Double Preconditioning for Gabor Frames

Peter Balazs*, *Member, IEEE,* Hans G. Feichtinger, Mario Hampejs and Günther Kracher

***Abstract*— We present an application of the general idea of preconditioning in the context of Gabor frames. While most (iterative) algorithms aim at a more or less costly exact numerical calculation of the inverse Gabor frame matrix, we propose here the use of "cheap methods" to find an approximation for it, based on (double) preconditioning. We thereby obtain good approximations of the true dual Gabor atom at low computational costs. Since the Gabor frame matrix commutes with certain time-frequency shifts it is natural to make use of diagonal and circulant preconditioners sharing this property. Part of the efficiency of the proposed scheme results from the fact that all the matrices involved share a well-known block matrix structure. At least, for the smooth Gabor atoms typically used, the combination of these two preconditioners leads consistently to good results. These claims are supported by numerical experiments in the second part of the paper. For numerical evaluations we introduce two new matrix norms, which can be calculated efficiently by exploiting the structure of the frame matrix.**

***Index Terms*— Block matrices; efficient algorithm; Gabor frame matrices; approximated dual windows; time-frequency analysis; matrix norms; discrete transforms; matrix inversion; EDICS : DSP-WAVL Wavelets theory and applications; DSP-FAST Fast algorithms for digital signal processing**

I. Introduction:

THE Short-time Fourier transform (STFT), also called Gabor-Transform in its sampled variant, is a well known, valuable tool for displaying the energy distribution of a signal f over the time-frequency plane. The equivalence between Gabor analysis and certain filter banks is a well-known fact [1]. For a number of applications for example in audio processing like time stretching without changing the frequency content [2], more complex modifications like psychoacoustical masking [3] or others ([4], [5], [6]), the time domain signal needs to be reconstructed using the time-frequency domain coefficients. The dual, atomic composition problem, building a given signal as a series using a time-frequency shifted window as building blocks (see e.g. [7]), is also needed in applications.

The main question is, how to find a Gabor analysis-synthesis system with perfect (or depending on the application a satisfactorily accurate) reconstruction in a numerical efficient way. Basic Gabor frame theory (cf. introduction of [8]) tells us, that when using the so called canonical *dual* Gabor atom $\tilde{g} = S^{-1}g$, perfect reconstruction is always achieved, if the *frame-operator* S (see Section II-B) is invertible. The dual atom is thus obtained by solving the equation $S\tilde{g} = g$, and to this end, the Neumann algorithm (see Figure 1) with relaxation parameter λ can be applied. If the inequality $\|Id - \lambda S\|_{Op} < 1$ holds, this algorithm converges, S is invertible and the algorithm approximates the dual Gabor atom $\tilde{g}$.

Instead of finding the canonical dual, other dual windows can be searched for, and sometimes they can be found in a numerically more efficient way (see [9]). But in general, the computation of a dual window can be very complicated and numerically inefficient. The Zak transform [10], [11] is heavily used for theoretic purposes, but its use for numerical calculations is limited [12]. The celebrated paper from Wexler and Raz [13] gives an important bi-orthogonality relation, which reduces the problem to a simple linear system. In order to find a very efficient algorithm, Qiu and Feichtinger use the sparse structure of the frame operator [14], which leads them to an algorithm for the inversion of the frame matrix with $O(abn)$ operations, where n is the signal length and a,b are the time and frequency shift parameters.

In this work another well known tool to speed up the convergence rate, namely, *preconditioning*, is used to further improve the numerical efficiency of this calculation. In our proposed method, we use a special invertible preconditioning matrix P, which makes $\|Id - PS\|$ small. Then, instead of $S\tilde{g} = g$, the equation $PS\tilde{g} = Pg$ is solved. The matrix $M = P \cdot S$ is therefore intended to be an approximation of the identity. If M is a reasonable good approximation, e.g., $\|Id - M\| < 0.1$, then only a few iterations are needed in order to find the true dual atom (up to precision limitations). Moreover, if M is a very good approximation, e.g., $\|Id - M\| \ll 0.1$, then the preconditioning matrix P can already be considered as a good approximation of the inverse matrix of S.

