Is Clustering Advantageous in Statistical Ill-Posed Linear Inverse Problems?

Rasika Rajapakshage and Marianna Pensky

Abstract—In many statistical linear inverse problems, one needs to recover classes of similar objects from their noisy images under an operator that does not have a bounded inverse. Problems of this kind appear in many areas of application. Routinely, in such problems clustering is carried out at a pre-processing step and then the inverse problem is solved for each of the cluster averages separately. As a result, the errors of the procedures are usually examined for the estimation step only. The objective of this paper is to examine, both theoretically and via simulations, the effect of clustering on the accuracy of the solutions of general ill-posed linear inverse problems. In particular, we assume that one observes $X_m = Af_m + \delta \epsilon_m$, $m = 1, \cdots, M$, where functions f_m can be grouped into K classes and one needs to recover a vector function $\mathbf{f} = (f_1, \cdots, f_M)^T$. We construct an estimator for $\mathbf{f}$ as a solution of a penalized optimization problem which corresponds to the clustering before estimation setting. We derive an oracle inequality for its precision and confirm that the estimator is minimax optimal or nearly minimax optimal up to a logarithmic factor of the number of observations. One of the advantages of our approach is that we do not assume that the number of clusters is known in advance. Subsequently, we compare the accuracy of the above procedure with the precision of estimation without clustering, and clustering following the recovery of each of the unknown functions separately. We conclude that clustering at the pre-processing step is beneficial when the problem is moderately ill-posed. It should be applied with extreme care when the problem is severely ill-posed.

Index Terms—Ill-posed linear inverse problem, clustering, oracle inequality, minimax convergence rates.

I. INTRODUCTION

IN THIS paper, we consider a set of general ill-posed linear inverse problems $Af_m = q_m$, $m = 1, \cdots, M$, where A is a bounded linear operator that does not have a bounded inverse and the right-hand sides q_m are measured with error. In particular, we assume that some of the objects f_m and hence q_m, are very similar to each other, so that they can be averaged and recovered together. As a result, one supposedly obtains estimators of f_j with smaller errors. The grouping is usually unknown (as well as the number of groups) and is carried out at a pre-processing step by applying one of the standard

clustering techniques with the number of clusters determined by trial and error. Subsequently, the objects in the same cluster are averaged and the errors of those aggregated objects are used as true errors in the analysis.

Problems of this kind appear in many areas of application such as astronomy (blurred images), econometrics (instrumental variables), medical imaging (tomography, dynamic contrast enhanced Computerized Tomography and Magnetic Resonance Imaging), finance (model calibration of volatility) and many others where similar objects are measured and can be recovered together. Indeed, clustering has been applied for decades to solve ill-posed inverse problems in pattern recognition [5], astronomy [22], astrophysics [14], pattern-based time series segmentation [10], medical imaging [9], elastography for computation of the unknown stiffness distribution [4] and for detecting early warning signs on stock market bubbles [18], to name a few. While in some other settings the main objective is finding group assignments, we are considering only applications where clustering is used merely as a denoising technique. In those applications, routinely, clustering is carried out at the pre-processing step and then the inverse problems are solved for each of the cluster averages separately. As a result, the errors of the procedures are usually examined for the estimation step only. The objective of this paper is to examine, both theoretically and via simulations, the effect of clustering on the accuracy of the solutions of general ill-posed linear inverse problems.

There exists immense literature on the statistical inverse problems (see, e.g., [1], [2], [6]–[8], [11], [20] and monographs [3], [13] and references therein, to name a few). However, to the best of our knowledge, the question about the effects of clustering in statistical inverse problems has never been investigated. Recently, as a part of a more general theory, the effect of clustering on the precision of recovery in multiple regression problems has been studied in [17]. Klopp *et al.* [17] concluded that, even under uncertainty, clustering improves the estimation accuracy. The goal of this paper is to extend this study to the ill-posed linear inverse problems setting.

In particular, we consider the following problem. Let $A : \mathcal{H}_1 \to \mathcal{H}_2$ be a known linear operator where $\mathcal{H}_1$ and $\mathcal{H}_2$ are Hilbert spaces with inner products $\langle \cdot, \cdot \rangle_{\mathcal{H}_1}$ and $\langle \cdot, \cdot \rangle_{\mathcal{H}_2}$, respectively. The objective is to recover functions $f_m \in \mathcal{H}_1$ from

$$X_m(x) = q_m(x) + \delta \epsilon_m(x), \quad (1.1)$$
$$q_m = Af_m, \ m = 1, \cdots, M,$$

where $\epsilon_m(x)$ are the independent white noise processes and the goal is to recover the vector function $f = (f_1, \cdots, f_M)^T$. Assume that observations are taken as functionals of X_m: for any $\psi \in \mathcal{H}_2$ one observes

$$\langle X_m, \psi \rangle = \langle Af_m, \psi \rangle + \delta\, \xi_m(\psi), \qquad (1.2)$$

where δ is noise level and $\xi_m(\psi)$ are zero mean Gaussian random variables with

$$\mathbb{E}[\xi_m(\psi_1)\xi_l(\psi_2)] = \begin{cases} \langle \psi_1, \psi_2 \rangle_{\mathcal{H}_2}, & m = l \\ 0, & m \neq l \end{cases} \qquad (1.3)$$

In what follows we consider the situation where, despite of M being large, there are only K types of functions $f_m(t)$. In particular, we assume that there exists a collection of functions $h_1(t), \ldots, h_K(t)$ such that $f_m(t) = h_K(t)$ for any m and some $k = z(m)$. In other words, one can define a clustering function $z = z(m)$, $m = 1, \ldots, M$, with values in $\{1, \ldots, K\}$ such that $f_m = h_{z(m)}$. We denote the clustering matrix corresponding to the clustering function $z(m)$ by $\mathbf{Z}$. Note that $\mathbf{Z} \in \{0, 1\}^{M \times K}$ and $\mathbf{Z}_{m,k} = 1$ if and only if $z(m) = k$, so that matrix $\mathbf{D}^2 = \mathbf{Z}^T \mathbf{Z}$ is diagonal.

If the function $z(m)$ were known, one could improve precision of estimating f_m by averaging the signals within clusters and construct the estimators $\hat{h}_k$ of the common cluster means, thus reducing the noise levels, and subsequently set $\hat{f}_m = \hat{h}_{z(m)}$. In reality, however, neither the true clustering matrix $\mathbf{Z}_*$, nor the true number of classes K_* are available, so they also need to be estimated.

Note that the objective is accurate estimation of functions f_m, $m = 1, \cdots, M$, rather than recovery of the clustering matrix $\mathbf{Z}$. Moreover, although a true clustering matrix $\mathbf{Z}_*$ always exists (if all functions f_m are different, one can choose $K_* = M$ and $\mathbf{Z}_* = \mathbf{I}_M$), one is not interested in finding $\mathbf{Z}_*$. Indeed, one would rather incur a small bias resulting from replacement of f_m by $h_k \approx f_m$ than obtain estimators with high variances, that are common in inverse problems where each function f_m is estimated separately. On the other hand, using the clustering procedure leads to one more type of errors that are due to erroneously pooling together estimators of functions f_m that belong to different classes, i.e., the errors due to mistakes in clustering.

The goal of this paper is the study of the theoretical recovery limits for the unknown functions f_m, $m = 1, \cdots, M$, when one applies clustering, thus taking advantage of the fact that some of the functions f_m are similar to each other, or ignores this knowledge and proceeds with estimation without clustering. In order to evaluate benefits of clustering, we formulate estimation with clustering problem as an optimization problem. One of the advantages of our approach is that we do not assume that the number of clusters is known in advance. Instead, we elicit the unknown number of clusters, the clustering matrix and the estimators of the unknown functions as a solution of a penalized optimization problem where a penalty is placed on the unknown number of clusters. For this reason, our analysis applies not only to an "ideal" (but usually impractical) situation when the number of clusters is known but to the realistic scenario when it is unknown.

In this paper we analyze the situation where clustering is done before estimation, at the pre-processing level, as it usually happens in many applications. The optimization problem in the paper corresponds to this scenario (specifically, to the K-means clustering setting), as well as our in-depth theoretical study which evaluates the precision of estimators with clustering and compares it to the estimation accuracy without clustering. In order to further assess benefits of clustering, we implement a numerical study and compare the estimators where clustering was carried out at the pre-processing level ("Clustering before") to the estimators where clustering was done post-estimation ("Clustering after") and the estimators without clustering ("No clustering"). We conclude that clustering at the pre-processing level improves estimation precision when the inverse problem is moderately ill-posed but brings no benefits (and can even increase estimation errors) if the problem is severely ill-posed.

The rest of the paper is organized as follows. In Section II, we introduce notations and assumptions and discuss optimization problem that delivers the estimator. Section III deals with quantification of estimation errors. In particular, Section III-A provides the oracle expression for the risk of an estimator obtained in Section II-D. Section III-B presents upper bounds for the risk under the assumptions in Section II-C. In order to ensure that the estimators in Section II-D are asymptotically optimal, in Section III-C we derive minimax lower bounds for the risk. Finally, Section III-D carries out theoretical comparison of estimation accuracy with and without clustering in asymptotic setting. Section IV performs a similar comparison via a simulation study for the case of finite-valued parameters. Finally, Section V contains in-depth discussion and recommendations about application of the pre-clustering in the linear ill-posed problems. Section VI contains proofs of all statements in the paper.

II. ASSUMPTIONS AND ESTIMATION

A. Notations

Below, we shall use the following notations. We denote $[m] = \{1, \cdots, m\}$. We denote vectors and matrices by bold letters. For any vector $\mathbf{a}$, we denote its l_2-norm by $\|\mathbf{a}\|$ and the l_0 norm, the number of non zero elements, by $\|\mathbf{a}\|_0$. For any matrix $\mathbf{A}$, we denote its Frobenius norm by $\|\mathbf{A}\|_F$, the operator norm by $\|\mathbf{A}\|_{op}$ and the span of the column space of matrix $\mathbf{A}$ by $\text{Span}(\mathbf{A})$. We denote the Hamming distance between matrices $\mathbf{A}_1$ and $\mathbf{A}_2$, the number of nonzero elements in $\mathbf{A}_1 - \mathbf{A}_2$, by $\|\mathbf{A}_1 - \mathbf{A}_2\|_H$. We denote the $(k \times k)$ identity matrix by $\mathbf{I}_k$ and drop subscript k when there is no uncertainty about the dimension. We denote the inner product and the corresponding norm in a Hilbert space $\mathcal{H}$ by $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ and $\|\cdot\|_{\mathcal{H}}$, respectively, and drop subscript $\mathcal{H}$ whenever there is no ambiguity. For any set S, we denote cardinality of S by $|S|$. We denote the set of all clustering matrices for grouping M objects into K classes by $\mathcal{M}(M, K)$. We denote $a_n \lesssim b_n$ if there exist $c < \infty$ independent of n such that $a_n \leq c b_n$ and $a_n \gtrsim b_n$ if there exist $c > 0$ independent of n such that $a_n \geq c b_n$. Also, $a_n \asymp b_n$ if simultaneously $a_n \lesssim b_n$ and $a_n \gtrsim b_n$. Finally, we use C as a generic absolute constant

independent of n, M and K, which can take different values in different places.

B. Reduction to the Matrix Model

Since observations are taken as linear functionals (1.2), the problem can be reduced to the so-called sequence model. For this purpose, the unknown functions are expanded over an orthonormal basis ϕ_j, $j = 1, 2, \cdots$, of $\mathcal{H}_1$ and the problem reduces to the recovery of the unknown coefficients of those functions. This is a common technique in the field of statistical inverse problems (see, e.g., Cavalier *et al.* (2002), Cavalier and Golubev (2006) and Knapik *et al.* (2011)). The orthonormal basis is commonly taken to be the eigenbasis of the operator A. Since the eigenbasis is often unknown, in this paper, we consider a wider variety of basis functions. Specifically, we assume that operator A allows a wavelet-vaguelette decomposition introduced by Donoho (1995). In particular, Donoho (1995) assumed that there exists an orthonormal basis ϕ_j, $j = 1, 2, \cdots$, of $\mathcal{H}_1$ and nearly orthogonal sets of functions $\psi_j, \eta_j \in \mathcal{H}_2$, $j = 1, 2, \cdots$, such that for some constants $\nu_j > 0$, and some absolute constants $0 < c_\psi, C_\psi, c_\eta, C_\eta < \infty$ independent of j, one has for any vector $\mathbf{a}$:

$$A\phi_j = \nu_j^{-1}\eta_j, \quad A^*\psi_j = \nu_j^{-1}\phi_j;$$
$$\langle \eta_{j_1}, \psi_{j_2} \rangle_{\mathcal{H}_2} = I(j_1 = j_2); \quad (2.1)$$
$$c_\psi^2 \|\mathbf{a}\|^2 \leq \|\sum_j a_j \psi_j\|^2 \leq C_\psi^2 \|\mathbf{a}\|^2,$$
$$c_\eta^2 \|\mathbf{a}\|^2 \leq \|\sum_j a_j \eta_j\|^2 \leq C_\eta^2 \|\mathbf{a}\|^2, \quad (2.2)$$

where $A^* : \mathcal{H}_2 \to \mathcal{H}_1$ is the linear operator conjugate to A and $I(\ldots)$ is the indicator function. The name was motivated by the fact that conditions (2.1) and (2.2) hold for a variety of linear operators such as convolution, numerical differentiation or Radon transform when $\{\phi_j\}$ is a wavelet basis (see also Abramovich and Silverman (1998)). Obviously, assumptions (2.1) and (2.2) are valid when $\{\phi_j\}$ is the eigenbasis of the operator A. Under conditions (2.1) and (2.2), any function f can be recovered from its image Af using reproducing formula

$$f = \sum_j \nu_j \langle Af, \psi_j \rangle \phi_j \quad (2.3)$$

which is analogous to the reproducing formula for the eigenbasis case.

We expand functions $f_m \in \mathcal{H}_1$ over the basis ϕ_j, $j = 1, \cdots$, and denote the matrix of coefficients by $\mathbf{G}$. Denote $\langle Af_m, \psi_j \rangle = \mathbf{Q}_{j,m}$, so that, by (2.3), for $j = 1, 2, \cdots$, $m = 1, \cdots, M$, one has

$$\mathbf{G}_{j,m} = \langle f_m, \phi_j \rangle = \nu_j \langle f_m, A^*\psi_j \rangle$$
$$= \nu_j \langle Af_m, \psi_j \rangle = \nu_j \mathbf{Q}_{j,m}. \quad (2.4)$$

Consider matrix of observations $\mathbf{Y}$ and matrix of errors $\mathbf{E}$ with respective components $\mathbf{Y}_{j,m} = \langle X_m, \psi_j \rangle$ and $\mathbf{E}_{j,m} = \xi_m(\psi_j)$ where $\xi_m(\psi)$ is defined in (1.3). Let $\mathbf{G}_*$ and $\mathbf{Q}_*$ be the true matrices of coefficients. Then, it follows from (1.1), (1.2) and (2.4) that elements $\mathbf{Y}_{j,m}$ of column m of matrix $\mathbf{Y}$ obey the sequence model for $j = 1, 2, \ldots, m = 1, \cdots, M$,

$$\mathbf{Y}_{j,m} = \nu_j^{-1}(\mathbf{G}_*)_{j,m} + \delta \mathbf{E}_{j,m}. \quad (2.5)$$

Here, $\mathbb{E}(\mathbf{E}_{j,m}) = 0$ and, by (1.3),

$$\mathbb{E}(\mathbf{E}_{j_1,m_1}\mathbf{E}_{j_2,m_2}) = \begin{cases} 0, & m_1 \neq m_2 \\ \langle \psi_{j_1}, \psi_{j_2} \rangle, & m_1 = m_2 \end{cases} \quad (2.6)$$

In order to make the model computationally convenient, we cut the sequence model at some index n where n is large enough to make the error, which is due to this reduction, negligibly small. Then, $j = 1, \ldots, n$, and $\mathbf{G}_*$, $\mathbf{Q}_*$, $\mathbf{Y}$ and $\mathbf{E}$ are $n \times M$ matrices, Also, it follows from (2.5) that

$$\mathbf{\Upsilon Y} = \mathbf{G}_* + \delta \mathbf{\Upsilon E}, \quad \mathbf{\Upsilon} = \text{diag}(\nu_1, \cdots, \nu_n). \quad (2.7)$$

We shall discuss the choice of n later in Section II-C.

Denote the matrix with elements $\mathbf{\Sigma}_{i,j} = \langle \psi_i, \psi_j \rangle$ by $\mathbf{\Sigma}$ and observe that (2.6) implies that

$$\mathbb{E}[(\mathbf{EE}^T)] = M \mathbf{\Sigma}, \quad \mathbb{E}(\mathbf{E}^T\mathbf{E}) = n \mathbf{I}_M. \quad (2.8)$$

Hence, matrix $\mathbf{E}$ has the matrix-variate normal distribution $\mathbf{E} \sim N(0, \mathbf{\Sigma} \otimes \mathbf{I}_M)$. Observe that the first relation in formula (2.2) implies that

$$\|\mathbf{\Sigma}\|_{op} \leq C_\psi^2. \quad (2.9)$$

C. Assumptions

Recall that functions f_m belong to K different groups, so that $f_m = h_k$ with $k = z(m)$ where $z = z(m)$ is a clustering function. Denote the matrix of coefficients of functions h_k in the basis ϕ_j by $\mathbf{\Theta}$, so that $\mathbf{\Theta}_{j,k} = \langle h_k, \phi_j \rangle$, $j = 1, \cdots, n$, $k = 1, \cdots, K$.

