Relaxing Tight Frame Condition in Parallel Proximal Methods for Signal Restoration *

Nelly Pustelnik† Jean-Christophe Pesquet, and Caroline Chaux. ‡

May 27, 2011

Abstract

A fruitful approach for solving signal deconvolution problems consists of resorting to a frame-based convex variational formulation. In this context, parallel proximal algorithms and related alternating direction methods of multipliers have become popular optimization techniques to approximate iteratively the desired solution. Until now, in most of these methods, either Lipschitz differentiability properties or tight frame representations were assumed. In this paper, it is shown that it is possible to relax these assumptions by considering a class of non necessarily tight frame representations, thus offering the possibility of addressing a broader class of signal restoration problems. In particular, it is possible to use non necessarily maximally decimated filter banks with perfect reconstruction, which are common tools in digital signal processing. The proposed approach allows us to solve both frame analysis and frame synthesis problems for various noise distributions. In our simulations, it is applied to the deconvolution of data corrupted with Poisson noise or Laplacian noise by using (non-tight) discrete dual-tree wavelet representations and filter bank structures.

1 Introduction

Many works in signal/image processing are concerned with data restoration problems. For such problems, the original data $y \in \ell^2(\mathbb{Z})$ is degraded by a stable convolutive operator $T: \ell^2(\mathbb{Z}) \to \ell^2(\mathbb{Z})$ and by a non-necessarily additive noise. The resulting observation model can be written as $z = D_{\alpha}(T\gamma)$ where $D_{\alpha}$ denotes the noise effect and $\alpha > 0$ is some related parameter (for

---

*Part of this work appeared in the conference proceedings of EUSIPCO 2010 [1]
†N. Pustelnik (Corresponding Author) is with the Université Bordeaux 1, Laboratoire IMS, CNRS-UMR 5218, 351 Cours de la Libération, 33405 Talence Cedex, France. Phone: +33 5 40 00 31 33, E-mail: nelly.pustelnik@ims-bordeaux.fr.
‡J.-C. Pesquet and C. Chaux are with the Université Paris-Est, LIGM, CNRS-UMR 8049, 77454 Marne-la-Vallée Cedex 2, France. Phone: +33 1 60 95 77 39, E-mail: {jean-christophe.pesquet,caroline.chaux}@univ-paris-est.fr.

1$\ell^2(\mathbb{Z})$ denotes the space of discrete-time real-valued signals defined on $\mathbb{Z}$ having a finite energy.
example, \( \alpha \) may represent the variance for Gaussian noise or the scaling parameter for Poisson noise. In this context, our objective is to recover a signal \( y \in \ell^2(Z) \), the closest possible to \( \tilde{y} \), from the observation vector \( z \) assumed to belong to \( \ell^2(Z) \) and available prior information (sparsity, positivity, . . . ). In early works, this problem was solved, mainly for Gaussian noise, by using Wiener filtering, or equivalently quadratic regularization techniques. Later, multiresolution analyses were used for denoising by applying a thresholding to the generated coefficients [2]. Then, in order to improve the denoising performance, redundant frame representations were substituted for wavelet bases [3, 4]. In [3, 6, 7, 8], authors considered convex optimization techniques to jointly address the effects of a noise and of a linear degradation within a convex variational framework. When the noise is Gaussian, the forward-backward (FB) algorithm [5] (also known as thresholded Landweber algorithm when the regularization term is an \( \ell_1 \)-norm [6, 7, 8]) and its extensions [9, 10] can be employed in the context of wavelet basis decompositions and its use can be extended to arbitrary frame representations [11]. However, in the context of a non-additive noise such as a Poisson noise or a Laplace noise, FB algorithm is no longer applicable due to the non-Lipschitz differentiability of the data fidelity term. Other convex optimization algorithms must be employed such as the Douglas-Rachford (DR) algorithm [12], the Parallel ProXimal Algorithm (PPXA) [13] or the Alternating Direction Method of Multipliers (ADMM) [14, 15]. These algorithms belong to the class of proximal algorithms and, for tractability issues, they often require to use tight frame representation for which closed forms of the involved proximity operators can be derived [13, 14, 15, 16]. The goal of this paper is to propose a way to relax the tight frame requirement by considering an appropriate class of frame representations.

