



HAL
open science

Model-assisted variable clustering: minimax-optimal recovery and algorithms

Florentina Bunea, Christophe Giraud, Xi Luo, Martin Royer, Nicolas Verzelen

► **To cite this version:**

Florentina Bunea, Christophe Giraud, Xi Luo, Martin Royer, Nicolas Verzelen. Model-assisted variable clustering: minimax-optimal recovery and algorithms. *Annals of Statistics*, 2020, 48 (1), pp.111-137. 10.1214/18-AOS1794 . hal-01769271

HAL Id: hal-01769271

<https://hal.science/hal-01769271>

Submitted on 18 Apr 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

MODEL ASSISTED VARIABLE CLUSTERING: MINIMAX-OPTIMAL RECOVERY AND ALGORITHMS

BY FLORENTINA BUNEA

Cornell University

BY CHRISTOPHE GIRAUD

CNRS, Université Paris Sud, Université Paris-Saclay

BY XI LUO

Brown University

BY MARTIN ROYER

CNRS, Université Paris Sud, Université Paris-Saclay

AND

BY NICOLAS VERZELEN

INRA

The problem of variable clustering is that of estimating groups of similar components of a p -dimensional vector $X = (X_1, \dots, X_p)$ from n independent copies of X . There exists a large number of algorithms that return data-dependent groups of variables, but their interpretation is limited to the algorithm that produced them. An alternative is model-based clustering, in which one begins by defining population level clusters relative to a model that embeds notions of similarity. Algorithms tailored to such models yield estimated clusters with a clear statistical interpretation. We take this view here and introduce the class of G -block covariance models as a background model for variable clustering. In such models, two variables in a cluster are deemed similar if they have similar associations with all other variables. This can arise, for instance, when groups of variables are noise corrupted versions of the same latent factor. We quantify the difficulty of clustering data generated from a G -block covariance model in terms of cluster proximity, measured with respect to two related, but different, cluster separation metrics. We derive minimax cluster separation thresholds, which are the metric values below which no algorithm can recover the model-defined clusters exactly, and show that they are different for the two metrics. We therefore develop two algorithms, COD and PECOK, tailored to G -block covariance models, and study their minimax-optimality with respect to each metric. Of independent interest is the fact that the analysis of the PECOK algorithm, which is based on a corrected convex relaxation of the popular K -means algorithm, provides the first statistical analysis of such algorithms for variable clustering. Additionally, we contrast our methods with another popular clustering method, spectral clustering, specialized to variable clustering, and show that ensuring exact cluster recovery via this method requires clusters to have a higher separation, relative to the minimax threshold. Extensive simulation studies, as well as our data analyses, confirm the applicability of our approach.

1. Introduction. The problem of variable clustering is that of grouping similar components of a p -dimensional vector $X = (X_1, \dots, X_p)$. These groups are referred to as clusters. In this work we investigate the problem of cluster recovery from a sample of n independent copies of X . Variable clustering has had a long history in a variety of fields, with important examples stemming from gene expression data [18, 22, 39] or protein profile data [7]. The solutions to this problem are typically algorithmic and entirely data based. They include applications of K -means, hierarchical clustering, spectral clustering, or versions of them. The statistical properties of these procedures have received a very limited amount of investigation. It is not currently known what probabilistic cluster models on X can be estimated by these popular techniques, or by their modifications. More generally, model-based variable clustering has received a limited amount of attention. One net advantage of model-based clustering is that population-level clusters are clearly defined, offering both interpretability of the clusters and a benchmark against which one can check the quality of a particular clustering algorithm.

In this work we propose the G -block covariance model as a flexible model for variable clustering and show that the clusters given by this model are uniquely defined. We then motivate and develop two algorithms tailored to the model, COD and PECOK, and analyze their respective performance in terms of exact cluster recovery, for minimally separated clusters, under appropriately defined cluster separation metrics.

1.1. *The G -block covariance model.* Our proposed model for variable clustering subsumes that the covariance matrix Σ of a centered random vector $X \in \mathbf{R}^p$ follows a block, or near-block, decomposition, with blocks corresponding to a partition $G = \{G_1, \dots, G_K\}$ of $\{1, \dots, p\}$. This structure of the covariance matrix has been observed to hold, empirically, in a number of very recent studies on the parcellation of the human brain, for instance [17, 19, 24, 38]. We further support these findings in Section 8, where we apply the clustering methods developed in this paper, tailored to G -block covariance models, for the clustering of brain regions.

To describe our model, we associate, to a partition G , a membership matrix $A \in \mathbf{R}^{p \times K}$ defined by $A_{ak} = 1$ if $a \in G_k$, and $A_{ak} = 0$ otherwise.

(A) The exact G -block covariance model. In view of the above discussion, clustering the variables (X_1, \dots, X_p) amounts to find a minimal (i.e. coarsest partition) G^* , such that two variables belong to the same cluster if they have the same covariance with all other variables. This implies that the covariance matrix Σ of X decomposes as

$$(1) \quad \Sigma = AC^*A^t + \Gamma,$$

where A is relative to G^* , C^* is a symmetric $K \times K$ matrix, and Γ a diagonal matrix. When a such a decomposition exists with the partition G^* , we say that $X \in \mathbf{R}^p$ follows an (exact) G^* -block covariance model.

(i) G -Latent Model. Such a structure arises, for instance, when components of X that belong to the same group can be decomposed into the sum between a common latent variable and an uncorrelated random fluctuation. Similarity within group is therefore given by association with the same unobservable source. Specifically, the exact block-covariance model (1) holds, with a diagonal matrix Γ , when

$$(2) \quad X_a = Z_{k(a)} + E_a,$$

with $\text{Cov}(Z_{k(a)}, E_a) = 0$, $\text{Cov}(Z) = C^*$, and the individual fluctuations E_a are uncorrelated, and thus E has diagonal covariance matrix Γ . The index assignment function $k : \{1, \dots, p\} \rightarrow$

$\{1, \dots, K\}$ is defined by $G_k = \{a : k(a) = k\}$. In practice, this model is used to justify the construction of a single variable that represents a cluster, the average of X_a , $a \in G_k$, viewed as an observable proxy of $Z_{k(a)}$. For example, a popular analysis approach for fMRI data, called region-of-interest (ROI) analysis [34], requires averaging the observations from multiple voxels (a imaging unit for a small cubic volume of the brain) within each ROI (or cluster of voxels) to produce new variables, each representing a larger and interpretable brain area. These new variables are then used for downstream analyses. From this perspective, model (2) can be used in practice, see, for example [6], as a building block in a data analysis based on cluster representatives, which in turn requires accurate cluster estimation. Indeed, data-driven methods for clustering either voxels into regions or regions into functional systems, especially based on the covariance matrix of X , is becoming increasingly important, see for example [17, 19, 35, 38]. Accurate data-driven clustering methods also enable studying the cluster differences across subjects [16] or experimental conditions [21].

(ii) The Ising Block Model. The Ising Block Model has been proposed in [8] for modelling social interactions, for instance political affinities. Under this model, the joint distribution of $X \in \{-1, 1\}^p$, a p -dimensional vector with binary entries, is given by

$$(3) \quad f(x) = \frac{1}{\kappa_{\alpha, \beta}} \exp \left[\frac{\beta}{2p} \sum_{a \sim b} x_a x_b + \frac{\alpha}{2p} \sum_{a \not\sim b} x_a x_b \right],$$

where the quantity $\kappa_{\alpha, \beta}$ is a normalizing constant, and the notation $a \sim b$ means that the elements are in the same group of the partition. The variables X_a may for instance represent the votes of U.S. senators on a bill [5]. For parameters $\alpha > \beta$, the density (3) models the fact that senators belonging to the same political group tend to share the same vote. By symmetry of the density f , the covariance matrix Σ of X decomposes as an exact block covariance model $\Sigma = AC^*A^t + \Gamma$ where Γ is diagonal. When all groups G_k^* have identical size, we have $C^* = (\omega_{in} - \omega_{out})I_K + \omega_{out}J$ and $\Gamma = (1 - \omega_{in})I$, where the $K \times K$ matrix J has all entries equal to 1, and I_K denotes the $K \times K$ identity matrix, and the quantities $\omega_{in}, \omega_{out}$ depend on α, β, p .

(B) The approximate G -block model. In many situations, it is more appealing to group variables that *nearly* share the same covariance with all the other variables. In that situation, the covariance matrix Σ would decompose as

$$(4) \quad \Sigma = ACA^t + \Gamma, \text{ where } \Gamma \text{ has small off-diagonal entries.}$$

Such a situation can arise, for instance when $X_a = (1 + \delta_a)Z_{k(a)} + E_a$, with $\delta_a = o(1)$ and the individual fluctuations E_a are uncorrelated, $1 \leq a \leq p$.

1.2. Our contribution. We assume that the data consist in i.i.d. observations $X^{(1)}, \dots, X^{(n)}$ of a random vector X with mean 0 and covariance matrix Σ . This work is devoted to the development of computationally feasible methods that yield estimates \widehat{G} of G^* , such that $\widehat{G} = G^*$, with high probability, when the clusters are minimally separated, and to characterize the minimal value of the cluster separation from a minimax perspective. The separation between clusters is a key element in quantifying the difficulty of a clustering task as, intuitively, well separated clusters should be easier to identify. We consider two related, but different, separation metrics, that can be viewed as canonical whenever Σ satisfies (4). Although all our results allow, and are proved, for small departures from the diagonal structure of Γ in (1), our main contribution can be best seen when Γ is a diagonal matrix. We focus on this case below, for clarity of exposition. The case of Γ being a perturbation of a diagonal matrix is treated in Section 6.

When Γ is diagonal, our target partition G^* can be easily defined. It is the unique minimal (with respect to partition refinement) partition G^* for which there is a decomposition $\Sigma = AC^*A^t + \Gamma$, with A associated to G^* . We refer to Section 2 for details. We observe in particular, that $\max_{c \neq a, b} |\Sigma_{ac} - \Sigma_{bc}| > 0$ if and only if X_a and X_b belong to different clusters in G^* .

This last remark motivates our first metric MCOV based on the following COVariance Difference (COD) measure

$$(5) \quad \text{COD}(a, b) := \max_{c \neq a, b} |\Sigma_{ac} - \Sigma_{bc}| \quad \text{for any } a, b = 1, \dots, p.$$

We use the notation $a \stackrel{G^*}{\sim} b$ whenever a and b belong to the same group G_k^* , for some k , in the partition G^* , and similarly $a \not\stackrel{G^*}{\sim} b$ means that there does not exist any group G_k^* of the partition G^* that contains both a and b . We define the MCOV metric as

$$(6) \quad \text{MCOV}(\Sigma) := \min_{a \not\stackrel{G^*}{\sim} b} \text{COD}(a, b).$$

The measure $\text{COD}(a, b)$ quantifies the similarity of the covariances that X_a and X_b have, respectively, with all other variables. From this perspective, the size of $\text{MCOV}(\Sigma)$ is a natural measure for the difficulty of clustering when analyzing clusters with components that are similar in this sense. Moreover, note that this metric is well defined even if C^* of model (1) is not semi-positive definite.

Another cluster separation metric appears naturally when we view model (1) as arising via model (2), or via small deviations from it. Then, clusters in (1) are driven by the latent factors, and intuitively they differ when the latent factors differ. Specifically, we define the "within-between group" covariance gap

$$(7) \quad \Delta(C^*) := \min_{j < k} (C_{kk}^* + C_{jj}^* - 2C_{jk}^*) = \min_{j < k} \mathbf{E} [(Z_j - Z_k)^2],$$

where the second equality holds whenever (2) holds. In the latter case, the matrix C^* , which is the covariance matrix of the latent factors, is necessarily semi-positive definite. Further, we observe that $\Delta(C^*) = 0$ implies $Z_j = Z_k$ a.s. Conversely, we prove in Corollary 1 of Section 2 that if the decomposition (1) holds with $\Delta(C^*) > 0$, then the partition related to A is the partition G^* described above. An instance of $\Delta(C^*) > 0$ corresponds to having the within group covariances stronger than those between groups. This suggests the usage of this metric $\Delta(C^*)$ for cluster analysis whenever, in addition to the general model formulation (1), we also expect clusters to have this property, which has been observed, empirically, to hold in applications. For instance, it is implicit in the methods developed by [17] for creating a human brain atlas by partitioning appropriate covariance matrices. We also present a neuroscience-based data example in 8.

Formally, the two metrics are connected via the following chain of inequalities, proved in Lemma 1 of Section 1.1 of the supplemental material [12], and valid as soon as the size of the smallest cluster is larger than one, Γ and C^* is semi-positive definite (for the last inequality)

$$(8) \quad 2\lambda_K(C^*) \leq \Delta(C^*) \leq 2\text{MCOV}(\Sigma) \leq 2\sqrt{\Delta(C^*)} \max_{k=1, \dots, K} \sqrt{C_{kk}^*}.$$

The first inequality shows that conditions on $\Delta(C^*)$ are weaker than conditions on the minimal eigenvalue $\lambda_K(C^*)$ of C^* . In order to preserve the generality of our model, we do not necessarily assume that $\lambda_K(C^*) > 0$, as we show that, for model identifiability, it is enough to have the weaker condition $\Delta(C^*) > 0$, when the two quantities differ.

The second inequality in (8) shows that $\Delta(C^*)$ and $\text{MCOD}(\Sigma)$ can have the same order of magnitude, whereas the third inequality shows that they can also differ in order, and $\Delta(C^*)$ can be as small as $\text{MCOD}^2(\Sigma)$, for small values of these metrics, which is our main focus. This suggests that different statistical assessments, and possibly different algorithms, should be developed for estimators of clusters defined by (1), depending on the cluster separation metric. To substantiate this intuition, we first derive, for each metric, the rate below which no algorithm can recover exactly the clusters defined by (1). We call this the minimax optimal threshold for cluster separation, and prove that it is different for the two metrics. We call an algorithm that can be proved to recover exactly clusters with separation above the minimax threshold a minimax optimal algorithm.

Theorem 1 in Section 3 shows that no algorithm can estimate consistently clusters defined by (1) if

$$(9) \quad \text{MCOD}(\Sigma) \lesssim \sqrt{\frac{\log(p)}{n}}.$$

Here and throughout this paper the symbol \lesssim is used whenever an inequality holds up to multiplicative constants, which are made precise in the statements of the theorems where such inequalities are proved. Theorem 2 in Section 3 shows that optimal separation distances with respect to the metric $\Delta(C^*)$ are sensitive to the size of the smallest cluster,

$$m^* = \min_{1 \leq k \leq K} |G_k^*|,$$

in that no algorithm can estimate consistently clusters defined by (1) when

$$(10) \quad \Delta(C^*) \lesssim \left(\sqrt{\frac{\log(p)}{nm^*}} \vee \frac{\log(p)}{n} \right).$$

The first term will be dominant whenever the smallest cluster has size $m^* < n/\log(p)$, which will be the case in most situations. The second term in (10) becomes dominant whenever $m^* > n/\log(p)$, which can also happen when p scales as n , and we have a few balanced clusters.

The PECOK algorithm is tailored to the $\Delta(C^*)$ metric, and is shown in Theorem 4 to be near-minimax optimal. For instance, for balanced clusters, exact recovery is guaranteed when $\Delta(C^*) \gtrsim \sqrt{\frac{K\sqrt{\log p}}{m^*n}} + \frac{K\sqrt{\log(p)}}{n}$. This differs by factors in K from the $\Delta(C^*)$ -minimax threshold, for general K , whereas it is of optimal order when K is a constant, or grows as slowly as $\log p$. A similar discrepancy between minimax lower bounds and the performance of polynomial-time estimators has also been pinpointed in network clustering via the stochastic block model [15] and in sparse PCA [9]. It has been conjectured that, when K increases with n , there exists a gap between the statistical boundary, i.e. the minimal cluster separation for which a statistical method achieves perfect clustering with high probability, and the polynomial boundary, i.e. the minimal cluster separation for which there exists a polynomial-time algorithm that achieves perfect clustering. Further investigation of this computational trade-off is beyond the scope of this paper and we refer to [15] and [9] for more details.

However, if we consider directly the metric $\text{MCOD}(\Sigma)$, and its corresponding, larger, minimax threshold, we derive the COD algorithm, which is minimax optimal with respect to $\text{MCOD}(\Sigma)$. In view of (8), it is also minimax optimal with respect to $\Delta(C^*)$, whenever there exist small clusters, the size of which does not change with n . The description of the two algorithms and theoretical properties are given in Sections 4 and 5, respectively, for exact block covariance models. Companions of these results, regarding the performance of the algorithms for approximate block covariance models are given in Section 6, in Theorem 6 and Theorem 7, respectively.

Table 1 below gives a snap-shot of our results, which for ease of presentation, correspond to the case of balanced clusters, with the same number of variables per cluster. We stress that neither our algorithms, nor our theory, is restricted to this case, but the exposition becomes more transparent in this situation.

Metric	Minimax threshold	PECOK	COD
$d_1 =: \Delta(C^*)$	$\sqrt{\frac{\log p}{mn}} + \frac{\log p}{n}$	Always near-minimax optimal w.r.t. d_1 .	Minimax optimal w.r.t. d_1 when m is constant.
$d_2 =: \text{MCOD}(\Sigma)$	$\sqrt{\frac{\log p}{n}}$	Minimax optimal w.r.t. d_2 when $m > n/\log(p)$ and $K = O(\log p)$.	Always minimax optimal w.r.t. d_2 .

TABLE 1

Algorithm performance relative to minimax thresholds of each metric

In this table m denotes the size of the smallest cluster in the partition. The performance of COD under d_1 follows from the second inequality in (8), whereas the performance of PECOK under d_2 follows from the last inequality in (8). The overall message transmitted by Table 1 and our analysis is that, irrespective of the separation metric, the COD algorithm will be most powerful whenever we expect to have at least one, possibly more, small clusters, a situation that is typically not handled well in practice by most of the popular clustering algorithms, see [11] for an in-depth review. The PECOK algorithm is expected to work best for larger clusters, in particular when there are no clusters of size one. We defer more comments on the relative numerical performance of the methods to the discussion Section 9.3.

We emphasize that both our algorithms are generally applicable, and our performance analysis is only in terms of the most difficult scenarios, when two different clusters are almost indistinguishable and yet, as our results show, consistently estimable. Our extensive simulation results confirm these theoretical findings.

We summarize below our key contributions.

- (1) **An identifiable model for variable clustering and metrics for cluster separation.** We advocate model-based variable clustering, as a way of proposing objectively defined and interpretable clusters. We propose identifiable G -block covariance models for clustering, and prove cluster identifiability in Proposition 1 of Section 2.
- (2) **Minimax lower bounds on cluster separation metrics for exact partition recovery.** Two of our main results are Theorem 2 and Theorem 1, presented in Section 3, in which we establish, respectively, minimax limits on the size of the $\Delta(C^*)$ -cluster separation and $\text{MCOD}(\Sigma)$ -cluster separation below which no algorithm can recover clusters defined by (1) consistently, from a sample of size n on X . To the best of our knowledge these are the first results of this type in variable clustering.
- (3) **Variable clustering procedures with guaranteed exact recovery of minimally separated clusters.** The results of (1) and (2) provide a much needed framework for motivating variable clustering algorithm development and for clustering algorithm assessments.

