

# ADAPTIVE GABOR FRAMES BY PROJECTION ONTO TIME-FREQUENCY SUBSPACES

Monika Dörfler<sup>\*</sup> and Gino Angelo Velasco<sup>†</sup>

<sup>\*</sup>NuHAG, Faculty of Mathematics, University of Vienna, Austria

<sup>†</sup>Institute of Mathematics, University of the Philippines Diliman, Quezon City, Philippines

## ABSTRACT

The application of adaptive, time-frequency based signal analysis has recently attracted increasing attention. While using adapted time-frequency atoms has shown promising results for example in audio processing, the reconstruction from the corresponding analysis coefficients usually exhibits significant error. In this contribution we propose a method to reduce the reconstruction error by using modified time-frequency atoms in the transition region between adjacent areas of time-frequency adaptation. The modification is obtained by projecting the relevant atoms onto a system of weighted vectors which are optimally concentrated inside the desired regions of adaptation. We give a theoretical derivation of the improvement of error and illustrate our method with numerical examples.

**Index Terms**— Gabor frames; adaptive representations; audio processing; time-frequency localization operator

## 1. INTRODUCTION AND OUTLINE

Processing audio signals is most often based on some kind of local Fourier analysis. The most classical tool is the phase vocoder, aka short-time Fourier transform [1–4]. The basic analysis consists of the multiplication of signal parts with a finite-length and sufficiently smooth time-window and subsequent Fourier transform. To obtain more flexibility with respect to the length of the localizing window and thus the time-resolution and, correspondingly, frequency resolution, transforms with locally varying window-lengths have been proposed, e.g. lapped transforms [5] or, more recently, the non-stationary Gabor transform [6].

However, in some cases, the thus-obtained flexibility is not sufficient; certain applications require the usage of *different* time-adaptive window systems in different frequency bands. Examples for these situations occurring in audio processing can be found in [7]. In the latter paper, the problem is approached by weighting the synthesis windows prior to reconstruction. Thereby, regions of overlap between adjacent frequency bands (2 in the case of [7]) are introduced and the transition is accomplished by using weights which sum up to 1. While the method is innovative and leads to significant improvements of processing results in certain application situations, the reconstruction of a given signal from the unaltered coefficients features rather high errors in the transition regions between adjacent frequency bands.

In this contribution we suggest a new method to obtain analysis-synthesis systems which provide desired resolution in various frequency bands and at the same time provide arbitrarily good reconstruction quality. These properties are achieved by relying on recently achieved results on time-frequency localization operators and frames obtained from them, cf. [8].

MD is funded by the Vienna Science and Technology Fund (WWTF) through project VRG12-009.

The two main contributions of this article are the following:

- We propose the application of *time-frequency localization operators* corresponding to the desired partition into frequency bands. At the same time, the windows can also change over time as desired. Then, by projecting the reconstruction atoms of interest onto subspaces generated by the eigenvectors which are best-concentrated in each frequency band and time interval, we naturally obtain a smooth transition between adjacent frequency bands and time intervals.
- We show how to pick the number of reconstruction atoms outside each time-frequency area of interest according to a prescribed tolerable error bound and provide a number of numerical results in comparison to the state-of-the-art method proposed in [7].

### 1.1. Previous Work

Truly flexible tilings of the time-frequency plane have frequently been addressed in the past decades. The earliest work tried to establish orthogonal bases with a prescribed time-frequency profile [9]. Partitions based on Gabor frames and information criteria were introduced in [10] and also applied in [11] and their existence under relatively general conditions was proved in [12]. Recently, the idea to obtain discrete tilings based of the spectral decomposition of continuous time-frequency localization operators, cf. [13], was introduced, see [8, 14].

In parallel to the endeavor to simultaneously obtain flexibility in time and frequency, the concept of *nonstationary Gabor frames* provides a convenient tool for obtaining adaptivity in either time or frequency, see [6, 15, 16].

