A workshop on optimization and operator theory
dedicated to Prof. Lev Bregman (Nov. 15 -17, 2021)

A constrained LiGME model for sparsity-rank aware least squares estimation problems*

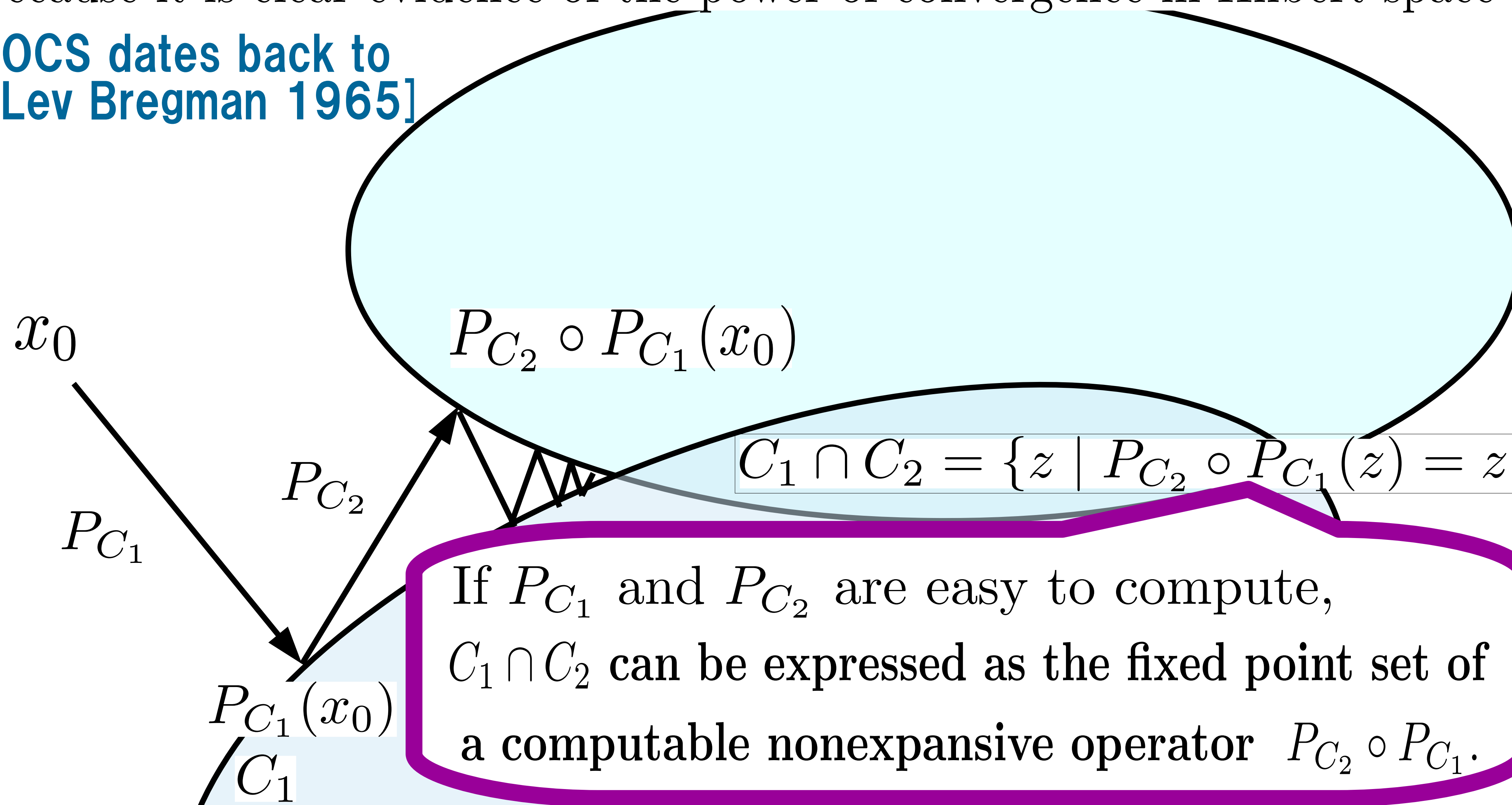
Isao Yamada

Tokyo Institute of Technology

* Based on a joint work with W. Yata & M. Yamagishi

My lecture "Applied Functional Analysis" in Tokyo Tech starts with POCS because it is clear evidence of the power of convergence in Hilbert space !

POCS dates back to [Lev Bregman 1965]



If P_{C_1} and P_{C_2} are easy to compute, $C_1 \cap C_2$ **can be expressed as the fixed point set of a computable nonexpansive operator** $P_{C_2} \circ P_{C_1}$.

The operator $P_{C_2} \circ P_{C_1}$ is **nonexpansive**, i.e.,

$$(x, y \in \mathcal{H}) \quad \|P_{C_2} \circ P_{C_1}(x) - P_{C_2} \circ P_{C_1}(y)\| \leq \|P_{C_1}(x) - P_{C_1}(y)\| \leq \|x - y\|$$

Many tasks in sparsity-rank-aware signal processing and machine learning have been formulated as

sparsity-rank-aware regularized least squares models

$$\underset{x \in \mathcal{X}}{\text{minimize}}\ J_{\Psi \circ \mathfrak{L}}(x) := \frac{1}{2}\|y - Ax\|_{\mathcal{Y}}^2 + \mu \Psi \circ \mathfrak{L}(x),\ \ \mu > 0, \qquad (1)$$

where $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$: finite dimensional real Hilbert spaces, $y \in \mathcal{Y}$, $A \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ (i.e. A is a bounded linear operator from $\mathcal{X}$ to $\mathcal{Y}$), $\mathfrak{L} \in \mathcal{B}(\mathcal{X}, \mathcal{Z})$ and $\Psi : \mathcal{Z} \to \mathbb{R}_+$ is a certain approximation of $\|\cdot\|_0$ ($\sharp$ of nonzero entries) or $\text{rank}(\cdot)$.