It is well known that recovery of an unknown function from noisy observations relies on the fact that it possesses some minimal level of smoothness. This smoothness usually manifests as gradual decline of coefficients of this function in some basis, so the coefficients decrease as one uses more and more complex basis functions. For this reason, we assume that h_k belong to a ball: $h_k \in \mathcal{S}(r, \mathcal{A})$, $k = 1, \ldots, K$, where

$$\mathcal{S}(r, \mathcal{A}) = \left\{ h = \sum_j \theta_j \phi_j : \sum_{j=1}^\infty |\theta_j|^2 j^{2r} \leq \mathcal{A}^2 \right\}. \quad (2.10)$$

If ϕ_j is the Fourier basis, then (2.10) defines a well known Sobolev ball. Formula (2.10) implies that

$$\sum_{j=1}^\infty |\mathbf{\Theta}_{j,k}|^2 j^{2\,r} \leq \mathcal{A}^2, \quad k = 1, \ldots, K. \quad (2.11)$$

If $r \geq 1/2$, then one can set the cut-off value to $n \approx \delta^{-2}$. Indeed, the error rate in the problem cannot be smaller than a parametric rate of $C\delta^2$ and (2.11) implies that the approximation error with this value of n will not exceed

$$\sum_{j=n+1}^\infty |\mathbf{\Theta}_{j,k}|^2 \leq n^{-2r} \sum_{j=1}^\infty |\mathbf{\Theta}_{j,k}|^2 j^{2r} \leq \mathcal{A}^2 n^{-2r} \leq \mathcal{A}^2 \delta^2 \quad (2.12)$$

In addition, it is well known ([23]) that, in the regression setting, the observational version of the white noise model (1.1) based on a sample of size n leads to $\delta = \sigma/\sqrt{n}$ where σ is the standard deviation of the noise.

Furthermore, since operator A does not have a bounded inverse, the values of ν_j in (2.1) are growing with j. While one can consider various scenarios, the standard assumption is that ν_j grow monotonically with j (see, e.g., Alquier *et al.* (2011)):

$$\aleph_1 j^\gamma \exp\left(\alpha j^\beta\right) \le |\nu_j| \le \aleph_2 j^\gamma \exp\left(\alpha j^\beta\right) \qquad (2.13)$$

for some absolute positive constants $\aleph_1$, $\aleph_2$ and nonnegative γ, α and β where $\beta = 0$ and $\gamma > 0$ whenever $\alpha = 0$. The problem (1.1) is called *moderately ill-posed* if $\alpha = 0$ and *severely ill-posed* if $\alpha > 0$.

D. Clustering and Estimation

In what follows, we denote the true quantities using the star symbol, i.e., K_* is the true number of clusters, $\mathbf{Z}_*$ is the true clustering matrix, $\mathbf{G}_*$, $\mathbf{Q}_*$ and $\mathbf{\Theta}_*$ are the true versions of matrices $\mathbf{G}$, $\mathbf{Q}$ and $\mathbf{\Theta}$ and so on. As it was indicated before, we choose $n = [\delta^{-2}]$, the largest integer that is no greater than δ^{-2}.

If $z : [M] \to [K]$ is the clustering function and $\mathbf{Z} \in \{0,1\}^{M \times K}$ is a clustering matrix, then $\mathbf{G}_{i,j} = \mathbf{\Theta}_{i,z(j)}$ for $i = 1, \ldots, n$, $j = 1, \ldots, M$. Therefore, if the clustering matrix $\mathbf{Z}$ were known, then one would repeat columns of matrix $\mathbf{\Theta}$ to obtain $\mathbf{G}$ and average columns of $\mathbf{G}$ to construct $\mathbf{\Theta}$. Specifically, $\mathbf{G} = \mathbf{\Theta}\mathbf{Z}^T$ and $\mathbf{\Theta} = \mathbf{G}\mathbf{Z}\mathbf{D}^{-2}$, where matrix $\mathbf{D}^2 = \mathbf{Z}^T\mathbf{Z}$ is diagonal.

Denote by $\mathbf{\Pi}_{\mathbf{Z},K}$ and $\mathbf{\Pi}_{\mathbf{Z},K}^\perp$ the projection matrices on the column space of matrix $\mathbf{Z}$ and on the orthogonal subspace, respectively:

$$\mathbf{\Pi}_{\mathbf{Z},K} = \mathbf{Z}(\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T, \quad \mathbf{\Pi}_{\mathbf{Z},K}^\perp = \mathbf{I}_M - \mathbf{\Pi}_{\mathbf{Z},K}. \qquad (2.14)$$

Here, we use index K to indicate that not only the clustering matrix $\mathbf{Z}$ but also the number of clusters K is unknown. The projection matrix $\mathbf{\Pi}_{\mathbf{Z},K}$ is such, that for any matrix $\mathbf{G} \in \mathbb{R}^{n \times M}$, $\mathbf{G}\mathbf{\Pi}_{\mathbf{Z},K}$ replaces each column of $\mathbf{G}_j$ of $\mathbf{G}$ by its average over all columns in cluster $z(j)$. Then, matrix $\mathbf{G}_*$ is such that $\mathbf{G}_* = \mathbf{G}_*\mathbf{\Pi}_{\mathbf{Z}_*,K_*}$ and, due to (2.7), if $\mathbf{Z}_*$ were known, it would seem to be reasonable to estimate $\mathbf{G}_*$ by $\mathbf{\Upsilon}\mathbf{Y}\mathbf{\Pi}_{\mathbf{Z}_*,K_*}$. It is well known, however, that this estimator is inadmissible and one needs to shrink or threshold elements of matrix $\mathbf{\Upsilon}\mathbf{Y}\mathbf{\Pi}_{\mathbf{Z}_*,K_*}$ to achieve an optimal bias-variance balance ([19], Section 11.2).

Observe that, since for the ill-posed inverse problems, the values of ν_j are growing with j due to equation (2.13), the elements $\mathbf{G}_{j,i} = \mathbf{\Theta}_{j,z(i)}$ of matrix G are harder and harder to recover as j is growing. On the other hand, condition (2.11) means that coefficients $\mathbf{\Theta}_{j,k}$ decrease rapidly as j increases, and hence, for large n, one does not need to keep all n coefficients for an accurate estimation of functions h_k (and therefore f_m). On the contrary, this will yield an estimator with a huge variance. For this reason, due to the fact that conditions (2.11) apply to all $k = 1, \cdots, K$ simultaneously, we need to choose a set $J \subseteq \{1, \ldots, n\}$ and set $\mathbf{\Theta}_{jk} = 0$ if $j \notin J$. Then, one has $\mathbf{G}_{j,m} = 0$ if $j \in J^c$ where the set J^c is complementary to J. In order to express the latter in a matrix form, we introduce matrix

$$\mathbf{W}_J = \mathrm{diag}(\mathbf{w}_1, \ldots, \mathbf{w}_n) \quad \text{with} \quad \mathbf{w}_j = \mathbb{I}(j \in J), \qquad (2.15)$$

and observe that, for any matrix $\mathbf{G}$, condition $(\mathbf{I}_n - \mathbf{W}_J)\mathbf{G} = \mathbf{0}$ ensures that $\mathbf{G}_{j,m} = 0$, $j \in J^c$.

Consider integer $K \in [M]$, set $\mathcal{M}(M,K)$ of clustering matrices that cluster M nodes into K groups and set $J \subseteq \{1, \ldots, n\}$. Then, the objective is to find matrices $\mathbf{G}$ and $\mathbf{Z} \in \mathcal{M}(M,K)$, a set J and an integer K:

$$(\hat{\mathbf{Z}}, \widehat{\mathbf{G}}, \hat{J}, \hat{K}) \in \operatorname*{argmin}_{\mathbf{Z},\mathbf{G},J,K} \{\|\mathbf{G} - \mathbf{\Upsilon}\mathbf{Y}\mathbf{\Pi}_{\mathbf{Z},K}\|_F^2$$
$$+ \|\mathbf{\Upsilon}\mathbf{Y}\mathbf{\Pi}_{\mathbf{Z},K}^\perp\|_F^2\} \qquad (2.16)$$
$$\text{subject to } (\mathbf{I}_n - \mathbf{W}_J)\mathbf{G} = \mathbf{0},$$

where $\mathbf{\Pi}_{\mathbf{Z},K}^\perp$ is defined in (2.14). The second term in (2.16) corresponds to the error of the K-means clustering of the matrix $\mathbf{\Upsilon}\mathbf{Y}$ while the first term quantifies the difference between the clustered version of data matrix $\mathbf{\Upsilon}\mathbf{Y}\mathbf{\Pi}_{\mathbf{Z},K}$ and the matrix $\mathbf{G}$.

Since $\|\mathbf{\Upsilon}\mathbf{Y}\mathbf{\Pi}_{\mathbf{Z},K}\|_F^2 + \|\mathbf{\Upsilon}\mathbf{Y}\mathbf{\Pi}_{\mathbf{Z},K}^\perp\|_F^2 = \|\mathbf{\Upsilon}\mathbf{Y}\|_F^2$ is independent of $\mathbf{G}$ and $\mathbf{Z}$, the problem can be re-written in an equivalent form as

$$(\hat{\mathbf{Z}}, \widehat{\mathbf{G}}, \hat{J}, \hat{K}) \in \operatorname*{argmin}_{\mathbf{Z},\mathbf{G},J,K} \{\Lambda(\mathbf{Z}, \mathbf{G}, J, K)\} \qquad (2.17)$$
$$\text{subject to } (\mathbf{I}_n - \mathbf{W}_J)\mathbf{G} = \mathbf{0}.$$

where

$$\Lambda(\mathbf{Z}, \mathbf{G}, J, K) = \|\mathbf{G}\|_F^2 - 2\mathrm{Tr}(\mathbf{Y}^T\mathbf{\Upsilon}\mathbf{G}\mathbf{\Pi}_{\mathbf{Z},K}). \qquad (2.18)$$

Note though that optimization problem (2.17) has a trivial solution: $K = M$, $J = [n]$, $\mathbf{Z} = \mathbf{I}_M$ and $\mathbf{G} = \mathbf{\Upsilon}\mathbf{Y}$.

In order to avoid this, we put a penalty on the value of K and the set J, and find $\mathbf{Z}, \mathbf{G}, J$ and K as a solution of the following optimization problem:

$$(\hat{\mathbf{Z}}, \widehat{\mathbf{G}}, \hat{J}, \hat{K}) \in \operatorname*{argmin}_{\mathbf{Z},\mathbf{G},J,K} \{\Lambda(\mathbf{Z}, \mathbf{G}, J, K) + \mathrm{Pen}(J, K)\}$$
$$\text{subject to } \mathbf{Z} \in \mathcal{M}(M,K), (\mathbf{I}_n - \mathbf{W}_J)\mathbf{G} = \mathbf{0}, \qquad (2.19)$$

where $\Lambda(\mathbf{Z}, \mathbf{G}, J, K)$ is defined in (2.18). Optimization procedure (2.19) leads to group thresholding of the rows of matrix $\mathbf{\Upsilon}\mathbf{Y}\mathbf{\Pi}_{\mathbf{Z},\mathbf{K}}$ due to the condition $(\mathbf{I}_n - \mathbf{W}_J)\mathbf{G} = \mathbf{0}$. Indeed, if $\hat{\mathbf{Z}}, \hat{J}$ and $\hat{K}$ were known, then it follows from (2.16) that $\widehat{\mathbf{G}}$ would be given by

$$\widehat{\mathbf{G}} = \mathbf{W}_{\hat{J}}\mathbf{\Upsilon}\mathbf{Y}\mathbf{\Pi}_{\hat{\mathbf{Z}},\hat{K}} \qquad (2.20)$$

and problem (2.19) can be presented as

$$(\hat{\mathbf{Z}}, \hat{J}, \hat{K}) \in \operatorname*{argmin}_{\mathbf{Z},J,K} \{\|(\mathbf{I} - \mathbf{W}_J)\mathbf{\Upsilon}\mathbf{Y}\mathbf{\Pi}_{\mathbf{Z},K}\|_F^2$$
$$+ \|\mathbf{\Upsilon}\mathbf{Y}\mathbf{\Pi}_{\mathbf{Z},K}^\perp\|_F^2 + \mathrm{Pen}(J, K)\} \qquad (2.21)$$
$$\text{subject to } \mathbf{Z} \in \mathcal{M}(M,K), J \subseteq [n], K \in [M].$$

Note that the objective function in (2.21) is a sum of two components: the first one is responsible for the best fitting of the matrix $\mathbf{\Upsilon}\mathbf{Y}\mathbf{\Pi}_{\mathbf{Z},K}$ when some of its rows are set to zero, the second one corresponds to the error of the K-means clustering of columns of matrix $\mathbf{\Upsilon}\mathbf{Y}$, while the penalty prevents over-fitting. The solution of the optimization problem relies on the K-means algorithm that is NP-hard but, however, is known

to provide very accurate results as long as initialization point is not too far from the true solution.

In practice, we shall solve optimization problem (2.21) separately for each $K \in [M]$ and then choose the value of K that delivers the smallest value in (2.21). We estimate the matrix of coefficients $\mathbf{G}$ by $\widehat{\mathbf{G}}$ defined in (2.20). After coefficients $\widehat{\mathbf{G}}$ are obtained, we estimate f_m, $m = 1, \ldots, M$, by

$$\hat{f}_m = \sum_{j \in J} \widehat{\mathbf{G}}_{j,m} \phi_j, \quad m = 1, \cdots, M. \tag{2.22}$$

The penalty in (2.19) and (2.21) should be chosen to exceed the random errors level with high probability. If the number of clusters K, the set J and the clustering matrix $\mathbf{Z}$ were known, then the penalty would be of the order of the variance term $K \sum_{j \in J} \nu_j^2$. However, since K, J and $\mathbf{Z}$ are unknown, we need to account for the uncertainty in estimation of those parameters by applying a union bound and, hence, adding the terms that are proportional to the log-cardinality of the sets of those parameters. Since one has n choices for K, K^M possible clustering arrangements and approximately $\exp\{|J| \ln(ne/|J|)\}$ sets J of cardinality $|J|$ for every $|J| = 1, \ldots, n$, we need to add a term proportional to $(\max_{j \in J} \nu_j^2) [M \ln K + |J| \ln(ne/|J|) + \ln(Mn)]$. Finally, we need to choose a constant τ and add a term proportional to $\max_{j \in J} \nu_j^2 \ln(\delta^{-\tau})$ to ensure that the upper bound holds with probability at least $1 - 2\delta^\tau$. By carefully evaluating the upper bounds for each of the components of the error, we derive the penalty

$$\mathrm{Pen}(J, K) = 2C_\psi^2 \delta^2 \left[26K \sum_{j \in J} \nu_j^2 + 39(\max_{j \in J} \nu_j^2) \left\{ M \ln K \right. \right.$$
$$\left. \left. + |J| \ln(ne/|J|) + \ln(Mn \delta^{-\tau}) \right\} \right] \tag{2.23}$$

where $n = [\delta^{-2}]$, C_ψ is defined in (2.9) and the choice of τ ensures that the upper bound for the error will hold with probability at least $1 - 2\delta^\tau$. Hence, in any real life setting, the constant τ should be such that this probability is large enough.

Penalty (2.23) consists of four terms. The first term, $26 K \sum_{j \in J} \nu_j^2$ represents the error of estimating $|J|$ coefficients for each of the distinct functions h_k, $k = 1, \ldots, K$. The second and the third terms account for the difficulty of clustering M functions into K classes and choosing a set $J \subset \{1, \ldots, n\}$. The last term is of the smaller asymptotic order, it offsets the error of the choice of K and also ensures that the oracle inequality holds with the probability at least $1 - 2\delta^\tau$. Observe that since the data is weighted by the diagonal matrix Υ in (2.7), the last three terms are weighted by $\max_{j \in J} \nu_j^2$.

The penalty (2.23) corresponds to the general model selection that does not rely on assumptions (2.10) and (2.13). If those conditions hold, the elements $(\mathbf{G}_*)_{j,m}$ are decreasing with j for every m, while the values of ν_j are increasing. Therefore, one should choose a set J of the form $J = \{1, \ldots, L\}$ for some $L \leq n$. Since the cardinality of the set of

possible L's is just n, this would lead to replacement of the term $|J| \ln(ne/|J|)$ in the penalty by merely $\ln n$ leading to

$$\overline{\mathrm{Pen}}(L, K) = 2C_\psi^2 \delta^2 \left[26K \sum_{j=1}^{L} \nu_j^2 + 39\nu_L^2 \left\{ M \ln K \right. \right.$$
$$\left. \left. + \ln(Mn \delta^{-\tau}) \right\} \right] \tag{2.24}$$

Remark 1 (Unknown Noise Level): The value of δ in (2.23) and (2.24) is usually unknown but can be easily obtained from data. Indeed, one can apply a wavelet transform to the original data matrix $\mathbf{Y}$, and then recover δ as the median of the absolute value of the wavelet coefficients at the highest resolution level divided by 0.6745 (see, e.g., Mallat (2009), Section 11.3). In fact, in our simulations, we treated δ as an unknown quantity and estimated it by this procedure.

