Frames of eigenspaces and localization of signal components

Monika Dörfler(1) and José Luis Romero(2)
Faculty of Mathematics, University of Vienna,
Nordbergstrasse 15,A-1090 Wien, Austria
Email: (1) Monika.Doerfler@univie.ac.at, (2) Jose.Luis.Romero@univie.ac.at.

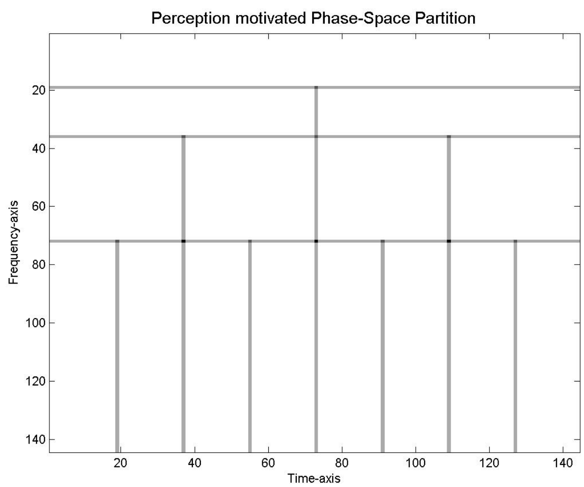

Fig. 1. Time-frequency partition with varying time-frequency bands

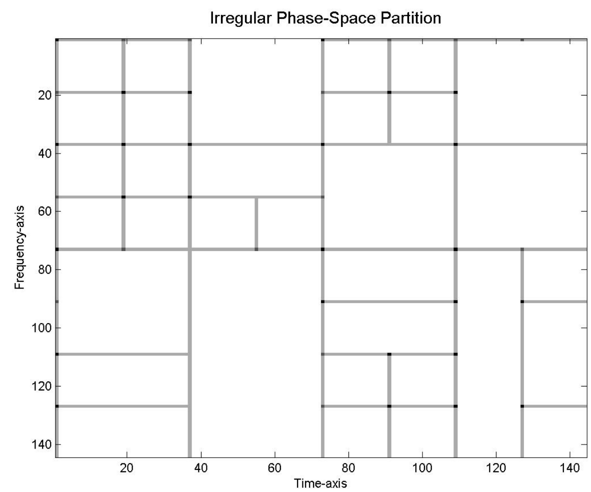

Fig. 2. Fully irregular time-frequency partition

***Abstract*—We present a construction of frames adapted to a given time-frequency cover and study certain computational aspects of it. These frames are based on a family of orthogonal projections that can be used to localize signals in the time-frequency plane. We compare the effect of the corresponding orthogonal projections to the traditional time-frequency masking.**

I. Introduction

When representing a signal in a time-frequency dictionary, the atoms are usually chosen as time-frequency shifts of a window along a lattice (Gabor frame). The choice of the lattice together with the characteristics (shape, width) of the basic window or family of windows determines the ability of the representation to localize certain signal components and, furthermore, the possibility to separate them. Various approaches have been taken to circumvent the restrictions possibly imposed by a rigid application of lattice structure (reassignment, adaptive frames) [1], [2], [6], [3], [17], giving time-frequency partitions consisting of frequency (resp. time) strips of varying widths (see figure 1)

In [10] we have presented a construction of frames whose spectrogram follows a prescribed time-frequency pattern. This pattern may be quite irregular and in particular does not need to be a Cartesian product of a time and a frequency partition (see figure 2).

This construction is achieved by selecting from each tile of the cover an orthonormal set of functions that maximizes its joint spectrogram within the tile. These functions are eigenfunctions of time-frequency localization operators (see below), whose concentration is no more restricted to be located at lattice points. By definition, the eigenfunctions corresponding to high eigenvalues of the localization operators, are maximally localized within a (weighted) subfamily of the time-frequency shifted atoms; thus, they provide potentially better localization in a certain time-frequency region than the time-frequency atoms themselves.

