sqrt{n}+\sqrt{\frac{\alpha+1}{c}\log(p)}\right)}\right\}$$

for any $C_{\mathfrak{s},n,p}$ *such that and*

$$C_{\mathfrak{s},n,p} \quad \leq \quad \min\left\{0.1 \cdot \frac{r}{\sqrt{n\ \left(\frac{\alpha\ (1-e^{-1})}{\vartheta_*\ C_\chi}\right)^{\frac{1}{n}}\left(\frac{1}{\log(p)^{\nu-1}}\right)^{\frac{1}{n}}}}; \frac{1}{2}\sqrt{\log(p)}\right\}.$$

Assumptions 2.4.7 *The support of* $\beta^*_{T^*}$ *is suffiently generic. More precisely, we will require that the support of* $\beta*$ *is random and uniformly distributed among subsets of* $\{1,\dots,p\}$ *with cardinal* s^**. The sign of* $\beta^*_{T^*}$ *is random with uniform distribution on* $\{-1,1\}^{s^*}$*.*

Remark 2.4.2 *This last assumption is a transposition to the proxy* β^* *of the conditions on* β *in [10].*

Assumptions 2.4.8 *Relationships between the constants.*

$$C_{col} \quad \geq \quad e^2(\alpha+1)\max\{\sqrt{C_{spar}}, C_\mu\}.$$

Assumptions 2.4.9 *Assume that*

$$p \leq 0.01 \cdot \rho^{1-(\alpha+1)\log(K)^2} \tag{2.12}$$

for the same ρ *as in Assumption 2.4.3.*

2.4.1.2 The general theorem

The main result of this paper is the following theorem.

Theorem 2.4.3 *Let* $\lambda = 2\sigma\sqrt{2\alpha\ \log(p)}$*. Assume that* X *is drawn from the Gaussian mixture model of Section 2.3.1 with* $\mathcal{K}$ *drawn uniformly at random among all possible index subsets of* $\{1,\dots,K\}$ *with cardinal* s^**. Let Assumptions 2.4.1, 2.4.2, 2.4.3, 2.4.4, 2.4.5,*

2.4.6, 2.4.7, 2.4.8 and 2.4.9 hold. Then, for all $r \in (0, 0.5)$, with probability at least $1 - C_\alpha(\rho^\alpha + p^{-\alpha})$, we have

$$\frac{1}{n}\|X\left(\hat{\beta}-\beta\right)\|_2^2 \leq \frac{3}{n}\lambda s^* \frac{1}{1-r^*_{\alpha,n,\rho}(r)}\left(\frac{3}{2}\lambda + \sqrt{1+r^*_{\alpha,n,\rho}(r)}\|X\left(\beta^*-\beta\right)\|_2\right) + \frac{1}{n}\|X\left(\beta^*-\beta\right)\|_2^2$$

with

$$r^*_{\alpha,n,\rho}(r) = r\left(\frac{1}{2} + 0.1\ C_{\alpha,n,\rho}\right)\left(2 + \frac{1}{2}r + 0.1\ r\ C_{\alpha,n,\rho}\right) \tag{2.13}$$

where [§] $C_{\alpha,n,\rho} = \sqrt{\alpha} + \sqrt{\frac{\log(n)}{\log(\rho^{-1})}}$.

Remark 2.4.4 *Notice that our result is of fast rate type but includes new additional terms involving the approximation error $\beta^* - \beta$. More precisely, the right hand side in* (2.13) *can be decomposed into two parts:*

- *the term*

$$\frac{9}{2n}\lambda^2 s^* \frac{1}{1-r^*_{\alpha,n,\rho}(r)},$$

  which is similar to the "fast rate term" in the standard incoherent case of Theorem 2.2.3.

- *the term*

$$\frac{3}{2}\frac{\lambda s^*}{1-r^*_{\alpha,n,\rho}(r)}\sqrt{1+r^*_{\alpha,n,\rho}(r)}\|X\left(\beta^*-\beta\right)\|_2 + \frac{1}{n}\|X\left(\beta^*-\beta\right)\|_2^2$$

  is not present in the standard analysis of the LASSO and depends on how well β can be approximated by β^*, and depends on the model and more precisely $\mathfrak{C}$ and β.

We now begin the proof of Theorem 2.4.3.

2.4.2 Preliminaries: Candès and Plan's conditions

The following proposition will be much used in the arguments.

Proposition 2.4.5 *We have the following properties:*

1.

$$\mathbb{P}\left(\|\mathfrak{C}^t_{\mathcal{K}_T}\mathfrak{C}_{\mathcal{K}_T} - \mathrm{Id}_s\| \geq \rho_{\mathfrak{C}_{\mathcal{K}_T}}\right) \leq \frac{216}{p^\alpha}. \tag{2.14}$$

[§] $\rho \in (0,1)$ is introduced in Assumption 2.4.9

2.

$$\mathbb{P}\left(\left\|X_T^t X_T - I\right\| \geq r^*_{\alpha,n,p}(r)\right) \leq \frac{219}{p^\alpha}. \tag{2.15}$$

where $r^*_{\alpha,n,p}(r)$ is defined by (2.8)

3.

$$\mathbb{P}\left(\|X^t z\|_\infty \geq \sigma\sqrt{2\alpha\ \log(p)}\right) \leq \frac{1}{p^\alpha}. \tag{2.16}$$

4.

$$\begin{aligned}\left\|X_{T^{*c}}^t X_{T^*}(X_{T^*}^t X_{T^*})^{-1} X_{T^*}^t z\right\|_\infty + \lambda\ \left\|X_{T^{*c}}^t X_{T^*}(X_{T^*}^t X_{T^*})^{-1}\mathrm{sgn}\left(\beta^*_{T^*}\right)\right\|_\infty \\ \leq \sigma\mathcal{C}_1 + \lambda\mathcal{C}_2\end{aligned} \tag{2.17}$$

where $\mathcal{C}_1$ and $\mathcal{C}_2$ are defined by (2.48) and (2.49).

Proof: See Appendix 2.5. □

2.4.3 The prediction bound

By definition, the LASSO estimator satisfies

$$\frac{1}{2}\|y - X\hat{\beta}\|_2^2 + \lambda\|\hat{\beta}\|_1 \ \leq\ \frac{1}{2}\|y - X\beta^*\|_2^2 + \lambda\|\beta^*\|_1. \tag{2.18}$$

One may introduce $X\beta$ in this expression and obtain

$$\frac{1}{2}\|y - X\beta + X(\beta - \hat{\beta})\|_2^2 + \lambda\|\hat{\beta}\|_1 \ \leq\ \frac{1}{2}\|y - X\beta + X(\beta - \beta^*)\|_2^2 + \lambda\|\beta^*\|_1,$$

from which we deduce

$$\begin{aligned}\frac{1}{2}\|X(\beta - \hat{\beta})\|_2^2 \ &\leq \langle y - X\beta, X(\hat{\beta} - \beta^*)\rangle - \lambda\left(\|\hat{\beta}\|_1 - \|\beta^*\|_1\right) \\ &\quad + \tfrac{1}{2}\|X(\beta - \beta^*)\|_2^2.\end{aligned} \tag{2.19}$$

Set $h^* := \hat{\beta} - \beta^*$. Using sparsity of β^*, we obtain that $h^*_{T^{*c}} = \hat{\beta}_{T^{*c}} - \beta^*_{T^{*c}} = \hat{\beta}_{T^{*c}}$. Thus, we have

$$\begin{aligned}\|\hat{\beta}\|_1 - \|\beta^*\|_1 \ &= \ \|\beta^* + h^*\|_1 - \|\beta^*\|_1 \\ &= \ \|{\beta^*}_{T^*} + {h^*}_{T^*}\|_1 + \|{\beta^*}_{T^{*c}} + {h^*}_{T^{*c}}\|_1 - \|{\beta^*}_{T^*}\|_1 \\ &= \ \|{\beta^*}_{T^*} + {h^*}_{T^*}\|_1 - \|{\beta^*}_{T^*}\|_1 + \|{h^*}_{T^{*c}}\|_1.\end{aligned}$$

Since, for any b with no zero component, the gradient of $\|\cdot\|_1$ at b is $\mathrm{sgn}(b)$, the subgradient inequality gives

$$\|\beta^*_{T^*} + h^*_{T^*}\|_1 \;\geq\; \|\beta^*_{T^*}\|_1 + \langle \mathrm{sgn}\left(\beta^*_{T^*}\right), h^*_{T^*}\rangle$$

and combining this latter inequality with (2.19), we obtain

$$\frac{1}{2}\|X(\beta-\hat{\beta})\|_2^2 \leq \langle y - X\beta, Xh^*\rangle - \lambda \langle \mathrm{sgn}\left(\beta^*_{T^*}\right), h^*_{T^*}\rangle - \lambda\|h^*_{T^{*c}}\|_1 + \frac{1}{2}\|X(\beta-\beta^*)\|_2^2. \tag{2.20}$$

Set $\gamma := \beta^* - \beta$ and $h := \hat{\beta} - \beta$. Using these notations, equation (2.20) may be written

$$\frac{1}{2}\|Xh\|_2^2 \;\leq\; \langle z, Xh^*\rangle - \lambda \langle \mathrm{sgn}\left(\beta^*_{T^*}\right), h^*_{T^*}\rangle - \lambda\|h^*_{T^{*c}}\|_1 + \frac{1}{2}\|X\gamma\|_2^2. \tag{2.21}$$

Using the fact that

$$\langle X^t z, h^*\rangle \;=\; \langle X^t_{T^*} z, h^*_{T^*}\rangle + \langle X^t_{T^{*c}} z, h^*_{T^{*c}}\rangle$$

and the following majorization based on (2.16)

$$\begin{aligned}\langle X^t_{T^{*c}} z, h^*_{T^{*c}}\rangle &\leq \|h^*_{T^{*c}}\|_1 \|X^t_{T^{*c}} z\|_\infty \\ &\leq \frac{1}{2}\,\lambda\,\|h^*_{T^{*c}}\|_1,\end{aligned}$$

we obtain that

$$\frac{1}{2}\|Xh\|_2^2 \;\leq\; \langle v, h^*_{T^*}\rangle - (1-\frac{1}{2})\,\lambda\|h^*_{T^{*c}}\|_1 + \frac{1}{2}\|X\gamma\|_2^2, \tag{2.22}$$

where $v := X^t_{T^*} z - \lambda\,\mathrm{sgn}\left(\beta^*_{T^*}\right)$.

Now, observe that

$$\begin{aligned}\langle v, h^*_{T^*}\rangle &= \langle v, (X^t_{T^*}X_{T^*})^{-1} X^t_{T^*}X_{T^*} h^*_{T^*}\rangle \\ &= \langle (X^t_{T^*}X_{T^*})^{-1} v, X^t_{T^*}X_{T^*} h^*_{T^*}\rangle \\ &= \underbrace{\langle (X^t_{T^*}X_{T^*})^{-1} v, X^t_{T^*}Xh^*\rangle}_{A_1} - \underbrace{\langle (X^t_{T^*}X_{T^*})^{-1} v, X^t_{T^*}X_{T^{*c}} h^*_{T^{*c}}\rangle}_{A_2}.\end{aligned}$$

Let us begin by studying A_2. We have that

$$\begin{aligned}A_2 &\geq -\|X^t_{T^{*c}}X_{T^*}(X^t_{T^*}X_{T^*})^{-1} v\|_\infty \|h^*_{T^{*c}}\|_1 \\ &\geq -\|X^t_{T^{*c}}X_{T^*}(X^t_{T^*}X_{T^*})^{-1} X^t_{T^*} z\|_\infty \|h^*_{T^{*c}}\|_1 \\ &\quad -\lambda\,\|X^t_{T^{*c}}X_{T^*}(X^t_{T^*}X_{T^*})^{-1}\mathrm{sgn}\left(\beta^*_{T^*}\right)\|_\infty \|h^*_{T^{*c}}\|_1 \\ &\geq -\left(\sigma\mathcal{C}_1 + \lambda\mathcal{C}_2\right)\|h^*_{T^{*c}}\|_1\end{aligned}$$

by (2.17). Thus

$$\langle v, h^*_{T^*}\rangle \;\leq\; A_1 + \left(\sigma\mathcal{C}_1 + \lambda\mathcal{C}_2\right)\|h^*_{T^{*c}}\|_1$$

and we deduce that

$$\frac{1}{2}\|Xh\|_2^2 \;\leq\; A_1 + \left(\sigma\mathcal{C}_1 + \lambda\mathcal{C}_2 - \frac{1}{2}\lambda\right)\|h^*_{T^{*c}}\|_1 + \frac{1}{2}\|X\gamma\|_2^2 \tag{2.23}$$

Let us now bound A_1 from above. We have that

$$A_1 \;\leq\; \underbrace{\|X^t_{T^*}Xh^*\|_\infty}_{B_1}\underbrace{\|(X^t_{T^*}X_{T^*})^{-1}v\|_1}_{B_2}$$

Firstly,

$$\begin{aligned} B_1 &\leq \|X^t_{T^*}(X\beta^* - y)\|_\infty + \|X^t_{T^*}(X\hat{\beta} - y)\|_\infty \\ &\leq \|X^t_{T^*}(X\gamma - z)\|_\infty + \|X^t_{T^*}(y - X\hat{\beta})\|_\infty \\ &\leq \frac{1}{2}\lambda + \|X^t_{T^*}X\gamma\|_\infty + \lambda \end{aligned}$$

where we used (2.16), and the optimality condition for the LASSO estimator ((2.5)). Secondly,

$$\begin{aligned} B_2 &\leq \sqrt{s^*}\|(X^t_{T^*}X_{T^*})^{-1}v\|_2 \\ &\leq \sqrt{s^*}\|(X^t_{T^*}X_{T^*})^{-1}\|\|v\|_2 \\ &\leq s^*\|(X^t_{T^*}X_{T^*})^{-1}\|\|v\|_\infty. \end{aligned}$$

Moreover, (2.15) gives $\|\left(X^t_{T^*}X_{T^*}\right)^{-1}\| \leq \frac{1}{1-r^*_{\alpha,n,p}(r)}$

$$\|v\|_\infty \leq \|X^t_{T^*}z\|_\infty + \lambda \leq \frac{3}{2}\,\lambda$$

Thus, we obtain that

$$A_1 \;\leq\; \frac{3}{2}\,\lambda\,s^*\,\frac{1}{1 - r^*_{\alpha,n,p}(r)}\left(\frac{3}{2}\,\lambda + \|X^t_{T^*}X\gamma\|_\infty\right)$$

and thus,

$$\begin{aligned} \frac{1}{2}\|Xh\|_2^2 \qquad &\leq \tfrac{3}{2}\,\lambda\,s^*\,\tfrac{1}{1-r^*_{\alpha,n,p}(r)}\left(\tfrac{3}{2}\,\lambda + \|X^t_{T^*}X\gamma\|_\infty\right) \\ &\quad + \left(\sigma\mathcal{C}_1 + \lambda\mathcal{C}_2 - \tfrac{1}{2}\lambda\right)\|h^*_{T^{*c}}\|_1 + \tfrac{1}{2}\|X\gamma\|_2^2. \end{aligned}$$

Since $\|X^t_{T^*}X\gamma\|_\infty \leq \|X^t_{T^*}X\gamma\|_2$ and since

$$\|X^t_{T^*}X\gamma\|_2 \leq \sqrt{1 + r^*_{\alpha,n,p}(r)}\|X\gamma\|_2,$$

we obtain

$$\begin{aligned} \frac{1}{2}\|Xh\|_2^2 \;&\leq \tfrac{3}{2}\lambda s^*\,\tfrac{1}{1-r^*_{\alpha,n,p}(r)}\left(\tfrac{3}{2}\,\lambda + \sqrt{1 + r^*_{\alpha,n,p}(r)}\|X\gamma\|_2\right) \\ &\quad + \left(\sigma\mathcal{C}_1 + \lambda\mathcal{C}_2 - \tfrac{1}{2}\lambda\right)\|h^*_{T^{*c}}\|_1 + \tfrac{1}{2}\|X\gamma\|_2^2. \end{aligned}$$

which completes the proof.

2.5 Checking the Candes-Plan conditions

The goal of this section is to Proposition 2.4.5 which gives a version of Candès and Plan's conditions adapted to our Gaussian mixture model.

2.5.1 Control of $\|E_{T^*}\|$

Consider the matrix $E_{T^*}^t$, whose columns are independent. We would like to bound its operator norm.

Lemma 2.5.1 *Let the event*

$$\mathcal{E}_\alpha^* = \bigcap_{j^* \in T^*} \left\{ \|E_{j^*}\|_2 \leq \mathfrak{s} \sqrt{n \left(\frac{\alpha \ (1 - e^{-1})}{\vartheta_* \ C_\chi} \right)^{\frac{1}{n}} \left(\frac{1}{\log(p)^{\nu - 1}} \right)^{\frac{1}{n}}} \right\}.$$

Then, $\mathbb{P}(\mathcal{E}_\alpha^*) \geq 1 - \rho^\alpha$.

Proof: Using the independence of the E_j, $j \in \mathcal{J}_{k_{j^*}}$, we have

$$\begin{aligned} \mathbb{P}\left(\|E_{j^*}\|_2 \geq u \right) &= \mathbb{P}\left(\min_{j \in \mathcal{J}_{k_{j^*}}} \|E_j\|_2 \geq u \right) \\ &= \prod_{j \in \mathcal{J}_{k_{j^*}}} \mathbb{P}\left(\|E_j\|_2^2 \geq u^2 \right), \\ &\leq \mathbb{P}\left(\|E_j\|_2^2 \geq u^2 \right)^{\min_{j^* \in T^*} |\mathcal{J}_{k_{j^*}}|}. \end{aligned}$$

We also have

$$\mathbb{P}\left(\|E_j\|_2^2 \geq u^2 \right) = 1 - \mathbb{P}\left(\|E_j\|_2^2 \leq u^2 \right).$$

On the other hand, as is well known, we have

$$\mathbb{P}\left(\frac{\|E_j\|_2^2}{\mathfrak{s}^2} \leq u^2 \right) \leq C_\chi \left(\frac{u^2}{n} \right)^n$$

for some positive constant C_χ. Thus, the union bound gives

$$\mathbb{P}\left(\max_{j^* \in T^*} \|E_{j^*}\|_2 \geq u \right) \leq s^* \left(1 - C_\chi \left(\frac{u^2}{n \ \mathfrak{s}^2} \right)^n \right)^{\min_{j^* \in T^*} |\mathcal{J}_{k_{j^*}}|}.$$

Let us tune u so that

$$s^* \left(1 - C_\chi \left(\frac{u^2}{n \ \mathfrak{s}^2} \right)^n \right)^{\min_{j^* \in T^*} |\mathcal{J}_{k_{j^*}}|} \leq \rho^\alpha$$

i.e.

$$u^2 \;\geq\; \frac{n\,\mathfrak{s}^2}{C_\chi^{\frac{1}{n}}}\left(1-\left((s^*)^{-1}\rho^\alpha\right)^{\frac{1}{\min\limits_{j^*\in T^*}|\mathcal{J}_{k_{j^*}}|}}\right)^{\frac{1}{n}}$$

and since $\min_{j^*\in T^*}|\mathcal{J}_{k_{j^*}}| \geq \vartheta_*\log(p)^\nu$ by (2.10),

$$u^2 \;\geq\; \frac{n\,\mathfrak{s}^2}{C_\chi^{\frac{1}{n}}}\left(1-\exp\left(-\frac{\alpha}{\vartheta_*\log(p)^{\nu-1}}-\frac{\log(s^*)}{\vartheta_*\log(p)^\nu}\right)\right)^{\frac{1}{n}}. \tag{2.24}$$

On $(0,1)$, we have

$$\exp(-z) \;\leq\; 1-(1-e^{-1})z$$

and thus,

$$u^2 \;\geq\; n\,\mathfrak{s}^2\,\left(\frac{\alpha\,(1-e^{-1})}{\vartheta_*\,C_\chi}\right)^{\frac{1}{n}}\left(\frac{1}{\log(p)^{\nu-1}}\right)^{\frac{1}{n}},$$

from which the desired estimate follows. □

Lemma 2.5.2 *We have*

$$\mathbb{P}\left(\left\|E_{T^*}^t\right\| \geq \mathfrak{s}K_{n,s^*} \mid \mathcal{E}_\alpha^*\right) \;\leq\; \frac{2}{p^\alpha}$$

where

$$K_{n,s^*} = \sqrt{n\left(\alpha\log(\rho^{-1})+\log(n)\right)\left(\frac{\alpha\,(1-e^{-1})}{\vartheta_*\,C_\chi}\right)^{\frac{1}{n}}\left(\frac{1}{\log(p)^{\nu-1}}\right)^{\frac{1}{n}}}. \tag{2.25}$$

Proof: Let us first notice that since $\|E_{T^*}\| = \|E_{T^*}^t\|$, we can write

$$\begin{aligned}\|E_{T^*}^t\| &= \sqrt{\|E_{T^*}E_{T^*}^t\|}\\ &= \sqrt{\left\|\sum_{j^*\in T^*}E_{j^*}E_{j^*}^t\right\|}\end{aligned}$$

This latter expression is well suited for our problem, since it is the norm of the sum of independent positive semi-definite random matrices. In order to apply this inequality, we need a bound on the norm of each summand. By Lemma 2.5.1, on $\mathcal{E}^*$, we have

$$\begin{aligned}\left\|E_{j^*}E_{j^*}^t\right\|_2 &= \|E_{j^*}\|_2^2\\ &\leq \mathfrak{s}^2\;n\;\left(\frac{\alpha\,(1-e^{-1})}{\vartheta_*\,C_\chi}\right)^{\frac{1}{n}}\left(\frac{1}{\log(p)^{\nu-1}}\right)^{\frac{1}{n}}.\end{aligned}$$

We also need a bound on the norm of the expectation. We have

$$\left\| \mathbb{E}\left[\sum_{j^* \in T^*} E_{j^*} E_{j^*}^t \mid \mathcal{E}_\alpha^* \right] \right\| \;=\; \left\| \sum_{j^* \in T^*} \mathbb{E}\left[E_{j^*} E_{j^*}^t \mid \mathcal{E}_\alpha^* \right] \right\|.$$

Due to rotational invariance, we have that the law of E_{j^*} is the same as the law of $D(\zeta)E_{j^*}$, where $\zeta_1, \ldots, \zeta_n$ are i.i.d. Rademacher ± 1 random variables independent from E_{j^*}. Thus, for $i \neq i'$,

$$\begin{aligned} \mathbb{E}\left[\zeta_i E_{i,j^*} \zeta_{i'} E_{i',j^*} \mid \mathcal{E}_\alpha^* \right] &= \mathbb{E}\left[\mathbb{E}\left[\zeta_i E_{i,j^*} \zeta_{i'} E_{i',j^*} \mid E_{i,j^*},\ E_{i',j^*} \right] | \mathcal{E}_\alpha^* \right] \\ &= 0. \end{aligned} \tag{2.26}$$

On the other hand, we have the following result.

Lemma 2.5.3 *We have*

$$\mathbb{E}\left[E_{i,j^*}^2 \mid \mathcal{E}_\alpha^* \right] \;\leq\; \mathfrak{s}^2 \left(\frac{\alpha\ (1 - e^{-1})}{\vartheta_*\ C_\chi} \right)^{\frac{1}{n}} \left(\frac{1}{\log(p)^{\nu-1}} \right)^{\frac{1}{n}}.$$

Proof: Due to rotational invariance of the law of E_{j^*} and the event $\mathcal{E}_\alpha^*$, we have

$$\mathbb{E}\left[E_{1,j^*}^2 \mid \mathcal{E}_\alpha^* \right] \;= \cdots = \; \mathbb{E}\left[E_{n,j^*}^2 \mid \mathcal{E}_\alpha^* \right].$$

Therefore,

$$\mathbb{E}\left[E_{i,j^*}^2 \mid \mathcal{E}_\alpha^* \right] \;\leq\; \frac{1}{n}\, \mathbb{E}\left[\sum_{i=1}^n E_{i,j^*}^2 \mid \mathcal{E}_\alpha^* \right]$$

and by the definition of $\mathcal{E}_\alpha^*$,

$$\mathbb{E}\left[E_{i,j^*}^2 \mid \mathcal{E}_\alpha^* \right] \;\leqslant\; \mathfrak{s}^2 \left(\frac{\alpha\ (1 - e^{-1})}{\vartheta_*\ C_\chi} \right)^{\frac{1}{n}} \left(\frac{1}{\log(p)^{\nu-1}} \right)^{\frac{1}{n}}.$$

□

Based on this lemma, and the fact that the matrix

$$\mathbb{E}\left[\sum_{j^* \in T^*} E_{j^*} E_{j^*}^t \mid \mathcal{E}_\alpha^* \right],$$

is diagonal by (2.26), we obviously obtain that

$$\left\| \mathbb{E}\left[\sum_{j^* \in T^*} E_{j^*} E_{j^*}^t \mid \mathcal{E}_\alpha^* \right] \right\| \;\leqslant\; \mathfrak{s}^2 \left(\frac{\alpha\ (1 - e^{-1})}{\vartheta_*\ C_\chi} \right)^{\frac{1}{n}} \left(\frac{1}{\log(p)^{\nu-1}} \right)^{\frac{1}{n}}.$$

With the bound on the norm of the expectation and on the variance in hand and obtain

$$\mathbb{P}\left(\left\|\sum_{j^*\in T^*} E_{j^*}E_{j^*}^t\right\| \geq u \mid \mathcal{E}_\alpha^*\right) \leq n\left(\frac{e\ \mathfrak{s}^2\ \left(\frac{\alpha\ (1-e^{-1})}{\vartheta_*\ C_\chi}\right)^{\frac{1}{n}}\left(\frac{1}{\log(p)^{\nu-1}}\right)^{\frac{1}{n}}}{u}\right)^{\frac{u}{\mathfrak{s}^2\ n\ \left(\frac{\alpha\ (1-e^{-1})}{\vartheta_*\ C_\chi}\right)^{\frac{1}{n}}\left(\frac{1}{\log(p)^{\nu-1}}\right)^{\frac{1}{n}}}}. \quad (2.27)$$

Let us finally tune u so that the right hand side term is less than ρ^α, i.e.

$$\log\left(\frac{e\ \mathfrak{s}^2\ \left(\frac{\alpha\ (1-e^{-1})}{\vartheta_*\ C_\chi}\right)^{\frac{1}{n}}\left(\frac{1}{\log(p)^{\nu-1}}\right)^{\frac{1}{n}}}{u}\right) \leq -\frac{\mathfrak{s}^2\ n\ \left(\frac{\alpha\ (1-e^{-1})}{\vartheta_*\ C_\chi}\right)^{\frac{1}{n}}\left(\frac{1}{\log(p)^{\nu-1}}\right)^{\frac{1}{n}}}{u}\ \left(\alpha\log(\rho^{-1})+\log(n)\right).$$

Take

$$u \ = \ \mathfrak{s}^2\ n\left(\alpha\log(\rho^{-1})+\log(n)\right)\left(\frac{\alpha\ (1-e^{-1})}{\vartheta_*\ C_\chi}\right)^{\frac{1}{n}}\left(\frac{1}{\log(p)^{\nu-1}}\right)^{\frac{1}{n}}. \quad (2.28)$$

Moreover, the value of u given by (2.28) is less than or equal to $\mathfrak{s}^2\ K_{n,s^*}^2$ with

$$K_{n,s^*} = \sqrt{n\left(\alpha\log(\rho^{-1})+\log(n)\right)\left(\frac{\alpha\ (1-e^{-1})}{\vartheta_*\ C_\chi}\right)^{\frac{1}{n}}\left(\frac{1}{\log(p)^{\nu-1}}\right)^{\frac{1}{n}}}. \quad (2.29)$$

This completes the proof. □

2.5.2 Important properties of $\mathfrak{C}$

The invertibily condition for (2.14) is a direct consequence of [31]. An alternative approach, based on the Matrix Chernov inequality is proposed in [17], with improved constants. We have in particular

Theorem 2.5.4 *[17, Theorem 1] Let $r \in (0, 1/2)$, $\alpha \geq 1$. Let Assumptions 2.4.1 and 2.4.3 hold with*

$$\mathrm{C}_{spar} = \frac{r^2}{4(1+\alpha)e^2}\ . \quad (2.30)$$

With $\mathcal{K} \subset \{1, \ldots, K\}$ chosen randomly from the uniform distribution among subsets with cardinality s^*, the following bound holds:

$$\mathbb{P}\left(\|\mathfrak{C}_{\mathcal{K}}^t\mathfrak{C}_{\mathcal{K}} - \mathrm{Id}_s\| \geq r\right) \ \leq \ \frac{216}{p^\alpha}. \quad (2.31)$$

Moreover, the following property will also be very useful.

Lemma 2.5.5 *(Adapted from [17, Lemma 5.3]) If* $v^2 \geq e\ s^*\ \|\mathfrak{C}\|/K_o$, *we have*

$$\mathbb{P}\left(\max_{k\in\mathcal{K}^c}\left\|\mathfrak{C}_{\mathcal{K}}^t\mathfrak{C}_k\right\|_2 \geq \frac{v}{1-r}\right) \;\leq\; K_o\left(e\frac{s^*\ \|\mathfrak{C}\|^2}{K_o\ v^2}\right)^{\frac{v^2}{\mu(\mathfrak{C})^2}}.$$

Based on this lemma, we easily get the following bound.

Lemma 2.5.6 *Take* $C_{col} \geq \sqrt{e^2(\alpha+1)}\max\{\sqrt{C_{spar}}, C_\mu\}$. *Then, we have*

$$\mathbb{P}\left(\max_{k\in\mathcal{K}^c}\left\|\mathfrak{C}_{\mathcal{K}}^t\mathfrak{C}_k\right\|_2 \geq \frac{C_{col}\cdot\sqrt{\log(\rho^{-1})}}{1-r}\right) \;\leq\; \frac{1}{\rho^{1-(\alpha+1)\log(K)^2}}. \tag{2.32}$$

Proof: Taking $v = C_{col}\cdot\sqrt{\log(\rho^{-1})}$, we obtain from Lemma 2.5.5

$$\mathbb{P}\left(\max_{k\in\mathcal{K}^c}\left\|\mathfrak{C}_{\mathcal{K}}^t\mathfrak{C}_k\right\|_2 \geq \frac{C_{col}\cdot\sqrt{\log(\rho^{-1})}}{1-r}\right) \;\leq\; K_o\left(e\frac{s^*\ \|\mathfrak{C}\|^2}{K_o\ C_{col}^2\cdot\log(\rho^{-1})}\right)^{\frac{C_{col}^2}{C_\mu^2}\log(\rho^{-1})\cdot\log(K)^2}.$$

Using Assumption 2.4.3, this gives

$$\mathbb{P}\left(\max_{k\in\mathcal{K}^c}\left\|\mathfrak{C}_{\mathcal{K}}^t\mathfrak{C}_k\right\|_2 \geq \frac{C_{col}\cdot\sqrt{\log(\rho^{-1})}}{1-r}\right) \leq K_o\left(e\ K\ \frac{C_{spar}}{K_o\ \log(K)\ C_{col}^2\cdot\log(\rho^{-1})}\right)^{\frac{C_{col}^2}{C_\mu^2}\log(\rho^{-1})\cdot\log(K)^2}$$

and using Assumption 2.4.4, we have

$$\mathbb{P}\left(\max_{k\in\mathcal{K}^c}\left\|\mathfrak{C}_{\mathcal{K}}^t\mathfrak{C}_k\right\|_2 \geq \frac{C_{col}\cdot\sqrt{\log(\rho^{-1})}}{1-r}\right) \;\leq\; K_o\left(e\ \frac{C_{spar}\ C_K}{K_o\ \log(K)\ C_{col}^2}\right)^{\frac{C_{col}^2}{C_\mu^2}\log(\rho^{-1})\cdot\log(K)^2}.$$

Since $C_{col} \geq \sqrt{e^2(\alpha+1)}\max\{\sqrt{C_{spar}}, C_\mu\}$, we get

$$K_o\left(e\ \frac{C_{spar}\ C_K}{K_o\ \log(K)\ C_{col}^2}\right)^{\frac{C_{col}^2}{C_\mu^2}\log(\rho^{-1})\cdot\log(K)^2}$$
$$\leqslant K_o\left(e\ \frac{C_K}{K_o\ \log(K)\ e^2\ (\alpha+1)}\right)^{e^2(\alpha+1)\log(\rho^{-1})\cdot\log(K)^2}.$$

Thus, we have

$$\mathbb{P}\left(\max_{k\in\mathcal{K}^c}\left\|\mathfrak{C}_{\mathcal{K}}^t\mathfrak{C}_k\right\|_2 \geq \frac{C_{col}\cdot\sqrt{\log(\rho^{-1})}}{1-r}\right) \leq K_o\,\frac{1}{(\rho^{-1})^{(\alpha+1)\log(K)^2}}\cdot\frac{1}{K_o^{(\alpha+1)\log(\rho^{-1})\log(K)^2}}.$$

and since $K_0 \leq \rho^{-1}$ by Assumption 2.4.3, we obtain

$$\mathbb{P}\left(\max_{k\in\mathcal{K}^c}\left\|\mathfrak{C}_{\mathcal{K}}^t\mathfrak{C}_k\right\|_2 \geq \frac{C_{col}\cdot\sqrt{\log(\rho^{-1})}}{1-r}\right) \leq \frac{1}{\rho^{1-(\alpha+1)\log(K)^2}}.$$

□

2.5.3 Similar properties for X_{T^*}

2.5.3.1 Control of $\|X_{T^*}^t X_{T^*} - I\|$

We have

$$\sigma_{\min}\left(X_{T^*}^t X_{T^*}\right) = \sigma_{\min}\left(\left(\mathfrak{C}_{\mathcal{K}_{T^*}} + E_{T^*}\right)^t D_*^2\left(\mathfrak{C}_{\mathcal{K}_{T^*}} + E_{T^*}\right)\right)$$

where D_* is a diagonal matrix whose diagonal elements are indexed by T^* and are defined by

$$D_{*,j^*,j^*} = \frac{1}{\left\|\mathfrak{C}_{k_{j^*}} + E_{j^*}\right\|_2},$$

for $j^* \in T^*$. By the definition of $\mathcal{E}_\alpha^*$, we have

$$\sigma_{\min}(D_*) \geq \frac{1}{1+\mathfrak{s}\sqrt{n\left(\frac{\alpha(1-e^{-1})}{\vartheta_*\ C_\chi}\right)^{\frac{1}{n}}\left(\frac{1}{\log(p)^{\nu-1}}\right)^{\frac{1}{n}}}}.$$

and

$$\sigma_{\max}(D_*) \leq \frac{1}{1-\mathfrak{s}\sqrt{n\left(\frac{\alpha(1-e^{-1})}{\vartheta_*\ C_\chi}\right)^{\frac{1}{n}}\left(\frac{1}{\log(p)^{\nu-1}}\right)^{\frac{1}{n}}}}.$$

By the triangular inequality,

$$\begin{aligned}\sigma_{\min}\left(X_{T^*}^t X_{T^*}\right) &\geq \sigma_{\min}\left(\mathfrak{C}_{\mathcal{K}_{T^*}}^t D_*^2\mathfrak{C}_{\mathcal{K}_{T^*}}\right) - 2\left\|\mathfrak{C}_{\mathcal{K}_{T^*}}^t D_*^2 E_{T^*}\right\| - \left\|E_{T^*}^t D_*^2 E_{T^*}\right\| \\ &\geq \frac{1-r}{\left(1+\mathfrak{s}\sqrt{n\left(\frac{\alpha(1-e^{-1})}{\vartheta_*\ C_\chi}\right)^{\frac{1}{n}}\left(\frac{1}{\log(p)^{\nu-1}}\right)^{\frac{1}{n}}}\right)^2} \\ &\quad - \frac{2\sqrt{1+r}\ \|E_{T^*}\| + \|E_{T^*}\|^2}{\left(1-\mathfrak{s}\sqrt{n\left(\frac{\alpha(1-e^{-1})}{\vartheta_*\ C_\chi}\right)^{\frac{1}{n}}\left(\frac{1}{\log(p)^{\nu-1}}\right)^{\frac{1}{n}}}\right)^2}.\end{aligned}$$

and

$$\begin{aligned}\sigma_{\max}\left(X_{T^*}^t X_{T^*}\right) &\leq \left\|\mathfrak{C}_{\mathcal{K}_{T^*}}^t D_*^2 \mathfrak{C}_{\mathcal{K}_{T^*}}\right\| + 2\left\|\mathfrak{C}_{\mathcal{K}_{T^*}}^t D_*^2 E_{T^*}\right\| + \left\|E_{T^*}^t D_*^2 E_{T^*}\right\| \\ &\leq \frac{(1+r)+2\sqrt{1+r}\ \|E_{T^*}\| + \|E_{T^*}\|^2}{\left(1-\mathfrak{s}\sqrt{n\left(\frac{\alpha(1-e^{-1})}{\vartheta_*\ C_\chi}\right)^{\frac{1}{n}}\left(\frac{1}{\log(p)^{\nu-1}}\right)^{\frac{1}{n}}}\right)^2}.\end{aligned}$$

Moreover, using Theorem 2.5.4 and Lemma 2.5.2, we obtain

$$\mathbb{P}\left(\left\|X_{T^*}^t X_{T^*} - I\right\| \geq r^* \mid \mathcal{E}_\alpha^*\right) \leq \frac{218}{p^\alpha}$$

with r^* given by

$$\begin{aligned}r^* = \max\Bigg\{ &\frac{(1+r)+2\sqrt{1+r}\ \mathfrak{s}K_{n,s^*} + \mathfrak{s}^2 K_{n,s^*}^2}{\left(1-\mathfrak{s}\sqrt{n\left(\frac{\alpha(1-e^{-1})}{\vartheta_*\ C_\chi}\right)^{\frac{1}{n}}\left(\frac{1}{\log(p)^{\nu-1}}\right)^{\frac{1}{n}}}\right)^2} - 1; \\ &1 - \Bigg(\frac{1-r}{\left(1+\mathfrak{s}\sqrt{n\left(\frac{\alpha(1-e^{-1})}{\vartheta_*\ C_\chi}\right)^{\frac{1}{n}}\left(\frac{1}{\log(p)^{\nu-1}}\right)^{\frac{1}{n}}}\right)^2} \\ &-\frac{2\sqrt{1+r}\ \mathfrak{s}K_{n,s^*} + \mathfrak{s}^2 K_{n,s^*}^2}{\left(1-\mathfrak{s}\sqrt{n\left(\frac{\alpha(1-e^{-1})}{\vartheta_*\ C_\chi}\right)^{\frac{1}{n}}\left(\frac{1}{\log(p)^{\nu-1}}\right)^{\frac{1}{n}}}\right)^2}\Bigg)\Bigg\}.\end{aligned} \tag{2.33}$$

Using (2.25) and Assumption (2.4.6), we have

$$\begin{aligned}\mathfrak{s}\ K_{n,s^*} &\leq C_{\mathfrak{s},n,p}\frac{\sqrt{\alpha\ \left(\frac{\alpha\ (1-e^{-1})}{\vartheta_*\ C_\chi}\right)^{\frac{1}{n}}\left(\frac{1}{\log(p)^{\nu-1}}\right)^{\frac{1}{n}}}}{\left(1+\sqrt{\frac{\alpha+1}{c}\frac{\log(p)}{n}}\right)} \\ &\quad + \frac{C_{\mathfrak{s},n,p}}{\sqrt{\log(p)}}\frac{\sqrt{\log(n)\ \left(\frac{\alpha\ (1-e^{-1})}{\vartheta_*\ C_\chi}\right)^{\frac{1}{n}}\left(\frac{1}{\log(\rho^{-1})^{\nu-1}}\right)^{\frac{1}{n}}}}{\left(1+\sqrt{\frac{\alpha+1}{c}\frac{\log(p)}{n}}\right)},\end{aligned}$$

and thus,

$$\begin{aligned}
\mathfrak{s}\ K_{n,s^*} &\leq C_{\mathfrak{s},n,p}\sqrt{\alpha\ \left(\frac{\alpha\ (1-e^{-1})}{\vartheta_*\ C_\chi}\right)^{\frac{1}{n}}\left(\frac{1}{\log(p)^{\nu-1}}\right)^{\frac{1}{n}}} \\
&\quad + C_{\mathfrak{s},n,p}\sqrt{\frac{\log(n)}{\log(\rho^{-1})}}\sqrt{\left(\frac{\alpha\ (1-e^{-1})}{\vartheta_*\ C_\chi}\right)^{\frac{1}{n}}\left(\frac{1}{\log(p)^{\nu-1}}\right)^{\frac{1}{n}}}, \\
&= C_{\mathfrak{s},n,p}\sqrt{\left(\frac{\alpha\ (1-e^{-1})}{\vartheta_*\ C_\chi}\right)^{\frac{1}{n}}\left(\frac{1}{\log(p)^{\nu-1}}\right)^{\frac{1}{n}}}\left(\sqrt{\alpha}+\sqrt{\frac{\log(n)}{\log(\rho^{-1})}}\right) \\
&\leq 0.1\ r\left(\sqrt{\alpha}+\sqrt{\frac{\log(n)}{\log(\rho^{-1})}}\right)
\end{aligned}$$

On the other hand,

$$\begin{aligned}
\mathfrak{s}\sqrt{n\left(\frac{\alpha(1-e^{-1})}{\vartheta_*\ C_\chi}\right)^{\frac{1}{n}}\left(\frac{1}{\log(p)^{\nu-1}}\right)^{\frac{1}{n}}} &\leq \frac{C_{\mathfrak{s},n,p}}{\sqrt{\log(\rho^{-1})}}\ \frac{\sqrt{\left(\frac{\alpha(1-e^{-1})}{\vartheta_*\ C_\chi}\right)^{\frac{1}{n}}\left(\frac{1}{\log(p)^{\nu-1}}\right)^{\frac{1}{n}}}}{\left(1+\sqrt{\frac{\frac{\alpha+1}{c}\log(p)}{n}}\right)} \\
&\leq \frac{C_{\mathfrak{s},n,p}}{\sqrt{\log(\rho^{-1})}}\ \sqrt{\left(\frac{\alpha(1-e^{-1})}{\vartheta_*\ C_\chi}\right)^{\frac{1}{n}}\left(\frac{1}{\log(p)^{\nu-1}}\right)^{\frac{1}{n}}}
\end{aligned}$$

which, by Assumption 2.4.6, gives

$$\mathfrak{s}\sqrt{n\left(\frac{\alpha(1-e^{-1})}{\vartheta_*\ C_\chi}\right)^{\frac{1}{n}}\left(\frac{1}{\log(p)^{\nu-1}}\right)^{\frac{1}{n}}} \leq 0.1\ r \tag{2.34}$$

Summing up, we get

$$\begin{aligned}
r^* &\leq (1+r)+2\sqrt{1+r}\cdot 0.1\cdot r\left(\sqrt{\alpha}+\sqrt{\frac{\log(n)}{\log(\rho^{-1})}}\right)+0.01\cdot r^2\left(\sqrt{\alpha}+\sqrt{\frac{\log(n)}{\log(\rho^{-1})}}\right)^2-1 \\
&= \left(\sqrt{1+r}+0.1\ r\ \mathrm{C}_{\alpha,n,\rho}\right)^2-1 \qquad (2.35) \\
&\leq r^*_{\alpha,n,\rho}(r)
\end{aligned}$$

with

$$r^*_{\alpha,n,\rho}(r)=r\left(\frac{1}{2}+0.1\ \mathrm{C}_{\alpha,n,\rho}\right)\left(2+\frac{1}{2}r+0.1\ r\ \mathrm{C}_{\alpha,n,\rho}\right) \tag{2.36}$$

where $\mathrm{C}_{\alpha,n,\rho}=\sqrt{\alpha}+\sqrt{\frac{\log(n)}{\log(\rho^{-1})}}$. Thus, using 2.43 and Lemma 2.5.1,

$$\mathbb{P}\left(\left\|X^t_{T^*}X_{T^*}-I\right\|\geq r^*_{\alpha,n,p}(r)\right) \leq \frac{218+1}{p^\alpha}, \tag{2.37}$$

2.5.3.2 Control of $\max_{k\in T^{*c}} \left\|X_{T^*}^t X_k\right\|_2$

By the triangular inequality, we have that

$$\begin{aligned}\max_{k\in T^{*c}} \left\|X_{T^*}^t X_k\right\|_2 &= \max_{k\in T^{*c}} \left\|\left(\mathfrak{C}_{\mathcal{K}} + E_{T^*}\right)^t D_*^2 \left(\mathfrak{C}_k + E_k\right)\right\|_2 \\ &\leq \Bigg(\max_{k\in T^{*c}} \left\|\mathfrak{C}_{\mathcal{K}}^t \mathfrak{C}_k\right\|_2 + \|\mathfrak{C}_{\mathcal{K}}\| \max_{k\in T^{*c}} \|E_k\|_2 \\ &\quad + \|E_{T^*}\| \max_{k\in T^{*c}} \|\mathfrak{C}_k\|_2 + \|E_{T^*}\| \max_{k\in T^{*c}} \|E_k\|_2 \Bigg) \|D_*^2\|. \end{aligned} \tag{2.38}$$

On the other hand, we have

$$\mathbb{P}\left(\max_{k\in\{1,\ldots,p\}} \|E_k\|_2 \geq \mathfrak{s}\left(\sqrt{n} + \sqrt{\frac{\alpha+1}{c}\log(p)}\right) \mid \mathcal{E}_\alpha^* \right) \;\leq\; \frac{C}{p^\alpha}. \tag{2.39}$$

Thus, using Lemma 2.5.6 and Lemma 2.5.2, we obtain

$$\begin{aligned}\mathbb{P}\Bigg(\max_{k\in T^{*c}} \left\|X_{T^*}^t X_k\right\|_2 \geq \Bigg(&\frac{C_{col}\cdot\sqrt{\log(\rho^{-1})}}{1-r} + \Big(\frac{1}{\sqrt{1-r}} + \mathfrak{s}K_{n,s^*}\Big)\mathfrak{s}\Big(\sqrt{n} + \sqrt{\frac{\alpha+1}{c}\log(p)}\Big) \\ &+ \mathfrak{s}K_{n,s^*}\frac{C_{col}\cdot\sqrt{\log(\rho^{-1})}}{(1-r)^{\frac{3}{2}}}\Bigg) \times \frac{1}{\left(1+\mathfrak{s}\sqrt{n\left(\frac{\alpha(1-e^{-1})}{\vartheta_*\ C_\chi}\right)^{\frac{1}{n}}\left(\frac{1}{\log(p)^{\nu-1}}\right)^{\frac{1}{n}}}\right)^2} \mid \mathcal{E}_\alpha^*\Bigg) \\ &\leq \frac{C+2}{p^\alpha} + \frac{2}{\rho^{1-(\alpha+1)\log(K)^2}}. \end{aligned} \tag{2.40}$$

Using the fact

$$\mathbb{P}(A) \leqslant \mathbb{P}(A\cap E) + \mathbb{P}(E^c)$$

with $A = \max_{k\in T^{*c}} \left\|X_{T^*}^t X_k\right\|_2$ and

$$E = \left(\max_{k\in T^{*c}} \left\|\mathfrak{C}_{\mathcal{K}}^t \mathfrak{C}_k\right\|_2 + \|\mathfrak{C}_{\mathcal{K}}\| \max_{k\in T^{*c}} \|E_k\|_2 + \|E_{T^*}\| \max_{k\in T^{*c}} \|\mathfrak{C}_k\|_2 + \|E_{T^*}\| \max_{k\in T^{*c}} \|E_k\|_2 \right) \|D_*^2\|.$$

Furthermore, since $A\subset E^c$ and $E = E_1\cap E_2\cap E_3\cap E_4$, we have by union bound,,

$$\mathbb{P}(A) \;\leq\; \mathbb{P}(E_1^c\cup E_2^c\cup E_3^c\cup E_4^c) \leq \mathbb{P}(E_1^c) + \mathbb{P}(E_2^c) + \mathbb{P}(E_3^c) + \mathbb{P}(E_4^c). \tag{2.41}$$

Since, by Assumption (2.4.6),

$$\left(\frac{1}{\sqrt{1-r}} + \mathfrak{s}K_{n,s^*}\right)\mathfrak{s}\left(\sqrt{n} + \sqrt{\frac{\alpha+1}{c}\log(p)}\right) \;\leq\; \left(\frac{1}{\sqrt{1-r}} + 0.1\ r\ \mathrm{C}_{\alpha,n,\rho}\right)\ \frac{C_{\mathfrak{s},n,p}}{\sqrt{\log(\rho^{-1})}},$$

we obtain

$$\mathbb{P}\Bigg(\max_{k\in T^{*c}}\left\|X_{T^*}^t X_k\right\|_2 \geq \Bigg(\frac{C_{col}\cdot\sqrt{\alpha\log(\rho^{-1})}}{1-r}+\left(\frac{1}{\sqrt{1-r}}+0.1\ r\ \mathrm{C}_{\alpha,n,\rho}\right)\frac{C_{\mathfrak{s},n,p}}{\sqrt{\log(\rho^{-1})}}$$
$$+0.1\ r\ \mathrm{C}_{\alpha,n,p}\cdot\frac{C_{col}\cdot\sqrt{\alpha\log(\rho^{-1})}}{(1-r)^{3/2}}\Bigg)\times\frac{1}{\left(1-\mathfrak{s}\sqrt{n\left(\frac{\alpha(1-e^{-1})}{\vartheta_*\ C_\chi}\right)^{\frac{1}{n}}\left(\frac{1}{\log(p)^{\nu-1}}\right)^{\frac{1}{n}}}\right)^2}\mid\mathcal{E}_\alpha^*\Bigg)$$
$$\leq\frac{C+2}{p^\alpha}+2\rho^{(\alpha+1)\log(K)^2-1}. \tag{2.42}$$

Moreover, for any event $\mathcal{A}$,

$$\mathbb{P}(\mathcal{A})\leq\mathbb{P}(\mathcal{A}\mid\mathcal{E}_\alpha)+\mathbb{P}(\mathcal{E}_\alpha^c), \tag{2.43}$$

and Lemma 2.5.1, we obtain

$$\mathbb{P}\Bigg(\max_{k\in T^{*c}}\left\|X_{T^*}^t X_k\right\|_2 \geq \Bigg(\frac{C_{col}\cdot\sqrt{\alpha\log(\rho^{-1})}}{1-r}+\left(\frac{1}{\sqrt{1-r}}+0.1\ r\ \mathrm{C}_{\alpha,n,p}\right)\frac{C_{\mathfrak{s},n,p}}{\sqrt{\log(\rho^{-1})}}$$
$$+0.1\ r\ \mathrm{C}_{\alpha,n,p}\cdot\frac{C_{col}\cdot\sqrt{\alpha\log(\rho^{-1})}}{(1-r)^{3/2}}\Bigg)\times\frac{1}{\left(1-\mathfrak{s}\sqrt{n\left(\frac{\alpha(1-e^{-1})}{\vartheta_*\ C_\chi}\right)^{\frac{1}{n}}\left(\frac{1}{\log(p)^{\nu-1}}\right)^{\frac{1}{n}}}\right)^2}\Bigg)$$
$$\leq\frac{C+3}{p^\alpha}+2\rho^{(\alpha+1)\log(K)^2-1}. \tag{2.44}$$

2.5.4 The last two inequalities

The proof of (2.16) is standard and, under Assumption 2.4.6, the proof of (2.17) can be proved using the ideas of [10, Section 3.3]. We give the proofs for the sake of completeness.