In the following, we consider two general convex minimization problems, which are useful to solve frame-based restoration problems formulated under a Synthesis Form (SF) or an Analysis Form (AF). The SF can be expressed as:

\[
\min_{y = F^*x, x \in \ell^2(Z)} \sum_{r=1}^{R} f_r(L_r y) + \sum_{s=1}^{S} g_s(x) \tag{1}
\]

and the AF is:

\[
\min_{y \in \ell^2(Z)} \sum_{r=1}^{R} f_r(L_r y) + \sum_{s=1}^{S} g_s(F y). \tag{2}
\]

\( F: \ell^2(Z) \to \ell^2(Z) \) (resp. \( F^*: \ell^2(Z) \to \ell^2(Z) \)) denotes the frame analysis (resp. synthesis) operator. For every \( r \in \{1, \ldots, R\} \), \( f_r: \ell^2(Z) \to [\infty, +\infty] \) is a convex, lower semicontinuous and proper function, \( L_r: \ell^2(Z) \to \ell^2(Z) \) is a stable convolutive operator, and for every \( s \in \{1, \ldots, S\} \), \( g_s: \ell^2(Z) \to [\infty, +\infty] \) is a convex, lower semicontinuous and proper function. In several works, SF has been preferred since AF appears to be more difficult to solve numerically [17, 18, 19, 20]. In the proposed framework, both approaches have a similar complexity.

This paper is organized as follows: in Section 2 the class of frames considered in this work is defined and their connections with filter bank structures is emphasized. In Section 3 we show how these (non necessarily tight) frames can be combined with parallel proximal algorithms in order to solve Problems (1) and (2). The proposed approach is also applicable to related augmented Lagrangian approaches. Finally, restoration results are provided in Section 4 for scenarios
involving Poisson noise or Laplace noise by using Dual-Tree Transforms (DTT) and filter bank representations.

Notation: Throughout this paper, \( \Gamma_0(\mathcal{H}) \) designates the class of lower semicontinuous convex functions \( \varphi \) defined on a real Hilbert space \( \mathcal{H} \) and taking their values in \( ]-\infty, +\infty[ \), which are proper in the sense that their domain \( \text{dom} \varphi = \{ u \in \mathcal{H} \mid \varphi(u) < +\infty \} \) is nonempty. If \( \varphi \in \Gamma_0(\mathcal{H}) \) has a unique minimizer, it is denoted by \( \text{argmin}_{u \in \mathcal{H}} \varphi(u) \). The interior of a set \( C \) is denoted by \( \text{int} C \).

2 Frame representations

2.1 Definitions

Physical properties of the target signal \( y \in \ell^2(\mathbb{Z}) \), such as sparsity or spatial regularity, may be suitably expressed in terms of its coefficients \( x = (x(k))_{k \in \mathbb{Z}} \in \ell^2(\mathbb{Z}) \) where \( y = \sum_{k=-\infty}^{+\infty} \langle y \mid e_k \rangle e_k \) and \( (e_k)_{k \in \mathbb{Z}} \) denotes a dictionary of signals in \( \ell^2(\mathbb{Z}) \). Such a dictionary constitutes a frame if there exist two constants \( \mu \) and \( \overline{\mu} \in ]0, +\infty[ \) such that

\[
(\forall y \in \ell^2(\mathbb{Z})) \quad \mu \|y\|^2 \leq \sum_{k=-\infty}^{+\infty} |\langle y \mid e_k \rangle|^2 \leq \overline{\mu} \|y\|^2. \tag{3}
\]

The associated frame operator is the injective linear operator defined as

\[
(\forall y \in \ell^2(\mathbb{Z})) \quad Fy = (\langle y \mid e_k \rangle)_{k \in \mathbb{Z}}, \tag{4}
\]

the adjoint of which is the surjective linear operator given by

\[
(\forall x = (x(k))_{k \in \mathbb{Z}} \in \ell^2(\mathbb{Z})) \quad F^*x = \sum_{k=-\infty}^{+\infty} x(k) e_k. \tag{5}
\]

When \( F^{-1} = F^* \), an orthonormal basis is obtained. Further constructions as well as a detailed account of frame theory in Hilbert spaces can be found in [21].

A tight frame is such that, for some \( \mu \in ]0, +\infty[ \), \( F^*F = \mu I \) where \( I \) denotes the identity operator. In Condition (3), this means that \( \mu = \overline{\mu} = \mu \). A simple example of a tight frame is the union of \( \mu \) orthonormal bases. Other examples of tight frames can be found in [22, 23, 24].

