



HAL
open science

Proximal optimization of OT-regularized cost functions for multichannel signal recovery

Jean Mallejac, Jérôme Idier, Charles Soussen, Xiangyi Li

► **To cite this version:**

Jean Mallejac, Jérôme Idier, Charles Soussen, Xiangyi Li. Proximal optimization of OT-regularized cost functions for multichannel signal recovery. 2024. hal-04488937

HAL Id: hal-04488937

<https://hal.science/hal-04488937v1>

Preprint submitted on 8 Mar 2024

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

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

Proximal optimization of OT-regularized cost functions for multichannel signal recovery

Jean Malléjac¹, Jérôme Idier², Charles Soussen¹, Xiangyi Li³

¹*Université Paris-Saclay, CNRS, CentraleSupélec, L2S, F-91192 Gif-sur-Yvette, France*

²*Nantes Université, Ecole Centrale Nantes, CNRS, LS2N, UMR 6004, F-44000 Nantes, France*

³*Aix Marseille Univ, CNRS, Centrale Marseille, Institut Fresnel, Marseille, France*

March 4, 2024

1 Introduction

1.1 Problem Statement

We consider a multichannel signal viewed as a sequence of monovariate signals $(\mathbf{y}_n)_{n \in \llbracket 1, N \rrbracket}$ with N the number of channels. Each signal in the sequence $\mathbf{y}_n \in \mathbb{R}^K$ can be represented in dictionary $A \in \mathbb{R}^{K \times M}$. The inverse problem aims to retrieve $(\mathbf{x}_n)_{n \in \llbracket 1, N \rrbracket}$ from given signal $(\mathbf{y}_n)_{n \in \llbracket 1, N \rrbracket}$ and dictionary A .

We propose two variational formulations which make use of Optimal Transport (OT) regularization. These are Quadratic Programs that we solve using the Condat-Vu's Primal-Dual algorithm.

1.2 Definition & notations

Notation:

- $\mathbf{1}_M$ is the vector of ones of size $M \times 1$
- $\langle \cdot, \cdot \rangle_F$ is the Frobenius inner product
- $A \otimes B$ is the Kronecker product between two matrices of arbitrary sizes
- $\mathbb{A} \bar{\times}_n \mathbf{b}$ is the n -mode vector product of tensor $\mathbb{A} \in \mathbb{R}^{I_1 \times \dots \times I_n \times \dots \times I_N}$ (of order N) with vector $\mathbf{b} \in \mathbb{R}^{I_n}$ [2].

$$(\mathbb{A} \bar{\times}_n \mathbf{b})_{i_1 \dots i_{n-1} i_{n+1} \dots i_N} = \sum_{i_n=1}^{I_n} a_{i_1 i_2 \dots i_K} \mathbf{b}_{i_n}.$$

Each mode- n fiber of tensor \mathbb{A} is multiplied by \mathbf{b} (inner product). The order of the resulting tensor $\mathbb{A} \bar{\times}_n \mathbf{b}$ is $N - 1$.

Definition 1.1 (Spectral Norm). *Let A be a matrix. The spectral norm $\|\cdot\|$ of A is defined as:*

$$\begin{aligned} \|A\| &= \sup \{ \|Ax\|_2 : \|x\|_2 \leq 1 \} \\ \|A\| &= \sqrt{\lambda_{\max}(A^\top A)} \end{aligned}$$

with $\lambda_{\max}(A^\top A)$ the largest eigenvalue of $A^\top A$ (always positive because $A^\top A$ is a PSD matrix).

Basic Properties:

- Given A, B matrices : $\|A \otimes B\| = \|A\| \|B\|$.
- Let A be a matrix, Q and P permutation matrices: $\|QAP\| = \|A\|$
- Let A, B, C, D be matrices with consistent size $(A \otimes B)(C \otimes D) = (AC \otimes BD)$

See reviews [1] & [2] for more details.

