Matrix denoising for weighted loss functions and heterogeneous signals

William Leeb*

Abstract

We consider the problem of estimating a low-rank matrix from a noisy observed matrix. Previous work has shown that the optimal method depends crucially on the choice of loss function. In this paper, we use a family of weighted loss functions, which arise naturally in many settings such as heteroscedastic noise, missing data, and submatrix denoising. However, weighted loss functions are challenging to analyze because they are not orthogonally-invariant. We derive optimal spectral denoisers for these weighted loss functions. By combining different weights, we then use these optimal denoisers to construct a new denoiser that exploits heterogeneity in the signal matrix to boost estimation with unweighted loss.

1 Introduction

This paper is concerned with the problem of low-rank matrix denoising: estimating a low-rank signal matrix $X$ from an observed matrix $Y = X + G$, where $G$ is a full-rank matrix of noise. This is a ubiquitous problem in statistical applications to which a fairly large literature has already been devoted.

The denoising problem is often studied in an asymptotic regime where the number of rows and columns of $X$ grow in proportion to each other; that is, if $X$ has $p$ rows and $n$ columns, both $p$ and $n$ grow to infinity, and their ratio $p/n$ converges to a definite limit. We assume, as is often the case, that the energy of the signal $X$ stays fixed with $p$ and $n$, while the energy in the noise matrix $G$ grows. There is an extensive literature on the asymptotic behavior of the singular vectors and singular values of such data matrices $Y$ in this high-dimensional, high-noise regime, which may be referred to as the spiked model [4, 3, 5, 37, 27, 7].

A standard procedure for the denoising problem is singular value shrinkage [40, 18, 17, 36, 16]. In its simplest incarnation, singular value shrinkage keeps the singular vectors of the observed matrix $Y$, while deflating its singular values to remove the effects of the noise. Most of the observed singular values will be set to 0, while the top $r$ singular values (if $X$ has rank $r$) are deflated to remove the influence of the noise.

We introduce a natural extension of singular value shrinkage, which we call spectral denoising. Previous work has derived optimal singular value shrinkers for a variety of orthogonally-invariant loss functions, such as Frobenius norm loss, operator norm loss, and nuclear norm loss [18]. In this paper, we will instead consider the class of weighted Frobenius loss functions – in which errors in different rows and columns incur different penalties – and derive optimal spectral denoisers for these losses. For uniform weights (i.e. unweighted Frobenius loss), optimal spectral denoising and optimal singular value shrinkage coincide. However, for general weighted loss functions spectral denoising will outperform singular value shrinkage.

Weighted loss functions arise naturally in a number of applications, such as denoising with heteroscedastic noise [34, 43], missing data and other linear deformations of $X$ [14], and submatrix denoising. However, weighted loss functions are mathematically challenging because they are not orthogonally-invariant. In order to derive optimal spectral denoisers for weighted loss functions, we extend the asymptotic theory of the classical spiked model, building on earlier work from [34].

In addition to deriving optimal spectral denoisers for weighted losses, we derive a new procedure for ordinary unweighted Frobenius loss. Prior work has shown that when the noise matrix $G$ is orthogonally-invariant, then for orthogonally-invariant loss functions singular value shrinkage is both minimax optimal [16] and optimal in expectation over a uniform prior on the signal singular vectors [40]. More informally,

*School of Mathematics, University of Minnesota, Twin Cities. Minneapolis, MN.
singular value shrinkage is optimal when we possess no prior knowledge on the signal. This leaves open the question of how to exploit structural information on the signal components.

As a byproduct of our analysis of weighted loss functions, we derive a new matrix denoising algorithm for ordinary, unweighted Frobenius loss that asymptotically is never worse than optimal singular value shrinkage, but may be significantly better. We call this method localized denoising. Localized denoising exploits heterogeneity in X’s singular vectors. Informally, this means that when certain blocks of coordinates of X are known to contain more of the signal’s energy than others, localized denoising will outperform shrinkage.

The remainder of the paper is structured as follows. Section 2 introduces the observation model; describes the class of spectral denoisers; introduces weighted loss functions and the associated estimation problem; and defines heterogeneity. Section 3 introduces the new mathematical theory for the spiked model with weights. Section 4 derives the optimal spectral denoiser for weighted loss. Section 5 introduces a certain subclass of spectral denoisers, called diagonal denoisers, which are more amenable to analysis. Section 6 introduces the family of localized denoisers for unweighted loss. Section 7 describes three applications of weighted loss functions: heteroscedastic noise, missing data, and submatrix denoising, and derives several theoretical results for these problems. Section 8 reports on numerical experiments. Finally, Section 9 makes several concluding remarks and discusses future research questions and potential applications.

2 Preliminaries

2.1 The observation model

We observe a p-by-n data matrix \( Y = X + G \), consisting of a low-rank signal matrix \( X \) and a full-rank isotropic Gaussian noise matrix \( G \). We will suppose \( X \) has SVD:

\[
X = \sum_{k=1}^{r} t_k u_k v_k^T,
\]

where the \( u_k \) and \( v_k \) are orthonormal vectors in \( \mathbb{R}^p \) and \( \mathbb{R}^n \), respectively, and \( t_1 > \cdots > t_r > 0 \). The entries of the noise matrix \( G \) are iid \( N(0, 1/n) \). We write the SVD of \( Y \) as

\[
Y = \sum_{k=1}^{\min(p,n)} \lambda_k \hat{u}_k \hat{v}_k^T
\]

where the \( \hat{u}_k \) and \( \hat{v}_k \) are orthonormal vectors in \( \mathbb{R}^p \) and \( \mathbb{R}^n \), respectively, and \( \lambda_1 \geq \cdots \geq \lambda_{\min(p,n)} \geq 0 \).

We let \( \Omega \) be a matrix with \( p \) rows, and \( \Pi \) be a matrix with \( n \) rows. In Section 2.3, these matrices will be used to define the loss function for estimating \( X \).

We will parametrize the problem size by the number of columns \( n \), and let the number of rows \( p = p_n \) grow with \( n \). Specifically, we will assume that the limit

\[
\gamma = \lim_{n \to \infty} \frac{p_n}{n}
\]

is well-defined and finite. In all statements where \( n \to \infty \), it will be implicitly assumed as well that \( p \to \infty \) and \( p/n \to \gamma \).

In order to have a well-defined asymptotic theory, we will need to assume that certain quantities have definite, finite limits. We define:

\[
\mu = \lim_{p \to \infty} \frac{1}{p} \text{tr}(\Omega^T \Omega)
\]

and

\[
\nu = \lim_{n \to \infty} \frac{1}{n} \text{tr}(\Pi^T \Pi).
\]

For each \( k = 1, \ldots, r \), we also define

\[
\alpha_k = \lim_{p \to \infty} \| \Omega u_k \|^2
\]
and

\[ \beta_k = \lim_{n \to \infty} \| \Pi v_k \|^2. \]  

(7)

We assume that these limits exist, and are finite and positive. Finally, since the matrices \( \Omega \) and \( \Pi \) grow as \( n \) grows, we assume that their operator norms remain bounded as \( n \to \infty \).

### 2.2 Heterogeneity, genericity, and weighted orthogonality

One of the aspects of the theory of matrix denoising we will explore is the role of the distribution of the signal matrix \( X \)'s singular vectors, \( u_1, \ldots, u_r \) and \( v_1, \ldots, v_r \). We will say that a unit vector \( x \in \mathbb{R}^m \) is \textit{generic} with respect to an \( m \times m \) positive-semidefinite matrix \( A \) if it satisfies

\[ x^T Ax \sim \frac{1}{m} \text{tr}(A), \]  

(8)

where “\( \sim \)” indicates that the difference between the two sides vanishes almost surely as \( m \to \infty \). By contrast, we say that \( x \) is \textit{heterogeneous} if it is not generic. Heterogeneity means that the energy of the vector \( x \) is not uniformly distributed across its coordinates in the eigenbasis of \( A \). Indeed, if \( A = \sum_{k=1}^m h_k w_k w_k^T \) is the eigendecomposition of \( A \), then

\[ x^T Ax = \sum_{k=1}^m h_k \langle x, w_k \rangle^2. \]  

(9)

If the energy of \( x \) were equally spread out across the \( w_k \), then \( \langle x, w_k \rangle \sim 1/\sqrt{m} \), and so \( x^T Ax \sim \text{tr}(A)/m \).

Given a collection of vectors \( x_1, \ldots, x_k \in \mathbb{R}^m \), we will say that they satisfy the \textit{weighted orthogonality} condition (or are \textit{weighted orthogonal}) with respect to a positive-semidefinite matrix \( A \) if

\[ x_i^T Ax_j \sim 0 \]  

whenever \( i \neq j \). In other words, the \( x_j \) are asymptotically orthogonal with respect to the weighted inner product defined by \( A \).

**Remark 1.** From the Hanson-Wright inequality \([21, 42, 39]\), random unit vectors \( x \) from suitably regular distributions are generic, with respect to any \( A \) with bounded operator norm. Furthermore, the weighted orthogonality condition will also hold for independent random unit vectors \( x_1, \ldots, x_k \) from a suitable distribution (see [6]).

### 2.3 Spectral denoisers and weighted loss functions

For the top \( r \) empirical singular vectors \( \hat{u}_1, \ldots, \hat{u}_r \) and \( \hat{v}_1, \ldots, \hat{v}_r \), define the matrices \( \hat{U} = [\hat{u}_1, \ldots, \hat{u}_r] \) and \( \hat{V} = [\hat{v}_1, \ldots, \hat{v}_r] \). Consider the class of estimators of \( X \) defined by

\[ S = \left\{ \hat{U} \hat{B} \hat{V}^T : \hat{B} \in \mathbb{R}^{r \times r} \right\}. \]  

(11)

Each matrix in \( S \) has the same singular subspaces as the observed matrix \( Y \). We will call \( S \) the family of \textit{spectral denoisers}.

We consider estimating the low-rank signal matrix \( X \) with respect to the \textit{weighted Frobenius loss} defined by

\[ L_n(\hat{X}, X) = \| \Omega(\hat{X} - X)\Pi^T \|_F^2, \]  

(12)

where \( \| \cdot \|_F \) denotes the matrix Frobenius norm. This type of loss function is used when the user pays different prices for errors in different rows and columns.

We now define the precise estimation problem we will consider. For any deterministic \( r \)-by-\( r \) matrix \( \hat{B} \), we define the asymptotic error

\[ L(\hat{U} \hat{B} \hat{V}^T, X) = \lim_{n \to \infty} L_n(\hat{U} \hat{B} \hat{V}^T, X). \]  

(13)
It will follow from our subsequent analysis that this asymptotic loss is well-defined almost surely. Our goal is then to find the matrix \( \hat{B} \) to minimize this loss, and show how \( \hat{B} \) may be consistently estimated from the observed matrix \( Y \). That is, we define

\[
\hat{B} = \operatorname{argmin}_{\hat{B} \in \mathbb{R}^{r \times r}} \mathcal{L}(\hat{U}\hat{B}\hat{V}^T, X)
\]

and define \( \hat{X} = \hat{U}\hat{B}\hat{V}^T \).

