Journal of Machine Learning Research 19 (2018) 1-49 Submitted 4/17; Revised 2/18; Published 9/18

Harmonic Mean Iteratively Reweighted Least Squares for Low-Rank Matrix Recovery

Christian Kümmerle C.KUEMMERLE@TUM.DE
Department of Mathematics
Technische Universität München
Boltzmannstr. 3, 85748 Garching/Munich, Germany

Juliane Sigl JULIANE.SIGL@MA.TUM.DE
Department of Mathematics
Technische Universität München
Boltzmannstr. 3, 85748 Garching/Munich, Germany

Editor: Benjamin Recht

Abstract

We propose a new iteratively reweighted least squares (IRLS) algorithm for the recovery of a matrix $X \in \mathbb{C}^{d_1 \times d_2}$ of rank $r \ll \min(d_1, d_2)$ from incomplete linear observations, solving a sequence of low complexity linear problems. The easily implementable algorithm, which we call harmonic mean iteratively reweighted least squares (`HM-IRLS`), optimizes a non-convex Schatten-p quasi-norm penalization to promote low-rankness and carries three major strengths, in particular for the matrix completion setting. First, we observe a remarkable global convergence behavior of the algorithm's iterates to the low-rank matrix for relevant, interesting cases, for which any other state-of-the-art optimization approach fails the recovery. Secondly, `HM-IRLS` exhibits an empirical recovery probability close to 1 even for a number of measurements very close to the theoretical lower bound $r(d_1 + d_2 - r)$, i.e., already for significantly fewer linear observations than any other tractable approach in the literature. Thirdly, `HM-IRLS` exhibits a locally superlinear rate of convergence (of order $2 - p$) if the linear observations fulfill a suitable null space property. While for the first two properties we have so far only strong empirical evidence, we prove the third property as our main theoretical result.

Keywords: Iteratively Reweighted Least Squares, Low-Rank Matrix Recovery, Matrix Completion, Non-Convex Optimization

1. Introduction

The problem of recovering a low-rank matrix from incomplete linear measurements or observations has gained considerable attention in the last few years due to the omnipresence of low-rank models in different areas of science and applied mathematics. Low-rank models arise in a variety of areas such as system identification (Liu et al., 2013; Liu and Vandenberghe, 2010), signal processing (Ahmed and Romberg, 2015), quantum tomography (Gross et al., 2010; Gross, 2011) and phase retrieval (Candès et al., 2013; Candès et al., 2013; Gross et al., 2015). An instance of this problem of particular importance, e.g., in recommender systems (Srebro et al., 2005; Goldberg et al., 1992; Candès and Recht, 2009), is the *matrix*

completion problem, where the measurements correspond to entries of the matrix to be recovered.

Although the low-rank matrix recovery problem is NP-hard in general, several tractable algorithms have been proposed that allow for provable recovery in many important cases. The *nuclear norm minimization* (`NNM`) approach (Fazel, 2002; Candès and Recht, 2009), which solves a surrogate semidefinite program, is particularly well-understood. For `NNM`, recovery guarantees have been shown for a number of measurements on the order of the information theoretical lower bound $r(d_1 + d_2 - r)$, if r denotes the rank of a $d_1 \times d_2$-matrix (Recht et al., 2010; Candès and Recht, 2009); i.e., for a number of measurements $m \geqslant \rho r(d_1 + d_2 - r)$ with some oversampling constant $\rho \geqslant 1$. Even though `NNM` is solvable in polynomial time, it can be computationally very demanding if the problem dimensions are large, which is the case in many potential applications. Another issue is that although the number of measurements necessary for successful recovery by nuclear norm minimization is of *optimal order*, it is not *optimal*. More precisely, it turns out that the oversampling factor ρ of nuclear norm minimization *needs to be much larger than the oversampling factor of some other, non-convex algorithmic approaches* (Zheng and Lafferty, 2015; Tanner and Wei, 2013).

These limitations of convex relaxation approaches have led to a rapidly growing line of research discussing the advantages of non-convex optimization for the low-rank matrix recovery problem (Jain et al., 2010; Tanner and Wei, 2013; Haldar and Hernando, 2009; Jain et al., 2013; Wen et al., 2012; Tanner and Wei, 2016; Vandereycken, 2013; Wei et al., 2016; Tu et al., 2016). For several of these non-convex algorithmic approaches, recovery guarantees comparable to those of `NNM` have been derived (Candès et al., 2015; Tu et al., 2016; Zheng and Lafferty, 2015; Sun and Luo, 2016). Their advantage is a higher empirical recovery rate and an often more efficient implementation. While there are some results about global convergence of first-order methods minimizing a non-convex objective (Ge et al., 2016; Bhojanapalli et al., 2016) so that a success of the method might not depend on a particular initialization, the assumptions of these results are not always optimal, e.g., in the scaling of the numbers of measurements m in the rank r (Ge et al., 2016, Theorem 5.3). In general, the success of many non-convex optimization approaches relies on a distinct, possibly expensive initialization step.

1.1 Contribution of this paper

In this spirit, we propose a new iteratively reweighted least squares (IRLS) algorithm for the low-rank matrix recovery problem[1] that strives to minimize a non-convex objective function based on the Schatten-p quasi-norm

$$\min_X \|X\|_{S_p}^p \text{ subject to } \Phi(X) = Y, \tag{1}$$

for $0 < p < 1$, where $\Phi : \mathbb{C}^{d_1 \times d_2} \to \mathbb{C}^m$ is the linear measurement operator and $Y \in \mathbb{C}^m$ is the data vector defining the problem. The overall strategy of the proposed IRLS algorithm is to mimic this minimization by a sequence of weighted least squares problems. This strategy

1. The algorithm and partial results were presented at the 12th International Conference on Sampling Theory and Applications in Tallinn, Estonia, July 3–7, 2017. The corresponding conference paper has been published in its proceedings (Kümmerle and Sigl, 2017).

is shared by the related previous algorithms of (Fornasier et al., 2011; Mohan and Fazel, 2012) which minimize (1) by defining iterates as

$$X^{(n+1)} = \min_X \|W_L^{(n)\,\frac{1}{2}} X\|_F^2 \text{ subject to } \Phi(X) = Y, \tag{2}$$

where $W_L^{(n)} \approx (X^{(n)} X^{(n)*})^{\frac{p-2}{2}}$ is a so-called *weight matrix* which *reweights* the quadratic penalty by operating on the column space of the matrix variable. Thus, we call this column-reweighting type of IRLS algorithms `IRLS-col`. Due to the inherent symmetry, it is evident to conceive, still in the spirit of (Fornasier et al., 2011; Mohan and Fazel, 2012), the algorithm `IRLS-row`

$$X^{(n+1)} = \min_X \|W_R^{(n)\,\frac{1}{2}} X^*\|_F^2 \text{ subject to } \Phi(X) = Y \tag{3}$$

with $W_R^{(n)} \approx (X^{(n)*} X^{(n)})^{\frac{p-2}{2}}$, which reweights the quadratic penalty by acting on the *row space* of the matrix variable. We note that even for square dimensions $d_1 = d_2$, `IRLS-col` and `IRLS-row` do not coincide.

In this paper, as an important innovation, we propose the use of a different type of weight matrices, which can be interpreted as the *harmonic mean* of the matrices $W_L^{(n)}$ and $W_R^{(n)}$ above. This motivates the name *harmonic mean iteratively reweighted least squares* (`HM-IRLS`) for the corresponding algorithm. The harmonic mean of the weight matrices of `IRLS-col` and of `IRLS-row` in `HM-IRLS` is able to use the information in both the column and the row space of the iterates, and it also gives rise to a *qualitatively better* behavior than the use of more obvious symmetrizations as, e.g., the arithmetic mean of weight matrices would allow for both in theory and in practice.

We argue that the choice of harmonic mean weight matrices as in `HM-IRLS` leads to an efficient algorithm for the low-rank matrix recovery problem with fast convergence and superior performance in terms of sample complexity, also compared to algorithms based on strategies different from IRLS.

On the one hand, we show that the accumulation points of the iterates of `HM-IRLS` converge to stationary points of a smoothed Schatten-p functional under the linear constraint, as it is known for, e.g., `IRLS-col`, c.f. (Fornasier et al., 2011; Mohan and Fazel, 2012). On the other hand, we extend the theoretical guarantees which are based on a Schatten-p *null space property* (NSP) of the measurement operator (Oymak et al., 2011; Foucart and Rauhut, 2013) to `HM-IRLS`.

Our main theoretical result is that `HM-IRLS` exhibits a locally superlinear convergence rate of order $2-p$ in the neighborhood of a low-rank matrix for the non-convexity parameter $0 < p < 1$ connected to the Schatten-p quasi-norm, if the measurement operator fulfills the mentioned NSP of sufficient order. For $p \ll 1$, this means that the convergence rate is *almost quadratic*.

Although parts of our theoretical results, as in the case of the IRLS algorithms algorithms of Fornasier et al. (2011) and Mohan and Fazel (2012), do not apply to the matrix completion setting, due to the popularity of the problem and for reasons of comparability with other algorithms, we conduct numerical experiments to explore the empirical performance of `HM-IRLS` also for this setting. Surprisingly enough we observe that the theoretical

results comply with our numerical experiments also for matrix completion. In particular, the theoretically predicted local convergence rate of order $2 - p$ can be observed very precisely for this important measurement model as well (see Figures 3 to 5).

This local superlinear convergence rate of `HM-IRLS` is unprecedented by previous IRLS variants such as `IRLS-col` or those that use the arithmetic mean of the one-sided weight matrices: this means that neither can a superlinear rate be verified numerically, nor is it possible to show such a rate by our proof techniques for any other previsouly considered IRLS variant designed for the low-rank matrix recovery problem.

To the best of our knowledge, `HM-IRLS` is the first algorithm for low-rank matrix recovery which achieves superlinear rate of convergence for low complexity measurements as well as for larger problems.

Additionally, we conduct extensive numerical experiments comparing the efficiency of `HM-IRLS` with previous IRLS algorithms as `IRLS-col`, Riemannian optimization techniques (Vandereycken, 2013), alternating minimization approaches (Haldar and Hernando, 2009; Tanner and Wei, 2016), algorithms based on iterative hard thresholding (Kyrillidis and Cevher, 2014; Blanchard et al., 2015), and others (Park et al., 2016), in terms of sample complexity, again for the important case of *matrix completion*.

The experiments lead to the following observation: `HM-IRLS` recovers low-rank matrices systematically with an optimal number of measurements that is very close to the theoretical lower bound on the number of measurements that is necessary for recovery with high empirical probability. We consider this result to be remarkable, as it means that for problems of moderate dimensionality (matrices of $\approx 10^7$ variables, e.g. $(d_1 \times d_2)$-matrices with $d_1 \approx d_2 \approx 3 \cdot 10^3$) *the proposed algorithm needs fewer measurements for the recovery of a low rank matrix than all the state-of-the-art algorithms we included in our experiments* (see Figure 6).

An important practical observation of `HM-IRLS` is that its performance is very robust to the choice of the initialization and that it can be used as a stand-alone algorithm to recover low-rank matrices also starting from a trivial initialization. This is suggested by our numerical experiments since even for random or adversary initializations, `HM-IRLS` converges to the low-rank matrix, even though it is based on an objective function which is highly non-convex. While a complete theoretical understanding of this behavior is not yet achieved, we regard the empirical evidence in a variety of interesting cases as strong. In this context, we consider a proof of the global convergence of `HM-IRLS` for non-convex penalizations under appropriate assumptions as an interesting open problem.

1.2 Organization of the paper

We proceed in the paper as follows. In Section 2, we introduce some notation to be used and provide some background about different reformulations of the Schatten-p quasi-norm in terms of weighted ℓ_2-norms. This leads to the derivation of the harmonic mean iteratively reweighted least squares (`HM-IRLS`) algorithm in Section 3. We present our main theoretical results, the convergence guarantees and the locally superlinear convergence rate for the algorithm in Section 4. Numerical experiments and comparisons to state-of-the-art methods for low-rank matrix recovery are carried out in Section 5. In Section 6, we interpret the algorithm's different steps as minimizations of an auxililary functional with respect to its

arguments and show theoretical guarantees for `HM-IRLS` extending similar guarantees for `IRLS-col`. After this, we detail the proof of the locally superlinear convergence rate under appropriate assumptions on the null space of the measurement operator.

In Appendix A, we provide a short overview about Kronecker and Hadamard products, and end with some deferred proofs in Appendix B and Appendix C.

2. Notation and background

2.1 General notation, Schatten-p and weighted norms

In this section, we explain some of the notation we use in the course of this paper.

The set of matrices $X \in \mathbb{C}^{d_1 \times d_2}$ is denoted by $M_{d_1 \times d_2}$. Unless stated otherwise, vectors $x \in \mathbb{C}^d$ are considered as column vectors. We also use the vectorized form $X_{\mathrm{vec}} = \left[X_1^T, \dots, X_j^T, \dots, X_{d_2}^T\right]^T \in \mathbb{C}^{d_1 d_2}$ of a matrix $X \in M_{d_1 \times d_2}$ with columns X_j, $j \in \{1, \dots, d_2\}$. The reverse recast of a vector $x \in \mathbb{C}^{d_1 d_2}$ into a matrix of dimension $d_1 \times d_2$ is denoted by $x_{\mathrm{mat}(d_1,d_2)} = [X_1, \dots, X_j, \dots, X_{d_2}]$, where $X_j = [x_{(d_1-1)\cdot j+1}, \dots, x_{(d_1-1)\cdot j+d_1}]^T$, $j = 1, \dots, d_2$ are column vectors, or X_{mat} if the dimensions are clear from the context. Obviously, it holds that $X = (X_{\mathrm{vec}})_{\mathrm{mat}}$.

The identity matrix in dimension $d \times d$ is denoted by $\mathbf{I}_d$. With $\mathbf{0}_{d_1 \times d_2} \in M_{d_1 \times d_2}$ and $\mathbf{1}_{d_1 \times d_2} \in M_{d_1 \times d_2}$ we denote the matrices with only 0- or 1-entries respectively. The set of Hermitian matrices is denoted by $H_{d \times d} := \{X \in M_{d \times d} \mid X = X^*\}$. We write $X^+ \in M_{d_1 \times d_2}$ for the Moore-Penrose inverse of the matrix $X \in M_{d_1 \times d_2}$.

Let $\mathcal{U}_d = \{U \in \mathbb{C}^{d \times d}; UU^* = \mathbf{I}_d\}$ denote the set of unitary matrices. Then the singular value decomposition of a matrix $X \in M_{d_1 \times d_2}$ can be written as $X = U\Sigma V^*$ with $U \in \mathcal{U}_{d_1}$, $V \in \mathcal{U}_{d_2}$ and $\Sigma \in M_{d_1 \times d_2}$, where Σ is diagonal and contains the singular values of X such that $\Sigma_{ii} = \sigma_i(X) \geqslant 0$ for $i \in \{1, \dots, \min(d_1, d_2)\}$. We define the *Schatten-p (quasi-)norm* of $X \in M_{d_1 \times d_2}$ as

$$\|X\|_{S_p} := \begin{cases} \operatorname{rank}(X), & \text{for } p = 0, \\ \left[\sum_{j=1}^{\min(d_1,d_2)} \sigma_j^p(X)\right]^{1/p}, & \text{for } 0 < p < \infty, \\ \sigma_{\max}(X), & \text{for } p = \infty. \end{cases} \tag{4}$$

Note that for $p = 1$, the Schatten-p norm is also called *nuclear norm*, written as $\|X\|_* := \|X\|_{S_1}$. The *trace* $\operatorname{tr}[X]$ of a matrix $X \in M_{d_1 \times d_2}$ is defined by the sum of its diagonal elements, $\operatorname{tr}[X] = \sum_{j=1}^{\min(d_1,d_2)} X_{jj}$. It can be seen that the p-th power of the Schatten-p norm coincides with $\|X\|_{S_p}^p = \operatorname{tr}\left[(X^*X)^{p/2}\right]$. The Schatten-2 norm is also called *Frobenius norm* and has the property that it is induced by the Frobenius scalar product $\langle X, Y\rangle_F = \operatorname{tr}[X^*Y]$, i.e., $\|X\|_F = \|X\|_{S_2} = \sqrt{\langle X, X\rangle_F}$. We define the *weighted Frobenius scalar product* of two matrices $X, Y \in M_{d_1 \times d_2}$ weighted by the the positive definite weight matrix $W \in H_{d_1 \times d_1}$ as $\langle X, Y\rangle_{F(W)} := \langle WX, Y\rangle_F = \langle X, WY\rangle_F$. This scalar product induces the *weighted Frobenius norm* $\|X\|_{F(W)} = \sqrt{\langle X, X\rangle_{F(W)}} = \sqrt{\operatorname{tr}[(WX)^*X]}$. It is clear that the Frobenius norm of a matrix X coincides with the ℓ_2-norm of its vectorization X_{vec}, i.e., $\|X\|_F = \|X_{\mathrm{vec}}\|_{\ell_2}$.

Similar to weighted Frobenius norms, we define the *weighted ℓ_2-scalar product* of vectors $x, y \in \mathbb{C}^d$ weighted by the positive definite weight matrix $W \in H_{d\times d}$ as $\langle x, y\rangle_{\ell_2(W)} = x^* W y = \overline{y^* W x}$ and its induced *weighted ℓ_2-norm* as $\|x\|_{\ell_2(W)} = \sqrt{x^* W x}$. We use the notation $X \succ 0$ for a positive definite matrix $X \in H_{d\times d}$. Furthermore, we denote the range of a linear map $\Phi : M_{d_1\times d_2} \to \mathbb{C}^m$ by $\mathrm{Ran}(\Phi) = \big\{Y \in \mathbb{C}^m;\ \text{there is } X \in M_{d_1\times d_2} \text{ such that } Y = \Phi(X)\big\}$ and its null space by $\mathcal{N}(\Phi) = \big\{X \in M_{d_1\times d_2}; \Phi(X) = 0\big\}$.

2.2 Problem setting and characterization of S_p- and reweighted Frobenius norm minimizers

Given a linear map $\Phi : M_{d_1\times d_2} \to \mathbb{C}^m$ such that $m \ll d_1 d_2$, we want to uniquely identify and reconstruct an unknown matrix X_0 from its linear image $Y := \Phi(X_0) \in \mathbb{C}^m$. However, basic linear algebra tells us that this is not possible without further assumptions, since Φ is not injective if $m < d_1 d_2$. Indeed, there is a $(d_1 d_2 - m)$-dimensional affine space $\{X_0\} + \mathcal{N}(\Phi)$ fulfilling the linear constraint

$$\Phi(X) = Y.$$

Nevertheless, under the additional assumption that the matrix $X_0 \in M_{d_1\times d_2}$ has rank $r < \min(d_1, d_2)$ and under appropriate assumptions on the map Φ, the recovery of X_0 is possible by solving the affine rank minimization problem

$$\min \mathrm{rank}(X) \text{ subject to } \Phi(X) = Y. \tag{5}$$

The unique solvability of (5) is given with high probability if, for example, Φ is a linear map whose matrix representation has i.i.d. Gaussian entries (Eldar et al., 2012) and $m = \Omega(r(d_1 + d_2))$. Unfortunately, solving (5) is intractable in general, but the works (Candès and Recht, 2009; Recht et al., 2010; Candès and Plan, 2011) suggest solving the tractable convex optimization program

$$\min \|X\|_{S_1} \text{ subject to } \Phi(X) = Y, \tag{6}$$

also called *nuclear norm minimization (`NNM`)*, as a proxy.

As discussed in the introduction, there are empirical as well as theoretical results (e.g., in (Daubechies et al., 2010; Chartrand, 2007)) coming from the related *sparse vector recovery* problem that suggest alternative relaxation approaches. These results indicate that it might be even more advantageous to solve the non-convex problem

$$\min F^p(X) := \|X\|_{S_p}^p \text{ subject to } \Phi(X) = Y, \tag{7}$$

for $0 < p < 1$, i.e., minimizing the p-th power of the Schatten-p quasi-norms under the affine constraint. Heuristically, the choice of $p < 1$ relatively small can be motivated by the observation that by the definition (4) of the Schatten-p quasi-norm

$$\|X\|_{S_p}^p \xrightarrow{p\to 0} \mathrm{rank}(X) =: \|X\|_{S_0}.$$

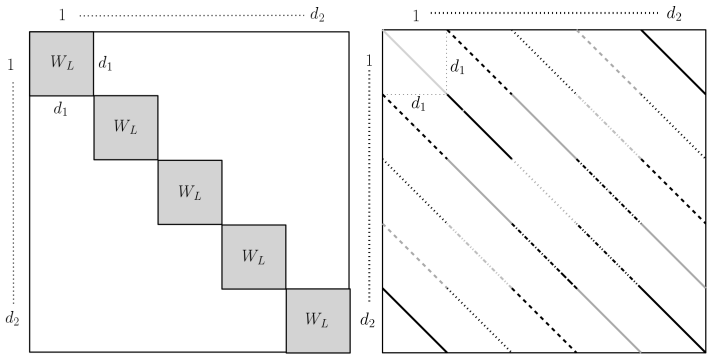
The above consideration suggests that the solution of (7) might be closer to (5) than (6) for small p. On the other hand, again, it is in general computationally intractable to find a global minimum of the non-convex optimization problem (7) if $p < 1$. Therefore it is a

natural and very relevant question to ask which optimization algorithm to use to find global minimizers of (7).

In this paper, we discuss an algorithm striving to solve (7) that is based on the following observations: Assume for the moment that we are given a square matrix $X \in M_{d_1 \times d_2}$ with $d_1 = d_2$ of full rank. Then, we can rewrite the p-th power of its Schatten-p quasi-norm as a squared weighted Frobenius norm, or, using Kronecker product notiation as explained in Appendix A, as a squared weighted ℓ_2-norm (if we use the vectorized notation X_{vec}): It turns out that

(i) $$\|X\|_{S_p}^p = \mathrm{tr}\big[(XX^*)^{\frac{p}{2}}\big] = \mathrm{tr}\big[(XX^*)^{\frac{p-2}{2}}(XX^*)\big] = \mathrm{tr}(W_L XX^*) = \|W_L^{\frac{1}{2}} X\|_F^2$$
$$= \|X\|_{F(W_L)}^2 = \|(\mathbf{I}_{d_2} \otimes W_L)^{\frac{1}{2}} X_{\mathrm{vec}}\|_{\ell_2}^2 = \|X_{\mathrm{vec}}\|_{\ell_2(\mathbf{I}_{d_2} \otimes W_L)}^2,$$
where W_L is the symmetric weight matrix $(XX^*)^{\frac{p-2}{2}}$ in $M_{d_1 \times d_1}$ and $\mathbf{I}_{d_2} \otimes W_L$ is the block diagonal weight matrix in $M_{d_1 d_2 \times d_1 d_2}$ with d_2 instances of W_L on the diagonal blocks, but also that

(ii) $$\|X\|_{S_p}^p = \mathrm{tr}\big[(X^*X)^{\frac{p}{2}}\big] = \mathrm{tr}\big[(X^*X)(X^*X)^{\frac{p-2}{2}}\big] = \mathrm{tr}(X^*XW_R) = \|XW_R^{\frac{1}{2}}\|_F^2$$
$$= \|X^*\|_{F(W_R)}^2 = \|(W_R \otimes \mathbf{I}_{d_1})^{\frac{1}{2}} X_{\mathrm{vec}}\|_{\ell_2}^2 = \|X_{\mathrm{vec}}\|_{\ell_2(W_R \otimes \mathbf{I}_{d_1})}^2,$$
where W_R is the symmetric weight matrix $(X^*X)^{\frac{p-2}{2}}$ in $M_{d_2 \times d_2}$. It follows from the definition of the Kronecker product that the weight matrix $W_R \otimes \mathbf{I}_{d_1} \in M_{d_1 d_2 \times d_1 d_2}$ is a block matrix of diagonal blocks of the type $\mathrm{diag}\big((W_R)_{ij}, \ldots, (W_R)_{ij}\big) \in M_{d_1 \times d_1}$, $i, j \in [d_2]$.



(a) $\mathbf{I}_{d_2} \otimes W_L$ (b) $W_R \otimes \mathbf{I}_{d_1}$

Figure 1: Sparsity structure of the weight matrices $\in M_{d_1 d_2 \times d_1 d_2}$

The sparsity structures of $\mathbf{I}_{d_2} \otimes W_L$ and $W_R \otimes \mathbf{I}_{d_1}$ are illustrated in Fig. 1. Note that a representation of $\|X\|_{S_p}^p$ by squares of Frobenius norms can be achieved by multiplying X by $W_L^{\frac{1}{2}}$ from the left in (i), or by $W_R^{\frac{1}{2}}$ from the right in (ii).

The above calculations are not well-defined if X is not of full rank or if $d_1 \neq d_2$, since in these cases at least one of the matrices $XX^* \in M_{d_1 \times d_1}$ or $X^*X \in M_{d_2 \times d_2}$ is singular, prohibiting the definition of the matrices $W_R = (X^*X)^{\frac{p-2}{2}}$ or $W_L = (XX^*)^{\frac{p-2}{2}}$ for $p < 2$. However, these issues can be overcome by introducing a smoothing parameter $\epsilon > 0$ and

smoothed weight matrices $W_L(X,\epsilon) \in M_{d_1\times d_1}$ and $W_R(X,\epsilon) \in M_{d_2\times d_2}$ defined by

$$W_L(X,\epsilon) := (XX^* + \epsilon^2 \mathbf{I}_{d_1})^{\frac{p-2}{2}}, \tag{8}$$

$$W_R(X,\epsilon) := (X^*X + \epsilon^2 \mathbf{I}_{d_2})^{\frac{p-2}{2}}. \tag{9}$$

Remark 1 *The weight matrices $W_L(X,\epsilon)$ and $W_R(X,\epsilon)$ are symmetric and positive definite.*

The possibility to rewrite the p-th power of the Schatten-p of a matrix as a squared weighted Frobenius norm gives rise to the general strategy of IRLS algorithms for low-rank matrix recovery: Weighted least squares problems of the type

$$\min_{\substack{X\in M_{d_1\times d_2}\\ \Phi(X)=Y}} \|X\|^2_{F(W_L)} \quad \text{or} \quad \min_{\substack{X\in M_{d_1\times d_2}\\ \Phi(X)=Y}} \|X^*\|^2_{F(W_R)}$$

are solved and weight matrices W_L are updated alternatingly, leading to the algorithms column-reweighting `IRLS-col` and row-reweighting `IRLS-row`, respectively (Mohan and Fazel, 2012; Fornasier et al., 2011).