2 Condat-Vu's Algorithm

Condat-Vu's algorithm [3] is a Primal-Dual algorithm to solve the following type of problem:

$$\min_{\mathbf{z}} h(\mathbf{z}) + f(\mathbf{z}) + \sum_{i=0}^{I-1} g_i(L_i \mathbf{z}), \quad (1)$$

with the following terms in each part of the splitting:

- h is convex and differentiable
- f and $(g_i)_{i \in \llbracket 0, I-1 \rrbracket}$ are convex
- $(L_i)_{i \in \llbracket 0, I-1 \rrbracket}$ are linear operators

Algorithm 1 Condat-Vu's algorithm

Initialize: $z^{(0)} = \mathbf{0}, \quad \forall i, u_i^{(0)} = \mathbf{0}$
for $k = 1, 2, \dots$ **do**

$$z^{(k+1)} = \text{prox}_{\tau f}(z^{(k)} - \tau \nabla h(z^{(k)}) - \tau \sum_{i=1}^I L_i^* u_i^{(k)})$$

for $i = 0, 1, \dots, I - 1$ **do**

$$u_i^{(k+1)} = \text{prox}_{\sigma_i g_i^*} \left(u_i^{(k)} + \sigma_i L_i (2z^{(k+1)} - z^{(k)}) \right)$$

end for
end for

Nesterov acceleration technique can be added to the algorithm.

Proposition 1. *Gradient step of algorithm should verify the condition expressed in [3, Th. 9.7] to ensure convergence of iterates:*

$$\forall i \in \llbracket 1, I \rrbracket, \quad \sigma_i \geq 0, \quad \tau \geq 0, \quad (2)$$

$$\tau \leq \frac{1}{\left\| Q + \sum_{i=0}^{I-1} \sigma_i L_i^\top L_i \right\|} \quad (3)$$

Using the triangular inequality, we propose a slightly more restrictive condition on gradient steps, easier to compute:

$$\tau \leq \frac{1}{\|Q\| + \sum_{i=0}^{I-1} \sigma_i \|L_i^\top L_i\|} \quad (4)$$

Clearly, (4) implies (3)

3 Problem adaptation to Condat-Vu's framework: First order

The first OT regularized problem results in the following Quadratic Program:

$$\min_{\substack{\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{R}_+^M \\ P_1, \dots, P_{N-1} \in \mathbb{R}_+^{M \times M}}} \sum_{n=1}^N \|\mathbf{y}_n - A\mathbf{x}_n\|_2^2 + \omega \sum_{n=1}^{N-1} \langle P_n, C \rangle_F, \quad (5a)$$

$$\mathbf{s.t.} \quad \forall n \in \llbracket 1, N \rrbracket, \quad \mathbf{1}_M^\top \mathbf{x}_n = 1, \quad (5b)$$

$$\forall n \in \llbracket 1, N-1 \rrbracket, \quad P_n \bar{\times}_2 \mathbf{1}_M = \mathbf{x}_n, \quad P_n \bar{\times}_1 \mathbf{1}_M = \mathbf{x}_{n+1}. \quad (5c)$$

Let us introduce global column vector \mathbf{z} , which is an unfolded version gathering all variables together:

$$\begin{aligned} \mathbf{x} &= \text{Vect}(\mathbf{x}_1, \dots, \mathbf{x}_N) \\ \mathbf{p} &= \text{Vect}(P_1, \dots, P_{N-1}) \\ \mathbf{z} &= \text{Vect}(\mathbf{x}, \mathbf{p}) \end{aligned}$$

Proposition 2. *Criteria of problem (5) can be reformulated with respect to global variable \mathbf{z} as $h(\mathbf{z})$:*

$$h(\mathbf{z}) = \frac{1}{2} \mathbf{z}^\top Q \mathbf{z} + \mathbf{q}^\top \mathbf{z} + \text{cst}, \quad (6)$$

with h quadratic and:

$$\begin{aligned} Q &= \begin{pmatrix} I_N \otimes 2A^\top A & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}, \\ \mathbf{q} &= \text{Vect}(-2A^\top \mathbf{y}_1, \dots, -2A^\top \mathbf{y}_N, \omega \mathbf{c}, \dots, \omega \mathbf{c}) \\ \text{cst} &= \sum_{n=1}^N \|\mathbf{y}_n\|_2 \end{aligned}$$