### 3 Asymptotic theory for the spiked model

In this section, we derive the asymptotic limits of inner products between the weighted population and empirical vectors. We define the cosines between the unweighted empirical and population vectors:

\[
c_{jk} = \lim_{p \to \infty} \langle \hat{u}_j, u_k \rangle, \quad \hat{c}_{jk} = \lim_{n \to \infty} \langle \hat{v}_j, v_k \rangle.
\]

Next we define weighted inner products between the population and empirical vectors:

\[
c_{\omega}^{\omega} = \lim_{p \to \infty} \langle \Omega\hat{u}_j, \Omega u_k \rangle, \quad \hat{c}_{\omega}^{\omega} = \lim_{n \to \infty} \langle \Pi\hat{v}_j, \Pi v_k \rangle.
\]

These are inner products with weight matrices \( \Omega^T\Omega \) and \( \Pi^T\Pi \), respectively. We also define the weighted inner products between the empirical singular vectors:

\[
d_{jk} = \lim_{p \to \infty} \langle \Omega \hat{u}_j, \Omega \hat{u}_k \rangle, \quad \hat{d}_{jk} = \lim_{n \to \infty} \langle \Pi \hat{v}_j, \Pi \hat{v}_k \rangle.
\]

Finally, we define weighted inner products between the population singular vectors:

\[
e_{jk} = \lim_{p \to \infty} \langle \Omega u_j, \Omega u_k \rangle, \quad \hat{e}_{jk} = \lim_{n \to \infty} \langle \Pi v_j, \Pi v_k \rangle.
\]

We will let \( c_{\omega}^{\omega} = c_{\omega k}^{\omega} \) and \( \hat{c}_{\omega}^{\omega} = \hat{c}_{\omega k}^{\omega} \), and similarly for the other terms with repeated subscripts.

**Remark 2.** In the notation of Section 2.1, specifically definitions (6) and (7), \( e_k = \alpha_k \) and \( \hat{e}_k = \beta_k \).

The first result provides formulas for \( c_{jk} \) and \( \hat{c}_{jk} \), and relates the singular values of \( X \) to those of \( Y \). It is well-known in the literature (see e.g., \([37, 7]\)).

**Proposition 3.1.** For \( 1 \leq j, k \leq r \),

\[
\lambda_j^2 = \begin{cases} 
(t_k^2 + 1) \left(1 + \frac{\gamma}{t_k}\right), & \text{if } t_k > \gamma^{1/4}, \\
(1 + \sqrt{\gamma})^2, & \text{if } t_k \leq \gamma^{1/4},
\end{cases}
\]

(19)

\[
c_{jk}^2 = \begin{cases} 
1 - \gamma/t_k^2 & \text{if } j = k \text{ and } t_k > \gamma^{1/4}, \\
1 + \gamma/t_k^2 & \text{if } j \neq k \text{ or } t_k \leq \gamma^{1/4},
\end{cases}
\]

(20)

and

\[
\hat{c}_{jk}^2 = \begin{cases} 
1 - \gamma/t_k^2 & \text{if } j = k \text{ and } t_k > \gamma^{1/4}, \\
1 + \gamma/t_k^2 & \text{if } j \neq k \text{ or } t_k \leq \gamma^{1/4}.
\end{cases}
\]

(21)

**Remark 3.** While the signs of \( c_k \) and \( \hat{c}_k \) are arbitrary (since the sign of a singular vector may be flipped), their product satisfies \( c_k\hat{c}_k \geq 0 \). We may therefore assume that \( c_k \geq 0 \) and \( \hat{c}_k \geq 0 \) (see, e.g., \([36]\)).

Now we describe the limits of the weighted inner products.
Theorem 3.2. Suppose $1 \leq j, k \leq r$. Then the following expressions hold almost surely:

$$c_{jk}^* = \begin{cases} e_{jk}e_j, & \text{if } t_j > \gamma^{1/4}, \\ 0, & \text{if } t_j \leq \gamma^{1/4}, \end{cases}$$

(22)

$$\tilde{c}_{jk}^* = \begin{cases} \hat{e}_{jk}\hat{e}_j, & \text{if } t_j > \gamma^{1/4}, \\ 0, & \text{if } t_j \leq \gamma^{1/4}, \end{cases}$$

(23)

$$d_{jk} = \begin{cases} c_k^2\alpha_k + s_k^2\mu, & \text{if } j = k \text{ and } t_k > \gamma^{1/4}, \\ e_{jk}c_jc_k, & \text{if } j \neq k \text{ and } \min\{t_j, t_k\} > \gamma^{1/4}, \\ 0, & \text{if } j \neq k \text{ and } \min\{t_j, t_k\} \leq \gamma^{1/4}, \end{cases}$$

(24)

$$\tilde{d}_{jk} = \begin{cases} \tilde{c}_k^2\beta_k + \tilde{s}_k^2\nu, & \text{if } j = k \text{ and } t_k > \gamma^{1/4}, \\ \hat{e}_{jk}\hat{e}_j\tilde{c}_k, & \text{if } j \neq k \text{ and } \min\{t_j, t_k\} > \gamma^{1/4}, \\ 0, & \text{if } j \neq k \text{ and } \min\{t_j, t_k\} \leq \gamma^{1/4}. \end{cases}$$

(25)

The proof of Theorem 3.2 may be found in Appendix A.

Remark 4. While the signs of inner products between singular vectors is arbitrary (since the signs of singular vectors may be chosen arbitrarily), Theorem 3.2 shows that the signs of $c_{jk}^*$, $\tilde{c}_{jk}^*$, $d_{jk}$ and $\tilde{d}_{jk}$ are determined by the signs of $e_{jk}$ and $\hat{e}_{jk}$.

4 Optimal spectral denoising

In this section, we give the expression for the asymptotically optimal spectral denoiser with respect to the weighted loss, and show how it may be consistently estimated from the observed matrix $Y$. We define the $r$-by-$r$ weighted inner product matrices $D = (d_{kl})$, $\tilde{D} = (\tilde{d}_{kl})$, $E = (e_{jk})$, $\hat{E} = (\hat{e}_{jk})$, $C = (c_{jk}^*)$, and $\tilde{C} = (\tilde{c}_{jk}^*)$, and the vector $t = (t_1, \ldots, t_r)^T$ of population singular values. Then:

Theorem 4.1. The optimal choice of $\tilde{B}$ is given by:

$$\tilde{B} = D^+ C\text{diag}(t)\tilde{C}^T\tilde{D}^+, \quad (26)$$

with weighted AMSE almost surely equal to

$$\lim_{n \to \infty} \|\Omega(\tilde{X} - X)\|^2_F = (\text{diag}(t)\hat{E} - C^T D^+ C\text{diag}(t)\tilde{C}^T \tilde{D}^+ \tilde{C}, \text{diag}(t))_F. \quad (27)$$

The proof of Theorem 4.1 may be found in Appendix B.

The weighted inner product matrices $D$, $\tilde{D}$, $E$, $\hat{E}$, $C$, and $\tilde{C}$ and the singular values $t_1, \ldots, t_r$ may be estimated using Proposition 3.1 and Theorem 3.2. First, from Proposition 3.1, we can estimate $t_k$, $c_k$ and $\hat{c}_k$, so long as $t_k > \gamma^{1/4}$, i.e. if $\lambda_k > 1 + \sqrt{\gamma}$. Specifically,

$$t_k = \sqrt{\frac{\lambda_k^2 - 1 - \gamma + \sqrt{(\lambda_k^2 - 1 - \gamma)^2 - 4\gamma}}{2}}, \quad (28)$$

$$c_k = \sqrt{\frac{1 - \gamma/t_k^2}{1 + \gamma/t_k^2}}, \quad (29)$$

and

$$\hat{c}_k = \sqrt{\frac{1 - \gamma/t_k^2}{1 + 1/t_k^2}}. \quad (30)$$
Algorithm 1 Optimal spectral denoising with weighted loss

1: Input: $Y$; weights $\Omega$ and $\Pi$
2: rank $r$ SVD of $Y$:
   \[
   \lambda_1 \geq \cdots \geq \lambda_r > 1 + \sqrt{\gamma} \\
   \hat{U} = [\hat{u}_1, \ldots, \hat{u}_r], \quad \hat{V} = [\hat{v}_1, \ldots, \hat{v}_r]
   \]
3: $\mu = \text{tr}(\Omega^T \Omega)/p$, $\nu = \text{tr}(\Pi^T \Pi)/n$
4: for $1 \leq k \leq r$:
   \[
   t_k = \sqrt{\frac{\lambda_k^2 - 1 - \gamma + \sqrt{(\lambda_k^2 - 1 - \gamma)^2 - 4\gamma}}{2}} \\
   c_k = \sqrt{\frac{1 - s_k^2}{1 + 1/s_k^2}}, \quad \hat{c}_k = \sqrt{\frac{1 - \overline{s}_k^2}{1 + 1/\overline{s}_k^2}} \\
   s_k = \frac{1 - \overline{c}_k^2}{\overline{c}_k^2}, \quad \overline{s}_k = \sqrt{1 - \overline{c}_k^2} \\
   d_k = \|\Omega \hat{u}_k\|^2, \quad \hat{d}_k = \|\Pi \hat{v}_k\|^2 \\
   e_k = (d_k - s_k^2 \mu)/c_k^2, \quad \hat{e}_k = (\hat{d}_k - \overline{s}_k^2 \nu)/\overline{c}_k^2 \\
   c_k^2 = e_k c_k, \quad \overline{c}_k^2 = \hat{e}_k \hat{c}_k
   \]
5: for $1 \leq j \neq k \leq r$:
   \[
   d_{jk} = \hat{u}_j^T \Omega^T \Omega \hat{u}_k, \quad \hat{d}_{jk} = \hat{v}_j^T \Pi^T \Pi \hat{v}_k \\
   e_{jk} = d_{jk}/(c_j c_k), \quad \hat{e}_{jk} = \hat{d}_{jk}/(\hat{c}_j \hat{c}_k), \quad j \neq k \\
   e_{jk}^2 = e_{jk} c_j, \quad \overline{e}_{jk}^2 = \overline{e}_{jk} \overline{c}_j
   \]
6: $D = (d_{jk}), \quad \hat{D} = (\hat{d}_{jk})$
7: $E = (e_{jk}), \quad \hat{E} = (\hat{e}_{jk})$
8: $C = (e_{jk}^2), \quad \hat{C} = (\overline{e}_{jk}^2)$
9: $t = (t_1, \ldots, t_r)^T$
10: $\tilde{\mathbf{B}} = D C \text{diag}(t) \hat{C}^T \hat{D}^+$
11: $\hat{X} = \hat{U} \hat{B} \hat{V}^T$
12: $\text{AMSE} = \langle (E \text{diag}(t) \tilde{E} - C^T D^+ C \text{diag}(t) \hat{C}^T \hat{D}^+ \tilde{C}, \text{diag}(t))_F$

Remark 5. It follows from Remark 3 that we can take both $c_k$ and $\hat{c}_k$ to be positive.

The values $d_{jk} = \hat{u}_j^T \Omega^T \Omega \hat{u}_j$ and $\hat{d}_{jk} = \hat{v}_j^T \Pi^T \Pi \hat{v}_k$ are directly estimable, as they are the weighted inner products between the empirical singular vectors. We then estimate $e_k = \alpha_k$ and $\hat{e}_k = \beta_k$, assuming $t_k > \gamma^{1/4}$:

\[
e_k = \frac{d_k - s_k^2 \mu}{c_k^2},
\]
and

\[
\hat{e}_k = \frac{\hat{d}_k - \overline{s}_k^2 \nu}{\overline{c}_k^2}.
\]

When $j \neq k$, we solve for $e_{jk}$ and $\hat{e}_{jk}$ by $e_{jk} = d_{jk}/(c_j c_k)$ and $\hat{e}_{jk} = \hat{d}_{jk}/(\hat{c}_j \hat{c}_k)$ (so long as $t_j$ and $t_k$ both exceed $\gamma^{1/4}$, i.e. $\lambda_j$ and $\lambda_k$ both exceed $1 + \sqrt{\gamma}$). Finally, for all $j, k$, we estimate $e_{jk}^2$ and $\overline{e}_{jk}^2$ by $e_{jk}^2 = e_{jk} c_j$ and $\overline{e}_{jk}^2 = \overline{e}_{jk} \overline{c}_j$.