Remark 2 (Different Smoothness for Different Clusters): One can consider a more general case where functions from different clusters have different smoothness levels. In this case, each function h_k has a corresponding set of nonzero coefficients J_k, $k = 1, \ldots, K$, which may be of the form $\{1, \ldots, L_k\}$. Consequently, the terms $K \sum_{j \in J} \nu_j^2$ and $K \sum_{j=1}^{L} \nu_j^2$ in the penalties (2.23) and (2.24) should be replaced by, respectively,

$$\sum_{k=1}^{K} \sum_{j \in J_k} \nu_j^2 \quad \text{and} \quad \sum_{k=1}^{K} \sum_{j=1}^{L_k} \nu_j^2.$$

Theoretical results for this case are a matter of a future investigation.

III. Estimation Error

A. The Oracle Inequality

The average error of estimating f_m by $\hat{f}_m$, $m = 1, \ldots, M$, is given by

$$R(\mathbf{f}, \hat{\mathbf{f}}) = M^{-1} \sum_{m=1}^{M} \| \hat{f}_m - f_m \|^2, \tag{3.1}$$

where $\mathbf{f}$ and $\hat{\mathbf{f}}$ are column vector with functional components f_m and $\hat{f}_m$, $m = 1, \ldots, M$, respectively. Due to the inequality (2.12), the errors of approximation of functions f_m by the n-term expansions over ϕ_j, $j = 1, \ldots, n$, are much smaller than the errors due to estimation or thresholding of the first n coefficients of these expansions. Therefore, the main portion of the error is due to $M^{-1} \| \widehat{\mathbf{G}} - \mathbf{G}_* \|_F^2$. The following statement places an upper bound on $\| \widehat{\mathbf{G}} - \mathbf{G}_* \|_F^2$.

Theorem 1: Let $(\widehat{\mathbf{Z}}, \widehat{\mathbf{G}}, \hat{J}, \hat{K})$ be a solution of optimization problem (2.19) with the penalty $\mathrm{Pen}(J, K)$ given by expression (2.23). Then, there exists a set $\Omega = \Omega(\tau)$ with $\mathbb{P}(\Omega) \geq 1 - 2\delta^\tau$ such that for every $\omega \in \Omega$ one has

$$\| \widehat{\mathbf{G}} - \mathbf{G}_* \|_F^2 \leq \min_{\mathbf{Z}, J, K} \left\{ 3 \| \mathbf{W}_J \mathbf{G}_* \mathbf{\Pi}_{\mathbf{Z}, K} - \mathbf{G}_* \|_F^2 \right.$$
$$\left. + 4 \mathrm{Pen}(J, K) \right\} \tag{3.2}$$

Moreover, if assumptions (2.10) and (2.13) hold and $(\widehat{\mathbf{Z}}, \widehat{\mathbf{G}}, \hat{L}, \hat{K})$ is a solution of optimization problem (2.19) with

$J = \{1,\ldots,L\}$ and the penalty $\text{Pen}(J,K)$ replaced with $\overline{\text{Pen}}(L,K)$ defined in (2.24), then, for $\omega \in \Omega$

$$\|\widehat{\mathbf{G}} - \mathbf{G}_*\|_F^2 \leq \min_{\mathbf{Z},L,K} \left\{ 3 \|\mathbf{W}_J \mathbf{G}_* \mathbf{\Pi}_{\mathbf{Z},K} - \mathbf{G}_*\|_F^2 \right.$$
$$\left. + 4 \overline{\text{Pen}}(L,K) \right\} \tag{3.3}$$

Theorem 1 provides an oracle inequality for $\|\widehat{\mathbf{G}} - \mathbf{G}_*\|_F^2$. The first term in expression (3.2) is the bias term that quantifies the error of approximation of matrix $\mathbf{G}_*$ when its columns are averaged over K clusters using matrix $\mathbf{Z}$ and one keeps only terms with $j \in J$ in the approximations of each of the K cluster means. This term is decreasing when K and $|J|$ are increasing. The second term, $\text{Pen}(J,K)$, is the variance term that represents the error of estimation for the particular choices of $\mathbf{Z}$, J and K. This term grows when K and $|J|$ are increasing. The error is provided by the best possible bias-variance balance in (3.2).

Since the right hand side in (3.2) is minimized over $\mathbf{Z}$ and K, if some of the functions h_k, $k = 1 \cdots, K$, are similar but not exactly identical to each other, it may be advantageous to place those functions in the same cluster, hence, reducing the variance component of the error. Our methodology will automatically take advantage of this opportunity. Note that the error bounds in (3.2) are non-asymptotic and are valid for any true matrix $\mathbf{G}_*$ and any relationship between K, M and δ.

While those results are very valuable, they do not allow to quantify the effect of clustering on estimation errors when δ is small and M is large, so that $\delta \to 0$, $M \to \infty$, and possibly $K \to \infty$. In the next section we shall investigate this issue under assumptions of Section II-C.

B. The Upper Bounds for the Estimation Error

In order to study particular scenarios, in what follows, we assume that ν_j satisfies condition (2.13). Assume that $h_k \in \mathcal{S}(r,\mathcal{A})$, $k = 1,\ldots,K_*$, where $\mathcal{S}(r,\mathcal{A})$ is defined in (2.10). Denote by $\mathbf{h}$ the functional column vector with components h_k, $k = 1,\ldots,K_*$. Consider the maximum risk of our estimator $\widehat{\mathbf{f}}$ over all $h_k \in \mathcal{S}(r,\mathcal{A})$, $k = 1,\ldots,K_*$, and all true clustering matrices $\mathbf{Z}_* \in \mathcal{M}(M,K_*)$

$$R(\widehat{\mathbf{f}}, \mathcal{S}(r,\mathcal{A}), M, K_*) = \max_{\mathbf{f},\mathbf{Z}_*} R(\mathbf{f},\widehat{\mathbf{f}}) \quad \text{subject to} \tag{3.4}$$

$$\mathbf{f} = \mathbf{Z}_* \mathbf{h}, \; h_k \in \mathcal{S}(r,\mathcal{A}), \; k = 1,\ldots,K_*, \; \mathbf{Z}_* \in \mathcal{M}(M,K_*),$$

where $\mathcal{S}(r,\mathcal{A})$ is defined in (2.10) and $\mathcal{M}(M,K_*)$ is the set of all clustering matrices that place M objects into K_* classes.

In what follows, we assume that both n and M are growing simultaneously, that is, $\ln M \asymp \ln(n)$. Note that this is a mild condition since it is satisfied when M is growing at a rate of any positive power of n or visa versa. Hence, due to $n \asymp \delta^{-2}$, we obtain

$$\ln(\delta^{-1}) \asymp \ln n \asymp \ln M \asymp \ln(Mn). \tag{3.5}$$

Observe that the first relation follows from the definition of n while the third one is the direct consequence of the second. Note also that the second assumption is both very mild and very natural. Since $\ln x$ grows very slowly with x, in practical terms, it merely states that both M and δ^{-1} tend to infinity.

The main consequence of the assumption (3.5) is that the terms $\ln(\delta^{-1})$, $\ln n$ and $\ln M$ become interchangeable up to a constant.

Then, application of the oracle inequality (3.2) with $|J| = L$ and $K = K_*$ provides the following upper bounds for the error.

Theorem 2: Let assumption (3.5) hold and ν_j, $j = 1,\cdots,n$, satisfy condition (2.13) with $r \geq 1/2$. Let $(\widehat{\mathbf{Z}}, \widehat{\mathbf{G}}, \hat{L}, \hat{K})$ be a solution of optimization problem (2.19) with the penalty given by expression (2.23). Then, with probability at least $1 - 2\delta^\tau$, one has

$$R(\widehat{\mathbf{f}}, \mathcal{S}(r,\mathcal{A}), M, K_*) \leq C\, R(M, K_*, \delta),$$

where the constant C depends on $\alpha, \beta, \gamma, r, \tau$ and $\mathcal{A}$ only and

$$R(M,K_*,\delta) = \left(\delta^2 \ln K_*\right)^{\frac{2r}{2r+2\gamma}} + \left(\delta^2 M^{-1} K_*\right)^{\frac{2r}{2r+2\gamma+1}}, \tag{3.6}$$

if $\alpha = \beta = 0$, and

$$R(M,K_*,\delta) = \left[\ln\left(\frac{1}{\delta^2 \ln K_*}\right)\right]^{-\frac{2r}{\beta}} + \left[\ln\left(\frac{M}{\delta^2 K_*}\right)\right]^{-\frac{2r}{\beta}}, \tag{3.7}$$

if $\alpha > 0, \beta > 0$.

The expressions in (3.6) and (3.7) are well defined if $K_* \geq 2$. If $K_* = 1$, then $\ln K_* = 0$ and the first terms in (3.6) and (3.7) are just equal to zero.

C. The Minimax Lower Bounds for the Risk

In order to show that the estimator developed in this paper is asymptotically near-optimal, below we derive minimax lower bounds for the risk over all $h_k \in \mathcal{S}(r,\mathcal{A})$, $k = 1,\ldots,K_*$, and all clustering matrices $\mathbf{Z}_* \in \mathcal{M}(M,K_*)$. For this purpose, we define the minimax risk as

$$R_{\min}(\mathcal{S}(r,\mathcal{A}), M, K_*) = \min_{\tilde{\mathbf{f}}} R(\tilde{\mathbf{f}}, \mathcal{S}(r,\mathcal{A}), M, K_*) \tag{3.8}$$

where $\tilde{\mathbf{f}}$ is any estimator of $\mathbf{f}$ on the basis of matrix of observations $\mathbf{Y}$.

Theorem 3: Let ν_j, $j = 1,\cdots$, satisfy condition (2.13) and $r \geq 1/2$. Then, with probability at least 0.1, one has

$$R_{\min}(\mathcal{S}(r,\mathcal{A}), M, K_*) \geq C R_{\min}(M, K_*, \delta) \tag{3.9}$$

where the constant C depends on α, β, γ, r and $\mathcal{A}$ only and

$$R_{\min}(M,K_*,\delta) = \tag{3.10}$$
$$\max\left\{\left(\delta^2 \ln K_*\right)^{\frac{2r}{2r+2\gamma}}, \; \left(\delta^2 M^{-1} K_*\right)^{\frac{2r}{2r+2\gamma+1}}\right\},$$

if $\alpha = \beta = 0$, and

$$R_{\min}(M,K_*,\delta) = \tag{3.11}$$
$$\max\left\{\left[\ln\left(\frac{1}{\delta^2 \ln K_*}\right)\right]^{-\frac{2r}{\beta}}, \; \left[\ln\left(\frac{M}{\delta^2 K_*}\right)\right]^{-\frac{2r}{\beta}}\right\},$$

if $\alpha > 0, \beta > 0$.

Observe that expressions for the upper and the lower bounds of the risk (3.6) and (3.10) in the case of $\alpha = \beta = 0$, and (3.7) and (3.11) in the case of $\alpha > 0, \beta > 0$ are identical, so our estimators are asymptotically optimal.

TABLE I

ESTIMATION AND CLUSTERING ERRORS FOR THE "CLUSTERING BEFORE", "CLUSTERING AFTER" AND 'NO CLUSTERING" SCENARIOS AVERAGED OVER 100 SIMULATION RUNS (THE STANDARD DEVIATIONS OF THE MEANS ARE IN PARENTHESES). RESULTS FOR THE SET OF FUNCTIONS (4.3) WITH THE $g_1(x)$ KERNEL IN (4.2) AND THE SAME SET OF NONZERO COEFFICIENTS FOR ALL FUNCTIONS

	$\lambda = 7$				
	Clustering Before		Clustering After		No Clustering
	Error	Miss-rate	Error	Miss-rate	
$SNR = 3$	0.0365(0.0262)	0.0090	0.0554(0.0083)	0.0068	0.0556(0.0001)
$SNR = 5$	0.0270(0.0015)	0.0000	0.0419(0.0095)	0.0070	0.0423(0.0001)
$SNR = 7$	0.0250(0.0084)	0.0031	0.0405(0.0000)	0.0000	0.0414(0.0000)
	$\lambda = 5$				
$SNR = 3$	0.0377(0.0038)	0.0000	0.0549(0.0059)	0.0033	0.0567(0.0002)
$SNR = 5$	0.0317(0.0016)	0.0000	0.0542(0.0000)	0.0000	0.0551(0.0000)
$SNR = 7$	0.0269(0.0014)	0.0000	0.0406(0.0000)	0.0000	0.0421(0.0001)
	$\lambda = 3$				
$SNR = 3$	0.0498(0.0398)	0.0106	0.0810(0.0217)	0.0133	0.0788(0.0001)
$SNR = 5$	0.0409(0.0237)	0.0036	0.0543(0.0000)	0.0000	0.0565(0.0002)
$SNR = 7$	0.0350(0.0241)	0.0026	0.0542(0.0000)	0.0000	0.0554(0.0001)

D. The Advantage of Clustering

Theorems 2 and 3 allow to answer the question whether clustering in linear ill-posed inverse problems improves the estimation accuracy as $M \to \infty$ and $\delta \to 0$. Indeed, solving problem (1.1) for each $m = 1, \cdots, M$ separately is equivalent to choosing $K = M = 1$ in the penalty. In this case, one obtains the following corollary.

Corollary 1: If each of the inverse problems is solved separately, where the penalty is of the form (2.23) with $K = M = 1$ and $J = \{1, \cdots, L\}$, then, with probability at least $1 - 2\delta^{\tau}$, the average estimation error $\tilde{R}(\delta)$ defined in (3.1) is bounded by

$$\tilde{R}(\delta) \asymp \begin{cases} \left[\delta^2\right]^{\frac{2r}{2\gamma+2r+1}}, & \text{if } \alpha = \beta = 0, \\ \left[\ln(\delta^{-1})\right]^{-\frac{2r}{\beta}}, & \text{if } \alpha > 0, \beta > 0. \end{cases} \quad (3.12)$$

If $r \geq 1/2$ and assumption (3.5) holds, then for $\delta \to 0$, $M \to \infty$, one has

$$\frac{R(M, K_*, \delta)}{\tilde{R}(\delta)} \asymp \begin{cases} 1 & \text{if } \alpha > 0, \beta > 0, \\ M^{-\frac{2r}{2\gamma+2r+1}}, & \text{if } \alpha = \beta = 0, K_* = 1 \\ \left(\frac{K_*}{M}\right)^{\frac{2r}{2\gamma+2r+1}} + \delta^{\frac{4r}{(2\gamma+2r+1)(2r+2\gamma)}} \ln(K_*) \\ \quad \text{if } \alpha = \beta = 0, K_* \geq 2. \end{cases} \quad (3.13)$$

Therefore, when $\delta \to 0$, $M \to \infty$, clustering is asymptotically advantageous if $\alpha = \beta = 0$.

IV. SIMULATIONS

In order to study finite sample properties of the proposed estimation procedure, we carried out a numerical study. In particular, we considered a periodic convolution equation $q = Ah = h * g$ with a kernel g that transforms into a product in the Fourier domain

$$\tilde{q}_j = \tilde{g}_j \tilde{h}_j, \quad \nu_j = 1/\tilde{h}_j, \; j = 1, \cdots, n, \quad (4.1)$$

where, for any function t, we denote its j-th Fourier coefficient by $\tilde{t}_j$. The periodic Fourier basis serves as the eigenbasis for this operator.
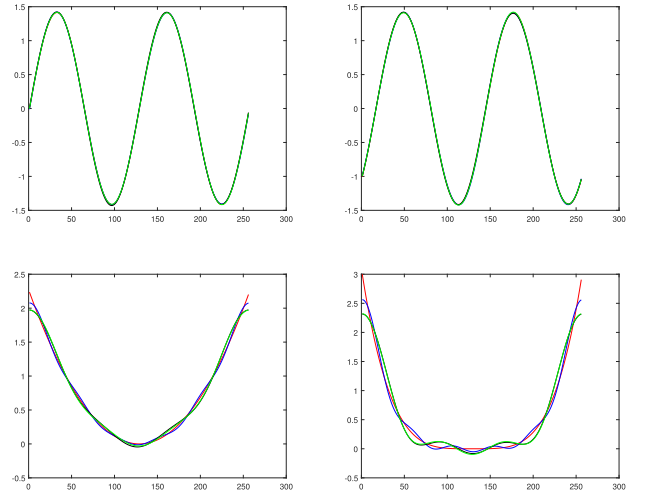

Fig. 1. True functions (red) and their estimators: "Clustering before" (blue), "Clustering after" (green) and "No clustering" (black). Results for the functions in (4.3) and the kernel g_1 in (4.2) with $\lambda = 3$ and SNR = 3. Top row: h_1 (left), h_2 (right). Bottom row: h_3 (left), h_4 (right).

We carried out simulations with the periodized versions of the following two kernels

$$g_1(x) = 0.5 \exp(-\lambda|x|), \quad g_2(x) = \exp(-\lambda x^2/2) \quad (4.2)$$

where $g_1(x)$ corresponds to the case of $\alpha = \beta = 0, \gamma = 2$ while $g_2(x)$ corresponds to $\alpha \propto 1/\lambda$, $\beta = 2$ in (2.13). Hence, the problem is moderately ill-posed with g_1 and severely ill-posed with g_2. In addition, recovery of the solution becomes easier as λ grows.