2.5.4.1 Control of $\|X_{T^{*c}}^t X_{T^*}(X_{T^*}^t X_{T^*})^{-1}X_{T^*}^t z\|_\infty$

For any $j\in T^{*c}$, by the results of section 2.5.3.2, we have

$$\|X_{T^*}(X_{T^*}^t X_{T^*})^{-1}X_{T^*}^t X_j^t\|_2\leq\frac{\sqrt{1+r_{\alpha,n,p}^*(r)}}{\left(1-r_{\alpha,n,p}^*(r)\right)\left(1-\mathfrak{s}\sqrt{n\left(\frac{\alpha(1-e^{-1})}{\vartheta_*\ C_\chi}\right)^{\frac{1}{n}}\left(\frac{1}{\log(p)^{\nu-1}}\right)^{\frac{1}{n}}}\right)^2}$$
$$\times\left(0.1\ r\ \mathrm{C}_{\alpha,n,p}\cdot\frac{C_{col}\cdot\sqrt{\log(\rho^{-1})}}{(1-r)^{3/2}}+\frac{C_{col}\cdot\sqrt{\log(\rho^{-1})}}{1-r}+\left(\frac{1}{\sqrt{1-r}}+0.1\ r\ \mathrm{C}_{\alpha,n,p}\right)\frac{C_{\mathfrak{s},n,p}}{\sqrt{\log(\rho^{-1})}}\right)$$

with probability at least $1-\left(\frac{C+3}{p^{\alpha}}+2\rho^{(\alpha+1)\log(K)^2-1}\right)$, we get

$$\mathbb{P}\left(X_j^t X_{T^*}(X_{T^*}^t X_{T^*})^{-1}X_{T^*}^t z\geq u\right)\leq$$

$$\frac{1}{2}\exp\left(-\frac{u^2\left(1-\mathfrak{s}\sqrt{n\left(\frac{\alpha(1-e^{-1})}{\vartheta_*\ C_\chi}\right)^{\frac{1}{n}}\left(\frac{1}{\log(p)^{\nu-1}}\right)^{\frac{1}{n}}}\right)^2}{2\sigma^2\frac{\sqrt{1+r^*_{\alpha,n,\rho}(r)}}{1-r^*_{\alpha,n,\rho}(r)}\left(\begin{array}{c}\left(0.1r\ \mathrm{C}_{\alpha,n,p}\cdot\frac{C_{col}\cdot\sqrt{\log(\rho^{-1})}}{(1-r)^{3/2}}+\frac{C_{col}\cdot\sqrt{\log(\rho^{-1})}}{1-r}\right)\\ +\left(\frac{1}{\sqrt{1-r}}+0.1r\ \mathrm{C}_{\alpha,n,\rho}\right)\frac{C_{\mathfrak{s},n,p}}{\sqrt{\log(\rho^{-1})}}\end{array}\right)}\right)+\frac{C+3}{p^{\alpha}}+\rho^{(\alpha+1)\log(K)^2-1}.$$

Taking u such that

$$\frac{1}{2}\exp\left(-\frac{u^2\left(1-\mathfrak{s}\sqrt{n\left(\frac{\alpha(1-e^{-1})}{\vartheta_*\ C_\chi}\right)^{\frac{1}{n}}\left(\frac{1}{\log(p)^{\nu-1}}\right)^{\frac{1}{n}}}\right)^2}{2\sigma^2\frac{\sqrt{1+r^*_{\alpha,n,\rho}(r)}}{1-r^*_{\alpha,n,\rho}(r)}\left(\begin{array}{c}\left(0.1r\ \mathrm{C}_{\alpha,n,p}\cdot\frac{C_{col}\cdot\sqrt{\log(\rho^{-1})}}{(1-r)^{3/2}}+\frac{C_{col}\cdot\sqrt{\log(\rho^{-1})}}{1-r}\right)\\ +\left(\frac{1}{\sqrt{1-r}}+0.1r\ \mathrm{C}_{\alpha,n,\rho}\right)\frac{C_{\mathfrak{s},n,p}}{\sqrt{\log(\rho^{-1})}}\end{array}\right)}\right)=\rho^{\alpha}$$

i.e.

$$u=\sqrt{\left(\alpha\log(\rho^{-1})-\log(2)\right)}\sqrt{\frac{2\sigma^2\frac{\sqrt{1+r^*_{\alpha,n,\rho}(r)}}{1-r^*_{\alpha,n,\rho}(r)}\left(\begin{array}{c}\left(0.1r\ \mathrm{C}_{\alpha,n,p}\cdot\frac{C_{col}\cdot\sqrt{\log(\rho^{-1})}}{(1-r)^{3/2}}+\frac{C_{col}\cdot\sqrt{\log(\rho^{-1})}}{1-r}\right)\\ +\left(\frac{1}{\sqrt{1-r}}+0.1r\ \mathrm{C}_{\alpha,n,\rho}\right)\frac{C_{\mathfrak{s},n,p}}{\sqrt{\log(\rho^{-1})}}\end{array}\right)}{\left(1-\mathfrak{s}\sqrt{n\left(\frac{\alpha(1-e^{-1})}{\vartheta_*\ C_\chi}\right)^{\frac{1}{n}}\left(\frac{1}{\log(p)^{\nu-1}}\right)^{\frac{1}{n}}}\right)^2}}$$

Using the union bound, we finally obtain

$$\mathbb{P}\left(\begin{array}{c}\|X_{T^{*c}}^t X_{T^*}(X_{T^*}^t X_{T^*})^{-1} X_{T^*}^t z\|_\infty \geq \\ \sqrt{(\alpha\log(\rho^{-1})-\log(2))}\sqrt{\dfrac{2\sigma^2\ \frac{\sqrt{1+r^*_{\alpha,n,\rho}(r)}}{1-r^*_{\alpha,n,p}(\rho)}\left(\begin{array}{c}0.1\ r\ \mathrm{C}_{\alpha,n,\rho}\cdot\frac{C_{col}\cdot\sqrt{\log(\rho^{-1})}}{(1-r)^{3/2}}+\frac{C_{col}\cdot\sqrt{\log(\rho^{-1})}}{1-r}\\ +\left(\frac{1}{\sqrt{1-r}}+0.1\ r\ \mathrm{C}_{\alpha,n,\rho}\right)\frac{C_{\mathfrak{s},n,p}}{\sqrt{\log(\rho^{-1})}}\end{array}\right)}{\left(1-\mathfrak{s}\sqrt{n\left(\frac{\alpha(1-e^{-1})}{\vartheta_*\ C_\chi}\right)^{\frac{1}{n}}\left(\frac{1}{\log(p)^{\nu-1}}\right)^{\frac{1}{n}}}\right)^2}}\end{array}\right)$$

$$\leq \frac{C+4}{p^{\alpha-1}} + p\ \rho^{(\alpha+1)\log(K)^2-1}. \tag{2.45}$$

2.5.4.2 Control of $\|X_{T^{*c}}^t X_{T^*}(X_{T^*}^t X_{T^*})^{-1}\mathrm{sgn}\,(\beta^*_{T^*})\|_\infty$

For any $j \in T^{*c}$, again by the results of section (C.2.2), we have

$$\|X_j^t X_{T^*}(X_{T^*}^t X_{T^*})^{-1}\|_2 \ \leq\ \frac{\left(\begin{array}{c}\left(0.1\ r\ \mathrm{C}_{\alpha,n,\rho}\cdot\frac{C_{col}\cdot\sqrt{\log(\rho^{-1})}}{(1-r)^{3/2}}\right.\\ \left.+\frac{C_{col}\cdot\sqrt{\log(p)}}{1-r}+\left(\frac{1}{\sqrt{1-r}}+0.1\ r\ \mathrm{C}_{\alpha,n,\rho}\right)\frac{C_{\mathfrak{s},n,p}}{\sqrt{\log(\rho^{-1})}}\right)\end{array}\right)}{\left(1-r^*_{\alpha,n,\rho}(r)\right)\left(1-\mathfrak{s}\sqrt{n\left(\frac{\alpha(1-e^{-1})}{\vartheta_*\ C_\chi}\right)^{\frac{1}{n}}\left(\frac{1}{\log(p)^{\nu-1}}\right)^{\frac{1}{n}}}\right)^2}$$

with probability at least $1-\left(\frac{C+3}{p^\alpha}+2\rho^{(\alpha+1)\log(K)^2-1}\right)$. Hoeffding's inequality gives

$$\begin{aligned}\mathbb{P}\left(X_j^t X_{T^*}(X_{T^*}^t X_{T^*})^{-1}\mathrm{sgn}\,(\beta^*_{T^*})\geq u\right) &\leq \frac{1}{2}\exp\left(-\frac{u^2}{2\ \|(X_{T^*}^t X_{T^*})^{-1}X_{T^*}^t X_j\|_2}\right)\\ &\leq \frac{1}{2}\exp\left(-\frac{u^2\left(1-\mathfrak{s}\sqrt{n\left(\frac{\alpha(1-e^{-1})}{\vartheta_*\ C_\chi}\right)^{\frac{1}{n}}\left(\frac{1}{\log(p)^{\nu-1}}\right)^{\frac{1}{n}}}\right)^2}{\frac{2}{1-r^*_{\alpha,n,\rho}(r)}\left(\begin{array}{c}\left(0.1r\ \mathrm{C}_{\alpha,n,p}\cdot\frac{C_{col}\cdot\sqrt{\log(\rho^{-1})}}{(1-r)^{3/2}}+\frac{C_{col}\cdot\sqrt{\log(\rho^{-1})}}{1-r}\right.\\ \left.+\left(\frac{1}{\sqrt{1-r}}+0.1r\ \mathrm{C}_{\alpha,n,\rho}\right)\frac{C_{\mathfrak{s},n,p}}{\sqrt{\log(\rho^{-1})}}\right)\end{array}\right)}\right)\\ &\quad+\frac{C+3}{p^\alpha}+2\rho^{(\alpha+1)\log(K)^2-1}.\end{aligned} \tag{2.46}$$

Choosing

$$u = \sqrt{\frac{\frac{2(\alpha\log(\rho^{-1})-\log(2))}{1-r^*_{\alpha,n,\rho}(r)}\left(\begin{array}{c}0.1\ r\ \mathrm{C}_{\alpha,n,\rho}\cdot\frac{C_{col}\cdot\sqrt{\log(\rho^{-1})}}{(1-r)^{3/2}}+\frac{C_{col}\cdot\sqrt{\log(\rho^{-1})}}{1-r}\\ +\left(\frac{1}{\sqrt{1-r}}+0.1\ r\ \mathrm{C}_{\alpha,n,\rho}\right)\frac{C_{\mathfrak{s},n,p}}{\sqrt{\log(\rho^{-1})}}\end{array}\right)}{\left(1-\mathfrak{s}\sqrt{n\left(\frac{\alpha(1-e^{-1})}{\vartheta_*\ C_\chi}\right)^{\frac{1}{n}}\left(\frac{1}{\log(p)^{\nu-1}}\right)^{\frac{1}{n}}}\right)^2}} \tag{2.47}$$

and applying the union bound, we obtain

$$\begin{aligned}&\mathbb{P}\Bigg(\ \left\|X^t_{T^{*c}}X_{T^*}(X^t_{T^*}X_{T^*})^{-1}\mathrm{sgn}\left(\beta^*_{T^*}\right)\right\|_\infty\\ &\geq\sqrt{\frac{\frac{2\left(\alpha\log(\rho^{-1})-\log(2)\right)}{1-r^*_{\alpha,n,\rho}(r)}\left(\begin{array}{c}0.1\ r\ \mathrm{C}_{\alpha,n,\rho}\cdot\frac{C_{col}\cdot\sqrt{\log(\rho^{-1})}}{(1-r)^{3/2}}+\frac{C_{col}\cdot\sqrt{\log(\rho^{-1})}}{1-r}\\ +\left(\frac{1}{\sqrt{1-r}}+0.1\ r\ \mathrm{C}_{\alpha,n,\rho}\right)\frac{C_{\mathfrak{s},n,p}}{\sqrt{\log(\rho^{-1})}}\end{array}\right)}{\left(1-\mathfrak{s}\sqrt{n\left(\frac{\alpha(1-e^{-1})}{\vartheta_*\ C_\chi}\right)^{\frac{1}{n}}\left(\frac{1}{\log(p)^{\nu-1}}\right)^{\frac{1}{n}}}\right)^2}}\ \Bigg)\\ &\leq\tfrac{C+4}{p^{\alpha-1}}+p\ \rho^{(\alpha+1)\log(K)^2-1}.\end{aligned}$$

2.5.4.3 Summing up

We obtain that

$$\begin{aligned}&\left\|X^t_{T^{*c}}X_{T^*}(X^t_{T^*}X_{T^*})^{-1}X^t_{T^*}z\right\|_\infty+\lambda\ \left\|X^t_{T^{*c}}X_{T^*}(X^t_{T^*}X_{T^*})^{-1}\mathrm{sgn}\left(\beta^*_{T^*}\right)\right\|_\infty\\ &\qquad\leq\sigma\mathcal{C}_1+\lambda\mathcal{C}_2\end{aligned}$$

where

$$\mathcal{C}_1=\sqrt{\left(\alpha\log(\rho^{-1})-\log(2)\right)}\sqrt{\frac{\sqrt{1+r^*_{\alpha,n,\rho}(r)}}{1-r^*_{\alpha,n,\rho}(r)}\left(\begin{array}{c}0.1\ r\ \mathrm{C}_{\alpha,n,\rho}\cdot\frac{C_{col}\cdot\sqrt{\log(\rho^{-1})}}{(1-r)^{3/2}}+\frac{C_{col}\cdot\sqrt{\log(\rho^{-1})}}{1-r}\\ +\left(\frac{1}{\sqrt{1-r}}+\frac{0.1\ r}{\sqrt{n}}\mathrm{C}_{\alpha,n,p}\right)\frac{C_{\mathfrak{s},n,p}}{\sqrt{\log(\rho^{-1})}}\end{array}\right)} \tag{2.48}$$

and

$$\mathcal{C}_2=\sqrt{\frac{2\left(\alpha\log(\rho^{-1})-\log(2)\right)}{1-r^*_{\alpha,n,\rho}(r)}\left(\begin{array}{c}0.1\ r\ \mathrm{C}_{\alpha,n,p}\cdot\frac{C_{col}\cdot\sqrt{\log(\rho^{-1})}}{(1-r)^{3/2}}+\frac{C_{col}\cdot\sqrt{\log(\rho^{-1})}}{1-r}\\ +\left(\frac{1}{\sqrt{1-r}}+0.1\ r\ \mathrm{C}_{\alpha,n,p}\right)\frac{C_{\mathfrak{s},n,p}}{\sqrt{\log(\rho^{-1})}}\end{array}\right)} \tag{2.49}$$

as announced. Choosing ρ such that Assumption 2.4.9 is satisfied and C_{col} is sufficiently small and n sufficiently large such that $\mathcal{C}_1$ and $\mathcal{C}_2$ be smaller than $1/16$ and the proof is completed.

2.6 Conclusion

The goal of this paper is to propose a sound study of the behavior of the LASSO algorithm for the linear model in the case where the design matrix is not satisfying the usual non-colinearity conditions that are enforced in standard analysis. We introduce a new model for the design matrix. In this new model, the columns are assumed to be drawn from a Gaussian mixture model where the centers of the mixture model, and them only, satisfy the incoherence condition. As a result, we are able to analyse an interesting example of applying the LASSO to a non-incoherent matrix and we obtain a performance bound. The price to pay for such a generality is that our prediction bounds hold with arbritrarily high but fixed probability, as compared with the incoherent setting where the the probability goes to one as p tends to $+\infty$.

Bibliography

[1] AlQuraishi, M. and McAdams, H., Direct inference of protein–DNA interactions using compressed sensing methods, PNAS 108, 14819 (2011).

[2] Becker, S., Bobin, J. and Candès, E. J., Nesta: a fast and accurate first-order method for sparse recovery. In press SIAM J. on Imaging Science.

[3] Bickel, P. J., Ritov, Y., Tsybakov, A. B. Simultaneous analysis of lasso and Dantzig selector. Ann. Statist. 37 (2009), no. 4, 1705–1732.

[4] Bock, M.E, Judge, G.G, Yancey, T.A. A simple form for the inverse moments of non-central χ^2 and F random variable and certain confluent hypergeometric functions. Journal of Econometrics 25 (1984); 217–234. North-Holland.

[5] Bousquet, O., A Bennett concentration inequality and its application to suprema of empirical processes, Comptes Rendus Mathematique, 334 (2002), no. 6, 495–500.

[6] Bunea, F., Tsybakov, A., and Wegkamp, M. (2007a). Sparsity oracle inequalities for the Lasso. Electron. J. Stat., 1 :169–194.

[7] Bunea, F., Honest variable selection in linear and logistic regression models via ℓ_1 and $\ell_1 + \ell_2$ penalization , the Electronic Journal of Statistics, (2008) Vol. 2, 1153-1194

[8] Candès, E., Compressive sampling, (2006) 3, International Congress of Mathematics, 1433–1452, EMS.

[9] Candès, E. J. The restricted isometry property and its implications for compressed sensing. C. R. Math. Acad. Sci. Paris 346 (2008), no.

[10] Candès, E. J. and Plan, Yaniv. Near-ideal model selection by ℓ_1 minimization. Ann. Statist. 37 (2009), no. 5A, 285–2177.

[11] Candès, E. and Tao T., Decoding by linear programming. IEEE Information Theory, 51 (2005) no. 12, 4203-4215.

[12] Candès, E. J. and Tao, T., The Dantzig Selector: statistical estimation when p is much larger than n. Ann. Stat.

[13] Chandrasekaran, V., Recht, B., Parrilo, P.A. and Willsky, A.S., 2012. The convex geometry of linear inverse problems. Foundations of Computational mathematics, 12(6), pp.805-849.

[14] Chrétien, S., 2009. An alternating l_1 approach to the compressed sensing problem, IEEE Signal Processing Letters, 17(2), pp.181-184.

[15] Chrétien, S. and Hero, A. and Perdry, H., 2012. Space alternating penalized Kullback proximal point algorithms for maximizing likelihood with nondifferentiable penalty, Annals of the Institute of Statistical Mathematics, 64(4), pp. 791–809.

[16] Chrétien, S. and Darses, Sparse recovery with unknown variance: a LASSO-type approach, ArXiv 2012.

[17] Chrétien, S. and Darses, S. Invertibility of random submatrix via tail decoupling and a matrix Chernoff inequality, Stat. and Prob. Lett. 82 (2012), no. 7, 1479-1487.

[18] Constructing message passing algorithms for compressed sensing D. L. Donoho, A. Maleki, and A. Montanari, submitted to IEEE Trans. Inf. Theory.

[19] Dossal, C., A necessary and sufficient condition for exact recovery by ℓ_1 minimization. http://hal.archives-ouvertes.fr/docs/00/16/47/38/PDF/DossalMinimisationl1.pdf

[20] Fuchs, J.J., On sparse representations in arbitrary redundant bases. IEEE Trans. Info. Th., 2002.

[21] Jaggi, M., 2013, February. Revisiting Frank-Wolfe: Projection-free sparse convex optimization. In International Conference on Machine Learning (pp. 427-435). PMLR.

[22] Kim, S.-J., Koh, K., Boyd, S. and Gorinevsky, D., SIAM Review, problems and techniques section, 51 , (2009), no. 2, 339–360.

[23] Lafond, J., Wai, H.T. and Moulines, E., 2015. On the online Frank-Wolfe algorithms for convex and non-convex optimizations. arXiv preprint arXiv:1510.01171.

[24] Massart, P., Concentration inequalities and model selection. Lectures from the 33rd Summer school on Probability Theory in Saint Flour. Lecture Notes in Mathematics, 1896. Springer Verlag (2007).

[25] Meinshausen, N. and Bühlmann, P., High-dimensional graphs and variable selection with the Lasso, Ann. Statist. 34 (2006), no. 3, 1436-1462.

[26] Neto, D. Sardy, S. and Tseng, P., l1-Penalized Likelihood Smoothing and Segmentation of Volatility Processes allowing for Abrupt Changes, Journal of Computational and Graphical Statistics, 21 (2012), no. 1, 217–233.

[27] Osborne, M.R., Presnell, B. and Turlach, B.A., A new approach to variable selection in least squares problems, IMA J. Numer. Anal. 20 (2000), no. 3, 389–403.

[28] Parikh, N. and Boyd, S., 2014. Proximal algorithms. Foundations and Trends in optimization, 1(3), pp.127-239.

[29] Rudelson, M. and Vershynin, R. Smallest singular value of a random rectangular matrix. Comm. Pure Appl. Math. 62 (2009), no. 12, 1707–20131739.

[30] Tibshirani, R. Regression shrinkage and selection via the LASSO, J.R.S.S. Ser. B, 58, no. 1 (1996), 267–288.

[31] Tropp, J. A. Norms of random submatrices and sparse approximation. C. R. Math. Acad. Sci. Paris 346 (2008), no. 23-24, 1271–1274.

[32] Tropp, J. A., User friendly tail bounds for sums of random matrices, http://arxiv.org/abs/1004.4389, (2010).

[33] van de Geer, S., High-dimensional generalized linear models and the Lasso. The Annals of Statistics 36, 614-645.

[34] Vershynin, R., Introduction to the non-asymptotic analysis of random matrices, Chapter 5 of the book Compressed Sensing, Theory and Applications, ed. Y. Eldar and G. Kutyniok. Cambridge University Press, 2012. pp. 210–268. [arXiv:1011.3027, Aug 2010].

[35] Wainwright, Martin J., Sharp thresholds for high-dimensional and noisy sparsity recovery using ℓ_1-constrained quadratic programming (Lasso). IEEE Trans. Inform. Theory 55 (2009), no. 5, 2183–2202.

[36] Zhang, T., 2009. Some sharp performance bounds for least squares regression with l1 regularization. The Annals of Statistics, 37(5A), pp.2109-2144.

Chapter 3

Spectrally Sparse Tensor Reconstruction in Optical Coherence Tomography Using Nuclear Norm Penalisation

3.1 Introduction

3.1.1 Motivations and Contributions

3D Image reconstruction from subsamples is a difficult problem at the intersection of the fields of inverse problem, computational statistics and numerical analysis. Following the Compressed Sensing paradigm [40], the problem can be tackled using sparsity priors in the case where sparsity can be proved present in the image to be recovered. Compressed sensing has evolved from the recovery of a sparse vector [41], to the recovery of spectrally sparse matrices [42]. The problem we are considering in the present paper was motivated by Optical Coherence Tomography (OCT), where a 3D volume is to be recovered from a small set of measurements along a ray across the volume, making the problem a kind of tensor completion problem.

The possibility of efficiently addressing the reconstruction problem in Compressed Sensing, Matrix Completion and their avatars greatly depend on formulating it as a problem which can be relaxed as a convex optimisation problem. For example, the sparsity of a vector is measured by the number of non-zero components in that vector. This quantity is a non-convex function of the vector but a surrogate can be easily found in the form the ℓ_1-norm. Spectral sparsity of a matrix is measured by the number of nonzero singular values and, similarly, a convex surrogate is the ℓ_1-norm of the vector of singular values, also called the *nuclear norm*.

The main problem with tensor recovery is that the spectrum is not a well defined quantity and several approaches have been proposed for defining it [43, 44]. Different nuclear norms have also been proposed in the literature, based on the various definitions for the spectrum [45, 46, 47]. A very interesting approach was proposed in References

[37, 38]. This approach is interesting in several ways:

- the 3D tensors are considered as matrix of 1D vectors (tubes) and the approach uses a tensor product similar to the classical matrix product, after replacing multiplication of entries by, for example, convolution of the tubes,
- the SVD is fast to compute,
- a specific and natural nuclear norm can be easily defined.

Motivated by the 3D reconstruction problem, our goal in the present paper is to study the natural nuclear norm penalised estimator for tensor completion problem. Our approach extends the method proposed in [39] to the framework developed by [38]. The main contribution of the paper is

- to present the most natural definition of the nuclear norm in the framework of Reference [38],
- to compute the subdifferential of this nuclear norm,
- to present a precise mathematical study of the nuclear norm penalised least-squares reconstruction method,
- to illustrate the efficiency of the approach on the OCT reconstruction problem with real data.

3.1.2 Background on Tensor Completion

3.1.2.1 Matrix Completion

After the many successes of Compressed Sensing in inverse problems, Matrix and tensor completion problems have recently taken the stage and become the focus of an extensive research activity. Completion problems have applications in collaborative filtering [48], Machine Learning [49, 50], sensor networks [51], subspace clustering [52], signal processing [47], and so forth. The problem is intractable if the matrix to recover does not have any particular structure. An important discovery in References [53, 36] is that when the matrix to be recovered has a low rank, then it can be recovered based on a few observations only [53] and using nuclear norm penalised estimation is a reasonably easy problem to solve.

The use of the nuclear norm as a convex surrogate for the rank was first proposed in Reference [54] and further analysed in a series of breakthrough papers [42, 36, 55, 56, 57, 39].

3.1.2.2 Tensor Completion

The matrix completion problem was recently generalised to the problem of tensor completion. A very nice overview of tensors and algorithms is given in Reference [58]. Tensors play a fundamental role in statistics [59] and more recently in Machine Learning [49]. It can be used for multidimensional time series [60], analysis of seismic data [45], Hidden Markov Models [49], Gaussian Mixture based clustering [61], Phylogenetics [62], and much more. Tensor completion is however a more difficult problem from many points of view. First, the rank is NP-hard to compute in full generality [44]. Second, many different Singular Value Decompositions are available. The Tucker decomposition [44] extends the useful sum of rank one decomposition to tensors but it is NP-hard to compute in general. An interesting algorithm was proposed in References [63, 64]; see also the very interesting Reference [65]. Another SVD was proposed in Reference [66] with the main advantage that the generalized singular vectors form an orthogonal family. The usual diagonal matrix is replaced with a so called "core" tensor with nice orthogonality properties and very simple structure in the case of symmetric [67] or orthogonally decomposable tensors [68].

Another very interesting approach, called the t-SVD was proposed recently in [38] for 3D tensors with applications to face recognition [69] and image deblurring, computerized tomography [70] and data completion [71]. In the t-SVD framework, one dimension is assumed to be of a different nature from the others, such as, for example, time. In this setting, the decomposition resembles the SVD closely by using a diagonal tensor instead of a diagonal matrix as for the standard 2D SVD. The t-SVD was also proved to be amenable to online learning [72]. One other interesting advantage of the t-SVD is that the nuclear norm is well defined and its subdifferential is easy to compute.

The t-SVD proposed in Reference [38] is a very attractive representation of tensors in many settings such as image analysis, multivariate time series, and so forth. Obviously, it is not so attractive for the study of symmetric moment or cumulant tensors as studied in Reference [49], but for large image sequences and times series, this approach seems to be extremely relevant.

3.1.3 Sparsity

One of the main discoveries of the past decade is that sparsity may help in certain contexts where data are difficult or expensive to acquire [40]. When the object A_0 to recover is a vector, the fact that it may be sparse in a certain dictionary will dramatically improve recovery, as demonstrated by the recent breakthroughs in high dimensional statistics on the analysis of estimators obtained via convex optimization, such as the LASSO [38, 73, 40], the Adaptive LASSO [74] or the Elastic Net [75]. When A_0 is a matrix, the property of having a low rank may be crucial for the recovery as proved in References [76, 39]. Extensions to the tensor case are studied in References [46, 77, 78]. Tensor completion using the t-SVD framework has been analyzed in Reference [71]. In particular, our results complement and improve on the results in Reference [71]. Reference [79] deserves special mention for using more advanced techniques based on real algebraic geometry and sums of

squares decompositions.

The usual estimator in a sparse recovery problem is a nuclear norm penalized least squares such as the one studied here and defined by (3.27). Several other types of nuclear norms have been used for tensor estimation and completion. In particular, several nuclear norms can be naturally defined such as in for example, References [46, 80, 81]. It is interesting to notice that, sparsity promoting penalization of the least squares estimator crucially relies on the structure of the subdifferential of the norm involved. See for instance References [39, 82] or Reference [83]. In Reference [46] a subset of the subdifferential is studied and then used in order to establish the efficiency of the penalized least squares approach. Another interesting approach is the one in [81] where an outer approximation of the subdifferential is given. In the matrix setting, the work in References [84, 85] are famous for providing a neat characterization of the subdifferential of matrix norms or more generally functions of the matrix enjoying enough symmetries. In the 3D or higher dimensional setting, however, the case is much less understood. The relationship between the tensor norms and the norms of the flattening are intricate although some good bounds relating one to the other can be obtained, as, for example, in Reference [86].

The extension of previous results on low rank matrix reconstruction to the tensor setting is nontrivial but is shown to be relatively easy to obtain once the appropriate background is given. In particular, our theoretical analysis will generalise the analysis in Reference [39]. In order to do this, we provide a complete characterisation of the subdifferential of the nuclear norm. Our results will be illustrated by computational experiments for solving a problem in Optical Coherence Tomography (OCT).

3.1.4 Plan of the Paper

Section 3.2 presents the necessary mathematical background about tensors and sparse recovery. Section 3.3 introduces the measurement model and the present our nuclear-norm penalised estimator. In Section 3.4 we prove our main theoretical result. In Section 3.5, our approach is finally illustrated in the context of Optical Coherence Tomography, and some computational results based on real data are presented.

3.2 Background on Tensors, t-SVD

In this section, we present what is meant by the notion of tensor and the various generalisations of the common objects in linear algebra to the tensor setting. In particular, we will introduce the Singular Value Decomposition proposed in Reference [38] and some associated Schatten norms.

3.2.1 Basic Notations for Third-Order Tensor

In this section, we recall the framework introduced by Kilmer and Martin [37, 38] for a very special class of tensors which is particularly adapted to our setting.

3.2.1.1 Slices and Transposition

For a third-order tensor A, its (i,j,k)th entry is denoted by A_{ijk}.

Definition 3.2.1 *The k^{th}- frontal slice of A is defined as*

$$A^{(k)} = A\left(:,:,k\right).$$

The j^{th}-transversal slice of A is defined as

$$\vec{A}^{(j)} = A\left(:,j,:\right).$$

A tubal scalar (t-scalar) is an element of $\mathbb{R}^{1\times 1\times n_3}$ and a tubal vector (t-vector) is an element of $\mathbb{R}^{n_1\times 1\times n_3}$

Definition 3.2.2 (Tensor transpose) *The conjugate transpose of a tensor $A \in \mathbb{R}^{n_1\times n_2\times n_3}$ tensor A^t obtained by conjugate transposing each of the frontal slice and then reversing the order of transposed frontal slices starting from the slide number 2 to the slice number n_3 and then appending the conjugate transposed frontal slice $A^{(1)^\top}$.*

Definition 3.2.3 (The "dot" product) *The dot product $A \cdot B$ between two tensors $A \in \mathbb{R}^{n_1\times n_2\times n_3}$ and $B \in \mathbb{R}^{n_2\times n_4\times n_3}$ is the tensor $C \in \mathbb{R}^{n_1\times n_4\times n_3}$ whose slice $C^{(n)}$ is the matrix product of the slice $A^{(n)}$ with the slice $B^{(n)}$:*

$$C^{(k)} := (A\cdot B)^{(k)} := A^{(k)}\ B^{(k)}, \quad k = 1,\ldots,n_3. \tag{3.1}$$

We will also need the canonical inner product.

Definition 3.2.4 (Inner product of tensors) *If A and B are third-order tensors of same size $n_1 \times n_2 \times n_3$, then the inner product between A and B is defined as the following (notice the normalization constant of FFT),*

$$\langle A, B\rangle = \sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3} A_{ijk}B_{ijk}. \tag{3.2}$$

3.2.1.2 Convolution and Fourier Transform

Definition 3.2.5 (t-product for circular convolution) *The t-product $A{*}B$ of $A \in \mathbb{R}^{n_1\times n_2\times n_3}$ and $B \in \mathbb{R}^{n_2\times n_4\times n_3}$ is an $n_1 \times n_4 \times n_3$ tensor whose (i,j)-th tube is given by*

$$C\left(i,j,:\right) = \sum_{k=1}^{n_2} A\left(i,k,:\right) * B\left(k,j,:\right), \tag{3.3}$$

where $*$ denotes the circular convolution between two cubes of same size.

Definition 3.2.6 (Identity tensor) *The identity tensor $J \in \mathbb{R}^{n_1\times n_1\times n_3}$ is defined to be a tensor whose first frontal slice J^1 is the $n_1 \times n_1$ identity matrix and all other frontal slices $J^i, i = 2,\ldots,n_3$ are zero.*

Definition 3.2.7 (Orthogonal tensor) *A tensor $Q \in \mathbb{R}^{n\times n\times n_3}$ is orthogonal if it satisfies*

$$Q^\top * Q = Q * Q^\top = J. \tag{3.4}$$

$\hat{A}$ is a tensor which is obtained by taking the Fast Fourier Transform (FFT) along the third dimension and we will use the following convention for Fast Transform along the 3rd dimension

$$\hat{A} = \text{fft}\,(A, [\,], 3)\,.$$

The one-dimensional FFT along the 3th-dimension is given

$$\hat{A}(j_1, j_2, k_3) = \sum_{j_3=1}^{n_3} A(j_1, j_2, j_3)\, \exp(-2\frac{i\pi j_3 k_3}{n_3}), \quad \forall\, j_1,\, j_2,\, 1 \leqslant j_1 \leqslant n_1,\, 1 \leqslant j_2 \leqslant n_2.$$

Naturally, one can compute A from $\hat{A}$ via $\text{ifft}\Big(\hat{A}, [\,], 3\Big)$ using the inverse FFT, and is defined:

$$A(j_1, j_2, k_3) = \sum_{j_3=1}^{n_3} \hat{A}(j_1, j_2, j_3)\, \exp(2\frac{i\pi j_3 k_3}{n_3}), \quad \forall\, j_1,\, j_2,\, 1 \leqslant j_1 \leqslant n_1,\, 1 \leqslant j_2 \leqslant n_2.$$

Definition 3.2.8 (Inverse of a tensor) *The inverse of a tensor $A\in \mathbb{R}^{n\times n\times n_3}$ is written as A^{-1} satisfying*

$$A^{-1} * A = A * A^{-1} = J. \tag{3.5}$$

where J is the identity tensor of size $n \times n \times n_3$.

Remark 3.2.9 *It is proved in Reference [38] that for any tensor $A \in \mathbb{R}^{n_1\times n_2\times n_3}$ and $B \in \mathbb{R}^{n_2\times n_4\times n_3}$, we have*

$$A * B = C \Leftrightarrow \hat{A} \cdot \hat{B} = \hat{C}.$$

3.2.2 The t-SVD

We finally arrive at the definition of the t-SVD.

Definition 3.2.10 (f-diagonal tensor) *Tensor A is called f-diagonal if each frontal slice $A^{(i)}$ is a diagonal matrix.*

Definition 3.2.11 (Tensor Singular Value Decomposition: t-SVD) *For $M \in \mathbb{R}^{n_1\times n_2\times n_3}$, the t-SVD of M is given by*

$$M = U * S * V^\top, \tag{3.6}$$

where U and V are orthogonal tensor of size $n_1 \times n_1 \times n_3$ and $n_2 \times n_2 \times n_3$ respectively. S is a rectangular f-diagonal tensor or size $n_1 \times n_2 \times n_3$, and the entries in S are called the singular values of M. This SVD can be obtained using the Fourier transform as follows:

$$\hat{M}^{(i)} = \hat{U}^{(i)} \cdot \hat{S}^{(i)} \cdot \Big(\hat{V}^{(i)}\Big)^\top. \tag{3.7}$$

This t-SVD is illustrated in Figure 4.1 below. Notice that the diagonal elements of S, i.e., $S(i,i,:)$ are tubal scalars as introduced in Definition 3.2.1. They will also be called *tubal eigenvalues*.

Definition 3.2.12 *The spectrum $\sigma(A)$ of the tensor A is the tubal vector given by*

$$\sigma(A)_i = S(i,i,:) \tag{3.8}$$

for $i = 1, \ldots, \min\{n_1, n_2\}$.

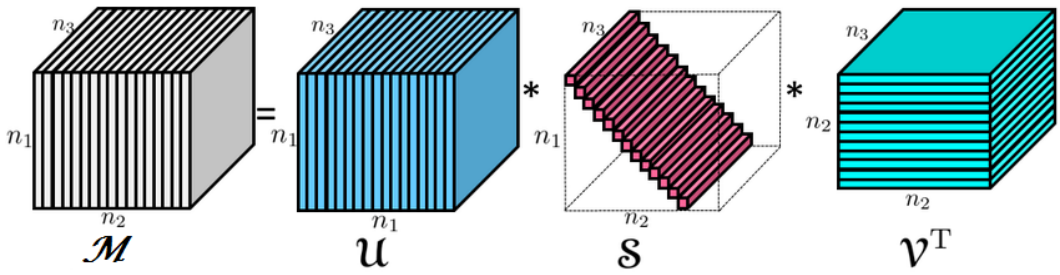

Figure 3.1 – The t-SVD of a tensor.

3.2.3 Some Natural Tensor Norms

Using the previous definitions, it is easy to define some generalisations of the usual matrix norms.

Definition 3.2.13 (Tensor Frobenius norm) *The induced Frobenius norm from the inner product defined above is given by,*

$$\|A\|_F = \langle A, A\rangle^{1/2} = \frac{1}{\sqrt{n_3}}\|\hat{A}\|_F = \sqrt{\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3} A_{ijk}^2}. \tag{3.9}$$

Definition 3.2.14 (Tensor spectral norm) *The tensor spectral norm $\|A\|_{S_\infty}$ is defined as follows:*

$$\|A\|_{S_\infty} = \max_i \|\sigma(A)_i\|_2 \tag{3.10}$$

where $\|.\|_2$ is the l_2-norm.

Proposition 3.2.15 *Let M be $n_1 \times n_2 \times n_3$ tensor. Therefore*

$$\|M\|_{S_\infty} = \|\mathcal{F}(M)\|_{S_\infty},$$

where $\mathcal{F}$ corresponds to the Fast Fourier Transform.

Definition 3.2.16 (Tubal nuclear norm) *The tensor nuclear norm of a tensor A denoted as $\|A\|_{\circledast}$ is the sum of singular values of all the frontal slices of A. Moreover,*

$$\begin{aligned}\|A\|_{\circledast} &= \sum_{i=1}^{\min\{n_1,n_2\}} \sqrt{\sum_{j=1}^{n_3} S(i,i,j)^2} \\ &= \sum_{i=1}^{\min\{n_1,n_2\}} \|\sigma(A)_i\|_2. \end{aligned} \tag{3.11}$$

Note that by Parseval's inequality

$$\sqrt{\sum_{j=1}^{n_3} S(i,i,j)^2} = \frac{1}{\sqrt{n_3}} \sqrt{\sum_{j=1}^{n_3} \hat{S}(i,i,j)^2}. \tag{3.12}$$

Therefore, it is equivalent to define the tubal-nuclear norm via in the Fourier domain. Recall moreover that the $\hat{S}(i,i,j)$ are all non-negative due to the fact that $\hat{U}^{(k)}\hat{S}^{(k)}\hat{V}^{(k)^\top}$ is the SVD of the k^{th} slice of A.

Proposition 3.2.17 (Trace duality property) *Let A, B be $n_1 \times n_2 \times n_3$ tensor. Therefore*

$$|\langle A, B\rangle| \leqslant \|A\|_{\circledast} \|B\|_{S_\infty}.$$

Proof: By Cauchy-Schwartz, we have

$$\begin{aligned} |\langle A, B\rangle| &= |\langle \mathcal{F}(A), \mathcal{F}(B)\rangle| \\ &= |\langle \mathcal{F}(U)\mathcal{F}(S)\mathcal{F}(V^\top), \mathcal{F}(B)\rangle| \\ &= \left| \sum_{i=1}^{n_3} \operatorname{tr}\left(\hat{S}^{(i)} \hat{V}^{(i)^\top} \mathcal{F}(B)^{(i)^\top} \hat{U}^{(i)} \right) \right| \\ &= \left| \sum_{i=1}^{n_3} \sum_{j=1}^{\min\{n_1,n_2\}} \hat{S}^{(i)}_{jj} \left(\hat{V}^{(i)^\top} \mathcal{F}(B)^{(i)^\top} \hat{U}^{(i)} \right)_{jj} \right| \\ &\leqslant \sum_{j=1}^{\min\{n_1,n_2\}} \left(\|\hat{S}_{jj}\|_2 \right)^{1/2} \left(\|(\hat{V}^\top \mathcal{F}(B)^t \hat{U})_{jj}\|_2 \right)^{1/2} \\ &\leqslant \sum_{j=1}^{\min\{n_1,n_2\}} \left(\|\hat{S}_{jj}\|_2 \right)^{1/2} \left(\|\mathcal{F}(B)_{jj}\|_2 \right)^{1/2}, \end{aligned}$$

taking the maximum of $\|\mathcal{F}(B)_{jj}\|_2$ and the sum the slices of $\left(\|\hat{S}_{jj}\|_2\right)^{1/2}$, and apply (4.15) and inverse of FFT, we obtain the result. □

Proposition 3.2.18 *Given tensor $A \in \mathbb{R}^{n_1 \times n_2 \times n_3}$. We have*

$$\|A\|_{\circledast} \leqslant \sqrt{rank(A)} \|A\|_F.$$

Proof: Again by Cauchy-Schwartz, we have

$$\begin{aligned}
\|A\|_{\circledast} &= \sum_{j=1}^{\min\{n_1,n_2\}} \|S(j,j,:)\|_2 \\
&= \sum_{j=1}^{\text{rank(A)}} \|S(j,j,:)\|_2 \\
&\leqslant \sqrt{\text{rank}(A)} \left(\sum_{j=1}^{\min\{n_1,n_2\}} \|S(j,j,;)\|_2^2 \right)^{1/2} \\
&\leqslant \sqrt{\text{rank}(A)} \|A\|_F.
\end{aligned}$$

□

Lemma 3.2.19 *We have*

$$\|P_{S_1^\perp} \hat{A}^\lambda P_{S_2^\perp}\|_{\circledast} = \max_{\|W\|_{S_\infty} \leqslant 1} \langle W, P_{S_1^\perp} \hat{A}^\lambda P_{S_2^\perp} \rangle.$$

Proof:

$$\begin{aligned}
\langle W, P_{S_1^\perp} \hat{A}^\lambda P_{S_2^\perp} \rangle &= \langle \mathcal{F}(W), \mathcal{F}(P_{S_1^\perp} \hat{A}^\lambda P_{S_2^\perp}) \rangle \quad \textit{denote } B = \mathcal{F}(P_{S_1^\perp} \hat{A}^\lambda P_{S_2^\perp}) \text{ and take this t-svd} \\
&= \sum_{i=1}^{n_3} \text{tr}\left(\hat{V}^{(i)^\top} \mathcal{F}(W)^{(i)^\top} \hat{U}^{(i)} \hat{S}^{(i)} \right) \\
&= \sum_{i=1}^{n_3} \sum_{j=1}^{\min\{n_1,n_2\}} \left(\hat{V}^{(i)^\top} \mathcal{F}(W)^{(i)^\top} \hat{U}^{(i)} \right)_{jj} \hat{S}_{jj}^{(i)}
\end{aligned}$$

If we take $\mathcal{F}(W)$ such as $\left(\hat{V}^\top \mathcal{F}(W)^\top \hat{U} \right)_{jj}$ is colinear with $\hat{S}_{jj}$, i.e, $\forall$, $j = 1, \ldots, \min\{n_1, n_2\}$

$$\left(\hat{V}^\top \mathcal{F}(W)^\top \hat{U} \right)_{jj} = \alpha_j \hat{S}_{jj} \Leftrightarrow \left(\hat{V} \hat{S}_W \hat{U}^\top \right)_{jj} = \alpha_j \hat{S}_{jj} \text{ with } \alpha_j \leqslant 1.$$

This means to solve the equation system to determine the α_j. Thus,

$$|\alpha_j \|\hat{S}_{jj}\|_2| = 1 \implies |\alpha_j| = \frac{1}{\|\hat{S}_{jj}\|_2}.$$

With this result, the remaining of proof follows directly from the proof of Proposition 4.3.17. □

3.2.4 Rank, Range and Kernel

The rank, the range and the kernel are extremely important notions for matrices. They will play a role in our analysis of the penalised least squares tensor recovery procedure as well.

As noticed in Reference [38], a tubal scalar may have all its entrees different from zero but still be non-invertible. According to the definition, a tubal scalar $a \in \mathbb{R}^{1\times 1\times n_3}$ is invertible if there exists a tubal scalar b such that $a * b = b * a = e$. Equivalently, the Fourier transform $\hat{a}$ of a has no coefficient equal to zero. We can define the tubal rank ρ_i of $S_{i,i,:}$ as the number of non-zero components of $\hat{S}(i,i,:)$. Then, the easiest way to define the rank of a tensor is by means of the notion of multirank as follows.

Definition 3.2.20 *The multirank of a tensor is the vector $(\rho_1, \ldots, \rho_r)$ where r is the number of nonzero tubal vectors $S(i,i,:)$, $i = 1, \ldots, \min\{n_1, n_2\}$.*

We now define the range of a tensor.

Definition 3.2.21 *Let j denote the number of invertible tubal eigenvalues and let k denote the number of nonzero noninvertible tubal eigenvalues. The range $\mathcal{R}(M)$ of a tensor $M \in \mathbb{R}^{n_1\times n_2\times n_3}$ is defined as*

$$\mathcal{R}(M) = \left\{ \vec{U}^{(1)} * c_1 + \cdots + \vec{U}^{(j+k)} * c_{j+k} \mid c_l \in \mathcal{R}ange(s_l * \cdot),\ l \in \{j+1, \cdots, j+k\} \right\}. \tag{3.13}$$

Definition 3.2.22 *Let j denote the number of invertible tubal eigenvalues. The kernel $\mathcal{K}(M)$ of a tensor $M \in \mathbb{R}^{n_1\times n_2\times n_3}$ is defined as*

$$\mathcal{K}(M) = \left\{ \vec{V}^{(j+1)} * c_1 + \cdots + \vec{V}^{(n_2)} * c_{n_2} \mid s_l * c_l = 0,\ l \in \{j+1, \cdots, j+n_2\} \right\}. \tag{3.14}$$

3.3 Measurement Model and the Estimator

3.3.1 The Observation Model

In the model considered hereafter, the observed data are Y_1,...,Y_n given by the following model

$$Y_i \ = \ \langle X_i, A_0 \rangle \ + \ \xi_i, \ \ i = 1, \ldots, n,$$

where the notation $\langle \cdot, \cdot \rangle$ stands for the canonical scalar product of tensors. This can be seen as a tensor regression problem X_i, $i = 1, \ldots, n$ are some tensors in $\mathbb{R}^{n_1\times n_2\times n_3}$ and ξ_i, $i = 1, \ldots, n$ are some independent zero mean random variables. Assume that the frontal faces $X^{(i)}$ are i.i.d uniformly distributed on the set

$$\mathcal{X} = \{e_j(n_1)e_k^\top(n_2), 1 \leqslant j \leqslant n_1, 1 \leqslant k \leqslant n_2\}, \tag{3.15}$$

where $e_k(n)$ are the canonical basis vectors in $\mathbb{R}^n$. Our goal is to recover the tensor A_0 based on the data Y_i, $i = 1, \ldots, n$ only for n as small as possible.

Definition 3.3.1 *For any tensors $A, B \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, we define the scalar product*

$$\langle A, B \rangle = \sum_{i_1=1,\ldots,n_1} \sum_{i_2=1,\ldots,n_2} \sum_{i_3=1,\ldots,n_3} A_{i_1,i_2,i_3} B_{i_1,i_2,i_3},$$

and the bilinear form

$$\langle A, B \rangle_{L_2(\Pi)} = \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}\left[\langle A, X_i \rangle \langle B, X_i \rangle\right]$$

Here $\Pi = \frac{1}{n}\sum_{i=1}^{n} \Pi_i$, where Π_i denotes the distribution of X_i. The corresponding semi-norm $\|A\|^2_{L_2(\Pi)}$ is given by

$$\|A\|^2_{L_2(\Pi)} = \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}\left[\langle A, X_i \rangle^2\right]$$

and will denote by M the tensor given by

$$M \;=\; \frac{1}{n} \sum_{i=1}^{n} \left(Y_i X_i - \mathbb{E}\left[Y_i X_i\right]\right).$$

3.3.2 The Estimator

The approach proposed in Reference [39] for low rank matrix estimation which will be extended to tensor estimation in the present paper consists in minimising

$$\hat{A}_\lambda \in \operatorname{argmin}_{A \in \mathbb{A}} \; L_n(A), \tag{3.16}$$

where

$$L_n(A) = \|A\|^2_{L_2(\Pi)} - \left\langle \frac{2}{n} \sum_{i=1}^{n} Y_i X_i, A \right\rangle + \lambda \; \|A\|_{\circledast}, \tag{3.17}$$

where $\|\cdot\|_{\circledast}$ is a tubal tensor nuclear norm that we will introduce in Definition 3.2.16.

Recall that nuclear norm penalisation is widely used in sparse estimation when the matrix to be recovered is low rank [40]. Following the success of the application of sparsity to low rank matrix recovery, several extensions of the matrix nuclear norm were proposed in the literature [47, 80, 46], etc. Another type of nuclear norm was proposed in Reference [71] based on the tubal framework of Kilmer [38]. Our estimator is another nuclear norm penalisation based estimator. As it will be explained in Section 3.2 below, the nuclear norm used in the present paper has some advantages over other norms in the context of tubal low rank tensors and the resulting estimator is most relevant in many applications where we want to recover a tensor which is the sum of a small number of rank-one tensors.

3.4 Main Results

In this section, we provide our main results. First, in Section 3.4.2, we give a complete characterisation of the subdifferential of the nuclear norm (Definition 3.2.16). Then, we propose a statistical study of the nuclear-norm penalised estimator $\hat{A}_\lambda$ in Section 3.4.3.

3.4.1 Preliminary Remarks

3.4.1.1 Orthogonal Invariance

It is easy to see that the t-nuclear norm is orthogonally invariant. Indeed, consider two orthogonal tensors $O_1 \in \mathbb{R}^{n_1 \times n_1 \times n_3}$, $O_2 \in \mathbb{R}^{n_2 \times n_2 \times n_3}$, $k = 1, \dots, n_3$. Since the product of two orthogonal tensors is itself orthogonal, we have

$$\sigma(A) = \sigma(O_1 * A * O_2^\top). \tag{3.18}$$

3.4.1.2 Support of a Tensor

Given that the t-svd of a tensor is

$$A = U * S * V^\top \tag{3.19}$$

$$= \sum_{k=1}^{\min\{n_1,n_2\}} \vec{U}^{(k)} * \vec{s}_{kk} * \vec{V}^{(k)^\top}, \tag{3.20}$$

with $(\vec{U}^{(1)}, \dots, \vec{U}^{(\min\{n_1,n_2\})})$ is a family of orthonormal matrix in $\mathbb{R}^{n_1 \times n_1}$, and $(\vec{V}^{(1)}, \dots, \vec{V}^{(\min\{n_1,n_2\})})$ a family of orthonormal matrix in $\mathbb{R}^{n_2 \times n_2}$ and $\vec{s}_{kk} = S(k,k,:)$ are the spectrum of A.

The support of A is the couple of linear vector spaces (S_1, S_2) of tubal tensors, where

- S_1 is the linear span of $(\vec{U}^{(1)}, \dots, \vec{U}^{(\min\{n_1,n_2\})})$ and
- S_2 is the linear span of $(\vec{V}^{(1)}, \dots, \vec{V}^{(\min\{n_1,n_2\})})$.

We also let $S_j^\perp$, $j = 1, 2$ to be the orthogonal complements of S_j and by P_{S_j}, $j = 1, 2$, the projector on linear vector subspace S of tubal tensors.

3.4.2 The Subdifferential of the t-nuclear Norm

Our first result is a characterisation of the subdifferential of $\|\cdot\|_\circledast$. Recall first the particular case of the matrix nuclear norm $\|\cdot\|_*$. By Corollary 2.5 in [85], we have

$$\partial \|\cdot\|_* = \{UV^t + W \mid \|W\| \leqslant 1,\ U^t W = 0,\ VW = 0\}.$$

This result is established in [85] using the Von-Neumann inequality and in particular the equality case of this inequality.

3.4.2.1 Von-Neumann's Inequality for Tubal Tensors

Theorem 3.4.1 *Let A, B be $n_1 \times n_2 \times n_3$ tensor. Therefore*

$$\langle A, B \rangle \leqslant \langle S^A, S^B \rangle \tag{3.21}$$

where S^A is a rectangular f-diagonal tensor, contains all the singular values of A. Equality holds in (3.21) *if and only if A, B have the same singular tensors.*

Proof: Let $\mathcal{F}$ denote the Fast Fourier Transform. We have

$$\begin{aligned}
\langle A, B \rangle &= \langle \mathcal{F}(A), \mathcal{F}(B) \rangle \\
&= \left\langle \mathcal{F}\left(U^A * S^A * \left(V^{A^\top}\right)\right), \mathcal{F}\left(U^B * S^B * \left(V^{B^\top}\right)\right) \right\rangle \\
&= \left\langle \mathcal{F}\left(U^A\right) \cdot \mathcal{F}\left(S^A\right) \cdot F\left(V^{A^\top}\right), \mathcal{F}\left(U^B\right) \cdot \mathcal{F}\left(S^B\right) \cdot \mathcal{F}\left(V^{B^\top}\right) \right\rangle \\
&= \sum_{k=1}^{n_3} \left\langle \mathcal{F}\left(U^A\right)^{(k)} \cdot \mathcal{F}\left(S^A\right)^{(k)} \cdot \left(\mathcal{F}\left(V^{A^\top}\right)\right)^{(k)}, \right. \\
&\qquad \left. \mathcal{F}\left(U^B\right)^{(k)} \cdot \mathcal{F}\left(S^B\right)^{(k)} \cdot \left(\mathcal{F}\left(V^{B^\top}\right)\right)^{(k)} \right\rangle.
\end{aligned}$$

Thus, using the Von Neumann inequality for matrices, we get

$$\begin{aligned}
&\leqslant \sum_{k=1}^{n_3} \langle \mathcal{F}\left(S^A\right)^{(k)}, \mathcal{F}\left(S^B\right)^{(k)} \rangle \\
&\leqslant \langle \mathcal{F}\left(S^A\right), \mathcal{F}\left(S^B\right) \rangle \\
&= \langle S^A, S^B \rangle,
\end{aligned}$$

where the last equality stems from the isometry property of Fast Fourier Transform. □

Notice that the Von-Neumann inequality was extended in [87] to general tensors and exploited in [88] for the computation of the subdifferential of some tensor norms. In comparison, the case of the t-nuclear norm only need an appropriate use of the matrix Von Neumann inequality.

3.4.2.2 Lewis's Characterization of the Subdifferential

Theorem 3.4.2 *Let $f : \mathbb{R}^{\min\{n_1,n_2\}} \mapsto \mathbb{R}$ be a function convex, so :*

$$(f \circ \sigma)^* = f^* \circ \sigma.$$

The proof is exactly the same as in [85], Theorem 2.4.

Theorem 3.4.3 *Let us suppose that the function $f : \mathbb{R}^{\min\{n_1,n_2\}} \longrightarrow \mathbb{R}$ is convex. Then, the tensor*

$$Y \in \partial (f \circ \sigma)(X) \text{ if and only if } \sigma(Y) \in \partial f(\sigma(X)). \tag{3.22}$$

The proof is the same as in [85], Corollary 2.5 where the matrix Von Neumann inequality (and more precisely, the exact characterization of the equality case) is replaced with Von Neumann inequality for tubal tensors given by Theorem 4.1.