This completes the derivation and estimation of the optimal spectral denoiser. The entire procedure is summarized in Algorithm 1.

5 Diagonal denoisers

In this section, we consider a subset of spectral denoisers, in which the matrix $\hat{B}$ is required to be diagonal. More precisely, we search for a vector $t = (t_1, \ldots, t_r)^T$ of real numbers, so that the estimator

\[
\hat{X}^{dd} = \hat{U} \text{diag}(t) \hat{V}^T = \sum_{k=1}^r t_k \hat{u}_k \hat{v}_k^T
\]
minimizes the AMSE $\mathcal{L}(\hat{X}_{dd}, X) = \lim_{n \to \infty} \|\Omega(\hat{X}_{dd} - X)\Pi^T\|_F^2$.

We will show the following result:

**Theorem 5.1.** Suppose that either $u_1, \ldots, u_r$ are weighted orthogonal with respect to $\Omega^T\Omega$, or $v_1, \ldots, v_r$ are weighted orthogonal with respect to $\Pi^T\Pi$. Suppose too that $t_k > \gamma^{1/4}$, $1 \leq k \leq r$. Then the singular values $t_k$, $1 \leq k \leq r$, of the optimal diagonal denoiser $\hat{X}^{dd}$ are given by

$$\hat{t}_k = t_k c_k \tilde{c}_k \cdot \frac{\alpha_k}{c_k^2 \alpha_k + s_k^2 \mu} \cdot \frac{\beta_k}{c_k^2 \beta_k + s_k^2 \nu},$$

with weighted AMSE almost surely equal to

$$\lim_{n \to \infty} \|\Omega(\hat{X} - X)\Pi^T\|_F^2 = \sum_{k=1}^r t_k^2 \alpha_k \beta_k \left(1 - c_k^2 e_{k}^2 \cdot \frac{\alpha_k}{c_k^2 \alpha_k + s_k^2 \mu} \cdot \frac{\beta_k}{c_k^2 \beta_k + s_k^2 \nu}\right).$$

If $u_1, \ldots, u_r$ and $v_1, \ldots, v_r$ are both weighted orthogonal with respect to $\Omega^T\Omega$ and $\Pi^T\Pi$, respectively, then $\hat{X} = \hat{X}_{dd}$. That is, the optimal spectral denoiser is equal to the optimal diagonal denoiser.

The proof of Theorem 5.1 is found in Appendix C.

**Remark 6.** Theorem 5.1 provides an explicit formula for the optimal diagonal denoiser when either the left or right singular vectors (or both) of $X$ satisfy the weighted orthogonality condition. This condition will hold, for instance, when the columns of $X$ are drawn iid from a random distribution in $\mathbb{R}^p$, as is often the case in applications.

**Remark 7.** The optimal diagonal denoiser will never have better asymptotic performance than the optimal spectral denoiser, since diagonal denoisers are a subset of spectral denoisers. However, Theorem 5.1 shows that under weighted orthogonality, the two methods coincide. We also remark that the weighted orthogonality condition can be tested from the observed matrix $Y$ using Theorem 3.2, as $e_{jk} = 0$ if and only if $d_{jk} = 0$.

**Remark 8.** As shown in [40, 18, 36], the optimal singular value shrinker for unweighted Frobenius loss has singular values

$$\tilde{t}^{shr}_k = t_k c_k \tilde{c}_k.$$  \hspace{1cm} (36)

Theorem 5.1 shows that when minimizing a weighted loss function under weighted orthogonality, the optimal singular values are of the form $t = \tilde{t}^{shr}_k \eta_k$, where $\eta_k$ is the “correction factor” given by

$$\eta_k = \frac{\alpha_k}{c_k^2 \alpha_k + s_k^2 \mu} \cdot \frac{\beta_k}{c_k^2 \beta_k + s_k^2 \nu}. \hspace{1cm} (37)$$

The factor $\eta_k$ can be interpreted as measuring the interaction between the singular vectors $u_k$ and $v_k$ of $X$ and the weight matrices $\Omega$ and $\Pi$. For instance, if both $u_k$ and $v_k$ are generic with respect to, respectively, $\Omega^T\Omega$ and $\Pi^T\Pi$ (in the sense defined in Section 2.2), then $\alpha_k = \mu$ and $\beta_k = \nu$, and consequently $\eta_k = 1$. In this case, the optimal singular value $\hat{t}_k$ for weighted Frobenius loss would be equal to the optimal singular value $\tilde{t}^{shr}_k$ for unweighted Frobenius loss.

### 5.1 Behavior of the optimal singular values

In this section, we study denoising of a single component; we consequently drop the subscript $k$. We describe certain aspects of the behavior of the optimal singular value $\hat{t}$. Our results are summarized in Propositions 5.2 and 5.3.

**Proposition 5.2.** If either $\alpha \leq \mu$ or $\beta \leq \nu$, then $\hat{t} \leq \lambda$; that is, the optimal separable denoiser shrinks the observed singular values. Conversely, for any fixed value of $t$, there are sufficiently large values of $\alpha$ and $\beta$ for which $\hat{t} > \lambda$. 


Figure 1: The optimal singular value $\hat{t}$, plotted as a function of the observed singular value $\lambda$ (left) and the population singular value $t$ (right), for varying values of $\alpha$ and $\beta$ and $\mu = \nu = 1$.

**Proposition 5.3.** If either $\alpha \leq \mu$ or $\beta \leq \nu$, then $\hat{t}$ increases monotonically as a function of $\lambda$.

The proofs of Propositions 5.2 and 5.3 may be found in Appendix D and Appendix E, respectively.

**Remark 9.** As described in Remark 8, the optimal singular value for unweighted Frobenius loss is $\hat{t}_{\text{shr}} = tc\hat{c}$, which is smaller than the observed singular value $\lambda$. Proposition 5.2 shows that with weighted Frobenius loss, this shrinkage phenomenon will only occur when either $\alpha$ or $\beta$ are sufficiently small.

**Remark 10.** The conclusion of Proposition 5.3 need not hold if $\alpha > \mu$ and $\beta > \mu$. To illustrate, in Figure 1 we plot the optimal $\hat{t}$, both as a function of the observed singular value $\lambda$ and the population singular value $t$, for various values of $\alpha$ and $\beta$ (and $\mu = \nu = 1$). The non-monotonicity is apparent when $\alpha = \beta = 10$.

## 6 Localized denoising

In this section, we revisit the problem of estimating $X$ using unweighted Frobenius loss. It is known that singular value shrinkage is the optimal method in this setting, both in the minimax sense and when averaging over a uniform prior on the $u_k$ and $v_k$ [16, 40]. The optimal singular value shrinkage estimator is given by:

$$\hat{X}_{\text{shr}} = \sum_{k=1}^r \hat{t}_{k}^{\text{shr}} \hat{u}_k \hat{v}_k^T,$$

(38)

where $\hat{t}_{k}^{\text{shr}} = t_k c_k \hat{c}_k$.

In this section, we introduce a new procedure which we call localized denoising. As we will show, localized denoising is asymptotically never worse than singular value shrinkage, and can perform better when the singular vectors of $X$ are heterogeneous.

### 6.1 Definition of localized denoising

To define localized denoising, take two expansions of the identity matrices $I_p$ and $I_n$ into sums of pairwise orthogonal projection matrices $\Omega_i \in \mathbb{R}^{p \times p}, 1 \leq i \leq I$, and $\Pi_j \in \mathbb{R}^{n \times n}, 1 \leq j \leq J$:

$$I_p = \sum_{i=1}^I \Omega_i, \quad I_n = \sum_{j=1}^J \Pi_j.$$

(39)

Here, $I$ and $J$ are fixed numbers. $\Omega_i = \Omega_i^T = \Omega_i^2$ and $\Omega_i \Omega_i = O_{p \times p}$ for $i \neq i'$; and similarly for the $\Pi_j$. 

8
We let \( \hat{X}_{loc}^{(i,j)} \) denote the optimal spectral denoiser with respect to the weight matrices \( \Omega_i \) and \( \Pi_j \). We then define the locally-denoised matrix:
\[
\hat{X}_{loc} = \sum_{i=1}^{I} \sum_{j=1}^{J} \Omega_i \hat{X}_{loc}^{(i,j)} \Pi_j.
\]

(40)

We summarize the localized denoising procedure in Algorithm 2.

**Algorithm 2** Localized denoising

1: Input: \( Y \); pairwise orthogonal projections \( \Omega_1, \ldots, \Omega_I \), \( \Pi_1, \ldots, \Pi_J \)
2: for \( 1 \leq i \leq I, 1 \leq j \leq J \):
   \( \hat{X}_{loc}^{(i,j)} \) is output of Algorithm 1 with weights \( \Omega_i \) and \( \Pi_j \)
   AMSE_{loc}^{(i,j)} is estimated mean squared error
3: \( \hat{X}_{loc} = \sum_{i=1}^{I} \sum_{j=1}^{J} \Omega_i \hat{X}_{loc}^{(i,j)} \Pi_j \)
4: AMSE_{loc} = \sum_{i=1}^{I} \sum_{j=1}^{J} \text{AMSE}_{loc}^{(i,j)}

6.2 Performance of localized denoising

We prove the following results, which compare the behavior of the localized denoiser \( \hat{X}_{loc} \) to the optimal singular value shrinker \( \hat{X}_{shr} \). We recall from Section 2.2 that \( u_k \) is *generic* with respect to \( \Omega_i \) if \( u_k^T \Omega_i u_k \sim \text{tr}(\Omega_i)/p \), and \( u_k \) is *heterogeneous* with respect to \( \Omega_i \) if it is not generic; similarly for \( v_k \) and \( \Pi_j \).

In both Theorem 6.1 and Theorem 6.2, we suppose \( \Omega_1, \ldots, \Omega_I \) and \( \Pi_1, \ldots, \Pi_J \) are pairwise orthogonal projections satisfying \( I_p = \sum_{i=1}^{I} \Omega_i \) and \( I_n = \sum_{j=1}^{J} \Pi_j \).

**Theorem 6.1.** The AMSE of \( \hat{X}_{loc} \) does not exceed the AMSE of \( \hat{X}_{shr} \):
\[
\| \hat{X}_{loc} - X \|_F^2 \leq \| \hat{X}_{shr} - X \|_F^2,
\]

where the inequality holds almost surely as \( n \to \infty \).

**Theorem 6.2.** Suppose that either \( u_1, \ldots, u_r \) are weighted orthogonal with respect to all \( \Omega_i \), or \( v_1, \ldots, v_r \) are weighted orthogonal with respect to all \( \Pi_j \). If some \( u_k \) is heterogeneous with respect to some \( \Omega_i \), or some \( v_k \) is heterogeneous with respect to some \( \Pi_j \), then the AMSE of \( \hat{X}_{loc} \) is strictly smaller than the AMSE of \( \hat{X}_{shr} \):
\[
\| \hat{X}_{loc} - X \|_F^2 < \| \hat{X}_{shr} - X \|_F^2,
\]

where the strict inequality holds almost surely as \( n \to \infty \).

The proofs of Theorems 6.1 and 6.2 are found in Appendix F and Appendix G, respectively.

**Remark 11.** As stated in Remark 6, the weighted orthogonality condition required by Theorem 6.2 will hold if, for example, the columns of \( X \) are drawn iid from some distribution in \( \mathbb{R}^p \), as is often the case in applications.