The aim of this article is to investigate the idea of *double preconditioning* of the frame operator S. This method was already suggested as an idea in the very last paragraphs of [12] and [15]. In this paper the double preconditioning method will be fully developed, examined and backed up with systematic experimental numerical data. This scheme relies, again, on the very special structure of the Gabor frame operator S, it is an a block-circulant matrix with b diagonal blocks. It is well known, that there are two extreme cases for this nice structure. (1) If the frequency sampling is dense enough and g has support inside an interval I, with the length $\leq \frac{n}{b}$, then S is a diagonal matrix. (2) If the time sampling is dense enough and $\hat{g}$ has compact support on an interval with length $\leq \frac{n}{a}$, then $\hat{S}$ is diagonal and therefore S is circulant. In both cases it is easy to find the inverse matrix [12]. If the window g is not

This work was partly supported by the European Union's Human Potential Programme, under contract HPRN-CT-2002-00285 (HASSIP).

P. Balazs is with the Acoustics Research Institute of the Austrian Academy of Sciences, Reichsratsstrasse 17, A-1010 Wien, Austria. (e-mail: Peter.Balazs@oeaw.ac.at; Telephone: +43 1 4277 29510; Fax: +43 1 4277 9295)

The other authors are with NuHAG, Faculty of Mathematics, University of Vienna, Nordbergstrasse 15, A-1090 Austria (e-mail: hans.feichtinger@univie.ac.at; mario.hampejs@siemens.at; gkracher@gmx.at; Telephone: +43 1 4277 50696; Fax: +43 1 4277 50690)

supported on I, then S becomes non-diagonal. However, if S is strictly diagonal dominated it is well known for $D = (d_{i,j})$, with $d_{i,j} = \delta_{i,j} S_{i,j}$ i.e. the best approximation of S by diagonal matrices, S^{-1} can be approximated well using the preconditioning matrix $P = D^{-1}$ (see the *Jacobi* method e.g., in [16] or [17]). An analogue property holds if $\hat{S}$ is strictly diagonal dominated, obtaining a circulant matrix as preconditioning matrix. If using these two preconditioning matrices at the same time, hence the name *double preconditioning*, we will get a new method.

The main observation is the fact that the use of double preconditioning often leads to better results than the use of single preconditioning. Moreover, in the cases where this is not true, the difference is in general not significant. This behavior is observed in numerical experiments. More precisely, we will first study single cases and then proceed with systematic experiments, where the efficiency of the double preconditioning method is investigated for different windows.

In this work we will also introduce two new norms, motivated by the need for a numerically efficient measure for evaluating the speed of convergence. For the Neumann algorithms the speed of convergence can be estimated by

$$\|x_k - \tilde{g}\| \leq C \cdot \|Id - \lambda S\|_{Op}^k \|x_0 - \tilde{g}\|. \tag{1}$$

It is well known, that the operator norm is not only a measure of the convergence speed but also for the condition of the problem. The condition number $\kappa(S) = \|S\|_{Op} \cdot \left\|S^{-1}\right\|_{Op}$ is a widely used tool for describing the numerical stability of a linear problem [18]. Because the calculation of the operator norm is computationally costly ([18]) we introduce two alternative norms. Both dominate the operator norm and have the extra benefit that they can be easily obtained from the (compressed) block-representations of the matrix S. There are algorithms which need good approximations of the upper frame bound respectively the operator norm of S, so the results for these norms are interesting independently of the double preconditioning idea in conjecture with such algorithms e.g. found in [19].

The paper is organized as follows. In Section II we will review basic facts about frames, discrete Gabor analysis, matrix algebras and algorithms. In Section III we will investigate the connection between two different representations of S as block matrices and introduce two norms. In Section IV we will review and extend the use of diagonal and circulant matrices as preconditioners for the Gabor frame operator. In Section V we will explain how to combine these preconditioners to invert the frame matrix S, and finally, in Section VI, we will demonstrate the efficiency of this idea.

