On identifying clusters from sum-of-norms clustering computation

Tao Jiang* Stephen Vavasis†

June 19, 2020

Abstract

Sum-of-norms clustering is a clustering formulation based on convex optimization that automatically induces hierarchy. Multiple algorithms have been proposed to solve the optimization problem: subgradient descent by Hocking et al. [6], ADMM and ADA by Chi and Lange [4], stochastic incremental algorithm by Panahi et al. [10] and semismooth Newton-CG augmented Lagrangian method by Yuan et al. [13]. All algorithms yield approximate solutions, even though an exact solution is demanded to determine the correct cluster assignment. The purpose of this paper is to close the gap between the output from existing algorithms and the exact solution to the optimization problem. We present a clustering test which identifies and certifies the correct cluster assignment from an approximate solution yielded by any primal-dual algorithm. The test may not succeed if the approximation is inaccurate. However, we show the correct cluster assignment is guaranteed to be found by a symmetric primal-dual path following algorithm after sufficiently many iterations, provided that the model parameter λ avoids a finite number of bad values. Numerical experiments are implemented to support our results.

1 Introduction

Clustering is a fundamental problem in unsupervised learning. The goal of clustering is to seek a partition of n points, $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n \in \mathbb{R}^d$, such that points in the same subset are closer to each other than those that are not. Clustering is usually formulated as a discrete optimization problem, which is combinatorially hard to solve and beset by nonoptimal local minimizers. Classical methods such as k-means and hierarchical clustering are prone to these issues.

*Department of Combinatorics & Optimization, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1, tao.jiang@uwaterloo.ca.

†Department of Combinatorics & Optimization, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1, vavasis@uwaterloo.ca. Research supported in part by a Discovery Grant from the Natural Science and Engineering Research Council (NSERC) of Canada.

Meanwhile, issues of hardness and suboptimality of many nonconvex optimization problems are resolved by convex relaxation. At an affordable computational cost, convex relaxation yields a good solution to the original problem. Pelckmans et al. [11], Hocking et al. [6], and Lindsten et al. [8] proposed the following convex formulation for the clustering problem:

$$\min_{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n\in\mathbf{R}^d} f'(\boldsymbol{x}) = \frac{1}{2}\sum_{i=1}^{n}\|\boldsymbol{x}_i-\boldsymbol{a}_i\|^2 + \lambda\sum_{1\le i<j\le n}\|\boldsymbol{x}_i-\boldsymbol{x}_j\|. \tag{1}$$

where $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ denote the given data and λ denotes the tuning parameter. The formulation (1) is best known as sum-of-norms (SON) clustering, convex clustering, or clusterpath clustering. The clusters are read from the optimizer of (1). Let $\boldsymbol{x}_1^*, \boldsymbol{x}_2^*, \ldots, \boldsymbol{x}_n^*$ denote the optimizer. Points i, i' are assigned to the same cluster if $\boldsymbol{x}_i^* = \boldsymbol{x}_{i'}^*$, and they are assigned to different clusters otherwise.

The optimizer must satisfy the following condition:

$$\boldsymbol{x}_i^* - \boldsymbol{a}_i + \lambda\sum_{j\neq i}\boldsymbol{w}_{ij}^* = \mathbf{0} \qquad \forall i = 1,\ldots,n, \tag{2}$$

where $\boldsymbol{w}_{ij}^*$ is the subgradients of the Euclidean norm $\|\boldsymbol{x}_i - \boldsymbol{x}_j\|$ satisfying

$$\boldsymbol{w}_{ij}^* = \begin{cases} \frac{\boldsymbol{x}_i^*-\boldsymbol{x}_j^*}{\|\boldsymbol{x}_i^*-\boldsymbol{x}_j^*\|}, & \text{for } \boldsymbol{x}_i^* \neq \boldsymbol{x}_j^*, \\ \text{arbitrary point in } B(\mathbf{0},1), & \text{for } \boldsymbol{x}_i^* = \boldsymbol{x}_j^*, \end{cases}$$

with $\boldsymbol{w}_{ij}^* = -\boldsymbol{w}_{ji}^*$ in the second case. We use $B(\boldsymbol{c}, r)$ to denote a closed Euclidean ball centered at $\boldsymbol{c}$ of radius r.

The first term of the objective function ensures $\boldsymbol{x}^*$ is a good approximation of the original data $\boldsymbol{a}$, while the second term penalizes the differences $\boldsymbol{x}_i^* - \boldsymbol{x}_{i'}^*$. As a result, the second term tends to make $\boldsymbol{x}_i^*$ equal to each other for many i. Furthermore, the tuning parameter λ controls the number of clusters indirectly. When $\lambda = 0$, each point is assigned to a cluster of its own. When λ is sufficiently large, all points are assigned to the same cluster.

In this paper, we only consider the l_2 norm. Nonetheless, the reader should be aware that many other norms such as l_1, l_∞, or the general l_p norms are also extensively studied in the literature of sum-of-norms clustering. Also, many researchers investigate the weighted penalty in the objective function. However, our result only applies to the case of unit weights.

Many algorithms, both primal-only and primal-dual methods, have been proposed to solve (1). Primal-only algorithms include subgradient descent by Hocking et al. [6] and stochastic incremental algorithm by Panahi et al. [10]. Primal-dual algorithms are also widely considered such as ADMM and ADA by Chi and Lange [3], and semismooth Newton-CG augmented Lagrangian method by Yuan et al. [13]. All these iterative algorithms yield only approximate solutions, even though exact knowledge of the optimizer is demanded to determine the clusters.

To identify the correct clusters from an approximate solution, authors in practice propose the following approximate test with an artificial tolerance, $\epsilon > 0$. If the approximate solution satisfies $\|\boldsymbol{x}_i - \boldsymbol{x}_{i'}\| \le \epsilon$, i, i' are assigned to the same cluster. Otherwise, i, i' are assigned to different clusters. Hence, the value of artificial tolerance is critical. Unfortunately, to the best of our knowledge, neither the value of the tolerance nor the approximate test itself has been rigorously justified. The test is not robust. Since the relation $\|\boldsymbol{x}_i - \boldsymbol{x}_{i'}\| \leq \epsilon$ is not transitive, it is not clear how the test would cluster points i, j, k if $\|\boldsymbol{x}_i - \boldsymbol{x}_j\| \leq \epsilon$, $\|\boldsymbol{x}_j - \boldsymbol{x}_k\| \leq \epsilon$, and $\|\boldsymbol{x}_i - \boldsymbol{x}_k\| > \epsilon$. The test may not be accurate. The clusters obtained by the approximate test could deviate from the clusters corresponding to the optimizer of (1). The inaccuracy may lead to the failure of known properties of sum-of-norms clustering such as the recovery of a mixture of Gaussians and the agglomeration property. It has been established that for the appropriate choice of λ, (1) exactly recovers a mixture of Gaussians due to Panahi et al. [10], Sun et al. [13], and Jiang et al. [7]. However, it is unknown if the recovery result still holds when the approximate test is applied. Hocking et al. [6] conjectured that sum-of-norms clustering is agglomerative in the sense that as λ increases, clusters may fuse but never break apart. The conjecture was proven by Chiquet, Gutierrez and Rigaill [5] with some techniques which may not be applicable when the approximate test is implemented. Thus the agglomeration property may no longer hold. The full agglomeration theorem is stated as follows.

Theorem 1. *If there is a C such that minimizer $\boldsymbol{x}^*$ of* (1) *at λ satisfies $\boldsymbol{x}_i^* = \hat{\boldsymbol{x}}$ for $i \in C$, $\boldsymbol{x}_i^* \neq \hat{\boldsymbol{x}}$ for $i \notin C$ for some $\hat{\boldsymbol{x}} \in \mathbf{R}^d$, then at any $\lambda' \geq \lambda$, there exists an $\hat{\boldsymbol{x}}' \in \mathbf{R}^d$ such that the minimizer of* (1)*, $\bar{\boldsymbol{x}}^*$, satisfies $\bar{\boldsymbol{x}}_i^* = \hat{\boldsymbol{x}}'$ for $i \in C$.*

Let *fusion values* denote the values of λ at which clusters fuse to form a larger cluster. According to Theorem 1, there are at most n fusion values.

The purpose of this paper is to present our clustering test and to justify it rigorously. The clustering test takes a primal and dual feasible solution for the second-order cone formulation of sum-of-norms clustering and attempts to determine all clusters. The test may report ‘success’ or ‘failure’. If the test reports ‘success’, all clusters are correctly identified and a certificate is produced. The test and the proof of correctness are stated in Section 4. The proof heavily relies on two sufficient conditions for clustering and distinct clustering, which are presented in Section 2. The test requires the knowledge of a primal and dual feasible solution for the conic formulation of sum-of-norms clustering, which can be constructed from the output of any primal-dual algorithm. The conic formulation and algorithms are stated in Section 3. If a primal-dual path following algorithm is used, the test is guaranteed to report ‘success’ after a finite number of iterations except the test may never report ‘success’ when λ is at a fusion value. These results are shown in Section 6. The proof of the theoretical guarantee is a result of the properties of the central path, which are stated in Section 5. In Section 7, we present a few computational experiments to verify our test in practice.

2 Sufficient conditions on clustering

Let $C \subseteq \{1, 2, ..., n\}$ denote a subset. To draw meaningful conclusions about C, we use the two sufficient conditions in this section to develop our test. Theorem 2 is due to Chiquet et al. [5], which is a sufficient condition for clustering. The reader may refer to the work by Jiang, Vavasis and Zhai [7] for an exposition of Theorem 2. Theorem 3 is a sufficient condition for distinct clustering; its proof is trivial.

Let $\boldsymbol{x}^*$ denote the optimal solution of (1), and let $\boldsymbol{x}$ denote the output of some primal-dual algorithm which solves (1).

Theorem 2. *Suppose there exist $\boldsymbol{q}_{ij}^*$ for all $i, j \in C, i \neq j$ solving the following system (3). Then there exists some $\hat{\boldsymbol{x}} \in \mathbf{R}^d$ such that the minimizer $\boldsymbol{x}^*$ of (1) satisfies $\boldsymbol{x}_i^* = \hat{\boldsymbol{x}}$ for $i \in C$, hence C is a cluster or part of a larger cluster.*

$$\begin{aligned} \boldsymbol{a}_i - \frac{1}{|C|}\sum_{l \in C} \boldsymbol{a}_l = \lambda \sum_{j \in C - \{i\}} \boldsymbol{q}_{ij}^*, \quad \forall i \in C, \\ \left\|\boldsymbol{q}_{ij}^*\right\| \leq 1, \quad \forall i, j \in C, i \neq j, \\ \boldsymbol{q}_{ij}^* = -\boldsymbol{q}_{ji}^*, \quad \forall i, j \in C, i \neq j. \end{aligned} \tag{3}$$

Theorem 3. *Define $\tau > 0$ such that the true optimizer and the approximate solution are at distance of at most τ away (i.e. $\|\boldsymbol{x} - \boldsymbol{x}^*\| \leq \tau$). If there exist $i, j \in C$ such that $\|\boldsymbol{x}_i - \boldsymbol{x}_j\| > 2\tau$, then C is not a cluster or part of a larger cluster.*

3 Feasibility and complementary slackness

In this section, we consider a second-order cone (SOCP) formulation of (1). Both feasibility and complementary slackness are stated. A second-order cone program can be directly solved by a feasible interior-point method. For infeasible algorithms such as the ADMM proposed by Chi and Lange [3], we construct a feasible solution for the SOCP from the outputs of such algorithms.

