A Convexly Constrained LiGME Model and Its Proximal Splitting Algorithm

Wataru Yata, Masao Yamagishi, and Isao Yamada
Dept. of Information and Communications Engineering
Tokyo Institute of Technology, Tokyo, Japan
Email: {yata,myamagi,isao}@sp.ict.e.titech.ac.jp

Abstract—For the sparsity-rank-aware least squares estimations, the LiGME (Linearly involved Generalized Moreau Enhanced) model was established recently in [Abe, Yamagishi, Yamada, 2020] to use certain nonconvex enhancements of linearly involved convex regularizers without losing their overall convexities. In this paper, for further advancement of the LiGME model by incorporating multiple a priori knowledge as hard convex constraints, we newly propose a convexly constrained LiGME (cLiGME) model. The cLiGME model can utilize multiple convex regularizers without losing their overall convexity of the problem (2), the LiGME model can admit multiple linearly involved convex functions \( \Psi \), satisfying coercivity (i.e., \( \lim_{\|z\| \to \infty} \Psi(z) = \infty \)) and \( \text{dom} (\Psi) = Z \), where \( \Psi_B \) is the generalized Moreau enhanced (GME) penalty

\[
\Psi_B(\cdot) := \Psi(\cdot) - \min_{v \in Z} \left[ \Psi(v) + \frac{1}{2} \| B (\cdot - v) \|^2_2 \right]
\]

with a finite dimensional real Hilbert space \( Z \) and a linear operator \( B : Z \to \tilde{Z} \). Indeed, if \( A^* A - \mu B^* B \leq 0 \), the overall convexity of \( J_{\Psi B \circ \mathcal{L}} \) is achieved. Compared with its special case \( (\Psi, \mathcal{L}) = (\| \cdot \|_1, \text{Id}) \) [17], [24] and variants [23], the LiGME model can admit multiple linearly involved convex constraints \( \Psi \).

1 A convex function \( \Psi : Z \to (-\infty, \infty] \) is said to be \( \Psi \in \Gamma_0(Z) \) if \( \Psi \) is proper (i.e., \( \text{dom}(\Psi) := \{ z \in Z | \Psi(z) < \infty \} \neq \emptyset \) and lower semicontinuous (i.e., \( (\forall a \in \mathbb{R}) \text{lev}_{\leq a} \Psi := \{ z \in Z | \Psi(z) \leq a \} \) is closed).

For \( \Psi \in \Gamma_0(Z) \), the proximity operator of \( \Psi \) is defined by

\[
\text{Prox}_\Psi : Z \to Z : x \mapsto \arg\min_{y \in Z} \left[ \Psi(y) + \frac{1}{2} \| x - y \|^2_Z \right].
\]
nonconvex penalties, by formulating it in a product space, without losing the overall convexity of $J_{\Psi B:L}$ and thus is applicable to a broader range of scenarios in the sparsity-rank-aware signal processing.

In contrast, as seen in the set theoretic estimation [4], [5], [19], [22], if we know a priori that our target vector $x^*$ in (1) satisfies $c_i x^* \in C_i$ ($i \in I$), where $I$ is a finite index set, then $C_i$ ($i \in I$) is a closed convex subset of a finite dimensional real Hilbert space $\mathcal{H}$, and $c_i : \mathcal{X} \rightarrow \mathcal{H}$ is a linear operator, we can incorporate the estimation achieved via the model (3) by incorporating multiple convex constraints as

$$
\min_{(i \in I)} \left\{ J_{\Psi B:0:L}(x) := \frac{1}{2} \| Ax - y \|^2_Y + \mu \Psi_B \circ \mathcal{L}(x) \right\}.
$$

(5)

Note that, as we will see in Lemma 1 the model (3) itself does not cover the model (5) because the condition $\text{dom}(\Psi) = Z$ does not allow to use, in the model (3), the indicator function

$$
\iota_{C_i}(x) := \begin{cases} 0 & (x \in C_i) \\ \infty & \text{(otherwise)}. \end{cases}
$$