Theorem 3.4.4 *Let r denote the number of $\vec{S}_X^{(k)}$ which are non-zero. The subdifferential of the t-nuclear norm is given by*

$$\partial\|X\|_{\circledast} = \Big\{U * D(\vec{\mu}) * V^t + W| \;\; \vec{\mu}_k = \vec{S}_X^{(k)}/\|\vec{S}_X^{(k)}\|_2, \;\; k = 1, \ldots, r, \\ \|\vec{\mu}_k\|_2 = 1, \;\; r < k \leqslant \min\{n_1, n_2\}\Big\}. \tag{3.23}$$

Proof: We only need to rewrite (3.22) using the well-known formula for the subdifferential of the Euclidean norm. We provide the details for the sake of completeness.

Let $\vec{S}_X$ be a t-vector in $\mathbb{R}^{n_1\times 1\times n_3}$.

$$\begin{aligned} \partial f(\sigma(X)) &= \partial\left(\sum_{k=1}^{\min\{n_1,n_2\}} \|\vec{S}_X^{(k)}\|_2\right) \\ &= \partial\|\vec{S}_X^{(1)}\|_2 \times \ldots \times \|\vec{S}_X^{(\min\{n_1,n_2\})}\|_2 \end{aligned} \tag{3.24}$$

where in second equality we use the result of [89], Chapter 16,l Section 1, Proposition 16.8.

As is well known, the subgradient of the l_2-norm is

$$\partial\|w\|_2 = \begin{cases} \left\{\dfrac{w}{\|w\|_2}\right\} & \text{if} \quad w \neq 0 \\ \\ \{z, \|z\|_2 \leqslant 1\} & \text{if} \quad w = 0 \end{cases}$$

and plugg this formula into (3.24). Therefore

$$\partial(f\circ\sigma)(X) = \Big\{U * D(\vec{\mu}) * V^\top | \, \vec{\mu} \in \partial f(\sigma(X)), U \in \mathbb{R}^{n_1\times n_1\times n_3}, V \in \mathbb{R}^{n_2\times n_2\times n_3}, \quad X = U * S * V^\top\Big\}. \tag{3.25}$$

Let T be the set of indices j which tubal scalar $\sigma(X)_j \neq 0$. Thus,

$$U * D(\vec{\mu}) * V^\top = U_T * D(\vec{\mu}_T) * V_T^\top + U_{T^c} * D(\vec{\mu}_{T^c}) * V_{T^c}^\top$$

Moreover,

$$\vec{\mu}_T = \begin{bmatrix} \dfrac{\vec{S}_X^{(1)}}{\|\vec{S}_X^{(1)}\|_2} \\ \vdots \\ \dfrac{\vec{S}_X^{(\min\{n_1,n_2\})}}{\|\vec{S}_X^{(\min\{n_1,n_2\})}\|_2} \end{bmatrix}. \tag{3.26}$$

Therefore, $U * D(\vec{\mu}) * V^\top$ is of the form

$$U_T * D(\vec{\mu}_T) * V_T^\top + W$$

with $W = U_{T^c} * D(\vec{\mu}_{T^c}) * V_{T^c}^\top$ and

$$\vec{\mu}_{T^c} \in \{z, \|z\|_2 \leqslant 1\}^{\times |T^c|}.$$

So we have,

$$\|W\|_{S_\infty} \leqslant 1,\ U_T^\top * W = 0 \text{ and } W * V_T = 0.$$

□

3.4.3 Error Bound

Our error bound on tubal-tensor recovery using the approach of Koltchinskii, Lounici and Tsybakov [39] is given in the following theorem.

Theorem 3.4.5 *Let $\mathbb{A} \subseteq \mathbb{R}^{n_1 \times n_2 \times n_3}$ be a convex set of tensors. Let $\hat{A}^\lambda$ be defined by*

$$\hat{A}^\lambda = \underset{A \in \mathbb{A}}{\text{Argmin}}\ \|A\|_{L_2(\Pi)}^2 - 2\left\langle \frac{1}{n}\sum_{i=1}^n Y_i X_i, A \right\rangle + \lambda \|A\|_{\circledast}, \tag{3.27}$$

where we recall that $\|\cdot\|_{\circledast}$ denotes the tensor nuclear norm (see Definition 3.2.16). Assume that there exists a constant $\rho > 0$ such that, for all tensor $A \in \mathbb{A} - \mathbb{A} := \{A_1 - A_2 : A_1, A_2 \in \mathbb{A}\}$,

$$\|A\|_{L_2(\Pi)}^2 \ \geq\ \rho^{-2} \|A\|_F^2. \tag{3.28}$$

If $\lambda \geq 2\ \|M\|_{S_\infty}$, then

$$\|\hat{A}^\lambda - A_0\|_{L_2(\Pi)}^2 \ \leqslant\ \inf_{A \in \mathbb{A}} \left[\|A - A_0\|_{L_2(\Pi)}^2 + \left(\frac{1+\sqrt{2}}{2}\right)^2 \rho^2 \lambda^2\ \text{rank}(A) \right].$$

Proof: We follow the proof of Reference [39]. We provide all the details for the sake of the completeness.

$$\hat{A}^\lambda \in \underset{A \in \mathbb{A}}{\text{Argmin}}\ L_n(\hat{A}^\lambda) = \|A\|_{L_2(\Pi)}^2 - 2\left\langle \frac{1}{n}\sum_{i=1}^n Y_i X_i, A \right\rangle + \lambda \|A\|_{\circledast}. \tag{3.29}$$

Let us compute the directional derivative of L_n at $\hat{A}^\lambda$.

$$DL_n(\hat{A}^\lambda; h) = \lim_{t \to 0} \frac{L_n(\hat{A}^\lambda + th) - L_n(\hat{A}^\lambda)}{t}.$$

The optimality condition of $\hat{A}^\lambda$ implies $DL_n(\hat{A}^\lambda; A - \hat{A}^\lambda) \geq 0, \quad \forall A \in \mathbb{A}$. Thus

$$\begin{aligned} L_n(\hat{A}^\lambda + th) &= \|\hat{A}^\lambda + th\|^2_{L_2(\Pi)} - 2\left\langle \frac{1}{n}\sum_{i=1}^n Y_i X_i, \hat{A}^\lambda + th \right\rangle + \lambda\|\hat{A}^\lambda + th\|_\circledast \\ &= \|A\|^2_{L_2(\Pi)} + 2t\langle \hat{A}^\lambda, h\rangle_{L_2(\Pi)} + t^2\|h\|^2_{L_2(\Pi)} - 2\left\langle \frac{1}{n}\sum_{i=1}^n Y_i X_i, \hat{A}^\lambda \right\rangle \\ &\quad -2t\left\langle \frac{1}{n}\sum_{i=1}^n Y_i X_i, h \right\rangle + \lambda\|\hat{A}^\lambda + th\|_\circledast \end{aligned}$$

$$\begin{aligned} \lim_{t\to 0} \frac{L_n(A+th) - L_n(A)}{t} &= 2\langle \hat{A}^\lambda, h\rangle_{L_2(\Pi)} - 2\left\langle \frac{1}{n}\sum_{i=1}^n Y_i X_i, h \right\rangle \\ &\quad + \lambda\left(\|\hat{A}^\lambda + th\|_\circledast - \|\hat{A}^\lambda\|_\circledast\right) \geq 0 \end{aligned}$$

for all $h \in T_{\mathbb{A}}(\hat{A}^\lambda)$ (the tangent cone to $\mathbb{A}$ at $\hat{A}^\lambda$). Taking $h = A - \hat{A}^\lambda$, we have,

$$2\langle \hat{A}^\lambda, A - \hat{A}^\lambda\rangle_{L_2(\Pi)} - 2\left\langle \frac{1}{n}\sum_{i=1}^n Y_i X_i, A - \hat{A}^\lambda \right\rangle + \lambda \lim_{t\to 0}\left(\frac{\|\hat{A}^\lambda + th\|_\circledast - \|\hat{A}^\lambda\|_\circledast}{t}\right) \geq 0. \quad (3.30)$$

On the other hand,

$$\begin{aligned} \lim_{t\to 0}\left(\frac{\|\hat{A}^\lambda + th\|_\circledast - \|\hat{A}^\lambda\|_\circledast}{t}\right) &= \max_{G\in\partial\|.\|(\hat{A}^\lambda)} \langle G, A - \hat{A}^\lambda\rangle \\ &= \langle \hat{G}^\lambda, A - \hat{A}^\lambda\rangle \end{aligned} \quad (3.31)$$

by compactness of $G \in \partial\|.\|_\circledast(\hat{A}^\lambda)$; see [90]. Combining (3.31) with (3.30), we obtain

$$2\langle \hat{A}^\lambda, \hat{A}^\lambda - A\rangle_{L_2(\Pi)} - 2\left\langle \frac{1}{n}\sum_{i=1}^n Y_i X_i, \hat{A}^\lambda - A \right\rangle + \lambda\langle \hat{G}, \hat{A}^\lambda - A\rangle \leq 0. \quad (3.32)$$

Consider an arbitrary tensor $A \in \mathbb{A}$ of tubal rank r with spectral representation

$$A = U * S * V^t = \sum_{j=1}^r \vec{U}^{(j)} * \vec{s}_{jj} * \vec{V}^{(j)^t} \quad (3.33)$$

where $\vec{s}_{jj} = S(j, j, :)$, and with support (S_1, S_2). Using that

$$\left\langle \frac{1}{n}\sum_{i=1}^n \mathbb{E}[Y_i X_i], \hat{A}^\lambda - A \right\rangle = \langle A_0, \hat{A}^\lambda - A\rangle_{L_2(\Pi)} \quad (3.34)$$

it thus follows from (3.32) that

$$2\langle \hat{A}^\lambda - A_0, \hat{A}^\lambda - A\rangle_{L_2(\Pi)} + \lambda\langle \hat{G}^\lambda - G, \hat{A}^\lambda - A\rangle \leqslant -\lambda\langle G, \hat{A}^\lambda - A\rangle + 2\langle M, \hat{A}^\lambda - A\rangle \quad (3.35)$$

with

$$M \;=\; \frac{1}{n}\sum_{i=1}^{n}\left(Y_iX_i - \mathbb{E}\left[Y_iX_i\right]\right).$$

By the monotonicity of subdifferentials of convex functions, we have $\langle \hat{G}-G, \hat{A}^\lambda - A\rangle \geq 0$ (cf. [91], Chapter 4, Section 3, Proposition 9). Therefore

$$2\langle \hat{A}^\lambda - A_0, \hat{A}^\lambda - A\rangle_{L_2(\Pi)} \leqslant -\lambda\langle G, \hat{A}^\lambda - A\rangle + 2\langle M, \hat{A}^\lambda - A\rangle. \tag{3.36}$$

Furthermore, by (3.23), the following representations holds

$$G = \sum_{j=1}^{r} \vec{U}^{(j)} * \vec{\mu}_j * \vec{V}^{(j)^\top} + P_{S_1^\perp} W P_{S_2^\perp}$$

where W is an arbitrary tensor with $\|W\|_{S_\infty} \leqslant 1$ and

$$\begin{aligned}\langle P_{S_1^\perp} W P_{S_2^\perp}, \hat{A}^\lambda - A\rangle &= \langle P_{S_1^\perp} W P_{S_2^\perp}, \hat{A}^\lambda\rangle \\ &= \langle W, P_{S_1^\perp}\hat{A}^\lambda P_{S_2^\perp}\rangle\end{aligned}$$

and using Proposition 4.3.18, we can choose W such that

$$\langle P_{S_1^\perp} W P_{S_2^\perp}, \hat{A}^\lambda - A\rangle = \|P_{S_1^\perp}\hat{A}^\lambda P_{S_2^\perp}\|_{\circledast},$$

where in the first equality we used that A has support (S_1, S_2). For this particular choice of W, (3.36) implies that

$$\begin{aligned}2\langle \hat{A}^\lambda - A_0, \hat{A}^\lambda - A\rangle_{L_2(\Pi)} + \lambda\|P_{S_1^\perp}\hat{A}^\lambda P_{S_2^\perp}\|_{\circledast} \leqslant -\lambda\left\langle \sum_{j=1}^{r} \vec{U}^{(j)} * \vec{\mu}_j * \vec{V}^{(j)^\top}, \hat{A}^\lambda - A\right\rangle \\ + 2\langle M, \hat{A}^\lambda - A\rangle\end{aligned} \tag{3.37}$$

Using the identity

$$2\langle \hat{A}^\lambda - A_0, \hat{A}^\lambda - A\rangle_{L_2(\Pi)} \;=\; \|\hat{A}^\lambda - A_0\|^2_{L_2(\Pi)} \;+\; \|\hat{A}^\lambda - A\|^2_{L_2(\Pi)} \;-\; \|A - A_0\|^2_{L_2(\Pi)}$$

and the fact that

$$\|\sum_{j=1}^{r} \vec{U}^{(j)} * \vec{\mu}_j * \vec{V}^{(j)^\top}\|_{S_\infty} = \|D(\vec{\mu})\|_{S_\infty}$$

$$\left\langle \sum_{j=1}^{r} \vec{U}^{(j)} * \vec{\mu}_j * \vec{V}^{(j)^\top}, \hat{A}^\lambda - A\right\rangle = \left\langle \sum_{j=1}^{r} \vec{U}^{(j)} * \vec{\mu}_j * \vec{V}^{(j)^\top}, P_{S_1}\left(\hat{A}^\lambda - A\right)P_{S_2}\right\rangle$$

we deduce from (3.37) that

$$\|\hat{A}^\lambda - A_0\|^2_{L_2(\Pi)} + \|\hat{A}^\lambda - A\|^2_{L_2(\Pi)} + \lambda\|P_{S_1^\perp}\hat{A}^\lambda P_{S_2^\perp}\|_\circledast \leqslant \|A - A_0\|^2_{L_2(\Pi)}$$

$$+ \lambda\|D(\vec{\mu})\|_\infty\|P_{S_1}(\hat{A}^\lambda - A)P_{S_2}\|_\circledast + 2\langle M, \hat{A}^\lambda - A\rangle. \tag{3.38}$$

Now, to find an upper bound on $2\langle M, \hat{A}^\lambda - A\rangle$, we use the following decomposition:

$$\begin{aligned}\langle \mathrm{M}, \hat{A}^\lambda - A\rangle &= \langle \mathcal{P}_A(\mathrm{M}), \hat{A}^\lambda - A\rangle + \langle P_{S_1^\perp}\mathrm{M}P_{S_2^\perp}, \hat{A}^\lambda - A\rangle \\ &= \langle \mathcal{P}_A(\mathrm{M}), \mathcal{P}_A(\hat{A}^\lambda - A)\rangle + \langle P_{S_1^\perp}\mathrm{M}P_{S_2^\perp}, \hat{A}^\lambda\rangle\end{aligned}$$

where $\mathcal{P}_A(M) = M - P_{S_1^\perp}MP_{S_2^\perp}$. This implies, due to the trace duality,

$$\begin{aligned}2|\langle \mathrm{M}, \hat{A}^\lambda - A\rangle| &= 2|\langle \mathcal{P}_A(M), \mathcal{P}_A(\hat{A}^\lambda - A)\rangle + \langle P_{S_1^\perp}\mathrm{M}P_{S_2^\perp}, \hat{A}^\lambda\rangle| \\ &\leqslant 2|\langle \mathcal{P}_A(M), \mathcal{P}_A(\hat{A}^\lambda - A)\rangle| + 2|\langle P_{S_1^\perp}\mathrm{M}P_{S_2^\perp}, \hat{A}^\lambda\rangle| \\ &\leqslant 2\|\mathcal{P}_A(M)\|_F\|\mathcal{P}_A(\hat{A}^\lambda - A)\|_F + 2\|P_{S_1^\perp}\mathrm{M}P_{S_2^\perp}\|_{S_\infty}\|P_{S_1^\perp}\hat{A}^\lambda P_{S_2^\perp}\|_\circledast \\ &\leqslant \Lambda\|\hat{A}^\lambda - A\|_F + \Gamma\|P_{S_1^\perp}\hat{A}^\lambda P_{S_2^\perp}\|_\circledast \end{aligned}\tag{3.39}$$

where

$$\Lambda = 2\|\mathcal{P}_A(M)\|_F, \text{ and } \Gamma = 2\|P_{S_1^\perp}\mathrm{M}P_{S_2^\perp}\|_{S_\infty}$$

Using that

$$\mathcal{P}_A(\mathrm{M}) = P_{S_1^\perp}\mathrm{M}P_{S_2} + P_{S_1}\mathrm{M} \text{ and } \mathrm{rank}(P_{S_j}) \leqslant \mathrm{rank}(A),\ j = 1, 2$$

we have with $\Delta = \|M\|_{S_\infty}$

$$\Lambda \leqslant 2\sqrt{\mathrm{rank}(\mathcal{P}_A)(\mathrm{M})}\|\mathrm{M}\|_{S_\infty} \leqslant 2\sqrt{2\mathrm{rank}(A)}\Delta \leqslant \sqrt{2\mathrm{rank}(A)}\lambda$$

Thus,

$$2|\langle \mathrm{M}, \hat{A}^\lambda - A\rangle| \leqslant \sqrt{2\mathrm{rank}(A)}\ \lambda\|\hat{A}^\lambda - A\|_F + 2\Delta\|P_{S_1^\perp}\hat{A}^\lambda P_{S_2^\perp}\|_\circledast \tag{3.40}$$

Due to Proposition 4.3.18, we have

$$\|P_{S_1}(\hat{A}^\lambda - A)P_{S_2}\|_\circledast \leqslant \sqrt{\mathrm{rank}(\hat{A}^\lambda - A)}\|P_{S_1}(\hat{A}^\lambda - A)P_{S_2}\|_F \ \leqslant \sqrt{\mathrm{rank}(\hat{A}^\lambda - A)}\|\hat{A}^\lambda - A\|_F \tag{3.41}$$

and using the assumption (3.28), it follows from (3.38) and (3.40) that

$$\begin{aligned}\|\hat{A}^\lambda - A_0\|^2_{L_2(\Pi)} + \|\hat{A}^\lambda - A\|^2_{L_2(\Pi)} + (\lambda - 2\Delta)\|P_{S_1^\perp}\hat{A}^\lambda P_{S_2^\perp}\|_\circledast \leqslant \|A - A_0\|^2_{L_2(\Pi)} + \\ + \left(\|D(\vec{\mu})\|_{S_\infty} + \sqrt{2}\right)\rho\lambda\sqrt{\mathrm{rank}(A)}\|\hat{A}^\lambda - A\|_{L_2(\Pi)}\end{aligned}\tag{3.42}$$

Using

$$\Big(\|D(\vec{\mu})\|_{S_\infty}+\sqrt{2}\Big)\,\rho\lambda\sqrt{\mathrm{rank}(A)}\|\hat{A}^\lambda-A\|_{L_2(\Pi)}-\|\hat{A}^\lambda-A\|^2_{L_2(\Pi)}$$
$$\leqslant\frac{1}{4}\Big(\|D(\vec{\mu})\|_{S_\infty}+\sqrt{2}\Big)^2\rho^2\lambda^2\mathrm{rank}(A),$$

we deduce from (3.42) that

$$\|\hat{A}^\lambda-A_0\|^2_{L_2(\Pi)}+(\lambda-2\Delta)\|P_{S_1^\perp}\hat{A}^\lambda P_{S_2^\perp}\|_\circledast\leqslant\|A-A_0\|^2_{L_2(\Pi)}$$
$$+\frac{1}{4}\Big(\|D(\vec{\mu})\|_{S_\infty}+\sqrt{2}\Big)^2\rho^2\lambda^2\mathrm{rank}(A)$$

□

In Section, 3.7, it is proved that we can set

$$\rho=\sqrt{mn_1n_2n_3^{3/2}} \tag{3.43}$$

and

$$\lambda=2n_3\|\mathcal{F}(A_0)\|_{S_\infty}\max\left\{\sqrt{\frac{t+\log(m)}{nn_1n_2}};\left(1+\frac{1}{\sqrt{n_1n_2}}\right)\frac{t+\log(m)}{n}\right\}. \tag{3.44}$$

3.5 Numerical Experiments

The proposed methods were numerically validated using 3D OCT images. OCT is widely studied as a medical imaging system in many clinical applications and fundamental research. In numerous clinical purposes, OCT is considered as a very interesting technique for in situ tissue characterization known as "*optical biopsy*" (in opposition to the conventional physical biopsy). OCT is operating under the principle of low coherence interferometry providing micro-meter spatial resolution at several MHz A-scan (1D optical core in z direction) acquisition rate. In these experiments, we found that depth was a different coordinate from the two other coordinates and the tubal SVD approach appeared particularly relevant. One way of circumventing the problem in the case where the three coordinates have the same properties is to perform the reconstruction three times using the proposed method and take the average of the results.

3.5.1 Benefits of Subsampling for OCT

The OCT imaging device, as the case of the most medical imaging systems, obeys two key requirements for successful application of compressed sensing methods: (1) medical imaging is naturally compressible by sparse coding in an appropriate transform domain (e.g.,

wavelet, shearlet transforms, etc.) and (2) OCT scanning system (e.g., Spectral-Domain OCT, the most used) naturally acquire encoded samples, rather than direct pixel samples (e.g., in spatial-frequency encoding). Therefore, the images resulting from Spectral-Domain OCT are sparse in native representation, hence yielding themselves well to various implementations of the ideas from Compressed Sensing theory.

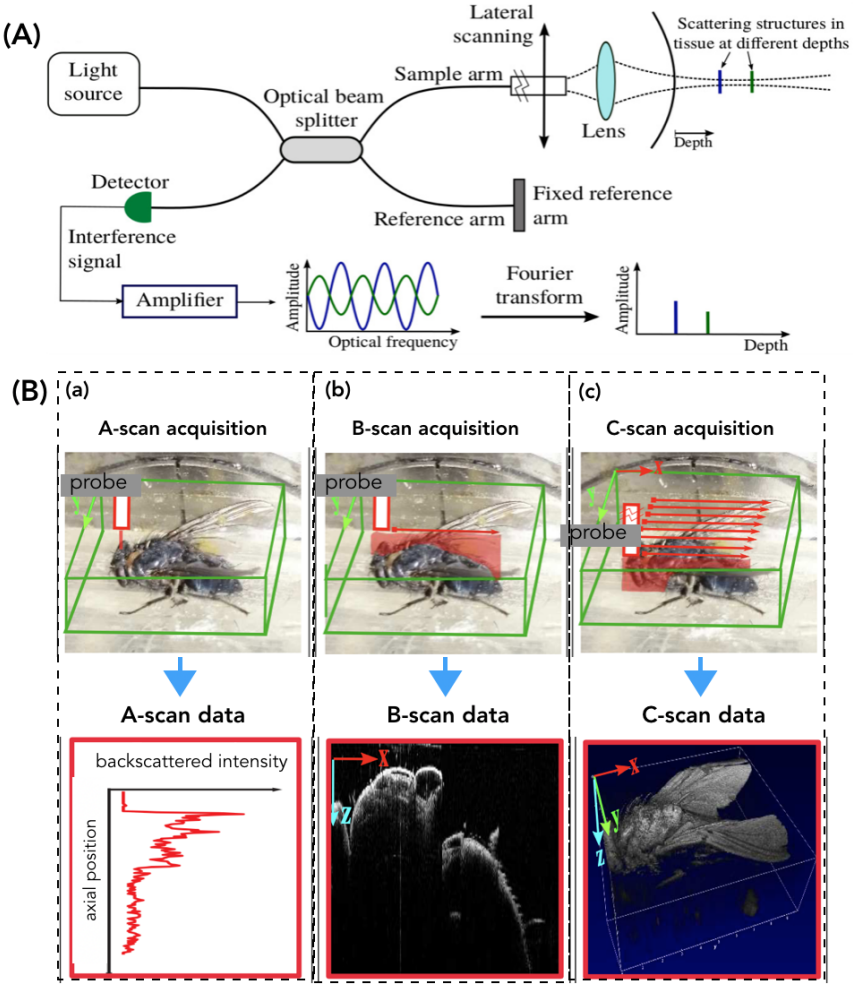

Figure 3.2 – Illustration of the Fourier-Domain Optical Coherence Tomography (OCT) operating. (**A**) the optical principle and (**B**) the different available acquisition modes: A-scan (1D optical core), B-Scan (2D image), and C-scan (volume).

Certainly, OCT enables high-speed A-scan and B-scan acquisitions (Figure 4.3B) but presents a serious issue when it comes to acquiring a C-scan volume. Therefore, especially in case of biological sample/tissue examination, using OCT volumes poses the problem of frame-rate acquisition (generating artifacts) as well as the data transfer i.e., several hundred Mo for each volume (Figure 4.3B).

Indeed, OCT volume data can be considered as a $n_1 \times n_2 \times n_3$ tensor of voxels. Thereby, the mathematical methods and materials related to tensors study are well suited for 3D

OCT data.

3.5.2 Experimental Set-up

An OCT imaging system (*Telesto-II 1325 nm* spectral domain OCT) from ThorLabs (Figure 3.3), is used to validate the proposed distortion models. Axial (resp. lateral) resolution is 5.5 µm (resp. 7 µm) and up to 3.5 mm depth. The *Telesto-II* allows a maximum field-of-view of 10×10×3.5 mm^3 with a maximum A-Scan (optical core) acquisition rate of 76 kHz.

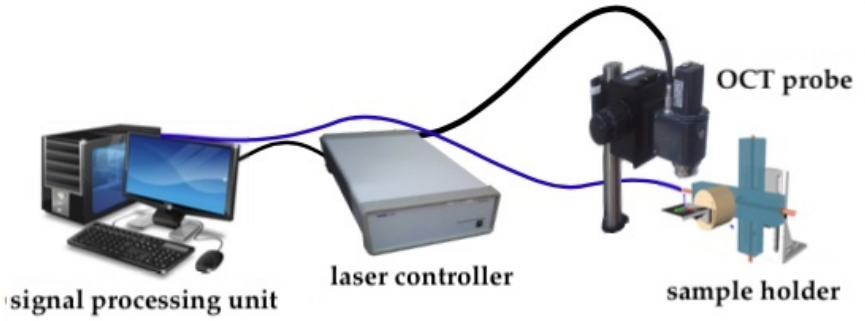

Figure 3.3 – Global view of the OCT acquisition setup.

3.5.3 Validation Scenario

The method proposed in this paper was implemented in a MatLab framework without taking into account the code optimization aspects as well as the time-computation. In order to validate experimentally the approach presented here, we acquired different OCT volumes (C-Scan images) of realistic biological samples: for example, a part of a grape (Figure 4.4(left)) and a part of a fish eye retina (Figure 4.4(right)). The size of the OCT volume used in these numerical validations, for both samples, is $A_{n\times m\times l} = 281 \times 281 \times 281$ voxels.

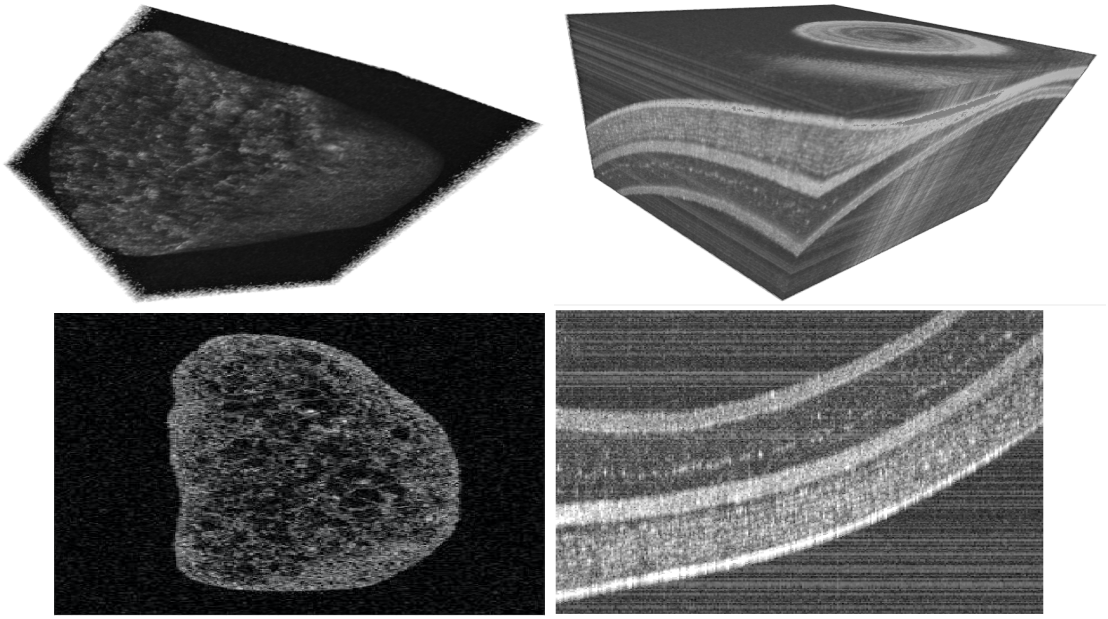

Figure 3.4 – Examples of the OCT volumes of biological samples used to validate the proposed method. (**first row**) the initial OCT volumes and (**second row**), B-scan images (100th vertical slice) taken from the initial volumes.

To access the performance of the proposed algorithm, we constructed several under-sampled 3D volume using 30%, 50%, 70%, and 90% of the original OCT volume. To do this, we created two types of 3D masks. The first consists of a pseudo-random binary masks M_v using a density random sampling in which the data are selected in a vertical manner (Figure 3.5(left)). For the second type of mask, instead of a vertical subsampling of the data, we designed oblique masks M_o as shown in Figure 3.5(right)) which are more appropriate in the case of certain imaging systems such as Magnetic Resonance Imaging (MRI), Computerized Tomography scan (CT-scan), and in certain instances, OCT imaging systems as well.

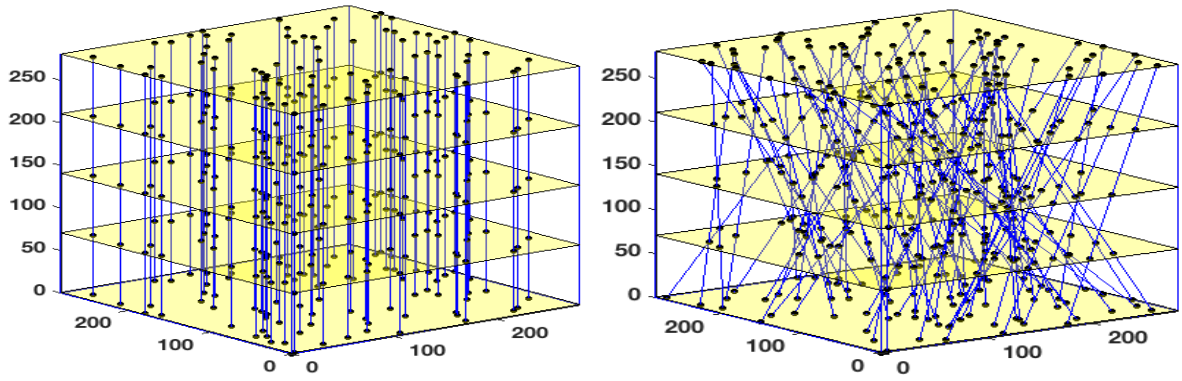

Figure 3.5 – Illustration of the implemented 3D masks used to subsampled the data and creating the 3D masks. (**Left**) a 3D mask allowing a vertical and random selection of the data, and (**right**) a 3D mask allowing an oblique selection of the subsampled data.

3.5.4 Obtained Results

The validation scenarios were carried out as follows—the studied method was applied to each OCT volume (i.e., grape or fish eye) using various subsampling rates (i.e., 30%, 50%, 70%, and 90%). Also, in each case, the vertical M_v or the oblique M_o 3D binary masks were used. The results obtained with our nuclear norm penalised reconstruction approach are discussed below.

3.5.4.1 Grape Sample

Figures 3.6 and 3.7 depict the different reconstructed OCT volumes using the oblique and the vertical binary masks. Instead of illustrating the fully reconstructed OCT volume, we choose to show 2D images (the 100th xy slice of the reconstructed volumes) for a better visualization, with the naked eye, of the quality of the obtained results. It can be highlighted that the sharpness of the boundary is well preserved; however, it loses some features in the zones where the image has low intensity. This is a common effect of most of the compressed sensing and inpainting methods. In order to improve the quality of the recovered data, conventional filters based post-processing could be also be used in order to enhance contrast.

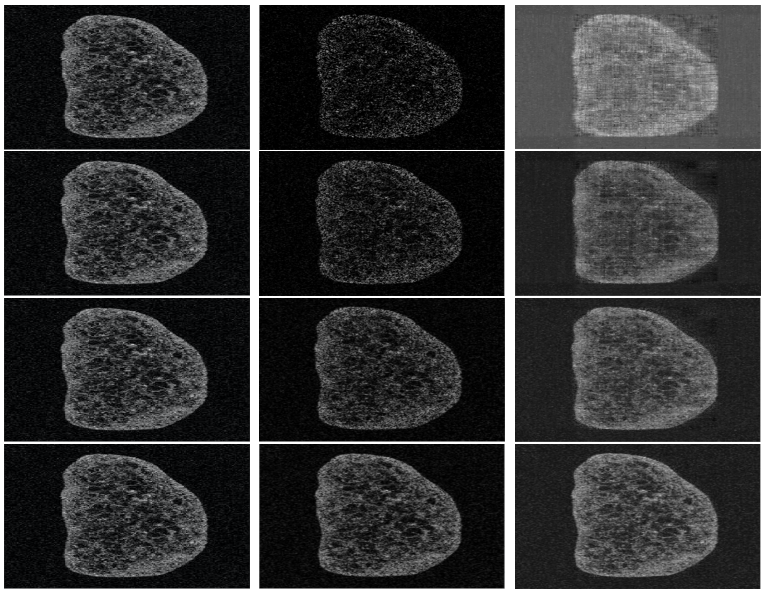

Figure 3.6 – [**sample: grape, mask: vertical**]—Reconstructed OCT images (only a 2D slice is shown in this example). Each row corresponds to an under-sampling rate: 30% (**1st row**), 50% (**2nd row**), 70% (**3th row**), and 90% (**4th row**). The first column represents the initial OCT image, the second column the under-sampled data used for the reconstruction, and the third column, the recovered OCT images.

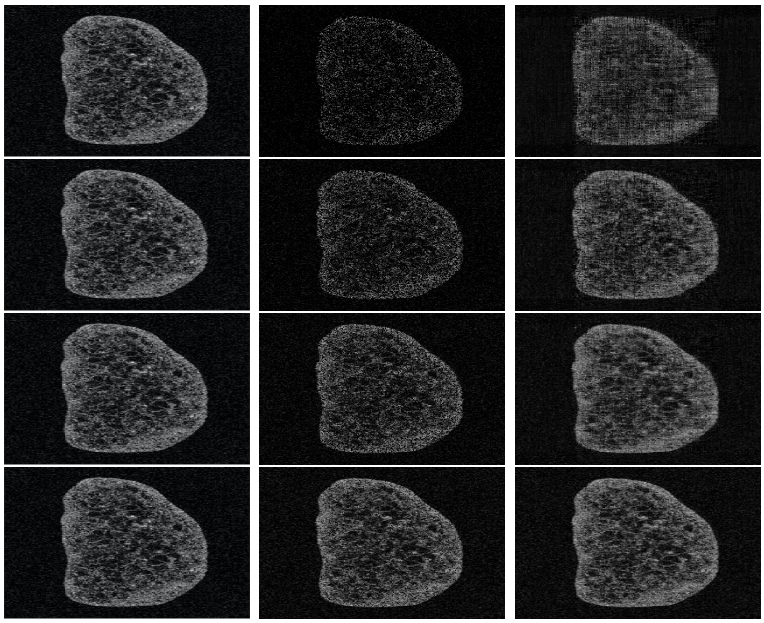

Figure 3.7 – [**sample: grape, mask: oblic**]—Reconstructed OCT images (only a 2D slice is shown in this example). Each row corresponds to an under-sampling rate: 30% (**1st row**), 50% (**2nd row**), 70% (**3th row**), and 90% (**4th row**). The first column represents the initial OCT image, the second column the under-sampled data used for the reconstruction, and the third column, the recovered OCT images.

3.5.4.2 Fish Eye Sample

we also performed validation experiments using an optical biopsy of a fish eye (see the images sequence depicted in Figures 3.8 and 3.9).

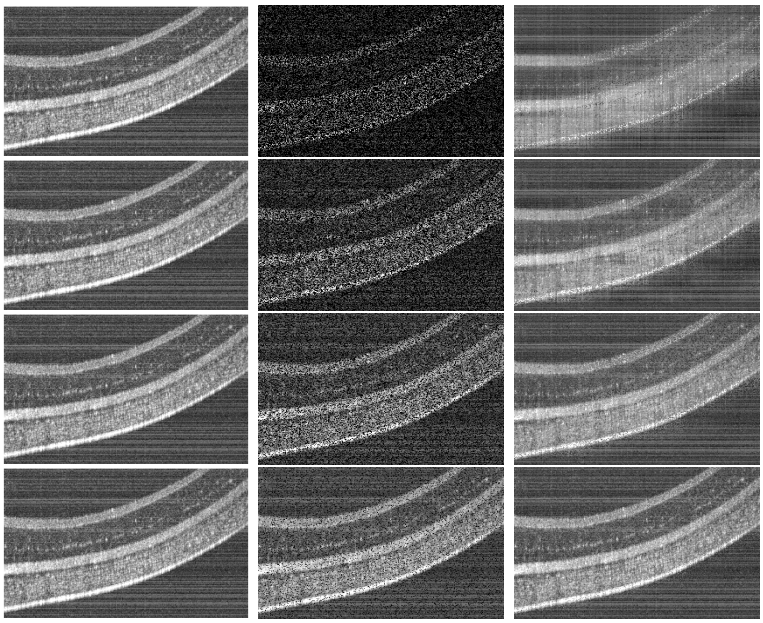

Figure 3.8 – [**sample: fish eye retina, mask: vertical**]—Reconstructed OCT images (only a 2D slice is shown in this example). Each row corresponds to an under-sampling rate: 30% (**1st row**), 50% (**2nd row**), 70% (**3th row**), and 90% (**4th row**). The first column represents the initial OCT image, the second column the under-sampled data used for the reconstruction, and the third column, the recovered OCT images.

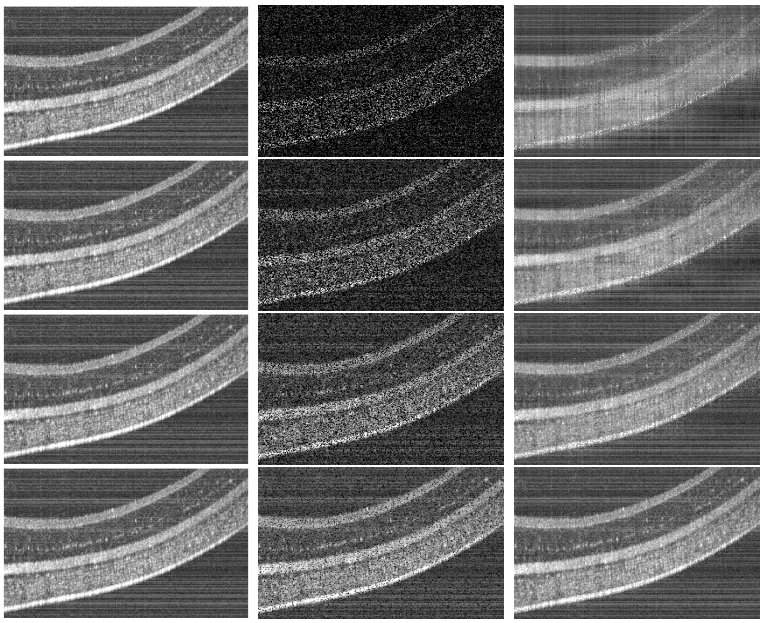

Figure 3.9 – [**sample: fish eye retina, mask: oblic**]—Reconstructed OCT images (only a 2D slice is shown in this example). Each line corresponds to an under-sampling rate: 30% (**1st line**), 50% (**2nd line**), 70% (**3th line**), and 90% (**4th line**). The first column represents the initial OCT image, the second column the under-sampled data used for the reconstruction, and the third column, the recovered OCT images.

3.5.5 The Singular Value Thresholding Algorithm

In this section, we describe the algorithm used for the computation of our estimator, namely a tubal tensor version of the fixed point algorithm of Reference [92]. This algorithm is a very simple and scalable iterative scheme which converges to the solution of (3.17). Each iteration consists of two successive steps:

- a singular value thresholding step where all tubal singular values with norm below the level 2λ are set to zero and the remaining larger singular values are being removed an offset 2λ.
- a relaxed gradient step.

In mathematical terms, the algorithms works as follows:

$$\left\{ \begin{array}{l} Z^{(l)} = \text{shrink}\left(A^{(l-1)}, \delta\lambda\right) \\ A^{(l)} = A^{(l-1)} + \delta\ \mathcal{P}_\Omega\left(\sum_{i=1}^n Y_i X_i - Z^{(l)}\right) \end{array} \right)$$

In the setting of our algorithm, the Shrinkage operator operates as follows:

- compute the circular Fourier transform $\mathcal{F}(A^{(l-1)})$ of the tubal components of $A^{(l-1)}$,

- compute the SVD of all the slices of $\mathcal{F}(A^{(l-1)})$ and forms the tubal singular values,
- sets to zero the tubal singular values whose ℓ_2-norm lies below the level $2\delta\lambda$ and shrink the other by $2\delta\lambda$,
- recompose the spectrally thresholded matrix and take the inverse Fourier transform of the tubal components.

On the other hand, $\mathcal{P}_\Omega$ is the operator that assigns to the entries indexed by Ω the observed values and leaves the other values unchanged.

The convergence analysis of [92] directly applies to our setting.

3.5.6 Analysis of the Numerical Results

In order to quantitatively assess the obtained results using different numerical validation scenarios and OCT images, we implemented two images similarity scores extensively employed in the image processing community. In particular, we use

- the Peak Signal Noise Ratio (PSNR) computed as follows

$$PSNR = 10\log_{10}\left(\frac{d^2}{MSE}\right) \tag{3.45}$$

where d is the maximal pixel value in the initial OCT image and the MSE (mean-squared error) is obtained by

$$MSE = \sum_{i=1}^{n}\sum_{j=1}^{m}\Big(I_o(i,j) - I_r(i,j)\Big)^2 \tag{3.46}$$

with I_o and I_r represent an initial 2D OCT slice (selected from the OCT volume) and the recovered one, respectively.

- The second image similarity score consists of the Structural Similarity Index (SIMM) which allows measuring the degree of similarity between two images. It is based on the computation of three values namely the luminance l, the contrast c and the structural aspect s. It is given by

$$SSIM = \Big(s\big(I_r, I_o)\Big)\Big(l\big(I_r, I_o\Big)\Big(c\big(I_r, I_o\big)\Big) \tag{3.47}$$

where,

$$l = \frac{2\mu_{I_r}\mu_{I_o} + C_1}{\mu_{I_r}^2 + \mu_{I_o}^2 + C_1}, \tag{3.48}$$

$$c = \frac{2\sigma_{I_r}\sigma_{I_o} + C_2}{\sigma_{I_r}^2 + \sigma_{I_o}^2 + C_2}, \tag{3.49}$$

$$s = \frac{2\sigma_{I_r,I_o} + C_3}{\sigma_{I_r}\sigma_{I_o} + C_3}, \tag{3.50}$$

with μ_{I_r}, μ_{I_o}, σ_{I_r}, σ_{I_o}, and μ_{I_r,I_o} are the local means, standard deviations, and cross-covariance for images I_r, I_o. The variables C_1, C_2, C_3 are used to stabilize the division with weak denominator.

3.5.6.1 Illustrating the Rôle of the Nuclear Norm Penalisation

In order to understand the rôle of the nuclear norm penalisation in the reconstruction, we have performed several experiments with different values of the hyper-parameter λ which are reported in Figure 3.10. This figure shows the performance of the estimator as a function of the ratio between the largest singular value and the smallest selected singular value. This ratio is an implicit function of λ, which is more intuitive than λ for interpreting the results. The smaller the ratio, the smaller the number of singular values incorporated in the estimation. The corresponding values are reported in Table 3.1.

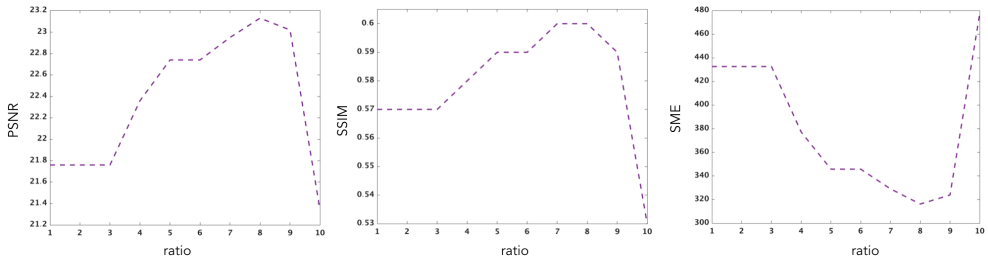

Figure 3.10 – Evolution of the PSNR, SSIM and SME performance criteria as a function of the ratio between the largest singular value and the smallest singular value selected by the penalised estimator.

Table 3.1 – SME, PSNR, SSIM for various values of the ratio.

Ratio	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
SME	432.65	432.65	432.65	377.12	345.76	345.75	329.06	316.23	323.87	477.43
PSNR	21.76	21.76	21.76	22.36	22.74	22.74	22.95	23.13	23.02	21.34
SSIM	0.57	0.57	0.57	0.58	0.59	0.59	0.60	0.60	0.59	0.53

The results of these experiments show that different weights for the nuclear norm penalisation imply different erros. In these experiments, one sees that the SME is optimised at a ration equal to 8 and the PSNR is maximised at 8 as well. The SSIM is maximum for values of the ratio equal to 7 and 8. Estimation which does not account for the intrinsic complexity of the object to recover will clearly fail to reconstruct the 3D images properly.

3.5.6.2 Performance Results

Tables 3.2 and 3.3 summarise the numerical values of MSE, PSNR, SSIM computed for each test, that is, using different undersampling rates and masks, for our two different test samples (grape or fish eye retina). The parameter λ was chosen using the simple and efficient method proposed in [93]. As expected the error decreases as a function of the percentage of observed pixels. The results also show that the estimator is not very sensitive to the orientation of the mask. There seems to be a phase transition after the 70% level, above which the reconstruction accuracy is suddenly improved in terms of PSNR and SSIM, but the method still works satisfactorily at smaller sampling rates.

Table 3.2 – **[sample: grape]:** Numerical values of the different image similarities (between initial image and reconstructed one): MSE, PSNR and SSIM.

Sample: Part of a Grape				
under-sampling rates	30%	50%	70%	90%
vertical masks				
MSE	536.68	316.23	185.83	62.34
PSNR	20.83	23.13	25.43	30.18
SSIM	0.39	0.59	0.77	0.92
oblic masks				
MSE	725.20	431.91	171.52	54.85
PSNR	19.52	21.77	25.78	30.73
SSIM	0.34	0.55	0.78	0.93

Table 3.3 – **[sample: fish eye retina]:** Numerical values of the different image similarities (between initial image and recontructed one): MSE, PSNR and SSIM.

Sample: Fish Eye Retina				
under-sampling rates	30%	50%	70%	90%
vertical masks				
MSE	829.96	570.81	333.13	109.65
PSNR	18.94	20.56	22.50	27.73
SSIM	000.51	0.65	0.78	0.91
oblic masks				
MSE	828.56	575.90	344.12	99.13
PSNR	18.94	20.52	22.76	28.16
SSIM	0.51	0.65	0.77	0.92

3.6 Conclusions and Perspectives

In this paper, we studied tensor completion problems based on the framework proposed by Kilmer et al. [38]. We provided some theoretical analysis of the nuclear norm penalised estimator. These theoretical results are validated numerically using realistic OCT data (volumes) of biological samples. These encouraging results with real datasets demonstrate the relevance of the low rank assumption for practical applications. Further research will be undertaken in devising fast algorithms and incorporating other penalties such as, e.g., based on sparsity of the shearlet transform [94].

3.7 Some Technical Results

3.7.1 Calculation of ρ and a value of λ such that $\|M\|_{S_\infty} \leq \lambda$ with high probability

3.7.1.1 Computation of ρ

Using the fact $n = m \times n_3$ where m is the number of pixels considered, we have with $i = \{j_1, j_2, j_3\}$

$$\begin{aligned}
\|A\|^2_{L_2(\Pi)} &= \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\left[\langle A, X_i\rangle^2\right] \\
&= \frac{1}{mn_3}\sum_{j_3=1}^{n_3}\sum_{j_1=1}^{n_1}\sum_{j_2=1}^{n_2}\langle A, X_{j_1,j_2,j_3}\rangle^2 \frac{1}{n_1 n_2} \\
&= \frac{1}{mn_3}\sum_{j_1=1}^{n_1}\sum_{j_2=1}^{n_2}\frac{1}{n_1 n_2}\sum_{j_3=1}^{n_3}\hat{A}^2_{j_1,j_2,j_3} \\
&= \frac{1}{mn_3}\sum_{j_1=1}^{n_1}\sum_{j_2=1}^{n_2}\frac{1}{n_1 n_2 n_3^{1/2}}\sum_{j_3=1}^{n_3}A^2_{j_1,j_2,j_3} \\
&= \frac{1}{mn_1 n_2 n_3^{3/2}}\|A\|^2_F.
\end{aligned}$$

Thus we can set

$$\rho = \sqrt{mn_1 n_2 n_3^{3/2}}. \tag{3.51}$$

3.7.1.2 Control of the Stochastic Error $\|M\|_{S_\infty}$

In this part, we will show that a possible value for the coefficient λ can be taken as

$$\lambda = 2n_3\|\mathcal{F}(A_0)\|_{S_\infty}\max\left\{\sqrt{\frac{t+\log(m)}{nn_1n_2}}; \left(1+\frac{1}{\sqrt{n_1n_2}}\right)\frac{t+\log(m)}{n}\right\}. \tag{3.52}$$

In order to prove this bound, we will need a large deviation inequality given by the following result

Proposition 3.7.1 (Theorem 1.6 in [95]) *Let $Z_1, \ldots, Z_n$ be independant random variables with dimension $n_1 \times n_2$ that satisfy $\mathbb{E}[Z_i] = 0$ and $\|Z_i\|_{S_\infty} \leqslant U$ almost surely for some constant U and all $i = 1, \ldots, n$. We define*

$$\sigma_Z = \max\left\{\left\|\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left[Z_{ij}Z_{ij}^\top\right]\right\|^{1/2}_{S_\infty}, \left\|\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left[Z_{ij}^\top Z_{ij}\right]\right\|^{1/2}_{S_\infty}\right\}$$

Then, for all $t > 0$, with probability at least $1 - e^{-t}$, we have

$$\left\| \frac{1}{n} \sum_{i=1}^{n} Z_i \right\|_{S_\infty} \leqslant 2 \max \left\{ \sigma_Z \sqrt{\frac{t + \log m}{n}}, U \frac{t + \log m}{n} \right\}$$

where $m = n_1 + n_2$.

The next lemma gives a bound of the stochastic error for tensor completion.