Proposition 3. *Problem (5) can be reformulated into:*

$$\min_z h(\mathbf{z}), \quad (7a)$$

$$\mathbf{s.t.} \quad \mathbf{z} \geq \mathbf{0}, \quad L_1 \mathbf{z} = \mathbf{1}_N, \quad L_0 \mathbf{z} = \mathbf{0}, \quad (7b)$$

with L_1 and L_0 linear operators. Then, it can be easily expressed in the Condat-Vu's framework:

$$\min_z h(\mathbf{z}) + f(\mathbf{z}) + g_0(L_0 \mathbf{z}) + g_1(L_1 \mathbf{z}), \quad (8)$$

with:

- $I = 2$,
- $g_0 = \iota_{\{\mathbf{0}\}}$, $g_1 = \iota_{\{\mathbf{1}_N\}}$ indicator functions on vectors,
- $f = \iota_{\mathbb{R}_+^z}$ indicator function on the positive orthant.

3.1 Expression of linear operators L_1 and L_0

Constraints (5b) can be expressed with respect to \mathbf{z} introducing operator L_1 :

$$L_1 \mathbf{z} = \begin{bmatrix} I_N \otimes \mathbf{1}_M^\top & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{p} \end{bmatrix} = \mathbf{1}_N,$$

Constraints (5c) are reformulated:

$$\begin{aligned} & \forall n \in \llbracket 1, N-1 \rrbracket, \quad P_n \bar{\times}_2 \mathbf{1}_M = \mathbf{x}_n, \quad P_n \bar{\times}_1 \mathbf{1}_M = \mathbf{x}_{n+1}, \\ \Leftrightarrow & \forall n \in \llbracket 1, N-1 \rrbracket, \quad S_2 \mathbf{p}_n = \mathbf{x}_n, \quad S_1 \mathbf{p}_n = \mathbf{x}_{n+1}, \\ \Leftrightarrow & \quad \quad \quad S_2 \mathbf{p} = M_2 \mathbf{x}, \quad S_1 \mathbf{p} = M_1 \mathbf{x}, \end{aligned}$$

with M_k selecting operators on \mathbf{x} :

$$M_1 = [\mathbf{0}, I_{(N-1)M}], \quad M_2 = [I_{(N-1)M}, \mathbf{0}],$$

with S_1 sum operator on the first dimension (rows) and S_2 sum operator on the second dimension (columns).

$$S_2 \mathbf{p} = \text{Vect}(S_2 \mathbf{p}_1, \dots, S_2 \mathbf{p}_{N-1}) \quad S_1 \mathbf{p} = \text{Vect}(S_1 \mathbf{p}_1, \dots, S_1 \mathbf{p}_{N-1})$$

Constraints (5c) are expressed with respect to \mathbf{z} introducing operator L_0 :

$$L_0 \mathbf{z} = \begin{bmatrix} -M_2 & S_2 \\ -M_1 & S_1 \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{p} \end{bmatrix} = \mathbf{0}_{2M(N-1)},$$

S_2 and M_2 are columns permutation of matrix S_1 and M_1 respectively. We have:

$$S_1 = I_{(N-1)M} \otimes \mathbf{1}_M^\top$$

Introducing $H = \begin{bmatrix} -M_1 & S_1 \end{bmatrix}$, there exist R a permutation matrix such that:

$$L_0 = \begin{bmatrix} HR \\ H \end{bmatrix}$$

3.2 Condition on gradient step

The condition (4) is adapted for the above problem (8):

$$\tau \leq \frac{1}{\|Q\| + \sigma_0 \|L_0^\top L_0\| + \sigma_1 \|L_1^\top L_1\|}. \quad (9)$$

3.2.1 Computation of $\|Q\|$

$$\|Q\| = \left\| \begin{pmatrix} I_N \otimes 2A^\top A & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \right\| = \|I_N \otimes 2A^\top A\| = \|I_N\| \|2A^\top A\| = 2 \|A^\top A\|.$$