On the other hand, the desire to accurately manipulate certain signal components while leaving surrounding ones as unbiased as possible, often arises in applications and corresponding analysis coefficients are modified accordingly. However, due to the uncertainty principle, the separation of components is usually not precise enough and thus these approaches typically lead to rather high reconstruction errors, as also reported in [7]. In this contribution, we thus propose a completely novel approach of a theoretically motivated adaptation of synthesis windows.

## 2. THEORETICAL DERIVATION

### 2.1. Time-frequency localization

We introduce the concept of localization in time-frequency in the continuous domain in order to avoid heavy notation. In Section 3 we show how the continuous setting can be mimicked in a discrete setting.

The short-time Fourier transform (STFT) of a function  $f \in L^2(\mathbb{R})$  is

a function on  $\mathbb{R}^2$  defined, by means of an adequate smooth and fast-decaying window function  $\varphi \in L^2(\mathbb{R})$ ,  $\|\varphi\|_2 = 1$ , e.g. a normalized Gaussian window, as

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

Motivated by the well-known fact, that  $f$  can be re-synthesized from its time-frequency content by,

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

the *time-frequency localization operator*  $H_\Omega$  is defined, for some compact set  $\Omega \subseteq \mathbb{R}^2$  in the time-frequency plane, by masking the coefficients in (1), cf. [13], i.e.

$$H_\Omega f(t) = \int_{\Omega} \mathcal{V}_\varphi f(x, \xi) \varphi(t-x) e^{2\pi i \xi t} dx d\xi. \quad (2)$$

$H_\Omega$  is self-adjoint and trace-class, so we can consider its spectral decomposition  $H_\Omega f = \sum_{k=1}^{\infty} \mu_k^\Omega \langle f, \phi_k^\Omega \rangle \phi_k^\Omega$ . Its eigenfunction  $\phi_1^\Omega$  corresponding to the largest eigenvalue  $\mu_1^\Omega$ , is optimally concentrated inside  $\Omega$  in the sense that

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

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

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

among all orthonormal sets of  $N$  functions  $\{\psi_j\}_{j=1, \dots, N}$  in  $L^2(\mathbb{R}^2)$ . In this sense, their time-frequency profile is optimally adapted to  $\Omega$ .

## 2.2. Gabor frames

We now obtain a collection of time-frequency shifted functions from sampling of the time-frequency plane. Given a window function  $g \in L^2(\mathbb{R})$  and a lattice  $\Lambda$  with TF parameters  $(t_0, \omega_0)$ , i.e.  $\Lambda = t_0\mathbb{Z} \times \omega_0\mathbb{Z}$ , we denote a time-frequency shift by  $\lambda = (\lambda_1, \lambda_2)$  by  $\pi(\lambda)$ , i.e.  $\pi(\lambda)g(t) = g(t - \lambda_1) e^{2\pi i \lambda_2 t}$ . Then, the collection  $\mathcal{G}(g, \Lambda) = \mathcal{G}(g, t_0, \omega_0) = \{\pi(\lambda)g : \lambda \in \Lambda\}$  is called a *Gabor system*. Associated with  $\mathcal{G}(g, \Lambda)$  are the *analysis operator*  $C$ , given by  $Cf = \{\langle f, \pi(\lambda)g \rangle : \lambda \in \Lambda\}$ , *synthesis operator*  $D = C^*$ , given by  $Dc = \sum_{\lambda \in \Lambda} c_\lambda \pi(\lambda)g$ ,  $c \in \ell^2$ , and the *frame operator*  $S$  given by  $Sf = DCf = \sum_{\lambda \in \Lambda} \langle f, \pi(\lambda)g \rangle \pi(\lambda)g$ . Equivalently, the Gabor system  $\mathcal{G}(g, \Lambda)$  is a *Gabor frame*, cf. [17], if  $S$  is invertible. In this case, there exists a dual frame  $\mathcal{G}(\tilde{g}, \Lambda)$  and reconstruction is possible via:

$$f = \sum_{\lambda \in \Lambda} \langle f, \pi(\lambda)g \rangle \pi(\lambda)\tilde{g} = \sum_{\lambda \in \Lambda} \langle f, \pi(\lambda)\tilde{g} \rangle \pi(\lambda)g. \quad (3)$$

The canonical dual window generating a dual frame is given by  $\tilde{g} = S^{-1}g$ .