7 Applications of weighted denoising

In this section, we will describe three examples in which weighted loss functions naturally arise: submatrix estimation, estimation with heteroscedastic noise, and estimation with missing data. In all three problems, we wish to estimate a low-rank matrix with respect to *unweighted* Frobenius loss, but other aspects of the statistical model lead us to perform denoising with respect to a weighted loss function. That is, even when our ultimate goal is unweighted estimation, an intermediate step of the estimation procedure requires the use of a weighted loss function.
7.1 Submatrix denoising

We suppose we observe a data matrix $Y = X + G$. However, instead of recovering the entire matrix $X$, our goal is to estimate only a $p_0$-by-$n_0$ submatrix of $X$, where $p_0/p \sim \mu$ and $n_0/n \sim \nu$. Denoting by $\Omega \in \mathbb{R}^{p_0 \times p}$ the rectangular coordinate selection operator for the $p_0$ rows of $X_0$, and $\Pi \in \mathbb{R}^{n_0 \times n}$ the rectangular coordinate selection operator for the $n_0$ rows of $X_0$, our goal is to estimate the submatrix $X_0 = \Omega X \Pi^T$.

One approach to this problem is to use weighted estimation procedure, estimating the entire matrix $X$ with respect to the weighted loss

$$L(\hat{X}, X) = \|\Omega (\hat{X} - X) \Pi^T\|_F^2.$$  

This loss function only penalizes errors in the $p_0$ rows and $n_0$ columns of interest. We let $\hat{X}$ denote the optimal spectral denoiser, which minimizes $L(\hat{X}, X)$. We then define our estimator of $X_0$ to be $\hat{X}_0 = \Omega \hat{X} \Pi^T$.

Another natural approach is to simply ignore the $p - p_0$ rows and $n - n_0$ columns outside of $X_0$, and denoise $X_0$ directly by optimal singular value shrinkage to the matrix $Y_0 = X_0 + G_0$ (defining $G_0 = \Omega G \Pi^T$). We let $\hat{X}_0^{shr}$ denote this estimator.

We will prove a result comparing these two estimators. Define $\alpha_k$ and $\beta_k$, $1 \leq k \leq r$, as in (6) and (7), respectively. Note that $p_0 = \text{tr}(\Omega^T \Omega)$, and $n_0 = \text{tr}(\Pi^T \Pi)$.

**Proposition 7.1.** Suppose $u_1, \ldots, u_r$ are weighted orthogonal with respect to $\Omega^T \Omega$, and $v_1, \ldots, v_r$ are weighted orthogonal with respect to $\Pi^T \Pi$. Suppose that for $1 \leq k \leq r$,

$$\alpha_k < \sqrt{\mu}, \quad \beta_k < \sqrt{\nu}.$$  

(44)

Then the AMSE of $\hat{X}_0$ is less than that of $\hat{X}_0^{shr}$; that is, 

$$\|\hat{X}_0 - X_0\|_F^2 < \|\hat{X}_0^{shr} - X_0\|_F^2,$$  

(45)

where the strict inequality holds almost surely in the limit $n \to \infty$.

The proof of Proposition 7.1 is found in Appendix H.

**Remark 12.** If the signal vectors $u_k$ and $v_k$ are generic with respect to $\Omega^T \Omega$ and $\Pi^T \Pi$, respectively, then $\alpha_k = \mu$ and $\beta_k = \nu$. By contrast, Proposition 7.1 requires the much weaker condition that $\alpha_k < \sqrt{\mu}$ and $\beta_k < \sqrt{\nu}$ (note that $\mu < \sqrt{\mu}$ and $\nu < \sqrt{\nu}$). Informally, even if the fraction of the signal’s energy contained in $X_0$ is disproportionately large, it still pays to denoise $X_0$ using the entire observed matrix $Y$, rather than the submatrix $Y_0$ alone. In other words, the $p - p_0$ rows and $n - n_0$ columns of $Y$ lying outside $Y_0$ serve to boost the signal in the submatrix $Y_0$.

**Remark 13.** It will follow from the proof of Proposition 7.1 that if $\alpha_k < \sqrt{\mu}$ and $\beta_k < \sqrt{\nu}$, then the singular vectors of the submatrix $X_0$ are better approximated by computing the singular vectors of the full matrix $Y$ and projecting onto the images of $\Omega$ and $\Pi$, respectively, rather than computing the singular vectors of the submatrix $Y_0$ itself. More precisely, if we define

$$u_k^0 = \frac{\Omega u_k}{\|\Omega u_k\|}, \quad v_k^0 = \frac{\Pi v_k}{\|\Pi v_k\|}, \quad 1 \leq k \leq r,$$  

(46)

then we will show that $u_k^0$ and $v_k^0$ are the singular vectors of $X_0$; and the vectors

$$\hat{u}_k^0 = \frac{\Omega \hat{u}_k}{\|\Omega \hat{u}_k\|}, \quad \hat{v}_k^0 = \frac{\Pi \hat{v}_k}{\|\Pi \hat{v}_k\|}, \quad 1 \leq k \leq r,$$  

(47)

are better correlated with $u_k^0$ and $v_k^0$, respectively, then are the singular vectors of $Y_0$. 


7.2 Doubly-heteroscedastic noise

We suppose we wish to estimate a low-rank matrix $X$ from an observed matrix $Y = X + N$, where $N$ is a heteroscedastic noise matrix of the form $N = A^{1/2}GB^{1/2}$, where $G$ has iid entries with distribution $N(0, 1/n)$, and $A \in \mathbb{R}^{p \times p}$ and $B \in \mathbb{R}^{n \times n}$ are diagonal matrices with positive diagonal entries.

We consider the following three-step procedure. First, we whiten the noise, replacing $Y$ by $\tilde{Y}$ defined by

$$\tilde{Y} = A^{-1/2}YB^{-1/2}. \tag{48}$$

We may write $\tilde{Y} = \tilde{X} + G$, where $\tilde{X} = A^{-1/2}XB^{-1/2}$ and $G$ has iid $N(0, 1/n)$ entries. Next, we apply some denoiser to $\tilde{Y}$ to estimate $\tilde{X}$; we denote this by $\psi(\tilde{Y})$, for $\psi$ tailored to removing white noise. Finally, we unwhiten $\psi(\tilde{Y})$ to obtain our final estimate $\hat{X} = A^{1/2}\psi(\tilde{Y})B^{1/2}$ of $X$.

The Frobenius loss between $\tilde{X}$ and $X$ may be written as follows:

$$\|\tilde{X} - X\|_F^2 = \|A^{1/2}\psi(\tilde{Y})B^{1/2} - A^{1/2}\tilde{X}B^{1/2}\|_F^2 = \|A^{1/2}(\psi(\tilde{Y}) - \tilde{X})B^{1/2}\|_F^2, \tag{49}$$

which is a weighted loss between $\psi(\tilde{Y})$ and $\tilde{X}$, with weight matrices $A^{1/2}$ and $B^{1/2}$. Consequently, the denoiser $\psi(\tilde{Y})$ should be chosen to minimize this weighted Frobenius loss. That is, if our goal is to estimate $X$ in Frobenius loss, then the denoiser we apply to $\tilde{Y}$ should not minimize the Frobenius loss with $\tilde{X}$, but rather the weighted Frobenius loss.

Remark 14. Doubly-heteroscedastic noise arises when we observe independently-drawn vectors of the form $y_j = x_j + b_j^{1/2} \varepsilon_j$, where

$$x_j = \sum_{k=1}^{r} t_k^{1/2} z_{jk} u_k \tag{50}$$

where $\varepsilon_i \sim N(0, A)$, and the $z_{jk}$ are independent random variables. Each observation has heteroscedastic noise, and also a different total noise level. In the case where $b_j = 1$ for all $j$, i.e. $B = I_n$, this is the model considered in [43] and [34]; here, we extend the model to permit each observation to have differing overall noise strength, which is also an extension of the model in [22, 23, 24] (where $A = I_p$). We note too that random matrices with a rank 1 variance profile have also been studied in [19, 20, 33].

Remark 15. The procedure of whitening, denoising, and unwhitening has been employed in a number of recent papers on the spiked model; see, for instance, [35, 34, 14]. In particular, [34] shows several advantages of working with the whitened matrix when the noise is one-sided, such as improved estimation of the singular vectors of $X$. In Section 7.2.1, namely Proposition 7.2, we show another benefit of noise whitening for doubly-heteroscedastic noise.

7.2.1 Whitening increases the SNR for generic signal matrices

We show that the whitening transformation increases a natural signal-to-noise ratio. We will assume throughout this section that the $u_k$ (respectively, $v_k$) are generic with respect to $A$ (respectively, $B$), and that they satisfy the pairwise orthogonality condition with respect to $A$ (respectively $B$). Writing the SVD of $X$ as

$$X = \sum_{k=1}^{r} t_k u_k v_k^T \tag{51}$$

we follow [34] and define the operator norm signal-to-noise ratio (SNR) for each component of the signal matrix $X$ by:

$$\text{SNR}_{\text{op}}^{(k)} = \frac{t_k^2}{\|N\|_{\text{op}}^2}, \tag{52}$$

which is the asymptotic ratio of the squared operator norm of each component $t_k^{1/2} u_k v_k^T$ of $X$ and the squared operator norm of the noise. $\text{SNR}_{\text{op}}^{(k)}$ measures the strength of the $k^{th}$ component relative to the strength of the noise matrix $N$. 

11
After whitening the noise matrix, the observation changes into \( \tilde{Y} = \tilde{X} + \tilde{G} \), where \( \tilde{X} \) is still rank \( r \), and in 1-1 correspondence with \( X \). We may write

\[
\tilde{X} = \sum_{k=1}^{r} t_k (A^{-1/2}u_k)(B^{-1/2}v_k)^T = \sum_{k=1}^{r} \tilde{t}_k \tilde{u}_k \tilde{v}_k^T,
\]

(53)

where \( \tilde{t}_k = t_k\|A^{-1/2}u_k\|\|B^{-1/2}v_k\|, \tilde{u}_k = A^{-1/2}u_k/\|A^{-1/2}u_k\|, \) and \( \tilde{v}_k = B^{-1/2}v_k/\|B^{-1/2}v_k\| \). Consequently, the operator norm SNR after whitening is:

\[
\overline{SNR}_{\text{op}}^{(k)} = \frac{\tilde{t}_k^2}{\|G\|_{\text{op}}^2}.
\]

(54)

We will prove the following result, which is a simple extension of one in [34] to the setting of doubly-heteroscedastic noise:

**Proposition 7.2.** Suppose \( A \) and \( B \) are not both equal to constant multiples of the identity. Then

\[
\overline{SNR}_{\text{op}}^{(k)} > SNR_{\text{op}}^{(k)}, \quad 1 \leq k \leq r,
\]

(55)

where the strict inequalities holds almost surely as \( n \to \infty \).

In other words, if the signal is generic, then the SNR increases after whitening the noise; energy is transferred from the noise component to the signal component. The proof of Proposition 7.2 is an extension of the analogous result from [34]. We provide a proof in Appendix I.

### 7.3 Matrices with missing/unobserved values

We consider the setting where \( X \) is a low-rank target matrix we wish to recover and \( G \) is a matrix of iid Gaussian \( N(0,1) \) entries, but rather than observe \( X + G \), we observe only some subset of the entries. The problem of estimating a matrix from a subset of its entries is known as matrix completion; it has been extensively studied, in both the noise-free and high-noise regimes [30, 28, 10, 38, 29, 31, 25, 11, 14].

In this section, we will adopt a heterogeneous, rank 1 sampling model, which is considered in [12]. Specifically, we suppose that the rows and columns are sampled independently, with row \( i \) sampled with probability \( q_i^r \), and column \( j \) sampled with probability \( q_j^c \); that is, entry \((i,j)\) of \( X + G \) is sampled with probability \( q_i^r q_j^c \). We observe \( F(X + G) \), where \( F : \mathbb{R}^{p \times n} \to \mathbb{R}^M \) is the subsampling operator, with \( M \) being the number of sampled entries.

