

Numerical Assessment of Alternating Projection Methods for Matrix Completion with Application to Sparse Image Reconstruction

Mattia Silei[1], Stefania Bellavia[2], and Simone Rebegoldi[3(✉)]

[1] Dipartimento di Matematica e Informatica, Università degli Studi di Firenze, viale G.B. Morgagni 45, 50134 Firenze, Italy
mattia.silei@unifi.it

[2] Dipartimento di Ingegneria Industriale, Università degli Studi di Firenze, Viale G.B. Morgagni 40, 50134 Firenze, Italy
stefania.bellavia@unifi.it

[3] Dipartimento di Scienze Fisiche, Informatiche e Matematiche, Università degli Studi di Modena e Reggio Emilia, Via Campi 213/B, 41125 Modena, Italy
simone.rebegoldi@unimore.it

Abstract. In this paper, we evaluate the numerical performance of the alternating projection method (APM) and a regularized variant of the same method (RAPM) for matrix completion. Both methods are based on the reformulation of matrix completion as a nonconvex feasibility problem. However, the regularized method shares global convergence guarantees even in the nonconvex setting, unlike its standard counterpart. Numerical experiments on randomly generated Gaussian matrices show that RAPM is much more robust with respect to the choice of the initial guess than APM is, as well as being insensitive to the regularization effect for a wide range of regularization parameters. Preliminary numerical results showing the effectiveness of RAPM on some sparse image reconstruction test problems are also presented.

Keywords: Alternating Projections · Matrix Completion · Image Reconstruction

1 Introduction

Matrix Completion (MC) consists in finding the missing entries of a matrix from the knowledge of a subset of its entries. Examples of MC include the collaborative filtering problem [26,27], data imputation [1,28], shape recovery and motion estimation in computer vision [15,18], signal and image denoising [21,23,30], and image inpainting [8,20].

Letting $M \in \mathbb{R}^{n_1 \times n_2}$ be the matrix whose entries need be recovered, and $\Omega \subseteq \{1, \dots, n_1\} \times \{1, \dots, n_2\}$ the subset of indices corresponding to the known entries of M, the MC problem amounts to finding a matrix $Z \in \mathbb{R}^{n_1 \times n_2}$ satisfying the constraint $Z_{i,j} = M_{i,j}$ for all $(i, j) \in \Omega$. As there exists an infinite number of matrices complying with this condition, MC is a severely ill-posed problem.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2026
O. Gervasi et al. (Eds.): ICCSA 2025 Workshops, LNCS 15892, pp. 430–446, 2026.
https://doi.org/10.1007/978-3-031-97638-4_27

However, in most applications, one can assume that the matrix to be recovered has low rank [11]. Therefore, we can formulate MC as the following affine rank minimization problem

$$\begin{aligned} \min \quad & \operatorname{rank}(Z) \\ \text{s.t.} \quad & Z_{i,j} = M_{i,j},\ \forall\ (i,j) \in \Omega. \end{aligned} \tag{1}$$

It is well known that problem (1) is NP-hard, thus making it computationally intractable when the matrix dimension is large [29]. In order to circumvent this issue, problem (1) has been addressed in several ways in the literature, e.g. by replacing the rank function with a convex function such as the nuclear norm [11], or by reformulating it as a semidefinite programming problem [8].

In this paper, we address problem (1) by reformulating it as a feasibility problem, that is, the problem of computing a point belonging to the intersection of two sets. Given an estimate r of the rank of the sought matrix, we introduce the following sets

$$\mathcal{C} = \{Z \in \mathbb{R}^{n_1 \times n_2} :\ Z_{i,j} = M_{i,j},\ \forall\ (i,j) \in \Omega\}, \tag{2}$$

$$\mathcal{C}_r = \{Z \in \mathbb{R}^{n_1 \times n_2} : \operatorname{rank}(Z) \leq r\}, \tag{3}$$

where $\mathcal{C}$ is the set of matrices that match the observed entries, and $\mathcal{C}_r$ denotes the rank level set, namely the set of matrices whose rank is smaller or equal than r. Note that $\mathcal{C}$ is an affine subset of $\mathbb{R}^{n_1 \times n_2}$, whereas $\mathcal{C}_r$ is nonconvex. Then, one can reformulate MC as the problem of finding a matrix in the intersection of $\mathcal{C}$ and $\mathcal{C}_r$, i.e.,

$$\text{find } Z \in \mathbb{R}^{n_1 \times n_2} \text{ such that } Z \in \mathcal{C} \cap \mathcal{C}_r. \tag{4}$$

This approach has already been explored in the literature, although perhaps not as extensively as other available approaches for MC. Some papers have implicitly employed formulation (4) for applications in signal processing and computer vision [15,25], but only few have explicitly formulated and addressed MC as a feasibility problem, see e.g. [9,16].

The most standard method for solving feasibility problems of the form (4) is the Alternating Projection Method (APM), which dates back to the seminal paper [22] by J. von Neumann, and consists in projecting back and forth onto the two sets until a point in the intersection is found. APM converges to a feasible point for any choice of the initial guess when the two sets are convex [4–7,13]. Nonetheless, when one of the sets is nonconvex, APM converges only locally and further hypotheses on the sets and their intersection need be imposed [24]. Note that APM has been previously applied to MC problems in the literature, see e.g. [9,15,16,25]. A regularized variant of APM (which we name RAPM) has been introduced for the general feasibility problem in [2,3]. Notably, RAPM converges globally even for nonconvex sets, by assuming that the underlying objective function satisfies the Kurdyka-Łojasiewicz inequality on its domain [2, Corollary 3.1]. To the best of our knowledge, RAPM has been applied to the MC problem (4) only in [9], but without an extensive study on its sensitivity with respect to the choice of its regularization parameters.

In this paper, we evaluate the practical performance of APM and RAPM applied to problem (4). First, we perform some numerical experiments on randomly generated Gaussian matrices. From this first set of experiments, we can appreciate the robustness of RAPM with respect to the choice of the regularization parameters; at the same time, we observe the numerical failure of APM for initial guesses sufficiently far away from the solution of the problem. On the one hand, this is in line with what the authors in [16] observed, namely that, for reasonable choices of the initial guess, APM converges and that theoretical conditions for convergence are much stricter than the practical ones. On the other hand, to the best of our knowledge, this is the first numerical example where APM applied to the MC problem actually converges only locally, showing an initial guess for which the method fails to converge. In conclusion, this example is representative of both the theoretical and practical aspects of APM convergence when applied to MC, while also showing the benefits of adopting a regularization approach in this context. As a second experiment, we apply RAPM in order to recover images in which only half of the pixels are observed. Such a task may arise due to a variety of factors, including but not limited to: sensor malfunctioning, the overimposition of text, the presence of scratches, or the detection of noise [12,31]. We refer to the considered problem as sparse image reconstruction, which falls under the broader term of image inpainting [12].