If  $S = AI$ , where  $I$  is the identity operator and  $A$  a positive constant, the Gabor frame is said to be *tight*. Inversion is done simply by  $S^{-1} = \frac{1}{A}I$ . Now, a tight Gabor frame can actually be obtained from any Gabor frame. If the Gabor system  $\mathcal{G}(g, \Lambda)$  is a frame, then it follows from positivity of the (self-adjoint) frame operator  $S$  that the operator  $S^{-1/2}$  is well defined, positive and self-adjoint, and

$$f = S^{-1/2}S(S^{-1/2}f) = \sum_{\lambda \in \Lambda} \langle f, \pi(\lambda)S^{-1/2}g \rangle \pi(\lambda)S^{-1/2}g. \quad (4)$$

Letting  $g_t = S^{-1/2}g$ , the system  $\mathcal{G}(g_t, \Lambda)$  is a tight frame and  $g_t$  is called the canonical tight window. We see from (4) that a tight frame has the advantage of using the same window for analysis and synthesis, a property that is useful in applications that involve resynthesis of processed analysis coefficients, e.g. masking, which we shall do in the succeeding sections.

## 2.3. Obtaining sets of localized, adapted atoms

Assume now that we are given a partition of  $\mathbb{R}^2$ , i.e. a set family of sets  $\Omega_\gamma \subset \mathbb{R}^2$  such that  $\sum_\gamma 1_{\Omega_\gamma} \equiv 1$ . Here,  $1_\Omega$  denotes the indicator function of the set  $\Omega$ . Then, from (1) and using the spectral decomposition of each  $H_{\Omega_\gamma}$ , we obtain

$$f = \sum_\gamma H_{\Omega_\gamma} f = \sum_\gamma \sum_{j=1}^{\infty} \mu_j^{\Omega_\gamma} \langle f, \phi_j^{\Omega_\gamma} \rangle \phi_j^{\Omega_\gamma}.$$

Now assume further, that a tight Gabor frame  $\mathcal{G}^\gamma(g^\gamma, \Lambda^\gamma)$  is assigned to each set  $\Omega_\gamma$ . Expanding  $f$  with respect to each of these frames, we obtain:

$$f = \sum_\gamma \sum_{\lambda \in \Lambda^\gamma} \langle f, \pi(\lambda)g^\gamma \rangle \sum_{j=1}^{\infty} \mu_j^{\Omega_\gamma} \langle \pi(\lambda)g^\gamma, \phi_j^{\Omega_\gamma} \rangle \phi_j^{\Omega_\gamma} \quad (5)$$

In a next step we make two observations:

- The largest eigenvalues  $\mu_j^{\Omega_\gamma}$  of a localization operator typically are close to one and then drop to zero very fast (in fact the sequence  $(\mu_j^{\Omega_\gamma})_j$  has exponential decay), cf. [18]. Precisely  $|\Omega_\gamma|$  eigenvalues lie above 0.5. Consequently, one can safely discard elements with index numbers  $j > N_\gamma$  for some  $N_\gamma > |\Omega_\gamma|$  in (5).
- On the other hand, the inner product  $\langle \pi(\lambda)g^\gamma, \phi_j^{\Omega_\gamma} \rangle$  can be shown to decay fast with respect to the distance of  $\lambda$  from  $\Omega_\gamma$ , e.g., for a Gaussian windows  $g^\gamma$ , the decay is exponential, while milder decay conditions lead to polynomial decay, cf. [13]. Therefore, all  $\pi(\lambda)g^\gamma$  with  $\text{dist}(\lambda, \Omega_\gamma) \geq b$  for some  $b$  can be omitted from the expansion (5).