In particular, they motivate a correction of a convex relaxation of the K -means algorithm, leading to our proposed PECOK procedure, based on Semi-Definite Programming (SDP). Theorem 4 shows it to be near-minimax optimal with respect to the $\Delta(C^*)$ metric. The PECOK - $\Delta(C^*)$ pairing is natural, as $\Delta(C^*)$ measures the difference of the "within cluster" signal relative to the "between clusters" signal, which is the idea that underlies K -means type procedures. To the best of our knowledge, this is the first work that explicitly shows what model-based clusters of variables can be estimated via K -means style methods, and assesses theoretically the quality of estimation.

Moreover, our work shows that the results obtained in [8], for the block Ising model, can be generalized to arbitrary values of K and unbalanced clusters.

The COD procedure is a companion of PECOK for clusters given by model (1), and is minimax optimal with respect to the $MCOD(\Sigma)$ cluster separation, as established in Theorem 1. Another advantage of COD is of computational nature, as SDP-based methods, although convex, can be computationally involved.

(4) Comparison with corrected spectral variable clustering methods. In Section 5.4, we connect PECOK with another popular algorithm, spectral clustering. We show that although it may be less computationally involved than PECOK, good cluster recovery can only be theoretically guaranteed for very well separated clusters, well above the minimax optimal threshold.

1.3. *Organization of the paper.* The rest of the paper is organized as follows:

Sections 1.4 and 1.5 contain the notation and distributional assumptions used throughout the paper.

For clarity of exposition, Sections 2 - 5 contain results established for model (1), when Γ is a diagonal matrix. Extensions to the case when Γ has small off-diagonal entries are presented in Section 6.

Section 2 shows that we have a uniquely defined target of estimation, the partition G^* .

Section 3 derives the minimax thresholds on the separation metrics $\Delta(C^*)$ and $MCOD(\Sigma)$, respectively, for estimating G^* consistently.

Section 4 is devoted to the COD algorithm, and its analysis.

Section 5 is devoted to the PECOK algorithm and its analysis.

Section 5.4 analyses spectral clustering for variable clustering, and compares it with PECOK.

Section 6 contains extensions to approximate G -block covariance models.

Section 7 explores the numerical performance of our methods, and Section 8 presents their application to the clustering of putative brain areas using a real fMRI data.

Section 9 contains a discussion of our results and overall recommendations regarding the usage of our methods. Given the space constraints, only some of our proofs are presented in the Appendix, while a large portion of them are included in the supplemental material [12].

1.4. *Notation.* We denote by \mathbf{X} the $n \times p$ matrix with rows corresponding to observations $X^{(i)} \in \mathbf{R}^p$, for $i = 1, \dots, n$. The sample covariance matrix $\hat{\Sigma}$ is defined by

$$\hat{\Sigma} = \frac{1}{n} \mathbf{X}^t \mathbf{X} = \frac{1}{n} \sum_{i=1}^n X^{(i)} (X^{(i)})^t.$$

Given a vector v and $q \geq 1$, $|v|_q$ stands for the ℓ_q norm. For a generic matrix M : $|M|_q$ denotes its the entry-wise ℓ_q norm, $\|M\|_{op}$ denotes its operator norm, and $\|M\|_F$ refers to the Frobenius norm. We use $M_{.a}$, $M_{b.}$, to denote the a -th column or, respectively, b -th row of a generic matrix M . The bracket $\langle \cdot, \cdot \rangle$ refers to the Frobenius scalar product. Given a matrix M , we denote $\text{supp}(M)$ its support, that is the set of indices (i, j) such that $M_{ij} \neq 0$. I denotes the identity matrix. We define the variation semi-norm of a diagonal matrix D as $|D|_V := \max_a D_{aa} - \min_a D_{aa}$. We use $B \succcurlyeq 0$ to denote a symmetric and positive semidefinite matrix.

Throughout this paper will make use of the notation c_1, c_2, \dots to denote positive constants independent of n, p, K, m . The same letter, for instance c_1 may be used in different statements and may denote different constants, which are made clear within each statement, when there is no possibility for confusion.

We use $[p]$ to denote the set $\{1, \dots, p\}$. We use the notation $a \stackrel{G}{\sim} b$ whenever $a, b \in G_k$, for the same k . Also, $m = \min_k |G_k|$ stands for the size of the smallest group of the partition G .

The notation \gtrsim and \lesssim is used for whenever the inequalities hold up to multiplicative numerical constants.

1.5. Distributional assumptions. For a p -dimensional random vector Y , its Orlicz norm is defined by $\|Y\|_{\psi_2} = \sup_{t \in \mathbf{R}^p: \|t\|_2=1} \inf\{s > 0 : \mathbb{E}[e^{(Z^t t)/s^2}] \leq 2\}$. Throughout the paper we will assume that X follows a sub-Gaussian distribution. Specifically, we use:

Assumption 1. (sub-Gaussian distributions) There exists $L > 0$ such that random vector $\Sigma^{-1/2}X$ satisfies $\|\Sigma^{-1/2}X\|_{\psi_2} \leq L$, where

Our class of distributions includes, in particular, that of bounded distributions, which may be of independent interest, as example (ii) illustrates. We will therefore also specialize some of our results to this case, in which case we will use directly

Assumption 1-bis. (Bounded distributions) There exists $M > 0$ such that $\max_{i=1, \dots, p} |X_i| \leq M$ almost surely.

Gaussian distributions satisfy Assumption 1 with $L = 1$. A bounded distribution is also sub-Gaussian, but the corresponding quantity L can be much larger than M , and sharper results can be obtained if Assumption 1-bis holds.

2. Cluster identifiability in G -block covariance models. To keep the presentation focused, we consider in sections 2–5 the model (1) with Γ diagonal. We treat the case corresponding to a diagonally dominant Γ in Section 6 below. In the sequel, it is assumed that $p > 2$.

We observe that if the decomposition (1) holds for a partition G , it also holds for any subpartition of G . It is natural therefore to seek the smallest (coarsest) of such partitions, that is the partition with the least number of groups for which (1) holds. Since the partition ordering is a partial order, the smallest partition is not necessarily unique. However, the following lemma shows that uniqueness is guaranteed for our model class.

LEMMA 1. Consider any covariance matrix Σ .

- (a) There exists a unique minimal partition G^* such that $\Sigma = ACA^t + \Gamma$ for some diagonal matrix Γ , some membership matrix A associated to G^* and some matrix C .
- (b) The partition G^* is given by the equivalence classes of the relation

$$(11) \quad a \equiv b \text{ if and only if } COD(a, b) := \max_{c \neq a, b} |\Sigma_{ac} - \Sigma_{bc}| = 0.$$

PROOF OF LEMMA 1. If decomposition $\Sigma = ACA^t + \Gamma$ holds with A related to a partition G , then we have $COD(a, b) = 0$ for any a, b belonging to the same group of G . Hence, each group G_k of G is included in one of the equivalence class of \equiv . As a consequence, G is a finer partition than G^* as defined in (b). Hence, G^* is the (unique) minimal partition such that decomposition $\Sigma = ACA^t + \Gamma$ holds. \square

As a consequence, the partition G^* is well-defined and is identifiable. Next, we discuss the definitions of MCOD and Δ metrics. For any partition G , we let $MCOD(\Sigma, G) := \min_{a \stackrel{G}{\sim} b} COD(a, b)$, where we recall that the notation $a \stackrel{G}{\sim} b$ means that a and b are not in a same group of the partition G . By definition of G^* , we notice that $MCOD(\Sigma, G^*) > 0$ and the next proposition shows that G^* is characterized by this property.

PROPOSITION 1. *Let G be any partition such that $\text{MCOD}(\Sigma, G) > 0$ and the decomposition $\Sigma = ACA^t + \Gamma$ holds with A associated to G . Then $G = G^*$.*

The proofs of this proposition and the following corollary are given in Section 1 of the supplemental material [12]. In what follows, we use the notation $\text{MCOD}(\Sigma)$ for $\text{MCOD}(\Sigma, G^*)$.

In general, without further restrictions on the model parameters, the decomposition $\Sigma = ACA^t + \Gamma$ with A relative to G^* is not unique. If, for instance Σ is the identity matrix I , then G^* is the complete partition (with p groups) and the decomposition (1) holds for any $(C, \Gamma) = (\lambda I, (1 - \lambda)I)$ with $\lambda \in \mathbf{R}$.

Recall that $m^* := \min |G_k^*|$ stands for the size of the smallest cluster. If we assume that $m^* > 1$ (no singleton), then Γ is uniquely defined. Besides, the matrix C in (1) is only defined up to a permutation of its rows and columns. In the sequel, we denote C^* any of these matrices C . When the partition contains singletons ($m^* = 1$), the matrix decomposition $\Sigma = ACA^t + \Gamma$ is made unique (up to a permutation of row and columns of C) by putting the additional constraint that the entries Γ_{aa} corresponding to singletons are equal to 0. Since the definition of $\Delta(C)$ is invariant with respect to permutation of rows and columns, this implies that $\Delta(C^*)$ is well-defined for any covariance matrix Σ .

For arbitrary Σ , $\Delta(C^*)$ is not necessarily positive. Nevertheless, if $\Delta(C^*) > 0$, then G^* is characterized by this property.

COROLLARY 1. *Let G be a partition such that $m = \min_k |G_k| \geq 2$, the decomposition $\Sigma = ACA^t + \Gamma$ holds with A associated to G and $\Delta(C) > 0$. Then $G = G^*$.*

As pointed in (7), in the latent model (2), $\Delta(C^*)$ is equal to the square of the minimal L^2 -norm between two latent variables. So, in this case, the condition $\Delta(C^*) > 0$ simply requires that all latent variables are distincts.

3. Minimax thresholds on cluster separation for perfect recovery. Before developing variable clustering procedures, we begin by assessing the limits of the size of each of the two cluster separation metrics below which no algorithm can be expected to recover the clusters perfectly. We denote by $m^* = \min_k |G_k^*|$ the size of the smallest cluster of the target partition G^* defined above. For $1 \leq m \leq p$ and $\eta, \tau > 0$, we consider the following sets of covariance matrices : $\mathcal{M}(m, \eta) := \{\Sigma : \text{MCOD}(\Sigma) > \eta |\Sigma|_\infty, m^* > m\}$ and $\mathcal{D}(m, \tau) := \{\Sigma : \Delta(C^*) > \tau |\Gamma|_\infty, m^* > m\}$. We use the notation \mathbb{P}_Σ to refer to the normal distribution with covariance Σ .

THEOREM 1. *There exists a positive constant c_2 such that, for any $1 \leq m \leq p/3$ and any η such that*

$$(12) \quad 0 \leq \eta < \eta^* := c_2 \sqrt{\frac{\log(p)}{n}},$$

we have $\inf_{\hat{G}} \sup_{\Sigma \in \mathcal{M}(m, \eta)} \mathbb{P}(\hat{G} \neq G^) \geq 1/7$, where the infimum is taken over all possible estimators.*

We also have:

THEOREM 2. *There exists a positive constant c_1 such that, for any $2 \leq m \leq p/2$ and any τ such that*

$$(13) \quad 0 \leq \tau < \tau^* := c_1 \left[\sqrt{\frac{\log(p)}{n(m-1)}} \vee \frac{\log(p)}{n} \right]$$

then $\inf_{\widehat{G}} \sup_{\Sigma \in \mathcal{D}(m, \tau)} \mathbb{P}_{\Sigma}[\widehat{G} \neq G^*] \geq 1/7$, where the infimum is taken over all estimators.

Theorems 2 and 1 show that if either metric falls below the thresholds in (13) or (12), respectively, the estimated partition \widehat{G} , irrespective of the method of estimation, cannot achieve perfect recovery with high-probability uniformly over the set $\mathcal{M}(m, \eta)$ or $\mathcal{D}(m, \tau)$. The proofs are given in Sections 2.1 and 2.2, respectively, of the supplement [12]. We note that $\Delta(C^*)$ minimax threshold takes into account the size m^* of the smallest cluster, and therefore the required cluster separation becomes smaller for large clusters. This is not the case for the second metric. The proof of (12) shows that even when we have $K = 3$ clusters, that are very large, of size $m^* = p/3$ each, the $\text{MCOD}(\Sigma)$ threshold does not decrease with m^* .

4. COD for variable clustering.

4.1. *COD Procedure.* We begin with a procedure that can be viewed as natural for model (1). It is based on the following intuition. Two indices a and b belong to the same cluster of G^* , if and only if $\text{COD}(a, b) = 0$, with COD defined in (11). Equivalently, a and b belong to the same cluster when

$$s\text{COD}(a, b) =: \max_{c \neq a, b} \frac{|\text{cov}(X_a - X_b, X_c)|}{\sqrt{\text{var}(X_b - X_a)\text{var}(X_c)}} = \max_{c \neq a, b} |\text{cor}(X_a - X_b, X_c)| = 0,$$

where $s\text{COD}$ stands for scaled COvariance Differences. In the following we work with this quantity, as it is scale invariant. It is natural to place a and b in the same cluster when the estimator $\widehat{s\text{COD}}(a, b)$ is below a certain threshold, where

$$(14) \quad \widehat{s\text{COD}}(a, b) := \max_{c \neq a, b} |\widehat{\text{cor}}(X_a - X_b, X_c)| = \max_{c \neq a, b} \left| \frac{\widehat{\Sigma}_{ac} - \widehat{\Sigma}_{bc}}{\sqrt{(\widehat{\Sigma}_{aa} + \widehat{\Sigma}_{bb} - 2\widehat{\Sigma}_{ab}) \widehat{\Sigma}_{cc}}} \right|.$$

We estimate the partition \widehat{G} according to the simple COD algorithm explained below. The algorithm does not require as input the specification of the number K of groups, which is automatically estimated by our procedure. Step 3(c) of the algorithm is called the “or” rule, and can be replaced with the “and” rule below, without changing the theoretical properties of our algorithm,

$$\widehat{G}_l = \left\{ j \in S : \widehat{s\text{COD}}(a_l, j) \vee \widehat{s\text{COD}}(b_l, j) \leq \alpha \right\}.$$

The numerical performance of these two rules are also very close through simulation studies, same as we reported on a related COD procedure on correlations [13]. Due to these small differences, we will focus on the “or” rule for the sake of space. The algorithmic complexity for computing $\widehat{\Sigma}$ is $O(p^2n)$ and the complexity of COD is $O(p^3)$, so the overall complexity of our estimation procedure is $O(p^2(p \vee n))$. The procedure is also valid when Γ has very small off-diagonal entries, and the results are presented in Section 6.

The COD Algorithm

- Input: $\widehat{\Sigma}$ and $\alpha > 0$
- Initialization: $S = \{1, \dots, p\}$ and $l = 0$
- Repeat: while $S \neq \emptyset$
 1. $l \leftarrow l + 1$
 2. If $|S| = 1$ Then $\widehat{G}_l = S$
 3. If $|S| > 1$ Then
 - (a) $(a_l, b_l) = \underset{a, b \in S, a \neq b}{\operatorname{argmin}} \widehat{s\text{COD}}(a, b)$
 - (b) If $\widehat{s\text{COD}}(a_l, b_l) > \alpha$ Then $\widehat{G}_l = \{a_l\}$
 - (c) If $\widehat{s\text{COD}}(a_l, b_l) \leq \alpha$ Then

$$\widehat{G}_l = \left\{ j \in S : \widehat{s\text{COD}}(a_l, j) \wedge \widehat{s\text{COD}}(b_l, j) \leq \alpha \right\}$$
 4. $S \leftarrow S \setminus \widehat{G}_l$
- Output: the partition $\widehat{G} = (\widehat{G}_l)_{l=1, \dots, k}$

4.2. *Perfect cluster recovery with COD for minimax optimal $\text{MCOD}(\Sigma)$ cluster separation.* Theorem 3 shows that the partition \widehat{G} produced by the COD algorithm has the property that $\widehat{G} = G^*$, with high probability, as soon as the separation $\text{MCOD}(\Sigma)$ between clusters exceeds its minimax optimal threshold established in Theorem 1 of the previous section.

THEOREM 3. *Under the distributional Assumption 1, there exists numerical constants $c_1, c_2 > 0$ such that, if*

$$\alpha \geq c_1 L^2 \sqrt{\frac{\log(p)}{n}}$$

and $\text{MCOD}(\Sigma) > 3\alpha|\Sigma|_\infty$, then we have exact cluster recovery with probability $1 - c_2/p$.

We recall that for Gaussian data, the constant $L = 1$. The proof is given in Section 3.1 of the supplement [12].

4.3. *A Data-driven Calibration procedure for COD.* The performance of the COD algorithm depends on the value of the threshold parameter α . Whereas Theorem 3 ensures that a good value for α is the order of $\sqrt{\log p/n}$, its optimal value depends on the actual distribution (at least through the subGaussian norm) and is unknown to the statistician. We propose below a new, fully data dependent, criterion for selecting α , and the corresponding partition \widehat{G} , from a set of candidate partitions \mathcal{G} . This criterion is based on data splitting: the estimators are built from a *training* sample, and then the selection involves an independent *test* sample (Hold-Out sample). The main task is to design a meaningful selection criterion. Let us consider two independent sample sets indexed by $i = 1, 2$, each of size $n/2$. The sample (1) will be a *training* dataset, and we denote by $\widehat{\mathcal{G}}^{(1)}$ a collection of partitions computed from sample (1), for instance via the COD algorithm with a varying threshold α . For any $a < b$, we set $\widehat{\Delta}_{ab}^{(i)} = \left[\widehat{\text{Cor}}^{(i)}(X_a - X_b, X_c) \right]_{c \neq a, b}$; $i = 1, 2$. Since

$\Delta_{ab} := [\text{Cor}(X_a - X_b, X_c)]_{c \neq a, b}$ equals zero if and only if $a \stackrel{\mathcal{G}}{\sim} b$, we want to select a partition G such that $\widehat{\Delta}_{ab}^{(2)} \mathbf{1}_{a \not\sim b}^G$ is a good predictor of Δ_{ab} . To implement this principle, it remains to evaluate

Δ_{ab} independently of $\widehat{\Delta}_{ab}^{(2)}$. For this evaluation, we propose to re-use the *training* sample (1) which has already been used to build the family of partitions $\widehat{\mathcal{G}}^{(1)}$. More precisely, we select $\widehat{G} \in \widehat{\mathcal{G}}^{(1)}$ by minimizing

$$\widehat{G} \in \operatorname{argmin}_{G \in \widehat{\mathcal{G}}^{(1)}} CV(G) \quad \text{with} \quad CV(G) = \sum_{a < b} \left[\left| \widehat{\Delta}_{ab}^{(2)} \mathbf{1}_{a \not\sim b} - \widehat{\Delta}_{ab}^{(1)} \right|_{\infty}^2 \right].$$

An unusual feature of the above criterion is that the *training* sample (1) is involved both in the *training* stage and in the *test* stage.

The following proposition assesses the performance of \widehat{G} . We need the following additional assumption.

(P1) If $Cor(X_a - X_b, X_c) = 0$ then $\mathbb{E} \widehat{Cor}(X_a - X_b, X_c) = 0$.

In general, the sample correlation is not an unbiased estimator of the population level correlation. Still, **(P1)** is satisfied when the data are normally distributed or in a latent model (2) when the noise variables E_a have a symmetric distribution. The next proposition provides guaranties for the CV criterion averaged over the Hold-Out sample $\mathbb{E}^{(2)}[CV(G)]$. The proof is given in Section 3.2 of the supplement [12].