2.2 A class of non necessarily tight frames

We consider a linear operator \( F \) which is basically obtained by cascading a non necessarily maximally decimated filter bank and a semi-orthogonal transform. A linear operator \( U : (\ell^2(\mathbb{Z}))^N \to (\ell^2(\mathbb{Z}))^Q \), with \( N \in \mathbb{N}^* \) and \( Q \in \mathbb{N}^* \), is said to be semi-orthogonal if there exists \( \mu_U \in ]0, +\infty[ \) such that \( U^*U = \mu_U I \). Recall that an analysis filter bank can be put under its polyphase form by
performing a polyphase decomposition $\Pi_D$ followed by a real MIMO (Multi-Input Multi-Output) filtering $V$ [25]:

- the polyphase decomposition $\Pi_D$ is an operator from $\ell^2(\mathbb{Z})$ to $(\ell^2(\mathbb{Z}))^D$ with $D \in \mathbb{N}^*$ such that, for every $y = (y(n))_{n \in \mathbb{Z}} \in \ell^2(\mathbb{Z})$, $\Pi_D y = (y^{(j)})_{1 \leq j \leq D}$ where $y^{(j)} = (y(Dn - j + 1))_{n \in \mathbb{Z}}$ is the $j$-th polyphase component of order $D$ of the signal $y$. The adjoint operator of $\Pi_D$ is given by $\Pi_D^* : (\ell^2(\mathbb{Z}))^D \rightarrow \ell^2(\mathbb{Z})$: $(y^{(j)})_{1 \leq j \leq D} \mapsto u = (u(n))_{n \in \mathbb{Z}}$ where, for every $n \in \mathbb{Z}$ and $j \in \{1, \ldots, D\}$, $u(Dn - j + 1) = y^{(j)}(n)$. So, $\Pi_D^*$ allows us to concatenate $D$ square summable sequences into a single one. It can be noticed that $\Pi_D^* \Pi_D = I$ and $\Pi_D \Pi_D^* = I$, which means that $\Pi_D$ is an isometry and $\Pi_D^{-1} = \Pi_D^*$.

- The MIMO filter $V$ is defined as

$$V = \begin{bmatrix} V_{1,1} & \ldots & V_{1,D} \\ \vdots & \ddots & \vdots \\ V_{N,1} & \ldots & V_{N,D} \end{bmatrix} \tag{6}$$

where, for every $i \in \{1, \ldots, N\}$ and $j \in \{1, \ldots, D\}$, $V_{i,j} : \ell^2(\mathbb{Z}) \rightarrow \ell^2(\mathbb{Z})$ is a SISO (Single-Input Single-Output) stable filter. Hence, the impulse response of this filter belongs to $\ell^1(\mathbb{Z})$ and its frequency response $\tilde{v}_{i,j}$ is a continuous function. In addition, it is assumed that $V$ is left invertible, that is: for every $\nu \in [-1/2, 1/2]$, the rank of the matrix $\hat{v}(\nu) = [\tilde{v}_{i,j}(\nu)]_{1 \leq i \leq N, 1 \leq j \leq D}$ is equal to $D$. The adjoint operator of $V$ is the $D \times N$ MIMO filter given by

$$V^* = \begin{bmatrix} V_{1,1}^* & \ldots & V_{N,1}^* \\ \vdots & \ddots & \vdots \\ V_{1,D}^* & \ldots & V_{N,D}^* \end{bmatrix} \tag{7}$$

where, for every $i \in \{1, \ldots, N\}$ and $j \in \{1, \ldots, D\}$, $V_{i,j}^*$ is the SISO filter with complex conjugate frequency response $\tilde{v}_{i,j}^*$.

We have then the following result (the proof is provided in Section 5.1):

**Proposition 2.1** The operator $F = \Pi_D^* U V \Pi_D$ is a frame operator with frame constants $\mu = \mu_U \inf_{\nu \in [-1/2, 1/2]} \sigma_{\min}(\nu)$ and $\overline{\mu} = \mu_U \sup_{\nu \in [-1/2, 1/2]} \sigma_{\max}(\nu)$, where, for every $\nu$, $\sigma_{\min}(\nu) \in ]0, +\infty[$ and $\sigma_{\max}(\nu) \in ]0, +\infty[$ are the minimum and maximum eigenvalues of $\hat{v}(\nu)^H \hat{v}(\nu)$. In addition, we have:

$$F^* F = \mu_U \Pi_D^* V^* V \Pi_D \tag{8}$$

The resulting frame is not necessarily tight. However, when $(\forall \nu \in [-1/2, 1/2]) \hat{v}(\nu)^H \hat{v}(\nu) = I$, the frame is tight. This includes paraunitary systems $V$ as particular cases when $D = N$.

Below, we provide examples of popular transforms in signal processing which belong to the class of considered frame representations.

