DESIGNING GABOR WINDOWS USING CONVEX OPTIMIZATION

NATHANAËL PERRAUDIN AND NICKI HOLIGHAUS, PETER L. SØNDERGAARD AND PETER BALAZS

Abstract. Redundant Gabor frames admit an infinite number of dual frames, yet only the canonical dual Gabor system, constructed from the minimal $\ell^2$-norm dual window, is widely used. This window function however, might lack desirable properties, such as good time-frequency concentration, small support or smoothness. We employ convex optimization methods to design dual windows satisfying the Wexler-Raz equations and optimizing various constraints. Numerical experiments suggest that alternate dual windows with considerably improved features can be found.

1. Introduction

The construction of good synthesis filterbanks is an important topic for signal processing. For perfect reconstruction filterbanks several methods have been proposed in the signal processing literature, relying on several approaches, like polyphase representation [39] or algebraic methods [32], also on frame theory [7], similar to our approach. Modulated cosine filterbanks, which are closely related to Gabor systems (sampled short-time Fourier transforms) [20] [19], are ubiquitous in signal processing. Gabor transforms represent a signal as superposition of translates and modulations of a single window function. A window function is particularly suitable for representation and processing if it is well-concentrated in time and frequency.

A signal can be reconstructed from its Gabor transform using a dual system with the same modulation and translation structure. Moreover, infinitely many such systems exist if the Gabor transform is redundant, which allows a choice of the synthesis system given other constraints [5]. Yet, traditional methods utilize the so-called canonical dual window, which is computed by applying the inverse frame operator to the analysis window. It is the only dual window that can be found by the application of a linear operator to the original window. Since, in this contribution, we are interested in finding a dual system with desirable properties given a prescribed analysis window, an alternative method is required.

As an example, a compactly supported window leads to perfect reconstruction with a finite number of computations. The canonical dual window is only guaranteed to be compactly supported in the special case where the number of frequency channels is greater or equal to the support the window
(known as the painless case), though. The technique presented in this paper allows to get rid of this limit. A method for computing dual windows satisfying specific support constraints was proposed by Strohmer [37], based on the Moore-Penrose pseudoinverse of a linear equation system describing duality and the support conditions. The method therein allows the use of other regularization constraints, if these can be expressed through a Hermitian positive definite matrix, e.g. weighted $\ell^2$-norms. Already Wexler and Raz [41] impose linear constraints to find alternative dual windows, while Daubechies [16] presents a formula for finding dual windows that are optimal in a modified $L^2$ sense.

We will introduce a more general method that allows for simple fine-tuning of multiple optimization criteria at once. In general, the construction of dual pairs of compactly supported Gabor windows has received much interest, see e.g. [10, 11] and references therein.

Since all dual windows perform perfect reconstruction from unmodified Gabor coefficients, the purpose of constructing alternative dual windows might not immediately be obvious beside the simple minimization of the support. However, if the coefficients are modified, e.g. through signal processing procedures such as frame multipliers [18, 3, 36], the shape of the dual window plays an important role in the quality and localization of the performed modifications, see e.g. Figure 1. Therefore, we desire to improve certain criteria over the canonical dual window.

We propose a convex optimization scheme that selects a dual window such that several regularization parameters of the user’s choice are optimized. As the solution of a convex optimization problem, the proposed window will be optimal with regards to the selected criteria, cf. Section 2.2, provided that the set of admissible dual windows is nonempty.

Recently, convex optimization in the context of audio signal processing has grown into a active field of research and in particular proximal splitting methods [13, 12, 14] have been used to great effect, e.g. in audio inpainting [2, 1] and sparse representation [25]. In those cases, optimization techniques are applied directly to the signal or its time-frequency representation. In this contribution, we apply optimization techniques to shape the building blocks of the time-frequency representation instead.

More similar to our approach, the authors in [26] solve a different convex optimization problem to find sparse dual systems. Using a weighted $\ell^1$-norm. They also provide a method to obtain compactly supported dual windows. This problem is optimized for the research of sparse windows and not suitable for other constraint such as smoothness. With the method proposed in this paper, similar results can be obtained, but sparse dual windows only form a particular example for its possible applications.

We also introduce a new way to design tight windows, based on the proposed convex optimization technique. However, this case is beyond the assumption of our methods, i.e. it is not a convex problem. Even though
Figure 1. Reconstruction from modified coefficients: (a) Synthetic test signal and its spectrogram (b) Analysis window used to compute the spectrogram and modified spectrogram. (c,d) 2 different dual windows recomputed spectrograms after synthesis with the corresponding dual window to the left. Note that the smearing effect depends on the concentration of the synthesis window and is considerably worse in the 2nd case.
we lack any theoretical guarantee, we observe experimentally that favorable results can be obtained.

The rest of this article is structured as follows. In the next section, we repeat the necessary background from Gabor and frame theory, as well as convex optimization. The convex optimization problem that forms the backbone of our method is described in detail in Section 3, where we also recall, respectively introduce, some of the regularization functions we have found most useful in our experiments and an algorithm designed to solve the problem. Finally, Section 4 presents a wealth of examples showing the effect of different regularizers and possible applications of our method.

Some content of this contribution has been presented as [29].

2. Preliminaries

In this contribution, we consider sampled functions, i.e. sequences \( f \) in \( \ell^2(\mathbb{Z}) \) or \( \mathbb{C}^L \). Here, the latter is interpreted as the space of \( L \)-periodic sequences with indices considered modulo \( L \). For such \( f \), we refer to the smallest closed interval containing all nonzero values of \( f \) as its support, denoted by \( \text{supp}(f) \).

Furthermore, we denote translation and modulation operators

\[
T_n f[l] = f[l - n] \quad \text{and} \quad M_\omega f[l] = f[l] e^{2\pi i \omega l},
\]

for \( f \in \ell^2(\mathbb{Z}) \), \( l, n \in \mathbb{Z} \) and \( \omega \in [0, 1) \). Their counterparts on \( \mathbb{C}^L \) are defined in the usual way with indexing modulo \( L \).

2.1. Gabor systems and frames. We define the Gabor system

\[
G(g, a, M) := (g_{m,n} = M_{m/M} T_{n a} g)_{n \in \mathbb{Z}, \ m = 0, \ldots, M - 1},
\]

for a given window \( g \in \ell^2(\mathbb{Z}) \), the hop size \( a \in \mathbb{Z} \) and the number of frequency bins \( M \in \mathbb{Z} \).

For any \( f \in \ell^2(\mathbb{Z}) \), the Gabor coefficients (or Gabor analysis) with respect to \( G(g, a, M) \) is given by

\[
(G f)[m + nM] = \langle f, g_{m,n} \rangle = \sum_{l \in \mathbb{Z}} f[l] \overline{g_{m,n}[l]},
\]

with the analysis operator \( G \) given as the infinite matrix \( G[m + nM, l] := G_{g,a,M}[m + nM, l] := g_{m,n}[l] \).

The coefficients \( (G f)[m+nM] \) can be interpreted as a sequence of Fourier transforms of the signal \( f \), localized by multiplication with a translate of the window \( g \).

Furthermore, we can perform Gabor synthesis of a coefficient sequence \( c \in \ell^2(\mathbb{Z}) \) with respect to \( G(g, a, M) \):

\[
f_{\text{syn}}[l] = (G^* c)[l] = \sum_{n \in \mathbb{Z}} \sum_{m = 0}^{M-1} c[m + nM] g_{m,n}[l],
\]
where $G^*$ denotes the transpose conjugate of $G$. A maximally redundant Gabor transform, setting $a = 1$, $M = L$, is also known as (unsampled) short-time Fourier transform (STFT). Only this highly overcomplete setup allows for straightforward inversion using the synthesis operator, i.e. $f = G^*Gf$, however. Otherwise, we require that $G(g, a, M)$ forms a stable, (over-)complete system, i.e. a frame.