2.3 Averaging of weight matrices

While the algorithms `IRLS-col` and `IRLS-row` provide a tractable local minimization strategy of smoothed Schatten-p functionals under the linear constraint, we argue that it is suboptimal to follow either one of the two approaches as they do not exploit the symmetry of the problem in an optimal way: They *either* use low-rank information in the column space *or* in the row space.

A first intuitive approach towards a symmetric exploitation of the low-rank structure is inspired by the following identity, by combing the calculations (i) and (ii) carried out in Section 2.2.

Lemma 2 *Let $0 < p \leqslant 2$ and $X \in M_{d_1\times d_2}$ with $d = d_1 = d_2$ be a matrix of full rank. Then*

$$\|X\|^p_{S_p} = \frac{1}{2}\left(\|W_L^{\frac{1}{2}}X\|_F^2 + \|XW_R^{\frac{1}{2}}\|_F^2\right) = \left\|\left(\frac{W_L\oplus W_R}{2}\right)^{\frac{1}{2}} X_{\mathrm{vec}}\right\|^2_{\ell_2} = \|X_{\mathrm{vec}}\|^2_{\ell_2(W_{(arith)})},$$

where

$$\frac{1}{2}\left(\mathbf{I}_{d_2}\otimes W_L + W_R\otimes \mathbf{I}_{d_1}\right) = \frac{W_L\oplus W_R}{2} =: W_{(arith)}$$

is the arithmetic mean matrix of the symmetric and positive definite weight matrices $\mathbf{I}_{d_2}\otimes W_L$ and $W_R\otimes \mathbf{I}_{d_1}$, $W_L := (XX^*)^{\frac{p-2}{2}}$, and $W_R := (X^*X)^{\frac{p-2}{2}}$.

Unfortunately, the introduction of arithmetic mean weight matrices does not prove to be particularly advantageous compared to one-sided reweighting strategies. Convincing improvements could be noted neither in numerical experiments nor in the theoretical investigations for the convergence rate of IRLS for low-rank matrix recovery, cf. also Section 5.2 and Remark 22.

In contrast, we want to promote the usage of the *harmonic mean of the weight matrices* $\mathbf{I}_{d_2} \otimes W_L$ and $W_R \otimes \mathbf{I}_{d_1}$, i.e., weight matrices of the type $2\left(W_R^{-1} \otimes \mathbf{I}_{d_1} + \mathbf{I}_{d_2} \otimes W_L^{-1}\right)^{-1} = 2\left(W_L^{-1} \oplus W_R^{-1}\right)^{-1} =: W_{(\mathrm{harm})}$. In the remaining parts of the paper, we explain why $W_{(\mathrm{harm})}$ is able to significantly outperform other weighting variants both theoretically and practically.

The following lemma verifies that the harmonic mean $W_{(\mathrm{harm})}$ of the weight matrices $\mathbf{I}_{d_2} \otimes W_L$ and $W_R \otimes \mathbf{I}_{d_1}$ leads to a legitimate reformulation of the Schatten-p quasi-norm power, as it we already saw for the arithmetic mean $W_{(\mathrm{arith})}$.

Lemma 3 *Let $0 < p \leqslant 2$ and $X \in \mathbb{C}^{d_1 \times d_2}$ with $d = d_1 = d_2$ be a full rank matrix. Then*

$$\|X\|_{S_p}^p = 2\left\|\left(W_L^{-1} \oplus W_R^{-1}\right)^{-\frac{1}{2}} X_{\mathrm{vec}}\right\|_{\ell_2}^2 = \|X_{\mathrm{vec}}\|_{\ell_2(W_{(harm)})}^2,$$

where

$$2\left(W_R^{-1} \otimes \mathbf{I}_{d_1} + \mathbf{I}_{d_2} \otimes W_L^{-1}\right)^{-1} = 2\left(W_L^{-1} \oplus W_R^{-1}\right)^{-1} =: W_{(harm)}$$

is the harmonic mean matrix of the symmetric and positive definite weight matrices $\mathbf{I}_{d_2} \otimes W_L$ and $W_R \otimes \mathbf{I}_{d_2}$, $W_L := (XX^*)^{\frac{p-2}{2}}$ and $W_R := (X^*X)^{\frac{p-2}{2}}$.

Proof Let $X = U\Sigma V^* = \sum_{i=1}^d \sigma_i u_i v_i^* \in M_{d\times d}$ be the singular value decomposition of X. Therefore for the vectorized version, $X_{\mathrm{vec}} = (V \otimes U)\Sigma_{\mathrm{vec}}$ holds true. By the definitions of W_L and W_R, we can write $W_L^{-1} = \sum_{i=1}^d \sigma_i^{2-p} u_i u_i^*$ and $W_R^{-1} = \sum_{i=1}^d \sigma_i^{2-p} v_i v_i^*$. Using the Kronecker sum inversion formula of Lemma 23 in Appendix A, we obtain

$$\begin{aligned}
\|X_{\mathrm{vec}}\|_{\ell_2(W_{(\mathrm{harm})})}^2 &= \|W_{(\mathrm{harm})}^{\frac{1}{2}} X_{\mathrm{vec}}\|_{\ell_2}^2 = 2\left\|\left(W_L^{-1} \oplus W_R^{-1}\right)^{-\frac{1}{2}} X_{\mathrm{vec}}\right\|_{\ell_2}^2 \\
&= 2\,\mathrm{tr}\left(\left(\left(W_L^{-1} \oplus W_R^{-1}\right)^{-1} X_{\mathrm{vec}}\right)_{\mathrm{mat}}^* X\right) \\
&= \sum_{i=1}^d \sum_{j=1}^d \sum_{k=1}^d \frac{2\sigma_k}{\sigma_i^{2-p} + \sigma_j^{2-p}} v_j v_i^* v_k u_k^* u_i u_i^* \sum_{l=1}^d \sigma_l u_l v_l^* \\
&= 2\left(\sum_{i=1}^d \frac{\sigma_i^2}{2\sigma_i^{2-p}}\right) = \|X\|_{S_p}^p,
\end{aligned}$$

which finishes the proof. ∎

3. Harmonic mean iteratively reweighted least squares algorithm

In this section, we use this idea to formulate a new iteratively reweighted least squares algorithm for low-rank matrix recovery. The so-called *harmonic mean iteratively reweighted least squares* algorithm (`HM-IRLS`) solves a sequence of weighted least squares problems to recover a low-rank matrix $X_0 \in M_{d_1\times d_2}$ from few linear measurements $\Phi(X_0) \in \mathbb{C}^m$. The weight matrices appearing in the least squares problems can be seen as the harmonic mean of the weight matrices in (8) and (9), i.e., the ones used by `IRLS-col` and `IRLS-row`.

More precisely, for $0 < p \leqslant 1$ and $d = \min(d_1, d_2)$, $D = \max(d_1, d_2)$, given a non-increasing sequence of non-negative real numbers $(\epsilon^{(n)})_{n=1}^{\infty}$ and the sequence of iterates $(X^{(n)})_{n=1}^{\infty}$ produced by the algorithm, we update our weight matrices such that

$$\widetilde{W}^{(n)} = 2\left[U^{(n)}(\bar{\Sigma}_{d_1}^{(n)})^{2-p}U^{(n)*} \oplus V^{(n)}(\bar{\Sigma}_{d_2}^{(n)})^{2-p}V^{(n)*}\right]^{-1}, \tag{10}$$

with the diagonal matrices $\bar{\Sigma}_{d_t}^{(n)} \in M_{d_t \times d_t}$ for $d_t = \{d_1, d_2\}$ such that

$$(\bar{\Sigma}_{d_t}^{(n)})_{ii} = \begin{cases} (\sigma_i(X^{(n)})^2 + \epsilon^{(n)2})^{\frac{1}{2}} & \text{if } i \leqslant d, \\ 0 & \text{if } d < i \leqslant D, \end{cases} \tag{11}$$

and the matrices $U^{(n)} \in \mathcal{U}_{d_1}$ and $V^{(n)} \in \mathcal{U}_{d_2}$ containing the left and right singular vectors of $X^{(n)}$ in its columns, respectively.

We note that this definition of $\widetilde{W}^{(n)}$ can be seen as a stabilized version of the harmonic mean weight matrix $W_{(\mathrm{harm})}$ of Lemma 3. This stabilization is necessary as $\widetilde{W}^{(n)}$ becomes very ill-conditioned as soon as some of the singular values of $X^{(n)}$ approach zero and, related to that, $(X^{(n)}X^{(n)*})^{\frac{2-p}{2}} \oplus (X^{(n)*}X^{(n)})^{\frac{2-p}{2}}$ would even be singular as soon as $X^{(n)}$ is not of full rank.

Additionally, for the formulation of the algorithm it is convenient to define the linear operator $(\widetilde{\mathcal{W}}^{(n)})^{-1} : M_{d_1 \times d_2} \to M_{d_1 \times d_2}$ for any $n \in \mathbb{N}$ such that

$$(\widetilde{\mathcal{W}}^{(n)})^{-1}(X) := \frac{1}{2}\left[U^{(n)}(\bar{\Sigma}_{d_1}^{(n)})^{2-p}U^{(n)*}X + XV^{(n)}(\bar{\Sigma}_{d_2}^{(n)})^{2-p}V^{(n)*}\right], \tag{12}$$

describing the operation of the inverse of $\widetilde{W}^{(n)}$ on $M_{d_1 \times d_2}$.

Finally, `HM-IRLS` can be formulated in pseudo code as follows.

Algorithm 1 Harmonic Mean IRLS for low-rank matrix recovery (`HM-IRLS`)

Input: A linear map $\Phi : M_{d_1 \times d_2} \to \mathbb{C}^m$, image $Y = \Phi(X_0)$ of the ground truth matrix $X_0 \in M_{d_1 \times d_2}$, rank estimate $\tilde{r}$, non-convexity parameter $0 < p \leqslant 1$.

Output: Sequence $(X^{(n)})_{n=1}^{n_0} \subset M_{d_1 \times d_2}$.

Initialize $n = 0$, $\epsilon^{(0)} = 1$ and $\widetilde{W}^{(0)} = \mathbf{I}_{d_1 d_2} \in M_{d_1 d_2 \times d_1 d_2}$.

repeat

$$X^{(n+1)} = \operatorname*{arg\,min}_{\Phi(X)=Y} \|X_{\mathrm{vec}}\|^2_{\ell_2(\widetilde{W}^{(n)})} = ((\widetilde{\mathcal{W}}^{(n)})^{-1}\big(\Phi^*\big((\Phi \circ ((\widetilde{\mathcal{W}}^{(n)})^{-1} \circ \Phi^*)^{-1}(Y))\big)\big), \tag{13}$$

$$\epsilon^{(n+1)} = \min\left(\epsilon^{(n)}, \sigma_{\tilde{r}+1}(X^{(n+1)})\right), \tag{14}$$

$$\widetilde{W}^{(n+1)} = 2\left[U^{(n+1)}(\bar{\Sigma}_{d_1}^{(n+1)})^{2-p}U^{(n+1)*} \oplus V^{(n+1)}(\bar{\Sigma}_{d_2}^{(n+1)})^{2-p}V^{(n+1)*}\right]^{-1}, \tag{15}$$

where $U^{(n+1)} \in \mathcal{U}_{d_1}$ and $V^{(n+1)} \in \mathcal{U}_{d_2}$ are matrices containing the left and right singular vectors of $X^{(n+1)}$ in its columns, and the $\bar{\Sigma}_{d_t}^{(n+1)}$ are defined for $t \in \{1, 2\}$ according to (11).

$$n = n + 1,$$

until *stopping criterion is met.*

Set $n_0 = n$.

From a practical point of view, it is beneficial that the explicit calculation of the very large weight matrices $\widetilde{W}^{(n)} \in H_{d_1 d_2 \times d_1 d_2}$ (cf. (15)) is not necessary in implementations of Algorithm 1. As suggested by formulas (12) and (13), it can be seen that just the operation of its inverse $(\widetilde{W}^{(n)})^{-1}$ is needed, which can be implemented by matrix-matrix multiplications on the space $M_{d_1 \times d_2}$: For matrices $X, \widetilde{X} \in M_{d_1 \times d_2}$, we have that $\widetilde{W}^{(n)} X_{\mathrm{vec}} = \widetilde{X}_{\mathrm{vec}}$ if and only if $X_{\mathrm{vec}} = (\widetilde{W}^{(n)})^{-1} \widetilde{X}_{\mathrm{vec}}$, which can be written in matrix variables as

$$X = \frac{1}{2}\left[U^{(n)} (\bar{\Sigma}_{d_1}^{(n)})^{2-p} U^{(n)*} \widetilde{X} + \widetilde{X} V^{(n)} (\bar{\Sigma}_{d_2}^{(n)})^{2-p} V^{(n)*}\right].$$

The last equivalence is due to the definitions of $\widetilde{W}^{(n)}$ and the Kronecker sum, cf. (15) and Appendix A.

Note that the smoothing parameters $\epsilon^{(n)}$ are chosen in dependence on a rank estimate $\tilde{r}$ here, which will be an important ingredient for the theoretical analysis of the algorithm. In practice, however, other choices of non-increasing sequences of non-negative real numbers $(\epsilon^{(n)})_{n=1}^{\infty}$ are possible and can as well lead to (a maybe even faster) convergence when tuned appropriately.

We refer to Section 5.4 for a further discussion of implementation details.

Example With a simple example, we illustrate the versatility of `HM-IRLS`: Let $d_1 = d_2 = 4$, and assume that we want to reconstruct the rank-1 matrix

$$X_0 = uv^* = \begin{pmatrix} 1 \\ 10 \\ -2 \\ 0.1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 10 & 20 & 30 & 40 \\ -2 & -4 & -6 & -8 \\ 0.1 & 0.2 & 0.3 & 0.4 \end{pmatrix}$$

from $m = d_f = r(d_1 + d_2 - r) = 7$ sampled entries $\Phi(X_0)$, where Φ is the linear map $\Phi : M_{4\times 4} \to \mathbb{C}^7$, $\Phi(X) = \big(X_{2,1},\ X_{4,1},\ X_{3,2},\ X_{4,2},\ X_{4,3},\ X_{1,4},\ X_{2,4}\big)$. Since the linear map Φ samples some entries of matrices in $M_{4\times 4}$ and does not see the others, this is an instance of the problem that is called *matrix completion*.

In general, reconstructing a $(d_1 \times d_2)$ rank-r matrix from $m = r(d_1 + d_2 - r)$ entries is a hard problem, as it is known that if $m < r(d_1 + d_2 - r)$, there is always more than one matrix X such that $\Phi(X) = \Phi(X_0)$, and even for equality, the property that Φ is invertible on (most) rank-r matrices might be hard to verify (Király et al., 2015).

It can be argued that the specific matrix completion problem we consider is in some sense a hard one, since, e.g., the deterministic sufficient condition for unique completability of (Pimentel-Alarcón et al., 2016, Theorem 2) is not fulfilled (less then 2 observed entries in the third column), and since the classical coherence parameters $\mu(u) = d_1 \max\limits_{1\leqslant i\leqslant 4} \frac{\|uu^* e_i\|_2^2}{\|u\|_2^4} \approx 3.81$ and $\mu(v) = d_2 \max\limits_{1\leqslant i\leqslant 4} \frac{\|vv^* e_i\|_2^2}{\|v\|_2^4} \approx 2.13$ that are used to analyze the behavior of many matrix completion algorithms (Candès and Recht, 2009; Jain et al., 2013) are quite large, with $\mu(u)$ being quite close to the maximal value of 4.

On the other hand, as the problem is small and X_0 has rank $r = 1$, it is possible to impute the missing values of

$$\begin{pmatrix} * & * & * & 4 \\ 10 & * & * & 40 \\ * & -4 & * & * \\ 0.1 & 0.2 & 0.3 & * \end{pmatrix}$$

by solving very simple linear equations, since, for example, $X_{4,4} = u_4 v_4$, $X_{2,1} = u_2 v_1$, $X_{2,4} = u_2 v_4$, and $X_{4,1} = u_4 v_1$, and therefore $X_{4,4} = \frac{X_{4,1} X_{2,4}}{X_{2,1}} = 0.4$. This shows that the only rank-1 matrix compatible with $\Phi(X_0)$ is X_0.

It turns out that—without using the combinatorial simplicity of the problem—the classical `NNM` does not solve the problem, as the nuclear norm minimizer (solution of (6) for $Y = \Phi(X_0)$) produced by the semidefinite program of the convex optimization package `CVX` (Grant and Boyd, 2014) converges to

$$\bar{X}_{\text{nuclear}} \approx \begin{pmatrix} 1 & 0.023 & 0.041 & 4 \\ 10 & 0.232 & 0.411 & 40 \\ -0.056 & -4 & -0.200 & -0.226 \\ 0.1 & 0.2 & 0.3 & 0.400 \end{pmatrix},$$

a matrix with $45.74 \approx \|\bar{X}_{\text{nuclear}}\|_{S_1} < \|X_0\|_{S_1} = \sigma_1(X_0) \approx 56.13$ and a relative Frobenius error of $\frac{\|\bar{X}_{\text{nuclear}} - X_0\|_F}{\|X_0\|_F} = 0.661$.

On the other hand, `HM-IRLS` is able to solve the problem—if p is chosen small enough—with very high precision already after few iterations, for example, up to a relative error of $4.18 \cdot 10^{-13}$ after 24 iterations if $p = 0.1$. This is in contrast to the behavior of `IRLS-col`, `IRLS-row` and also to the behavior of `AM-IRLS`, the IRLS variant that uses weight matrices derived from the *arithmetic mean* of the weights of `IRLS-col` and `IRLS-row`, cf. Lemma 2. The iterates $X^{(n)}$ for iteration $n = 2000$ of these algorithms exhibit relative errors of 0.240, 0.489 and 0.401, respectively, for the choice of $p = 0.1$. Furthermore, there is no choice of p that would lead to a convergence to X_0.

To understand this very different behavior, we note that the n-th iterate of any of the four IRLS variants can be written, using Appendix A, in a concise way as

$$X^{(n+1)} = \operatorname*{arg\,min}_{\Phi(X)=Y} \langle X_{\text{vec}}, W^{(n)} X_{\text{vec}} \rangle, \tag{16}$$

where

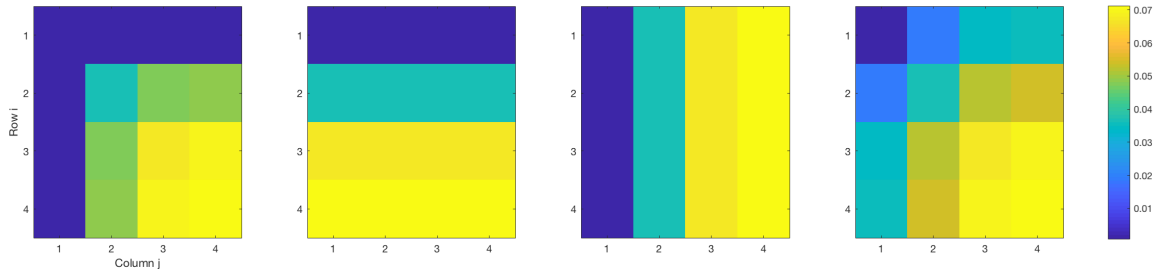
$$\langle X_{\text{vec}}, W^{(n)} X_{\text{vec}} \rangle = \langle X, U^{(n)} \big[H^{(n)} \circ (U^{(n)*} X V^{(n)}) \big] V^{(n)*} \rangle_F = \sum_{i,j=1}^{4} H_{ij}^{(n)} |\langle u_i^{(n)}, X v_j^{(n)} \rangle|^2 \tag{17}$$

with $X^{(n)} = U^{(n)} \Sigma^{(n)} V^{(n)*} = \sum_{i=1}^{4} \sigma_i^{(n)} u_i^{(n)} v_i^{(n)}$ being the SVD of $X^{(n)}$, and

$$H_{i,j}^{(n)} = \begin{cases} 2\Big[\big((\sigma_i^{(n)})^2 + (\epsilon^{(n)})^2 \big)^{\frac{2-p}{2}} + \big((\sigma_j^{(n)})^2 + (\epsilon^{(n)})^2 \big)^{\frac{2-p}{2}} \Big]^{-1} & \text{for } \texttt{HM-IRLS}, \\ \big((\sigma_i^{(n)})^2 + (\epsilon^{(n)})^2 \big)^{\frac{p-2}{2}} & \text{for } \texttt{IRLS-col}, \\ \big((\sigma_j^{(n)})^2 + (\epsilon^{(n)})^2 \big)^{\frac{p-2}{2}} & \text{for } \texttt{IRLS-row}, \text{ and} \\ 0.5 \cdot \Big[\big((\sigma_i^{(n)})^2 + (\epsilon^{(n)})^2 \big)^{\frac{p-2}{2}} + \big((\sigma_i^{(n)})^2 + (\epsilon^{(n)})^2 \big)^{\frac{p-2}{2}} \Big] & \text{for } \texttt{AM-IRLS}, \end{cases}$$

for $i, j \in \{1, 2, 3, 4\}$ and $\epsilon^{(n)} = \min(\sigma_2^{(n)}, \epsilon^{(n-1)})$.

The values of the matrix $H^{(1)}$ of weight coefficients after the first iteration in the above example are visualized in Figure 2, for each of the four IRLS versions above.



(a) `HM-IRLS` (b) `IRLS-col` (c) `IRLS-row` (d) `AM-IRLS`

Figure 2: Values of the matrix $H^{(1)}$ of "weight coefficients" corresponding to the orthonormal basis $(u_i^{(1)} v_j^{(1)*})_{i,j=1}^4$ after the first iteration in the example

The intuition for the superior behavior of `HM-IRLS` is now the following: Since large entries of $H^{(n)}$ *penalize* the corresponding parts of the space $M_{d_1 \times d_2} = \operatorname{span}\{u_i^{(n)} v_j^{(n)*}, i \in [d_1], j \in [d_2]\}$ in the minimization problem (16), large areas of *blue* and *dark blue* in Figure 2 indicate a benign optimization landscape where the minimizer $X^{(n+1)}$ of (16) *is able to improve considerably* on the previous iterate $X^{(n)}$.

In particular, it can be seen that in the case of `HM-IRLS`, the penalties on the *whole direct sum of column and row space of the best rank-r approximation of* $X^{(n)}$

$$T^{(n)} := \left\{ \begin{pmatrix} u_1^{(n)}, & \dots, & u_r^{(n)} \end{pmatrix} Z_1^* + Z_2 \begin{pmatrix} v_1^{(n)}, & \dots, & v_r^{(n)} \end{pmatrix}^* : Z_1 \in M_{d_1 \times r}, Z_2 \in M_{d_2 \times r} \right\},$$

are small compared to the other penalites, since the coefficients of $H^{(1)}$ corresponding to $T^{(1)}$ are exactly the ones in the first row and first column of the (4×4) matrices in Figure 2—a contrast that becomes more and more pronounced as $X^{(n)}$ approaches the rank-r ground truth X_0 (with $r = 1$ in the example).

On the other hand, `IRLS-col`, `IRLS-row` and `AM-IRLS` only have small coefficients on smaller parts of $T^{(n)}$, which, from a global perspective, explains why their usage might lead to non-global minima of the Schatten-p objective.

We note that the space $T^{(n)}$ plays also an important role in Riemannian optimization approaches for matrix recovery problems (see Vandereycken, 2013), since it is also the tangent space of the smooth manifold of rank-r matrices at the best rank-r approximation of $X^{(n)}$.

4. Convergence results

In the following part, we state our main theoretical results about convergence properties of the algorithm `HM-IRLS`. Furthermore, their relation to existing results for `IRLS-col` and `IRLS-row` is discussed.

It cannot be expected that a low-rank matrix recovery algorithm like `HM-IRLS` succeeds to converge to a low-rank matrix without any assumptions on the measurement operator Φ that defines the recovery problem (5). For the purpose of the convergence analysis of `HM-IRLS`, we introduce the following strong *Schatten-p null space property* (Fornasier et al., 2011; Oymak et al., 2011; Foucart and Rauhut, 2013).

Definition 4 (Strong Schatten-p null space property) *Let $0 < p \leqslant 1$. We say that a linear map $\Phi : M_{d_1 \times d_2} \to \mathbb{C}^m$ fulfills the strong Schatten-p null space property (Schatten-p NSP) of order r with constant $0 < \gamma_r \leqslant 1$ if*

$$\left(\sum_{i=1}^{r} \sigma_i^2(X) \right)^{p/2} < \frac{\gamma_r}{r^{1-\frac{p}{2}}} \left(\sum_{i=r+1}^{d} \sigma_i^p(X) \right) \tag{18}$$

for all $X \in \mathcal{N}(\Phi) \backslash \{0\}$.

Intuitively explained, if a map Φ fulfills the strong Schatten-p null space property of order r, there are no rank-r matrices in the null space and all the elements of the null space must not have a quickly decaying spectrum.

Null space properties have already been used to guarantee the success of nuclear norm minimization (6), or Schatten-1 minimization in our terminology, for solving the low-rank matrix recovery problem (Recht et al., 2011).

