A Convex Approach to K-Means Clustering and Image Segmentation

Laurent Condat

CNRS, GIPSA-Lab, Univ. Grenoble Alpes, 38000 Grenoble, France
Contact: see `http://www.gipsa-lab.grenoble-inp.fr/~laurent.condat/`

Abstract. A new convex formulation of data clustering and image segmentation is proposed, with fixed number K of regions and possible penalization of the region perimeters. So, this problem is a spatially regularized version of the K-means problem, a.k.a. piecewise constant Mumford–Shah problem. The proposed approach relies on a discretization of the search space; that is, a finite number of candidates must be specified, from which the K centroids are determined. After reformulation as an assignment problem, a convex relaxation is proposed, which involves a kind of $l_{1,\infty}$ norm ball. A splitting of it is proposed, so as to avoid the costly projection onto this set. Some examples illustrate the efficiency of the approach.

Keywords: Image segmentation · Piecewise constant Mumford–Shah problem · K-means · Convex relaxation

1 Introduction

Data partitioning, or clustering, aims at decomposing a set of elements into groups, so as to minimize some notion of intra-group dissimilarity [1, 2]. Thus, the classical K-*means* problem [3], consists in partitioning N points of $\mathbb{R}^d$ into K groups, by minimizing the sum of squared distances from every point to the nearest *centroid*, which is the center of mass of a group. For scalar data ($d = 1$), the K-means problem can be solved exactly and efficiently using dynamic programming [4, 5]. By contrast, when $d \geq 2$, it is generally NP-hard [6, 7]. An application is color image quantization [8,9]: one looks for the palette of K colors representing at best a given image; in that case, the points are the pixel values in $\mathbb{R}^3$, corresponding to the coordinates in some color space.

A fundamental problem in image processing and vision, which is even more difficult, is image segmentation: one wants to decompose an image of N pixels into K regions, corresponding to the objects of the scene, by favoring, in addition to intra-region similarity, some notion of spatial homogeneity [10, 11]. We consider in this article the NP-hard *piecewise-constant Mumford–Shah problem* [11, 12]: spatial homogeneity is obtained by penalizing the sum of the region perimeter.

In general terms, the considered problem can be formalized as follows. The data $y = (y_n)_{n \in \Omega}$ is a 1-D signal of domain $\Omega = \{1, \ldots, N\}$ or a 2-D image of domain $\Omega = \{1, \ldots, N_1\} \times \{1, \ldots, N_2\}$ (having $N = N_1 N_2$ pixels), with values y_n in $\mathbb{R}^d$, endowed with the Euclidean norm. Given an integer $K \geq 2$, one wants to partition Ω into K regions[1] Ω_k (so $\bigcup_{k=1}^{K} \Omega_k = \Omega$ and $\Omega_k \cap \Omega_{k'} = \emptyset$, for all $k \neq k'$), and to find the corresponding centroids $c_k \in \mathbb{R}^d$, so as to

$$\underset{(\Omega_k)_{k=1}^{K}, (c_k)_{k=1}^{K}}{\text{minimize}} \frac{1}{2} \sum_{k=1}^{K} \sum_{n \in \Omega_k} \|y_n - c_k\|^2 + \frac{\lambda}{2} \sum_{k=1}^{K} \text{per}(\Omega_k), \tag{1}$$

where per denotes the perimeter and $\lambda \geq 0$ is a parameter controlling the level of spatial regularization. When $\lambda = 0$, this is exactly the K-means problem; then the geometry of the domain and the indexing order do not play any role and one can think in terms of partitioning the point cloud $(y_n)_{n \in \Omega}$ in $\mathbb{R}^d$ into K groups, whose c_k are the means.

We can define the quantized or segmented signal or image $x = (x_n)_{n \in \Omega}$, with $x_n = c_k$ if $n \in \Omega_k$. So, x is a piecewise constant approximation of y, taking at most K different values. If $\lambda = 0$, we can express the problem (1) as:

$$\underset{x \in (\mathbb{R}^d)^\Omega}{\text{minimize}} \ \frac{1}{2} \|y - x\|_2^2 \quad \text{s.t.} \quad |\{x_n \ : \ n \in \Omega\}| \leq K, \tag{2}$$

where $\|y - x\|_2^2 = \sum_{n \in \Omega} \|y_n - x_n\|^2$ and $|\cdot|$ denotes the cardinality of a set. Moreover, in the case of a 1-D signal, $\frac{1}{2} \sum_{k=1}^{K} \text{per}(\Omega_k)$ is equal to $|\{n \ : \ x_n \neq x_{n+1}\}|$, the number of jumps in the signal[2]. So, in 1-D, we can express the problem (1) as:

$$\underset{x \in (\mathbb{R}^d)^N}{\text{minimize}} \ \frac{1}{2} \|y - x\|_2^2 + \lambda |\{n = 1, \ldots, N-1 \ : \ x_n \neq x_{n+1}\}|$$

$$\text{s.t.} \quad |\{x_n \ : \ n = 1, \ldots, N\}| \leq K. \tag{3}$$

The definition of the discrete perimeter in 2-D, based on finite differences, is much more problematic [13, 14].

We can note that, if the regions Ω_k are fixed, the centroids c_k, solutions to the problem (1), are the means of the elements of the regions: $c_k = \frac{1}{|\Omega_k|} \sum_{n \in \Omega_k} y_n$.

---

[1] The number of regions is actually at most K, and not exactly K, because some regions Ω_k could be empty. This is never the case in practical applications.

[2] We assume symmetric boundary conditions, so the boundary of the domain Ω is not counted in the perimeter.

Conversely, if the c_k are fixed, i) if $\lambda = 0$, we simply get the regions as the Voronoi cells $\Omega_k = \{n \in \Omega \ : \ k = \arg\min_{k'} \|y_n - c_{k'}\|\}$; ii) if $\lambda > 0$, there exist efficient methods to solve the problem, by convex relaxation [13–15] or by graph cuts [16], see also [17]. Therefore, a strategy consists in alternating between updating the Ω_k at fixed c_k, and the other way around. In the case $\lambda = 0$, this yields exactly the classical K-means algorithm, due to Lloyd [18]; it must be distinguished from the K-means problem (1), for which it is a heuristic. It converges to a local minimum of the problem, but is very dependent on the initialization [3, 9].