In this paper, we call the problem (5) a convexly constrained Linearly involved Generalized Moreau Enhanced (cLiGME) model. This problem can be reformulated as simpler models (3) and (7). The overall convexity of $J_{\Psi B:0:L}$ is guaranteed, by Proposition 1 under a certain condition regarding a tunable matrix $B$. We also present a way to design such a matrix $B$ satisfying the overall convexity condition (see Proposition 2 and Proposition 3). Moreover, as an extension of [1] Theorem 1, we present a novel proximal splitting type algorithm of guaranteed convergence to a global minimizer of the function $J_{\Psi B:0:L}(x)$; see Algorithm

$$
\min_{x \in \mathcal{P}} \frac{1}{2} \| y - Ax \|^2_Y + \mu \Psi_B \circ \mathcal{L}(x).
$$

(6)

Remark 1 (The model (5) is a specialization of the model (6)). By using $(3_i, C_i, \bar{C}_i)$ ($i \in I$) in [3], define a new real Hilbert spaces $\mathcal{H}_i := \mathcal{X} \times \mathcal{Z}$, a new linear operator $c_i : \mathcal{X} \rightarrow \mathcal{H}_i : x \mapsto (c_i x, c(y))$ and $B_{c_i} : B + \mathcal{O}_3 : \mathcal{C} \rightarrow \mathcal{Z} = (\mathcal{Z}(122), \mathcal{Z}(123))$, and a new convex function $\Psi_i := \Psi \circ \mathcal{L} \in \Gamma(0, \mathcal{Z}_c)$ by $\Psi_i : \mathcal{Z} \times 3 \rightarrow (\mathcal{H}_i)_{x \in I}$ and a new closed convex set $C_i := \mathcal{X} \times \mathcal{C}$. Then, we have the equivalence

$$
\mathcal{C} \times \mathcal{Z}_c \equiv (\forall i \in I) \mathcal{C}_i \times \mathcal{Z}_c.
$$

We use the following constraint-free reformulation of the model (6).

Lemma 1 (A constraint-free reformulation of the cLiGME model). In Problem 7 define new Hilbert spaces $\mathcal{Z}_c := \mathcal{H} \times 3$ and $\mathcal{Z} := \mathcal{H} \times 3$, new linear operators $B_{c_i} : \mathcal{X} \rightarrow \mathcal{Z} \times 3 : \mathcal{Z} \rightarrow \mathcal{Z} = (\mathcal{Z}(122), \mathcal{Z}(123))$, and a new convex function $\Psi_i := \Psi \circ \mathcal{L} \in \Gamma(0, \mathcal{Z}_c)$ by $\Psi_i : \mathcal{Z} \times 3 \rightarrow (\mathcal{H}_i)_{x \in I}$ and a new closed convex set $C_i := \mathcal{X} \times \mathcal{C}$. Then, the cLiGME model (6) can be formulated equivalently as

$$
\min_{x \in \mathcal{X}} \frac{1}{2} \| y - Ax \|^2_Y + \mu \Psi_B \circ \mathcal{L}(x),
$$

(7)

where

$$
\Psi_{B_{c_i}}(\cdot) := \Psi(\cdot) - \min_{v \in \mathcal{Z}_c} \left\{ \Psi(v) + \frac{1}{2} \left\| B_{c_i} \cdot (v - \cdot) \right\|_{\mathcal{Z}_c}^2 \right\}.
$$

(8)

(9)

The overall convexity condition for the cLiGME model is given below.

Proposition 1 (Overall convexity condition for the cLiGME model). Let $A, B, \mu \in B(\mathcal{X}, \mathcal{Y}) \times B(\mathcal{Z}, \mathcal{Z}) \times \mathcal{R}_{++}$. For three conditions ($C_1$) $A^* A - \mu \mathcal{S}_B B^* B \succeq \mathcal{O}_N$, ($C_2$) $J_{\Psi B:0:L} \min_{x \in \mathcal{X}} \left\{ \frac{1}{2} \| A \|^2_2 + \mu \Psi_B \circ \mathcal{L} \right\}$, ($C_3$) $J_{\Psi B:0:L} \min_{x \in \mathcal{X}} \left\{ \frac{1}{2} \| A \|^2_2 + \mu \Psi_B \circ \mathcal{L} \right\}$, the relation ($C_1$) $\iff$ ($C_2$) $\iff$ ($C_3$) holds.