3.1 Second-order cone formulation

We first present the equivalent SOCP formulation to (1)

$$\min_{\boldsymbol{x},\boldsymbol{y},\boldsymbol{z},s,u,t} \quad f(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, s, u, t) = \sum_{i=1}^{n} s_i + \lambda \sum_{1 \leq i < j \leq n} t_{ij} \tag{4a}$$

$$\text{s.t} \quad \boldsymbol{x}_i - \boldsymbol{x}_j - \boldsymbol{y}_{ij} = \mathbf{0}\ , \quad \forall 1 \leq i < j \leq n\ , \tag{4b}$$

$$\boldsymbol{x}_i - \boldsymbol{z}_i = \boldsymbol{a}_i\ , \quad \forall i = 1, \ldots, n\ , \tag{4c}$$

$$s_i - u_i = 1\ , \quad \forall i = 1, \ldots, n\ , \tag{4d}$$

$$t_{ij} \geq \|\boldsymbol{y}_{ij}\|\ , \quad \forall 1 \leq i < j \leq n\ , \tag{4e}$$

$$s_i \geq \left\| \begin{pmatrix} \boldsymbol{z}_i \\ u_i \end{pmatrix} \right\|\ , \quad \forall i = 1, \ldots, n\ . \tag{4f}$$

The SOCP formulation of the dual problem is as follows.

$$\max_{\boldsymbol{\delta},\boldsymbol{\beta},\gamma} \quad h(\boldsymbol{\delta},\boldsymbol{\beta},\gamma) = \sum_{i=1}^{n} \boldsymbol{a}_i^T \boldsymbol{\beta}_i + \sum_{i=1}^{n} \gamma_i \tag{5a}$$

$$\text{s.t} \quad -\sum_{j=1}^{i-1} \boldsymbol{\delta}_{ji} + \sum_{j=i+1}^{n} \boldsymbol{\delta}_{ij} + \boldsymbol{\beta}_i = \boldsymbol{0} \ , \quad \forall i = 1, \ldots, n \ , \tag{5b}$$

$$\lambda \geq \|\boldsymbol{\delta}_{ij}\| \ , \quad \forall 1 \leq i < j \leq n \ , \tag{5c}$$

$$1 - \gamma_i \geq \left\| \begin{pmatrix} \boldsymbol{\beta}_i \\ \gamma_i \end{pmatrix} \right\| \ , \quad \forall i = 1, \ldots, n \ . \tag{5d}$$

Both primal and dual problems are feasible, and Slater condition holds for both problems. Consider the following primal and dual feasible solution:

$$\boldsymbol{x}_i = \boldsymbol{a}_i, \boldsymbol{z}_i = \boldsymbol{0}, s_i = 1, u_i = 0, \ \forall i = 1, \ldots, n; \quad \boldsymbol{y}_{ij} = \boldsymbol{a}_i - \boldsymbol{a}_j, t_{ij} = \|\boldsymbol{a}_i - \boldsymbol{a}_j\| + 1, \ \forall 1 \leq i < j \leq n.$$

$$\boldsymbol{\delta}_{ij} = \boldsymbol{0}, \ \forall 1 \leq i < j \leq n; \quad \boldsymbol{\beta}_i = \boldsymbol{0}, \gamma_i = 0, \ \forall i = 1, \ldots, n.$$

which is a also primal and dual Slater point. Hence, strong duality holds since the problem is formulated as convex optimization.

For the clustering test, we require primal-dual feasibility for the above primal and dual SOCP. Such a primal and dual feasible solution can be obtained by applying a feasible primal-dual interior-point method to the problem above. Each iterate of the algorithm is feasible, so is the output. Nevertheless, the output may not be feasible for our SOCP when a general primal-dual algorithm is used to solve for (1). Luckily, given that the output is close to the feasible set, we are able to find a small perturbation on the output to attain feasibility. The rest of the section elaborates on the perturbation and validates feasibility for the perturbed solution.

Now let us consider a general primal-dual algorithm which solves (1) and yields an output that is either in or close to the feasible set. The dual of (1) is as follows.

$$\max_{\boldsymbol{\delta}} \quad h'(\boldsymbol{\delta}) = -\frac{1}{2} \sum_{i=1}^{n} \left\| \sum_{j=1}^{i-1} \boldsymbol{\delta}_{ji} - \sum_{j=i+1}^{n} \boldsymbol{\delta}_{ij} \right\|^2 - \sum_{1 \leq i < j \leq n} \langle \boldsymbol{\delta}_{ij}, \boldsymbol{a}_i - \boldsymbol{a}_j \rangle \tag{6a}$$

$$\text{s.t} \quad \|\boldsymbol{\delta}_{ij}\| \leq \lambda \ , \quad \forall 1 \leq i < j \leq n \ . \tag{6b}$$

Notice the formulation is equivalent to the SOCP dual after $\boldsymbol{\beta}$ and γ are eliminated. Let $(\boldsymbol{x}, \boldsymbol{\delta})$ denote the output yielded by the primal-dual algorithm. To construct a feasible solution from $(\boldsymbol{x}, \boldsymbol{\delta})$, we first update $\boldsymbol{\delta}$ as follows

$$\boldsymbol{\delta}_{ij} \leftarrow \begin{cases} \frac{\lambda \boldsymbol{\delta}_{ij}}{\|\boldsymbol{\delta}_{ij}\|}, & \text{if } \|\boldsymbol{\delta}_{ij}\| > \lambda, \\ \boldsymbol{\delta}_{ij}, & \text{otherwise.} \end{cases}$$

The updated $\boldsymbol{\delta}_{ij}$ has norm no more than λ. Notice that the perturbation is small provided that the dual solution was already close to the feasible set. Next, define the following variables:

$$\begin{aligned}
\boldsymbol{y}_{ij} &= \boldsymbol{x}_i - \boldsymbol{x}_j, \quad \forall 1 \le i < j \le n,\\
\boldsymbol{z}_i &= \boldsymbol{x}_i - \boldsymbol{a}_i, \quad \forall i = 1,\dots,n,\\
s_i &= \frac{1}{2}(1 + \|\boldsymbol{z}_i\|^2), \quad \forall i = 1,\dots,n,\\
u_i &= \frac{1}{2}(-1 + \|\boldsymbol{z}_i\|^2), \quad \forall i = 1,\dots,n,\\
t_{ij} &= \|\boldsymbol{y}_{ij}\|, \quad \forall 1 \le i < j \le n,\\
\boldsymbol{\beta}_i &= \sum_{j=1}^{i-1} \boldsymbol{\delta}_{ji} - \sum_{j=i+1}^{n} \boldsymbol{\delta}_{ij}, \quad \forall i = 1,\dots,n,\\
\gamma_i &= \frac{1}{2}(1 - \|\boldsymbol{\beta}_i\|^2), \quad \forall i = 1,\dots,n.
\end{aligned}$$

It can be easily verified that these newly defined variables, $\boldsymbol{x}$ from the primal-dual algorithm and the updated $\boldsymbol{\delta}$ form a primal and dual feasible solution for the SOCP. The original objective function value at $(\boldsymbol{x}, \boldsymbol{\delta})$ and the SCOP objective function value at the updated solution differ by a constant $\frac{n}{2}$:

$$\begin{aligned}
f'(\boldsymbol{x}) &= \frac{1}{2}\sum_{i=1}^{n} \|\boldsymbol{x}_i - \boldsymbol{a}_i\|^2 + \lambda \sum_{1\le i<j\le n} \|\boldsymbol{y}_{ij}\|\\
&= \sum_{i=1}^{n} s_i - \frac{n}{2} + \lambda \sum_{1\le i<j\le n} t_{ij}\\
&= f(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, s, u, t) - \frac{n}{2}
\end{aligned} \tag{7}$$

$$\begin{aligned}
h'(\boldsymbol{\delta}) &= -\frac{1}{2}\sum_{i=1}^{n} \|\sum_{j=i+1}^{n} \boldsymbol{\delta}_{ij} - \sum_{j=1}^{i-1} \boldsymbol{\delta}_{ji}\|^2 - \sum_{1\le i<j\le n} \langle \boldsymbol{\delta}_{ij}, \boldsymbol{a}_i - \boldsymbol{a}_j\rangle\\
&= \sum_{i=1}^{n} \gamma_i - \frac{n}{2} + \sum_{i=1}^{n} \langle \boldsymbol{a}_i, \boldsymbol{\beta}_i\rangle\\
&= h(\boldsymbol{\delta}, \boldsymbol{\beta}, \gamma) - \frac{n}{2}
\end{aligned} \tag{8}$$

3.2 Complementary slackness

Let $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, s, u, t, \boldsymbol{\delta}, \boldsymbol{\beta}, \gamma)$ be a primal and dual feasible solution for the SOCP formulation of sum-of-norms clustering. Let us define $\boldsymbol{\epsilon}^{ij} = \begin{pmatrix} \epsilon_1^{ij} \\ \boldsymbol{\epsilon}_2^{ij} \end{pmatrix}$ for all $1 \le i < j \le n$ and $\boldsymbol{\sigma}^i = \begin{pmatrix} \sigma_1^i \\ \boldsymbol{\sigma}_2^i \\ \sigma_3^i \end{pmatrix}$ for all $i = 1, \dots, n$ as follows:

$$t_{ij}\lambda + \boldsymbol{y}_{ij}^T\boldsymbol{\delta}_{ij} = \epsilon_1^{ij}, \quad \forall 1 \le i < j \le n, \tag{9}$$
$$t_{ij}\boldsymbol{\delta}_{ij} + \lambda\boldsymbol{y}_{ij} = \boldsymbol{\epsilon}_2^{ij}, \quad \forall 1 \le i < j \le n, \tag{10}$$
$$s_i(1-\gamma_i) + \boldsymbol{z}_i^T\boldsymbol{\beta}_i + u_i\gamma_i = \sigma_1^i, \quad \forall i = 1,\ldots,n, \tag{11}$$
$$s_i\boldsymbol{\beta}_i + (1-\gamma_i)\boldsymbol{z}_i = \boldsymbol{\sigma}_2^i, \quad \forall i = 1,\ldots,n, \tag{12}$$
$$s_i\gamma_i + (1-\gamma_i)u_i = \sigma_3^i, \quad \forall i = 1,\ldots,n. \tag{13}$$

At the optimizer, there hold $\boldsymbol{\epsilon} = \mathbf{0}, \boldsymbol{\sigma} = \mathbf{0}$ by KKT conditions. The system of equalities above becomes the complementary slackness condition. At an approximate solution, the right-hand sides $\boldsymbol{\epsilon}, \boldsymbol{\sigma}$ are non-zero. If $\boldsymbol{\epsilon}^{ij} = \begin{pmatrix} \mu' \\ \mathbf{0} \end{pmatrix}, \boldsymbol{\sigma}^i = \begin{pmatrix} \mu' \\ \mathbf{0} \\ 0 \end{pmatrix}$ for all $i = 1,\ldots,n$ and for all $1 \le i < j \le n$, we refer the corresponding solution as a μ'-centered solution. Otherwise, an upper bound on general right-hand sides $\boldsymbol{\epsilon}, \boldsymbol{\sigma}$ can be derived from the duality gap:

$$\begin{aligned}
&f'(\boldsymbol{x}) - h'(\boldsymbol{\delta}) \\
=&f(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z},s,u,t) - h(\boldsymbol{\delta},\boldsymbol{\beta},\gamma) \quad \text{(By (7) and (8))} \\
=&\sum_{i=1}^n s_i - \sum_{i=1}^n \gamma_i + \sum_{i=1}^n \langle \boldsymbol{x}_i - \boldsymbol{a}_i, \boldsymbol{\beta}_i\rangle + \lambda \sum_{1\le i<j\le n} t_{ij} - \sum_{i=1}^n \langle \boldsymbol{x}_i, \boldsymbol{\beta}_i\rangle \quad \text{(By adding and subtracting } \sum_{i=1}^n \langle \boldsymbol{x}_i, \boldsymbol{\beta}_i\rangle) \\
=&\sum_{i=1}^n s_i - \sum_{i=1}^n \gamma_i + \sum_{i=1}^n \langle \boldsymbol{x}_i - \boldsymbol{a}_i, \boldsymbol{\beta}_i\rangle + \lambda \sum_{1\le i<j\le n} t_{ij} - \sum_{i=1}^n \langle \boldsymbol{x}_i, \sum_{j=1}^{i-1} \boldsymbol{\delta}_{ji} - \sum_{j=i+1}^{n} \boldsymbol{\delta}_{ij}\rangle \quad \text{(By (5b))} \\
=&\sum_{i=1}^n s_i - \sum_{i=1}^n \gamma_i + \sum_{i=1}^n \langle \boldsymbol{x}_i - \boldsymbol{a}_i, \boldsymbol{\beta}_i\rangle + \lambda \sum_{1\le i<j\le n} t_{ij} - \sum_{1\le i<j\le n} \langle \boldsymbol{x}_j - \boldsymbol{x}_i, \boldsymbol{\delta}_{ij}\rangle \quad \text{(By expanding the summation)} \\
=&\sum_{i=1}^n (s_i - \gamma_i + \langle \boldsymbol{x}_i - \boldsymbol{a}_i, \boldsymbol{\beta}_i\rangle) + \sum_{1\le i<j\le n} (\lambda t_{ij} + \langle \boldsymbol{y}_{ij}, \boldsymbol{\delta}_{ij}\rangle) \quad \text{(By (4b))} \\
=&\sum_{i=1}^n (s_i(1-\gamma_i) + \langle \boldsymbol{z}_i, \boldsymbol{\beta}_i\rangle + u_i\gamma_i) + \sum_{1\le i<j\le n} (\lambda t_{ij} + \langle \boldsymbol{y}_{ij}, \boldsymbol{\delta}_{ij}\rangle) \quad \text{(By (4c), (4d))} \\
=&\sum_{i=1}^n \sigma_1^i + \sum_{1\le i<j\le n} \epsilon_1^{ij}
\end{aligned}$$

Each term in the both summations is non-negative as shown below:

$$\sigma_1^i = s_i(1-\gamma_i) + \langle \boldsymbol{z}_i, \boldsymbol{\beta}_i\rangle + u_i\gamma_i = \frac{1}{2}(\|\boldsymbol{z}_i\|^2 + 2\langle \boldsymbol{z}_i, \boldsymbol{\beta}_i\rangle + \|\boldsymbol{\beta}_i\|^2) \ge 0, \quad \forall i = 1,\ldots,n,$$

$$\epsilon_1^{ij} = \lambda t_{ij} + \langle \boldsymbol{y}_{ij}, \boldsymbol{\delta}_{ij}\rangle \ge \lambda t_{ij} - \|\boldsymbol{y}_{ij}\|\,\|\boldsymbol{\delta}_{ij}\| \ge \lambda t_{ij} - \lambda\,\|\boldsymbol{y}_{ij}\| = 0, \quad \forall 1 \le i < j \le n.$$

Define $\mu := f'(\boldsymbol{x}) - h'(\boldsymbol{\delta})$ to be the duality gap at the feasible solution. Combined with the non-negativity condition, $\sigma_1^i, \epsilon_1^{ij}$ satisfy $\sigma_1^i \le \mu$ for all $i = 1,\ldots,n$ and $\epsilon_1^{ij} \le \mu$ for all $1 \le i < j \le n$. At termination, the duality gap μ at the feasible solution is small, which implies the right-hand sides $\sigma_1^i, \epsilon_1^{ij}$ are also well bounded.