Globally convex methods, with high complexity, have been proposed to solve the problem (1), in the particular cases $K = 2$ or $K = 4$ [19–21]. The author is not aware of a generic method to approximate the global minimum of (1); existing convex relaxations of the K-means problem are discussed in Sect. 3. In this work, the problem is addressed by **discretizing** the search space of the centroids: we fix a set $\Gamma = \{a_m\}_{m=1}^M$ of M points in $\mathbb{R}^d$, called the *candidates*, and the centroids are constrained to belong to Γ, instead of to the whole space $\mathbb{R}^d$. Typically, $K \ll M$. We propose a convex formulation of the problem and an algorithm to solve it in Sect. 2. We illustrate the good performances of the approach in Sect. 4.

2 Proposed Method

2.1 Problem Reformulation by Lifting

In the segmented signal or image x, every element x_n is one of the centroids c_k, which is itself one of the candidates a_m. Thus, we can reformulate the problem (1) as an equivalent problem, whose unknown is the assignment array z, which has one more dimension than y, indexed by $m = 1, \ldots, M$; this is called *lifting* [14, 22]. For every $n \in \Omega$ and $m = 1, \ldots, M$, $z_{m,n}$ is equal to 1 if $x_n = a_m$ and to 0 else. Each vector $z_{:,n} = (z_{m,n})_{m=1}^M$ belongs to the set $\mathcal{A}$ of binary assignment vectors, i.e. vectors with elements in $\{0, 1\}$ whose sum is 1.

We retrieve x from z by a simple summation:

$$x_n = \sum_{m=1}^M z_{m,n} a_m, \quad \forall n \in \Omega. \tag{4}$$

Moreover, the data fidelity term in (1) can be rewritten as

$$\frac{1}{2} \sum_{n \in \Omega} \|x_n - y_n\|^2 = \frac{1}{2} \sum_{n \in \Omega} \sum_{m=1}^M z_{m,n} w_{m,n}, \tag{5}$$

or, in short, $\frac{1}{2}\|x - y\|_2^2 = \langle z, w \rangle$, where

$$w_{m,n} = \tfrac{1}{2}\|y_n - a_m\|^2. \tag{6}$$

By using the coarea formula, according to which the total variation (TV) of the indicator function of a set (1 inside, 0 outside) is equal to the perimeter of that set [14], the regularization term in (1) can be rewritten as

$$\frac{\lambda}{2} \sum_{k=1}^{K} \mathrm{per}(\Omega_k) = \frac{\lambda}{2} \sum_{m=1}^{M} \mathrm{TV}(z_{m,:}), \tag{7}$$

where TV is some discrete form of the TV [13] and $z_{m,:} = (z_{m,n})_{n \in \Omega}$ is a scalar signal or image of same domain Ω as y.

It remains to reformulate the constraint that the number of regions Ω_k is at most K or, equivalently, that x takes its values in only K among the M candidates, i.e. $|\{x_n \,:\, n \in \Omega\}| \leq K$. We have the following property:

Proposition 1 *The assignment array $z \in \mathcal{A}^\Omega$ corresponds, by (4), to a signal or image $x \in \Gamma^\Omega$ taking at most K distinct values, if and only if*

$$\|z\|_{1,\infty} \leq K, \tag{8}$$

where[3] $\|z\|_{1,\infty} = \sum_{m=1}^{M} \max_{n \in \Omega} z_{m,n}$.

Indeed, since z takes its values in $\{0,1\}$, a candidate a_m is assigned to at least one point x_n, and therefore is one of the centroids c_k, if and only if $z_{m,:}$ contains at least one 1; that is, if and only if $\max_{n \in \Omega} z_{m,n} = 1$.

Hence, we can rewrite the problem (1), with discrete search space Γ for the centroids, as

$$\underset{z \in \mathcal{A}^\Omega}{\text{minimize}} \ \langle z, w \rangle + \frac{\lambda}{2} \sum_{m=1}^{M} \mathrm{TV}(z_{m,:}) \ \ \text{s.t.} \ \ \|z\|_{1,\infty} \leq K. \tag{9}$$

2.2 Convex Relaxation of the Problem

The problem (9) is not convex; more precisely, the functions and sets are all convex, except the set $\mathcal{A}$ of binary assignment vectors. So, we consider a convex relaxation, obtained by replacing $\mathcal{A}$ by its convex hull, which is the simplex Δ, i.e. the set of vectors with nonnegative elements whose sum is 1 [23]. Let us introduce the ball $\mathcal{B} = \{s \in \mathbb{R}^{M \times \Omega} \,:\, \|s\|_{1,\infty} \leq K\}$, and the convex indicator function $\imath_\mathcal{E}$ of a convex set $\mathcal{E}$, which takes the value 0 if its variable belongs to $\mathcal{E}$ and $+\infty$ else. The proposed convex problem is then:

$$\underset{z \in \mathbb{R}^{M \times \Omega}}{\text{minimize}} \ \langle z, w \rangle + \sum_{n \in \Omega} \imath_\Delta(z_{:,n}) + \frac{\lambda}{2} \sum_{m=1}^{M} \mathrm{TV}(z_{m,:}) + \imath_\mathcal{B}(z). \tag{10}$$

For conveniency, we denote by Δ^Ω the set of arrays of same size à z whose columns are on the simplex, so that $\imath_{\Delta^\Omega}(z) = \sum_{n \in \Omega} \imath_\Delta(z_{:,n})$.

---

[3] In this paper, we make an abuse of the terms $l_{1,\infty}$ norm and ball: the elements of z are nonnegative, so there is no need to take their absolute values.

We can note that another convex relaxation of the perimeter term, which is better, when z belongs to Δ^Ω, than the one in (7) we use in this paper, has been proposed in [14]. We do not consider it here because of its higher computational complexity, but there would be no difficulty in using it in our context.

The projection onto the simplex can be performed efficiently [23], see code on the author's webpage. However, the projection onto $\mathcal{B}$, which can also be performed exactly in finite time [24], is very costly. That is why we propose a (dual) splitting [25] of the maximum function:

Proposition 2 *The maximum function of a vector, or more generally of an array with N elements, s can be expressed as an infimal convolution [25]:*

$$\max_{n=1,\ldots,N} s_n = \min_{q \in \mathbb{R}} q/\sqrt{\mu N} + \imath_{\leq 0}(s - S^* q/\sqrt{\mu N}), \tag{11}$$

where $\mu > 0$ is some fixed constant, $\imath_{\leq 0}$ is the indicator function of the cone of arrays with nonpositive elements, S is the linear operator, which maps an array to the sum of its elements, and its adjoint operator S^* duplicates a real number into an array of same size as s with N identical elements.