Following the approach from [14], we consider the backprojected matrix \( Y = F^*(F(X + G)) \in \mathbb{R}^{p \times n} \), in which the unobserved entries are replaced by 0’s. We write \( Y = F^*(F(X)) + F^*(F(G)) \). We show that asymptotically, \( F^*(F(X)) \) behaves like the matrix \( PXQ \). More precisely, we have the following result:

**Proposition 7.3.** Suppose \( \max_{1 \leq k \leq r} \|u_k\|_\infty \|v_k\|_\infty = o(n^{-1/2}) \). Then in the limit \( p/n \to \gamma \),

\[
\|F^*(F(X)) - PXQ\|_{\text{op}} \to 0
\]

(56)

almost surely.

The proof of Proposition 7.3 may be found in Appendix J. It is a straightforward generalization of the analogous one-sided result in [14].

Now, let \( N = F^*(F(G)) \). Writing \( N = (N_{ij}) \), we have \( N_{ij} = \delta_{ij}G_{ij} \), where \( \delta_{ij} \) is 1 if entry \((i,j)\) is sampled, and 0 otherwise. Then \( N_{ij} \) has variance \( q_i^r q_j^c \). Consequently, we can whiten the noise matrix by applying \( P^{-1/2} \) and \( Q^{-1/2} \); Proposition 7.2 suggests this will improve estimation of the matrix. To that end, we define

\[
\tilde{Y} = P^{-1/2}YQ^{-1/2} = \tilde{X} + \tilde{G},
\]

(57)

where \( \tilde{X} = P^{-1/2}F^*(F(X))Q^{-1/2} \), and \( \tilde{G} = P^{-1/2}NQ^{-1/2} \). Then \( \tilde{G} \) is a random matrix where each entry has mean zero and variance 1.
From Proposition 7.3, asymptotically the matrix \( \tilde{X} \) behaves like \( P^{1/2}XQ^{1/2} \). Consequently, denoising \( \tilde{Y} \) produces an estimate of \( P^{1/2}XQ^{1/2} \), not \( X \) itself. Therefore, to estimate \( X \) we should perform denoising to \( \tilde{Y} \) with respect to the weighted loss function

\[
L(\hat{X}, X) = \|P^{-1/2}(\hat{X} - X)Q^{-1/2}\|_F^2,
\]

with weight matrices \( P^{-1/2} \) and \( Q^{-1/2} \). We then apply \( P^{-1/2} \) and \( Q^{-1/2} \) to the resulting matrix, to obtain an estimator of \( X \) itself. We are therefore led naturally to the use of a weighted loss function.

\section{Numerical results}

We report on numerical simulations demonstrating the performance of the algorithms from this paper. In Section 8.1, we compare localized denoising to global singular value shrinkage. In Section 8.2, we report on spectral denoising for submatrix denoising, as described in Section 7.1. In Section 8.3, we report on spectral denoising for doubly-heteroscedastic noise, as described in Section 7.2. In Section 8.4, we report on spectral denoising method for missing data, as described in Section 7.3.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{localized_denoising_vs_global_shrinkage.png}
\caption{Localized denoising versus singular value shrinkage.}
\end{figure}

\subsection{Localized denoising}

We evaluate the performance of localized denoising, the method described in Section 6. For comparison, we also apply optimal singular value shrinkage to the entire matrix \( X \). Theorem 6.2 predicts that localized denoising will outperform shrinkage when the matrix \( X \) is heterogeneous, i.e. when the energy in the singular vectors is not uniformly distributed across the coordinates.

We ran the following experiment. We generated a rank 1 signal matrix \( X \) of size 1000-by-2000. The singular value of \( X \) is taken to be \( \sqrt{1/2} + 5 \). The left singular vector is piecewise constant on the first 500 and last 500 coordinates, and the right singular vector is piecewise constant on the first 1000 and last 1000 coordinates. The values of the singular vectors are chosen so that the upper left 500-by-1000 submatrix \( X_0 \) contains a specified fraction \( f \in (0, 1) \) of the total energy of \( X \); that is, \( f = \|X_0\|_F^2/\|X\|_F^2 \).

For each value of \( f \), we repeat the experiment 50 times, and for each method the errors are averaged over these 50 runs. In Figure 3, we plot the logarithm of the MSE against \( f \). For all values of \( f \) outside of 0.25, spectral denoising outperforms shrinkage, because the matrix is heterogenous with respect to the coordinate projections used in localized denoising. At \( f \approx 0.25 \), the two methods perform approximately the same, since the energy is uniformly distributed throughout all coordinates of the matrix. This is consistent with the theory from Section 6.
8.2 Submatrix denoising

We evaluate the performance of spectral denoising for estimating a submatrix $X_0$ contained within a larger matrix $X$. We employ spectral denoising with coordinate-selection weight matrices, as described in Section 7.1. For comparison, we also apply optimal singular value shrinkage to the submatrix $X_0$ alone.

We ran the following experiment, whose setup is similar to that from Section 8.1. We generated a rank 1 signal matrix $X$ of size 1000-by-2000. The singular value of $X$ is taken to be $\sqrt{1/2} + 2$. The left singular vector is piecewise constant on the first 500 and last 500 coordinates, and the right singular vector is piecewise constant on the first 1000 and last 1000 coordinates. The values of the singular vectors are chosen so that the upper left 500-by-1000 submatrix $X_0$ contains a specified fraction $f \in (0, 1)$ of the total energy of $X$; that is, $f = \|X_0\|_F^2/\|X\|_F^2$.

For each value of $f$, we repeat the experiment 50 times, and for each method the errors are averaged over these 50 runs. In Figure 3, we plot the logarithm of the MSE against $f$. For most values of $f$, spectral denoising outperforms singular value shrinkage. At a certain large value of $f$, the submatrix $X_0$ contains a sufficiently large fraction of the energy that spectral denoising performs worse than singular value shrinkage. In other words, when $X_0$ contains almost all of the energy of the signal, it is harmful rather than helpful to use the parts of the matrix outside $X_0$.

Figure 3: Optimal spectral denoising versus singular value shrinkage on the submatrix.

Figure 4: Optimal singular value denoising with bi-whitening versus OptShrink.
8.3 Doubly-heteroscedastic noise

We evaluate the performance of spectral denoising for denoising matrices with doubly-heteroscedastic noise. We employ the whitening procedure described in Section 7.2. We ran the following experiment. We generated a rank 5 signal matrix $X$ of size 500-by-1000. The singular values of $X$ are taken to be $\sqrt{1/2 + k}$, $k = 1, \ldots, 5$. Both the left and right singular vectors of $X$ are chosen uniformly randomly.

For a specified $\kappa \geq 1$, we generated row and column diagonal covariance matrices $A$ and $B$, each with condition number $\kappa$, and with linearly-spaced eigenvalues. The eigenvalues of $A$ and $B$ are normalized to sum to 1, which ensures that the total energy of the noise is constant. We then generated the noise matrix $A^{1/2}GB^{1/2}$, where $G$ has iid Gaussian entries. We observe the matrix $X + A^{1/2}GB^{1/2}$.

For each draw of this data, we perform the denoising scheme based on whitening, weighted spectral denoising, and unwhitening, as described in Section 7.2. For comparison, we also apply the OptShrink algorithm [36], which optimally denoises singular values for colored noise matrices (without whitening the noise). The experiment is repeated 50 times for each value of $\kappa$, and the errors averaged over these 50 runs.

In Figure 4, we plot the errors of these two schemes as a function of the condition number $\kappa$. Whereas OptShrink’s performance is unaffected by the heteroscedasticity (because the energy of the noise is constant), the performance of spectral denoising with whitening improves as $\kappa$ increases. For all values of $\kappa$, spectral denoising with whitening outperforms OptShrink.

![Figure 5: Optimal spectral denoising versus nuclear norm regularized least squares.](image)

8.4 Missing data

We test the performance of singular value denoising on a missing data problem, following the procedure described in Section 7.3. We compare our method to nuclear-norm regularized least-squares [10], which estimates $X$ by:

$$\hat{X}_{\text{nuc}} = \arg\min_{\hat{X} \in \mathbb{R}^{p \times n}} \left\{ \frac{1}{2} \| \mathcal{F}(\hat{X}) - y \|^2 + \theta \| P^{1/2} \hat{X}Q^{1/2} \|_* \right\}. \quad (59)$$

Here, $\| \cdot \|_*$ denotes the nuclear norm; $\mathcal{F}: \mathbb{R}^{p \times n} \to \mathbb{R}^M$ is the projection operator onto the $M$ observed samples; $P$ and $Q$ are the diagonal matrices of sampling probabilities for rows and columns, respectively; and $y \in \mathbb{R}^M$ is the vector of observed entries. We weight the nuclear norm by the square root of the sampling probabilities, as is suggested in [12]. Following [10], we choose the parameter $\theta$ so that when $y$ is pure noise, $\hat{X}_{\text{nuc}}$ is set to zero. It follows from the KKT conditions for the problem [9] that this is equivalent to:

$$\theta = \| P^{-1/2} \mathcal{F}^*(y)Q^{-1/2} \|_* \quad (60)$$
Since the matrix $P^{-1/2}F^*(y)Q^{-1/2}$ has independent entries with variance 1, we may set $\theta = 1 + \sqrt{\gamma}$. We solve the minimization (59) using the algorithm in [26].

We generate a rank 5 signal matrix $X$ of size 200-by-400, with singular values $\sqrt{1/2 + 200k}$, $k = 1, \ldots, 5$. Both the left and right singular vectors of $X$ are chosen uniformly randomly. We add to $X$ a Gaussian noise matrix $G$, where each entry has variance $\sigma^2$ for a specified value of $\sigma$. The matrix $X + G$ is then subsampled using row and column sampling probabilities each equispaced between 0.3 and 0.7.

For each value of the noise level $\sigma$, we repeat the experiment 50 times and average the errors. In Figure 5, we plot the logarithm of the MSE against $\log(\sigma)$. When $\sigma$ is large, spectral denoising outperforms nuclear-norm regularized least-squares, whereas in the small $\sigma$ regime nuclear-norm regularized least-squares is the better method. This is consistent with the experimental results from [14].

**Remark 16.** The error curve for the spectral denoiser in Figure 5 has a slight “dent” at large $\sigma$. This is due to the instability of rank estimation in low SNR. In our experiment, we keep the components whose singular values exceed the bulk edge $1 + \sqrt{\gamma}$ plus a small correction factor $\epsilon$; in future work a more principled choice of rank estimation might be employed, such as an adaptation of the method in [32] to non-Gaussian noise. We note that rank estimation in spiked models has been the subject of recent investigations [15, 13].

9 Conclusion

This paper has introduced a family of spectral denoisers for low-rank matrix estimation, which generalizes the standard method of singular value shrinkage. We have derived optimal spectral denoisers for weighted loss functions. By judiciously combining these denoisers for different weights we contructed the method of localized denoising. Localized denoising outperforms singular value shrinkage when the signal matrix is heterogeneous. We also introduced applications of weighted loss functions to heteroscedastic noise, missing data, and submatrix denoising.

The theory we have developed holds for Gaussian noise matrices $G$. However, one of the proposed applications of our theory is to problems with missing data — and more generally, the linearly-transformed spiked model described in [14] — for which the effective noise term is typically not Gaussian. While numerical simulations show agreement with the predicted theory even for non-Gaussian noise, an open question is to extend our theoretical results to these more general noise models. This will allow us to justify the use of our denoisers for the linearly-transformed spiked model.