Although we carried out simulations for a much wider sets of parameters, here we report the results for two series of simulations with $n = 256$, $M = 60$ and $K = 4$. In the first batch, we considered a set of smooth spatially homogeneous test functions

$$l_1(x) = \sin(4\pi x), \; l_2(x) = \sin(4\pi(x - 1/16)), \quad (4.3)$$
$$l_3(x) = (x - 0.5)^2, \; l_4(x) = (x - 0.5)^4,$$

coefficients of which follow the assumption (2.11). For this set, we used Fourier basis ϕ_j, $j = 1, \cdots, n$, that diagonalizes

TABLE II

ESTIMATION AND CLUSTERING ERRORS FOR THE "CLUSTERING BEFORE", "CLUSTERING AFTER" AND "NO CLUSTERING" SCENARIOS AVERAGED OVER 100 SIMULATION RUNS (THE STANDARD DEVIATIONS OF THE MEANS ARE IN PARENTHESES). RESULTS FOR THE SET OF FUNCTIONS (4.3) WITH THE $g_2(x)$ KERNEL IN (4.2) AND THE SAME SET OF NONZERO COEFFICIENTS FOR ALL FUNCTIONS

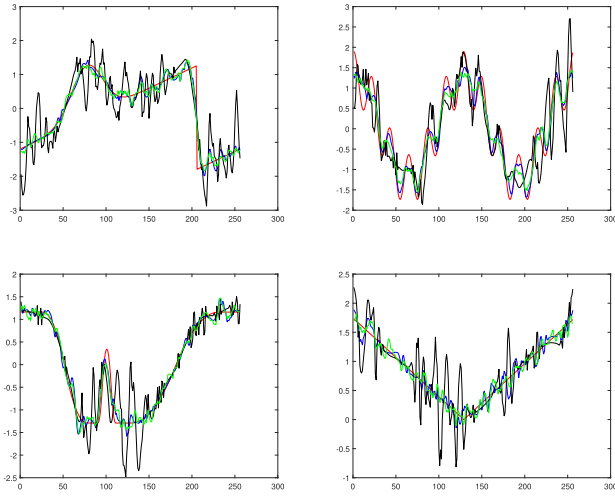|	Clustering Before		Clustering After		No Clustering
	Error	Miss-rate	Error	Miss-rate	
$\lambda = 15$					
$SNR = 5$	0.1568(0.0684)	0.0623	0.1258(0.0175)	0.0071	0.1252(0.0002)
$SNR = 7$	0.1516(0.0640)	0.0521	0.1252(0.0063)	0.0180	0.1245(0.0001)
$SNR = 10$	0.1307(0.0342)	0.0128	0.1237(0.0000)	0.0000	0.1241(0.0000)
$\lambda = 12$					
$SNR = 5$	0.2336(0.0770)	0.1601	0.1609(0.0173)	0.0398	0.1659(0.0034)
$SNR = 7$	0.2186(0.0759)	0.1303	0.1602(0.0080)	0.0413	0.1620(0.0025)
$SNR = 10$	0.1938(0.0660)	0.0758	0.1583(0.0058)	0.0211	0.1592(0.0017)
$\lambda = 10$					
$SNR = 5$	0.5419(0.0707)	0.2513	0.7933(0.1005)	0.2796	0.7448(0.0000)
$SNR = 7$	0.5196(0.0128)	0.2331	0.7678(0.0733)	0.2693	0.7448(0.0000)
$SNR = 10$	0.5078(0.0212)	0.1715	0.4853(0.0084)	0.0430	0.4849(0.0037)

Fig. 2. True functions (red) and their estimators: "Clustering before" (blue), "Clustering after" (green) and "No clustering" (black). Results for the functions in (4.4) and the kernel g_1 in (4.2) with $\lambda = 3$ and SNR = 3. Top row: h_1 (left), h_2 (right). Bottom row: h_3 (left), h_4 (right).

the problem. Moreover, since the functions are spatially homogeneous, they can be well estimated when the same set J of nonzero coefficients is used for all four functions. In the second round, we expanded our study to the set of spatially inhomogeneous functions

$$l_1(x) = l_B(x), \ l_2(x) = l_W(x), \quad (4.4)$$
$$l_3(x) = l_P(x), \ l_4(x) = |x - 0.5|$$

where $l_B(x)$, $l_W(x)$ and $l_P(x)$ are the *blip, wave* and *parabolas* introduced by Donoho and Johnstone [12]. In this case, Fourier basis does not allow accurate estimation, hence, we used the Daubechies 8 wavelet basis as ϕ_j, $j = 1, \cdots, n$, for which conditions (2.1) and (2.2) hold with ν_j given in (4.1) (see, e.g., [1]). Although the second example does not follow our assumptions, it shows that our conclusions are true even in the situation when those assumptions are violated. In particular, we used a different set of nonzero coefficients J_k for l_k, $k = 1, \ldots, 4$, for the functions in (4.4). We sampled the test functions on the equispaced grid on the interval $[0, 1]$ and

scaled them to have norms $\sqrt{n}$, obtaining $h_k = c_k l_k$ where $c_k = \sqrt{n}/\|l_k\|$, $k = 1, \ldots, 4$. Note that, while the functions in Set 1 (4.3) are simpler and easier to recover, they are less distinct and harder to cluster since l_1 is similar to l_2 and l_3 is similar to l_4. On the other hand, while it is easier to distinguish between images of functions in Set 2 (4.4), they are more difficult to estimate. For each of the test functions h_k, $k = 1, \cdots, K$, we evaluated $u_k = (Ah)_k$, and sampled those functions on the grid of n equispaced points j/n, $j = 1, \cdots, n$, on the interval $[0, 1]$, obtaining vectors $\mathbf{h}_k$ and $\mathbf{u}_k$, $k = 1, \cdots, K$. Furthermore, we generated a clustering function $z : M \to K$ that places M objects into K classes, M/K into each class at random. We obtained the true matrices $\mathbf{F}, \mathbf{Q} \in \mathbb{R}^{n \times M}$ with the columns $\mathbf{h}_{z(m)}$ and $\mathbf{u}_{z(m)}$, $m = 1, \cdots, M$, respectively. Finally, we generated data $\mathbf{X}$ by adding independent Gaussian noise with the standard deviation σ to every element in $\mathbf{Q}$. We found σ by fixing the Signal-to-Noise Ratio (SNR) and choosing $\sigma = \text{std}(\mathbf{F})/SNR$, where $\text{std}(\mathbf{F})$ is the standard deviation of the matrix $\mathbf{F}$ reshaped as a vector. In what follows, we considered several noise scenarios: SNR = 3, 5 and 7 for g_1 and SNR = 5, 7, and 10 for g_2. In our study we treat K as known and compare the estimators where clustering was carried out at pre-processing level ("Clustering before") to the estimators where clustering was done post-estimation ("Clustering after") and estimators without clustering ("No clustering").

For the "Clustering before" setting, we applied clustering directly to the elements of matrix $\mathbf{Y}$. As it follows from equation (2.21), the matrix $\hat{\mathbf{Z}} \in \mathcal{M}(M, K)$ which minimizes the objective function is a solution of the K-means clustering problem. Subsequently, we found matrix $\mathbf{\Pi}_{\hat{\mathbf{Z}}, K}$ and, following equation (2.20), estimated $\mathbf{G}_*$ by $\widehat{\mathbf{G}} = \mathbf{W}_{\hat{J}} \mathbf{\Upsilon} \mathbf{Y} \mathbf{\Pi}_{\hat{\mathbf{Z}}, K}$. For the set of functions (4.3), the set $\hat{J}$ was obtained by applying hard thresholding to the rows of the matrix $\mathbf{\Upsilon} \mathbf{Y} \mathbf{\Pi}_{\hat{\mathbf{Z}}, K}$, while for the set of functions (4.4), we applied hard hard thresholding to each of the elements of the matrix $\mathbf{\Upsilon} \mathbf{Y} \mathbf{\Pi}_{\hat{\mathbf{Z}}, K}$. Finally, the estimator $\widehat{\mathbf{F}}$ of the matrix $\mathbf{F}_*$ is obtained by applying the inverse Fourier transform (in the case of the functions in (4.3)) or the inverse wavelet transform (in the case of the functions in (4.4)) to the columns of matrix $\widehat{\mathbf{G}}$. For the "Clustering after" setting, we first constructed the "No clustering" estimator $\check{\mathbf{G}}$

TABLE III

ESTIMATION AND CLUSTERING ERRORS FOR THE "CLUSTERING BEFORE", "CLUSTERING AFTER" AND "NO CLUSTERING" SCENARIOS AVERAGED OVER 100 SIMULATION RUNS (THE STANDARD DEVIATIONS OF THE MEANS ARE IN PARENTHESES). RESULTS FOR THE SET OF FUNCTIONS (4.4) WITH THE $g_1(x)$ KERNEL IN (4.2) AND UNIQUE SET OF NONZERO COEFFICIENTS FOR EACH OF THE FUNCTIONS

	Clustering Before		Clustering After		No Clustering
$\lambda = 7$					
	Error	Miss-rate	Error	Miss-rate	
$SNR = 3$	0.1364(0.0055)	0.0000	0.2650(0.0609)	0.0250	0.2810(0.0056)
$SNR = 5$	0.1190(0.0039)	0.0000	0.2187(0.0661)	0.0180	0.2470(0.0030)
$SNR = 7$	0.1033(0.0058)	0.0000	0.1700(0.0599)	0.0205	0.2004(0.0039)
$\lambda = 5$					
$SNR = 3$	0.1480(0.0077)	0.0000	0.2845(0.1095)	0.0731	0.3569(0.0091)
$SNR = 5$	0.1186(0.0053)	0.0000	0.2322(0.1117)	0.0610	0.2719(0.0040)
$SNR = 7$	0.1026(0.0045)	0.0000	0.1632(0.0656)	0.0221	0.2169(0.0042)
$\lambda = 3$					
$SNR = 3$	0.1806(0.0110)	0.0000	0.2932(0.1309)	0.1010	0.4831(0.0092)
$SNR = 5$	0.1442(0.0061)	0.0000	0.2326(0.1199)	0.0690	0.3250(0.0053)
$SNR = 7$	0.1310(0.0047)	0.0000	0.2149(0.1207)	0.0718	0.2542(0.0042)

TABLE IV

ESTIMATION AND CLUSTERING ERRORS FOR THE "CLUSTERING BEFORE", "CLUSTERING AFTER" AND "NO CLUSTERING" SCENARIOS AVERAGED OVER 100 SIMULATION RUNS (THE STANDARD DEVIATIONS OF THE MEANS ARE IN PARENTHESES). RESULTS FOR THE SET OF FUNCTIONS (4.4) WITH THE $g_2(x)$ KERNEL IN (4.2) AND UNIQUE SET OF NONZERO COEFFICIENTS FOR EACH OF THE FUNCTIONS

	Clustering Before		Clustering After		No Clustering
$\lambda = 15$					
	Error	Miss-rate	Error	Miss-rate	
$SNR = 5$	0.3709(0.0000)	0.0000	0.3709(0.0000)	0.0000	0.3714(0.0001)
$SNR = 7$	0.3708(0.0000)	0.0000	0.3708(0.0000)	0.0000	0.3711(0.0000)
$SNR = 10$	0.3708(0.0000)	0.0000	0.3708(0.0000)	0.0000	0.3710(0.0000)
$\lambda = 12$					
$SNR = 5$	0.3768(0.0009)	0.0000	0.3768(0.0009)	0.0000	0.3810(0.0011)
$SNR = 7$	0.3766(0.0006)	0.0000	0.3780(0.0137)	0.0036	0.3787(0.0006)
$SNR = 10$	0.3765(0.0004)	0.0000	0.3785(0.0202)	0.0035	0.3776(0.0004)
$\lambda = 10$					
$SNR = 5$	0.4876(0.0049)	0.0000	0.4933(0.0294)	0.0141	0.4940(0.0050)
$SNR = 7$	0.4869(0.0035)	0.0000	0.4869(0.0035)	0.0000	0.4903(0.0035)
$SNR = 10$	0.4872(0.0027)	0.0000	0.4872(0.0027)	0.0000	0.4888(0.0027)

of matrix $\mathbf{G}_*$ by thresholding elements of the columns of the matrix $\mathbf{\Upsilon Y}$ in equation (2.7), and then obtained the estimator $\check{\mathbf{F}}$ of matrix $\mathbf{F}_*$ by applying the inverse Fourier or wavelet transform to the columns of matrix $\check{\mathbf{G}}$. Finally, the "Clustering after" estimator of $\mathbf{F}_*$ is obtained by applying the K-means clustering procedure to the columns of matrix $\check{\mathbf{F}}$.

Tables I–IV report simulations results for the three clustering scenarios above ("Clustering before", "Clustering after" and "No clustering"), for each of the sets of test functions in (4.3) and (4.4) and for each of the two kernels in (4.2) with various values of λ. In the Tables, we display the accuracies of the three estimators where the precision of an estimator $\widehat{\mathbf{F}}$ is measured by the Frobenius norms of its error

$$\Delta = \Delta(\widehat{\mathbf{F}}) = \|\widehat{\mathbf{F}} - \mathbf{F}\|_F / \sqrt{Mn}. \qquad (4.5)$$

In addition, we report the proportion of erroneously clustered nodes ("Miss-rate") for the "Clustering before"and the "Clustering after" estimators.

We ought to point out that the "Clustering before" estimation procedure is much more computationally efficient since it does not require to recover M unknown functions separately which is necessary for the "Clustering after" and "No clustering" procedures.

V. CONCLUSION

In this paper, we investigate theoretically and via a limited simulation study, the effect of clustering on the accuracy of recovery in ill-posed linear inverse problems. As we have stated earlier, in many applications leading to such problems, clustering is carried out at a pre-processing step and later is totally forgotten when it comes to error evaluation. Our main objective has been to evaluate what effect clustering at the pre-processing step has on the precision of the resulting estimators.

It appears that benefits of pre-clustering depend significantly on the nature of the inverse problem at hand. If the problem is moderately ill-posed (kernel g_1 in (4.2), $\alpha = \beta = 0$), then, as Corollary 1 shows, the "Clustering Before" estimator has asymptotically smaller errors than the "No Clustering" estimator when the number of functions and the sample size grow. Tables I and II, corresponding to this case, confirm that, for the finite number of functions and moderate sample size, the "Clustering before" procedure delivers better precision than the "Clustering after" and "No clustering" techniques. Furthermore, the "Clustering before" estimation has profound computational benefits since one needs to recover K unknown functions instead of M. Moreover, the advantages of clustering at pre-processing step become more prominent when the

problem is less ill-posed (larger λ). Indeed, in the case when the problem is not ill-posed ($\alpha = \beta = \gamma = 0$ in (2.13)), as findings of Klopp *et al.* [17] show, clustering always improves estimation precision.

The situation changes drastically if the inverse problem is severely ill-posed (kernel g_2 in (4.2), $\alpha > 0, \beta > 0$). Our theoretical results indicate that clustering, in this case, does not improve the estimation precision as the number of functions and the sample size grow. These findings are consistent with the simulation study. In the case of functions in (4.4), Table IV implies that the precisions of all three methodology are approximately the same, and the estimation errors are high even when clustering errors are small or zero. This is due to the fact that the reduction in the noise level due to clustering is not sufficient to counteract the ill-posedness of the problem and, thus, it does not lead to a meaningful improvement in estimation accuracy. Table III, that reports on the simulations with functions in (4.3), presents an even more grim picture. Since functions in the set (4.3) resemble each other to start with and convolutions with the kernel g_2 make them to appear even more similar, "Clustering before" procedure leads to relatively high clustering errors that, in turn, produce higher estimation errors than the "Clustering after" and "No clustering" techniques.

In conclusion, clustering at the pre-processing step is beneficial when the problem is moderately ill-posed. It should be applied with extreme care when the problem is severely ill-posed.

VI. Proofs

A. Proof of the Oracle Inequality

Proof of Theorem 1: The proof of the inequality (3.2) is based on the standard techniques for proofs of oracle inequalities. We use optimization problem (2.19) to present the left-hand side as a sum of the error of any estimator plus the random error term followed by the difference between the penalty terms. Later on, we upper-bound the random error term for any number of classes K, any clustering matrix Z and any set J. After that, we take a union bound over all possible K, Z and J to obtain an upper bound for the probability that the error exceeds certain threshold. The novelty of the proof lies in the fact that we are using a vectorization of the model which allows us to attain the upper bounds.