$^2$A new Hilbert space $\mathcal{H}_i := \mathcal{X} \times \mathcal{C}_i \times \mathcal{Z}_i$ is equipped with the addition $\mathcal{H}_i \times \mathcal{H}_i := \mathcal{X} \times \mathcal{Y} \times \mathcal{X} \times \mathcal{Y}$, the scalar product $\mathcal{H}_i \times \mathcal{H}_i := \mathcal{X} \times \mathcal{Y}$, the inner product $\langle \cdot, \cdot \rangle_L : \mathcal{H}_i \times \mathcal{H}_i \rightarrow \mathcal{C}_i \times \mathcal{Y}$, and its induced norm $\| \cdot \|_L : \mathcal{H}_i \rightarrow \mathcal{R}_{++}$. For a matrix $A, A^T$ denotes the transpose of $A$ and $A^\dagger$ the Moore-Penrose pseudoinverse of $A$. We use $I_m \in \mathcal{R}_{m \times n}$ to denote the identity matrix. We also use $O_{m \times n} \in \mathcal{R}_{m \times n}$ for the zero matrix.

III. cLiGME MODEL AND ITS PROXIMAL SPLITTING ALGORITHM

We start with some reformulations of (5).

A. cLiGME model

Problem 1 (cLiGME model). Let $(\mathcal{X}, \langle \cdot, \cdot \rangle_{\mathcal{X}}, \| \cdot \|_{\mathcal{X}})$, $(\mathcal{Y}, \langle \cdot, \cdot \rangle_{\mathcal{Y}}, \| \cdot \|_{\mathcal{Y}})$, $(\mathcal{Z}, \langle \cdot, \cdot \rangle_{\mathcal{Z}}, \| \cdot \|_{\mathcal{Z}})$, $(\mathcal{Z}, \langle \cdot, \cdot \rangle_{\mathcal{Z}}, \| \cdot \|_{\mathcal{Z}})$, and $(\mathcal{Z}, \langle \cdot, \cdot \rangle_{\mathcal{Z}}, \| \cdot \|_{\mathcal{Z}})$ be finite dimensional real Hilbert spaces and $\mathcal{C} \subset \mathcal{Z}$ be a nonempty closed convex set. Let $\Psi \in \Gamma(0, \mathcal{Z})$ be coercive with $\text{dom} \Psi = \mathcal{Z}$. Let $(A, B, \mathcal{L}, \mathcal{C}, \mu) \in B(\mathcal{X}, \mathcal{Y}) \times B(\mathcal{Z}, \mathcal{Z}) \times B(\mathcal{C}, \mathcal{Z})$ and $\mathcal{C} \cap \mathcal{K} \neq \emptyset$. Then, we consider a constrained cLiGME (cLiGME) model:

$$
\min_{x \in \mathcal{X}} \frac{1}{2} \| y - Ax \|^2_Y + \mu \Psi_B \circ \mathcal{L}(x).
$$

(6)
B. A proximal splitting algorithm for cLiGME model

Our target is the following convex optimization problem.

**Problem 2.** Suppose $\Psi \in \Gamma_0$ satisfies even symmetry $\Psi \circ (-\text{Id}) = \Psi$ and is proximizable (i.e., $\text{Prox}_{\gamma \Psi}$ is available as a computable operator for every $\gamma \in \mathbb{R}_{++}$) and assume $\mathcal{C} \subset \mathcal{Z}$ is a closed convex subset onto which the metric projection $P_{\mathcal{C}}(x) : \mathcal{Z} \mapsto \mathcal{C}$ is computable. Then, for $(A, \mathcal{L}, \mathcal{B}, \mathcal{C}, y, \mu) \in \mathcal{B}(\mathcal{X}, \mathcal{Y}) \times \mathcal{B}(\mathcal{X}, \mathcal{Z}) \times \mathcal{B}(\mathcal{Z}, \mathcal{E}_c) \times \mathcal{Y} \times \mathbb{R}_{++}$ satisfying $A^* - \mu B^* B \mathcal{L} \geq O_{\mathcal{X}}$ and $0, \mathcal{C} \in \mathcal{C} = \mathcal{C}(\mathcal{X} - \mathcal{Y})$, find $x^* \in S := \arg \min_{x \in \mathcal{X}} J_{\Psi \cdot \mu \mathcal{L} \cdot \mathcal{C}}(x) = \arg \min_{x \in \mathcal{X}} J_{\Psi \cdot \mu \mathcal{L} \cdot \mathcal{C}}(x)$.