Since the frames introduced in [10] are constructed by choosing a finite number of eigenfunctions from each localization operator corresponding to a partition of the time-frequency plane, they produce a resolution of the identity by orthogonal projections. This means that replacing the usual time-frequency masking operators by certain orthogonal projections does not lead to loss of information, provided that the projection is chosen judiciously.

In this article we consider certain computational aspects of the construction of frames adapted to time-frequency covers and compare the effect of the corresponding orthogonal projections to the traditional time-frequency masking.

II. Time-frequency localization

A. Localization operators

The short-time Fourier transform (STFT) of a distribution $f \in \mathcal{S}'(\mathbb{R}^d)$ is a function defined on $\mathbb{R}^d \times \mathbb{R}^d$ defined, by means of an adequate smooth and fast-decaying window function $\varphi \in \mathcal{S}(\mathbb{R}^d)$, as

$$\mathcal{V}_\varphi f(z) = \int_{\mathbb{R}^d} f(t)\overline{\varphi(t-x)}e^{-2\pi i \xi t}dt, \quad z=(x,\xi) \in \mathbb{R}^d \times \mathbb{R}^d.$$

The number $\mathcal{V}_\varphi f(x,\xi)$ represents the influence of the frequency ξ near x. The distribution f can be re-synthesized from its time-frequency content by,

$$f(t) = \int_{\mathbb{R}^d \times \mathbb{R}^d} \mathcal{V}_\varphi f(x,\xi)\varphi(t-x)e^{2\pi i \xi t}dxd\xi. \tag{1}$$

Given a compact set $\Omega \subseteq \mathbb{R}^{2d}$ in the time-frequency plane, the *time-frequency localization operator* L_Ω is defined by masking the coefficients in (1), cf. [5], i.e.

$$L_\Omega f(t) = \int_\Omega \mathcal{V}_\varphi f(x,\xi)\varphi(t-x)e^{2\pi i \xi t}dxd\xi. \tag{2}$$

L_Ω is self-adjoint and trace-class, so we can consider its spectral decomposition

$$L_\Omega f = \sum_{k=1}^{\infty} \lambda_k^\Omega \langle f, \phi_k^\Omega \rangle \phi_k^\Omega.$$

The first eigenfunction, ϕ_1^Ω, is optimally concentrated inside Ω in the following sense,

$$\int_\Omega \left|\mathcal{V}_\varphi \phi_1^\Omega(z)\right|^2 dz = \max_{\|f\|_2=1} \int_\Omega \left|\mathcal{V}_\varphi f(z)\right|^2 dz.$$

More generally, the first N eigenfunctions of H_Ω form an orthonormal set in $L^2(\mathbb{R}^d)$ that maximizes the quantity

$$\sum_{j=1}^{N} \int_\Omega |\mathcal{V}_\varphi \phi_j^\Omega(z)|^2 dz,$$

among all orthonormal sets of N functions in $L^2(\mathbb{R}^d)$. In this sense, their time-frequency profile is optimally adapted to Ω.

B. Time-Frequency areas of interest

The shape of the time-frequency areas one may be interested to localize in, will usually depend on the application and the characteristics of the underlying class of signals. Typically, one may consider rectangles of different eccentricities in order to be able to focus on signal components showing a more transient or more harmonic characteristic. Examples are depicted in Figure 3. In some applications, one may be interested in more exotic shapes, such as triangular, cf. Figure 4, for example to account for the spectral roll-off in instrumental sounds, or chirped components, cf. Figure 5, which are also omnipresent in both speech and music signals.

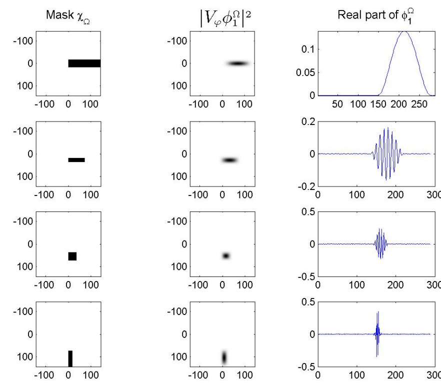

Fig. 3. Four different rectangular masks in time-frequency domain and the first eigenfunctions of the corresponding localization operators. Middle plots show the absolute value squared of the STFT and right plots show the real part.