Note that it follows from the optimization problem (2.19) that for any fixed $\mathbf{G}, \mathbf{Z}, J$ and K one has

$$\|\widehat{\mathbf{G}}\|_F^2 - 2\mathrm{Tr}(\mathbf{Y}^T \mathbf{\Upsilon} \widehat{\mathbf{G}} \mathbf{\Pi}_{\hat{\mathbf{Z}}, \hat{K}}) + \mathrm{Pen}(\hat{J}, \hat{K})$$
$$\leq \|\mathbf{G}\|_F^2 - 2\mathrm{Tr}(\mathbf{Y}^T \mathbf{\Upsilon} \mathbf{G} \mathbf{\Pi}_{\mathbf{Z}, K}) + \mathrm{Pen}(J, K).$$

Then, adding and subtracting $\mathbf{G}_*$, we obtain

$$\|\widehat{\mathbf{G}} - \mathbf{G}_*\|_F^2 + \|\mathbf{G}_*\|_F^2 + 2\mathrm{Tr}((\widehat{\mathbf{G}} - \mathbf{G}_*)^T \mathbf{G}_*)$$
$$- 2\mathrm{Tr}(\mathbf{Y}^T \mathbf{\Upsilon} \widehat{\mathbf{G}} \mathbf{\Pi}_{\hat{\mathbf{Z}}, \hat{K}}) + \mathrm{Pen}(\hat{J}, \hat{K}) \leq$$
$$\|\mathbf{G} - \mathbf{G}_*\|_F^2 + \|\mathbf{G}_*\|_F^2 + 2\mathrm{Tr}((\mathbf{G} - \mathbf{G}_*)^T \mathbf{G}_*)$$
$$- 2\mathrm{Tr}(\mathbf{Y}^T \mathbf{\Upsilon} \mathbf{G} \mathbf{\Pi}_{\mathbf{Z}, K}) + \mathrm{Pen}(J, K).$$

Combine the trace product terms and recall that, due to equation (2.5), $\mathbf{Y} = \mathbf{\Upsilon}^{-1} \mathbf{G}_* + \delta \mathbf{E}$. Hence, the last inequality

yields

$$\|\widehat{\mathbf{G}} - \mathbf{G}_*\|_F^2 \leq \|\mathbf{G} - \mathbf{G}_*\|_F^2 + 2\,\delta\,\mathrm{Tr}[\mathbf{E}^T \mathbf{\Upsilon}(\widehat{\mathbf{G}} - \mathbf{G})] \quad (6.1)$$
$$+ \mathrm{Pen}(J, K) - \mathrm{Pen}(\hat{J}, \hat{K})$$

We choose $\mathbf{G} = \mathbf{W}_J \mathbf{G}_* \mathbf{\Pi}_{\mathbf{Z}, K}$ and, in order to analyze the cross term $\mathrm{Tr}[\mathbf{E}^T \mathbf{\Upsilon}(\widehat{\mathbf{G}} - \mathbf{G})]$, we use vectorization of the model. For this purpose, we choose $\mathbf{S}$ such that $\mathbf{\Sigma} = \mathbf{S}\mathbf{S}^T$ and denote

$$\mathbf{\Pi}_{\hat{\mathbf{Z}}, \hat{K}, \hat{J}} = (\mathbf{\Pi}_{\hat{\mathbf{Z}}, \hat{K}} \otimes \mathbf{W}_{\hat{J}}), \mathbf{\Pi}_{\mathbf{Z}, K, J} = (\mathbf{\Pi}_{\mathbf{Z}, K} \otimes \mathbf{W}_J) \quad (6.2)$$
$$\hat{\mathbf{g}} = \mathrm{vec}(\widehat{\mathbf{G}}), \quad \mathbf{g} = \mathrm{vec}(\mathbf{G}), \quad \boldsymbol{\epsilon} = \mathrm{vec}(\mathbf{E}),$$
$$\mathbf{\Gamma} = (\mathbf{I}_M \otimes \mathbf{\Upsilon}), \quad \boldsymbol{\eta} = (\mathbf{I}_M \otimes \mathbf{S}^{-1})\boldsymbol{\epsilon}. \quad (6.3)$$

By definition of the matrix-variate normal distribution (Theorem 2.3.1 of Gupta and Nagar (2000)) and (2.8), we derive that

$$\boldsymbol{\epsilon} \sim N(0, \mathbf{\Sigma} \otimes \mathbf{I}_M) \quad (6.4)$$

Then, $\mathbb{E}(\boldsymbol{\eta}\boldsymbol{\eta}^T) = \mathbf{I}_{nM}$, so that, $\boldsymbol{\eta} \sim N(0, \mathbf{I}_{nM})$, where $\boldsymbol{\epsilon}$ is defined in (6.3) and $\|\mathbf{S}\|_{op} \leq C_\psi$. Then, equation (2.7) can be re-written as

$$\mathbf{\Gamma}\mathbf{y} = \mathbf{g}_* + \delta\,\mathbf{\Gamma}\,(\mathbf{I}_M \otimes \mathbf{S})\boldsymbol{\eta}. \quad (6.5)$$

Observe that by Theorem 1.2.22 of Gupta and Nagar (2000), one has

$$\hat{\mathbf{g}} = \mathrm{vec}(\mathbf{W}_{\hat{J}} \mathbf{\Upsilon} \mathbf{Y} \mathbf{\Pi}_{\hat{\mathbf{Z}}, \hat{K}}) = \mathbf{\Pi}_{\hat{\mathbf{Z}}, \hat{K}, \hat{J}} \mathbf{\Gamma}\mathbf{y}, \quad \mathbf{g} = \mathbf{\Pi}_{\mathbf{Z}, K, J} \mathbf{g}_*$$

and $\mathrm{Tr}[\mathbf{E}^T \mathbf{\Upsilon}(\widehat{\mathbf{G}} - \mathbf{G})] = \boldsymbol{\eta}^T(\mathbf{I}_M \otimes \mathbf{S}^T \mathbf{\Upsilon})(\mathbf{\Pi}_{\hat{\mathbf{Z}}, \hat{K}, \hat{J}} \mathbf{\Gamma}\mathbf{y} - \mathbf{\Pi}_{\mathbf{Z}, K, J} \mathbf{g}_*)$. Now (6.1) can be rewritten in a vector form as

$$\|\hat{\mathbf{g}} - \mathbf{g}_*\|^2 \leq \|\mathbf{g} - \mathbf{g}_*\|^2 + \Delta + \mathrm{Pen}(J, K) - \mathrm{Pen}(\hat{J}, \hat{K}) \quad (6.6)$$

where

$$\Delta = 2\,\delta\boldsymbol{\eta}^T(\mathbf{I}_M \otimes \mathbf{S}^T \mathbf{\Upsilon})(\mathbf{\Pi}_{\hat{\mathbf{Z}}, \hat{K}, \hat{J}} \mathbf{\Gamma}\mathbf{y} - \mathbf{\Pi}_{\mathbf{Z}, K, J} \mathbf{g}_*) \quad (6.7)$$
$$= \Delta_1 + \Delta_2$$

with

$$\Delta_1 = 2\,\delta\boldsymbol{\eta}^T(\mathbf{I}_M \otimes \mathbf{S}^T \mathbf{\Upsilon})(\mathbf{\Pi}_{\hat{\mathbf{Z}}, \hat{K}, \hat{J}}(\mathbf{\Gamma}\mathbf{y} - \mathbf{g}_*)),$$
$$\Delta_2 = 2\,\delta\boldsymbol{\eta}^T(\mathbf{I}_M \otimes \mathbf{S}^T \mathbf{\Upsilon})(\mathbf{\Pi}_{\hat{\mathbf{Z}}, \hat{K}, \hat{J}} - \mathbf{\Pi}_{\mathbf{Z}, K, J})\mathbf{g}_*. \quad (6.8)$$

Derivation of upper bounds for Δ_1 and Δ_2 is based on the following lemma.

Lemma 1: Let K, J be fixed, $\hat{J}$ be an arbitrary random subset of $\{1, \ldots, n\}$ and $\hat{K}$ be a random integer between 1 and M. Let $\mathbf{Z} \in \mathcal{M}(M, K)$ and $\widehat{\mathbf{Z}} \in \mathcal{M}(M, \hat{K})$ be a fixed and a random clustering matrix, respectively. Denote the projection matrices on the column spaces of matrices $\mathbf{Z}$ and $\widehat{\mathbf{Z}}$ by, respectively, $\mathbf{\Pi}_{\mathbf{Z}, K}$ and $\mathbf{\Pi}_{\hat{\mathbf{Z}}, \hat{K}}$. Let $\mathbf{S}$ be a matrix with $\|\mathbf{S}\|_{op} \leq C_\psi$ and $\boldsymbol{\eta} \sim N(0, \mathbf{I}_{nM})$. Then, for any $\tau > 0$, there exist sets $\Omega_{1\tau}$ and $\Omega_{2\tau}$ with $\mathbb{P}(\Omega_{1\tau}) \geq 1 - \delta^\tau$ and

$\mathbb{P}(\Omega_{2\tau}) \geq 1 - \delta^{\tau}$ such that

$$\|(\mathbf{\Pi}_{\mathbf{Z},K} \otimes (\mathbf{W}_J \mathbf{\Upsilon S}))\boldsymbol{\eta}\|^2 \leq 2KC_{\psi}^2 (\sum_{j \in J} \nu_j^2) \qquad (6.9)$$

$$+ 3C_{\psi}^2 (\max_{j \in J} \nu_j^2) \tau \ln(\delta^{-1}), \quad \forall \omega \in \Omega_{1\tau};$$

$$\|(\mathbf{\Pi}_{\hat{\mathbf{Z}},\hat{K}} \otimes (\mathbf{W}_{\hat{j}} \mathbf{\Upsilon S}))\boldsymbol{\eta}\|^2 \leq 2\hat{K} C_{\psi}^2 (\sum_{j \in \hat{J}} \nu_j^2)$$

$$+ 3C_{\psi}^2 (\max_{j \in \hat{J}} \nu_j^2) \Big\{ M \ln \hat{K} + |\hat{J}| \ln(ne/|\hat{J}|) \qquad (6.10)$$

$$+ \ln(Mn) + \tau \ln(\delta^{-1}) \Big\} \quad \forall \omega \in \Omega_{2\tau}.$$

Moreover, if $J = \{1, \ldots, L\}$ is fixed and $\hat{J} = \{1, \ldots, \hat{L}\}$ for some random integer $\hat{L} \geq 1$, then

$$\|(\mathbf{\Pi}_{\mathbf{Z},K} \otimes (\mathbf{W}_J \mathbf{\Upsilon S}))\boldsymbol{\eta}\|^2 \leq 2KC_{\psi}^2 \sum_{j=1}^{L} \nu_j^2 \qquad (6.11)$$

$$+ 3C_{\psi}^2 \tau \ln(\delta^{-1}) \nu_L^2, \quad \forall \omega \in \Omega_{1\tau};$$

$$\|(\mathbf{\Pi}_{\hat{\mathbf{Z}},\hat{K}} \otimes (\mathbf{W}_{\hat{j}} \mathbf{\Upsilon S}))\boldsymbol{\eta}\|^2 \leq 2\hat{K} C_{\psi}^2 \sum_{j=1}^{L} \nu_j^2 \qquad (6.12)$$

$$+ 3C_{\psi}^2 \nu_L^2 \Big\{ M \ln \hat{K} + \ln(Mn) + \tau \ln(\delta^{-1}) \Big\} \quad \forall \omega \in \Omega_{2\tau}.$$

In what follows, we carry out only the proof of the upper bound (3.2) that takes place for a generic set J. The proof of the upper bound (3.3) can be obtained from the proof below with minimal modifications.

Note that Δ_1 can be re-written as $\Delta_1 = 2\delta^2 \boldsymbol{\eta}^T (\mathbf{I}_M \otimes \mathbf{S}^T \mathbf{\Upsilon})(\mathbf{\Pi}_{\hat{\mathbf{Z}},\hat{K}} \otimes \mathbf{W}_{\hat{j}})(\mathbf{I}_M \otimes \mathbf{\Upsilon S})\boldsymbol{\eta}$. Due to $\mathbf{\Gamma y} - \mathbf{g}_* = \delta \mathbf{\Gamma \epsilon}$ and (6.3), we obtain $\Delta_1 = 2\delta^2 \|(\mathbf{\Pi}_{\hat{\mathbf{Z}},\hat{K}} \otimes (\mathbf{W}_{\hat{j}} \mathbf{\Upsilon S}))\boldsymbol{\eta}\|^2$. Therefore, by (6.10), we obtain that for $\omega \in \Omega_{2\tau}$

$$|\Delta_1| \leq 2\delta^2 C_{\psi}^2 \Big[2\hat{K} \sum_{j \in \hat{J}} \nu_j^2 \qquad (6.13)$$

$$+ 3(\max_{j \in \hat{J}} \nu_j^2) \Big\{ M \ln \hat{K} + |\hat{J}| \ln(ne/|\hat{J}|) + \ln(Mn\delta^{-\tau}) \Big\} \Big]$$

In order to construct an upper bound for Δ_2, consider the following sets

$$\tilde{J} = J \cup \hat{J}, \quad J_1 = J \cap \hat{J}, \qquad (6.14)$$

$$J_2 = J^c \cap \hat{J}, \quad J_3 = \hat{J}^c \cap J.$$

The sets J_1, J_2 and J_3 are non-overlapping and $\tilde{J} = J_1 \cup J_2 \cup J_3$. Furthermore, consider matrix $\tilde{\mathbf{Z}}$ that includes all linearly independent columns in matrices $\mathbf{Z}_K$ and $\hat{\mathbf{Z}}_{\hat{K}}$, so that $\mathrm{Span}\{\tilde{\mathbf{Z}}\} = \mathrm{Span}\{\mathbf{Z}_K, \hat{\mathbf{Z}}_{\hat{K}}\}$. Let $\tilde{K}$ be the number of columns of matrix $\tilde{\mathbf{Z}}$. Then, one has

$$\mathbf{\Pi}_{\hat{\mathbf{Z}},\hat{K}} \mathbf{\Pi}_{\tilde{\mathbf{Z}},\tilde{K}} = \mathbf{\Pi}_{\tilde{\mathbf{Z}},\tilde{K}} \mathbf{\Pi}_{\hat{\mathbf{Z}},\hat{K}} = \mathbf{\Pi}_{\hat{\mathbf{Z}},\hat{K}},$$

$$\mathbf{W}_J = \mathbf{W}_{J_1} + \mathbf{W}_{J_3}, \quad \mathbf{W}_{\hat{j}} = \mathbf{W}_{J_1} + \mathbf{W}_{J_2},$$

$$\mathbf{W}_{\tilde{j}} = \mathbf{W}_{J_1} + \mathbf{W}_{J_2} + \mathbf{W}_{J_3}.$$

In order to obtain an upper bound for Δ_2 defined in (6.8), note that using notations above, we can rewrite Δ_2 as

$$\Delta_2 = 2\delta \boldsymbol{\eta}^T (\mathbf{I}_M \otimes \mathbf{S}^T \mathbf{\Upsilon})[(\mathbf{\Pi}_{\hat{\mathbf{Z}},\hat{K}} \otimes \mathbf{W}_{J_2}) + (\mathbf{\Pi}_{\hat{\mathbf{Z}},\hat{K}} \otimes \mathbf{W}_{J_1})$$

$$- (\mathbf{\Pi}_{\mathbf{Z},K} \otimes \mathbf{W}_{J_1}) - (\mathbf{\Pi}_{\mathbf{Z},K} \otimes \mathbf{W}_{J_3})]\mathbf{g}_*$$

$$= 2\delta \boldsymbol{\eta}^T (\mathbf{I}_M \otimes \mathbf{S}^T \mathbf{\Upsilon})[(\mathbf{\Pi}_{\hat{\mathbf{Z}},\hat{K}} \otimes \mathbf{W}_{J_2}) + (\mathbf{\Pi}_{\tilde{\mathbf{Z}},\tilde{K}} \otimes \mathbf{W}_{J_1})$$

$$+ (\mathbf{\Pi}_{\mathbf{Z},K} \otimes \mathbf{W}_{J_3})][(\mathbf{\Pi}_{\hat{\mathbf{Z}},\hat{K}} \otimes \mathbf{W}_{J_2})$$

$$+ (\mathbf{\Pi}_{\hat{\mathbf{Z}},\hat{K}} \otimes \mathbf{W}_{J_1}) - (\mathbf{\Pi}_{\mathbf{Z},K} \otimes \mathbf{W}_{J_1}) - (\mathbf{\Pi}_{\mathbf{Z},K} \otimes \mathbf{W}_{J_3})]\mathbf{g}_*$$

$$= 2\delta \boldsymbol{\eta}^T (\mathbf{I}_M \otimes \mathbf{S}^T \mathbf{\Upsilon})[(\mathbf{\Pi}_{\hat{\mathbf{Z}},\hat{K}} \otimes \mathbf{W}_{J_2})$$

$$+ (\mathbf{\Pi}_{\tilde{\mathbf{Z}},\tilde{K}} \otimes \mathbf{W}_{J_1}) + (\mathbf{\Pi}_{\mathbf{Z},K} \otimes \mathbf{W}_{J_3})][(\mathbf{\Pi}_{\hat{\mathbf{Z}},\hat{K},\hat{j}} - \mathbf{\Pi}_{\mathbf{Z},K,J})]\mathbf{g}_*$$

Using Cauchy inequality and $2ab \leq 4\,a^2 + b^2/4$, we obtain

$$|\Delta_2| \leq |\Delta_{2,1}| + |\Delta_{2,2}|, \qquad (6.15)$$

$$|\Delta_{2,2}| = 0.25 \|(\mathbf{\Pi}_{\hat{\mathbf{Z}},\hat{K},\hat{j}} \mathbf{g}_* - \mathbf{\Pi}_{\mathbf{Z},K,J} \mathbf{g}_*)\|^2,$$

$$|\Delta_{2,1}| = 4\delta^2 \|[(\mathbf{\Pi}_{\hat{\mathbf{Z}},\hat{K}} \otimes \mathbf{W}_{J_2}) + (\mathbf{\Pi}_{\tilde{\mathbf{Z}},\tilde{K}} \otimes \mathbf{W}_{J_1})$$

$$+ (\mathbf{\Pi}_{\mathbf{Z},K} \otimes \mathbf{W}_{J_3})](\mathbf{I}_M \otimes \mathbf{\Upsilon S})\boldsymbol{\eta}\|^2.$$