The paper is organized as follows. In Sect. 2, we present APM and RAPM applied to the MC problem given in (4). Section 3 is devoted to the description of the numerical results on the two considered tests. In Sect. 4, we give the conclusions and propose some directions for future work.

2 Alternating Projection Methods for Matrix Completion

The aim of this section is to describe APM and its regularized variant RAPM when applied to the MC problem (4). Furthermore, we discuss their convergence according to some existing results in the literature, which hold thanks to the analytical properties satisfied by the sets $\mathcal{C}$ and $\mathcal{C}_r$.

2.1 Notations and Basic Definitions

The symbol $\|\cdot\|$ denotes the Frobenius norm on the matrix space $\mathbb{R}^{n_1\times n_2}$.

We write $(M)_\Omega \in \mathbb{R}^{n_1\times n_2}$ to denote the matrix obtained by keeping the entries of $M \in \mathbb{R}^{n_1\times n_2}$ with indices in $\Omega \subseteq \{1,\dots,n_1\}\times\{1,\dots,n_2\}$ and setting to 0 the others, i.e.,

$$((M)_\Omega)_{i,j} = \begin{cases} M_{i,j}, & \text{if } (i,j)\in\Omega \\ 0, & \text{otherwise,} \end{cases}$$

and similarly we define $(M)_{\bar{\Omega}}$ with $\bar{\Omega}$ the complementary of Ω.

Given a non-empty, closed set $\mathcal{A} \subseteq \mathbb{R}^n$, we denote with $\iota_{\mathcal{A}} : \mathbb{R}^n \to \bar{\mathbb{R}}$ the indicator function of the set $\mathcal{A}$, namely,

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

We recall that the projection operator $\mathcal{P}_{\mathcal{A}} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ onto the set $\mathcal{A}$ is defined by

$$\mathcal{P}_{\mathcal{A}}(x) = \operatorname*{argmin}_{y \in \mathcal{A}} \frac{1}{2}\|y - x\|^2 = \operatorname*{argmin}_{y \in \mathbb{R}^n} \frac{1}{2}\|y - x\|^2 + \iota_{\mathcal{A}}(y).$$

A set $\mathcal{A} \subseteq \mathbb{R}^n$ is called *semialgebraic* if it assumes the form [10]

$$A = \bigcup_{i=1}^{p} \bigcap_{j=1}^{q} \{x \in \mathbb{R}^n : \ f_{i,j}(x) = 0, \ g_{i,j}(x) < 0\}, \tag{5}$$

where $f_{i,j}, g_{i,j} : \mathbb{R}^n \to \mathbb{R}$ are polynomial functions for all $1 \leq i \leq p$, $1 \leq j \leq q$. The proximal operator of a proper, lower semicontinuous function $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ is denoted by $\operatorname{prox}_f : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ and defined as

$$\operatorname{prox}_f(x) = \operatorname*{argmin}_{y \in \mathbb{R}^n} \frac{1}{2}\|y - x\|^2 + f(y).$$

Given a matrix $Z \in \mathbb{R}^{n_1 \times n_2}$, we denote its Singular Value Decomposition (SVD) as

$$Z = U \Sigma V^T,$$

where $U \in \mathbb{R}^{n_1 \times n_1}$, $V \in \mathbb{R}^{n_2 \times n_2}$ are orthogonal matrices and $\Sigma \in \mathbb{R}^{n_1 \times n_2}$ is the rectangular diagonal matrix with the diagonal entries $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_{\min(n_1,n_2)} \geq 0$ denoting the singular values of Z. Given $1 \leq r \leq \min\{n_1, n_2\}$, we consider the matrix partitions $U = [U_1 \ U_2]$, $V = [V_1 \ V_2]$, and $\Sigma = \operatorname{diag}(\Sigma_1, \Sigma_2)$, with $U_1 \in \mathbb{R}^{n_1 \times r}, U_2 \in \mathbb{R}^{n_1 \times (n_1 - r)}, V_1 \in \mathbb{R}^{n_2 \times r}, V_2 \in \mathbb{R}^{n_2 \times (n_2 - r)}, \Sigma_1 \in \mathbb{R}^{r \times r}, \Sigma_2 \in \mathbb{R}^{(n_1 - r) \times (n_2 - r)}$. With these partitions, we can write

$$Z = U_1 \Sigma_1 V_1^T + U_2 \Sigma_2 V_2^T = [Z]_1 + [Z]_2.$$

The matrix $[Z]_1 = U_1 \Sigma_1 V_1^T$ is the truncated SVD to the first r singular values, also referred to as r-truncated SVD.

2.2 Methods and Related Convergence Results

Let $\mathcal{C}$ be the set given in (2), i.e., the set of matrices matching the observed entries of M. It is easy to see that the projection operator onto $\mathcal{C}$ is given by

$$\mathcal{P}_{\mathcal{C}}(Z) = (M)_{\Omega} + (Z)_{\bar{\Omega}}. \tag{6}$$

Hence, the projection onto $\mathcal{C}$ is known in closed form and is cheap, as it consists of $n_1 \times n_2$ element-wise trivial operations.

Regarding the rank level set $\mathcal{C}_r$ given in (3), the well-known result by Eckart and Young [14, Theorem 2.4.8] guarantees that the projection onto $\mathcal{C}_r$ can be computed in closed form. In particular, given a matrix $Z \in \mathbb{R}^{n_1 \times n_2}$ with $r < \text{rank}(Z) \leq \min\{n_1, n_2\}$, the nearest matrix of rank r to Z (with respect to $\|\cdot\|$) is given by its r-truncated SVD, i.e.,

$$\mathcal{P}_{\mathcal{C}_r}(Z) = [Z]_1 = U_1 \Sigma_1 V_1^T. \tag{7}$$

Now consider the MC problem (4) and the related minimization problem

$$\underset{(X,Y)\in\mathbb{R}^n \times \mathbb{R}^n}{\text{argmin}} \quad L(X,Y) \equiv \frac{1}{2}\|X - Y\|^2 + \iota_{\mathcal{C}}(X) + \iota_{\mathcal{C}_r}(Y). \tag{8}$$

The standard APM for solving (4) can be seen as the standard alternating minimization (or Gauss-Seidel) method applied to problem (8), where the function is cyclically minimized with respect to each block of variables, while keeping the other one fixed. Given $Y^0 \in \mathcal{C}_r$, it generates two sequences $\{X^k\}_{k\in\mathbb{N}} \subseteq \mathcal{C}$ and $\{Y^k\}_{k\in\mathbb{N}} \subseteq \mathcal{C}_r$ defined as follows:

$$X^{k+1} \in \mathcal{P}_{\mathcal{C}}(Y^k) = \underset{X\in\mathbb{R}^{n_1\times n_2}}{\text{argmin}} \; L(X, Y^k) = \underset{X\in\mathbb{R}^n}{\text{argmin}} \; \frac{1}{2}\|X - Y^k\|^2 + \iota_{\mathcal{C}}(X) \tag{9}$$

$$Y^{k+1} \in \mathcal{P}_{\mathcal{C}_r}(X^{k+1}) = \underset{Y\in\mathbb{R}^{n_1\times n_2}}{\text{argmin}} \; L(X^{k+1}, Y) = \underset{Y\in\mathbb{R}^{n_1\times n_2}}{\text{argmin}} \; \frac{1}{2}\|X^{k+1} - Y\|^2 + \iota_{\mathcal{C}_r}(Y). \tag{10}$$