This paper is planned as a first of a series on this and related topics and therefore, in some cases, material has been included with this perspective in mind.

II. Preliminaries And Notations

A. Matrices

We regard vectors (e.g., discrete signals) $x = (x_0, x_1, \ldots, x_{n-1}) \in \mathbb{C}^n$ as periodic functions on $\mathbb{Z}$ (with period n), so $x_{i+k\cdot n} = x_i$ for all $i,k \in \mathbb{Z}$. On this vector space we have a (Euclidean) norm $\|x\|$ which is induced by the scalar product $\langle x,y \rangle = \sum_{i=0}^{n-1} x_i \overline{y}_i$. Every linear operator $A : \mathbb{C}^n \to \mathbb{C}^m$ can be identified with a matrix vector multiplication

$$A(x) = A \cdot x = \left(\sum_{j=0}^{n-1} a_{i,j} x_j \right)_{i=0..m-1},$$

where $A = (a_{i,j})_{m,n}$ is an $m \times n$ matrix, symbolically $A \in M_{m,n}$. We will sometimes use the canonical basis elements $\delta_k = (0, \ldots, 0, 1, 0, \ldots, 0)$, where only one, the k-th entry is non-zero and equal to one. The notation A^T will signify the transpose of the matrix A.

1) Matrix Norms And Spaces: Approximating the inverse of a given matrix, we will need some measure of how ‘good’ the approximation is. To this end, we use some matrix norms, each of them having different advantages. We will introduce two new norms in Section III-A , but the following are well-known:

Definition 1: Let $A = (a_{i,j})_{m,n}$ be an m by n matrix, then

$$\|A\|_{Op} = \sup_{x \in \mathbb{C}^n : \|x\|=1} \{\|A \cdot x\|\}$$

is the *operator norm*. Also,

$$\|A\|_{fro} = \sqrt{\sum_{i=0}^{n-1} \sum_{j=0}^{m-1} \left|a_{i,j}\right|^2}$$

is the *Frobenius* or *Hilbert-Schmidt norm*.

A generalization of the Hilbert-Schmidt norm is the so-called *mixed-norm*, which is defined by

$$\|A\|_{p,q} = \left(\sum_{j=0}^{m-1} \left(\sum_{i=0}^{n-1} \left|a_{i,j}\right|^p \right)^{\frac{q}{p}} \right)^{\frac{1}{q}}$$

The definition above extends in a natural way to infinity as follows:

$$\|A\|_{\infty,q} = \left(\sum_{j=0}^{n-1} \left(\max_{i=0,\ldots,n-1} \left\{ \left|a_{i,j}\right| \right\} \right)^q \right)^{\frac{1}{q}}$$

The use of the operator norm is the natural way to measure the quality of an approximation as it satisfies $\|A \cdot x\| \leq \|A\|_{Op} \|x\|$, for all $x \in \mathbb{C}^n$. Another important application of this norm is the condition-number for invertible matrices, $\kappa(A) = \left\|A^{-1}\right\|_{Op} \cdot \|A\|_{Op}$, which measures the stability of a linear equation system. The problem with the operator norm is that its computation is very costly. It can be shown that the operator norm of a self-adjoint operator is equal to its largest eigenvalue, and the calculation of the eigenvalues of an operator is numerically very expensive, even if using elaborated methods [18]. Therefore, we will consider other norms, which are numerically easier to determine, see Definition 13.

The Frobenius norm can be defined by the Hilbert-Schmidt inner product, $\|A\|_{fro} = \sqrt{\langle A,A \rangle_{HS}}$, where for $A,B \in M_{m,n}$

$$\langle A,B \rangle_{HS} = \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} a_{i,j} \overline{b}_{i,j}$$

Together with this norm, the space of all $m \times n$ matrices $M_{m,n}$ forms a Hilbert space. This provides us with a number of

Hilbert space tools like orthonormal bases and the uniqueness of the best approximation on subspaces. The space $M_{m,n}$ is isomorphic to $\mathbb{C}^{m\cdot n}$ (for example by writing the columns one below each other) and the Hilbert-Schmidt inner product coincides with the ordinary scalar product on $\mathbb{C}^{m\cdot n}$.