A Gabor system constitutes a Gabor frame [9] if constants $0 < A \leq B < \infty$ exist, such that

$$(4) \quad A\|f\|_2^2 \leq \|Gf\|_2^2 \leq B\|f\|_2^2, \text{ for all } f \in \ell^2(\mathbb{Z}).$$

In that case, the closed linear span of its elements equals $\ell^2(\mathbb{Z})$ and every sequence $f \in \ell^2(\mathbb{Z})$ can be written as

$$f = G^*_{g,a,M}c$$

for some coefficient sequence $c \in \ell^2(\mathbb{Z})$.

Moreover, the signal can be reconstructed from the Gabor coefficients by means of a dual Gabor frame $G(h, a, M)$, i.e. $f = G_{h,a,M}^* (G_{g,a,M} f)$. Since frames are mutually dual, also $f = G_{g,a,M}^* (G_{h,a,M} f)$ holds for all $f \in \ell^2(\mathbb{Z})$, showing that the sequence $c$ in (5) can be chosen to be the Gabor coefficients with respect to $G(h, a, M)$. If $G(g, a, M)$ is redundant, then the dual window $h \in \ell^2(\mathbb{Z})$ is not unique.

Instead, the space of dual windows forms an affine subspace of $\ell^2(\mathbb{Z})$. It equals the solution set of the Wexler-Raz (WR) equations [11, 12], that provide a necessary and sufficient condition for a function $h \in \ell^2(\mathbb{Z})$ to be a dual Gabor window for $G(g, a, M)$. They are given by

$$M \left\langle h, g[-nm]e^{2\pi im/a} \right\rangle = \delta[n]\delta[m], \quad \text{for } m = 0, \ldots, a - 1, \ n \in \mathbb{Z}. \quad (6)$$

Here, $\delta$ denotes the Kronecker delta. With the analysis matrix $G^\circ = G_{g,M,a}$, i.e. switching the role of $a$ and $M$, they can be stated as

$$G^\circ h = \frac{a}{M} \delta, \quad (7)$$

While any of these dual windows facilitates perfect reconstruction from unmodified coefficients, their features influence synthesis results from processed coefficients as illustrated by Figure 4.

A particular dual window is the so-called canonical dual window, obtained by applying the inverse frame operator to the frame elements. The frame operator is given by the concatenation $S_{g,a,M} := G_{g,a,M}^* G_{g,a,M}$ of the analysis and synthesis operators.

The frame operator is invertible if and only if $G(g, a, M)$ forms a frame, in which case the canonical dual frame is given by $G(\gamma, a, M)$, where

$$\gamma = S^{-1}_{g,a,M}g$$
is the canonical dual window. The synthesis operator of $\gamma$ coincides with the pseudoinverse of the original analysis operator, i.e. $G_{\gamma,a,M} = G^\dagger$. Consequently, the inversion formula reads
\[ f[l] = \sum_{m,n} \langle f, g_{m,n} \rangle \gamma_{m,n}[l] = G^\dagger G f[l]. \]
Several approaches exist for finding the canonical dual efficiently, cf. \cite{6, 22}. Unless the length of the window $g$ is less than or equal to the number of channels $M$ (sometimes referred to as the \textit{painless} case), the canonical dual is most often infinitely long. This has made the construction where the length of the window equals the number of channels omnipresent in signal processing, to the point where these two numbers are not distinguished.

We note that the canonical dual window $\gamma$ minimizes both the $\ell^2$-norm as well as the $\ell^2$ distance to $g$ among all duals, see e.g \cite[Prop. 7.6.2]{20}. However, alternate duals might possess properties preferable to those of the canonical dual, e.g. shorter support, better localization or smoothness.

In the following sections, we will select dual windows optimizing certain desired properties and/or satisfying a support constraint from the solution set of the WR equations (6) by means of convex optimization.

2.2. Proximal splitting. The convex optimization problems we consider are of the form
\[ \text{minimize} \sum_{i=1}^{K} f_i(x), \]
where the $f_i$ are convex functions. Note that if at least one function $f_i$ is not differentiable, it is not possible to apply smooth optimization techniques \cite{21}. A simple example for a nondifferentiable choice of $f_i$, is given by $\ell^1$ regularization. Proximal splitting methods \cite{13}, on the other hand may still apply, so long as the $f_i$ are at least lower-semi-continuous. More precisely, we require $f_i \in \Gamma_0(\mathbb{R}^L)$ for all $i$, where
\[ \Gamma_0(\mathbb{R}^L) = \{ f : \mathbb{R}^L \to \mathbb{R} : f \text{ lower semi-continuous, convex and proper} \}. \]

\textbf{Definition 1.} The proximity operator of a function $f \in \Gamma_0(\mathbb{R}^L)$ is defined by
\[ \text{prox}_f(y) := \arg \min_{x \in \mathbb{R}^L} \left\{ \frac{1}{2} \|y - x\|_2^2 + f(x) \right\}. \]
Since $f$ is convex, the minimization problem in (10) has a unique solution for every $y \in \mathbb{R}^L$ and consequently $\text{prox}_f : \mathbb{R}^L \to \mathbb{R}^L$ is well-defined.

More information on the properties of proximity operators can be found in \cite{31, 27, 13}.
In many cases, we would like to restrict that set of admissible functions to a subset $C$ of $\mathbb{R}^L$, i.e. find the optimal solution to (9) considering only functions in $C$. This is achieved by the indicator function. For any nonempty, closed and convex set $C \subset \mathbb{R}^L$, the indicator function $i_C$ of $C$ is defined as

\[
i_C : \mathbb{R}^L \to \{0, +\infty\} : x \mapsto \begin{cases} 0, & \text{if } x \in C \\ +\infty, & \text{otherwise} \end{cases}
\]

The corresponding proximity operator is given by the projection onto the set $C$:

\[
P_C(y) = \arg\min_{x \in \mathbb{R}^L} \left\{ \frac{1}{2} \|y - x\|^2 + i_C(x) \right\} = \arg\min_{x \in C} \left\{ \|y - x\|^2 \right\}
\]

Such restrictions are called constraints and can be given, e.g. by a set of linear equations that the solution is required to satisfy. We rewrite an optimization problem on $C$ into a problem of the form (9):

\[
\begin{align*}
\arg\min_{x \in C} \sum_{i=1}^K \lambda_i f_i(x) &= \arg\min_{x \in \mathbb{R}^L} \sum_{i=1}^K \lambda_i f_i(x) + i_C.
\end{align*}
\]

We also call $C = \{ x \in \mathbb{R}^L : x \text{ satisfies the constraints} \}$ the set of admissible points. If $C$ is nonempty and convex, Equation (12) has a solution for any given choice of regularization parameters $\lambda_i$, uniquely determined if at least one $f_i$ is strictly convex.

### 3. Design of Optimal Dual Windows

Suppose that for a a given Gabor frame $G(g, a, M)$, we are looking for a dual window which is concentrated in time and in frequency. In this section, we transform those requirements into a convex optimization problem, subsequently solved by algorithms proven to converge to an optimal solution. That solution is the desired dual window. The set of eligible functions is described by a set of constraints, i.e. compulsory requirements on the solution. In this case these are duality and possibly support constraints. Both are given by a linear equation system, the WR equations (6) in the case of duality. Note that the solution set of a linear equation system is always convex.