We note that the definitions of Schatten-p null space properties are quite analogous to the ℓ_p-null space property in classical compressed sensing (Foucart and Rauhut, 2013, Theorem 4.9), applied to the vector of singular values. In particular, (18) implies that

$$\sum_{i=1}^{r} \sigma_i^p(X) < \sum_{i=r+1}^{d} \sigma_i^p(X) \qquad \text{for all } X \in \mathcal{N}(\Phi) \backslash \{0\}, \tag{19}$$

since $\|X\|_{S_p} \leqslant r^{1/p-1/2} \|X\|_{S_2}$ for X that is rank-r. This, in turn, ensures the existence of unique solutions to (7) if $Y = \Phi(X_0)$ are the measurements of a low-rank matrix X_0.

Proposition 5 (Foucart (2018)) *Let $\Phi : M_{d_1 \times d_2} \to \mathbb{C}^m$ be a linear map, let $0 < p \leqslant 1$ and $r \in \mathbb{N}$. Then every matrix $X_0 \in M_{d_1 \times d_2}$ such that $\operatorname{rank}(X_0) \leqslant r$ and $\Phi(X_0) = Y \in \mathbb{C}^m$ is the* unique *solution of Schatten-p minimization* (7) *if and only if Φ fulfills* (19)*.*

Remark 6 *The sufficiency of the Schatten-p NSP* (19) *in Proposition 5 has already been pointed out by Oymak et al. (2011). The necessity as stated in the theorem, however, is due to a recent generalization of Mirsky's singular value inequalities to concave functions (Audenaert, 2014; Foucart, 2018).*

It can be seen that the (weak) Schatten-p NSP of (19) is a *stronger* property for larger p in the sense that if $0 < p' \leqslant p \leqslant 1$, the Schatten-$p$ property implies the Schatten-p' property. Very related to this, it can be seen that for any $0 < p \leqslant 1$, the strong Schatten-p null space property is implied by a sufficiently small *rank restricted isometry constant* δ_r, which is a classical tool in the analysis of low-rank matrix recovery algorithms (Recht et al., 2010; Candès and Plan, 2011).

Definition 7 (Restricted isometry property (RIP)) *The* restricted isometry constant $\delta_r > 0$ of order r *of the linear map* $\Phi : M_{d_1 \times d_2} \to \mathbb{C}^m$ *is defined as the smallest number such that*

$$(1 - \delta_r)\|X\|_F^2 \leqslant \|\Phi(X)\|_{\ell_2}^2 \leqslant (1 + \delta_r)\|X\|_F^2$$

for all matrices $X \in M_{d_1 \times d_2}$ *of rank at most* r.

Indeed, it follows from the proof of Chavez-Dominguez and Kutzarova (2015, Theorem 4.1) that a restricted isometry constant of order $2r$ such that $\delta_{2r} < \frac{2}{\sqrt{2}+3} \approx 0.4531$ implies the strong Schatten-p NSP of order r with a constant $\gamma_r < 1$ for any $0 < p \leqslant 1$. More precisely, it can be seen that $\delta_{2r} < \frac{2}{\sqrt{2}+3}$ implies that the strong Schatten-p NSP (18) of order r holds with the constant $\gamma_r = \frac{(\sqrt{2}+1)^p}{2^p} \frac{\delta_{2r}^p}{(1-\delta_{2r})^p}$.

Linear maps that are instances drawn from certain random models are known to fulfill the restricted isometry property with high probability if the number of measurements is sufficiently large (Davenport and Romberg, 2016), and, a fortiori, the Schatten-p null space property. In particular, this is true for *(sub-)Gaussian* linear measurement maps $\Phi : M_{d_1 \times d_2} \to \mathbb{C}^m$ whose matrix representation is such that

$$\frac{1}{\sqrt{m}}\widetilde{\Phi} \in \mathbb{C}^{m \times d_1 d_2}, \text{ where } \widetilde{\Phi} \text{ has i.i.d. standard (sub-)Gaussian entries,} \tag{20}$$

as it is summarized in the following lemma.

Lemma 8 *For any* $0 < p \leqslant 1$, $0 < \gamma < 1$ *and any (sub-)Gaussian random operator* $\Phi : M_{d_1 \times d_2} \to \mathbb{C}^m$ *(e.g. as defined in* (20)*), there exist constants* $C_1 > 1$, $C_2 > 0$ *such that if* $m \geqslant C_1 r(d_1 + d_2)$, *the strong Schatten-p null space property* (18) *of order* r *with constant* $\gamma_r < \gamma$ *is fulfilled with probability at least* $1 - e^{-C_2 m}$.

4.1 Local convergence for $p < 1$

In this section, we provide a convergence analysis for `HM-IRLS` covering several aspects. We show that the algorithm converges to stationary points of a smoothed Schatten-p functional g_ϵ^p as in (21) without any additional assumptions on the measurement map Φ. Such guarantees have already been obtained for IRLS algorithms with one-sided reweighting as `IRLS-col` and `IRLS-row`, in particular for $p = 1$ by Fornasier et al. (2011) and for $0 < p \leqslant 1$ by Mohan and Fazel (2012).

Beyond that, assuming the measurement operator fulfills an appropriate Schatten-p null space property as defined in Definition 4, we show the a-posteriori exact recovery statement that `HM-IRLS` converges to the low-rank matrix X_0 if $\lim_{n \to \infty} \epsilon_n = 0$, which only was shown for one-sided `IRLS` for the case $p = 1$ by Fornasier et al. (2011).

Moreover, we provide a local convergence guarantee stating that `HM-IRLS` recovers the low-rank matrix X_0 if we obtain an iterate $X^{(\bar{n})}$ that is close enough to X_0, which is novel for IRLS algorithms.

Let $0 < p \leqslant 1$ and $\epsilon > 0$. To state the theorem, we introduce the *ϵ-perturbed Schatten-p functional* $g_\epsilon^p : M_{d_1 \times d_2} \to \mathbb{R}_{\geqslant 0}$ such that

$$g_\epsilon^p(X) = \sum_{i=1}^{d} (\sigma_i(X)^2 + \epsilon^2)^{\frac{p}{2}}, \tag{21}$$

where $\sigma(X) \in \mathbb{R}^d$ denotes the vector of singular values of $X \in M_{d_1 \times d_2}$.

Theorem 9 *Let $\Phi : M_{d_1 \times d_2} \to \mathbb{C}^m$ be a linear operator and $Y \in \mathrm{Ran}(\Phi)$ a vector in its range. Let $(X^{(n)})_{n \geqslant 1}$ and $(\epsilon^{(n)})_{n \geqslant 1}$ be the sequences produced by Algorithm 1 for input parameters Φ, Y, r and $0 < p \leqslant 1$, let $\epsilon = \lim_{n \to \infty} \epsilon^{(n)}$.*

(i) If $\epsilon = 0$ and if Φ fulfills the strong Schatten-p NSP (18) of order r with constant $0 < \gamma_r < 1$, then the sequence $(X^{(n)})_{n \geqslant 1}$ converges to a matrix $\bar{X} \in M_{d_1 \times d_2}$ of rank at most r that is the unique minimizer of the Schatten-p minimization problem (7). Moreover, there exists an absolute constant $\hat{C} > 0$ such that for any X with $\Phi(X) = Y$ and any $\widetilde{r} \leqslant r$, it holds that

$$\|X - \bar{X}\|_F^p \leqslant \frac{\hat{C}}{r^{1-p/2}} \beta_{\widetilde{r}}(X)_{S_p},$$

where $\hat{C} = \frac{2^{p+1} \gamma_r^{1-p/2}}{1-\gamma_r}$ and $\beta_{\widetilde{r}}(X)_{S_p}$ is the best rank-$\widetilde{r}$ Schatten-p approximation error of X, i.e.,

$$\beta_{\widetilde{r}}(X)_{S_p} := \inf \left\{ \|X - \widetilde{X}\|_{S_p}^p,\ \widetilde{X} \in M_{d_1 \times d_2} \text{ has rank } \widetilde{r} \right\}. \tag{22}$$

(ii) If $\epsilon > 0$, then each accumulation point $\bar{X}$ of $(X^{(n)})_{n \geqslant 1}$ is a stationary point of the ϵ-perturbed Schatten-p functional g_ϵ^p of (21) under the linear constraint $\Phi(X) = Y$. If additionally $p = 1$, then $\bar{X}$ is the unique global minimizer of g_ϵ^p.

(iii) Assume that there exists a matrix $X_0 \in M_{d_1 \times d_2}$ with $\Phi(X_0) = Y$ such that $\mathrm{rank}(X_0) = r \leqslant \frac{\min(d_1, d_2)}{2}$, a constant $0 < \zeta < 1$ and an iteration $\bar{n} \in \mathbb{N}$ such that

$$\|X^{(\bar{n})} - X_0\|_{S_\infty} \leqslant \zeta \sigma_{\widetilde{r}}(X_0)$$

and $\epsilon^{\bar{n}} = \sigma_{r+1}(X^{\bar{n}})$. If Φ fulfills the strong Schatten-p NSP of order $2r$ with $\gamma_{2r} < 1$ and if the condition number $\kappa = \frac{\sigma_1(X_0)}{\sigma_r(X_0)}$ of X_0 and ζ are sufficiently small (see condition (25) and formula (26)), then

$$X^{(n)} \to X_0 \quad \text{for } n \to \infty.$$

It is important to note that by using Lemma 8, it follows that the assertions of Theorem 9(i) and (iii) hold for (sub-)Gaussian operators (20) with high probability in the regime of measurements of optimal sample complexity order. In particular, there exist constant oversampling factors $\rho_1, \rho_2 \geqslant 1$ such that the assertions of (i) and (iii) hold with high probability if $m > \rho_k r(d_1 + d_2)$, $k \in \{1, 2\}$, respectively.

Remark 10 *However, if $m < d_1 d_2$, null space property-type assumptions as (18) or (19) do not hold for the important case of matrix completion-type measurements (Candès and Recht, 2009), where $\Phi(X)$ is given as m sample entries*

$$\Phi(X)_\ell = X_{i_\ell, j_\ell}, \qquad \ell = 1, \ldots, m, \tag{23}$$

and $(i_\ell, j_\ell) \in [d_1] \times [d_2]$ *for all* $\ell \in [m]$*, of the matrix* $X \in M_{d_1 \times d_2}$*, which also were considered in the example of Section 3.*

This means that parts (i) and (iii) of Theorem 9 do, unfortunately, not apply for matrix completion measurements, which define a very relevant class of low-rank matrix recovery problems. This problem is shared by any existing theory for IRLS algorithms for low-rank matrix recovery (Fornasier et al., 2011; Mohan and Fazel, 2012). However, in Section 5, we provide strong numerical evidence that `HM-IRLS` *exhibits properties as predicted by (i) and (iii) of Theorem 9 even for the matrix completion setting. We leave the extension of the theory of* `HM-IRLS` *to matrix completion measurements as an open problem to be tackled by techniques different from uniform null space properties (Davenport and Romberg, 2016, Section V).*

4.2 Locally superlinear convergence rate for $p < 1$

Next, we state the second main theoretical result of this paper, Theorem 11. It shows that in a neighborhood of a low-rank matrix X_0 that is compatible with the measurement vector Y, the algorithm `HM-IRLS` converges to X_0 with a *convergence rate* that is *superlinear of the order* $2 - p$, if the operator Φ fulfills an appropriate Schatten-p null space property.

Theorem 11 (Locally Superlinear Convergence Rate) *Assume that the linear map* $\Phi : M_{d_1 \times d_2} \to \mathbb{C}^m$ *fulfills the strong Schatten-p NSP of order* $2r$ *with constant* $\gamma_{2r} < 1$ *and that there exists a matrix* $X_0 \in M_{d_1 \times d_2}$ *with* $\operatorname{rank}(X_0) = r \leqslant \frac{\min(d_1, d_2)}{2}$ *such that* $\Phi(X_0) = Y$*, let* Φ, Y, r *and* $0 < p \leqslant 1$ *be the input parameters of Algorithm 1. Moreover, let* $\kappa = \frac{\sigma_1(X_0)}{\sigma_r(X_0)}$ *be the condition number of* X_0 *and* $\eta^{(n)} := X^{(n)} - X_0$ *be the error matrices of the* n*-th output of Algorithm 1 for* $n \in \mathbb{N}$*.*
Assume that there exists an iteration $\bar{n} \in \mathbb{N}$ *and a constant* $0 < \zeta < 1$ *such that*

$$\|\eta^{(\bar{n})}\|_{S_\infty} \leqslant \zeta \sigma_r(X_0) \tag{24}$$

and $\epsilon^{(\bar{n})} = \sigma_{r+1}(X^{(\bar{n})})$*. If additionally the condition number* κ *and* ζ *are small enough, or more precisely, if*

$$\mu \|\eta^{(\bar{n})}\|_{S_\infty}^{p(1-p)} < 1 \tag{25}$$

with the constant

$$\mu := 2^{5p}(1 + \gamma_{2r})^p \Big(\frac{\gamma_{2r}(3 + \gamma_{2r})(1 + \gamma_{2r})}{(1 - \gamma_{2r})} \Big)^{2-p} \Big(\frac{d - r}{r} \Big)^{2 - \frac{p}{2}} r^p \frac{\sigma_r(X_0)^{p(p-1)}}{(1 - \zeta)^{2p}} \kappa^p \tag{26}$$

then

$$\|\eta^{(n+1)}\|_{S_\infty} \leqslant \mu^{1/p} \left(\|\eta^{(n)}\|_{S_\infty} \right)^{2-p} \quad \text{and} \quad \|\eta^{(n+1)}\|_{S_p} \leqslant \mu^{1/p} \left(\|\eta^{(n)}\|_{S_p} \right)^{2-p}$$

for all $n \geqslant \bar{n}$*.*

We think that the result of Theorem 11 is remarkable, since there are only few low-rank recovery algorithms which exhibit either theoretically or practically verifiable superlinear convergence rates. In particular, although the algorithms of Mishra et al. (2013)

and `NewtonSLRA` of Schost and Spaenlehauer (2016) do show superlinear convergence rates, the first is not competitive to `HM-IRLS` in terms of sample complexity and the second has neither applicable theoretical guarantees for most of the interesting problems nor the ability of solving medium size problems.

Remark 12 *It is interesting to compare Theorem 11 with a related result for an IRLS algorithm for the sparse vector recovery problem in Daubechies et al. (2010, Theorem 7.9). We observe that while the statement describes the observed rates of convergence very accurately (cf. Section 5.2), the assumption* (25) *on the neighborhood that enables convergence of a rate* $2-p$ *is more pessimistic than our numerical experiments suggest. Our experiments confirm that the local convergence rate of order* $2-p$ *also holds for matrix completion measurements, where the assumption of a Schatten-p null space property fails to hold, cf. Section 5.*

4.3 Discussion and comparison with existing IRLS algorithms

Optimally, we would like to have a statement in Theorem 9 about the accumulation points $\bar{X}$ being *global minimizers* of g_ϵ^p, instead of mere stationary points (Fornasier et al., 2011, Theorem 6.11), (Daubechies et al., 2010, Theorem 5.3). A statement that strong is, unfortunately, difficult to achieve due to the non-convexity of the Schatten-p quasi-norm and of the ϵ-perturbed version g_ϵ^p. Nevertheless, our theorems can be seen as analogues of Daubechies et al. (2010, Theorem 7.7), which discusses the convergence properties of an IRLS algorithm for sparse recovery based on ℓ_p-minimization with $p < 1$.

As already mentioned in previous sections, Fornasier et al. (2011) and Mohan and Fazel (2012) proposed IRLS algorithms for low-rank matrix recovery and analysed their convergence properties. The algorithm of Fornasier et al. (2011) corresponds (almost) to `IRLS-col` with $p = 1$ as explained in Section 3. In this context, Theorem 9 recovers the results of Fornasier et al. (2011, Theorem 6.11(i-ii)) for $p = 1$ and generalize them, with weaker conclusions due to the non-convexity, to the cases $0 < p < 1$. The algorithm `IRLS-`p of Mohan and Fazel (2012) is similar to the former, but differs in the choice of the ϵ-smoothing and also covers non-convex choices $0 < p < 1$. However, we note that in the non-convex case, its convergence result (Mohan and Fazel, 2012, Theorem 5.1) corresponds to Theorem 9(ii), but does not provide statements similar to (i) and (iii) of Theorem 9.

Theorem 11 with its analysis of the convergence rate is new in the sense that to the best of our knowledge, there are no convergence rate proofs for IRLS algorithms for the low-rank matrix recovery problem in the literature. Indeed, we refer to Remark 22 in Section 6.3 for an explanation why the variants of Fornasier et al. (2011) and Mohan and Fazel (2012) cannot exhibit superlinear convergence rates, unlike `HM-IRLS`.

We also note that there is a close connection between the statements of Theorems 9 and 11 and results that were obtained for an IRLS algorithm dedicated to the sparse vector recovery problem in Daubechies et al. (2010, Theorems 7.7 and 7.9).

5. Numerical experiments

In this section, we demonstrate first that the superlinear convergence rate that was proven theoretically for Algorithm 1 (`HM-IRLS`) in Theorem 11 can indeed be accurately verified in

numerical experiments, even beyond measurement operators fulfilling the strong null space property, and compare its performance to other variants of `IRLS`.

In Section 5.3, we then examine the recovery performance of `HM-IRLS` for the matrix completion setting with the performance of other state-of-the-art algorithms comparing the measurement complexities that are needed for successful recovery for many random instances.

The numerical experiments are conducted on Linux and Mac systems with MATLAB R2017b. An implementation of `HM-IRLS` for matrix completion including code reproducing many conducted experiments is available at `https://github.com/ckuemmerle/hm_irls`.

5.1 Experimental setup

In the experiments, we sample $(d_1 \times d_2)$ dimensional ground truth matrices X_0 of rank r such that $X_0 = U\Sigma V^*$, where $U \in \mathbb{R}^{d_1 \times r}$ and $V \in \mathbb{R}^{d_2 \times r}$ are independent matrices with i.i.d. standard Gaussian entries and $\Sigma \in \mathbb{R}^{r \times r}$ is a diagonal matrix with i.i.d. standard Gaussian diagonal entries, independent from U and V.

We recall that a rank-r matrix $X \in M_{d_1 \times d_2}$ has $d_f = r(d_1 + d_2 - r)$ degrees of freedom, which is the theoretical lower bound on the number of measurements that are necessary for exact reconstruction (Candès and Plan, 2011). The random measurement setting we use in the experiments can be described as follows: We take measurements of matrix completion type, sampling $m = \lfloor \rho d_f \rfloor$ entries of X_0 uniformly over its $d_1 d_2$ indices to obtain $Y = \Phi(X_0)$. Here, ρ is such that $\frac{d_1 d_2}{d_f} \geqslant \rho \geqslant 1$ and parametrizes the difficulty of the reconstruction problem, from very hard problems for $\rho \approx 1$ to easier problems for larger ρ.

However, this uniform sampling of Φ could yield instances of measurement operators whose information content is not large enough to ensure well-posedness of the corresponding low-rank matrix recovery problem, even if $\rho > 1$. More precisely, it is impossible to recover a matrix exactly if the number of revealed entries in any row or column is smaller than its rank r, which is explained and shown in the context of the proof of Pimentel-Alarcón et al. (2016, Theorem 1).

Thus, in order to provide for a sensible measurement model for small ρ, we exclude operators Φ that sample fewer than r entries in any row or column. Therefore, we adapt the uniform sampling model such that operators Φ are discarded and sampled again until the requirement of at least r entries per column and row is met and recovery can be achieved from a theoretical point of view.

We note that the described phenomenon is very related to the fact that matrix completion recovery guarantees for the uniform sampling model require at least one additional log factor, i.e., they require at least $m \geqslant \log(\max(d_1, d_2))d_f$ sampled entries (Davenport and Romberg, 2016, Section V).

While we detail the experiments for the matrix completion measurement setting just described in the remaining section, we add that Gaussian measurement models also lead to very similar results in experiments.

5.2 Convergence rate comparison with other IRLS algorithms

In this subsection, we vary the Schatten-p parameter between 0 and 1 and compare the corresponding convergence behavior of `HM-IRLS` with the IRLS variant `IRLS-col`, which

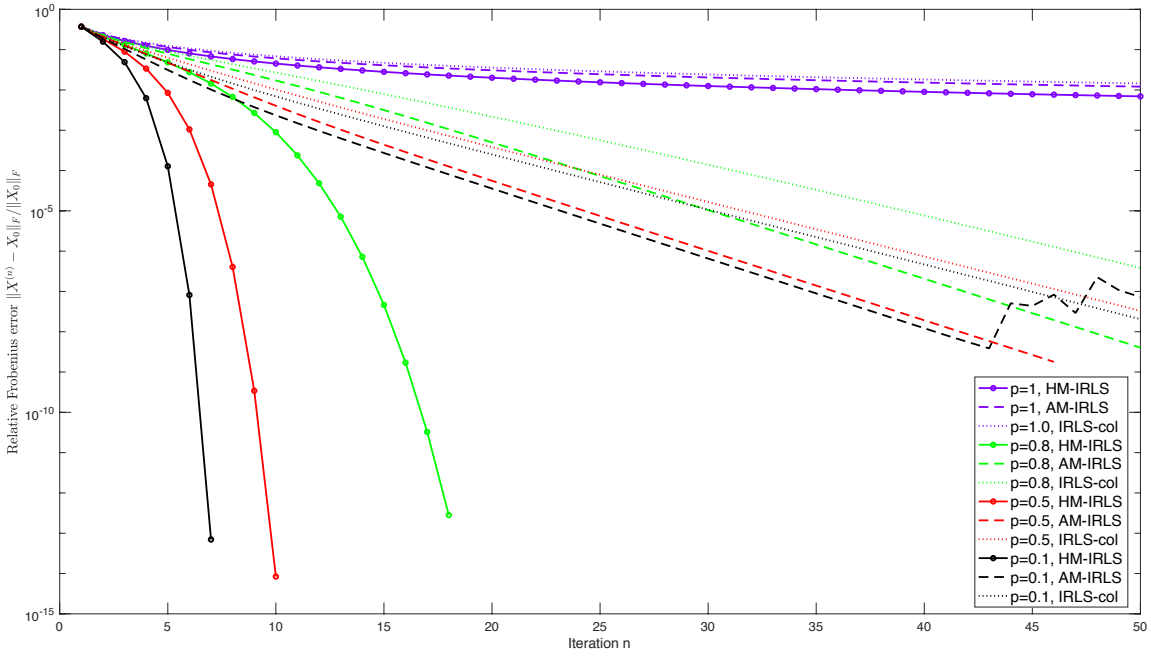


Figure 3: Relative Frobenius errors as a function of the iteration n for oversampling factor $\rho = 2$ (easy problem).

performs the reweighting just in the column space, and with the arithmetic mean variant `AM-IRLS`. The latter two coincide with Algorithm 1 except that the weight matrices are chosen as described in Equation (17) in Section 3.

We note that `IRLS-col` is very similar to the IRLS algorithms of Fornasier et al. (2011) and Mohan and Fazel (2012) and differs from them basically just in the choice of the ϵ-smoothing. We present the experiments with `IRLS-col` to isolate the influence of the weight matrix type, but very similar results can be observed for the algorithms of Fornasier et al. (2011) and Mohan and Fazel (2012).[2]

In the matrix completion setup of Section 5.1, we choose $d_1 = d_2 = 40$, $r = 10$ and distinguish easy, hard and very hard problems corresponding to oversampling factors ρ of 2.0, 1.2 and 1.0, respectively. The algorithms are provided with the ground truth rank r and are stopped whenever the relative change of Frobenius norm $\|X^{(n)} - X^{(n-1)}\|_F / \|X^{(n-1)}\|_F$ drops below the threshold of 10^{-10} or a maximal iteration of iterations $n_{\max}$ is reached.

5.2.1 Convergence rates

First, we study the behavior of the three IRLS algorithms for the easy setting of an oversampling factor of $\rho = 2$, which means that $\frac{2r(d_1+d_2-r)}{d_1 d_2} = 0.875$ of the entries are sampled, and parameters $p \in \{0.1, 0.5, 0.8, 1\}$.

In Figure 3, we observe that for $p = 1$, `HM-IRLS`, `AM-IRLS` and `IRLS-col` have a quite similar behavior, as the relative Frobenius errors $\|X^{(n)} - X_0\|_F / \|X_0\|_F$ decrease only slowly,

2. Implementations of the mentioned authors' algorithms were downloaded from `https://faculty.washington.edu/mfazel/` and `https://github.com/rward314/IRLSM`, respectively.