Applying Cauchy inequality to the term $\Delta_{2,1}$ and using that $J_2 \subseteq \hat{J}$ and $J_3 \subseteq J$ we rewrite

$$|\Delta_{2,1}| \leq 12\delta^2 \Big[\|(\mathbf{\Pi}_{\hat{\mathbf{Z}},\hat{K}} \otimes (\mathbf{W}_{\hat{j}} \mathbf{\Upsilon S}))\boldsymbol{\eta}\|^2$$

$$+ \|(\mathbf{\Pi}_{\tilde{\mathbf{Z}},\tilde{K}} \otimes (\mathbf{W}_{J_1} \mathbf{\Upsilon S}))\boldsymbol{\eta}\|^2 + \|(\mathbf{\Pi}_{\mathbf{Z},K} \otimes (\mathbf{W}_J \mathbf{\Upsilon S}))\boldsymbol{\eta}\|^2 \Big]$$

The upper bounds for the first and the third term in the inequality above can be obtained directly from Lemma 1. For the second term, note that since $\tilde{K} \leq K + \hat{K}$ and $J_1 \subseteq J$ and $J_1 \subseteq \hat{J}$ for any $\omega \in \Omega_{1\tau} \cap \Omega_{2\tau}$ one has

$$\|(\mathbf{\Pi}_{\tilde{\mathbf{Z}},\tilde{K}} \otimes (\mathbf{W}_{J_1} \mathbf{\Upsilon S}))\boldsymbol{\eta}\|^2 \leq C_{\psi}^2 \Bigg[2K \sum_{j \in J} \nu_j^2$$

$$+ 2\hat{K} \sum_{j \in \hat{J}} \nu_j^2 + 3(\max_{j \in \hat{J}} \nu_j^2) \Big\{ M \ln \hat{K} \qquad (6.16)$$

$$+ |\hat{J}| \ln\left(\frac{ne}{|\hat{J}|}\right) + \ln(Mn) + \tau \ln(\delta^{-1}) \Big\} \Bigg]$$

due to

$$\tilde{K} \sum_{j \in J_1} \nu_j^2 \leq K \sum_{j \in J} \nu_j^2 + \hat{K} \sum_{j \in \hat{J}} \nu_j^2.$$

Combining (6.16) with equations (6.9) and (6.10), we obtain for any $\omega \in \Omega_{1\tau} \cap \Omega_{2\tau}$

$$|\Delta_{2,1}| \leq 12\delta^2 C_{\psi}^2 \Bigg[4\hat{K} \sum_{j \in \hat{J}} \nu_j^2 + 4K \sum_{j \in J} \nu_j^2$$

$$+ 3(\max_{j \in J} \nu_j^2)(\tau \ln n) + 6(\max_{j \in \hat{J}} \nu_j^2) \Big\{ M \ln \hat{K} \qquad (6.17)$$

$$+ |\hat{J}| \ln\left(\frac{ne}{|\hat{J}|}\right) + \ln(Mn) + \tau \ln(\delta^{-1}) \Big\} \Bigg]$$

Now consider $|\Delta_{2,2}|$ defined in (6.15). Rewrite $|\Delta_{2,2}|$ as $|\Delta_{2,2}| = 0.25 \|(\mathbf{\Pi}_{\hat{\mathbf{Z}},\hat{K},\hat{j}} \mathbf{g}_* - \mathbf{g}_*) - (\mathbf{\Pi}_{\mathbf{Z},K,J} \mathbf{g}_* - \mathbf{g}_*)\|^2$, so that

$$|\Delta_{2,2}| \leq 0.5 \|(\mathbf{\Pi}_{\hat{\mathbf{Z}},\hat{K},\hat{j}} \mathbf{g}_* - \mathbf{g}_*)\|^2 + 0.5 \|(\mathbf{\Pi}_{\mathbf{Z},K,J} \mathbf{g}_* - \mathbf{g}_*)\|^2.$$

Since $\hat{\mathbf{g}} = \mathbf{\Pi}_{\hat{\mathbf{Z}}, \hat{K}, \hat{J}} \mathbf{\Gamma} \mathbf{y}$ and

$$\|(\mathbf{\Pi}_{\hat{\mathbf{Z}}, \hat{K}, \hat{J}} \mathbf{\Gamma} \mathbf{y} - \mathbf{g}_*)\|^2 = \|(\mathbf{\Pi}_{\hat{\mathbf{Z}}, \hat{K}, \hat{J}} (\mathbf{g}_* + \delta \, \mathbf{\Gamma} \boldsymbol{\epsilon}) - \mathbf{g}_*)\|^2$$
$$= \|(\mathbf{I} - \mathbf{\Pi}_{\hat{\mathbf{Z}}, \hat{K}, \hat{J}}) \mathbf{g}_*\|^2 + \delta^2 \|\mathbf{\Pi}_{\hat{\mathbf{Z}}, \hat{K}, \hat{J}} \, \mathbf{\Gamma} \boldsymbol{\epsilon}\|^2,$$

we derive

$$\|\hat{\mathbf{g}} - \mathbf{g}_*\|^2 \geq \|\mathbf{\Pi}_{\hat{\mathbf{Z}}, \hat{K}, \hat{J}} \mathbf{g}_* - \mathbf{g}_*\|^2 \qquad (6.18)$$

Taking into account that $\mathbf{g} = \mathbf{\Pi}_{\mathbf{Z}, K, J} \mathbf{g}_*$, so that $\|\mathbf{g} - \mathbf{g}_*\|^2 = \|\mathbf{\Pi}_{\mathbf{Z}, K, J} \mathbf{g}_* - \mathbf{g}_*\|^2$, we obtain

$$|\Delta_{2,2}| \leq 0.5 \|\hat{\mathbf{g}} - \mathbf{g}_*\|^2 + 0.5 \|\mathbf{g} - \mathbf{g}_*\|^2. \qquad (6.19)$$

By combining upper bounds of Δ_1, $\Delta_{2,1}$ and $\Delta_{2,2}$, we derive from (6.13) and (6.17)–(6.19) that for any $\omega \in \Omega_{1\tau} \cap \Omega_{2\tau}$, an upper bound for Δ can be written as

$$|\Delta| \leq 0.5 \|\hat{\mathbf{g}} - \mathbf{g}_*\|^2 + 0.5 \|\mathbf{g} - \mathbf{g}_*\|^2$$
$$+ 2 \delta^2 C_\psi^2 \left\{ 26 \hat{K} \sum_{j \in \hat{J}} \nu_j^2 + 24 K \sum_{j \in J} \nu_j^2 \right. \qquad (6.20)$$
$$+ 39 \left(\max_{j \in \hat{J}} \nu_j^2 \right) \left[M \ln \hat{K} + |\hat{J}| \ln \left(\frac{ne}{|\hat{J}|} \right) \right.$$
$$+ \ln(Mn) + \tau \ln(\delta^{-1}) \Big] + 18 (\max_{j \in J} \nu_j^2) \tau \ln(\delta^{-1}) \bigg\}$$

Since it follows from (6.3) that $\|\widehat{\mathbf{G}} - \mathbf{G}_*\|_F^2 = \|\hat{\mathbf{g}} - \mathbf{g}_*\|^2$, we obtain from (6.6) that for any $\mathbf{G} = \mathbf{\Pi}_{\mathbf{Z}, K, J} \mathbf{G}_*$ on the set $\Omega_{1\tau} \cap \Omega_{2\tau}$ one has

$$\|\widehat{\mathbf{G}} - \mathbf{G}_*\|_F^2 \leq 3 \|\mathbf{G} - \mathbf{G}_*\|_F^2 + 2 \delta^2 C_\psi^2 \left\{ 48 K \sum_{j \in J} \nu_j^2 \right.$$
$$+ 36 (\max_{j \in J} \nu_j^2) \tau \ln \delta^{-1} + 52 \hat{K} \sum_{j \in \hat{J}} \nu_j^2 \qquad (6.21)$$
$$+ 78 (\max_{j \in \hat{J}} \nu_j^2) \left[M \ln \hat{K} + |\hat{J}| \ln \left(\frac{ne}{|\hat{J}|} \right) \right.$$
$$+ \ln(Mn) + \tau \ln \delta^{-1} \Big] \bigg\} + 2 [\mathrm{Pen}(J, K) - \mathrm{Pen}(\hat{J}, \hat{K})]$$

Choose $\mathrm{Pen}(J, K)$ in the form (2.23) and note that all terms containing $\hat{J}$ and $\hat{K}$ in (6.21) cancel. Finally we obtained for any $\mathbf{G} = \mathbf{W}_J \mathbf{G}_* \mathbf{\Pi}_{\mathbf{Z}, K}$ that with probability at least $1 - 2\delta^\tau$

$$\|\widehat{\mathbf{G}} - \mathbf{G}_*\|_F^2 \leq 3 \|\mathbf{G} - \mathbf{G}_*\|_F^2 +$$
$$2 \delta^2 C_\psi^2 \left\{ 48 K \sum_{j \in J} \nu_j^2 + 36 (\max_{j \in J} \nu_j^2) \tau \ln n \right\} + 2 \, \mathrm{Pen}(J, K)$$

which yields (3.2).

B. Proof of the Upper Bounds for the Error

Proof of Theorem 2: Since, when j is growing, coefficients Θ_{jk} are decreasing while the values of ν_j are increasing according to (2.13), the optimal set J is of the form $J = \{1, \cdots, L\}$, so that $|J| = L$. Then, we find $(\hat{\mathbf{Z}}, \widehat{\mathbf{G}}, \hat{L}, \hat{K})$ as a solution of optimization problem (2.19) with the penalty given by expression (2.24).

Note that for the true number of classes K_* with $N_k, k = 1, \ldots, K_*$ elements in each class, $\mathbf{G}$ are coefficients of each

f_m and $\boldsymbol{\Theta}$ is the clustered version of those coefficients. It follows from (2.4) that

$$R(\hat{f}, \mathcal{S}(r, \mathcal{A}), M, K_*) \leq M^{-1} \|\widehat{\mathbf{G}} - \mathbf{G}_*\|_F^2 \qquad (6.22)$$
$$+ M^{-1} \sum_{k=1}^{K_*} N_k \sum_{j=n+1}^{\infty} \boldsymbol{\Theta}_{jk}^2.$$

Therefore, application of the upper bound (3.3) with a generic L, $\mathbf{Z} = \mathbf{Z}_*$, $K = K_*$, where $\mathbf{Z}_*$ and K_* are respectively the true clustering matrix and the true number of classes, yields

$$M^{-1} \|\widehat{\mathbf{G}} - \mathbf{G}_*\|_F^2 \leq 3 M^{-1} \|\mathbf{W}_J \mathbf{G}_* \mathbf{\Pi}_{\mathbf{Z}_*, K_*} - \mathbf{G}_*\|_F^2$$
$$+ 4 M^{-1} \overline{\mathrm{Pen}}(L, K_*) \qquad (6.23)$$

where $\overline{\mathrm{Pen}}(L, K)$ is defined in (2.24). Observe that

$$\|\mathbf{W}_J \mathbf{G}_* \mathbf{\Pi}_{\mathbf{Z}_*, K_*} - \mathbf{G}_*\|_F^2 = \|(\mathbf{W}_J - \mathbf{I}_n) \mathbf{G}_*\|_F^2 \qquad (6.24)$$
$$= \sum_{k=1}^{K^*} N_k \sum_{j=L+1}^{n} \boldsymbol{\Theta}_{jk}^2$$

where N_k is the number of functions $f_m = h_k$ in the cluster k, $k = 1, \cdots, K^*$, and $\boldsymbol{\Theta}_{jk}$ are the true coefficients of those functions. Hence, it follows from (2.11) that

$$\sum_{j=L+1}^{n} \boldsymbol{\Theta}_{jk}^2 \leq \mathcal{A}^2 \, L^{-2r}. \qquad (6.25)$$

Since $\sum_{k=1}^{K^*} N_k = M$, (6.24) and (6.25) yield

$$\|\mathbf{W}_J \mathbf{G}_* \mathbf{\Pi}_{\mathbf{Z}_*, K_*} - \mathbf{G}_*\|_F^2 \leq \mathcal{A}^2 \, M L^{-2r} \qquad (6.26)$$

Moreover, it follows from (2.12) that

$$M^{-1} \sum_{k=1}^{K_*} N_k \sum_{j=n+1}^{\infty} \boldsymbol{\Theta}_{jk}^2 \leq \mathcal{A}^2 n^{-2r} \asymp \delta^2,$$

so that the last term in (6.22) is smaller than $C \, R(M, K_*, \delta)$.

Now, consider the second term in (6.23). Due to the condition (2.13), one obtains

$$\nu_L^2 \leq \aleph_2^2 \, L^{2\gamma} \exp \left(2\alpha L^\beta\right), \quad \sum_{j=1}^{L} \nu_j^2 \leq \aleph_2^2 \, L^{2\gamma+1} \exp \left(2\alpha L^\beta\right).$$

Denote

$$R_1 \equiv R_1(K_*, \delta) \asymp K_*, \qquad (6.27)$$
$$R_2 \equiv R_2(M, K_*, \delta) \asymp M \ln K_* + \ln(\delta^{-1}).$$

Therefore, it follows from (2.23) and (3.2) that, under condition (3.5),

$$\frac{\|\widehat{\mathbf{G}} - \mathbf{G}_*\|_F^2}{M} \leq \tilde{C} \min_L \left\{ L^{-2r} + \qquad (6.28) \right.$$
$$\frac{\delta^2 L^{2\gamma} \exp(2\alpha L^\beta)}{M} [L R_1(K_*, \delta) + R_2(M, K_*, \delta)] \right\}$$

where $R_1(K_*, \delta)$ and $R_2(M, K_*, \delta)$ are defined in (6.27) and $\tilde{C}$ depends only on μ, $\mathcal{A}$, $\aleph_2$, C_ψ^2 and is independent of M, L, δ and K_*.

In order to find the minimum of the right hand side of (6.28), denote

$$R(L, M, K_*, \delta) = \delta^2 M^{-1} \exp\left(2\alpha L^\beta\right) \left[L^{2\gamma+1} R_1 \quad (6.29)\right.$$
$$\left. + L^{2\gamma} R_2\right] + L^{-2r}$$

and observe that

$$M^{-1} \|\widehat{\mathbf{G}} - \mathbf{G}_*\|_F^2 \leq \tilde{C} \min_L R(L, M, K_*, \delta) \qquad (6.30)$$

where L_{opt} is the value of L minimizing the right hand side of (6.28). Denote

$$L_{1,opt} = \underset{L}{\operatorname{argmin}}[L^{-2r} + \delta^2 M^{-1} \exp\left(2\alpha L^\beta\right) L^{2\gamma+1} R_1],$$
$$(6.31)$$
$$L_{2,opt} = \underset{L}{\operatorname{argmin}}[L^{-2r} + \delta^2 M^{-1} \exp\left(2\alpha L^\beta\right) L^{2\gamma} R_2].$$
$$(6.32)$$

It is easy to see that since the first terms in expressions (6.31) and (6.32) are decreasing in L while the second terms are increasing, the values $L_{1,opt}$ and $L_{2,opt}$ are such that those terms are equal to each other up to a multiplicative constant. Then, $R(L_{opt}, M, K_*, \delta) = \max\left\{L_{1,opt}^{-2r}, L_{2,opt}^{-2r}\right\}$, and, due to $\max(a, b) \asymp a + b$ for positive a and b, we obtain

$$R(L_{opt}, M, K_*, \delta) \asymp L_{1,opt}^{-2r} + L_{2,opt}^{-2r}. \qquad (6.33)$$

Consider two cases.
Case 1: $\alpha = \beta = 0$. Direct calculations yield

$$L_{1,opt} \asymp \left(M^{-1} \delta^2 R_1\right)^{-\frac{1}{2\gamma+2r+1}}, \ L_{2,opt} \asymp \left(M^{-1} \delta^2 R_2\right)^{-\frac{1}{2\gamma+2r}},$$

so that, due to (6.27),

$$L_{1,opt} = (M^{-1} \delta^2 K_*)^{-\frac{1}{2\gamma+2r+1}},$$
$$L_{2,opt} = [\delta^2(\ln K_* + M^{-1} \ln \delta^{-1})]^{-\frac{1}{2\gamma+2r}}$$

Then, by (6.33),

$$R(L_{opt}, M, K_*, \delta) \asymp (M^{-1} \delta^2 K_*)^{\frac{2r}{2\gamma+2r+1}} \qquad (6.34)$$
$$+ [\delta^2(\ln K_* + M^{-1} \ln \delta^{-1})]^{\frac{2r}{2\gamma+2r}}.$$

Now, in order to obtain the expression (3.6), note that if $K_* \geq 2$, then $\ln K_*$ dominates $M^{-1} \ln \delta^{-1}$. If $K_* = 1$, then (6.34) can be re-written as

$$R(L_{opt}, M, K_*, \delta) \asymp \left(\frac{\delta^2}{M}\right)^{\frac{2r}{2\gamma+2r+1}} [1+$$
$$\left(\frac{\delta^2}{M}\right)^{\frac{2r}{(2\gamma+2r+1)(2r+2\gamma)}} (\ln \delta^{-1})^{\frac{2r}{2\gamma+2r}}]$$
$$\asymp \left(\frac{\delta^2 K_*}{M}\right)^{\frac{2r}{2\gamma+2r+1}},$$

which yields (3.6).
Case 2: $\alpha > 0, \beta > 0$. Minimizing expressions in (6.31) and (6.32), we obtain

$$L_{i,opt} \asymp \left\{\left[\ln\left(\frac{M}{\delta^2 R_i}\right)\right]\right\}^{\frac{1}{\beta}}, \quad i = 1, 2,$$

If $K_* \geq 2$, then $R_2 \geq R_1$. Taking into account that, under assumption (3.5), for large M and small δ, $\ln\left(M\delta^{-2} \ln M\right) \asymp \ln\left(M\delta^{-2}\right)$ and $\ln(Mn) \asymp \ln M$, we obtain

$$L_{1,opt} = \min\left\{\left[\ln\left(\frac{M}{\delta^2 K_*}\right)\right]; \left[\ln\left(\frac{M}{\delta^2 \ln M}\right)\right]\right\}^{\frac{1}{\beta}}$$
$$\asymp \left[\ln\left(\frac{M}{\delta^2 K_*}\right)\right]^{\frac{1}{\beta}}.$$

Similarly,

$$L_{2,opt} = \min\left\{\left[\ln\left(\frac{1}{\delta^2 \ln K_*}\right)\right]; \left[\ln\left(\frac{M}{\delta^2 \ln M}\right)\right]\right\}^{\frac{1}{\beta}}$$
$$\asymp \left[\ln\left(\frac{1}{\delta^2 \ln K_*}\right)\right]^{\frac{1}{\beta}},$$

which, together with (6.30) and (6.33), yield the expression (3.7). One can easily check that the case of $K_* = 1$ leads to the same results.