Furthermore, objective functions $f_i$ are chosen in order to select a dual windows in the set $C_{\text{dual}}$ (or $C_{\text{dual}} \cap C_{\text{supp}}$) that is optimal with regards to certain criteria, see Section 3.1 Those properties can be balanced by regularizations parameter $\lambda_i$ in order to fine-tune the window. A list of common objective functions with their effect is given in Table 1.

When computing dual Gabor windows for a frame $G(g, a, M)$, with finite $L_g = |I_g|$ and $I_g = \text{supp}(g)$ the support of $g$, the most general problem we
Consider is

\[
\arg \min_{x \in \mathcal{C}_{dual}} \sum_{i=1}^{K} f_i(x).
\]

Here, \( \mathcal{C}_{dual} \) is the set of functions \( x \in \ell^2(\mathbb{Z}) \) that satisfy the WR equations. In general, a solution of the above problem will have infinite support and thus cannot be computed in finite time. Therefore, we add an additional support constraint.

This support constraint serves two distinct purposes. Firstly, restricting the support of the solution to a length \( L \), with \( a, M \mid L \) and \( L \geq L_g \) allows computing a solution in finite time since both the number of relevant, i.e. not trivially solved, WR equations as well as the number of computations in each optimization step is finite. Secondly, we can restrict the support further, to a length of \( L_h < L \) to design dual windows with small support.

Let \( I_h \) be the interval of length \( L_h \) on which we wish to constrain the support of the solution, then the final convex problem is of the form

\[
\arg \min_{x \in \mathcal{C}_{dual} \cap \mathcal{C}_{supp}} \sum_{i=1}^{K} f_i(x),
\]

where \( \mathcal{C}_{supp} \) is the set of functions supported on \( I_h \). Since \( \mathcal{C}_{supp} = \{ x \in \ell^2(\mathbb{Z}) : Vx = 0 \} \), where the square matrix \( V \) is given by

\[
V_{i,j} = \begin{cases} 
1 & \text{if } i = j \text{ and } i \notin I_h, \\
0 & \text{else},
\end{cases}
\]

the set \( \mathcal{C}_{supp} \) is also convex. Consequently, \( \mathcal{C}_{dual} \cap \mathcal{C}_{supp} \) is also a convex set. From here on, we assume \( \mathcal{C}_{dual} \cap \mathcal{C}_{supp} \neq \emptyset \) for the rest of this section.

**Remark 1.** It is known \cite{37} that the WR equations are linearly independent. The same can easily be seen for the equations describing the support set \( \mathcal{C}_{supp} \). Therefore, controlling the number of equations is sufficient for ensuring solvability. However, when jointly considering both equation systems, we have observed linear dependencies for nonrandom analysis windows. Thus, this restriction is not optimal. An investigation of this issue is planned for a later contribution.

**Compactly supported dual windows:** We will now discuss in detail, how to construct compactly supported dual windows. As before, we assume \( L_g \) and \( L_h = |I_h| \) to be finite, where \( I_h \) is the desired support of the dual solution window. To this end, we determine which WR equations are trivially satisfied. Furthermore, we will show that duality of compactly supported Gabor windows on \( \ell^2(\mathbb{Z}) \) is equivalent to duality on \( \mathbb{C}^{L_0} \) and sufficient for duality on \( \mathbb{C}^{L_1} \), where \( L_0 > L_g + L_h \) and \( L_1 \geq \max(L_g, L_h) \) are multiples of \( a, M \).

**Lemma 1.** Let \( I_g, I_h \) be intervals of length \( L_g \) and \( L_h \), without loss of generality centered around 0. For any Gabor system \( \mathcal{G}(g, a, M) \) with \( \text{supp}(g) \subseteq I_g \)
and any $h \in \ell^2(\mathbb{Z})$ with $\text{supp}(h) \subseteq I_h$, all but $a(2\lceil \frac{L_g+L_h}{2M} \rceil - 1)$ of the WR equations are trivially satisfied. Moreover, if and only if $g, h$ are Gabor dual windows on $\ell^2(\mathbb{Z})$ for $a, M$, so are $g_{\text{fin}}[l] = \sum_{k \in \mathbb{Z}} g[l-kL]$ for $l = 0, \ldots, L-1$ and $h_{\text{fin}}$, defined analogous, on $\mathbb{C}^L$, for any $L > L_g + L_h$ with $a, M \mid L$.

**Proof.** To prove the first part, note that $I_g \cap (I_h + nM) = \emptyset$ for $|n| \geq \frac{L_g+L_h}{2M}$. Therefore $\langle h, M_{\text{ma}^{-1}}T_{nMg} \rangle = 0$ for all such $n$ and only $a(2\lceil \frac{L_g+L_h}{2M} \rceil - 1)$ equations remain. However, since $L > L_g + L_h$ and $M \mid L$,

$$\langle h_{\text{fin}}, \exp(2\pi im \cdot /a)T_{nMg_{\text{fin}}} \rangle_{\mathbb{C}^L} = \langle h, M_{\text{ma}^{-1}}T_{nMg} \rangle_{\ell^2}$$

holds for all $m = 0, \ldots, a-1$ and $n = -\lceil \frac{L}{2M} \rceil, \ldots, \lfloor \frac{L}{2M} \rfloor - 1$, showing that $g_{\text{fin}}, h_{\text{fin}}$ are dual on $\mathbb{C}^L$. Note that translation on the left side of the equation is circular. The equation above covers all WR equations of $\mathbb{C}^L$ and, since $L/M - 1 \geq 2\lceil \frac{L_g+L_h}{2M} \rceil$, all WR equations on $\ell^2$ not trivially satisfied, proving equivalence. 

The solutions of the nontrivial equations from the WR equation system \cite{6} form a convex set denoted by $C_{\text{dual}}$.

It is easily checked that, in addition to the equivalence proven above, duality of the windows $g, h \in \ell^2(\mathbb{Z})$ above is sufficient for duality of $g_{\text{fin}}, h_{\text{fin}}$ on $\mathbb{C}^L$, if at least $L \geq L_g, L_h$ and $a, M \mid L$. In that case, we have

$$\delta[n]\delta[m] = \langle h_{\text{fin}}, g_{\text{fin}}[-nM]e^{2\pi im/a} \rangle = \begin{cases} 
\langle h, M_{m/a}(T_{nMg} + T_{nM+Lg}) \rangle \quad &\text{if } nM \leq L/2, \\
\langle h, M_{m/a}(T_{nMg} + T_{nM-Lg}) \rangle \quad &\text{if } nM > L/2.
\end{cases}$$

**Compactly supported duals by truncation:** In \cite{9}, Strohmer proposed a simple algorithm for the computation of compactly supported dual windows, which we will call the truncation method. Note that adding a support constraint to the equation system given by the WR matrix \cite{9} is equivalent to deleting the corresponding columns and only computing the values that are possibly nonzero. The method proposed by Strohmer is based on this truncated WR matrix, considering only the nontrivial equations, see the previous section. The resulting equation system is then solved by computing the pseudoinverse, obtaining the least-squares solution. While the resulting windows satisfy the duality conditions, they are not very smooth and indeed show some discontinuity-like behavior, see also Figure \cite{9}(e,f). One of the goals of this contribution is the improvement of these undesirable effects.