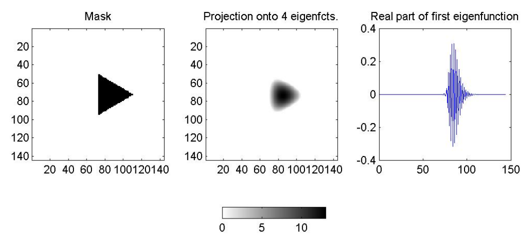

Fig. 4. Triangular-shaped mask, absolute value squared of the STFT of projection of random noise onto most localized resulting eigenfunctions, real part of most concentrated eigenfunction.

III. Frames of eigenfunctions

We now present the main result on the construction of frames adapted to a cover and then explore certain computational aspects of it. The proof of the following theorem can be found in [10], together with an extended discussion on its quantitative aspects (see also [8], [15], [9], [16]).

Theorem 1: Let $\{\Omega_\gamma : \gamma \in \Gamma\}$ be a cover of $\mathbb{R}^{2d}$ such that

$$B_r(\gamma) \subseteq \Omega_\gamma \subseteq B_R(\gamma), \quad \text{with } \Gamma \text{ a lattice and } R \geq r > 0.$$

Then, there exists a constant $C > 0$ such that for every choice of N_γ, $C\,|\Omega_\gamma| \leq N_\gamma \leq N < \infty$, the family of functions

$$\left\{\phi_k^{\Omega_\gamma} : \gamma \in \Gamma, 1 \leq k \leq N_\gamma\right\}$$

is a frame of $L^2(\mathbb{R}^d)$.

A. Computing the eigenfunctions in each tile

In practice we work with a discrete realization of L_Ω given by

$$H_{\mathbf{m},\mathbf{\Lambda}} f = \sum_{\lambda \in \Lambda} m(\lambda) \langle f, \pi(\lambda) g \rangle \pi(\lambda) g, \tag{3}$$

where

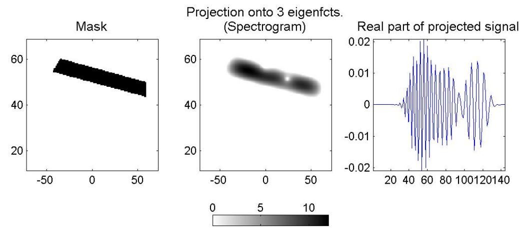

Fig. 5. Chirp-shaped mask, absolute value squared of the STFT of projection of random noise onto most localized resulting eigenfunctions, real part of the projection.

- $\Lambda \subseteq \mathbb{R}^{2d}$ is a lattice,
- $\{\pi(\lambda)g = e^{2\pi i\lambda_2 \cdot} g(\cdot - \lambda_1) : \lambda = (\lambda_1, \lambda_2) \in \Lambda\}$ is a tight Gabor frame of $L^2(\mathbb{R}^d)$.
- $\mathbf{m} = (\mathbf{m}_\lambda)_{\lambda\in\mathbf{\Lambda}}$ is a bounded sequence of complex numbers.

The operator $H_{\mathbf{m},\mathbf{\Lambda}}$ is called a Gabor multiplier with mask $\mathbf{m}$. If we let $m(\lambda) := 1$, if $\lambda \in \Omega$ and 0 otherwise, then $H_{\mathbf{m},\mathbf{\Lambda}}$ is a discretization of the operator L_Ω in (2).