We thus choose an appropriate  $N_\gamma$ , an extension size or overlap  $b$  and set  $\Omega_\gamma^* = \Omega_\gamma \cup \{z \in \mathbb{R}^2 \setminus \Omega_\gamma : \text{dist}(z, \Omega_\gamma) < b\}$ . We then propose to use the following approximate reconstruction formula:

$$\begin{aligned} \tilde{f} &= \sum_\gamma \sum_{\lambda \in \Lambda^\gamma \cap \Omega_\gamma^*} \langle f, \pi(\lambda)g^\gamma \rangle \sum_{j=1}^{N_\gamma} \mu_j^{\Omega_\gamma} \langle \pi(\lambda)g^\gamma, \phi_j^{\Omega_\gamma} \rangle \phi_j^{\Omega_\gamma} \quad (6) \\ &= \sum_\gamma \sum_{\lambda \in \Lambda^\gamma \cap \Omega_\gamma^*} \langle f, \pi(\lambda)g^\gamma \rangle P_{N_\gamma}(\pi(\lambda)g^\gamma) \quad (7) \end{aligned}$$

## 2.4. Error estimates

**Proposition 1.** *Let a partition of  $\mathbb{R}^2$  be given and let the windows  $g^\gamma$  satisfy a joint polynomial decay condition of the form  $|\mathcal{V}_\varphi g^\gamma(z - \lambda_\gamma)| \leq C \frac{1}{1+|z-\lambda_\gamma|^s}$  for all  $z \in \mathbb{R}^2$  and all  $\lambda_\gamma \in \Lambda^\gamma$ . Let  $\tilde{f}$  be the approximate reconstruction of  $f$  in (6). Then, the reconstruction error is bounded by  $\|\tilde{f} - f\|_2 \leq \sum_\gamma \text{err}_\gamma \cdot \|f\|_2$ , where for all  $\gamma$*

and some  $0 < \delta < 1$  the following estimate holds:

$$\begin{aligned} \text{err}_\gamma \leq & \left( \sum_{j=N_\gamma+1}^{\infty} (\mu_j^{\Omega_\gamma})^2 \right)^{\frac{1}{2}} + \\ & + \left( C \sum_{\lambda \notin \Lambda^\gamma \cap \Omega_\gamma^*} \left( 1 + (1 - \delta) \left( \inf_{z \in \Omega} |z - \lambda|^{\frac{\delta}{2}} \right) \right)^{-1} \right) \end{aligned} \quad (8)$$

It should be noted that the sum of  $\text{err}_\gamma$  over all  $\gamma$  can be shown to be finite for appropriate choice of  $N_\gamma$  and  $\Omega_\gamma^*$ , cf. [8, 12]. Here, we prefer to state the explicit *local* errors, since their expression is more informative in showing directly the influence of the parameters  $N_\gamma$  and  $b$ .

### 3. DERIVED ALGORITHM

#### 3.1. Computation of $P_N$

To obtain the eigenvectors and  $\lambda$ -values needed for the approximation in (6), we work with discrete versions of the localization operators  $H_{\Omega_\gamma}$ . To this end, consider the tight Gabor frame  $(g_t, \Lambda)$ . We define the Gabor multiplier  $H_{m_\gamma, \Lambda}$  as follows:

$$H_{m_\gamma, \Lambda} f = \sum_{\lambda \in \Lambda} m_\gamma(\lambda) \langle f, \pi(\lambda) g_t \rangle \pi(\lambda) g_t, \quad (9)$$

where the masks  $m_\gamma$  are obtained by letting  $m_\gamma(\lambda) := 1$ , if  $\lambda \in \Omega_\gamma$  and 0 otherwise. Then  $H_{m_\gamma, \Lambda}$  is a discretization of the operator  $H_{\Omega_\gamma}$  in (2) and it can be shown that its spectral decomposition accurately approximates  $H_{\Omega_\gamma}$  for sufficiently dense lattice  $\Lambda$ , cf. [19, 20].