Lemma 3.7.2 *Let $X^{(i)}$ be i.i.d uniformly distributed on $\mathcal{X}$. Then for any $t > 0$ with probability at least $1 - \exp(-t)$, we have*

$$\|M\|_{S_\infty} \leqslant 2n_3 \|\mathcal{F}(A_0)\|_{S_\infty} \max \left\{ \sqrt{\frac{t + \log(m)}{nn_1n_2}}; \left(1 + \frac{1}{\sqrt{n_1n_2}}\right) \frac{t + \log(m)}{n} \right\}. \tag{3.53}$$

In order to prove this lemma, we will need the following lemma

Lemma 3.7.3 *We draw $X_i^{(j)}$ with a uniform random position on $\{1, \ldots, n_1\} \times \{1, \ldots, n_2\}$ with null entries everywhere except one input equals 1 at the position (k, l). Observe that*

$$\|\mathcal{F}(X_i^{(j)})\|_{S_\infty} = 1, \quad \left\| \mathbb{E}\left[\mathcal{F}(X_i^{(j)})\right] \right\|_{S_\infty} = \sqrt{\frac{1}{n_1n_2}} \tag{3.54}$$

and

$$\sigma_Z = \max \left\{ \left\| \mathbb{E}\left[ZZ^\top\right] \right\|_{S_\infty}^{1/2}, \left\| \mathbb{E}\left[Z^\top Z\right] \right\|_{S_\infty}^{1/2} \right\} \leqslant \frac{\|\mathcal{F}(A_0)\|_{S_\infty}}{\sqrt{n_1n_2}}. \tag{3.55}$$

Proof: Let $X_i^{(j)}$ of the form

$$X_i^{(j)} = \begin{bmatrix} 0 & 0 & \ldots & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \ldots & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & \ldots & 0 \end{bmatrix}. \tag{3.56}$$

Determine its Fourier transform and spectral norm of expectation of $\mathcal{F}(X_i^{(j)})$. Thus,

$$\mathcal{F}(X_i^{(j)}) = \begin{bmatrix} 0 & 0 & \ldots & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \ldots & 0 & e^{-2i\pi j \frac{k_3}{n_3}} & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & \ldots & 0 \end{bmatrix}. \tag{3.57}$$

Moreover,

$$\begin{aligned}\mathbb{E}\left[\mathcal{F}(X_i^{(j)})\right] &= \frac{1}{n_1 n_2}\begin{bmatrix} e^{-2i\pi j\frac{1}{n_3}} & e^{-2i\pi j\frac{2}{n_3}} & \dots & e^{-2i\pi j\frac{n_2}{n_3}} \\ e^{-2i\pi j\frac{n_2+1}{n_3}} & \dots & \dots & e^{-2i\pi j\frac{n_2+n_2}{n_3}} \\ e^{-2i\pi j\frac{n_2+n_2+1}{n_3}} & \dots & \dots & e^{-2i\pi j\frac{n_2+n_2+n_2}{n_3}} \\ \vdots & \vdots & \vdots & \vdots \\ e^{-2i\pi j\frac{(n_1-2)\times n_2+1}{n_3}} & \dots & \dots & e^{-2i\pi j\frac{(n_1-1)\times n_2}{n_3}} \\ e^{-2i\pi j\frac{(n_1-1)\times n_2+1}{n_3}} & \dots & \dots & e^{-2i\pi j\frac{n_1\times n_2}{n_3}} \end{bmatrix} \\ &= DLN,\end{aligned}$$

with

$$D = \begin{bmatrix} 1 & & & \\ & \left(e^{\frac{-2i\pi j}{n_3}}\right)^{n_2} & & \\ & & \ddots & 0 \\ & 0 & & \\ & & & \left(e^{\frac{-2i\pi j}{n_3}}\right)^{(n_1-1)n_2} \end{bmatrix}, \quad L = \begin{bmatrix} 1 \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ 1 \end{bmatrix}$$

and

$$N = \begin{bmatrix} e^{\frac{-2i\pi j}{n_3}} & \left(e^{\frac{-2i\pi j}{n_3}}\right)^2 & \dots & \left(e^{\frac{-2i\pi j}{n_3}}\right)^{n_2} \end{bmatrix}. \tag{3.58}$$

So, we have $\left\|\mathcal{F}(X_i^{(j)})\right\|_{S_\infty} = 1$, and

$$\begin{aligned}\left\|\mathbb{E}\left[\mathcal{F}(X_i^{(j)})\right]\right\|_{S_\infty} &= \frac{1}{n_1 n_2}\left\|DLN\right\| \\ &\leqslant \frac{1}{n_1 n_2}\left\|D\right\|_{S_\infty}\left\|L\right\|_2\left\|N\right\|_2 \\ &\leqslant \frac{1}{\sqrt{n_1 n_2}}\end{aligned}$$

□

Proof of Lemma 3.7.2. Consider the tensor completion under uniform sampling at random with Gaussian error. Recall that in this case we assume that the pairs (X_i, Y_i) are

i.i.d. Using the fact $Y_i = \langle X_i, A_0\rangle + \varepsilon_i$ and $\mathbb{E}\left[\varepsilon_i X_i | X_i\right] = 0$, we have

$$\begin{aligned}
\|M\|_{S_\infty} &= \left\| \frac{1}{n}\sum_{i=1}^{n} \left(Y_i X_i - \mathbb{E}\left[Y_i X_i\right]\right) \right\|_{S_\infty} \\
&= \left\| \frac{1}{n}\sum_{i=1}^{n} \left(\langle X_i, A_0\rangle + \varepsilon_i\right) X_i - \mathbb{E}\left[\langle X_i, A_0\rangle X_i\right] \right\|_{S_\infty} \\
&\leqslant \left\| \frac{1}{n}\sum_{i=1}^{n} \langle X_i, A_0\rangle X_i - \mathbb{E}\left[\langle X_i, A_0\rangle X_i\right] \right\|_{S_\infty} + \left\| \frac{1}{n}\sum_{i=1}^{n} \varepsilon_i X_i \right\|_{S_\infty} \\
&= \Lambda_1 + \Lambda_2.
\end{aligned}$$

In the following, we treat the terms Λ_1 and Λ_2 separately in the lemma and proposition below. Before proceeding, notice that Λ_1 is the Schatten norm of a quadratic function of $X_1, \ldots, X_n$. Furthermore, $\langle X_i, A_0\rangle$ is the Fourier Transform of the tube (i_1, i_2) for the frequency k_3. Using the Property 4.3.15, we have

$$\begin{aligned}
\Lambda_1 &= \left\| \frac{1}{n}\sum_{i=1}^{n} \mathcal{F}(X_i)\langle \mathcal{F}(X_i), \mathcal{F}(A_0)\rangle - \mathbb{E}\left[\langle \mathcal{F}(X_i), \mathcal{F}(A_0)\rangle \mathcal{F}(X_i)\right] \right\|_{S_\infty} \\
&= \left\| \sum_{k_3=1}^{n_3} \left(\frac{1}{n}\sum_{i=1}^{n} \mathcal{F}(X_i)^{(k_3)}\langle \mathcal{F}(X_i), \mathcal{F}(A_0)\rangle - \mathbb{E}\left[\langle \mathcal{F}(X_i), \mathcal{F}(A_0)\rangle \mathcal{F}(X_i)^{(k_3)}\right] \right) \right\|_{S_\infty} \\
&\leqslant \sum_{k_3=1}^{n_3} \left\| \frac{1}{n}\sum_{i=1}^{n} \mathcal{F}(X_i)^{(k_3)}\langle \mathcal{F}(X_i), \mathcal{F}(A_0)\rangle - \mathbb{E}\left[\langle \mathcal{F}(X_i), \mathcal{F}(A_0)\rangle \mathcal{F}(X_i)^{(k_3)}\right] \right\|_{S_\infty}.
\end{aligned}$$

Define the operator $\Gamma^{(j)}$ which takes the slice j of a tensor (i.e tensor $\mathcal{F}(X_i)$) and puts it in the same place in a zero tensor. The following proposition

Proposition 3.7.4 *Let T be a null tensor except at slice j. Therefore*

$$\|T\|_{S_\infty} = \|T^{(j)}\|_{S_\infty}. \tag{3.59}$$

Using 3.59, we have

$$\Lambda_1 \leqslant n_3 \left\| \frac{1}{n}\sum_{i=1}^{n} \mathcal{F}(X_i)^{(j)}\langle \mathcal{F}(A_0), \mathcal{F}(X_i)\rangle - \mathbb{E}\left[\langle \mathcal{F}(A_0), \mathcal{F}(X_i)\rangle \mathcal{F}(X_i)^{(j)}\right] \right\|_{S_\infty}.$$

Lemma 3.7.5 *Let $X_i^{(j)}$ be i.i.d a uniform random position on $\{1, \ldots, n_1\} \times \{1, \ldots, n_2\}$. Then, for all $t > 0$, with probability at least $1 - e^{-t}$, we have*

$$\Lambda_1 \leqslant 2\max\left\{ \left\|\mathcal{F}(A_0)\right\|_{S_\infty} \sqrt{\frac{t + \log(m)}{n_1 n_2}} n, \|\mathcal{F}(A_0)\|_{S_\infty} \left(1 + \sqrt{\frac{1}{n_1 n_2}}\right) \frac{t + \log(m)}{n} \right\}. \tag{3.60}$$

Proof: To prove this lemma, we apply the Proposition 3.7.1 for the random variables $Z_{ij} = \mathcal{F}(X_i)^{(j)}\langle \mathcal{F}(A_0), \mathcal{F}(X_i)\rangle - \mathbb{E}\left[\langle \mathcal{F}(A_0), \mathcal{F}(X_i)\rangle \mathcal{F}(X_i)^{(j)}\right]$. Moreover, using the facts

- $\mathbb{E}\left[Z_{ij}\right] = \mathbb{E}\left[\mathcal{F}(X_i)^{(j)}\langle \mathcal{F}(A_0), \mathcal{F}(X_i)\rangle - \mathbb{E}\left[\langle \mathcal{F}(A_0), \mathcal{F}(X_i)\rangle \mathcal{F}(X_i)^{(j)}\right]\right] = 0.$
- $$\begin{aligned}\|Z_{ij}\|_{S_\infty} &= \mathcal{F}(X_i)^{(j)}\langle \mathcal{F}(A_0), \mathcal{F}(X_i)\rangle - \mathbb{E}\left[\langle \mathcal{F}(A_0), \mathcal{F}(X_i)\rangle \mathcal{F}(X_i)^{(j)}\right] \|_{S_\infty} \\ &\leqslant \|\mathcal{F}(X_i)^{(j)}\|_{S_\infty} |\langle \mathcal{F}(A_0), \mathcal{F}(X_i)\rangle| \\ &\quad + |\langle \mathcal{F}(A_0), \mathcal{F}(X_i)\rangle| \|\mathbb{E}\left[\mathcal{F}(X_i)^{(j)}\right] \|_{S_\infty} \\ &\leqslant \|\mathcal{F}(X_i)^{(j)}\|_{S_\infty} \|\mathcal{F}(A_0)\|_{S_\infty} \|\mathcal{F}(X_i)^{(j)}\|_{S_\infty} \\ &\quad + \|\mathcal{F}(A_0)\|_{S_\infty} \|\mathcal{F}(X_i)^{(j)}\|_{S_\infty} \|\mathbb{E}\left[\mathcal{F}(X_i)^{(j)}\right] \|_{S_\infty} \\ &\leqslant \|\mathcal{F}(A_0)\|_{S_\infty} \left(1 + \sqrt{\frac{1}{n_1 n_2}}\right)\end{aligned}$$
- $$\begin{aligned}\mathbb{E}\left[ZZ^\top\right] &= \mathbb{E}\left[\langle \mathcal{F}(A_0), \mathcal{F}(X_i)\rangle^2 \mathcal{F}(X_i)^{(i)} \mathcal{F}(X_i)^{(i)^\top}\right] \\ &\quad - \mathbb{E}\left[\langle \mathcal{F}(A_0), \mathcal{F}(X_i)\rangle \mathcal{F}(X_i)^{(i)}\right] \mathbb{E}\left[\langle \mathcal{F}(A_0), \mathcal{F}(X_i)\rangle \mathcal{F}(X_i)^{(i)}\right]^\top\end{aligned}$$

  and

  $$\begin{aligned}\mathbb{E}\left[Z^\top Z\right] &= \mathbb{E}\left[\langle \mathcal{F}(A_0), \mathcal{F}(X_i)\rangle^2 \mathcal{F}(X_i)^{(i)^\top} \mathcal{F}(X_i)^{(i)}\right] \\ &\quad - \mathbb{E}\left[\langle \mathcal{F}(A_0), \mathcal{F}(X_i)\rangle \mathcal{F}(X_i)^{(i)}\right]^\top \mathbb{E}\left[\langle \mathcal{F}(A_0), \mathcal{F}(X_i)\rangle \mathcal{F}(X_i)^{(i)}\right].\end{aligned}$$

  Using the duality trace and Jensen's inequality, we have

  $$\begin{aligned}\sigma_Z^2 &\leqslant \max\left\{\left\|\mathbb{E}\left[\langle \mathcal{F}(A_0), \mathcal{F}(X_i)\rangle^2 \mathcal{F}(X_i)^{(i)} \mathcal{F}(X_i)^{(i)^\top}\right] \right\|_{S_\infty}, \right. \\ &\quad \left. \left\|\mathbb{E}\left[\langle \mathcal{F}(A_0), \mathcal{F}(X_i)\rangle^2 \mathcal{F}(X_i)^{(i)^\top} \mathcal{F}(X_i)^{(i)}\right] \right\|_{S_\infty}\right\} \\ &\leqslant \left\|\mathcal{F}(A_0)\right\|_{S_\infty}^2 \left\|\mathcal{F}(X_i)\right\|_{S_\infty}^2 \mathbb{E}\left[\left\|\mathcal{F}(X_i)\right\|_{S_\infty}^2\right] \\ &\leqslant \frac{\left\|\mathcal{F}(A_0)\right\|_{S_\infty}^2}{n_1 n_2}.\end{aligned}$$

Thus, (3.60) follows from (3.7.1). □

Now, we study the bound of Λ_2. For this purpose, we use the proposition below is an immediate consequence of the matrix Gaussian's inequality of Theorem 1.6 of [95].

Proposition 3.7.6 *Let $Z_1, \ldots, Z_n$ be independent random variables with dimensions $n_1 \times n_2$, and $\{\varepsilon_i\}$ be a finite sequence of independent standart normal. Define*

$$\sigma := \max\left\{\left\|Z_i Z_i^\top\right\|_{S_\infty}, \left\|Z_i^\top Z_i\right\|_{S_\infty}\right\}.$$

Then, for all $t \geqslant 0$,

$$\left\|\frac{1}{n}\sum_{i=1}^{n}\varepsilon_i Z_i\right\|_{S_\infty} \leqslant 2\sigma\sqrt{\frac{t+\log(m)}{n}}$$

where $m = n_1 + n_2$.

Using this fact and (3.59), we have

$$\begin{aligned}
\Lambda_2 &= \left\|\frac{1}{n}\sum_{i=1}^{n}\varepsilon_i X_i\right\|_{S_\infty} \\
&= \left\|\frac{1}{n}\sum_{i=1}^{n}\sum_{k=1}^{n_3}\mathcal{F}(\varepsilon_i)^{(k)}\mathcal{F}(X_i)^{(k)}\right\|_{S_\infty} \\
&\leqslant n_3\left\|\frac{1}{n}\sum_{i=1}^{n}\mathcal{F}(\varepsilon_i)^{(j)}\mathcal{F}(X_i)^{(j)}\right\|_{S_\infty} \\
&\leqslant 2\,n_3\,\sigma\sqrt{\frac{t+\log(m)}{n}}
\end{aligned}$$

and the proof is completed. □

Bibliography

[1] Candès, E.J.; Recht, B. *Exact matrix completion via convex optimization*, Found. Comput. Math, 2009, **9**, no. 717–772.

[2] Kilmer, M.E.; Martin, C.D. *Factorization strategies for third-order tensors*, Linear Algebra Its Appl, 2011, **435**, no. 641–658.

[3] Kilmer, M.E.; Braman, K.; Hao, N.; Hoover, R.C. *Third-order tensors as operators on matrices: A theoretical and computational framework with applications in imaging*, Siam J. Matrix Anal. Appl, 2013,**34**, no. 148–172.

[4] Koltchinskii, V.; Lounici, K.; Tsybakov, A.B. *Nuclear-norm penalization and optimal rates for noisy low-rank matrix completion*, Ann. Stat, 2011, **39**, no. 2302–2329.

[5] Candès, E.J. *Mathematics of sparsity (and a few other things)*, In Proceedings of the International Congress of Mathematicians, Seoul, Korea, 13–21 August 2014.

[6] Candès, E.J.; Romberg, J.; Tao, T. *Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information*, IEEE Trans. Inf. Theory, 2006, **52**, no. 489–509.

[7] Recht, B.; Fazel, M.; Parrilo, P.A. *Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization*, SIAM Rev, 2010, **52**, no. 471–501.

[8] Hackbusch, W. *Tensor Spaces and Numerical Tensor Calculus*; Springer Science & Business Media: Berlin, Germany, 2012; **42**.

[9] Landsberg, J.M. *Tensors: Geometry and Applications*, American Mathematical Society: Providence, RI, USA, 2012; **128**.

[10] Sacchi, M.D.; Gao, J.; Stanton, A.; Cheng, J. *Tensor Factorization and its Application to Multidimensional Seismic Data Recovery*, In SEG Technical Program Expanded Abstracts, Society of Exploration Geophysicists: Houston, TX, USA, 2015.

[11] Yuan, M.; Zhang, C.H. *On tensor completion via nuclear norm minimization*, Found. Comput. Math, 2016, **16**, no. 1031–1068.

[12] Signoretto, M.; Van de Plas, R.; De Moor, B.; Suykens, J.A. *Tensor versus matrix completion: a comparison with application to spectral data*, Signal Process. Lett. IEEE, 2011, **18**, no. 403–406.

[13] Candes, E.J.; Plan, Y. *Matrix completion with noise*, Proc. IEEE, 2010, **98**, no. 925–936.

[14] Anandkumar, A.; Ge, R.; Hsu, D.; Kakade, S.M.; Telgarsky, M. *Tensor decompositions for learning latent variable models*, J. Mach. Learn. Res, 2014, **15**, no. 2773–2832.

[15] Signoretto, M.; Dinh, Q.T.; De Lathauwer, L.; Suykens, J.A. *Learning with tensors: a framework based on convex optimization and spectral regularization*, Mach. Learn, 2014, **94**, no. 303–351.

[16] So, A.M.C.; Ye, Y. *Theory of semidefinite programming for sensor network localization*, Math. Program, 2007, **109**, no. 367–384.

[17] Eriksson, B.; Balzano, L.; Nowak, R. *High-rank matrix completion and subspace clustering with missing data*, arXiv, 2011, arXiv:1112.5629.

[18] Singer, A.; Cucuringu, M. *Uniqueness of low-rank matrix completion by rigidity theory*, SIAM J. Matrix Anal. Appl., 2010, **31**, no. 1621–1641.

[19] Fazel, M. *Matrix Rank Minimization with Applications*, Ph.D. Thesis, Stanford University, Stanford, CA, USA, 2002.

[20] Gross, D. *Recovering low-rank matrices from few coefficients in any basis*, Inf. Theory IEEE Trans., 2011, **57**, no. 1548–1566.

[21] Recht, B. *A simpler approach to matrix completion*, J. Mach. Learn. Res., 2011, **12**, no. 3413–3430.

[22] Candès, E.J.; Tao, T. *The power of convex relaxation: Near-optimal matrix completion*, Inf. Theory IEEE Trans, 2010, **56**,no. 2053–2080.

[23] Kolda, T.G.; Bader, B.W. *Tensor decompositions and applications*, SIAM Rev., 2009, **51**,no. 455–500.

[24] McCullagh, P. *Tensor Methods in Statistics*, Chapman and Hall: London, UK, 1987; **161**.

[25] Rogers, M.; Li, L.; Russell, S.J. *Multilinear Dynamical Systems for Tensor Time Series*, Adv. Neural Inf. Process. Syst., 2013, **26**, no. 2634–2642

[26] Ge, R.; Huang, Q.; Kakade, S.M. *Learning mixtures of gaussians in high dimensions*, arXiv, 2015, arXiv:1503.00424.

[27] Mossel, E.; Roch, S. *Learning nonsingular phylogenies and hidden Markov models*, In Proceedings of the Thirty-Seventh Annual ACM Symposium on Theory of Computing, Baltimore, MD, USA,**22–24** May, 2005; no. 366–375.

[28] Jennrich, R. *A generalization of the multidimensional scaling model of Carroll and Chang*, UCLA Work. Pap. Phon., 1972, **22**, no. 45–47.

[29] Kroonenberg, P.M.; De Leeuw, J. *Principal component analysis of three-mode data by means of alternating least squares algorithms*, Psychometrika, 1980, **45**,no. 69–97.

[30] Bhaskara, A.; Charikar, M.; Moitra, A.; Vijayaraghavan, A. *Smoothed analysis of tensor decompositions*, k In Proceedings of the 46th Annual ACM Symposium on Theory of Computing, New York, NY, USA, 2014; pp. 594–603.

[31] De Lathauwer, L.; De Moor, B.; Vandewalle, J. *A multilinear singular value decomposition*, SIAM J. Matrix Anal. Appl., 2000, **21**,no. 1253–1278.

[32] Kolda, T.G. *Symmetric Orthogonal Tensor Decomposition is Trivial*, arXiv, 2015, arXiv:1503.01375.

[33] Robeva, E.; Seigal, A. *Singular Vectors of Orthogonally Decomposable Tensors*, arXiv, 2016, arXiv:1603.09004.

[34] Hao, N.; Kilmer, M.E.; Braman, K.; Hoover, R.C. *Facial recognition using tensor-tensor decompositions*, SIAM J. Imaging Sci., 2013, **6**,no. 437–463.

[35] Semerci, O.; Hao, N.; Kilmer, M.E.; Miller, E.L. *Tensor-based formulation and nuclear norm regularization for multienergy computed tomography*, Image Process. IEEE Trans., 2014, **23**,no. 1678–1693.

[36] Zhang, Z.; Aeron, S. *Exact tensor completion using t-svd*, IEEE Trans. Signal Process, 2017, **65**,no. 1511–1526.

[37] Pothier, J.; Girson, J.; Aeron, S. *An algorithm for online tensor prediction*, arXiv, 2015, arXiv:1507.07974.

[38] Tibshirani, R. *Regression shrinkage and selection via the lasso*, J. R. Stat. Soc. Ser. (Methodol.), 1996, **58**, no. 267–288.

[39] Bickel, P.J.; Ritov, Y.; Tsybakov, A.B. *Simultaneous analysis of Lasso and Dantzig selector*, Ann. Stat., 2009, **37**, no. 1705–1732.

[40] Candès, E.J.; Plan, Y.; others. *Near-ideal model selection by l1 minimization*, Ann. Stat., 2009, **37**, no. 2145–2177.

[41] Zou, H. *The adaptive lasso and its oracle properties*, J. Am. Stat. Assoc., 2006, **101**, no. 1418–1429.

[42] Zou, H.; Hastie, T. *Regularization and variable selection via the elastic net*, J. R. Stat. Soc. Ser. (Stat. Methodol.), 2005, **67**,no. 301–320.

[43] Gaïffas, S.; Lecué, G. *Weighted algorithms for compressed sensing and matrix completion*, arXiv, 2011, arXiv:1107.1638.

[44] Xu, Y. *On Higher-order Singular Value Decomposition from Incomplete Data*, arXiv, 2014, arXiv:1411.4324.

[45] Raskutti, G.; Yuan, M. *Convex Regularization for High-Dimensional Tensor Regression*, arXiv, 2015, arXiv:1512.01215.

[46] Barak, B.; Moitra, A. *Noisy Tensor Completion via the Sum-of-Squares Hierarchy*, arXiv, 2015, arXiv:1501.06521.

[47] Gandy, S.; Recht, B.; Yamada, I. *Tensor completion and low-n-rank tensor recovery via convex optimization*, Inverse Probl., 2011, **27**, no. 025010.

[48] Mu, C.; Huang, B.; Wright, J.; Goldfarb, D. *Square Deal: Lower Bounds and Improved Convex Relaxations for Tensor Recovery*, In Proceedings of the International Conference on Machine Learning, Beijing, China, 21–26 June 2014.

[49] Amelunxen, D.; Lotz, M.; McCoy, M.B.; Tropp, J.A. *Living on the edge: Phase transitions in convex programs with random data* Inf. Inference, 2014, **3**, pp.224–294.

[50] Lecué, G.; Mendelson, S. *Regularization and the small-ball method I: sparse recovery*, arXiv, 2016, arXiv:1601.05584.

[51] Watson, G.A. *Characterization of the subdifferential of some matrix norms*, Linear Algebra Its Appl., 1992, **170**, no. 33–45.

[52] Lewis, A.S. *The convex analysis of unitarily invariant matrix functions*, J. Convex Anal., 1995, **2**, no. 173–183.

[53] Hu, S. *Relations of the nuclear norm of a tensor and its matrix flattenings*, Linear Algebra Its Appl., 2015, **478**, no. 188–199.

[54] Chrétien, S.; Wei, T. *Von Neumann's inequality for tensors*, arXiv, 2015, arXiv:1502.01616

[55] Chrétien, S.; Wei, T. *On the subdifferential of symmetric convex functions of the spectrum for symmetric and orthogonally decomposable tensors*, Linear Algebra Its Appl., 2018, **542**, no. 84–100.

[56] Heinze H.Bauschke, P.L. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Canada Mathematical Society, Ottawa, ON, Canada, 2011.

[57] Borwein, J.; Lewis, A.S. *Convex Analysis and Nonlinear Optimization: Theory and Examples*, Springer Science & Business Media: Berlin, Germany, 2010.

[58] Jean-Pierre Aubin, I.E. *Applied Nonlinear Analysis*, Dover Publications: Mineola, NY, USA, 2006, **518**.

[59] Ma, S.; Goldfarb, D.; Chen, L. *Fixed point and Bregman iterative methods for matrix rank minimization.*, Math. Program., 2011, **128**, no. 321–353.

[60] Chrétien, S.; Lohvithee, M.; Sun, W.; Soleimani, M. *Efficient hyper-parameter selection in Total Variation-penalised XCT reconstruction using Freund and Shapire's Hedge approach*, Mathematics, 2020, **8**,no. 493.

[61] Guo, W.; Qin, J.; Yin, W. *A new detail-preserving regularization scheme* SIAM J. Imaging Sci., 2014, **7**, no. 1309–1334.

[62] Tropp, J.A. *User-Frindly tail bounds for sums of random matrices*, arXiv, 2010, arXiv:1004.4389v7.

Chapter 4

Low tubal rank tensor recovery using the Bürer-Monteiro factorisation approach. Application to optical coherence tomography

4.1 Introduction

Tensor completion has many applications in various fields of engineering and data science, such as Markov Field analysis, signal and image processing, etc [96], seismic data reconstruction [97], health data analytics [98], compression of hyperspectral images [99], 3D image and video reconstruction from subsampled measurements [101], [100]. From the mathematical viewpoint, tensor completion is currently a very active research trend as well [58], [102], [103], etc. Tensor completion relies on the often-encountered property that the sought after tensor is low rank, which is the case in many different applications; see [104], [49], [96], etc. Previous work on *tensor completion* based on partial sampling includes [80], [105], [106], [107], [108], [109], [18], [110], etc.

One of the main ingredients of both theoretical analysis and practical implementations of tensor completion is the singular value decomposition (SVD), which has been recently extended from the matrix to the tensor setting in various directions depending on the properties one intends to preserve [66], [58], [96], [21], etc. One particular approach relies on expressing the original tensor as a sum of rank-1 tensors [58], a construct which loses the orthogonality property of the singular vectors as compared with the matrix SVD. Conversely, the multilinear SVD of [66] enforces this orthogonality property of singular vectors at the price of losing the standard notion of scalar singular values and replacing it with the notion of core tensors. Another trend is the recent tubal-SVD of [21], which extends the matrix notion of SVD to the higher dimensional setting by generalising matrix products from 2D arrays of scalars to 2D arrays of "tubes" and which uses a specific "tubal" product. As for [66], the singular values are no longer scalars, but become higher dimensional "tubal singular values". As a consequence of the diversity of approaches to the

construction of the tensor SVD, the notion of rank is very specific to the definition of the SVD used in the application of interest.

On the computational side, the estimation of low rank tensors has been a topic of increasing interest, due to the high dimensionality of the problem. The approach developed in [111] is an example of efficient algorithm with guaranteed complexity. Convex relaxations based on various extensions of the nuclear norm penalisation approach, originally devised in the matrix case in [42], have also been proposed lately in the tensor case; see e.g. [42], [109], [18], etc. Tighter relaxations have also been proposed and precisely studied in [105], [110]. Factorisation based methods form another group of methods which have been successfully employed in [112], however without clear mathematical underpinnings. In the case of tubal SVD-based approaches, some recent works include [113], [114] where iterative methods are implemented achieving high practical efficiency, while leaving the question of establishing the convergence rigorously to further investigation. In [115], an interesting nuclear norm penalised approach is studied and an upper bound is given in the statistical reconstruction error.

The goal of the present paper is to study the factorisation-based approach to the low tubal-rank tensor recovery problem. This approach, also known as the Bürer-Monteiro factorisation approach in the Semi-Definite Programming literature [116], [117], has proved very efficient in practice but also amenable to thorough theoretical analysis [117], [118] in the matrix setting. The main contribution of the present paper is a proof that similar results also hold for the low tubal-rank tensor completion problem, after appropriately generalising some of the main ingredients from [118]. In particular, we prove that all the local minimisers of the least-squares cost function applied to the couple resulting from the factorisation, are in fact global minimisers. The main consequence of this analysis is that one can safely run a gradient-like or alternating optimisation algorithm for the reconstruction of the low tubal-rank tensor of interest. We also illustrate our theoretical results with promising numerical experiments for a challenging OCT reconstruction problem.

We now define more precisely the mathematical problem addressed in this paper.

4.1.1 The Tensor Completion problem

All the notations and concepts about tensors are collected in 4.3.

Tensor Completion is a generalisation of Matrix Completion, an extensively studied problem in data science which went through a sudden surge of interest triggered by the Netflix context [57]. The tensor completion problem is the one of recovering a tensor $\mathcal{M}^\star$ a small number of components of which are observed uniformly at random. More precisely, given $\Omega \subset \{1, \dots, n_1\} \times \{1, \dots, n_2\}$ to be a set of indices for the observed entries, we define

for any tensor $\mathcal{M}$, the observed tensor $\mathcal{M}_\Omega$ by

$$[\mathcal{M}_\Omega]_{ijk} = \begin{cases} \mathcal{M}_{ijk} & \text{for all } k \in \{1, \dots, n_3\}, \text{ if } \Big(i, j\Big) \in \Omega \\ \\ 0 & \text{otherwise.} \end{cases}$$

The probability of selecting one entry will be denoted by $p = 1/(n_1 \times n_2)$.

This recovery problem is hopeless in the general setting, but in many areas of engineering, social, biological etc sciences, the tensor to be recovered can be assumed to be low rank, where by rank, one refers to the appropriate notion among the ones that have been devised in the literature on the mathematics of tensors [58], [96].

Under a low tensor rank assumption, one way of recovering the original $\mathcal{M}^\star \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is to solve the following optimization problem

$$\min_{\mathcal{M} \in \mathbb{R}^{n_1 \times n_2 \times n_3}} f(\mathcal{M}) = \frac{1}{2p} \|\mathcal{M} - \mathcal{M}^\star\|_\Omega^2 \quad \text{s.t.} \quad \text{rank}(\mathcal{M}) = r \tag{4.1}$$

for some $r \geq 0$, where $\|\cdot\|_\Omega$ denotes the Frobenius norm restricted to the components indexed by Ω. The main difficulty in addressing this optimisation problem resides in handling the rank constraint, even in the case of 2D tensors, i.e. matrices. In the matrix case, one efficient method which has gained a lot of interest lately is the Bürer-Monteiro factorisation approach [116] which, in the symmetric case $n_1 = n_2 = n$, takes the following form

$$\min_{\mathcal{U} \in \mathbb{R}^{n \times r \times n_3}} f(\mathcal{U}) = \frac{1}{2p} \|\mathcal{U} * \mathcal{U}^\top - \mathcal{M}^\star\|_\Omega^2. \tag{4.2}$$

Recent work by [118] showed that the Bürer Monteiro approach is able to recover low rank matrices when the observation set up is of the Compressed Sensing type. The work in [118] also showed that for matrix completion, a penalisation term is required to be added to the least squares functional in order to ensure successful recovery via the following optimisation problem:

$$\min_{\mathcal{U} \in \mathbb{R}^{n \times r \times n_3}} \frac{1}{2p} \|\mathcal{U} * \mathcal{U}^\top - \mathcal{M}^\star\|_\Omega^2 + Q(\mathcal{U}) \tag{4.3}$$

where Q will be specified later.

The goal of the present paper is to study an extension of the penalised Bürer-Monteiro least squares problem (4.3) in the setting of tensor recovery. As in [118] both the symmetric and non-symmetric settings will be studied.

4.1.2 Plan of the paper

This chapter is organized as follows. Section 4.2 introduces some notations and preliminaries of optimality conditions. In Section 4.3 we define the background of several algebraic structures of third-order tensors. Section 4.4 presents the main recovery results. In Section 4.5 we demonstrate the effectiveness of low rank tensor reconstruction for the problem of Optical Coherence Tomography. In Section 4.7 we prove our main theoretical results.

4.2 Preliminaries

In this section we provide all the preliminary technical results and notations which will be used in our analysis.

4.2.1 Notations

In this paper, we focus on real valued third-order tensors in the space $\mathbb{R}^{n_1 \times n_2 \times n_3}$. We use n_1, n_2, n_3 for tensor dimensions, x for vectors and $X \in \mathbb{R}^{n_1 \times n_2}$ for matrices. Tensors are denoted by calligraphic letters, i.e $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$. For a vector x, $\|x\|_2$ denotes its ℓ_2 norm and for a matrix X we use $\|X\|_F$ to denote its Frobenius norm. For a tensor $\mathcal{A}$, we use $\|\mathcal{A}\|_F$ to denote its Frobenius norm which is the square root of the sum of its squared components; see 4.3 for further details. Throughout the paper, we use $\mathcal{M}^\star$ to denote the original low rank solution to be recovered and we denote by $\sigma_1^\star$ its largest singular value, by $\sigma_r^\star$ its r-th singular value. The ratio $\kappa^\star = \sigma_1^\star / \sigma_r^\star$ be called the condition number.

Given a transformation $\mathcal{H}$, we use the notations $\mathcal{M} : \mathcal{H} : \mathcal{N}$ to denote the quadratic form $\langle \mathcal{M}, \mathcal{H}(\mathcal{N}) \rangle$.

4.2.2 Optimality Conditions

Suppose we are optimizing a function $f(\mathbf{x})$ with no constraints on $\mathbf{x}$. For a point $\mathbf{x}$ to be a local minimum, it must satisfy the first and second order necessary conditions, i.e. $\nabla f(\mathbf{x}) = 0$ ans $\nabla^2 f(\mathbf{x}) \succeq 0$.

If one of the two conditions is not verified, i.e., if we are not a local minimum, it is always possible to follow the gradient and reduce the value of the function. In this case, [119] defines a **strict-saddle** property, which is a quantitative version of the optimality conditions.

Definition 4.2.1 *We say f satisfies the $\Big(\theta, \gamma, \zeta\Big)$-strict saddle property if for any point $\boldsymbol{x}$ at least one of the following is true:*

1. $\|\nabla f(x)\| \geqslant \theta$.
2. $\lambda_{\min}\Big(\nabla^2 f(x)\Big) \leqslant -\gamma$.
3. *x is ζ-close to $\mathcal{X}^\star$ where $\mathcal{X}^\star$ is the set of a local minima.*

This definition intuitively says that any point at which the gradient of f is small, is either close to a local minimiser, a local maximiser or a saddle point with at least one significantly negative eigenvalue.

4.3 Background on tensors and the tubal algebra

4.3.1 Basic Notations for Third-order Tensor

In this section, we recall the framework introduced by Kilmer and Martin [33] and [21] for a very special class of tensors which is particularly adapted to our setting.

4.3.1.1 Slices and transposition

A third-order tensor are represented as A and its $\big(i,j,k\big)$th entry is denoted by A_{ijk}.

Definition 4.3.1 *The k^{th}- frontal slice of A is defined as*

$$A^{\Big(k\Big)} = A\Big(:,:,k\Big).$$

The j^{th}-transversal slice of A is defined as

$$\vec{A}^{\Big(j\Big)} = A\Big(:,j,:\Big).$$

A tubal scalar (t-scalar) is an element of $\mathbb{R}^{1\times 1\times n_3}$ and a tubal vector (t-vector) is an element of $\mathbb{R}^{n_1\times 1\times n_3}$

Definition 4.3.2 *(Tensor transpose) The conjugate transpose of a tensor $A \in \mathbb{R}^{n_1\times n_2\times n_3}$ tensor A^t obtained by conjugate transposing each of the frontal slice and then reversing the order of transposed frontal slices starting from the slide number 2 to the slice number n_3 and then appending the conjugate transposed frontal slice $A^{(1)^\top}$.*

Definition 4.3.3 *(The "dot" product) The dot product $A \cdot B$ between two tensors $A \in \mathbb{R}^{n_1\times n_2\times n_3}$ and $B \in \mathbb{R}^{n_2\times n_4\times n_3}$ is the tensor $C \in \mathbb{R}^{n_1\times n_4\times n_3}$ whose slice $C^{(n)}$ is the matrix product of the slice $A^{(n)}$ with the slice $B^{(n)}$:*

$$C^{(k)} := (A\cdot B)^{(k)} := A^{(k)}\ B^{(k)}, \quad k = 1,\ldots,n_3. \tag{4.4}$$

We will also need the canonical inner product.

Definition 4.3.4 *(Inner product of tensors) If A and B are third-order tensors of same size $n_1 \times n_2 \times n_3$, then the inner product between A and B is defined as the following (notice the normalization constant of FFT),*

$$\langle A, B\rangle = \sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3} A_{ijk}B_{ijk}. \tag{4.5}$$

4.3.1.2 Convolution and Fourier transform

Definition 4.3.5 *(t-product for circular convolution) The t-product $A * B$ of $A \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ and $B \in \mathbb{R}^{n_2 \times n_4 \times n_3}$ is an $n_1 \times n_4 \times n_3$ tensor whose $\big(i, j\big)$-th tube is given by*

$$C\big(i, j, :\big) = \sum_{k=1}^{n_2} A\big(i, k, :\big) * B\big(k, j, :\big) \tag{4.6}$$

where $*$ denotes the circular convolution between two cubes of same size.

Definition 4.3.6 *(Identity tensor) The identity tensor $J \in \mathbb{R}^{n_1 \times n_1 \times n_3}$ is defined to be a tensor whose first frontal slice J^1 is the $n_1 \times n_1$ identity matrix and all other frontal slices $J^i, i = 2, \ldots, n_3$ are zero.*

Definition 4.3.7 *(Orthogonal tensor) A tensor $Q \in \mathbb{R}^{n \times n \times n_3}$ is orthogonal if it satisfies*

$$Q^\top * Q = Q * Q^\top = J. \tag{4.7}$$

The tensor $\hat{A}$ is a tensor which is obtained by taking the Fast Fourier Transform (FFT) along the third dimension and we will use the following convention for Fast Transform along the 3rd dimension

$$\hat{A} = \text{fft}\big(A, [\,], 3\big).$$

The one-dimensional FFT along the 3th-dimension is given

$$\hat{A}(j_1, j_2, k_3) = \sum_{j_3=1}^{n_3} A(j_1, j_2, j_3) \exp\Big(-2\frac{i\pi j_3 k_3}{n_3}\Big),$$

for all j_1, j_2, $1 \leqslant j_1 \leqslant n_1$, $1 \leqslant j_2 \leqslant n_2$. Naturally, one can compute A from $\hat{A}$ via $\text{ifft}\big(\hat{A}, [\,], 3\big)$ using the inverse FFT, which is defined as:

$$A(j_1, j_2, k_3) = \sum_{j_3=1}^{n_3} \hat{A}(j_1, j_2, j_3) \exp\Big(2\frac{i\pi j_3 k_3}{n_3}\Big),$$

for all j_1, j_2, $1 \leqslant j_1 \leqslant n_1$, $1 \leqslant j_2 \leqslant n_2$.

Definition 4.3.8 *(Inverse of a tensor) The inverse of a tensor $A \in \mathbb{R}^{n \times n \times n_3}$ is written as A^{-1} satisfying*

$$A^{-1} * A = A * A^{-1} = J \tag{4.8}$$

where J is the identity tensor of size $n \times n \times n_3$.

Remark 4.3.9 *It is proved in [21] that for any tensor $A \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ and $B \in \mathbb{R}^{n_2 \times n_4 \times n_3}$, we have*

$$A * B = C \Leftrightarrow \hat{A} \cdot \hat{B} = \hat{C}.$$

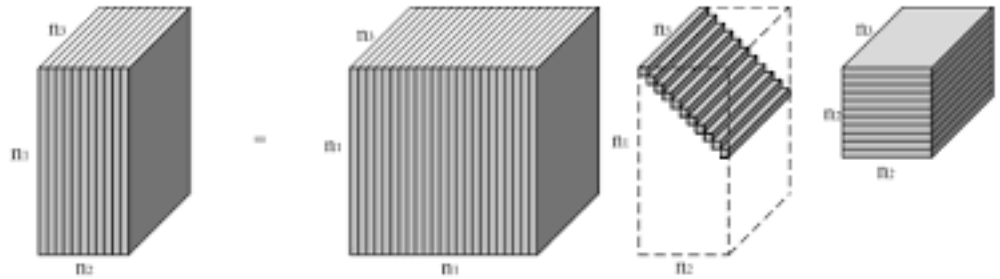

Figure 4.1 – The t-SVD of a tensor

4.3.2 The t-SVD

We finally arrive at the definition of the t-SVD.

Definition 4.3.10 *(f-diagonal tensor) Tensor A is called f-diagonal if each frontal slice $A^{(i)}$ is a diagonal matrix.*

Definition 4.3.11 *(Tensor Singular Value Decomposition: t-SVD) For $M \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, the t-SVD of M is given by*

$$M = U * S * V^\top \tag{4.9}$$

where U and V are orthogonal tensor of size $n_1 \times n_1 \times n_3$ and $n_2 \times n_2 \times n_3$ respectively. S is a rectangular f-diagonal tensor or size $n_1 \times n_2 \times n_3$, and the entries in S are called the singular values of M. This SVD can be obtained using the Fourier transform as follows:

$$\hat{M}^{(i)} = \hat{U}^{(i)} \cdot \hat{S}^{(i)} \cdot \left(\hat{V}^{(i)}\right)^\top. \tag{4.10}$$

This t-SVD is illustrated in Figure 4.1 below. Notice that the diagonal elements of S, i.e. $S(i,i,:)$ are tubal scalars as introduced in Definition 4.3.1. They will also be called *tubal eigenvalues*.

Definition 4.3.12 *The spectrum $\sigma(A)$ of the tensor A is the tubal vector given by*

$$\sigma(A)_i = S(i,i,:) \tag{4.11}$$

for $i = 1, \ldots, \min\{n_1, n_2\}$.

4.3.3 Some natural Tensor Norms

Using the previous definitions, it is easy to define some generalisations of the usual matrix norms.

Definition 4.3.13 *(Tensor Frobenius norm) The induced Frobenius norm from the inner product defined above is given by,*

$$\|A\|_F = \langle A, A\rangle^{1/2} = \frac{1}{\sqrt{n_3}}\|\hat{A}\|_F = \sqrt{\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3} A_{ijk}^2}. \tag{4.12}$$

Definition 4.3.14 *(Tensor spectral norm) The tensor spectral norm $\|A\|_\infty$ is defined as follows:*

$$\|A\|_\infty = \max_i \|\sigma(A)_i\|_2 \tag{4.13}$$

where $\|.\|_2$ is the l_2-norm.

Proposition 4.3.15 *Let M be $n_1 \times n_2 \times n_3$ tensor. Therefore*

$$\|M\|_\infty = \|\mathcal{F}(M)\|_\infty$$

where $\mathcal{F}$ corresponds to the Fast Fourier Transform.

Definition 4.3.16 *(Tubal nuclear norm) The tensor nuclear norm of a tensor A denoted as $\|A\|_\circledast$ is the sum of singular values of all the frontal slices of A. Moreover,*

$$\begin{aligned}\|A\|_\circledast &= \sum_{i=1}^{\min\{n_1,n_2\}} \sqrt{\sum_{j=1}^{n_3} S(i,i,j)^2} \\ &= \sum_{i=1}^{\min\{n_1,n_2\}} \|\sigma(A)_i\|_2.\end{aligned} \tag{4.14}$$

Note that by Parseval's equality

$$\sqrt{\sum_{j=1}^{n_3} S(i,i,j)^2} = \frac{1}{\sqrt{n_3}}\sqrt{\sum_{j=1}^{n_3} \hat{S}(i,i,j)^2}. \tag{4.15}$$

Therefore, it is equivalent to define the tubal-nuclear norm via in the Fourier domain. Recall moreover that the $\hat{S}(i,i,j)$ are all non-negative due to the fact that $\hat{U}^{(k)}\hat{S}^{(k)}\hat{V}^{(k)^t}$ is the SVD of the k^{th} slice of A.

Proposition 4.3.17 *(Trace duality property) Let A, B be $n_1 \times n_2 \times n_3$ tensor. Therefore*

$$|\langle A, B\rangle| \leqslant \|A\|_\circledast \|B\|_\infty.$$

Proof: By Cauchy-Schwartz, we have

$$\begin{aligned}
|\langle A, B\rangle| &= |\langle \mathcal{F}(A), \mathcal{F}(B)\rangle| \\
&= |\langle \mathcal{F}(U)\mathcal{F}(S)\mathcal{F}(V^\top), \mathcal{F}(B)\rangle| \\
&= \left|\sum_{i=1}^{n_3} \mathrm{tr}\Big(\hat{S}^{(i)}\hat{V}^{(i)^\top}\mathcal{F}(B)^{(i)^\top}\hat{U}^{(i)}\Big)\right| \\
&= \left|\sum_{i=1}^{n_3}\sum_{j=1}^{\min\{n_1,n_2\}} \hat{S}_{jj}^{(i)}\Big(\hat{V}^{(i)^\top}\mathcal{F}(B)^{(i)^\top}\hat{U}^{(i)}\Big)_{jj}\right| \\
&\leqslant \sum_{j=1}^{\min\{n_1,n_2\}} \Big(\|\hat{S}_{jj}\|_2\Big)^{1/2}\Big(\|(\hat{V}^\top\mathcal{F}(B)^t\hat{U})_{jj}\|_2\Big)^{1/2} \\
&\leqslant \sum_{j=1}^{\min\{n_1,n_2\}} \Big(\|\hat{S}_{jj}\|_2\Big)^{1/2}\Big(\|\mathcal{F}(B)_{jj}\|_2\Big)^{1/2}
\end{aligned}$$

taking the maximum of $\|\mathcal{F}(B)_{jj}\|_2$ and the sum the slices of $\Big(\|\hat{S}_{jj}\|_2\Big)^{1/2}$, and apply (4.15) and inverse of FFT, we obtain the result. □

Proposition 4.3.18 *Given tensor $A \in \mathbb{R}^{n_1\times n_2\times n_3}$. We have*

$$\|A\|_\circledast \leqslant \sqrt{\mathrm{rank}(A)}\|A\|_F.$$

Proof: Again by Cauchy-Schwartz, we have

$$\begin{aligned}
\|A\|_\circledast &= \sum_{j=1}^{\min\{n_1,n_2\}} \|S(j,j,:)\|_2 \\
&= \sum_{j=1}^{\mathrm{rank(A)}} \|S(j,j,:)\|_2 \\
&\leqslant \sqrt{\mathrm{rank}(A)}\Big(\sum_{j=1}^{\min\{n_1,n_2\}} \|S(j,j,;)\|_2^2\Big)^{1/2} \\
&\leqslant \sqrt{\mathrm{rank}(A)}\|A\|_F.
\end{aligned}$$

□

4.3.4 Rank, Range and Kernel

The rank, the range and the kernel are extremely important notions for matrices. They will play a role in our analysis of the penalised least squares tensor recovery procedure as well.

As noticed in [21], a tubal scalar may have all its entrees different from zero but still be non-invertible. According to the definition, a tubal scalar $a \in \mathbb{R}^{1\times 1\times n_3}$ is invertible if there

exists a tubal scalar b such that $a * b = b * a = e$. Equivalently, the Fourier transform $\hat{a}$ of a has no coefficient equal to zero. We can define the tubal rank ρ_i of $S_{i,i,:}$ as the number of non-zero components of $\hat{S}(i,i,:)$. Then, the easiest way of defining the rank of a tensor is by means of the notion of multirank as follows.

Definition 4.3.19 *The multirank of a tensor is the vector $(\rho_1, \ldots, \rho_r)$ where r is the number of nonzero tubal vectors $S(i,i,:)$, $i = 1, \ldots, \min\{n_1, n_2\}$ where comes from the t-SVD of M and r is also called the rank of the tensor M.*

We now define the range of a tensor.

Definition 4.3.20 *Let j denote the number of invertible tubal eigenvalues and let k denote the number of nonzero non-invertible tubal eigenvalues. The range $\mathcal{R}(M)$ of a tensor $M \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is defined as*

$$\mathcal{R}(M) = \{\vec{U}^{(1)} * c_1 + \cdots + \vec{U}^{(j+k)} * c_{j+k} \mid c_l \in \mathcal{R}ange(s_l * \cdot), \\ l \in \{j+1, \cdots, j+k\}\}.$$

Definition 4.3.21 *Let j denote the number of invertible tubal eigenvalues. The kernel $\mathcal{K}(M)$ of a tensor $M \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is defined as*

$$\mathcal{K}(M) = \{\vec{V}^{(j+1)} * c_1 + \cdots + \vec{V}^{(n_2)} * c_{n_2} \mid s_l * c_l = 0,\ l \in \{j+1, \cdots, j+n_2\}\}.$$

4.4 Main result

In this section, we present our main contribution to the analysis of the Bürer-Monteiro approach to the tensor completion problem. For this purpose, let $\mathcal{M}^\star = \mathcal{U}^\star * \mathcal{V}^{\star^\top}$ denote the factorisation of $\mathcal{M}^*$, and for any variable tensor $\mathcal{M}$, we will use the similar factorisation $\mathcal{M} = \mathcal{U} * \mathcal{V}^\top$. We can now define the following objective function of $\mathcal{M}$ expressed as a function of $(\mathcal{U}, \mathcal{V})$:

$$f\big(\mathcal{U}, \mathcal{V}\big) = 2\big(\mathcal{M} - \mathcal{M}^\star\big) : \mathcal{H}_0 : \big(\mathcal{M} - \mathcal{M}^\star\big) + \frac{1}{2}\|\mathcal{U}^\top * \mathcal{U} - \mathcal{V}^\top * \mathcal{V}\|_F^2 \\ + Q_0\big(\mathcal{U}, \mathcal{V}\big). \tag{4.16}$$

The asymmetric problem can easily be reduced to a symmetric problem as follows. Suppose $\mathcal{M}^\star$ is the optimal solution and its t-SVD is $\mathcal{X}^\star * \mathcal{D}^\star * \mathcal{Y}^{\star^\top}$. Let $\mathcal{U}^\star = \mathcal{X}^\star * \left(\mathcal{D}^\star\right)^{\frac{1}{2}}$, $\mathcal{V}^\star = \mathcal{Y}^\star * \left(\mathcal{D}^\star\right)^{\frac{1}{2}}$ and $\mathcal{M} = \mathcal{U} * \mathcal{V}^\top$ is the current point, we reduce the problem into a symmetric case by using following notations.

$$\mathcal{W} = \begin{pmatrix} \mathcal{U} \\ \mathcal{V} \end{pmatrix}, \mathcal{W}^\star = \begin{pmatrix} \mathcal{U}^\star \\ \mathcal{V}^\star \end{pmatrix}, \mathcal{N} = \mathcal{W} * \mathcal{W}^\top, \mathcal{N}^\star = \mathcal{W}^\star * \mathcal{W}^{\star^\top}. \tag{4.17}$$

In the sequel, we define $\Delta = \mathcal{W} - \mathcal{W}^\star * \mathcal{R}$.

We will also transform the Hessian operator to operate $(n_1 + n_2) \times r \times n_3$ tensors. For this purpose, define the tensors $\mathcal{H}_1$ and $\mathcal{G}$ such that for all $(\mathcal{U}, \mathcal{V})$ we have:

$$\begin{aligned}\mathcal{N} : \mathcal{H}_1 : \mathcal{N} &= \mathcal{M} : \mathcal{H}_0 : \mathcal{M} \\ \mathcal{N} : \mathcal{G} : \mathcal{N} &= \|\mathcal{U}^\top * \mathcal{U} - \mathcal{V}^\top * \mathcal{V}\|_F^2\end{aligned}$$

where we recall that $\mathcal{N}$ is a function of $(\mathcal{U}, \mathcal{V})$. Now, let $Q\big(\mathcal{W}\big) := Q_0\big(\mathcal{U}, \mathcal{V}\big)$ and we can rewrite the objective function $f\big(\mathcal{W}\big)$ as

$$\frac{1}{2}\Big[\big(\mathcal{N} - \mathcal{N}^\star\big) : 4\mathcal{H}_1 : \big(\mathcal{N} - \mathcal{N}^\star\big) + \mathcal{N} : \mathcal{G} : \mathcal{N}\Big] + Q\big(\mathcal{W}\big). \tag{4.18}$$

The main result of this paper is the following theorem.

Theorem 4.4.1 *Let* $d = \max\{n_1, n_2\}$. *Recall that the probability of selecting one entry is denoted by* $p = 1/(n_1 \times n_2)$. *Assume that* μ *and* r *satisfy the following inequality*

$$p \geqslant c_1 \frac{\mu^4 r^6 \big(\kappa^\star\big)^6 \log(d)}{\min\{n_1, n_2\}},$$

for some positive constant c_1. *Choose*

$$\alpha_1^2 = \Theta\Big(\frac{\mu r \sigma_1^\star}{n_1}\Big),\ \alpha_2^2 = \Theta\Big(\frac{\mu r \sigma_1^\star}{n_2}\Big)$$

and

$$\lambda_1 = \Theta\Big(\frac{n_1}{\mu r \kappa^\star}\Big),\ \lambda_2 = \Theta\Big(\frac{n_2}{\mu r \kappa^\star}\Big).$$

Then with probability at least

$$1 - 2n_1 n_3 \exp\left\{-pn_2\Big(\Big(1 + \frac{t}{pn_2}\Big)\ln\Big(1 + \frac{t}{pn_2}\Big) - \frac{t}{pn_2}\Big)\right\},$$

for tensor completion objective (4.18) we have

- *All local minima satisfy* $\mathcal{U} * \mathcal{V}^\top = \mathcal{M}^\star$;
- *the function is* $\Big(\varepsilon, c_1\big(\sigma_r^\star\big), C\big(\frac{\varepsilon}{\sigma_r^\star}\big)\Big)$*-strict saddle for polynomially small* ε.

for some positives constants C, c_1.

Proof: See 4.7.3. □

4.5 Numerical validation on medical images

In this section, we present some numerical results validating our approach on medical images and volumes. Our experiments were performed on Optical Coherence Tomography (OCT) images, also called "*optical biopsies*" used by clinicians to perform micrometric (at cellular level) characterization of biological tissues in both *in situ* and *ex vivo* settings. Application of OCT in different medical setups such as ophthalmology, dermatology, cardiovascular surgery, etc, is usually considered of high clinical value. However, *in situ* acquisition of high resolution and 3-dimensional optical biopsies is well known to be very challenging in practice. Some well known drawbacks of using OCT for such medical applications are: long acquisition times (generating artefacts, e.g., under physiological disturbances) for full-resolution volume acquisition. Moreover, preprocessing/processing, transfer and storage of very large datasets (up to 10 Go for a full resolution OCT volume) is one of the main limitations for using OCT-based optical biopsies in some medical applications of interest. The subsampling approach together with the efficient factorisation-based optimisation method proposed in the present paper aim at circumventing these issues.