Given an operator $H_{\mathbf{m},\mathbf{\Lambda}}$ defined in (3), mapping $L^2(\mathbb{R}^d)$ into itself, we denote $K = \#\operatorname{supp}(\mathbf{m})$, assume that K is finite, and write $H_\mathbf{m}$ as a composition of the operator $G_{\sqrt{\mathbf{m}}} : f \mapsto [\sqrt{m(\lambda)}\langle f, \pi(\lambda)g\rangle]_{\lambda\in\Lambda\cap\operatorname{supp}(\mathbf{m})}$, mapping $L^2(\mathbb{R}^d)$ into $\mathbb{C}^K$ and its adjoint $G^*_{\sqrt{\mathbf{m}}}$.
Both $G_{\sqrt{\mathbf{m}}}$ and $G^*_{\sqrt{\mathbf{m}}}$ are finite-rank operators and can be written in their singular value decomposition:

$$G_{\sqrt{\mathbf{m}}} = \sum_{j=1}^{K} s_j \langle \cdot, v_j\rangle_{L^2} u_j, \tag{4}$$

$$G^*_{\sqrt{\mathbf{m}}} = \sum_{j=1}^{K} s_j \langle \cdot, u_j\rangle_{\mathbb{C}^K} v_j. \tag{5}$$

Then, applying $G^*_{\sqrt{\mathbf{m}}}$ to u_k yields $G^*_{\sqrt{\mathbf{m}}} \cdot u_k = s_k \cdot v_k$ and thus the eigenfunctions v_j of $H_{\mathbf{m},\mathbf{\Lambda}}$ may be obtained from the eigenfunctions of the Gramian operator $\Gamma_m := G_{\sqrt{\mathbf{m}}} \cdot G^*_{\sqrt{\mathbf{m}}}$ by

$$v_j = \frac{1}{s_j} \cdot G^*_{\sqrt{\mathbf{m}}} \cdot u_j,\ j = 1, \ldots, K. \tag{6}$$

In typical applications, where $H_{\mathbf{m},\mathbf{\Lambda}}$ is a matrix whose size depends on the signal length, the size of the corresponding Gramian matrix is $K \times K$ with K being the size of the support of the mask (or, in the case of 0/1-masks, the support of Ω) which is usually small enough for the computation of the spectral decomposition to be a feasible task. Furthermore, in (6) only the eigenfunctions corresponding to relevant eigenvalues s_j^2 need to be computed.

B. Computing the whole frame

Section III-A deals with the computation of the relevant eigenfunctions for each individual tile of the cover. To compute the whole frame we use the following observation based on the so-called covariance of the Short-Time Fourier transform.

Lemma 1: If $\Omega' = \Omega + z_0$, for some $z_0 \in \mathbb{R}^{2d}$. Then the eigenfunctions of L_Ω and $L_{\Omega'}$ are related by

$$\phi_k^{\Omega'} = \pi(z_0)\phi_k^{\Omega}, \quad k \geq 1.$$

where $\pi(x, w)f(t) = e^{2\pi iwt} f(t - x)$. Hence, if the cover $\{\Omega_\gamma : \gamma \in \Gamma\}$ in Theorem 1 consists of translates of N basic tiles $\Omega^1, \ldots, \Omega^N$,

$$\Omega_\gamma = \Omega^{k_\gamma} + z_\gamma, \quad 1 \leq k_\gamma \leq N, z_\gamma \in \mathbb{R}^{2d},$$

then only N sets of eigenfunctions need to be computed.

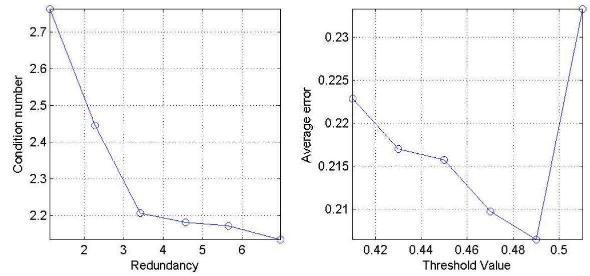

Fig. 6. *Evaluation of the procedure to obtain frames adapted to a given time-frequency partition.*

C. The number of eigenfunctions and the resulting frame quality

In order to test the performance of the procedure described in Theorem 1 and Lemma 1 for the generation of a new frame, we generated random partitions of the time-frequency plane, consisting of three different rectangular shapes, thus in the spirit of the example shown in Figure 2. Then, the eigenvalues of the resulting spectral decomposition were thresholded by 6 different values between 0.51 and 0.41 and the corresponding eigenfunctions were used to generate time-frequency frames with redundancies between 1.15 and 7. The condition numbers of the resulting frames are shown in Figure 6, as well as the average error, when the corresponding frame operators are applied to (1000 realizations of) random noise. Interestingly, while the condition number of the resulting frame improves for increased redundancy, the optimal approximation of the identity seems to be obtained for a threshold very close to 0.5. This agrees with the observation that the number of eigenvalues above 0.5 is given by the volume of the localization area [14], [11], [7], [12]. This effect can be circumvented by renormalizing the eigenfunctions to its corresponding eigenvalue (see [10]).