We now have $\sigma_1^i, \epsilon_1^{ij}$ upper bounded in terms of μ, and the remainder of the section is to establish upper bounds on $\left\|\boldsymbol{\epsilon}_2^{ij}\right\|$ and $\left\|\begin{pmatrix}\boldsymbol{\sigma}_2^i \\ \sigma_3^i\end{pmatrix}\right\|$. In fact, in (14) and (15) below, we show that both are upper bounded by $O(\sqrt{\mu})$. Consider a general setting of second-order cone programming.

Lemma 4. *Let $(\boldsymbol{x}, \boldsymbol{z})$ denote a primal and dual feasible solution for a second-order cone program where $\boldsymbol{x} = \begin{pmatrix} x_0 \\ \bar{\boldsymbol{x}} \end{pmatrix}, \boldsymbol{z} = \begin{pmatrix} z_0 \\ \bar{\boldsymbol{z}} \end{pmatrix}$. If $\boldsymbol{x}^T\boldsymbol{z} \leq \mu$, then $\|z_0\bar{\boldsymbol{x}} + x_0\bar{\boldsymbol{z}}\| \leq \sqrt{2x_0z_0\mu}$.*

Proof. If $x_0 = 0$, then $\|\bar{\boldsymbol{x}}\| \leq x_0 = 0$ by feasibility assumption. Hence, $\bar{\boldsymbol{x}} = \mathbf{0}$, which implies $\|z_0\bar{\boldsymbol{x}} + x_0\bar{\boldsymbol{z}}\| = 0 \leq \sqrt{2x_0z_0\mu}$. If $z_0 = 0$, then $\boldsymbol{z}$ satisfies $\|z_0\bar{\boldsymbol{x}} + x_0\bar{\boldsymbol{z}}\| = 0 \leq \sqrt{2x_0z_0\mu}$ by the same argument.

Otherwise, $x_0 > 0, z_0 > 0$, and we derive the following inequalities

$$\begin{aligned}
\boldsymbol{x}^T\boldsymbol{z} = x_0z_0 + \bar{\boldsymbol{x}}^T\bar{\boldsymbol{z}} \leq \mu \\
\Rightarrow 1 + \frac{\bar{\boldsymbol{x}}^T}{x_0}\frac{\bar{\boldsymbol{z}}}{z_0} \leq \frac{\mu}{x_0z_0} \quad (\text{Since } x_0 > 0, z_0 > 0) \\
\Rightarrow \left\|\frac{\bar{\boldsymbol{x}}}{x_0} + \frac{\bar{\boldsymbol{z}}}{z_0}\right\|^2 = \left\|\frac{\bar{\boldsymbol{x}}}{x_0}\right\|^2 + \left\|\frac{\bar{\boldsymbol{z}}}{z_0}\right\|^2 + 2\frac{\bar{\boldsymbol{x}}^T}{x_0}\frac{\bar{\boldsymbol{z}}}{z_0} \leq 2 - 2 + \frac{2\mu}{x_0z_0} \quad (\text{Since } x_0 \geq \|\bar{\boldsymbol{x}}\|, z_0 \geq \|\bar{\boldsymbol{z}}\|) \\
\Rightarrow \left\|\frac{\bar{\boldsymbol{x}}}{x_0} + \frac{\bar{\boldsymbol{z}}}{z_0}\right\| \leq \sqrt{\frac{2\mu}{x_0z_0}} \\
\Rightarrow \|z_0\bar{\boldsymbol{x}} + x_0\bar{\boldsymbol{z}}\| \leq \sqrt{2x_0z_0\mu}.
\end{aligned}$$

□

Let $\bar{\boldsymbol{a}} := \frac{1}{n}\sum_{i=1}^n \boldsymbol{a}_i$ denote the centroid of all data points. Let $\boldsymbol{x}_1' := \boldsymbol{x}_2' := ... := \boldsymbol{x}_n' := \bar{\boldsymbol{a}}$. Then the primal objective value of the original sum-of-norms formulation at $\boldsymbol{x}'$ is

$$f'(\boldsymbol{x}') = \frac{1}{2}\sum_{l=1}^n \|\bar{\boldsymbol{a}} - \boldsymbol{a}_l\|^2.$$

Let $\boldsymbol{\delta}_{ij}' = \mathbf{0}$ for all $1 \leq i < j \leq n$. Then $\boldsymbol{\delta}'$ is a feasible solution to the dual problem of the original formulation and the dual objective value at $\boldsymbol{\delta}'$ is

$$h'(\boldsymbol{\delta}') = 0.$$

Let f^* and h^* denote the primal and dual optimal values of the SOCP respectively, which must satisfy the following inequality by strong duality:

$$\frac{n}{2} = h'(\boldsymbol{\delta}') + \frac{n}{2} \leq h^* = f^* \leq f'(\boldsymbol{x}') + \frac{n}{2} = \frac{1}{2}\sum_{l=1}^n \|\bar{\boldsymbol{a}} - \boldsymbol{a}_l\|^2 + \frac{n}{2},$$

where the term $\frac{n}{2}$ comes from (7) and (8). At the feasible solution $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, s, u, t, \boldsymbol{\delta}, \boldsymbol{\beta}, \gamma)$, the objective value is at a distance of at most μ away from the optimal value, which implies

$$\sum_{i=1}^n s_i + \lambda \sum_{1 \leq i < j \leq n} t_{ij} \leq f^* + \mu \leq \frac{1}{2}\sum_{l=1}^n \|\bar{\boldsymbol{a}} - \boldsymbol{a}_l\|^2 + \frac{n}{2} + \mu,$$

which is rearranged to

$$\sum_{i=1}^{n}\left(s_i - \frac{1}{2}\right) + \lambda \sum_{1\le i<j\le n} t_{ij} \le \frac{1}{2}\sum_{l=1}^{n}\left\|\bar{\boldsymbol{a}} - \boldsymbol{a}_l\right\|^2 + \mu.$$

Moreover, by feasibility, $s_i \ge \frac{1}{2}$ holds for all $i = 1, \ldots, n$ and $t_{ij} \ge 0$ holds for all $1 \le i < j \le n$. Hence,

$$s_i \le \frac{1}{2}\sum_{l=1}^{n}\left\|\bar{\boldsymbol{a}} - \boldsymbol{a}_l\right\|^2 + \frac{1}{2} + \mu, \quad t_{ij} \le \frac{1}{\lambda}\left(\frac{1}{2}\sum_{l=1}^{n}\left\|\bar{\boldsymbol{a}} - \boldsymbol{a}_l\right\|^2 + \mu\right).$$

As $t_{ij}\lambda + \boldsymbol{y}_{ij}^T\boldsymbol{\delta}_{ij} = \epsilon_1^{ij} \le \mu$, $\left\|\boldsymbol{\epsilon}_2^{ij}\right\|$ has the following upper bound by Lemma 4

$$\left\|\boldsymbol{\epsilon}_2^{ij}\right\| = \left\|t_{ij}\boldsymbol{\delta}_{ij} + \lambda\boldsymbol{y}_{ij}\right\| \le \sqrt{2t_{ij}\lambda\mu} \le \sqrt{\sum_{l=1}^{n}\left\|\bar{\boldsymbol{a}} - \boldsymbol{a}_l\right\|^2\mu + 2\mu^2}. \tag{14}$$

Similarly, at the feasible solution, the dual objective value is at a distance of at most μ away from the optimal value, which implies

$$\sum_{i=1}^{n}\boldsymbol{a}_i^T\boldsymbol{\beta}_i + \sum_{i=1}^{n}\gamma_i \ge h^* - \mu \ge \frac{n}{2} - \mu,$$

which is rearranged to

$$\sum_{i=1}^{n}\left(\frac{1}{2} - \gamma_i\right) \le \sum_{i=1}^{n}\boldsymbol{a}_i^T\boldsymbol{\beta}_i + \mu.$$

By feasibility, $\frac{1}{2} - \gamma_i \ge 0$, which implies

$$1 - \gamma_i \le \frac{1}{2} + \sum_{i=1}^{n}\boldsymbol{a}_i^T\boldsymbol{\beta}_i + \mu.$$

Since $\lambda \ge \|\boldsymbol{\delta}_{ij}\|$, $\|\boldsymbol{\beta}_i\|$ satisfies

$$\|\boldsymbol{\beta}_i\| = \left\|\sum_{j=1}^{i-1}\boldsymbol{\delta}_{ji} - \sum_{j=i+1}^{n}\boldsymbol{\delta}_{ij}\right\| \le (n-1)\lambda.$$

By Cauchy-Schwartz inequality,

$$\boldsymbol{a}_i^T\boldsymbol{\beta}_i \le \|\boldsymbol{a}_i\| \cdot \|\boldsymbol{\beta}_i\| \le (n-1)\lambda\left\|\boldsymbol{a}_i\right\|.$$

Therefore, $1 - \gamma_i$ satisfies

$$1 - \gamma_i \le \frac{1}{2} + \sum_{l=1}^{n}(n-1)\lambda\left\|\boldsymbol{a}_l\right\| + \mu.$$

Since $s_i(1-\gamma_i)+\boldsymbol{z}_i^T\boldsymbol{\beta}_i+u_i\gamma_i=\sigma_1^i$, $\left\|\begin{pmatrix}\boldsymbol{\sigma}_2^i\\ \sigma_3^i\end{pmatrix}\right\|$ has the following upper bound by Lemma 4

$$\begin{aligned}\left\|\begin{pmatrix}\boldsymbol{\sigma}_2^i\\ \sigma_3^i\end{pmatrix}\right\| &= \left\|\begin{pmatrix}s_i\boldsymbol{\beta}_i+(1-\gamma_i)\boldsymbol{z}_i\\ s_i\gamma_i+(1-\gamma_i)u_i\end{pmatrix}\right\| \le \sqrt{2s_i(1-\gamma_i)\mu}\\ &\le \sqrt{2\cdot\left(\frac{1}{2}\sum_{l=1}^n\|\bar{\boldsymbol{a}}-\boldsymbol{a}_l\|^2+\frac{1}{2}+\mu\right)\cdot\left(\frac{1}{2}+\sum_{l=1}^n(n-1)\lambda\|\boldsymbol{a}_l\|+\mu\right)\cdot\mu} \qquad (15)\\ &\le \sqrt{\left(\sum_{l=1}^n\|\bar{\boldsymbol{a}}-\boldsymbol{a}_l\|^2+1+2\mu\right)\cdot\left(\frac{1}{2}+\sum_{l=1}^n(n-1)\lambda\|\boldsymbol{a}_l\|+\mu\right)\mu}.\end{aligned}$$

4 Clustering test

Given a primal and dual feasible solution $(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z},s,u,t,\boldsymbol{\delta},\boldsymbol{\beta},\gamma)$ with duality gap μ, we find candidate clusters as follows. First, select an index i from $\{1,\ldots,n\}$ arbitrarily. Construct a ball of radius $\mu^{0.75}$ about $\boldsymbol{x}_i$. Create a candidate cluster with all indices k such that $\boldsymbol{x}_k$ is located in the ball about $\boldsymbol{x}_i$ (i.e. $\{k:\|\boldsymbol{x}_i-\boldsymbol{x}_k\|\le\mu^{3/4}\}$). Now find an index j that is not in any candidate cluster and construct a ball about $\boldsymbol{x}_j$. Repeat until all data points are used up.