In applications  $H_{m_\gamma, \Lambda}$  is a matrix whose size depends on the signal length  $L$  and it is not feasible to obtain its eigenvectors directly. However, the size of the corresponding *Gramian matrix*, defined in (10) below, is  $K \times K$  with  $K$  being the number of the lattice points  $\lambda_\gamma$  inside the support of the mask  $m_\gamma$ , which is usually small enough for the computation of the spectral decomposition to be a feasible task. Writing  $H_{m_\gamma, \Lambda}$  as a composition of the operator  $G_{\sqrt{m_\gamma}} : f \mapsto [\sqrt{m_\gamma}(\lambda) \langle f, \pi(\lambda) g \rangle]_{\lambda \in \Lambda \cap \text{supp}(m)}$ , mapping  $\mathbb{C}^L$  into  $\mathbb{C}^K$ , and its adjoint  $G_{\sqrt{m_\gamma}}^*$ , the eigenfunctions of  $H_{m_\gamma, \Lambda} = G_{\sqrt{m_\gamma}}^* \cdot G_{\sqrt{m_\gamma}}$  may be obtained from the eigenfunctions of the Gramian matrix

$$\Gamma_{m_\gamma} := G_{\sqrt{m_\gamma}} \cdot G_{\sqrt{m_\gamma}}^* \quad \text{by} \quad (10)$$

$$\phi_j^{\Omega_\gamma} = \frac{1}{s_j} \cdot G_{\sqrt{m_\gamma}}^* \cdot u_j, \quad j = 1, \dots, K, \quad (11)$$

where  $G_{\sqrt{m_\gamma}} f = \sum_{j=1}^K s_j \langle f, \phi_j^{\Omega_\gamma} \rangle_{\mathbb{C}^L} u_j$  is the singular value decomposition of  $G_{\sqrt{m_\gamma}}$ . Furthermore, in (11) only the largest  $N_\gamma$  eigenfunctions  $u_j$  need to be computed.

#### 3.2. Choosing $N_\gamma$ and $\Omega^*$

For each  $\gamma$ ,  $N_\gamma$  eigenfunctions  $\{\phi_j^{\Omega_\gamma}\}_j$  of  $H_{m_\gamma, \Lambda}$ , associated to the eigenvalues  $\mu_j^{\Omega_\gamma}$  greater than a threshold  $t_\gamma$  must be chosen. In practice, the  $\Omega_\gamma$  often are of the same area, and we just take the same value of  $N_\gamma$  for each  $\gamma$ . Choosing  $N_\gamma$  such that  $\mu_{N_\gamma}^{\Omega_\gamma} < 10^{-m}$ , the first expression in the error estimate (8) is bounded by  $10^{-m}$  due to the exponential decay of the eigenvalues.

Second, we choose a rectangular extension  $\Omega_k^*$  of  $\Omega_k$  by increasing the sides of  $\Omega_k$  also by a margin of size  $b$ , such that in the second expression of (8), the value  $\inf_{z \in \Omega} |z - \lambda|$  is sufficiently big for all  $\lambda \notin \Lambda^\gamma \cap \Omega_\gamma^*$ .

## 4. NUMERICAL EXPERIMENTS

To provide a numerical evidence of our concept, we look at examples in the finite discrete case  $\mathbb{C}^L$ ,  $L = 144$ . The time-frequency plane will be partitioned into four parts, dividing the time axis at  $t_{\text{cut}} = L/2$ , and the frequency axis into bands corresponding to the frequencies above and below  $\omega_{\text{cut}} = L/4$ . We note that these frequency bands extend to the negative frequencies in a symmetric manner about the frequency 0.

The following tight Gabor frames will then be associated to the four regions:

1.  $\mathcal{G}(g_t^1, 12, 4)$  at the region  $\Omega_1$  (lower frequency region and time  $t \leq L/2$ );
2.  $\mathcal{G}(g_t^2, 16, 6)$  at the region  $\Omega_2$  (lower frequency region and time  $t > L/2$ );
3.  $\mathcal{G}(g_t^3, 8, 16)$  at the region  $\Omega_3$  (higher frequency region and time  $t \leq L/2$ ); and
4.  $\mathcal{G}(g_t^4, 9, 12)$  at the region  $\Omega_4$  (higher frequency region and time  $t > L/2$ ).

The signal will be analyzed using these tight Gabor frames, and applied with weighted functions or restricted over regions that cover our partitions. Reconstruction is performed via the method introduced in [7] and the proposed method, respectively. We then compare the approximation quality from the two methods.