C. Proofs of the Minimax Lower Bounds for the Error

Proof of Theorem 3: Since the estimation error is comprised of the error due to non-parametric estimation and to clustering, we consider two cases here.
Lower bound for the error due to clustering.
Let $K \geq 2$ be the fixed number of classes. Consider a subset $\mathcal{Z}(M, K) \subset \mathcal{M}(M, K)$ of the set of all clustering matrices which contains all matrices that cluster $\frac{M}{K}$ vectors into each class. By Lemma 5 in Pensky (2019) with $\gamma = 1$, obtain that the cardinality of the set $\mathcal{Z}(M, K)$ is

$$|\mathcal{Z}(M, K)| = M!\Big/[(M/K)!]^K \geq \exp\left(M \ln K/4\right) \quad (6.35)$$

Let set J be of the form $J = \{L_1, \ldots, L_2\}$ where $1 \leq L_1 < L_2 \leq n$ and $n = [\delta^{-2}]$. Choose $\boldsymbol{\Theta}_{jk} = 0$ if $j \notin J$. In what follows, we use the Packing Lemma (Lemma 4 of Pensky (2019)):
Lemma 2 (The Packing Lemma): Let $\mathcal{Z}(M, K) \subseteq \mathcal{M}(M, K)$ be a collection of clustering matrices and q be a positive constant. Then, there exists a subset $\mathcal{S}_{M,K}(q) \subset \mathcal{Z}(M, K)$ such that for $\mathbf{Z}_1, \mathbf{Z}_2 \in \mathcal{S}_{M,K}(q)$ one has $\|\mathbf{Z}_1 - \mathbf{Z}_2\|_H = \|\mathbf{Z}_1 - \mathbf{Z}_2\|_F^2 \geq q$ and $\ln |\mathcal{S}_{M,K}(q)| \geq \ln |\mathcal{Z}(M, K)| - q \ln(MKe/q)$.
Apply this lemma with $q = dM$, $0 < d < 1/4$. Then, by (6.35), derive

$$\ln |\mathcal{S}_{M,K}\left(dM\right)| \geq M \left[\ln K - 4\, d \ln(Ke/d)\right]/4.$$

Use the following statement:
Lemma 3: If $K \geq 2$ and d is such that

$$d - d \ln d \leq (\ln 2)/9, \quad d \leq 1/9, \qquad (6.36)$$

then $\ln K - 4d \ln(Ke/d) \geq (\ln K)/9$.
It is easy to calculate that, e.g., $d = 0.0147$ satisfies the condition (6.36). Then, for d obeying (6.36), one has

$$\ln |\mathcal{S}_{M,K}(dM)| \geq \frac{M}{36} \ln K, \qquad (6.37)$$
$$\|\mathbf{Z}_1 - \mathbf{Z}_2\|_H \geq dM \text{ for any } \mathbf{Z}_1, \mathbf{Z}_2 \in \mathcal{S}_{M,K}(dM), \ \mathbf{Z}_1 \neq \mathbf{Z}_2$$

Consider a collection of binary vectors $\boldsymbol{\omega} \in \{0,1\}^{|J|}$. By Varshamov-Gilbert bound lemma, there exists a subset $\mathcal{W}$ of those vectors such that, for any $\boldsymbol{\omega}, \boldsymbol{\omega}' \in \mathcal{W}$ such that $\boldsymbol{\omega} \neq \boldsymbol{\omega}'$ one has $\|\boldsymbol{\omega} - \boldsymbol{\omega}'\|_H \geq |J|/8$ and $\ln |\mathcal{W}| \geq |J| \ln(2)/8$. Choose a subset $\mathcal{W}_K$ of $\mathcal{W}$ such that $|\mathcal{W}_K| = K$. This is possible if $K \leq 2^{|J|/8}$ which is equivalent to $|J| \geq 8 \ln K / \ln 2$. Consider a set of vectors $\mathbf{w} \in \{0,1\}^n$ obtained by packing $\boldsymbol{\omega}$ with zeros for components not in J. Then

$$\mathcal{W}_K = \{\mathbf{w}_1, \ldots, \mathbf{w}_K \in \{0,1\}^n : \|\mathbf{w}_i\|_0 \leq |J|, \quad (6.38)$$
$$\|\mathbf{w}_i - \mathbf{w}_j\|_0 \geq |J|/8, \ i \neq j\}$$

Define matrix $\mathbf{W}$ with columns $\mathbf{w}_k$, $k = 1, \ldots, K$. Finally, form the set $\mathcal{G}_{M,K}$ of matrices $\mathbf{G}$ of the form

$$\mathcal{G}_{M,K} = \left\{\mathbf{G} \in R^{n \times M} : \mathbf{G} = \theta \mathbf{W} \mathbf{Z}^T, \mathbf{Z} \in \mathcal{S}_{M,K}(dM)\right\}$$

where d satisfies (6.36) and $\theta > 0$ depends on M, δ and K. Note that, due to (6.37), one has

$$\ln |\mathcal{G}_{M,K}| \geq (M \ln K)/36 \quad (6.39)$$

Let $\mathbf{Z}_1, \mathbf{Z}_2 \in \mathcal{S}_{M,K}$ be two clustering matrices. Set $\mathbf{G}_1 = \theta \mathbf{W} \mathbf{Z}_1^T$ $\mathbf{G}_2 = \theta \mathbf{W} \mathbf{Z}_2^T$, so that $\mathbf{G}_1, \mathbf{G}_2 \in \mathcal{G}_{M,K}$. Since for any i, i' one has $\|\mathbf{w}_i - \mathbf{w}_{i'}\|_0 = \|\mathbf{w}_i - \mathbf{w}_{i'}\|^2$, derive that

$$\|\theta \mathbf{W} (\mathbf{Z}_1 - \mathbf{Z}_2)^T\|_F^2 = \sum_{m=1}^{M} \sum_{j=1}^{n} \theta^2 \left[(\mathbf{w}_{z_1(m)})_j - (\mathbf{w}_{z_2(m)})_j\right]^2$$
$$= \theta^2 \sum_{m=1}^{M} \|\mathbf{w}_{z_1(m)} - \mathbf{w}_{z_2(m)}\|^2 \quad (6.40)$$
$$\geq \#\{m : z_1(m) \neq z_2(m)\} \theta^2 |J|/8.$$

On the other hand, observe that for $\mathbf{Z}_1, \mathbf{Z}_2 \in \mathcal{S}_{M,K}$ one has

$$\#\{m : z_1(m) \neq z_2(m)\} = 0.5 \|\mathbf{Z}_1 - \mathbf{Z}_2\|_H \geq dM/2.$$

Therefore, the last two inequalities yield for any $\mathbf{G}_1, \mathbf{G}_2 \in \mathcal{G}_{M,K}$

$$\|\mathbf{G}_1 - \mathbf{G}_2\|_F^2 \geq d \theta^2 |J| M/16. \quad (6.41)$$

Now, it is easy to calculate that for any $\mathbf{G}_1, \mathbf{G}_2 \in \mathcal{G}_{M,K}$ and the corresponding probability measures $P_{\mathbf{G}_1}$ and $P_{\mathbf{G}_2}$ associated with $\mathbf{Y} = \boldsymbol{\Upsilon}^{-1} \mathbf{G}_i + \delta \mathbf{E}$, $i = 1, 2$, in (2.5), one has the following inequality for the Kullback-Leibler divergence between $P_{\mathbf{G}_1}$ and $P_{\mathbf{G}_2}$:

$$K(P_{\mathbf{G}_1}, P_{\mathbf{G}_2}) \leq \frac{1}{2\delta^2 C_\psi^2} \|\boldsymbol{\Upsilon}^{-1} (\mathbf{G}_2 - \mathbf{G}_1)\|_F^2 \quad (6.42)$$

Since $\mathbf{G}_1 = \theta \mathbf{W} \mathbf{Z}_1$, $\mathbf{G}_2 = \theta \mathbf{W} \mathbf{Z}_2$, we obtain

$$\|\boldsymbol{\Upsilon}^{-1} (\mathbf{G}_2 - \mathbf{G}_1)\|_F^2 \leq \theta^2 \|\mathbf{Z}_2 - \mathbf{Z}_1\|_{op}^2 \|\boldsymbol{\Upsilon}^{-1} \mathbf{W}\|_F^2 \quad (6.43)$$

Note that $\mathcal{S}_{M,K}(dM) \subset \mathcal{Z}(M,K)$, so that for any $\mathbf{Z} \in \mathcal{S}_{M,K}(dM)$ one has $\mathbf{Z}^T \mathbf{Z} = (M/K) \mathbf{I}_K$, hence $\|\mathbf{Z}\|_{op} = \sqrt{M/K}$. Then, $\|\mathbf{Z}_1 - \mathbf{Z}_2\|_{op}^2 \leq 4M/K$. Also, due to $J = \{L_1, \ldots, L_2\}$ and condition (2.13), one has

$$\sum_{j \in J} \nu_j^{-2} \leq \aleph_1^{-2} |J| L_1^{-2\gamma} \exp\left(-2\alpha L_1^\beta\right). \quad (6.44)$$

Since $\|\boldsymbol{\Upsilon}^{-1} \mathbf{W}\|_F^2 \leq \sum_{k=1}^{K} \sum_{j \in J} \nu_j^{-2}$, obtain

$$K(P_{\mathbf{G}_1}, P_{\mathbf{G}_2}) \leq \frac{2\theta^2 |J| M}{\delta^2 \aleph_1^2 C_\psi^2} L_1^{-2\gamma} \exp\left(-2\alpha L_1^\beta\right). \quad (6.45)$$

Finally, due to condition (2.11), one needs $\theta^2 \sum_{j \in J} (j+1)^{2r} \leq \mathcal{A}^2$, so that we can choose

$$\theta^2 = \mathcal{A}^2 |J|^{-1} L_2^{-2r}. \quad (6.46)$$

In order to apply Theorem 2.5 of Tsybakov (2009) with $\alpha = 1/9$, we need $K(P_{\mathbf{G}_1}, P_{\mathbf{G}_2}) \leq \ln |\mathcal{G}_{M,K}|/9$ which, due to (6.37), is guaranteed by

$$\frac{\theta^2 |J|}{\delta^2 \aleph_1^2 C_\psi^2} L_1^{-2\gamma} \exp\left(-2\alpha L_1^\beta\right) \leq \frac{\ln K}{648}. \quad (6.47)$$

If inequality (6.47) holds, then application of Theorem 2.5 of Tsybakov (2009) yields that, with probability at least 0.1, one has (3.9) where, due to (3.1) and (6.41),

$$R_{\min}(M, K_*, \delta) = \theta^2 |J|. \quad (6.48)$$

Consider $L_1 = L/2 + 1$ and $L_2 = L$, so that

$$\theta^2 \asymp L^{-(2r+1)}, \quad R_{\min}(M, K_*, \delta) \asymp L^{-2r}. \quad (6.49)$$

If $\alpha = 0$, $\beta = 0$, then, by (6.49), inequality (6.47) holds if $L \asymp (\delta^2 \ln K)^{-\frac{1}{2r+2\gamma}}$. Hence,

$$R_{\min}(M, K_*, \delta) \gtrsim (\delta^2 \ln K_*)^{\frac{2r}{2r+2\gamma}}. \quad (6.50)$$

If $\alpha > 0$, $\beta > 0$, then inequality (6.47) holds if $L^{-(2\gamma+2r)} \exp(-2\alpha L^\beta) \lesssim \delta^2 \ln K$, so that $L \asymp \left[\ln\left(\frac{1}{\delta^2 \ln K}\right)\right]^{\frac{1}{\beta}}$. Therefore,

$$R_{\min}(M, K_*, n) \gtrsim \left[\ln\left(\frac{1}{\delta^2 \ln K_*}\right)\right]^{-\frac{2r}{\beta}}. \quad (6.51)$$

Lower bound for the error due to estimation.
Let, as before, $n = [\delta^{-2}]$ and $J = \{L_1, \ldots, L_2\}$ where $1 \leq L_1 < L_2 \leq n$. Consider a set of binary vectors $\boldsymbol{\omega} \in \{0,1\}^{|J|K}$ and set $N = |J|K$. Complete vectors $\boldsymbol{\omega}$ with zeros to obtain vectors $\mathbf{w} \in \{0,1\}^{nK}$. By Varshamov-Gilbert lemma, there exists a subset $\mathcal{B}$ of those vectors such that for any $\mathbf{w}, \mathbf{w}' \in \mathcal{B}$ such that $\mathbf{w} \neq \mathbf{w}'$ one has $\|\mathbf{w} - \mathbf{w}'\|_H \geq N/8$ and $\ln |\mathcal{B}| \geq N \ln(2)/8$. Pack vectors $\mathbf{w}$ into matrices $\mathbf{W} \in \{0,1\}^{n \times K}$. Denote the set of those matrices by $\mathcal{W}$ and observe that

$$\|\mathbf{W}_1 - \mathbf{W}_2\|_F^2 \geq N/8, \ \forall \mathbf{W}_1, \mathbf{W}_2 \in \mathcal{W}, \ \mathbf{W}_1 \neq \mathbf{W}_2;$$
$$\ln |\mathcal{W}| \geq (N \ln 2)/8. \quad (6.52)$$

Let $\mathbf{Z}$ be the clustering matrix that corresponds to uniform sequential clustering, M/K vectors per class. Finally, form the set $\mathcal{G}_{M,K}$ of matrices $\mathbf{G}$ of the form

$$\mathcal{G}_{M,K} = \left\{\mathbf{G} \in R^{M \times K} : \mathbf{G} = \theta \mathbf{W} \mathbf{Z}^T, \ \mathbf{W} \in \mathcal{W}\right\}$$

where $\theta > 0$ depends on M, δ and K. Then, for any $\mathbf{G}_1, \mathbf{G}_2 \in \mathcal{G}_{M,K}$, $\mathbf{G}_1 \neq \mathbf{G}_2$, due to $\mathbf{Z}^T \mathbf{Z} = (M/K) \mathbf{I}_K$ and (6.52), obtain

$$\|(\mathbf{G}_1 - \mathbf{G}_2)\|_F^2 = \theta^2 \|(\mathbf{W}_1 - \mathbf{W}_2) \mathbf{Z}^T\|_F^2 = \quad (6.53)$$
$$\frac{\theta^2 M}{K} \|\mathbf{W}_1 - \mathbf{W}_2\|_F^2 \geq \frac{\theta^2 M N}{8K}$$

Now, since $\mathbf{G}_1 = \theta \mathbf{W}_1 \mathbf{Z}$ and $\mathbf{G}_2 = \theta \mathbf{W}_2 \mathbf{Z}$, using formula (6.42), derive that

$$K(P_{\mathbf{G}_1}, P_{\mathbf{G}_2}) \leq \frac{\theta^2}{2\delta^2 C_\psi^2} \|\boldsymbol{\Upsilon}^{-1} (\mathbf{W}_2 - \mathbf{W}_1)\|_F^2 \|\mathbf{Z}\|_{op}^2$$

Recalling that $\|\mathbf{Z}\|_{op}^2 = M/K$ and $\|\mathbf{\Upsilon}^{-1}(\mathbf{W}_2 - \mathbf{W}_1)\|_F^2 \leq \sum_{k=1}^{K}\sum_{j\in J}\nu_j^{-2}$, and using (6.44), arrive at

$$K\left(P_{\mathbf{G}_1}, P_{\mathbf{G}_2}\right) \leq \frac{M\theta^2}{2\delta^2\,\aleph_1^2\,C_\psi^2}\,|J|\,L_1^{-2\gamma}\exp\left(-2\alpha L_1^\beta\right).$$

In order to apply Theorem 2.5 of Tsybakov (2009) with $\alpha = 1/9$, we need $K\left(P_{\mathbf{G}_1}, P_{\mathbf{G}_2}\right) \leq (1/9)\ln|\mathcal{G}_{M,K}|$ which, due to (6.52), is guaranteed by

$$\frac{\theta^2 M}{\delta^2\aleph_1^2 C_\psi^2}L_1^{-2\gamma}\exp\left(-2\alpha L_1^\beta\right) \leq \frac{K}{36}. \qquad (6.54)$$

If inequality (6.54) holds, then application of Theorem 2.5 of Tsybakov (2009) yields that, with probability at least 0.1, one has (3.9), where, due to (3.1) and (6.53),

$$R_{\min}(M, K_*, \delta) \gtrsim \theta^2|J| \qquad (6.55)$$

Now, as before, we consider two choices of L_1 and L_2: $L_1 = L_2 = L$ and $L_1 = L/2+1$, $L_2 = L$ leading to the values of θ^2 given by (6.49). Again, we consider the cases of $\alpha = \beta = 0$ and $\alpha > 0$, $\beta > 0$ separately.