The next theorem presents an iterative algorithm of guaranteed convergence to a global minimizer of Problem 2.

**Theorem 1** (Averaged nonexpansive operator $T_{\text{cLiGME}}$ and its fixed point approximation). In Problem 2 let $(\mathcal{H} := \mathcal{X} \times \mathcal{Z} \times \mathcal{E}_c, \langle \cdot, \cdot \rangle_{\mathcal{H}}, \| \cdot \|_{\mathcal{H}})$ be a real Hilbert space, Define $T_{\text{cLiGME}} : \mathcal{H} \mapsto \mathcal{H} : (v, x, w) \mapsto (\xi, \zeta, \eta)$ with $(\sigma, \tau, \kappa) \in \mathbb{R}_{++} \times \mathbb{R}_{++}$ by

$$\xi := \left( \begin{array}{c}
\text{Id} - \frac{1}{\sigma} (A^* A - \mu \mathcal{L}^* B \mathcal{L}) \quad \eta := (\text{Id} - \text{Prox}_{\Psi \cdot \mu \mathcal{L} \cdot \mathcal{C}}) \left( 2 \mathcal{L} \xi - \mathcal{L} x + w \right),
\end{array} \right)$$

where $\text{Prox}_{\Psi \cdot \mu \mathcal{L} \cdot \mathcal{C}}(w_1, w_2) = (\text{Prox}_{\Psi}(w_1), P_{\mathcal{C}}(w_2))$. Then, (a) The solution set $S$ of Problem 2 can be expressed as $S = \Xi (\text{Fix}(T_{\text{cLiGME}}))$ and $\Xi : \mathcal{H} \mapsto \mathcal{X} \times \mathcal{Z} \times \mathcal{E}_c$ is defined by $(x, v, w) \mapsto (x, v, w)$ and $T_{\text{cLiGME}} := \{(x, v, w) \in \mathcal{H} | (x, v, w) = T_{\text{cLiGME}} (x, v, w)\}$.

(b) Choose $(\sigma, \tau, \kappa) \in \mathbb{R}_{++} \times \mathbb{R}_{++} \times (1, \infty)$ satisfying

$$\left\{ \begin{array}{l}
\text{Id} - \frac{\kappa}{\sigma} A^* A - \mu \mathcal{L}^* \mathcal{L} \geq O_{\mathcal{X}} \quad \tau \geq \frac{4}{2 + \kappa} \mu \| B \|^2_{\text{op}},
\end{array} \right.$$ \hspace{1cm} (10)

Then,

$$\Psi := \left[ \begin{array}{c}
\text{Id} - \frac{\kappa}{\sigma} A^* A - \mu \mathcal{L}^* \mathcal{L} \quad \tau \text{Id} \\
- \mu B^* B \mathcal{L} \quad \text{Id}
\end{array} \right] > O_{\mathcal{H}} \quad \text{with} \quad \Psi := \left[ \begin{array}{c}
\text{Id} - \frac{\kappa}{\sigma} A^* A - \mu \mathcal{L}^* \mathcal{L} \quad \tau \text{Id} \\
- \mu B^* B \mathcal{L} \quad \text{Id}
\end{array} \right] > O_{\mathcal{H}}$$ \hspace{1cm} (11)

Algorithm 1 below is made based on Theorem 1.

Algorithm 1 for Problem 2

Choose $(x_0, v_0, w_0) \in \mathcal{H}$. Let $(\sigma, \tau, \kappa) \in \mathbb{R}_{++} \times \mathbb{R}_{++} \times (1, \infty)$ satisfying (10).

Define $\Psi$ as (11). Choose $(\sigma, \tau, \kappa) \in \mathbb{R}_{++} \times \mathbb{R}_{++} \times (1, \infty)$ satisfying (10).