Sparsity-aware convexly regularized least squares

Lasso [Tibshirani `96] ,
TV [ROF`92],
[Daubechies et al `04] ...

→ **Nonconvex regularization via Moreau enhancement**

MC [Zhang `10], GMC [Selesnick `17]
LiGME [AYY `20] ...

→ **This study: cLiGME model**

[YYY MLSP`21]

Set theoretic estimation with convex projections → cLiGME model

[J.von Neumann `30], [Bregman `65], [Youla-Webb`82], [Combettes `93], [Censor-Elfving `94], [Bauschke-Borwein`96], [Deutsch `00], [Byrne `04]...

LiGME model

In sparsity-rank-aware least squares estimation models,

$$\underset{x\in\mathcal{X}}{\text{minimize}}\ J_{\Psi\circ\mathfrak{L}}(x) := \frac{1}{2}\|y - Ax\|_{\mathcal{Y}}^2 + \mu\Psi\circ\mathfrak{L}(x),$$

Best convex approx. of $\|\cdot\|_0$ / $\mathrm{rank}(\cdot)$

$\|\cdot\|_1$ and $\|\cdot\|_{\mathrm{nuc}}$ have been used as the standards of Ψ.

➡ For nonconvex enhancement of $\Psi \in \Gamma_0(\mathcal{Z})$ while achieving the overall convexity of $J_{\Psi\circ\mathfrak{L}}$,

proper+lower semicontinuous +convex functions

LiGME model [J.Abe, M.Yamagishi, IY, Inverse Problems 2020]

$$\underset{x\in\mathcal{X}}{\text{minimize}}\ J_{\Psi_B\circ\mathfrak{L}}(x) := \frac{1}{2}\|y - Ax\|_{\mathcal{Y}}^2 + \mu\Psi_B\circ\mathfrak{L}(x)$$

(B is a tuning parameter for Linearly involved Generalized Moreau Enhancement of Ψ)

NOTE: LiGME extends MC [Zhang '10] and GMC [Selesnick '17] ($\Psi = \|\cdot\|_1$, $\mathfrak{L} = \mathrm{Id}$).

Good News 1

Generalized Moreau enhancement Ψ_B of $\Psi \in \Gamma_0(\mathcal{Z})$ bridges the gap between naive discrete measures and their convex envelopes

Generalized Moreau enhancement of convex regularizers

Let $\mathcal{Z}$ and $\widetilde{\mathcal{Z}}$ be Hilbert spaces.

Then for given $\Psi \in \Gamma_0(\mathcal{Z})$ [prox-friendly, coercive, $\operatorname{dom}\Psi = \mathcal{Z}$] and $B \in \mathcal{B}(\mathcal{Z}, \widetilde{\mathcal{Z}})$,

$$\Psi_B(\cdot) := \Psi(\cdot) - \min_{v \in \mathcal{Z}} \left[\Psi(v) + \frac{1}{2} \| B(\cdot - v) \|_{\widetilde{\mathcal{Z}}}^2 \right],$$ Nonconvex

Moreau-Yosida like regularization of $\Psi(\cdot)$

Example 1 Let $\mathcal{Z} = \widetilde{\mathcal{Z}} := \mathbb{R}^n$, $\Psi := \| \cdot \|_1$ and $B := \frac{1}{\sqrt{\gamma}}\mathrm{Id}$ for $\gamma \in \mathbb{R}_{++}$.

For $(x_1, \ldots, x_n) \in \mathbb{R}^n$, we have
$$\begin{cases} \lim_{\gamma \downarrow 0} \frac{2}{\gamma} (\| \cdot \|_1)_{\frac{1}{\sqrt{\gamma}}\mathrm{Id}} (x_1, \ldots, x_n) = \|(x_1, \ldots, x_n)\|_0, \\ (\| \cdot \|_1)_O (x_1, \ldots, x_n) = \|(x_1, \ldots, x_n)\|_1 \end{cases}$$

Example 2 Let $\mathcal{Z} = \widetilde{\mathcal{Z}} := \mathbb{R}^{m \times n}$, $\Psi := \| \cdot \|_{\mathrm{nuc}}$ and $B := \frac{1}{\sqrt{\gamma}}\mathrm{Id}$ for $\gamma \in \mathbb{R}_{++}$.

For $X \in \mathbb{R}^{m \times n}$, we have
$$\begin{cases} \lim_{\gamma \downarrow 0} \frac{2}{\gamma} (\| \cdot \|_{\mathrm{nuc}})_{\frac{1}{\sqrt{\gamma}}\mathrm{Id}} (X) = \mathrm{rank}(X), \\ (\| \cdot \|_{\mathrm{nuc}})_O (X) = \|X\|_{\mathrm{nuc}} \end{cases}$$

Minimax-Concave (MC) penalty [C.-H. Zhang 2010]