Since $\mathcal{C}_r$ is nonconvex, the global convergence of method (9)-(10) to a solution of problem (4) can not be guaranteed. On the other hand, local convergence results for APM applied to the MC problem have been provided in [9,16]. Note that both results rely on the well-known fact that $\mathcal{C}_r$ is prox-regular at all matrices $\bar{X} \in \mathbb{R}^{n_1\times n_2}$ with $\text{rank}(\bar{X}) = r$, which means that the projection operator $\mathcal{P}_{\mathcal{C}_r}$ is single-valued around $\bar{X}$, see [19]. However, the theorem in [16] is based on a previous result given in [17, Theorem 5.16] requiring that the two sets have a linearly regular intersection, a property that the authors in [16] do not explicitly prove for the sets $\mathcal{C}$ and $\mathcal{C}_r$. On the other hand, the authors in [9] deduce the local convergence of APM applied to (4), as well as its convergence rate, by relying on a result given in [24], which requires that the sets intersect separably; as noted in [9], this property is satisfied by $\mathcal{C}$ and $\mathcal{C}_r$ as they are both semialgebraic. For the sake of completeness, the local convergence result and the convergence rate given in [9] are reported below.

Theorem 1. *[9, Theorem 4.3] Suppose that there exists $\bar{X} \in \mathcal{C} \cap \mathcal{C}_r$ such that rank$(\bar{X}) = r$. Let $\{X^k\}_{k\in\mathbb{N}} \subseteq \mathbb{R}^{n_1\times n_2}$ and $\{Y^k\}_{k\in\mathbb{N}} \subseteq \mathbb{R}^{n_1\times n_2}$ be the sequences generated by the method given in (9)-(10). Then, if (X_0, Y_0) is sufficiently close to the intersection of $\mathcal{C}$ and $\mathcal{C}_r$, then the sequences $\{X^k\}_{k\in\mathbb{N}}$ and $\{Y^k\}_{k\in\mathbb{N}}$ converge to a matrix $Z \in \mathbb{R}^{n_1\times n_2}$ with $Z \in \mathcal{C} \cap \mathcal{C}_r$. Furthermore, there exists $\omega \in (0, 2)$ such that $\|X^k - Z\| = \mathcal{O}(k^{-\frac{2-\omega}{2\omega}})$ and $\|Y^k - Z\| = \mathcal{O}(k^{-\frac{2-\omega}{2\omega}})$.*

In [2,3], the authors propose a regularized modification of APM, obtained by proximal modification of the two alternating steps. Given two strictly positive sequences $\{\lambda_k\}_{k\in\mathbb{N}}$ and $\{\mu_k\}_{k\in\mathbb{N}}$, the regularized APM generates two sequences $\{X^k\}_{k\in\mathbb{N}} \subseteq \mathcal{C}$ and $\{Y^k\}_{k\in\mathbb{N}} \subseteq \mathcal{C}_r$ such that:

$$X^{k+1} \in \operatorname{prox}_{\lambda_k L(\cdot,Y^k)}(X^k) = \operatorname*{argmin}_{X\in\mathbb{R}^{n_1\times n_2}} \frac{1}{2}\|X-Y^k\|^2 + \frac{1}{2\lambda_k}\|X-X^k\|^2 + \iota_{\mathcal{C}}(X),$$

$$Y^{k+1} \in \operatorname{prox}_{\mu_k L(X^{k+1},\cdot)}(Y^k) = \operatorname*{argmin}_{Y\in\mathbb{R}^{n_1\times n_2}} \frac{1}{2}\|Y-X^{k+1}\|^2 + \frac{1}{2\mu_k}\|Y-Y^k\|^2 + \iota_{\mathcal{C}_r}(Y).$$

Some easy algebraic manipulations show that the matrices X^{k+1} and Y^{k+1} can be interpreted as projections of weighted means of the current iterates onto $\mathcal{C}$ and $\mathcal{C}_r$, respectively, namely:

$$X^{k+1} \in \mathcal{P}_{\mathcal{C}}\left(\frac{1}{\lambda_k+1}X^k + \frac{\lambda_k}{\lambda_k+1}Y^k\right), \quad Y^{k+1} \in \mathcal{P}_{\mathcal{C}_r}\left(\frac{1}{\mu_k+1}Y^k + \frac{\mu_k}{\mu_k+1}X^{k+1}\right).$$

The regularized APM (RAPM) is sketched in Algorithm 1.

Algorithm 1. Regularized Alternating Projection Method (RAPM) for MC problem (4)

Choose $X^0 \in \mathcal{C}$, $Y^0 \in \mathcal{C}_r$, $0 < \lambda_- < \lambda_+$, $0 < \mu_- < \mu_+$, $\{\lambda_k\}_{k\in\mathbb{N}} \subseteq [\lambda_-, \lambda_+]$, $\{\mu_k\}_{k\in\mathbb{N}} \subseteq [\mu_-, \mu_+]$.
For $k = 0, 1, \ldots$

Step 1 Compute $X^{k+1} \in \mathcal{P}_{\mathcal{C}}\left(\frac{1}{\lambda_k+1}X^k + \frac{\lambda_k}{\lambda_k+1}Y^k\right)$.
Step 2 Compute $Y^{k+1} \in \mathcal{P}_{\mathcal{C}_r}\left(\frac{1}{\mu_k+1}Y^k + \frac{\mu_k}{\mu_k+1}X^{k+1}\right)$.

The authors in [9] deduce the convergence properties of Algorithm 1 by relying on the convergence analysis of RAPM available in [2], which is carried out by assuming that the function L satisfies the so-called Kurdyka-Łojasiewicz (KL) inequality [10] on its domain. As explained in [9], the KL inequality holds for the MC problem (4) thanks to the fact that both sets $\mathcal{C}$ and $\mathcal{C}_r$ are semialgebraic. Unlike APM, the RAPM algorithm enjoys both local and global convergence guarantees, as well as a convergence rate result on the gap between the iterates and the limit point. We summarize the result derived in [9] below.