Case 1: $\alpha = 0$, $\beta = 0$, $L_1 = L/2 + 1$, $L_2 = L$, $|J| = L/2$.
Since $L_1 \asymp L_2 \asymp |J| \asymp L$, inequality (6.54) holds if $L \asymp \left(\delta^2 M^{-1} K\right)^{-\frac{1}{2r+2\gamma+1}}$ and

$$R_{\min}(M, K_*, \delta) \gtrsim \left(\delta^2\,M^{-1}\,K\right)^{\frac{2r}{2r+2\gamma+1}}. \qquad (6.56)$$

Case 2: $\alpha > 0$, $\beta > 0$, $L_1 = L_2 = L$, $|J| = 1$.
Plugging the first expression from (6.49) into (6.54), derive that $L^{-(2\gamma+2r)}\exp\left(-2\alpha L^\beta\right) \lesssim \delta^2 M^{-1}K$, so that $L \asymp \left[\ln\left(\frac{M}{\delta^2 K}\right)\right]^{\frac{1}{\beta}}$. Therefore,

$$R_{\min}(M, K_*, \delta) \gtrsim \left[\ln\left(\frac{M}{\delta^2 K}\right)\right]^{-\frac{2r}{\beta}} \qquad (6.57)$$

Now, in order to obtain the expressions for the lower bounds, we find the maximum of (6.50) and (6.56) if $\alpha = 0$, $\beta = 0$, and of (6.51) and (6.57) if $\alpha > 0$, $\beta > 0$.

D. Proofs of the Comparison of the Risks With and Without Clustering

Proof of Corollary 1: First observe that expressions (3.12) are obtained directly from (3.6) and (3.7) by setting $M = K_* = 1$ since all functions belong to the same Sobolev ball (2.10). In order to compare the upper bounds (3.6) and (3.7) obtained with clustering with the upper bound (3.12) derived without clustering, we consider several cases.

Case 1 $\alpha = 0$, $\beta = 0$.
Expressions in (3.13) are obtain by direct evaluation. Note that the second expression in the case of $K_* \geq 2$ tends to zero as $M \to \infty$ since, due to (3.5), $\ln K_* \leq \ln M \asymp \ln \delta^{-1}$.
Case 2 $\alpha > 0$, $\beta > 0$.
Note that, due to the condition (3.5),

$$\ln(\delta^{-2}) \leq \ln(M\delta^{-2}K_*^{-1}) \leq \ln M + \ln(\delta^{-2}) \asymp \ln(\delta^{-2}),$$

Also, for $K_* \geq 2$ and $\delta^{-2} \geq e$, due to $\ln x \leq x/2$ for $x \geq 1$, obtain

$$\ln\left(\delta^{-2}\ln(K_*^{-1})\right) = \ln(\delta^{-2}) - \ln\ln K_* \geq$$
$$\ln(\delta^{-2}) - 0.5\ln(\delta^{-2}) \asymp \ln(\delta^{-2}),$$

which completes the proof.

E. Proofs of Supplementary Statements

Proof of Lemma 1: Proof of Lemma 1 is based on the following statement provided in Gendre(2014)

Lemma 4 (Gendre (2014)): Let $\mathbf{A} \in R^{p\times p}$ be a fixed matrix and $\boldsymbol{\epsilon} \sim N(0, \mathbf{I}_p)$. Then, for any $x > 0$ one has

$$\mathbb{P}\left\{\|\mathbf{A}\boldsymbol{\epsilon}\|^2 \geq \operatorname{Tr}(\mathbf{A}^T\mathbf{A}) + 2\sqrt{\|\mathbf{A}\|_{op}^2\operatorname{Tr}(\mathbf{A}^T\mathbf{A})\,x} \right. \qquad (6.58)$$
$$\left. +2\|\mathbf{A}\|_{op}^2 x\right\} \leq e^{-x}$$

Note that, due to $2ab \leq a^2 + b^2$, probability (6.58) can be re-written as

$$\mathbb{P}(\|\mathbf{A}\boldsymbol{\epsilon}\|^2 \geq 2\|\mathbf{A}\|_F^2 + 3\|\mathbf{A}\|_{op}^2\,x) \leq e^{-x} \qquad (6.59)$$

Consider $\|[\mathbf{\Pi}_{\mathbf{Z},K}\otimes(\mathbf{W}_J\mathbf{\Upsilon S})]\boldsymbol{\eta}\|^2$ with $\mathbf{Z}, J, K$ fixed. Note that, due to $\|\mathbf{\Pi}_{\mathbf{Z},K}\|_{op}^2 = 1$, $\|\mathbf{S}\|_{op}^2 \leq C_\psi^2$, $\|\mathbf{W}_J\mathbf{\Upsilon}\|_{op}^2 = \max_{j\in J}\nu_j^2$ and $\|\mathbf{W}_J\mathbf{\Upsilon}\|_F^2 = \sum_{j\in J}\nu_j^2$, one has

$$\|(\mathbf{\Pi}_{\mathbf{Z},K}\otimes(\mathbf{W}_J\mathbf{\Upsilon S}))\|_{op}^2 \leq \qquad (6.60)$$
$$\|\mathbf{\Pi}_{\mathbf{Z},K}\|_{op}^2\|\mathbf{W}_J\mathbf{\Upsilon}\|_{op}^2\|\mathbf{S}\|_{op}^2 \leq C_\psi^2\max_{j\in J}\nu_j^2$$

$$\|(\mathbf{\Pi}_{\mathbf{Z},K}\otimes(\mathbf{W}_J\mathbf{\Upsilon S}))\boldsymbol{\eta}\|_F^2 \leq \qquad (6.61)$$
$$\|\mathbf{\Pi}_{\mathbf{Z},K}\|_F^2\|\mathbf{W}_J\mathbf{\Upsilon}\|_F^2\|\mathbf{S}\|_{op}^2 \leq KC_\psi^2\sum_{j\in J}\nu_j^2$$

Now applying inequality (6.59) to $\|[\mathbf{\Pi}_{\mathbf{Z},K}\otimes(\mathbf{W}_J\mathbf{\Upsilon S})]\boldsymbol{\eta}\|^2$ where $\boldsymbol{\eta} \sim N(0, \mathbf{I}_{nM})$, we obtain for any $x > 0$

$$\mathbb{P}\left\{\|(\mathbf{\Pi}_{\mathbf{Z},K}\otimes(\mathbf{W}_J\mathbf{\Upsilon S}))\boldsymbol{\eta}\|^2 \geq \right.$$
$$2\|(\mathbf{\Pi}_{\mathbf{Z},K}\otimes(\mathbf{W}_J\mathbf{\Upsilon S}))\|_F^2 +$$
$$\left. 3\|(\mathbf{\Pi}_{\mathbf{Z},K}\otimes(\mathbf{W}_J\mathbf{\Upsilon S}))\|_{op}^2\,x\right\} \leq \qquad (6.62)$$
$$\mathbb{P}\left\{\|(\mathbf{\Pi}_{\mathbf{Z},K}\otimes(\mathbf{W}_J\mathbf{\Upsilon S}))\boldsymbol{\eta}\|^2 - \right.$$
$$\left. C_\psi^2\left[2\,K\sum_{j\in J}\nu_j^2 + 3x\max_{j\in J}\nu_j^2\right] \geq 0\right\} \leq e^{-x}.$$

Setting $x = \tau\ln(\delta^{-1})$ yields (6.9). Inequality (6.11) follows from (6.9) since ν_j are growing with j and $J = \{1, \ldots, L\}$.

In order to prove inequality (6.10), note that for

$$x(M, K, |J|, s) = M\ln K + |J|\ln(ne/|J|) + \ln(Mn) + s,$$

due to $\ln\binom{n}{j} \leq j\ln(\frac{ne}{j})$, one has

$$\sum_{\mathbf{Z},K,J}e^{-x(M,K,|J|,s)} \equiv$$
$$\sum_{K=1}^{M}\sum_{j=1}^{n}\sum_{|J|=j}\sum_{\mathbf{Z}\in\mathcal{M}(M,K)}e^{-x(M,K,j,s)} = \qquad (6.63)$$
$$\sum_{K=1}^{M}\sum_{j=1}^{n}\binom{n}{j}K^M e^{-x(M,K,j,s)} \leq$$
$$\sum_{K=1}^{M}\sum_{j=1}^{n}\left(\frac{ne}{j}\right)^j K^M e^{-x(M,K,j,s)} \leq e^{-s}$$

Therefore, by (6.62) and (6.63), we obtain

$$\mathbb{P}\left(\|(\mathbf{\Pi}_{\hat{\mathbf{Z}},\hat{K}} \otimes (\mathbf{W}_{\hat{J}}\Upsilon\mathbf{S}))\boldsymbol{\eta}\|^2 - \right.$$
$$2\|(\mathbf{\Pi}_{\hat{\mathbf{Z}},\hat{K}} \otimes (\mathbf{W}_{\hat{J}}\Upsilon\mathbf{S}))\|_F^2 - $$
$$\left. 3\|(\mathbf{\Pi}_{\hat{\mathbf{Z}},\hat{K}} \otimes (\mathbf{W}_{\hat{J}}\Upsilon\mathbf{S}))\|_{op}^2 \, x(M,\hat{K},|\hat{J}|,s) \geq 0\right) \leq$$
$$\sum_{\mathbf{Z},K,J} \mathbb{P}\left(\|(\mathbf{\Pi}_{\mathbf{Z},K} \otimes (\mathbf{W}_J\Upsilon\mathbf{S}))\boldsymbol{\eta}\|^2 - \right.$$
$$C_\psi^2 \left[2K\sum_{j\in J}\nu_j^2 + 3\,x(M,K,|J|,s)\left(\max_{j\in J}\nu_j^2\right)\right] \geq 0\right) \leq$$
$$\sum_{\mathbf{Z},K,J} e^{-x(M,K,|J|,s)} \leq e^{-s}.$$

Setting $s = \tau \ln(\delta^{-1})$ yields (6.10).

Similarly, in order to prove (6.12), choose $J = \{1, \ldots, L\}$, $x(M,K,|J|,s) = M\ln K + \ln(Mn) + s$, and replace (6.63) by

$$\sum_{\mathbf{Z},K,J} e^{-x(M,K,|J|,s)} \equiv \sum_{K=1}^{M}\sum_{L=1}^{n}\sum_{\mathbf{Z}\in\mathcal{M}(M,K)} e^{-x(M,K,L,s)}$$
$$\leq \sum_{K=1}^{M} n\,K^M e^{-x(M,K,L,s)} \leq e^{-s}$$

Proof of Lemma 3: By using (6.36), $K \geq 2$ and $0 < d \leq 1/9$

$$\ln K - 4d\ln(Ke/d) = \ln K - 4[d\ln(K) + d - d\ln d]$$
$$\geq \ln K - 4d\ln K - \frac{4}{9}\ln 2 \geq \frac{5}{9}\ln K - \frac{4}{9}\ln K \geq \frac{\ln K}{9}.$$

REFERENCES

[1] F. Abramovich, "Wavelet decomposition approaches to statistical inverse problems," *Biometrika*, vol. 85, no. 1, pp. 115–129, Mar. 1998.

[2] F. Abramovich, D. De Canditiis, and M. Pensky, "Solution of linear ill-posed problems by model selection and aggregation," *Electron. J. Statist.*, vol. 12, no. 1, pp. 1822–1841, 2018.

[3] P. Alquier, E. Gautier, and G. Stoltz, *Inverse Problems and High-Dimensional Estimation*. Berlin, Germany: Springer-Verlag, 2011.

[4] A. Arnold, S. Reichling, O. T. Bruhns, and J. Mosler, "Efficient computation of the elastography inverse problem by combining variational mesh adaption and a clustering technique," *Phys. Med. Biol.*, vol. 55, no. 7, pp. 2035–2056, Apr. 2010.

[5] J. C. Bezdek and S. K. Pal, *Fuzzy Models for Pattern Recognition Methods That Search for Structures in Data*. New York, NY, USA: IEEE Press, 1992.

[6] N. Bissantz, T. Hohage, A. Munk, and F. Ruymgaart, "Convergence rates of general regularization methods for statistical inverse problems and applications," *SIAM J. Numer. Anal.*, vol. 45, no. 6, pp. 2610–2636, Jan. 2007.

[7] G. Blanchard, M. Hoffmann, and M. Reiß, "Early stopping for statistical inverse problems via truncated SVD estimation," *Electron. J. Statist.*, vol. 12, no. 2, pp. 3204–3231, 2018.

[8] A. Cohen, M. Hoffmann, and M. Reiss, "Adaptive wavelet Galerkin methods for linear inverse problems," *SIAM J. Numer. Anal.*, vol. 42, no. 4, pp. 1479–1501, Jan. 2004.

[9] F. Comte, C.-A. Cuenod, M. Pensky, and Y. Rozenholc, "Laplace deconvolution on the basis of time domain data and its application to dynamic contrast-enhanced imaging," *J. Roy. Stat. Soc., B, Stat. Methodol.*, vol. 79, no. 1, pp. 69–94, Jan. 2017.

[10] Z. Deng, F.-L. Chung, and S. Wang, "Clustering-inverse: A generalized model for pattern-based time series segmentation," *J. Intell. Learn. Syst. Appl.*, vol. 3, no. 1, pp. 26–36, 2011.

[11] D. L. Donoho, "Nonlinear solution of linear inverse problems by Wavelet–Vaguelette decomposition," *Appl. Comput. Harmon. Anal.*, vol. 2, no. 2, pp. 101–126, Apr. 1995.

[12] D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–456, 1994.

[13] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems*. Amsterdam, The Netherlands: Kluwer, 2000.

[14] D. Fraix-Burnet and S. Girard, *Statistics for Astrophysics Clustering and Classification*. Les Ulis, France: EDP Sciences, 2016.

[15] X. Gendre, "Model selection and estimation of a component in additive regression," *ESAIM, Probab. Statist.*, vol. 18, pp. 77–116, 2014, doi: 10.1051/ps/2012028.

[16] A. K. Gupta and D. K. Nagar, *Matrix Variate Distributions*. Boca Raton, FL, USA: Chapman & Hall, 1999.

[17] O. Klopp, Y. Lu, A. B. Tsybakov, and H. H. Zhou, "Structured matrix estimation and completion," *Bernoulli*, vol. 25, no. 4B, pp. 3883–3911, Nov. 2019.

[18] E. Kürüm, G.-W. Weber, and C. Iyigun, "Early warning on stock market bubbles via methods of optimization, clustering and inverse problems," *Ann. Oper. Res.*, vol. 260, nos. 1–2, pp. 293–320, Jan. 2018.

[19] S. Mallat, *A Wavelet Tour of Signal Processing. The Sparse Way*, 3rd ed. New York, NY, USA: Academic, 2009.

[20] M. Pensky, "Solution of linear ill-posed problems using overcomplete dictionaries," *Ann. Statist.*, vol. 44, no. 4, pp. 1739–1764, Aug. 2016.

[21] M. Pensky, "Dynamic network models and graphon estimation," *Ann. Statist.*, vol. 47, no. 4, pp. 2378–2403, 2019.

[22] J. L. Starck, E. Pantin, and F. Murtagh, "Deconvolution in astronomy: A review," *Publications Astronomical Soc. Pacific*, vol. 114, no. 800, pp. 1051–1069, 2002.

[23] A. B. Tsybakov, *Introduction to Nonparametric Estimation*. New York, NY, USA: Springer, 2009.

Rasika Rajapakshage received the M.S. degree in mathematics from the University of North Florida in 2014, and the Ph.D. degree in mathematics from the University of Central Florida in 2019. He is currently a Postdoctoral Research Associate at the Department of Mathematics and Statistics, University of Wyoming. His research interests include statistical inverse problems, medical imaging, machine learning, and big data.

Marianna Pensky is currently a Professor with the Department of Mathematics, University of Central Florida (UCF). She has authored or coauthored more than 100 peer-reviewed journal and conference papers. She is an Associate Editor of the *Journal of the American Statistical Association*, *Journal of Statistical Planning and Inference*, and *Journal of Nonparametric Statistics*, and an Elected Member of the Institute of Mathematical Statistics and the International Statistical Institute.