2) Matrix Fourier Transformation: The notion of Fourier transformation can be easily extended to matrices (see [12] or [3]) as follows:

Definition 2: Let $A \in M_{m,n}$.The *Matrix Fourier Transformation (MFT)* of A, in symbols $\hat{A}$, is defined by

$$\hat{A} = F_m \circ A \circ F_n^*$$

where F_n is the *FFT-matrix*, $(F_n)_{k,l} = \frac{1}{\sqrt{n}} \cdot e^{-\frac{2\pi ikl}{n}}$.

3) Preconditioning: An alternative way to solve a linear system of equations $Ax = b$ consists in solving the system $PAx = Pb$ for a properly chosen matrix P. To this end, the matrix P should be chosen according to the following criteria:

1) P should be constructed within few operations,
2) P should be able to be stored in an efficient way,
3) $\kappa(PA) \ll \kappa(A)$ and
4) there is an algorithm for solving $PAx = Pb$ with better convergence properties.

The first two criteria are intended to keep the number of operations and memory requirements below those of the non-preconditioned system. The third criterion is intended to improve the numeric stability of the system. A sufficient condition for the third criterion is a clustered spectrum as $\kappa(A) = \frac{\sigma_n}{\sigma_1}$ where σ_n and σ_1 are the largest and smallest singular values, respectively. A clustered spectrum also yields a faster convergence (see [20], [21]).

B. Frames

Let us give short summary of frame theory on Hilbert spaces. For a more detailed presentation see [22], [23] or [24].

Definition 3: The sequence $\mathcal{G} = (g_k | k \in K) \subseteq \mathcal{H}$ is called a *frame* for the Hilbert space $\mathcal{H}$, with inner product $\langle .,. \rangle$, if constants $A, B > 0$ exist, such that

$$A \cdot \|f\|^2 \leq \sum_{k\in K} |\langle f, g_k \rangle|^2 \leq B \cdot \|f\|^2 \ \ \forall \ f \in \mathcal{H} \tag{2}$$

The constants A and B are called *lower frame bound* and *upper frame bound*, respectively. The best possible constants are called *the* frame bounds.

Definition 4: Let $\mathcal{G} = (g_k | k \in K)$ be a frame. The operator $S_{\mathcal{G}} : \mathcal{H} \to \mathcal{H}$ defined by

$$S_{\mathcal{G}}(f) = \sum_{k\in K} \langle f, g_k \rangle \cdot g_k \ \forall f \in \mathcal{H}$$

is called the *frame operator*.

For every frame, the frame operator is self-adjoint, positive and invertible. We will simply denote the frame operator by S, when there is no risk of ambiguity.

Proposition 5: Let $\mathcal{G} = (g_k)$ be a frame for $\mathcal{H}$ with frame bounds A, B. Then $\tilde{\mathcal{G}} = (\tilde{g}_k) := \left(S_{\mathcal{G}}^{-1} g_k \right)$ is a frame with frame bounds B^{-1}, $A^{-1} > 0$, the so-called *canonical dual frame*. Moreover, every $f \in \mathcal{H}$ has expansions

$$f = \sum_{k\in K} \left\langle f, S_{\mathcal{G}}^{-1} g_k \right\rangle g_k \text{ and } f = \sum_{k\in K} \langle f, g_k \rangle S_{\mathcal{G}}^{-1} g_k$$

where both sums converge unconditionally in $\mathcal{H}$, meaning that the convergence does not depend on the order of the elements $\{g_k\}$. The inverse of the frame operator $S_{\mathcal{G}}$ associated to a given frame (g_k) equals the frame operator associated to the dual frame, i.e., $S_{\mathcal{G}}^{-1} = S_{\tilde{\mathcal{G}}}$.

In the discrete, finite-dimensional case, $\mathcal{H} = \mathbb{C}^n$, a sequence is a frame if and only if it spans $\mathcal{H}$. In this case the optimal frame bounds equal the maximal and minimal eigenvalues of the frame operator, which are equal to $\frac{1}{\|S^{-1}\|_{Op}}$ and $\|S\|_{Op}$.