3.2.2 Computation of $\|L_1^\top L_1\|$

Lemma 1. Let $\boldsymbol{\rho}_M \in \{-1, 1\}^M$ be a vector of size M whose components are either -1 or 1 . then:

$$\|\boldsymbol{\rho}_M \boldsymbol{\rho}_M^\top\| = M$$

Proof. Notice that:

$$\forall \mathbf{x}, \quad \|\boldsymbol{\rho}_M \boldsymbol{\rho}_M^\top \mathbf{x}\|_2 = \|\boldsymbol{\rho}_M \langle \boldsymbol{\rho}_M, \mathbf{x} \rangle\|_2 = |\langle \boldsymbol{\rho}_M, \mathbf{x} \rangle| \|\boldsymbol{\rho}_M\|_2$$

Using the Cauchy-Schwarz inequality, we get:

$$\|\boldsymbol{\rho}_M \boldsymbol{\rho}_M^\top \mathbf{x}\|_2 \leq \|\boldsymbol{\rho}_M\|_2^2 \|\mathbf{x}\|_2$$

As $\|\boldsymbol{\rho}_M\|_2^2 = M$, we have:

$$\|\boldsymbol{\rho}_M \boldsymbol{\rho}_M^\top \mathbf{x}\|_2 \leq M \|\mathbf{x}\|_2 \quad (10)$$

Clearly, $\boldsymbol{\rho}_M \boldsymbol{\rho}_M^\top$ is a PSD matrix. Thus, all eigenvalues of $\boldsymbol{\rho}_M \boldsymbol{\rho}_M^\top$ are positive. According to (10), the eigenvalues are bound by M . Also, note that:

$$\boldsymbol{\rho}_M \boldsymbol{\rho}_M^\top \boldsymbol{\rho}_M = \boldsymbol{\rho}_M \langle \boldsymbol{\rho}_M, \boldsymbol{\rho}_M \rangle = \boldsymbol{\rho}_M \|\boldsymbol{\rho}_M\|_2^2 = M \boldsymbol{\rho}_M$$

Thus, M is the largest eigenvalue of $\boldsymbol{\rho}_M \boldsymbol{\rho}_M^\top$. ■

$$\begin{aligned} \|L_1^\top L_1\| &= \left\| \begin{bmatrix} (I_N \otimes \mathbf{1}_M^\top)^\top I_N \otimes \mathbf{1}_M^\top & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \right\| \\ &= \|(I_N \otimes \mathbf{1}_M^\top)^\top (I_N \otimes \mathbf{1}_M^\top)\| \end{aligned}$$

Using Lemma 1:

$$\|(I_N \otimes \mathbf{1}_M^\top)^\top I_N \otimes \mathbf{1}_M^\top\| = \|I_N\| \|\mathbf{1}_M \mathbf{1}_M^\top\| = \|\mathbf{1}_M \mathbf{1}_M^\top\| = M,$$

consequently:

$$\|L_1^\top L_1\| = M. \quad (11)$$

3.2.3 Computation of $\|L_0^\top L_0\|$

Let $E \in \mathbb{N}$ be an arbitrary dimension. We introduce matrix: $H_{E,M} = \begin{bmatrix} I_E & I_E \otimes -\mathbf{1}_M^\top \\ I_E \otimes -\mathbf{1}_M & I_E \otimes \mathbf{1}_M \mathbf{1}_M^\top \end{bmatrix}$.

Lemma 2. *The spectral norm of matrix $H_{E,M}$ is equal to $M + 1$.*

Proof. There exist permutation matrices R_1 and R_2 on rows and columns respectively such that:

$$\begin{aligned} H_{E,M} &= \begin{bmatrix} I_E & I_E \otimes -\mathbf{1}_M^\top \\ I_E \otimes -\mathbf{1}_M & I_E \otimes \mathbf{1}_M \mathbf{1}_M^\top \end{bmatrix} \\ H_{E,M} &= \begin{bmatrix} I_E & I_E \otimes -\mathbf{1}_M^\top \\ I_E \otimes -\mathbf{1}_M & I_E \otimes \mathbf{1}_M \otimes \mathbf{1}_M^\top \end{bmatrix} \\ H_{E,M} &= \begin{bmatrix} I_E & I_E \otimes -\mathbf{1}_M^\top \\ R_1(-\mathbf{1}_M \otimes I_E) & R_1(\mathbf{1}_M \otimes I_E \otimes \mathbf{1}_M^\top) \end{bmatrix} \\ H_{E,M} &= \begin{bmatrix} I_E & \mathbf{0} \\ \mathbf{0} & R_1 \end{bmatrix} \begin{bmatrix} I_E & I_E \otimes -\mathbf{1}_M^\top \\ -\mathbf{1}_M \otimes I_E & \mathbf{1}_M \otimes I_E \otimes \mathbf{1}_M^\top \end{bmatrix} \\ H_{E,M} &= \begin{bmatrix} I_E & \mathbf{0} \\ \mathbf{0} & R_1 \end{bmatrix} \begin{bmatrix} I_E & (-\mathbf{1}_M^\top \otimes I_E) R_2 \\ -\mathbf{1}_M \otimes I_E & (\mathbf{1}_M \otimes \mathbf{1}_M^\top \otimes I_E) R_2 \end{bmatrix} \\ H_{E,M} &= \begin{bmatrix} I_E & \mathbf{0} \\ \mathbf{0} & R_1 \end{bmatrix} \begin{bmatrix} I_E & -\mathbf{1}_M^\top \otimes I_E \\ -\mathbf{1}_M \otimes I_E & \mathbf{1}_M \mathbf{1}_M^\top \otimes I_E \end{bmatrix} \begin{bmatrix} I_E & \mathbf{0} \\ \mathbf{0} & R_2 \end{bmatrix} \end{aligned}$$

As spectral norm is invariant by permutation:

$$\begin{aligned} \|H_{E,M}\| &= \left\| \begin{bmatrix} I_E & \mathbf{0} \\ \mathbf{0} & R_1 \end{bmatrix} \begin{bmatrix} I_E & -\mathbf{1}_M^\top \otimes I_E \\ -\mathbf{1}_M \otimes I_E & \mathbf{1}_M \mathbf{1}_M^\top \otimes I_E \end{bmatrix} \begin{bmatrix} I_E & \mathbf{0} \\ \mathbf{0} & R_2 \end{bmatrix} \right\| \\ \|H_{E,M}\| &= \left\| \begin{bmatrix} I_E & -\mathbf{1}_M^\top \otimes I_E \\ -\mathbf{1}_M \otimes I_E & \mathbf{1}_M \mathbf{1}_M^\top \otimes I_E \end{bmatrix} \right\| \\ &= \left\| \begin{bmatrix} 1 & -\mathbf{1}_M^\top \\ -\mathbf{1}_M & \mathbf{1}_M \mathbf{1}_M^\top \end{bmatrix} \otimes I_E \right\| \\ &= \left\| \begin{bmatrix} 1 & -\mathbf{1}_M^\top \\ -\mathbf{1}_M & \mathbf{1}_M \mathbf{1}_M^\top \end{bmatrix} \right\| \|I_E\| \end{aligned}$$

Using Lemma 1:

$$\|H_{E,M}\| = \left\| \begin{bmatrix} 1 & -\mathbf{1}_M^\top \\ -\mathbf{1}_M & \mathbf{1}_M \mathbf{1}_M^\top \end{bmatrix} \right\| = \|\boldsymbol{\rho}_{M+1} \boldsymbol{\rho}_{M+1}^\top\| = M + 1$$

■

$$\begin{aligned} \|L_0^\top L_0\| &= \left\| \begin{bmatrix} H^\top & (HR)^\top \end{bmatrix} \begin{bmatrix} H \\ HR \end{bmatrix} \right\| \\ \|L_0^\top L_0\| &= \|H^\top H + R^\top H^\top HR\| \\ \|L_0^\top L_0\| &\leq \|H^\top H\| + \|R^\top H^\top HR\| \end{aligned}$$