### 2.2.1 Example 1 – Dual-tree transforms

In order to obtain low redundancy representations, frames such as the 2D $M$-band DTT have been proposed [26, 27, 28]. The real (resp. complex) DTT consists of performing $N = 2$ (resp.
$N = 4$) $M$-band orthonormal wavelet decompositions in parallel where, for every $i \in \{1, \ldots, N\}$, $U_i : \ell^2(\mathbb{Z}) \to \ell^2(\mathbb{Z})$ denotes the $i$-th orthogonal wavelet transform. An orthogonal combination of the subbands modeled by $\Phi : (\ell^2(\mathbb{Z}))^N \to (\ell^2(\mathbb{Z}))^N$ is applied to ensure directionality properties. $(U_i)_{1 \leq i \leq N}$ and $\Phi$ are related to $U$ by the following relation:

$$U = \Phi \begin{bmatrix} U_1 & 0 & \cdots & 0 \\ 0 & U_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & U_N \end{bmatrix}.$$  \hfill (9)

Since $\Phi^*\Phi = \Phi\Phi^* = I$ and, for every $i \in \{1, \ldots, N\}$, $U_i^*U_i = U_iU_i^* = I$, $U$ is an orthogonal transform. In the general form of DTT, each orthonormal wavelet decomposition is preceded by a prefiltering stage related to the discretization process. More precisely, it takes the form:

$$V = \begin{bmatrix} V_{1,1}^* \cdots V_{N,1}^* \end{bmatrix}^*. \hfill (10)$$

The variables $Q$ and $D$ as defined before are thus equal to $N$ and $1$, respectively. The left invertibility condition here reduces to the fact that the frequency response $|\hat{\nu}_{1,1}|^2 + \cdots + |\hat{\nu}_{N,1}|^2$ does not vanish. Note that, due to the presence of prefilters, the discrete DTT is not a tight frame in general. Furthermore, an extended class of DTTs can be obtained by using tight frame overcomplete representations $(U_i)_{1 \leq i \leq N}$ having the same frame constant $\mu_U$ (e.g. redundant wavelet representations derived from orthogonal filters after appropriate renormalization), in which case $U$ as defined by (9) is a semi-orthogonal transform.

#### 2.2.2 Example 2 – Filter banks

Analysis filter banks with perfect reconstruction correspond to the case when $U = I$ and $Q = N$. $D$ is the decimation factor and $N$ is the number of channels. The redundancy introduced by such a filter bank structure is $N/D \geq 1$. More details about filter bank design can be found in [29]. Lapped transforms [30] can also be implemented with filter bank structures. As already mentioned, filter banks do not constitute tight frames in general. Note that if $D = 1$, a fully undecimated filter bank is obtained.

Tight frame representations have been widely used for data recovery by using convex optimization methods. In the next section, we recall some convex optimization tools and show the relevance of the considered class of frames in recent optimization approaches.

### 3 Use of parallel proximal algorithms

A number of algorithms such as the alternating split Bregman algorithm [31], augmented Lagrangian techniques [14] and parallel proximal methods [32] have been recently proposed to address possibly nonsmooth convex optimization problems encountered in the solution of restoration problems. Here, we will focus on this specific class of algorithms which have proven to be useful
for solving problems like (1) and (2) when a tight frame representation is employed. A common feature of the aforementioned algorithms is that they require a large-size linear system to be solved at each iteration. When non tight frame representations are used, the computational cost of the associated inversion may become prohibitive. We will see however that for the class of frames introduced in Section 2, this inversion can be performed in an efficient manner. This fact will be demonstrated by considering an extension of the Parallel ProXimal Algorithm (PPXA), hereafter designated as PPXA+. As shown in [32], PPXA+ is closely related to the alternating direction method of multipliers and its parallel extensions [15].

We recall that the proximity operator [33] of a function \( \varphi \in \Gamma_0(\mathcal{H}) \) is defined as

\[
\text{prox}_{\varphi} : \mathcal{H} \to \mathcal{H} : v \mapsto \arg \min_{u \in \mathcal{H}} \frac{1}{2} \|u - v\|^2 + \varphi(u).
\] (11)

It can be observed that when \( \varphi \) is the indicator function \( \iota_C \) of a nonempty closed convex subset \( C \) of \( \mathcal{H} \), i.e. it takes on the value 0 in \( C \) and \( +\infty \) in \( \mathcal{H} \setminus C \), \( \text{prox}_{\iota_C} \) reduces to the projector \( P_C \) onto \( C \). Other examples of proximity operators corresponding to potential functions of standard log-concave univariate probability densities have been listed in [5, 11, 13]. The proximity operators employed in the experimental part of this paper, are recalled below.