There is a number of algorithms for inverting the frame operator. A well known algorithm is the Neumann algorithm described in Figure 1 for obtaining $S^{-1}f$.

Initialization:

$$x_0 = f, h_0 = f, A = I - S$$

Iteration :

- Set $h_{k+1} = Ah_k$;
- Set $x_{k+1} = x_k + h_{k+1}$;
- Check exit condition $\|x_{k+1} - x_k\| < \varepsilon$.

Fig. 1. The Neumann algorithm

If in the Neumann algorithm a relaxation parameter $\lambda = \frac{2}{A+B}$ is used in front of S, it becomes the so-called *frame algorithm*. The relaxation parameter is needed for the convergence of the frame algorithm. Its calculation requires the computation of the frame bounds, which, as mentioned above, are numerically costly to compute. In order to deal with this drawback, for example the conjugate gradient algorithm was proposed (see [16] or [24]). The algorithm proposed in this article avoids this drawback, as well.

C. Gabor Analysis

We start with a summary of the main results on Gabor analysis. For a more detailed presentation see [8] or [23].

Recall that for any non-zero window function g and a signal f the STFT can be defined as $\mathcal{V}_g(f)(t,\omega) = \langle f, M_\omega T_t g \rangle$ using the translation operator $T_\tau f(z) = f(z - \tau)$ and the modulation operator $M_\omega f(t) = f(t) e^{2\pi i \omega t}$. In $L^2(\mathbb{R}^d)$, the space of square-integrable functions from $\mathbb{R}^d$ to $\mathbb{C}$, we have

$$\mathcal{V}_g(f)(t,\omega) = \int_{\mathbb{R}^d} f(x) \overline{g(x-t)} e^{-2\pi i \omega x} dx$$

Definition 6: For a non-zero function g (the *window*) and parameters $\alpha, \beta > 0$, the set of time-frequency shifts of g

$$\mathcal{G}(g, \alpha, \beta) = \{M_{\beta n} T_{\alpha k} g : k, n \in \mathbb{Z}^d\}$$

is called a *Gabor system*. Moreover, if $\mathcal{G}(g, \alpha, \beta)$ is a frame, it is called a *Gabor frame*. We will denote its frame operator by S_g.

Gabor frames are the appropriate tools for time-frequency analysis, as the analysis operator of a Gabor frame i.e. $f \mapsto \left(\left\langle f, T_{\alpha k} M_{\beta n} g \right\rangle \right)_{k,n}$, is just the STFT of f with window g, sampled at the time frequency points $(\alpha k, \beta n)$. The dual frame of a Gabor frame is a Gabor system again, which is generated by the *dual window* $\tilde{g} = S_g^{-1} g$ and by the same parameters α and β. We call the set of time frequency points $\left\{ (\alpha k, \beta n) \,\middle|\, k, n \in \mathbb{Z}^d \right\}$ the *lattice* of $\mathcal{G}(g, \alpha, \beta)$.

1) Discrete Gabor Expansion: For a good introduction to Gabor analysis in the discrete finite-dimensional case see [12].

From now on, we will consider the Hilbert space $\mathbb{C}^n$, and restrict the lattice parameter, now called a and b, to factors of n such that the numbers $\tilde{a} = \frac{n}{a}$ and $\tilde{b} = \frac{n}{b}$ are integers. In this case, the modulation and time shift operators are discretized, i.e., $T_l x = (x_{n-l}, x_{n-l+1}, \ldots, x_0, x_1, \ldots, x_{n-l-1})$ and $M_k x = \left(x_0 \cdot W_n^0, x_1 \cdot W_n^{1 \cdot k}, \ldots, x_{n-1} \cdot W_n^{(n-1)k} \right)$ with $W_n = e^{\frac{2\pi i}{n}}$. Recall that we are regarding all vectors as periodic, so the translation is a cyclic operator. We will consider the Gabor system $\mathcal{G}(g, a, b) = \left\{ M_{bl} T_{ak} g : k = 0, \ldots, \tilde{a} - 1; l = 0, \ldots, \tilde{b} - 1 \right\}$. Notice that this is equivalent to sampling with sampling period T and setting $\omega = \frac{k}{nT}$ and $\tau = l \cdot T$.