Theorem 2. *[9, Theorem 4.4] Let $\{X^k\}_{k\in\mathbb{N}}$ and $\{Y^k\}_{k\in\mathbb{N}}$ be the sequences generated by Algorithm 1.*

(i) (Global convergence) Either $\|(X^k, Y^k)\| \to \infty$, or the sequence $\{(X^k, Y^k)\}_{k\in\mathbb{N}}$ converges to a stationary point of the function L.
(ii) (Local convergence) Let $(\bar{X}, \bar{Y})$ be any point such that $\|\bar{X}-\bar{Y}\| = \min\{\|X-Y\| : X \in \mathcal{C}, Y \in \mathcal{C}_r\}$. If (X^0, Y^0) is sufficiently close to $(\bar{X}, \bar{Y})$, then the

whole sequence $\{(X^k, Y^k)\}_{k\in\mathbb{N}}$ *converges to a point* (X_∞, Y_∞) *such that* $\|X_\infty - Y_\infty\| = \min\{\|X - Y\| : \ X \in \mathcal{C}, \ Y \in \mathcal{C}_r\}$.

(iii) (Rate of convergence) Suppose that the sequence $\{(X^k, Y^k)\}_{k\in\mathbb{N}}$ *converges to* $(X^*, Y^*) \in \mathcal{C} \times \mathcal{C}_r$. *Then, the sequence* $\{(X^k, Y^k)\}_{k\in\mathbb{N}}$ *satisfies one of the following: it terminates in a finite number of iterations; there exists* $\tau \in [0, 1)$ *such that* $\|(X^k - Y^k) - (X^*, Y^*)\| = \mathcal{O}(\tau^k)$*; there exists* $\theta \in (\frac{1}{2}, 1)$ *such that* $\|(X^k - Y^k) - (X^*, Y^*)\| = \mathcal{O}\left(k^{-(1-\theta)/(2\theta-1)}\right)$.

3 Numerical Experiments

In this section, we aim at investigating the numerical behaviour of APM and RAPM and whether the theoretical results reflect the practical performances of the two methods. First, we report some numerical results on MC with random Gaussian matrices; then, we show how APM and RAPM perform on MC problems arising in sparse image reconstruction.

The set of experiments on random Gaussian matrices is concerned with the effect of regularization on reconstruction accuracy. We recall that regularizing is equivalent to projecting a convex combination of two consecutive points of the generated sequences. Then, it is evident that this slows down the convergence of RAPM with respect to APM. Moreover, the larger the values of λ, μ the more negligible the regularization, and hence the more the RAPM's convergence history is expected to resemble the APM's one. On the other hand, regularization was introduced to overcome the local convergence of APM. In [16], the authors commented that APM resulted robust to the initial point choice and that it could be considered to converge globally in practice. Interestingly, our experiments show that APM converges only locally, albeit with a very large convergence radius, while regularization further enlarges the convergence radius of RAPM.

The second set of experiments regards the problem of sparse image reconstruction, i.e., restoring an image in the presence of missing pixels. A grayscale image can be seen as a matrix, where each pixel corresponds to an entry of the matrix in the interval $[0, 1]$. In this sense, recovering the missing pixels of an image can be seen as an instance of the MC problem; hence, it can be approached with APM and RAPM. In this context, the low rank assumption corresponds to the hypothesis that the information in the image depends on a small number of parameters, which is a common assumption in the literature.

Before showing the results, we give some more details on the experimental settings.

3.1 Experimental Settings

We recall that a square matrix $M \in \mathbb{R}^{n\times n}$ with rank r has $d_r = r(2n-r)$ degrees of freedom. For the MC problem, we consider to observe m entries uniformly sampled at random. Then we denote by $p = m/n^2$ the proportion of observed entries and $over = m/d_r$ the *oversampling ratio*, that is, how many observed entries there are per degree of freedom. As observed in [20], the oversampling

ratio is an index of the "hardness" of the MC problem at hand: the larger it is, the easier the MC problem is. Then, in the following experiments we specify the triple $(n, r, over)$ to denote the configuration of the MC problem considered.

Another important role is played by the matrix family considered for the experiments. For the first set of experiments, we use the family of Gaussian matrices, which is very common in the MC literature. They are fixed-rank random matrices and are built as $M = LR^T$, where $L, R \in \mathbb{R}^{n\times r}$ are dense matrices with entries sampled from a normal distribution. For the experiments on image reconstruction, we use the pepper and starfish images (see Fig. 4 on the left) which are taken from a widely used set in the literature [31].

A maximum number of iterations k_f and a tolerance tol on the distance between two iterates of the generated sequences are fixed to stop the iterations of the two methods. In particular, given two matrices X and Y, we denote

$$d(X, Y) = \frac{\|X - Y\|}{n} \tag{11}$$

and use $d(X^k, Y^k)$ to monitor the convergence of the methods. For evaluating the final reconstruction, we consider the *mean squared error* and the *Peak Signal to Noise Ratio*, a metric especially designed for images. In particular, the mean squared error can be computed either on a subset of the entries, e.g., Ω, or on the whole matrix to evaluate the final reconstruction. The metrics are defined as follows:

$$MSE(Y^k) = \frac{\|Y^k - M\|^2}{n^2} \tag{12}$$

$$MSE_\Omega(Y^k) = \frac{\|(Y^k - M)_\Omega\|^2}{|\Omega|} \tag{13}$$

$$PSNR(X^{k_f}) = 10 \log_{10}\left(\frac{255^2}{MSE(X^{k_f})}\right). \tag{14}$$

In the following experiments, we compare the reconstruction metrics provided by the last iterate of each method either after performing a fixed number of iterations or stopped whenever $d(X^k, Y^k) \leq tol$, where tol is a given tolerance and $d(X, Y)$ is defined in (11).

Finally, in all of the experiments, the regularization parameters of RAPM are set as constant values, namely $\lambda_k = \lambda$ and $\mu_k = \mu$ for all $k \in \mathbb{N}$.

3.2 Regularization Effect

In the first experiment, we consider Gaussian matrices with $n = 500$, $r = 10$, $over = 4$ and evaluate the effect of regularization on the final accuracy after $k_f = 500$ iterations for 10 random runs. As starting points we use $X^0 = (M)_\Omega, Y^0 = \mathcal{P}_{\mathcal{C}_r}(X^0)$, (only Y^0 for APM). We report the results in Fig. 1. On the left, we show the quantity $\log(REC(Y^{k_f}_{RAPM})/REC(Y^{k_f}_{APM}))$ (mean over 10 runs) for different values of λ, μ. We note that a value of 0 means that RAPM reached

the same accuracy as APM, while the larger the value, the worse the accuracy compared to APM. First, we note that in this case the heatmap is symmetric: this means that regularization on the two sets has the same effect. Moreover, it is evident that for smaller values of parameters, (that is, more regularization) the final accuracy decreases. This can be explained by the fact that regularization slows down convergence and then, final accuracy with fixed number of iterations decreases. This idea is clear in the right plot of Fig. 1, where $MSE_\Omega(Y^k)$ is plotted along iterations for APM (black) and RAPM equipped with different values for the regularization parameters and setting $\lambda = \mu$. The lines represent the mean over 10 runs while the bands correspond to the min-max range.