For the approximate reconstruction [7], weight functions  $W_T^1$  and  $W_T^2$ , depending only on time, and  $W_F^1$  and  $W_F^2$ , depending only on frequency, shall be applied to the analysis coefficients. These weight functions are defined as follows:

$$W_T^1(t) := \begin{cases} 1 & \text{if } 1 \leq t \leq t_1 \\ \frac{t-t_2}{t_1-t_2} & \text{if } t_1 \leq t \leq t_2 \\ 0 & \text{elsewhere} \end{cases}$$

with  $t_1 \leq t_{\text{cut}} \leq t_2$ ,  $W_T^2 := 1 - W_T^1$ , i.e.  $W_T^1(t) + W_T^2(t) = 1$  for each  $t$ ,

$$W_F^1(\omega) := \begin{cases} 1 & \text{if } -\omega_1 \leq \omega \leq \omega_1 \\ \frac{\omega-\omega_2}{\omega_1-\omega_2} & \text{if } \omega_1 \leq \omega \leq \omega_2 \\ \frac{\omega+\omega_2}{\omega_2-\omega_1} & \text{if } -\omega_2 \leq \omega \leq -\omega_1 \\ 0 & \text{elsewhere} \end{cases}$$

where  $\omega_1 \leq \omega_{\text{cut}} \leq \omega_2$ ,  $W_F^2 := 1 - W_F^1$ , i.e.  $W_F^1(\omega) + W_F^2(\omega) = 1$  for each  $\omega$ . Figure 1 shows the four weight functions. We note that varying the  $t_i$  and  $\omega_i$  amounts to varying the overlap of the weight functions. In the experiment, the overlap value  $b := t_2 - t_{\text{cut}} = t_{\text{cut}} - t_1$  for the weight function in time shall also be used for the weight function in frequency so that  $b = \omega_2 - \omega_{\text{cut}} = \omega_{\text{cut}} - \omega_1$ .

Recall that in [7], the reconstruction formula is given by

$$\tilde{f}_W = \sum_{k=1}^4 \sum_{\lambda \in \Lambda^k} W_{TF}^k(t, \omega) \langle f, \pi(\lambda) g_t^k \rangle \pi(\lambda) g_t^k, \quad (12)$$

where  $W_{TF}^k$  corresponds to  $W_T^1 \cdot W_F^1$  for  $k = 1$ ,  $W_T^2 \cdot W_F^1$  for  $k = 2$ ,  $W_T^1 \cdot W_F^2$  for  $k = 3$ , and  $W_T^2 \cdot W_F^2$  for  $k = 4$ .

We now compare the errors in approximating  $f$  using the methods described above. Figure 2 shows the average of the root mean

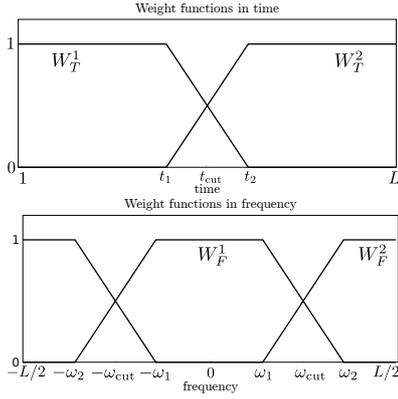


Fig. 1. Weight functions  $W_T^1$ ,  $W_T^2$ ,  $W_F^1$ , and  $W_F^2$ .

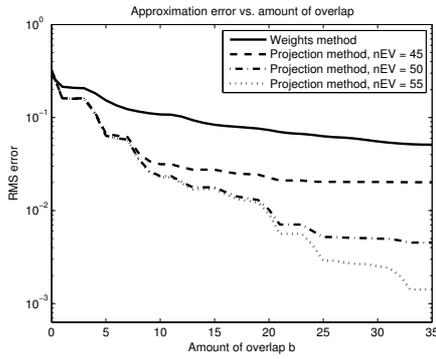


Fig. 2. Approximation error vs. amount of overlap.