We can note that, in Proposition 2, the norm of the linear operator $(s, q) \mapsto s - S^* q/\sqrt{\mu N}$ is $1 + 1/\mu$.

Let us introduce the constraint set $\mathcal{C} = \left\{ s \in \mathbb{R}^M \ : \ \sum_{m=1}^M s_m \leq K\sqrt{\mu N} \right\}$. The convex problem we propose to solve is then:

$$\operatorname*{minimize}_{z \in \mathbb{R}^{M \times \Omega}, q \in \mathbb{R}^M} \langle z, w \rangle + \imath_{\Delta^\Omega}(z) + \imath_\mathcal{C}(q) + \frac{\lambda}{2} \sum_{m=1}^M \mathrm{TV}(z_{m,:})$$

$$+ \sum_{m=1}^M \imath_{\leq 0}\left(z_{m,:} - S^* q_m/\sqrt{\mu N}\right). \tag{12}$$

We can remark that the use of $\|z\|_{1,\infty}$ to control the number of regions has been proposed in the same context in [26], but as a penalty and not as a constraint. Handling a $l_{1,\infty}$ constraint is a priori more difficult than regularizing with the $l_{1,\infty}$ norm, but we have seen that the splitting technique in Proposition 2 ends in the convex optimization problem (12), involving only simple terms.

2.3 Proposed Algorithm

In the following, we assume that the 'isotropic' form of the discrete TV [13] is used. There would be no difficulty in using instead the form proposed in [13], for better quality but higher computational complexity. As usual, we express the isotropic TV as the $l_{1,2}$ norm composed with a linear operator D of finite differences [13, 27]. More precisely, in the case of a 2-D image with domain $\Omega = \{1, \ldots, N_1\} \times \{1, \ldots, N_2\}$, we introduce $D : \mathbb{R}^{M \times \Omega} \to \mathbb{R}^{M \times \Omega \times 2}, z \mapsto v$, with $v_{m,n,1} = z_{m,(n_1+1,n_2)} - z_{m,n}$ and $v_{m,n,2} = z_{m,(n_1,n_2+1)} - z_{m,n}$, for every $m = 1, \ldots, M$ and $n = (n_1, n_2) \in \Omega$ (using symmetric boundary conditions). Note

that the operator norm of D is 8. We also introduce $\mathcal{V} = \{v \in \mathbb{R}^{M \times \Omega \times 2}$: $v_{m,n,1}^2 + v_{m,n,2}^2 \leq \lambda^2/4, \ \forall m = 1, \ldots, M, \ n \in \Omega\}$.

In the case of a 1-D signal with domain $\Omega = \{1, \ldots, N\}$, some simplifications can be made. We can set $D : \mathbb{R}^{M \times N} \to \mathbb{R}^{M \times (N-1)}, z \mapsto v$, with $v_{m,n} = z_{m,n+1} - z_{m,n}$, for every $m = 1, \ldots, M$ and $n = 1, \ldots, N - 1$, and $\mathcal{V} = \{v \in \mathbb{R}^{M \times (N-1)}$: $|v_{m,n}| \leq \lambda/2, \ \forall m = 1, \ldots, M, \ n = 1, \ldots, N - 1\}$. Note that the operator norm of D is 4 in this case, so we can set $\sigma_v := (1 - \gamma)/\tau/4$ in Algorithm 1 below.

We denote by $\mathrm{P}_{\mathcal{E}}$ the projection onto a set $\mathcal{E}$, by $\iota_{\geq 0}$ and $P_{\geq 0}$ the indicator function of and the projection onto the cone of arrays with nonnegative elements, respectively. We extend the summation operator S to any array $u \in \mathbb{R}^{M \times \Omega}$ of same size as z, to mean summation with respect to the index $n \in \Omega$; that is, $Su = (Su_{m,:})_{m=1}^M \in \mathbb{R}^M$. Consequently, the last term in (12) can be rewritten more shortly as $\sum_{m=1}^M \iota_{\leq 0}(z_{m,:} - S^*q_m/\sqrt{\mu N}) = \iota_{\leq 0}(z - S^*q/\sqrt{\mu N})$.

The proposed algorithm is the over-relaxed version [28] of the Chambolle–Pock algorithm [27], applied to the problem (12), viewed as the sum of one function and two functions composed with linear operators, with the pair (z, q) as variable. With the proposed range of parameters, it is proved to converge to a solution of (12).

Algorithm 1
Input : w, K, λ, M, N, μ. Output : estimate $z^{(i+\frac{1}{2})}$ of a solution to (12).
Choose $\rho \in [1, 2)$, $\tau > 0$, $\gamma \in (0, 1)$, and the initial estimates $z^{(0)}$, $q^{(0)}$, $u^{(0)}$, $v^{(0)}$.
Set $\sigma_u := \gamma/\tau/(1 + 1/\mu)$, $\sigma_v := (1 - \gamma)/\tau/8$.
Iterate: for $i = 0, 1, \ldots$

$$
\begin{aligned}
z^{(i+\frac{1}{2})} &:= \mathrm{P}_{\Delta^\Omega}\big(z^{(i)} - \tau(u^{(i)} + w + D^*v^{(i)})\big), \\
q^{(i+\frac{1}{2})} &:= \mathrm{P}_{\mathcal{C}}\big(q^{(i)} + (\tau/\sqrt{\mu N})Su^{(i)}\big), \\
u^{(i+\frac{1}{2})} &:= \mathrm{P}_{\geq 0}\big(u^{(i)} + \sigma_u\big(2z^{(i+\frac{1}{2})} - z^{(i)} - S^*(2q^{(i+\frac{1}{2})} - q^{(i)})/\sqrt{\mu N}\big)\big), \\
v^{(i+\frac{1}{2})} &:= \mathrm{P}_{\mathcal{V}}\big(v^{(i)} + \sigma_v D(2z^{(i+\frac{1}{2})} - z^{(i)})\big), \\
z^{(i+1)} &:= z^{(i)} + \rho(z^{(i+\frac{1}{2})} - z^{(i)}), \\
q^{(i+1)} &:= q^{(i)} + \rho(q^{(i+\frac{1}{2})} - q^{(i)}), \\
u^{(i+1)} &:= u^{(i)} + \rho(u^{(i+\frac{1}{2})} - u^{(i)}), \\
v^{(i+1)} &:= v^{(i)} + \rho(v^{(i+\frac{1}{2})} - v^{(i)}).
\end{aligned}
$$

The memory size for z and the dual variables u and v is $O(NM)$; it is $O(M)$ for q. The complexity of P_{Δ^Ω} using the default sorting strategy is $O(NM \log M)$ [23]; it can be reduced to $O(NM)$ using a linear-time median-finding subroutine [23]. The complexity of the other operations is $O(NM)$, so the overall complexity of every iteration of the algorithm is $O(NM \log M)$. The parameters μ, τ, γ, ρ influence the convergence speed. They must be chosen on a case-by-case basis, but as a first step, one might consider $\rho = 1.9$, $\mu = 1$, $\gamma = 0.01$.