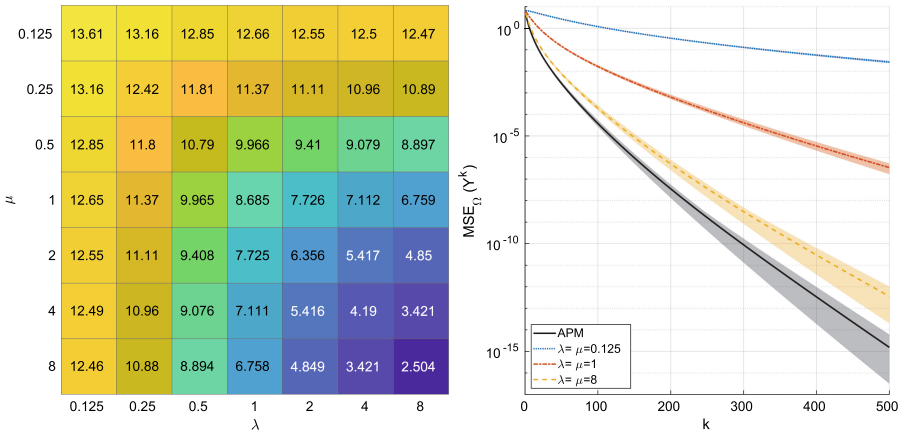

Fig. 1. (Left) Heatmap of the quantity $\log(MSE(Y^{k_f}_{RAPM})/MSE(Y^{k_f}_{APM}))$ (mean over 10 runs) for a grid of values of λ, μ. (Right) Evolution of $MSE_\Omega(Y^k)$ for APM (black) and RAPM for $\lambda = \mu = 0.125, 1, 8$. Lines represent mean over 10 runs, while bands correspond to the min-max range.

In the second experiment we test the convergence radius of APM and RAPM. Since the previous problem was "easy", for this experiment a different configuration is considered. In particular, we choose $n = 200$, $r = 20$, $over = 4.2$ and the starting points are defined as follow: we sample a matrix $Z \sim U([-1, 1])$ and normalize it such that it has unitary norm. Then we consider $X^0 = \mathcal{P}_{\mathcal{C}}(RZ)$, $Y^0 = \mathcal{P}_{\mathcal{C}_r}(RZ)$ for values of R ranging from $R = 10^4$ up to $R = 4 \cdot 10^4$ and we count how many runs out of 10 reach $MSE_\Omega(Y^k) \leq 10^{-6}$. Results are shown in Fig. 2 on the left. We see that APM (black circles) converges 100% of runs till $R = 1.5 \cdot 10^4$ and after that rapidly decreases. On the other hand, RAPM performs better for some combinations of λ, μ and the best one is $(4, 0.5)$ (yellow diamonds), which always converges up to $R = 3 \cdot 10^4$. In Fig. 2 (right) we show the percentage of successful runs for fixed radius $R = 3 \cdot 10^4$ while in Fig. 3 we show the convergence of 5 runs always for $R = 3 \cdot 10^4$. We interestingly note that in this case the heatmap is no longer symmetric and that different levels of regularization on the 2 sets have opposite effects. In particular, we note that small

λ (more regularization on $\mathcal{C}$) reach worse accuracies while better performances are obtained for $\lambda > \mu$. Then, the conclusion is twofold: first, regularization increases in practice the convergence radius; second, the suggestion seems to be to regularize more on $\mathcal{C}_r$ with respect to $\mathcal{C}$. Looking at the plots in Fig. 3, despite the runs' behavior varies a lot, we can observe what follows. First, all of the methods show a stepwise behavior but less pronounced with regularization. We can think of the methods as stalling near local minima without being able to converge to the global minimum. Secondly, even when APM converges, RAPM with the choice $(\lambda, \mu) = (4, 0.5)$ converges faster, making clear that the regularization helps the convergence in this case. Moreover, regularization reduces the stepwise behavior both in terms of the number of iterations and in number of stalling iterations. All of these aspects are further evidence of the effectiveness of regularization.

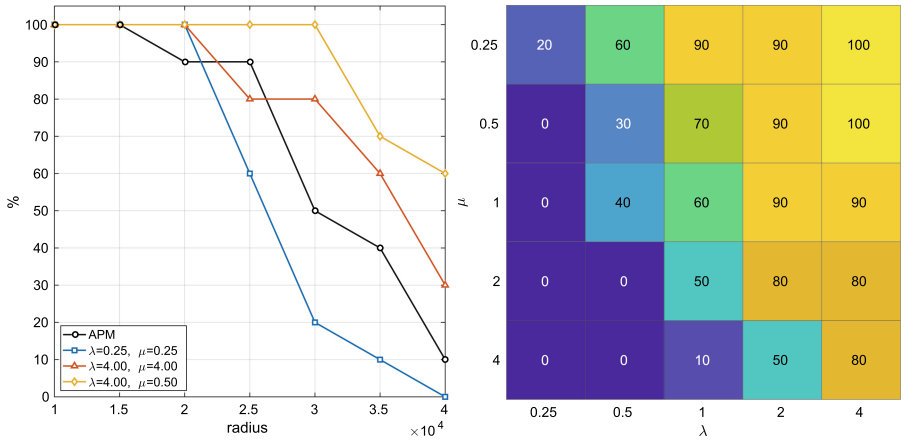

Fig. 2. (Left) Percentage of successful runs (out of 10) versus the radius R. (Right) Percentage of successful runs (out of 10) for fixed radius $R = 3 \cdot 10^4$ on a grid of regularization parameters. A run is considered successful if $MSE_\Omega(Y^k) < 10^{-6}$.

We conclude noting that the setting of the experiment is pretty "extreme" in the sense that we made an unusual choice for the starting points to enforce the observed behavior. On the other hand, we can say that APM performs well on Gaussian matrices with the classic choice of the starting point.