IV. Frames of eigenspaces

Theorem 1 can be interpreted in the following way. For each $\gamma \in \Gamma$ let V_γ be the subspace spanned by the first N_γ eigenfunction $\{\phi_1^\Omega, \dots, \phi_{N_\gamma}^\Omega\}$ and let P_γ be the corresponding orthogonal projection. Then,

$$\|f\|_2^2 \approx \sum_{\gamma \in \Gamma} \|P_\gamma f\|_2^2, \quad f \in L^2(\mathbb{R}^d).$$

This means that $\{V_\gamma : \gamma \in \Gamma\}$ is a *fusion frame* in the sense of [4]. In certain situations, using the projection P_γ may be preferable to masking the coefficients with a multiplier like the one in (3).

A. Cutting with reduced spilling

Denoising by time-frequency masking is a ubiquitous method in signal restoration, cf. [18]. However, in dependence on the time-frequency concentration of the window used to obtain the time-frequency representation used, this method leads to significant spilling of energy outside the region of relevant signal components. Applying projection onto significant eigenfunctions of a time-frequency multiplier instead of applying the multiplier itself, can ameliorate this bias. An example for this is shown in Figure 7. Here, a Hann window h was chosen as a reference signal, while the analysis window is still a Gaussian window. The signal h was disturbed by additive white noise, with a signal to noise ratio (SNR) of $3.5dB$ to obtain the noisy signal h_n. Then, the original signal was recovered by either applying a Gabor multiplier derived from a $0/1$-mask on the estimated region, with an underlying Gabor frame of redundancy 16, and, on the other hand, the projection onto the eigenfunctions corresponding to eigenvalues close to 1. The average achieved SNR (over a 1000 noise-realizations) was $12.5dB$ for the projection approach and $11.3dB$ for the plain Gabor multiplier.

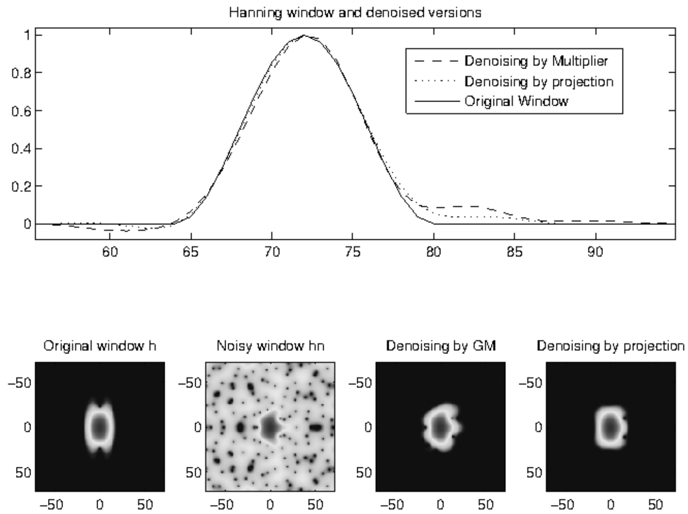

Fig. 7. *Denoising a disturbed signal by either time-frequency masking (GM) or projection onto relevant eigenfunctions. The upper plot shows the time-domain signals. The lower plots show the db-values of the spectrograms of the original window h, its noisy version h_n and the two denoised signals.*

V. Conclusion and Perspectives

In this article we have presented a new method to obtain frames adapted to a given partition of the time-frequency plane and addressed certain computational aspects of it. We also showed that using projections onto the space spanned by the first eigenfunctions corresponding to the Gabor multiplier of a certain localization region can yield better results than applying the Gabor multiplier itself. These are preliminary results that must be evaluated more extensively and in particular given a proof of concept by means of application to real-life data.