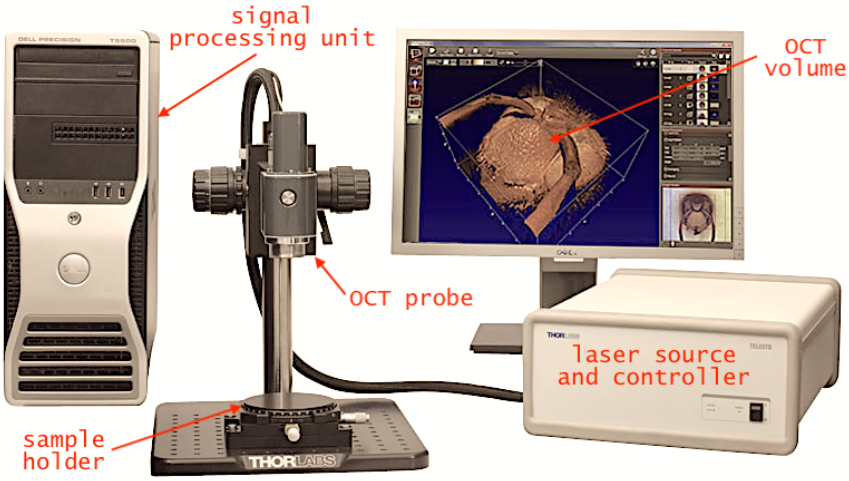

Figure 4.2 – Photography of our OCT imaging system.

This section discusses different setups using progressive subsampling rate ranging from 20% to 80%. In Section 4.5.1, we present of our spectral domain OCT system (the most popular marketed OCT systems). In Section 4.5.2, we describe the different experimental scenarios. Our cmputational results are presented in Section 4.5.3

4.5.1 OCT Imaging System

As mentioned above, OCT is a well-established medical imaging technique (e.g., for optical biopsy-based diagnosis) that uses a light wave to capture 3-dimensional images of a light-diffusing material (e.g., biological tissue) with a micrometer (1μm) resolution [120]. OCT is uses low coherence interferometric technique at near-infrared wavelength. Indeed, light

absorption of imaged biological tissues is limited in near-infrared light wavelength range, which restricts penetration up to about 1mm. This technique is thus halfway between ultrasonic (resolution of 150μm, penetration of 10cm) and confocal microscopy (resolution of 0.5μm, penetration of 200μm).

The OCT imaging technique allows to retrieve three types of information. Firstly, each position of the light spot on the imaged tissue gives the reflectivity profile (z axis), called A-scan, which can contain information about the structure and spatial dimensions of the sample under study. Secondly, a 2-dimensional slice ($x-z$ axes scan) of the sample (transverse tomography), called B-scan, can be obtained by combining series of A-scan profiles. Finally, combining successive B-scan cross-sections allows acquiring volumetric OCT data ($x-y-z$ axes scan), called C-scan.

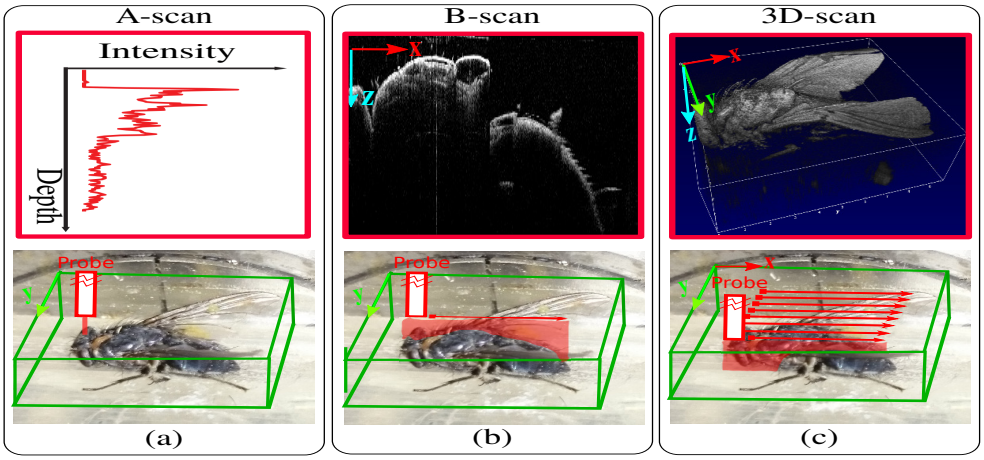

Figure 4.3 – Available acquisition modes in an OCT imaging system: (a) A-scan, (b) B-Scan, and (c) C-scan.

The acquisition of the different types of OCT signals (i.e., A-scan, B-scan, and C-scan) is performed sequentially by moving the light spot on the imaged sample. In other words, it is possible to acquire each single data independently of the others. In particular, the 2 degrees-of-freedom galvanometers integrated in the OCT probe makes it possible to optimise the sampling using any prespecified geometrically constrained protocol. As a result, one of the great advantages of OCT is that it is ideally suited to geometric subsampling in the spirit of compressed sensing.

4.5.2 Validation Scenarios

The developed materials and methods were implemented in a MatLab framework without taking into account code optimization aspects nor time-computation. The numerical validation of the methods was achieved using two optical biopsies acquired on biological samples: a piece of a grape (Fig. 4.4(left)), a sample of fish eye retina (Fig. 4.4(right))

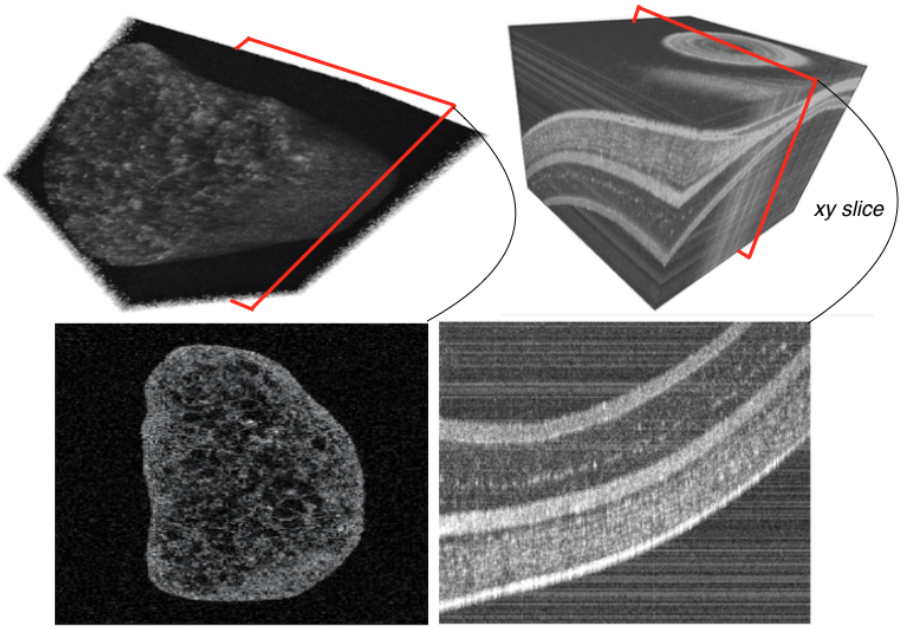

Figure 4.4 – Examples of the OCT volumes of biological samples used to validate the proposed method. (first row): the initial OCT volumes and (second row), B-scan images (100^{th} vertical slice) taken from the original volumes.

recorded from a commercial OCT device*. Both optical biopsies (considered as low-tubal-rank tensors) have equal size $A_{n_1 \times n_2 \times n_3} = 281 \times 281 \times 281$ voxels. Different scenarios were considered to assess the performance of our algorithm. The sampling rates used for these experiments ranged from 20% to 80% (with a step of 10%) and formed masks that were applied to the original volume (we randomly pick 20% to 80% pixels from the original tensors). Finally, we set the maximum iteration number to be $i_{max} = 10$.

4.5.3 Obtained Results

Note that instead of illustrating the fully reconstructed OCT volume, we choose to show 2D images (the 100^{th} xz slice of the reconstructed volumes) for a better visualization, with the naked eye, of the quality of the obtained results as can be shown in Fig. 4.6. Again, it can be noticed that the sharpness of the boundary is will preserved. Furthermore, the recovered data can be improved such as using conventional filters based post-processing methods.

*The Telesto-II from Thorlabs (https://www.thorlabs.com/thorproduct.cfm?partnumber=TELESTO-II)

Retina

The retina is chosen because for its translucent characteristics offering very good conditions for acquiring high resolution OCT (optical biopsy) images/volumes and interesting signal-to-noise ratio comparing to most OCT images. The reconstructed data from the subsampled volumes are depicted in Fig. 4.5

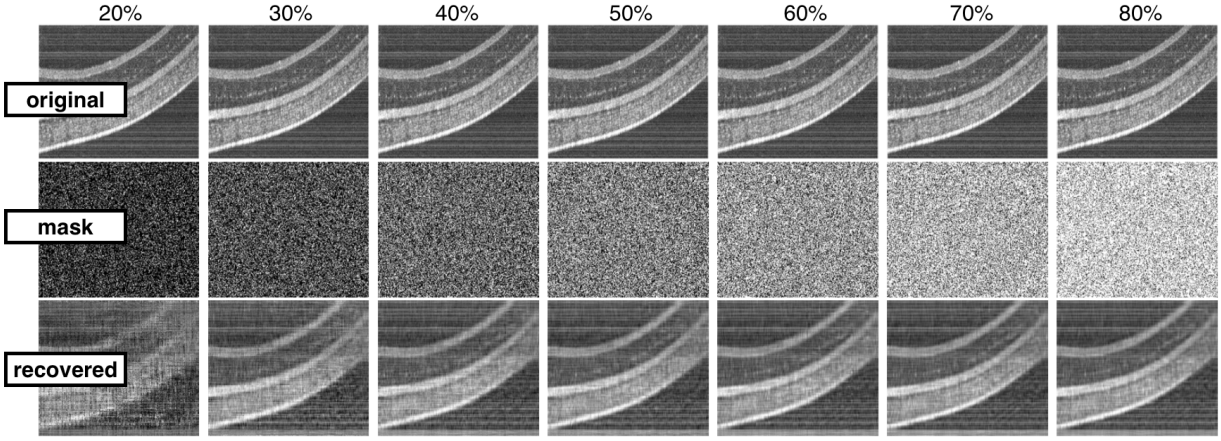

Figure 4.5 – [**sample: retina**] - Reconstructed OCT images (only a 2D slice is shown in this example). First row corresponds to the original slice, second row the subsampled data (ranging from 20% to 80% with a step of 10%) to be reconstructed, and third row the reconstructed slices.

Grape

As mentioned above, the second OCT volume used to assess the performance of the algorithm is recorded by imaging a part of a grape. Even, the grape is also a translucent medium, the signal-to-noise ratio is less important than the one obtained by imaging the retina. The validation scenario is still the same as for the first test, i.e., different subsampling OCT volumes were built using 20% to 80% (with a step of 10%) of the original data. Once again, it can be noticed that the sharpness of the boundary is preserved. Furthermore, the recovered data can be improved such as using conventional filters based post-processing methods.

4.5.4 Evaluation Scores

To quantitatively assess the results obtained with the proposed methods, we implemented two criteria: 1) the peak signal noise ratio (PSNR), and 2) the structural similarity index (SSIM), which are generally used by the compressed sensing and more widely the image processing communities [121].

- The Peak Signal Noise Ratio (PSNR) computed as follows

$$PSNR = 10 \log_{10} \left(\frac{d^2}{MSE} \right) \tag{4.19}$$

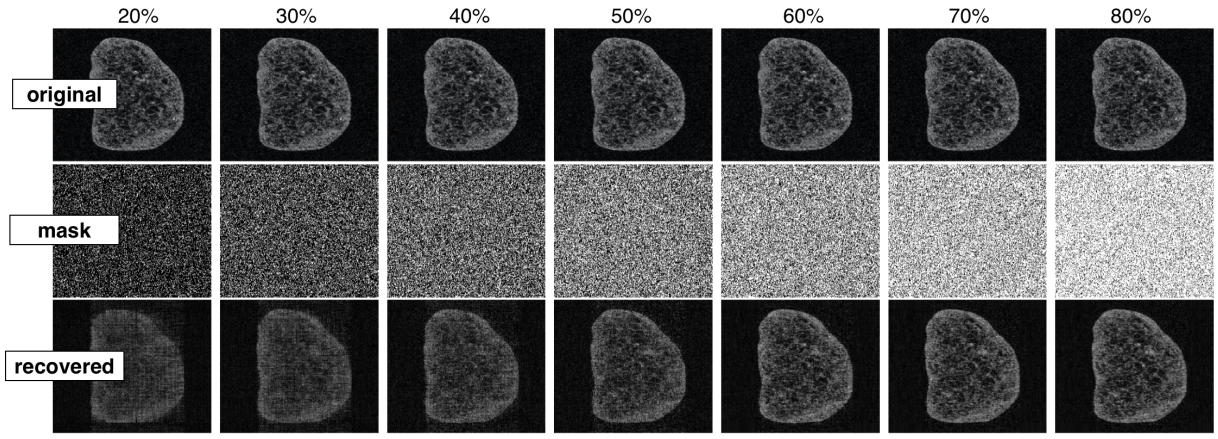

Figure 4.6 – [**sample: grape**] - Reconstructed OCT images (only a 2D slice is shown in this example). First row corresponds to the original slice, second row the subsampled data (ranging from 20% to 80% with a step of 10%) to be reconstructed, and third row the reconstructed slices.

where d is the maximal pixel value in the initial OCT image and the MSE (mean-squared error) is obtained by

$$MSE = \sum_{i=1}^{n}\sum_{j=1}^{m}\Big(I_o(i,j) - I_r(i,j)\Big)^2 \tag{4.20}$$

with I_o and I_r represent an initial 2D OCT slice (selected from the OCT volume) and the recovered one, respectively.

- The Structural Similarity Index (SIMM) which allows measuring the degree of similarity between two images. It is based on the computation of three values namely the luminance l, the contrast c and the structural aspect s. It is given by

$$SSIM = \Big(s\big(I_r, I_o)\Big)\Big(l\big(I_r, I_o\Big)\Big(c\big(I_r, I_o\big)\Big) \tag{4.21}$$

where,

$$l = \frac{2\mu_{I_r}\mu_{I_o} + C_1}{\mu_{I_r}^2 + \mu_{I_o}^2 + C_1}, \tag{4.22}$$

$$c = \frac{2\sigma_{I_r}\sigma_{I_o} + C_2}{\sigma_{I_r}^2 + \sigma_{I_o}^2 + C_2}, \tag{4.23}$$

$$s = \frac{2\sigma_{I_r,I_o} + C_3}{\sigma_{I_r}\sigma_{I_o} + C_3}, \tag{4.24}$$

with μ_{I_r}, μ_{I_o}, σ_{I_r}, σ_{I_o}, and μ_{I_r,I_o} are the local means, standard deviations, and cross-covariance for images I_r, I_o. The variables C_1, C_2, C_3 are used to stabilize the division with weak denominator.

Tables 4.1 summarizes the numerical values the PSNR and SSIM computed for each test. The obtained numerical results for both evaluation scores clearly demonstrate the relevance of the proposed approach for this type of images/volumes. As expected, increasing the number of samples significantly improves the quality scores, however, using only 20% sampled data gives unexpectedly good and exploitable recovery. In the range from 30% to 80%, the reconstructed data are faithful to the original ones.

Table 4.1 – Numerical values of the SSIM and PSNR scores.

sample 1: eye							
subsampling rate	20%	30%	40%	50%	60%	70%	80%
PSNR	14.20	17.73	18.44	18.70	18.87	18.99	19.19
SSIM	00.13	00.29	00.34	00.36	00.38	00.38	00.39
sample 1: grape							
subsampling rate	20%	30%	40%	50%	60%	70%	80%
PSNR	19.24	19.69	19.64	20.30	22.07	23.58	24.44
SSIM	00.20	00.25	00.30	00.38	00.37	00.43	00.46

4.5.5 Impact of the Initialization Parameters on the Quality of the Reconstruction

Note that two initialization parameters have influence on the quality of the reconstruction. They concern the number of iterations "i" and the tubal rank "r". First, in the numerical validation discussed above, both the number of iterations i and tubal rank r were, respectively, fixed to $i_{max} = 20$ and $r = 20$.

4.5.5.1 Number of iterations i

In this section, we varied the values of these parameters and for each pair (i, r), we computed both the PSNR and SSIM values. As can be seen in Fig. 4.7, the best reconstruction (for $r = 20$) was obtained using only few iterations i.e., $i = 5$.

4.5.5.2 Tubal rank r

According to the above statement, we fixed the iterations number $i = 5$, and we varied the tubal rank r. As can be shown in Fig. 4.8, the best similarity scores (PSNR and SIMM) are obtained for a $r = 80$.

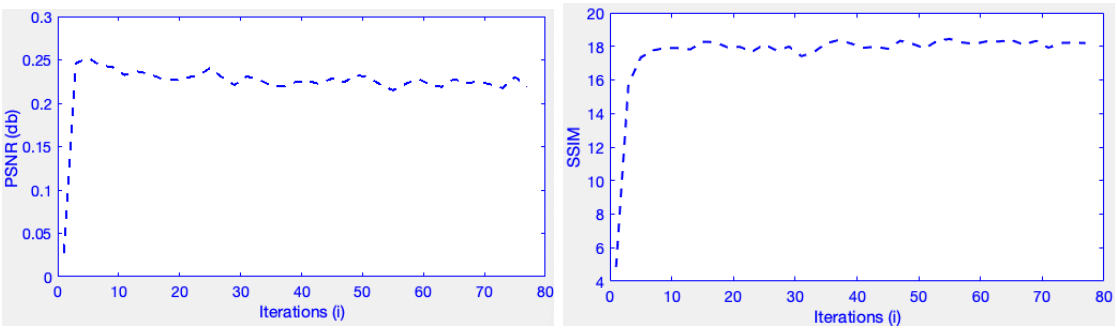

Figure 4.7 – Representation of both the PSNR (right) and the SSIM (left) criteria in function of the number of iterations i.

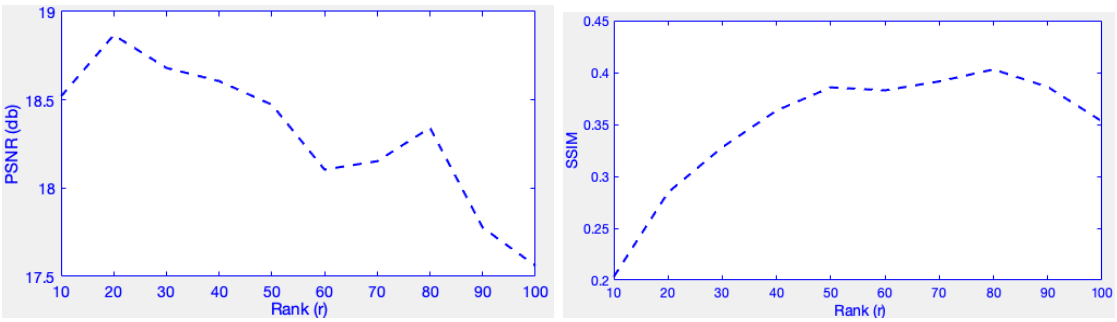

Figure 4.8 – Representation of both the PSNR (right) and the SSIM (left) criteria in function of the tubal rank values r

4.6 Conclusion and Perspectives

In this paper, we studied a low tubal-rank tensor completion problem using non-convex optimisation, as initially proposed in [118]. A theoretical extension of the analysis in [119] was provided in order to address the important tubal tensor case. The theoretical results were validated numerically using real data, i.e., OCT volumes acquired in biological samples (a retina and a grape). The obtained results are encouraging and demonstrate the performance of the low tubal-rank tensor completion problem.

Further work will consist in the validation of the method in a physical imaging device. In order to achieve this, it will be important to consider a GPU implementation of the algorithm in order to address the real-time processing inherent challenges. Additional research work can be undertaken in adapting the algorithm to an online setting where the hyperparameters can be learned using e.g. the approach of [36].

Bibliography

[1] A. Moitra, *Algorithmic aspects of machine learning*, Cambridge University Press, 2018.

[2] N. Kreimer, A. Stanton, and M. D Sacchi, *Tensor completion based on nuclear norm minimization for 5d seismic data reconstruction*, Geophysics, **78**, 2013, no. 6, V273–V284.

[3] Y. Wang, R. Chen, J. Ghosh, J. C Denny, A. Kho, Y. Chen, B. A Malin, and J. Sun, *Rubik: Knowledge guided tensor factorization and completion for health data analytics*, Int. Conf. on Knowledge Discovery and Data Mining, ACM, 2015, pp. 1265–1274.

[4] N. Li and B. Li, *Tensor completion for on-board compression of hyperspectral images*, IEEE Int. Conf. on Image Processing.

[5] W. Hu, D. Tao, W. Zhang, Y. Xie, and Y. Yang, *The twist tensor nuclear norm for video completion*, IEEE Trans. on Neural Networks and Learning Systems, **28**, 2017, no. 12, 2961–2973.

[6] Ji Liu, Przemyslaw Musialski, Peter Wonka, and Jieping Ye, *Tensor completion for estimating missing values in visual data*, IEEE Trans. on Pattern Analysis and Machine Intelligence, **35**, 2013, no. 1, 208–220.

[7] T. G Kolda and B. W Bader, *Tensor decompositions and applications*, SIAM review **51**, 2009, no. 3, 455–500.

[8] Andrzej Cichocki, Danilo Mandic, Lieven De Lathauwer, Guoxu Zhou, Qibin Zhao, Cesar Caiafa, and Huy Anh Phan, *Tensor decompositions for signal processing applications: From two-way to multiway component analysis*, IEEE Signal Processing Mag., **32**, 2015, no. 2, 145–163.

[9] Nicholas D Sidiropoulos, Lieven De Lathauwer, Xiao Fu, Kejun Huang, Evangelos E Papalexakis, and Christos Faloutsos, *Tensor decomposition for signal processing and machine learning*, IEEE Trans. on Signal Processing, **65**, 2017, no. 13, 3551–3582.

[10] H. Sedghi, M. Janzamin, and A. Anandkumar, *Provable tensor methods for learning mixtures of generalized linear models*, Artificial Intelligence and Statistics, 2016, pp. 1223–1231.

[11] A. Anandkumar, R. Ge, D. Hsu, D. M Kakade, and M. Telgarsky, *Tensor decompositions for learning latent variable models*, The Journal of Machine Learning Research, ,**15** 2014, no. 1, 2773–2832.

[12] S. Gandy, B. Recht, and I. Yamada, *Tensor completion and low-n-rank tensor recovery via convex optimization*, Inverse Problems **27**, 2011, no. 2, 025010.

[13] B. Romera-Paredes and M. Pontil, *A new convex relaxation for tensor completion*, Advances in Neural Information Processing Systems, 2013, pp. 2967–2975.

[14] J. A. Bazerque, G. Mateos, and G. B. Giannakis, *Rank regularization and bayesian inference for tensor completion and extrapolation*, IEEE Trans. on Signal Processing **61**, 2013, no. 22, 5689–5703.

[15] D. Kressner, M. Steinlechner, and B. Vandereycken, *Low-rank tensor completion by riemannian optimization*, BIT Numerical Mathematics, **54**, 2014, no. 2, 447–468.

[16] M. Filipović and A. Jukić, *Tucker factorization with missing data with application to low-rank tensor completion*, Multidimensional Systems and Signal Processing, **26**, 2015, no. 3, 677–692.

[17] S. Chrétien and T. Wei, *Convex recovery of tensors using nuclear norm penalization*, Int. Conf. on Latent Variable Analysis and Signal Separation, Springer, 2015, pp. 360–367.

[18] M. Yuan and C.-H. Zhang, *On tensor completion via nuclear norm minimization*, Foundations of Computational Mathematics, **16**, 2016, no. 4, 1031–1068.

[19] S. Chrétien and T. Wei, *Sensing tensors with gaussian filters*, IEEE Trans. on Information Theory, **63**, 2017, no. 2, 843–852.

[20] L. De Lathauwer, B. De Moor, and J. Vandewalle, *A multilinear singular value decomposition*, SIAM Jounal on Matrix Analysis and Applications, **21**, 2000, no. 4, 1253–1278.

[21] M. E Kilmer, K. Braman, N. Hao, and R. C Hoover, *Third-order tensors as operators on matrices: A theoretical and computational framework with applications in imaging*, SIAM Journal on Matrix Analysis and Applications, **34**, 2013, no. 1, 148–172.

[22] A. Anandkumar, R. Ge, and M. Janzamin, *Guaranteed non-orthogonal tensor decomposition via alternating rank*-1 *updates*, arXiv, preprint arXiv:1402.5180, 2014.

[23] B. Recht, M. Fazel, and P. A Parrilo, *Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization*, SIAM review, **52**, 2010, no. 3, 471–501.

[24] Yuanyuan Liu and Fanhua Shang, *An efficient matrix factorization method for tensor completion*, IEEE Signal Processing Letters, **20** (2013), no. 4, 307–310.

[25] X.-Y. Liu, S. Aeron, V. Aggarwal, and X. Wang, *Low-tubal-rank tensor completion using alternating minimization*, Modeling and Simulation for Defense Systems and Applications XI, vol. 9848, Int. Society for Optics and Photonics, 2016, p. 984809.

[26] D. Wei, A. Wang, X. Feng, B. Wang, and B. Wang, *Tensor completion based on triple tubal nuclear norm*, Algorithms, **11**, 2018, no. 7, 94.

[27] A. Wang, Z. Lai, and Z. Jin, *Noisy low-tubal-rank ten-8sor completion.* Neurocomputing, 330:267-279, 2019.

[28] S. Burer and R. DC Monteiro, *A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization*, Mathematical Programming, **95**, 2003, no. 2, 329–357.

[29] N. Boumal, V. Voroninski, and A. Bandeira, *The non-convex burer-monteiro approach works on smooth semidefinite programs*, Advances in Neural Information Processing Systems, 2016, pp. 2757–2765.

[30] Rong Ge, Chi Jin, and Yi Zheng, *No spurious local minima in nonconvex low rank problems: A unified geometric analysis*, Int. Conf. on Machine Learning, 2017, pp. 1233–1242.

[31] Candès, Emmanuel J and Tao, Terence, The power of convex relaxation: Near-optimal matrix completion, IEEE Transactions on Information Theory, 56, no. 5, 2053–2080, 2010.

[32] Rong Ge, Furong Huang, Chi Jin, and Yang Yuan, *Escaping From Saddle Points — Online Stochastic Gradient for Tensor Decomposition*, arXiv, preprint arXiv:1503.02101, 2015.

[33] M. E Kilmer and C. D Martin, *Factorization strategies for third-order tensors*, Linear Algebra and its Applications, **435** 2011, no. 3, 641–658.

[34] Fujimoto, James G, Pitris, Costas, Boppart, Stephen A and Brezinski, Mark E, Optical coherence tomography: an emerging technology for biomedical imaging and optical biopsy, Neoplasia, **2**, 1-2, 9–25,2000.

[35] Miao J, Huo D, and Wilson DL., *Quantitative image quality evaluation of MR images using perceptual difference models*, Med Phys., **35**, 2008, no. 6, 2541-2553.

[36] Chretien, Stephane and Lohvithee, Manasavee and Sun, Wenjuan and Soleimani, Manuchehr, *Efficient hyper-parameter selection in Total Variation-penalised XCT reconstruction using Freund and Shapire's Hedge approach*, Mathematics, **8**, 2020, no. 4, 493.

[37] Candès, Emmanuel and Recht Benjamin, *Exact matrix ompletion via convex optimization*, Communications of ACM, **55**, 2012, no. 6, 111–119.

[38] Zemin Zhang and Shuchin Aeron, *Exact tensor completion using t-svd*, IEEE Transactions on Signal Processing, **65**, 2017, no. 6, 1511–1526.

4.7 Proofs of our results

4.7.1 Concentration Inequalities for matrix completion

For matrix completion, we need different concentration inequalities for different kinds of matrices. The first type of matrix lies a tangent space and is proved in [37].

Lemma 4.7.1 *[37] Let $d = \max\{n_1, n_2\}$. Define the subspace*

$$\mathcal{T} = \{M \in \mathbb{R}^{n_1 \times n_2} \,|M = U^\star X^\top + Y V^{\star^\top}, \text{ for some } X \in \mathbb{R}^{n_1 \times r}, Y \in \mathbb{R}^{n_2 \times r}\}.$$

Then, for any $\delta > 0$, as long as sample rate $p \geqslant \Omega\Big(\frac{\mu r}{\delta^2 d}\log(d)\Big)$, we will have:

$$\|\frac{1}{p}\mathcal{P}_\mathcal{T}\mathcal{P}_\Omega\mathcal{P}_\mathcal{T} - \mathcal{P}_\mathcal{T}\| \leqslant \delta.$$

For arbitrary low rank matrix, we will need the following lemma.

Lemma 4.7.2 *Suppose that $\Omega \subset [n_1] \times [n_2]$ is the set of edges of a random bipartite graph with $\Big(n_1, n_2\Big)$ nodes, where any pair nodes on different side is connected with probability p. Let $d = \max\{n_1, n_2\}$, then there exists two universal constants c_1, c_2, for any $\delta > 0$ such that for $p \geqslant c_1 \frac{\log\Big(d\Big)}{\min\{n_1, n_2\}}$, then with probability at least $1 - d^{-4}$, we have for any $x, y \in \mathbb{R}^n$:*

$$\frac{1}{p}\sum_{\Big(i,j\Big)\in\Omega} x_i y_j \leqslant \|x\|_1\|y\|_1 + c_2\sqrt{\frac{d}{p}}\,\|x\|_2\|y\|_2.$$

This theorem implies following:

Lemma 4.7.3 *Let $d = \max\{n_1, n_2\}$. There exists universal constant c_1, c_2, for any $\delta > 0$ so that if $p \geqslant c_1 \frac{\log\Big(d\Big)}{\min\{n_1, n_2\}}$ then with probability at least $1 - \frac{1}{2}d^{-4}$, we have for any matrices $X, Y \in \mathbb{R}^{d\times r}$:*

$$\frac{1}{p}\|XY^\top\|_\Omega^2 \leqslant \|X\|_F^2\|Y\|_F^2 + c_2\sqrt{\frac{d}{p}}\|X\|_F\|Y\|_F \,.\, \max_{1\leqslant i\leqslant d}\|e_i^\top X\| \,.\, \max_{1\leqslant j\leqslant d}\|e_j^\top Y\|.$$

On the other hand, for all low rank matrices we also have the following which is tighter for incoherent matrices.

Lemma 4.7.4 *[118]Let $d = \max\{n_1, n_2\}$, then with at least probability $1 - e^{\Omega\Big(d\Big)}$ over random choice of Ω, we have for any rank $2r$ matrices $A \in \mathbb{R}^{n_1 \times n_2}$:*

$$\Big|\frac{1}{p}\|\mathcal{P}_\Omega\Big(A\Big)\|_\Omega^2 - \|A\|_F^2\Big| \leqslant C\Big(\frac{d\ r\log\Big(d\Big)}{p}\|A\|_\infty^2 + \sqrt{\frac{d\ r\log\Big(d\Big)}{p}}\|A\|_F\|A\|_\infty\Big)$$

for some positive constant C.

Finally, for a matrix with each entry randomly sampled independently with small probability p, next theorem says with high probability, no row can have too many non-zero entries.

Lemma 4.7.5 *Let Ω_i denote the support of Ω on the i-th row, let $d = \max\{n_1, n_2\}$. Assume $pn_2 \geqslant \log\Big(2d\Big)$, then with probability at least*

$$1 - 2n_1 \exp\left\{-pn_2\Big(\Big(1 + \frac{t}{pn_2}\Big)\ln\Big(1 + \frac{t}{pn_2}\Big) - \frac{t}{pn_2}\Big)\right\}$$

over random choice of Ω, we have for all $i \in [n_1]$ simultaneously:

$$|\Omega_i| \leqslant C\ pn_2,$$

for some positive constant C.

4.7.2 The case of symmetric positive definite problems

We start with the simple definition of tubal symmetry for tensors.

Definition 4.7.6 *[21] $\mathcal{A} \in \mathbb{R}^{n\times n\times n_3}$ is a symmetric positive definite if $\hat{\mathcal{A}}^{(i)}$ are Hermitian positive definite for $i = 1, \ldots, n_3$ where $\hat{\mathcal{A}}$ is the Fast Fourier Transform (FFT) of tensor $\mathcal{A}$.*

In the following, we assume that the tensor $\mathcal{M}^\star = \mathcal{U}^\star * \Big(\mathcal{U}^\star\Big)^\top$ is symmetric and positive semi-definite with $\mathcal{U} \in \mathbb{R}^{n\times r\times n_3}$. The goal is to find the unknown tensor $\mathcal{U}^\star$ solution the following non-convex optimization problem

$$\min_{\mathcal{M}\in\mathbb{R}^{n\times n\times n_3}} \frac{1}{2}\Big(\mathcal{M} - \mathcal{M}^\star\Big) : \mathcal{H} : \Big(\mathcal{M} - \mathcal{M}^\star\Big) \text{ s.t. } \operatorname{rank}(\mathcal{M}) = r \tag{4.25}$$

where the rank of $\mathcal{M}$ is defined in Section 4.3.4. Using the factorization idea of Burer and Monteiro [116], the corresponding unconstrained optimization problem with regularization Q can be written as

$$\min_{\mathcal{U}\in\mathbb{R}^{n\times n\times n_3}} \frac{1}{2}\Big(\mathcal{U} * \mathcal{U}^\top - \mathcal{M}^\star\Big) : \mathcal{H} : \Big(\mathcal{U} * \mathcal{U}^\top - \mathcal{M}^\star\Big) + Q(\mathcal{U}). \tag{4.26}$$

We now present the concept of "direction of improvement" which was introduced in [118].

Definition 4.7.7 *(Direction of improvement) Let* $\mathcal{U}, \mathcal{U}^\star \in \mathbb{R}^{n\times r\times n_3}$*, define*

$$\Delta = \mathcal{U} - \mathcal{U}^\star * \mathcal{R}$$

where $\mathcal{R} \in \mathbb{R}^{r\times r\times n_3}$ *is defined as*

$$\mathcal{R} = \underset{\mathcal{Z}^\top * \mathcal{Z} = \mathcal{Z} * \mathcal{Z}^\top = \mathcal{J}}{argmin} \|\mathcal{U} - \mathcal{U}^\star * Z\|_F^2.$$

The direction of improvement is clearly the best direction towards the ground truth solution and the first set to take if one wants to improve the objective value. The direction of improvement is intrumental for proving Lemma 4.7.8, which is key to our analysis. This lemma Our version is an adaptation of [118, Lemma 7] to the case of low rank tubal tensor factorisation in the sense proposed by Kilmer. The main technical difficulty of adapting the proof of [118, Lemma 7] is to decouple the slices of the tensor in order to arrive at the same type of computations as in the original version of the result. This is achieved by taking the Fourier transform along the tubes.

Lemma 4.7.8 *(Main) Let* Δ *be defined as in (4.7.7) and* $\mathcal{M} = \mathcal{U} * \mathcal{U}^\top$*. Then, for any* $\mathcal{U} \in \mathbb{R}^{n\times r\times n_3}$*, we have*

$$\begin{aligned}\Delta : \nabla^2 f(\mathcal{U}) : \Delta = \Delta * \Delta^\top : \mathcal{H} : \Delta * \Delta^\top - 3(\mathcal{M} - \mathcal{M}^\star) : \mathcal{H} : (\mathcal{M} - \mathcal{M}^\star) \\ + 4\langle \nabla f(\mathcal{U}), \Delta\rangle + [\Delta : \nabla^2 Q(\mathcal{U}) : \Delta - 4\langle \nabla Q(\mathcal{U}), \Delta\rangle].\end{aligned} \tag{4.27}$$

Proof: We have

$$\begin{aligned} f(\mathcal{U}) &= \frac{1}{2}\left(\mathcal{U} * \mathcal{U}^\top - \mathcal{M}^\star\right) : \mathcal{H} : \left(\mathcal{U} * \mathcal{U}^\top - \mathcal{M}^\star\right) + Q(\mathcal{U}) \\ &= \frac{1}{2}\left\langle \mathcal{U} * \mathcal{U}^\top - \mathcal{M}^\star, \mathcal{H}\Big(\mathcal{U} * \mathcal{U}^\top - \mathcal{U}^\star\Big)\right\rangle + Q(\mathcal{U}) \end{aligned}$$

and we therefore get

$$\begin{aligned} f(\mathcal{U}) &= \frac{1}{2}\left\langle \mathcal{F}(\mathcal{U} * \mathcal{U}^\top - \mathcal{M}^\star), \mathcal{F}(\mathcal{H}(\mathcal{U} * \mathcal{U}^\top - \mathcal{M}^\star))\right\rangle + Q(\mathcal{U}) \\ &= \frac{1}{2}\left\langle \mathcal{F}(\mathcal{U}) * \mathcal{F}(\mathcal{U}^\top) - \mathcal{F}(\mathcal{M}^\star), \mathcal{F}(\mathcal{H}) * \mathcal{F}(\mathcal{U}) * \mathcal{F}(\mathcal{U}^\top) - \mathcal{F}(\mathcal{H}) * \mathcal{F}(\mathcal{M}^\star)\right\rangle + Q(\mathcal{U}) \end{aligned}$$

which gives

$$\begin{aligned} &= \frac{1}{2}\sum_{k=1}^{n_3}\Big\langle \mathcal{F}\Big(\mathcal{U}\Big)^{(k)}\mathcal{F}\Big(\mathcal{U}^\top\Big)^{(k)} - \mathcal{F}\Big(\mathcal{M}^\star\Big)^{(k)}, \mathcal{F}\Big(\mathcal{H}\Big)^{(k)}\mathcal{F}\Big(\mathcal{U}\Big)^{(k)}\mathcal{F}\Big(\mathcal{U}^\top\Big)^{(k)} \\ &\quad - \mathcal{F}\Big(\mathcal{H}\Big)^{(k)}\mathcal{F}\Big(\mathcal{M}^\star\Big)^{(k)}\Big\rangle + Q(\mathcal{U}) \end{aligned}$$

and thus

$$f(\mathcal{U}) = \frac{1}{2}\sum_{k=1}^{n_3}\Big\langle \mathcal{F}(\mathcal{U})^{(k)}\mathcal{F}(\mathcal{U}^\top)^{(k)} - \mathcal{F}(\mathcal{M}^\star)^{(k)}, \mathcal{F}(\mathcal{H})^{(k)}\Big(\mathcal{F}(\mathcal{U})^{(k)}\mathcal{F}(\mathcal{U}^\top)^{(k)} - \mathcal{F}(\mathcal{M}^\star)^{(k)}\Big)\Big\rangle + Q(\mathcal{U})$$

$$= \underbrace{\frac{1}{2}\sum_{k=1}^{n_3}\begin{bmatrix} \Big(\mathcal{F}\Big(\mathcal{U}\Big)^{(k)}\mathcal{F}\Big(\mathcal{U}^\top\Big)^{(k)} - \mathcal{F}\Big(\mathcal{M}^\star\Big)^{(k)}\Big) \\ : \mathcal{F}\Big(\mathcal{H}\Big)^{(k)} : \\ \Big(\mathcal{F}\Big(\mathcal{U}\Big)^{(k)}\mathcal{F}\Big(\mathcal{U}^\top\Big)^{(k)} - \mathcal{F}\Big(\mathcal{M}^\star\Big)^{(k)}\Big) \end{bmatrix}}_{=G(\mathcal{U})} + Q(\mathcal{U}).$$

Using the fact, with for any $\mathcal{Z} \in \mathbb{R}^{n\times r\times n_3}$, we have:

$$\langle \nabla f(\mathcal{U}), \mathcal{Z}\rangle = \langle \nabla G(\mathcal{U}), \mathcal{Z}\rangle + \langle \nabla Q(\mathcal{U}), \mathcal{Z}\rangle$$

and

$$\mathcal{Z} : \nabla^2 f(\mathcal{U}) : \mathcal{Z} = \mathcal{Z} : \nabla^2 G(\mathcal{U}) : \mathcal{Z} + \mathcal{Z} : \nabla^2 Q(\mathcal{U}) : \mathcal{Z}.$$

So, we need to compute $\langle \nabla G(\mathcal{U}), \mathcal{Z}\rangle$ and $\mathcal{Z} : \nabla^2 G(\mathcal{U}) : \mathcal{Z}$. For this, by expanding the fact, for any $\mathcal{Z} \in \mathbb{R}^{n\times r\times n_3}$, we know that:

$$G\Big(\mathcal{U}+\mathcal{Z}\Big) = G\Big(\mathcal{U}\Big) + \langle \nabla G\Big(\mathcal{U}\Big), \mathcal{Z}\rangle + \frac{1}{2}\,\mathcal{Z} : \nabla^2 G\Big(\mathcal{U}\Big) : \mathcal{Z} + O\Big(\|\mathcal{Z} * \mathcal{Z}^\top\|^2\Big).$$

We obtain,

$$\begin{aligned}\langle \nabla G(\mathcal{U}), \mathcal{Z}\rangle &= \sum_{k=1}^{n_3}\left\{\begin{array}{c} \Big[\mathcal{F}(\mathcal{U})^{(k)}\mathcal{F}(\mathcal{U}^\top)^{(k)} - \mathcal{F}(\mathcal{M}^\star)^{(k)}\Big] \\ : \mathcal{F}(\mathcal{H})^{(k)} : \\ \Big[\mathcal{F}(\mathcal{U})^{(k)}\mathcal{F}(\mathcal{Z}^\top)^{(k)} + \mathcal{F}(\mathcal{Z})^{(k)}\mathcal{F}(\mathcal{U}^\top)^{(k)}\Big] \end{array}\right\} \\ &= \Big(\mathcal{F}(\mathcal{M}) - \mathcal{F}(\mathcal{M}^\star)\Big) : \mathcal{F}(\mathcal{H}) : \Big(\mathcal{F}(\mathcal{U}) * \mathcal{F}(\mathcal{Z}^\top) + \mathcal{F}(\mathcal{Z}) * \mathcal{F}(\mathcal{U}^\top)\Big) \\ &= \Big(\mathcal{M} - \mathcal{M}^\star\Big) : \mathcal{H} : \Big(\mathcal{U} * \mathcal{Z}^\top + \mathcal{Z} * \mathcal{U}^\top\Big) \end{aligned} \tag{4.28}$$

and

$$\begin{aligned}\mathcal{Z} : \nabla^2 G(\mathcal{U}) : \mathcal{Z} &= \sum_{k=1}^{n_3}\left\{\begin{array}{c} \Big[\mathcal{F}(\mathcal{U})^{(k)}\mathcal{F}(Z^\top)^{(k)} + \mathcal{F}(\mathcal{Z})^{(k)}\mathcal{F}(\mathcal{U}^\top)^{(k)}\Big] \\ : \mathcal{F}(\mathcal{H})^{(k)} : \\ \Big[\mathcal{F}(\mathcal{U})^{(k)}\mathcal{F}(\mathcal{Z}^\top)^{(k)} + \mathcal{F}(\mathcal{Z})^{(k)}\mathcal{F}(\mathcal{U}^\top)^{(k)}\Big] \end{array}\right\} \\ &\quad + 2\Big[\mathcal{F}(\mathcal{U})^{(k)}\mathcal{F}(\mathcal{U}^\top)^{(k)} - \mathcal{F}(\mathcal{M}^\star)^{(k)}\Big] : \mathcal{F}(\mathcal{H})^{(k)} : \Big[\mathcal{F}(\mathcal{Z})^{(k)}\mathcal{F}(\mathcal{Z}^\top)^{(k)}\Big] \\ &= \Big(\mathcal{F}(\mathcal{U}) * \mathcal{F}(\mathcal{Z}^\top) + \mathcal{F}(\mathcal{Z}) * \mathcal{F}(\mathcal{U}^\top)\Big) : \mathcal{F}(\mathcal{H}) : \Big(\mathcal{F}(\mathcal{U}) * \mathcal{F}(\mathcal{Z}^\top) + \mathcal{F}(\mathcal{Z}) * \mathcal{F}(\mathcal{U}^\top)\Big) \\ &\quad + 2\Big(\mathcal{F}(\mathcal{M}) - \mathcal{F}(\mathcal{M}^\star)\Big) : \mathcal{F}(\mathcal{H}) : \mathcal{F}(\mathcal{Z}) * \mathcal{F}(\mathcal{Z}^\top) \\ &= \Big(\mathcal{U} * \mathcal{Z}^\top + \mathcal{Z} * \mathcal{U}^\top\Big) : \mathcal{H} : \Big(\mathcal{U} * \mathcal{Z}^\top + \mathcal{Z} * \mathcal{U}^\top\Big) \\ &\quad + 2\Big(\mathcal{M} - \mathcal{M}^\star\Big) : \mathcal{H} : \mathcal{Z} * \mathcal{Z}^\top. \end{aligned} \tag{4.29}$$

In the last equality of (4.28) and (4.29), we use the linearity of Fourier transform and the inverse of FFT. Let $\mathcal{Z} = \Delta = \mathcal{U} - \mathcal{U}^\star * \mathcal{R}$ and $\mathcal{M} - \mathcal{M}^\star + \Delta * \Delta^\top = \mathcal{U} * \Delta^\top + \Delta * \mathcal{U}^\top$. Indeed,

$$\mathcal{U} * \Delta^\top + \Delta * \mathcal{U}^\top = \mathcal{U} * \mathcal{U}^\top - \mathcal{U} * \mathcal{R}^\top * \mathcal{U}^{\star^\top} + \mathcal{U} * \mathcal{U}^\top - \mathcal{U}^\star * \mathcal{R} * \mathcal{U}^\top$$

and using that $\mathcal{R} * \mathcal{R}^\top = \mathcal{J}$, where $\mathcal{J}$ is a identity tensor, we have

$$\begin{aligned}\mathcal{M} - \mathcal{M}^\star + \Delta * \Delta^\top &= \mathcal{U} * \mathcal{U}^\top - \mathcal{U}^\star * \mathcal{U}^{\star^\top} + \big(\mathcal{U} - \mathcal{U}^\star * \mathcal{R}\big) * \big(\mathcal{U} - \mathcal{U}^\star * \mathcal{R}\big)^\top \\ &= \mathcal{U} * \mathcal{U}^\top - \mathcal{U} * \mathcal{R}^\top * \mathcal{U}^{\star^\top} + \mathcal{U} * \mathcal{U}^\top - \mathcal{U}^\star * \mathcal{R} * \mathcal{U}^\top.\end{aligned}$$

Using

$$\big(\mathcal{M} - \mathcal{M}^\star\big) : \mathcal{H} : \big(\mathcal{M} - \mathcal{M}^\star + \Delta * \Delta^\top\big) = \langle \nabla f(\mathcal{U}), \Delta\rangle - \langle \nabla Q(\mathcal{U}), \Delta\rangle,$$

we have

$$\begin{aligned}\Delta : \nabla^2 f(\mathcal{U}) : \Delta =& \big(\mathcal{U} * \Delta^\top + \Delta * \mathcal{U}^\top\big) : \mathcal{H} : \big(\mathcal{U} * \Delta^\top + \Delta * \mathcal{U}^\top\big) \\ &+2\big(\mathcal{M} - \mathcal{M}^\star\big) : \mathcal{H} : \Delta * \Delta^\top + \Delta : \nabla^2 Q(\mathcal{U}) : \Delta \\ =& \big(\mathcal{M} - \mathcal{M}^\star + \Delta * \Delta^\top\big) : \mathcal{H} : \big(\mathcal{M} - \mathcal{M}^\star + \Delta * \Delta^\top\big) \\ &+2\big(\mathcal{M} - \mathcal{M}^\star\big) : \mathcal{H} : \Delta * \Delta^\top + \Delta : \nabla^2 Q(\mathcal{U}) : \Delta \\ =& \Delta * \Delta^\top : \mathcal{H} : \Delta * \Delta^\top + \big(\mathcal{M} - \mathcal{M}^\star\big) : \mathcal{H} : \big(\mathcal{M} - \mathcal{M}^\star\big) \\ &+4\big(\mathcal{M} - \mathcal{M}^\star\big) : \mathcal{H} : \Delta * \Delta^\top + \Delta : \nabla^2 Q(\mathcal{U}\mathcal{M}) : \Delta \\ =& \Delta * \Delta^\top : \mathcal{H} : \Delta * \Delta^\top - 3\big(\mathcal{M} - \mathcal{M}^\star\big) : \mathcal{H} : \big(\mathcal{M} - \mathcal{M}^\star\big) \\ &+4\big(\mathcal{M} - \mathcal{M}^\star\big) : \mathcal{H} : \big(\mathcal{M} - \mathcal{M}^\star + \Delta * \Delta^\top\big) + \Delta : \nabla^2 Q(\mathcal{U}) : \Delta \\ =& \Delta * \Delta^\top : \mathcal{H} : \Delta * \Delta^\top - 3(\mathcal{M} - \mathcal{M}^\star) : \mathcal{H} : (\mathcal{M} - \mathcal{M}^\star) \\ &+4\langle \nabla f(\mathcal{U}), \Delta\rangle + [\Delta : \nabla^2 Q(\mathcal{U}) : \Delta - 4\langle \nabla Q(\mathcal{U}), \Delta\rangle].\end{aligned}$$

□

Using the previous lemma, we are now able to prove the following result.

Lemma 4.7.9 *Given tensors $\mathcal{U}, \mathcal{U}^\star \in \mathbb{R}^{n\times r\times n_3}$. Let $\mathcal{M} = \mathcal{U} * \mathcal{U}^\top$, $\mathcal{M}^\star = \mathcal{U}^\star * \mathcal{U}^{\star^\top}$, and Δ be defined as in* (4.7.7)*, then we have*

$$\|\Delta * \Delta^\top\|_F^2 \leqslant 2\|\mathcal{M} - \mathcal{M}^\star\|_F^2 \quad \textit{and} \quad \sigma_r^\star \|\Delta\|_F^2 \leqslant \frac{1}{2\big(\sqrt{2}-1\big)}\|\mathcal{M} - \mathcal{M}^\star\|_F^2.$$

Proof: We begin to show that

$$\mathcal{U}^\top * \mathcal{U}^\star * \mathcal{R}_\mathcal{U} \text{ is a symmetric } \textbf{PSD} \text{ tensor.} \tag{4.30}$$

where $\mathcal{R}_\mathcal{U} = \underset{\mathcal{R}*\mathcal{R}^\top=\mathcal{R}^\top*\mathcal{R}=\mathcal{J}}{\text{Argmin}} \|\mathcal{U}-\mathcal{U}^\star*\mathcal{R}\|_F^2$. By developping the Frobenius norm and letting the t-SVD of ${\mathcal{U}^\star}^\top * \mathcal{U}$ be $\mathcal{A}*\mathcal{D}*\mathcal{B}^\top$, we have:

$$\|\mathcal{U}-\mathcal{U}^\star*\mathcal{R}\|_F^2 = \|\mathcal{U}*\mathcal{U}^\top\|_F^2 - 2\langle \mathcal{U}, \mathcal{U}^\star*\mathcal{R}\rangle + \|\mathcal{U}^\star*{\mathcal{U}^\star}^\top\|_F^2.$$

Hence,

$$\begin{aligned}
\underset{\mathcal{R}*\mathcal{R}^\top=\mathcal{R}^\top*\mathcal{R}=\mathcal{J}}{\text{Argmin}} \|\mathcal{U}-\mathcal{U}^\star*\mathcal{R}\|_F^2 &= \underset{\mathcal{R}*\mathcal{R}^\top=\mathcal{R}^\top*\mathcal{R}=\mathcal{J}}{\text{Argmin}} - \langle \mathcal{U}, \mathcal{U}^\star * R\rangle \\
&= \underset{\mathcal{R}*\mathcal{R}^\top=\mathcal{R}^\top*\mathcal{R}=\mathcal{J}}{\text{Argmin}} - \langle \mathcal{F}(\mathcal{U}), \mathcal{F}(\mathcal{U}^\star)*\mathcal{F}(\mathcal{R})\rangle \\
&= \underset{\mathcal{R}*\mathcal{R}^\top=\mathcal{R}^\top*\mathcal{R}=\mathcal{J}}{\text{Argmin}} - \sum_{k=1}^{n_3} \langle \mathcal{F}(\mathcal{U})^{(k)}, \mathcal{F}(\mathcal{U}^\star)^{(k)}\mathcal{F}(\mathcal{R})^{(k)}\rangle \\
&= \underset{\mathcal{R}*\mathcal{R}^\top=\mathcal{R}^\top*\mathcal{R}=J}{\text{Argmin}} - \sum_{k=1}^{n_3} \text{trace}\Big(\mathcal{F}(\mathcal{U})^{(k)^\top}\mathcal{F}(\mathcal{U}^\star)^{(k)}\mathcal{F}(\mathcal{R})^{(k)}\Big) \\
&= \underset{\mathcal{R}*\mathcal{R}^\top=\mathcal{R}^\top*\mathcal{R}=\mathcal{J}}{\text{Argmin}} - \sum_{k=1}^{n_3} \text{trace}\Big(\mathcal{F}\Big(\mathcal{D}\Big)^{(k)}\mathcal{F}\Big(\mathcal{A}^\top\Big)^{(k)}\mathcal{F}(\mathcal{R})^{(k)}\mathcal{F}\Big(\mathcal{B}\Big)^{(k)}\Big).
\end{aligned}$$

Since $\mathcal{A}$, $\mathcal{R}$ and $\mathcal{B}$ are orthogonal tensors, then $\mathcal{F}\Big(\mathcal{A}^\top\Big)^{(k)}$, $\mathcal{F}(\mathcal{R})^{(k)}$ and $\mathcal{F}\Big(\mathcal{B}\Big)^{(k)}$ are orthogonal matrices. For any orthogonal tensor $\mathcal{T}$, we have

$$\begin{aligned}
\text{trace}\Big(\mathcal{D}*\mathcal{T}\Big) = \langle \mathcal{D}, \mathcal{T}\rangle &= \langle \mathcal{F}(\mathcal{D}), \mathcal{F}(\mathcal{T})\rangle \\
&= \sum_{k=1}^{n_3} \text{trace}\Big(\mathcal{F}(\mathcal{D})^{(k)}\mathcal{F}(\mathcal{T})^{(k)}\Big) \\
&= \sum_{k=1}^{n_3}\sum_{i=1}^{r} \mathcal{F}(\mathcal{D})_{ii}^{(k)}\mathcal{F}(\mathcal{T})_{ii}^{(k)} \\
&\leqslant \sum_{k=1}^{n_3}\sum_{i=1}^{r} \mathcal{F}(\mathcal{D})_{ii}^{(k)}
\end{aligned}$$

where the last inequality uses the fact that $\mathcal{F}(\mathcal{D})_{ii}^{(k)}$ is a positive singular values and $\mathcal{T}$ is an orthogonal tensor thus $\mathcal{F}(\mathcal{T})_{ii}^{(k)} \leqslant 1$. This implies that the maximum of $\mathcal{F}(\mathcal{D})_{ii}^{(k)}\mathcal{F}(\mathcal{T})_{ii}^{(k)}$ is attained at $\mathcal{T}=\mathcal{J}$. In other words, the minimum is attained when

$$-\sum_{k=1}^{n_3} \text{trace}\Big(\mathcal{F}\Big(\mathcal{D}\Big)^{(k)}\mathcal{F}\Big(\mathcal{A}^\top\Big)^{(k)}\mathcal{F}(\mathcal{R})^{(k)}\mathcal{F}\Big(\mathcal{B}\Big)^{(k)}\Big)$$

is attained when

$$\mathcal{R} = \mathcal{A}*\mathcal{B}^\top.$$

Finally, since

$$\mathcal{U}^\top * \mathcal{U}^\star * \mathcal{R}_\mathcal{U} = \mathcal{B} * \mathcal{D} * \mathcal{A}^\top * \mathcal{R}_\mathcal{U} = \mathcal{B} * \mathcal{D} * \mathcal{B}^\top$$

we get that $\mathcal{U}^\top * \mathcal{U}^\star * \mathcal{R}_\mathcal{U}$ is a symmetric PSD tensor and the proof is completed. □

The following technical result follows the lines of the analysis provided in [118] and shows how one can control the factorisation of differences using the differences of factorisations.