3.3 Sparse Image Reconstruction

We now turn our attention to the problem of image reconstruction with missing pixels, i.e., sparse image reconstruction. We consider the pepper and starfish images that are shown in Fig. 4 (top and bottom left). We process them dividing each pixel value by 255, obtaining 256×256 matrices with entries in $[0, 1]$. The images are full rank but we will show that they can be well approximated by

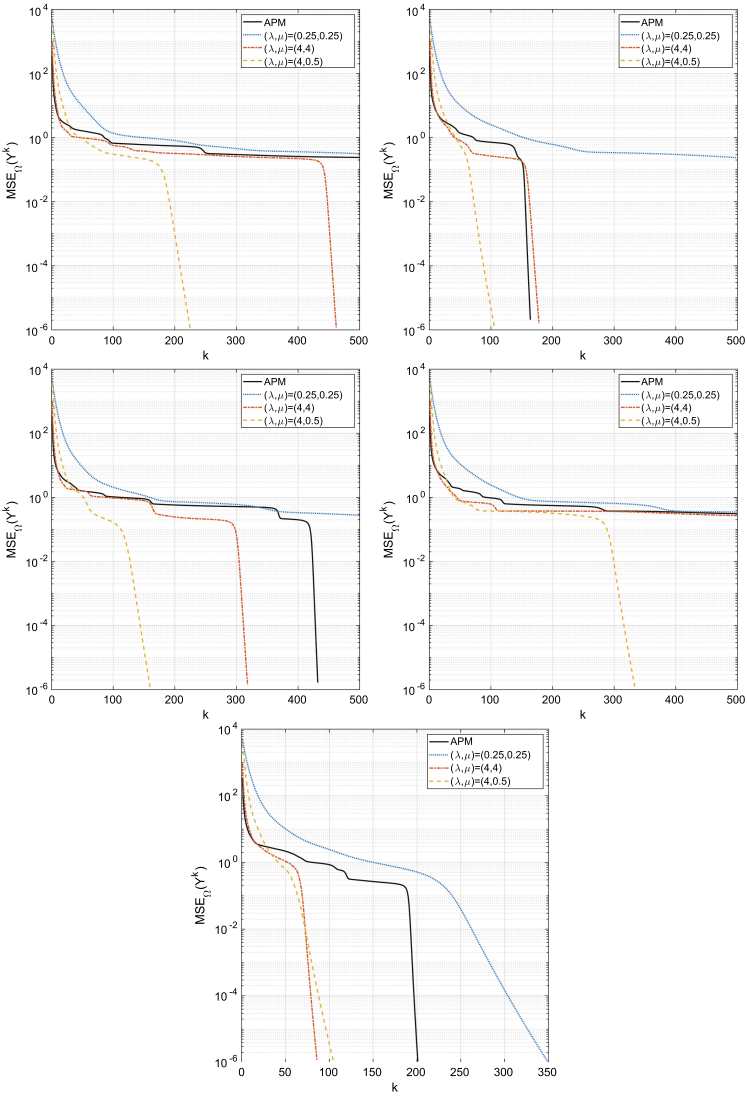

Fig. 3. Plot of the MSE in 5 runs of APM with $R = 3 \cdot 10^4$. We can observe the stepwise behavior of the methods and that regularization reduces the phenomenon. At the same time, it is evident that the configuration $(\lambda, \mu) = (4, 0.5)$ speeds up the convergence in all the reported runs.

small rank matrices, justifying the use of alternating projection methods in our experiments. We observed 50% of the entries sampled uniformly at random and the resulting subsampled images are shown in Fig. 4 (top and bottom right).

Since the rank is unknown, we need to devise a procedure to select the appropriate rank. We use a typical machine learning approach, that is, we keep 10%

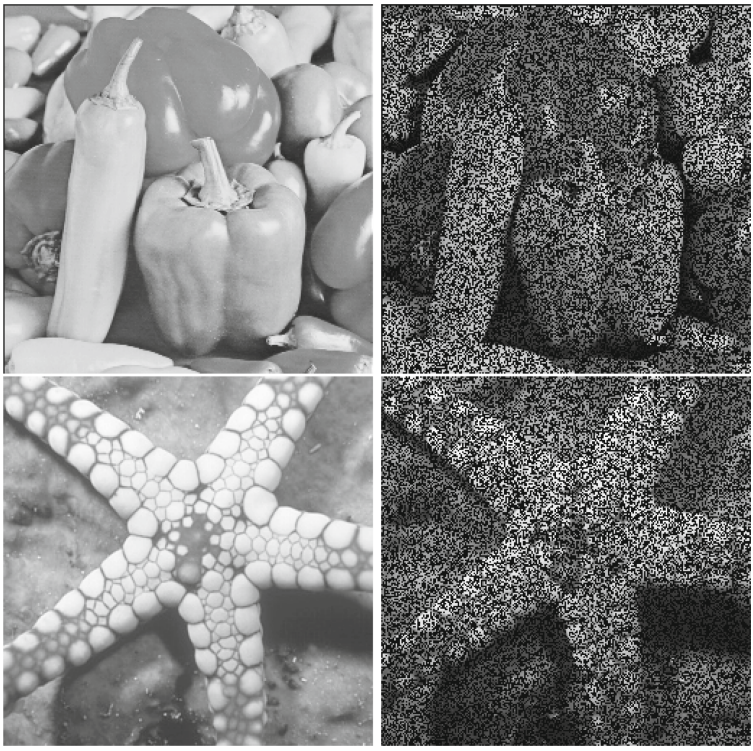

Fig. 4. Original images (left) and subsampled images with 50% of observed pixels (right). Missing pixels are in black.

of the observed entries as validation set, then we run APM for different values of the rank r with $k_f = 250$ and $tol = 10^{-4}$. At termination, we measure the error on the training and validation sets. The results obtained with the two images are shown in Fig. 5. We see that error on the training set (Ω_{tr}) always decreases as the rank increases, on the other hand, the error on the validation set (Ω_{va}) first decreases and then increases, and it can be interpreted as overfitting. In Fig. 6–7, reconstructions for different values of the rank are shown together with the PNSR value. The optimal value of the rank is $r = 20$ and for larger values ($r = 60, 100$) some artifacts appear in the final reconstruction due to the overfitting phenomenon.

With the described procedure we select the rank that achieves the best error on the validation set (that is $r = 20$ for both images) and apply RAPM for different values of λ, μ, with $k_f = 250$ and $tol = 10^{-4}$. The final PSNR is shown in Table 1. Comparing values in the Tables with the PSNR obtained by APM with the optimal choice of the rank (see Figs. 6–7), we see that regularization does not help much in achieving a faster convergence . This is similar to what

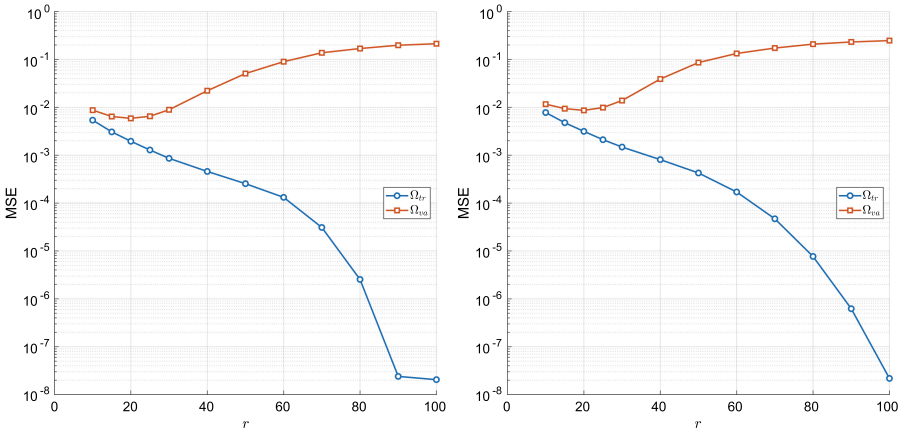

Fig. 5. Results of procedure for selecting the best rank for the two test images.