References

[1] A. Aldroubi, C. Cabrelli, and U. Molter. Wavelets on irregular grids with arbitrary dilation matrices, and frame atoms for $L^2(R^d)$. *Appl. Comput. Harmon. Anal.*, Special Issue on Frames II.:119–140, 2004.

[2] F. Auger and P. Flandrin. Improving the readability of time-frequency and time-scale representations by the reassignment method *IEEE Trans. Sig. Proc.*, 43(5):1068 –1089, May 1995.

[3] D. Jones and R. G. Baraniuk. A simple scheme for adapting time-frequency representations. *IEEE Trans. Signal Process.*, 42(12):3530–3535, Dec 1994.

[4] P. G. Casazza and G. Kutyniok. Frames of subspaces. In *Wavelets, Frames and Operator Theory*, volume 345 of *Contemp. Math.*, pages 87–113. Amer. Math. Soc., Providence, RI, 2004.

[5] I. Daubechies. Time-frequency localization operators: a geometric phase space approach. *IEEE Trans. Inform. Theory*, 34(4):605–612, July 1988.

[6] I. Daubechies, J. Lu, and H.-T. Wu. Synchrosqueezed wavelet transforms: an empirical mode decomposition-like tool. *Appl. Comput. Harmon. Anal.*, 30(2):243–261, 2011.

[7] F. DeMari, H. G. Feichtinger, and K. Nowak. Uniform eigenvalue estimates for time-frequency localization operators. *J. London Math. Soc.*, 65(3):720–732, 2002.

[8] M. Dörfler, H. G. Feichtinger, and K. Gröchenig. Time-frequency partitions for the Gelfand triple $(S_0, L^2, S_0{}')$. *Math. Scand.*, 98(1):81–96, 2006.

[9] M. Dörfler and K. Gröchenig. Time-frequency partitions and characterizations of modulations spaces with localization operators. *J. Funct. Anal.*, 260(7):1903 – 1924, 2011.

[10] M. Dörfler and J. L. Romero. Frames adapted to a phase-space cover. *Preprint.*, arXiv:1207.5383.

[11] H. G. Feichtinger and K. Nowak. A Szegö-type theorem for Gabor-Toeplitz localization operators. *Michigan Math. J.*, 49(1):13–21, 2001.

[12] H. G. Feichtinger and K. Nowak. A first survey of Gabor multipliers. In H. G. Feichtinger and T. Strohmer, editors, *Advances in Gabor Analysis*, Appl. Numer. Harmon. Anal., pages 99–128. Birkhäuser, 2003.

[13] F. Jaillet and B. Torrésani. Time–frequency jigsaw puzzle: adaptive multiwindow and multilayered Gabor expansions. *Int. J. Wavelets Multiresolut. Inf. Process.*, 2:293–316, 2007.

[14] J. Ramanathan and P. Topiwala. Time-frequency localization and the spectrogram. *Appl. Comput. Harmon. Anal.*, 1(2):209–215, 1994.

[15] J. L. Romero. Surgery of spline-type and molecular frames. *J. Fourier Anal. Appl.*, 17:135 – 174, 2011.

[16] J. L. Romero. Characterization of coorbit spaces with phase-space covers. *J. Funct. Anal.*, 262(1):59–93, 2012.

[17] P. J. Wolfe, D. Rudoy, and B. Prabahan. Superposition frames for adaptive time-frequency analysis and fast reconstruction. *IEEE Trans. Signal Process.*, 58:2581–2596, 2010.

[18] P. Wolfe and S. Godsill. Efficient alternatives to the Ephraim and Malah suppression rule for audio signal enhancement. *EURASIP J. Appl. Signal Process.*, 2003(10):1043–1051, 2003.

Monika Dörfler was supported by the WWTF project Audio-Miner (MA09-024). José Luis Romero gratefully acknowledges support from the Austrian Science Fund (FWF):[P22746-N13].