The current paper has focused on theoretical and algorithmic development. In future work, we plan to apply optimal spectral denoising to problems where related but suboptimal methods have previously been employed. This includes the problems of denoising and deconvolution of images from cryoelectron microscopy [8]; 3D reconstruction of heterogeneous molecules from noisy images [1]; and denoising XFEL images [35, 44].

Acknowledgements

I thank Elad Romanov and Amit Singer for stimulating discussions related to this work, and Edgar Dobriban for valuable feedback on an earlier version of this manuscript. I acknowledge support from the NSF BIGDATA program IIS 1837992 and BSF award 2018230.

References


A Proof of Theorem 3.2

The proof of Theorem 3.2 is similar to the analysis found in [34], in that it rests on the same decomposition of the empirical singular vectors $\mathbf{u}_j$ and $\mathbf{v}_j$ into the signal and residual components. If $\mathbf{a}$ and $\mathbf{b}$ are vectors of the same dimension, we will write $\mathbf{a} \sim \mathbf{b}$ as a short-hand for $\|\mathbf{a} - \mathbf{b}\| \to 0$ almost surely as $p, n \to \infty$. The statements are symmetric in the left and right singular vectors, so for compactness we will only prove them for the left ones. The proofs for the other side are identical.

Because the noise matrix $\mathbf{G}$ has an isotropic distribution, we can write:

$$\hat{\mathbf{u}}_j \sim c_j \mathbf{u}_j + s_j \tilde{\mathbf{u}}_j,$$

where $\hat{\mathbf{u}}_j$ is a unit vector that is uniformly random over the sphere in the subspace orthogonal to $\mathbf{u}_1, \ldots, \mathbf{u}_r$ (see [37]). Because $\tilde{\mathbf{u}}_j$ is uniformly random, it is asymptotically orthogonal to any independent unit vector $\mathbf{w}$; that is,

$$\tilde{\mathbf{u}}_j^T \mathbf{w} \sim 0.$$

Furthermore, $\tilde{\mathbf{u}}_j$ satisfies the normalized trace formula, namely if $\mathbf{A}$ is any matrix with bounded operator norm, then

$$\tilde{\mathbf{u}}_j^T \mathbf{A} \tilde{\mathbf{u}}_j \sim \frac{1}{p} \text{tr} (\mathbf{A}).$$

We refer the reader to [6, 21, 42, 39] for details. We will use (62) and (63) repeatedly. Furthermore, when $j \neq k$ it follows from Lemma A.2 in [34] that

$$\tilde{\mathbf{u}}_j^T \mathbf{A} \tilde{\mathbf{u}}_k \sim 0.$$

Applying $\Omega$ to each side of (61), we have:

$$\Omega \hat{\mathbf{u}}_j \sim c_j \Omega \mathbf{u}_j + s_j \Omega \tilde{\mathbf{u}}_j,$$

The proofs of the identities in Theorem 3.2 now follow by manipulating the asymptotic equation (65) appropriately, in conjunction with (62), (63) and (64).

We first show the formulas for $c_{jk}$. We take inner products of each side of (65) with $\Omega \mathbf{u}_k$:

$$c_{jk} \sim \langle \Omega \hat{\mathbf{u}}_j, \Omega \mathbf{u}_k \rangle \sim c_j \langle \Omega \mathbf{u}_j, \Omega \mathbf{u}_k \rangle + s_j \langle \Omega \hat{\mathbf{u}}_j, \Omega \mathbf{u}_k \rangle \sim c_j \langle \Omega \mathbf{u}_j, \Omega \mathbf{u}_k \rangle \sim c_j e_{jk},$$

where we have used (62).

To derive the formula for $d_j$, we take the squared norm of each side of (65):

$$d_j \sim \|\Omega \hat{\mathbf{u}}_j\|^2 \sim c_j^2 \|\Omega \mathbf{u}_j\|^2 + s_j^2 \|\Omega \tilde{\mathbf{u}}_j\|^2 \sim c_j^2 \alpha_j + s_j^2 \mu.$$

The first asymptotic equivalence follows from (62), and the second from (63).

Finally, we derive the formula for $d_{jk}$, $j \neq k$. From (65), we have

$$\langle \Omega \hat{\mathbf{u}}_j, \Omega \hat{\mathbf{u}}_k \rangle \sim c_j c_k \langle \Omega \mathbf{u}_j, \Omega \mathbf{u}_k \rangle + s_j s_k \langle \Omega \hat{\mathbf{u}}_j, \Omega \mathbf{u}_k \rangle + s_j c_k \langle \Omega \hat{\mathbf{u}}_j, \Omega \hat{\mathbf{u}}_k \rangle + c_j s_k \langle \Omega \mathbf{u}_j, \Omega \mathbf{u}_k \rangle.$$

From (62) and (64), the terms involving $\hat{\mathbf{u}}_j$ and $\tilde{\mathbf{u}}_k$ vanish, and we are left with

$$d_{jk} \sim \langle \Omega \hat{\mathbf{u}}_j, \Omega \mathbf{u}_k \rangle \sim c_j c_k \langle \Omega \mathbf{u}_j, \Omega \mathbf{u}_k \rangle \sim c_j c_k e_{jk}.$$

This completes the proof of Theorem 3.2.
\section*{B Proof of Theorem 4.1}

The target matrix $X$ may be written

$$X = \sum_{k=1}^{r} t_k u_k v_k^T = U \text{diag}(t) V^T,$$

and our estimate $\hat{X}$ is of the form

$$\hat{X} = \hat{U}\hat{V} \hat{V}^T,$$

where $U = [u_1, \ldots, u_r]$, $V = [v_1, \ldots, v_r]$, $\hat{U} = [\hat{u}_1, \ldots, \hat{u}_r]$, $\hat{V} = [\hat{v}_1, \ldots, \hat{v}_r]$, and $t = (t_1, \ldots, t_r)^T$.

Define $W = \Omega U$, $Z = IV$, $\hat{W} = \Omega \hat{U}$, and $\hat{Z} = IV$. We may then write the weighted loss as follows:

$$L(\hat{X}, X) = \|\Omega(\hat{X} - X)\Pi T\|^2_F = \|\Omega\hat{X}\Pi T - \Omega X\Pi T\|^2_F = \|\hat{W}\hat{B}\hat{Z}^T - W\text{diag}(t)Z^T\|^2_F,$$

which is the \textit{unweighted} Frobenius loss between $\hat{W}\hat{B}\hat{Z}^T$ and $W\text{diag}(t)Z^T$. Continuing, we have:

$$L(\hat{X}, X) = \|W\text{diag}(t)Z^T - \hat{W}\hat{B}\hat{Z}^T\|_F^2 = \|W\text{diag}(t)Z^T\|_F^2 + \|\hat{W}\hat{B}\hat{Z}^T\|_F^2 - 2\langle W\text{diag}(t)Z^T, \hat{W}\hat{B}\hat{Z}^T \rangle_F$$

$$= \langle W^T W\text{diag}(t)Z, \text{diag}(t) \rangle_F + \langle \hat{W}^T \hat{W}\hat{B}\hat{Z}^T \hat{Z}, \hat{B} \rangle_F - 2\langle \hat{W}^T W\text{diag}(t)Z^T \hat{Z}, \hat{B} \rangle_F$$

$$\sim \langle E\text{diag}(t)E, \text{diag}(t) \rangle_F + \langle \hat{D}\hat{D}, \hat{D} \rangle_F - 2\langle C\text{diag}(t)C^T, \hat{C} \rangle_F. \quad (73)$$

Defining the operator $\mathcal{T}$ by $\mathcal{T}(\hat{B}) = \hat{D}\hat{B}\hat{D}^+$, the pseudoinverse of $\mathcal{T}$ is given by $\mathcal{T}^+(B) = D^+BD^+$. Consequently, the choice of $\hat{B}$ that minimizes $L(\hat{X}, X)$ is given by:

$$\hat{B} = D^+ C\text{diag}(t)C^T \hat{D}^+. \quad (74)$$

The error may then be evaluated by substituting this expression for $\hat{B}$ into (73), completing the proof.

\section*{C Proof of Theorem 5.1}

Under weighted orthogonality, $e_{jk} = \hat{e}_{jk} = 0$ whenever $j \neq k$, and so $d_{jk} = \hat{d}_{jk} = \hat{e}_{jk} = 0$ when $j \neq k$ as well. Consequently, the matrices $E$, $\hat{E}$, $D$, $\hat{D}$, $C$, and $\hat{C}$ are diagonal. The optimal $\hat{B}$ is given by:

$$\hat{B} = D^+ C\text{diag}(t)C^T \hat{D}^+, \quad (75)$$

which is also diagonal, with diagonal entries

$$\hat{t}_k = \frac{t_k c_k \hat{e}_k^2}{d_k d_k} = \frac{t_k c_k \hat{e}_k \hat{e}_k}{(c_k^2 \alpha_k + s_k^2 \mu)(c_k^2 \beta_k + s_k^2 \nu)} = t_k \frac{c_k \hat{e}_k}{(c_k^2 \alpha_k + s_k^2 \mu)(c_k^2 \beta_k + s_k^2 \nu)}, \quad (76)$$

which is the desired expression.

\section*{D Proof of Proposition 5.2}

Suppose a coordinate has signal strength $t = t_k$ (we drop the subscript as we are only considering one component). We may assume without loss of generality (and by rescaling $\alpha$ and $\beta$) that $\mu = \nu = 1$. Consequently, the optimal singular value is equal to:

$$\hat{t} = tc\hat{e} \cdot \frac{\alpha}{c^2 \alpha + s^2} \cdot \frac{\beta}{c^2 \beta + s^2}, \quad (77)$$

20
By taking $\alpha$ and $\beta$ sufficiently large, this value can be made arbitrarily close to
\[
\frac{t}{\tilde{c}} = \frac{t\sqrt{(1 + \gamma/t^2)(1 + 1/t^2)}}{1 - \gamma/t^4} = \frac{\lambda}{1 - \gamma/t^4} > \lambda.
\] (78)

That is, the optimal singular value $\hat{t}$ will be greater than the observed singular value $\lambda$ in this parameter regime.

On the other hand, if $\beta \leq 1 = \nu$, we have:
\[
\hat{t} = \frac{1}{\tilde{c}} \frac{t}{c} \frac{\alpha}{\alpha c^2 + s^2} \frac{\beta}{\beta \tilde{c}^2 + s^2} \leq \frac{1}{\lambda} \frac{t}{c} \frac{\sqrt{t^2 + 1}((t^2 + \gamma) t^2 + 1)}{t^2 + 1} \leq 1,
\] (79)

which shows that $\hat{t} \leq \lambda$. A nearly identical proof works if $\alpha \leq \mu$. This completes the proof.