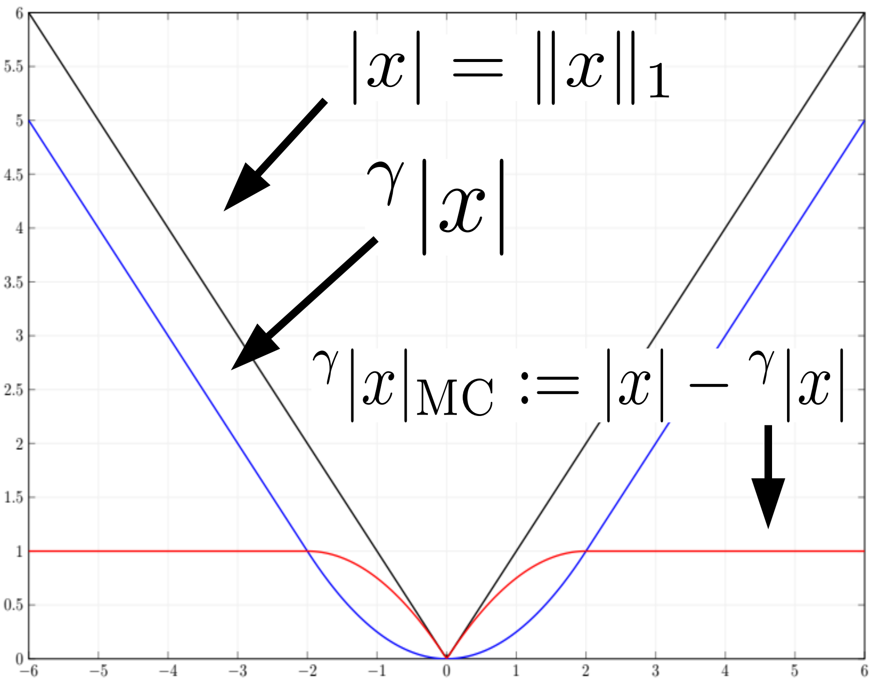
is a simplest 1D example of LiGME function $(\Psi = |\cdot| \text{ and } \mathfrak{L} = \mathrm{Id})$

Moreau envelope of $|\cdot|$

P.J.Huber, Ann. Math. Statist.'64

$$ {}^{\gamma}|x| := \min_{v\in\mathbb{R}} \left[|v| + \frac{1}{2\gamma}|x-v|^2 \right] = \begin{cases} \frac{1}{2\gamma}x^2, & \text{if } |x| \le \gamma \,; \\ |x| - \frac{1}{2}\gamma, & \text{otherwise.} \end{cases} $$

converges pointwise to $|\cdot|$ as $\gamma \downarrow 0$.



C.-H.Zhang, Ann. Statist.'10

Minimax-Concave penalty

$$ {}^{\gamma}|x|_{\mathrm{MC}} := |x| - {}^{\gamma}|x| = \begin{cases} |x| - \frac{1}{2\gamma}x^2, & \text{if } |x| \le \gamma \,; \\ \frac{1}{2}\gamma, & \text{otherwise.} \end{cases} $$

has been proposed as a nearly unbiased **nonconvex enhancement of the best convex sparsity promoting regularizer** l_1-norm $\|\cdot\|_1$

LiGME is a Unified + Linearly involved extension

[Abe-Yamagishi-IY (Inverse Problems '20)]

For $\mathcal{X}, \mathcal{Z}, \widetilde{\mathcal{Z}}$: Hilbert spaces and $\Psi \in \Gamma_0(\mathcal{Z})$ [prox-friendly, coercive, $\operatorname{dom}\Psi = \mathcal{Z}$],

$$\left(B \in \mathcal{B}(\mathcal{Z}, \widetilde{\mathcal{Z}}),\ \mathfrak{L} \in \mathcal{B}(\mathcal{X}, \mathcal{Z})\right) \qquad \Psi_B \circ \mathfrak{L} : \mathcal{X} \to \mathbb{R} \qquad \textbf{(LiGME)}$$

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

Moreau-Yosida like regularization of $\Psi(\cdot)$

Generalized Minimax-Concave (GMC) Penalty

[I. Selesnick *IEEE T-SP,* 2017]

$$(B \in \mathbb{R}^{m\times l}) \qquad (\|\cdot\|_1)_B(\mathbf{z}) := \|\mathbf{z}\|_1 - \min_{\mathbf{v}\in\mathbb{R}^l}\left[\|\mathbf{v}\|_1 + \frac{1}{2}\|B(\mathbf{z}-\mathbf{v})\|^2\right] \qquad \textbf{(GMC)}$$

Minimax-Concave (MC) Penalty [C.-H. Zhang, Ann. Statist.'10]

$$(\gamma \in \mathbb{R}_{++}) \qquad {}^{\gamma}|z|_{\mathrm{MC}} : \mathbb{R} \to \mathbb{R} :\ z \mapsto |z| - \min_{v\in\mathbb{R}}\left[|v| + \frac{1}{2\gamma}|z-v|^2\right] \qquad \textbf{(MC)}$$

Good News 2

Thanks to the great freedom in the choice of $B \in \mathcal{B}(\mathcal{Z}, \widetilde{\mathcal{Z}})$, $\Psi_B \circ \mathfrak{L}$ can achieve flexibly the desired overall convexity !

Linearly involved Generalized Moreau Enhanced model

$$\underset{x \in \mathcal{X}}{\text{minimize}}\ J_{\Psi_B \circ \mathfrak{L}}(x) := \frac{1}{2}\|y - Ax\|^2 + \mu \Psi_B \circ \mathfrak{L}(x), \ \mu > 0, \qquad (2)$$

where $\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \widetilde{\mathcal{Z}}$: Hilbert spaces, $y \in \mathcal{Y}$, $A \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$, $\mathfrak{L} \in \mathcal{B}(\mathcal{X}, \mathcal{Z})$ and

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

with $\Psi \in \Gamma_0(\mathcal{Z})$ [prox-friendly, coercive, $\operatorname{dom}\Psi = \mathcal{Z}$] and $B \in \mathcal{B}(\mathcal{Z}, \widetilde{\mathcal{Z}})$.

Overall Convexity Condition for (2)

$$A^*A - \mu \mathfrak{L}^* B^* B \mathfrak{L} \succeq O \Rightarrow J_{\Psi_B \circ \mathfrak{L}} \in \Gamma_0(\mathcal{X})$$

In particular, if $\Psi \in \Gamma_0(\mathcal{Z})$ satisfies the condition as a norm of vector space $\mathcal{Z}$,

$$A^*A - \mu \mathfrak{L}^* B^* B \mathfrak{L} \succeq 0 \Leftrightarrow J_{\Psi_B \circ \mathfrak{L}} \in \Gamma_0(\mathcal{X})$$ [Abe, Yamagishi, IY (Inverse Problems 2020)]

How can we apply LiGME model ?

$$\underset{x \in \mathcal{X}}{\text{minimize}}\ J_{\Psi_B \circ \mathfrak{L}}(x) := \frac{1}{2}\|y - Ax\|^2 + \mu \Psi_B \circ \mathfrak{L}(x), \ \ \mu > 0, \quad (2)$$

$$\textbf{where} \quad \Psi_B(\cdot) := \Psi(\cdot) - \min_{v \in \mathcal{Z}} \left[\Psi(v) + \frac{1}{2}\|B(\cdot - v)\|^2 \right].$$

Q: Can we choose B flexibly to ensure overall-convexity of (2) ?