In the case of clustering or quantization, i.e. $\lambda = 0$, the algorithm can be simplified as:

Algorithm 2

Input : w, K, M, N, μ. Output : estimate $z^{(i+\frac{1}{2})}$ of a solution to (12).
Choose $\rho \in [1,2)$, $\tau > 0$, and the initial estimates $z^{(0)}$, $q^{(0)}$, $u^{(0)}$. Set $\sigma :=$
$1/\tau/(1+1/\mu)$.
Iterate: for $i = 0, 1, \ldots$

$$
\begin{aligned}
&z^{(i+\frac{1}{2})} := \mathrm{P}_{\Delta^\Omega}\big(z^{(i)} - \tau(u^{(i)} + w)\big), \\
&q^{(i+\frac{1}{2})} := \mathrm{P}_{\mathcal{C}}\big(q^{(i)} + (\tau/\sqrt{\mu N})Su^{(i)}\big), \\
&u^{(i+\frac{1}{2})} := \mathrm{P}_{\geq 0}\big(u^{(i)} + \sigma\big(2z^{(i+\frac{1}{2})} - z^{(i)} - S^*(2q^{(i+\frac{1}{2})} - q^{(i)})/\sqrt{\mu N}\big)\big), \\
&z^{(i+1)} := z^{(i)} + \rho(z^{(i+\frac{1}{2})} - z^{(i)}), \\
&q^{(i+1)} := q^{(i)} + \rho(q^{(i+\frac{1}{2})} - q^{(i)}), \\
&u^{(i+1)} := u^{(i)} + \rho(u^{(i+\frac{1}{2})} - u^{(i)}).
\end{aligned}
$$

To study the numerical convergence of Algorithm 1, it is useful to consider the dual problem [25] of the primal problem (12):

$$
\underset{u\in\mathbb{R}^{M\times\Omega},v\in\mathbb{R}^{M\times\Omega\times 2}}{\text{minimize}} \sum_{n\in\Omega} \max_{m=1,\ldots,M} (-u - D^*v - w)_{m,n} + \iota_{\mathcal{C}^*}(Su) + K(Su)_1
$$
$$
+ \iota_{\geq 0}(u) + \iota_{\mathcal{V}}(v), \tag{13}
$$

where $\mathcal{C}^* = \{s \in \mathbb{R}^M \ : \ s_1 = \cdots = s_M \geq 0\}$. At convergence, the primal and the dual cost values are opposite to each other [25]. That is, for a solution z to (12) and a solution pair (u,v) to (13),

$$
\Psi^\infty = \langle z, w\rangle + \frac{\lambda}{2} \sum_{m=1}^{M} \mathrm{TV}(z_{m,:}) = \sum_{n\in\Omega} \min_{m=1,\ldots,M} (u + D^*v + w)_{m,n} - K(Su)_1. \tag{14}
$$

The iterate $z^{(i+\frac{1}{2})}$ does not belong to the $l_{1,\infty}$ ball $\mathcal{B}$, so its cost value

$$
\Psi^{\mathrm{p},i} = \langle z^{(i+\frac{1}{2})}, w\rangle + \frac{\lambda}{2} \sum_{m=1}^{M} \mathrm{TV}(z_{m,:}^{(i+\frac{1}{2})}), \tag{15}
$$

which tends to Ψ^∞ as $i \to +\infty$, is not guaranteed to be larger than Ψ^∞. Concerning the dual variables, $u^{(i+\frac{1}{2})} \geq 0$ and $v^{(i+\frac{1}{2})} \in \mathcal{V}$, but $Su^{(i+\frac{1}{2})}$ does not belong to $\mathcal{C}^*$. However, taking the maximum over m of $Su^{(i+\frac{1}{2})}$ to evaluate the cost yields a valid lower bound of Ψ^∞; that is,

$$
\Psi^{\mathrm{d},i} = \sum_{n\in\Omega} \min_{m=1,\ldots,M} (u^{(i+\frac{1}{2})} + D^*v^{(i+\frac{1}{2})} + w^{(i+\frac{1}{2})})_{m,n} - K \max_{m=1,\ldots,M} \sum_{n\in\Omega} u_{m,n}^{(i+\frac{1}{2})}
$$
$$
\tag{16}
$$

is $\leq \Psi^\infty$ and tends to Ψ^∞ as $i \to +\infty$. Hence, a good way to test the convergence of Algorithm 1 is to check that

$$
\sum_{m=1}^{M} \max_{n\in\Omega} z_{m,n}^{(i+\frac{1}{2})} - K \leq \epsilon_1 \tag{17}
$$

and

$$\Psi^{\mathrm{p},i} - \Psi^{\mathrm{d},i} \leq \epsilon_2, \tag{18}$$

for small constants ϵ_1 and ϵ_2.

In all the examples of Sect. 4, it took several tenths of minutes to generate the results with the proposed algorithms, run several thousands of iterations.

3 Prior Work on the K-Means Problem

The convex relaxation described in the previous section turns out to be well known in operations research for the closely related K-*median*, a.k.a. p-median, problem, which is the same as the K-means problem, with the halved squared Euclidean distance replaced by the Euclidean distance or any cost function with metric properties [29]. The search space for the cluster centers is discretized, as well. The classical example is a *facility location problem*: given N clients $\{y_n\}_{n=1}^N$ and a set $\Gamma = \{a_m\}_{m=1}^M$ of potential facilities, open K facilities so as to minimize the sum of the cost for each client of using its nearest open facility; typically this cost is the distance $w_{m,n} = \|y_n - a_m\|$. The 0-1 integer formulation and its convex relaxation as the linear program (12) (with $\lambda = 0$, $\mu = 1/N$, and the simplex constraint split into nonnegativity and sum-to-one constraints) can be traced back to half a century ago [30]. The variable q_m say whether facility m has to be opened or not and $z_{m,n}$ is the fraction of the demand of client n that is supplied by facility m. When the solution is not binary, several rounding strategies have been proposed [31].