It should be noted that Strohmer’s method is not restricted to support constraints, but can be used to compute directly dual windows $h \in C_{\text{dual}}$ that minimize $\|Rh\|_2$, for some Hermitian positive definite matrix $R$. Naturally, solving this problem in finite time requires a support constraint, as in the previous section.
### Table 1. Some regularization functions

| Function | Effect on the signal | \( \text{prox}_{\mu f}(y) \) |
|----------|----------------------|-----------------|
| \( \mu |x|_1 \) | sparse representation in time | \( \text{soft}_\mu(y) = \text{sgn}(y)(|y| - \mu)_+ \) |
| \( \mu |F x|_1 \) | sparse representation in frequency | \( \mathcal{F}^{-1}\{ \text{soft}_\mu(f y) \} \) |
| \( \mu \| \nabla x \|_2^2 \) | smoothen (time)/concentrate (frequency) | \( \mathcal{F}^{-1}\{ \frac{f \|\nabla x\|_2^2}{1 + \lambda} \} \text{ with } \lambda(\ell) = 2 - 2 \cos(2\pi \ell/L) \) |
| \( \mu \| x \|_0 \) | Concentrate (time and frequency) | \( \mathcal{G}^{-1}\{ \text{soft}_\mu(\mathcal{G} y) \} \), with \( \mathcal{G} \) the STFT |
| \( \mu \| z \|_2^2 \) | spread values more evenly | \( (1 + \mu \mathbf{A}^T \mathbf{A})^{-1}(y + \mu \mathbf{A}^T z) \) |
| \( \mu |\mathbf{A} x - z|_2^2 \) | find \( \mathbf{A} x \) close to \( z \) | \( \text{prox}_{\mu |w|_2^2}^1(y) \), where \( w = \frac{1}{\sqrt{T}}[-L/2,-L/2-1,\ldots,0,\ldots,L/2] \) |
| \( \mu \text{var}(x) \) | Concentrate the signal in time | \( \text{prox}_{\mu |w|_2^2}^1(y) \) |
| \( \mu \text{var}(F x) \) | Concentrate the signal in frequency | \( \text{prox}_{\mu |w|_2^2}^1(y) \) |
| \( \mu \text{var}(|F x|^2) \) | Concentrate the signal in frequency | \( \text{prox}_{\mu |w|_2^2}^1(y) \) |
| \( i_C(x) \) | force \( x \in \mathcal{C} \) | \( P_{\mathcal{C}}(y) \) |

3.1. **Functionals and proximal operators.** In order to tune the solution of a convex optimization problem \([12]\) towards the properties we desire, e.g. time-frequency localization in the case of dual Gabor windows, we have to select *objective functions* \( f_i \), also called *regularizers*, that promote these properties. The purpose of this section is to review some relevant regularizers and their effect on the solution window.

If more than one objective function is selected, their effects can be weighed against each other by tuning the regularization parameters \( \lambda_i \) appropriately. Intuitively, each regularization parameter should be chosen proportional to the importance of the properties it promotes. However, by definition, different regularizers will scale differently, which has to be taken into account when setting \( \lambda_i \). Table 1 lists the regularizers we consider for this work, their effects and proximity operators. Most of the listed regularizers are widely used because they provide important properties. Moreover, their proximity operators are easily computed, therefore allowing their efficient application in proximal splitting algorithms.

Most of the regularizers we use are variations on the widely used \( \ell^1 \)- and \( \ell^2 \)-norm objective functions. In particular, \( \ell^1 \)-norm minimization is considered, whenever sparsity is desired. It forms the convex relaxation of the \( \ell^0 \) minimization problem and its solution is therefore the optimal convex approximation of a truly sparse solution. Under some specific conditions \([26, 8]\), it can be shown to be equivalent or at least close to sparsity optimization.

Therefore, the solution of \( \ell^1 \) minimization can be expected to yield a solution that is at least approximately sparse, i.e. has few large values. This norm jointly used with a linear operator can also provide sparsity in the image of the operator. In time-frequency literature, modulation space norms,
i.e. $\ell^p$ norms on the short-time Fourier coefficients are frequently used to measure joint time-frequency localization, see e.g. [20, 17]. In particular $\|x\|_{S_0} = \|G_{g,1,L}x\|_1$, where $g$ is a Gaussian function, is considered a quality measure for window functions. The reader interested in $\ell^1$ minimization by means of proximal splitting is referred to [15, 13, 8, 38] for more information.

Minimization of the $\ell^2$ norm on the other hand has the tendency to spread values evenly over the whole signal range. The associated objective function is not only convex, but also smooth, admitting gradient descent approaches for minimization. It is traditionally used as a data fidelity term, i.e. the solution is expected to be close, in the $\ell^2$-norm sense, to a given estimate.

Similar to the $\ell^1$-norm, the $\ell^2$-norm can be used jointly with linear operator. In particular, we propose to use the gradient operator to ensure smoothness in the time domain. In our experiments, we observed that gradient $\ell^2$ minimization smoothens out large jumps in the solution. Consequently, we obtain a solution well-concentrated in the frequency domain. Similarly, optimizing the gradient $\ell^2$-norm of the Fourier transform provides localization in time.

Yet another measure for localization is variance, either in magnitude or energy. We use the following definition of the (magnitude) variance of a signal $x \in \mathbb{R}^L$:

$$\text{var}(x) = \frac{1}{\sqrt{L}} \sum_{i=-L/2}^{L/2-1} (i - \bar{x})^2|x_i|$$

with the mean value $\bar{x} = \frac{1}{L} \sum_{i=-L/2}^{L/2-1} i|x_i|$. To compute the energy variance of $x$ instead, consider $\text{var}(f^2)$. It can easily be seen that these variances are nothing but weighted $\ell^1$- or $\ell^2$-norms. Therefore, minimization using proximal splitting is very efficient. Furthermore, we usually desire symmetric analysis and synthesis windows, such that fixing $\bar{x} = 0$ is no restriction and magnitude variance is equal to $\ell^1$-norm with the weight vector $\frac{1}{\sqrt{L}} \left(-(L/2)^2, \ldots, (L/2 - 1)^2\right)$. Similarly, energy variance equals squared $\ell^2$-norm with weight vector $\frac{1}{\sqrt{L}} \left(-L/2, \ldots, L/2 - 1\right)$.

3.2. Solving the problem. After the selection of suitable constraints, e.g. duality and support, and objective functions, we have formulated a well-defined convex optimization problem. Depending on the constraints and regularizers chosen, various methods exist for computing the solution. For the purpose of demonstration and providing a complete working framework, we decided to use and present the parallel proximal algorithm (PPXA, Algorithm [1]). However, our aim is not to provide the most efficient solution to Equation (9), but to present a novel application of convex optimization methods.

Alternative algorithms, e.g. generalized forward backward [30] or simultaneous-direction method of multipliers (SDMM) [33, 13], might prove more efficient. Existing optimization toolboxes like CVX (for MATLAB) can also be used.
If at least one of the regularizers has a well defined gradient, the generalized forward backward algorithm can be used, providing considerable speedup in many cases.

The parallel proximal algorithm minimizes the functions $f_i$ in \([12]\) iteratively, employing the corresponding proximal operator.