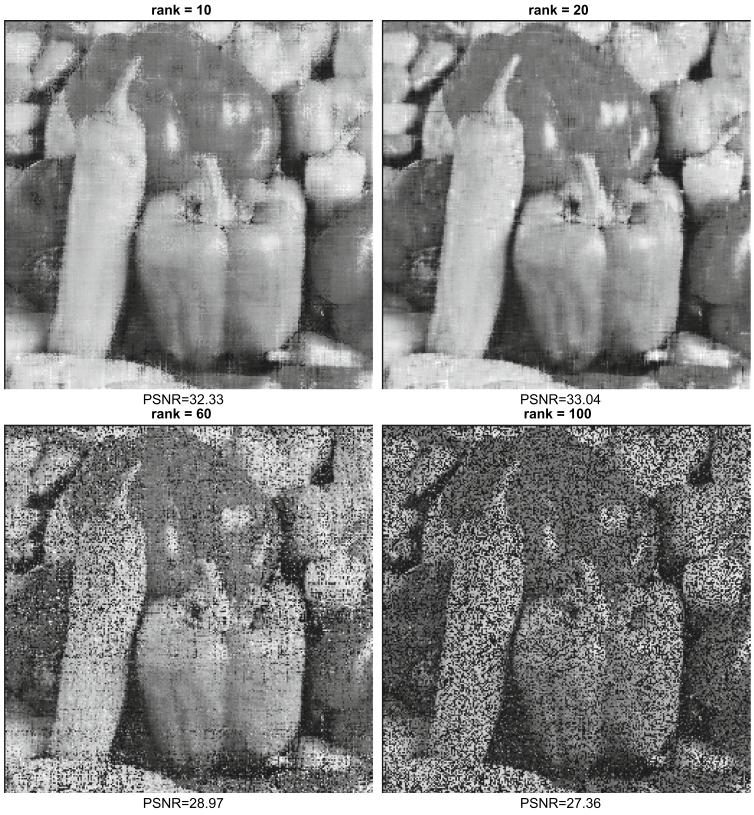

Fig. 6. Reconstructions of the pepper image for 4 different rank values $r = 10, 20, 60, 100$, with the corresponding PSNR values. The rank $r = 20$ is the optimal value selected by the procedure based on the validation set.

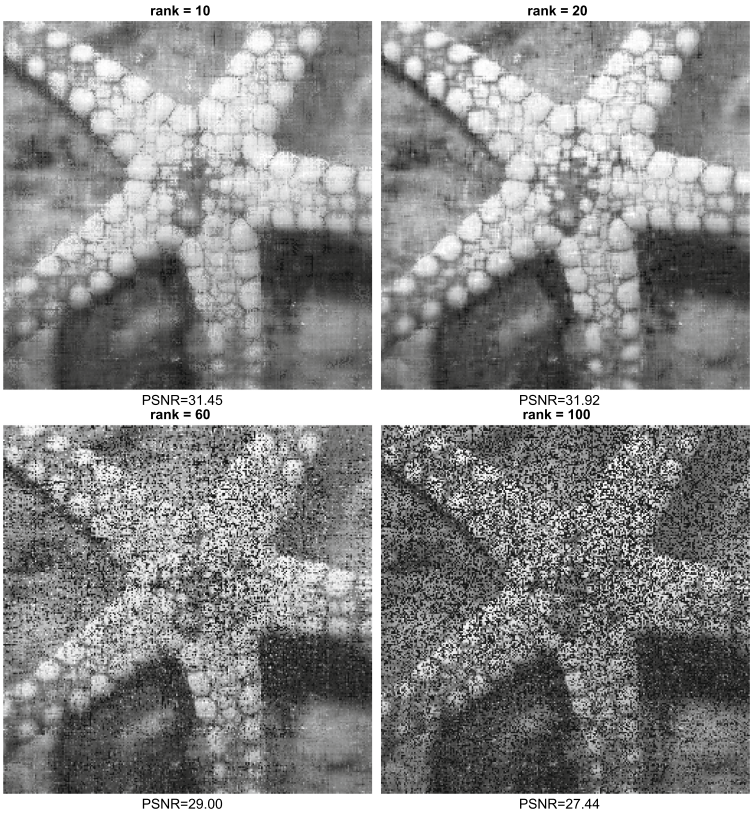

Fig. 7. Reconstructions of the starfish image for 4 different rank values $r = 10, 20, 60, 100$, with the corresponding PSNR values. The rank $r = 20$ is the optimal value selected by the procedure based on the validation set.

Table 1. PSNR values of RAPM for different choices of the regularization parameters (λ, μ) applied to pepper (left) and starfish (right) images.

μ \ λ	1/16	1/4	1	4	16
1/16	26.51	27.55	27.95	28.06	28.09
1/4	27.49	29.66	30.29	30.43	30.46
1	27.88	30.30	30.70	30.71	30.72
4	27.98	30.43	30.72	30.73	30.73
16	28.01	30.45	30.72	30.73	30.73

μ \ λ	1/16	1/4	1	4	16
1/16	26.71	27.41	27.63	27.697	27.72
1/4	27.37	28.69	29.14	29.24	29.27
1	27.60	29.16	29.50	29.52	29.52
4	27.67	29.26	29.51	29.52	29.52
16	27.68	29.28	29.50	29.52	29.51

we noted about the first experiments, where we commented that regularization resulted only in slowing RAPM with respect to APM.

4 Conclusions

In this paper, we present an extended study of the numerical performance of APM and its regularized variant (RAPM) when applied to the MC problem. The numerical experiments presented herein serve to augment the existing compendium on alternating projection methods applied to affine rank minimization problems. In the experiments on Gaussian matrices, a study on the effect of regularization is conducted, which demonstrates that for "easy" problems regularization has similar effects on the two sets; however, for "difficult" problems, it is preferable to regularize more before projecting onto the nonconvex set while reducing regularization on the affine set. Furthermore, a counterexample is presented, demonstrating that APM converges locally in practice. In the context of sparse image reconstruction, it is observed that APM, despite its simplicity, attains satisfactory results in the final reconstruction. Furthermore, a machine learning-inspired procedure is employed for selecting the optimal rank value for the image of interest. This approach enables the application of alternating projection methods even in real-case scenarios where the desired matrix rank is not known in advance.