E Proof of Proposition 5.3

Without loss of generality, we will assume $\mu = \nu = 1$. We consider the functions $c(t) = \sqrt{(1 - \gamma/t^4)/(1 + \gamma/t^2)}$ and $\tilde{c}(t) = \sqrt{(1 - \gamma/t^4)/(1 + 1/t^2)}$. Define the functions $\varphi(t)$ and $\psi(t)$ by
\[
\varphi(t) = \frac{\alpha c(t)}{\alpha c(t)^2 + 1 - c(t)^2}
\] (80)
and
\[
\psi(t) = \frac{\beta \tilde{c}(t)}{\beta \tilde{c}(t)^2 + 1 - \tilde{c}(t)^2}.
\] (81)

Then we may write the optimal singular value $\hat{t}$ as a function $f(t)$ as follows:
\[
f(t) = t\varphi(t)\psi(t).
\] (82)

Let us assume that $\alpha \leq 1$; the proof for $\beta \leq 1$ will be nearly identical. We wish to show that $f'(t) \geq 0$, for $t > \gamma^{1/4}$. We have
\[
\frac{f'(t)}{f(t)} = \frac{\varphi'(t)}{\varphi(t)} + \frac{\psi'(t)}{\psi(t)} + \frac{1}{t},
\] (83)

and since $f(t) > 0$, we must show that the right side is positive. It is straightforward to verify that
\[
\varphi'(t) = \frac{\alpha c(t)[1 - (\alpha - 1)c(t)]}{[1 + (\alpha - 1)c(t)]^2}
\] (84)
from which it follows that
\[
\frac{\varphi'(t)}{\varphi(t)} = \frac{c'(t) 1 - (\alpha - 1)c(t)^2}{c(t) 1 + (\alpha - 1)c(t)^2} \geq \frac{c'(t)}{c(t)}.
\] (85)

Similarly, we can show
\[
\frac{\psi'(t)}{\psi(t)} = \frac{\tilde{c}'(t) 1 - (\beta - 1)\tilde{c}(t)^2}{\tilde{c}(t) 1 + (\beta - 1)\tilde{c}(t)^2} \geq -\frac{\tilde{c}'(t)}{\tilde{c}(t)}.
\] (86)

Consequently, it is enough to show
\[
\frac{c'(t)}{c(t)} - \frac{\tilde{c}'(t)}{\tilde{c}(t)} + \frac{1}{t} \geq 0.
\] (87)
Direction computation shows
\[
\frac{c'(t)}{c(t)} = \frac{\gamma t^4 + 2t^2 + \gamma}{t(t^2 + \gamma)(t^4 - \gamma)}
\] (88)

and
\[
\frac{c'(t)}{c(t)} = \frac{t^4 + 2\gamma t^2 + \gamma}{t(t^2 + 1)(t^4 - \gamma)}. \tag{89}
\]

Substituting (88) and (89) into the left side of (87) and multiplying by \(t(t^2 + \gamma)(t^2 + 1)\), we get:
\[
t(t^2 + \gamma)(t^2 + 1) \left( \frac{c'(t)}{c(t)} - \frac{c'(t)}{c(t)} + \frac{1}{t} \right)
= t(t^2 + \gamma)(t^2 + 1) \left( \frac{t^4 + 2\gamma t^2 + \gamma}{t(t^2 + \gamma)(t^4 - \gamma)} - \frac{t^4 + 2\gamma t^2 + \gamma}{t(t^2 + 1)(t^4 - \gamma)} + \frac{1}{t} \right)
= t^4 + 2\gamma t^2 + \gamma > 0,
\] (90)

which is the desired result.

**F Proof of Theorem 6.1**

We denote by \(\hat{t}_{i}^\text{shr}, \ldots, \hat{t}_{r}^\text{shr}\) the singular values of \(\hat{X}^\text{shr}\), and \(\hat{t}^\text{shr} = (\hat{t}_{1}^\text{shr}, \ldots, \hat{t}_{r}^\text{shr})^T\). We may then write
\[
\hat{X}^\text{shr} = \hat{U}\text{diag}(\hat{t}^\text{shr})\hat{V}^T.
\] (91)

This is a spectral denoiser (in the set \(S\)), and hence its weighted loss with weights \(\Omega_i\) and \(\Pi_j\) cannot be less than that of the optimal spectral denoiser \(\hat{X}^\text{loc}_{(i,j)}\). That is,
\[
\|\Omega_i(\hat{X}^\text{loc}_{(i,j)} - X)\Pi_j^T\|_F^2 \leq \|\Omega_i(\hat{X}^\text{shr} - X)\Pi_j^T\|_F^2
\] (92)

Because the \(\Omega_i\) and \(\Pi_j\) are pairwise orthogonal projections which sum to the identity, the total Frobenius loss can be decomposed:
\[
\|\hat{X}^\text{loc} - X\|_F^2 = \sum_{i=1}^{I} \sum_{j=1}^{J} \|\Omega_i(\hat{X}^\text{loc} - X)\Pi_j^T\|_F^2 = \sum_{i=1}^{I} \sum_{j=1}^{J} \|\Omega_i(\hat{X}^\text{loc}_{(i,j)} - X)\Pi_j^T\|_F^2
\leq \sum_{i=1}^{I} \sum_{j=1}^{J} \|\Omega_i(\hat{X}^\text{shr} - X)\Pi_j^T\|_F^2 = \|\hat{X}^\text{shr} - X\|_F^2,
\] (93)

which is the desired inequality.

**G Proof of Theorem 6.2**

For \(1 \leq k \leq r\), \(1 \leq i \leq I\), and \(1 \leq j \leq J\), let \(\alpha_k^{(i)} = \|\Omega_i u_k\|^2\), \(\mu^{(i)} = \text{tr}(\Omega_i) / p\), \(\beta_k^{(j)} = \|\Pi_j v_k\|^2\), and \(\nu^{(j)} = \text{tr}(\Pi_j) / n\). Then
\[
\sum_{i=1}^{I} \alpha_k^{(i)} = \sum_{i=1}^{I} \mu^{(i)} = \sum_{j=1}^{J} \beta_k^{(j)} = \sum_{j=1}^{J} \nu^{(j)} = 1.
\] (94)

Let \(\hat{X}_{(i,j)}^\text{dd}\) be the optimal diagonal denoiser with weights \(\Omega_i\) and \(\Pi_j\). Because of the weighted orthogonality condition, Theorem 5.1 states that the AMSE for \(\hat{X}_{(i,j)}^\text{dd}\) is
\[
\|\Omega_i(\hat{X}_{(i,j)}^\text{dd} - X)\Pi_j^T\|_F^2 = \sum_{k=1}^{r} c_k^2 \alpha_k^{(i)} \beta_k^{(j)} \left( 1 - c_k^2 c_k^2 \frac{\alpha_k^{(i)}}{s_k^2 \mu^{(i)}} \cdot \frac{\beta_k^{(j)}}{s_k^2 \nu^{(j)}} \right).
\] (95)
Since \( \hat{X}_{(i,j)}^{\text{loc}} \) minimizes the weighted error with weights \( \Omega_i \) and \( \Pi_j \), we have:

\[
\| \hat{X}_{(i,j)}^{\text{loc}} - X \|_F^2 = \sum_{i=1}^{I} \sum_{j=1}^{J} |\Omega_i (\hat{X}_{(i,j)}^{\text{loc}} - X) \Pi_j^T |^2_F = \sum_{i=1}^{I} \sum_{j=1}^{J} |\Omega_i (\hat{X}_{(i,j)}^{\text{loc}} - X) \Pi_j^T |^2_F
\]

\[
\leq \sum_{i=1}^{I} \sum_{j=1}^{J} |\Omega_i (\hat{X}_{(i,j)}^{\text{dd}} - X) \Pi_j^T |^2_F
\]

\[
= \sum_{i=1}^{I} \sum_{j=1}^{J} r^2 \sum_{k=1}^{r} \sum_{i=1}^{J} \sum_{k=1}^{r} \left( \frac{1}{c_k^2} \frac{c_i^2}{c_k^2 c_i^2 + s_k^2 \mu(i)} \cdot \frac{\beta(i) \beta(j)}{c_k^2 \beta_k + s_k^2 \nu(j)} \right).
\]

\[
= \frac{r t_k^2}{c_k^2} \sum_{j=1}^{J} \sum_{i=1}^{I} \sum_{k=1}^{r} \left( \frac{1}{c_i^2} \frac{c_i^2}{c_i^2 \mu(i)} + s_k^2 \mu(i) \right) \cdot \frac{\beta(i) \beta(j)}{c_k^2 \beta_k + s_k^2 \nu(j)}.
\]

(96)

On the other hand, the error obtained by \( \hat{X}^{\text{shr}} \) is equal to

\[
\| \hat{X}^{\text{shr}} - X \|_F^2 = \sum_{k=1}^{I} c_k^2 (1 - c_k^2 c_k^2).
\]

(97)

Comparing (96) and (97), the result will follow if we can show that for each \( 1 \leq k \leq r \),

\[
\sum_{i=1}^{I} \sum_{j=1}^{J} \frac{(\alpha(i))^2}{c_i^2 \alpha(i) + s_k^2 \mu(i)} \cdot \frac{\beta(j)}{c_k^2 \beta_k + s_k^2 \nu(j)} \geq 1,
\]

(98)

where the inequality is strict so long as one of \( u_k \) or \( v_k \) is not generic with respect to some \( \Omega_i \) or \( \Pi_j \); or equivalently, either \( \alpha(i) \neq \mu(i) \) for some \( i \), or \( \beta(j) \neq \nu(j) \) for some \( j \). Because

\[
\sum_{i=1}^{I} \sum_{j=1}^{J} \frac{(\alpha(i))^2}{c_i^2 \alpha(i) + s_k^2 \mu(i)} \cdot \frac{\beta(j)}{c_k^2 \beta_k + s_k^2 \nu(j)} = \left( \sum_{i=1}^{I} \frac{(\alpha(i))^2}{c_i^2 \alpha(i) + s_k^2 \mu(i)} \right) \cdot \left( \sum_{j=1}^{J} \frac{\beta(j)^2}{c_k^2 \beta_k + s_k^2 \nu(j)} \right),
\]

(99)

it is enough to show that

\[
\sum_{i=1}^{I} \frac{(\alpha(i))^2}{c_i^2 \alpha(i) + s_k^2 \mu(i)} \geq 1,
\]

(100)

with the inequality being strict so long as \( \alpha(i) \neq \mu(i) \) for some \( i \).

For each \( 1 \leq i \leq I \), let \( r_i = \alpha(i) / \mu(i) \). Then

\[
\sum_{i=1}^{I} \frac{(\alpha(i))^2}{c_i^2 \alpha(i) + s_k^2 \mu(i)} = \sum_{i=1}^{I} \mu(i) \frac{r_i^2}{c_k^2 r_i + s_k^2}.
\]

(101)

The function \( F(r) = r^2 / (c_k^2 r + s_k^2) \) is convex. Since \( \sum_{i=1}^{I} \mu(i) = 1 \), Jensen’s inequality implies

\[
\sum_{i=1}^{I} \mu(i) \frac{r_i^2}{c_k^2 r_i + s_k^2} = \sum_{i=1}^{I} \mu(i) F(r_i) \geq F \left( \sum_{i=1}^{I} \mu(i) r_i \right) = F \left( \sum_{i=1}^{I} \alpha(i) \right) = F(1) = 1,
\]

(102)

which is the desired inequality. The inequality will be strict so long as \( r_i = \alpha(i) / \mu(i) \) is not constantly equal to 1 over \( i \), or equivalently if \( \alpha(i) \neq \mu(i) \) for some \( i \). This is the desired result.
H Proof of Proposition 7.1

Since $Y_0 = \Omega Y \Pi^T$ has only $n_0$ columns, to ensure that the scaling of the noise matches that of the standard spiked model, we must multiply it by $\sqrt{n/n_0} = 1/\sqrt{\nu}$. We define $Y_0 = Y_0/\sqrt{\nu}$ and $\tilde{X}_0 = X_0/\sqrt{\nu}$. Then $Y_0$ follows a standard spiked model with signal matrix $\tilde{X}_0$.

For $1 \leq k \leq r$, we let $u_k$ and $v_k$ denote the $k^{th}$ singular vectors of $X$; $\hat{u}_k$ and $\hat{v}_k$ denote the $k^{th}$ singular vectors of $Y$; $u_k^0$ and $v_k^0$ denote the $k^{th}$ singular vectors of $X_0$ (and $\tilde{X}_0$); and $\hat{u}_k^0$ and $\hat{v}_k^0$ denote the $k^{th}$ singular vectors of $Y_0$ (and $\tilde{Y}_0$). We let $t_0, \ldots, t_r$ denote the singular values of $\tilde{X}_0$. We also let $\gamma_0 = p_0/n_0 = (\mu/\nu)\gamma$ be the aspect ratio of the submatrix.