Suppose $\mathfrak{L} \in \mathbb{R}^{l \times n}$ **satisfies** $\mathrm{rank}(\mathfrak{L}) = l$.

Choose a nonsingular $\widetilde{\mathfrak{L}} \in \mathbb{R}^{n \times n}$**, s.t.,** $\left[\ O_{l \times (n-l)} \quad \mathbf{I}_l \ \right] \widetilde{\mathfrak{L}} = \mathfrak{L}$.

$B_\theta := \sqrt{\theta/\mu}\Lambda^{1/2}U^T, \ \theta \in [0,1]$, **ensures the convexity of** $J_{\Psi_{B_\theta} \circ \mathfrak{L}}$,

$$\textbf{where } U\Lambda U^T := \widetilde{A}_2^\top \widetilde{A}_2 - \widetilde{A}_2^\top \widetilde{A}_1 \left(\widetilde{A}_1^\top \widetilde{A}_1\right)^\dagger \widetilde{A}_1^\top \widetilde{A}_2 \in \mathbb{R}^{l \times l}$$

$$\textbf{is the EVD with } A\left(\widetilde{\mathfrak{L}}\right)^{-1} = \left[\ \widetilde{A}_1 \quad \widetilde{A}_2 \ \right].$$

J. Abe, M. Yamagishi, I. Yamada,
“Linearly involved generalized Moreau enhanced models and their proximal splitting algorithm under overall convexity condition,” Inverse Problems, (36pp), 2020.

Good News 3

Through a product space reformulation, the LiGME model (2) covers the following seemingly much more general model:

$$\underset{x \in \mathcal{X}}{\text{minimize}} \; J_{\Psi_B \circ \mathfrak{L}}(x) := \frac{1}{2}\|y - Ax\|^2 + \sum_{i=1}^{\mathcal{M}} \mu_i \Psi_{B_i}^{<i>} \circ \mathfrak{L}_i(x) \qquad (3)$$

where $\mathcal{X}, \mathcal{Y}, \mathcal{Z}_i, \widetilde{\mathcal{Z}}_i \; (i = 1, 2, \ldots, \mathcal{M})$, $\mathcal{Z} := \mathcal{Z}_1 \times \cdots \times \mathcal{Z}_{\mathcal{M}}, \; \widetilde{\mathcal{Z}} := \widetilde{\mathcal{Z}}_1 \times \cdots \times \widetilde{\mathcal{Z}}_{\mathcal{M}}$: Hilbert spaces,

$$\Psi_{B_i}^{<i>}(\cdot) := \Psi^{<i>}(\cdot) - \min_{v \in \mathcal{Z}_i} \left[\Psi^{<i>}(v) + \frac{1}{2}\|B_i(\cdot - v)\|^2 \right],$$ Nonconvex

with $\Psi^{<i>} \in \Gamma_0(\mathcal{Z}_i)$ [prox-friendly, coercive, $\operatorname{dom} \Psi^{<i>} = \mathcal{Z}_i$] and $B_i \in \mathcal{B}(\mathcal{Z}_i, \widetilde{\mathcal{Z}}_i)$.

Set theoretic estimation with multiple convex projections - A powerful mathematical idea pioneered by Lev Bregman

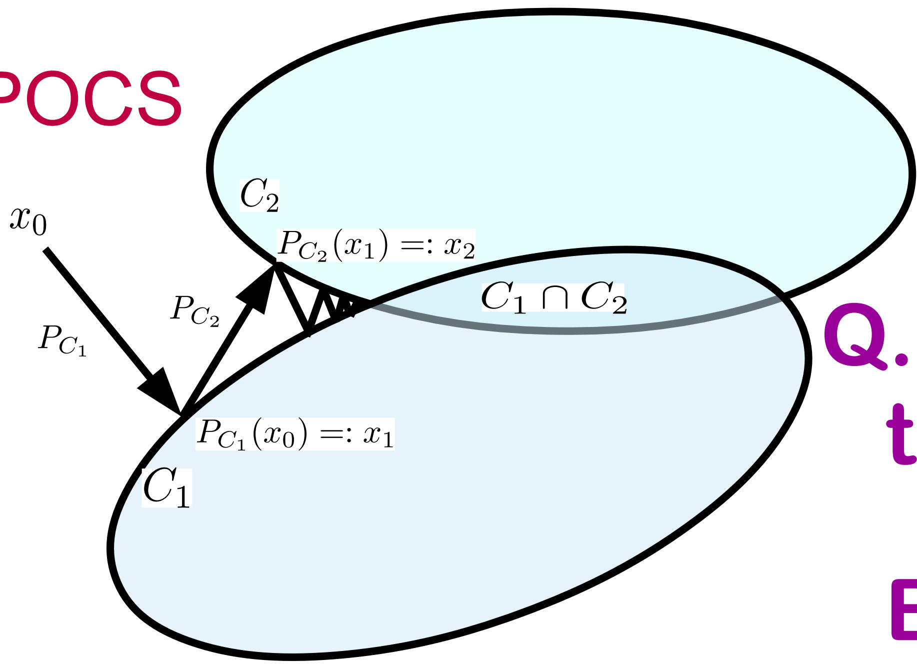
Split convex feasibility problem (e.g., [Censor-Elfving 1994])

$$\text{Find } x^\star \in \mathcal{X} \text{ s.t. } \mathfrak{C}_j x^\star \in C_j \ (1 \leq j \leq N),$$

where $\mathcal{X}, \mathfrak{Z}_j$: real Hilbert spaces, $\mathfrak{C}_j \in \mathcal{B}(\mathcal{X}, \mathfrak{Z}_j)$, and $C_j \subset \mathfrak{Z}_j$ are simple closed convex sets meaning that metric projections P_{C_j} are assumed computable.