Acknowledgments. The research that led to the present paper was partially supported by INDAM-GNCS through Progetti di Ricerca 2025 and by PNRR - Missione 4 Istruzione e Ricerca - Componente C2 Investimento 1.1, Fondo per il Programma Nazionale di Ricerca e Progetti di Rilevante Interesse Nazionale (PRIN) funded by the European Commission under the NextGeneration EU programme, projects: "Advanced optimization METhods for automated central veIn Sign detection in multiple sclerosis from magneTic resonAnce imaging (AMETISTA)", code: P2022J9SNP, MUR D.D. financing decree n. 1379 of 1st September 2023 (CUP E53D23017980001), "Numerical Optimization with Adaptive Accuracy and Applications to Machine Learning", code: 2022N3ZNAX MUR D.D. financing decree n. 973 of 30th June 2023 (CUP B53D23012670006), project "Inverse problems in PDE: theoretical and numerical analysis", code: 2022B32J5C, MUR D.D. financing decree n. 973 of 30th June 2023 (CUP B53D23009200006), and by Partenariato esteso FAIR "Future Artificial Intelligence Research" Project "Mathematical and Physical approaches to innovative Machine Learning technologies (MaPLe)".

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Agarwal, A., Amjad, M.J., Shah, D., Shen, D.: Model agnostic time series analysis via matrix estimation. In: Proceedings of the ACM on Measurement and Analysis of Computing Systems, vol. 2, pp. 1–39. ACM New York, NY, USA (2018)
2. Attouch, H., Bolte, J., Redont, P., Soubeyran, A.: Proximal alternating minimization and projection methods for nonconvex problems: an approach based on the Kurdyka-Łojasiewicz inequality. Math. Oper. Res. **35**(2), 438–457 (2010)

3. Attouch, H., Redont, P., Soubeyran, A.: A new class of alternating proximal minimization algorithms with costs-to-move. SIAM J. Optim. **18**(3), 1061–1081 (2007). https://doi.org/10.1137/060657248, https://doi.org/10.1137/060657248
4. Bauschke, H.H., Borwein, J.M., Lewis, A.S.: The method of cyclic projections for closed convex sets in Hilbert space. Contemp. Math. **204**, 1–38 (1997)
5. Bauschke, H.H., Borwein, J.M.: On the convergence of von Neumann's alternating projection algorithm for two sets. Set-Valued Anal. **1**, 185–212 (1993)
6. Bauschke, H.H., Borwein, J.M.: On projection algorithms for solving convex feasibility problems. SIAM Rev. **38**(3), 367–426 (1996)
7. Behling, R., Bello-Cruz, Y., Santos, L.R.: Infeasibility and error bound imply finite convergence of alternating projections. SIAM J. Optim. **31**(4), 2863–2892 (2021). https://doi.org/10.1137/20M1358669, https://doi.org/10.1137/20M1358669
8. Bellavia, S., Gondzio, J., Porcelli, M.: A relaxed interior point method for low-rank semidefinite programming problems with applications to matrix completion. J. Sci. Comput. **89**(2), 46 (2021)
9. Bellavia, S., Rebegoldi, S., Silei, M.: An Inexact Alternating Projection Method with Application to Matrix Completion. Optimization Online (2025)
10. Bolte, J., Daniilidis, A., Lewis, A.: The Łojasiewicz inequality for nonsmooth subanalytic functions with applications to subgradient dynamical systems. SIAM J. Optim. **17**(4), 1205–1223 (2007)
11. Candes, E., Recht, B.: Exact matrix completion via convex optimization. Commun. ACM **55**(6), 111–119 (2012)
12. Elharrouss, O., Almaadeed, N., Al-Maadeed, S., Akbari, Y.: Image inpainting: a review. Neural Process. Lett. **51**, 2007–2028 (2020)
13. Escalante, R., Raydan, M.: Alternating Projection Methods. Society for Industrial and Applied Mathematics, Philadelphia, PA (2011).https://doi.org/10.1137/9781611971941, https://epubs.siam.org/doi/abs/10.1137/9781611971941
14. Golub, G.H., Van Loan, C.F.: Matrix Computations. JHU Press (2013)
15. Guerreiro, R., Aguiar, P.: Estimation of rank deficient matrices from partial observations: two-step iterative algorithms. In: International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition, pp. 450–466. Springer (2003)
16. Jiang, X., Zhong, Z., Liu, X., So, H.C.: Robust matrix completion via alternating projection. IEEE Signal Process. Lett. **24**(5), 579–583 (2017)
17. Lewis, A., Luke, D.R., Malick, J.: Local linear convergence of alternating and averaged nonconvex projections. Found. Comput. Math. **9**(4), 485–513 (2009)
18. Li, K., Dai, Q., Xu, W., Yang, J., Jiang, J.: Three-dimensional motion estimation via matrix completion. IEEE Trans. Syst. Man Cybern. **42**(2), 539–551 (2012)
19. Luke, D.R.: Prox-regularity of rank constraint sets and implications for algorithms. J. Math. Imag. Vis. **47**, 231–238 (2013)
20. Ma, S., Goldfarb, D., Chen, L.: Fixed point and Bregman iterative methods for matrix rank minimization. Math. Program. **128**(1), 321–353 (2011)
21. Nejati, M., Samavi, S., Derksen, H., Najarian, K.: Denoising by low-rank and sparse representations. J. Vis. Commun. Image Represent. **36**, 28–39 (2016)
22. von Neumann, J.: Functional operators, Annals of Mathematics Studies, vol. 22. Princeton University Press, Princeton, NJ (2010)
23. Nguyen, L.T., Kim, J., Shim, B.: Low-rank matrix completion: a contemporary survey. IEEE Access **7**, 94215–94237 (2019)
24. Noll, D., Rondepierre, A.: On local convergence of the method of alternating projections. Found. Comput. Math. **16**, 425–455 (2016)

25. Oropeza, V., Sacchi, M.: Simultaneous seismic data denoising and reconstruction via multichannel singular spectrum analysis. Geophysics **76**(3), 25–32 (2011)
26. Ramlatchan, A., Yang, M., Liu, Q., Li, M., Wang, J., Li, Y.: A survey of matrix completion methods for recommendation systems. Big Data Mining Anal. **1**(4), 308–323 (2018)
27. Rennie, J., Srebro, N.: Fast maximum margin matrix factorization for collaborative prediction. In: Proceedings of the International Conference of Machine Learning, pp. 713–719 (2005)
28. Silei, M., Bellavia, S., Superchi, F., Bianchini, A.: Recovering corrupted data in wind farm measurements: a matrix completion approach. Energies **16**(4) (2023). https://doi.org/10.3390/en16041674, https://www.mdpi.com/1996-1073/16/4/1674
29. Vandenberghe, L., Boyd, S.: Semidefinite Programming. SIAM Rev. **38**(1), 49–95 (1996)
30. Wang, H., Cai, J., Wang, T., Wei, K.: Fast Cadzow's algorithm and a gradient variant. J. Sci. Comput. **88**(2), 41 (2021)
31. Xu, X., Geng, S.: Blind image inpainting with mixture noise using ℓ_0 and total regularization. Comput. Math. Meth. Med. **2022**(1), 3180612 (2022)