Lemma 4.7.10 *Let $\mathcal{U}$ and $\mathcal{Y}$ be two $n \times n \times n_3$ tensors. Let $\mathcal{U}^\top * \mathcal{Y} = \mathcal{Y}^\top * \mathcal{U}$ be a PSD tensor. Then,*

$$\|\Big(\mathcal{U} - \mathcal{Y}\Big) * \Big(\mathcal{U} - \mathcal{Y}\Big)^\top\|_F^2 \leqslant \|\mathcal{U} * \mathcal{U}^\top - \mathcal{Y} * \mathcal{Y}^\top\|_F^2.$$

Proof: Let $\Delta = \mathcal{U} - \mathcal{Y}$, and we have

$$\begin{aligned}
\|\mathcal{U} * \mathcal{U}^\top - \mathcal{Y} * \mathcal{Y}^\top\|_F^2 =& \|\mathcal{U} * \Delta^\top + \Delta * \mathcal{U}^\top - \Delta * \Delta^\top\|_F^2 \\
=& \text{trace}\Big(\Delta * \mathcal{U}^\top * \mathcal{U} * \Delta^\top + \Delta * \mathcal{U}^\top * \Delta * \mathcal{U}^\top \\
& - \Delta * \mathcal{U}^\top * \Delta * \Delta^\top + \mathcal{U} * \Delta^\top * \mathcal{U} * \Delta^\top\Big) \\
& + \text{trace}\Big(\mathcal{U} * \Delta^\top * \Delta * \mathcal{U}^\top - \mathcal{U} * \Delta^\top * \Delta * \Delta^\top \\
& - \Delta * \Delta^\top * \mathcal{U} * \Delta^\top - \Delta * \Delta^\top * \Delta * \mathcal{U}^\top\Big) \\
& + \text{trace}\Big(\Delta * \Delta^\top * \Delta * \Delta^\top\Big).
\end{aligned}$$

On the other hand,

$$\begin{aligned}
\text{trace}\Big(\Delta * \mathcal{U}^\top * \mathcal{U} * \Delta^\top\Big) &= \langle \mathcal{F}\Big(\Delta\Big) \cdot \mathcal{F}\Big(\mathcal{U}^\top\Big), \mathcal{F}\Big(\Delta\Big) \cdot \mathcal{F}\Big(\mathcal{U}^\top\Big)\rangle \\
&= \sum_{k=1}^{n_3} \text{trace}\Big(\mathcal{F}\Big(\Delta\Big)^{(k)} \mathcal{F}\Big(\mathcal{U}^\top\Big)^{(k)} \mathcal{F}\Big(\mathcal{U}\Big)^{(k)} \mathcal{F}\Big(\Delta^\top\Big)^{(k)}\Big) \\
&= \sum_{k=1}^{n_3} \text{trace}\Big(\mathcal{F}\Big(\mathcal{U}^\top\Big)^{(k)} \mathcal{F}\Big(\mathcal{U}\Big)^{(k)} \mathcal{F}\Big(\Delta^\top\Big)^{(k)} \mathcal{F}\Big(\Delta\Big)^{(k)}\Big) \\
&= \text{trace}\Big(\mathcal{U}^\top * \mathcal{U} * \Delta^\top * \Delta\Big)
\end{aligned}$$

$$\begin{aligned}
\text{trace}\Big(\Delta * \mathcal{U}^\top * \Delta * \Delta^\top\Big) &= \langle \mathcal{F}\Big(\Delta\Big) \cdot \mathcal{F}\Big(\mathcal{U}^\top\Big), \mathcal{F}\Big(\Delta^\top\Big) \cdot \mathcal{F}\Big(\Delta\Big)\rangle \\
&= \sum_{k=1}^{n_3} \text{trace}\Big(\mathcal{F}\Big(\Delta\Big)^{(k)} \mathcal{F}\Big(\mathcal{U}^\top\Big)^{(k)} \mathcal{F}\Big(\Delta\Big)^{(k)} \mathcal{F}\Big(\Delta^\top\Big)^{(k)}\Big) \\
&= \sum_{k=1}^{n_3} \text{trace}\Big(\mathcal{F}\Big(\mathcal{U}^\top\Big)^{(k)} \mathcal{F}\Big(\Delta\Big)^{(k)} \mathcal{F}\Big(\Delta^\top\Big)^{(k)} \mathcal{F}\Big(\Delta\Big)^{(k)}\Big) \\
&= \text{trace}\Big(\mathcal{U}^\top * \Delta * \Delta^\top * \Delta\Big).
\end{aligned}$$

In a similar manner, we get

- $\operatorname{trace}\big(\Delta * \mathcal{U}^\top * \Delta * \Delta^\top\big) = \operatorname{trace}\big(\mathcal{U} * \Delta^\top * \Delta * \Delta^\top\big) = \operatorname{trace}\big(\Delta * \Delta^\top * \mathcal{U} * \Delta^\top\big) = \operatorname{trace}\big(\Delta * \Delta^\top * \Delta * \mathcal{U}^\top\big)$
- $\operatorname{trace}\big(\Delta * \mathcal{U}^\top * \mathcal{U} * \Delta^\top\big) = \operatorname{trace}\big(\mathcal{U} * \Delta^\top * \Delta * \mathcal{U}^\top\big)$
- $\operatorname{trace}\big(\Delta * \mathcal{U}^\top * \Delta * \mathcal{U}^\top\big) = \operatorname{trace}\big(\mathcal{U} * \Delta^\top * \mathcal{U} * \Delta^\top\big)$.

Therefore using that $\mathcal{U}^\top * \mathcal{Y} = \mathcal{Y}^\top * \mathcal{U}$, we have

$$
\begin{aligned}
&\|\mathcal{U} * \mathcal{U}^\top - \mathcal{Y} * \mathcal{Y}^\top\|_F^2 \\
&= \operatorname{trace}\big(2\mathcal{U}^\top * \mathcal{U} * \Delta^\top * \Delta + \Delta^\top * \Delta * \Delta^\top * \Delta + \mathcal{U}^\top * \Delta * \mathcal{U}^\top * \Delta \\
&\qquad\qquad - 4\mathcal{U}^\top * \Delta * \Delta^\top * \Delta\big) \\
&= \operatorname{trace}\big(2\mathcal{U}^\top * \mathcal{U} * \Delta^\top * \Delta + 2\big(\mathcal{U}^\top * \Delta\big)^2 + \big(\Delta^\top * \Delta\big)^2 - 4\mathcal{U}^\top * \Delta * \Delta^\top * \Delta\big) \\
&= \operatorname{trace}\big(2\mathcal{U}^\top * \big(\mathcal{U} - \Delta\big) * \Delta^\top * \Delta + \big(\frac{1}{\sqrt{2}}\Delta^\top * \Delta - \sqrt{2}\mathcal{U}^\top * \Delta\big)^2 \\
&\qquad\qquad + \frac{1}{2}\big(\Delta^\top * \Delta\big)^2\big) \\
&\geqslant \operatorname{trace}\big(2\mathcal{U}^\top * \mathcal{Y} * \Delta^\top * \Delta \\
&\qquad\qquad + \frac{1}{2}\big(\Delta^\top * \Delta\big)^2\big) \\
&\geqslant \frac{1}{2}\|\Delta * \Delta^\top\|_F^2
\end{aligned}
$$

where the last inequality is a consequence of the fact that $\mathcal{U}^\top * \mathcal{Y}$ is a positive semi-definite tensor. □

The next lemma will also be key.

Lemma 4.7.11 *Let $\mathcal{U}$ and $\mathcal{Y}$ be two $n \times n \times n_3$ tensors. Let $\mathcal{U}^\top * \mathcal{Y} = \mathcal{Y}^\top * \mathcal{U}$ be a PSD tensor. Then,*

$$
\begin{aligned}
\sigma_{\min}\big(\mathcal{U}^\top * \mathcal{U}\big)\|\mathcal{U} - \mathcal{Y}\|_F^2 &\leqslant \|\big(\mathcal{U} - \mathcal{Y}\big) * \mathcal{U}^\top\|_F^2 \\
&\leqslant \frac{1}{2\big(\sqrt{2}-1\big)}\|\mathcal{U} * \mathcal{U}^\top - \mathcal{Y} * \mathcal{Y}^\top\|_F^2.
\end{aligned}
$$

Proof: Let $\Delta = \mathcal{U} - \mathcal{Y}$, and we have

$$\begin{aligned}
&\|\mathcal{U} * \mathcal{U} - \mathcal{Y} * \mathcal{Y}^\top\|_F^2 = \|\mathcal{U} * \Delta^\top + \Delta * \mathcal{U}^\top - \Delta * \Delta^\top\|_F^2 \\
&= \operatorname{trace}\Big(2\mathcal{U}^\top * \mathcal{U} * \Delta^\top * \Delta + 2\Big(\mathcal{U}^\top * \Delta\Big)^2 + \Big(\Delta^\top * \Delta\Big)^2 - 4\mathcal{U}^\top * \Delta * \Delta^\top * \Delta\Big) \\
&\geqslant \operatorname{trace}\Big(\Big(4 - 2\sqrt{2}\Big)\mathcal{U}^\top * \mathcal{Y} * \Delta^\top * \Delta + 2\Big(\sqrt{2} - 1\Big)\mathcal{U}^\top * \mathcal{U} * \Delta^\top * \Delta\Big) \\
&\geqslant 2\Big(\sqrt{2} - 1\Big)\|\mathcal{U} * \Delta^\top\|_F^2.
\end{aligned}$$

where the last inequality uses the positive semidefiniteness of $\mathcal{U}^\top * \mathcal{Y}$. □

Combining 4.7.10 and 4.7.11, it now within reach to obtain Lemma 4.7.9, after replacing $\mathcal{U}$ by $\mathcal{U}^\star * \mathcal{R}_\mathcal{U}$ and $\mathcal{Y}$ by $\mathcal{U}$.

Let us now turn to clarifying the interaction between the Hessian and the regulariser. The necessity of using a penalisation (regularisation) comes from the deficiency of the Hessian operator $\mathcal{H}$ in preserving the norm of all low rank tubal tensors. A standard approach to making the Bürer Monteiro successful is to impose some incoherence on the matrix to be reconstructed such as proposed in the following definition.

Definition 4.7.12 *[71] Let $\mathcal{M} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ and its t-SVD of the form $\mathcal{M} = \mathcal{X} * \mathcal{D} * \mathcal{Y}^\top$. Let $r = rank(\mathcal{M})$. Then, $\mathcal{M}$ is said to satisfy the tensor incoherence property with parameter $\mu > 0$, if*

$$\max_{i=1,\dots,n_1} \|e_i^\top * \mathcal{X}\|_F \leqslant \sqrt{\frac{\mu r}{n_1}}$$

$$\max_{j=1,\dots,n_2} \|e_j^\top * \mathcal{Y}\|_F \leqslant \sqrt{\frac{\mu r}{n_2}}$$

where e_i is the $n_1 \times 1 \times n_3$ column basis with $e_{i11} = 1$ and e_j is the $n_2 \times 1 \times n_3$ column basis with $e_{j11} = 1$.

In the following, we will assume that our unknown low rank tensor $\mathcal{M}^\star$ is μ-incoherent.

In the non-convex problem, we try to make sure that the decomposition $\mathcal{U} * \mathcal{U}^\top$ is also incoherent by adding a regularizer of [118], that penalize the function objective when some row of $\mathcal{F}\Big(\mathcal{U}\Big)^{(k)}, k = 1 \dots, n_3$ is too large.

$$Q\Big(\mathcal{U}\Big) = \lambda \sum_{k=1}^{n_3} \sum_{i=1}^{n} \left(\|\mathcal{F}\Big(e_i^\top\Big)^{(k)} \mathcal{F}\Big(\mathcal{U}\Big)^{(k)}\|_2 - \alpha \right)_+^4.$$

Here λ, α are parameters that we choose later, $(x)_+ = \max\{x, 0\}$. By adding this regularizer, we can transform the objective function to the unconstrained form

$$\min_{\mathcal{U} \in \mathbb{R}^{n \times r \times n_3}} \quad \frac{1}{2p}\|\mathcal{U} * \mathcal{U}^\top - \mathcal{M}^\star\|_\Omega^2 + Q(\mathcal{U}). \tag{4.31}$$

Using this fact we begin to show that the regularizer ensures that all rows of $\mathcal{F}(\mathcal{U})^{(k)}, k = 1 \dots, n_3$ are small.

We now study the properties of the gradient and Hessian of the regularizer Q:

Lemma 4.7.13 *The gradient and the Hessian of the regularizer*

$$Q\big(\mathcal{U}\big) = \lambda \sum_{k=1}^{n_3} \sum_{i=1}^{n} \Big(\|\mathcal{F}\Big(e_i^\top\Big)^{(k)} \mathcal{F}\big(\mathcal{U}\big)^{(k)}\|_2 - \alpha \Big)_+^4$$

is:

$$\begin{aligned} \langle \nabla Q(\mathcal{U}), \mathcal{Z} \rangle = 4\lambda \sum_{k=1}^{n_3} \sum_{i=1}^{n} \Big(\Big(\|\mathcal{F}\Big(e_i^\top\Big)^{(k)} \mathcal{F}\big(\mathcal{U}\big)^{(k)}\|_2 - \alpha \Big)_+^3 \\ \times \frac{\mathcal{F}\Big(e_i^\top\Big)^{(k)} \mathcal{F}\big(\mathcal{U}\big)^{(k)} \mathcal{F}\big(\mathcal{Z}\big)^{(k)^\top} \mathcal{F}\Big(e_i\Big)^{(k)}}{\|\mathcal{F}\Big(e_i^\top\Big)^{(k)} \mathcal{F}\big(\mathcal{U}\big)^{(k)}\|_2} \Big). \end{aligned} \tag{4.32}$$

$$\begin{aligned} &\mathcal{Z} : \nabla^2 Q(\mathcal{U}) : \mathcal{Z} = \\ &12\lambda \sum_{k=1}^{n_3} \sum_{i=1}^{n} \Big(\Big(\|\mathcal{F}\Big(e_i^\top\Big)^{(k)} \mathcal{F}\big(\mathcal{U}\big)^{(k)}\|_2 - \alpha \Big)_+^2 \Big(\frac{\mathcal{F}\Big(e_i^\top\Big)^{(k)} \mathcal{F}\big(\mathcal{U}\big)^{(k)} \mathcal{F}\big(\mathcal{Z}\big)^{(k)^\top} \mathcal{F}\Big(e_i\Big)^{(k)}}{\|\mathcal{F}\Big(e_i^\top\Big)^{(k)} \mathcal{F}\big(\mathcal{U}\big)^{(k)}\|_2} \Big)^2 \Big) \\ &+ 4\lambda \sum_{k=1}^{n_3} \sum_{i=1}^{n} \Big(\Big(\|\mathcal{F}\Big(e_i^\top\Big)^{(k)} \mathcal{F}\big(\mathcal{U}\big)^{(k)}\|_2 - \alpha \Big)_+^3 \\ &\times \frac{\|\mathcal{F}\Big(e_i^\top\Big)^{(k)} \mathcal{F}\big(\mathcal{U}\big)^{(k)}\|_2^2 \, \|\mathcal{F}\Big(e_i^\top\Big)^{(k)} \mathcal{F}\big(\mathcal{Z}\big)^{(k)}\|_2^2 - \Big(\mathcal{F}\Big(e_i^\top\Big)^{(k)} \mathcal{F}\big(\mathcal{U}\big)^{(k)} \mathcal{F}\big(Z\big)^{(k)^\top} \mathcal{F}\Big(e_i\Big)^{(k)}\Big)^2}{\|\mathcal{F}\Big(e_i^\top\Big)^{(k)} \mathcal{F}\big(\mathcal{U}\big)^{(k)}\|_2^3} \Big). \end{aligned} \tag{4.33}$$

Proof: Let

$$\varphi\big(\mathcal{U}\big) = \sum_{k=1}^{n_3} \sum_{i=1}^{n} h_i\Big(\mathcal{F}\big(\mathcal{U}\big)^{(k)} + t\mathcal{F}\big(\mathcal{Z}\big)^{(k)}\Big) - h_i\big(\mathcal{U}\big)$$

where

$$h_i(\mathcal{M}) = \Big(\|\mathcal{F}\Big(e_i^\top\Big)^{(k)} \mathcal{F}\big(\mathcal{M}\big)^{(k)}\|_2 - \alpha \Big)_+^4.$$

We will have to determine the directional derivative of φ in the direction of $\mathcal{F}(\mathcal{Z})^{(k)}$ for $k = 1, \ldots, n_3$. Suppose that $\|e_i^\top * \mathcal{U}\|_F \geqslant \alpha$, so for all sufficiently small t and for any $k = 1, \ldots, n_3$, we have

$$\left(\|\mathcal{F}\Big(e_i^\top\Big)^{(k)} \mathcal{F}\big(\mathcal{U}\big)^{(k)}\|_2 - \alpha \right)_+^4 = \left(\|\mathcal{F}\Big(e_i^\top\Big)^{(k)} \mathcal{F}\big(\mathcal{U}\big)^{(k)}\| - \alpha \right)^4.$$

Hence, we have

$$h_i(\mathcal{U}) = g(f_i(\mathcal{U})) \text{ with } g : x \mapsto x^4$$

as well as

$$f_i(\mathcal{U}) = \|\mathcal{F}\big(e_i^\top\big)^{(k)}\mathcal{F}\big(\mathcal{U}\big)^{(k)}\|_2 - \alpha.$$

$$\partial h_i(\mathcal{U}) = \partial g(f_i(\mathcal{U})) \cdot \partial f_i(\mathcal{U})$$

and

$$\partial^2 h_i(\mathcal{U}) = \partial^2(g(f_i(\mathcal{U}))) \cdot \partial f_i(\mathcal{U}) + \partial g(f_i(\mathcal{U})) \cdot \partial^2 f_i(\mathcal{U})$$

with

$$\partial g(f_i(\mathcal{U})) = 4\Big(\|\mathcal{F}\big(e_i^\top\big)^{(k)}\mathcal{F}\big(\mathcal{U}\big)^{(k)}\|_2 - \alpha\Big)^3,$$

$$\partial f_i\big(\mathcal{U}\big) = \frac{\mathcal{F}\big(e_i^\top\big)^{(k)}\mathcal{F}\big(\mathcal{U}\big)^{(k)}\mathcal{F}\big(\mathcal{Z}\big)^{(k)^\top}\mathcal{F}\big(e_i\big)^{(k)}}{\|\mathcal{F}\big(e_i^\top\big)^{(k)}\mathcal{F}\big(\mathcal{U}\big)^{(k)}\|_2}$$

and

$$\partial^2(g(f_i(\mathcal{U}))) = 12\Big(\|\mathcal{F}\big(e_i^\top\big)^{(k)}\mathcal{F}\big(\mathcal{U}\big)^{(k)}\|_2 - \alpha\Big)^2 \frac{\mathcal{F}\big(e_i^\top\big)^{(k)}\mathcal{F}\big(\mathcal{U}\big)^{(k)}\mathcal{F}\big(\mathcal{Z}\big)^{(k)^\top}\mathcal{F}\big(e_i\big)^{(k)}}{\|\mathcal{F}\big(e_i^\top\big)^{(k)}\mathcal{F}\big(\mathcal{U}\big)^{(k)}\|_2}$$

and thus,

$$\partial^2 f_i(\mathcal{U}) = \frac{\|\mathcal{F}\big(e_i^\top\big)^{(k)}\mathcal{F}\big(\mathcal{U}\big)^{(k)}\|_2^2\|\mathcal{F}\big(e_i^\top\big)^{(k)}\mathcal{F}\big(\mathcal{Z}\big)^{(k)}\|_2^2 - \Big(\mathcal{F}\big(e_i^\top\big)^{(k)}\mathcal{F}\big(\mathcal{U}\big)^{(k)}\mathcal{F}\big(\mathcal{Z}\big)^{(k)^\top}\mathcal{F}\big(e_i\big)^{(k)}\Big)^2}{\|\mathcal{F}\big(e_i^\top\big)^{(k)}\mathcal{F}\big(\mathcal{U}\big)^{(k)}\|_2^3}$$

With this result in hand, the remainder of the proof follows in a straightforward manner. □

Lemma 4.7.14 *There exists an absolute constant c, such that when the probability p satisfies*

$$p > c_1\frac{\mu r\ \log(n)}{n},\ \alpha^2 = \Theta\big(\frac{\mu r\sigma_1^\star}{n}\big)\ \textit{and}\ \lambda = \Theta\big(\frac{n}{\mu r\kappa^\star}\big),$$

we have for any U with $\|\nabla f(\mathcal{U})\|_F \leqslant \varepsilon$ for any polynomial small ε, with probability at least

$$1 - 2nn_3\exp\left\{-pn\Big(\Big(1+\frac{t}{pn}\Big)\ln\Big(1+\frac{t}{pn}\Big) - \frac{t}{pn}\Big)\right\},$$

$$\begin{aligned}\max_{1\leqslant i\leqslant n}\|e_i^\top * \mathcal{U}\|_F^2 &= \max_{1\leqslant i\leqslant n}\|\mathcal{F}(e_i^\top)\cdot\mathcal{F}(\mathcal{U})\|_F^2 \\ &= \max_{1\leqslant i\leqslant n}\sum_{k=1}^{n_3}\|\mathcal{F}\big(e_i^\top\big)^{(k)}\mathcal{F}\big(\mathcal{U}\big)^{(k)}\|_2^2 \leqslant C\ n_3\frac{(\mu r)^{1.5}\kappa^\star\sigma_1^\star}{n}\end{aligned}$$

for some constant positive C.

Proof: We first show that the regulariser forces the tensor $\mathcal{U}$ to have small rows, i.e, prove the Lemma 4.7.14. By Lemma 4.7.13, we know that:

$$\nabla Q(\mathcal{U}) = 4\lambda \sum_{k=1}^{n_3} \sum_{i=1}^{n} \Big(\Big(\|\mathcal{F}\Big(e_i^\top\Big)^{(k)} \mathcal{F}\Big(\mathcal{U}\Big)^{(k)}\|_2 - \alpha \Big)_+^3 \times \frac{\mathcal{F}\Big(e_i\Big)^{(k)} \mathcal{F}\Big(e_i^\top\Big)^{(k)} \mathcal{F}\Big(\mathcal{U}\Big)^{(k)}}{\|\mathcal{F}\Big(e_i^\top\Big)^{(k)} \mathcal{F}\Big(\mathcal{U}\Big)^{(k)}\|_2} \Big). \tag{4.34}$$

Using this formula, we have

$$\begin{aligned} \nabla f\Big(\mathcal{U}\Big) &= \frac{2}{p}\Big(M - M^\star\Big)_\Omega * \mathcal{U} + \nabla Q(\mathcal{U}) \\ &= \frac{2}{p} \sum_{k=1}^{n_3} \Big(\mathcal{F}\Big(\mathcal{M}\Big)^{(k)} - \mathcal{F}\Big(\mathcal{M}^\star\Big)^{(k)}\Big)_\Omega \mathcal{F}\Big(\mathcal{U}\Big)^{(k)} + \nabla Q(\mathcal{U}). \end{aligned} \tag{4.35}$$

Let us study the potential consequence of having $\|e_{i^\star}^\top * \mathcal{U}\|_F \geqslant 2\alpha$. Consider the gradient along $e_{i^\star} * e_{i^\star}^\top * \mathcal{U}$ direction. Since $\|\nabla f(\mathcal{U})\|_F \leqslant \varepsilon$, we have

$$\langle \nabla f(\mathcal{U}), e_{i^\star} * e_{i^\star}^\top * \mathcal{U} \rangle = \langle e_{i^\star}^\top * \nabla f(\mathcal{U}), e_{i^\star}^\top * \mathcal{U} \rangle \leqslant \varepsilon \|e_{i^\star}^\top * \mathcal{U}\|_F.$$

Therefore, with probability larger than

$$1 - 2nn_3 \exp\left\{ -pn\Big(\Big(1 + \frac{t}{pn}\Big) \ln\Big(1 + \frac{t}{pn}\Big) - \frac{t}{pn}\Big) \right\},$$

using equalities (4.34) and (4.35) the followings holds:

$$\varepsilon \|e_{i^\star}^\top * \mathcal{U}\|_F = \varepsilon \|\mathcal{F}\Big(e_{i^\star}^\top\Big) \cdot \mathcal{F}\Big(\mathcal{U}\Big)\|_F = \varepsilon \sum_{k=1}^{n_3} \|\mathcal{F}\Big(e_{i^\star}^\top\Big)^{(k)} \mathcal{F}\Big(\mathcal{U}\Big)^{(k)}\|_2.$$

Now, for any $k = 1, \dots, n_3$, we have

$$
\begin{aligned}
&\varepsilon \|\mathcal{F}\Big(e_{i^\star}^\top\Big)^{(k)} \mathcal{F}\Big(\mathcal{U}\Big)^{(k)}\|_2 \\
&\overset{\#}{\geqslant} 4\lambda\Big(\|\mathcal{F}(e_{i^\star}^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2 - \alpha\Big)_+^3 \|\mathcal{F}(e_{i^\star}^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2 \\
&- \frac{2}{p}\langle \mathcal{F}\Big(e_{i^\star}^\top\Big)^{(k)}\Big(\mathcal{F}(\mathcal{M}^\star)^{(k)}\Big)_\Omega, \mathcal{F}\Big(e_{i^\star}^\top\Big)^{(k)}\Big(\mathcal{F}(\mathcal{U})^{(k)}\mathcal{F}(\mathcal{U}^\top)^{(k)}\Big)_\Omega\rangle \\
&\overset{\#\#}{\geqslant} \frac{\lambda}{2}\|\mathcal{F}(e_{i^\star}^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2^4 \\
&- 2\cdot\frac{1}{\sqrt{p}}\|\mathcal{F}(e_{i^\star}^\top)^{(k)}\Big(\mathcal{F}(\mathcal{M}^\star)^{(k)}\Big)_\Omega\|_2 \cdot \frac{1}{\sqrt{p}}\|\mathcal{F}(e_{i^\star}^\top)^{(k)}\Big(\mathcal{F}(\mathcal{U})^{(k)}\mathcal{F}(\mathcal{U}^\top)^{(k)}\Big)_\Omega\|_2 \\
&\overset{\#\#\#}{=} \frac{\lambda}{2}\|\mathcal{F}(e_{i^\star}^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2^4 \\
&- \frac{2}{\sqrt{p}}\sqrt{\sum_{k=1}^{n_3}\Big\|\mathcal{F}\Big(e_{i^\star}^\top\Big)^{(k)}(\mathcal{F}\Big(\mathcal{M}^\star\Big)^{(k)})_\Omega\Big\|_2^2} \times \frac{1}{\sqrt{p}}\sqrt{\sum_{k=1}^{n_3}\|\mathcal{F}\Big(e_{i^\star}^\top\Big)^{(k)}\Big(\mathcal{F}\Big(\mathcal{U}\Big)^{(k)}\mathcal{F}\Big(\mathcal{U}^\top\Big)^{(k)}\Big)_\Omega\|_2^2} \\
&\overset{\#\#\#\#}{\geqslant} \frac{\lambda}{2}\|\mathcal{F}(e_{i^\star}^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2^4 \\
&- 2\sum_{k=1}^{n_3}\sqrt{1+0.01}\|\mathcal{F}\Big(e_{i^\star}^\top\Big)^{(k)}\mathcal{F}\Big(\mathcal{M}^\star\Big)^{(k)}\|_2 \cdot C\sqrt{n}\sum_{k=1}^{n_3}\|\mathcal{F}\Big(\mathcal{U}\Big)^{(k)}\mathcal{F}\Big(\mathcal{U}^\top\Big)^{(k)}\|_\infty \\
&\overset{\#\#\#\#\#}{\geqslant} \frac{\lambda}{2}\|\mathcal{F}(e_{i^\star}^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2^4 - C\sqrt{\mu r}\sigma_1^\star\sum_{k=1}^{n_3}\|\mathcal{F}\Big(e_{i^\star}^\top\Big)^{(k)}\mathcal{F}\Big(\mathcal{U}\Big)^{(k)}\|_2^2
\end{aligned}
$$

where we used the relation

$$
\begin{aligned}
&\langle \mathcal{F}\Big(e_{i^\star}^\top\Big)^{(k)}\Big(\mathcal{F}\Big(\mathcal{U}\Big)^{(k)}\mathcal{F}\Big(\mathcal{U}^\top\Big)^{(k)}\Big)_\Omega \mathcal{F}\Big(\mathcal{U}\Big)^{(k)}, \mathcal{F}\Big(e_{i^\star}^\top\Big)^{(k)}\mathcal{F}\Big(\mathcal{U}\Big)^{(k)}\rangle \\
&\qquad = \|\mathcal{F}\Big(e_{i^\star}^\top\Big)^{(k)}\Big(\mathcal{F}\Big(\mathcal{U}\Big)^{(k)}\mathcal{F}\Big(\mathcal{U}^\top\Big)^{(k)}\Big)_\Omega\|_2^2 \geqslant 0
\end{aligned}
$$

in (#); the Cauchy-Schwartz inequality in (##); the isometry of the FFT in (###); (4.7.1) and (4.7.5) in (####) and the μ-incoherence of $\mathcal{M}^\star$ in (#####). Therefore, we obtain:

$$
\|e_{i^\star}^\top * \mathcal{U}\|_F^3 \leqslant C n_3 \frac{\sqrt{\mu r}\ \sigma_1^\star}{\lambda}\|e_{i^\star}^\top * \mathcal{U}\|_F + \frac{2\varepsilon}{\lambda}.
$$

By choosing ε sufficiently small we can impose $\Big(\varepsilon/\lambda\Big)^{\frac{2}{3}} \leqslant \dfrac{\sqrt{\mu r}\ \sigma_1^\star}{\lambda}$ and obtain

$$
\max_{1\leqslant i\leqslant n}\|e_i^\top * \mathcal{U}\|_F^2 \leqslant c\max\left\{\alpha^2, n_3\frac{\sqrt{\mu r}\ \sigma_1^\star}{\lambda}\right\}.
$$

Finally, substituting our choice of α^2 and λ, the proof is completed. □

We now show that the Hessian operator satisfies that when $\mathcal{U}$ and $\mathcal{U}^\star$ are not close to each other, the terms involving the Hessian operator $\mathcal{H}$ in Equation (4.27) are significantly negative.

Lemma 4.7.15 *When the probability* $p \geqslant c_1\Big(\frac{\mu^3 r^4 \Big(\kappa^\star\Big)^4 \log n}{n}\Big)$, *by choosing* $\alpha^2 = \Theta\Big(\frac{\mu r \sigma_1^\star}{n}\Big)$ *and* $\lambda = \Theta\Big(\frac{n}{\mu r \kappa^\star}\Big)$ *with probability at least*

$$1 - 2nn_3 \exp\left\{-pn\Big(\Big(1+\frac{t}{pn}\Big)\ln\Big(1+\frac{t}{pn}\Big) - \frac{t}{pn}\Big)\right\},$$

for all $\mathcal{U}$ *with* $\|\nabla f(\mathcal{U})\|_F \leqslant \varepsilon$ *for polynomially small* ε, *we have*

$$\begin{aligned}\Delta * \Delta^\top : \mathcal{H} : \Delta * \Delta^\top - 3(\mathcal{M} - \mathcal{M}^\star) : \mathcal{H} : (\mathcal{M} - \mathcal{M}^\star) \\ \leqslant -0.3 \sum_{k=1}^{n_3} \sigma_r^\star \|\mathcal{F}\Big(\Delta\Big)^{(k)}\|_F^2.\end{aligned}$$

Proof: Introduce

$$\Delta = \mathcal{U} - \mathcal{U}^\star.$$

Note that when Δ is not incoherent, the Hessian will still preserve norm for frontal faces like $\mathcal{F}(\Delta)^{(k)}\mathcal{F}(\mathcal{U}^\top)^{(k)}$, but but it will not necessarily preserve the norm of frontal faces such as $\mathcal{F}(\Delta)^{(k)}\mathcal{F}(\Delta^\top)^{(k)}$. Hence, we use different concentration lemmas in different cases.

First with the choice of α, λ and using Lemma 4.7.14 we know that with probability larger than

$$1 - 2n \exp\left\{-pn\Big(\Big(1+\frac{t}{pn}\Big)\ln\Big(1+\frac{t}{pn}\Big) - \frac{t}{pn}\Big)\right\},$$

the maximum the Euclidean norm of any row of $\mathcal{F}\Big(\mathcal{U}\Big)^{(k)}$ for $k = 1, \ldots, n_3$ is small as well:

$$\max_{1 \leqslant i \leqslant n} \|\mathcal{F}\Big(e_i^\top\Big)^{(k)} \mathcal{F}\Big(\mathcal{U}\Big)^{(k)}\|_2^2 \leqslant C\ \frac{(\mu r)^{1.5} \kappa^\star \sigma_1^\star}{n}.$$

Let us now split the analysis into two cases.

Case 1: $\|\mathcal{F}(\Delta)^{(k)}\|_F^2 \leqslant \sigma_r^\star/4$, for any $k = 1, \ldots, n_3$.

In this case, Δ is small and $\Delta * \Delta^\top$ is small too but $\mathcal{H}$ not preserve norm very well for frontal slides $\mathcal{F}\Big(\Delta\Big)^{(k)}\mathcal{F}\Big(\Delta^\top\Big)^{(k)}$. Using the choice of p and by Lemma 4.7.1, we have

$$\frac{1}{p}\|\mathcal{U}^\star * \Delta^\top\|_\Omega^2 \geqslant \Big(1-\delta\Big)\|\mathcal{U}^\star * \Delta^\top\|_F^2 \geqslant \Big(1-\delta\Big)\sigma_r^\star \|\Delta\|_F^2.$$

On the other hand,

$$\frac{1}{p}\|\Delta * \Delta^\top\|_\Omega^2 = \sum_{k=1}^{n_3} \frac{1}{p}\|\mathcal{F}(\Delta)^{(k)}\mathcal{F}(\Delta^\top)^{(k)}\|_\Omega^2.$$

Using Lemma 4.7.3, for any $k = 1, \ldots, n_3$, we have for some positive constant C:

$$\begin{aligned}\frac{1}{p}\|\mathcal{F}(\Delta)^{(k)}\mathcal{F}(\Delta^\top)^{(k)}\|_\Omega^2 &\leqslant \|\mathcal{F}(\Delta)^{(k)}\|_F^4 + C\sqrt{\frac{n}{p}}\cdot\frac{(\mu r)^{1.5}\kappa^\star\sigma_1^\star}{n}\|\mathcal{F}(\Delta)^{(k)}\|_F^2 \\ &\leqslant \|\mathcal{F}(\Delta)^{(k)}\|_F^4 + \frac{\sigma_r^\star}{4}\|\mathcal{F}(\Delta)^{(k)}\|_F^2 \\ &\leqslant \frac{\sigma_r^\star}{2}\|\mathcal{F}(\Delta)^{(k)}\|_F^2.\end{aligned}$$

Using these facts, we obtain

$$\begin{aligned}&\Delta * \Delta : \mathcal{H} : \Delta * \Delta^\top - 3\Big(\mathcal{M} - \mathcal{M}^\star\Big) : \mathcal{H} : \Big(\mathcal{M} - \mathcal{M}^\star\Big) \\ =&\sum_{k=1}^{n_3} \mathcal{F}(\Delta)^{(k)}\mathcal{F}(\Delta^\top)^{(k)} : \mathcal{F}(\mathcal{H})^{(k)} : \mathcal{F}(\Delta)^{(k)}\mathcal{F}(\Delta^\top)^{(k)} \\ &-3\Big(\mathcal{F}(\mathcal{M})^{(k)} - \mathcal{F}(\mathcal{M}^\star)^{(k)}\Big) : \mathcal{F}(\mathcal{H})^{(k)} : \Big(\mathcal{F}(\mathcal{M})^{(k)} - \mathcal{F}(\mathcal{M}^\star)^{(k)}\Big) \\ =&\sum_{k=1}^{n_3} \mathcal{F}(\Delta)^{(k)}\mathcal{F}(\Delta^\top)^{(k)} : \mathcal{F}(\mathcal{H})^{(k)} : \mathcal{F}(\Delta)^{(k)}\mathcal{F}(\Delta^\top)^{(k)} \\ &-3\Big(\mathcal{F}(\mathcal{U}^\star)^{(k)}\mathcal{F}(\Delta^\top)^{(k)} + \mathcal{F}(\Delta)^{(k)}\mathcal{F}(\mathcal{U}^{\star\top})^{(k)} + \mathcal{F}(\Delta)^{(k)}\mathcal{F}(\Delta^\top)^{(k)}\Big) : \mathcal{F}(\mathcal{H})^{(k)} : \\ &\Big(\mathcal{F}(\mathcal{U}^\star)^{(k)}\mathcal{F}(\Delta^\top)^{(k)} + \mathcal{F}(\Delta)^{(k)}\mathcal{F}(\mathcal{U}^{\star\top})^{(k)} + \mathcal{F}(\Delta)^{(k)}\mathcal{F}(\Delta^\top)^{(k)}\Big) \\ \leqslant&\sum_{k=1}^{n_3} -12\ \mathcal{F}(\mathcal{U}^\star)^{(k)}\mathcal{F}(\Delta^\top)^{(k)} : \mathcal{F}(\mathcal{H})^{(k)} : \mathcal{F}(\Delta)^{(k)}\mathcal{F}(\Delta^\top)^{(k)} \\ &-12\ \mathcal{F}(\mathcal{U}^\star)^{(k)}\mathcal{F}(\Delta^\top)^{(k)} : \mathcal{F}(\mathcal{H})^{(k)} : \mathcal{F}(\mathcal{U}^\star)^{(k)}\mathcal{F}(\Delta^\top)^{(k)} \\ \leqslant&\sum_{k=1}^{n_3} -\frac{12}{p}\Big(\|\mathcal{F}(\mathcal{U}^\star)^{(k)}\mathcal{F}(\Delta^\top)^{(k)}\|_\Omega^2 - \|\mathcal{F}(\mathcal{U}^\star)^{(k)}\mathcal{F}(\Delta^\top)^{(k)}\|_\Omega\|\mathcal{F}(\Delta)^{(k)}\mathcal{F}(\Delta^\top)^{(k)}\|_\Omega\Big) \\ \leqslant&\sum_{k=1}^{n_3} -12\sqrt{1-\delta}\Big(\sqrt{1-\delta} - \sqrt{2/3}\Big)\sigma_r^\star\|\mathcal{F}(\Delta)^{(k)}\|_F^2\end{aligned}$$

where we use the fact

$$\frac{1}{p}\ \|\mathcal{F}(\Delta)^{(k)}\mathcal{F}(\Delta^\top)^{(k)}\|_\Omega^2 \leqslant \frac{\sigma_r^\star}{2}\ \|\mathcal{F}(\Delta)^{(k)}\|_F^2$$

and

$$-\frac{1}{p}\ \|\mathcal{F}(U^\star)^{(k)}\mathcal{F}(\Delta)^{(k)^\top}\|_\Omega^2 \leqslant -\Big(1-\delta\Big)\sigma_r^\star\ \|\mathcal{F}(\Delta)^{(k)}\|_F^2.$$

Thus, taking $p \geqslant c_1\frac{\mu^3 r^4\Big(\kappa^\star\Big)^4\log n}{n}$, we get

$$\Delta * \Delta : \mathcal{H} : \Delta * \Delta^\top - 3\Big(\mathcal{M} - \mathcal{M}^\star\Big) : \mathcal{H} : \Big(\mathcal{M} - \mathcal{M}^\star\Big) \leqslant \sum_{k=1}^{n_3} -1.2\sigma_r^\star\|\mathcal{F}(\Delta)^{(k)}\|_F^2.$$

Case 2: $\|\mathcal{F}(\Delta)^{(k)}\|_F^2 \geqslant \frac{\sigma_r^\star}{4}$, for any $k = 1, \ldots, n_3$.

Using Lemma (4.7.4) with high probability and with the choice of p that we have just made, we have

$$\begin{aligned}
\frac{1}{p}\|\Delta * \Delta^\top\|_\Omega^2 &= \sum_{k=1}^{n_3} \frac{1}{p}\|\mathcal{F}(\Delta)^{(k)}\mathcal{F}(\Delta^\top)^{(k)}\|_\Omega^2 \\
&\leqslant \sum_{k=1}^{n_3} \|\mathcal{F}(\Delta)^{(k)}\mathcal{F}(\Delta^\top)^{(k)}\|_F^2 \\
&\quad + C\Big(\frac{nr\log(n)}{p}\|\mathcal{F}(\Delta)^{(k)}\mathcal{F}(\Delta^\top)^{(k)}\|_\infty^2 \\
&\quad + \sqrt{\frac{nr\log(n)}{p}}\|\mathcal{F}(\Delta)^{(k)}\mathcal{F}(\Delta^\top)^{(k)}\|_F\|\mathcal{F}(\Delta)^{(k)}\mathcal{F}(\Delta^\top)^{(k)}\|_\infty\Big) \\
&\leqslant \sum_{k=1}^{n_3} \|\mathcal{F}(\Delta)^{(k)}\mathcal{F}(\Delta^\top)^{(k)}\|_F^2 \\
&\quad + C\Big(\frac{nr\log(n)}{p}\cdot\frac{\big(\mu r\big)^3\big(\kappa^\star\sigma_1^\star\big)^2}{n^2} + \sqrt{\frac{nr\log(n)}{p}\cdot\frac{\big(\mu r\big)^3\big(\kappa^\star\sigma_1^\star\big)^2}{n^2}}\|\mathcal{F}(\Delta)^{(k)}\|_F^2\Big) \\
&\leqslant \sum_{k=1}^{n_3} \|\mathcal{F}\Big(\Delta\Big)^{(k)}\mathcal{F}(\Delta^\top)^{(k)}\|_F^2 + \frac{\big(\sigma_r^\star\big)^2}{80} + \frac{\sigma_r^\star}{20}\|\mathcal{F}(\Delta)^{(k)}\|_F^2 \\
&\leqslant \sum_{k=1}^{n_3} \|\mathcal{F}(\Delta)^{(k)}\mathcal{F}(\Delta^\top)^{(k)}\|_F^2 + 0.1\sigma_r^\star\|\mathcal{F}(\Delta)^{(k)}\|_F^2.
\end{aligned}$$

In the second inequality, we used

$$\|\mathcal{F}(\Delta)^{(k)}\mathcal{F}(\Delta^\top)^{(k)}\|_\infty^2 \leqslant C\frac{\big(\mu r\big)^3\big(\kappa^\star\sigma_1^\star\big)^2}{n^2}$$

for some positive constant C. In the third inequality, we use for some positive constant c_1

$$p \geqslant c_1\frac{\mu^3 r^4\big(\kappa^\star\big)^4\log n}{n} \text{ and } \kappa = \sigma_1^\star/\sigma_r^\star.$$

Again, using Lemma 4.7.4, we have that, with high probability,

$$
\begin{aligned}
\frac{1}{p}\|\mathcal{M}-\mathcal{M}^\star\|_\Omega^2 &= \sum_{k=1}^{n_3}\frac{1}{p}\|\mathcal{F}(\mathcal{M})^{(k)}-\mathcal{F}(\mathcal{M}^\star)^{(k)}\|_\Omega^2 \\
&\geqslant \sum_{k=1}^{n_3}\|\mathcal{F}(\mathcal{M})^{(k)}-\mathcal{F}(\mathcal{M}^\star)^{(k)}\|_F^2 \\
&+ C\Big(\frac{nr\log(n)}{p}\|\mathcal{F}(\mathcal{M})^{(k)}-\mathcal{F}(\mathcal{M}^\star)^{(k)}\|_\infty^2 \\
&+ \sqrt{\frac{nr\log(n)}{p}}\|\mathcal{F}(\mathcal{M})^{(k)}-\mathcal{F}(\mathcal{M}^\star)^{(k)}\|_F\|\mathcal{F}(\mathcal{M})^{(k)}-\mathcal{F}(\mathcal{M}^\star)^{(k)}\|_\infty\Big) \\
&\geqslant \sum_{k=1}^{n_3}\|\mathcal{F}(M)^{(k)}-\mathcal{F}(\mathcal{M}^\star)^{(k)}\|_F^2 \\
&+ C\Big(\frac{nr\log(n)}{p}\cdot\frac{\Big(\mu r\Big)^3\Big(\kappa^\star\sigma_1^\star\Big)^2}{n^2} \\
&+ \sqrt{\frac{nr\log(n)}{p}\cdot\frac{\Big(\mu r\Big)^3\Big(\kappa^\star\sigma_1^\star\Big)^2}{n^2}}\|\mathcal{F}(\mathcal{M})^{(k)}-\mathcal{F}(\mathcal{M}^\star)^{(k)}\|_F\Big) \\
&\geqslant \sum_{k=1}^{n_3}\|\mathcal{F}(\mathcal{M})^{(k)}-\mathcal{F}(\mathcal{M}^\star)^{(k)}\|_F^2-\frac{\Big(\sigma_r^\star\Big)^2}{80}-\frac{\sigma_r^\star}{20}\|\mathcal{F}(\mathcal{M})^{(k)}-\mathcal{F}(\mathcal{M}^\star)^{(k)}\|_F \\
&\geqslant \sum_{k=1}^{n_3}0.95\ \|\mathcal{F}(\mathcal{M})^{(k)}-\mathcal{F}(\mathcal{M}^\star)^{(k)}\|_F^2-0.1\ \sigma_r^\star\|\mathcal{F}(\Delta)^{(k)}\|_F^2.
\end{aligned}
$$

The third inequality, again we use $p \geqslant c_1\Big(\frac{\mu^3 r^4\Big(\kappa^\star\Big)^4\log n}{n}\Big)$ and $\kappa = \sigma_1^\star/\sigma_r^\star$. This facts implies

$$
\begin{aligned}
&\Delta * \Delta^\top : \mathcal{H} : \Delta * \Delta^\top - 3(\mathcal{M}-\mathcal{M}^\star) : \mathcal{H} : (\mathcal{M}-\mathcal{M}^\star) \\
&= \sum_{k=1}^{n_3}\frac{1}{p}\Big(\|\mathcal{F}(\Delta)^{(k)}\mathcal{F}(\Delta^\top)^{(k)}\|_\Omega^2-3\ \|\mathcal{F}(\mathcal{M})^{(k)}-\mathcal{F}(\mathcal{M}^\star)^{(k)}\|_\Omega^2\Big) \\
&\leqslant \sum_{k=1}^{n_3}\|\mathcal{F}(\Delta)^{(k)}\mathcal{F}(\Delta^\top)^{(k)}\|_F^2+0.1\sigma_r^\star\|\mathcal{F}(\Delta)^{(k)}\|_F^2 \\
&- 3\Big(0.95\|\mathcal{F}(\mathcal{M})^{(k)}-\mathcal{F}(\mathcal{M}^\star)^{(k)}\|_F^2-0.1\,\sigma_r^\star\|\mathcal{F}(\Delta)^{(k)}\|_F^2\Big) \\
&\leqslant \sum_{k=1}^{n_3}-0.85\|\mathcal{F}(\mathcal{M})^{(k)}-\mathcal{F}(\mathcal{M}^\star)^{(k)}\|_F^2+0.4\,\sigma_r^\star\|\mathcal{F}(\Delta)^{(k)}\|_F^2 \\
&\leqslant \sum_{k=1}^{n_3}-0.3\,\sigma_r^\star\|\mathcal{F}(\Delta)^{(k)}\|_F^2.
\end{aligned}
$$

The two last inequalities, we use the two bounds of 4.7.9. □

Now, we need to bound the terms with the regularizer in (4.27).