PROPOSITION 2. *Assume that the distributional Assumption 1 and **(P1)** hold. Then, there exists a constant $c_1 > 0$ such that, when $MCOD(\Sigma) > c_1 |\Sigma|_{\infty} L^2 \sqrt{\log(p)/n}$, we have*

$$(15) \quad \mathbb{E}^{(2)}[CV(G^*)] \leq \min_{G \in \widehat{\mathcal{G}}^{(1)}} \mathbb{E}^{(2)}[CV(G)],$$

both with probability larger than $1 - 4/p$ and in expectation with respect to $\mathbb{P}^{(1)}$.

Under the condition $MCOD(\Sigma) > c_1 |\Sigma|_{\infty} L^2 \sqrt{\log(p)/n}$, Theorem 3 ensures that G^* belongs to $\widehat{\mathcal{G}}^{(1)}$ with high probability, whereas (15) suggests that the CV criterion is minimized at G^* .

If we consider a CV algorithm based on $\widehat{COD}(a, b)$ instead of $\widehat{sCOD}(a, b)$, then we can obtain a counterpart of Proposition 2 without requiring the additional assumption **(P1)**. Still, we favor the procedure based on $\widehat{sCOD}(a, b)$ mainly for its scale-invariance property.

5. Penalized convex K -means: PECOK.

5.1. *PECOK Algorithm.* Motivated by the fact that the COD algorithm is minimax optimal with respect to the $MCOD(\Sigma)$ metric, but not necessarily with respect to the $\Delta(C^*)$ metric (unless the size of the smallest cluster is constant), we propose below an alternative procedure, that adapts to this metric. Our second method is a natural extension of one of the most popular clustering strategies. When we view the G -block covariance model as arising via the latent factor representation in **(i)** in the Introduction, the canonical clustering approach would be via the K -means algorithm [29], which is NP-hard [4]. Following Peng and Wei [32], we consider a convex relaxation of it, which is computationally feasible in polynomial time. We argue below that, for estimating clusters given by (1), one needs to further tailor it to our model. The statistical analysis of the modified procedure is the first to establish consistency of variable clustering via K -means type procedures, to the best of our knowledge.

The estimator offered by the standard K -means algorithm, with the number K of groups of G^* known, is

$$(16) \quad \widehat{G} \in \operatorname{argmin}_G \operatorname{crit}(\mathbf{X}, G) \quad \text{with} \quad \operatorname{crit}(\mathbf{X}, G) = \sum_{a=1}^p \min_{k=1, \dots, K} \|\mathbf{X}_{:a} - \bar{\mathbf{X}}_{G_k}\|^2,$$

and $\bar{\mathbf{X}}_{G_k} = |G_k|^{-1} \sum_{a \in G_k} \mathbf{X}_{:a}$.

For a partition G , let us introduce the corresponding partnership matrix B by

$$(17) \quad B_{ab} = \begin{cases} \frac{1}{|G_k|} & \text{if } a \text{ and } b \text{ are in the same group } G_k, \\ 0 & \text{if } a \text{ and } b \text{ are in different groups.} \end{cases}$$

we observe that $B_{ab} > 0$ if and only if $a \stackrel{G}{\sim} b$. In particular, there is a one-to-one correspondence between partitions G and their corresponding partnership matrices. It is shown in Peng and Wei [32] that the collection of such matrices B is described by the collection \mathcal{O} of orthogonal projectors fulfilling $\text{tr}(B) = K$, $B1 = 1$ and $B_{ab} \geq 0$ for all a, b .

Theorem 2.2 in Peng and Wei [32] shows that solving the K -means problem is equivalent to finding the global maximum

$$(18) \quad \bar{B} = \underset{B \in \mathcal{O}}{\text{argmax}} \langle \hat{\Sigma}, B \rangle,$$

and then recovering \hat{G} from \bar{B} .

The set of orthogonal projectors is not convex, so, following Peng and Wei [32], we consider a convex relaxation \mathcal{C} of \mathcal{O} obtained by relaxing the condition "B orthogonal projector", by "B positive semi-definite", leading to

$$(19) \quad \mathcal{C} := \left\{ B \in \mathbf{R}^{p \times p} : \begin{array}{l} \bullet B \succcurlyeq 0 \text{ (symmetric and positive semidefinite)} \\ \bullet \sum_a B_{ab} = 1, \forall b \\ \bullet B_{ab} \geq 0, \forall a, b \\ \bullet \text{tr}(B) = K \end{array} \right\}.$$

Thus, the (uncorrected) convex relaxation of K -means is equivalent with finding

$$(20) \quad \tilde{B} = \underset{B \in \mathcal{C}}{\text{argmax}} \langle \hat{\Sigma}, B \rangle.$$

To assess the relevance of this estimator, we first study its behavior at the population level, when $\hat{\Sigma}$ is replaced by Σ in (20). Indeed, if the minimizer of our criterion does not recover the true partition at the population level, we cannot expect it to be consistent, even in a large sample asymptotic context (fixed p , n goes to infinity). We recall that $|\Gamma|_V := \max_a \Gamma_{aa} - \min_a \Gamma_{aa}$.

PROPOSITION 3. *Assume that $\Delta(C^*) > 2|\Gamma|_V/m^*$. Then, $B^* = \underset{B \in \mathcal{O}}{\text{argmax}} \langle \Sigma, B \rangle$. If $\Delta(C^*) > 7|\Gamma|_V/m^*$, then $B^* = \underset{B \in \mathcal{C}}{\text{argmax}} \langle \Sigma, B \rangle$.*

For $\Delta(C^*)$ large enough, the population version of convexified K -means recovers B^* . The next proposition illustrates that the condition $\Delta(C^*) > 2|\Gamma|_V/m^*$ for population K -means is in fact necessary.

PROPOSITION 4. *Consider the model (1) with*

$$C^* = \begin{bmatrix} \alpha & 0 & 0 \\ 0 & \beta & \beta - \tau \\ 0 & \beta - \tau & \beta \end{bmatrix}, \quad \Gamma = \begin{bmatrix} \gamma_+ & 0 & 0 \\ 0 & \gamma_- & 0 \\ 0 & 0 & \gamma_- \end{bmatrix}, \quad \text{and } |G_1^*| = |G_2^*| = |G_3^*| = m^*.$$

The population maximizer $B_\Sigma = \underset{B \in \mathcal{O}}{\text{argmax}} \langle \Sigma, B \rangle$ is not equal to B^ as soon as $2\tau = \Delta(C^*) < \frac{2}{m^*} |\Gamma|_V$.*

The two propositions above are proved in Appendix A.1. As a consequence, when Γ is not proportionnal to the identity matrix, the population minimizers based on K -means and convexified K -means do not necessary recover the true partition even when the "within-between group" covariance gap is strictly positive. This undesirable behavior of K -means is not completely unexpected as K -means is a quantization algorithm which aims to find for clusters of similar width, instead of "homogeneous" clusters. Hence, we need to modify it for our purpose.

This leads us to suggesting a population level correction in Proposition 3. Indeed, as a direct Corollary of Proposition 3, we have

$$B^* = \operatorname{argmin}_{B \in \mathcal{C}} \langle \Sigma - \Gamma, B \rangle ,$$

as long as $\Delta(C^*) > 0$. This suggests the following **Penalized Convex K-means** (PECOK) algorithm, in three steps. The main step 2 produces an estimator \widehat{B} of B from which we derive the estimated partition \widehat{G} . We summarize this below.

The PECOK algorithm

Step 1. Estimate Γ by $\widehat{\Gamma}$.

Step 2. Estimate B^* by $\widehat{B} = \operatorname{argmax}_{B \in \mathcal{C}} \left(\langle \widehat{\Sigma}, B \rangle - \langle \widehat{\Gamma}, B \rangle \right)$.

Step 3. Estimate G^* by applying a clustering algorithm to the columns of \widehat{B} .

The required inputs for Step 2 of our algorithm are: (i) $\widehat{\Sigma}$, the sample covariance matrix; (ii) $\widehat{\Gamma}$, the estimator produced at Step 1; and (iii) K , the number of groups. Our only requirement on the clustering algorithm applied in Step 3 is that it succeeds to recover the partition G^* when applied to true partnership matrix B^* . The standard K -means algorithm [29] seeded with K distinct centroids, `kmeans++` [3], or any approximate K -means as defined in (28) in Section 5.4, fulfill this property. This step is done at no additional statistical accuracy cost, as shown in Corollary 4 below.

We view the term $\langle \widehat{\Gamma}, B \rangle$ as a penalty term on B , with data dependent weights $\widehat{\Gamma}$. Therefore, the construction of an accurate estimator $\widehat{\Gamma}$ of Γ is a crucial step for guaranteeing the statistical optimality of the PECOK estimator.

5.2. *Construction of $\widehat{\Gamma}$.* Estimating Γ before estimating the partition itself is a non-trivial task, and needs to be done with care. We explain our estimation below and analyze it in Proposition 8 in Appendix A.4. We show that this estimator of Γ is appropriate whenever Γ is a diagonal matrix (or diagonally dominant, with small off-diagonal entries). For any $a, b \in [p]$, define

$$(21) \quad V(a, b) := \max_{c, d \in [p] \setminus \{a, b\}} \frac{\left| (\widehat{\Sigma}_{ac} - \widehat{\Sigma}_{ad}) - (\widehat{\Sigma}_{bc} - \widehat{\Sigma}_{bd}) \right|}{\sqrt{\widehat{\Sigma}_{cc} + \widehat{\Sigma}_{dd} - 2\widehat{\Sigma}_{cd}}},$$

with the convention $0/0 = 0$. Guided by the block structure of Σ , we define

$$b_1(a) := \operatorname{argmin}_{b \in [p] \setminus \{a\}} V(a, b) \quad \text{and} \quad b_2(a) := \operatorname{argmin}_{b \in [p] \setminus \{a, b_1(a)\}} V(a, b),$$

to be two elements "close" to a , that is two indices $b_1 = b_1(a)$ and $b_2 = b_2(a)$ such that the empirical covariance difference $\widehat{\Sigma}_{b_i c} - \widehat{\Sigma}_{b_i d}$, $i = 1, 2$, is most similar to $\widehat{\Sigma}_{ac} - \widehat{\Sigma}_{ad}$, for all variables c and d not

equal to a or b_i , $i = 1, 2$. It is expected that $b_1(a)$ and $b_2(a)$ either belong to the same group as a , or belong to some "close" groups. Then, our estimator $\widehat{\Gamma}$ is a diagonal matrix, defined by

$$(22) \quad \widehat{\Gamma}_{aa} = \widehat{\Sigma}_{aa} + \widehat{\Sigma}_{b_1(a)b_2(a)} - \widehat{\Sigma}_{ab_1(a)} - \widehat{\Sigma}_{ab_2(a)}, \quad \text{for } a = 1, \dots, p.$$

Intuitively, $\widehat{\Gamma}_{aa}$ should be close to $\Sigma_{aa} + \Sigma_{b_1(a)b_2(a)} - \Sigma_{ab_1(a)} - \Sigma_{ab_2(a)}$, which is equal to Γ_{aa} in the favorable event where both $b_1(a)$ and $b_2(a)$ belong to the same group as a .

In general, $b_1(a)$ and $b_2(a)$ cannot be guaranteed to belong to the same group as a . Nevertheless, these two surrogates $b_1(a)$ and $b_2(a)$ are close enough to a so that $|\widehat{\Gamma}_{aa} - \Gamma_{aa}|$ to be at most of the order of $|\Gamma|_\infty \sqrt{\log(p)/n}$ in ℓ^∞ -norm, as shown in Proposition 8 in Appendix A.4. In the next subsection, we show that $\widehat{\Gamma}$ is good enough to ensure that PECOK perfectly recovers G^* under minimal separation condition.

Note that PECOK requires the knowledge of the true number K of groups. When the number K of groups itself is unknown, we can modify the PECOK criterion by adding a penalty term as explained in a previous version of our work [14, Sec. 4]. Alternatively, we propose in Section 7 a simple cross-validation procedure.

5.3. *Perfect cluster recovery with PECOK for near-minimax Δ -cluster separation.* We show in this section that the PECOK estimator recovers the clusters exactly, with high probability, at a near-minimax separation rate with respect to the $\Delta(C^*)$ metric.

THEOREM 4. *There exist c_1, c_2, c_3 three positive constants such that the following holds. Let $\widehat{\Gamma}$ be any estimator of Γ , such that $|\widehat{\Gamma} - \Gamma|_V \leq \delta_{n,p}$ with probability $1 - c_1/p$. Then, under Assumption 1, and when $L^4 \log(p) \leq c_3 n$ and*

$$(23) \quad \Delta(C^*) \geq c_L \left[\|\Gamma\|_\infty \left\{ \sqrt{\frac{\log p}{m^* n}} + \sqrt{\frac{p}{nm^{*2}}} + \frac{\log(p)}{n} + \frac{p}{nm^*} \right\} + \frac{\delta_{n,p}}{m^*} \right],$$

then $\widehat{B} = B^*$, with probability higher than $1 - c_1/p$. Here, c_L is a positive constant that only depends on L in Assumption 1. In particular, if $\widehat{\Gamma}$ is the estimator (22), the same conclusion holds with probability higher than $1 - c_2/p$ when

$$(24) \quad \Delta(C^*) \geq c_L \|\Gamma\|_\infty \left\{ \sqrt{\frac{\log p}{m^* n}} + \sqrt{\frac{p}{nm^{*2}}} + \frac{\log(p)}{n} + \frac{p}{nm^*} \right\}.$$

The proof is given in Appendix A.3.

Remark 1. We left the term $\delta_{n,p}$ explicit in (23) in order to make clear how the estimation of Γ affects the cluster separation $\Delta(C^*)$ metric. Without a correction (i.e. taking $\widehat{\Gamma} = 0$), the term $\delta_{n,p}/m^*$ equals $|\Gamma|_V/m^*$ which is non zero (and does not decrease in a high-sample asymptotic) unless Γ has equal diagonal entries. This phenomenon is consistent with the population analysis in the previous subsection. Display (24) shows that the separation condition can be much decreased with the correction. In particular, for balanced clusters, exact recovery is guaranteed when

$$(25) \quad \Delta(C^*) \geq c_L \left[\sqrt{\frac{K \vee \log p}{m^* n}} + \frac{K \vee \log p}{n} \right],$$

for an appropriate constant $c_L > 0$. In view of Theorem 2 the rate is minimax optimal when the number of clusters $K \leq \log(p)$. For an even larger number of clusters ($K \geq \log(p)$), the rate is

only minimax up to some loss. For instance, if the clusters are balanced, we possibly lose a factor $K/\log(p)$ relative to the optimal rate. As discussed in the introduction, this gap is possibly due to a computational barrier and we refer to [15] for more details.

Bounded variables X also follow subGaussian distribution. Nevertheless, the corresponding subGaussian norm L may be large and Theorem 4 can sometimes be improved, as in Theorem 5 below, proved in Appendix A.3

THEOREM 5. *There exist c_1, c_2, c_3 three positive constants such that the following holds. Let $\widehat{\Gamma}$ be any estimator of Γ , such that $|\widehat{\Gamma} - \Gamma|_V \leq \delta_{n,p}$ with probability $1 - c_1/p$. Then, under Assumption 1-bis, and*

$$(26) \quad \Delta(C^*) \geq c_2 \left[M \|\Gamma\|_\infty^{1/2} \sqrt{\frac{p \log(p)}{nm^{*2}}} + M^2 \frac{p \log(p)}{nm^*} + \frac{\delta_{n,p}}{m^*} \right].$$

then $\widehat{B} = B^*$, with probability higher than $1 - c_1/p$.

When we choose $\widehat{\Gamma}$ as in (22), the term $\delta_{n,p}/m^*$ can be simplified as under Assumption 1, see Proposition 8 in Appendix A.4. For balanced clusters, Condition (26) can be simplified in

$$\Delta(C^*) \geq c_2 \left[M \|\Gamma\|_\infty^{1/2} \sqrt{\frac{K \log(p)}{nm^*}} + M^2 \frac{K \log(p)}{n} + \frac{\delta_{n,p}}{m^*} \right].$$

In comparison to (25), the condition does not depend anymore on the subGaussian nom L , but the term $K \vee \log(p)$ has been replaced by $K \log(p)$.

Remark 2. For the Ising Block Model (3) with K balanced groups, we have $M = 1$ and $p = m^*K$, $C^* = (\omega_{in} - \omega_{out})I_K + \omega_{out}J$ and $\Gamma = (1 - \omega_{in})I_K$. As a consequence, no diagonal correction is needed, that is we can take $\widehat{\Gamma} = 0$, and since $|\Gamma|_V = 0$, we have $\delta_{n,p} = 0$. Then, for K balanced groups, condition (26) simplifies to

$$(\omega_{in} - \omega_{out}) \gtrsim K \sqrt{\frac{\log(p)}{np}} + \frac{K \log(p)}{n}$$

In the specific case $K = 2$, we recover (up to numerical multiplicative constants) the optimal rate proved in [8]. Our procedure and analysis provide a generalization of these results, as they are valid for general K and Theorem 5 also allows for unequal groups.

5.4. *A comparison between PECOK and Spectral Clustering.* In this section we discuss connections between the clustering methods introduced above and spectral clustering, a method that has become popular in network clustering. When used for variable clustering, uncorrected spectral clustering consists in applying a clustering algorithm, such as K -means, on the rows of the $p \times K$ -matrix obtained by retaining the K leading eigenvectors of $\widehat{\Sigma}$.

SC algorithm

1. Compute \widehat{V} , the matrix of the K leading eigenvectors of $\widehat{\Sigma}$
2. Estimate G^* by applying a (rotation invariant) clustering method to the rows of \widehat{V} .

First, we recall the premise of spectral clustering, adapted to our context. For G^* -block covariance models as (1), we have $\Sigma - \Gamma = AC^*A^t$. Let U be the $p \times K$ matrix collecting the K leading

eigenvectors of $\Sigma - \Gamma$. It has been shown, see e.g. Lemma 2.1 in Lei and Rinaldo [27], that a and b belong to the same cluster if and only if $U_{a\cdot} = U_{b\cdot}$ and if and only if $[UU^t]_{a\cdot} = [UU^t]_{b\cdot}$. Arguing as in Peng and Wei [32], we have the following.

LEMMA 2. *SC algorithm is equivalent to the following algorithm:*

Step 1. Find $\bar{B} = \operatorname{argmax}\{\langle \hat{\Sigma}, B \rangle : \operatorname{tr}(B) = K, I \succcurlyeq B \succcurlyeq 0\}$.

Step 2. Estimate G^* by applying a (rotation invariant) clustering method to the rows of \bar{B} .