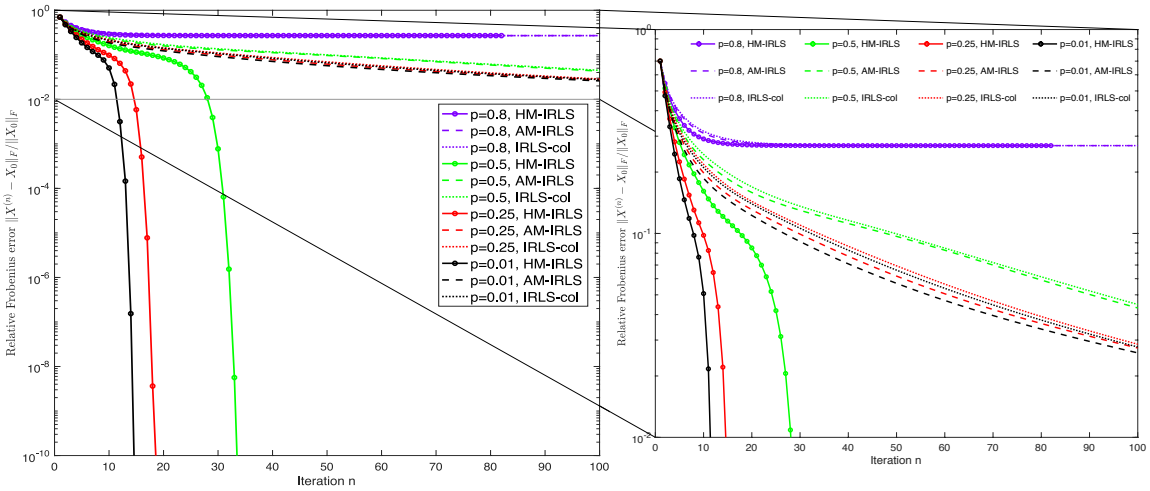


Figure 4: Relative Frobenius errors as a function of the iteration n for oversampling factor $\rho = 1.2$ (hard problem). Left column: y-range $[10^{-10}; 10^{0}]$. Right column: Enlarged section of left column corresponding to y-range of $[10^{-2}; 10^{0}]$.

i.e., even a linear rate is hardly identifiable. For choices $p < 1$ that correspond to non-convex objectives, we observe a very fast, superlinear convergence of `HM-IRLS`, as the iterates $X^{(n)}$ converge up to a relative error of less than 10^{-12} within fewer than 20 iterations for $p \in \{0.8, 0.5, 0.1\}$. Precise calculations verify that the rate of convergences are indeed of order $2 - p$, the order predicted by Theorem 11. We note that this fast convergence rate not only kicks in locally, but starting from the very first iteration.

On the other hand, it is easy to see that `AM-IRLS` and `IRLS-col` converge *linearly, but not superlinearly* to the ground truth X_0 for $p \in \{0.8, 0.5, 0.1\}$. The linear rate of `AM-IRLS` is slightly better than the one of `IRLS-col`, but the numerical stability of `AM-IRLS` deteriorates for $p = 0.1$ close to the ground truth (after iteration 43). This is due to a bad conditioning of the quadratic problems as the $X^{(n)}$ are close to rank-r matrices. In contrast, no numerical instability issues can be observed for `HM-IRLS`.

For the hard matrix completion problems with oversampling factor of $\rho = 1.2$, we observe that for $p = 0.8$, the three algorithms typically do not converge to ground truth. This can be seen in the example that is shown in Figure 4, where `HM-IRLS`, `AM-IRLS` and `IRLS-col` all exhibit a relative error of 0.27 after 100 iterations. We do not visualize the result for $p = 1$, as the iterates of the three algorithms do not converge to the ground truth either, which is to be expected: In some sense, they implement nuclear norm minimization, which is typically not able to recover a low-rank matrix from measurements with an oversampling factor as small as $\rho = 1.2$ (Donoho et al., 2013). The dramatic difference in behavior between `HM-IRLS` and the other approaches becomes very apparent for more non-convex choices of $p \in \{0.01, 0.25, 0.5\}$, where the former converges up to a relative Frobenius error of less than 10^{-10} within 15 to 35 iterations, while the others do not reach a relative error of 10^{-2} even after 100 iterations. For `HM-IRLS`, the convergence of order $2 - p$ can be very well locally observed also here, it just takes some iterations until the superlinear convergence begins, which is due to the increased difficulty of the recovery problem.

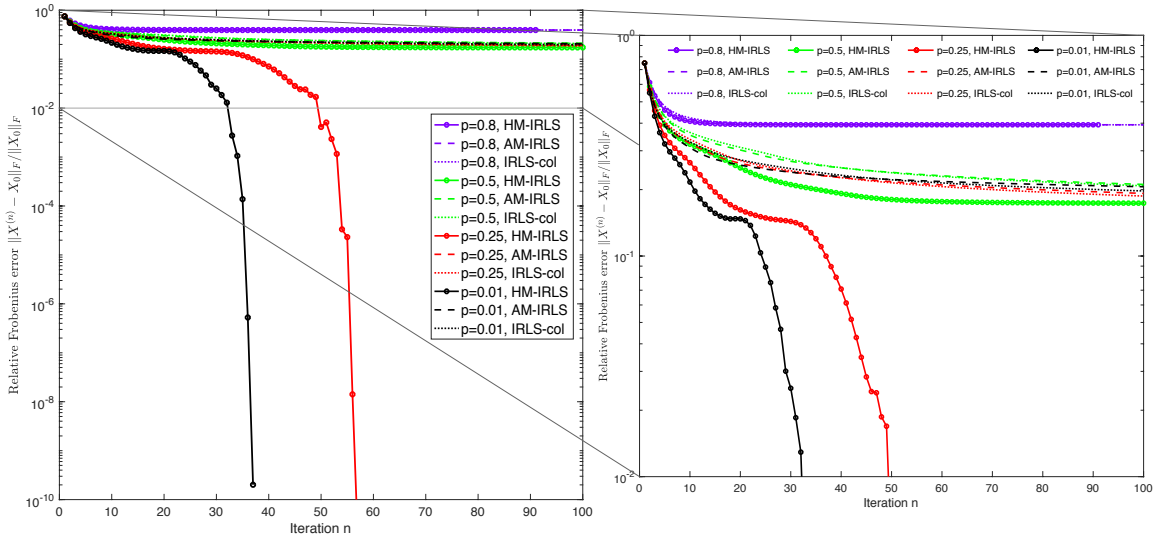


Figure 5: Relative Frobenius errors as a function of the iteration n for oversampling factor $\rho = 1.0$ (very hard problem). Left column: y-range $[10^{-10}; 10^{0}]$. Right column: Enlarged section of left column corresponding to y-range of $[10^{-2}; 10^{0}]$.

Finally, we see in the example shown in Figure 5 that even for the very hard problems where $\rho = 1$, which means that the number of sampled entries corresponds exactly to the degrees of freedom $r(d_1 + d_2 - r)$, `HM-IRLS` can be successful to recover the rank-r matrix if the parameter p is chosen small enough (here: $p \leqslant 0.25$). This is not the case for the algorithms `AM-IRLS` and `IRLS-col`.

5.2.2 `HM-IRLS` as the best extension of IRLS for sparse recovery

We summarize that among the three variants `HM-IRLS`, `AM-IRLS` and `IRLS-col`, only `HM-IRLS` is able to solve the low-rank matrix recovery problem for very low sample complexities corresponding to $\rho \approx 1$. Furthermore, it is the only IRLS algorithm for low-rank matrix recovery that exhibits a superlinear rate of convergence at all.

It is worthwhile to compare the properties of `HM-IRLS` with the behavior of the IRLS algorithm of Daubechies et al. (2010) designed to solve the sparse vector recovery problem by mimicking ℓ_p-minimization for $0 < p \leqslant 1$. While neither `IRLS-col` nor `AM-IRLS` are able to generalize the superlinear convergence behavior of Daubechies et al. (2010) (which is illustrated in Figure 8.3 of the same paper) to the low-rank matrix recovery problem, `HM-IRLS` is, as can be seen in Figures 3 to 5.

Taking the theoretical guarantees as well as the numerical evidence into account, we claim that ***`HM-IRLS`*** *is the presently best extension of IRLS for vector recovery in Daubechies et al. (2010) to the low-rank matrix recovery setting*, providing a substantial improvement over the reweighting strategies of Fornasier et al. (2011) and Mohan and Fazel (2012).

Moreover, we mention two observations which suggest that `HM-IRLS` has in some sense even more favorable properties than the algorithm of Daubechies et al. (2010): First, the discussion of Daubechies et al. (2010, Section 8) states that a superlinear convergence can only be observed locally after a considerable amount of iterations with just a linear error

decay. In contrast to that, HM-IRLS exhibits a superlinear error decay quite early (i.e., for example as early as after two iterations), at least if the sample complexity is large enough, cf. Figure 3.

Secondly, it can be observed that the convergence of the algorithm of Daubechies et al. (2010) to a sparse vector often breaks down if p is smaller than 0.5 (Daubechies et al., 2010, Section 8). In contrast to that, we observe that HM-IRLS does not suffer from this loss of global convergence for $p \ll 0.5$. Thus, a choice of very small parameters $p \approx 0.1$ or smaller is suggested as such a choice is accompanied by a very fast convergence.

5.3 Recovery performance compared to state-of-the-art algorithms

After comparing the performance of HM-IRLS with other IRLS variants, we now conduct experiments to compare the empirical performance of HM-IRLS also to that of low-rank matrix recovery algorithms different from IRLS.

To obtain a comprehensive picture, we consider not only the IRLS variants AM-IRLS and IRLS-col, but a variety of state-of-the-art methods in the experiments, as the Riemannian optimization technique Riemann_Opt (Vandereycken, 2013), the alternating minimization approaches AltMin (Haldar and Hernando, 2009), ASD (Tanner and Wei, 2016) and BFGD (Park et al., 2016), and finally the algorithms Matrix ALPS II (Kyrillidis and Cevher, 2014) and CGIHT_Matrix (Blanchard et al., 2015), which are based on iterative hard thresholding. As the IRLS variants we consider, all these algorithms use knowledge about the actual ground truth rank r.

In the experiments, we examine the empirical recovery probabilities of the different algorithms systematically for varying oversampling factors ρ, determining the difficulty of the low-rank recovery problem as the sample complexity fulfills $m = \lfloor \rho d_f \rfloor$. We recall that a large parameter ρ corresponds to an easy reconstruction problem, while a small ρ, e.g., $\rho \approx 1$, defines a very hard problem.

We choose $d_1 = d_2 = 100$ and the $r = 8$ as parameter of the experimental setting, conducting the experiments to recover rank-8 matrices $X_0 \in \mathbb{R}^{100\times 100}$. We remain in the matrix completion measurement setting described in Section 5.1, but sample now 150 random instances of X_0 and Φ for different numbers of measurements varying between $m_{\min} = 1500$ to $m_{\max} = 4000$. This means that the oversampling factor ρ increases from $\rho_{\min} = 0.975$ to $\rho_{\max} = 2.60$. For each algorithm, a successful recovery of X_0 is defined as a relative Frobenius error $\|X^{\text{out}} - X_0\|_F / \|X_0\|_F$ of the matrix X^{out} returned by the algorithm of smaller than 10^{-3}. The algorithms are run until stagnation of the iterates or until the maximal number of iterations $n_{\max} = 3000$ is reached. The number $n_{\max}$ is chosen large enough to ensure that a recovery failure is not due to a lack of iterations.

In the experiments, except for AltMin, for which we used our own implementation, we used implementations provided by the authors of the corresponding papers for the respective algorithms, using default input parameters provided by the authors. The respective code sources can be found in the references.

5.3.1 Beyond the state-of-the-art performance of HM-IRLS

The results of the experiment can be seen in Figure 6. We observe that HM-IRLS exhibits a very high empirical recovery probability for $p = 0.1$ and $p = 0.5$ as soon as the sample

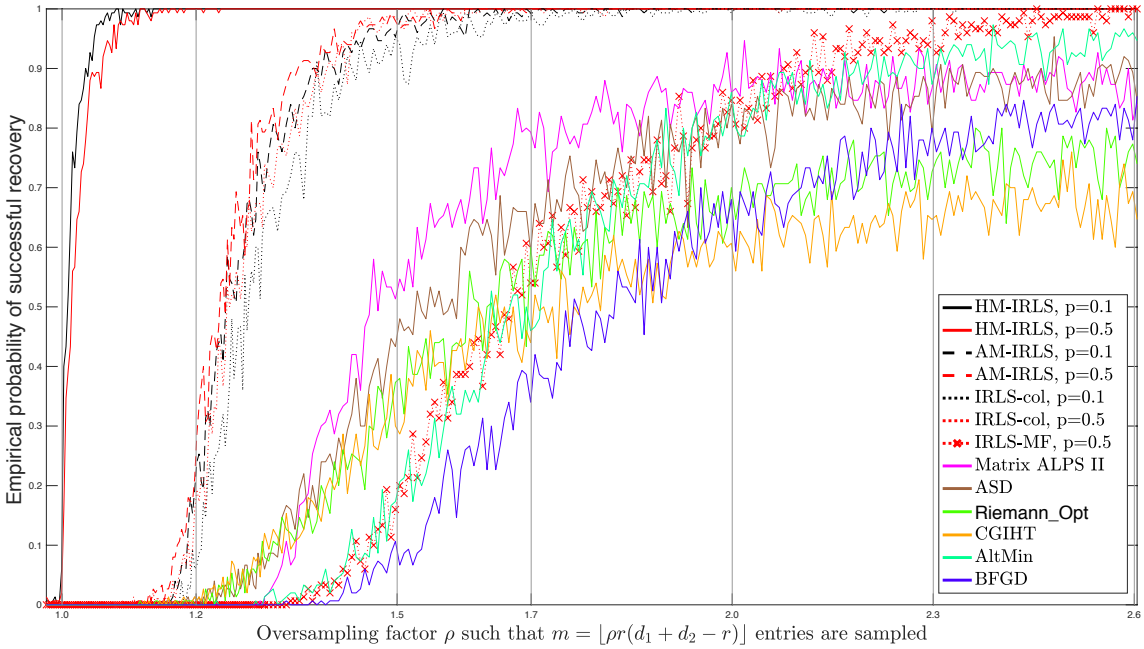


Figure 6: Comparison of empirical success rates of state-of-the-art algorithms, as a function of the oversampling factor ρ

complexity parameter ρ is slightly larger than 1.0, which means that $m = \lfloor \rho r(d_1 + d_2 - r) \rfloor$ measurements suffice to recover $(d_1 \times d_2)$-dimensional rank-r matrices with ρ close to 1. This is very close to the information theoretical lower bound of $d_f = r(d_1 + d_2 - r)$. Very interestingly, it can be observed that the empirical recovery probability reaches almost 1 already for an oversampling factor of $\rho \approx 1.1$, and remains at exactly 1 starting from $\rho \approx 1.2$.

Relatively good success rates can also be observed for the algorithms `AM-IRLS` and `IRLS-col` for non-convex parameter choices $p \in \{0.1, 0.5\}$, reaching an empirical success probability of almost 100% at around $\rho = 1.5$. `AM-IRLS` performs only marginally better than the classical IRLS strategy `IRLS-col`, which are both outperformed considerably by `HM-IRLS`. It is important to note that in accordance to what was observed in Section 5.2, in the successful instances, the error threshold that defines successful recovery is achieved already after a few dozen iterations for `HM-IRLS`, while typically only after several or many hundreds for `AM-IRLS` and `IRLS-col`. Furthermore, it is interesting to observe that the algorithm `IRLS-MF`, which corresponds to the variant studied and implemented by Mohan and Fazel (2012) and differs from `IRLS-col` mainly only in the choice of the ϵ-smoothing (14), has a considerably worse performance than the other IRLS methods. This is plausible since the smoothing influences severely the optimization landscape of the objective to be minimized.

The strong performance of `HM-IRLS` is in stark contrast to the behavior of all the algorithms that are based on different approaches than IRLS and that we considered in our experiments. They basically never recover any rank-r matrix if $\rho < 1.2$, and most of the algorithms need a sample complexity parameter of $\rho > 1.7$ to exceed a empirical recovery probability of a mere 0.5. A success rate of close to 0.8 is reached not before raising ρ above

2.0 in our experimental setting, and also only for a subset of the comparison algorithms, in particular for `Matrix ALPS II, ASD, AltMin`. The empirical probability of 1 is only reached for some of the IRLS methods, and not for any competing method in our experimental setting, even for rather large oversampling factors such as $\rho = 2.5$. While we do not rule out that a possible parameter tuning could improve the performance of any of the algorithms slightly, we conclude that for hard matrix completion problems, the experimental evidence for the vast differences in the recovery performance of `HM-IRLS` compared to other methods is very apparent.

Thus, our observation is that the proposed `HM-IRLS` algorithm *recovers low-rank matrices systematically with nearly the optimal number of measurements and needs fewer measurements than all the state-of-the-art algorithms we included in our experiments*, if the non-convexity parameter p is chosen such that $p \ll 1$.

We also note that the very sharp phase transition between failure and success that can be observed in Figure 6 for `HM-IRLS` indicates that the sample complexity parameter ρ is indeed the major variable determining the success of `HM-IRLS`. In contrast, the wider phase transitions for the other algorithms suggest that they might depend more on other factors, as the realizations of the random sampling model and the interplay of measurement operator Φ and ground truth matrix X_0.

Another conclusion that can be drawn from the empirical recovery probability of 1 is that, despite the severe non-convexity of the underlying Schatten-p quasi-norm for, e.g., $p = 0.1$, `HM-IRLS` with the initialization of $X^{(1)}$ as the Frobenius norm minimizer does not get stuck in stationary points if the oversampling factor is large enough. Further experiments conducted with random initializations as well as severely adversary initializations, e.g., with starting points chosen in the orthogonal complement of the spaces spanned by the singular vectors of the ground truth matrix X_0, lead to comparable results. Therefore, we claim that `HM-IRLS` exhibits a global convergence behavior in interesting application cases and for oversampling factor ranges for which competing non-convex low-rank matrix recovery algorithms fail to succeed. We consider a theoretical investigation of such behavior as an interesting open problem to explore.

5.4 Computational complexity

While the harmonic mean weight matrix $\widetilde{W}^{(n)}$, cf. (15), is an inverse of a $(d_1 d_2 \times d_1 d_2)$-matrix and therefore in general a dense $(d_1 d_2 \times d_1 d_2)$-matrix, it is important to note that it never has to be computed explicitly in an implementation of `HM-IRLS`; neither is it necessary to compute its inverse $(\widetilde{W}^{(n)})^{-1} = \frac{1}{2}\left(U^{(n)}(\bar{\Sigma}^{(n)})^{2-p}U^{(n)*} \oplus V^{(n)}(\bar{\Sigma}^{(n)})^{2-p}V^{(n)*}\right)$ explicitly.

Indeed, as it can be seen in (13) and by the definition of the Kronecker sum (55), the harmonic mean weight matrix appears just as the linear operator $(\mathcal{W}^{(n)})^{-1}$ on the space of matrices $M_{d_1 \times d_2}$, whose action consists of a left- and right-sided matrix multiplication, cf. (12). Therefore, the application of $(\mathcal{W}^{(n)})^{-1}$ is $O(d_1 d_2 (d_1 + d_2))$ by the naive matrix multiplication algorithm, and can be easily parallelized.

While this useful observation is helpful for the implementation of `HM-IRLS`, it is not true for `AM-IRLS`, as the action of $(W^{(n)}_{(\mathrm{arith})})^{-1}$, the inverse of the arithmetic mean weight matrix at iteration n, is not representable as a sum of left- and right-sided matrix multiplication.

This means that even the execution of a fixed number of iterations of HM-IRLS is faster than computational advantage over AM-IRLS.

The cost to compute $\Phi \circ \widetilde{\mathcal{W}}^{(n)-1} \circ \Phi^* \in M_{m\times m}$ depends on the linear measurement operator Φ. In the matrix completion setting (23), no additional arithmetic operations have to be performed, as Φ is a just a selection operator in this case, and for HM-IRLS, this means that $\Phi \circ \widetilde{\mathcal{W}}^{(n)-1} \circ \Phi^*$ is a sparse matrix.

Thus, the algorithm HM-IRLS consists of basically of two computational steps per iteration: The computation of the SVD of the $d_1 \times d_2$-matrix $X^{(n)}$ and the solution of the linearly constrained least squares problem in (13). The first is of time complexity $O(d_1 d_2 \min(d_1, d_2))$. The time complexity of the second depends on Φ, but is dominated by the inversion of a symmetric, $m \times m$ sparse linear system in the matrix completion setting, if m is the number of given entries. This has a worst case time complexity of $O(\max(d_1, d_2)^3 r^3)$ if ρ is just a constant oversampling factor.

For the matrix completion case, this allows us to recover low-rank matrices up to, e.g., $d_1 = d_2 = 3000$ on a single machine given very few entries with HM-IRLS.

Acceleration possibilities and extensions

To tackle higher dimensionalities in reasonable runtimes, a key strategy could be to address the computational bottleneck of HM-IRLS, the solution of the $m \times m$ linear system in (13), by using iterative methods. For IRLS algorithms designed for the related sparse recovery problem, the usage of conjugate gradient (CG) methods is discussed in Fornasier et al. (2016). By coupling the accuracy of the CG solutions to the outer IRLS iteration and using appropriate preconditioning, the authors obtain a competitive solver for the sparse recovery problem, also providing a convergence analysis. Similar ideas could be used for an acceleration of HM-IRLS.

It is interesting to see if further computational improvements can be achieved by combining the ideas of HM-IRLS with the usage of truncated and randomized SVDs (Halko et al., 2011), replacing the full SVDs of the $X^{(n)}$ that are needed to define the linear operator $(\mathcal{W}^{(n)})^{-1}$ in Algorithm 1.

6. Theoretical analysis

For the theoretical analysis of HM-IRLS, we introduce the following auxiliary functional $\mathcal{J}_p$, leading to a variational interpretation of the algorithm. In the whole section, we denote $d = \min(d_1, d_2)$ and $D = \max(d_1, d_2)$.

Definition 13 *Let $0 < p \leqslant 1$. Given a full rank matrix $Z \in M_{d_1\times d_2}$, let*

$$\widetilde{W}(Z) := 2\big[\mathbf{I}_{d_2} \otimes (ZZ^*)^{\frac{1}{2}}\big]\Big[(ZZ^*)^{\frac{1}{2}} \oplus (Z^*Z)^{\frac{1}{2}}\Big]^{-1}\big[(Z^*Z)^{\frac{1}{2}} \otimes \mathbf{I}_{d_1}\big] \in H_{d_1 d_2 \times d_1 d_2}$$

be the harmonic mean matrix $\widetilde{W}$ associated to Z.

We define the auxiliary functional $\mathcal{J}_p : M_{d_1\times d_2} \times \mathbb{R}_{\geqslant 0} \times M_{d_1\times d_2} \to \mathbb{R}_{\geqslant 0}$ *as*

$$\mathcal{J}_p(X, \epsilon, Z) := \begin{cases} \frac{p}{2}\|X_{\mathrm{vec}}\|^2_{\ell_2(\widetilde{W}(Z))} + \frac{\epsilon^2 p}{2}\sum_{i=1}^{d}\sigma_i(Z) + \frac{2-p}{2}\sum_{i=1}^{d}\sigma_i(Z)^{\frac{p}{(p-2)}} & \text{if } \operatorname{rank}(Z) = d, \\ +\infty & \text{if } \operatorname{rank}(Z) < d. \end{cases}$$

We note that the matrix $\widetilde{W}$ of Definition 13 is just the harmonic mean of the matrices $\widetilde{W}_1 := \mathbf{I}_{d_2} \otimes (ZZ^*)^{\frac{1}{2}}$ and $\widetilde{W}_2 = (Z^*Z)^{\frac{1}{2}} \otimes \mathbf{I}_{d_1}$, as introduced in Section 2.3, if $(ZZ^*)^{\frac{1}{2}}$ *and* $(Z^*Z)^{\frac{1}{2}}$ are positive definite. Indeed, in this case, $(ZZ^*)^{\frac{1}{2}} \oplus (Z^*Z)^{\frac{1}{2}} = \widetilde{W}_1 + \widetilde{W}_2$ is invertible and as $(A^{-1} + B^{-1})^{-1} = A(A+B)^{-1}B$ for any positive definite matrices A and B of the same dimensions,

$$\widetilde{W}(Z) = 2\widetilde{W}_1\big(\widetilde{W}_1 + \widetilde{W}_2\big)^{-1}\widetilde{W}_2 = 2(\widetilde{W}_1^{-1} + \widetilde{W}_2^{-1})^{-1}. \tag{27}$$

We use the more general definition $\widetilde{W}(Z)$ as it is well-defined for any full-rank $Z \in M_{d_1 \times d_2}$ and as it allows to handle the case of non-square matrices, i.e., the case $d_1 \neq d_2$, as in this case $(ZZ^*)^{\frac{1}{2}}$ or $(Z^*Z)^{\frac{1}{2}}$ has to be singular. Using the Moore-Penrose pseudo inverse $\widetilde{W}_1^+$ and $\widetilde{W}_2^+$ of the matrices $\widetilde{W}_1$ and $\widetilde{W}_2$, we can rewrite $\widetilde{W}(Z)$ from Definition 13 as

$$\widetilde{W}(Z) = 2\widetilde{W}_1\big(\widetilde{W}_1 + \widetilde{W}_2\big)^{-1}\widetilde{W}_2 = 2(\widetilde{W}_1^{+} + \widetilde{W}_2^{+})^{-1}.$$

With the auxiliary functional $\mathcal{J}_p$ at hand, we can interpret Algorithm 1 as an alternating minimization of the functional $\mathcal{J}_p(X, \epsilon, Z)$ with respect to its arguments X, ϵ and Z.

In the following, we derive the formula (15) for the weight matrix $\widetilde{W}^{(n+1)}$ as the evaluation $\widetilde{W}^{(n+1)} = \widetilde{W}\big(Z^{(n+1)}\big)$ of $\widetilde{W}$ from Definition 13 at the minimizer

$$Z^{(n+1)} = \operatorname*{arg\,min}_{Z \in M_{d_1 \times d_2}} \mathcal{J}_p(X^{(n+1)}, \epsilon^{(n+1)}, Z), \tag{28}$$

with the minimizer being unique. Similarly, formula (13) can be interpreted as

$$X^{(n+1)} = \operatorname*{arg\,min}_{\substack{X \in M_{d_1 \times d_2} \\ \Phi(X) = Y}} \|X_{\mathrm{vec}}\|^2_{\ell_2(\widetilde{W}(Z^{(n)}))} = \operatorname*{arg\,min}_{\substack{X \in M_{d_1 \times d_2} \\ \Phi(X) = Y}} \mathcal{J}_p(X, \epsilon^{(n)}, Z^{(n)}). \tag{29}$$

These observations constitute the starting point of the convergence analysis of Algorithm 1, which is detailed subsequently after the verification of the optimization steps.