Simplest case, where $\mathfrak{Z}_1 = \mathfrak{Z}_2$ and $\mathfrak{C}_1 = \mathfrak{C}_2 = \mathrm{Id}$ dates back to 1965 [Bregman '65].

POCS



Q. Why do not we integrate these two powerful ideas ?

But how ?

cLiGME model (proposed model)

To integrate the LiGME and the Set Theoretic Estimation,
we newly propose

cLiGME model

$$\underset{\mathfrak{C}x\in\mathbf{C}}{\text{minimize}}\ J_{\Psi_B\circ\mathfrak{L}}(x) := \frac{1}{2}\|y - Ax\|_{\mathcal{Y}}^2 + \mu\Psi_B\circ\mathfrak{L}(x),\ \ \mu > 0, \qquad (\star)$$

LiGME regularizer

where $\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \widetilde{\mathcal{Z}}, \mathfrak{Z}$: real Hilbert spaces, $y \in \mathcal{Y}$, $A \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$, $\mathfrak{L} \in \mathcal{B}(\mathcal{X}, \mathcal{Z})$,

$\mathfrak{C} \in \mathcal{B}(\mathcal{X}, \mathfrak{Z})$, $\mathbf{C} \subset \mathfrak{Z}$: simple closed convex, $0_{\mathfrak{Z}} \in \mathrm{ri}(\mathbf{C} - \mathrm{ran}\,\mathfrak{C})$, and

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

Nonconvex

with $\Psi \in \Gamma_0(\mathcal{Z})$[proximable, coercive, even symmetry, $\mathrm{dom}(\Psi) = \mathcal{Z}$] and $B \in \mathcal{B}(\mathcal{Z}, \widetilde{\mathcal{Z}})$.

$\Gamma_0(\mathcal{Z})$: set of all proper lower semicontinuous convex functions over $\mathcal{Z}$, $0_{\mathfrak{Z}}$: zero vector in $\mathfrak{Z}$

The cLiGME $(\star)$ with $\mathfrak{C} = \mathrm{Id}$ and $\mathbf{C} = \mathcal{X}$ reproduces the LiGME model [AYY '20].

At a glance, the model $(\star)$ seems to cover only a single constraint case.

cLiGME covers multiple regularizers and constraints

Through a product space reformulation, the cLiGME model ($\star$) can deal with multiple linearly involved convex constraints

$$\underset{(1\leq j\leq N)\ \mathfrak{C}_j x\in C_j}{\text{minimize}} \quad J_{\Psi_B\circ\mathfrak{L}}(x) := \frac{1}{2}\|y - Ax\|_{\mathcal{Y}}^2 + \sum_{i=1}^{M}\mu_i \Psi_{B^{\langle i\rangle}}^{\langle i\rangle}\circ\mathfrak{L}_i(x)$$

where $\mathcal{X}, \mathcal{Y}, \mathcal{Z}_i, \widetilde{\mathcal{Z}}_i, \mathfrak{Z}_j$: real Hilbert spaces, $y\in\mathcal{Y}$, $A\in\mathcal{B}(\mathcal{X},\mathcal{Y})$, $\mathfrak{L}_i\in\mathcal{B}(\mathcal{X},\mathcal{Z}_i)$,

$\mathfrak{C}_j\in\mathcal{B}(\mathcal{X},\mathfrak{Z}_j)$, $C_j\subset\mathfrak{Z}_j$: simple closed convex, $0_{\mathfrak{Z}_j}\in\mathrm{ri}(C_j - \mathrm{ran}\,\mathfrak{C}_j)$, and

$$\Psi_{B^{\langle i\rangle}}^{\langle i\rangle}(\cdot) := \Psi^{\langle i\rangle}(\cdot) - \min_{v\in\mathcal{Z}_i}\left[\Psi^{\langle i\rangle}(v) + \frac{1}{2}\left\|B^{\langle i\rangle}(\cdot - v)\right\|_{\widetilde{\mathcal{Z}}_i}^2\right]$$

with $\Psi^{\langle i\rangle}\in\Gamma_0(\mathcal{Z}_i)$[proximable, coercive, even symmetry, $\mathrm{dom}(\Psi^{\langle i\rangle}) = \mathcal{Z}_i$] and $B^{\langle i\rangle}\in\mathcal{B}(\mathcal{Z}_i,\widetilde{\mathcal{Z}}_i)$.

Good News 4

$$\underset{\mathfrak{C}x \in \mathbf{C}}{\text{minimize}}\; J_{\Psi_B \circ \mathfrak{L}}(x) := \frac{1}{2}\|y - Ax\|_{\mathcal{Y}}^2 + \mu \Psi_B \circ \mathfrak{L}(x), \;\; \mu > 0, \qquad (\star)$$

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

⇩

Although Ψ_B is nonsmooth and nonconvex, under mild conditions, we can express the set of all globally optimal solutions in terms of the fixed-point set of computable nonexpansive operator in a certain Hilbert space and therefore can solve $(\star)$.

W. Yata, M. Yamagishi, IY,
“A constrained Liniearly-involved-Generalized-Moreau-Enhanced model and its proximal splitting algorithm,” IEEE MLSP, Oct., 2021.