The spectral norm is invariant by permutation. Thus, we have:

$$\|L_0^\top L_0\| \leq 2 \|H^\top H\|$$

$$\begin{aligned} H^\top H &= \begin{bmatrix} M_1^\top M_1 & -M_1^\top S_1 \\ -S_1^\top M_1 & S_1^\top S_1 \end{bmatrix} \\ \|H^\top H\| &= \left\| \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & I_{(N-1)M} & -I_{(N-1)M} \otimes \mathbf{1}_M^\top \\ \mathbf{0} & -I_{(N-1)M} \otimes \mathbf{1}_M & I_{(N-1)M} \otimes \mathbf{1}_M \mathbf{1}_M^\top \end{bmatrix} \right\| \\ \|H^\top H\| &= \left\| \begin{bmatrix} I_{(N-1)M} & I_{(N-1)M} \otimes -\mathbf{1}_M^\top \\ I_{(N-1)M} \otimes -\mathbf{1}_M & I_{(N-1)M} \otimes \mathbf{1}_M \mathbf{1}_M^\top \end{bmatrix} \right\| \end{aligned}$$

We have:

$$\|H^\top H\| = \|H_{E,M}\| \quad \text{with } E = (N-1)M$$

Using Lemma 2:

$$\|H^\top H\| = \|H_{E,M}\| = M + 1$$

and:

$$\|L_0^\top L_0\| \leq 2(M + 1) \quad (12)$$

3.2.4 Condition on gradient step for problem (5)

Condition below is a weaker sufficient condition than (9)

$$\tau \leq \frac{1}{2 \|A^\top A\| + \sigma_1 M + 2\sigma_0(M + 1)} \quad (13)$$

4 Second order problem

$$\min_{\substack{\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{R}_+^M \\ P_1, \dots, P_{N-1} \in \mathbb{R}_+^{M \times M} \\ \mathbb{P}_1, \dots, \mathbb{P}_{N-2} \in \mathbb{R}_+^{M \times M \times M}}} \sum_{n=1}^N \|\mathbf{y}_n - A \mathbf{x}_n\|_2^2 + \omega \sum_{n=1}^{N-2} \langle \mathbb{P}_n, \mathbf{C} \rangle_F, \quad (14a)$$

$$\text{s.t. } \forall n \in \llbracket 1, N \rrbracket, \quad \mathbf{1}_M^\top \mathbf{x}_n = 1, \quad (14b)$$

$$\forall n \in \llbracket 1, N-1 \rrbracket, \quad P_n \bar{\times}_2 \mathbf{1}_M = \mathbf{x}_n, \quad P_n \bar{\times}_1 \mathbf{1}_M = \mathbf{x}_{n+1}. \quad (14c)$$

$$\forall n \in \llbracket 1, N-2 \rrbracket, \quad \mathbb{P}_n \bar{\times}_3 \mathbf{1}_M = P_n, \quad \mathbb{P}_n \bar{\times}_1 \mathbf{1}_M = P_{n+1}. \quad (14d)$$

5 Problem adaptation to Condat-Vu's framework: Second order

Let us introduce global variable \mathbf{z} :

$$\begin{aligned}\mathbf{x} &= \text{Vect}(\mathbf{x}_1, \dots, \mathbf{x}_N) \\ \mathbf{p} &= \text{Vect}(P_1, \dots, P_{N-1}) \\ \mathbb{p} &= \text{Vect}(\mathbb{P}_1, \dots, \mathbb{P}_{N-2}) \\ \mathbf{z} &= \text{Vect}(\mathbf{x}, \mathbf{p}, \mathbb{p})\end{aligned}$$

Proposition 4. *Criteria h can be reformulated with respect to global variable \mathbf{z} :*

$$h(\mathbf{z}) = \frac{1}{2} \mathbf{z}^\top Q \mathbf{z} + \mathbf{q}^\top \mathbf{z} + cst, \quad (15)$$

with quadratic, linear and constant term respectively:

$$\begin{aligned}Q &= \begin{pmatrix} I_N \otimes 2A^\top A & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix}, \\ \mathbf{q} &= (-2\mathbf{y}_1^\top A, \dots, -2\mathbf{y}_N^\top A, 0 \dots 0, \omega \text{Vect}(\mathbb{C})^\top, \dots, \omega \text{Vect}(\mathbb{C})^\top)^\top \\ cst &= \sum_{n=1}^N \mathbf{y}_n \mathbf{y}_n^\top\end{aligned}$$