The connection between (unpenalized) PECOK and spectral clustering now becomes clear. The (unpenalized) PECOK estimator \tilde{B} (20) involves the calculation of

$$(27) \quad \tilde{B} = \operatorname{argmax}_B \{\langle \hat{\Sigma}, B \rangle : B1 = 1, B_{ab} \geq 0, \operatorname{tr}(B) = K, B \succcurlyeq 0\}.$$

Since the matrices B involved in (27) are doubly stochastic, their eigenvalues are smaller than 1 and hence (27) is equivalent to $\tilde{B} = \operatorname{argmax}_B \{\langle \hat{\Sigma}, B \rangle : B1 = 1, B_{ab} \geq 0, \operatorname{tr}(B) = K, I \succcurlyeq B \succcurlyeq 0\}$. Note then that \bar{B} can be viewed as a less constrained version of \tilde{B} , in which \mathcal{C} is replaced by $\bar{\mathcal{C}} = \{B : \operatorname{tr}(B) = K, I \succcurlyeq B \succcurlyeq 0\}$, where we have dropped the $p(p+1)/2$ constraints given by $B1 = 1$, and $B_{ab} \geq 0$. We show in what follows that the possible computational gains resulting from such a strategy may result in severe losses in the theoretical guarantees for exact partition recovery. In addition, the proof of Lemma 2 shows that $\bar{B} = \hat{V}\hat{V}^t$, so, contrary to \tilde{B} , the estimator \bar{B} is (almost surely) never equal to B^* .

In view of this connection between Spectral clustering and unpenalized PECOK and on the fact that the population justification of spectral clustering deals with the spectral decomposition of $\Sigma - \Gamma$, this leads to propose the following corrected version of the algorithm based on $\tilde{\Sigma} := \hat{\Sigma} - \hat{\Gamma}$.

CSC algorithm

1. Compute \hat{U} , the matrix of the K leading eigenvectors of $\tilde{\Sigma} := \hat{\Sigma} - \hat{\Gamma}$
2. Estimate G^* by clustering the rows of \hat{U} , via an η -approximation of K -means (28).

For $\eta > 1$, an η -approximation of K -means is a clustering algorithm producing a partition \hat{G} such that

$$(28) \quad \operatorname{crit}(\hat{U}^t, \hat{G}) \leq \eta \min_G \operatorname{crit}(\hat{U}^t, G),$$

with $\operatorname{crit}(\cdot, \cdot)$ the K -means criterion (16). Although solving K -means is NP-Hard [4], there exist polynomial time approximate K -means algorithms, see Kumar *et al.* [25]. As a consequence of the above discussion, the first step of CSC can be interpreted as a relaxation of the program associated to PECOK estimator \tilde{B} .

In the sequel, we provide some results for CSC procedure. To simplify the presentation, we assume in the following that all the groups have the same size $|G_1^*| = \dots = |G_K^*| = m = p/K$. We emphasize that this information is not required by either PECOK or CSC, or in the proof of Proposition 5 below. We only use it here to illustrate the issues associated with CSC in a way that is not cluttered by unnecessary notation. We denote by \mathcal{S}_K the set of permutations on $\{1, \dots, K\}$ and we denote by

$$\bar{L}(\hat{G}, G^*) = \min_{\sigma \in \mathcal{S}_K} \sum_{k=1}^K \frac{|G_k^* \setminus \hat{G}_{\sigma(k)}|}{m}$$

the sum of the ratios of miss-assigned variables with indices in G_k^* . In the previous sections, we studied perfect recovery of G^* , which would correspond to $\bar{L}(\hat{G}, G^*) = 0$. We give below conditions under which $\bar{L}(\hat{G}, G^*) \leq \rho$, for an appropriate quantity $\rho < 1$, and we show that very small values of ρ require large cluster separation, possibly much larger than the minimax optimal rate. We begin with a general theorem pertaining to partial partition recovery by CSC, under restrictions on the smallest eigenvalue $\lambda_K(C^*)$ of C^* .

PROPOSITION 5. *We let $Re(\Sigma) = tr(\Sigma)/\|\Sigma\|_{op}$ denote the effective rank of Σ . There exist $c_{\eta,L} > 0$ only depending on η and L and a numerical constant c_1 such that the following holds under Assumption 1. For any $0 < \rho < 1$, if*

$$(29) \quad \lambda_K(C^*) \geq \frac{c_{\eta,L}\sqrt{K}\|\Sigma\|_{op}}{m^*\sqrt{\rho}} \sqrt{\frac{Re(\Sigma) \vee \log(p)}{n}},$$

then $\bar{L}(\hat{G}, G^*) \leq \rho$, with probability larger than $1 - c_1/p$.

The proof extends the arguments of [27], initially developed for clustering procedures in stochastic block models, to our context. Specifically, we relate the error $\bar{L}(\hat{G}, G^*)$ to the noise level, quantified in this problem by $\|\tilde{\Sigma} - AC^*A^t\|_{op}$. We then employ the results of [23] to show that this operator norm can be controlled, with high probability, which leads to the conclusion of the theorem.

We observe that $\Delta(C^*) \geq 2\lambda_K(C^*)$, so the lower bound (29) on $\lambda_K(C^*)$ enforces the same lower-bound on $\Delta(C^*)$. As n goes to infinity, the right hand side of Condition (29) goes to zero, and CSC is therefore consistent in a large sample asymptotic. In contrast, we emphasize that (uncorrected) SC algorithm is not consistent as can be shown by a population analysis similar to that of Proposition 4.

To further facilitate the comparison between CSC and PECOK, we discuss both the conditions and the conclusion of this theorem in the simple setting where $C^* = \tau I_K$ and $\Gamma = I_p$. Then, the cluster separation measures coincide up to a factor 2, $\Delta(C^*) = 2\lambda_K(C^*) = 2\tau$.

COROLLARY 2 (Illustrative example: $C^* = \tau I_K$ and $\Gamma = I_p$). *There exist three positive numerical constants $c_{\eta,L}$, $c'_{\eta,L}$ and c_3 such that the following holds under Assumption 1. For any $0 < \rho < 1$, if*

$$(30) \quad \rho \geq c_{\eta,L} \left[\frac{K^2}{n} + \frac{K \log(p)}{n} \right] \quad \text{and} \quad \tau \geq c'_{\eta,L} \left[\frac{K^2}{\rho n} \vee \frac{K}{\sqrt{\rho n m}} \right],$$

then $\bar{L}(\hat{G}, G^*) \leq \rho$, with probability larger than $1 - c_3/p$.

Recall that, as a benchmark, Theorem 4 above states that, when \hat{G} is obtained via the PECOK algorithm, and if $\tau \gtrsim \sqrt{\frac{K \vee \log p}{mn}} + \frac{\log(p) \vee K}{n}$, then $\bar{L}(\hat{G}, G^*) = 0$, or equivalently, $\hat{G} = G^*$, with high probability. We can therefore provide the following comparison.

- If we consider ρ as a user specified small value, independent of n or p , and if the number of groups K is either a constant or grows at most as $\log p$, then the size of the cluster separation given by either Condition (30) or by PECOK are essentially the same, up to unavoidable $\log p$ factors. The difference is that, in this regime, CSC guarantees recovery up to a fixed, small, fraction of mistakes, whereas PECOK guarantees exact recovery.

- Although perfect recovery, with high probability, cannot be guaranteed for CSC, we could be close to it by requiring ρ to be close to zero. In this case, the distinctions between Conditions (30) and that for PECOK become much more pronounced.
- When we move away from the case $C^* = \tau I_K$, the comparison becomes even less favorable to CSC. For instance, when $\Gamma = I$ and $C^* = \tau I_K + \alpha J$, with J being the matrix with all entries equal to one, as in the Ising Block model discuss page 16. Notice that in this case we continue to have $\Delta(C^*) = 2\lambda_K(C^*) = 2\tau$. Then, for a given, fixed, value of ρ and K fixed, condition (29) requires the cluster separation

$$\tau \gtrsim \frac{\alpha \sqrt{\log(p)}}{\sqrt{n\rho}},$$

which is independent of m , unlike the minimax cluster separation rate that we established in Theorem 2 above. Therefore, the correction strategies employed in SBM are not directly transferable to variable clustering, which further supports the merits our PECOK method.

All the results of this section are proved in Section 5 of the supplement [12].

6. Approximate G-block covariance models. In the previous sections, we have proved that under some separation conditions, COD and PECOK procedures are able to exactly recover the partition G^* . However, in practical situations, the separation conditions may not be met. Besides, if the entries of Σ have been modified by an infinitesimal perturbation, then the corresponding partition G^* would consist of p singletons.

As a consequence, it may be more realistic and more appealing from a practical point of view to look for a partition $G[K]$ with $K < |G^*|$ groups such that Σ is close to a matrix of the form $ACA^t + \Gamma$ where Γ is diagonal and A is associated to $G[K]$. This is equivalent to considering a decomposition $\Sigma = ACA^t + \Gamma$ with Γ non-diagonal, where the non-diagonal entries of Γ are small. In the sequence, we write $R = \Gamma - \text{Diag}(\Gamma)$ for the matrix of the off-diagonal elements of Γ and $D = \text{Diag}(\Gamma)$ for the diagonal matrix given by the diagonal of Γ .

In the next subsection, we discuss under which conditions the partition $G[K]$ is identifiable and then, we prove that COD and PECOK are able to recover these partitions.

6.1. Identifiability of approximate G-block covariance models. When Γ is allowed to be not exactly equal to a diagonal matrix, we encounter a further identifiability issue, as a generic matrix Σ may admit many decompositions $\Sigma = ACA^t + \Gamma$. In fact, such a decomposition holds for any membership matrix A and any matrix C if we define $\Gamma = \Sigma - ACA^t$. So we need to specify the kind of decomposition that we are looking for. For K being fixed, we would like to consider the partition G with K clusters that maximizes the distance between groups (e.g. $\text{MCOD}(\Sigma, G)$) while having the smallest possible noise term $|R|_\infty$. Unfortunately, such a partition G does not necessarily exist and is not necessarily unique. Let us illustrate this situation with a simple example.

Example. Assume that Σ is given by $\Sigma = \begin{bmatrix} 2r & 0 & 0 \\ 0 & 2r & 0 \\ 0 & 0 & 2r \end{bmatrix} + I_p$, with $r > 0$, with the convention that each entry corresponds to a block of size 2. Considering partitions with 2 groups and allowing Γ to be non diagonal, we can decompose Σ using different partitions. For instance

$$\Sigma = \underbrace{\begin{bmatrix} 2r & 0 & 0 \\ 0 & r & r \\ 0 & r & r \end{bmatrix}}_{=A_1 C_1 A_1^t} + \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ 0 & r & -r \\ 0 & -r & r \end{bmatrix}}_{=\Gamma_1} + I_p = \underbrace{\begin{bmatrix} r & r & 0 \\ r & r & 0 \\ 0 & 0 & 2r \end{bmatrix}}_{=A_2 C_2 A_2^t} + \underbrace{\begin{bmatrix} r & -r & 0 \\ -r & r & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{=\Gamma_2} + I_p.$$

Importantly, the two decompositions correspond to two different partitions G_1 and G_2 and both decompositions have $|R_i|_\infty = r$ and $\text{MCOD}(\Sigma, G_i) = 2r = 2|R|_\infty$, for $i = 1, 2$. In addition, no decomposition $\Sigma = ACA^t + D + R$ with associated partition in 2 groups, satisfies $\text{MCOD}(\Sigma, G) > 2r$ or $|R|_\infty < r$. As a consequence, there is no satisfying way to define a unique partition maximizing $\text{MCOD}(\Sigma, G)$, while having $|R|_\infty$ as small as possible. We show below that the cutoff $\text{MCOD}(\Sigma, G) > 2|R|_\infty$ is actually sufficient for partition identifiability.

For this, let us define $\mathcal{P}_j(\Sigma, K)$, $j = \{1, 2\}$ as the set of quadruplets (A, C, D, R) such that $\Sigma = ACA^t + D + R$, with A a membership matrix associated to a partition G with K groups with $\min_k |G_k| \geq j$, and D and R defined as above. Hence \mathcal{P}_1 corresponds to partitions without restrictions on the minimum group size. For instance, singletons are allowed. In contrast \mathcal{P}_2 only contains partitions without singletons. We define

$$\begin{aligned}\rho_1(\Sigma, K) &= \max \{ \text{MCOD}(\Sigma, G) / |R|_\infty : (A, C, D, R) \in \mathcal{P}_1(\Sigma, K) \text{ and } G \text{ associated to } A \}, \\ \rho_2(\Sigma, K) &= \max \{ \Delta(C) / |R|_\infty : (A, C, D, R) \in \mathcal{P}_2(\Sigma, K) \}.\end{aligned}$$

We view ρ_1 and ρ_2 as respective measures of ‘‘purity’’ of the block structure of Σ .

PROPOSITION 6.

- (i) Assume that $\rho_1(\Sigma, K) > 2$. Then, there exists a unique partition G such that there exists a decomposition $\Sigma = ACA^t + \Gamma$, with A associated to G and $\text{MCOD}(\Sigma, G) > 2|R|_\infty$. We denote by $G_1[K]$ this partition.
- (ii) Assume that $\rho_2(\Sigma, K) > 8$. Then, there exists a unique partition G with $\min_k |G_k| \geq 2$, such that there exists a decomposition $\Sigma = ACA^t + \Gamma$, with A associated to G and $\Delta(C) > 8|R|_\infty$. We denote by $G_2[K]$ this partition.
- (iii) In addition, if both $\rho_1(\Sigma, K) > 2$ and $\rho_2(\Sigma, K) > 8$, then $G_1[K] = G_2[K]$.

The conditions $\rho_1(\Sigma, K) > 2$ and $\rho_2(\Sigma, K) > 8$ are minimal for defining uniquely the partition $G_1[K]$. For ρ_1 , this has been illustrated in the example above the proposition. For ρ_2 , we provide a counter example when $\rho_2(\Sigma, K) = 8$ in Section 1.3 of the supplement [12]. The proof of Proposition 6. is given in Section 1.2 of [12].

The conclusion of Proposition 6 does essentially revert to that of Proposition 1 of Section 2 as soon as $|R|_\infty$ is small enough respective to the cluster separation sizes. Denoting K^* the number of groups of G^* , we observe that $G_1[K^*] = G^*$ and $G_2[K^*] = G^*$ if $m^* \geq 2$. Besides, $\rho_1(\Sigma, K) = \rho_2(\Sigma, K) = 0$ for $K > K^*$. For $K < K^*$ and when $G_1[K]$ (resp. $G_2[K]$) are well defined, then the partition $G_1[K]$ (resp. $G_2[K]$) is coarser than G^* . In other words, $G_1[K]$ is derived from G^* by merging groups G_k^* thereby increasing $\text{MCOD}(\Sigma, G)$ (resp. $\Delta(C)$) while requiring $|R|_\infty$ to be small enough.

We point out that, in general, there is no unique decomposition $\Sigma = ACA^t + \Gamma$ with A associated to $G_2[K]$, even when $\min_k |G_2[K]_k| \geq 2$. Actually, it can be possible to change some entries of C and R , while keeping $C + R$, $\Delta(C)$ and $|R|_\infty$ unchanged.

6.2. *The COD algorithm for approximate G-block covariance models.* We show below that the COD algorithm is still applicable if Σ has small departures from a block structure. We set $\lambda_{\min}(\Sigma)$ for the smallest eigenvalue of Σ .

THEOREM 6. *Under the distributional Assumption 1, there exist numerical constants $c_1, c_2 > 0$ such that the following holds for all $\alpha \geq c_1 L^2 \sqrt{\frac{\log p}{n}}$. If, for some partition G and decomposition $\Sigma = ACA^t + R + D$, we have*

$$(31) \quad |R|_\infty \leq \frac{\lambda_{\min}(\Sigma)}{2\sqrt{2}}\alpha \quad \text{and} \quad \text{MCOD}(\Sigma, G) > 3\alpha|\Sigma|_\infty,$$

then COD recovers G with probability higher than $1 - c_2/p$.

The proof is given in Section 3.1 of the supplement [12]. If G satisfies the assumptions of Theorem 6, then it follows from Proposition 6 that $G = G_1[K]$ for some $K > 0$. First, consider the situation where the tuning parameter α is chosen to be of the order $\sqrt{\log(p)/n}$. If $\text{MCOD}(\Sigma, G^*) \geq 3\alpha|\Sigma|_\infty$, then COD selects G^* with high probability. If $\text{MCOD}(\Sigma, G^*)$ is smaller than this threshold, then no procedure is able to recover G^* with high probability (Theorem 1). Nevertheless, COD is able to recover a coarser partition $G_1[K]$ whose corresponding MCOD metric $\text{MCOD}(\Sigma, G)$ is higher than the threshold $3\alpha|\Sigma|_\infty$ and whose matrix R is small enough. For larger α , then COD recovers a coarser partition G (corresponding to $G_1[K]$ with a smaller K) whose corresponding approximation $|R|_\infty$ is allowed to be larger.

6.3. The PECOK algorithm for approximate G -block covariance models. In this subsection, we investigate the behavior of PECOK under the approximate G -block models. The number K of groups being fixed, we assume that $\rho_2(\Sigma, K) > 8$ so that $G_2[K]$ is well defined. We shall prove that PECOK recovers $G_2[K]$ with high probability. By abusing the notation, we denote in this subsection G^* for the target partition $G_2[K]$, B^* for the associated partnership matrix and $(A, C^*, D, R) \in \mathcal{P}_2(\Sigma, K)$ any decomposition of Σ maximizing $\Delta(C)/|R|_\infty$.

Similarly to Proposition 3, we first provide sufficient conditions on C^* under which a population version of PECOK can recover the true partition.

PROPOSITION 7. *If, $\Delta(C^*) > \frac{7|D|_V + 2\|R\|_{op}}{m} + 3|R|_\infty$, then $B^* = \operatorname{argmin}_{B \in \mathcal{C}} \langle \Sigma, B \rangle$.*

COROLLARY 3. *If $\Delta(C^*) > 3|R|_\infty + \frac{2\|R\|_{op}}{m}$, then $B^* = \operatorname{argmin}_{B \in \mathcal{C}} \langle \Sigma - D, B \rangle$.*

In contrast to the exact G -block model, the cluster distance $\Delta(C^*)$ now needs to be larger than $|R|_\infty$ for the population version to recover the true partition. The $|R|_\infty$ condition is fact necessary as discussed in subsection 6.1. In comparison to the necessary conditions discussed in subsection 6.1, there is an additional $\|R\|_{op}/m$ term. The proofs are given in Appendix A.2.

We now examine the behavior of PECOK when we specify the estimator $\widehat{\Gamma}$ to be as in (22). Note that in this approximate block covariance setting, the diagonal estimator $\widehat{\Gamma}$ is in fact an estimator of the diagonal matrix D . In order to derive deviation bounds for our estimator $\widehat{\Gamma}$, we need the following diagonal dominance assumption.

Assumption 2: (diagonal dominance of Γ) The matrix $\Gamma = D + R$ fulfills

$$(32) \quad \Gamma_{aa} \geq 3 \max_{c:c \neq a} |\Gamma_{ac}| \quad (\text{or equivalently} \quad D_{aa} \geq 3 \max_{c:c \neq a} |R_{ac}|).$$

The next theorem states that PECOK estimator \widehat{B} recovers the groups under similar conditions to that of Theorem 4 if R is small enough. The proof is given in Appendix A.3, with proofs of intermediate results given in Section 4 of the supplement [12].