**Algorithm 1** Parallel proximal algorithm (PPXA)

1. **Initialize** $\epsilon \in [0, 1]$, $\bar{\gamma} > 0$, $(\omega_i)_{1 \leq i \leq K} \in [0, 1]^K$ with $\sum_{i=1}^{K} \omega_i = 1$, $y_{1,0} \in \mathbb{R}^L$, $y_{K,0} \in \mathbb{R}^L$
2. **Fix** $\lambda \in [\epsilon, 2 - \epsilon[$
3. $x_0 \leftarrow \sum_{i=1}^{K} \omega_i y_{1,0}$
4. for $n = 1, 2, \ldots$ do
   1. for $i = 1, \ldots, K$ do
      1. $p_{i,n} \leftarrow \text{prox}_{\bar{\gamma} f_i/\omega_i} (y_{i,n})$
   2. end for
   3. $p_n \leftarrow \sum_{i=1}^{K} \omega_i p_{i,n}$
   4. for $i = 1, \ldots, K$ do
      1. $y_{i,n+1} \leftarrow y_{i,n} + \lambda (2p_n - x_n - p_{i,n})$
   2. end for
   3. $x_{n+1} \leftarrow x_n + \lambda (p_n - x_n)$
   4. end for

3.3. **Tight windows.** Instead of computing the dual window to a previously selected analysis window, we now aim to compute a tight Gabor frame with good properties, given the parameters $a, M$. A system $\mathcal{G}(h, a, M)$ is a tight frame, if it is a dual frame to itself. This is no longer a convex problem, since the set of tight windows for $a, M$ is not a convex set. Indeed, it is the solution set associated to the nonlinear equation system

$$\frac{M}{a} \left< h, h[\cdot - nM]e^{2\pi im/a} \right> = \delta[n] \delta[m],$$

obtained by setting $g = h$ in the WR equations \([6]\).

In this situation, the framework proposed before is no longer guaranteed to converge to an optimal solution or, in fact, to converge at all. We have observed experimentally, see also Experiment 5 in Section \([4]\), that good solutions are obtained nonetheless, with the algorithm seemingly converging to a local minimum. Therefore, we propose our method as a heuristic scheme to find optimized, yet possibly not optimal, tight Gabor frames. Before that can be done, 2 issues remain to be solved.

Firstly, since the problem is not convex, the starting value is very important. Good results have been obtained choosing a window function that is already close to what we aim for, i.e. provides good frame bounds for the Gabor parameters $a, M$ and shows the properties we wish to promote in the tight window.
Secondly, and more importantly, we need to define the projection operator to the set of tight windows $C_{\text{tight}}$:
\[
P_{C_{\text{tight}}}(y) = \arg \min_{x \in C_{\text{tight}}} \|x - y\|_2.
\]

Luckily, this projection can be computed via the formula
\[
P_{C_{\text{tight}}}(y) = S_{y,a,M}^{-1/2} y,
\]
where $S_{y,a,M}$ is the frame operator with respect to $G(y,a,M)$, cf. [23]. The result of $S_{y,a,M}^{-1/2} y$ is called the canonical tight window and efficient algorithms for its computation are freely available, e.g., in the LTFAT toolbox [35], see also [24].

4. Numerical experiments

Simulations were performed using the LTFAT [35] and the UNLocBoX Matlab toolbox. A reproducible research addendum is available in \url{http://unlocbox.sourceforge.net/rr/gdwuco}. We refer to this address as the webpage.

Experiment 1 - Effect of the objective function: We wish to demonstrate that different optimization criteria lead to widely different solution windows, with each criterion promoting a certain shape or behavior in the optimized dual. To emphasize this effect, we apply the regularizer functions introduced in the previous section to the same tight starting system. More precisely, our setup considers $G(g, 15, 120)$, i.e., a system with redundancy 8, where $g$ is a Itersine window [40] of length $L_g = 120$ samples, see Figure 2(a,b). The (continuous) Itersine window is defined as
\[
g(t) = \begin{cases} 
\sin \left(0.5\pi \cos(\pi t)^2\right) & \text{if } t \in [-1/2, 1/2], \\
0 & \text{else}.
\end{cases}
\]

This window yields a tight frame for half-overlap and shows a better frequency decay and concentration, the latter in terms of main lobe width, than the more widely used cosine window.

In this experiment, we fixed $L$ to 240, therefore forcing $L_h \leq 240$, but not adding further support constraints. This setup, in particular its high redundancy, allows us to shape the dual windows rather freely for different objective functions, therefore producing characteristic examples. Since the Itersine window forms a tight frame for the chosen parameters we do not show the canonical dual. Moreover, the $\ell^2$-norm minimizing dual window coincides with the canonical dual and consequently with the original window up to normalization, cf. Section 2.4. All presented dual windows are assumed to perform perfect reconstruction. More precisely, using [22 Eq. (60)], the maximum relative reconstruction error can be shown to be of the order of precision of the machine, i.e., around $10^{-15}$. 
Due to numerical issues of proximal splitting methods, the timestep $\tilde{\gamma}$ (See Algorithm 1), has influence on the solution. In particular, there is a tradeoff between precision and convergence speed. Our contribution is intended as conceptual, rather than algorithmic work therefore we do not discuss those issues in detail.

Concentration in time and in frequency is a very important property for Gabor synthesis. Indeed when modifying the coefficients in the Gabor domain, the time-frequency spread of this operation will strongly depend on the time-frequency concentration of the synthesis window. To improve concentration in one domain, we will smooth the function in the other. For instance, the smoothest function in time is also the most concentrated in frequency. It can be computed by minimizing the function $\|\nabla x\|_2^2$. In a same manner, the smoothest function in frequency (and most concentrated in time) is the solution of the minimizing the function $\|\nabla Fx\|_2^2$. Result are presented in Figure 2 (c,d) and (e,f).

Simultaneous time-frequency concentration can be obtained using the $S_0$ norm as regularizer. However, minimizing the norm is computationally less
efficient than optimizing smoothness in time and frequency as described
above. The result is presented in Figure 2 (g,h).

As convex relaxation of the \( \ell^0 \) problem, we also consider the \( \ell^1 \)-norm
minimal dual window. It should be noted however, that minimizing the \( \ell^1 \)-
norm does not guarantee a solution with small support. Even if the solution
has few nonzero values, there is no theoretical guarantee that those zeros
are concentrated, i.e. contained in a small interval. The example shown
in Figure 2 (i,j) is produced by considering pure \( \ell^1 \) minimization. In order
to analyze the result, we begin by setting to zero all coefficients smaller
than the tolerance of our algorithm (-100 dB). The resulting number
of nonzero coefficients is 32 which is very close to the number of WR equations
\( \frac{L}{M} a = 30 \). However the length of the support is 82 which is far greater than
minimum possible compact support (\( L_{\text{min}} = \frac{L}{M} a = 30 \)). This problem stems
from the fact that the \( \ell^1 \) minimization provides no localization.

However, this localization can be obtained with other functionals. In
particular optimizing the variance can be of great interest. In Figure 2
(m,n) and (o,p), the variance in time and in frequency is optimized. For
more information about the variance and its optimization, see Experiment
2.

![Figure 3](image)

**Figure 3.** Result of optimization function \( \|x - gsh\|_2^2 \): respectively the windows \( gsh \) and the result of optimization. : (a)(b) A sine. (c)(d) An impulse (Kronecker delta).

Instead of optimizing a function that promotes certain abstract features,
free design can be performed using the objective function \( \|x - gsh\|_2^2 \), where
\( gsh \) is used as a model for the dual window. The algorithm then returns the
dual window closest to \( g_{sh} \) in the \( \ell^2 \)-norm sense. For \( x \) to have similar shape as \( g_{sh} \), correct normalization of \( g_{sh} \) is important. We have observed that choosing \( g_{sh} \) such that \( \| g_{sh} \|_2 = \| \gamma \|_2 \), where \( \gamma \) is the canonical dual, yields good results. Therefore, we chose that approach for the examples presented in Figure 3(a,b) and (c,d).