Choose a nonsingular $\mathcal{L} \in \mathbb{R}^{n \times n}$ satisfying $\text{rank} \mathcal{L} = l$. Choose a nonsingular $\mathcal{L} \in \mathbb{R}^{n \times n}$ satisfying

C. How to choose $B$ to achieve the overall convexity

A matrix $B$ achieving the overall convexity of $J_{\Psi \cdot \mu \mathcal{L} \cdot \mathcal{C}}$ can be designed in exactly same way as [1] Proposition 2 for (7).

**Proposition 2** (A design of $B$ to ensure the overall convexity condition in Proposition 1). In Problem 2 let $(\mathcal{X}, \mathcal{Y}, \mathcal{Z}) = (\mathbb{R}^n, \mathbb{R}^m, \mathbb{R}^l)$, $(A, \mathcal{L}, \mu) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times \kappa} \times \mathbb{R}_{++}$ and $\text{rank} \mathcal{L} = l$. Choose a nonsingular $\mathcal{L} \in \mathbb{R}^{n \times n}$ satisfying $\text{rank} \mathcal{L} = l$. Then,

$$B = B_{\theta} := \sqrt{\theta} \mu A^{1/2} U^T \in \mathbb{R}^{l \times l}, \theta \in [0, 1]$$ \hspace{1cm} (12)

ensures $J_{\Psi \cdot \mu \mathcal{L} \cdot \mathcal{C}} \in \Gamma_0(\mathbb{R}^n)$, where

$$[\tilde{A}_1 \quad \tilde{A}_2] = A(\mathcal{L})^{-1} \quad \text{and} \quad U \tilde{A} \tilde{A}^T := \tilde{A}_2 \tilde{A}_2 + \tilde{A}_2 \tilde{A}_1 \left( \tilde{A}_1 \tilde{A}_1 \right)^{-1} \tilde{A}_1 \tilde{A}_2 \quad \tilde{A}_2 \text{is an eigendecomposition}$$ \hspace{1cm} (13)

In the following proposition, (a) shows that the cLiGME model (6) admits multiple penalties and (b) shows how to design $B_{\theta}$ for multiple penalties.

**Proposition 3** (Multiple penalties for the cLiGME model).
(a) The sum of multiple LiGME penalties can be modeled as a single cLiGME penalty on product space. Let \( Z_i \) and \( \tilde{Z}_i \) be finite dimensional real Hilbert spaces \((i = 1, 2, 3, \ldots, M)\). Define a new real Hilbert spaces \( Z := \bigotimes_{i=1}^{M} Z_i \) and \( \tilde{Z} := \bigotimes_{i=1}^{M} \tilde{Z}_i \). A new linear operator \( B : Z \to X \) and \( \tilde{B} : \tilde{Z} \to \tilde{X} \) are defined with considering additional multiple convex constraints, respectively (see Proposition 1). Then, we have

\[
\Psi_{B_x} \circ \mathcal{L}_c = \left( \sum_{i=1}^{M} \mu_i \left( \Psi^{(i)}_{B^{(i)}} \right) \right) \circ \mathcal{L}_c + (i\mathcal{C} \circ \mathcal{C}). \tag{14}
\]

(b) A design of \( B^{(i)} \) to achieve optimal spaceabilities in Prop. 1. Let \((X, Y, Z_1, Z_2) = (\mathbb{R}^n, \mathbb{R}^m, \mathbb{R}^1)\). \((A, Z_1, \mu_i) \in \mathcal{B}(\mathbb{R}^n, \mathbb{R}^m) \times \mathcal{B}(X, Y, Z_1) \times \mathbb{R} \) and \( rank(Z_1) = i \) \((i = 1, 2, 3, \ldots, M)\). For each \( i = 1, 2, \ldots, M \), choose nonsingular \( \Sigma_1 \in \mathbb{R}^{n \times n} \) satisfying \( X \in \mathbb{R}^n \times (n-i) \mathcal{I} \), \( \bar{Z}_i \) and \( \omega_i \in \mathbb{R}^{n+i} \) satisfying \( \sum_{i=1}^{M} \omega_i = 1 \). For each \( i = 1, 2, \ldots, M \), apply Proposition 2 to \( \left( \sqrt{\mu_i} A, \Sigma_1, \mu_i \right) \) to obtain \( B^{(i)} \in \mathbb{R}^{n \times 1} \) satisfying \( \sqrt{\mu_i} A - \mu_i \Sigma_1 \geq X \). Then, for \( \theta = (\theta_1, \theta_2, \ldots, \theta_M) \), \( B_\theta : \bigotimes_{i=1}^{M} \mathbb{R}^1 \to \bigotimes_{i=1}^{M} \mathbb{R}^{1+i} \) ensures \( J_{\Psi_{B_\theta} \circ \mathcal{L}_c} \in \Gamma_0(\mathbb{R}^n) \) (see Lemma 7).