If the output of the primal-dual algorithm is not feasible for the second-order cone program, we construct a feasible solution as described in the previous section. With the feasible solution, we define

$$\boldsymbol{g}_{ij}:=\begin{cases}-\boldsymbol{\delta}_{ij}, & \text{if } i<j,\\ \boldsymbol{\delta}_{ji}, & \text{if } j<i.\end{cases}$$

For any candidate cluster C, compute $\boldsymbol{q}_{ij}:=\boldsymbol{g}_{ij}+\frac{1}{m}\cdot(\boldsymbol{x}_i-\boldsymbol{x}_j-\boldsymbol{\omega}_i+\boldsymbol{\omega}_j)+\frac{1}{m}\sum_{k\notin C}(\boldsymbol{g}_{ik}-\boldsymbol{g}_{jk})$ for all $i,j\in C, i\ne j$, denoted as *Chiquet-Gutierrez-Rigail (CGR) subgradients*. Check if the following two conditions hold:

***CGR subgradient condition*:** All CGR subgradients $\boldsymbol{q}_{ij}$ satisfy the CGR inequality $\|\boldsymbol{q}_{ij}\|\le\lambda$.

***Separation condition*:** All candidate clusters are separated at distance of at least 2τ, where $\tau=\sqrt{2\mu}$.

If both conditions hold for all candidate clusters, then the test terminates and reports 'success'. Each candidate cluster is a real cluster given by the optimal solution, thus all clusters are correctly identified. The $\boldsymbol{q}_{ij}$'s serve as certificates. If either condition fails for any candidate cluster, the test reports 'failure'. One has to run more iterations of the algorithm to decrease the duality gap μ. Repeat the process until the test reports 'success'. Note that this test is algorithm-independent, but it does require the algorithm be of primal-dual type.

The test is validated by our two sufficient conditions presented in Section 2. The CGR subgradients condition certifies that each cluster we identify is indeed a cluster or part of a larger cluster by Theorem 2. This is presented in Section 4.1. The separation condition certifies that there is no super-cluster with more than one cluster we identify by Theorem 3, as shown in 4.2. Therefore, we determine all clusters correctly when the test succeeds.

4.1 CGR subgradients and clustering corollary

Let $C \subseteq [n]$ denote a subset of points. Let $m := |C|$ denote the cardinality of C.

Lemma 5. *For all* $i, j \in C, i \neq j$, *define* $\boldsymbol{q}_{ij} := \boldsymbol{g}_{ij} + \frac{1}{m} \cdot (\boldsymbol{x}_i - \boldsymbol{x}_j - \boldsymbol{\omega}_i + \boldsymbol{\omega}_j) + \frac{1}{m} \sum_{k \notin C} (\boldsymbol{g}_{ik} - \boldsymbol{g}_{jk})$. *Then* $\boldsymbol{q}_{ij}$ *satisfies*

$$\boldsymbol{a}_i - \bar{\boldsymbol{a}} = \sum_{j \in C \setminus \{i\}} \boldsymbol{q}_{ij}, \quad \forall i \in C \tag{16}$$

$$\boldsymbol{q}_{ij} = -\boldsymbol{q}_{ji}, \quad \forall i, j \in C, i \neq j, \tag{17}$$

where $\bar{\boldsymbol{a}} = \frac{1}{m} \sum_{i \in C} \boldsymbol{a}_i$.

Proof. Substitute the primal constraint (4d) into the perturbed complementary slackness (13) to obtain the following equality of γ_i and s_i

$$1 - \gamma_i = s_i - \sigma_3^i, \quad \forall i = 1, \ldots, n.$$

Substitute the equality above into (12) and divide both sides by s_i to obtain the following equation of $\boldsymbol{\beta}_i$ in terms of $\boldsymbol{z}_i$

$$\boldsymbol{\beta}_i = -\boldsymbol{z}_i + \boldsymbol{\omega}_i, \quad \forall i = 1, \ldots, n.$$

Notice that the operation is valid because $s_i \geq \frac{1}{2}$ by the primal constraint (4d) and (4f). Substitute the primal constraint (4c) and the equality above into the dual constraint (5b) yielding

$$-\sum_{j=1}^{i-1} \boldsymbol{\delta}_{ji} + \sum_{j=i+1}^{n} \boldsymbol{\delta}_{ij} - \boldsymbol{x}_i + \boldsymbol{a}_i + \boldsymbol{\omega}_i = \boldsymbol{0}, \quad \forall i = 1, \ldots, n.$$

With the definition of $\boldsymbol{g}_{ij}$, the equality above is rewritten as

$$-\boldsymbol{x}_i + \boldsymbol{a}_i + \boldsymbol{\omega}_i - \sum_{j \neq i} \boldsymbol{g}_{ij} = \boldsymbol{0}, \quad \forall i = 1, \ldots, n. \tag{18}$$

By (18), we have the following equality holds for all $i \in C$

$$-\boldsymbol{x}_i + \boldsymbol{a}_i + \boldsymbol{\omega}_i - \sum_{j \in C \setminus \{i\}} \boldsymbol{g}_{ij} - \sum_{k \notin C} \boldsymbol{g}_{ik} = \boldsymbol{0}. \tag{19}$$

Sum (19) over all $i \in C$ and divide the new equality by m to obtain

$$-\frac{1}{m} \sum_{i \in C} \boldsymbol{x}_i + \bar{\boldsymbol{a}} + \frac{1}{m} \sum_{i \in C} \boldsymbol{\omega}_i - \frac{1}{m} \sum_{i \in C} \sum_{k \notin C} \boldsymbol{g}_{ik} = \boldsymbol{0}. \tag{20}$$

Change the index in (20) from i to j. Subtract (20) from (19) to obtain

$$-\boldsymbol{x}_i + \frac{1}{m} \sum_{j \in C} \boldsymbol{x}_j + \boldsymbol{a}_i - \bar{\boldsymbol{a}} + \boldsymbol{\omega}_i - \frac{1}{m} \sum_{j \in C} \boldsymbol{\omega}_j - \sum_{j \in C \setminus \{i\}} \boldsymbol{g}_{ij} + \frac{1}{m} \sum_{j \in C} \sum_{k \notin C} (\boldsymbol{g}_{jk} - \boldsymbol{g}_{ik}) = \boldsymbol{0}, \quad \forall i \in C,$$

which is rearranged to

$$\begin{aligned}\boldsymbol{a}_i - \bar{\boldsymbol{a}} &= \boldsymbol{x}_i - \frac{1}{m}\sum_{j\in C}\boldsymbol{x}_j - \boldsymbol{\omega}_i + \frac{1}{m}\sum_{j\in C}\boldsymbol{\omega}_j + \sum_{j\in C\setminus\{i\}}\boldsymbol{g}_{ij} + \frac{1}{m}\sum_{j\in C}\sum_{k\notin C}(\boldsymbol{g}_{ik} - \boldsymbol{g}_{jk}) \\ &= \sum_{j\in C\setminus\{i\}}\left[\frac{1}{m}(\boldsymbol{x}_i - \boldsymbol{x}_j - \boldsymbol{\omega}_i + \boldsymbol{\omega}_j) + \boldsymbol{g}_{ij} + \frac{1}{m}\sum_{k\notin C}(\boldsymbol{g}_{ik} - \boldsymbol{g}_{jk})\right] \\ &= \sum_{j\in C\setminus\{i\}}\boldsymbol{q}_{ij} \quad \text{(By definition)}, \quad \forall i \in C.\end{aligned}$$

Moreover, by the definition of $\boldsymbol{q}_{ij}$, we observe the following property for all $i, j \in C, i \neq j$

$$\boldsymbol{q}_{ij} = \boldsymbol{g}_{ij} + \frac{1}{m}\cdot(\boldsymbol{x}_i - \boldsymbol{x}_j - \boldsymbol{\omega}_i + \boldsymbol{\omega}_j) + \frac{1}{m}\sum_{k\notin C}(\boldsymbol{g}_{ik} - \boldsymbol{g}_{jk}) = -\boldsymbol{g}_{ji} - \frac{1}{m}\cdot(\boldsymbol{x}_j - \boldsymbol{x}_i - \boldsymbol{\omega}_j + \boldsymbol{\omega}_i) - \frac{1}{m}\sum_{k\notin C}(\boldsymbol{g}_{jk} - \boldsymbol{g}_{ik}) = -\boldsymbol{q}_{ji}$$

□

Corollary 5.1. *If $\|\boldsymbol{q}_{ij}\| \leq \lambda$ holds for all $i \neq j, i, j \in C$ where C is a candidate cluster, then C is a cluster or part of a larger cluster.*

The proof of the corollary follows trivially by Theorem 2.

4.2 Duality gap and distinct clustering corollary

As derived earlier in Section 3.2, the duality gap at the feasible solution is:

$$f(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, s, u, t) - h(\boldsymbol{\delta}, \boldsymbol{\beta}, \gamma) = \sum_{i=1}^{n} s_i + \lambda\sum_{1\leq i<j\leq n} t_{ij} - \sum_{i=1}^{n}\boldsymbol{a}_i^T\boldsymbol{\beta}_i - \sum_{i=1}^{n}\gamma_i = \sum_{i=1}^{n}\sigma_1^i + \sum_{1\leq i<j\leq n}\epsilon_1^{ij} =: \mu. \tag{21}$$

By the property of strong convexity of f', we have

$$\frac{1}{2}\|\boldsymbol{x} - \boldsymbol{x}^*\|^2 \leq f'(\boldsymbol{x}) - f'(\boldsymbol{x}^*) = f(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, s, u, t) - f(\boldsymbol{x}^*, \boldsymbol{y}^*, \boldsymbol{z}^*, s^*, u^*, t^*),$$

which is further bounded as follows by weak duality

$$\frac{1}{2}\|\boldsymbol{x} - \boldsymbol{x}^*\|^2 \leq f(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, s, u, t) - h(\boldsymbol{\delta}, \boldsymbol{\beta}, \gamma). \tag{22}$$

Then, for any primal-dual algorithm, the distance between the approximate solution and the optimizer is given by

$$\tau = \|\boldsymbol{x} - \boldsymbol{x}^*\| \leq \sqrt{2\mu} \tag{23}$$

Corollary 5.2. *Let C denote a candidate cluster. If for all $i \in C, j \notin C$ it holds that $\|\boldsymbol{x}_i - \boldsymbol{x}_j\| > 2\tau$ where $\tau = \sqrt{2\mu}$ for any primal-dual algorithm, then there does not exist a super-cluster which strictly contains C.*

The proof follows directly from Theorem 3.

5 Properties of the central path

In this section, we explore the properties of the central path for a symmetric primal-dual path following algorithm. These properties play a fundamental role in the proof of our main theorem in Section 6. In the main theorem, we state that if a symmetric primal-dual path following algorithm is used, our clustering test will eventually succeed after a finite number of iterations when λ is not at any fusion value. The proof of the ultimate success relies on the linear convergence to the optimal primal-dual pair, which will be shown to be satisfied in the remainder of this section.

Even though there are very few theorems about the central path of second-order-cone programming in literature, there are established theorems from semidefinite programming, which specialize in SOCP. The following theorem states that the μ'-centered iterates converge to the analytic center superlinearly.

Theorem 6 (Luo et al [9]). *Assume the semidefinite program has a strictly complementary solution and the iterates of the algorithm converge tangentially to the central path. Let $(X(\mu'), Z(\mu'))$ denote a μ'-centered primal-dual pair. Let (X^a, Z^a) denote the analytic centers of the primal and dual optimal sets. Let $\mu' \in (0,1)$ be the central path parameter. There holds*

$$\|X(\mu') - X^a\| = O(\mu'), \quad \|Z(\mu') - Z^a\| = O(\mu').$$

Assume a primal-dual path following algorithm satisfying the assumptions of Luo et al. is applied. To employ Theorem 6, we show that our SOCP has a strictly complementary optimizer. Notice that this statement is not necessarily true for all values of λ. One has to assume λ is not at any fusion value. When λ is exactly at a fusion value λ^*, strict complementarity may fail. The failure is not surprising since any arbitrarily small negative perturbation $\lambda^*+\epsilon$ yields a different clustering. In other words, complete cluster identification for these fusion values is ill-posed. Thus it is unreasonable to expect an algorithm which satisfies a guarantee for such a problem. There are at most n fusion values as a result of Theorem 1.