6.1 Optimization of $\mathcal{J}_p$ with respect to Z and X

We fix $X \in M_{d_1 \times d_2}$ with singular value decomposition $X = \sum_{i=1}^{d} \sigma_i u_i v_i^*$, where $u_i \in \mathbb{C}^{d_1}$ $v_i \in \mathbb{C}^{d_2}$ are the left and right singular vectors respectively and $\sigma_i = \sigma_i(X)$ denote its singular values for $i \in [d]$.

Our objective in the following is the justification of formula (15). To yield the building blocks of the weight matrix $\widetilde{W}^{(n+1)}$, we consider the minimization problem

$$\operatorname*{arg\,min}_{Z \in M_{d_1 \times d_2}} \mathcal{J}_p(X, \epsilon, Z) \tag{30}$$

for $\epsilon > 0$.

Lemma 14 *The unique minimizer of* (30) *is given by*

$$Z_{\mathrm{opt}} = \sum_{i=1}^{d} (\sigma_i(X)^2 + \epsilon^2)^{\frac{p-2}{2}} u_i v_i^*.$$

Furthermore, the value of $\mathcal{J}_p$ at the minimizer Z_{opt} is

$$\mathcal{J}_p(X, \epsilon, Z_{\mathrm{opt}}) = \sum_{i=1}^{d} (\sigma_i(X)^2 + \epsilon^2)^{\frac{p}{2}} =: g_\epsilon^p(X) \tag{31}$$

for $p > 0$.

The proof of Lemma 14 is detailed in Appendix B.

Remark 15 *We note that the value of $\mathcal{J}_p(X, \epsilon, Z_{\mathrm{opt}})$ can be interpreted as a smooth ϵ-perturbation of a p-th power of a Schatten-p quasi-norm of the matrix X. In fact, for $\epsilon = 0$ we have*

$$\mathcal{J}_p(X, 0, Z_{\mathrm{opt}}) = \|X\|_{S_p}^p = g_0^p(X).$$

Now, we show that our definition rule (13) of $X^{(n+1)}$ in Algorithm 1 can be interpreted as a minimization of the auxiliary functional $\mathcal{J}_p$ with respect to the variable X. Additionally, this minimization step can be formulated as the solution of a weighted least squares problem with weight matrix $\widetilde{W}^{(n)}$. This is summarized in the following lemma.

Lemma 16 *Let $0 < p \leqslant 1$. Given a full-rank matrix $Z \in M_{d_1 \times d_2}$, let $\widetilde{W}(Z) := 2([(ZZ^*)^{\frac{1}{2}}]^+ \oplus [(Z^*Z)^{\frac{1}{2}}]^+)^{-1} \in H_{d_1 d_2 \times d_1 d_2}$ be the matrix from Definition 13 and $\mathcal{W}^{-1} : M_{d_1 \times d_2} \to M_{d_1 \times d_2}$ the linear operator of its inverse*

$$\mathcal{W}^{-1}(X) := \frac{1}{2}\left[[(ZZ^*)^{\frac{1}{2}}]^+ X + X[(Z^*Z)^{\frac{1}{2}}]^+\right].$$

Then the matrix

$$X_{\mathrm{opt}} = \left(\mathcal{W}^{-1} \circ \Phi^* \circ (\Phi \circ \mathcal{W}^{-1} \circ \Phi^*)^{-1}\right)(Y) \in M_{d_1 \times d_2}$$

is the unique minimizer of the optimization problems

$$\operatorname*{arg\,min}_{\Phi(X)=Y} \mathcal{J}_p(X, \epsilon, Z) = \operatorname*{arg\,min}_{\Phi(X)=Y} \|X_{\mathrm{vec}}\|^2_{\ell_2(\widetilde{W})}. \tag{32}$$

Moreover, a matrix $X_{\mathrm{opt}} \in M_{d_1 \times d_2}$ is a minimizer of the minimization problem (32) if and only if it fulfills the property

$$\langle \widetilde{W}(Z)(X_{\mathrm{opt}})_{\mathrm{vec}}, H_{\mathrm{vec}} \rangle_{\ell_2} = 0 \quad \text{for all} \quad H \in \mathcal{N}(\Phi) \quad \text{and} \quad \Phi(X_{\mathrm{opt}}) = Y. \tag{33}$$

In Appendix B, the interested reader can find a sketch of the proof of this lemma.

6.2 Basic properties of the algorithm and convergence results

In the following subsection, we will have a closer look at Algorithm 1 and point out some of its properties, in particular, the boundedness of the iterates $(X^{(n)})_{n \in \mathbb{N}}$ and the fact that two consecutive iterates are getting arbitrarily close as $n \to \infty$. These results will be used to show convergence and to determine the rate of convergence of Algorithm 1 under conditions determined along the way.

Lemma 17 *Let $(X^{(n)}, \epsilon^{(n)})_{n\in\mathbb{N}}$ be the sequence of iterates and smoothing parameters of Algorithm 1. Let $X^{(n)} = \sum_{i=1}^{d} \sigma_i^{(n)} u_i^{(n)} v_i^{(n)*}$ be the SVD of the n-th iterate $X^{(n)}$. Let $(Z^{(n)})_{n\in\mathbb{N}}$ be a corresponding sequence such that*

$$Z^{(n)} = \sum_{i=1}^{d} (\sigma_i^{(n)2} + \epsilon^{(n)2})^{\frac{p-2}{2}} u_i^{(n)} v_i^{(n)*}$$

for $n \in \mathbb{N}$. Then the following properties hold:

(a) $\mathcal{J}_p(X^{(n)}, \epsilon^{(n)}, Z^{(n)}) \geqslant \mathcal{J}_p(X^{(n+1)}, \epsilon^{(n+1)}, Z^{(n+1)})$ *for all* $n \geqslant 1$,

(b) $\|X^{(n)}\|_{S_p}^p \leqslant \mathcal{J}_p(X^{(1)}, \epsilon^{(0)}, Z^{(0)}) =: \mathcal{J}_{p,0}$ *for all* $n \geqslant 1$,

(c) The iterates $X^{(n)}, X^{(n+1)}$ come arbitrarily close as $n \to \infty$, i.e., $\lim_{n\to\infty} \|(X^{(n)} - X^{(n+1)})_{\mathrm{vec}}\|_{\ell_2}^2 = 0$.

At this point we notice that, assuming $X^{(n)} \to \bar{X}$ and $\epsilon^{(n)} \to \bar{\epsilon}$ for $n \to \infty$ with the limit point $(\bar{X}, \bar{\epsilon}) \in M_{d_1\times d_2} \times \mathbb{R}_{\geqslant 0}$, it would follow that

$$\mathcal{J}_p(X^{(n)}, \epsilon^{(n)}, Z^{(n)}) \to g_{\bar{\epsilon}}^p(\bar{X})$$

for $n \to \infty$ by equation (31).

Now, let $\epsilon > 0$, a measurement vector $Y \in \mathbb{C}^m$ and the linear operator Φ be given and consider the optimization problem

$$\min_{\substack{X\in M_{d_1\times d_2} \\ \Phi(X)=Y}} g_\epsilon^p(X) \tag{34}$$

with $g_\epsilon^p(X) = \sum_{i=1}^{d} (\sigma_i(X)^2 + \epsilon^2)^{\frac{p}{2}}$ and $\sigma_i(X)$ being the i-th singular value of X, cf. (31). If $g_\epsilon^p(X)$ is non-convex, which is the case for $p < 1$, one might practically only be able to find critical points of the problem.

Lemma 18 *Let $X \in M_{d_1\times d_2}$ be a matrix with the SVD such that $X = \sum_{i=1}^{d} \sigma_i u_i v_i^*$, let $\epsilon > 0$. If we define*

$$\widetilde{W}(X, \epsilon) = 2\Big[\Big(\sum_{i=1}^{d} (\sigma_i^2 + \epsilon^2)^{\frac{2-p}{2}} u_i u_i^* \Big) \oplus \Big(\sum_{i=1}^{d} (\sigma_i^2 + \epsilon^2)^{\frac{2-p}{2}} v_i v_i^* \Big) \Big]^{-1} \in H_{d_1 d_2 \times d_1 d_2},$$

then $\widetilde{W}(X^{(n)}, \epsilon^{(n)}) = \widetilde{W}^{(n)}$, with $\widetilde{W}^{(n)}$ defined as in Algorithm 1, cf. (10). Furthermore, X is a critical point of the optimization problem (34) if and only if

$$\langle \widetilde{W}(X, \epsilon) X_{\mathrm{vec}}, H_{\mathrm{vec}} \rangle_{\ell_2} = 0 \quad \text{for all} \quad H \in \mathcal{N}(\Phi) \quad \text{and} \quad \Phi(X) = Y. \tag{35}$$

In the case that g_ϵ^p is convex, i.e., if $p = 1$, (35) implies that X is the unique minimizer of (34).

Now, we have some basic properties of the algorithm at hand that allow us, together with the strong nullspace property in Definition 4 to carry out the proof of the convergence result in Theorem 9. The proof is sketched in Appendix C using the results above.

6.3 Locally superlinear convergence

In the proof of Theorem 11 we use the following bound on perturbations of the singular value decomposition, which is originally due to Wedin (1972). It bounds the alignment of the subspaces spanned by the singular vectors of two matrices by their norm distance, given a gap between the first singular values of the one matrix and the last singular values of the other matrix that is sufficiently pronounced.

Lemma 19 (Wedin's bound (Stewart, 2006)) *Let X and $\bar{X}$ be two matrices of the same size and their singular value decompositions*

$$X = \begin{pmatrix} U_1 & U_2 \end{pmatrix} \begin{pmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{pmatrix} \begin{pmatrix} V_1^* \\ V_2^* \end{pmatrix} \quad \text{and} \quad \bar{D} = \begin{pmatrix} \bar{U}_1 & \bar{U}_2 \end{pmatrix} \begin{pmatrix} \bar{\Sigma}_1 & 0 \\ 0 & \bar{\Sigma}_2 \end{pmatrix} \begin{pmatrix} \bar{V}_1^* \\ \bar{V}_2^* \end{pmatrix},$$

where the submatrices have the sizes of corresponding dimensions. Suppose that δ, α satisfying $0 < \delta \leqslant \alpha$ are such that $\alpha \leqslant \sigma_{\min}(\Sigma_1)$ and $\sigma_{\max}(\bar{\Sigma}_2) < \alpha - \delta$. Then

$$\|\bar{U}_2^* U_1\|_{S_\infty} \leqslant \sqrt{2} \frac{\|X - \bar{X}\|_{S_\infty}}{\delta} \text{ and } \|\bar{V}_2^* V_1\|_{S_\infty} \leqslant \sqrt{2} \frac{\|X - \bar{X}\|_{S_\infty}}{\delta}. \tag{36}$$

As a first step towards the proof of Theorem 11, we show the following lemma.

Lemma 20 *Let $(X^{(n)})_n$ be the output sequence of Algorithm 1 for parameters Φ, Y, r and $0 < p \leqslant 1$, and $X_0 \in M_{d_1 \times d_2}$ be a matrix such that $\Phi(X_0) = Y$.*

(i) Let $\eta_{2r}^{(n+1)}$ be the best rank-$2r$ approximation of $\eta^{(n+1)} = X^{(n+1)} - X_0$. Then

$$\|\eta^{(n+1)} - \eta_{2r}^{(n+1)}\|_{S_p}^{2p} \leqslant 2^{2-p} \Bigg(\sum_{i=r+1}^{d} \left(\sigma_i^2(X^{(n)}) + \epsilon^{(n)2}\right)^{\frac{p}{2}} \Bigg)^{2-p} \|\eta_{\mathrm{vec}}^{(n+1)}\|_{\ell_2(\widetilde{W}^{(n)})}^{2p},$$

where $\widetilde{W}^{(n)}$ denotes the harmonic mean weight matrix from (10).

(ii) Assume that the linear map $\Phi : M_{d_1 \times d_2} \to \mathbb{C}^m$ fulfills the strong Schatten-p NSP of order $2r$ with constant $\gamma_{2r} < 1$. Then

$$\|\eta^{(n+1)}\|_{S_2}^{2p} \leqslant 2^p \frac{\gamma_{2r}^{2-p}}{r^{2-p}} \Bigg(\sum_{i=r+1}^{d} \left(\sigma_i^2(X^{(n)}) + \epsilon^{(n)2}\right)^{\frac{p}{2}} \Bigg)^{2-p} \|\eta_{\mathrm{vec}}^{(n+1)}\|_{\ell_2(\widetilde{W}^{(n)})}^{2p}. \tag{37}$$

(iii) Under the same assumption as for (ii), it holds that

$$\|\eta^{(n+1)}\|_{S_p}^{2p} \leqslant (1 + \gamma_{2r})^2 2^{2-p} \Bigg(\sum_{i=r+1}^{d} \left(\sigma_i^2(X^{(n)}) + \epsilon^{(n)2}\right)^{\frac{p}{2}} \Bigg)^{2-p} \|\eta_{\mathrm{vec}}^{(n+1)}\|_{\ell_2(\widetilde{W}^{(n)})}^{2p}.$$

Proof (i) Let the $X^{(n)} = \widetilde{U}^{(n)} \Sigma^{(n)} \widetilde{V}^{(n)*}$ be the (full) singular value decomposition of $X^{(n)}$, i.e., $\widetilde{U}^{(n)} \in \mathcal{U}_{d_1}$ and $\widetilde{V}^{(n)} \in \mathcal{U}_{d_2}$ are unitary matrices and $\Sigma^{(n)} = \mathrm{diag}(\sigma_1(X^{(n)}), \ldots, \sigma_r(X^{(n)})) \in M_{d_1 \times d_2}$. We define $U_T^{(n)} \in M_{d_1 \times r}$ as the matrix of the first r columns of $\widetilde{U}^{(n)}$ and $U_{T_c}^{(n)} \in$

$M_{d_1\times(d_1-r)}$ as the matrix of its last d_1-r columns, so that $\widetilde{U}^{(n)} = \begin{pmatrix} U_T^{(n)} & U_{T_c}^{(n)} \end{pmatrix}$, and similarly $V_T^{(n)}$ and $V_{T_c}^{(n)}$.

As $\mathbf{I}_{d_1} = U_T^{(n)}U_T^{(n)*} + U_{T_c}^{(n)}U_{T_c}^{(n)*}$ and $\mathbf{I}_{d_2} = V_T^{(n)}V_T^{(n)*} + V_{T_c}^{(n)}V_{T_c}^{(n)*}$, we note that

$$U_{T_c}^{(n)}U_{T_c}^{(n)*}\eta^{(n+1)}V_{T_c}^{(n)}V_{T_c}^{(n)*} = \eta^{(n+1)} - U_T^{(n)}U_T^{(n)*}\eta^{(n+1)} + U_{T_c}^{(n)}U_{T_c}^{(n)*}\eta^{(n+1)}V_T^{(n)}V_T^{(n)*},$$

while $U_T^{(n)}U_T^{(n)*}\eta^{(n+1)} + U_{T_c}^{(n)}U_{T_c}^{(n)*}\eta^{(n+1)}V_T^{(n)}V_T^{(n)*}$ has a rank of at most $2r$. This implies that

$$\|\eta^{(n+1)} - \eta_{2r}^{(n+1)}\|_{S_p} \leqslant \|U_{T_c}^{(n)}U_{T_c}^{(n)*}\eta^{(n+1)}V_{T_c}^{(n)}V_{T_c}^{(n)*}\|_{S_p} = \|U_{T_c}^{(n)*}\eta^{(n+1)}V_{T_c}^{(n)}\|_{S_p}. \tag{38}$$

Using the definitions of $\widetilde{U}^{(n)}$ and $\widetilde{V}^{(n)}$, we write the harmonic mean weight matrices of the n-th iteration (10) as

$$\widetilde{W}^{(n)} = 2\big(\widetilde{V}^{(n)}\otimes\widetilde{U}^{(n)}\big)\big(\bar{\Sigma}_{d_1}^{(n)2-p}\oplus\bar{\Sigma}_{d_2}^{(n)2-p}\big)^{-1}\big(\widetilde{V}^{(n)}\otimes\widetilde{U}^{(n)}\big)^*, \tag{39}$$

where $\bar{\Sigma}_{d_1}^{(n)} \in M_{d_1\times d_1}$ and $\bar{\Sigma}_{d_2}^{(n)} \in M_{d_2\times d_2}$ are the diagonal matrices with the smoothed singular values of $X^{(n)}$ from (11), but filled up with zeros if necessary. Using the abbreviation

$$\Omega := (\widetilde{V}^{(n)}\otimes\widetilde{U}^{(n)})^*\widetilde{W}^{(n)\frac{1}{2}}\eta_{\mathrm{vec}}^{(n+1)} \in \mathbb{C}^{d_1d_2}, \tag{40}$$

we rewrite

$$\begin{aligned}\eta_{\mathrm{vec}}^{(n+1)} &= \widetilde{W}^{(n)-\frac{1}{2}}\widetilde{W}^{(n)\frac{1}{2}}\eta_{\mathrm{vec}}^{(n+1)} = 2^{-1/2}\big(\widetilde{V}^{(n)}\otimes\widetilde{U}^{(n)}\big)\big(\bar{\Sigma}_{d_1}^{(n)2-p}\oplus\bar{\Sigma}_{d_2}^{(n)2-p}\big)^{1/2}\Omega \\ &= 2^{-1/2}\big(\widetilde{V}^{(n)}\otimes\widetilde{U}^{(n)}\big)\left[(\mathbf{I}_{d_2}\otimes\bar{\Sigma}_{d_1}^{(n)\frac{2-p}{2}})D_L + (\bar{\Sigma}_{d_2}^{(n)\frac{2-p}{2}}\otimes\mathbf{I}_{d_1})D_R\right]\Omega\end{aligned} \tag{41}$$

with the diagonal matrices $D_L, D_R \in M_{d_1d_2\times d_1d_2}$ such that

$$(D_L)_{i+(j-1)d_1,i+(j-1)d_1} = \Big(1+\Big(\frac{\sigma_j^2(X^{(n)})+\epsilon^{(n)2}}{\sigma_i^2(X^{(n)})+\epsilon^{(n)2}}\Big)^{\frac{2-p}{2}}\Big)^{-1/2}$$

and

$$(D_R)_{i+(j-1)d_1,i+(j-1)d_1} = \Big(\Big(\frac{\sigma_i^2(X^{(n)})+\epsilon^{(n)2}}{\sigma_j^2(X^{(n)})+\epsilon^{(n)2}}\Big)^{\frac{2-p}{2}}+1\Big)^{-1/2}$$

for $i\in[d_1]$ and $j\in[d_2]$. This can be seen from the definitions of the Kronecker product $\otimes$ and the Kronecker sum $\oplus$ (cf. Appendix A), as

$$\begin{aligned}&\Big(\big(\bar{\Sigma}_{d_1}^{(n)2-p}\oplus\bar{\Sigma}_{d_2}^{(n)2-p}\big)^{1/2}\Big)_{i+(j-1)d_1,i+(j-1)d_1} = (s_i+s_j)^{1/2} \\ &= s_i(s_i+s_j)^{-1/2} + s_j(s_i+s_j)^{-1/2} = s_i^{1/2}(1+\frac{s_j}{s_i})^{-1/2} + s_j^{1/2}(\frac{s_i}{s_j}+1)^{-1/2}\end{aligned}$$

if s_ℓ denotes the ℓ-th diagonal entry of $\bar{\Sigma}_{d_2}^{(n)2-p}$ and $\bar{\Sigma}_{d_1}^{(n)2-p}$ for $\ell\in[\max(d_1,d_2)]$.

If we write $\bar{\Sigma}_{d_1,T_c^2}^{(n)\frac{2-p}{2}} \in M_{(d_1-r)\times(d_1-r)}$ for the diagonal matrix containing the $d_1 - r$ last diagonal elements of $\bar{\Sigma}_{d_1}^{(n)2-p}$ and $\bar{\Sigma}_{d_2,T_c^2}^{(n)\frac{2-p}{2}} \in M_{(d_1-r)\times(d_1-r)}$ for the diagonal matrix containing the $d_2 - r$ last diagonal elements of $\bar{\Sigma}_{d_2}^{(n)2-p}$, it follows from (41) that

$$\begin{aligned}\|U_{T_c}^{(n)*}\eta^{(n+1)}V_{T_c}^{(n)}\|_{S_p}^p &= 2^{-\frac{p}{2}}\Big\|U_{T_c}^{(n)*}\widetilde{U}^{(n)}\Big[\bar{\Sigma}_{d_1}^{(n)\frac{2-p}{2}}(D_L\Omega)_{\mathrm{mat}} + (D_R\Omega)_{\mathrm{mat}}\bar{\Sigma}_{d_2}^{(n)\frac{2-p}{2}}\Big]\widetilde{V}^{(n)*}V_{T_c}^{(n)}\Big\|_{S_p}^p \\ &\leqslant 2^{-\frac{p}{2}}\Big\|\bar{\Sigma}_{d_1,T_c^2}^{(n)\frac{2-p}{2}}\big[(D_L\Omega)_{\mathrm{mat}}\big]_{T_c,T_c}\Big\|_{S_p}^p + \Big\|\big[(D_R\Omega)_{\mathrm{mat}}\big]_{T_c,T_c}\bar{\Sigma}_{d_2,T_c^2}^{(n)\frac{2-p}{2}}\Big\|_{S_p}^p\end{aligned}$$

with the notation that M_{T_c,T_c} denotes the submatrix of M which contains the intersection of the last $d_1 - r$ rows of M with its last $d_2 - r$ columns.

Now, Hölder's inequality for Schatten-p quasi-norms (e.g., Gohberg et al. (2000, Theorem 11.2)) can be used to see that

$$\Big\|\bar{\Sigma}_{d_1,T_c^2}^{(n)\frac{2-p}{2}}\big[(D_L\Omega)_{\mathrm{mat}}\big]_{T_c,T_c}\Big\|_{S_p}^p \leqslant \Big\|\bar{\Sigma}_{T_c}^{(n)\frac{2-p}{2}}\Big\|_{S_{\frac{2p}{2-p}}}^p \Big\|\big[(D_L\Omega)_{\mathrm{mat}}\big]_{T_c,T_c}\Big\|_{S_2}^p. \tag{42}$$

Inserting the definition

$$\Big\|\bar{\Sigma}_{T_c}^{(n)\frac{2-p}{2}}\Big\|_{S_{\frac{2p}{2-p}}}^p = \Big(\sum_{i=r+1}^{d}\big(\sigma_i^2(X^{(n)}) + \epsilon^{(n)2}\big)^{\frac{2p(2-p)}{(2-p)4}}\Big)^{\frac{2-p}{2}} = \Big(\sum_{i=r+1}^{d}\big(\sigma_i^2(X^{(n)}) + \epsilon^{(n)2}\big)^{\frac{p}{2}}\Big)^{\frac{2-p}{2}}$$

allows us to rewrite the first factor, while the second factor can be bounded by

$$\begin{aligned}\big\|\big[(D_L\Omega)_{\mathrm{mat}}\big]_{T_c,T_c}\big\|_{S_2}^p &\leqslant \big\|(D_L\Omega)_{\mathrm{mat}}\big\|_{S_2}^p \leqslant \|\Omega_{\mathrm{mat}}\|_{S_2}^p = \|(\widetilde{V}^{(n)}\otimes\widetilde{U}^{(n)})^*\widetilde{W}^{(n)\frac{1}{2}}\eta_{\mathrm{vec}}^{(n+1)}\|_{\ell_2}^p \\ &= \|\widetilde{W}^{(n)\frac{1}{2}}\eta_{\mathrm{vec}}^{(n+1)}\|_{\ell_2}^p = \|\eta_{\mathrm{vec}}^{(n+1)}\|_{\ell_2(\widetilde{W}^{(n)})}^p,\end{aligned}$$

as the matrix $D_L \in M_{d_1d_2\times d_1d_2}$ from (41) fulfills $\|D_L\|_{S_\infty} \leqslant 1$ since its entries are bounded by 1; we also recall the definition (40) of Ω and that $\widetilde{V}^{(n)}$ and $\widetilde{U}^{(n)}$ are unitary.

The term $\Big\|\big[(D_R\Omega)_{\mathrm{mat}}\big]_{T_c,T_c}\bar{\Sigma}_{d_2,T_c^2}^{(n)\frac{2-p}{2}}\Big\|_{S_p}^p$ in the bound of $\|U_{T_c}^{(n)*}\eta^{(n+1)}V_{T_c}^{(n)}\|_{S_p}^p$ can be estimated analogously. Combining this with (38), we obtain

$$\|\eta^{(n+1)} - \eta_{2r}^{(n+1)}\|_{S_p}^{2p} \leqslant 2^{-p}\Big(2\Big(\sum_{i=r+1}^{d}\big(\sigma_i^2(X^{(n)}) + \epsilon^{(n)2}\big)^{\frac{p}{2}}\Big)^{\frac{2-p}{2}}\Big)^2\|\eta_{\mathrm{vec}}^{(n+1)}\|_{\ell_2(\widetilde{W}^{(n)})}^{2p},$$

concluding the proof of statement (i).