In the case the candidates coincide with the data, i.e. $M = N$ and $\Gamma = y$, the K-means or K-median problem is often called the K-*medoid(s)* problem [32]; a centroid c_k is then constrained to be an exemplar from within the dataset and is called a medoid. The convex problem (12) can be simplified, because there is no need to introduce the auxiliary variables q_m: the maximum element of $z_{m,:}$ is $z_{m,m}$. Therefore, the problem can be rewritten as: minimize$_{z \in \mathbb{R}^{N \times N}}$ $\langle z, w \rangle$ subject to $\mathrm{tr}(z) = K$ and $0 \leq z_{m,n} \leq z_{m,m}$ and $\sum_m z_{m,n} = 1$, for every m, n.

There does exist a convex relaxation of the K-means problem, which does not require discretizing the search space of the centroids. Indeed, in the solution, we have the identity $\sum_{n \in \Omega_k} \|y_n - c_k\|^2 = \sum_{(n,n') \in (\Omega_k)^2} \|y_n - y_{n'}\|^2/(2|\Omega_k|)$. Consequently, we can define the symmetric affinity matrix h of size $N \times N$, with $h_{n,n'} = \{1/|\Omega_k|$ if n and n' belong to the same cluster Ω_k in the solution, 0 else$\}$. Every column of h is on the simplex, h is positive semidefinite, its trace and rank are K. We can then reformulate the problem as minimizing $\langle h, w \rangle$ over the set of such matrices. Linear programming and semidefinite programming convex relaxations of this problem have been proposed [33–35]. They are less efficient in practice than the convex formulation (12) [35].

Another convex relaxation of the K-means problem consists in minimizing $\|x - y\|_2^2/2 + \lambda \sum_{n=1}^N \sum_{n'>n} \|x_n - x_{n'}\|_p$ over x, for some $p \in [1, +\infty]$ and some regularization parameter λ [36–40]. Then a cluster consists in all the points y_n corresponding to the same point x_n. This is a hierarchical approach: if two points are in the same cluster, they remain so when λ increases.

We can notice that the alternatives to the convex formulation (12) mentioned in this section require to store matrices of size $N \times N$ and to perform operations with complexity $O(N^2)$; this is not feasible when N is a typical image size.

4 Experimental Validation

4.1 K-Means Clustering

We consider the K-means problem, i.e. (1) with $\lambda = 0$, applied to the dataset A1 from `https://cs.joensuu.fi/sipu/datasets/` [41], to partition the point cloud y of $N = 3000$ points in dimension $d = 2$ into $K = 20$ clusters.

In a first experiment, the set of candidates Γ is a uniform grid of 80×40 points, from which we kept only the $M = 2280$ points in the convex hull of y. Remarkably, Algorithm 2 converges to the exact solution of the problem (12), up to machine precision, in a finite number of iterations. It is also remarkable that the solution z, represented in Fig. 1 (a), is binary, i.e. with values in $\{0, 1\}$; this can be easily verified by testing that $\min_{n \in \Omega} \max_{m=1}^{M} z_{m,n} = 1$. Therefore, the solutions to the convex problem (10) and to the nonconvex problem (9) coincide, and the proposed method yields the *global* solution of the K-means problem (for this choice of Γ).

In a second experiment, we consider the K-medoids case: $M = N$ and $\Gamma = y$. In this case too, Algorithm 2 converges in finite time to the exact solution of the problem, represented in Fig. 1 (b), which is binary. So, the global solution of the K-medoids problem has been found.

We can note that it is easy to design counter-examples, where the solution of the convex problem (12) is not binary and does not yield a solution to the nonconvex K-means problem.

4.2 Color Image Quantization

We consider the color image quantization problem [9]. It is an instance of the K-means problems, which consists in partitioning the pixel values $y_n \in \mathbb{R}^3$ of an image, supposed to be coordinates in some color space. We consider here the CIELAB color space, because the Euclidean distance in it approximately matches the human perceptual distance. We first construct a palette of $M = 279$ colors, shown in Fig. 2, obtained by sampling the CIELAB space on a body centered cubic lattice. Indeed, this lattice is the one minimizing the quantization error; that is, for a given sampling density, or size of the Voronoi cell, the average squared distance between any point and the closest point in the lattice is minimized [42]. We consider three images and, for each, the three cases $K = 4$, $K = 5$, $K = 6$. The results are shown in Fig. 3. In the nine cases, Algorithm 2 converges after a finite number of iterations to the exact solution of (12), which is binary. So, the obtained images are the global solutions of the nonconvex color image quantization problem.
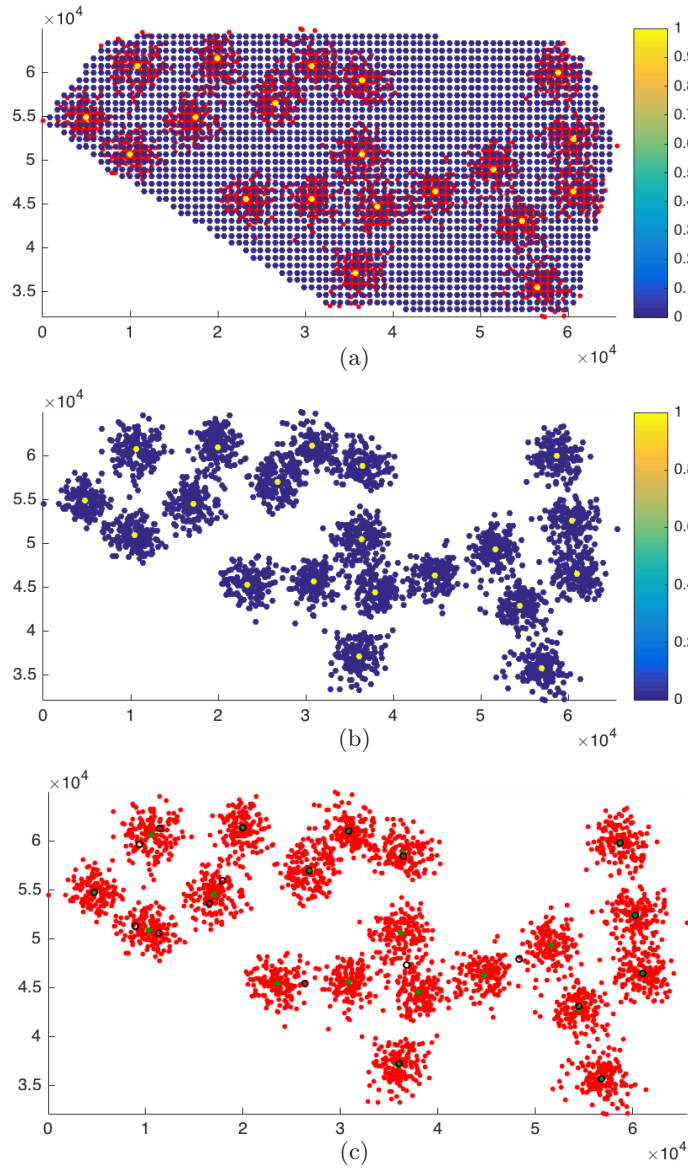
Fig. 1. Clustering of $N = 3000$ points, in red, into $K = 20$ groups. In (a), the M candidates a_m are on a uniform grid and in (b), they coincide with the data points. In both cases, they are represented with a color corresponding to the value $\max_{n \in \Omega} z_{m,n}$, where z is the solution to (12), with $\lambda = 0$. It turns out that z is binary, so it is a solution to the nonconvex problem (9). In (c), the centroids found by the kmeans algorithm of MATLAB, with default random initialization, are shown in black, whereas the centroids found by the same algorithm, but initialized with the centroids found by the proposed method, in yellow in (b), are shown in green. The latter are the global solutions of the K-means problem (1).