(by extending a theorem in [J. Abe, M. Yamagishi, IY, 2020])

Theorem 1 Assume $\Psi \circ (-\mathrm{Id}) = \Psi$, $1/2\|y - A(\cdot)\|^2 + \mu \Psi_B \circ \mathfrak{L}(\cdot) \in \Gamma_0(\mathcal{X})$. Let $\mathcal{Z}_c := \mathcal{Z} \times \mathfrak{Z}$, $\mathcal{H} := \mathcal{X} \times \mathcal{Z} \times \mathcal{Z}_c$ and $\mathfrak{L}_c : \mathcal{X} \to \mathcal{Z}_c : x \mapsto (\mathfrak{L}(x), \mathfrak{C}(x))$.

Define $T_{\mathrm{cLiGME}} : \mathcal{H} \to \mathcal{H} : (x, v, w) \mapsto (\xi, \zeta, \eta)$ by

$$\xi := \left[\mathrm{Id} - \frac{1}{\sigma}\left(A^*A - \mu \mathfrak{L}^* B^* B \mathfrak{L}\right)\right] x - \frac{\mu}{\sigma}\mathfrak{L}^* B^* B v - \frac{\mu}{\sigma}\mathfrak{L}_c^* w + \frac{1}{\sigma} A^* y$$

$$\zeta := \mathrm{Prox}_{\frac{\mu}{\tau}\Psi}\left[\frac{2\mu}{\tau} B^* B \mathfrak{L}\xi - \frac{\mu}{\tau} B^* B \mathfrak{L} x + \left(\mathrm{Id} - \frac{\mu}{\tau} B^* B\right) v\right]$$

$$\eta := \left(\mathrm{Id} - \mathrm{Prox}_{\Psi \oplus \iota_C}\right)\left(2\mathfrak{L}_c \xi - \mathfrak{L}_c x + w\right),$$

$\mathrm{Prox}_{\Psi \oplus \iota_C}(w_1, w_2) = (\mathrm{Prox}_{\Psi}(w_1), P_C(w_2))$

where for any $\kappa > 1$, $(\sigma, \tau) \in \mathbb{R}_{++} \times \mathbb{R}_{++}$ is chosen to satisfy

$$\begin{cases} \sigma \mathrm{Id} - \frac{\kappa}{2} A^* A - \mu \mathfrak{L}_c^* \mathfrak{L}_c \succ O_{\mathcal{X}} \\ \tau \geq \left(\frac{\kappa}{2} + \frac{2}{\kappa}\right) \mu \|B\|_{\mathrm{op}}^2. \end{cases}$$

Under the Overall Convexity Condition,

$$\operatorname{argmin} \left[J_{\Psi_B \circ \mathfrak{L}} + \iota_C \circ \mathfrak{C}\right] = \Xi\left(\mathrm{Fix}(T_{\mathrm{cLiGME}})\right),$$

where $\Xi : \mathcal{X} \times \mathcal{Z} \times \mathcal{Z}_c \to \mathcal{X} : (x, v, w) \mapsto x$

and $\mathrm{Fix}(T_{\mathrm{cLiGME}}) := \{(x, v, w) \mid T_{\mathrm{cLiGME}}(x, v, w) = (x, v, w)\}$

$$\mathfrak{P} := \begin{bmatrix} \sigma \mathrm{Id} & -\mu \mathfrak{L}^* B^* B & -\mu \mathfrak{L}_c^* \\ -\mu B^* B \mathfrak{L} & \tau \mathrm{Id} & O_{\mathcal{B}(\mathcal{Z}_c, \mathcal{Z})} \\ -\mu \mathfrak{L}_c & O_{\mathcal{B}(\mathcal{Z}, \mathcal{Z}_c)} & \mu \mathrm{Id} \end{bmatrix} \succ O$$ **and**

$T_{\mathrm{cLiGME}} : \mathcal{H} (:= \mathcal{X} \times \mathcal{Z} \times \mathcal{Z}_c) \to \mathcal{H}$ **is** $\frac{\kappa}{2\kappa - 1}$**-averaged nonexpansive in the Hilbert space** $(\mathcal{H}, \langle \cdot, \cdot \rangle_{\mathfrak{P}}, \| \cdot \|_{\mathfrak{P}})$**, i.e.,**

$(\forall \mathbf{x}, \mathbf{y} \in \mathcal{H})$

$$\|T_{\mathrm{cLiGME}}(\mathbf{x}) - T_{\mathrm{cLiGME}}(\mathbf{y})\|_{\mathfrak{P}}^2 \leq \|\mathbf{x} - \mathbf{y}\|_{\mathfrak{P}}^2 - \frac{\kappa - 1}{\kappa} \|(\mathrm{Id} - T_{\mathrm{cLiGME}})(\mathbf{x}) - (\mathrm{Id} - T_{\mathrm{cLiGME}})(\mathbf{y})\|_{\mathfrak{P}}^2$$

For any initial point $(x_0, v_0, w_0) \in \mathcal{X} \times \mathcal{Z} \times \mathcal{Z}_c$**, the sequence** $(x_n, v_n, w_n)_{n \in \mathbb{N}} \subset \mathcal{X} \times \mathcal{Z} \times \mathcal{Z}_c$ **generated by**

[Krasnosel'skiĭ-Mann]

$$(x_{n+1}, v_{n+1}, w_{n+1}) := T_{\mathrm{cLiGME}}(x_n, v_n, w_n)$$

converges to a point $(x^\star, v^\star, w^\star) \in \mathrm{Fix}(T_{\mathrm{cLiGME}})$ **and**

$$\lim_{n \to \infty} x_n = x^\star \in \mathrm{argmin}\ [J_{\Psi_B \circ \mathfrak{L}} + \iota_{\mathbf{C}} \circ \mathfrak{C}]$$

Algorithm 1 for cLiGME model

Choose $(x_0, v_0, w_0) \in \mathscr{H}$. $\mathscr{X} \times \mathscr{Z} \times \mathscr{Z}_c =: \mathscr{H}$ where $\mathscr{Z}_c := \mathscr{Z} \times \mathfrak{Z}$