(ii) Using the strong Schatten-p null space property (18) of order $2r$ and that $\eta^{(n+1)} \in \mathcal{N}(\Phi)$, we estimate

$$\begin{aligned}\|\eta^{(n+1)}\|_{S_2}^{2p} &= \big(\|\eta_{2r}^{(n+1)}\|_{S_2}^2 + \|\eta^{(n+1)} - \eta_{2r}^{(n+1)}\|_{S_2}^2\big)^p \leqslant \Big(\frac{\gamma_{2r}^{2/p} + \gamma_{2r}^{2/p-1}}{(2r)^{2/p-1}}\|\eta^{(n+1)} - \eta_{2r}^{(n+1)}\|_{S_p}^2\Big)^p \\ &\leqslant \frac{\gamma_{2r}^{2-p}(\gamma_{2r}+1)^p}{(2r)^{2-p}}\|\eta^{(n+1)} - \eta_{2r}^{(n+1)}\|_{S_p}^{2p} \leqslant 2^p\frac{\gamma_{2r}^{2-p}}{2^{2-p}r^{2-p}}\|\eta^{(n+1)} - \eta_{2r}^{(n+1)}\|_{S_p}^{2p},\end{aligned}$$

where we use in the second inequality a version of Stechkin's lemma (Kabanava et al., 2016, Lemma 3.1), which leads to the estimate

$$\|\eta^{(n+1)} - \eta_{2r}^{(n+1)}\|_{S_2}^2 \leqslant \frac{\|\eta_{2r}^{(n+1)}\|_{S_2}^{2-p}}{(2r)^{2-p}} \|\eta^{(n+1)} - \eta_{2r}^{(n+1)}\|_{S_p}^p \leqslant \frac{\gamma_{2r}^{2/p-1}}{(2r)^{2/p-1}} \|\eta^{(n+1)} - \eta_{2r}^{(n+1)}\|_{S_p}^2.$$

Combining the estimate for $\|\eta^{(n+1)}\|_{S_2}^{2p}$ with statement (i), this results in

$$\|\eta^{(n+1)}\|_{S_2}^{2p} \leqslant 2^p \frac{\gamma_{2r}^{2-p}}{r^{2-p}} \Bigg(\sum_{i=r+1}^{d} \big(\sigma_i^2(X^{(n)}) + \epsilon^{(n)2}\big)^{\frac{p}{2}} \Bigg)^{2-p} \|\eta_{\mathrm{vec}}^{(n+1)}\|_{\ell_2(\widetilde{W}^{(n)})}^{2p},$$

which shows statement (ii).

(iii) For the third statement, we use the strong Schatten-p NSP (18) to see that

$$\|\eta^{(n+1)}\|_{S_p}^p = \|\eta_{2r}^{(n+1)}\|_{S_p}^p + \|\eta^{(n+1)} - \eta_{2r}^{(n+1)}\|_{S_p}^p \leqslant (1+\gamma_{2r}) \|\eta^{(n+1)} - \eta_{2r}^{(n+1)}\|_{S_p}^p,$$

and combine this with statement (i). ■

Lemma 21 *Let $(X^{(n)})_n$ be the output sequence of Algorithm 1 with parameters Φ, Y, r and $0 < p \leqslant 1$, and $\widetilde{W}^{(n)}$ be the harmonic mean weight matrix matrix (10) for $n \in \mathbb{N}$. Let $X_0 \in M_{d_1 \times d_2}$ be a rank-r matrix such that $\Phi(X_0) = Y$ with condition number $\kappa := \frac{\sigma_1(X_0)}{\sigma_r(X_0)}$.*

(i) If (24) is fulfilled for iteration n, then $\eta^{(n+1)} = X^{(n)} - X_0$ fulfills

$$\Big\|\eta_{\mathrm{vec}}^{(n+1)}\Big\|_{\ell_2(\widetilde{W}^{(n)})}^{2p} \leqslant \frac{4^p r^{p/2} \sigma_r(X_0)^{p(p-1)}}{(1-\zeta)^{2p}} \kappa^p \frac{\|\eta^{(n)}\|_{S_\infty}^{2p-p^2}}{(\epsilon^{(n)})^{2p-p^2}} \|\eta^{(n+1)}\|_{S_2}^p.$$

(ii) Under the same assumption as for (i), it holds that

$$\Big\|\eta_{\mathrm{vec}}^{(n+1)}\Big\|_{\ell_2(\widetilde{W}^{(n)})}^{2p} \leqslant \frac{7^p r^{p/2} \max(r, d-r)^{p/2} \sigma_r(X_0)^{p(p-1)}}{(1-\zeta)^{2p}} \kappa^p \frac{\|\eta^{(n)}\|_{S_\infty}^{2p-p^2}}{(\epsilon^{(n)})^{2p-p^2}} \|\eta^{(n+1)}\|_{S_\infty}^p.$$

Proof (i) Recall that $X^{(n+1)} = \underset{\Phi(X)=Y}{\arg\min} \|X_{\mathrm{vec}}\|_{\ell_2(\widetilde{W}^{(n)})}^2$ is the minimizer of the weighted least squares problem with weight matrix $\widetilde{W}^{(n)}$. As $\eta^{(n+1)} = X^{(n+1)} - X_0$ is in the null space of the measurement map Φ, it follows from Lemma 16 that

$$0 = \langle \widetilde{W}^{(n)} X_{\mathrm{vec}}^{(n+1)}, \eta_{\mathrm{vec}}^{(n+1)} \rangle = \langle \widetilde{W}^{(n)} (\eta^{(n+1)} + X_0)_{\mathrm{vec}}, \eta_{\mathrm{vec}}^{(n+1)} \rangle,$$

which is equivalent to

$$\Big\|\eta_{\mathrm{vec}}^{(n+1)}\Big\|_{\ell_2(\widetilde{W}^{(n)})}^2 = \langle \widetilde{W}^{(n)} \eta_{\mathrm{vec}}^{(n+1)}, \eta_{\mathrm{vec}}^{(n+1)} \rangle = -\langle \widetilde{W}^{(n)} (X_0)_{\mathrm{vec}}, \eta_{\mathrm{vec}}^{(n+1)} \rangle.$$

Using Hölder's inequality, we can therefore estimate

$$\begin{aligned}\left\|\eta_{\mathrm{vec}}^{(n+1)}\right\|^2_{\ell_2(\widetilde{W}^{(n)})} &= -\langle \widetilde{W}^{(n)}(X_0)_{\mathrm{vec}}, \eta_{\mathrm{vec}}^{(n+1)}\rangle_{\ell_2} = -\langle [\widetilde{W}^{(n)}(X_0)_{\mathrm{vec}}]_{\mathrm{mat}}, \eta^{(n+1)}\rangle_F \\ &\leqslant \left\|\left[\widetilde{W}^{(n)}(X_0)_{\mathrm{vec}}\right]_{\mathrm{mat}}\right\|_{S_2} \|\eta^{(n+1)}\|_{S_2}.\end{aligned} \tag{43}$$

To bound the first factor, we first rewrite the action of $\widetilde{W}^{(n)}$ on X_0 in the matrix space as

$$\begin{aligned}\left[\widetilde{W}^{(n)}(X_0)_{\mathrm{vec}}\right]_{\mathrm{mat}} &= 2[(\widetilde{V}^{(n)} \otimes \widetilde{U}^{(n)})\big(\bar{\Sigma}_{d_1}^{(n)2-p} \oplus \bar{\Sigma}_{d_2}^{(n)2-p}\big)^{-1}(\widetilde{V}^{(n)} \otimes \widetilde{U}^{(n)})^*(X_0)_{\mathrm{vec}}]_{\mathrm{mat}} = \\ &= \widetilde{U}^{(n)}\big(H^{(n)} \circ (\widetilde{U}^{(n)*} X_0 \widetilde{V}^{(n)})\big)\widetilde{V}^{(n)*},\end{aligned}$$

using (39) and Lemma 20 about the action of inverses of Kronecker sums, with the notation that $H^{(n)} \in M_{d_1 \times d_2}$ such that

$$H_{ij}^{(n)} = 2\left[\mathbb{1}_{\{i \leqslant d\}}(\sigma_i^2(X^{(n)}) + \epsilon^{(n)2})^{\frac{2-p}{2}} + \mathbb{1}_{\{j \leqslant d\}}(\sigma_j^2(X^{(n)}) + \epsilon^{(n)2})^{\frac{2-p}{2}}\right]^{-1}$$

for $i \in [d_1]$, $j \in [d_2]$, where $\mathbb{1}_{\{i \leqslant d\}} = 1$ if $i \leqslant d$ and $\mathbb{1}_{\{i \leqslant d\}} = 0$ otherwise. This enables us to estimate

$$\left\|\left[\widetilde{W}^{(n)}(X_0)_{\mathrm{vec}}\right]_{\mathrm{mat}}\right\|^2_{S_2} = \left\|\widetilde{U}^{(n)}\big(H^{(n)} \circ (\widetilde{U}^{(n)*} X_0 \widetilde{V}^{(n)})\big)\widetilde{V}^{(n)*}\right\|^2_{S_2} = \left\|H^{(n)} \circ (\widetilde{U}^{(n)*} X_0 \widetilde{V}^{(n)})\right\|^2_{S_2}$$

$$\begin{aligned}&= \left\|H^{(n)} \circ \begin{pmatrix} U_T^{(n)*} X_0 V_T^{(n)} & U_T^{(n)*} X_0 V_{T_c}^{(n)} \\ U_{T_c}^{(n)*} X_0 V_T^{(n)} & U_{T_c}^{(n)*} X_0 V_{T_c}^{(n)} \end{pmatrix}\right\|^2_{S_2} \\ &= \left\|H_{T,T}^{(n)} \circ (U_T^{(n)*} X_0 V_T^{(n)})\right\|^2_{S_2} + \left\|H_{T,T_c}^{(n)} \circ (U_T^{(n)*} X_0 V_{T_c}^{(n)})\right\|^2_{S_2} \\ &+ \left\|H_{T_c,T}^{(n)} \circ (U_{T_c}^{(n)*} X_0 V_T^{(n)})\right\|^2_{S_2} + \left\|H_{T_c,T_c}^{(n)} \circ (U_{T_c}^{(n)*} X_0 V_{T_c}^{(n)})\right\|^2_{S_2},\end{aligned} \tag{44}$$

using the notation from the proof of Lemma 20. To bound the first summand, we calculate

$$\begin{aligned}&\left\|H_{T,T}^{(n)} \circ (U_T^{(n)*} X_0 V_T^{(n)})\right\|_{S_2} \leqslant \left\|H_{T,T}^{(n)} \circ (U_T^{(n)*} X^{(n)} V_T^{(n)})\right\|_{S_2} + \left\|H_{T,T}^{(n)} \circ (-U_T^{(n)*} \eta^{(n)} V_T^{(n)})\right\|_{S_2} \\ &\leqslant \left\|H_{T,T}^{(n)} \circ \Sigma_T^{(n)}\right\|_{S_2} + \left\|H_{T,T}^{(n)} \circ (U_T^{(n)*} \eta^{(n)} V_T^{(n)})\right\|_{S_2} \\ &\leqslant \left(\sum_{i=1}^r \frac{\sigma_i^2(X^{(n)})}{\big(\sigma_i^2(X^{(n)}) + \epsilon^{(n)2}\big)^{2-p}}\right)^{1/2} + \max_{i,j=1}^r |H_{i,j}^{(n)}| \|U_T^{(n)*} \eta^{(n)} V_T^{(n)}\|_{S_2} \\ &\leqslant \sqrt{r}\sigma_r^{p-1}(X^{(n)}) + (\sigma_r^2(X^{(n)}) + \epsilon^{(n)2})^{\frac{p-2}{2}} \|U_T^{(n)*} \eta^{(n)} V_T^{(n)}\|_{S_2} \\ &\leqslant \sqrt{r}\sigma_r^{p-1}(X^{(n)}) + \sigma_r^{p-2}(X^{(n)})\sqrt{r}\|\eta^{(n)}\|_{S_\infty} = \sqrt{r}\sigma_r^{p-2}(X^{(n)})\big[\sigma_r(X^{(n)}) + \|\eta^{(n)}\|_{S_\infty}\big],\end{aligned}$$

denoting $\Sigma_T^{(n)} = \mathrm{diag}(\sigma_i(X^{(n)}))_{i=1}^r$ and that the matrices $U_T^{(n)}$ and $V_T^{(n)}$ contain the first r left resp. right singular vectors of $X^{(n)}$ in the second inequality, together with the estimates $\|X\|_{S_1} \leqslant \sqrt{r}\|X\|_{S_2} \leqslant r\|X\|_{S_\infty}$ for $(r \times r)$-matrices X.

With the notations $s_r^0 := \sigma_r(X_0)$ and $s_1^0 := \sigma_1(X_0)$, we note that

$$\sigma_r(X^{(n)}) \geqslant s_r^0(1-\zeta),$$

as the assumption (24) implies that

$$s_r^0 = \sigma_r(X_0) = \sigma_r(X^{(n)} - \eta^{(n)}) \leqslant \sigma_r(X^{(n)}) + \sigma_1(\eta^{(n)}) \leqslant \sigma_r(X^{(n)}) + \zeta s_r^0,$$

using Bernstein (2009, Proposition 9.6.8) in the first inequality.

Therefore, we can bound the first summand of (44) such that

$$\left\| H_{T,T}^{(n)} \circ (U_T^{(n)*} X_0 V_T^{(n)}) \right\|_{S_2} \leqslant \sqrt{r}(s_r^0(1-\zeta))^{p-2}[s_r^0(1-\zeta)+\zeta s_r^0] = \sqrt{r}(s_r^0)^{p-1}(1-\zeta)^{p-2}. \quad (45)$$

For the second summand in the estimate of $\left\| \left[\widetilde{W}^{(n)}(X_0)_{\mathrm{vec}}\right]_{\mathrm{mat}} \right\|_{S_2}^2$, similar arguments and again assumption (24) are used to compute

$$\begin{aligned}
&\left\| H_{T,T_c}^{(n)} \circ (U_T^{(n)*} X_0 V_{T_c}^{(n)}) \right\|_{S_2} \leqslant \left\| H_{T,T_c}^{(n)} \circ \overbrace{(U_T^{(n)*} X^{(n)} V_{T_c}^{(n)})}^{=0} \right\|_{S_2} + \\
&\left\| H_{T,T_c}^{(n)} \circ (U_T^{(n)*} \eta^{(n)} V_{T_c}^{(n)}) \right\|_{S_2} \leqslant \max_{\substack{i\in[r] \\ j\in\{r+1,\dots,d_2\}}} |H_{i,j}^{(n)}| \|U_T^{(n)*} \eta^{(n)} V_{T_c}^{(n)}\|_{S_2} \\
&\leqslant \frac{2\|U_T^{(n)*} \eta^{(n)} V_{T_c}^{(n)}\|_F}{\left[(\sigma_r(X^{(n)})^2 + \epsilon^{(n)2})^{\frac{2-p}{2}}\right]} \leqslant 2\sigma_r(X^{(n)})^{p-2} \|U_T^{(n)*} \eta^{(n)} V_{T_c}^{(n)}\|_{S_2} \\
&\leqslant 2\sqrt{r}(s_r^0(1-\zeta))^{p-2} \|\eta^{(n)}\|_{S_\infty} \leqslant 2\zeta\sqrt{r}(s_r^0)^{p-1}(1-\zeta)^{p-2}.
\end{aligned} \quad (46)$$

From exactly the same arguments it follows that also

$$\left\| H_{T_c,T}^{(n)} \circ (U_{T_c}^{(n)*} X_0 V_T^{(n)}) \right\|_{S_2} \leqslant 2\zeta\sqrt{r}(s_r^0)^{p-1}(1-\zeta)^{p-2}. \quad (47)$$

It remains to bound the last summand $\left\| H_{T_c,T_c}^{(n)} \circ (U_{T_c}^{(n)*} X_0 V_{T_c}^{(n)}) \right\|_{S_2}^2$. We see that

$$\begin{aligned}
&\left\| H_{T_c,T_c}^{(n)} \circ (U_{T_c}^{(n)*} X_0 V_{T_c}^{(n)}) \right\|_{S_2} \leqslant \max_{\substack{i\in\{r+1,\dots,d_1\} \\ j\in\{r+1,\dots,d_2\}}} |H_{i,j}^{(n)}| \|U_{T_c}^{(n)*} X_0 V_{T_c}^{(n)}\|_{S_2} \\
&\leqslant (\epsilon^{(n)})^{p-2} \|U_{T_c}^{(n)*} X_0 V_{T_c}^{(n)}\|_{S_2} \leqslant (\epsilon^{(n)})^{p-2} \|U_{T_c}^{(n)*} U_T^0\|_{S_\infty} \|S^0\|_{S_2} \|V_T^{0*} V_{T_c}^{(n)}\|_{S_\infty} \\
&\leqslant (\epsilon^{(n)})^{p-2} \frac{\sqrt{2}\|\eta^{(n)}\|_{S_\infty}}{(1-\zeta)s_r^0} \sqrt{r} s_1^0 \frac{\sqrt{2}\|\eta^{(n)}\|_{S_\infty}}{(1-\zeta)s_r^0} = 2\sqrt{r}\|\eta^{(n)}\|_{S_\infty}^2 (\epsilon^{(n)})^{p-2}(1-\zeta)^{-2}(s_r^0)^{-1}\frac{s_1^0}{s_r^0},
\end{aligned} \quad (48)$$

where Hölder's inequality for Schatten norms was used in the third inequality. In the fourth inequality, Wedin's singular value perturbation bound of Lemma 19 is used with the choice $Z = X_0$, $\bar{Z} = X^{(n)}$, $\alpha = s_r^0$ and $\delta = (1-\zeta)s_r^0$, and finally $\epsilon^{(n)} \leqslant \zeta s_r^0$ in the last inequality, which is implied by the rule (14) for $\epsilon^{(n)}$ together with assumption (24).

Summarizing the estimates (45)–(48), we conclude that

$$\begin{aligned}\Big\|\big[\widetilde{W}^{(n)}(X_0)_{\mathrm{vec}}\big]_{\mathrm{mat}}\Big\|_{S_2}^2 &\leqslant \frac{r(s_r^0)^{2p-2}}{(1-\zeta)^{4-2p}}\bigg[1+8\zeta^2+4\frac{\|\eta^{(n)}\|_{S_\infty}^4}{(1-\zeta)^{2p}}(\epsilon^{(n)})^{2p-4}(s_r^0)^{-2p}\left(\frac{s_1^0}{s_r^0}\right)^2\bigg]\\ &= \frac{r(s_r^0)^{2p-2}}{(1-\zeta)^4}\bigg[(1+8\zeta^2)(1-\zeta)^{2p}+4\frac{\|\eta^{(n)}\|_{S_\infty}^{4-2p}}{(\epsilon^{(n)})^{4-2p}}\frac{\|\eta^{(n)}\|_{S_\infty}^{2p}}{(s_r^0)^{2p}}\left(\frac{s_1^0}{s_r^0}\right)^2\bigg]\\ &\leqslant \frac{r(s_r^0)^{2p-2}}{(1-\zeta)^4}\bigg[9+4\frac{\|\eta^{(n)}\|_{S_\infty}^{4-2p}}{(\epsilon^{(n)})^{4-2p}}\zeta^{2p}\kappa^2\bigg] \leqslant \frac{13r(s_r^0)^{2p-2}}{(1-\zeta)^4}\bigg[\frac{\|\eta^{(n)}\|_{S_\infty}^{4-2p}}{(\epsilon^{(n)})^{4-2p}}\kappa^2\bigg],\end{aligned}$$

as $0<\zeta<1$, $\epsilon^{(n)} \leqslant \sigma_{r+1}(X^{(n)}) = \|X_{T_c}^{(n)}\|_{S_\infty} \leqslant \|\eta^{(n)}\|_{S_\infty}$ and using the assumption (24) in the second inequality. This concludes the proof of Lemma 21(i) together with inequality (43) as $13^{p/2} \leqslant 16^{p/2} = 4^p$.

(ii) For the second statement of Lemma 21, we proceed similarly as before, but note that by Hölder's inequality, also

$$\Big\|\eta_{\mathrm{vec}}^{(n+1)}\Big\|_{\ell_2(\widetilde{W}^{(n)})}^2 \leqslant \big\|\big[\widetilde{W}^{(n)}(X_0)_{\mathrm{vec}}\big]_{\mathrm{mat}}\big\|_{S_1}\|\eta^{(n+1)}\|_{S_\infty}, \tag{49}$$

cf. (43). Furthermore

$$\begin{aligned}\big\|\big[\widetilde{W}^{(n)}(X_0)_{\mathrm{vec}}\big]_{\mathrm{mat}}\big\|_{S_1} \leqslant{}& \Big\|H_{T,T}^{(n)}\circ(U_T^{(n)*}X_0V_T^{(n)})\Big\|_{S_1} + \Big\|H_{T,T_c}^{(n)}\circ(U_T^{(n)*}X_0V_{T_c}^{(n)})\Big\|_{S_1}\\ &+ \Big\|H_{T_c,T}^{(n)}\circ(U_{T_c}^{(n)*}X_0V_T^{(n)})\Big\|_{S_1} + \Big\|H_{T_c,T_c}^{(n)}\circ(U_{T_c}^{(n)*}X_0V_{T_c}^{(n)})\Big\|_{S_1}.\end{aligned} \tag{50}$$

The four Schatten-1 norms can then be estimated by $\max(r,(d-r))^{1/2}$ times the corresponding Schatten-2 norms. Using then again inequalities (45)–(48), we conclude the proof of (ii). ∎

We proceed now to the proof of Theorem 11.

Proof First we note that

$$\left(\sum_{i=r+1}^{d}\big(\sigma_i^2(X^{(n)})+\epsilon^{(n)2}\big)^{\frac{p}{2}}\right)^{2-p} \leqslant 2^{p-\frac{p^2}{2}}(d-r)^{2-p}\sigma_{r+1}(X^{(n)})^{p(2-p)} \tag{51}$$

as $\epsilon^{(n)} \leqslant \sigma_{r+1}(X^{(n+1)})$ due to the choice of $\epsilon^{(n)}$ in (14). We proceed by induction over $n \geqslant \bar{n}$. Theorem 20(ii) and Theorem 21(ii) imply together with (51) that for $n=\bar{n}$,

$$\begin{aligned}\|\eta^{(n+1)}\|_{S_\infty}^p \leqslant \frac{\|\eta^{(n+1)}\|_{S_2}^{2p}}{\|\eta^{(n+1)}\|_{S_\infty}^p} &\leqslant 2^p\gamma_{2r}^{2-p}2^{p-\frac{p^2}{2}}\Big(\frac{d-r}{r}\Big)^{2-p/2}\frac{7^p r^p(s_r^0)^{p(p-1)}}{(1-\zeta)^{2p}}\kappa^p\|\eta^{(n)}\|_{S_\infty}^{2p-p^2}\\ &\leqslant 2^{5p}\gamma_{2r}^{2-p}\Big(\frac{d-r}{r}\Big)^{2-p/2}\frac{r^p(s_r^0)^{p(p-1)}}{(1-\zeta)^{2p}}\kappa^p\|\eta^{(n)}\|_{S_\infty}^{p(2-p)}\end{aligned} \tag{52}$$

as $\sigma_{r+1}(X^{(n)}) = \epsilon^{(n)}$ by assumption for $n=\bar{n}$.

Similarly, by Lemma 20(iii), Lemma 21(ii) and (51), the error in the Schatten-p quasi-norm fulfills

$$\|\eta^{(n+1)}\|_{S_p}^{2p} \leqslant (1+\gamma_{2r})^2 2^{2+2p}\big(d-r\big)^{2-p}\frac{r^{p/2}(s_r^0)^{p(p-1)}}{(1-\zeta)^{2p}}\kappa^p\|\eta^{(n)}\|_{S_\infty}^{p(2-p)}\|\eta^{(n+1)}\|_{S_2}^p \tag{53}$$

for $n=\bar{n}$. Using the strong Schatten-p null space property of order $2r$ for the operator Φ, we see with the arguments of the proof of Lemma 20(ii) that

$$\|\eta^{(n)}\|_{S_\infty}^p \leqslant \|\eta^{(n)}\|_{S_2}^p \leqslant \frac{2^{p-1}\gamma_{2r}^{1-p/2}}{r^{1-p/2}}\|\eta^{(n)}\|_{S_p}^p$$

and also $\|\eta^{(n+1)}\|_{S_2}^p \leqslant \frac{2^{p-1}\gamma_{2r}^{1-p/2}}{r^{1-p/2}}\|\eta^{(n+1)}\|_{S_p}^p$. Inserting this in (53) and dividing by $\|\eta^{(n+1)}\|_{S_p}^p$, we obtain

$$\|\eta^{(n+1)}\|_{S_p}^p \leqslant 2^{4p}(1+\gamma_{2r})^2\gamma_{2r}^{2-p}\Big(\frac{d-r}{r}\Big)^{2-p}\frac{r^{p/2}(s_r^0)^{p(p-1)}}{(1-\zeta)^{2p}}\kappa^p\|\eta^{(n)}\|_{S_\infty}^{p(1-p)}\|\eta^{(n)}\|_{S_p}^p.$$

Under the assumption that (25) holds, it follows from this and (52) that

$$\|\eta^{(n+1)}\|_{S_\infty}^p \leqslant \|\eta^{(n)}\|_{S_\infty}^p \text{ and } \|\eta^{(n+1)}\|_{S_p}^p \leqslant \|\eta^{(n)}\|_{S_p}^p \tag{54}$$

for $n=\bar{n}$, which also entails the statement of Theorem 11 for this iteration.

Let now $n' > \bar{n}$ such that (54) is true for all n with $n' > n \geqslant \bar{n}$.

If $\sigma_{r+1}(X^{(n')}) \leqslant \epsilon^{(n'-1)}$, then $\epsilon^{(n')} = \sigma_{r+1}(X^{(n')})$ and the arguments from above show (54) also in the case $n = n'$.