**Experiment 2 - Optimizing time-frequency concentration I:** In this experiment, we aim to model time-frequency concentration through the combination of 2 objective functions. Since concentration in time is closely related to smoothness in frequency and vice-versa, we select the two smoothing functionals: \( \| \nabla x \|_2^2 \) and \( \| \nabla \mathcal{F} x \|_2^2 \) and balance them to form the optimization problem:

\[
\arg \min_{x \in \mathcal{C}_{dual}} \| \nabla \mathcal{F} x \|_2^2 + 5 \| \nabla x \|_2^2.
\]

We apply this optimization scheme in the setup proposed in Experiment 1, to obtain a dual window with good concentration and smoothness in both time and frequency. The solution we obtain is shown in Figure 4(a,b). In comparison to the canonical dual, Figure 2(a,b), the solution window is more concentrated in time, with slightly broader spread in frequency but considerably improved decay. In fact, it closely resembles a Gaussian function. To illustrate the improved time-frequency concentration and its use in Gabor synthesis, Figure 5 shows the differences of synthesizing a group of nonzero coefficients with the 2 alternative windows. Note that these coefficients are not the result of an analysis, but are freely chosen in the time-frequency plane. Optimally, to be able to perform precise time-frequency processing using a Gabor analysis/synthesis scheme, the spectrogram of the synthesized coefficients in this example should have as little spread as possible. Note though, that the actual spread shown combines the leaking caused by the synthesis window with the smearing effect caused by the analysis. To minimize the latter, we have decided to used a well-concentrated window for analysis. Therefore, all spectrograms show a full STFT, computed using a Nuttall [28] window of length 60. The continuous Nuttall window is defined as

\[
g(t) = \begin{cases} 
\sum_{k=0}^{3} c_k \cos(2\pi k t) & \text{if } t \in [-1/2, 1/2], \\
0 & \text{else},
\end{cases}
\]

where \( c_0 = 0.355768 \), \( c_1 = 0.487396 \), \( c_2 = 0.144232 \) and \( c_3 = 0.012604 \). See also Figure 9(a,b) for a sampled Nuttall window.

We observe that, indeed, the window computed with our optimization scheme shows improved concentration, in particular in time.

**Experiment 3 - Optimizing time-frequency concentration II:**

We present the construction of time-frequency concentrated dual Gabor windows for a system of fixed length \( L = 500 \). Our setup considers
Figure 4. Result of an optimization problem respectively in time and frequency domain (a)(b) $\|\nabla F_{x}\|_{2}^{2} + 5\|\nabla x\|_{2}^{2}$.

Figure 5. Result of different synthesis of Gabor coefficients. Top: original Gabor coefficients. Bottom: Spectrogram of the synthesis using the canonical dual window (left) and spectrogram of the analysis coefficient using the alternate dual window from Figure 4 (right).

$G(g, 25, 125)$, i.e. a system with redundancy 5, where $g$ is a rectangular window of length $L_g = 125$ samples. This construction allows the comparison of alternative time-frequency localization measures and their effect on the solution window. The system forms a tight frame, hence the canonical dual window is given by the same, badly localized window, up to normalization. Consequently, this setup is an ideal sandbox for experimenting with various localization measures.
We apply 3 different methods for measuring joint time-frequency localization:

- As in Experiment 2, we combine the $\ell^2$-norms of the gradient of $x$ and of $\mathcal{F}x$, to optimize the smoothness in the time and the frequency domain.
- Variance is associated with the Heisenberg uncertainty principle, a classical measure for time-frequency localization. While we cannot directly optimize the Heisenberg uncertainty (the product $\text{var}(x) \text{var}(\mathcal{F}x)$), the sum of variances provides a very similar measure. We consider both the magnitude variance, as in Heisenberg’s uncertainty, as well as energy variance.
- The $S_0$-norm of $x$, or equivalently $\ell^1$-norm of its STFT, is widely used in Gabor analysis to measure the spread of $x$ in the time-frequency plane. Since truly sparse short-time Fourier transforms do not exist, it promotes functions that are concentrated around few points in the time-frequency plane. Since concentrated windows have small spread, they also have a small $S_0$-norm. However, small spread is not a guarantee for concentration. To further improve the concentration, we also consider to weight the $S_0$-norm as with the following:

$$W(f, t) = \sqrt{x^2(t) + y^2(f)},$$

with $x(i) = y(i)$ being the $i$-th entry of the vector

$$\frac{1}{\sqrt{L}} (-L/2, -L/2 - 1, ..., 0, ..., L/2).$$

To show the effect of the gradient and variance measures when applied in a single domain, Figure 6 shows the results obtained for optimizing the concentration in time only. Since the analysis system forms a tight frame, the canonical dual is a rectangular windows of length 125. In this example, magnitude variance minimization provides the strongest concentration in terms of support, albeit at the cost of the clear, slim peak obtained from the very similar energy variance and gradient optimization. The results for a similar experiment, optimizing frequency concentration can be found on the webpage.

In fact, we are more interested in simultaneous time-frequency concentration. Therefore, we performed another experiment comparing the time-frequency localization measures proposed above. Figures 7 and 8 present the results of the experiment. The advantage of the gradient and variance approaches is the simplicity of tuning the tradeoff between time and frequency concentration, which involves only the appropriate choice of regularization parameter. On the other hand, only $S_0$-norm based optimization schemes truly measure concentration on the whole time frequency plane, instead of its one dimensional projection on the time, respectively frequency domain.
The results show that pure $S_0$-norm optimization is insufficient to obtain good concentration. Weighted $S_0$-norm and magnitude variance optimization on the other hand, provide very good results. The magnitude variance shows slightly better concentration in time, while the weighted $S_0$-norm solution is superior to all other measures in terms of time-frequency decay. As observed for the previous experiment, energy variance and gradient approaches produce very similar results, providing slightly worse time-frequency concentration compared to magnitude variance and weighted $S_0$-norm. While the experiment suggests weighted $S_0$-norm and magnitude variance as the most promising approaches, this experiment should not be considered an exhaustive study of the various time-frequency localization measures. More experiments are required to study the different approaches in detail.
Figure 8. Spectrograms of the resulting dual windows. Note the good decay of the weighted $S_0$-norm

**Experiment 4 - Smooth compactly supported duals:** Now, we want to construct a dual Gabor window with short support in a nonpainless setup, i.e. for a system with few frequency channels. This has previously been attempted in [37]. However, the solutions obtained by the truncation method proposed therein, cf. Section 3, are often badly localized in frequency due to the tendency of the truncation method to yield nonsmooth solutions, i.e. solutions with “jumps” or discontinuity-like behavior in time.

Here, we present the construction of a dual window that combines short support with reasonable smoothness and therefore frequency concentration. We start from a Gabor system $G(g, 30, 60)$ with redundancy 2. The analysis window $g$ is chosen as a Nuttall window of length $L_g = 120$ samples.

We desire a dual window $h$ with the same support as $g$, i.e. $L_h = L_g = 120$. Therefore, the set of admissible windows is given by $C_{\text{dual}} \cap C_{\text{supp}}$, where $C_{\text{supp}} = \{x \in \mathbb{R}^L : \text{supp}(x) \subseteq [-60, 59]\}$. Furthermore, we aim to
achieve localization and smoothness by selecting the regularization functions  
\[ f_1 = \| \cdot \|_1, f_2 = \| F(\cdot) \|_1, f_3 = \| \nabla (\cdot) \|_2^2 \text{ and } f_4 = \| \nabla F(\cdot) \|_2^2. \]

The results in Figure 9(c,d) show the optimal dual window with regards to the regularization parameters \( \lambda_1 = \lambda_2 = 0.001 \) and \( \lambda_3 = \lambda_4 = 1 \). Those values were chosen experimentally, balancing the effect of the regularization functions described in Table 1. As reference, we included the least-squares solution provided by the truncation method, see Figure 9(e,f).