THEOREM 7. *There exist c_1, c_2, c_L, c'_L four positive constants such that the following holds. Under Assumptions 1 and 2, and when $L^4 \log(p) \leq c_1 n$ and*

$$(33) \quad |R|_\infty + \frac{\sqrt{|R|_\infty |D|_\infty} + \|R\|_{op}}{m} \leq c_L \|\Gamma\|_{op} \left\{ \sqrt{\frac{\log p}{mn}} + \sqrt{\frac{p}{nm^2}} + \frac{\log(p)}{n} + \frac{p}{nm} \right\}$$

we have $\widehat{B} = B^*$, with probability higher than $1 - c_2/p$, as soon as

$$(34) \quad \Delta(C^*) \geq c'_L \left[\|\Gamma\|_{op} \left\{ \sqrt{\frac{\log p}{mn}} + \sqrt{\frac{p}{nm^2}} + \frac{\log(p)}{n} + \frac{p}{nm} \right\} \right],$$

So, as long as $|R|_\infty$ and $\|R\|_{op}$ are small enough so that (33) are satisfied, PECOK algorithm will correctly identify the target partition G^* at the Δ -(near) optimal minimax level (34). A counterpart of Theorem 7 for Assumption 1-bis is provided in Appendix A.3.

7. Simulation results. In this section we verify numerically our theoretical findings and also illustrate the finite sample performance of our methods. The implementation of PECOK can be found at github.com/martinroyer/pecok/ and that of COD at CRAN.R-project.org/package=cord.

7.1. *Simulation design.* Recall our G -latent covariance $\Sigma = ACA' + \Gamma$. Under various scenarios of A and Γ to be described momentarily, we consider the following models for C :

- Model 1: $C = B^T B$ where B is a random $(K - 1) \times K$ matrix with independent entries. Each entry takes the value $+1$ and -1 with equal probability $0.5 \times K^{-1/2}$, and the value 0 with probability $1 - K^{-1/2}$.
- Model 2: $C = C' - 0.001I$ where C' is generated by Model 1.

The matrix C is positive semi-definite in Model 1 and negative definite in Model 2. In the first two simulation scenarios (referred to as M1 and M2 thereafter), we set C derived from Models 1–2 respectively, and specify A to correspond to $K = 10$ equal-size groups of variables (or equivalently $m = p/K$).

In a third scenario (M1S), we specify A such that it corresponds to the existence of 5 singletons, which are variables that form their own groups of size 1, respectively, and the remaining $K - 5$ groups have equal-size, while C is the same as M1.

In these first three scenarios, we employ diagonal $\Gamma = D$ where the p diagonal entries of D are random permutations of $\{0.5, 0.5 + 1.5/(p - 1), \dots, 2\}$.

In the fourth scenario (M1P), we consider the approximate G -block model by setting $\Gamma = D + R$ where $R = 0.1 \cdot U^t U / \max(U^t U)$ and U is a $p \times p$ matrix with iid random entries from a uniform distribution on $[-1, 1]$. We run these four scenarios for two representative $p = 200$ and $p = 1600$, and for sample sizes $n = 100, 300, \dots, 900$ unless noted otherwise. All simulations are repeated 100 times.

The goal of our methods is to create sub-groups of vectors of dimension n , from a given collection of p vectors of observations, each of dimension n . This task can be viewed as that of clustering p objects in \mathbb{R}^n . The existing data clustering algorithms are not tailored to recovering groups with this structure, but they can serve as comparative methods. We thus compare the performance of COD and PECOK with three popular clustering algorithms: K-means, Hierarchical Clustering (HC) and spectral clustering. We apply K-means on the columns of the $n \times p$ matrix of n observations, and use the negative correlation as distance matrix in HC. The spectral clustering algorithm is discussed in Section 5.4, which does not correct for Γ . We use the standard K-means and HC algorithms in R, and we also implement our COD algorithm in R. We also include a variant of the COD algorithm suggested by a reviewer. This variant applies the connected component algorithm (as implemented in CRAN package `igraph` available at <https://CRAN.R-project.org/package=igraph>) to a graph converted from our proposed COD metric matrix thresholded at level α , which will be referred to as COD-cc. The spectral clustering algorithm is based on the widely used Python package `scikit-learn`, and we implement our PECOK algorithm also using Python.

The three competing methods require specification of the number of groups K , and we will use the true $K = 10$ to evaluate their oracle performance. For our proposed methods, we use the CV approach proposed in Section 4.3 to determine K in a data-adaptive fashion. We select either K from a grid in PECOK or the K -related threshold α also from a grid in COD, using the two independent datasets of size n proposed in Section 4.3. The grid for PECOK is $2, 4, \dots, 40$. Since the theoretical choice of α is proportional to $n^{-1/2} \log^{1/2} p$, we use a grid of $\alpha / (n^{-1/2} \log^{1/2} p) = 0.25, 0.5, \dots, 5$ in COD.

7.2. Exact recovery performance and comparison. Figure 1 shows the average percentages of exact recovery across 100 runs by K-means, HC, COD, COD-cc, and PECOK when n varies. Under one setting with a very large $p = 1600$, PECOK did not complete computation within two weeks, and thus its numerical performance for large $p = 1600$ was not reported in the figure. COD and COD-cc clearly outperform all other competing methods (K-means, HC and spectral clustering) when n is about 300 or larger in all the scenarios. K-means and HC, even with the oracle choice of K and large $n = 900$, fail to recover the true groups exactly. COD-cc is better than COD for small $n = 100$, there are almost no finite sample differences between COD and COD-cc for $n = 900$, which is consistent with our theory showing that they share the same rates described before. COD-cc and COD have similar performance across almost all models, and COD-cc achieves close to 100% recovery for smaller samples than COD under the model with singletons (M1S). Except for the model containing singletons (M1S), PECOK has the best performance for small $n = 100$ and $p = 200$, and achieves close to 100% as COD and COD-cc for larger n . Under the singleton scenario M1S, COD-cc has the best performance for all n , while the difference between COD-cc and COD vanishes when $n = 900$. Under this model, the recovery percentages of PECOK increase with n but only reaches about 40% when $n = 900$ and $p = 200$. This is consistent with our theoretical results that PECOK is not expected to work well in the presence of singletons, while COD adapts to this situation. We also note that the competing methods used for comparison are able to recover clusters very close to the truth (measured by the adjusted rand index, or ARI), see Section 7 of the supplemental material [12] for this partial recovery comparison.

7.3. The importance of correcting for Γ in PECOK. The step 1 of our PECOK algorithm is to estimate and correct for Γ . We illustrate the importance of this step by comparing its performance with two closely related methods, Γ -uncorrected PECOK and K-means, neither of which corrects for Γ . The Γ -uncorrected PECOK algorithm simply replaces the estimated $\hat{\Gamma}$ in the step 1 of PECOK by a zero matrix. We use the true K as input to these two methods to assess their oracle performance under the true K , while the parameter K in PECOK is selected by CV as described before. To fix ideas, we use scenario M1 described before and set $p = 200$. Figure 2(a) shows that the exact recovery percentages of K-means are close to zero, and those of Γ -uncorrected PECOK are smaller than 30%, across all n . After correcting for Γ , PECOK yields close to 100% exact recovery when n increases to 300.

7.4. Comparison under varying m . In this section we illustrate the finite sample performance of PECOK when m varies. We use the simulation scenario M1 under $p = 200$, and consider $m = 50$ and $m = 20$, for $n = 60, 80, 100, 150$, to compare the increasing trend in performance before reaching 100% when n reaches 200 as shown in Figure 1. The rest of the simulation set-up is the same as that of the previous section. As predicted by our theory, Figure 2(b) shows the percentages of exact recovery is better for a larger $m = 50$, compared with $m = 20$, for the same n .

Fig 1: Percentages of exact recovery by K-means (K-Oracle, medium green lines, triangle points), HC (H-Oracle, light blue lines, cross points), spectral clustering (S-Oracle, light yellow lines, up-side-down triangle points), COD (dark red lines, circle points), COD-cc (light orange lines, plus points), PECOK (light pink lines, diamond points) across 100 runs of 4 scenarios described in the main text, when $p = 200$ (solid lines) and $p = 1600$ (dashed lines). All standard errors are smaller than 5%.

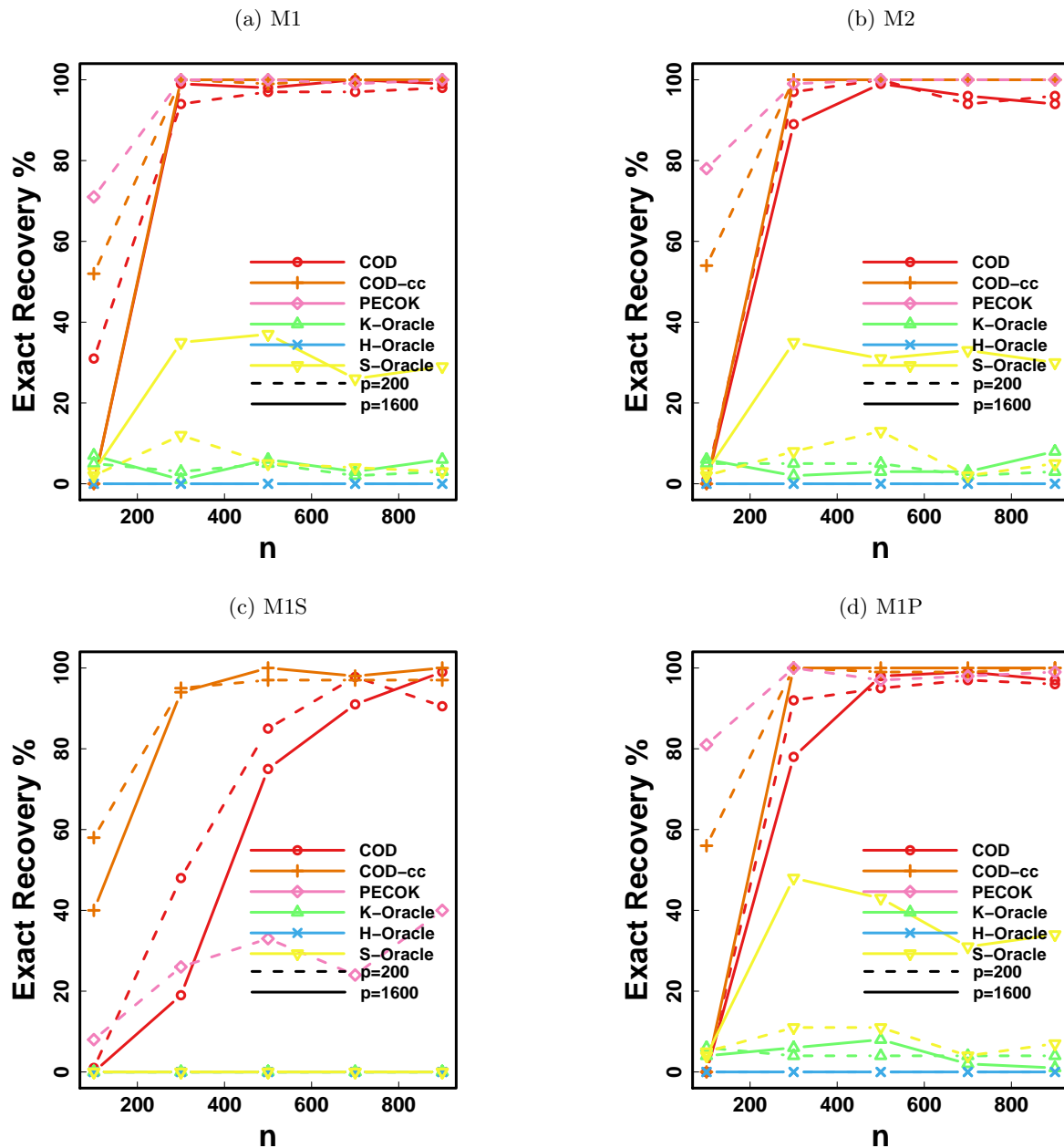
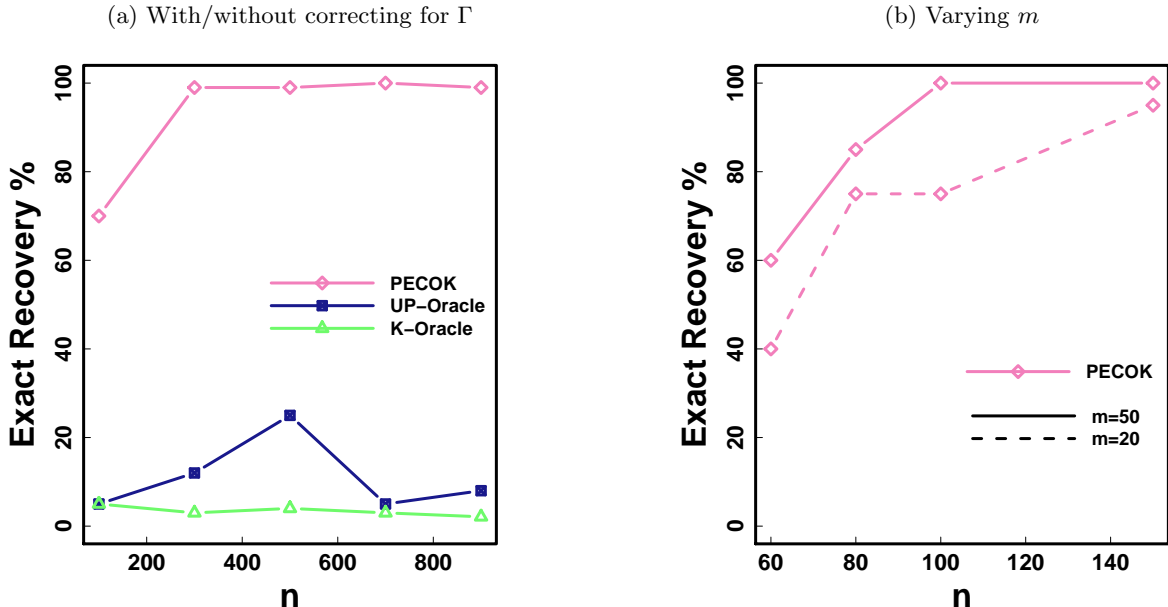


Fig 2: Comparison of Exact recovery percentages across 100 runs. (a) The parameter K in K-means (K-Oracle, medium green lines, triangle points) and Γ -uncorrected PECOK (UP-Oracle, navy blue lines, square points) are set to the true K while PECOK (light pink lines, diamond points) selects K using our CV criterion. (b) The exact recovery percentages of PECOK are shown under $m = 50$ (solid line) and $m = 20$ (dash line). All standard errors are smaller than 5%.



8. Data analysis. Using functional MRI data, [35] found that the human brain putative areas are organized into clusters, sometimes referred to as networks or functional systems. We use a publicly available fMRI dataset to illustrate the clusters recovered by different methods. The dataset was originally published in [37] and is publicly available from Open fMRI (<https://openfmri.org/datasets>) under the accession number ds000007. We will focus on analyzing two scan sessions from subject 1 under a visual-motor stop/go task (task 1). Before performing the analysis, we follow the preprocessing steps suggested by [37], and we follow [35] to subsample the whole brain data using $p = 264$ putative areas, see Section 8 of the supplementary materials [12] for details. This subject was also scanned in two separate sessions, and each session yielded $n = 180$ samples for each putative area.

We apply our CV approach described in Section 4.3 to these two session data. Using the first scan session data only, we first estimate \hat{G} using COD and COD-CC on a fine grid of $\alpha = c\sqrt{\log(p)/n}$ where $c = 0.5, 0.6, \dots, 3$. For a fair comparison, we set K in PECOK to be the same as the resulting K 's found by COD. We then use the second session data to evaluate the CV loss $CV(G)$ given in Section 4.3. Among our methods (COD, COD-CC, and PECOK), COD yields the smallest CV loss when $K = 142$. We thus first focus on illustrating the COD clusters here. Table 2 lists the largest cluster of putative areas recovered by COD and their functional classification based on prior knowledge. Most of these areas are classified to be related to visual, motor, and task functioning, which is consistent with the implication of our experimental task that requires the subject to perform motor responses based on visual stimuli. Figure 3(a) plots the locations of these coordinates on a standard brain template. It shows that our COD cluster appears to come mostly from approximately symmetric locations from the left and right hemisphere, though we do not

enforce this brain function symmetry in our algorithm. Note that the original coordinates in [35] are not sampled with exact symmetry from both hemispheres of the brain, and thus we do not expect exact symmetric locations in the resulting clusters based on these coordinates.

Because there are no gold standards for partitioning the brain, we follow common practice and use a prediction criterion to further compare the clustering performance of different methods. For a fair comparison, we also estimate \hat{G} using K-means, HC, and spectral clustering on the same resulting K 's found by COD. The prediction criterion is as follows. We first compute the covariance matrices \hat{S}_1 and \hat{S}_2 from the first and second session data respectively. For a grouping estimate \hat{G} , we use the following loss to evaluate its performance

$$(35) \quad \left\| \hat{S}_2 - \Upsilon \left(\hat{S}_1, \hat{G} \right) \right\|_F.$$

where block averaging operator $\Upsilon(R, G)$ produces a G -block structured matrix based on \hat{G} . For any $a \in G_k$ and $b \in G_{k'}$, the output matrix entry $[\Upsilon(R, G)]_{ab}$ is given by

$$[\Upsilon(R, G)]_{ab} = \begin{cases} |G_k|^{-1} (|G_k| - 1)^{-1} \sum_{i,j \in G_k, i \neq j} R_{ij} & \text{if } a \neq b \text{ and } k = k' \\ |G_k|^{-1} |G_{k'}|^{-1} \sum_{i \in G_k, j \in G_{k'}} R_{ij} & \text{if } a \neq b \text{ and } k \neq k' \\ 1 & \text{if } a = b. \end{cases}$$

In essence, this operator smooths over the matrix entries with indices in the same group, and one may expect that such smoothing over variables in the true cluster will reduce the loss (35) while smoothing over different clusters will increase the loss.

Figure 3(b) compares the prediction loss values under different group sizes for each method. This shows that our CV approach for COD indeed selects a value $K = 142$ that is immediately next to a slightly larger one ($K = 206$), the latter having the smallest prediction loss, near the bottom plateau. However, the differences are almost negligible. This suggests that our CV criterion, which comes with theoretical guarantees, also provides good prediction performance in this real data example, while selecting a slightly smaller K , as desired, since this makes the resulting clusters easier to describe and interpret.

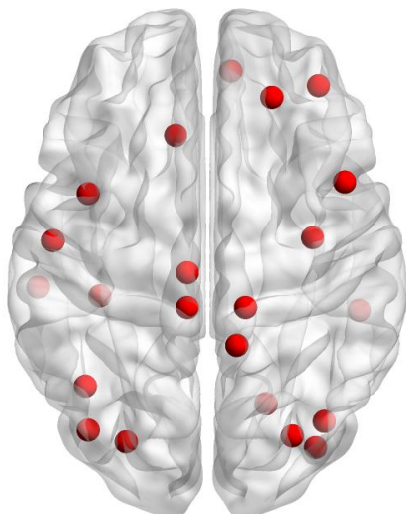
Regardless of the choice of K or α , Figure 3(b) also shows that COD almost always yields the smallest prediction loss for a wide range of K , while PECOK does slightly better when K is between 5 and 10. Though COD-CC has large losses for medium or small K , its performance is very close to the best performer COD near $K = 146$. Kmeans in this example is the closest competing method, while the other two methods (HC and SC) yield larger losses across the choices of K .

9. Discussion. In this section, we discuss some related models and give an overall recommendation on the usage of our methods.