If $t_1, \ldots, t_r$ are the singular values of the full $p$-by-$n$ signal matrix $X$, then we may write the rescaled submatrix $X_0$ as

$$\tilde{X}_0 = \Omega \Xi \Pi^T / \sqrt{\nu} = \frac{1}{\sqrt{\nu}} \sum_{k=1}^{r} t_k \Omega u_k v_k^T \Pi^T = \sum_{k=1}^{r} t_k \sqrt{\frac{\alpha_k \beta_k}{\nu}} \frac{\Omega u_k}{\|\Omega u_k\|} \left( \frac{\Pi v_k}{\|\Pi v_k\|} \right)^T. \tag{103}$$

Because the $\Omega u_k$ and $\Pi v_k$ are assumed to be pairwise orthogonal, (103) is the SVD of $\tilde{X}_0$. Consequently:

$$u_k^0 = \frac{\Omega u_k}{\|\Omega u_k\|}, \quad v_k^0 = \frac{\Pi v_k}{\|\Pi v_k\|}, \quad t_k^0 = t_k \sqrt{\frac{\alpha_k \beta_k}{\nu}}. \tag{104}$$

We define the cosines

$$c_k^0 = \langle \hat{u}_k, u_k^0 \rangle, \quad c_k^0 = \langle \hat{v}_k, v_k^0 \rangle. \tag{105}$$

Following Remark 3, we may assume that the singular vectors have been chosen so that both $c_k^0$ and $c_k^0$ are non-negative. Then the AMSE obtained by first applying optimal singular value shrinkage to $Y_0$, and then rescaling by $\nu$, is

$$\|\tilde{X}_0^{shr} - X_0\|_F^2 \sim \nu \sum_{k=1}^{r} (\hat{c}_k^0)^2 (1 - (c_k^0 c_k^0)^2) \sim \sum_{k=1}^{r} t_k^2 \alpha_k \beta_k (1 - (c_k^0 c_k^0)^2). \tag{106}$$

We now turn to the weighted estimator $\tilde{X}_0 = \Omega \tilde{X} \Pi^T$. From the weighted orthogonality condition, $\tilde{X} = \tilde{X}^{dd}$, the optimal diagonal denoiser. From Theorem 5.1, the AMSE of $\tilde{X}_0$ may be written

$$\|\tilde{X}_0 - X_0\|_F^2 = \|\Omega (\tilde{X}^{shr} - X)\|_F^2 \sim \sum_{k=1}^{r} t_k^2 \alpha_k \beta_k \left( 1 - c_k^2 c_k^2 \alpha_k \beta_k + s_k^2 \frac{\alpha_k \beta_k}{\nu} \right). \tag{107}$$

Comparing (106) and (107), the result will be proven if we can show

$$(c_k^0)^2 < c_k^2, \quad \frac{\alpha_k}{c_k^2 \alpha_k + s_k^2 \mu}, \quad 1 \leq k \leq r, \tag{108}$$

and

$$(c_k^0)^2 < c_k^2, \quad \frac{\beta_k}{c_k^2 \beta_k + s_k^2 \nu}, \quad 1 \leq k \leq r. \tag{109}$$

By the symmetry in the problem, it is enough to prove (108). Because we are working with each singular component separately, we will drop the subscript $k$. From Proposition 3.1, the formula for $(c^0)^2$ is given by

$$(c^0)^2 = \begin{cases} \frac{1 - \gamma_0 t^0}{1 + \gamma_0 t^0 / \mu^2}, & \text{if } t^0 \geq \gamma_0^{1/4}, \\ 0, & \text{if } t^0 \leq \gamma_0^{1/4}. \end{cases} \tag{110}$$

If $t^0 \leq \gamma_0^{1/4}$, then (108) is trivial. Consequently, we assume $t^0 > \gamma_0^{1/4}$. Because $t^0 = t \sqrt{\alpha \beta / \nu}$ and $\gamma_0 = \gamma \mu / \nu$, this is equivalent to the condition

$$t^4 > \gamma \frac{\mu \nu}{\alpha^2 \beta^2}. \tag{111}$$
Defining $R = \gamma/t^4$, we may consequently assume that
\[ R < \frac{\alpha^2\beta^2}{\mu\nu} \leq 1. \quad (112) \]

We may rewrite $(c^0)^2$ in terms of $R$ as follows:
\[ (c^0)^2 = \frac{\alpha^2\beta^2 - R\mu\nu}{\alpha^2\beta^2 + t^2\alpha\beta\mu R}. \quad (113) \]

From the formula
\[ c^2 = \frac{1 - \gamma/t^4}{1 + \gamma/t^2}, \quad (144) \]

we may rewrite the right side of (108) as
\[ \frac{c^2\alpha}{c^2\alpha + s^2\mu} = \frac{\alpha^2\beta^2 - R\mu\nu + R(\mu\nu - \alpha^2\beta^2)}{\alpha^2\beta^2 + t^2\alpha\beta\mu R + \alpha\beta R[t^2\mu\beta + \mu\beta - \alpha\beta - t^2\mu]}. \quad (115) \]

Comparing (113) to (115), the inequality (108) is equivalent to showing
\[ (\alpha^2\beta^2 - R\mu\nu)[t^2\mu\beta + \mu\beta - \alpha\beta - t^2\mu] < (\mu\nu - \alpha^2\beta^2)(\alpha\beta + t^2\mu R). \quad (116) \]

Because each side is affine linear in $R$, and $0 \leq R \leq \alpha^2\beta^2/(\mu\nu)$, it is enough to verify (116) at $R = 0$ and $R = \alpha^2\beta^2/(\mu\nu)$. When $R = \alpha^2\beta^2/(\mu\nu)$, the left side of (116) is 0, whereas the right side is non-negative because $\alpha < \sqrt{\nu}$ and $\beta < \sqrt{\mu}$, verifying the inequality in this case. When $R = 0$, the difference between the right side and left side of (116), divided by $\alpha\beta\mu$, is equal to
\[ v - \alpha^2\beta^2/\mu - \alpha\beta[t^2\beta + \beta - \alpha\beta/\mu - t^2] = v - \alpha\beta[t^2\beta + \beta - t^2] \]
\[ = v - \alpha\beta^2t^2 - \alpha\beta^2 + \alpha\beta t^2 \]
\[ = t^2\alpha\beta(1 - \beta) + v - \alpha\beta^2. \quad (117) \]

Since $\beta^2 \leq v \leq 1$ and $\alpha \leq 1$, this expression is positive, verifying (116) and completing the proof.

I Proof of Proposition 7.2

To prove (55), we begin by deriving a lower bound on the operator norm of the noise matrix $N = A^{1/2}GB^{1/2}$. We let $a$ and $b$ be unit vectors so that $G^\top b = \|G\|_{\text{op}} a$. Then
\[ \|GB^{1/2}\|_{\text{op}} \geq \|B^{1/2}G^\top b\| = \|G\|_{\text{op}} \|B^{1/2}a\|. \quad (118) \]

Next, we take unit vectors $c$ and $d$ so that $GB^{1/2}d = \|GB^{1/2}\|_{\text{op}} c$. Then we have
\[ \|N\|_{\text{op}}^2 \geq \|A^{1/2}GB^{1/2}d\|^2 = \|GB^{1/2}\|_{\text{op}}^2 \|A^{1/2}c\|^2 \geq \|G\|_{\text{op}}^2 \cdot \|B^{1/2}a\|^2 \cdot \|A^{1/2}c\|^2. \quad (119) \]

Since the distribution of $G$ is orthogonally-invariant, the distributions of $a$ and $c$ are uniform over the unit spheres in $\mathbb{R}^p$ and $\mathbb{R}^n$, respectively. Consequently, $\|B^{1/2}a\|^2 \sim \text{tr}(B)/n$ and $\|A^{1/2}c\|^2 \sim \text{tr}(A)/p$. Therefore,
\[ \|N\|_{\text{op}}^2 \geq (\text{tr}(A)/p) \cdot (\text{tr}(B)/n) \cdot \|G\|_{\text{op}}^2 \sim (\text{tr}(A)/p) \cdot (\text{tr}(B)/n) \cdot (1 + \sqrt{\gamma})^2, \quad (120) \]

where the inequality holds almost surely in the large $p$, large $n$ limit. We have used that the spectral norm of $G$ converges almost surely to $1 + \sqrt{\gamma}$; see, e.g., [2]. Furthermore, we also have
\[ t_k^2 = t_k^2 \|A^{-1/2}u_k\|^2 \|B^{-1/2}v_k\|^2 \sim t_k^2 \cdot \frac{1}{p} \text{tr}(A^{-1}) \cdot \frac{1}{n} \text{tr}(B^{-1}). \quad (121) \]

Now we can show the improvement in operator norm SNR. We have:
\[ \text{SNR}_{\text{op}}^{(k)} = \frac{t_k^2}{\|N\|_{\text{op}}^2} \leq \frac{t_k^2}{(\text{tr}(A)/p) \cdot (\text{tr}(B)/n) \cdot (1 + \sqrt{\gamma})^2} \]
\[ < \frac{t_k^2 \cdot (\text{tr}(A^{-1})/p) \cdot (\text{tr}(B^{-1})/n)}{(1 + \sqrt{\gamma})^2} = \frac{t_k^2}{\|G\|_{\text{op}}^2} = \text{SNR}_{\text{op}}^{(k)}, \quad (122) \]

where we have used Jensen’s inequality for the strict inequality. This completes the proof of (55).
Proof of Proposition 7.3

Let $\delta_{ij}$ be 1 if entry $(i,j)$ is sampled, and 0 otherwise. Then $\delta_{ij} \sim \text{Bernoulli}(p_i q_j)$. Let $\Delta = (\delta_{ij})$; then $F^*(F(X)) = \Delta \odot X$, where $\odot$ denotes the Hadamard product. Let $q_r = (q_r^1, \ldots, q_r^p)^T$ and $q_c = (q_c^1, \ldots, q_c^n)^T$.

The matrix $\Delta - q_r q_c^T$ is a random matrix with mean zero, whose entries are uniformly bounded. It follows from Corollary 2.3.5 of [41] that $\|\Delta - q_r q_c^T\|_{\text{op}} / \sqrt{n} \leq A$ a.s. as $n \to \infty$, for some constant $A > 0$.

We may write $PXQ = X \odot (q_r q_c^T)$, and consequently $\Delta \odot X = PXQ = (\Delta - q_r q_c^T) \odot X$. Since $X = \sum_{k=1}^r t_k u_k v_k^T$, it is enough to show that

$$\| (\Delta - q_r q_c^T) \odot u_k v_k^T \|_{\text{op}} \to 0$$

almost surely, for each $k$.

Suppose $a$ and $b$ are two unit vectors. Then

$$|a^T (\Delta - q_r q_c^T) \odot u_k v_k^T b| = |(a \odot u_k)^T (\Delta - q_r q_c^T)(b \odot v_k)| \leq A \sqrt{n} \|a \odot u_k\| \|b \odot v_k\|$$

almost surely as $n \to \infty$.

Now, since $a$ is a unit vector,

$$\|a \odot u_k\| = \sqrt{\sum_{j=1}^p a_j^2 u_{jk}^2} \leq \|u_k\|_{\infty}$$

and similarly,

$$\|b \odot v_k\| \leq \|v_k\|_{\infty}.$$  

Since $\max_{1 \leq k \leq r} \|u_k\|_{\infty} \|v_k\|_{\infty} = o(n^{-1/2})$, the result follows.