square (RMS) of the error given by

$$\text{err}(f_{\text{rec}}) = \frac{\|f - f_{\text{rec}}\|_2}{\|f\|_2} = \sqrt{\frac{\sum_{n=1}^L (f[n] - f_{\text{rec}}[n])^2}{\sum_{n=1}^L (f[n])^2}},$$

of 50 random signals against the amount of overlap  $b$ . The solid line is from the weight function method in [7] while the non-solid lines result from the proposed projection method. Each of the non-solid lines uses a different number of eigenfunctions: 45, 50, and 55 eigenfunctions, corresponding to the eigenvalue thresholds  $t_\gamma = 0.1016, 0.0243, \text{ and } 0.0040$ , respectively. In both methods, we see the dependence of the approximate reconstruction on the overlap amount. In the case of our proposed method, the second term in (8) approaches 0 as the overlap, or margin,  $b$  increases. Moreover, the projection method has the added possibility of improving the approximation error by increasing the number of eigenfunctions in the reconstruction. The dependence of the reconstruction error on the number of eigenfunctions in the subspace is depicted in Figure 3.

Finally, we point out, that the separation between the distinct regions chosen for the different desired resolutions, that is,  $\Omega_\gamma$ ,  $\gamma = 1, \dots, 4$ , is much sharper using the projection method. This fact is illustrated in Figure 4, where we show the results of applying one of the local systems to random white noise. Depicted are the spectrograms of the results for the systems corresponding to low frequencies, first signal part and high frequencies, second signal part,

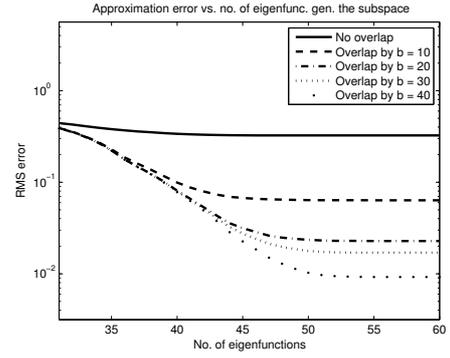


Fig. 3. Approximation error vs. number of eigenfunctions in the subspace.

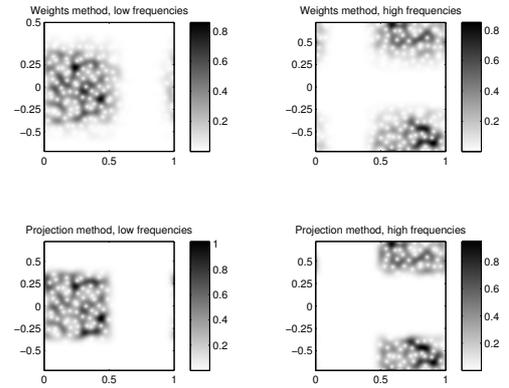


Fig. 4. Concentration of local systems within  $\Omega_\gamma$ . The spectrograms of local systems applied to random noise are shown.

respectively. For both methods, the set of parameters providing the best approximation quality is used. It can clearly be seen, that the projection method significantly reduces the spill outside the region of interest which is quite considerable in the weight function method.

## 5. DISCUSSION AND PERSPECTIVES

In this contribution, we introduced an innovative method for obtaining time-frequency frames with desired local properties. While our experiments provide only toy-examples for signals of short length, all methods can be efficiently extended to realistic signal lengths. The implementation of the necessary routines and their subsequent evaluation on a database of realistic signals is part of the future work on the topic. However, already the preliminary results show the promising potential of the proposed method: it provides arbitrarily good approximation quality while conserving the good localization property. This is in clear contrast to previously existing methods. On the other hand, the computational effort is significantly higher; accurate evaluation in a realistic scenario will be provided in a future contribution. However, for many of the applications of interest, real-time processing is not an issue and thus, higher computational cost can be accepted for the sake of better processing results.