Minimizing the selected regularization functions improves upon the desired features, in particular smoothness and localization by means of minimizing \( f_3 \) and \( f_4 \). The addition of \( f_1 \) and \( f_2 \) is supposed to suppress the tendency of the solution to have \( M \)-like shape, i.e. multiple peaks. This is unwanted as it leads to windows with ambiguous temporal or frequency position. Heuristically, minimizing the \( l^1 \)-norm pushes all big coefficients to similar values, therefore achieving the suppression of multiple significant peaks.

The solution provided is assumed to perform perfect reconstruction on any signal with admissible length greater or equal to \( L \). As in the previous experiments, the maximum relative reconstruction error (4.5e^{-14}) is of the order of the machine precision.

To guarantee the canonical dual window to be supported on \( L_h = L_g \), we would be required to increase the number of frequency channels to \( M \geq 120 \), putting us in the painless case. Therefore the redundancy is increased twofold, which is an unwanted side effect. Alternatively, in this setting, we could decide to keep the parameters \( a = 30, M = 60 \) fixed, but decrease the window size to \( L_g \leq 60 \). However, this construction provides a system with a more than 8 times larger frame bound ratio. Consequently, the resulting canonical dual window \( h \), shown in Figure 10, shows bad frequency behavior and an undesirable, \( M \)-like shape in time. In contrast, the method proposed in this manuscript allows the use of nicely shaped, compactly supported dual Gabor windows at low redundancies, without the strong restrictions of the painless case.

**Experiment 5 - Improvement of classical tight frame constructions:** Especially for frames with small redundancy, it has been observed that a tradeoff between localization and smoothness in time-frequency exists between the analysis and dual windows. From the perspective of tight frames, this has the consequence that the available tight prototypes show, in comparison, suboptimal time-frequency concentration. In this final experiment, we will propose 2 approaches for the improvement of a popular tight frame construction. In particular, the Itersine window, already introduced in Experiment 1, forms a tight frame in a half-overlap situation when sufficient frequency channels are used. The Itersine window has been proposed as an improvement of the rectangular and cosine windows, both widely used for tight frames with redundancy 2, providing better smoothness and time-frequency concentration than either.
Figure 9. Experiments. (a) Analysis window in time. (b) Analysis window in frequency. (c) Synthesis window in time. (d) Synthesis window in frequency. (e) Truncation method in time. (f) Truncation method in frequency.

Figure 10. Half-overlap painless case construction ($G(g, 30, 60), L_g = 60$): Canonical dual window in time (a) and in frequency (b).
Consequently, the starting point of our experiment is a Gabor system of the form $G(g, 30, 60)$, i.e. a system with redundancy 2, with an Itersine window $g$ of length $L_g = 60$. Our aim is to construct both an alternative tight frame, as well as a pair of dual windows, such that all 3 obtained prototypes show improved time-frequency concentration over the Itersine window $g$.

First, we construct an improved tight window. To that end, we use the heuristic method for tight window design proposed in Section 3. Additionally, to gain some design freedom, we allow the tight window $g_t$ to have support length $L_{g_t} \leq 360$. We then attempt to solve the optimization problem

$$\arg \min_{x \in C_{\text{tight}} \cap C_{\text{supp}}} \lambda_1 \|\nabla F x\|_2^2 + \lambda_2 \|\nabla x\|_2^2$$

The gradient functionals have been selected to promote a window that is smooth and well-localized in both domains. We initialized the algorithm with the Itersine window $g$. Despite the proposed optimization problem not being convex, we obtain a tight window with roughly the same concentration in time, but improved decay and sidelobe attenuation in frequency, see the comparison in Figure 11. The tight window shown was obtained for the regularization parameters $\lambda_1 = 1$, $\lambda_2 = 5$.

![Figure 11](image)

**Figure 11.** $G(g, 30, 60)$, $L_g = 60$, compactly supported tight windows: $L_g = 360$.

In the second part of the experiment, we construct a pair of compactly supported dual windows $g_0, h$ such that $G(g_0, 30, 60)$ and $G(h, 30, 60)$ form dual frames. Again, we allow some freedom for the support, choosing for $g_0$ a Nuttall window of length $L_{g_0} = 120$. This window is slightly broader than the Itersine window $g$ in time, but very good decay and localization in frequency, see Figure 12. For the dual window, we consider $L_h \leq 360$ as in the tight case and the convex optimization problem

$$\arg \min_{x \in C_{\text{dual}} \cap C_{\text{supp}}} \mu_1 \|\nabla F x\|_2^2 + \mu_2 \|\nabla x\|_2^2$$

The dual window we obtain (using the regularization parameters $\mu_1 = 0.1$, $\mu_2 = 1$), shows roughly the same time localization as the Nuttall window
or the tight window obtained before and is quite similar to the latter in frequency. The result is shown in Figure 12.

![Figure 12. Construction of a pair of dual FIR windows both beating the 'Itersine' \((G(y,30,60), \text{Itersine } L_q = 60, \text{Nuttall } L_{g_0} = 120, \text{Dual window } L_h = 120)\): windows in time and frequency](image)

In terms of joint time-frequency localization, we strongly feel that all 3 prototypes constructed in this experiment show a considerable improvement upon the Itersine window and therefore also over the cosine window.

5. Conclusion

In this contribution, we have proposed an algorithm for the design of Gabor dual windows relying on standard techniques from convex optimization. The proposed method extends previous approaches and provides a framework for tuning various properties of the synthesis system. It can also be used for the construction of tight windows, albeit only in a heuristic fashion without mathematical convergence guarantees.

Our method can be applied in various situations to construct dual frames with properties more important for application than minimal \(\ell^2\)-norm.

Therefore, we introduced several useful regularizer functionals and discussed their effect on the shape and behavior of the dual window. In an extensive experiment section, we demonstrated these effects and their application to design pairs of dual windows with good properties. In particular, we use smoothness regularizers in time and frequency to construct a dual window with improved time-frequency concentration compared to the canonical dual. We compare this measure of time-frequency concentration to alternate approaches based on variance or \(S_0\)-norm, both of which perform similar or even better than the smoothness prior in our experiments.

Through a combination of smoothness and \(\ell^1\) priors, we show that our method can improve previous results by Strohmer [37] on the construction of dual windows with short support.

Finally, a classical setup for tight Gabor frames is examined and improvements are proposed based on our framework for dual and tight windows, respectively.
Future work will further be concerned with applying the findings herein to frames with a different structure, e.g. nonstationary Gabor frames [4].

**Acknowledgement**

This work was supported by the Austrian Science Fund (FWF) START-project FLAME (“Frames and Linear Operators for Acoustical Modeling and Parameter Estimation”; Y 551-N13).

**References**

[1] A. Adler, V. Emiya, M. Jafari, M. Elad, R. Gribonval, and M. Plumbley. A constrained matching pursuit approach to audio declipping. In IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pages 329–332. IEEE, 2011.

[2] A. Adler, V. Emiya, M. Jafari, M. Elad, R. Gribonval, and M. Plumbley. Audio inpainting. IEEE Transactions on Audio, Speech, and Language Processing, 20(3):922–932, 2012.