Fig. 2. Palette of $M = 279$ colors obtained by sampling the CIELAB color space on a body centered cubic lattice. The two last orange and red colors have been added manually.

4.3 Image Segmentation

We now consider the segmentation problem, i.e. (1) with $\lambda > 0$, using the same images as in Sect. 4.2 and Fig. 3 (a), the palette Γ of $M = 279$ candidate colors in Fig. 2, and $\lambda = 500$. The segmented images, solution to (12) and (4), are shown in Fig. 4. z is never binary in this context and a small blur is present at edges. Indeed, each value $z_{m,n}$ can be interpreted as the proportion of the candidate a_m required to represent the pixel value x_n at location $n \in \Omega$; for a pixel at an edge between two regions, it is natural that it is soft-classified, instead of being fully assigned to one or the other region.

The method succeeds in providing images made of K colors in the palette, for the *sunflower* with $K = 4$ and $K = 6$, the *ladybug* with $K = 5$ and $K = 6$, and the *parrot* with $K = 4$ and $K = 5$; that is, $\max_{n \in \Omega} z_{m,n} = 1$ for K indexes m, and $z_{m,n} = 0$ for all the other m and all $n \in \Omega$. In these cases, up to the blur at the edges, which can be removed by rounding z to make it binary, we can consider that we have obtained the *global* solution to the segmentation problem (1). For the *sunflower* with $K = 5$, the orange at the center of the large sunflower is actually a mixture of 59% of the color $m = 80$ and 41% of the color $m = 92$. Similarly, for the *ladybug* with $K = 4$, there are two pure colors, i.e. $\max_{n \in \Omega} z_{m,n} = 1$ for $m = 155$ and $m = 107$, and mixtures of four colors, with $\max_{n \in \Omega} z_{m,n}$ equal to 0.88 for $m = 23$ and $m = 84$, and equal to 0.12 for $m = 31$ and $m = 83$. In both cases, this is not an issue, since if one really wants to identify the K colors of the palette adapted to represent the image, a post-processing step keeping the K colors with the largest value $\max_{n \in \Omega} z_{m,n}$ would be appropriate. However, in the last case of the *parrot* with $K = 6$, the method fails to provide an image with 6 dominant colors. Indeed, $\max_{n \in \Omega} z_{m,n}$ is equal to 1, 1, 1, 1, 0.66, 0.34, 0.34, 0.34, 0.33, for $m = 261, 255, 63, 18, 155, 60, 66, 58, 68$, respectively. There is no obvious rounding procedure to keep 6 out of these 9 colors. Thus, the proposed method succeeds eight times and fails once, in the nine examples considered.

We can remark that the proposed approach, which estimates the K colors and the corresponding regions with low perimeter jointly, performs better than a two-step strategy, that would first estimate the K colors using quantization and then solve the segmentation problem restricted to these $M = K$ colors. Indeed, we can see that the proposed approach yields different orange and green colors for the *sunflower* with $K = 4$, in Fig. 4 (c), from the ones in Fig. 3 (d). The dark gray for the *parrot* with $K = 4$, in Fig. 4 (c), is also different from the dark

green in Fig. 3 (d). So, in these two examples, the two-step strategy would have failed to provide an image with the appropriate colors.
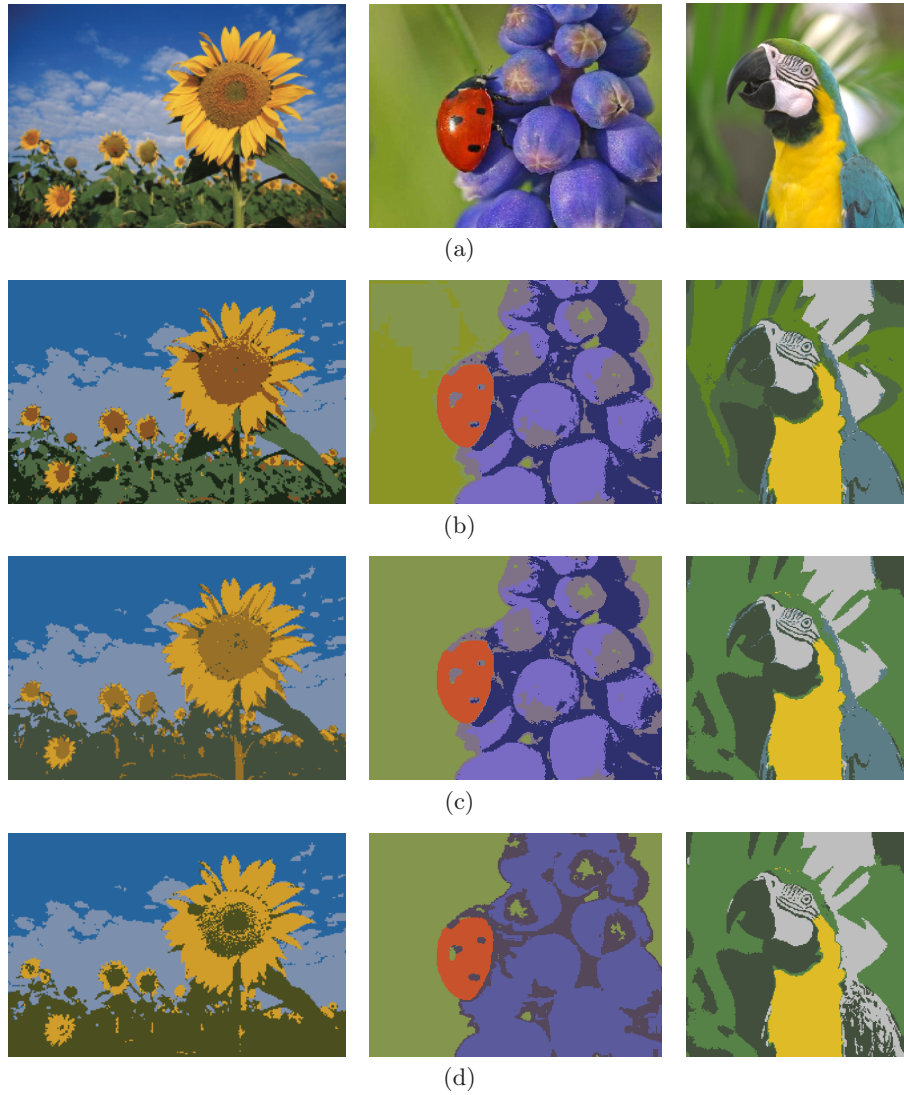
Fig. 3. In (a), original *sunflower*, *ladybug*, *parrot* images y, of size 254×168, 298×228, and 200×199, respectively. In (b)–(d), quantized images x solution to (12) ($\lambda = 0$) and (4), with $K = 6$, $K = 5$, $K = 4$, respectively, with the palette Γ of candidate colors shown in Fig. 2.