## 6. REFERENCES

- [1] M. Dolson, "The phase vocoder: a tutorial," *Computer Musical Journal*, vol. 10, no. 4, pp. 11–27, 1986.
- [2] M. R. Portnoff, "Time–frequency representation of digital signals and systems based on short-time Fourier analysis," *IEEE Trans. Acoustics, Speech and Signal Processing*, vol. 28, no. 1, pp. 55–69, 1980.
- [3] L. R. Rabiner and J. B. Allen, "Short-time Fourier analysis techniques for FIR system identification and power spectrum estimation," *IEEE Trans. Acoustics, Speech and Signal Processing*, vol. 27, no. 2, pp. 182–192, 1979.
- [4] M. Dörfler, "Time-frequency Analysis for Music Signals. A Mathematical Approach," *Journal of New Music Research*, vol. 30, no. 1, pp. 3–12, 2001.
- [5] H. Malvar, *Signal Processing with Lapped Transforms*. Boston, MA: Artech House, xvi, 1992.
- [6] P. Balazs, M. Dörfler, F. Jaillet, N. Holighaus, and G. A. Velasco, "Theory, implementation and applications of nonstationary Gabor Frames," *J. Comput. Appl. Math.*, vol. 236, pp. 1481–1496, 2011.
- [7] M. Liuni, A. Robel, E. Matusiak, M. Romito, and X. Rodet, "Automatic adaptation of the time-frequency resolution for sound analysis and re-synthesis," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 21, no. 5, pp. 959–970, 2013.
- [8] M. Dörfler and J. L. Romero, "Frames adapted to a phase-space cover," *Constr. Approx.*, vol. to appear, 2012.
- [9] C. Herley, K. Ramchandran, M. Vetterli, and J. Kovacevic, "Tilings of the time-frequency plane: Construction of arbitrary orthogonal bases and fast tiling algorithms," *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3341–3359, 1993.
- [10] M. Dörfler, "Quilted Gabor frames - A new concept for adaptive time-frequency representation," *Advances in Applied Mathematics*, vol. 47, no. 4, pp. 668 – 687, Oct. 2011.
- [11] F. Jaillet and B. Torrèsani, "Timefrequency jigsaw puzzle: adaptive multiwindow and multilayered Gabor expansions," *Int. J. Wavelets Multiresolut. Inf. Process.*, vol. 2, pp. 293–316, 2007.
- [12] J. L. Romero, "Surgery of spline-type and molecular frames," *J. Fourier Anal. Appl.*, vol. 17, pp. 135 – 174, 2011.
- [13] I. Daubechies, "Time-frequency localization operators: a geometric phase space approach," *IEEE Trans. Inform. Theory*, vol. 34, no. 4, pp. 605–612, July 1988.
- [14] M. Dörfler and J. L. Romero, "Frames of eigenfunctions and localization of signal components," *Proceedings Sampta13*, July 2013.
- [15] N. Holighaus, M. Dörfler, G. A. Velasco, and T. Grill, "A framework for invertible, real-time constant-Q transforms," *IEEE Trans. Audio Speech Lang. Process.*, vol. 21, no. 4, pp. 775 –785, 2013.
- [16] T. Necciari, P. Balazs, N. Holighaus, and P. Sondergaard, "The ERBlet transform: An auditory-based time-frequency representation with perfect reconstruction," in *Proceedings of the 38th International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2013)*, 2013, pp. 498–502.
- [17] H. G. Feichtinger and T. Strohmer, *Gabor Analysis and Algorithms. Theory and Applications*. Boston: Birkhäuser, 1998.
- [18] F. DeMari, H. G. Feichtinger, and K. Nowak, "Uniform eigenvalue estimates for time-frequency localization operators," *J. London Math. Soc.*, vol. 65, no. 3, pp. 720–732, 2002.
- [19] H. G. Feichtinger and K. Nowak, "A first survey of Gabor multipliers," in *Advances in Gabor Analysis*, ser. Appl. Numer. Harmon. Anal., H. G. Feichtinger and T. Strohmer, Eds. Birkhäuser, 2003, pp. 99–128.
- [20] M. Dörfler and H. G. Feichtinger, "Orthogonal projections derived from localization operators," in *Proc. Conf. EUSIPCO (Sept. 2004, TU Vienna)*, 2004, pp. 1195—1198.