9.1. Comparison with Stochastic Block Model. The problem of variable clustering that we consider in this work is fundamentally different from that of variable clustering from *network data*. The latter, especially in the context of the Stochastic Block Model (SBM), has received a large amount of attention over the past years, for instance [1, 15, 20, 26–28, 31]. The most important difference stems from the nature of the data: the data analyzed via the SBM is a $p \times p$ binary matrix \mathbf{A} , called the adjacency matrix, with entries assumed to have been generated as independent Bernoulli random variables; its expected value is assumed to have a block structure. In contrast, the data matrix \mathbf{X} generated from a G -block covariance is a $n \times p$ matrix with real entries, and rows viewed as i.i.d copies of a p -dimensional vector X with mean zero and dependent entries. The covariance matrix Σ of X is assumed to have (up to the diagonal) a block structure.

Fig 3: (a) Plot of the coordinates of the largest COD cluster overlaid over a standard brain template. The coordinates are shown as red balls. (b) Comparison of COD, COD-cc, PECOK, K-means, HC, and SC using the Frobenius prediction loss criterion (35) where the groups are estimated by these methods respectively.

(a) A brain cluster by COD



(b) Prediction loss

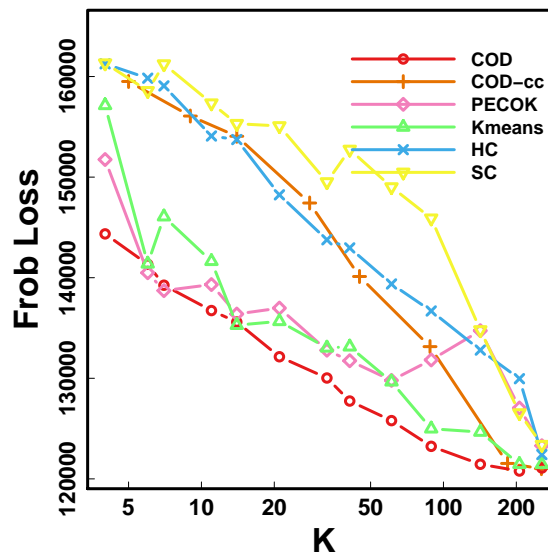


TABLE 2

MNI coordinates (x, y, z , in mm) of the largest COD group and their functioning classification.

X	Y	Z	Function	X	Y	Z	Function
40	-72	14	visual	-7	-21	65	motor
-28	-79	19	visual	-7	-33	72	motor
20	-66	2	visual	13	-33	75	motor
29	-77	25	visual	10	-46	73	motor
37	-81	1	visual	36	-9	14	motor
47	10	33	task	-53	-10	24	motor
-41	6	33	task	-37	-29	-26	uncertain
38	43	15	task	52	-34	-27	uncertain
-41	-75	26	default	-58	-26	-15	uncertain
8	48	-15	default	-42	-60	-9	attention
22	39	39	default	-11	26	25	saliency

Need for a correction. Even though the analysis of the methods in our setting would differ from the SBM setting, we could have applied available clustering procedures tailored for SBMs to the empirical covariance matrix $\widehat{\Sigma} = \mathbf{X}^t\mathbf{X}/n$ by treating it as some sort of weighted adjacency matrix. It turns out that applying verbatim the spectral clustering procedure of Lei and Rinaldo [27] or the SDP such as the ones in [2] would lead to poor results. The main reason for this is that, in our setting, we need to **correct** both the spectral algorithm and the SDP to recover the correct clusters (Section 5). Second, the SDPs studied in the SBM context (such as those of [2]) do not handle properly groups with different and unknown sizes, contrary to our SDP. To the best of our knowledge, our SDP (without correction) has only been independently studied by Mixon et al. [30] in the context of Gaussian mixtures.

Analysis of the SDP. As for the mathematical arguments, our analysis of the SDP in our on covariance-type model differs from that in mean-type models partly because of the the presence of non-trivial cross-product terms. Instead of relying on dual certificates arguments as in other work such as [33], we directly investigate the primal problem and combine different duality-norm bounds. The crucial step is the Lemma 5 which allows to control the Frobenius inner product by a (unusual) combination of ℓ^1 and spectral control. In our opinion, our approach is more transparent than dual certificates techniques, especially in the presence of a correction $\widehat{\Gamma}$ and allows for the attainment of optimal convergence rates.

9.2. Extension to other Models. The general strategy of correcting a convex relaxation of K -means can be applied to other models. In [36], one of the authors has adapted the PECOK algorithm to the clustering problem of mixture of subGaussian distributions. In particular, in the high-dimensional setting where the correction plays a key role, [36] obtains sharper separation conditions dependencies than in state-of-the-art clustering procedures [30]. Extensions to model-based overlapping clustering are beyond the scope of this paper, but we refer to [10] for recent results.

9.3. Practical recommendations. Based on our extensive simulation studies, we conclude this section with general recommendations on the usage of our proposed algorithms.

If p is moderate in size, and if there are reasons to believe that no singletons exist in a particular application, or if they have been removed in a pre-processing step, we recommend the usage of the PECOK algorithm, which is numerically superior to existing methods: exact recovery can be reached for relatively small sample sizes. COD is also very competitive, but requires a slightly larger sample size to reach the same performance as PECOK. The constraint on the size of p reflects the existing computational limits in state-of-the-art algorithms for SDP, not the statistical capabilities of the procedure, the theoretical analysis of which being one of the foci of this work.

If p is large, we recommend COD-type algorithms. Since COD is optimization-free, it scales very well with p , and only requires a moderate sample size to reach exact cluster recovery. Moreover, COD adapts very well to data that contains singletons and, more generally, to data that is expected to have many inhomogeneous clusters.

APPENDIX A: RESULTS FOR THE PECOK ESTIMATOR

In order to avoid notational clutter, we write G for G^* and m for m^* for the entirety of this section.

A.1. The motivation for a K -means correction: proof of Propositions 3 and 4 .

A.1.1. *Proofs of Proposition 3.* The basis of this proof is the following lemma.

LEMMA 3. *The collection \mathcal{C} contains only one matrix whose support is included in $\text{supp}(B^*)$, that is*

$$\mathcal{C} \cap \{B, \text{supp}(B) \subset \text{supp}(B^*)\} = \{B^*\} .$$

PROOF. Consider any matrix $B \in \mathcal{C}$ whose support is included in $\text{supp}(B^*)$. Since $B1 = 1$, it follows that each submatrix $B_{G_k G_k}$ is symmetric doubly stochastic. Since $B_{G_k G_k}$ is also positive semidefinite, we have

$$\text{tr}(B_{G_k G_k}) \geq \|B_{G_k G_k}\|_{op} \geq 1^t B_{G_k G_k} 1 / |G_k| = 1 .$$

As $B \in \mathcal{C}$, we have $\text{tr}(B) = K$, so all the submatrices $B_{G_k G_k}$ have a unit trace. Since $\|B_{G_k G_k}\|_{op} \geq 1$, this also enforces that $B_{G_k G_k}$ contains only one non-zero eigenvalue and that a corresponding eigenvector is the constant vector 1. As a consequence, $B_{G_k G_k} = 11^t / |G_k|$ for all $k = 1, \dots, K$ and $B = B^*$. \square

As a consequence of Lemma 3, in order to prove Proposition 3 we only need to prove that

$$\langle \Sigma, B^* - B \rangle > 0, \text{ for all } B \in \mathcal{C} \text{ (resp. } \mathcal{O} \text{) such that } \text{supp}(B) \not\subseteq \text{supp}(B^*).$$

We have

$$\langle \Sigma, B^* - B \rangle = \langle AC^* A^t, B^* - B \rangle + \langle \Gamma, B^* - B \rangle.$$

Define the p -dimensional vector v by $v = \text{diag}(AC^* A^t)$. Since $B1 = 1$ for all $B \in \mathcal{C}$, we have $\langle v1^t + 1v^t, B^* - B \rangle = 0$. Hence, we have

$$\begin{aligned} \langle AC^* A^t, B^* - B \rangle &= \langle AC^* A^t - \frac{1}{2}(v1^t + 1v^t), B^* - B \rangle \\ &= \sum_{j,k} \sum_{a \in G_j, b \in G_k} \left(C_{jk}^* - \frac{C_{jj}^* + C_{kk}^*}{2} \right) (B_{ab}^* - B_{ab}) \\ &= \sum_{j \neq k} \sum_{a \in G_j, b \in G_k} \left(\frac{C_{jj}^* + C_{kk}^*}{2} - C_{jk}^* \right) B_{ab} \\ (36) \quad &= \sum_{j \neq k} \left(\frac{C_{jj}^* + C_{kk}^*}{2} - C_{jk}^* \right) |B_{G_j G_k}|_1, \end{aligned}$$

where $B_{G_j G_k} = [B_{ab}]_{a \in G_j, b \in G_k}$. Next lemma lower bounds $\langle \Gamma, B^* - B \rangle$ for $B \in \mathcal{O}$. It is stated below and proved at page 30.

LEMMA 4. *For any matrix B belonging to \mathcal{O} and any diagonal matrix Γ ,*

$$(37) \quad \langle \Gamma, B^* - B \rangle \geq -\frac{\|\Gamma\|_V}{m} \sum_{k \neq j} |B_{G_j G_k}|_1 .$$

Hence, combining (36) and Lemma 4, we obtain

$$\langle \Sigma, B^* - B \rangle \geq \sum_{j \neq k} \left(\frac{C_{jj}^* + C_{kk}^*}{2} - C_{jk}^* - \frac{\|\Gamma\|_V}{m} \right) |B_{G_j G_k}|_1 ,$$

for all $B \in \mathcal{O}$. The condition $\Delta(C^*) > \frac{2\|\Gamma\|_V}{m}$ enforces that if $\text{supp}(B) \not\subseteq \text{supp}(B^*)$ then $\langle \Sigma, B^* - B \rangle > 0$. This proves the first claim of Proposition 3.

To show the counterpart of this result that corresponds to replacing \mathcal{O} by \mathcal{C} , we invoke the lemma stated below and proved at page 30.

LEMMA 5. *For any $p \times p$ symmetric matrix S , we have for any $B \in \mathcal{C}$*

$$|\langle S, B^* - B \rangle| \leq 2 \left[\sum_{j \neq k} |B_{G_j G_k}|_1 \right] \left(\frac{\|S\|_{op}}{2m} + 3|B^* S|_\infty \right).$$

Define the diagonal matrix $D = (\max_a \Gamma_{aa} + \min_a \Gamma_{aa})I/2$. Since $\text{tr}(B) = \text{tr}(B^*)$, we have $\langle \Gamma, B^* - B \rangle = \langle \Gamma - D, B^* - B \rangle$. The matrix $S = \Gamma - D$ satisfies $\|S\|_{op} = \|\Gamma\|_V/2$ and $|B^* S|_\infty \leq \|\Gamma\|_V/(2m)$. Applying Lemma 5 to S , we obtain

$$|\langle \Gamma, B^* - B \rangle| \leq \frac{7}{2m} \|\Gamma\|_V \left[\sum_{j \neq k} |B_{G_j G_k}|_1 \right].$$

Hence, together with (36), we obtain

$$\langle \Sigma, B^* - B \rangle \geq \sum_{j \neq k} \left(\frac{C_{jj}^* + C_{kk}^*}{2} - C_{jk}^* - \frac{7\|\Gamma\|_V}{2m} \right) |B_{G_j G_k}|_1,$$

for all $B \in \mathcal{C}$. The condition $\Delta(C^*) > \frac{7\|\Gamma\|_V}{m}$ enforces that if $\text{supp}(B) \not\subseteq \text{supp}(B^*)$ then $\langle \Sigma, B^* - B \rangle > 0$, which proves the second claim of Proposition 3. To complete the proof of Proposition 3 it remains to prove the two lemmas stated above.

PROOF OF LEMMA 4. By definition of B^* and since $\text{tr}(B) = \text{tr}(B^*) = K$, we have

$$\begin{aligned} \langle \Gamma, B^* - B \rangle &= \langle \Gamma - (\max_b \Gamma_{bb})I, B^* - B \rangle = \sum_{a=1}^p (\Gamma_{aa} - (\max_b \Gamma_{bb})) \left[\frac{1}{|G_{k(a)}|} - B_{aa} \right] \\ (38) \quad &\geq \sum_{a=1}^p -\|\Gamma\|_V \left[\frac{1}{|G_{k(a)}|} - B_{aa} \right]_+ \end{aligned}$$

Since B belongs to \mathcal{O} , each row sums to one and each B_{ab} is either equal to 0 or to B_{aa} . Thus,

$$\sum_{b \notin G_{k(a)}} B_{ab} = 1 - \sum_{b \in G_{k(a)}} B_{ab} \geq [1 - |G_{k(a)}| B_{aa}]_+$$

which implies

$$\left[\frac{1}{|G_{k(a)}|} - B_{aa} \right]_+ \leq \frac{1}{m} \sum_{b \notin G_{k(a)}} B_{ab}$$

Coming back to (38), this gives us $\langle \Gamma, B^* - B \rangle \geq -\frac{\|\Gamma\|_V}{m} \sum_{k \neq j} |B_{G_j G_k}|_1$. \square

PROOF OF LEMMA 5. Observe first that B^* is a projection matrix that induces the following decomposition of S .

$$S = B^* S + S B^* - B^* S B^* + (I - B^*) S (I - B^*).$$

By the definition of the inner product, followed by the triangle inequality, and since $(I - B^*)B^* = 0$, we further have

$$(39) \quad \begin{aligned} |\langle S, B^* - B \rangle| &\leq 3|B^*S|_\infty |B^*(B^* - B)|_1 + |\langle (I - B^*)S(I - B^*), B^* - B \rangle| \\ &= 3|B^*S|_\infty |B^*(B^* - B)|_1 + |\langle S, (I - B^*)B(I - B^*) \rangle|. \end{aligned}$$

Relying on the duality of the nuclear $\|\cdot\|_*$ and operator $\|\cdot\|_{op}$ norms, we have

$$|\langle S, (I - B^*)B(I - B^*) \rangle| \leq \|S\|_{op} \|(I - B^*)B(I - B^*)\|_*.$$

We begin by bounding the nuclear norm $\|(I - B^*)B(I - B^*)\|_*$. Since $(I - B^*)B(I - B^*) \in \mathcal{S}^+$, we have

$$\|(I - B^*)B(I - B^*)\|_* = \text{tr}((I - B^*)B(I - B^*)) = \langle I - B^*, B(I - B^*) \rangle = \langle I - B^*, B \rangle.$$

Using the fact that the sum of each row of B is 1 and $\text{tr}(B) = K$, we have

$$(40) \quad \begin{aligned} \|(I - B^*)B(I - B^*)\|_* = \langle I - B^*, B \rangle &= \text{tr}(B) - \sum_{k=1}^K \sum_{a,b \in G_k} \frac{B_{ab}}{|G_k|} \\ &= K - K + \sum_{k \neq j} \sum_{a \in G_k, b \in G_j} \frac{B_{ab}}{|G_k|} \\ &\leq \frac{1}{m} \sum_{k \neq j} |B_{G_j G_k}|_1. \end{aligned}$$

Next, we simplify the expression of $|B^*(B^* - B)|_1 = |B^*(I - B)|_1$.

$$\begin{aligned} |B^*(I - B)|_1 &= \sum_{j \neq k} \sum_{a \in G_j, b \in G_k} |(B^*B)_{ab}| + \sum_{k=1}^K \sum_{a,b \in G_k} |[B^*(I - B)]_{ab}| \\ &= \sum_{j \neq k} \sum_{a \in G_j, b \in G_k} \frac{1}{|G_j|} \sum_{c \in G_j} B_{cb} + \sum_{k=1}^K \sum_{a,b \in G_k} \frac{1}{|G_k|} \left| 1 - \sum_{c \in G_k} B_{cb} \right| \\ &= 2 \sum_{j \neq k} |B_{G_j G_k}|_1, \end{aligned}$$

where we used again $B1 = 1$ and that the entries of B are nonnegative. Gathering the above bounds together with (39) yields the desired result. This completes the proof of this result and of Proposition 3. \square

A.1.2. Proof of Proposition 4. By symmetry, we can assume that the true partition matrix B^* is diagonal block constant. Define the partition matrix $B_1 := \begin{bmatrix} 2/m & 0 & 0 \\ 0 & 2/m & 0 \\ 0 & 0 & 1/(2m) \end{bmatrix}$ where the first two blocks are of size $m/2$ and the the last block has size $2m$. The construction of the matrix B_1 amounts to merging groups G_2 and G_3 , and to splitting G_1 into two groups of equal size. Then,

$$\langle \Sigma, B^* \rangle = \gamma_+ + 2\gamma_- + m \text{tr}(C^*), \quad \langle \Sigma, B_1 \rangle = 2\gamma_+ + \gamma_- + m \text{tr}(C^*) - m\tau.$$

As a consequence, $\langle \Sigma, B_1 \rangle < \langle \Sigma, B^* \rangle$ if and only if $\tau > \frac{\gamma_+ - \gamma_-}{m}$.

A.2. Analysis of the population version under the approximate model: proofs of Proposition 7 and Corollary 3. In this subsection, we prove Proposition 7 and Corollary 3. As a consequence of Lemma 3 above, we only need to prove that,

$$\langle \Sigma, B^* - B \rangle > 0, \quad \text{for all } B \in \mathcal{C} \text{ such that } \text{supp}(B) \not\subseteq \text{supp}(B^*).$$

We have

$$\langle \Sigma, B^* - B \rangle = \langle AC^*A^t, B^* - B \rangle + \langle D, B^* - B \rangle + \langle R, B^* - B \rangle.$$

Define the p -dimensional vector v by $v = \text{diag}(AC^*A^t)$. Since $B1 = 1$ for all $B \in \mathcal{C}$, we have $\langle v1^t + 1v^t, B^* - B \rangle = 0$. Hence, we have

$$\begin{aligned} \langle AC^*A^t, B^* - B \rangle &= \langle AC^*A^t - \frac{1}{2}(v1^t + 1v^t), B^* - B \rangle \\ &= \sum_{j,k} \sum_{a \in G_j, b \in G_k} \left(C_{jk}^* - \frac{C_{jj}^* + C_{kk}^*}{2} \right) (B_{ab}^* - B_{ab}) \\ &= \sum_{j \neq k} \sum_{a \in G_j, b \in G_k} \left(\frac{C_{jj}^* + C_{kk}^*}{2} - C_{jk}^* \right) B_{ab} \\ (41) \quad &= \sum_{j \neq k} \left(\frac{C_{jj}^* + C_{kk}^*}{2} - C_{jk}^* \right) |B_{G_j G_k}|_1, \end{aligned}$$

where $B_{G_j G_k} = [B_{ab}]_{a \in G_j, b \in G_k}$. From Lemma 5 we get

$$|\langle D, B^* - B \rangle| \leq \frac{7}{2m} |D|_V \sum_{j \neq k} |B_{G_j G_k}|_1$$

and

$$|\langle R, B^* - B \rangle| \leq \left(\frac{3}{2} |R|_\infty + \frac{\|R\|_{op}}{m} \right) \sum_{j \neq k} |B_{G_j G_k}|_1.$$

Combining the two last inequalities with (41) gives the Proposition. The Corollary follows.