(a)

(b)

(c)
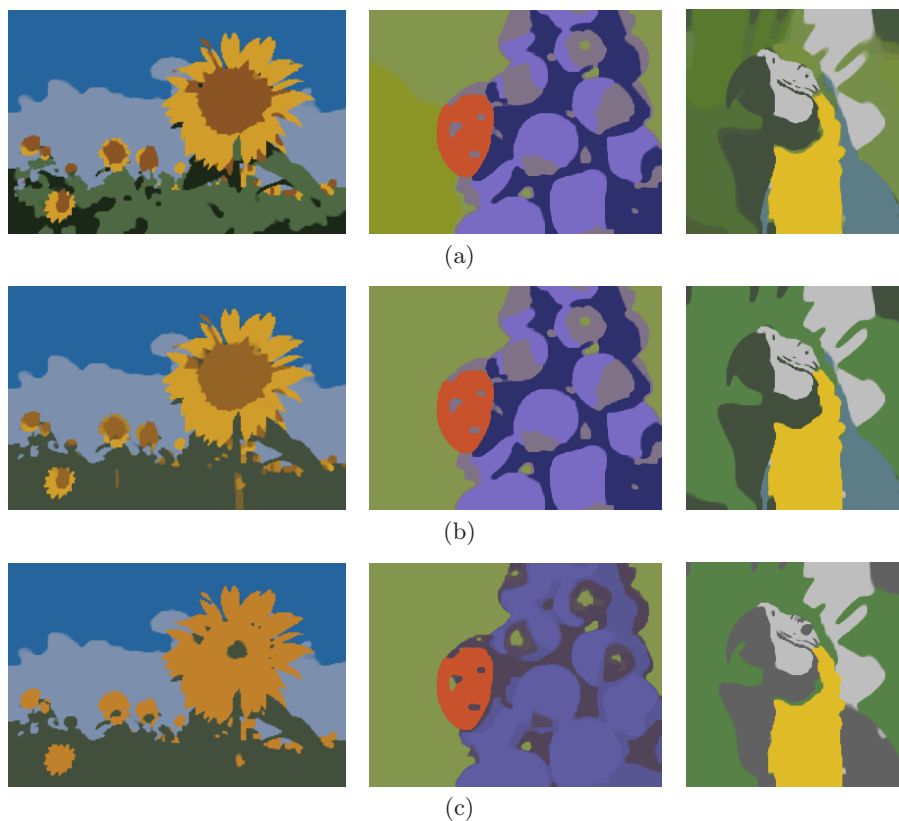
Fig. 4. Segmented images x, solution to (12) and (4), with y in Fig. 3 (a), $\lambda = 500$, and the palette Γ of candidate colors shown in Fig. 2. In (a), (b), (c), $K = 6$, $K = 5$, $K = 4$, respectively.

References

1. A. K. Jain, M. N. Murty, and P. J. Flynn, "Data clustering: A review," *ACM Computing Surveys*, vol. 31, 1999.
2. G. Gan, C. Ma, and J. Wu, *Data Clustering: Theory, Algorithms, and Applications*, ASA-SIAM Series on Statistics and Applied Probability. SIAM, Philadelphia, PA, USA, 2007.
3. D. Steinley, "K-means clustering: A half-century synthesis," *Br. J. Math. Stat. Psychol.*, vol. 59, no. 1, pp. 1–34, May 2006.
4. X. Wu, "Optimal quantization by matrix searching," *Journal of Algorithms*, vol. 12, no. 4, pp. 663–673, Dec. 1991.
5. F. K Soong and B. H. Juang, "Optimal quantization of LSP parameters," *IEEE Trans. Speech Audio Process.*, vol. 1, no. 1, pp. 15–24, Jan. 1993.
6. D. Aloise, A. Deshpande, P. Hansen, and P. Popat, "NP-hardness of Euclidean sum-of-squares clustering," *Mach. Learn.*, vol. 75, no. 2, pp. 245–248, May 2009.

7. M. Mahajan, P. Nimbhorkar, and K. Varadarajan, "The planar k-means problem is NP-hard," *Theoretical Computer Science*, vol. 442, pp. 13–21, July 2012, Special Issue on the Workshop on Algorithms and Computation (WALCOM 2009).

8. L. Brun and A. Trémeau, *Digital Color Imaging Handbook*, CRC Press, 2012, Chapter "Color quantization", pp. 589–638.

9. M. E. Celebi, "Improving the performance of k-means for color quantization," *Image and Vision Computing*, vol. 29, no. 4, pp. 260–271, Mar. 2011.

10. D. Cremers, M. Rousson, and R. Deriche, "A review of statistical approaches to level set segmentation: Integrating color, texture, motion and shape," *International Journal of Computer Vision*, vol. 72, pp. 195–215, 2007.

11. L. Bar, T. F. Chan, G. Chung, M. Jung, N. Kiryati, N. Sochen, and L. A. Vese, "Mumford and Shah model and its applications to image segmentation and image restoration," in *Handbook of Mathematical Methods in Imaging*, O. Scherzer, Ed. 2015, Springer.