**Example 3.1 (soft-thresholding rule)** Let \( \alpha > 0 \), and set \( \varphi : \mathbb{R} \to \mathbb{R} : \xi \mapsto \alpha |\xi| \). Then, for every \( \xi \in \mathbb{R} \), \( \text{prox}_{\varphi} \xi = \text{sign}(\xi) \max\{|\xi| - \alpha, 0\} \).

**Example 3.2 (Poisson minus likelihood function)** [11]

Let \( \alpha > 0 \), \( \chi \geq 0 \), and set

\[
\varphi : \mathbb{R} \to ]-\infty, +\infty] : \xi \mapsto \begin{cases} 
-\chi \ln(\xi) + \alpha \xi, & \text{if } \chi > 0 \text{ and } \xi > 0; \\
\alpha \xi, & \text{if } \chi = 0 \text{ and } \xi \geq 0; \\
+\infty, & \text{otherwise}. 
\end{cases}
\] (12)

Then, for every \( \xi \in \mathbb{R} \),

\[
\text{prox}_{\varphi} \xi = \frac{\xi - \alpha + \sqrt{\|\xi - \alpha\|^2 + 4\chi}}{2}. 
\] (13)

One of the difficulties in the resolution of problems like (1) or (2) is that they involve linear operators. Unfortunately, the proximity operator of the composition of a linear operator and a convex function takes a closed form expression only under restrictive assumptions as stated below:

**Proposition 3.3** [12]

Let \( \mathcal{G} \) be a real Hilbert space, \( \varphi \in \Gamma_0(\mathcal{G}) \) and let \( L : \mathcal{G} \to \mathcal{H} \) denotes a bounded linear operator. Suppose that \( LL^* = \chi I \), for some \( \chi \in ]0, +\infty[ \). Then, \( \varphi \circ L \in \Gamma_0(\mathcal{H}) \) and \( \text{prox}_{\varphi \circ L} = I + \chi^{-1}L^*(\text{prox}_{\chi \varphi} - I)L \).

When dealing with linear operators \( L \) such that \( LL^* \neq \chi I \) for any \( \chi > 0 \), proximal algorithms requiring the inversion of a linear operator at each iteration can however be designed. For example, Algorithm 1 (resp. Algorithm 2) can be applied to Problem 1 (resp. Problem 2).
these algorithms, the sequences \((a_{r,\ell})_{1 \leq r \leq R}\) and \((b_{s,\ell})_{1 \leq s \leq S}\) model possible numerical errors in the computation of the proximity operators at iteration \(\ell\).

**Algorithm 1**

Initialization
\[
(\eta_r)_{1 \leq r \leq R} \in [0, +\infty[^R, (\kappa_s)_{1 \leq s \leq S} \in [0, +\infty[^S ; \; (v_{r,0})_{1 \leq r \leq R} \in (\ell^2(\mathbb{Z}))^R, (w_{s,0})_{1 \leq s \leq S} \in (\ell^2(\mathbb{Z}))^S \\
x_0 = \arg \min_{u \in \ell^2(\mathbb{Z})} \sum_{r=1}^{R} \eta_r \| L_r F^* u - v_{r,0} \|^2 + \sum_{s=1}^{S} \kappa_s \| u - w_{s,0} \|^2 
\]

For \(\ell = 0, 1, \ldots\)
\[
\text{For } r = 1, \ldots, R \quad p_{r,\ell} = \text{prox}_{\frac{1}{\eta_r} v_{r,\ell}} + a_{r,\ell} \\
\text{For } s = 1, \ldots, S \quad r_{s,\ell} = \text{prox}_{\frac{1}{\kappa_s} w_{s,\ell}} + b_{s,\ell} \\
\lambda_\ell \in [0, 2[ \\
c_\ell = \arg \min_{u \in \ell^2(\mathbb{Z})} \sum_{r=1}^{R} \eta_r \| L_r F^* u - p_{r,\ell} \|^2 + \sum_{s=1}^{S} \kappa_s \| u - r_{s,\ell} \|^2 \\
\text{For } r = 1, \ldots, R \quad v_{r,\ell+1} = v_{r,\ell} + \lambda_\ell (L_r F^* (2c_\ell - x_\ell) - p_{r,\ell}) \\
\text{For } s = 1, \ldots, S \quad w_{s,\ell+1} = w_{s,\ell} + \lambda_\ell (2c_\ell - x_\ell - r_{s,\ell}) \\
x_{\ell+1} = x_\ell + \lambda_\ell (c_\ell - x_\ell)
\]