Proposition 5. *Problem (14) can be reformulated into:*

$$\min_z h(z), \quad (16a)$$

$$\mathbf{s.t.} \quad z \geq 0 \quad L'_1 z = \mathbf{1}_N, \quad L_{01} z = \mathbf{0}, \quad L_{02} z = \mathbf{0}, \quad (16b)$$

with L'_1, L_{01} and L_{02} linear operators. Then, it can be easily expressed in the Condat-Vu's framework:

$$\min_z h(z) + f(z) + g_1(L'_1 z) + g_0(L_{01} z) + g'_0(L_{02} z), \quad (17)$$

- $I = 3$,
- $g_1 = \iota_{\mathbf{1}_N}, g_0 = \iota_{\mathbf{0}}, g'_0 = \iota_{\mathbf{0}}$ indicator functions on vectors,
- $f = \iota_{\mathbb{R}_+}$ indicator function on the positive orthant.

5.0.1 Expression of linear operator L'_1, L_{01} and L_{02}

Constraints (5b) can be expressed with respect to \mathbf{z} introducing operator L_1 :

$$L'_1 \mathbf{z} = \begin{bmatrix} I_N \otimes \mathbf{1}_M^\top & \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{p} \\ \mathbb{p} \end{bmatrix} = \mathbf{1}_N,$$

with: $I_N \otimes \mathbf{1}_M^\top = I_N \otimes \mathbf{1}_M^\top$.

Constraints (5c) are equivalent to constraints (14c). Thus L_{01} can be expressed with operator L_0 :

$$L_{01} \mathbf{z} = \begin{bmatrix} L_0 & | & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{p} \\ \mathbb{p} \end{bmatrix}$$

Constraints (14d) are reformulated:

$$\begin{aligned} & \forall n \in \llbracket 1, N-2 \rrbracket, \quad \mathbb{P}_n \bar{\times}_3 \mathbf{1}_M = P_n, \quad \mathbb{P}_n \bar{\times}_1 \mathbf{1}_M = P_{n+1}, \\ \Leftrightarrow & \forall n \in \llbracket 1, N-2 \rrbracket, \quad \Sigma_2 \mathbb{P}_n = \mathbf{p}_n, \quad \Sigma_1 \mathbb{P}_n = \mathbf{p}_{n+1}, \\ & \Leftrightarrow \quad \tilde{\Sigma}_3 \mathbb{P} = K_3 \mathbf{p}, \quad \tilde{\Sigma}_3 \mathbb{P} = K_1 \mathbf{p}. \end{aligned}$$

with:

$$K_1 = [\mathbf{0}, I_{(N-2)M^2}], \quad K_2 = [I_{(N-2)M^2}, \mathbf{0}],$$

and:

$$\tilde{\Sigma}_1 = I_{(N-2)} \otimes \Sigma_1 = I_{(N-2)M^2} \otimes \mathbf{1}_M^\top, \quad \tilde{\Sigma}_3 = I_{(N-2)} \otimes \Sigma_3 = I_{(N-2)} \otimes \mathbf{1}_M^\top \otimes I_{M^2},$$

Constraints (5c) are expressed with respect to \mathbf{z} introducing operator L_0 :

$$L_{02} \mathbf{z} = \begin{bmatrix} \mathbf{0} & -K_3 & \tilde{\Sigma}_3 \\ \mathbf{0} & -K_1 & \tilde{\Sigma}_1 \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{p} \\ \mathbb{P} \end{bmatrix} = \mathbf{0}_{2M^2(N-2)},$$