12. D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems," *Communications on Pure and Applied Mathematics*, vol. 42, pp. 577–685, 1989.

13. L. Condat, "Discrete total variation: New definition and minimization," *SIAM J. Imaging Sciences*, vol. 10, no. 3, pp. 1258–1290, 2017.

14. A. Chambolle, D. Cremers, and T. Pock, "A convex approach to minimal partitions," *SIAM J. Imaging Sci.*, vol. 5, no. 4, pp. 1113–1158, 2012.

15. N. Pustelnik and L. Condat, "Proximity operator of a sum of functions; application to depth map estimation," *IEEE Signal Process. Lett.*, vol. 24, no. 12, pp. 1827–1831, Dec. 2017.

16. J. Yuan, E. Bae, X.-C. Tai, and Y. Boykov, "A continuous max-flow approach to Potts model," in *Proc. of ECCV*, 2010, pp. 379–392.

17. C. Zach, C. Häne, and M. Pollefeys, "What is optimized in convex relaxations for multilabel problems: Connecting discrete and continuously inspired MAP inference," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 1, pp. 157–170, Jan. 2014.

18. S. Lloyd, "Least squares quantization in PCM," *IEEE Trans. Inf. Theory*, vol. 28, no. 2, pp. 129–136, 1982.

19. E. S. Brown, T. F. Chan, and X. Bresson, "Completely convex formulation of the Chan–Vese image segmentation model," *Int. J. Comput. Vis.*, vol. 98, no. 1, pp. 103–121, 2012.

20. E. Bae, J. Yuan, and X.-C. Tai, "Simultaneous convex optimization of regions and region parameters in image segmentation models," in *Innovations for Shape Analysis: Models and Algorithms*, M. Breuß, A. Bruckstein, and P. Maragos, Eds. 2013, pp. 421–438, Springer Berlin Heidelberg.

21. E. Bae and X.-C. Tai, "Efficient global minimization methods for image segmentation models with four regions," *J. Math. Imaging and Vision*, vol. 51, no. 1, pp. 71–97, Jan. 2015.

22. T. Pock, D. Cremers, H. Bischof, and A. Chambolle, "Global solutions of variational models with convex regularization," *SIAM J. Imaging Sci.*, vol. 3, no. 4, pp. 1122–1145, 2010.

23. L. Condat, "Fast projection onto the simplex and the l1 ball," *Math. Program. Series A*, vol. 158, no. 1, pp. 575–585, July 2016.

24. A. Quattoni, X. Carreras, M. Collins, and T. Darrell, "An efficient projection for l1,∞ regularization," in *Proc. of ICML*, Montreal, Canada, June 2009, pp. 857–864.

25. H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer, New York, 2011.

26. J. Yuan, E. Bae, Y. Boykov, and X.-C. Tai, "A continuous max-flow approach to minimal partitions with label cost prior," in *Proc. of Scale Space and Variational Methods in Computer Vision (SSVM)*, 2011, pp. 279–290.

27. A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *J. Math. Imaging Vision*, vol. 40, no. 1, pp. 120–145, May 2011.

28. L. Condat, "A primal-dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms," *J. Optim. Theory Appl.*, vol. 158, no. 2, pp. 460–479, 2013.

29. J. Reese, "Solution methods for the p-median problem: An annotated bibliography," *Networks*, vol. 48, no. 3, pp. 125–142, Oct. 2006.

30. M. L. Balinski, "On finding integer solutions to linear programs," in *Proceedings of the I.B.M. Scientific Computing Symposium on Combinatorial Problems*, 1966, pp. 225–248.

31. S. Li and O. Svensson, "Approximating k-median via pseudo-approximation," in *Proc. of the forty-fifth annual ACM symposium on Theory of computing (STOC'13)*, Palo Alto, California, USA, June 2013, pp. 901–910.

32. M. Van der Laan, K. Pollard, and J. Bryan, "A new partitioning around medoids algorithm," *Journal of Statistical Computation and Simulation*, vol. 73, no. 8, pp. 575–584, 2003.

33. J. Peng and Y. Xia, "A new theoretical framework for K-means-type clustering," in *Foundations and Advances in Data Mining*, W. Chu and T. Young Lin, Eds., pp. 79–96. Springer Berlin Heidelberg, 2005.

34. J. Peng and Y. Wei, "Approximating K-means-type clustering via semidefinite programming," *SIAM J. Optim.*, vol. 18, no. 1, pp. 186–205, 2007.

35. P. Awasthi, A. S. Bandeira, M. Charikar, R. Krishnaswamy, S. Villar, and R. Ward, "Relax, no need to round: Integrality of clustering formulations," in *Proc. of the 2015 Conference on Innovations in Theoretical Computer Science (ITCS)*, Rehovot, Israel, Jan. 2015, pp. 191–200.

36. K. Pelckmans, J. De Brabanter, J.A.K. Suykens, and B. De Moor, "Convex clustering shrinkage," in *Proc. of Workshop on Statistics and Optimization of Clustering Workshop (PASCAL)*, London, UK, July 2005.

37. T. Hocking, J.-P. Vert, F. Bach, and A. Joulin, "Clusterpath: an algorithm for clustering using convex fusion penalties," in *Proc. of the 28th International Conference on Machine Learning (ICML)*, Bellevue, USA, June 2011, pp. 745–752.

38. F. Lindsten, H. Ohlsson, and L. Ljung, "Clustering using sum-of-norms regularization: With application to particle filter output computation," in *Proc. of Statistical Signal Processing Workshop (SSP)*, Nice, France, June 2011, pp. 201–204.

39. C. Zhu, H. Xu, C. Leng, and S. Yan, "Convex optimization procedure for clustering: theoretical revisit," in *Proc. of NIPS*, Montreal, Canada, Dec. 2014, pp. 1619–1627.

40. E. C. Chi and K. Lange, "Splitting methods for convex clustering," *J. Comput. Graph. Stat.*, vol. 24, no. 4, pp. 994–1013, 2015.

41. I. Kärkkäinen and P. Fränti, "Dynamic local search algorithm for the clustering problem," Tech. Rep. A-2002-6, Department of Computer Science, University of Joensuu, Joensuu, Finland, 2002.

42. E. S. Barnes and N. J. A. Sloane, "The optimal lattice quantizer in three dimensions," *SIAM. J. on Algebraic and Discrete Methods*, vol. 4, no. 1, pp. 30–41, 1983.