5.1 Strict complementarity

By specializing the definition of strict complementarity in SDP to SOCP [1], a primal and dual feasible solution satisfies strict complementarity if and only if

$$t_{ij} + \lambda > \|\boldsymbol{y}_{ij} + \boldsymbol{\delta}_{ij}\|, \quad \forall 1 \le i < j \le n, \tag{24}$$

$$s_i + 1 - \gamma_i > \left\| \begin{pmatrix} \boldsymbol{z}_i + \boldsymbol{\beta}_i \\ u_i + \gamma_i \end{pmatrix} \right\|, \quad \forall i = 1, ..., n \tag{25}$$

Let $\lambda > 0$ be a parameter value for which fusion does not occur and let λ_1, λ_2 be the two successive fusion values such that $\lambda \in (\lambda_1, \lambda_2)$. Note that it is possible for $\lambda_1 = 0$ or $\lambda_2 = \infty$. We will show there exists a strict complementary primal-dual solution at λ.

Let $(\boldsymbol{x}', \boldsymbol{y}', \boldsymbol{z}', s', u', t', \boldsymbol{\delta}', \boldsymbol{\beta}', \gamma')$ denote optimal primal and dual solutions at λ_1. Let $C_1, C_2, ..., C_K$ denote the clusters identified by the optimal solutions above. When $\lambda_1 = 0$, there are n clusters, and each cluster is a singleton set. When λ_1 is the largest fusion value, there is only one cluster containing all n points. Define $\boldsymbol{g}'$ with $\boldsymbol{\delta}'$ as before

$$\boldsymbol{g}'_{ij} := \begin{cases} -\boldsymbol{\delta}'_{ij}, & \text{if } i < j, \\ \boldsymbol{\delta}'_{ji}, & \text{if } j < i. \end{cases}$$

By Chiquet et al. [5], the dual solutions satisfy

$$\boldsymbol{a}_i - \bar{\boldsymbol{a}}_k = \sum_{j \in C_k - \{i\}} \boldsymbol{g}'_{ij}, \quad \forall i \in C_k, k \in [K], \qquad \text{and} \qquad \left\|\boldsymbol{g}'_{ij}\right\| = \left\|\boldsymbol{\delta}'_{ij}\right\| \leq \lambda_1, \quad \forall i \neq j \tag{26}$$

where $\bar{\boldsymbol{a}}_k := \frac{1}{|C_k|} \sum_{i \in C_k} \boldsymbol{a}_i$. Consider the following optimization problem

$$\min_{\boldsymbol{p}_1, ..., \boldsymbol{p}_K \in \mathbf{R}^d} \frac{1}{2} \sum_{k=1}^{K} |C_k| \left\|\boldsymbol{p}_k - \bar{\boldsymbol{a}}_k\right\|^2 + \lambda \sum_{1 \leq k < k' \leq K} |C_k| \cdot |C_{k'}| \left\|\boldsymbol{p}_k - \boldsymbol{p}_{k'}\right\|. \tag{27}$$

Let $\boldsymbol{p}$ denote the optimal solution of (27).

Lemma 7. *Vector $\boldsymbol{p}$ satisfies $\boldsymbol{p}_k \neq \boldsymbol{p}_{k'}$ for all $k, k' \in [K], k \neq k'$.*

Proof. For the purpose of contradiction, we may assume there exist $\hat{k} \neq \hat{k}'$ such that $\boldsymbol{p}_{\hat{k}} = \boldsymbol{p}_{\hat{k}'}$. Let $\boldsymbol{x}^*_i = \boldsymbol{p}_k, \forall i \in C_k, k \in [K]$. By the first-order optimality condition of (27) at $\boldsymbol{p}$, there exist $\boldsymbol{g}_{kk'} \in \partial \|\boldsymbol{p}_k - \boldsymbol{p}_{k'}\|$ for all $k \neq k'$ such that

$$\begin{aligned} \mathbf{0} &= \boldsymbol{p}_k - \bar{\boldsymbol{a}}_k + \lambda \cdot |C_{K'}| \cdot \sum_{k' \neq k} \boldsymbol{g}_{kk'} \\ &= \boldsymbol{p}_k - \boldsymbol{a}_i + \boldsymbol{a}_i - \bar{\boldsymbol{a}}_k + \lambda \cdot |C_{K'}| \cdot \sum_{k' \neq k} \boldsymbol{g}_{kk'} \\ &= \boldsymbol{p}_k - \boldsymbol{a}_i + \sum_{j \in C_k - \{i\}} \boldsymbol{g}'_{ij} + \lambda \cdot |C_{K'}| \cdot \sum_{k' \neq k} \boldsymbol{g}_{kk'} \\ &= \boldsymbol{x}^*_i - \boldsymbol{a}_i + \sum_{j \in C_k - \{i\}} \boldsymbol{g}'_{ij} + \lambda \cdot |C_{K'}| \cdot \sum_{k' \neq k} \boldsymbol{g}_{kk'}, \end{aligned}$$

satisfying (2) at i. As $i \in C_k, k \in [K]$ are chosen arbitrarily, the equality (2) holds for all i hence $\boldsymbol{x}^*$ is an optimal solution to (1). From the assumption on $\boldsymbol{p}$ and the agglomerative properties of the clusterpath, cluster $C_{\hat{k}}, C_{\hat{k}'}$ merge at $\lambda' \in (\lambda_1, \lambda]$, which contradicts our choice of λ_2. That concludes our proof. □

By Lemma 7, the objective function is differentiable at $\boldsymbol{p}$. Hence, there holds

$$|C_k|(\boldsymbol{p}_k - \bar{\boldsymbol{a}}_k) + \lambda |C_k| \cdot |C_{k'}| \frac{\boldsymbol{p}_k - \boldsymbol{p}_{k'}}{\|\boldsymbol{p}_k - \boldsymbol{p}_{k'}\|} = \mathbf{0}, \quad \forall k \in [K]. \tag{28}$$

Define the following solutions:

$$
\begin{aligned}
\boldsymbol{x}_i^* &= \boldsymbol{p}_k, \quad \forall i \in C_k, k \in [K] \\
\boldsymbol{y}_{ij}^* &= \boldsymbol{x}_i^* - \boldsymbol{x}_j^*, \quad \forall 1 \leq i < j \leq n \\
\boldsymbol{z}_i^* &= \boldsymbol{x}_i^* - \boldsymbol{a}_i, \quad \forall i = 1, \ldots, n, \\
s_i^* &= \frac{1}{2}(1 + \|\boldsymbol{z}_i^*\|^2), \quad \forall i = 1, \ldots, n \\
u_i^* &= \frac{1}{2}(-1 + \|\boldsymbol{z}_i^*\|^2), \quad \forall i = 1, \ldots, n \\
t_{ij}^* &= \|\boldsymbol{y}_{ij}^*\|, \quad \forall 1 \leq i < j \leq n \\
\boldsymbol{\delta}_{ij}^* &= \begin{cases} \boldsymbol{\delta}'_{ij}, & \text{if } i < j \text{ and } i, j \in C_k \\ \lambda \frac{\boldsymbol{x}_j^* - \boldsymbol{x}_i^*}{\|\boldsymbol{x}_j^* - \boldsymbol{x}_i^*\|}, & \text{otherwise} \end{cases} \qquad \forall 1 \leq i < j \leq n \\
\boldsymbol{\beta}_i^* &= -\boldsymbol{z}_i^*, \quad \forall i = 1, \ldots, n \\
\gamma_i^* &= \frac{1}{2}(1 - \|\boldsymbol{\beta}_i^*\|^2), \quad \forall i = 1, \ldots, n
\end{aligned} \tag{29}
$$

Lemma 8. *The solutions defined above are optimal for the second-order cone program at* λ*.*

Proof. By construction, the primal constraints (4b), (4c), (4d), (4e), (4f), the dual constraints (5c), (5d), and the complementary slackness conditions (9), (10), (11), (12) and (13) with $\boldsymbol{\epsilon} = \mathbf{0}, \boldsymbol{\sigma} = \mathbf{0}$ are automatically satisfied. It remains to check if these solutions satisfy (5b).

Verification for (5b)**:** For any $i \in C_k$ with some $k \in [K]$, (5b) is rewritten as follows due to (26) and (28)

$$
\begin{aligned}
& - \sum_{j=1}^{i-1} \boldsymbol{\delta}_{ji}^* + \sum_{j=i+1}^{n} \boldsymbol{\delta}_{ij}^* + \boldsymbol{\beta}_i^* \\
= & - \sum_{j<i, j \in C_k - \{i\}} \boldsymbol{\delta}'_{ji} + \sum_{j>i, j \in C_k - \{i\}} \boldsymbol{\delta}'_{ij} + \lambda \sum_{k \neq k'} |C_{k'}| \frac{\boldsymbol{p}_{k'} - \boldsymbol{p}_k}{\|\boldsymbol{p}_{k'} - \boldsymbol{p}_k\|} + \boldsymbol{a}_i - \boldsymbol{x}_i^* \\
= & - \sum_{j \in C_k - \{i\}} \boldsymbol{g}'_{ij} + \lambda \sum_{k \neq k'} |C_{k'}| \frac{\boldsymbol{p}_{k'} - \boldsymbol{p}_k}{\|\boldsymbol{p}_{k'} - \boldsymbol{p}_k\|} + \boldsymbol{a}_i - \boldsymbol{x}_i^* \\
= & \bar{\boldsymbol{a}}_k - \boldsymbol{a}_i + \lambda \sum_{k \neq k'} |C_{k'}| \frac{\boldsymbol{p}_{k'} - \boldsymbol{p}_k}{\|\boldsymbol{p}_{k'} - \boldsymbol{p}_k\|} + \boldsymbol{a}_i - \boldsymbol{p}_i \\
= & \bar{\boldsymbol{a}}_k + \lambda \sum_{k \neq k'} |C_{k'}| \frac{\boldsymbol{p}_{k'} - \boldsymbol{p}_k}{\|\boldsymbol{p}_{k'} - \boldsymbol{p}_k\|} - \boldsymbol{p}_i \\
= & \mathbf{0}.
\end{aligned}
$$

By KKT conditions, the solutions defined above form an optimal primal-dual pair.

□

Lemma 9. *The solutions defined above are strictly complementary.*

Proof. The strict complementarity is equivalent to (24) and (25), which can be easily checked as shown below

Verification for (24)**:** Let $1 \leq i < j \leq n$. If $\boldsymbol{y}_{ij}^* = \boldsymbol{0}$, then there exists some $k \in [K]$ such that $i, j \in C_k$. By definition, $t_{ij}^* = 0$ and $\boldsymbol{\delta}_{ij}^* = \boldsymbol{\delta}_{ij}$. Notice that $\boldsymbol{\delta}_{ij}$ is the optimal dual solution of (1) at λ_1, then it satisfies $\|\boldsymbol{\delta}_{ij}\| \leq \lambda_1 < \lambda$ by the definition of λ. Hence,

$$t_{ij}^* + \lambda = \lambda > \|\boldsymbol{\delta}_{ij}\| = \|\boldsymbol{\delta}_{ij}^*\| = \|\boldsymbol{y}_{ij}^* + \boldsymbol{\delta}_{ij}^*\|.$$

If $\boldsymbol{y}_{ij}^* \neq \boldsymbol{0}$, then there exist $k, k' \in [K]$ such that $i \in C_k, j \in C_{k'}$ and $k \neq k'$. By definition, $t_{ij}^* = \|\boldsymbol{y}_{ij}^*\| = \|\boldsymbol{p}_k - \boldsymbol{p}_{k'}\|$ and $\boldsymbol{\delta}_{ij}^* = \lambda \frac{\boldsymbol{p}_{k'} - \boldsymbol{p}_k}{\|\boldsymbol{p}_{k'} - \boldsymbol{p}_k\|}$. Hence,

$$t_{ij}^* + \lambda = \|\boldsymbol{p}_k - \boldsymbol{p}_{k'}\| + \lambda > |\|\boldsymbol{p}_k - \boldsymbol{p}_{k'}\| - \lambda| = \left\| \boldsymbol{p}_k - \boldsymbol{p}_{k'} + \lambda \frac{\boldsymbol{p}_{k'} - \boldsymbol{p}_k}{\|\boldsymbol{p}_{k'} - \boldsymbol{p}_k\|} \right\| = \left\| \boldsymbol{y}_{ij}^* + \boldsymbol{\delta}_{ij}^* \right\|.$$

Verification for (25)**:** Let $i \in [n]$. By construction,

$$s_i^* + 1 - \gamma_i^* = \|\boldsymbol{z}_i^*\|^2 + 1 > 0 = \left\| \begin{pmatrix} \boldsymbol{z}_i^* - \boldsymbol{z}_i^* \\ -\frac{1}{2}(1 - \|\boldsymbol{z}_i\|^2) + \frac{1}{2}(1 - \|\boldsymbol{z}_i\|^2) \end{pmatrix} \right\| = \left\| \begin{pmatrix} \boldsymbol{z}_i^* + \boldsymbol{\beta}_i^* \\ u_i^* + \gamma_i^* \end{pmatrix} \right\|$$

Since the indices are chosen arbitrarily, the solutions defined above are strictly complementary. □

6 Test Guarantee

In Section 4, we validated our test theoretically in the sense that if the test succeeds, it is guaranteed that the correct clusters are found. In this section, we show that the test succeeds after a finite number of iterations of a certain interior point method, provided that λ is not at any fusion value. Specifically, we prove that the two conditions in our test are guaranteed to hold for a primal-dual path following algorithm satisfying the assumptions of Luo et al. [9] when the duality gap μ is sufficiently small.