Otherwise $\sigma_{r+1}(X^{(n')}) > \epsilon^{(n'-1)}$ and there exists $n' > n'' \geqslant \bar{n}$ such that $\epsilon^{(n')} = \epsilon^{(n'')} = \sigma_{r+1}(X^{(n'')})$. Then

$$\|\eta^{(n'+1)}\|_{S_\infty}^p \leqslant 14^p\frac{\gamma_{2r}^{2-p}}{r^{2-p}}\bigg[\sum_{i=r+1}^d\Big(\frac{\sigma_i^2(X^{(n')})}{\epsilon^{(n'')2}}+1\Big)^{\frac{p}{2}}\bigg]^{2-p}\frac{r^{p/2}\max(r,d-r)^{p/2}}{(s_r^0)^{p(1-p)}(1-\zeta)^{2p}}\kappa^p\|\eta^{(n')}\|_{S_\infty}^{p(2-p)}$$

and we compute

$$\begin{aligned}
&\bigg[\sum_{i=r+1}^d\Big(\frac{\sigma_i^2(X^{(n')})}{\epsilon^{(n'')2}}+1\Big)^{\frac{p}{2}}\bigg]^{2-p} \leqslant \bigg[\sum_{i=r+1}^d\frac{\sigma_i^p(X^{(n')})}{\epsilon^{(n'')p}}+(d-r)\bigg]^{2-p}\\
&\leqslant \bigg[\frac{\|\eta^{(n')}\|_{S_p}^p}{\epsilon^{(n'')p}}+(d-r)\bigg]^{2-p} \leqslant \bigg[\frac{\|\eta^{(n'')}\|_{S_p}^p}{\epsilon^{(n'')p}}+(d-r)\bigg]^{2-p}\\
&\leqslant \bigg[\frac{2(1+\gamma_{2r})\|X_{T_c}^{(n'')}\|_{S_p}^p}{(1-\gamma_{2r})\epsilon^{(n'')p}}+(d-r)\bigg]^{2-p} \leqslant \bigg(\frac{3+\gamma_{2r}}{1-\gamma_{2r}}\bigg)^{2-p}(d-r)^{2-p},
\end{aligned}$$

using that X_0 is a matrix of rank at most r in the second inequality, the inductive hypothesis in the third inequality and an analogue of (61) for a Schatten-p quasi-norm on the left hand side (cf. Kabanava et al. (2016, Lemma 3.2) for the corresponding result for $p=1$) in the last inequality. The latter argument uses the assumption on the null space property. This shows that

$$\|\eta^{(n'+1)}\|_{S_\infty}^p \leqslant \mu\|\eta^{(n')}\|_{S_\infty}^{p(2-p)}$$

for

$$\widetilde{\mu} := 2^{4p}\gamma_{2r}^{2-p}\Big(\frac{(3+\gamma_{2r})(d-r)}{(1-\gamma_{2r})r}\Big)^{2-p}\frac{r^{p/2}(s_r^0)^{p(p-1)}}{(1-\zeta)^{2p}}\kappa^p \max\Big(2^p(d-r)^{\frac{p}{2}}, (1+\gamma_{2r})^2\Big),$$

and $\|\eta^{(n'+1)}\|_{S_\infty}^p \leqslant \|\eta^{(n')}\|_{S_\infty}^p$ under the assumption (25) of Theorem 11, as $\widetilde{\mu} \leqslant \mu$ with μ as in (26). Indeed, it holds that $\widetilde{\mu} \leqslant \mu$ since

$$\max\Big(2^p(d-r)^{\frac{p}{2}}, (1+\gamma_{2r})^2\Big)\Big(\frac{d-r}{r}\Big)^{2-p} r^{p/2} \leqslant 2^p(1+\gamma_{2r})^2\Big(\frac{d-r}{r}\Big)^{2-p/2} r^p.$$

The same argument shows that $\|\eta^{(n'+1)}\|_{S_p}^p \leqslant \|\eta^{(n')}\|_{S_p}^p$, which finishes the proof. ■

Remark 22 *We note that the weight matrices of the previous IRLS approaches* `IRLS-col` *and* `IRLS-row` *Fornasier et al. (2011); Mohan and Fazel (2012) at iteration n could be expressed in our notation as*

$$\mathbf{I}_{d_2} \otimes W_L^{(n)} := \mathbf{I}_{d_2} \otimes U^{(n)}(\bar{\Sigma}_{d_1}^{(n)})^{p-2}U^{(n)*}$$

and

$$W_R^{(n)} \otimes \mathbf{I}_{d_1} := V^{(n)}(\bar{\Sigma}_{d_2}^{(n)})^{p-2}V^{(n)*} \otimes \mathbf{I}_{d_1},$$

respectively, cf. Section 2.2, if $X^{(n)} = U^{(n)}\Sigma^{(n)}V^{(n)*} = U_T^{(n)}\Sigma_T^{(n)}V_T^{(n)*} + U_{T_c}^{(n)}\Sigma_{T_c}^{(n)}V_{T_c}^{(n)*}$ is the SVD of the iterate $X^{(n)}$ with $U_T^{(n)}$ and $V_T^{(n)}$ containing the r first left- and right singular vectors.

Now let

$$T^{(n)} := \big\{U_T^{(n)}Z_1^* + Z_2V_T^{(n)*} : Z_1 \in M_{d_1\times r}, Z_2 \in M_{d_2\times r}\big\}$$

be the tangent space of the smooth manifold of rank-r matrices at the best rank-r approximation $U_T^{(n)}\Sigma_T^{(n)}V_T^{(n)*}$ of $X^{(n)}$, or, put differently, the direct sum of the row and column spaces of $U_T^{(n)}\Sigma_T^{(n)}V_T^{(n)*}$.

The fact that left- or right-sided weight matrices do not lead to algorithms with superlinear convergence rates for $p < 1$ can be explained by noting that there are always parts of the space $T^{(n)}$ that are equipped with too large weights if $X^{(n)} = U^{(n)}\Sigma^{(n)}V^{(n)*}$ is already approximately low-rank. In particular, proceeding as in (44), we obtain for $\mathbf{I}_{d_2} \otimes W_L^{(n)}$

$$\begin{aligned}\Big\|\big[\mathbf{I}_{d_2} \otimes W_L^{(n)}(X_0)_{\mathrm{vec}}\big]_{\mathrm{mat}}\Big\|_{S_2}^2 =& \Big\|\big(\bar{\Sigma}_T^{(n)}\big)^{p-2}U_T^{(n)*}X_0V_T^{(n)}\Big\|_{S_2}^2 + \Big\|\big(\bar{\Sigma}_T^{(n)}\big)^{p-2}U_T^{(n)*}X_0V_{T_c}^{(n)}\Big\|_{S_2}^2 \\ &+\Big\|\big(\bar{\Sigma}_{T_c}{}^{(n)}\big)^{p-2}U_{T_c}^{(n)*}X_0V_T^{(n)}\Big\|_{S_2}^2 + \Big\|\big(\bar{\Sigma}_{T_c}^{(n)}\big)^{p-2}U_{T_c}^{(n)*}X_0V_{T_c}^{(n)}\Big\|_{S_2}^2\end{aligned}$$

if $\bar{\Sigma}_T^{(n)}$ denotes the diagonal matrix with the first r non-zero entries of $\bar{\Sigma}_{d_1}^{(n)}$ and $\bar{\Sigma}_{T_c}^{(n)}$ the one of the remaining entries.

Here, the third of the four summands would become too large for $p < 1$ to allow for a superlinear convergence when the last $d-r$ singular values of $X^{(n)}$ approach zero. An analogous argument can be used for the right-sided weight matrix $W_R^{(n)} \otimes \mathbf{I}_{d_1}$ and, notably, also for arithmetic mean weight matrices $W_{(arith)}^{(n)} = \mathbf{I}_{d_2} \otimes W_L^{(n)} + W_R^{(n)} \otimes \mathbf{I}_{d_1}$, cf. Section 2.3.

Acknowledgments

The two authors acknowledge the support and hospitality of the Hausdorff Research Institute for Mathematics (HIM) during the early stage of this work within the HIM Trimester Program "Mathematics of Signal Processing". C.K. is supported by the German Research Foundation (DFG) in the context of the Emmy Noether Junior Research Group "Randomized Sensing and Quantization of Signals and Images" (KR 4512/1-1) and the ERC Starting Grant "High-Dimensional Sparse Optimal Control" (HDSPCONTR - 306274). J.S. is supported by the DFG through the D-A-CH project no. I1669-N26 and through the international research training group IGDK 1754 "Optimization and Numerical Analysis for Partial Differential Equations with Nonsmooth Structures". The authors thank Ke Wei for providing code of his implementations. They also thank Massimo Fornasier for helpful discussions.

Appendix A. Kronecker and Hadamard products

For two matrices $A = (a_{ij})_{i\in[d_1],j\in[d_3]} \in \mathbb{C}^{d_1\times d_3}$ and $B \in \mathbb{C}^{d_2\times d_4}$, we call the matrix representation of their tensor product with respect to the standard bases the *Kronecker product* $A\otimes B \in \mathbb{C}^{d_1\cdot d_2\times d_3\cdot d_4}$. By its definition, $A\otimes B$ is a block matrix of $d_2 \times d_4$ blocks whose block of index $(i, j) \in [d_1] \times [d_3]$ is the matrix $a_{ij}B \in \mathbb{R}^{d_2\times d_4}$. This implies, e.g., for $A \in \mathbb{C}^{d_1\times d_3}$ with $d_1 = 2$ and $d_3 = 3$ that

$$A \otimes B = \left[\begin{array}{ccc} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{array}\right] \otimes B = \left[\begin{array}{c|c|c} a_{11}B & a_{12}B & a_{13}B \\ \hline a_{21}B & a_{22}B & a_{23}B \end{array}\right].$$

The Kronecker product is useful for the elegant formulation of matrix equations involving left and right matrix multiplications with the variable X, as

$$AXB^* = Y \qquad \text{if and only if} \qquad (B \otimes A)X_{\mathrm{vec}} = Y_{\mathrm{vec}}.$$

We define the *Hadamard product* $A \circ B \in \mathbb{C}^{d_1\times d_2}$ of two matrices $A \in \mathbb{C}^{d_1\times d_2}$ and $B \in \mathbb{C}^{d_1\times d_2}$ as their entry-wise product

$$(A \circ B)_{i,j} = A_{i,j}B_{i,j}$$

with $i \in [d_1]$ and $j \in [d_2]$. The Hadamard product is also known as *Schur product* in the literature.

Furthermore, if $d_1 = d_3$ and $d_2 = d_4$, we define the *Kronecker sum* $A \oplus B \in \mathbb{C}^{d_1d_2\times d_1d_2}$ of two matrices $A \in \mathbb{C}^{d_1\times d_1}$ and $B \in \mathbb{C}^{d_2\times d_2}$ as the matrix

$$A \oplus B = (\mathbf{I}_{d_2} \otimes A) + (B \otimes \mathbf{I}_{d_1}). \tag{55}$$

Note that equations of the form $AX + XB^* = Y$ can be rewritten as

$$(A \oplus B)X_{\mathrm{vec}} = Y_{\mathrm{vec}},$$

using again the vectorizations of X and Y. An explicit formula that expresses the inverse $(A \oplus B)^{-1}$ of the Kronecker sum $A \oplus B$ is provided by the following lemma.

Lemma 23 (Jameson (1968)) *Let $A \in H_{d_1 \times d_1}$ and $B \in H_{d_2 \times d_2}$, where one of the matrices is positive definite and the other positive semidefinite. If we denote the singular vectors of A by $u_i \in \mathbb{C}^{d_1}$, $i \in [d_1]$, its singular values by σ_i, $i \in [d_1]$ and the singular vectors resp. values of B by $v_j \in \mathbb{C}^{d_2}$ resp. μ_j, $j \in [d_2]$, then*

$$(A \oplus B)^{-1} = \sum_{i=1}^{d_1} \sum_{j=1}^{d_2} \frac{v_j v_j^* \otimes u_i u_i^*}{\sigma_i + \mu_j}. \tag{56}$$

Furthermore, the action of $(A \oplus B)^{-1}$ on the matrix space $M_{d_1 \times d_2}$ can be written as

$$\big[(A \oplus B)^{-1} Z_{\mathrm{vec}}\big]_{\mathrm{mat}} = U\big(H \circ (U^* Z V)\big)V^*. \tag{57}$$

for $Z \in M_{d_1 \times d_2}$, $U = [u_1, \ldots, u_{d_1}]$, and $V = [v_1, \ldots, v_{d_2}]$ and the matrix $H \in M_{d_1 \times d_2}$ with the entries $H_{i,j} = (\sigma_i + \mu_j)^{-1}$, $i \in [d_1]$, $j \in [d_2]$.

Appendix B. Proofs of preliminary statements in Section 6

B.1 Proof of Lemma 14: Main part

First, we define the function

$$f^p_{X,\epsilon}(Z) = \mathcal{J}_p(X, \epsilon, Z) = \begin{cases} \frac{p}{2}\|X_{\mathrm{vec}}\|^2_{\ell_2(\widetilde{W}(Z))} + \frac{\epsilon^2 p}{2} \sum_{i=1}^{d} \sigma_i(Z) + \frac{2-p}{2} \sum_{i=1}^{d} \sigma_i(Z)^{\frac{p}{(p-2)}} & \text{if } \operatorname{rank}(Z) = d, \\ +\infty & \text{if } \operatorname{rank}(Z) < d, \end{cases}$$

for $X \in M_{d_1 \times d_2}$, $\epsilon > 0$ fixed and with $Z \in M_{d_1 \times d_2}$ as its only argument. We note that the set of minimizers of $f^p_{X,\epsilon}(Z)$ does not contain an instance Z with rank smaller than d as the value of $f^p_{X,\epsilon}(Z)$ is infinite at such points and, therefore, it is sufficient to search for minimizers on the set $\Omega = \{Z \in M_{d_1 \times d_2} | \operatorname{rank}(Z) = d\}$ of matrices with rank d. We observe that the set Ω is an open set and that we have that

(a) $f^p_{X,\epsilon}(Z)$ is lower semi-continuous, which means that any sequence $(Z^k)_{k\in\mathbb{N}}$ with $Z^k \overset{k\to\infty}{\longrightarrow} Z$ fulfills $\liminf_{k\to\infty} f^p_{X,\epsilon}(Z^k) \geqslant f^p_{X,\epsilon}(Z)$,

(b) $f^p_{X,\epsilon}(Z) \geqslant \alpha$ for all $Z \in M_{d_1 \times d_2}$ for some constant α,

(c) $f^p_{X,\epsilon}(Z)$ is coercive, i.e., for any sequence $(Z^k)_{k\in\mathbb{N}}$ with $\|Z^k\|_F \overset{k\to\infty}{\longrightarrow} \infty$, we have $f^p_{X,\epsilon}(Z^k) \overset{k\to\infty}{\longrightarrow} \infty$.

Property (a) is true as $f^p_{X,\epsilon}(Z))|_\Omega$ is a concatenation of an indicator function of an open set, which is lower semi-continuous and a sum of continuous functions on Ω. Property (b) is obviously true for the choice $\alpha = 0$.

To justify point (c), we note that $f^p_{X,\epsilon}(Z) > \frac{\epsilon^2 p}{2} \sum_{i=1}^{d} \sigma_i(Z) = \frac{\epsilon^2 p}{2}\|Z\|_{S_1} \geqslant \frac{\epsilon^2 p}{2}\|Z\|_F$ and therefore, coercivity is clear from its definition. As a consequence from (a) and (c), it is

also true that the level sets $L_C = \left\{ Z \in M_{d_1 \times d_2} | f^p_{X,\epsilon}(Z) \leqslant C \right\}$ are closed and bounded and, therefore, compact.

Via the direct method of the calculus of variations, we conclude from the properties (a)–(c) that $f^p_{X,\epsilon}(Z)$ has at least one global minimizer belonging to the set of critical points of $f^p_{X,\epsilon}(Z)$ (Dacorogna, 1989, Theorem 1).
To characterize the set of critical points of $f^p_{X,\epsilon}(Z)$, its derivative with respect to Z is calculated explicitly and equated with zero in Subsection B.2. The solution of the resulting equation reveals that $Z_{\text{opt}} = \sum_{i=1}^d (\sigma_i^2(X) + \epsilon^2)^{\frac{p-2}{2}} u_i v_i^* =: \sum_{i=1}^d \widetilde{\sigma}_i u_i v_i^*$ is the only critical point and consequently the unique global minimizer of $f^p_{X,\epsilon}(Z)$. We define the matrices $W^L_{\text{opt}} := \sum_{i=1}^d \widetilde{\sigma}_i u_i u_i^*$ and $W^R_{\text{opt}} := \sum_{i=1}^d \widetilde{\sigma}_i v_i v_i^*$, and note that $\widetilde{W}(Z_{\text{opt}}) = 2\big((W^R_{\text{opt}})^{-1} \oplus (W^L_{\text{opt}})^{-1}\big)^{-1}$ with Definition 13. To verify the second part of the theorem, we simply plug the optimal solution Z_{opt} into the functional $\mathcal{J}_p$ and compute using (56) that

$$
\begin{aligned}
\mathcal{J}_p(X, \epsilon, Z_{\text{opt}}) &= \frac{p}{2} \|X_{\text{vec}}\|^2_{\ell_2(\widetilde{W}(Z_{\text{opt}}))} + \frac{\epsilon^2 p}{2} \sum_{i=1}^d \widetilde{\sigma}_i + \frac{2-p}{2} \sum_{i=1}^d \widetilde{\sigma}_i^{\frac{p}{p-2}} \\
&= \frac{p}{2} \sum_{i=1}^d \left[\sigma_i^2(X) (u_i^* \otimes v_i^*) 2 \left(\sum_{k=1}^{d_2} \sum_{j=1}^{d_1} \frac{u_k u_k^* \otimes v_j v_j^*}{\widetilde{\sigma}_k^{-1} + \widetilde{\sigma}_j^{-1}} \right) (u_i \otimes v_i) \right]_{ii} + \frac{\epsilon^2 p}{2} \sum_{i=1}^d \widetilde{\sigma}_i + \frac{2-p}{2} \sum_{i=1}^d \widetilde{\sigma}_i^{\frac{p}{p-2}} \\
&= \frac{p}{2} \sum_{i=1}^d (\sigma_i^2(X) + \epsilon^2) \widetilde{\sigma}_i + \frac{2-p}{2} \sum_{i=1}^d \widetilde{\sigma}_i^{\frac{p}{p-2}} \\
&= \frac{p}{2} \sum_{i=1}^d (\sigma_i^2(X) + \epsilon^2)(\sigma_i^2(X) + \epsilon^2)^{\frac{p-2}{2}} + \frac{2-p}{2} \sum_{i=1}^d (\sigma_i^2(X) + \epsilon^2)^{\frac{p}{2}} \\
&= \sum_{i=1}^d (\sigma_i^2(X) + \epsilon^2)^{\frac{p}{2}}.
\end{aligned}
$$

B.2 Proof of Lemma 14: Critical points of $f^p_{X,\epsilon}$

Let us without loss of generality consider the case $d = d_1 = d_2$ and define

$$\Omega = \{ Z \in M_{d \times d} \text{ s.t. } \operatorname{rank}(Z) = d \}.$$

As already mentioned in (27), the harmonic mean matrix $\widetilde{W}(Z)$ can then be rewritten as

$$\widetilde{W}(Z) = 2\widetilde{W}_1 \big(\widetilde{W}_1 + \widetilde{W}_2\big)^{-1} \widetilde{W}_2 = 2(\widetilde{W}_1^{-1} + \widetilde{W}_2^{-1})^{-1}$$

for $Z \in \Omega$ with the definitions $\widetilde{W}_1 := \mathbf{I}_d \otimes (ZZ^*)^{\frac{1}{2}}$ and $\widetilde{W}_2 = (Z^*Z)^{\frac{1}{2}} \otimes \mathbf{I}_d$. For $Z \in \Omega$, we reformulate the auxiliary functional such that

$$
\begin{aligned}
f^p_{X,\epsilon}(Z) = \mathcal{J}^p(X, \epsilon, Z) &= \frac{p}{2} \|X_{\text{vec}}\|^2_{\ell_2(\widetilde{W}(Z))} + \frac{\epsilon^2 p}{2} \sum_{i=1}^d \sigma_i(Z) + \frac{2-p}{2} \sum_{i=1}^d \sigma_i(Z)^{\frac{p}{(p-2)}} \\
&= \frac{p}{2} \|X_{\text{vec}}\|^2_{\ell_2(\widetilde{W}(Z))} + \frac{\epsilon^2 p}{2} \|(Z^*Z)^{1/2}\|_F^2 + \frac{2-p}{2} \|(Z^*Z)^{\frac{p}{2(p-2)}}\|_F^2.
\end{aligned}
$$

To identify the set of critical points of $f^p_{X,\epsilon}(Z)$ located in Ω, we compute its derivative with respect to Z using the derivative rules (7), (12), (13), (15), (16), (18), (20) in Chapter 8.2

and Theorem 3 in Chapter 8.4 of (Magnus and Neudecker, 1999) in the following. Using the notation of Magnus and Neudecker (1999), we calculate

$$\begin{aligned}\partial f^p_{X,\epsilon}(Z) = -\frac{p}{2}\operatorname{tr}\left(X^*_{\mathrm{vec}}\widetilde{W}\partial\widetilde{W}^{-1}\widetilde{W}X_{\mathrm{vec}}\right) + \frac{p\epsilon^2}{4}\left(\operatorname{tr}\left(Z(Z^*Z)^{-\frac{1}{2}}\partial Z^*\right) + \operatorname{tr}((Z^*Z)^{-\frac{1}{2}}Z^*\partial Z)\right)\\ -\frac{p}{4}\left(\operatorname{tr}\left(Z(Z^*Z)^{\frac{4-p}{2(p-2)}}\partial Z^*\right) + \operatorname{tr}((Z^*Z)^{\frac{4-p}{2(p-2)}}Z^*\partial Z)\right)\end{aligned}$$

where

$$\begin{aligned}\partial\widetilde{W}^{-1} = \frac{1}{2}\partial\left[(ZZ^*)^{-\frac{1}{2}}\oplus(Z^*Z)^{-\frac{1}{2}}\right] = -\frac{1}{4}\left[\left((Z^*Z)^{-\frac{3}{2}}Z^*\partial Z + \partial Z^*Z(Z^*Z)^{-\frac{3}{2}}\right)\otimes\mathbf{I}_{d_1}\right]\\ -\frac{1}{4}\left[\mathbf{I}_{d_2}\otimes\left(\partial Z(ZZ^*)^{-\frac{3}{2}}Z^* + (ZZ^*)^{-\frac{3}{2}}Z\partial Z^*\right)\right].\end{aligned} \tag{58}$$

We can reformulate the first term as follows using the cyclicity of the trace,

$$\begin{aligned}-\frac{p}{2}\operatorname{tr}\left(X^*_{\mathrm{vec}}\widetilde{W}\partial\widetilde{W}^{-1}\widetilde{W}X_{\mathrm{vec}}\right) = \frac{p}{8}\Big[&\operatorname{tr}\left((\widetilde{W}X_{\mathrm{vec}})^*_{\mathrm{mat}}(\widetilde{W}X_{\mathrm{vec}})_{\mathrm{mat}}(Z^*Z)^{-\frac{3}{2}}Z^*\partial Z\right)\\ &+\operatorname{tr}\left(Z(Z^*Z)^{-\frac{3}{2}}(\widetilde{W}X_{\mathrm{vec}})^*_{\mathrm{mat}}(\widetilde{W}X_{\mathrm{vec}})_{\mathrm{mat}}\partial Z^*\right)\\ &+\operatorname{tr}\left(Z^*(ZZ^*)^{-\frac{3}{2}}(\widetilde{W}X_{\mathrm{vec}})_{\mathrm{mat}}(\widetilde{W}X_{\mathrm{vec}})^*_{\mathrm{mat}}\partial Z\right)\\ &+\operatorname{tr}\left((\widetilde{W}X_{\mathrm{vec}})_{\mathrm{mat}}(\widetilde{W}X_{\mathrm{vec}})^*_{\mathrm{mat}}(ZZ^*)^{-\frac{3}{2}}Z\partial Z^*\right)\Big].\end{aligned}$$

To determine the critical points of $f^p_{X,\epsilon}(Z)$, we summarize the calculations above, rearrange the terms and equate the derivative with zero, such that

$$\begin{aligned}\partial f^p_{X,\epsilon}(Z) &= \frac{p}{8}\operatorname{tr}\Big(\Big[(\widetilde{W}X_{\mathrm{vec}})^*_{\mathrm{mat}}(\widetilde{W}X_{\mathrm{vec}})_{\mathrm{mat}}(Z^*Z)^{-\frac{3}{2}}Z^* + Z^*(ZZ^*)^{-\frac{3}{2}}(\widetilde{W}X_{\mathrm{vec}})_{\mathrm{mat}}(\widetilde{W}X_{\mathrm{vec}})^*_{\mathrm{mat}}\\ &\qquad +2\epsilon^2(Z^*Z)^{-\frac{1}{2}}Z^* - 2(Z^*Z)^{\frac{4-p}{2(p-2)}}Z^*\Big]\partial Z\Big)\\ &\quad +\frac{p}{8}\operatorname{tr}\Big(\Big[Z(Z^*Z)^{-\frac{3}{2}}(\widetilde{W}X_{\mathrm{vec}})^*_{\mathrm{mat}}(\widetilde{W}X_{\mathrm{vec}})_{\mathrm{mat}} + (\widetilde{W}X_{\mathrm{vec}})_{\mathrm{mat}}(\widetilde{W}X_{\mathrm{vec}})^*_{\mathrm{mat}}(ZZ^*)^{-\frac{3}{2}}Z\\ &\qquad +2\epsilon^2Z(Z^*Z)^{-\frac{1}{2}} - 2Z(Z^*Z)^{\frac{4-p}{2(p-2)}}\Big]\partial Z^*\Big)\\ &=: \frac{p}{8}\operatorname{tr}(A\partial Z) + \frac{p}{8}\operatorname{tr}(A^*\partial Z^*) = \frac{p}{8}\operatorname{tr}((A\oplus A)\partial Z) = 0,\end{aligned}$$

where

$$\begin{aligned}A = \Big[&(\widetilde{W}X_{\mathrm{vec}})^*_{\mathrm{mat}}(\widetilde{W}X_{\mathrm{vec}})_{\mathrm{mat}}(Z^*Z)^{-\frac{3}{2}}Z^* + Z^*(ZZ^*)^{-\frac{3}{2}}(\widetilde{W}X_{\mathrm{vec}})_{\mathrm{mat}}(\widetilde{W}X_{\mathrm{vec}})^*_{\mathrm{mat}}\\ &+2\epsilon^2(Z^*Z)^{-\frac{1}{2}}Z^* - 2(Z^*Z)^{\frac{4-p}{2(p-2)}}Z^*\Big].\end{aligned} \tag{59}$$

and hence an easy calculation as in (Duchi) gives

$$\frac{\partial f^p_{X,\epsilon}(Z)}{\partial Z} = \frac{\frac{p}{8}\operatorname{tr}((A\oplus A)\partial Z)}{\partial Z} = \frac{p}{8}(A\oplus A) = 0.$$

Now we have to find Z such that $A\oplus A = 0$. This implies that all eigenvalues of $A\oplus A = A\otimes\mathbf{I}_d + \mathbf{I}_d\otimes A$ are equal to zero. The eigenvalues of the Kronecker sum of two

matrices A_1 and A_2 with eigenvalues λ_s and μ_t with $s,t \in [d]$ are the sum of the eigenvalues $\lambda_s + \mu_t$. As in our case $A = A_1 = A_2$ this means that all eigenvalues of A itself have to be zero. This is only possible if A is the zero matrix.