Let $(\sigma, \tau, \kappa) \in \mathbb{R}_{++} \times \mathbb{R}_{++} \times (1, \infty)$ satisfying (3.5).

$$\begin{cases} \sigma \mathrm{Id} - \frac{\kappa}{2} A^* A - \mu \mathfrak{L}_c^* \mathfrak{L}_c \succ \mathrm{O}_{\mathscr{X}} \\ \tau \geq \left(\frac{\kappa}{2} + \frac{2}{\kappa}\right) \mu \|B\|_{\mathrm{op}}^2. \end{cases} \tag{3.5}$$

Define $\mathfrak{P}$ as (3.6).

$$\mathfrak{P} := \begin{bmatrix} \sigma \mathrm{Id} & -\mu \mathfrak{L}^* B^* B & -\mu \mathfrak{L}_c^* \\ -\mu B^* B \mathfrak{L} & \tau \mathrm{Id} & \mathrm{O}_{\mathscr{B}(\mathscr{Z}_c, \mathscr{Z})} \\ -\mu \mathfrak{L}_c & \mathrm{O}_{\mathscr{B}(\mathscr{Z}, \mathscr{Z}_c)} & \mu \mathrm{Id} \end{bmatrix} \succ \mathrm{O}_{\mathscr{H}} \tag{3.6}$$

$k \leftarrow 0$.

do

$x_{k+1} \leftarrow \left[\mathrm{Id} - \frac{1}{\sigma}(A^*A - \mu \mathfrak{L}^* B^* B \mathfrak{L})\right] x_k - \frac{\mu}{\sigma} \mathfrak{L}^* B^* B v_k - \frac{\mu}{\sigma} \mathfrak{L}_c^* w_k + \frac{1}{\sigma} A^* y$

$v_{k+1} \leftarrow \mathrm{Prox}_{\frac{\mu}{\tau}\Psi}\left[\frac{2\mu}{\tau} B^* B \mathfrak{L} x_{k+1} - \frac{\mu}{\tau} B^* B \mathfrak{L} x_k + \left(\mathrm{Id} - \frac{\mu}{\tau} B^* B\right) v_k\right]$

$w_{k+1} \leftarrow \mathrm{Prox}_{\Psi^*}(2\mathfrak{L}_c x_{k+1} - \mathfrak{L}_c x_k + w_k)$

$k \leftarrow k+1$

while $\|(x_k, v_k, w_k) - (x_{k-1}, v_{k-1}, w_{k-1})\|_{\mathfrak{P}}$ is not sufficiently small

return x_k

For any (x_0, v_0, w_0), the sequence $(x_k)_{k\in\mathbb{N}}$ generated by Algorithm 1 converges to a globally optimal solution of the cLiGME model.

The derivation of Theorem 1 and Algorithm 1 is inspired by Condat's primal-dual algorithm [Condat 2013] and is essentially based on the so-called forward-backward splitting method.

Since Condat's primal-dual algorithm was proposed for minimization of sum of linearly involved convex terms, it is not directly applicable to the cLiGME model involving nonconvex functions.

Numerical experiments (Piecewise constant image restoration)

$$\underset{(1\leq j\leq N)\ \mathfrak{C}_j x\in C_j}{\text{minimize}} \quad J_{\Psi_B\circ\mathfrak{L}}(x) := \frac{1}{2}\|y-Ax\|_{\mathcal{Y}}^2 + \sum_{i=1}^{M}\mu_i \Psi_{B^{\langle i\rangle}}^{\langle i\rangle}\circ\mathfrak{L}_i(x)$$

We applied the proposed model to

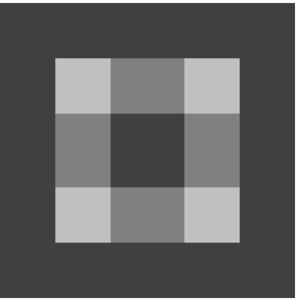
$$y = Ax^\star + \varepsilon \quad \begin{cases} x^\star : \text{vectorization of piecewise constant matrix} \\ A : \text{blur matrix} \\ \varepsilon : \text{AWG to achieve SNR (20 dB)} \end{cases}$$

To promote piecewise constantness, we used two regularizers

Type 1 : anisotropic TV (convex) : $\|\cdot\|_1\circ D$ (D is a 2dim difference operator)

Type 2 : LiGME - nonconvex enhancement of the anisotropic TV

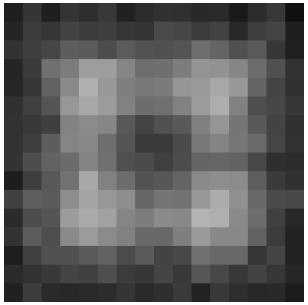
(i) each entry in $x^\star$ belongs to $[0.25, 0.75]$
(ii) every entry in the background is common (but unknown) valued



Original image $x^\star$

we compared Type 1 and Type 2 with constraints

♣ : no constraint, ♢ : constraint (i),
♡ : constraint (ii), ♠ : constraints (i) and (ii)