Theorem 10. *If λ is not a fusion value, then there exists $\mu_0 > 0$ such that both CGR subgradients and separation conditions in the test are satisfied for any duality gap $\mu \leq \mu_0$ for a primal-dual path following algorithm satisfying the assumptions of Luo et al. [9].*

Let $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, s, u, t, \boldsymbol{\delta}, \boldsymbol{\beta}, \gamma)$ denote a primal and dual feasible solution. Let $C_1, C_2, ..., C_K$ denote the clusters obtained at optimum. Let $\mu' \in (0, 1)$ denote the central path parameter and let μ denote the duality gap at the feasible solution. By Theorem 6, there hold

$$\|\boldsymbol{x}(\mu') - \boldsymbol{x}^a\| = O(\mu'), \quad \|\boldsymbol{\delta}(\mu) - \boldsymbol{\delta}^a\| = O(\mu')$$

where $\boldsymbol{x}(\mu'), \boldsymbol{\delta}(\mu')$ are μ'-centered solutions and $\boldsymbol{x}^a, \boldsymbol{\delta}^a$ are the analytic centers of the primal and dual optimal sets respectively. Moreover, since the iterates converge tangentially to the central path, we may assume the size of the central path neighborhood to be as follows

$$\|\boldsymbol{x} - \boldsymbol{x}(\mu')\| = O(\mu'), \quad \|\boldsymbol{\delta} - \boldsymbol{\delta}(\mu')\| = O(\mu').$$

Luo et al. [9] validated the assumption above for their interior point algorithm, which is a generalization of the Mizuno-Todd-Ye predictor-corrector method for linear programming. Combine the two sets of equations above and employ the triangle inequality to obtain

$$\|\boldsymbol{x} - \boldsymbol{x}^a\| = O(\mu'), \quad \|\boldsymbol{\delta} - \boldsymbol{\delta}^a\| = O(\mu').$$

As the duality gap μ is of linear order of the central path parameter μ', the equalities above are rewritten as

$$\|\boldsymbol{x} - \boldsymbol{x}^a\| = O(\mu), \quad \|\boldsymbol{\delta} - \boldsymbol{\delta}^a\| = O(\mu).$$

Define $p, p' \geq 0$ such that $\|\boldsymbol{x}_i - \boldsymbol{x}_i^a\| \leq p\mu$ for all i and $\left\|\boldsymbol{\delta}_{ij} - \boldsymbol{\delta}_{ij}^a\right\| \leq p'\mu$ for all distinct pairs (i, j). Then, for all distinct pairs (i, j) in any cluster C_k, there holds $\|\boldsymbol{x}_i - \boldsymbol{x}_j\| \leq 2p\mu$. Moreover, define $q > 0$ such that all $\boldsymbol{x}_i^a$'s in different clusters are at least q apart, which implies that $\boldsymbol{x}_i$'s in different clusters are separated by a distance of at least $q - 2p\mu$. We may assume the duality gap satisfies $\mu < \frac{q}{2p}$. Notice that this assumption is guaranteed to be true after a finite number of iterations.

Let $C := C_k$ for some $k \in [K]$. By Lemma 4, there hold $\left\|\boldsymbol{\epsilon}_2^{ij}\right\| \leq \sqrt{\sum_{l=1}^n \|\bar{\boldsymbol{a}} - \boldsymbol{a}_l\|^2 \mu + 2\mu^2}$ for all $i < j, C \cap \{i, j\} \neq \emptyset, [n]\backslash C \cap \{i, j\} \neq \emptyset$ and

$$\left\|\begin{pmatrix} \boldsymbol{\sigma}_2^i \\ \sigma_3^i \end{pmatrix}\right\| \leq \sqrt{\left(\sum_{l=1}^n \|\bar{\boldsymbol{a}} - \boldsymbol{a}_l\|^2 + 1 + 2\mu\right) \cdot \left(\frac{1}{2} + \sum_{l=1}^n (n-1)\lambda \|\boldsymbol{a}_l\| + \mu\right) \mu}$$

for all i.

6.1 Bound $\|\boldsymbol{\delta}_{ij}\|$

Lemma 11. *For all $i, j \in C, i \neq j$, the following inequality holds*

$$\|\boldsymbol{\delta}_{ij}\| \leq \lambda - r + p'\mu$$

where $r := \min_{l \neq l', l, l' \in C_k, k \in [K]} (\lambda - \|\boldsymbol{\delta}_{ll'}^a\|) > 0$.

Proof. Let $i, j \in C$ and $i \neq j$. By the definition of analytic center and strict complementarity,

$$\|\boldsymbol{\delta}_{ll'}^a\| < \lambda,$$

holds for all $l \neq l', l, l' \in C_k, k \in [K]$. Hence, $r > 0$ by definition. Moreover, r also satisfies

$$\left\|\boldsymbol{\delta}_{ij}^a\right\| \leq \lambda - r, \quad \forall i, j \in C, i \neq j.$$

Since $\left\|\boldsymbol{\delta}_{ij} - \boldsymbol{\delta}_{ij}^a\right\| \leq p'\mu$, we obtain

$$\|\boldsymbol{\delta}_{ij}\| \leq \lambda - r + p'\mu, \quad \forall i, j \in C, i \neq j.$$

□

6.2 Bound $\|\boldsymbol{g}_{ik} - \boldsymbol{g}_{jk}\|$

Lemma 12. *For all $i, j \in C$ and $k \notin C$, the following inequality holds*

$$\|\boldsymbol{g}_{ik} - \boldsymbol{g}_{jk}\| \leq \frac{4\lambda p\mu}{q-2p\mu} + \frac{2\sqrt{\sum_{l=1}^{n} \|\bar{\boldsymbol{a}} - \boldsymbol{a}_l\|^2 \mu + 2\mu^2}}{q-2p\mu} + \frac{\mu}{q-2p\mu}$$

Proof. Let $i, j \in C$ and $k \notin C$. Without loss of generality, we may assume $i < j < k$. Hence, $\boldsymbol{g}_{ik} = -\boldsymbol{\delta}_{ik}, \boldsymbol{g}_{jk} = -\boldsymbol{\delta}_{jk}$. By (10), we derive

$$t_{ik}\boldsymbol{\delta}_{ik} - t_{jk}\boldsymbol{\delta}_{jk} = -\lambda \boldsymbol{y}_{ik} + \lambda \boldsymbol{y}_{jk} + \boldsymbol{\epsilon}_2^{ik} - \boldsymbol{\epsilon}_2^{jk} = -\lambda(\boldsymbol{x}_i - \boldsymbol{x}_j) + \boldsymbol{\epsilon}_2^{ik} - \boldsymbol{\epsilon}_2^{jk}. \tag{30}$$

Adding the term $(t_{jk} - t_{ik})\boldsymbol{\delta}_{jk}$ to both sides of the equality to obtain

$$t_{ik}(\boldsymbol{\delta}_{ik} - \boldsymbol{\delta}_{jk}) = (t_{jk} - t_{ik})\boldsymbol{\delta}_{jk} - \lambda(\boldsymbol{x}_i - \boldsymbol{x}_j) + \boldsymbol{\epsilon}_2^{ik} - \boldsymbol{\epsilon}_2^{jk}.$$

Notice that $t_{ik} \geq \|\boldsymbol{y}_{ik}\| = \|\boldsymbol{x}_i - \boldsymbol{x}_k\| \geq q - 2p\mu > 0$ by the primal constraint (4e) and our assumption on the duality gap. Divide the equality above by t_{ik} to obtain

$$\boldsymbol{\delta}_{ik} - \boldsymbol{\delta}_{jk} = \frac{t_{jk} - t_{ik}}{t_{ik}}\boldsymbol{\delta}_{jk} - \frac{\lambda(\boldsymbol{x}_i - \boldsymbol{x}_j)}{t_{ik}} + \frac{\boldsymbol{\epsilon}_2^{ik} - \boldsymbol{\epsilon}_2^{jk}}{t_{ik}}.$$

Substitute the definition of $\boldsymbol{g}$ into the equality above to obtain

$$\boldsymbol{g}_{ik} - \boldsymbol{g}_{jk} = \frac{t_{ik} - t_{jk}}{t_{ik}}\boldsymbol{\delta}_{jk} + \frac{\lambda(\boldsymbol{x}_i - \boldsymbol{x}_j)}{t_{ik}} - \frac{\boldsymbol{\epsilon}_2^{ik} - \boldsymbol{\epsilon}_2^{jk}}{t_{ik}}. \tag{31}$$

By the perturbed complementary slackness (9), the primal constraint (4e) and the Cauchy-Schwarz inequality, we derive the following inequality

$$\epsilon_1^{ik} = t_{ik}\lambda + \boldsymbol{y}_{ik}^T\boldsymbol{\delta}_{ik} \geq t_{ik}\lambda - \|\boldsymbol{y}_{ik}\| \cdot \|\boldsymbol{\delta}_{ik}\| \geq t_{ik}\lambda - \|\boldsymbol{y}_{ik}\| \cdot \lambda,$$

which yields an upper bound on t_{ik}

$$t_{ik} \leq \|\boldsymbol{y}_{ik}\| + \frac{\epsilon_1^{ik}}{\lambda}.$$

Combined with the primal constraint (4e) at t_{jk} and the triangle inequality, we obtain the following

$$t_{ik} - t_{jk} \leq \|\boldsymbol{y}_{ik}\| + \frac{\epsilon_1^{ik}}{\lambda} - \|\boldsymbol{y}_{jk}\| \leq \|\boldsymbol{y}_{ik} - \boldsymbol{y}_{jk}\| + \frac{\epsilon_1^{ik}}{\lambda} = \|\boldsymbol{x}_i - \boldsymbol{x}_j\| + \frac{\epsilon_1^{ik}}{\lambda}. \tag{32}$$

The same inequality holds for $t_{jk} - t_{ik}$ due to the symmetry of (32). By (31), (32) and triangle inequality, the norm bound of $\boldsymbol{g}_{ik} - \boldsymbol{g}_{jk}$ is as follows

$$\begin{aligned}
\|\boldsymbol{g}_{ik} - \boldsymbol{g}_{jk}\| &\leq \frac{|t_{ik} - t_{jk}| \cdot \|\boldsymbol{\delta}_{jk}\|}{t_{ik}} + \frac{\lambda \|\boldsymbol{x}_i - \boldsymbol{x}_j\|}{t_{ik}} + \frac{\left\|\boldsymbol{\epsilon}_2^{ik}\right\| + \left\|\boldsymbol{\epsilon}_2^{jk}\right\|}{t_{ik}} \quad \text{(By triangle inequality)} \\
&\leq \frac{\|\boldsymbol{x}_i - \boldsymbol{x}_j\| + \frac{\epsilon_1^{ik}}{\lambda}}{t_{ik}} \|\boldsymbol{\delta}_{jk}\| + \frac{\lambda \|\boldsymbol{x}_i - \boldsymbol{x}_j\|}{t_{ik}} + \frac{\left\|\boldsymbol{\epsilon}_2^{ik}\right\| + \left\|\boldsymbol{\epsilon}_2^{jk}\right\|}{t_{ik}} \quad \text{(By (32))} \\
&\leq \frac{2\lambda \|\boldsymbol{x}_i - \boldsymbol{x}_j\|}{t_{ik}} + \frac{\left\|\boldsymbol{\epsilon}_2^{ik}\right\| + \left\|\boldsymbol{\epsilon}_2^{jk}\right\|}{t_{ik}} + \frac{\epsilon_1^{ik}}{t_{ik}} \quad \text{(By (4e) and (30))}.
\end{aligned}$$