Let $Z = U\Sigma V^* \in M_{d\times d}$ with $U, V \in \mathcal{U}_d$ and $\Sigma \in M_{d\times d}$, where $\Sigma = \operatorname{diag}(\sigma)$ is a diagonal matrix with *ascending* entries. We define the matrix $H = H_{i,j} = \frac{2}{\sigma_i^{-1}+\sigma_j^{-1}}$ for $i = 1,\dots,d, j = 1,\dots,d$ corresponding to the result of reshaping the diagonal of $2(\Sigma \oplus \Sigma)$ into a $d \times d$-matrix. Using (57), we can express $(\widetilde{W}X_{\mathrm{vec}})_{\mathrm{mat}} = U\big(H \circ (U^*XV)\big)V^*$ and denote $B := H \circ (U^*XV)$.

Plugging the decomposition $Z = U\Sigma V^*$ into (59), we can therefore calculate

$$\begin{aligned}
\mathrm{A} = 0 &\Leftrightarrow (UBV^*)^*(UBV^*)(V\Sigma^2V^*)^{-3/2}(U\Sigma V^*)^* + (U\Sigma V^*)^*(U\Sigma^2U^*)^*)^{-3/2}(UBV^*)(UBV^*)^* \\
&\quad + 2\epsilon^2(V\Sigma^2V^*)^{-1/2}(U\Sigma V^*)^* - 2(V\Sigma^2V^*)^{\frac{4-p}{2(p-2)}}(U\Sigma V^*)^* = 0 \\
&\Leftrightarrow VB^*B\Sigma^{-2}U^* + V\Sigma^{-2}BB^*U^* + 2\epsilon^2V\mathbf{I}_dU^* - 2V\Sigma^{\frac{2}{p-2}}U^* = 0 \\
&\Leftrightarrow B^*B\Sigma^{-2} + \Sigma^{-2}BB^* + 2\epsilon^2\mathbf{I}_d - 2\Sigma^{\frac{2}{p-2}} = 0.
\end{aligned} \tag{60}$$

We now note that $2\epsilon^2\mathbf{I}_d - 2\Sigma^{\frac{2}{p-2}}$ is diagonal and therefore, $B^*B\Sigma^{-2} + \Sigma^{-2}BB^*$ is diagonal as well. Moreover, observe that $B^*B + \Sigma^{-2}BB^*\Sigma^2$ is again a diagonal matrix and has a symmetric first summand B^*B. As the sum or difference of symmetric matrices is again symmetric also the second summand $\Sigma^{-2}BB^*\Sigma^2$ has to be symmetric, i.e., $\Sigma^{-2}BB^*\Sigma^2 = (\Sigma^{-2}BB^*\Sigma^2)^* = \Sigma^2BB^*\Sigma^{-2}$. We conclude that it has to hold that $BB^*\Sigma^4 = \Sigma^4BB^*$ and hence Σ^4 and BB^* commute.

This is only possible if either Σ is a multiple of the identity or if BB^* is diagonal. Assuming the first case, (60) would imply that also BB^* and B^*B have to be a multiple of the identity. Therefore, this first case, where Σ is a multiple of the identity is a special case of the second possible scenario, where BB^* is diagonal. Hence, it suffices to further consider the more general second case. (Considerations for B^*B can be carried out analogously.)

Diagonality of BB^* only occurs if B is either orthonormal or diagonal. Assuming orthonormality would lead to contradictions with the equations in (60). Hence $B = H \circ (U^*XV)$ can only be diagonal.

Let now be $X = \bar{U}\bar{S}\bar{V}^*$ the singular value decomposition of X. As H has no zero entries due to the full rank of W, this implies the diagonality of $U^*\bar{U}\bar{S}\bar{V}^*V$. Consequently, U and V can only be chosen such that $P = [U^*\bar{U}]_{d\times d}$ and $P^* = [\bar{V}^*V]_{d\times d}$ for a permutation matrix $P \in \mathcal{U}_d$. The reshuffled indexing corresponding to P is denoted by $p(i) \in [d]$ for $i \in [d]$. Bearing in mind that $H_{ii} = \sigma_i$ for $i \in [d]$, we obtain

$$\begin{aligned}
&(H \circ (P\bar{S}P^*))^*(H \circ (P\bar{S}P^*))\Sigma^{-2} + \Sigma^{-2}(H \circ (P\bar{S}P^*))(H \circ (P\bar{S}P^*))^* + 2\epsilon^2\mathbf{I}_d - 2\Sigma^{\frac{2}{p-2}} = 0 \\
\Leftrightarrow\quad & 2\bar{s}_{p(i)}^2 + 2\epsilon^2 = 2\sigma_i^{\frac{2}{p-2}} \text{ for all } i \in [d] \\
\Leftrightarrow\quad & \sigma_i = (\bar{s}_{p(i)}^2 + \epsilon^2)^{\frac{p-2}{2}} \text{ for all } i \in [d].
\end{aligned}$$

As the diagonal of Σ was assumed to have ascending entries and the diagonal of $\bar{S}$ has descending entries, the permutation matrix P has to be equal to the identity matrix. From $P = \mathbf{I}_d$, it follows that $U = \bar{U}$ and $V = \bar{V}$ and hence $\Sigma = (\bar{S}^2 + \epsilon^2\mathbf{I}_d)^{\frac{p-2}{2}}$.

We summarize our calculations by stating that

$$Z_{\mathrm{opt}} = \bar{U}\Sigma\bar{V}^* = \bar{U}(\bar{S}^2 + \epsilon^2\mathbf{I}_d)^{\frac{p-2}{2}}\bar{V}^*$$

is the only critical point of $f^p_{X,\epsilon}$ on the domain Ω.

The results extend for the case $d_1 \neq d_2$, where the definition of $\widetilde{W}(Z)$ is adapted by introducing the Moore-Penrose pseudo inverse of $(ZZ^*)^{1/2}$

$$\widetilde{W}(Z) = 2\widetilde{W}_1\big(\widetilde{W}_1 + \widetilde{W}_2\big)^{-1}\widetilde{W}_2 = 2(\widetilde{W}_1^{+} + \widetilde{W}_2^{-1})^{-1}.$$

The corresponding derivative rule as pointed out in Theorem 5 in Chapter 8.4 of Magnus and Neudecker (1999) can be used for the calculation in (58).

B.3 Proof of Lemma 16

The equality of the optimization problems (32) can easily be seen by the fact that only the first summand of $\mathcal{J}_p(X, \epsilon, Z)$ depends on X. Now, it is important to show first that $\widetilde{W}(Z) = 2([(Z^*Z)^{\frac{1}{2}}]^+ \oplus [(ZZ^*)^{\frac{1}{2}}]^+)^{-1}$ is positive definite as minimizing $\mathcal{J}_p(X, \epsilon, Z)$ then reduces to minimizing a quadratic form. Let $Z = \sum_{i=1}^d \sigma_i u_i v_i^*$, where u_i, v_i for $i \in [d]$ are the left and right singular vector respectively and σ_i for $i \in [d]$ are the singular values of Z. Since $Z^*Z = \sum_{i=1}^d \sigma_i^2 v_i v_i^* \succeq 0$, also the generalized inverse root fulfills $[(ZZ^*)^{\frac{1}{2}}]^+ \succeq 0$ and for $ZZ^* = \sum_{i=1}^d \sigma_i^2 u_i u_i^* \succeq 0$, it follows that $[(ZZ^*)^{\frac{1}{2}}]^+ \succeq 0$. We stress that at least one of the matrices $(ZZ^*)^{\frac{1}{2}}$ and $(Z^*Z)^{\frac{1}{2}}$ is positive definite and hence also $\widetilde{W}(Z) \succ 0$. With the fact that $\widetilde{W}(Z) \succ 0$, the statement can be proven analogously to the results in (Fornasier et al., 2011, Lemma 5.1).

B.4 Proof of Lemma 17

(a) With the minimization property that defines $X^{(n+1)}$ in (29), the inequality $\epsilon^{(n+1)} \leqslant \epsilon^{(n)}$, and the minimization property that defines $Z^{(n+1)}$ in (28) and Lemma 14, the monotonicity follows from

$$\begin{aligned}\mathcal{J}_p(X^{(n)}, \epsilon^{(n)}, Z^{(n)}) &\geqslant \mathcal{J}_p(X^{(n+1)}, \epsilon^{(n)}, Z^{(n)}) \geqslant \mathcal{J}_p(X^{(n+1)}, \epsilon^{(n+1)}, Z^{(n)}) \\ &\geqslant \mathcal{J}_p(X^{(n+1)}, \epsilon^{(n+1)}, Z^{(n+1)}).\end{aligned}$$

(b) Using Theorem 14 and the monotonicity property of (a) for all $n \in \mathbb{N}$, we see that

$$\|X^{(n)}\|_{S_p}^p \leqslant g^p_{\epsilon^{(n)}}(X^{(n)}) = \mathcal{J}_p(X^{(n)}, \epsilon^{(n)}, Z^{(n)}) \leqslant \mathcal{J}_p(X^{(1)}, \epsilon^{(0)}, Z^{(0)}).$$

(c) The proof follows analogously to (Fornasier et al., 2011, Proposition 6.1) where only the technical calculation to bound $\sigma_1^p\big((\widetilde{W}^{(n)})^{-1}\big)$ requires to take into account that the spectrum of a Kronecker sum $A \oplus B$ consists of the pairwise sum of the spectra of A and B (Bernstein, 2009, Proposition 7.2.3).

B.5 Proof of Lemma 18

The first statement $\widetilde{W}(X^{(n)}, \epsilon^{(n)}) = \widetilde{W}^{(n)}$ is clear from the definition of $\widetilde{W}(X, \epsilon)$ and (10). To show the necessity of (35), let $X \in M_{d_1 \times d_2}$ be a critical point of (34). Without loss

of generality, let us assume that $d_1 \leqslant d_2$. In this case, a short calculation shows that $g_\epsilon^p(X) = \operatorname{tr}\left[(XX^* + \epsilon^2 \mathbf{I}_{d_1})^{p/2}\right]$. It follows from the matrix derivative rules of Magnus and Neudecker (1999, (7),(15),(18),(20) of Chapter 8.2) that

$$\nabla g_\epsilon^p(X) = p(XX^* + \epsilon^2 \mathbf{I}_{d_1})^{\frac{p-2}{2}} X = p \sum_{i=1}^{d} (\sigma_i^2 + \epsilon^2)^{\frac{p-2}{2}} \sigma_i u_i v_i^*,$$

using the singular value decomposition $X = \sum_{i=1}^{d} \sigma_i u_i v_i^*$ in the last equality. Using the Kronecker sum inversion formula (56), we see that $\nabla g_\epsilon^p(X) = p\big[\widetilde{W}(X,\epsilon) X_{\mathrm{vec}}\big]_{\mathrm{mat}}$. The proof can be continued analogously to (Daubechies et al., 2010, Lemma 5.2).

Appendix C. Proof of Theorem 9

For statement (i) of the convergence result of Algorithm 1, we use the following *reverse triangle inequalities* implied by the strong Schatten-p NSP: Let $X, X' \in M_{d_1 \times d_2}$ such that $\Phi(X - X') = 0$. Then

$$\|X' - X\|_F^p \leqslant \frac{2^p \gamma_r^{1-p/2}}{r^{1-p/2}} \frac{1}{1-\gamma_r} \left(\|X'\|_{S_p}^p - \|X\|_{S_p}^p + 2\beta_r(X)_{S_p} \right), \tag{61}$$

where $\beta_r(X)_{S_p}$ is defined in (22). This inequality can be proven using an adaptation of the proof of the corresponding result for ℓ_p-minimization in (Gao et al., 2015, Theorem 13) and the generalization of Mirksy's singular value inequality to concave functions (Audenaert, 2014; Foucart, 2018). Furthermore, the proof of the similar statement in (Kabanava et al., 2016, Theorem 12) can be adapted to show (61).

The further part of the proof of (i) as well as (ii) follow analogously to (Fornasier et al., 2011, Theorem 6.11) and (Daubechies et al., 2010, Theorem 5.3) using the preliminary results deduced in Section 6.

Statement (iii) is a direct consequence of Theorem 11, which is proven in Section 6.3.

References

A. Ahmed and J. Romberg. Compressive multiplexing of correlated signals. *IEEE Trans. Inf. Theory*, 61(1):479–498, 2015.

K. M. R. Audenaert. A generalisation of Mirsky's singular value inequalities. preprint, arXiv:1410.4941 [math.FA], 2014.

D. S. Bernstein. *Matrix Mathematics: Theory, Facts, and Formulas (Second Edition)*. Princeton University Press, 2009.

S. Bhojanapalli, B. Neyshabur, and N. Srebro. Global optimality of local search for low rank matrix recovery. In *Advances in Neural Information Processing Systems (NIPS)*, pages 3873–3881, 2016.

J. D. Blanchard, J. Tanner, and K. Wei. CGIHT: conjugate gradient iterative hard thresholding for compressed sensing and matrix completion. *Inf. Inference*, 4(4):289–327, 2015.

E. J. Candès and Y. Plan. Tight oracle inequalities for low-rank matrix recovery from a minimal number of noisy random measurements. *IEEE Trans. Inf. Theory*, 57(4):2342–2359, April 2011.

E. J. Candès and B. Recht. Exact matrix completion via convex optimization. *Found. Comput. Math.*, 9(6):717–772, 2009.

E. J. Candès, Y. Eldar, T. Strohmer, and V. Voroninski. Phase retrieval via matrix completion. *SIAM J. Imag. Sci.*, 6(1):199–225, 2013.

E. J. Candès, T. Strohmer, and V. Voroninski. PhaseLift: Exact and Stable Signal Recovery from Magnitude Measurements via Convex Programming. *Commun. Pure Appl. Math.*, 66(8):1241–1274, 2013.

E. J. Candès, X. Li, and M. Soltanolkotabi. Phase retrieval via Wirtinger flow: Theory and algorithms. *IEEE Trans. Inf. Theory*, 61(4):1985–2007, 2015.

R. Chartrand. Exact reconstructions of sparse signals via nonconvex minimization. *IEEE Signal Process. Lett.*, 14:707–710, 2007.

J. A. Chavez-Dominguez and D. Kutzarova. Stability of low-rank matrix recovery and its connections to Banach space geometry. *J. Math. Anal. Appl.*, 427(1):320–335, 2015.

B. Dacorogna. *Direct Methods in the Calculus of Variations*. Springer, New York, 1989.

I. Daubechies, R. DeVore, M. Fornasier, and C.S. Güntürk. Iteratively reweighted least squares minimization for sparse recovery. *Commun. Pure Appl. Math.*, 63:1–38, 2010.

M. A. Davenport and J. Romberg. An overview of low-rank matrix recovery from incomplete observations. *IEEE J. Sel. Topics Signal Process.*, 10:608–622, 06 2016.

D. L. Donoho, M. Gavish, and A. Montanari. The phase transition of matrix recovery from Gaussian measurements matches the minimax MSE of matrix denoising. *Proc. Nat. Acad. Sci. U.S.A.*, 110(21):8405–8410, 2013.

J. Duchi. Properties of the Trace and Matrix Derivatives. Available electronically at `https://web.stanford.edu/~jduchi/projects/matrix_prop.pdf`.

Y.C. Eldar, D. Needell, and Y. Plan. Uniqueness conditions for low-rank matrix recovery. *Appl. Comput. Harmon. Anal.*, 33(2):309–314, 2012.

M. Fazel. *Matrix rank minimization with applications*. Ph.D. Thesis, Electrical Engineering Department, Stanford University, 2002.

M. Fornasier, H. Rauhut, and R. Ward. Low-rank matrix recovery via iteratively reweighted least squares minimization. *SIAM J. Optim.*, 21(4):1614–1640, 2011. code from `https://github.com/rward314/IRLSM`].

M. Fornasier, S. Peter, H. Rauhut, and S. Worm. Conjugate gradient acceleration of iteratively re-weighted least squares methods. *Comput. Optim. Appl.*, 65(1):205–259, 2016.

S. Foucart. Concave Mirsky Inequality and Low-Rank Recovery. *SIAM J. Matrix Anal. Appl.*, 39(1):99–103, 2018.

S. Foucart and H. Rauhut. *A Mathematical Introduction to Compressive Sensing*. Applied and Numerical Harmonic Analysis. Birkhäuser/Springer, New York, 2013.

Y. Gao, J. Peng, S. Yue, and Y. Zhao. On the null space property of ℓ_q-minimization for $0 < q \leqslant 1$ in compressed sensing. *J. Funct. Spaces*, 2015:4203–4215, 2015.

R. Ge, J. D. Lee, and T. Ma. Matrix completion has no spurious local minimum. In *Advances in Neural Information Processing Systems (NIPS)*, pages 2973–2981, 2016.

I. Gohberg, S. Goldberg, and N. Krupnik. *Traces and determinants of linear operators*, volume 116 of *Operator Theory: Advances and Applications*. Birkhäuser, Basel, 2000.

D. Goldberg, D. Nichols, B. M. Oki, and D. Terry. Using collaborative filtering to weave an information tapestry. *Commun. ACM*, 35(12):61–70, 1992.

M. Grant and S. Boyd. CVX: Matlab software for disciplined convex programming, version 2.1. `http://cvxr.com/cvx`, March 2014.

D. Gross. Recovering low-rank matrices from few coefficients in any basis. *IEEE Trans. Inf. Theory*, 57(3):1548–1566, 2011.

D. Gross, Y.-K. Liu, S. T. Flammia, S. Becker, and J. Eisert. Quantum state tomography via compressed sensing. *Phys. Rev. Lett.*, 105:150401, 2010.

D. Gross, F. Krahmer, and R. Kueng. A partial derandomization of phaselift using spherical designs. *J. Fourier Anal. Appl.*, 21(2):229–266, 2015.

J. P. Haldar and D. Hernando. Rank-constrained solutions to linear matrix equations using powerfactorization. *IEEE Signal Process. Lett.*, 16(7):584–587, July 2009. [using `AltMin` (Alternating Minimization) algorithm].

N. Halko, P. G. Martinsson, and J. A. Tropp. Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions. *SIAM Rev.*, 53 (2):217–288, 2011.

P. Jain, Raghu M., and I. S. Dhillon. Guaranteed rank minimization via singular value projection. In *Advances in Neural Information Processing Systems (NIPS)*, pages 937–945, 2010.

P. Jain, P. Netrapalli, and S. Sanghavi. Low-rank matrix completion using alternating minimization. In *Proc. ACM Symp. Theory Comput. (STOC)*, pages 665–674, Palo Alto, CA, USA, June 2013.

A. Jameson. Solution of the equation ax + xb = c by inversion of an m x m or n x n matrix. *SIAM J. Appl. Math.*, 16(5):1020–1023, 1968.

M. Kabanava, R. Kueng, H. Rauhut, and U. Terstiege. Stable low-rank matrix recovery via null space properties. *Inf. Inference*, 5(4):405–441, 2016.

F. J. Király, L. Theran, and R. Tomioka. The Algebraic Combinatorial Approach for Low-Rank Matrix Completion. *J. Mach. Learn. Res.*, 16:1391–1436, 2015.

A. Kyrillidis and V. Cevher. Matrix recipes for hard thresholding methods. *J. Math. Imaging Vision*, 48(2):235–265, 2014. [using `Matrix ALPS II` ('Matrix ALgrebraic PursuitS II') algorithm, code from `http://akyrillidis.github.io/projects/`].

C. Kümmerle and J. Sigl. Harmonic Mean Iteratively Reweighted Least Squares for low-rank matrix recovery. In *12th International Conference on Sampling Theory and Applications (SampTA)*, pages 489–493, 2017.

Z. Liu and L. Vandenberghe. Interior-point method for nuclear norm approximation with application to system identification. *SIAM J. Matrix Anal. Appl.*, 31(3):1235–1256, 2010.

Z. Liu, A. Hansson, and L. Vandenberghe. Nuclear norm system identification with missing inputs and outputs. *Systems Control Lett.*, 62(8):605–612, 2013.

J.R. Magnus and H. Neudecker. *Matrix Differential Calculus with Applications in Statistics and Econometrics*. Wiley Series in Probability and Statistics. Wiley, 1999.

B. Mishra, G. Meyer, F. Bach, and R. Sepulchre. Low-rank optimization with trace norm penalty. *SIAM J. Optim.*, 23(4):2124–2149, 2013.

K. Mohan and M. Fazel. Iterative reweighted algorithms for matrix rank minimization. *J. Mach. Learn. Res.*, 13(1):3441–3473, 2012. [using `IRLS-MF` ('IRLS-p') algorithm, code from `https://faculty.washington.edu/mfazel/`].

S. Oymak, K. Mohan, M. Fazel, and B. Hassibi. A simplified approach to recovery conditions for low rank matrices. In *Proceedings of the IEEE International Symposium on Information Theory (ISIT)*, pages 2318–2322, 2011.

D. Park, A. Kyrillidis, C. Caramanis, and S. Sanghavi. Finding Low-rank Solutions to Matrix Problems, Efficiently and Provably. preprint, arXiv:1606.03168 [math.OC], [using `BFGD` ('Bi-Factored Gradient Descent') algorithm, code from `http://akyrillidis.github.io/projects/`], 2016.

D.L. Pimentel-Alarcón, N. Boston, and R. D. Nowak. A Characterization of Deterministic Sampling Patterns for Low-Rank Matrix Completion. preprint, arXiv:1503.02596v3 [stat.ML], October 2016.

B. Recht, M. Fazel, and P. A. Parrilo. Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM Rev.*, 52(3):471–501, 2010.

B. Recht, W. Xu, and B. Hassibi. Null space conditions and thresholds for rank minimization. *Math. Program.*, 127(1):175–202, 2011.

É. Schost and P.-J. Spaenlehauer. A quadratically convergent algorithm for structured low-rank approximation. *Found. Comput. Math.*, 16(2):457–492, 2016.

N. Srebro, J. Rennie, and T. S. Jaakkola. Maximum-margin matrix factorization. In *Advances in Neural Information Processing Systems (NIPS)*, pages 1329–1336, 2005.

M. Stewart. Perturbation of the SVD in the presence of small singular values. *Linear Algebra Appl.*, 419(1):53–77, 2006.

R. Sun and Z. Q. Luo. Guaranteed matrix completion via non-convex factorization. *IEEE Trans. Inf. Theory*, 62(11):6535–6579, 2016.

J. Tanner and K. Wei. Normalized Iterative Hard Thresholding for Matrix Completion. *SIAM J. Sci. Comput.*, 35(5):S104–S125, 2013.

J. Tanner and K. Wei. Low rank matrix completion by alternating steepest descent methods. *Appl. Comput. Harmon. Anal.*, 40(2):417–429, 2016. [using `ASD` ('Alternating Steepest Descent') algorithm, code from `https://www.math.ucdavis.edu/~kewei/publications.html`].

S. Tu, R. Boczar, M. Simchowitz, M. Soltanolkotabi, and B. Recht. Low-rank Solutions of Linear Matrix Equations via Procrustes Flow. In *Proceedings of The 33rd International Conference on Machine Learning*, volume 48, pages 964–973, 2016.

B. Vandereycken. Low-rank matrix completion by Riemannian optimization. *SIAM J. Optim.*, 23(2):1214–1236, 2013. [using `Riemann_Opt` ('Riemannian Optimization') algorithm, code from `http://www.unige.ch/math/vandereycken/matrix_completion.html`].

P.-Å. Wedin. Perturbation bounds in connection with singular value decomposition. *BIT*, 12(1):99–111, 1972.

K. Wei, J.-F. Cai, T. F. Chan, and S. Leung. Guarantees of Riemannian Optimization for Low Rank Matrix Recovery. *SIAM J. Matrix Anal. Appl.*, 37(3):1198–1222, 2016.

Z. Wen, W. Yin, and Y. Zhang. Solving a low-rank factorization model for matrix completion by a nonlinear successive over-relaxation algorithm. *Math. Program. Comput.*, 4(4):333–361, 2012.

Q. Zheng and J. Lafferty. A convergent gradient descent algorithm for rank minimization and semidefinite programming from random linear measurements. In *Advances in Neural Information Processing Systems (NIPS)*, pages 109–117, 2015.