IV. NUMERICAL EXPERIMENTS

We conducted numerical experiments based on a scenario of image restoration for piecewise constant \( N \)-by-\( N \) image, which is the same as [1] Sec. 4.2 for space saving purpose but with considering additional multiple convex constraints. Set \( (X, Y, Z_1, Z_2) = (\mathbb{R}^{N^2}, \mathbb{R}^{N^2}, \mathbb{R}^{N(N-1)}, \mathbb{R}^{N(N-1)}) \). \( N = 16 \), \( \Psi^{(1)} = \Psi^{(2)} = L1 \) \( \mu_1 = \mu_2 = 1 \) and \( \mathcal{D} := \mathcal{D} := \mathcal{D}_V := \mathcal{D}_{\mathcal{D}_V} \) \( \mathcal{D}_V \in \mathbb{R}^{N(N-1) \times N^2} \) is the horizontal difference operator and \( \mathcal{D}_V \in \mathbb{R}^{N(N-1) \times N^2} \) is the horizontal difference operator used in [1] (43), respectively (see Proposition 3[4]). The blur matrix \( A \in \mathbb{R}^{N^2 \times N^2} \) is also defined as in [1] (44) and (45). The observation vector \( y \in \mathbb{R}^{N^2} \) is assumed to satisfy the linear regression model: \( y = Ax + \varepsilon \), where \( x^* \in \mathbb{R}^{N^2} \) is the vectorization\(^5\) of a piecewise constant image shown in Figure 3(a) and \( \varepsilon \in \mathbb{R}^{N^2} \) is the additive white Gaussian noise. As available a priori knowledge is found, e.g., in blind deconvolution \([10]\), the signal-to-noise ratio (SNR) is set by \( 10 \log_{10} \left( \frac{\|x^*\|}{\|\varepsilon\|} \right) = 20 \) [dB]. By letting \( X := \mathcal{Z}_1 = \mathcal{Z}_2, \mathcal{C}_1 := \mathcal{I}_2, \mathcal{C}_2 := \left[ \mathcal{C}_T^T, \mathcal{C}_2^T \right]^T \),

\[
C_1 := \left\{ x \in \mathbb{R}^{N^2} : \|x - \mathcal{Z}_1\| \leq 0.25 \right\},
\]

\[
C_2 := \left\{ \text{vec}(X) \in \mathbb{R}^{N^2} : \|x - \mathcal{Z}_2\| \leq 3 \right\}.
\]

Figure 3 shows the dependency of recovery performance on the parameter \( \mu \) in Problem 2. The mean squared error (MSE) is defined by the average of squared error (SE): \( \|x_k - x^*\|_X^2 \) over 100 independent realizations of the additive white Gaussian noise \( \varepsilon \). We see that constraints are effective to improve the estimation and cLiGME model in (16) can achieve better estimation than TV model in (15).

Figure 2 shows convergence performances observed over 5000 iterations. Common weights are used as \( \mu = \mu_{TV} := 0.013 \) for (15) and \( \mu = \mu_{cLiGME} := 0.03 \) for (16), where these

\(^5\)The vectorization of a matrix is the mapping:

\[
\text{vec} : \mathbb{R}^{m \times n} \to \mathbb{R}^m : A \mapsto \left[ a_1^T, a_2^T, \ldots, a_n^T \right]^T,
\]

where \( a_i \) is the \( i \)-th column vector of \( A \) for each \( i \in \{ 1, 2, \ldots, n \} \).
We proposed a convexly constrained LiGME (cLiGME) model to broaden applicability of the LiGME model. The cLiGME model allows to use multiple convex constraints. We also proposed a proximal splitting type algorithm for cLiGME model. Numerical experiments show the efficacy of the proposed model and the proposed algorithm.