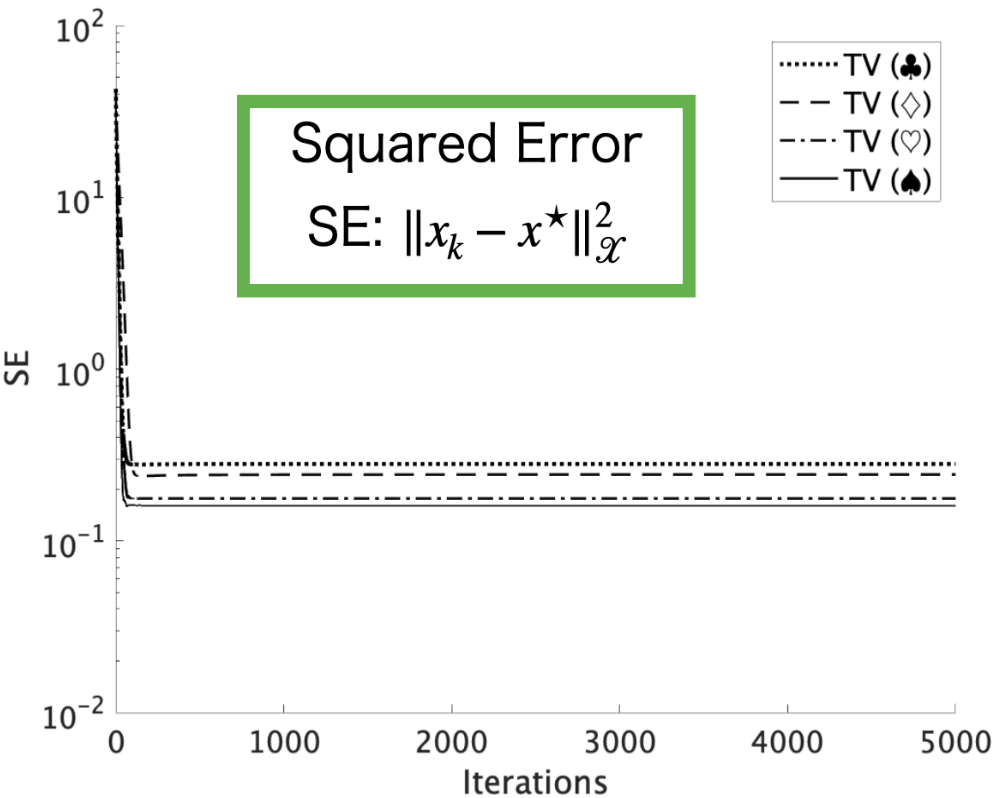


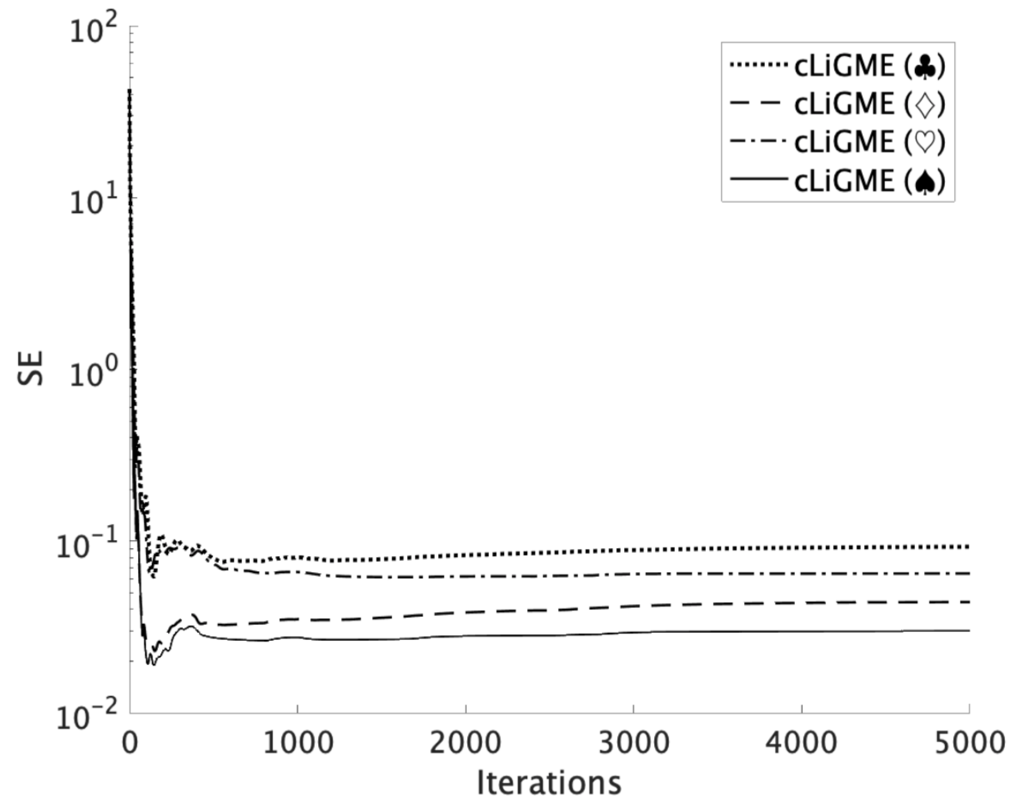
Blurred noisy image y

Numerical experiments (Piecewise constant image restoration)

1. With same constraints, the cLiGME model achieves better estimation than the anisotropic TV model.

2. With multiple constraints, LiGME model is improved by cLiGME model.





Conclusion

1. Mainly for sparsity-rank aware signal processing, the LiGME model presents a mathematically sound nonconvex enhancement of the convexly regularized least squares models.

2. For broader applications, the cLiGME model has been designed by integrating main ideas in set-theoretic estimation and in LiGME model.

Recent related references

[AYY `20] J. Abe, M. Yamagishi and I. Yamada,
"Linearly involved generalized Moreau enhanced models and their proximal splitting algorithm under overall convexity condition," Inverse Problems, (36pp), 2020.

[CYY `21] Y. Chen, M. Yamagishi and I. Yamada,
`A Linearly involved Generalized Moreau Enhancement of l_2,1-norm with application to weighted group sparse classification,'' Algorithms 2021, 14(11), 312.

[YYY `21] W. Yata, M. Yamagishi and I. Yamada,
"A constrained Linearly involved Generalized Moreau Enhanced model and Its proximal splitting algorithm," IEEE MLSP 2021, 2021 (Oct).

[ZY `21] Yi Zhang and I. Yamada,
"DC-LiGME: An efficient algorithm for improved convex sparse regularization," 55th Asilomar Conference on Signals, Systems and Computer, 2021 (Nov).