[3] P. Balazs. Basic definition and properties of Bessel multipliers. *Journal of Mathematical Analysis and Applications*, 325(1):571–585, January 2007.

[4] P. Balazs, M. Dörfler, N. Holighaus, F. Jaillet, and G. Velasco. Theory, implementation and applications of nonstationary Gabor frames. *Journal of Computational and Applied Mathematics*, 236(6):1481–1496, 2011.

[5] P. Balazs, M. Dörfler, M. Kowalski, and B. Torrésani. Adapted and adaptive linear time-frequency representations: a synthesis point of view. *IEEE Signal Processing Magazine (special issue: Time-Frequency Analysis and Applications)*, to appear:--, 2013.

[6] P. Balazs, H. G. Feichtinger, M. Hampejs, and G. Kracher. Double preconditioning for Gabor frames. *IEEE T. Signal. Proces.*, 54(12):4597–4610, December 2006.

[7] H. Böckl and F. Hlawatsch. Oversampled cosine modulated filter banks with perfect reconstruction. IEEE Trans. Circuits and Systems II, 45(8):1057–1071, Aug. 1998.

[8] S. S. Chen, D. L. Donoho, and M. A. Saunders. Atomic Decomposition by Basis Pursuit. *SIAM Journal on Scientific Computing*, 20(1):33–61, Jan. 1998.

[9] O. Christensen. *An Introduction to Frames and Riesz Bases*. Birkhäuser, 2003.

[10] O. Christensen, H. Kim, and R. Y. Kim. Gabor windows supported on $[-1, 1]$ and compactly supported dual windows. *Appl. Comput. Harmon. Anal.*, 28(1):89 – 103, January 2010.

[11] O. Christensen, H. Kim, and R. Y. Kim. Gabor windows supported on $[-1, 1]$ and dual windows with small support. *Adv. Comput. Math.*, 36(4):525–545, 2012.

[12] P. Combettes and J. Pesquet. A Douglas–Rachford splitting approach to nonsmooth convex variational signal recovery. *IEEE Journal of Selected Topics in Signal Processing*, 1(4):564–574, 2007.

[13] P. Combettes and J. Pesquet. Proximal splitting methods in signal processing. *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, pages 185–212, 2011.

[14] P. Combettes and V. Wajs. Signal recovery by proximal forward-backward splitting. *Multiscale Modeling & Simulation*, 4(4):1168–1200, 2005.

[15] P. L. Combettes and J.-C. Pesquet. Proximal thresholding algorithm for minimization over orthonormal bases. *SIAM Journal on Optimization*, 18(4):1351–1376, 2007.

[16] I. Daubechies, H. J. Landau, and Z. Landau. Gabor Time-Frequency Lattices and the Wexler-Raz Identity. *Journal of Fourier Analysis and Applications*, 1(4):437–478, Nov. 1994.

[17] H. G. Feichtinger. On a new Segal algebra. *Monatsh. Math.*, 92:269–289, 1981.
[18] H. G. Feichtinger and K. Nowak. A first survey of Gabor multipliers, chapter 5, pages 99–128, 2003.
[19] H. G. Feichtinger and T. Strohmer, editors. Gabor Analysis and Algorithms. Birkhäuser, Boston, 1998.
[20] K. Gröchenig. Foundations of Time-Frequency Analysis. Appl. Numer. Harmon. Anal. Birkhäuser Boston, Boston, MA, 2001.
[21] W. W. Hager and H. Zhang. A survey of nonlinear conjugate gradient methods. Pacific journal of Optimization, 2006.
[22] A. Janssen and P. L. Søndergaard. Iterative algorithms to approximate canonical Gabor windows: Computational aspects. J. Fourier Anal. Appl., 13(2):211–241, 2007.
[23] A. Janssen and T. Strohmer. Characterization and computation of canonical tight windows for gabor frames. Journal of Fourier Analysis and Applications, 8(1):1–28, 2002.
[24] A. J. E. M. Janssen and T. Strohmer. Characterization and computation of canonical tight windows for Gabor frames. J. Fourier Anal. Appl., 8(1):1–28, January 2002.
[25] M. Kowalski, K. Siedenburg, and M. Dörfler. Social sparsity! neighborhood systems enrich structured shrinkage operators. Signal Processing, IEEE Transactions on, 61(10):2498–2511, 2013.
[26] S. Li, Y. Liu, and T. Mi. Sparse Dual Frames and Dual Gabor Functions of Minimal Time and Frequency Supports. Journal of Fourier Analysis and Applications, pages 1–29, 2012.
[27] B. Martinet. Détermination approchée d’un point fixe d’une application pseudo-contractante. cas de l’application prox. CR Acad. Sci. Paris Ser. AB, 274:A163–A165, 1972.
[28] A. Nuttall. Some windows with very good sidelobe behavior. IEEE Trans. Acoust. Speech Signal Process., 29(1):84–91, 1981.
[29] N. Perraudin, N. Holighaus, P. Soendergaard, and P. Balazs. Gabor dual windows using convex optimization. In Proceedings of the 10th International Conference on Sampling theory and Applications (SAMPTA 2013), 2013.
[30] H. Raguet, J. Fadili, and G. Peyre. Generalized Forward-Backward Splitting. arXiv.org, Aug. 2011.
[31] R. Rockafellar. Monotone operators and the proximal point algorithm. SIAM Journal on Control and Optimization, 14(5):877–898, 1976.
[32] S. Samadi, M. O. Ahmad, and M. N. S. Swamy. Characterization of nonuniform perfect-reconstruction filterbanks using unit-step signal. IEEE Transactions on Signal Processing, 52(9):2490–2499, 2004.
[33] S. Setzer, G. Steidl, and T. Teuber. Deblurring poissonian images by split bregman techniques. Journal of Visual Communication and Image Representation, 21(3):193 – 199, 2010.
[34] P. L. Søndergaard. Gabor frames by Sampling and Periodization. Adv. Comput. Math., 27(4):355 –373, 2007.
[35] P. L. Søndergaard, B. Torrésani, and P. Balazs. The Linear Time Frequency Analysis Toolbox. International Journal of Wavelets, Multiresolution Analysis and Information Processing, 10(4), 2012.
[36] D. T. Stoeva and P. Balazs. Invertibility of multipliers. Applied and Computational Harmonic Analysis, 33(2):292–299, 2012.
[37] T. Strohmer. Numerical algorithms for discrete Gabor expansions. In Feichtinger and Strohmer [19], chapter 8, pages 267–294.
[38] E. van den Berg and M. P. Friedlander. Probing the Pareto Frontier for Basis Pursuit Solutions. SIAM Journal on Scientific Computing, 31(2):890–912, Jan. 2009.
[39] M. Vetterli. Filter banks allowing perfect reconstruction. Signal Process., 10:219–244, 1986.
[40] E. Wesfreid and M. Wickerhauser. Adapted local trigonometric transforms and speech processing. *IEEE Trans. Signal Process.*, 41(12):3596–3600, 1993.

[41] J. Wexler and S. Raz. Discrete Gabor expansions. *Signal Processing*, 21(3):207–220, Nov. 1990.

Nathanaël Perraudin is with the Signal Processing Laboratory 2, École polytechnique fédérale de Lausanne, CH-1015 Lausanne, Switzerland

E-mail address: nathanael.perraudin@epfl.ch

Nicki Holighaus, Peter L. Søndergaard and Peter Balazs are with the Acoustics Research Institute, Austrian Academy of Sciences, Wohllebengasse 12–14, 1040 Vienna, Austria