Lemma 4.7.16 *By choosing* $\alpha^2 = \Theta\Big(\frac{\mu r \sigma_1^\star}{n}\Big)$ *and* $\lambda\alpha^2 \leqslant C\ \sigma_r^\star$ *for some positive constant* C*, we have:*

$$\frac{1}{4}\Big[\Delta : \nabla^2 Q\Big(\mathcal{U}\Big) : \Delta - 4\langle \nabla Q\Big(\mathcal{U}\Big), \Delta\rangle\Big] \leqslant 0.1\sigma_r^\star \sum_{k=1}^{n_3} \|\mathcal{F}\Big(\Delta\Big)^{(k)}\|_F^2.$$

Proof: We know that:

$$\langle \nabla Q(\mathcal{U}), \mathcal{Z}\rangle = 4\lambda \sum_{k=1}^{n_3}\sum_{i=1}^{n} \Big(\|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2 - \alpha\Big)_+^3 \frac{\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\mathcal{F}(\mathcal{Z}^\top)^{(k)}\mathcal{F}(e_i)^{(k)}}{\|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2}$$

$$\begin{aligned}\mathcal{Z} : \nabla^2 Q(\mathcal{U}) : \mathcal{Z} &= 12\lambda \sum_{k=1}^{n_3}\sum_{i=1}^{n} \Big(\|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2 - \alpha\Big)_+^2 \Big(\frac{\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\mathcal{F}(\mathcal{Z}^\top)^{(k)}}{\|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2}\Big)^2 \\ &+ 4\lambda \sum_{k=1}^{n_3}\sum_{i=1}^{n} \Big(\|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2 - \alpha\Big)_+^3 \\ &\times \frac{\|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2^2 \|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{Z})^{(k)}\|_2^2 - \Big(\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\mathcal{F}(\mathcal{Z}^\top)^{(k)}\mathcal{F}(e_i)^{(k)}\Big)^2}{\|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2^3}.\end{aligned}$$

Using this facts with $\mathcal{Z} = \Delta = \mathcal{U} - \mathcal{U}^\star * \mathcal{R}$, we have:

$$\begin{aligned}&\frac{1}{4}\Big[\Delta : \nabla^2 Q\Big(\mathcal{U}\Big) : \Delta - 4\langle \nabla Q\Big(\mathcal{U}\Big), \Delta\rangle\Big] \\ &= \underbrace{\lambda \sum_{k=1}^{n_3}\sum_{i=1}^{n} \frac{\Big(\|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2 - \alpha\Big)_+^3 \times \|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2^2 \|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\Delta)^{(k)}\|_2^2 - \Big(\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\mathcal{F}(\Delta^\top)^{(k)}\mathcal{F}(e_i)^{(k)}\Big)^2}{\|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2^3}}_{=A_1} \\ &+ \underbrace{3\lambda \sum_{k=1}^{n_3}\sum_{i=1}^{n} \Big(\|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2 - \alpha\Big)_+^2 \Big(\frac{\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\mathcal{F}(\Delta^\top)^{(k)}\mathcal{F}(e_i)^{(k)}}{\|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2}\Big)^2}_{=A_2} \\ &- \underbrace{4\lambda \sum_{k=1}^{n_3}\sum_{i=1}^{n} \Big(\|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2 - \alpha\Big)_+^3 \frac{\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\mathcal{F}(\Delta^\top)^{(k)}\mathcal{F}(e_i)^{(k)}}{\|\mathcal{F}(e_i)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2}}_{=A_3}.\end{aligned}$$

Furthermore, using the incoherence property of $\mathcal{M}^\star$, we have for any $k = 1, \ldots, n_3$

$$\begin{aligned}\|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)} - \mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\Delta)^{(k)}\|_2 &= \|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U}^\star)^{(k)}\mathcal{F}(\mathcal{R})^{(k)}\|_2 \\ &= \|\mathcal{F}(e_i^\top)\mathcal{F}(\mathcal{U}^\star)^{(k)}\|_2 \\ &\leqslant \sqrt{\frac{\mu r \sigma_1^\star}{n}}.\end{aligned}$$

By choosing $\alpha > C_1\sqrt{\frac{\mu r \sigma_1^\star}{n}}$ for some large constant C_1 and when $\|\mathcal{F}(e_i^\top)\mathcal{F}(\mathcal{U})^{(k)}\|_2 - \alpha > 0$, we have for any $k = 1, \ldots, n_3$

$$\begin{aligned}
&\mathcal{F}\Big(e_i^\top\Big)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\mathcal{F}(\Delta^\top)^{(k)}\mathcal{F}(e_i)^{(k)} \\
=&\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\mathcal{F}(\mathcal{U}^\top)^{(k)}\mathcal{F}(e_i)^{(k)} - \mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\mathcal{F}(\mathcal{R}^\top)^{(k)}\mathcal{F}(\mathcal{U}^{\star\top})^{(k)}\mathcal{F}(e_i)^{(k)} \\
\geqslant&\|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2^2 - \|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2\|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U}^\star)^{(k)}\|_2 \\
\geqslant&\Big(1 - \frac{1}{C_1}\Big)\|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2^2.
\end{aligned}$$

The last inequality, we use the fact

$$\|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U}^\star)^{(k)}\|_2 < \frac{\alpha}{C_1}$$

and

$$\alpha < \|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2 \;\Rightarrow \|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U}^\star)^{(k)}\|_2 < \frac{1}{C_1}\,\|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2.$$

Further, we have

$$\begin{aligned}
&\|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2\|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\Delta)^{(k)}\|_2 \\
\leqslant&\|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2\Big(\|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2 + \|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U}^\star)^{(k)}\|_2\Big) \\
\leqslant&\Big(1 + \frac{1}{C_1}\Big)\|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2^2.
\end{aligned}$$

Now, we need to bound the summation $A_1 + A_2 + A_3$ as follows to get a bound $A_1 + 0.1A_3$ and $A_2 + 0.9A_3$. Thus,

$$\begin{aligned}
A_1 + 0.1A_3 =& \lambda\sum_{k=1}^{n_3}\sum_{i=1}^{n}\Bigg(\frac{\begin{array}{c}\Big(\|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2 - \alpha\Big)^3 \\ \times \\ \|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2^2\|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\Delta)^{(k)}\|_2^2 \\ -\Big(\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\mathcal{F}(\Delta^\top)^{(k)}\mathcal{F}(e_i)^{(k)}\Big)^2\end{array}}{\|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2^3}\Bigg) \\
&- 0.4\lambda\sum_{k=1}^{n_3}\sum_{i=1}^{n}\Big(\|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2 - \alpha\Big)^3\frac{\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\mathcal{F}(\Delta^\top)^{(k)}\mathcal{F}(e_i)^{(k)}}{\|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|} \\
\leqslant& \lambda\sum_{k=1}^{n_3}\sum_{i=1}^{n}\Big(\|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2 - \alpha\Big)^3\|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2 \\
&\times\Bigg[\Big(1 + \frac{1}{C_1}\Big)^2 - \Big(1 - \frac{1}{C_1}\Big)^2 - 0.4\Big(1 - \frac{1}{C_1}\Big)\Bigg] \\
<& 0.
\end{aligned}$$

Moreover,

$$
\begin{aligned}
&A_2 + 0.9A_3 \\
&= 3\lambda \sum_{k=1}^{n_3} \sum_{i=1}^{n} \Big(\|\mathcal{F}(e_i^\top)^{(k)} \mathcal{F}(\mathcal{U})^{(k)}\|_2 - \alpha \Big)^2 \Big(\frac{\mathcal{F}(e_i^\top)^{(k)} \mathcal{F}(\mathcal{U})^{(k)} \mathcal{F}(\Delta^\top)^{(k)} \mathcal{F}(e_i)^{(k)}}{\|\mathcal{F}(e_i^\top)^{(k)} \mathcal{F}(\mathcal{U})^{(k)}\|_2} \Big)^2 \\
&- 3.6\lambda \sum_{k=1}^{n_3} \sum_{i=1}^{n} \Big(\|\mathcal{F}(e_i^\top)^{(k)} \mathcal{F}(\mathcal{U})^{(k)}\|_2 - \alpha \Big)^3 \frac{\mathcal{F}(e_i^\top)^{(k)} \mathcal{F}(\mathcal{U})^{(k)} \mathcal{F}(\Delta^\top)^{(k)} \mathcal{F}(e_i)^{(k)}}{\|\mathcal{F}(e_i^\top)^{(k)} \mathcal{F}(\mathcal{U})^{(k)}\|_2}
\end{aligned}
$$

Denote i-the summand of the frontal faces of $A_2 + 0.9A_3$ as $A_2 + 0.9A_3 = \sum_{k=1}^{n_3} \sum_{i=1}^{n} B_i^k$, with $B_i = A_2^{(i)} + 0.9\, A_3^{(i)}$.

Case 1: for i such that $\|\mathcal{F}\Big(e_i^\top\Big)^{(k)} \mathcal{F}\Big(U\Big)^{(k)}\|_2 \geqslant 9\alpha$ and $C_1 \geqslant 100$, we have:

$$
\begin{aligned}
B_i^k &= 3\lambda \Big(\|\mathcal{F}(e_i^\top)^{(k)} \mathcal{F}(\mathcal{U})^{(k)}\|_2 - \alpha \Big)^2 \frac{\mathcal{F}(e_i^\top)^{(k)} \mathcal{F}(\mathcal{U})^{(k)} \mathcal{F}(\Delta^\top)^{(k)} \mathcal{F}(e_i)^{(k)}}{\|\mathcal{F}(e_i^\top)^{(k)} \mathcal{F}(\mathcal{U})^{(k)}\|_2} \\
&\times \left[\frac{\mathcal{F}(e_i^\top)^{(k)} \mathcal{F}(\mathcal{U})^{(k)} \mathcal{F}(\Delta^\top)^{(k)} \mathcal{F}(e_i)^{(k)}}{\|\mathcal{F}(e_i^\top)^{(k)} \mathcal{F}(\mathcal{U})^{(k)}\|_2} - 1.2 \Big(\|\mathcal{F}(e_i^\top)^{(k)} \mathcal{F}(\mathcal{U})^{(k)}\|_2 - \alpha \Big) \right] \\
&\leqslant 3\lambda \Big(\|\mathcal{F}(e_i^\top)^{(k)} \mathcal{F}(\mathcal{U})^{(k)}\|_2 - \alpha \Big)^2 \frac{\mathcal{F}(e_i^\top)^{(k)} \mathcal{F}(\mathcal{U})^{(k)} \mathcal{F}(\Delta^\top)^{(k)} \mathcal{F}(e_i)^{(k)}}{\|\mathcal{F}(e_i^\top)^{(k)} \mathcal{F}(\mathcal{U})^{(k)}\|_2} \\
&\times \left[\Big(1 + \frac{1}{C_1}\Big) \|\mathcal{F}(e_i^\top)^{(k)} \mathcal{F}(\mathcal{U})^{(k)}\|_2 - 1.2 \Big(\|\mathcal{F}(e_i^\top)^{(k)} \mathcal{F}(\mathcal{U})^{(k)}\|_2 - \alpha \Big) \right] \\
&\leqslant 0.
\end{aligned}
$$

Because:

- $\Big(\|\mathcal{F}(e_i^\top)^{(k)} \mathcal{F}(\mathcal{U})^{(k)}\|_2 - \alpha \Big)^2 > 0$
- $\dfrac{\mathcal{F}(e_i^\top)^{(k)} \mathcal{F}(\mathcal{U})^{(k)} \mathcal{F}(\Delta^\top)^{(k)} \mathcal{F}(e_i)^{(k)}}{\|\mathcal{F}(e_i^\top)^{(k)} \mathcal{F}(\mathcal{U})^{(k)}\|_2} \leqslant \Big(1 + \dfrac{1}{C_1}\Big) \|\mathcal{F}(e_i^\top)^{(k)} \mathcal{F}(\mathcal{U})^{(k)}\|_2 \geqslant 0$
- $\left[\Big(1 + \dfrac{1}{C_1}\Big) \|\mathcal{F}(e_i^\top)^{(k)} \mathcal{F}(\mathcal{U})^{(k)}\|_2 - 1.2 \Big(\|\mathcal{F}(e_i^\top)^{(k)} \mathcal{F}(\mathcal{U})^{(k)}\|_2 - \alpha \Big) \right] \leqslant 0$

Case 2: for i such that $\alpha < \|\mathcal{F}(e_i^\top)^{(k)} \mathcal{F}(\mathcal{U})^{(k)}\|_2 < 9\alpha$, we call this set

$I = \{i|\, \alpha < \|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}\big(\mathcal{U}\big)^{(k)}\|_2 < 9\alpha\}$ and we have for each frontal face:

$$\begin{aligned}
\sum_{i\in I} B_i^{(k)} &\leqslant 3\lambda \sum_{i\in I} \Big(\|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2 - \alpha\Big)_+^2 \frac{\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\mathcal{F}(\Delta^\top)^{(k)}\mathcal{F}(e_i)^{(k)}}{\|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2} \\
&\times \bigg[\Big(1+\frac{1}{C_1}\Big)\|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2 - 1.2\Big(\|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2 - \alpha\Big)_+\bigg] \\
&\leqslant 3\lambda \sum_{i\in I} \Big(\|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2 - \alpha\Big)_+^2 \frac{\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\mathcal{F}(\Delta^\top)^{(k)}\mathcal{F}(e_i)^{(k)}}{\|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2} \\
&\times \Big(1+\frac{1}{C_1}\Big)\|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2 \\
&\leqslant 3\lambda \sum_{i\in I} \Big(\|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2 - \alpha\Big)_+^2 \Big(1+\frac{1}{C_1}\Big)^2 \|\mathcal{F}(e_i^\top)^{(k)}\mathcal{F}(U)^{(k)}\|_2^2 \\
&\leqslant 3\lambda\ |I|\ 64\ \alpha^2.\Big(1+\frac{1}{C_1}\Big)^2\ 81\ \alpha^2 \\
&\leqslant 3\ 10^4\ |I|\ \lambda\ \alpha^4.
\end{aligned}$$

In sum, we obtain:

$$\frac{1}{4}\Big[\Delta : \nabla^2 Q\big(\mathcal{U}\big) : \Delta - 4\langle \nabla Q\big(\mathcal{U}\big), \Delta\rangle\Big] \leqslant c_2\, \lambda\, |I|\alpha^4$$

for some large constant c_2. Finally, remains to determine with the property of the set I on each front face:

$$\begin{aligned}
\sigma_r^\star \|\mathcal{F}(\Delta)^{(k)}\|_F^2 =& \sigma_r^\star \|\mathcal{F}\big(\mathcal{U}\big)^{(k)} - \mathcal{F}\big(\mathcal{U}^\star\big)^{(k)}\mathcal{F}\big(\mathcal{R}\big)^{(k)}\|_F^2 \\
&= \sigma_r^\star\Big(\|\mathcal{F}\big(\mathcal{U}\big)^{(k)}\|_F^2 + \|\mathcal{F}\big(\mathcal{U}^\star\big)^{(k)}\mathcal{F}\big(\mathcal{R}\big)^{(k)}\|_F^2 \\
&\quad - 2\langle \mathcal{F}\big(\mathcal{U}\big)^{(k)}, \mathcal{F}\big(\mathcal{U}^\star\big)^{(k)}\mathcal{F}\big(\mathcal{R}\big)^{(k)}\rangle\Big) \\
&\geqslant \sigma_r^\star \|\mathcal{F}\big(\mathcal{U}\big)^{(k)}\|_F^2 \\
&= \sigma_r^\star \sum_{i\in I} \|\mathcal{F}\big(e_i^\top\big)^{(k)}\mathcal{F}\big(\mathcal{U}\big)^{(k)}\|_2^2 \\
&\geqslant \sigma_r^\star \alpha^2 |I|.
\end{aligned}$$

Therefore, as long as $\lambda\alpha^2 \leqslant \sigma_r^\star/C_2$ for some large constant C_2 (which is satisfied by our choice of λ) we obtain:

$$\frac{1}{4}\Big[\Delta : \nabla^2 Q\big(\mathcal{U}\big) : \Delta - 4\langle \nabla Q\big(U\big), \Delta\rangle\Big] \leqslant 0.1 \sum_{k=1}^{n_3} \sigma_r^\star \|\mathcal{F}(\Delta)^{(k)}\|_F^2.$$

□

Combing these lemmas, we are now ready to prove the main theorem for symmetric tensor completion.

Theorem 4.7.17 *Assume that μ and r satisfy the following inequality*

$$p \geqslant c_1 \frac{\mu^3 r^4 \big(\kappa^\star\big)^4 \log(n)}{n},$$

for some positive constant and choose

$$\alpha^2 = \Theta\big(\frac{\mu r \sigma_1^\star}{n}\big) \text{ and } \lambda = \Theta\big(\frac{n}{\mu r \kappa^\star}\big).$$

Then with probability at least

$$1 - 2nn_3 \exp\left\{-pn\Big(\big(1+\frac{t}{pn}\big)\ln\big(1+\frac{t}{pn}\big) - \frac{t}{pn}\Big)\right\},$$

we have

- *All local minima of* (4.31) *satisfy* $\mathcal{U} * \mathcal{U}^\top = \mathcal{M}^\star$*;*
- *the function is* $\big(\varepsilon, \Omega\big(\sigma_r^\star\big), O\big(\frac{\varepsilon}{\sigma_r^\star}\big)\big)$*-strict saddle for polynomially small* ε*.*

Proof: We know by 4.7.15:

$$\Delta * \Delta^\top : \mathcal{H} : \Delta * \Delta^\top - 3(\mathcal{M} - \mathcal{M}^\star) : \mathcal{H} : (\mathcal{M} - \mathcal{M}^\star) \leqslant -0.3 \sum_{k=1}^{n_3} \sigma_r^\star \|\mathcal{F}(\Delta)^{(k)}\|_F^2.$$

Further, by 4.7.16, we have:

$$\frac{1}{4}\big[\Delta : \nabla^2 Q\big(\mathcal{U}\big) : \Delta - 4\langle \nabla Q\big(\mathcal{U}\big), \Delta\rangle\big] \leqslant 0.1\sigma_r^\star \sum_{k=1}^{n_3} \|\mathcal{F}(\Delta)^{(k)}\|_F^2.$$

Using this facts, for any $\mathcal{U}$ with small gradient satisfying $\|\nabla f\big(\mathcal{U}\big)\|_F \leqslant \varepsilon$, we have

$$\Delta : \nabla^2 f\big(\mathcal{U}\big) : \Delta \leqslant -0.2 \sum_{k=1}^{n_3} \sigma_r^\star \|\mathcal{F}(\Delta)^{(k)}\|_F^2 + 4\,\varepsilon \|\Delta\|_F.$$

That is, if $\mathcal{U}$ is not close to $\mathcal{U}^\star$, that is, $\|\Delta\|_F \geqslant \dfrac{40\varepsilon}{\sigma_r^\star}$, we have

$$\begin{aligned}
\Delta : \nabla^2 f\big(\mathcal{U}\big) : \Delta &\leqslant -0.2 \sum_{k=1}^{n_3} \sigma_r^\star \|\mathcal{F}(\Delta)^{(k)}\|_F^2 + 0.1\sigma_r^\star \|\Delta\|_F \\
&\leqslant -0.2 \sum_{k=1}^{n_3} \sigma_r^\star \|\mathcal{F}(\Delta)^{(k)}\|_F^2 + 0.1\sigma_r^\star \sqrt{\sum_{k=1}^{n_3} \|\mathcal{F}(\Delta)^{(k)}\|_F^2} \\
&\leqslant -0.1 \sum_{k=1}^{n_3} \sigma_r^\star \|\mathcal{F}(\Delta)^{(k)}\|_F^2.
\end{aligned}$$

This proves $\big(\varepsilon, 0.1\sigma_r^\star, \frac{40\varepsilon}{\sigma_r^\star}\big)$-strict saddle property. By taking $\varepsilon = 0$, then all stationary points with $\|\Delta\|_F \neq 0$ are saddle points. This means all local minima are global minima (satisfying $\mathcal{U} * \mathcal{U}^\top = \mathcal{M}^\star$), which finishes the proof. □

4.7.3 Proof of Theorem 4.4.1

The proof is split into two steps.

4.7.3.1 Study of the Hessian

Furthermore, we have $Q\big(\mathcal{W}\big) = Q_1\big(\mathcal{U}\big) + Q_2\big(\mathcal{V}\big)$.

Lemma 4.7.18 *Let Δ, $\mathcal{N}$, $\mathcal{N}^\star$ be defined as in Definition **??**. Then, for any $\mathcal{W} \in \mathbb{R}^{(n_1+n_2)\times r\times n_3}$, the Hessian of the objective (4.18) satisfies:*

$$\begin{aligned}\Delta : \nabla^2 f(\mathcal{W}) : \Delta \leqslant \Delta * \Delta^\top : \mathcal{H} : \Delta * \Delta^\top - 3(\mathcal{N} - \mathcal{N}^\star) : \mathcal{H} : (\mathcal{N} - \mathcal{N}^\star) \\ +4\langle \nabla f(\mathcal{W}), \Delta\rangle + [\Delta : \nabla^2 Q(\mathcal{W}) : \Delta - 4\langle \nabla Q(\mathcal{W}), \Delta\rangle]\end{aligned} \tag{4.36}$$

where

$$\mathcal{H} = 4\mathcal{H}_1 + \mathcal{G}.$$

Further, if $\mathcal{H}_0$ satisfies

$$\mathcal{M} : \mathcal{H}_0 : \mathcal{M} \in \big(1 \pm \delta\big)\|\mathcal{M}\|_F^2$$

for some tensor $\mathcal{M} = \mathcal{U} * \mathcal{V}^\top$, let $\mathcal{W}$ and $\mathcal{N}$ be defined as in (4.17), then

$$\mathcal{N} : \mathcal{H} : \mathcal{N} \in \big(1 \pm 2\delta\big)\|\mathcal{N}\|_F^2.$$

Proof: We know that the objective function with $\mathcal{N} = \mathcal{W} * \mathcal{W}^\top$ is:

$$\begin{aligned} f\big(\mathcal{W}\big) &= \frac{1}{2}\Big[\big(\mathcal{N} - \mathcal{N}^\star\big) : 4\mathcal{H}_1 : \big(\mathcal{N} - \mathcal{N}^\star\big) + \mathcal{N} : \mathcal{G} : \mathcal{N}\Big] + Q\big(\mathcal{W}\big) \\ &= \frac{1}{2}\sum_{k=1}^{n_3}\left\{\begin{array}{c} \big(\mathcal{F}\big(\mathcal{W}\big)^{(k)}\mathcal{F}\big(\mathcal{W}^\top\big)^{(k)} - \mathcal{F}\big(\mathcal{N}^\star\big)^{(k)}\big) \\ : 4\mathcal{F}\big(\mathcal{H}_1\big)^{(k)} : \\ \big(\mathcal{F}\big(\mathcal{W}\big)^{(k)}\mathcal{F}\big(\mathcal{W}^\top\big)^{(k)} - \mathcal{F}\big(\mathcal{N}^\star\big)^{(k)}\big) \end{array}\right\} \\ &\quad + \mathcal{F}\big(\mathcal{W}\big)^{(k)}\mathcal{F}\big(\mathcal{W}^\top\big)^{(k)} : \mathcal{F}\big(\mathcal{G}\big)^{(k)} : \mathcal{F}\big(\mathcal{W}\big)^{(k)}\mathcal{F}\big(\mathcal{W}^\top\big)^{(k)} + Q\big(\mathcal{W}\big). \end{aligned}$$

Determine the gradient and the Hessian of f.

$$f\big(\mathcal{W}\big) = G\big(\mathcal{W}\big) + Q\big(\mathcal{W}\big)$$

with

$$\begin{aligned} G\big(\mathcal{W}\big) = \frac{1}{2}\sum_{k=1}^{n_3}\left\{\begin{array}{c} \big(\mathcal{F}\big(\mathcal{W}\big)^{(k)}\mathcal{F}\big(\mathcal{W}^\top\big)^{(k)} - \mathcal{F}\big(\mathcal{N}^\star\big)^{(k)}\big) \\ : 4\mathcal{F}\big(\mathcal{H}_1\big)^{(k)} : \\ \big(\mathcal{F}\big(\mathcal{W}\big)^{(k)}\mathcal{F}\big(\mathcal{W}^\top\big)^{(k)} - \mathcal{F}\big(\mathcal{N}^\star\big)^{(k)}\big) \end{array}\right\} \\ + \mathcal{F}\big(\mathcal{W}\big)^{(k)}\mathcal{F}\big(\mathcal{W}^\top\big)^{(k)} : \mathcal{F}\big(\mathcal{G}\big)^{(k)} : \mathcal{F}\big(W\big)^{(k)}\mathcal{F}\big(\mathcal{W}^\top\big)^{(k)}. \end{aligned}$$

Using that, for any $\mathcal{Z} \in \mathbb{R}^{(n_1+n_2)\times r\times n_3}$

$$\begin{aligned}\langle \nabla f(\mathcal{W}), \mathcal{Z}\rangle &= \langle \nabla G(\mathcal{W}), \mathcal{Z}\rangle + \langle \nabla Q(\mathcal{W}), \mathcal{Z}\rangle \\ \mathcal{Z} : \nabla^2 f(\mathcal{W}) : \mathcal{Z} &= \mathcal{Z} : \nabla^2 G(\mathcal{W}) : \mathcal{Z} + \mathcal{Z} : \nabla^2 Q(W) : \mathcal{Z}.\end{aligned}$$

We now need to compute $\langle \nabla G(\mathcal{W}), \mathcal{Z}\rangle$ and $\mathcal{Z} : \nabla^2 G(\mathcal{W}) : \mathcal{Z}$. For this purpose, using the fact, for any $\mathcal{Z} \in \mathbb{R}^{\binom{n_1+n_2}{}\times r\times n_3}$, we obtain that:

$$G\big(\mathcal{W}+\mathcal{Z}\big) = G\big(\mathcal{W}\big) + \langle \nabla G\big(\mathcal{W}\big), \mathcal{Z}\rangle + \frac{1}{2}\mathcal{Z} : \nabla^2 G\big(\mathcal{W}\big) : \mathcal{Z} + O\big(\|\mathcal{Z} * \mathcal{Z}^\top\|^2\big).$$

we obtain:

$$\begin{aligned}\langle \nabla G(\mathcal{W}), \mathcal{Z}\rangle &= \sum_{k=1}^{n_3} \left\{ \begin{array}{c} \big(\mathcal{F}(\mathcal{W})^{(k)}\mathcal{F}(\mathcal{W}^\top)^{(k)} - \mathcal{F}(\mathcal{N}^\star)^{(k)}\big) \\ : 4\,\mathcal{F}(\mathcal{H}_1)^{(k)} : \\ \big(\mathcal{F}(\mathcal{W})^{(k)}\mathcal{F}(\mathcal{Z}^\top)^{(k)} + \mathcal{F}(\mathcal{Z})^{(k)}\mathcal{F}(\mathcal{W}^\top)^{(k)}\big) \end{array} \right\} \\ &+ \sum_{k=1}^{n_3} \mathcal{F}\big(\mathcal{W}\big)^{(k)}\mathcal{F}\big(\mathcal{W}^\top\big)^{(k)} : \mathcal{F}\big(\mathcal{G}\big)^{(k)} : \big(\mathcal{F}(\mathcal{W})^{(k)}\mathcal{F}(\mathcal{Z}^\top)^{(k)} + \mathcal{F}(\mathcal{Z})^{(k)}\mathcal{F}(\mathcal{W}^\top)^{(k)}\big) \\ &= \big(\mathcal{N}-\mathcal{N}^\star\big) : 4\,\mathcal{H}_1 : \big(\mathcal{W} * \mathcal{Z}^\top + \mathcal{Z} * \mathcal{W}^\top\big) + \mathcal{N} : \mathcal{G} : \big(\mathcal{W} * \mathcal{Z}^\top + \mathcal{Z} * \mathcal{W}^\top\big)\end{aligned}$$

and

$$\begin{aligned}\mathcal{Z} : \nabla^2 G(\mathcal{W}) : \mathcal{Z} &= \sum_{k=1}^{n_3} \left\{ \begin{array}{c} \big(\mathcal{F}(\mathcal{W})^{(k)}\mathcal{F}(\mathcal{Z}^\top)^{(k)} + \mathcal{F}(\mathcal{Z})^{(k)}\mathcal{F}(\mathcal{W}^\top)^{(k)}\big) \\ : \big(4\,\mathcal{F}(\mathcal{H}_1)^{(k)} + \mathcal{F}\big(\mathcal{G}\big)^{(k)}\big) : \\ \big(\mathcal{F}(\mathcal{W})^{(k)}\mathcal{F}(\mathcal{Z}^\top)^{(k)} + \mathcal{F}(\mathcal{Z})^{(k)}\mathcal{F}(\mathcal{W}^\top)^{(k)}\big) \end{array} \right\} \\ &+ \sum_{k=1}^{n_3} \left\{ \begin{array}{c} 2\big(\mathcal{F}(\mathcal{N})^{(k)} - \mathcal{F}(\mathcal{N}^\star)^{(k)}\big) \\ : 4\,\mathcal{F}(\mathcal{H}_1)^{(k)} : \\ \mathcal{F}(\mathcal{Z})^{(k)}\mathcal{F}(\mathcal{Z}^\top)^{(k)} \end{array} \right\} \\ &+ \sum_{k=1}^{n_3} 2\ \mathcal{F}(\mathcal{N})^{(k)} : \mathcal{F}(\mathcal{G})^{(k)} : \mathcal{F}(\mathcal{Z})^{(k)}\mathcal{F}(\mathcal{Z}^\top)^{(k)} \\ &= \big(\mathcal{W} * \mathcal{Z}^\top + \mathcal{Z} * \mathcal{W}^\top\big) : \big(4\,\mathcal{H}_1 + \mathcal{G}\big) : \big(\mathcal{W} * \mathcal{Z}^\top + \mathcal{Z} * \mathcal{W}^\top\big) \\ &+ 2\,\big(\mathcal{N}-\mathcal{N}^\star\big) : 4\mathcal{H}_1 : \mathcal{Z} * \mathcal{Z}^\top + 2\mathcal{N} : \mathcal{G} : \mathcal{Z} * \mathcal{Z}^\top.\end{aligned}$$

Let $\mathcal{Z} = \Delta = \mathcal{W} - \mathcal{W}^\star * \mathcal{R}$ and $\mathcal{H} = 4\mathcal{H}_1 + \mathcal{G}$. By noting that: $\mathcal{N} - \mathcal{N}^\star + \Delta * \Delta^\top = \mathcal{W} * \Delta^\top + \Delta * \mathcal{W}^\top$, we have

$$\begin{aligned}
\langle \nabla f(\mathcal{W}), \Delta \rangle &= \Big(\mathcal{N} - \mathcal{N}^\star\Big) : 4\mathcal{H}_1 : \Big(\mathcal{W} * \Delta^\top + \Delta * \mathcal{W}^\top\Big) + \mathcal{N} : \mathcal{G} : \Big(\mathcal{W} * \Delta^\top + \Delta * \mathcal{W}^\top\Big) \\
&\quad + \langle \nabla Q(\mathcal{W}), \Delta \rangle \\
&= \Big(\mathcal{N} - \mathcal{N}^\star\Big) : \mathcal{H} : \Big(\mathcal{W} * \Delta^\top + \Delta * \mathcal{W}^\top\Big) - \Big(\mathcal{N} - \mathcal{N}^\star\Big) : \mathcal{G} : \Big(\mathcal{W} * \Delta^\top + \Delta * \mathcal{W}^\top\Big) \\
&\quad + \mathcal{N} : \mathcal{G} : \Big(\mathcal{W} * \Delta^\top + \Delta * \mathcal{W}^\top\Big) + \langle \nabla Q(\mathcal{W}), \Delta \rangle \\
&= \Big(\mathcal{N} - \mathcal{N}^\star\Big) : \mathcal{H} : \Big(\mathcal{W} * \Delta^\top + \Delta * \mathcal{W}^\top\Big) + \mathcal{N}^\star : \mathcal{G} : \Big(\mathcal{N} - \mathcal{N}^\star + \Delta * \Delta^\top\Big) \\
&\quad + \langle \nabla Q(\mathcal{W}), \Delta \rangle \\
&= \Big(\mathcal{N} - \mathcal{N}^\star\Big) : \mathcal{H} : \Big(\mathcal{W} * \Delta^\top + \Delta * \mathcal{W}^\top\Big) + 2\,\mathcal{N}^\star : \mathcal{G} : \mathcal{N} + \langle \nabla Q(\mathcal{W}), \Delta \rangle.
\end{aligned}$$

The last equality, we expand $\mathcal{N} - \mathcal{N}^\star + \Delta * \Delta^\top$ and we use that

$$\mathcal{N}^\star : \mathcal{G} : \mathcal{N}^\star = \mathcal{N}^\star : \mathcal{G} : \mathcal{W}^\star * \mathcal{W}^\top = 0$$

due to fact

$$\mathcal{U}^{\star^\top} * \mathcal{U}^\star = \mathcal{V}^{\star^\top} * \mathcal{V}^\star.$$

Now, the Hessian along the direction Δ is:

$$\begin{aligned}
\Delta : \nabla^2 f(\mathcal{W}) : \Delta = &\Big(\mathcal{W} * \Delta^\top + \Delta * \mathcal{W}^\top\Big) : \mathcal{H} : \Big(\mathcal{W} * \Delta^\top + \Delta * \mathcal{W}^\top\Big) \\
&+ 2\Big(\mathcal{N} - \mathcal{N}^\star\Big) : 4\mathcal{H}_1 : \Delta * \Delta^\top \\
&+ 2\,\mathcal{N} : \mathcal{G} : \Delta * \Delta^\top + \Delta : \nabla^2 Q(\mathcal{W}) : \Delta.
\end{aligned} \tag{4.37}$$

We are interested in the first term of (4.37) with

$$\mathcal{N} - \mathcal{N}^\star + \Delta * \Delta^\top = \mathcal{W} * \Delta^\top + \Delta * \mathcal{W}^\top,$$

we have:

$$\begin{aligned}
&\Big(\mathcal{W} * \Delta^\top + \Delta * \mathcal{W}^\top\Big) : \mathcal{H} : \Big(\mathcal{W} * \Delta^\top + \Delta * \mathcal{W}^\top\Big) \\
&= \sum_{k=1}^{n_3} \left\{ \begin{array}{c} \Big(\mathcal{F}(\mathcal{W})^{(k)} \mathcal{F}(\Delta^\top)^{(k)} + \mathcal{F}(\Delta)^{(k)} \mathcal{F}(\mathcal{W}^\top)^{(k)}\Big) \\ : \mathcal{F}(\mathcal{H})^{(k)} : \\ \Big(\mathcal{F}(\mathcal{W})^{(k)} \mathcal{F}(\Delta^\top)^{(k)} + \mathcal{F}(\Delta)^{(k)} \mathcal{F}(\mathcal{W}^\top)^{(k)}\Big) \end{array} \right\} \\
&= \sum_{k=1}^{n_3} \mathcal{F}(\Delta)^{(k)} \mathcal{F}(\Delta^\top)^{(k)} : \mathcal{F}(\mathcal{H})^{(k)} : \mathcal{F}(\Delta)^{(k)} \mathcal{F}(\Delta^\top)^{(k)} \\
&\quad + 2\Big(\mathcal{F}(\mathcal{N})^{(k)} - \mathcal{F}(\mathcal{N}^\star)^{(k)}\Big) : \mathcal{F}(\mathcal{H})^{(k)} : \Big(\mathcal{F}(\mathcal{W})^{(k)} \mathcal{F}(\Delta^\top)^{(k)} + \mathcal{F}(\Delta)^{(k)} \mathcal{F}(\mathcal{W}^\top)^{(k)}\Big) \\
&\quad - \Big(\mathcal{F}(\mathcal{N})^{(k)} - \mathcal{F}(\mathcal{N}^\star)^{(k)}\Big) : \mathcal{F}(\mathcal{H})^{(k)} : \Big(\mathcal{F}(\mathcal{N})^{(k)} - \mathcal{F}(\mathcal{N}^\star)^{(k)}\Big) \\
&= \sum_{k=1}^{n_3} \mathcal{F}(\Delta)^{(k)} \mathcal{F}(\Delta^\top)^{(k)} : \mathcal{F}(\mathcal{H})^{(k)} : \mathcal{F}(\Delta)^{(k)} \mathcal{F}(\Delta^\top)^{(k)} \\
&\quad - \Big(\mathcal{F}(\mathcal{N})^{(k)} - \mathcal{F}(\mathcal{N}^\star)^{(k)}\Big) : \mathcal{F}(\mathcal{H})^{(k)} : \Big(\mathcal{F}(\mathcal{N})^{(k)} - \mathcal{F}(\mathcal{N}^\star)^{(k)}\Big) \\
&\quad - 4\mathcal{F}(\mathcal{N}^\star)^{(k)} : \mathcal{F}(\mathcal{G})^{(k)} : \mathcal{F}(\mathcal{N})^{(k)} + 2\langle \nabla f(\mathcal{W}), \Delta \rangle - 2\langle \nabla Q(\mathcal{W}), \Delta \rangle.
\end{aligned}$$

For the sum of second and third terms of (4.37), we have:

$$
\begin{aligned}
&2\Big(\mathcal{N}-\mathcal{N}^\star\Big) : 4\mathcal{H}_1 : \Delta * \Delta^\top + 2\mathcal{N} : \mathcal{G} : \Delta * \Delta^\top \\
&= \sum_{k=1}^{n_3} 2\Big(\mathcal{F}(\mathcal{N})^{(k)} - \mathcal{F}(\mathcal{N}^\star)^{(k)}\Big) : 4\,\mathcal{F}(\mathcal{H}_1)^{(k)} : \mathcal{F}(\Delta)^{(k)}\mathcal{F}(\Delta^\top)^{(k)} \\
&+ 2\,\mathcal{F}(\mathcal{N})^{(k)} : \mathcal{F}(\mathcal{G})^{(k)} : \mathcal{F}(\Delta)^{(k)}\mathcal{F}(\Delta^\top)^{(k)} \\
&= \sum_{k=1}^{n_3} 2\Big(\mathcal{F}(\mathcal{N})^{(k)} - \mathcal{F}(\mathcal{N}^\star)^{(k)}\Big) : \mathcal{F}(\mathcal{H})^{(k)} : \mathcal{F}(\Delta)^{(k)}\mathcal{F}(\Delta^\top)^{(k)} \\
&+ 2\,\mathcal{F}(\mathcal{N}^\star)^{(k)} : \mathcal{F}(\mathcal{G})^{(k)} : \mathcal{F}(\Delta)^{(k)}\mathcal{F}(\Delta^\top)^{(k)} \\
&= \sum_{k=1}^{n_3} -2\Big(\mathcal{F}(\mathcal{N})^{(k)} - \mathcal{F}(\mathcal{N}^\star)^{(k)}\Big) : \mathcal{F}(\mathcal{H})^{(k)} : \Big(\mathcal{F}(\mathcal{N})^{(k)} - \mathcal{F}(\mathcal{N}^\star)^{(k)}\Big) \\
&\quad + 2\Big(\mathcal{F}(\mathcal{N})^{(k)} - \mathcal{F}(\mathcal{N}^\star)^{(k)}\Big) : \mathcal{F}(\mathcal{H})^{(k)} : \Big(\mathcal{F}(\mathcal{W})^{(k)}\mathcal{F}(\Delta^\top)^{(k)} + \mathcal{F}(\Delta)^{(k)}\mathcal{F}(\mathcal{W}^\top)^{(k)}\Big) \\
&\quad + 2\,\mathcal{F}(\mathcal{N}^\star)^{(k)} : \mathcal{F}(\mathcal{G})^{(k)} : \mathcal{F}(\Delta)^{(k)}\mathcal{F}(\Delta^\top)^{(k)} \\
&= \sum_{k=1}^{n_3} -2\Big(\mathcal{F}(N)^{(k)} - \mathcal{F}(N^\star)^{(k)}\Big) : \mathcal{F}(\mathcal{H})^{(k)} : \Big(\mathcal{F}(N)^{(k)} - \mathcal{F}(N^\star)^{(k)}\Big) \\
&- 2\mathcal{F}(N^\star)^{(k)} : \mathcal{F}(\mathcal{G})^{(k)} : \mathcal{F}(N)^{(k)} + 2\langle \nabla f(\mathcal{W}), \Delta\rangle - 2\langle \nabla Q(\mathcal{W}), \Delta\rangle.
\end{aligned}
$$

To sum up, we have

$$
\begin{aligned}
&\Delta : \nabla^2 f(\mathcal{W}) : \Delta \\
&= \sum_{k=1}^{n_3} \mathcal{F}(\Delta)^{(k)}\mathcal{F}(\Delta^\top)^{(k)} : \mathcal{F}(\mathcal{H})^{(k)} : \mathcal{F}(\Delta)^{(k)}\mathcal{F}(\Delta^\top)^{(k)} \\
&- 3\Big(\mathcal{F}(\mathcal{N})^{(k)} - \mathcal{F}(\mathcal{N}^\star)^{(k)}\Big) : \mathcal{F}(\mathcal{H})^{(k)} : \Big(\mathcal{F}(\mathcal{N})^{(k)} - \mathcal{F}(\mathcal{N}^\star)^{(k)}\Big) \\
&\quad - 6\mathcal{F}(\mathcal{N}^\star)^{(k)} : \mathcal{F}(\mathcal{G})^{(k)} : \mathcal{F}(\mathcal{N})^{(k)} \\
&+ 4\langle \nabla f(\mathcal{W}), \Delta\rangle + \Big[\Delta : \nabla^2 Q(\mathcal{W}) : \Delta - 4\langle \nabla Q(\mathcal{W}), \Delta\rangle\Big] \\
&\leqslant \sum_{k=1}^{n_3} \mathcal{F}(\Delta)^{(k)}\mathcal{F}(\Delta^\top)^{(k)} : \mathcal{F}(\mathcal{H})^{(k)} : \mathcal{F}(\Delta)^{(k)}\mathcal{F}(\Delta)^{(k)^\top} \\
&- 3\Big(\mathcal{F}(\mathcal{N})^{(k)} - \mathcal{F}(\mathcal{N}^\star)^{(k)}\Big) : \mathcal{F}(\mathcal{H})^{(k)} : \Big(\mathcal{F}(\mathcal{N})^{(k)} - \mathcal{F}(\mathcal{N}^\star)^{(k)}\Big) \\
&+ 4\langle \nabla f(\mathcal{W}), \Delta\rangle + \Big[\Delta : \nabla^2 Q(\mathcal{W}) : \Delta - 4\langle \nabla Q(\mathcal{W}), \Delta\rangle\Big].
\end{aligned}
$$

The rest of the proof follows from the following lemma [†]. □

Lemma 4.7.19 *Let for any* $k = 1, \ldots, n_3$

$$
\mathcal{F}(\mathcal{N})^{(k)} := \begin{bmatrix} \mathcal{F}(\mathcal{U})^{(k)} \cdot \mathcal{F}(\mathcal{U}^\top)^{(k)} & \mathcal{F}(\mathcal{U})^{(k)} \cdot \mathcal{F}(\mathcal{V}^\top)^{(k)} \\ \mathcal{F}(\mathcal{V})^{(k)} \cdot \mathcal{F}(\mathcal{U}^\top)^{(k)} & \mathcal{F}(\mathcal{V})^{(k)} \cdot \mathcal{F}(\mathcal{V}^\top)^{(k)} \end{bmatrix} \in \mathbb{R}^{(n_1+n_2)\times(n_1+n_2)}. \tag{4.38}
$$

[†]Lemma 19 in [118] is false as stated

If $\mathcal{F}(\mathcal{H}_0)^{(k)}$ satisfies:

$$\Big(1-\delta\Big)\|\mathcal{F}(\mathcal{U})^{(k)}\cdot\mathcal{F}(\mathcal{V}^\top)^{(k)}\|_F^2 \leqslant \left\{\begin{array}{c}\mathcal{F}(\mathcal{U})^{(k)}\cdot\mathcal{F}(\mathcal{V}^\top)^{(k)}\\ :\mathcal{F}(\mathcal{H}_0)^{(k)}:\\ \mathcal{F}(\mathcal{U})^{(k)}\cdot\mathcal{F}(\mathcal{V}^\top)^{(k)}\end{array}\right\} \leqslant \Big(1+\delta\Big)\|\mathcal{F}(\mathcal{U})^{(k)}\cdot\mathcal{F}(\mathcal{V}^\top)^{(k)}\|_F^2. \tag{4.39}$$

Then, we have

$$\Big(1-2\delta\Big)\|\mathcal{F}(\mathcal{N})^{(k)}\|_F^2 \leqslant \mathcal{F}(\mathcal{N})^{(k)}:\mathcal{F}(\mathcal{H})^{(k)}:\mathcal{F}(\mathcal{N})^{(k)} \leqslant \Big(1+2\delta\Big)\|\mathcal{F}(\mathcal{N})^{(k)}\|_F^2.$$

Proof: Knowing that $\mathcal{H}_0$ preserves the norm $\mathcal{M}$, which is the off-diagonal of $\mathcal{N}$ and $\mathcal{G}$ the norm of the diagonal components of $\mathcal{N}$, we have, for any $k=1,\dots,n_3$

$$\begin{aligned}
&\mathcal{F}(\mathcal{N})^{(k)}:\mathcal{F}(\mathcal{H})^{(k)}:\mathcal{F}(\mathcal{N})^{(k)}\\
&=\mathcal{F}(\mathcal{N})^{(k)}:4\mathcal{F}(\mathcal{H}_1)^{(k)}:\mathcal{F}(\mathcal{N})^{(k)}+\mathcal{F}(\mathcal{N})^{(k)}:\mathcal{F}(\mathcal{G})^{(k)}:\mathcal{F}(\mathcal{N})^{(k)}\\
&=4\,\mathcal{F}(\mathcal{U})^{(k)}\cdot\mathcal{F}(\mathcal{V}^\top)^{(k)}:\mathcal{F}(\mathcal{H}_0)^{(k)}:\mathcal{F}(\mathcal{U})^{(k)}\cdot\mathcal{F}(\mathcal{V}^\top)^{(k)}\\
&\quad+\Big(\|\mathcal{F}(\mathcal{U})^{(k)}\cdot\mathcal{F}(\mathcal{U}^\top)^{(k)}\|_F^2+\|\mathcal{F}(\mathcal{V})^{(k)}\cdot\mathcal{F}(\mathcal{V}^\top)^{(k)}\|_F^2\\
&\quad-2\langle\mathcal{F}(\mathcal{U})^{(k)}\cdot\mathcal{F}(\mathcal{U}^\top)^{(k)},\mathcal{F}(\mathcal{V})^{(k)}\cdot\mathcal{F}(\mathcal{V}^\top)^{(k)}\rangle\Big)\\
&=4\,\mathcal{F}(\mathcal{U})^{(k)}\cdot\mathcal{F}(\mathcal{V}^\top)^{(k)}:\mathcal{F}(\mathcal{H}_0)^{(k)}:\mathcal{F}(\mathcal{U})^{(k)}\cdot\mathcal{F}(\mathcal{V}^\top)^{(k)}\\
&\quad+\Big(\|\mathcal{F}(\mathcal{U})^{(k)}\cdot\mathcal{F}(\mathcal{U}^\top)^{(k)}\|_F^2+\|\mathcal{F}(\mathcal{V})^{(k)}\cdot\mathcal{F}(\mathcal{V}^\top)^{(k)}\|_F^2\\
&\quad-2\|\mathcal{F}(\mathcal{U})^{(k)}\cdot\mathcal{F}(\mathcal{V}^\top)^{(k)}\|_F^2\Big).
\end{aligned}$$

Using (4.39), we obtain by calculating:

$$\Big(1-2\delta\Big)\|\mathcal{F}(\mathcal{N})^{(k)}\|_F^2 \leqslant \mathcal{F}(\mathcal{N})^{(k)}:\mathcal{F}(\mathcal{H})^{(k)}:\mathcal{F}(\mathcal{N})^{(k)} \leqslant \Big(1+2\delta\Big)\|\mathcal{F}(\mathcal{N})^{(k)}\|_F^2.$$

□

4.7.3.2 End of the proof

We first prove that the regularisation enforces that the rows of $\mathcal{F}\big(\mathcal{U}\big)^{(k)}$, $\mathcal{F}\big(\mathcal{V}\big)^{(k)}$, for $k=1\dots n_3$, cannot be too large.

Lemma 4.7.20 *Let $d=\max\{n_1,n_2\}$, there exists an absolute constant c_1, when sample rate*

$$p>c_1\frac{\mu r}{\min\{n_1,n_2\}}\log(d),$$

$\alpha_1^2=\Theta\big(\frac{\mu r\sigma_1^\star}{n_1}\big),\alpha_2^2=\Theta\big(\frac{\mu r\sigma_1^\star}{n_2}\big)$ and $\lambda_1=\Theta\big(\frac{n_1}{\mu r\kappa^\star}\big),\lambda_2=\Theta\big(\frac{n_2}{\mu r\kappa^\star}\big)$, we have for any $\mathcal{W}$ with $\|\nabla f(\mathcal{W})\|_F\leqslant\varepsilon$ for any polynomially small ε, with probability at least

$$1-2n_1n_3\exp\left\{-pn_2\Big(\Big(1+\frac{t}{pn_2}\Big)\ln\Big(1+\frac{t}{pn_2}\Big)-\frac{t}{pn_2}\Big)\right\}$$

$$\max_{1\leqslant i\leqslant n_1} \|e_i^\top * \mathcal{U}\|_F^2 \leqslant C\ n_3 \frac{\mu^2 r^{2.5} (\kappa^\star)^2 \sigma_1^\star}{n_1},$$

$$\max_{1\leqslant j\leqslant n_2} \|e_j^\top * \mathcal{V}\|_F^2 \leqslant C\ n_3 \frac{\mu^2 r^{2.5} (\kappa^\star)^2 \sigma_1^\star}{n_2}$$

for some constant positive C.

Proof: In this proof, by symmetry, without loss of generality, we can assume that for any $k = 1, \ldots, n_3$

$$\sqrt{n_1} \max_{1\leqslant i\leqslant n_1} \|\mathcal{F}\Big(e_i^\top\Big)^{(k)} \mathcal{F}\Big(\mathcal{U}\Big)^{(k)}\|_2 \geqslant \sqrt{n_2} \max_{1\leqslant j\leqslant n_2} \|\mathcal{F}\Big(e_j^\top\Big)^{(k)} \mathcal{F}\Big(\mathcal{V}\Big)^{(k)}\|_2.$$

By calculating the gradient, we can write the gradient as:

$$\nabla f(\mathcal{W}) = \frac{4}{p} \begin{pmatrix} \Big(\mathcal{M} - \mathcal{M}^\star\Big)_\Omega * \mathcal{V} \\ \Big(\mathcal{M} - \mathcal{M}^\star\Big)_\Omega * \mathcal{U} \end{pmatrix} + \begin{pmatrix} \mathcal{U} * \Big(\mathcal{U}^\top * \mathcal{U} - \mathcal{V}^\top * \mathcal{V}\Big) \\ \mathcal{V} * \Big(\mathcal{V}^\top * \mathcal{V} - \mathcal{U}^\top * \mathcal{U}\Big) \end{pmatrix} + \nabla Q(\mathcal{W})$$

where

$$\begin{aligned} \nabla Q(\mathcal{W}) = 4\lambda_1 \sum_{k=1}^{n_3} \sum_{i=1}^{n_1} \Big(\|\mathcal{F}\Big(e_i^\top\Big)^{(k)} \mathcal{F}\Big(\mathcal{W}\Big)^{(k)}\|_2 - \alpha_1\Big)_+^3 \\ \times \frac{\mathcal{F}\Big(e_i\Big)^{(k)} \mathcal{F}\Big(e_i^\top\Big)^{(k)} \mathcal{F}\Big(\mathcal{W}\Big)^{(k)}}{\|\mathcal{F}\Big(e_i\Big)^{(k)} \mathcal{F}\Big(\mathcal{W}\Big)^{(k)}\|_2^2} \\ + 4\lambda_2 \sum_{k=1}^{n_3} \sum_{i=n_1+1}^{n_2} \Big(\|\mathcal{F}\Big(e_i^\top\Big)^{(k)} \mathcal{F}\Big(\mathcal{W}\Big)^{(k)}\|_2 - \alpha_2\Big)_+^3 \\ \times \frac{\mathcal{F}\Big(e_i\Big)^{(k)} \mathcal{F}\Big(e_i^\top\Big)^{(k)} \mathcal{F}\Big(\mathcal{W}\Big)^{(k)}}{\|\mathcal{F}\Big(e_i\Big)^{(k)} \mathcal{F}\Big(\mathcal{W}\Big)^{(k)}\|_2^2}. \end{aligned}$$

Using the fact $\langle \nabla Q(\mathcal{W}), \mathcal{W} \rangle \geqslant 0$, thus, for any point $\mathcal{W}$ with gradient $\|\nabla f(\mathcal{W})\|_F \leqslant \varepsilon$, we

have:

$$
\begin{aligned}
\varepsilon\|\mathcal{W}\|_F &\geqslant \langle \nabla f(\mathcal{W}), \mathcal{W}\rangle \\
&= \|\mathcal{U}^\top * \mathcal{U} - \mathcal{V}^\top * \mathcal{V}\|_F^2 + \frac{4}{p}\Big\langle \big(\mathcal{M} - \mathcal{M}^\star\big)_\Omega, \mathcal{M}\Big\rangle + \Big\langle \nabla Q(\mathcal{W}), \mathcal{W}\Big\rangle \\
&\geqslant \|\mathcal{U}^\top * \mathcal{U} - \mathcal{V}^\top * \mathcal{V}\|_F^2 - \frac{4}{p}\Big\langle \big(\mathcal{M}^\star\big)_\Omega, \big(\mathcal{M}\big)_\Omega\Big\rangle \\
&\geqslant \|\mathcal{U}^\top * \mathcal{U} - \mathcal{V}^\top * \mathcal{V}\|_F^2 - 4\cdot\frac{1}{\sqrt{p}}\|\mathcal{M}^\star\|_\Omega \cdot \frac{1}{\sqrt{p}}\|\mathcal{M}\|_\Omega \\
&= \|\mathcal{U}^\top * \mathcal{U} - \mathcal{V}^\top * \mathcal{V}\|_F^2 - 4\cdot\frac{1}{\sqrt{p}}\sqrt{\sum_{k=1}^{n_3}\|\mathcal{F}\Big(\mathcal{M}^\star\Big)^{(k)}\|_\Omega^2}\cdot\frac{1}{\sqrt{p}}\sqrt{\sum_{k=1}^{n_3}\|\mathcal{F}\Big(\mathcal{M}\Big)^{(k)}\|_\Omega^2} \\
&\geqslant \|\mathcal{U}^\top * \mathcal{U} - \mathcal{V}^\top * \mathcal{V}\|_F^2 - C\ \sqrt{n_1 n_2}\sum_{k=1}^{n_3}\|\mathcal{F}\Big(\mathcal{M}^\star\Big)^{(k)}\|_F \cdot \sum_{k=1}^{n_3}\|\mathcal{F}\Big(M\Big)^{(k)}\|_\infty
\end{aligned}
$$

where in the last inequality, we use Lemma 4.7.1 and Lemma 4.7.5. Let

$$
i^\star = \underset{1\leqslant i\leqslant n_1}{\arg\max}\ \|e_i^\top * \mathcal{U}\|_F \tag{4.40}
$$

and

$$
j^\star = \underset{1\leqslant j\leqslant n_2}{\arg\max}\ \|e_j^\top * \mathcal{V}\|_F. \tag{4.41}
$$

Using these facts and recalling that, by assumption,

- $\sqrt{n_1}\|\mathcal{F}(e_{i^\star}^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2 \geqslant \sqrt{n_2}\|\mathcal{F}(e_{j^\star}^\top)^{(k)}\mathcal{F}(\mathcal{V})^{(k)}\|_2$
- $\|\mathcal{M}^\star\|_F = \|\mathcal{F}(\mathcal{M}^\star)\|_F = \sqrt{\sum_{k=1}^{n_3}\|\mathcal{F}(\mathcal{M}^\star)^{(k)}\|_F^2} \leqslant n_3\sigma_1^\star\sqrt{r}$
- $\|\mathcal{M}\|_\infty = \|\mathcal{F}(\mathcal{M})\|_\infty \leqslant \sum_{k=1}^{n_3}\|\mathcal{F}(e_{i^\star}^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2\|\mathcal{F}(e_{j^\star}^\top)^{(k)}\mathcal{F}(\mathcal{V})^{(k)}\|_2$

we have, for some positive constant C:

$$
\begin{aligned}
\|\mathcal{U}^\top * \mathcal{U} - \mathcal{V}^\top * \mathcal{V}\|_F^2 \leqslant C\ n_1 n_3\sigma_1^\star\sqrt{r}\sum_{k=1}^{n_3}\|\mathcal{F}(e_{i^\star}^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2^2 \\
+ C\ \varepsilon d\|e_{i^\star}^\top * \mathcal{U}\|_F,
\end{aligned} \tag{4.42}
$$

where in the second term of (4.42), we use:

$$
\begin{aligned}
\|\mathcal{W}\|_F &\leqslant \|\mathcal{U}\|_F + \|\mathcal{V}\|_F \\
&\leqslant \sum_{k=1}^{n_3} \|\mathcal{F}(\mathcal{U})^k\|_F + \|\mathcal{F}(\mathcal{V})^k\|_F \\
&\leqslant \sum_{k=1}^{n_3} \sqrt{n_1} \max_i \|\mathcal{F}(e_i^\top)^{(k)} \mathcal{F}(\mathcal{U})^{(k)}\|_2 + \sqrt{n_2} \max_j \|\mathcal{F}(e_j^\top)^{(k)} \mathcal{F}(\mathcal{V})^{(k)}\|_2 \\
&\leqslant 2 \sum_{k=1}^{n_3} \sqrt{n_1}\, \|\mathcal{F}(e_{i^\star}^\top)^{(k)} \mathcal{F}(\mathcal{U})^{(k)}\|_2 \\
&\leqslant d \sum_{k=1}^{n_3} \|\mathcal{F}(e_{i^\star}^\top)^{(k)} \mathcal{F}(\mathcal{U})^{(k)}\|_2 \text{ where } d = \max\{n_1, n_2\} \\
&= d \|e_{i^\star}^\top * \mathcal{U}\|_F.
\end{aligned}
$$

In the case $\|e_{i^\star}^\top * \mathcal{U}\|_2 \geqslant 2\alpha_1$, consider $\langle e_{i^\star}^\top * \nabla f(\mathcal{U}), e_{i^\star}^\top * \mathcal{U} \rangle$ as:

$$
\begin{aligned}
&\varepsilon \|e_{i^\star}^\top * \mathcal{U}\|_F \\
&\geqslant \langle e_{i^\star}^\top * \nabla f(\mathcal{U}), e_{i^\star}^\top * \mathcal{U} \rangle \\
&= \left\langle e_{i^\star}^\top * \Big(\frac{4}{p} \big(\mathcal{M} - \mathcal{M}^\star \big)_\Omega * \mathcal{V} + \mathcal{U} * \big(\mathcal{U}^\top * \mathcal{U} - \mathcal{V}^\top * \mathcal{V} \big) + \nabla Q_1(\mathcal{U}) \Big), e_{i^\star}^\top * \mathcal{U} \right\rangle \\
&\geqslant \frac{\lambda_1}{2} \Big(\|e_{i^\star}^\top * \mathcal{U}\|_F - \alpha_1 \Big)_+^3 \|e_{i^\star}^\top * \mathcal{U}\|_F - \frac{4}{p} \langle e_{i^\star}^\top * (\mathcal{M}^\star)_\Omega, e_{i^\star}^\top * \big(\mathcal{M} \big)_\Omega \rangle \\
&- \|\mathcal{U}^\top * \mathcal{U} - \mathcal{V}^\top * \mathcal{V}\|_F \|e_{i^\star}^\top * \mathcal{U}\|_F^2 \\
&\geqslant \frac{\lambda_1}{2} \|e_{i^\star}^\top * \mathcal{U}\|_F^4 - 4 \frac{1}{\sqrt{p}} \|e_{i^\star}^\top * (\mathcal{M}^\star)_\Omega\| \cdot \frac{1}{\sqrt{p}} \|e_{i^\star}^\top * \big(\mathcal{M} \big)_\Omega\| - \|\mathcal{U}^\top * \mathcal{U} - \mathcal{V}^\top * \mathcal{V}\|_F \cdot \|e_{i^\star}^\top * \mathcal{U}\|_F^2 \\
&\geqslant \frac{\lambda_1}{2} \|e_{i^\star}^\top * \mathcal{U}\|_F^4 - \|\mathcal{U}^\top * \mathcal{U} - \mathcal{V}^\top * \mathcal{V}\|_F \cdot \|e_{i^\star}^\top * \mathcal{U}\|_F^2 \\
&- 4\sqrt{1 + 0.01} \sum_{k=1}^{n_3} \|\mathcal{F}(e_{i^\star}^\top)^{(k)} \mathcal{F}(\mathcal{M}^\star)^{(k)}\|_2 \cdot C \sqrt{n_2} \sum_{k=1}^{n_3} \|\mathcal{F}(\mathcal{M})^{(k)}\|_\infty \\
&\geqslant \frac{\lambda_1}{2} \|e_{i^\star}^\top * \mathcal{U}\|_F^4 - \|\mathcal{U}^\top * \mathcal{U} - \mathcal{V}^\top * \mathcal{V}\|_F \cdot \|e_{i^\star}^\top * \mathcal{U}\|_F^2 \\
&\quad - C \sqrt{\mu r} \sigma_1^\star \sum_{k=1}^{n_3} \|\mathcal{F}(e_{i^\star}^\top)^{(k)} \mathcal{F}(\mathcal{U})^{(k)}\|_2^2
\end{aligned}
$$

where in the last inequality, we use (4.7.1) and (4.7.5). Further, using (4.42), we have:

$$\begin{aligned}\lambda_1 \sum_{k=1}^{n_3} \|\mathcal{F}(e_{i^\star}^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2^3 \leqslant 2\varepsilon &+ C\ \sqrt{\mu r}\sigma_1^\star \sum_{k=1}^{n_3} \|\mathcal{F}(e_{i^\star}^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2 \\ &+ C\ \sqrt{n_1 n_3 \sigma_1^\star} r^{\frac{1}{4}} \sum_{k=1}^{n_3} \|\mathcal{F}(e_{i^\star}^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2^2 \\ &+ C\sqrt{\varepsilon d} \sum_{k=1}^{n_3} \|\mathcal{F}(e_{i^\star}^\top)^{(k)}\mathcal{F}(\mathcal{U})^{(k)}\|_2^{\frac{3}{2}}.\end{aligned}$$

By choosing ε to be polynomially small, we have:

$$\begin{aligned}\sqrt{\frac{n_1}{n_2}} \max_{1\leqslant j\leqslant n_2} \|e_{j^\star}^\top * \mathcal{V}\|_F &\leqslant \max_{1\leqslant i\leqslant n_1} \|e_{i^\star}^\top * \mathcal{U}\|_F^2 \\ &\leqslant c \max\left\{\alpha_1^2, \frac{\sqrt{\mu r}.\sigma_1^\star}{\lambda_1}, \frac{n_1 n_3 \sigma_1^\star \sqrt{r}}{\lambda_1^2}\right\}\end{aligned}$$

for some positive constant c. Finally, substituting the choice of α^2 and λ_1, we finished the proof. □

Now, we prove, that the Hessian $\mathcal{H}$ terms in (4.36) is negative when $\mathcal{W} \neq \mathcal{W}^\star$.