**Algorithm 2**

Initialization
\[
(\eta_r)_{1 \leq r \leq R} \in [0, +\infty[^R, (\kappa_s)_{1 \leq s \leq S} \in [0, +\infty[^S ; \; (v_{r,0})_{1 \leq r \leq R} \in (\ell^2(\mathbb{Z}))^R, (w_{s,0})_{1 \leq s \leq S} \in (\ell^2(\mathbb{Z}))^S \\
y_0 = \arg \min_{u \in \ell^2(\mathbb{Z})} \sum_{r=1}^{R} \eta_r \| L_r u - v_{r,0} \|^2 + \sum_{s=1}^{S} \kappa_s \| F u - w_{s,0} \|^2 
\]

For \(\ell = 0, 1, \ldots\)
\[
\text{For } r = 1, \ldots, R \quad p_{r,\ell} = \text{prox}_{\frac{1}{\eta_r} v_{r,\ell}} + a_{r,\ell} \\
\text{For } s = 1, \ldots, S \quad r_{s,\ell} = \text{prox}_{\frac{1}{\kappa_s} w_{s,\ell}} + b_{s,\ell} \\
\lambda_\ell \in [0, 2[ \\
c_\ell = \arg \min_{u \in \ell^2(\mathbb{Z})} \sum_{r=1}^{R} \eta_r \| L_r u - p_{r,\ell} \|^2 + \sum_{s=1}^{S} \kappa_s \| F u - r_{s,\ell} \|^2 \\
\text{For } r = 1, \ldots, R \quad v_{r,\ell+1} = v_{r,\ell} + \lambda_\ell (L_r (2c_\ell - y_\ell) - p_{r,\ell}) \\
\text{For } s = 1, \ldots, S \quad w_{s,\ell+1} = w_{s,\ell} + \lambda_\ell (F (2c_\ell - y_\ell) - r_{s,\ell}) \\
y_{\ell+1} = y_\ell + \lambda_\ell (c_\ell - y_\ell)
\]

The convergence of the sequence \((x_\ell)_{\ell \in \mathbb{Z}}\) (resp. \((y_\ell)_{\ell \in \mathbb{Z}}\)) generated by Algorithm 1 (resp. Algorithm 2) to an optimal solution of the related optimization problem is guaranteed under the following technical assumptions (see [32] for more details):

**Assumption 3.4**

1. \((\bigcap_{r=1}^{R} \text{int dom } f_r \circ L_r F^*) \cap \left( \bigcap_{s=1}^{S-1} \text{int dom } g_s \right) \cap \text{dom } g_S \neq \emptyset\)

   (resp. \((\bigcap_{r=1}^{R} \text{int dom } f_r \circ L_r) \cap \left( \bigcap_{s=1}^{S-1} \text{int dom } g_s \circ F \right) \cap \text{dom } g_S \circ F \neq \emptyset\).)

2. There exists \(\Lambda \in [0, 2[\) such that \((\forall \ell \in \mathbb{N}), \; \Lambda \leq \lambda_{\ell+1} \leq \lambda_\ell\).

3. \((\forall r \in \{1, \ldots, R\}), \sum_{\ell=0}^{+\infty} \| a_{r,\ell} \| < +\infty\) and \((\forall s \in \{1, \ldots, S\}), \sum_{\ell=0}^{+\infty} \| b_{s,\ell} \| < +\infty\).
Note that quadratic minimizations need to be performed in the initialization step and in the computation of the intermediate variable \( c_\ell \) at iteration \( \ell \). These amount to inverting a large-size linear operator which is untractable for arbitrary frames. We will now show that the class of frame operators \( F \) introduced in Section 2 allows us to overcome this difficulty.

To do so, recall that the convolutive operators \( (L_r)_{1 \leq r \leq R} \) can be put under a polyphase form:

\[
L_r = W_r \Pi_D \quad (\forall r \in \{1, \ldots, R\})
\]

where \( W_r = [W_{r,1}, \ldots, W_{r,D}] \) is a MISO (Multi-Input Single-Output) filter (for every \( j \in \{1, \ldots, D\} \), \( W_{r,j} : \ell^2(\mathbb{Z}) \to \ell^2(\mathbb{Z}) \) is a SISO filter). Now, invoking Proposition 2.1 and making use of Woodbury identity yield:

- for SF (Algorithm 1)

\[
\kappa \left( \sum_{r=1}^{R} \eta_r F L^*_r L_r + \kappa I \right)^{-1} = I - F \left( \sum_{r=1}^{R} \eta_r L^*_r L_r \right) \left( \kappa I + F^* F \left( \sum_{r=1}^{R} \eta_r L^*_r L_r \right) \right)^{-1} F^*
\]

\[
= I - F \Pi_D \left( \sum_{r=1}^{R} \eta_r W^*_r W_r \right) \Pi_D \left( \kappa I + \mu_U \Pi_D^* V^* V \left( \sum_{r=1}^{R} \eta_r W^*_r W_r \right) \Pi_D \right)^{-1} F^*
\]

\[
= I - F \Pi_D \left( \sum_{r=1}^{R} \eta_r W^*_r W_r \right) \left( \kappa I + \mu_U V^* V \left( \sum_{r=1}^{R} \eta_r W^*_r W_r \right) \right)^{-1} \Pi_D F^*,
\]

(15)

- for AF (Algorithm 2)

\[
\left( \sum_{r=1}^{R} \eta_r L^*_r L_r + \kappa F^* F \right)^{-1} = \Pi_D \left( \sum_{r=1}^{R} \eta_r W^*_r W_r + \kappa \mu_U V^* V \right)^{-1} \Pi_D,
\]

(16)

where \( \kappa = \sum_{s=1}^{S} \kappa_s \). The inversions in (15) (resp. (16)) can be performed by noticing that, as \( (W_{r,j})_{1 \leq r \leq R} \) and \( V \) are multivariate filters with frequency responses \( (\hat{\omega}_r)_{1 \leq r \leq R} \) and \( \hat{v} \),

\[
\kappa I + \mu_U V^* V \left( \sum_{r=1}^{R} \eta_r W^*_r W_r \right) \quad (resp. \quad \sum_{r=1}^{R} \eta_r W^*_r W_r + \kappa \mu_U V^* V)
\]

is a MIMO filter with frequency response: for every \( \nu \in [-1/2, 1/2] \),

\[
\kappa I + \mu_U \hat{v}(\nu)^H \hat{v}(\nu) \left( \sum_{r=1}^{R} \eta_r \hat{w}_r(\nu)^H \hat{w}_r(\nu) \right) \quad (resp. \quad \sum_{r=1}^{R} \eta_r \hat{w}_r(\nu)^H \hat{w}_r(\nu) + \kappa \mu_U \hat{v}(\nu)^H \hat{v}(\nu)).
\]

Hence, by resorting to Fast Fourier Transform implementations, the problem reduces to the inversion of \( D \times D \) matrices.

4 Experimental results

We apply the proposed optimization method to the restoration of images degraded by a blur and a Poisson (resp. Laplace) noise. For simplicity sake, the proposed formulation has been presented
in the 1D case, but its 2D extension is straightforward. The scaling factor of the Poisson (resp. Laplace) noise is $\alpha = 0.1$ (resp. $\alpha = 10$). The data fidelity term is related to the Poisson (resp. Laplace) antilog-likelihood, which corresponds to the potential function $\varphi$ defined by (12) (resp. by Example 3.1). In the considered examples, the regularization term simply is an $\ell_1$-norm modeling the sparsity of $\mathbf{y}$ through a frame representation. Two different frames are considered: a (Complex) Dual-Tree Transform, presented in Section 2.2.1, and an eigenfilter bank decomposition [34, 35], which is a special case of the subband structure presented in Section 2.2.2. We constrain the data values to belong to $[0, 255]$, so defining a closed convex constraint set $C$. The considered SF (resp. AF) problem is a particular case of Problem (1) (resp. (2)) where $R = 2$, $S = 1$, $f_1 = \varphi$, $f_2 = \iota_C$, and $g_1 = \tau \parallel \cdot \parallel_1$ with $\tau > 0$. The last function corresponds to the regularization term operating in the frame domain. In the considered problems, $L_1 = T$ and $L_2 = I$. The proximity operators associated to $f_1$, $f_2$, and $g_1$ are derived from Example 3.2, the projection onto $C$, and Example 3.1. In our simulations, the parameter $\tau$ is empirically chosen to maximize the signal-to-noise-ratio (SNR). In general, the value of this parameter is not the same for FA and FS.

Note that an alternative to the proposed approach consists of resorting to primal-dual algorithms [36, 37, 38, 39, 40]. These algorithms are appealing as they do not require any operator inversion and they can thus be employed with arbitrary frames. However, after appropriate choices for the weights $(\eta_r)_{1 \leq r \leq R}$, $(\kappa_s)_{1 \leq s \leq S}$, and the relaxation parameter $\lambda_{\ell}$, the proposed approach appeared to be faster. For example, similar frequency domain implementations of the Forward Backward Forward (FBF) algorithm were observed to be about twenty times slower than the proposed method in term of iterations and computation time.