REFERENCES

[1] J. Abe, M. Yamagishi, and I. Yamada, "Linearly involved generalized Moreau enhanced models and their proximal splitting algorithm under overall convexity condition," Inverse Problems, vol. 36, no. 3, 035012(36pp.), 2020. (See also [arXiv:1910.10337])
[2] A. Blake and A. Zisserman, Visual Reconstruction. MIT Press, 1987.
[3] E. J. Candès, J. K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," Commun. Pure Appl. Math., vol. 59 no. 8, pp. 1207–1223, 2006.
[4] Y. Censor, S. A. Zenios, Parallel Optimization: Theory, Algorithm, and Applications. Oxford University Press, 1997.
[5] P. L. Combettes, "Foundation of set theoretic estimation," Proc. IEEE, vol. 81, no. 2, pp. 182–208, 1993.
[6] D. L. Donoho, "Compressed sensing," IEEE Trans. Inform. Theory, vol. 52, no. 4, pp. 1289–1306, 2006.
[7] M. Elad, Sparse and Redundant Representations, Springer, 2010.
[8] G. T. Herman, Image Reconstruction from Projections. Academic Press, 1980.
[9] I. Daubechies, M. Defrise and C. Mol, "An Iterative Thresholding Algorithm for Linear Inverse Problems with a Sparsity Constraint," Commun. on Pure Appl. Math., vol. 57, pp. 1413–57, 2004.
[10] D. Kundur and D. Hatzinakos, "A novel blind deconvolution scheme for image restoration using recursive filtering," IEEE Trans. Signal Process., vol. 46, no. 2, pp. 375-390, 1998.
[11] B. K. Natarajan, "Sparse approximate solutions to linear systems," SIAM J. Comput., vol. 24, no. 2, pp. 227–234, 1995.
[12] T. T. Nguyen, C. Soussen, J. Idier and E. Djerroune, "NP-hardness of f0 minimization problems: revision and extension to the non-negative setting," 2019 13th International conference on Sampling Theory and Applications (SampTA), Bordeaux, France, 2019, pp. 1-4.
[13] M. Nikolova, "Estimation of binary images by minimizing convex criteria," In IEEE Trans. Image Process., pages 108–112, 1998.
[14] M. Nikolova, "Markovian reconstruction using a GNC approach," IEEE Trans. Image Process., vol. 8, no. 9, pp. 1204–1220, 1999.
[15] L. I. Rudin, S. Osher and E. Fatemi, "Nonlinear total variation based noise removal algorithms," Physica D: Nonlinear Phenomena, vol. 60, no. 1–4, pp. 259-268 1992.
[16] M. Nikolova, "Energy minimization methods," In O. Scherzer, editor, Hand-book of Mathematical Methods in Imaging, pages 138–186, Springer, 2011.
[17] I. Selesnick, "Sparse regularization via convex analysis," IEEE Trans. Signal Process., vol. 65 no. 17, pp. 4481–4494, 2017.
[18] J. L. Starck, F. Murtagh and J. Fadili, Sparse Image and Signal Processing: Wavelets and Related Geometric Multiscale Analysis, 2nd ed., Cambridge: Cambridge University Press, 2015.
[19] H. Stark, Image Recovery: Theory and Application. Academic Press, 1987.
[20] S. Theodoridis, Machine Learning: A Bayesian and Optimization Perspective. 2nd ed., Academic Press, 2020.
[21] R. Tibshirani, "Regression shrinkage and selection via the lasso," J. R. Statist. Soc. B, vol. 58, no. 1, pp. 267–288, 1996.
[22] D. C. Youla and H. Webb. "Image Restoration by the Method of Convex Projections: Part 1-Theory," IEEE Trans. Med. Imag., vol. 1, no. 2, pp. 81-94, 1982.
[23] L. Yin, A. Parekh, and J. Selesnick, "Stable principal component pursuit via convex analysis," IEEE Trans. Signal Process., vol. 67, no. 10, pp. 2595–2607, 2019.
[24] C.-H. Zhang, "Nearly unbiased variable selection under minimax concave penalty," Ann. Statist., vol. 38, no. 2, pp. 894–942, 2010.