Lemma 4.7.21 *Let $d = \max\{n_1, n_2\}$, when that μ and r satisfy the inequality $p \geqslant \Omega\Big(\frac{\mu^4 r^6 (\kappa^\star)^6 \log d}{\min\{n_1,n_2\}}\Big)$, by choosing $\alpha_1^2 = \Theta\Big(\frac{\mu r \sigma_1^\star}{n_1}\Big)$, $\alpha_2^2 = \Theta\Big(\frac{\mu r \sigma_1^\star}{n_2}\Big)$ and $\lambda_1 = \Theta\Big(\frac{n_1}{\mu r \kappa^\star}\Big)$, $\lambda_2 = \Theta\Big(\frac{n_2}{\mu r \kappa^\star}\Big)$ with probability at least*

$$1 - 2n_1 n_3 \exp\left\{-pn_2\Big(\Big(1 + \frac{t}{pn_2}\Big)\ln\Big(1 + \frac{t}{pn_2}\Big) - \frac{t}{pn_2}\Big)\right\}$$

for all $\mathcal{W}$ with $\|\nabla f(\mathcal{W})\|_F \leqslant \varepsilon$ and for polynomially small ε, we have

$$\Delta * \Delta^\top : \mathcal{H} : \Delta * \Delta^\top - 3(\mathcal{M} - \mathcal{M}^\star) : \mathcal{H} : (\mathcal{M} - \mathcal{M}^\star) \leqslant -0.3 \sum_{k=1}^{n_3} \sigma_r^\star \|\mathcal{F}\Big(\Delta\Big)^{(k)}\|_F^2.$$

Proof: As the symmetric case, we are interested in studying the two cases on the norm of Δ and we use the different inequalities of concentrations. Using the choices of α and λ and (4.7.20), we know when ε is polynomially small with high probability:

$$\max_{1\leqslant i\leqslant n_1} \|e_i^\top * \mathcal{U}\|_F^2 \leqslant C\ n_3 \frac{\mu^2 r^{2.5} (\kappa^\star)^2 \sigma_1^\star}{n_1},$$

and

$$\max_{1\leqslant j\leqslant n_2} \|e_j^\top * \mathcal{V}\|_F^2 \leqslant C\ n_3 \frac{\mu^2 r^{2.5} (\kappa^\star)^2 \sigma_1^\star}{n_2}.$$

In the following, we denote for any $k = 1, \dots, n_3$,

$$\mathcal{F}(\Delta)^{(k)} = \Big(\mathcal{F}(\Delta_{\mathcal{U}}^\top)^{(k)}, \mathcal{F}(\Delta_{\mathcal{V}}^\top)^{(k)}\Big)^\top$$

and we have

$$\|\mathcal{F}(\Delta_{\mathcal{U}})^{(k)}\|_F \leqslant \|\mathcal{F}(\Delta)^{(k)}\|_F$$

and

$$\|\mathcal{F}(\Delta_{\mathcal{V}})^{(k)}\|_F \leqslant \|\mathcal{F}(\Delta)^{(k)}\|_F.$$

We now split the analysis into two cases.

Case 1: $\|\mathcal{F}(\Delta)^{(k)}\|_F^2 \leqslant \frac{\sigma_r^\star}{40}$, for any $k = 1, \dots, n_3$. By (4.7.1) and (4.7.19), we have:

$$\frac{1}{p}\|\mathcal{W}^\star * \Delta^\top\|_\Omega^2 \geqslant \Big(1 - 2\delta\Big)\|\mathcal{W}^\star * \Delta^\top\|_F^2 \geqslant \Big(1 - 2\delta\Big)\sigma_r^\star\|\Delta\|_F^2.$$

Furthermore, we know:

$$\frac{1}{p}\|\Delta_{\mathcal{U}} * \Delta_{\mathcal{V}}^\top\|_\Omega^2 = \frac{1}{p}\sum_{k=1}^{n_3}\|\mathcal{F}(\Delta_{\mathcal{U}})^{(k)}\mathcal{F}(\Delta_{\mathcal{V}}^\top)^{(k)}\|_\Omega^2.$$

By (4.7.3) and with the choice of p, for any $k = 1, \dots, n_3$, we have:

$$\begin{aligned}
\frac{1}{p}\|\mathcal{F}(\Delta_{\mathcal{U}})^{(k)}\mathcal{F}(\Delta_{\mathcal{V}}^\top)^{(k)}\|_\Omega^2 &\leqslant \Big(1+\delta\Big)\|\mathcal{F}(\Delta_{\mathcal{U}})^{(k)}\|_F^2\|\mathcal{F}(\Delta_{\mathcal{V}}^\top)^{(k)}\|_F^2 \\
&\quad + C\sqrt{\frac{d}{p}}\cdot\frac{\mu^2 r^{2.5}(\kappa^\star)^2\sigma_1^\star}{\sqrt{n_1 n_2}}\|\mathcal{F}(\Delta_{\mathcal{U}})^{(k)}\|_F\|\mathcal{F}(\Delta_{\mathcal{V}})^{(k)}\|_F \\
&\leqslant \Big(1+\delta\Big)\|\mathcal{F}(\Delta)^{(k)}\|_F^4 + \frac{\sigma_r^\star}{4}\|\mathcal{F}(\Delta)^{(k)}\|_F^2 \\
&\leqslant \frac{\sigma_r^\star}{20}\|\mathcal{F}(\Delta)^{(k)}\|_F^2.
\end{aligned}$$

So, using that for $k = 1, \dots, n_3$,

$$\|\mathcal{F}(\Delta_{\mathcal{U}})^{(k)}\|_F^2 \leqslant \|\mathcal{F}(\Delta)^{(k)}\|_F^2 \leqslant \frac{\sigma_r^\star}{40}$$

and

$$\|\mathcal{F}(\Delta_{\mathcal{V}})^{(k)}\|_F^2 \leqslant \|\mathcal{F}(\Delta)^{(k)}\|_F^2 \leqslant \frac{\sigma_r^\star}{40}$$

we obtain:

$$\begin{aligned}
&\Delta * \Delta^\top : \mathcal{H} : \Delta * \Delta^\top \\
&= \frac{4}{p}\|\Delta_{\mathcal{U}} * \Delta_{\mathcal{V}}^\top\|_\Omega^2 + \Big(\|\Delta_{\mathcal{U}} * \Delta_{\mathcal{U}}^\top\|_F^2 + \|\Delta_{\mathcal{V}} * \Delta_{\mathcal{V}}^\top\|_F^2 - 2\|\Delta_{\mathcal{U}} * \Delta_{\mathcal{V}}^\top\|_F^2\Big) \\
&\leqslant \sum_{k=1}^{n_3}\frac{1}{4}\sigma_r^\star\|\mathcal{F}\Big(\Delta\Big)^{(k)}\|_F^2.
\end{aligned}$$

Using this facts, we obtain

$$
\begin{aligned}
&\Delta * \Delta^\top : \mathcal{H} : \Delta * \Delta^\top - 3\Big(\mathcal{N} - \mathcal{N}^\star\Big) : \mathcal{H} : \Big(\mathcal{N} - \mathcal{N}^\star\Big) \\
&= \sum_{k=1}^{n_3} \mathcal{F}(\Delta)^{(k)} \mathcal{F}(\Delta^\top)^{(k)} : \mathcal{F}(\mathcal{H})^{(k)} : \mathcal{F}(\Delta)^{(k)} \mathcal{F}(\Delta^\top)^{(k)} \\
&- 3\Big(\mathcal{F}(\mathcal{N})^{(k)} - \mathcal{F}(\mathcal{N}^\star)^{(k)}\Big) : \mathcal{F}(\mathcal{H})^{(k)} : \Big(\mathcal{F}(N)^{(k)} - \mathcal{F}(\mathcal{N}^\star)^{(k)}\Big) \\
&= \sum_{k=1}^{n_3} \mathcal{F}(\Delta)^{(k)} \mathcal{F}(\Delta^\top)^{(k)} : \mathcal{F}(\mathcal{H})^{(k)} : \mathcal{F}(\Delta)^{(k)} \mathcal{F}(\Delta^\top)^{(k)} \\
&- \left\{ \begin{array}{c} 3\Big(\mathcal{F}(\mathcal{W}^\star)^{(k)} \mathcal{F}(\Delta^\top)^{(k)} + \mathcal{F}(\Delta)^{(k)} \mathcal{F}(\mathcal{W}^{\star^\top})^{(k)} + \mathcal{F}(\Delta)^{(k)} \mathcal{F}(\Delta^\top)^{(k)}\Big) \\ : \mathcal{F}(\mathcal{H})^{(k)} : \\ \Big(\mathcal{F}(\mathcal{W}^\star)^{(k)} \mathcal{F}(\Delta^\top)^{(k)} + \mathcal{F}(\Delta)^{(k)} \mathcal{F}(\mathcal{W}^{\star^\top})^{(k)} + \mathcal{F}(\Delta)^{(k)} \mathcal{F}(\Delta^\top)^{(k)}\Big) \end{array} \right\} \\
&\leqslant \sum_{k=1}^{n_3} -12\Big(\mathcal{F}(\mathcal{W}^\star)^{(k)} \mathcal{F}(\Delta^\top)^{(k)} : \mathcal{F}(\mathcal{H})^{(k)} : \mathcal{F}(\Delta)^{(k)} \mathcal{F}(\Delta^\top)^{(k)} \\
&+ \mathcal{F}(\mathcal{W}^\star)^{(k)} \mathcal{F}(\Delta^\top)^{(k)} : \mathcal{F}(\mathcal{H})^{(k)} : \mathcal{F}(\mathcal{W}^\star)^{(k)} \mathcal{F}(\Delta^\top)^{(k)}\Big) \\
&\leqslant \sum_{k=1}^{n_3} -\frac{12}{p}\Big(\|\mathcal{F}(\mathcal{W}^\star)^{(k)} \mathcal{F}(\Delta^\top)^{(k)}\|_\Omega^2 - \|\mathcal{F}(\mathcal{W}^\star)^{(k)} \mathcal{F}(\Delta^\top)^{(k)}\|_\Omega \|\mathcal{F}(\Delta)^{(k)} \mathcal{F}(\Delta^\top)^{(k)}\|_\Omega\Big) \\
&= \sum_{k=1}^{n_3} -\frac{12}{p}\|\mathcal{F}(\mathcal{W}^\star)^{(k)} \mathcal{F}(\Delta^\top)^{(k)}\|_\Omega \Big(\|\mathcal{F}(\mathcal{W}^\star)^{(k)} \mathcal{F}(\Delta^\top)^{(k)}\|_\Omega - \|\mathcal{F}(\Delta)^{(k)} \mathcal{F}(\Delta^\top)^{(k)}\|_\Omega\Big) \\
&\leqslant \sum_{k=1}^{n_3} -12\sqrt{1-2\delta}\Big(\sqrt{1-2\delta} - \sqrt{1/4}\Big)\sigma_r^\star \|\mathcal{F}(\Delta)^{(k)}\|_F^2 \\
&\leqslant -1.2 \sum_{k=1}^{n_3} \sigma_r^\star \|\mathcal{F}(\Delta)^{(k)}\|_F^2.
\end{aligned}
$$

Case 2: $\|\mathcal{F}(\Delta)^{(k)}\|_F^2 \geqslant \frac{\sigma_r^\star}{40}$, for any $k = 1, \ldots, n_3$. By Lemma 4.7.4 with high probability

with the choice of p, we have:

$$
\begin{aligned}
\frac{1}{p}\|\Delta_{\mathcal{U}} * \Delta_{\mathcal{V}}^{\top}\|_{\Omega}^2 &= \sum_{k=1}^{n_3} \frac{1}{p}\|\mathcal{F}(\Delta_{\mathcal{U}})^{(k)}\mathcal{F}(\Delta_{\mathcal{V}}^{\top})^{(k)}\|_{\Omega}^2 \\
&\leqslant \sum_{k=1}^{n_3} \|\mathcal{F}(\Delta_{\mathcal{U}})^{(k)}\mathcal{F}(\Delta_{\mathcal{V}}^{\top})^{(k)}\|_F^2 \\
&+ C\Big(\frac{dr\log(d)}{p}\ \|\mathcal{F}(\Delta_{\mathcal{U}})^{(k)}\mathcal{F}(\Delta_{\mathcal{V}}^{\top})^{(k)}\|_{\infty}^2 \\
&+ \sqrt{\frac{dr\log(d)}{p}}\|\mathcal{F}(\Delta_{\mathcal{U}})^{(k)}\mathcal{F}(\Delta_{\mathcal{V}}^{\top})^{(k)}\ \|_F\ \|\mathcal{F}(\Delta_{\mathcal{U}})^{(k)}\mathcal{F}(\Delta_{\mathcal{V}}^{\top})^{(k)}\|_{\infty}\Big) \\
&\leqslant \sum_{k=1}^{n_3} \|\mathcal{F}(\Delta_{\mathcal{U}})^{(k)}\mathcal{F}(\Delta_{\mathcal{V}}^{\top})^{(k)}\|_F^2 \\
&+ C\Big(\frac{dr\log(d)}{p}\cdot\frac{\mu^4 r^5(\kappa^\star)^4(\sigma_1^\star)^2}{n_1 n_2} + \sqrt{\frac{dr\log(d)}{p}\cdot\frac{\mu^4 r^5(\kappa^\star)^4(\sigma_1^\star)^2}{n_1 n_2}}\ \|\mathcal{F}(\Delta)^{(k)}\|_F^2\Big) \\
&\leqslant \sum_{k=1}^{n_3} \|\mathcal{F}(\Delta_{\mathcal{U}})^{(k)}\mathcal{F}(\Delta_{\mathcal{V}}^{\top})^{(k)}\|_F^2 + \frac{\left(\sigma_r^\star\right)^2}{1000} + \frac{\sigma_r^\star}{1000}\ \|\mathcal{F}(\Delta)^{(k)}\|_F^2 \\
&\leqslant \sum_{k=1}^{n_3} \|\mathcal{F}(\Delta_{\mathcal{U}})^{(k)}\mathcal{F}(\Delta_{\mathcal{V}}^{\top})^{(k)}\|_F^2 + 0.01\sigma_r^\star\ \|\mathcal{F}(\Delta)^{(k)}\|_F^2.
\end{aligned}
$$

Again using Lemma 4.7.4 with high probability, we have

$$\begin{aligned}
\frac{1}{p}\|\mathcal{M}-\mathcal{M}^\star\|_\Omega^2 &= \sum_{k=1}^{n_3}\frac{1}{p}\|\mathcal{F}(\mathcal{M})^{(k)}-\mathcal{F}(\mathcal{M}^\star)^{(k)}\|_\Omega^2 \\
&\geqslant \sum_{k=1}^{n_3}\|\mathcal{F}(\mathcal{M})^{(k)}-\mathcal{F}(\mathcal{M}^\star)^{(k)}\|_F^2 \\
&+ C\Big(\frac{d\,r\log(d)}{p}\|\mathcal{F}(\mathcal{M})^{(k)}-\mathcal{F}(\mathcal{M}^\star)^{(k)}\|_\infty^2 \\
&+ \sqrt{\frac{d\,r\log(d)}{p}}\|\mathcal{F}(\mathcal{M})^{(k)}-\mathcal{F}(\mathcal{M}^\star)^{(k)}\|_F\times\|\mathcal{F}(\mathcal{M})^{(k)}-\mathcal{F}(\mathcal{M}^\star)^{(k)}\|_\infty\Big) \\
&\geqslant \sum_{k=1}^{n_3}\|\mathcal{F}(\mathcal{M})^{(k)}-\mathcal{F}(\mathcal{M}^\star)^{(k)}\|_F^2 \\
&+ C\Big(\frac{d\,r\log(d)}{p}\cdot\frac{\mu^4 r^5(\kappa^\star)^4(\sigma_1^\star)^2}{n_1n_2} \\
&+ \sqrt{\frac{d\,r\log(d)}{p}\cdot\frac{\mu^4 r^5(\kappa^\star)^4(\sigma_1^\star)^2}{n_1n_2}}\|\mathcal{F}(\mathcal{M})^{(k)}-\mathcal{F}(\mathcal{M}^\star)^{(k)}\|_F\Big) \\
&\geqslant \sum_{k=1}^{n_3}\|\mathcal{F}(\mathcal{M})^{(k)}-\mathcal{F}(\mathcal{M}^\star)^{(k)}\|_F^2-\frac{\left(\sigma_r^\star\right)^2}{1000} \\
&- \frac{\sigma_r^\star}{1000}\|\mathcal{F}(\mathcal{M})^{(k)}-\mathcal{F}(\mathcal{M}^\star)^{(k)}\|_F \\
&\geqslant \sum_{k=1}^{n_3}0.95\|\mathcal{F}(\mathcal{M})^{(k)}-\mathcal{F}(\mathcal{M}^\star)^{(k)}\|_F^2-0.01\sigma_r^\star\|\mathcal{F}(\Delta)^{(k)}\|_F^2.
\end{aligned}$$

Thus, we get:

$$\begin{aligned}
&\Delta*\Delta^\top:\mathcal{H}:\Delta*\Delta^\top-3(\mathcal{N}-\mathcal{N}^\star):\mathcal{H}:(\mathcal{N}-\mathcal{N}^\star) \\
&=\sum_{k=1}^{n_3}\mathcal{F}(\Delta)^{(k)}\mathcal{F}(\Delta^\top)^{(k)}:\mathcal{F}(\mathcal{H}):\mathcal{F}(\Delta)^{(k)}\mathcal{F}(\Delta^\top)^{(k)} \\
&-3\Big(\mathcal{F}(\mathcal{N})^{(k)}-\mathcal{F}(\mathcal{N}^\star)^{(k)}\Big):\mathcal{F}(\mathcal{H}):\Big(\mathcal{F}(\mathcal{N})^{(k)}-\mathcal{F}(\mathcal{N}^\star)^{(k)}\Big) \\
&\leqslant\sum_{k=1}^{n_3}\|\mathcal{F}(\Delta)^{(k)}\mathcal{F}(\Delta^\top)^{(k)}\|_F^2+0.04\sigma_r^\star\|\mathcal{F}(\Delta)^{(k)}\|_F^2 \\
&-\Big(0.98\|\mathcal{F}(\mathcal{N})^{(k)}-\mathcal{F}(\mathcal{N}^\star)^{(k)}\|_F^2-0.04\,\sigma_r^\star\|\mathcal{F}(\Delta)^{(k)}\|_F^2\Big) \\
&\leqslant\sum_{k=1}^{n_3}0.94\|\mathcal{F}(\mathcal{N})^{(k)}-\mathcal{F}(\mathcal{N}^\star)^{(k)}\|_F^2+0.12\,\sigma_r^\star\|\mathcal{F}(\Delta)^{(k)}\|_F^2 \\
&\leqslant\sum_{k=1}^{n_3}-0.3\,\sigma_r^\star\|\mathcal{F}(\Delta)^{(k)}\|_F^2
\end{aligned}$$

and the proof is completed. □

Now, it remains to bound the regularizer the terms of the (4.36).

Lemma 4.7.22 *By choosing* $\alpha_1^2 = \Theta\big(\frac{\mu r \sigma_1^\star}{n_1}\big), \alpha_2^2 = \Theta\big(\frac{\mu r \sigma_1^\star}{n_2}\big)$ *and*

$$\lambda_1\ \alpha_1^2 \leqslant C_2\ \sigma_r^\star,\ \lambda_2\ \alpha_2^2 \leqslant C_2\ \sigma_r^\star,$$

for some positive constant C_2*, we have:*

$$\frac{1}{4}\big[\Delta : \nabla^2 Q\big(\mathcal{W}\big) : \Delta - 4\langle \nabla Q\big(\mathcal{W}\big), \Delta\rangle\big] \leqslant 0.1\,\sigma_r^\star \sum_{k=1}^{n_3} \|\mathcal{F}\big(\Delta\big)^{(k)}\|_F^2.$$

Proof: Define

$$Q_1\big(\mathcal{U}\big) = \lambda_1 \sum_{k=1}^{n_3}\sum_{i=1}^{n_1} \Big(\|\mathcal{F}\big(e_i^\top\big)^{(k)}\mathcal{F}\big(\mathcal{U}\big)^{(k)}\| - \alpha_1\Big)_+^4$$

and

$$Q_2\big(\mathcal{V}\big) = \lambda_2 \sum_{k=1}^{n_3}\sum_{j=1}^{n_2} \Big(\|\mathcal{F}\big(e_j^\top\big)^{(k)}\mathcal{F}\big(\mathcal{V}\big)^{(k)}\| - \alpha_2\Big)_+^4$$

and

$$Q\big(\mathcal{W}\big) = Q_1\big(\mathcal{U}\big) + Q_2\big(\mathcal{V}\big).$$

Proceeding as in the proof of Lemma (4.7.16) of the symmetric case, we obtain:

$$\frac{1}{4}\big[\Delta_{\mathcal{U}} : \nabla^2 Q_1\big(\mathcal{U}\big) : \Delta_{\mathcal{U}} - 4\langle \nabla Q_1\big(\mathcal{U}\big), \Delta_{\mathcal{U}}\rangle\big] \leqslant 0.1 \sum_{k=1}^{n_3} \sigma_r^\star \|\mathcal{F}(\Delta_{\mathcal{U}})^{(k)}\|_F^2$$

$$\frac{1}{4}\big[\Delta_{\mathcal{V}} : \nabla^2 Q_2\big(\mathcal{V}\big) : \Delta_{\mathcal{V}} - 4\langle \nabla Q_2\big(\mathcal{V}\big), \Delta_{\mathcal{V}}\rangle\big] \leqslant 0.1 \sum_{k=1}^{n_3} \sigma_r^\star \|\mathcal{F}(\Delta_{\mathcal{V}})^{(k)}\|_F^2$$

Then, using

$$\|\mathcal{F}(\Delta_{\mathcal{U}})^{(k)}\|_F^2 \leqslant \|\mathcal{F}(\Delta)^{(k)}\|_F^2$$

and

$$\|\mathcal{F}(\Delta_{\mathcal{V}})^{(k)}\|_F^2 \leqslant \|\mathcal{F}(\Delta)^{(k)}\|_F^2,$$

we have:

$$\frac{1}{4}\big[\Delta : \nabla^2 Q\big(\mathcal{W}\big) : \Delta - 4\langle \nabla Q\big(\mathcal{W}\big), \Delta\rangle\big] \leqslant 0.1\,\sigma_r^\star \sum_{k=1}^{n_3} \|\mathcal{F}(\Delta)^{(k)}\|_F^2.$$

□

Chapter 5

Conclusion générale et perspectives

L'objectif de la thèse était de généraliser la pénalisation ℓ_1 dans les problèmes d'estimation sous hypothèse de parcimonie. Pour cela,

- nous avons proposé une étude solide du comportement de l'algorithme LASSO pour le modèle linéaire en affaiblissant les hypothèses sur la matrice design, habituellement utilisées pour l'analyse du LASSO dans le cas où les colonnes de la matrice X sont sélectionnées à partir d'un mélange fini de K distributions gaussienne mais par contre les centres du modèle de mélange satisfont la condition d'incohérence et nous avons obtenu une borne de prédiction de l'erreur. Le prix à payer pour une telle généralité est que nos bornes de prédiction tiennent avec une probabilité petite qui coïncide avec les travaux précédents dans le cas p tendant vers $+\infty$ uniquement lorsque $\mathfrak{s} = 0$ (i.e. répétitions exacte des colonnes de $\mathfrak{C}$),

- nous avons étendu les problèmes de complétion matricielle à ceux de complétion tensorielle de faible rang en nous appuyant sur les travaux proposé par Kilmer et al. Nous avons fourni une analyse théorique de l'estimateur pénalisé par la norme nucléaire. Ces résultats théoriques sont validés numériquement en utilisant des données de Tomographie par Cohérence Optique (OCT) sur des échantillons biologiques. Ces résultats encourageants avec des ensembles de données réelles démontrent la pertinence de l'hypothèse du rang faible pour les applications pratiques. D'autres recherches seront entreprises pour concevoir des algorithmes rapides et incorporer d'autres pénalités comme, par exemple, celles basées sur la sparsité de la transformée de Shearlet.

- nous avons étendu les problèmes de complétion matricielle à ceux de complétion tensorielle de faible rang en nous appuyant sur les travaux proposé par Kilmer et al. Nous avons fourni une analyse théorique de l'estimateur pénalisé par la norme nucléaire. Ces résultats théoriques sont validés numériquement en utilisant des données de Tomographie par Cohérence Optique (OCT) sur des échantnce du problème de complétion de tenseurs à faible rang tubal. Les travaux ultérieurs consisteront

à valider la méthode dans un dispositif d'imagerie physique. Pour ce faire, il sera important d'envisager une mise en oeuvre de l'algorithme sur GPU afin de relever les défis inhérents au traitement en temps réel. Des travaux de recherche supplémentaires peuvent être entrepris pour adapter l'algorithme à un cadre en ligne où les hyperparamètres peuvent être appris en utilisant, par exemple, l'approche de [36].

Bibliography

[1] AlQuraishi, M. and McAdams, H., Direct inference of protein–DNA interactions using compressed sensing methods, PNAS 108, 14819 (2011).

[2] Becker, S., Bobin, J. and Candès, E. J., Nesta: a fast and accurate first-order method for sparse recovery. In press SIAM J. on Imaging Science.

[3] Bickel, P. J., Ritov, Y., Tsybakov, A. B. Simultaneous analysis of lasso and Dantzig selector. Ann. Statist. 37 (2009), no. 4, 1705–1732.

[4] Bock, M.E, Judge, G.G, Yancey, T.A. A simple form for the inverse moments of non-central χ^2 and F random variable and certain confluent hypergeometric functions. Journal of Econometrics 25 (1984); 217–234. North-Holland.

[5] Bousquet, O., A Bennett concentration inequality and its application to suprema of empirical processes, Comptes Rendus Mathematique, 334 (2002), no. 6, 495–500.

[6] Bunea, F., Tsybakov, A., and Wegkamp, M. (2007a). Sparsity oracle inequalities for the Lasso. Electron. J. Stat., 1 :169–194.

[7] Bunea, F., Honest variable selection in linear and logistic regression models via ℓ_1 and $\ell_1 + \ell_2$ penalization , the Electronic Journal of Statistics, (2008) Vol. 2, 1153-1194

[8] Candès, E., Compressive sampling, (2006) 3, International Congress of Mathematics, 1433–1452, EMS.

[9] Candès, E. J. The restricted isometry property and its implications for compressed sensing. C. R. Math. Acad. Sci. Paris 346 (2008), no.

[10] Candès, E. J. and Plan, Yaniv. Near-ideal model selection by ℓ_1 minimization. Ann. Statist. 37 (2009), no. 5A, 285–2177.

[11] Candès, E. and Tao T., Decoding by linear programming. IEEE Information Theory, 51 (2005) no. 12, 4203-4215.

[12] Candès, E. J. and Tao, T., The Dantzig Selector: statistical estimation when p is much larger than n. Ann. Stat.

[13] Chandrasekaran, V., Recht, B., Parrilo, P.A. and Willsky, A.S., 2012. The convex geometry of linear inverse problems. Foundations of Computational mathematics, 12(6), pp.805-849.

[14] Chrétien, S., 2009. An alternating l_1 approach to the compressed sensing problem, IEEE Signal Processing Letters, 17(2), pp.181-184.

[15] Chrétien, S. and Hero, A. and Perdry, H., 2012. Space alternating penalized Kullback proximal point algorithms for maximizing likelihood with nondifferentiable penalty, Annals of the Institute of Statistical Mathematics, 64(4), pp. 791–809.

[16] Chrétien, S. and Darses, Sparse recovery with unknown variance: a LASSO-type approach, ArXiv 2012.

[17] Chrétien, S. and Darses, S. Invertibility of random submatrix via tail decoupling and a matrix Chernoff inequality, Stat. and Prob. Lett. 82 (2012), no. 7, 1479-1487.

[18] Constructing message passing algorithms for compressed sensing D. L. Donoho, A. Maleki, and A. Montanari, submitted to IEEE Trans. Inf. Theory.

[19] Dossal, C., A necessary and sufficient condition for exact recovery by ℓ_1 minimization. http://hal.archives-ouvertes.fr/docs/00/16/47/38/PDF/DossalMinimisationl1.pdf

[20] Fuchs, J.J., On sparse representations in arbitrary redundant bases. IEEE Trans. Info. Th., 2002.

[21] Jaggi, M., 2013, February. Revisiting Frank-Wolfe: Projection-free sparse convex optimization. In International Conference on Machine Learning (pp. 427-435). PMLR.

[22] Kim, S.-J., Koh, K., Boyd, S. and Gorinevsky, D., SIAM Review, problems and techniques section, 51 , (2009), no. 2, 339–360.

[23] Lafond, J., Wai, H.T. and Moulines, E., 2015. On the online Frank-Wolfe algorithms for convex and non-convex optimizations. arXiv preprint arXiv:1510.01171.

[24] Massart, P., Concentration inequalities and model selection. Lectures from the 33rd Summer school on Probability Theory in Saint Flour. Lecture Notes in Mathematics, 1896. Springer Verlag (2007).

[25] Meinshausen, N. and Bühlmann, P., High-dimensional graphs and variable selection with the Lasso, Ann. Statist. 34 (2006), no. 3, 1436-1462.

[26] Neto, D. Sardy, S. and Tseng, P., l1-Penalized Likelihood Smoothing and Segmentation of Volatility Processes allowing for Abrupt Changes, Journal of Computational and Graphical Statistics, 21 (2012), no. 1, 217–233.

[27] Osborne, M.R., Presnell, B. and Turlach, B.A., A new approach to variable selection in least squares problems, IMA J. Numer. Anal. 20 (2000), no. 3, 389–403.

[28] Parikh, N. and Boyd, S., 2014. Proximal algorithms. Foundations and Trends in optimization, 1(3), pp.127-239.

[29] Rudelson, M. and Vershynin, R. Smallest singular value of a random rectangular matrix. Comm. Pure Appl. Math. 62 (2009), no. 12, 1707–20131739.

[30] Tibshirani, R. Regression shrinkage and selection via the LASSO, J.R.S.S. Ser. B, 58, no. 1 (1996), 267–288.

[31] Tropp, J. A. Norms of random submatrices and sparse approximation. C. R. Math. Acad. Sci. Paris 346 (2008), no. 23-24, 1271–1274.

[32] van de Geer, S., High-dimensional generalized linear models and the Lasso. The Annals of Statistics 36, 614-645.

[33] Vershynin, R., Introduction to the non-asymptotic analysis of random matrices, Chapter 5 of the book Compressed Sensing, Theory and Applications, ed. Y. Eldar and G. Kutyniok. Cambridge University Press, 2012. pp. 210–268. [arXiv:1011.3027, Aug 2010].

[34] Wainwright, Martin J., Sharp thresholds for high-dimensional and noisy sparsity recovery using ℓ_1-constrained quadratic programming (Lasso). IEEE Trans. Inform. Theory 55 (2009), no. 5, 2183–2202.

[35] Zhang, T., 2009. Some sharp performance bounds for least squares regression with l1 regularization. The Annals of Statistics, 37(5A), pp.2109-2144.

[36] Candès, E.J.; Recht, B. *Exact matrix completion via convex optimization*, Found. Comput. Math, 2009, **9**, no. 717–772.

[37] Kilmer, M.E.; Martin, C.D. *Factorization strategies for third-order tensors*, Linear Algebra Its Appl, 2011, **435**, no. 641–658.

[38] Kilmer, M.E.; Braman, K.; Hao, N.; Hoover, R.C. *Third-order tensors as operators on matrices: A theoretical and computational framework with applications in imaging*, Siam J. Matrix Anal. Appl, 2013,**34**, no. 148–172.

[39] Koltchinskii, V.; Lounici, K.; Tsybakov, A.B. *Nuclear-norm penalization and optimal rates for noisy low-rank matrix completion*, Ann. Stat, 2011, **39**, no. 2302–2329.

[40] Candès, E.J. *Mathematics of sparsity (and a few other things)*, In Proceedings of the International Congress of Mathematicians, Seoul, Korea, 13–21 August 2014.

[41] Candès, E.J.; Romberg, J.; Tao, T. *Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information*, IEEE Trans. Inf. Theory, 2006, **52**, no. 489–509.

[42] Recht, B.; Fazel, M.; Parrilo, P.A. *Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization*, SIAM Rev, 2010, **52**, no. 471–501.

[43] Hackbusch, W. *Tensor Spaces and Numerical Tensor Calculus*; Springer Science & Business Media: Berlin, Germany, 2012; **42**.

[44] Landsberg, J.M. *Tensors: Geometry and Applications*, American Mathematical Society: Providence, RI, USA, 2012; **128**.

[45] Sacchi, M.D.; Gao, J.; Stanton, A.; Cheng, J. *Tensor Factorization and its Application to Multidimensional Seismic Data Recovery*, In SEG Technical Program Expanded Abstracts, Society of Exploration Geophysicists: Houston, TX, USA, 2015.

[46] Yuan, M.; Zhang, C.H. *On tensor completion via nuclear norm minimization*, Found. Comput. Math, 2016, **16**, no. 1031–1068.

[47] Signoretto, M.; Van de Plas, R.; De Moor, B.; Suykens, J.A. *Tensor versus matrix completion: a comparison with application to spectral data*, Signal Process. Lett. IEEE, 2011, **18**, no. 403–406.

[48] Candes, E.J.; Plan, Y. *Matrix completion with noise*, Proc. IEEE, 2010, **98**, no. 925–936.

[49] Anandkumar, A.; Ge, R.; Hsu, D.; Kakade, S.M.; Telgarsky, M. *Tensor decompositions for learning latent variable models*, J. Mach. Learn. Res, 2014, **15**, no. 2773–2832.

[50] Signoretto, M.; Dinh, Q.T.; De Lathauwer, L.; Suykens, J.A. *Learning with tensors: a framework based on convex optimization and spectral regularization*, Mach. Learn, 2014, **94**, no. 303–351.

[51] So, A.M.C.; Ye, Y. *Theory of semidefinite programming for sensor network localization*, Math. Program, 2007, **109**, no. 367–384.

[52] Eriksson, B.; Balzano, L.; Nowak, R. *High-rank matrix completion and subspace clustering with missing data*, arXiv, 2011, arXiv:1112.5629.

[53] Singer, A.; Cucuringu, M. *Uniqueness of low-rank matrix completion by rigidity theory*, SIAM J. Matrix Anal. Appl., 2010, **31**, no. 1621–1641.

[54] Fazel, M. *Matrix Rank Minimization with Applications*, Ph.D. Thesis, Stanford University, Stanford, CA, USA, 2002.

[55] Gross, D. *Recovering low-rank matrices from few coefficients in any basis*, Inf. Theory IEEE Trans., 2011, **57**, no. 1548–1566.

[56] Recht, B. *A simpler approach to matrix completion*, J. Mach. Learn. Res., 2011, **12**, no. 3413–3430.

[57] Candès, E.J.; Tao, T. *The power of convex relaxation: Near-optimal matrix completion*, Inf. Theory IEEE Trans, 2010, **56**,no. 2053–2080.

[58] Kolda, T.G.; Bader, B.W. *Tensor decompositions and applications*, SIAM Rev., 2009, **51**,no. 455–500.

[59] McCullagh, P. *Tensor Methods in Statistics*, Chapman and Hall: London, UK, 1987; **161**.

[60] Rogers, M.; Li, L.; Russell, S.J. *Multilinear Dynamical Systems for Tensor Time Series*, Adv. Neural Inf. Process. Syst., 2013, **26**, no. 2634–2642

[61] Ge, R.; Huang, Q.; Kakade, S.M. *Learning mixtures of gaussians in high dimensions*, arXiv, 2015, arXiv:1503.00424.

[62] Mossel, E.; Roch, S. *Learning nonsingular phylogenies and hidden Markov models*, In Proceedings of the Thirty-Seventh Annual ACM Symposium on Theory of Computing, Baltimore, MD, USA,**22–24** May, 2005; no. 366–375.

[63] Jennrich, R. *A generalization of the multidimensional scaling model of Carroll and Chang*, UCLA Work. Pap. Phon., 1972, **22**, no. 45–47.

[64] Kroonenberg, P.M.; De Leeuw, J. *Principal component analysis of three-mode data by means of alternating least squares algorithms*, Psychometrika, 1980, **45**,no. 69–97.

[65] Bhaskara, A.; Charikar, M.; Moitra, A.; Vijayaraghavan, A. *Smoothed analysis of tensor decompositions*, k In Proceedings of the 46th Annual ACM Symposium on Theory of Computing, New York, NY, USA, 2014; pp. 594–603.

[66] De Lathauwer, L.; De Moor, B.; Vandewalle, J. *A multilinear singular value decomposition*, SIAM J. Matrix Anal. Appl., 2000, **21**,no. 1253–1278.

[67] Kolda, T.G. *Symmetric Orthogonal Tensor Decomposition is Trivial*, arXiv, 2015, arXiv:1503.01375.

[68] Robeva, E.; Seigal, A. *Singular Vectors of Orthogonally Decomposable Tensors*, arXiv, 2016, arXiv:1603.09004.

[69] Hao, N.; Kilmer, M.E.; Braman, K.; Hoover, R.C. *Facial recognition using tensor-tensor decompositions*, SIAM J. Imaging Sci., 2013, **6**,no. 437–463.

[70] Semerci, O.; Hao, N.; Kilmer, M.E.; Miller, E.L. *Tensor-based formulation and nuclear norm regularization for multienergy computed tomography*, Image Process. IEEE Trans., 2014, **23**,no. 1678–1693.

[71] Zhang, Z.; Aeron, S. *Exact tensor completion using t-svd*, IEEE Trans. Signal Process, 2017, **65**,no. 1511–1526.

[72] Pothier, J.; Girson, J.; Aeron, S. *An algorithm for online tensor prediction*, arXiv, 2015, arXiv:1507.07974.

[73] Bickel, P.J.; Ritov, Y.; Tsybakov, A.B. *Simultaneous analysis of Lasso and Dantzig selector*, Ann. Stat., 2009, **37**, no. 1705–1732.

[74] Zou, H. *The adaptive lasso and its oracle properties*, J. Am. Stat. Assoc., 2006, **101**, no. 1418–1429.

[75] Zou, H.; Hastie, T. *Regularization and variable selection via the elastic net*, J. R. Stat. Soc. Ser. (Stat. Methodol.), 2005, **67**,no. 301–320.

[76] Gaïffas, S.; Lecué, G. *Weighted algorithms for compressed sensing and matrix completion*, arXiv, 2011, arXiv:1107.1638.

[77] Xu, Y. *On Higher-order Singular Value Decomposition from Incomplete Data*, arXiv, 2014, arXiv:1411.4324.

[78] Raskutti, G.; Yuan, M. *Convex Regularization for High-Dimensional Tensor Regression*, arXiv, 2015, arXiv:1512.01215.

[79] Barak, B.; Moitra, A. *Noisy Tensor Completion via the Sum-of-Squares Hierarchy*, arXiv, 2015, arXiv:1501.06521.

[80] Gandy, S.; Recht, B.; Yamada, I. *Tensor completion and low-n-rank tensor recovery via convex optimization*, Inverse Probl., 2011, **27**, no. 025010.

[81] Mu, C.; Huang, B.; Wright, J.; Goldfarb, D. *Square Deal: Lower Bounds and Improved Convex Relaxations for Tensor Recovery*, In Proceedings of the International Conference on Machine Learning, Beijing, China, 21–26 June 2014.

[82] Amelunxen, D.; Lotz, M.; McCoy, M.B.; Tropp, J.A. *Living on the edge: Phase transitions in convex programs with random data* Inf. Inference, 2014, **3**, pp.224–294.

[83] Lecué, G.; Mendelson, S. *Regularization and the small-ball method I: sparse recovery*, arXiv, 2016, arXiv:1601.05584.

[84] Watson, G.A. *Characterization of the subdifferential of some matrix norms*, Linear Algebra Its Appl., 1992, **170**, no. 33–45.

[85] Lewis, A.S. *The convex analysis of unitarily invariant matrix functions*, J. Convex Anal., 1995, **2**, no. 173–183.

[86] Hu, S. *Relations of the nuclear norm of a tensor and its matrix flattenings*, Linear Algebra Its Appl., 2015, **478**, no. 188–199.

[87] Chrétien, S.; Wei, T. *Von Neumann's inequality for tensors*, arXiv, 2015, arXiv:1502.01616

[88] Chrétien, S.; Wei, T. *On the subdifferential of symmetric convex functions of the spectrum for symmetric and orthogonally decomposable tensors*, Linear Algebra Its Appl., 2018, **542**, no. 84–100.

[89] Heinze H.Bauschke, P.L. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Canada Mathematical Society, Ottawa, ON, Canada, 2011.

[90] Borwein, J.; Lewis, A.S. *Convex Analysis and Nonlinear Optimization: Theory and Examples*, Springer Science & Business Media: Berlin, Germany, 2010.

[91] Jean-Pierre Aubin, I.E. *Applied Nonlinear Analysis*, Dover Publications: Mineola, NY, USA, 2006, **518**.

[92] Ma, S.; Goldfarb, D.; Chen, L. *Fixed point and Bregman iterative methods for matrix rank minimization.*, Math. Program., 2011, **128**, no. 321–353.

[93] Chrétien, S.; Lohvithee, M.; Sun, W.; Soleimani, M. *Efficient hyper-parameter selection in Total Variation-penalised XCT reconstruction using Freund and Shapire's Hedge approach*, Mathematics, 2020, **8**,no. 493.

[94] Guo, W.; Qin, J.; Yin, W. *A new detail-preserving regularization scheme* SIAM J. Imaging Sci., 2014, **7**, no. 1309–1334.

[95] Tropp, J.A. *User-Frindly tail bounds for sums of random matrices*, arXiv, 2010, arXiv:1004.4389v7.

[96] A. Moitra, *Algorithmic aspects of machine learning*, Cambridge University Press, 2018.

[97] N. Kreimer, A. Stanton, and M. D Sacchi, *Tensor completion based on nuclear norm minimization for 5d seismic data reconstruction*, Geophysics, **78**, 2013, no. 6, V273–V284.

[98] Y. Wang, R. Chen, J. Ghosh, J. C Denny, A. Kho, Y. Chen, B. A Malin, and J. Sun, *Rubik: Knowledge guided tensor factorization and completion for health data analytics*, Int. Conf. on Knowledge Discovery and Data Mining, ACM, 2015, pp. 1265–1274.

[99] N. Li and B. Li, *Tensor completion for on-board compression of hyperspectral images*, IEEE Int. Conf. on Image Processing.

[100] W. Hu, D. Tao, W. Zhang, Y. Xie, and Y. Yang, *The twist tensor nuclear norm for video completion*, IEEE Trans. on Neural Networks and Learning Systems, **28**, 2017, no. 12, 2961–2973.

[101] Ji Liu, Przemyslaw Musialski, Peter Wonka, and Jieping Ye, *Tensor completion for estimating missing values in visual data*, IEEE Trans. on Pattern Analysis and Machine Intelligence, **35**, 2013, no. 1, 208–220.

[102] Andrzej Cichocki, Danilo Mandic, Lieven De Lathauwer, Guoxu Zhou, Qibin Zhao, Cesar Caiafa, and Huy Anh Phan, *Tensor decompositions for signal processing applications: From two-way to multiway component analysis*, IEEE Signal Processing Mag., **32**, 2015, no. 2, 145–163.

[103] Nicholas D Sidiropoulos, Lieven De Lathauwer, Xiao Fu, Kejun Huang, Evangelos E Papalexakis, and Christos Faloutsos, *Tensor decomposition for signal processing and machine learning*, IEEE Trans. on Signal Processing, **65**, 2017, no. 13, 3551–3582.

[104] H. Sedghi, M. Janzamin, and A. Anandkumar, *Provable tensor methods for learning mixtures of generalized linear models*, Artificial Intelligence and Statistics, 2016, pp. 1223–1231.

[105] B. Romera-Paredes and M. Pontil, *A new convex relaxation for tensor completion*, Advances in Neural Information Processing Systems, 2013, pp. 2967–2975.

[106] J. A. Bazerque, G. Mateos, and G. B. Giannakis, *Rank regularization and bayesian inference for tensor completion and extrapolation*, IEEE Trans. on Signal Processing **61**, 2013, no. 22, 5689–5703.

[107] D. Kressner, M. Steinlechner, and B. Vandereycken, *Low-rank tensor completion by riemannian optimization*, BIT Numerical Mathematics, **54**, 2014, no. 2, 447–468.

[108] M. Filipović and A. Jukić, *Tucker factorization with missing data with application to low-rank tensor completion*, Multidimensional Systems and Signal Processing, **26**, 2015, no. 3, 677–692.

[109] S. Chrétien and T. Wei, *Convex recovery of tensors using nuclear norm penalization*, Int. Conf. on Latent Variable Analysis and Signal Separation, Springer, 2015, pp. 360–367.

[110] S. Chrétien and T. Wei, *Sensing tensors with gaussian filters*, IEEE Trans. on Information Theory, **63**, 2017, no. 2, 843–852.

[111] A. Anandkumar, R. Ge, and M. Janzamin, *Guaranteed non-orthogonal tensor decomposition via alternating rank-*1 *updates*, arXiv, preprint arXiv:1402.5180, 2014.

[112] Yuanyuan Liu and Fanhua Shang, *An efficient matrix factorization method for tensor completion*, IEEE Signal Processing Letters, **20** (2013), no. 4, 307–310.

[113] X.-Y. Liu, S. Aeron, V. Aggarwal, and X. Wang, *Low-tubal-rank tensor completion using alternating minimization*, Modeling and Simulation for Defense Systems and Applications XI, vol. 9848, Int. Society for Optics and Photonics, 2016, p. 984809.

[114] D. Wei, A. Wang, X. Feng, B. Wang, and B. Wang, *Tensor completion based on triple tubal nuclear norm*, Algorithms, **11**, 2018, no. 7, 94.

[115] A. Wang, Z. Lai, and Z. Jin, *Noisy low-tubal-rank ten-8sor completion.* Neurocomputing, 330:267-279, 2019.

[116] S. Burer and R. DC Monteiro, *A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization*, Mathematical Programming, **95**, 2003, no. 2, 329–357.

[117] N. Boumal, V. Voroninski, and A. Bandeira, *The non-convex burer-monteiro approach works on smooth semidefinite programs*, Advances in Neural Information Processing Systems, 2016, pp. 2757–2765.

[118] Rong Ge, Chi Jin, and Yi Zheng, *No spurious local minima in nonconvex low rank problems: A unified geometric analysis*, Int. Conf. on Machine Learning, 2017, pp. 1233–1242.

[119] Rong Ge, Furong Huang, Chi Jin, and Yang Yuan, *Escaping From Saddle Points — Online Stochastic Gradient for Tensor Decomposition*, arXiv, preprint arXiv:1503.02101, 2015.

[120] Fujimoto, James G, Pitris, Costas, Boppart, Stephen A and Brezinski, Mark E, Optical coherence tomography: an emerging technology for biomedical imaging and optical biopsy, Neoplasia, **2**, 1-2, 9–25,2000.

[121] Miao J, Huo D, and Wilson DL., *Quantitative image quality evaluation of MR images using perceptual difference models*, Med Phys., **35**, 2008, no. 6, 2541-2553.