Since $i, j \in C$ and $k \notin C$, there hold $t_{ik} \geq \|\boldsymbol{y}_{ik}\| = \|\boldsymbol{x}_i - \boldsymbol{x}_k\| \geq q - 2p\mu$ and $\|\boldsymbol{x}_i - \boldsymbol{x}_j\| \leq 2p\mu$. Moreover, there also hold $\epsilon_1^{ik} \leq \mu$, $\left\|\boldsymbol{\epsilon}_2^{ik}\right\| \leq \sqrt{\sum_{l=1}^n \|\bar{\boldsymbol{a}} - \boldsymbol{a}_l\|^2 \mu + 2\mu^2}$ and $\left\|\boldsymbol{\epsilon}_2^{jk}\right\| \leq \sqrt{\sum_{l=1}^n \|\bar{\boldsymbol{a}} - \boldsymbol{a}_l\|^2 \mu + 2\mu^2}$. Hence, $\|\boldsymbol{g}_{ik} - \boldsymbol{g}_{jk}\|$ is further upper bounded as follows

$$\|\boldsymbol{g}_{ik} - \boldsymbol{g}_{jk}\| \leq \frac{4\lambda p\mu}{q - 2p\mu} + \frac{2\sqrt{\sum_{l=1}^n \|\bar{\boldsymbol{a}} - \boldsymbol{a}_l\|^2 \mu + 2\mu^2}}{q - 2p\mu} + \frac{\mu}{q - 2p\mu} \tag{33}$$

□

6.3 Bound $\|\boldsymbol{\omega}_i\|$

Lemma 13. *For all $i \in C$, it holds*

$$\|\boldsymbol{\omega}_i\| \leq 2\sqrt{2\left(\sum_{l=1}^n \|\bar{\boldsymbol{a}} - \boldsymbol{a}_l\|^2 + 1 + 2\mu\right) \cdot \left(\frac{1}{2} + \sum_{l=1}^n (n-1)\lambda \|\boldsymbol{a}_l\| + \mu\right)\mu}.$$

Proof. Let $i \in C$. By definition,

$$\boldsymbol{\omega}_i = \frac{\sigma_3^i}{s_i}\boldsymbol{z}_i + \frac{1}{s_i}\boldsymbol{\sigma}_2^i.$$

By the primal constraint (4f), we have

$$\|\boldsymbol{z}_i\| \leq \sqrt{2s_i - 1} \leq \sqrt{2s_i}, \quad s_i \geq \frac{1}{2},$$

which implies

$$\frac{\|\boldsymbol{z}_i\|}{s_i} \leq \sqrt{\frac{2}{s_i}} \leq \sqrt{4} = 2, \quad \frac{1}{s_i} \leq 2.$$

Coupled with triangle inequality, these two inequalities yield

$$\|\boldsymbol{\omega}_i\| \leq \frac{\|\boldsymbol{z}_i\|}{s_i}\sigma_3^i + \frac{1}{s_i}\left\|\boldsymbol{\sigma}_2^i\right\| \leq 2\sigma_3^i + 2\left\|\boldsymbol{\sigma}_2^i\right\|.$$

Moreover, since $\left\|\begin{pmatrix}\boldsymbol{\sigma}_2^i \\ \sigma_3^i\end{pmatrix}\right\| \leq \sqrt{\left(\sum_{l=1}^n \|\bar{\boldsymbol{a}} - \boldsymbol{a}_l\|^2 + 1 + 2\mu\right) \cdot \left(\frac{1}{2} + \sum_{l=1}^n (n-1)\lambda \|\boldsymbol{a}_l\| + \mu\right)\mu}$ holds for any $i \in [n]$ by Lemma 4 and the duality gap,

$$(\sigma_3^i)^2 + \left\|\boldsymbol{\sigma}_2^i\right\|^2 \leq \left(\sum_{l=1}^n \|\bar{\boldsymbol{a}} - \boldsymbol{a}_l\|^2 + 1 + 2\mu\right) \cdot \left(\frac{1}{2} + \sum_{l=1}^n (n-1)\lambda \|\boldsymbol{a}_l\| + \mu\right)\mu.$$

which implies the following inequality since $(a+b)^2 \leq 2a^2 + 2b^2$

$$(\sigma_3^i + \left\|\boldsymbol{\sigma}_2^i\right\|)^2 \leq 2 \cdot \left(\sum_{l=1}^n \|\bar{\boldsymbol{a}} - \boldsymbol{a}_l\|^2 + 1 + 2\mu\right) \cdot \left(\frac{1}{2} + \sum_{l=1}^n (n-1)\lambda \|\boldsymbol{a}_l\| + \mu\right)\mu.$$

Therefore, the following holds as $i \in C$ is chosen arbitrarily:

$$\|\boldsymbol{\omega}_i\| \leq 2\sigma_3^i + 2\left\|\boldsymbol{\sigma}_2^i\right\| \leq 2\sqrt{2 \cdot \left(\sum_{l=1}^n \|\bar{\boldsymbol{a}} - \boldsymbol{a}_l\|^2 + 1 + 2\mu\right) \cdot \left(\frac{1}{2} + \sum_{l=1}^n (n-1)\lambda \|\boldsymbol{a}_l\| + \mu\right)\mu}.$$

□

6.4 Bound the CGR subgradient

Lemma 14. *For all $i, j \in C$ and $i < j$, there holds*

$$\begin{aligned}\|\boldsymbol{q}_{ij}\| \leq& \lambda - r + p'\mu + \frac{1}{m} \cdot \left(2p\mu + 4\sqrt{2 \cdot \left(\sum_{l=1}^n \|\bar{\boldsymbol{a}} - \boldsymbol{a}_l\|^2 + 1 + 2\mu\right) \cdot \left(\frac{1}{2} + \sum_{l=1}^n (n-1)\lambda \|\boldsymbol{a}_l\| + \mu\right)\mu}\right) \\ &+ \frac{n-m}{m}\left(\frac{4\lambda p\mu}{q - 2p\mu} + \frac{2\sqrt{\sum_{l=1}^n \|\bar{\boldsymbol{a}} - \boldsymbol{a}_l\|^2 \mu + 2\mu^2}}{q - 2p\mu} + \frac{\mu}{q - 2p\mu}\right).\end{aligned}$$

Proof. Let $i, j \in C$ and $i < j$. By triangle inequality,

$$\|\boldsymbol{q}_{ij}\| \leq \|\boldsymbol{\delta}_{ij}\| + \frac{1}{m} \cdot (\|\boldsymbol{x}_i - \boldsymbol{x}_j\| + \|\boldsymbol{\omega}_i\| + \|\boldsymbol{\omega}_j\|) + \frac{1}{m}\sum_{k \notin C} \|\boldsymbol{g}_{ik} - \boldsymbol{g}_{jk}\|.$$

With the assumptions on the distance between points,

$$\|\boldsymbol{x}_i - \boldsymbol{x}_j\| \leq 2p\mu.$$

By Lemma 11, 12, Lemma 13 and the inequality above, we obtain

$$\begin{aligned}\|\boldsymbol{q}_{ij}\| \leq& \lambda - r + p'\mu + \frac{1}{m} \cdot \left(2p\mu + 4\sqrt{2 \cdot \left(\sum_{l=1}^n \|\bar{\boldsymbol{a}} - \boldsymbol{a}_l\|^2 + 1 + 2\mu\right) \cdot \left(\frac{1}{2} + \sum_{l=1}^n (n-1)\lambda \|\boldsymbol{a}_l\| + \mu\right)\mu}\right) \\ &+ \frac{n-m}{m}\left(\frac{4\lambda p\mu}{q - 2p\mu} + \frac{2\sqrt{\sum_{l=1}^n \|\bar{\boldsymbol{a}} - \boldsymbol{a}_l\|^2 \mu + 2\mu^2}}{q - 2p\mu} + \frac{\mu}{q - 2p\mu}\right).\end{aligned} \tag{34}$$

□

6.5 Proof of the main theorem

Proof. We rewrite (34) with $O(\cdot)$ notation to obtain the following inequality

$$\|\boldsymbol{q}_{ij}\| \leq \lambda - r + O(\sqrt{\mu}), \quad \forall i, j \in C_k, i \neq j, k \in [K],$$

since $C = C_k$ is an arbitrarily cluster. As $r > 0$ by Lemma 11, there exists $\mu_1 > 0$ such that for all $\mu \leq \mu_1$, $\|\boldsymbol{q}_{ij}\| \leq \lambda$ holds for all $i, j \in C_k, i \neq j, k \in [K]$. Here concludes the proof of the CGR subgradient condition.

Since $q > 0$, there exists $\mu_2 > 0$ such that $2\sqrt{2\mu_2} < q - 2p\mu_2$. Hence, for all $\mu \leq \mu_2$, all clusters are separated at distance of at least $2\sqrt{2\mu_2}$. Here concludes the proof of the second condition.

Let $\mu_0 = \min\{\mu_1, \mu_2\}$, then both CGR subgradient and separation conditions are satisfied for any $\mu \leq \mu_0$. □

7 Computational experiments

In this section, we conduct experiments in which a Chi-Lange ADMM solver [3] and our clustering test for sum-of-norms clustering are applied to a simulated dataset of two normally distributed half moons. We intend to answer the following questions: (1) How does the performance of our test depend on λ? and (2) How does the recovery of two half moons depend on λ?

Our algorithm is implemented in Julia [2]. It terminates if the clustering test succeeds, or if the maximum number of iterations is reached. In the algorithm, the code tests for clustering every t iterations of the ADMM solver. The value of t is taken to be 8 in our experiment. At the end of every t iterations, the solver yields a primal solution and a dual solution, from which our algorithm constructs a primal and dual feasible pair for the SOCP formulation by (29). With the feasible solutions, the algorithm then creates candidate clusters, computes duality gap and constructs CGR subgradients. The code checks for the CGR subgradient condition and separation condition. If both conditions hold, the clustering test reports 'success'. Otherwise, the code runs t more iterations of the ADMM solver and repeats the clustering test. The detailed algorithm is outlined as follows. Each iteration of the ADMM solver is of complexity $O(n^2 d)$.

Algorithm 1: Find clusters

$C \leftarrow \{1, \ldots, n\}$;
$k \leftarrow 1$;
while $C \neq \emptyset$ **do**
  Choose $i \in C$ arbitrarily;
  Create a cluster $R_k \leftarrow \{j : \|\boldsymbol{x}_i - \boldsymbol{x}_j\| \leq \mu^{3/4}\}$ (including i itself);
  Delete all these points in R_k from C;
  $k \leftarrow k + 1$;
Return candidate clusters $\{R_1, R_2, \ldots, R_{K'}\}$;

Algorithm 2: An ADMM algorithm with our clustering test

Result: Clustering assignment
Initialize $(\boldsymbol{x}, \boldsymbol{\delta})$;
while *clustering test fails or maximum number of iterations is not reached* **do**
  for $l = 1, 2, \ldots, t$ **do**
    ADMM updates by Chi and Lange [3];
  end
  Construct a feasible solution for SOCP by (29) from the current ADMM iterate;
  Compute the duality gap μ;
  Run Algorithm 1 to find clusters $\{R_1, R_2, \ldots, R_{K'}\}$;
  Compute CGR subgradients from dual variables for $\{R_1, R_2, \ldots, R_{K'}\}$;
  Check the CGR subgradient condition; Check that no two clusters are distance $\leq 2\sqrt{2\mu}$ of each other;
  Mark the clustering test 'success' if both conditions pass and mark it 'failure' otherwise;
end
Return recovered clusters $\{R_1, R_2, \ldots, R_{K'}\}$.

To assess the performance of recovery, we employ the Rand index by Rand [12]. Rand index is a measure which specifically evaluates the performance of clustering. It compares two clusterings $\{R_1, \ldots, R_{K'}\}$ and $\{V_1, \ldots, V_K\}$ in a pairwise manner. If a pair of data points are placed in the same cluster in both clusterings, or if a pair of data points are placed in different clusters in both clusterings, then this pair is called a similar assignment and it contributes to the measure of similarity between two clusterings. We define the following two sets of similar assignments on all distinct pairs of instances:

$$\begin{aligned}
S &:= \{(i,j) : 1 \leq i < j \leq n \text{ such that there exist } m, m' \text{ satisfying } i, j \in V_m \cap R_{m'}\}, \\
D &:= \{(i,j) : 1 \leq i < j \leq n \text{ such that } i \in V_{m_1}, j \in V_{m_2}, m_1 \neq m_2, \text{ and} \\
&\qquad\quad i \in R_{m'_1}, j \in R_{m'_2}, m'_1 \neq m'_2\}.
\end{aligned}$$

Then Rand index is defined as the fraction of all distinct pairs which are similar assignments:

$$R = \frac{|S| + |D|}{\binom{n}{2}},$$

where $|\cdot|$ denotes the cardinality function. The value of R ranges from 0 to 1. When $R = 0$, two clusterings are completely dissimilar. When $R = 1$, two clusterings are identical. A higher Rand index indicates a higher level of similarity. A random assignment to clusters in the case of equally sized clusters, $K = 2$ yields expected Rand index of 0.5.