A.3. Exact recovery with PECOK: approximate model. Proofs of Theorems 7, 4 and 5. The conclusion of Theorems 7 and 4 follows by combining the conclusion of Theorems 8 and 9, stated below and proved in the next subsection, with the conclusion of Proposition 8 stated and proved in Appendix A.4 below.

Specifically: Theorem 8, specialized to $R = 0$ and coupled with (i) of Proposition 8 proves Theorem 4. Theorem 9, specialized to $R = 0$ and coupled with (ii) of Proposition 8 proves Theorem 5. Finally, Theorem 8 coupled with (i) of Proposition 8 proves Theorem 7. If we combine Theorem 8 with (ii) of Proposition 8 we obtained a version of Theorem 7 for bounded variables, which we do not state, for space limitations.

The following theorem examines the behavior of PECOK under the general model (4).

THEOREM 8. *There exist c_1, \dots, c_3 three positive constants such that the following holds. Let $\widehat{\Gamma}$ be any estimator of D , such that $|\widehat{\Gamma} - D|_V \leq \delta_{n,p}$ with probability $1 - c_1/(2p)$. Then, under Assumption 1, and when $L^4 \log(p) \leq c_3 n$ and*

$$(42) \quad \Delta(C^*) \geq c_L \left[\|\Gamma\|_{op} \left\{ \sqrt{\frac{\log p}{mn}} + \sqrt{\frac{p}{nm^2}} + \frac{\log(p)}{n} + \frac{p}{nm} \right\} + \frac{\delta_{n,p} + \|R\|_{op}}{m} + |R|_\infty \right],$$

we have $\widehat{B} = B^$, with probability higher than $1 - c_1/p$.*

The following theorem examines the behavior of PECOK under the general model (4) and when the variables are bounded.

THEOREM 9. *There exist c_1, c_2 two positive constants such that the following holds. Let $\widehat{\Gamma}$ be any estimator of D , such that $|\widehat{\Gamma} - D|_V \leq \delta_{n,p}$ with probability $1 - c_1/(2p)$. Then, under Assumption 1-bis, and when*

$$(43) \quad \Delta(C^*) \geq c_2 \left[M \|\Gamma\|_{op}^{1/2} \sqrt{\frac{p \log(p)}{nm^2}} + M^2 \frac{p \log(p)}{nm} + \frac{\delta_{n,p} + \|R\|_{op}}{m} + |R|_\infty \right],$$

we have $\widehat{B} = B^*$, with probability higher than $1 - c_1/p$.

A.3.1. Proofs of Theorems 8 and 9. In contrast to other SDP analyses performed for other models [30], our proof does not rely on dual certificates techniques. Instead of that, we directly investigate the primal problem and combine different duality-norm bounds. In our opinion, this makes the arguments more transparent. The two key ingredients are Lemmas 3 and 5 above.

Given $k, l \in [K]$, we define $\Delta_{kl}(C^*) = C_{kk}^* + C_{ll}^* - 2C_{kl}^*$. As a consequence of Lemma 3 page 29, we only need to prove that

$$(44) \quad \langle \widehat{\Sigma} - \widehat{\Gamma}, B^* - B \rangle > 0, \quad \text{for all } B \in \mathcal{C} \text{ such that } \text{supp}(B) \not\subseteq \text{supp}(B^*),$$

with high probability.

We begin by introducing some notation. For any $k, l \in [K]$, we denote $m_k = |G_k|$ the size of group G_k and

$$(45) \quad \gamma_{kl} = \frac{1}{m_k m_l} \sum_{a \in G_k, b \in G_l} \Gamma_{ab}.$$

Recall that \mathbf{X} denotes the $n \times p$ matrix of observations and we set $\underline{\mathbf{Z}} = \mathbf{X}A^t(A^tA)^{-1}$. We have the decomposition

$$\mathbf{X} = \mathbf{X}B^* + \mathbf{X}(I - B^*) =: \underline{\mathbf{Z}}A^t + \underline{\mathbf{E}}$$

with $\text{Cov}(A\underline{\mathbf{Z}}, \underline{\mathbf{E}}) = B^*\Gamma(I - B^*)$, $\text{Cov}(\underline{\mathbf{E}}) = (I - B^*)\Gamma(I - B^*)$, and $\text{Cov}(\underline{\mathbf{Z}}_k, \underline{\mathbf{Z}}_l) = C_{kl}^* + \gamma_{kl}$. Note that in the latent model $X_a = Z_{k(a)} + E_a$, the random variables $\underline{\mathbf{Z}}_k$ and $\underline{\mathbf{E}}_a$ differ from Z_k and E_a .

Our first goal is to decompose $\widehat{\Sigma} - \widehat{\Gamma}$ in such a way that the distance $|\underline{\mathbf{Z}}_{:k} - \underline{\mathbf{Z}}_{:j}|_2^2$ becomes evident. To this end, recall that $n\widehat{\Sigma} = \mathbf{X}^t\mathbf{X}$ and let us define $\widetilde{\Gamma} = \frac{1}{n}\underline{\mathbf{E}}^t\underline{\mathbf{E}}$. Hence, we have

$$n\widehat{\Sigma} = A\underline{\mathbf{Z}}^t\underline{\mathbf{Z}}A^t + n\widetilde{\Gamma} + A(\underline{\mathbf{Z}}^t\underline{\mathbf{E}}) + (\underline{\mathbf{E}}^t\underline{\mathbf{Z}})A^t.$$

Using the fact that for any vectors v_1 and v_2 we have $|v_1 - v_2|_2^2 = |v_1|_2^2 + |v_2|_2^2 - 2v_1^t v_2$, we can write

$$[A\underline{\mathbf{Z}}^t\underline{\mathbf{Z}}A^t]_{ab} = \frac{1}{2}|[A\underline{\mathbf{Z}}^t]_{a:}|_2^2 + \frac{1}{2}|[A\underline{\mathbf{Z}}^t]_{b:}|_2^2 - \frac{1}{2}|[A\underline{\mathbf{Z}}^t]_{a:} - [A\underline{\mathbf{Z}}^t]_{b:}|_2^2,$$

for any $1 \leq a, b \leq p$. We also observe that

$$[A(\underline{\mathbf{Z}}^t\underline{\mathbf{E}}) + (\underline{\mathbf{E}}^t\underline{\mathbf{Z}})A^t]_{ab} = [(A\underline{\mathbf{Z}}^t)_{a:} - (A\underline{\mathbf{Z}}^t)_{b:}][\underline{\mathbf{E}}_{b:} - \underline{\mathbf{E}}_{a:}] + [A\underline{\mathbf{Z}}^t\underline{\mathbf{E}}]_{aa} + [A\underline{\mathbf{Z}}^t\underline{\mathbf{E}}]_{bb}.$$

Define the $p \times p$ matrix W by

$$(46) \quad W_{ab} := n(\widehat{\Sigma}_{ab} - \widehat{\Gamma}_{ab}) - \frac{1}{2} |[\mathbf{AZ}^t]_{a:}|_2^2 - \frac{1}{2} |[\mathbf{AZ}^t]_{b:}|_2^2 - [\mathbf{AZ}^t \mathbf{E}]_{aa} - [\mathbf{AZ}^t \mathbf{E}]_{bb}.$$

Combining the four displays above we have

$$(47) \quad W = W_1 + W_2 + W_3 + n(\Gamma - \widehat{\Gamma}),$$

with

$$(48) \quad (W_1)_{ab} := -\frac{1}{2} |[\mathbf{AZ}^t]_{a:}|_2^2 - nB^* \Gamma B^*,$$

$$(49) \quad (W_2)_{ab} := [(\mathbf{AZ}^t)_{a:} - (\mathbf{AZ}^t)_{b:}] [\mathbf{E}_{b:} - \mathbf{E}_{a:}] - n[B^* \Gamma (I - B^*) + (I - B^*) \Gamma B^*]_{ab},$$

and

$$(50) \quad W_3 = n\widetilde{\Gamma} - n(I - B^*) \Gamma (I - B^*),$$

for any $1 \leq a, b \leq p$. Observe from (46) that $W - n(\widehat{\Sigma} - \widehat{\Gamma})$ is a sum of four matrices, two of which are of the type $1v_1^t$, and two of the type $v_2 1^t$, for some vectors $v_1, v_2 \in \mathcal{R}^p$. Since for any two matrices B_1 and B_2 in \mathcal{C} , we have $B_1 1 = B_2 1 = 1$, it follows that

$$\langle W - n(\widehat{\Sigma} - \widehat{\Gamma}), B_1 - B_2 \rangle = 0.$$

As a consequence and using the decomposition (47), proving (44) reduces to proving

$$(51) \quad \langle W_1 + W_2 + W_3 + n(\Gamma - \widehat{\Gamma}), B^* - B \rangle > 0, \quad \text{for all } B \in \mathcal{C} \text{ such that } \text{supp}(B) \not\subseteq \text{supp}(B^*).$$

We will analyze the inner product between $B^* - B$ and each of the four matrices in (51) separately in the following lemmas. Their proofs are given after the proof of this theorem.

The matrix W_1 contains the information about the clusters, as we explain below. Note that for two variables a and b belonging to the same group G_k , $(W_1)_{ab} = -n\gamma_{kk}$. As a consequence, $\langle W_1, B^* \rangle = -n \sum_k m_k \gamma_{kk}$. For two variables a and b belonging to two different groups G_j and G_k , $(W_1)_{ab} = -|\mathbf{z}_{:j} - \mathbf{z}_{:k}|_2^2 / 2 - n\gamma_{jk}$. In the sequel, we denote by B_{G_j, G_k} the submatrix $(B_{ab})_{a \in G_j, b \in G_k}$. Since all the entries of B are nonnegative, and $B1 = 1$,

$$\begin{aligned} -\langle W_1, B \rangle &= \frac{1}{2} \sum_{j \neq k} |\mathbf{z}_{:j} - \mathbf{z}_{:k}|_2^2 |B_{G_j G_k}|_1 + \sum_{j, k} n\gamma_{jk} |B_{G_j G_k}|_1, \\ &= \frac{1}{2} \sum_{j \neq k} [|\mathbf{z}_{:j} - \mathbf{z}_{:k}|_2^2 + 2n\gamma_{jk}] |B_{G_j G_k}|_1 + n \sum_k \gamma_{kk} \left(m_k - \sum_{j: j \neq k} |B_{G_j G_k}|_1 \right) \\ &= \frac{1}{2} \sum_{j \neq k} (|\mathbf{z}_{:j} - \mathbf{z}_{:k}|_2^2 - n\gamma_{kk} - n\gamma_{jj} + 2n\gamma_{jk}) |B_{G_j G_k}|_1 + n \sum_k m_k \gamma_{kk}. \end{aligned}$$

Hence, we obtain

$$(52) \quad \langle W_1, B^* - B \rangle = \frac{1}{2} \sum_{j \neq k} \left[|\mathbf{z}_{:j} - \mathbf{z}_{:k}|_2^2 - n\gamma_{jj} - n\gamma_{kk} + 2n\gamma_{jk} \right] |B_{G_j G_k}|_1,$$

Each of the random variables $|\mathbf{z}_{:j} - \mathbf{z}_{:k}|_2^2$ is a quadratic form of independent random variables. As a consequence, we can apply Hanson-Wright inequalities, of the type stated in Lemma 10 of the supplement [12] to simultaneously control all these quantities. This leads us to Lemmas 6 - 9, proved in the supplement [12].

LEMMA 6. *Under either Assumption 1 and Condition (42) or Assumption 1-bis and Condition (43), it holds with probability higher than $1 - 1/p$, that*

$$(53) \quad \langle W_1, B^* - B \rangle \geq \sum_{j \neq k} n \frac{\Delta_{jk}(C^*)}{4} |B_{G_j G_k}|_1,$$

simultaneously for all matrices $B \in \mathcal{C}$.

We will analyze below the three remaining cross products.

LEMMA 7. *Under Assumption 1, there exists an event of probability larger than $1 - 2/p$ such that the following holds simultaneously for all $B \in \mathcal{C}$*

$$(54) \quad |\langle W_2, B^* - B \rangle| \leq c_1 L^2 \sum_{j \neq k} \left[\sqrt{\Delta_{jk}(C^*)} |\Gamma|_\infty + \frac{|D|_\infty}{\sqrt{m}} + |R|_\infty + |R|_\infty^{1/2} |D|_\infty^{1/2} \right] \left[\sqrt{n \log(p)} + \log(p) \right] |B_{G_j G_k}|_1,$$

Under Assumption 1-bis, there exists an event of probability larger than $1 - 2/p$ such that the following holds simultaneously for all $B \in \mathcal{C}$

$$(55) \quad |\langle W_2, B^* - B \rangle| \leq c'_1 M \sum_{j \neq k} \left[\sqrt{n \log(p) \left[\Delta_{k(a)k(b)}(C^*) + \frac{|D|_\infty}{m} + |R|_\infty \right]} + M \log(p) \right] |B_{G_j G_k}|_1.$$

It remains to control the term W_3 corresponding to the empirical covariance matrix of the noise $\underline{\mathbf{E}}$. This is the main technical difficulty in this proof.

LEMMA 8. *Under Assumption 1, it holds with probability higher than $1 - 1/p$ that*

$$(56) \quad |\langle W_3, B^* - B \rangle| \leq c_L \|\Gamma\|_{op} \left(\sqrt{\frac{np}{m^2}} + \frac{p}{m} \right) \sum_{j \neq k} |B_{G_j G_k}|_1,$$

simultaneously over all matrices $B \in \mathcal{C}$. Here, c_L is a constant that only depends on $L > 0$.

Under Assumption 1-bis, it holds with probability higher than $1 - 1/p$ that

$$(57) \quad |\langle W_3, B^* - B \rangle| \leq c_2 M \left[\sqrt{\frac{np \|\Gamma\|_{op} \log(p)}{m^2}} + \frac{pM \log(p)}{m} \right] \sum_{j \neq k} |B_{G_j G_k}|_1,$$

simultaneously over all matrices $B \in \mathcal{C}$.

Finally, we control the last term $\langle n(\Gamma - \hat{\Gamma}), B^* - B \rangle$ with the next lemma.

LEMMA 9. *It holds that*

$$(58) \quad |\langle \Gamma - \hat{\Gamma}, B^* - B \rangle| \leq c_3 \left[\frac{|D - \hat{\Gamma}|_V + \|R\|_{op}}{m} + |R|_\infty \right] \sum_{j \neq k} |B_{G_j G_k}|_1$$

simultaneously over all matrices $B \in \mathcal{C}$.

End of the proof of Theorem 8. Under Assumption 1, we combine (53), (54), (56) and (58) and the assumption $L^4 \log(p) \leq c_4 n$ and obtain that, with probability larger than $1 - c/p$,

$$(59) \quad \begin{aligned} \frac{1}{n} \langle W, B^* - B \rangle &\geq \sum_{j \neq k} \left[\frac{\Delta_{jk}(C^*)}{4} - c_1 L^2 \sqrt{\Delta_{jk}(C^*) |\Gamma|_\infty \frac{\log(p)}{n}} \right. \\ &\quad \left. - c_2 L^2 \left[\frac{|D|_\infty}{\sqrt{m}} + |R|_\infty + |R|_\infty^{1/2} |D|_\infty^{1/2} \right] \sqrt{\frac{\log(p)}{n}} - c_L \|\Gamma\|_{op} \left[\sqrt{\frac{p}{nm^2}} + \frac{p}{nm} \right] \right. \\ &\quad \left. - c_3 \left[\frac{|D - \hat{\Gamma}|_V + \|R\|_{op}}{m} + |R|_\infty \right] |B_{G_j G_k}|_1 \right], \end{aligned}$$

simultaneously for all $B \in \mathcal{C}$. Condition (42) enforces that, for each (j, k) , the term in the bracket of (59) is positive. Hence, with probability at least $1 - c_1/p$, the Inequality (51) holds since any matrix $B \in \mathcal{C}$ whose support is not included in $\text{supp}(B^*)$ satisfies $|B_{G_j G_k}|_1 > 0$ for some $j \neq k$. ■

End of the proof of Theorem 9. Let us now assume that Assumption 1-bis holds. Combining (52), (55), (57) and (58) we obtain that, with probability larger than $1 - c/p$,

$$(60) \quad \begin{aligned} \frac{1}{n} \langle W, B^* - B \rangle &\geq \sum_{j \neq k} \left[\frac{\Delta_{jk}(C^*)}{4} - c_1 M \sqrt{\Delta_{jk}(C^*) \frac{\log(p)}{n}} \right. \\ &\quad \left. - c_2 M \sqrt{\frac{|D|_\infty}{m} + |R|_\infty} \sqrt{\frac{\log(p)}{n}} - c_3 M^2 \frac{p \log(p)}{nm} - c_4 M \|\Gamma\|_{op}^{1/2} \sqrt{\frac{p \log(p)}{nm^2}} \right. \\ &\quad \left. - c_5 \left[\frac{|D - \hat{\Gamma}|_V + \|R\|_{op}}{m} + |R|_\infty \right] |B_{G_j G_k}|_1 \right], \end{aligned}$$

simultaneously for all $B \in \mathcal{C}$. Condition (43) enforces that, for each (j, k) , the term in the bracket of (60) is positive and as previously that (51) holds with probability at least $1 - c_3/p$. ■

A.4. Guarantees for the estimator (22) of Γ . Proposition 8 controls the estimation error of estimator $\hat{\Gamma}$ defined by (22) under both the exact model and the approximate block G -block model (4). We set $v^2 = \min_{c \neq d} \text{Var}(X_c - X_d)$.