Figure 1 shows the restoration results for a cropped version of the “Barbara” image in the presence of Poisson noise and a uniform blur of size $5 \times 5$. We adopt a SF criterion and we consider a tight version (i.e., for every $i \in \{1, \ldots, 4\}$, $V_{i,1} = I$) as well as a non-tight version of Complex DTT. The Complex DTT [28] is computed using symlets of length 6 over 3 resolution levels. In order to efficiently perform the inversions in (15) and (16), fast Fourier diagonalization techniques have been employed. The use of the non-tight Complex DTT including prefilters allows us to improve the quality of the results both visually and in terms of SNR and SSIM [41]. Figure 2 displays a second restoration example for a cropped version of the “Straw” image in the presence of Laplace noise and a uniform blur of size $5 \times 5$. AF results are presented by using a DTT and an eigenfilter bank ($D = 8$ and $N = 14$) computed from the degraded image. This formulation leads to better results than those obtained with SF. Significant gains in favour of the eigenfilter bank can be observed.
Figure 1: Cropped versions of Barbara image (size $128 \times 128$). Restored images using SF and complex DTT.

Figure 2: Cropped versions of Straw image (size $128 \times 128$). Restored images using AF considering DTT and eigenfilter banks.
5 Appendix

5.1 Proof of Proposition 2.1

Eq. (8) is a direct consequence of the fact that $U$ is semi-orthogonal and $\Pi_\mathbf{Q}$ is an isometry. By using Parseval’s formula, for every $y \in \ell^2(\mathbb{Z})$,

$$
\|Fy\|^2 = \mu_U \|V_\mathbf{D}y\|^2 = \mu_U \int_{-1/2}^{1/2} \|\hat{\mathbf{v}}(\nu)\hat{\mathbf{y}}(\nu)\|^2 d\nu
$$

(18)

where $\hat{\mathbf{y}}(\nu) = [\hat{y}^{(1)}(\nu), \ldots, \hat{y}^{(D)}(\nu)]^\top$, and, for every $j \in \{1, \ldots, D\}$, $\hat{y}^{(j)}$ is the Fourier transform of the discrete signal $y^{(j)}$. By using the fact that, for every $\nu \in [-1/2, 1/2]$, $\sigma_{\min}(\nu)\|\hat{\mathbf{y}}(\nu)\|^2 \leq \|\hat{\mathbf{v}}(\nu)\hat{\mathbf{y}}(\nu)\|^2 \leq \sigma_{\max}(\nu)\|\hat{\mathbf{y}}(\nu)\|^2$ we get

$$
\mu_U \inf_{\nu \in [-1/2, 1/2]} \sigma_{\min}(\nu) \int_{-1/2}^{1/2} \|\hat{\mathbf{y}}(\nu)\|^2 d\nu \leq \|Fy\|^2 \leq \mu_U \sup_{\nu \in [-1/2, 1/2]} \sigma_{\min}(\nu) \int_{-1/2}^{1/2} \|\hat{\mathbf{y}}(\nu)\|^2 d\nu.
$$

(19)

Besides, we have

$$
\int_{-1/2}^{1/2} \|\hat{\mathbf{y}}(\nu)\|^2 d\nu = \sum_{j=1}^{D} \int_{-1/2}^{1/2} |\hat{y}^{(j)}(\nu)|^2 d\nu = \sum_{j=0}^{D-1} \sum_{n=-\infty}^{+\infty} |y(Dn-j)|^2 = \sum_{n=-\infty}^{+\infty} |y(n)|^2 = \|y\|^2.
$$

(20)

Now, it can be noticed that since $\mathbf{v}$ is continuous, $\sigma_{\min}$ and $\sigma_{\max}$ are continuous functions too. By invoking Weierstrass theorem, there thus exist $\nu \in [-1/2, 1/2]$ and $\mathbf{v} \in [-1/2, 1/2]$ such that $\inf_{\nu \in [-1/2, 1/2]} \sigma_{\min}(\nu) = \sigma_{\min}(\mathbf{v})$ and $\sup_{\nu \in [-1/2, 1/2]} \sigma_{\max}(\nu) = \sigma_{\max}(\mathbf{v})$. Finally, $\sigma_{\min}(\mathbf{v}) > 0$ since $\rank(\hat{\mathbf{v}}(\nu)^\top \hat{\mathbf{v}}(\nu)) = \rank(\hat{\mathbf{v}}(\nu)) = D$.

References