The experiment is conducted on a simulated dataset of two normally distributed half moons with 500 instances. The angle of two half moons follows a Gaussian distribution with a mean of 0 and a standard deviation of $\frac{\pi}{6}$. A random noise which follows a two-dimensional Gaussian distribution with a mean of 0 and a standard deviation of 0.05 displaces the instances from the moons. Fifty linearly spaced values of λ are taken from the range $[10^{-8}, 0.00496]$. The range is determined empirically. Furthermore, the maximum number of iterations is chosen to be 50,000. It took approximately 15 hours total on an Intel Xeon processor single-threaded to complete the experiment.

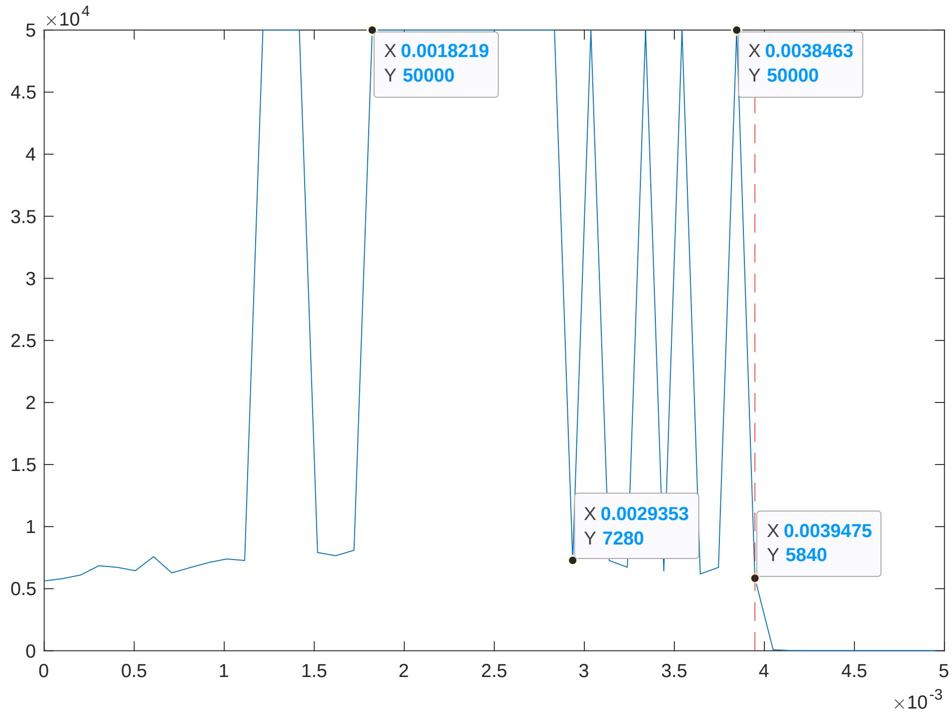

Figure 1: Iteration counts versus λ

Our first objective is to evaluate the performance of our clustering test. At 32 out of 50 values of λ, the clustering test succeeds before the maximum number of iterations is reached. When λ is in the range between $\lambda = 0.0018219$ and $\lambda = 0.0028341$, the algorithm repeatedly reaches the iteration threshold before the test succeeds as shown in Figure 1. The performance is interpretable with theories discussed earlier. The clustering test is not guaranteed to succeed when λ is at a fusion value, and the test performs poorly near a fusion value as shown in Figure 1. When $n = 500$, there are at most 500 fusion values. All fusion values are in the range between $\lambda = 0.00040$ and $\lambda = 0.00405$ as observed in the experiment. Hence, fusion occurs frequently, and massive fusion values are located densely in a small region. Thus, in our experiment, it is very likely that the λ we pick is near or at a fusion value, which leads to poor performance of our clustering test.

We anticipate that the clustering test improves with fewer data points, and it is indeed the case. The same experiment is also implemented for 200 instances generated from two normally distributed half moons with the same parameters. At 89 out of 100 values of λ, our clustering test succeeds before the maximum number of iterations is reached.

The experiment also attempts to explore the relationship between λ value and the recovery of half moons. To evaluates the recovery, we compute the Rand index with the recovered clustering and the generative clustering. The figure below shows Rand index against λ values. The value of Rand index increases monotonically and peaks at $\lambda = 0.00395$, where the clustering test succeeds and the Rand index achieves a value of 0.949.

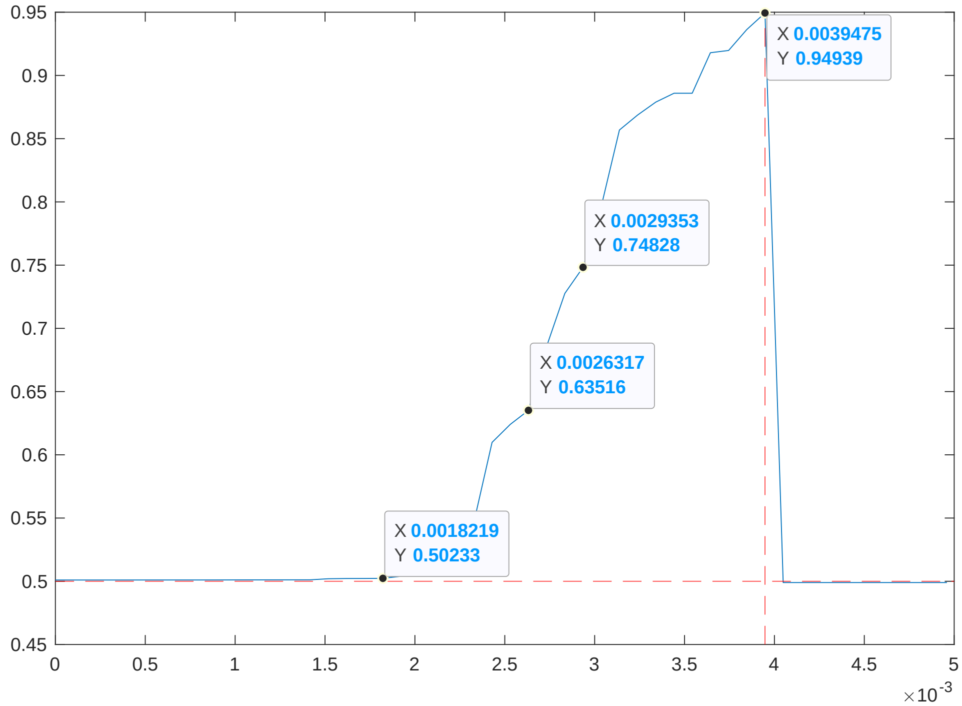

Figure 2: Rand index versus λ

To illustrate the clustering at $\lambda = 0.00395$, we also plot the two half moons and color the clusters. Red instances belong to one cluster, and blue instances belong to another cluster. Yellow instances are assigned to clusters of singleton points, and they are identified as noises.

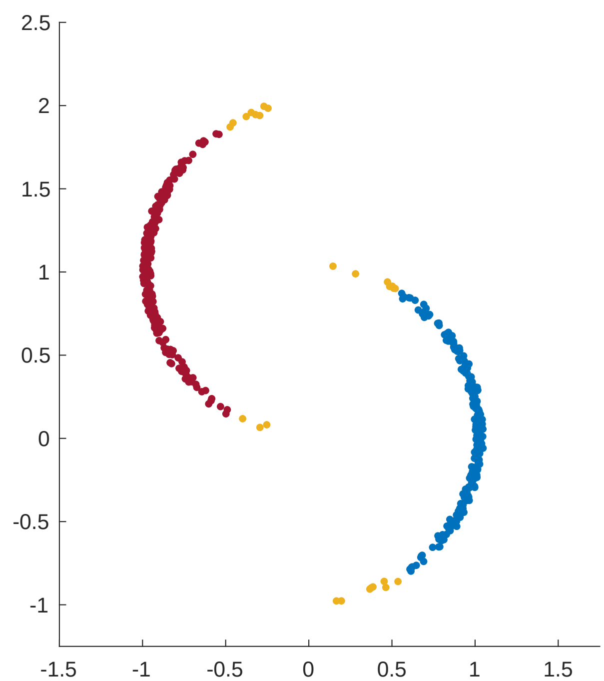

Figure 3: Labeled points with clustering at $\lambda = 0.00395$

Sum-of-norms clustering with equal weights performs poorly on standard half moons [3] and normally distributed half moons with large standard deviation. To resolve the issue, many authors such as Sun et al. [13] apply exponentially decaying weights to the sum-of-norms clustering. The exponentially decaying weight of pair (i, j) is determined by the distance between original data $\boldsymbol{a}_i$ and $\boldsymbol{a}_j$. The weight is set to zero if j is not among $i's$ k-nearest neighbors. Otherwise, the weight is computed as follows

$$w_{ij} = \exp(-\phi \left\| a_i - a_j \right\|^2)$$

where ϕ is a nonnegative parameter. Assigning weights in this manner implicitly imposes a prior hypothesis that the nearest-neighbor structure corresponds to true clustering, which is certainly the case for the standard half-moon data set. Chi and Lange [3] assess the effect of the number of nearest neighbors k and the parameter ϕ on SON clustering with numerical experiments on a half-moon dataset of 100 points. Setting $k = 10$ and $\phi = 0.5$ yields the best clustering. Choosing $k = 50$ and $\phi = 0$ results in a similar clustering pattern to our experiment: clusters only form until late then all points quickly coalesce to one cluster. At any value of λ, SON clustering could not identify two half moons with a high accuracy. When $k = 10$ and $\phi = 0$, or $k = 50$ and $\phi = 0.5$, SON clustering correctly identifies clusters for the easier points but fails to cluster points located at the lower tip of the right moon and the upper tip of the left moon.

8 Discussion

We proposed a test to determine all clusters from an approximate solution yielded from any primal-dual type method. If the test reports 'success', then the clusters are correctly identified. Moreover, if a primal-dual path following method that maintains close proximity to the central path is used, the test is guaranteed to report 'success' after a finite number of iterations at non-fusion values of λ, where strict complementarity holds. A few natural questions concerning strict complementarity and the test itself are (1) Is there a rigorous test that works when strict complementarity fails? (2) What is the complexity for our clustering test since it depends on the choice of λ values? (3) Is the test guaranteed to work for a general primal-dual algorithm? (4) Can one identify clusters correctly from a primal-only algorithm?

References

[1] F. Alizadeh and D. Goldfarb. Second-order cone programming. *Mathematical Programming*, 95, 12 2001.

[2] J. Bezanson, A. Edelman, S. Karpinski, and V.B. Shah. Julia: A fresh approach to numerical computing. *SIAM Rev.*, 59(1):65–98, 2017.

[3] Eric C. Chi and Kenneth Lange. Splitting methods for convex clustering. *Journal of Computational and Graphical Statistics*, 24(4):994–1013, 2015.

[4] Eric C. Chi and Stefan Steinerberger. Recovering trees with convex clustering. *SIAM Journal on Mathematics of Data Science*, 1(3):383–407, 2019.

[5] J. Chiquet, P. Gutierrez, and G. Rigaill. Fast tree inference with weighted fusion penalties. *Journal of Computational and Graphical Statistics*, 26:205–216, 2017.

[6] T. Hocking, A. Joulin, F. Bach, and J.-P. Vert. Clusterpath: An algorithm for clustering using convex fusion penalties. In *International Conference on Machine Learning*, 2011.

[7] T. Jiang, S. Vavasis, and C. W. Zhai. Recovery of a mixture of gaussians by sum-of-norms clustering. `https://arxiv.org/abs/1902.07137`, 2019.

[8] F. Lindsten, H. Ohlsson, and L. Ljung. Clustering using sum-of-norms regularization: With application to particle filter output computation. In *IEEE Statistical Signal Processing Workshop (SSP)*, 2011.

[9] Zhi-Quan Luo, Jos F. Sturm, and Shuzhong Zhang. Superlinear convergence of a symmetric primal-dual path following algorithm for semidefinite programming. *SIAM Journal on Optimization*, 8(1):59–81, 1998.

[10] A. Panahi, D. Dubhashi, F. Johansson, and C. Bhattacharyya. Clustering by sum of norms: Stochastic incremental algorithm, convergence and cluster recovery. *Journal of Machine Learning Research*, 70, 2017.

[11] Kristiaan Pelckmans, Joseph De Brabanter, Bart De Moor, and Johan A. K. Suykens. Convex clustering shrinkage. In *PASCAL Workshop on Statistics and Optimization of Clustering*, 2005.

[12] William M. Rand. Objective criteria for the evaluation of clustering methods. *Journal of the American Statistical Association*, 66(336):846–850, 1971.

[13] Defeng Sun, Kim-Chuan Toh, and Yancheng Yuan. Convex clustering: Model, theoretical guarantee and efficient algorithm, 2018.