PROPOSITION 8. *Assume that Γ either*

- (a) *is diagonal;*
- (b) *or fulfills the diagonal dominance assumption (32).*

Assume also that $\Delta(C^) \geq 0$. Then, the two following results holds.*

(i) *Under Assumption 1, there exist three numerical constants $c_1 - c_3$ such that when $m \geq 3$ and $L^4 \log(p) \leq c_1 n$, with probability larger than $1 - c_3/p$, the estimator $\hat{\Gamma}$ defined by (22) satisfies*

$$(61) \quad |\hat{\Gamma} - \Gamma|_V \leq 2|\hat{\Gamma} - \Gamma|_\infty \leq c_2 \left(\sqrt{|R|_\infty |\Gamma|_\infty} + |\Gamma|_\infty L^2 \sqrt{\frac{\log(p)}{n}} \right).$$

(ii) *Under Assumption 1-bis, there exist three numerical constants $c_1 - c_3$ such that when $m \geq 3$ and $\log(p) \leq c_1 (v/M)^2 n$, with probability larger than $1 - c_3/p$, the estimator $\hat{\Gamma}$ defined by (22) satisfies*

$$(62) \quad |\hat{\Gamma} - \Gamma|_V \leq 2|\hat{\Gamma} - \Gamma|_\infty \leq c_2 \left(\sqrt{|R|_\infty |\Gamma|_\infty} + M \sqrt{\frac{|\Gamma|_\infty \log(p)}{n}} + M^2 \frac{\log(p)}{n} \right).$$

PROOF OF PROPOSITION 8. To ease the presentation of this proof, we introduce the new notation

$$ne_1(a) := \operatorname{argmin}_{b \in [p] \setminus \{a\}} V(a, b) \quad \text{and} \quad ne_2(a) := \operatorname{argmin}_{b \in [p] \setminus \{a, ne_1(a)\}} V(a, b).$$

(i) We start with the first part of the proposition. Let a, b_1, b_2 be three different indices. Under Assumption 1, the Corollary 1 of Hanson-Wright inequality given in the supplemental material [12] gives that with probability at least $1 - p^{-4}$

$$\begin{aligned} \left| \frac{1}{n} \langle \mathbf{X}_{:a} - \mathbf{X}_{:b_1}, \mathbf{X}_{:a} - \mathbf{X}_{:b_2} \rangle - \operatorname{Cov}(X_a - X_{b_1}, X_a - X_{b_2}) \right| \\ \leq cL^2 \sqrt{\operatorname{Var}(X_a - X_{b_1}) \operatorname{Var}(X_a - X_{b_2})} \left(\sqrt{\frac{\log(p)}{n}} + \frac{\log(p)}{n} \right). \end{aligned}$$

Applying the inequality $2ab \leq a^2 + b^2$, and a union bound, we obtain that the inequalities

$$\begin{aligned} \left| \frac{1}{n} \langle \mathbf{X}_{:a} - \mathbf{X}_{:b_1}, \mathbf{X}_{:a} - \mathbf{X}_{:b_2} \rangle - \operatorname{Cov}(X_a - X_{b_1}, X_a - X_{b_2}) \right| \\ \leq c'L^2 [\operatorname{Var}(X_a - X_{b_1}) + \operatorname{Var}(X_a - X_{b_2})] \left(\sqrt{\frac{\log(p)}{n}} + \frac{\log(p)}{n} \right) \end{aligned}$$

hold simultaneously over all triplets of different indices a, b_1, b_2 , with probability $1 - 1/p$. Decomposing these variance and covariance terms, we obtain

$$\begin{aligned} \operatorname{Cov}(X_a - X_{b_1}, X_a - X_{b_2}) \\ = D_{aa} + \frac{1}{2} (\Delta_{k(a)k(b_1)}(C^*) + \Delta_{k(a)k(b_2)}(C^*) - \Delta_{k(b_2)k(b_1)}(C^*)) + R_{b_1 b_2} - R_{ab_1} - R_{ab_2}, \end{aligned}$$

and $\operatorname{Var}(X_a - X_b) = D_{aa} + D_{bb} + \Delta_{k(a)k(b)}(C^*) - 2R_{ab}$. Hence

$$(63) \quad \operatorname{Var}(X_a - X_{b_i}) \leq 2|D|_\infty + 2|R|_\infty + |\Delta_{k(a)k(b_i)}(C^*)|,$$

(64)

$$\left| \operatorname{Cov}(X_a - X_{b_1}, X_a - X_{b_2}) - D_{aa} \right| \leq 3|R|_\infty + \frac{|\Delta_{k(a)k(b_1)}(C^*)| + |\Delta_{k(a)k(b_2)}(C^*)| + |\Delta_{k(b_1)k(b_2)}(C^*)|}{2}.$$

For $i = 1, 2$, write $t_i := |\Delta_{k(a)k(ne_i(a))}(C^*)|$ and $t_{12} = |\Delta_{k(ne_1(a))k(ne_2(a))}(C^*)|$. Since $\log(p) \leq c_1 L^{-4} n$, the previous inequalities entail that

$$(65) \quad \left| \widehat{\Gamma}_{aa} - D_{aa} \right| \leq c \left(|R|_\infty + (t_1 + t_2 + t_{12}) + L^2 |\Gamma|_\infty \sqrt{\frac{\log(p)}{n}} \right)$$

with probability at least $1 - 1/p$. As a consequence, we only have to prove that t_1, t_2 and t_{12} are smaller than $c \left(\sqrt{|R|_\infty |\Gamma|_\infty} + L^2 |\Gamma|_\infty \sqrt{\log(p)/n} \right)$ with probability at least $1 - c/p$. We focus on t_1 , the arguments for t_2 and t_{12} being similar.

First note that $t_1 = 0$ if $k(a) = k(ne_1(a))$ so that we may assume henceforth that $k(a) \neq k(ne_1(a))$. We have the following.

LEMMA 10. *Assume that Γ either*

- (a) is diagonal;
(b) or fulfills the diagonal dominance assumption (32).

Assume also that $\Delta(C^*) \geq 0$. Then, there exists a numerical constant c_1 , such that, outside an event of probability less than $1/p^2$, we have

(i) under Assumption 1,

$$(66) \quad |\langle \mathbf{X}_{:a} - \mathbf{X}_{:ne_1(a)}, \mathbf{X}_{:c} - \mathbf{X}_{:d} \rangle| \leq c_1 \left(\sqrt{n|R|_\infty} + L^2|\Gamma|_\infty^{1/2} \sqrt{\log(p)} \right) |\mathbf{X}_{:c} - \mathbf{X}_{:d}|_2,$$

simultaneously over all $c, d \neq (a, ne_1(a))$;

(ii) under Assumption 1-bis,

$$(67) \quad |\langle \mathbf{X}_{:a} - \mathbf{X}_{:ne_1(a)}, \mathbf{X}_{:c} - \mathbf{X}_{:d} \rangle| \leq c_1 \left(\sqrt{n|R|_\infty} + M\sqrt{\log(p)} \right) |\mathbf{X}_{:c} - \mathbf{X}_{:d}|_2,$$

simultaneously over all $c, d \neq (a, ne_1(a))$.

Similar bounds also hold for $ne_2(a)$ instead of $ne_1(a)$.

This lemma is proved in the supplement [12]. Below, c' denotes a numerical constant, whose value may vary from line to line.

For any c and d , the variance of $X_c - X_d$ is less than $\Delta_{k(c)k(d)}(C^*) + 2|D|_\infty + 2|R|_\infty$. As a consequence, Hanson-Wright inequality together with an union bound over all $c, d \in [p]$ and the condition $L^4 \log(p) \leq c_1 n$ leads to

$$|\mathbf{X}_{:c} - \mathbf{X}_{:d}|_2 \leq c' \sqrt{n[4|\Gamma|_\infty + \Delta_{k(c)k(d)}(C^*)]},$$

simultaneously over all $c \neq d$, with probability $1 - 1/p^2$. Take c and d any two indices such that $k(a) = k(c)$ and $k(ne_1(a)) = k(d)$. So combined with Lemma 10, we get with probability at least $1 - 2p^{-2}$

$$(68) \quad |\langle \mathbf{X}_{:a} - \mathbf{X}_{:ne_1(a)}, \mathbf{X}_{:c} - \mathbf{X}_{:d} \rangle| \leq c' \left(\sqrt{n|R|_\infty} + L^2|\Gamma|_\infty^{1/2} \sqrt{\log(p)} \right) \sqrt{n[|\Gamma|_\infty + t_1]}.$$

Let us now lower bound the left hand-side of the above inequality. For any c in the same group as a and b in the same group as d , we have $\mathbb{E}[\langle X_a - X_b, X_c - X_d \rangle] = \Delta_{k(b)k(a)}(C^*) + R_{ac} + R_{bd} - R_{bc} - R_{ab}$. Therefore, Corollary 1 of the Hanson-Wright inequality given in the supplement [12] yields, with probability at least $1 - p^4$,

$$|\langle \mathbf{X}_{:a} - \mathbf{X}_{:b}, \mathbf{X}_{:c} - \mathbf{X}_{:d} \rangle| \geq n\Delta_{k(b)k(a)}(C^*) - 4n|R|_\infty - c'L^2\sqrt{n \log(p)}[\Delta_{k(b)k(a)}(C^*) + |\Gamma|_\infty].$$

As a consequence, for any c and d such that $k(a) = k(c)$ and $k(ne_1(a)) = k(d)$, we get from $L^4 \log(p) \leq c_1 n$ and a union bound

$$|\langle \mathbf{X}_{:a} - \mathbf{X}_{:ne_1(a)}, \mathbf{X}_{:c} - \mathbf{X}_{:d} \rangle| \geq nt_1/2 - 4n|R|_\infty - c'L^2|\Gamma|_\infty \sqrt{n \log(p)}$$

with probability $1 - 1/p^2$. Gathering the previous bound with (68), Condition (66), Assumption (32) and $L^4 \log(p) \leq n$, we conclude that

$$t_1 \leq c' \left[|R|_\infty^{1/2} |\Gamma|_\infty^{1/2} + L^2 |\Gamma|_\infty \sqrt{\frac{\log(p)}{n}} \right]$$

simultaneously for all a , with probability $1 - c_3/p$. Together with (65), this concludes the proof of the first part. \square

Acknowledgements. We thank Andrea Montanari for pointing to us the reference [30]. The project is partially supported by the CNRS PICS grant HighClust. Christophe Giraud is partially supported by the LabEx LMH, ANR-11-LABX-0056-LMH. Martin Royer is supported by an IDEX Paris-Saclay IDI grant, ANR-11-IDEX-0003-02. Xi Luo is partially supported by NSF-DMS 1557467, NIH R01EB022911, P01AA019072, P20GM103645, P30AI042853, and S10OD016366. Florentina Bunea is partially supported by NSF-DMS 1712709

REFERENCES

- [1] Emmanuel Abbe and Colin Sandon. Community detection in general stochastic block models: Fundamental limits and efficient algorithms for recovery. In *Foundations of Computer Science (FOCS), 2015 IEEE 56th Annual Symposium on*, pages 670–688, 2015.
- [2] A. A. Amini and E. Levina. On semidefinite relaxations for the block model. *ArXiv e-prints*, June 2014.
- [3] David Arthur and Sergei Vassilvitskii. K-means++: The advantages of careful seeding. In *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '07, pages 1027–1035, Philadelphia, PA, USA, 2007. Society for Industrial and Applied Mathematics.
- [4] Pranjal Awasthi, Moses Charikar, Ravishankar Krishnaswamy, and Ali Kemal Sinop. The hardness of approximation of euclidean k-means. In *Symposium on Computational Geometry*, 2015.
- [5] Onureena Banerjee, Laurent El Ghaoui, and Alexandre d’Aspremont. Model selection through sparse maximum likelihood estimation for multivariate gaussian or binary data. *Journal of Machine learning research*, 9(Mar):485–516, 2008.
- [6] Pierre Bellec, Vincent Perlbarg, Saâd Jbabdi, Mélanie Péligrini-Issac, Jean-Luc Anton, Julien Doyon, and Habib Benali. Identification of large-scale networks in the brain using fmri. *Neuroimage*, 29(4):1231–1243, 2006.
- [7] Juliana S. Bernardes, Fabio RJ Vieira, Lygia MM Costa, and Gerson Zaverucha. Evaluation and improvements of clustering algorithms for detecting remote homologous protein families. *BMC Bioinformatics*, 16(1):1–14, 2015.
- [8] Q. Berthet, P. Rigollet, and P. Srivastava. Exact recovery in the ising blockmodel. *Annals of Statistics (to appear)*, page arXiv:1612.03880, 2018.
- [9] Quentin Berthet and Philippe Rigollet. Complexity theoretic lower bounds for sparse principal component detection. In Shai Shalev-Shwartz and Ingo Steinwart, editors, *Proceedings of the 26th Annual Conference on Learning Theory*, volume 30 of *Proceedings of Machine Learning Research*, pages 1046–1066, Princeton, NJ, USA, 12–14 Jun 2013. PMLR.
- [10] M. Bing, F. Bunea, Y. Ning, and M. Wegkamp. Adaptive estimation in structured factor models with applications to overlapping clustering. *ArXiv e-prints*, 2018.
- [11] Charles Bouveyron and Camille Brunet-Saumard. Model-based clustering of high-dimensional data: A review. *Computational Statistics & Data Analysis*, 71:52–78, 2014.
- [12] F. Bunea, C. Giraud, X. Luo, M. Royer, and N. Verzelen. Supplement to: Model assisted variable clustering: minimax-optimal recovery and algorithms. *Preprint*, 2018.
- [13] Florentina Bunea, Christophe Giraud, and Xi Luo. Minimax optimal variable clustering in g -models via cord. *arXiv preprint arXiv:1508.01939*, 2015.
- [14] Florentina Bunea, Christophe Giraud, Martin Royer, and Nicolas Verzelen. Pecok: a convex optimization approach to variable clustering. *arXiv preprint arXiv:1606.05100*, 2016.
- [15] Yudong Chen and Jiaming Xu. Statistical-computational tradeoffs in planted problems and submatrix localization with a growing number of clusters and submatrices. *Journal of Machine Learning Research*, 17(27):1–57, 2016.
- [16] M Chong, C Bhushan, AA Joshi, S Choi, JP Haldar, DW Shattuck, RN Spreng, and RM Leahy. Individual parcellation of resting fmri with a group functional connectivity prior. *NeuroImage*, 156:87–100, 2017.
- [17] R Cameron Craddock, G Andrew James, Paul E Holtzheimer, Xiaoping P Hu, and Helen S Mayberg. A whole brain fmri atlas generated via spatially constrained spectral clustering. *Human brain mapping*, 33(8):1914–1928, 2012.
- [18] Nicolas Frei dit Frey, Ana Victoria Garcia, Jean Bigeard, Rim Zaag, Eduardo Bueso, Marie Garmier, Stéphanie Pateyron, Marie-Ludivine de Tauzia-Moreau, Véronique Brunaud, Sandrine Balzergue, Jean Colcombet, Sébastien Aubourg, Marie-Laure Martin-Magniette, and Heribert Hirt. Functional analysis of arabidopsis-immune-related mapks uncovers a role for mpk3 as negative regulator of inducible defences. *Genome Biology*, 15(6):1–22, 2014.
- [19] M.F. Glasser, T.S. Coalson, E.C. Robinson, C.D. Hacker, J. Harwell, E. Yacoub, K. Ugurbil, J. Andersson, C.F. Beckmann, M. Jenkinson, S.M. Smith, and D.C. Van Essen. A multi-modal parcellation of human cerebral cortex. *Nature*, 536:171–178, 2016.

- [20] Olivier Guédon and Roman Vershynin. Community detection in sparse networks via grothendieck’s inequality. *arXiv preprint arXiv:1411.4686*, 2014.
- [21] George Andrew James, Onder Hazaroglu, and Keith A Bush. A human brain atlas derived via n-cut parcellation of resting-state and task-based fmri data. *Magnetic resonance imaging*, 34(2):209–218, 2016.
- [22] Daxin Jiang, Chun Tang, and Aidong Zhang. Cluster analysis for gene expression data: a survey. *IEEE Transactions on Knowledge and Data Engineering*, 16(11):1370–1386, Nov 2004.
- [23] Vladimir Koltchinskii and Karim Lounici. Concentration inequalities and moment bounds for sample covariance operators. *Bernoulli*, to appear.
- [24] Ru Kong, Jingwei Li, Nanbo Sun, Mert Sabuncu, Hesheng Liu, Andrew Schaefer, Xi-Nian Zuo, Avram Holmes, Simon Eickhoff, and Thomas Yeo. Spatial topography of individual-specific cortical networks predicts human cognition, personality and emotion. <https://www.biorxiv.org/content/early/2018/01/31/213041>, 2018.
- [25] A. Kumar, Y. Sabharwal, and S. Sen. A simple linear time $(1 + \epsilon)$ -approximation algorithm for k-means clustering in any dimensions. In *Foundations of Computer Science, 2004. Proceedings. 45th Annual IEEE Symposium on*, pages 454–462, Oct 2004.
- [26] Can M Le, Elizaveta Levina, and Roman Vershynin. Optimization via low-rank approximation for community detection in networks. *arXiv preprint arXiv:1406.0067*, 2014.
- [27] Jing Lei and Alessandro Rinaldo. Consistency of spectral clustering in stochastic block models. *Ann. Statist.*, 43(1):215–237, 2015.
- [28] Jing Lei and Lingxue Zhu. A generic sample splitting approach for refined community recovery in stochastic block models. *arXiv preprint arXiv:1411.1469*, 2014.
- [29] S. Lloyd. Least squares quantization in pcm. *IEEE Trans. Inf. Theor.*, 28(2):129–137, September 1982.
- [30] D. G. Mixon, S. Villar, and R. Ward. Clustering subgaussian mixtures by semidefinite programming. *ArXiv e-prints*, February 2016.
- [31] Elchanan Mossel, Joe Neeman, and Allan Sly. Consistency thresholds for binary symmetric block models. *arXiv preprint arXiv:1407.1591*, 2014.
- [32] Jiming Peng and Yu Wei. Approximating k-means-type clustering via semidefinite programming. *SIAM J. on Optimization*, 18(1):186–205, February 2007.
- [33] A. Perry and A. S. Wein. A semidefinite program for unbalanced multisection in the stochastic block model. *ArXiv e-prints*, July 2015.
- [34] Russell A Poldrack. Region of interest analysis for fmri. *Social cognitive and affective neuroscience*, 2(1):67–70, 2007.
- [35] Jonathan D Power, Alexander L Cohen, Steven M Nelson, Gagan S Wig, Kelly Anne Barnes, Jessica A Church, Alecia C Vogel, Timothy O Laumann, Fran M Miezin, Bradley L Schlaggar, et al. Functional network organization of the human brain. *Neuron*, 72(4):665–678, 2011.
- [36] M. Royer. Adaptive clustering through semidefinite programming. *Advances in Neural Information Processing Systems (NIPS)*, 2017.
- [37] Gui Xue, Adam R Aron, and Russell A Poldrack. Common neural substrates for inhibition of spoken and manual responses. *Cerebral Cortex*, 18(8):1923–1932, 2008.
- [38] B.T. Yeo, F.M. Krienen, J. Sepulcre, M.R. Sabuncu, D. Lashkari, M. Hollinshead, J.L. Roffman, J.W. Smoller, L. Z?lle, J.R. Polimeni, B. Fischl, H. Liu, and R.L. Buckner. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, 106:1125–1165, 2011.
- [39] Rim Zaag, Jean-Philippe Tamby, Cécile Guichard, Zakia Tariq, Guillem Rigaiil, Etienne Delannoy, Jean-Pierre Renou, Sandrine Balzergue, Tristan Mary-Huard, Sébastien Aubourg, Marie-Laure Martin-Magniette, and Véronique Brunaud. Gem2net: from gene expression modeling to -omics networks, a new catdb module to investigate *Arabidopsis thaliana* genes involved in stress response. *Nucleic Acids Research*, 43(Database-Issue):1010–1017, 2015.

DEPARTMENT OF STATISTICAL SCIENCE
 CORNELL UNIVERSITY
 ITHACA, NY 14853-2601, USA
 E-MAIL: fb238@cornell.edu

DEPARTMENT OF BIostatISTICS
 BROWN UNIVERSITY
 PROVIDENCE, RI 02903, USA
 E-MAIL: xi.rossi.luo@gmail.com

LABORATOIRE DE MATHÉMATIQUES D’ORSAY
 UNIV. PARIS-SUD, CNRS, UNIVERSITÉ PARIS-SACLAY
 F-91405 ORSAY, FRANCE
 E-MAIL: christophe.giraud@math.u-psud.fr

LABORATOIRE DE MATHÉMATIQUES D’ORSAY
 UNIV. PARIS-SUD, CNRS, UNIVERSITÉ PARIS-SACLAY
 F-91405 ORSAY, FRANCE
 E-MAIL: martin.royer@math.u-psud.fr

INRA
 UMR 729 MISTEA, F-34060 MONTPELLIER, FRANCE
 E-MAIL: nicolas.verzelen@supagro.inra.fr