$\tilde{\Sigma}_2$ and K_2 are columns permutation of matrix $\tilde{\Sigma}_1$ and K_1 respectively. Introducing $H' = \begin{bmatrix} \mathbf{0} & -K_1 & \tilde{\Sigma}_1 \end{bmatrix}$, there exist R' a permutation matrix such that:

$$L_{02} = \begin{bmatrix} H' R' \\ H' \end{bmatrix}$$

5.1 Condition on gradient step

The condition (4) is adapted for problem (17):

$$\tau \leq \frac{1}{\|Q\| + \sigma_1 \|L_1'^\top L_1'\| + \sigma_{01} \|L_{01}^\top L_{01}\| + \sigma_{02} \|L_{02}^\top L_{02}\|} \quad (18)$$

5.1.1 Computation of $\|Q\|$

$$\|Q\| = \left\| \begin{pmatrix} I_N \otimes 2A^\top A & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} \right\| = 2 \|A^\top A\|,$$

5.1.2 Computation of $\|L_1'^\top L_1'\|$

$$\|L_1'^\top L_1'\| = \left\| \begin{bmatrix} (I_N \otimes \mathbf{1}_M^\top)^\top I_N \otimes \mathbf{1}_M^\top & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \right\| = \|(I_N \otimes \mathbf{1}_M^\top)^\top I_N \otimes \mathbf{1}_M^\top\| = M,$$

5.1.3 Computation of $\|L_{01}^\top L_{01}\|$

$$\|L_{01}^\top L_{01}\| = \left\| \begin{bmatrix} L_0^\top L_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \right\| = \|L_0^\top L_0\| \leq 2(M+1),$$

5.1.4 Computation of $\|L_{02}^\top L_{02}\|$

Thus, we have:

$$\|L_{02}^\top L_{02}\| \leq 2 \|H'^\top H'\|$$

$$H'^\top H' = \begin{bmatrix} K_1^\top K_1 & -K_1^\top \tilde{\Sigma}_1 \\ -\tilde{\Sigma}_1^\top M_1 & \tilde{\Sigma}_1^\top \tilde{\Sigma}_1 \end{bmatrix}$$

$$\|H'^\top H'\| = \left\| \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & I_{(N-2)M^2} & -I_{(N-2)M^2} \otimes \mathbf{1}_M^\top \\ \mathbf{0} & -I_{(N-2)M^2} \otimes \mathbf{1}_M & I_{(N-2)M^2} \otimes \mathbf{1}_M \mathbf{1}_M^\top \end{bmatrix} \right\|$$

$$\|H'^\top H'\| = \|H_{E,M}\| \quad \text{with } E = (N-2)M^2$$

Recall for matrix $H_{E,M}$: $H_{E,M} = \begin{bmatrix} I_E & I_E \otimes -\mathbf{1}_M^\top \\ I_E \otimes -\mathbf{1}_M & I_E \otimes \mathbf{1}_M \mathbf{1}_M^\top \end{bmatrix}$

Using Lemma 2, we have:

$$\|H'^\top H'\| = \|H_{E,M}\| = M + 1$$

and:

$$\|L_{02}^\top L_{02}\| \leq 2(M + 1)$$

5.1.5 Condition on gradient steps for problem (14)

Condition below is a weaker sufficient condition than (18)

$$\tau \leq \frac{1}{2\|A^\top A\| + \sigma_1 M + (\sigma_{01} + \sigma_{02})2(M + 1)} \quad (19)$$

References

- [1] N. D. Sidiropoulos, L. De Lathauwer, X. Fu, K. Huang, E. E. Papalexakis, and C. Faloutsos, “Tensor decomposition for signal processing and machine learning”, *IEEE Transactions on Signal Processing*, vol. 65, no. 13, pp. 3551–3582, 2017.
- [2] T. G. Kolda and B. W. Bader, “Tensor decompositions and applications”, *SIAM Review*, vol. 51, no. 3, pp. 455–500, 2009.
- [3] L. Condat, D. Kitahara, A. Contreras, and A. Hirabayashi, “Proximal splitting algorithms for convex optimization: A tour of recent advances, with new twists”, *SIAM Review*, vol. 65, no. 2, pp. 375–435, 2023.