In the discrete, finite-dimensional case, the Gabor frame operator has a very special structure, the matrix S is zero except in every $\tilde{b}$-th side-diagonals and these side-diagonals are periodic with period a. This property can be directly seen by using the *Walnut representation* [25] of the Gabor frame matrix $S = (S_{p,q})_{n,n}$:

Theorem 7:

$$S_{p,q} = \begin{cases} \tilde{b} \sum\limits_{k=0}^{\tilde{a}-1} \overline{g}_{p-ak} \cdot g_{q-ak} & \text{for } p - q \equiv 0 \text{ mod } \tilde{b} \\ 0 & \textit{otherwise} \end{cases}$$

This means S can be represented as a special block matrix, both as a block circulant matrix and as matrix with diagonal blocks [3]. Matrices with this structure are called *Gabor-type* [26] or *Walnut* matrices. There is a smaller matrix describing S uniquely by using only the first a entries of the non-zero side-diagonals [26]:

Definition 8: Let $\mathcal{G}(g, a, b)$ be a Gabor frame, and S be its associated frame operator. Let $B = (B_{i,j})_{b,a}$ be the $b \times a$ matrix given by

$$B_{i,j} = S_{i \cdot \tilde{b} + j, j}$$

We call B a *'non-zero' block matrix* [26] or the *auto-correlation matrix* [23] of the Gabor system $\mathcal{G}(g, a, b)$.

The auto-correlation matrix B enjoys the following useful properties: S is diagonal if and only if B is zero except in the first row, and S is circulant if and only if the rows of B are constant. Combining Definition 8 and Theorem 7 the non-zero block matrix can be expressed as

$$B_{i,j} = \tilde{b} \sum_{k=0}^{\tilde{a}-1} \overline{g}_{i\tilde{b}+j-ak} \cdot g_{j-ak}. \tag{3}$$

Thus the reconstruction can be done using the non-zero block matrix:

$$(Sx)_j = \sum_{p=0}^{b-1} x_{j+p\tilde{b}} \cdot B_{p,(j \bmod a)} \tag{4}$$

It is possible to realize the multiplication of two Gabor matrices by using only 'non-zero' block matrices [26]. This leads to a very efficient algorithm with $O(a\ b\ log(b))$ operations [27].

In addition to the auto-correlation matrix defined above, there is another "small" ($b \times a$) matrix, which fully describes the frame matrix S [19]:

Definition 9: The *Janssen* -matrix of S is the $a \times b$ matrix $J = (J_{k,l})_{a,b}$, given by

$$J_{k,l} = \frac{n}{a \cdot b} \cdot \mathcal{V}_g(\gamma) \left(l\tilde{b}, k\tilde{a} \right).$$

As the the set of time-frequency shift matrices, $\left\{ M_i \cdot T_j \,|\, i, j = 0, \ldots, n-1 \right\}$ form an orthogonal system in $M_{n,n}$ with the Hilbert-Schmidt inner product, it can be shown [15], that the frame matrix S can be represented by the following expansion:

Definition 10: We call

$$S_{g,\gamma} = \sum_{k=0}^{a-1} \sum_{l=0}^{b-1} J_{k,l} M_{k\tilde{a}} T_{l\tilde{b}} \tag{5}$$

the *Janssen-representation* of S.

2) Higher Dimensional Approach: In this paper, we will mostly use one-dimensional spaces. In Section VI-E the method proposed in this paper is applied to images and therefore to higher dimensional data. For $\mathbb{R}^m$ we will only use separable windows g, which means that there exist one-dimensional functions g_i, $i = i, \ldots, m$, such that

$$g = g_1 \otimes g_2 \otimes \ldots \otimes g_m,$$

which is the short notation for

$$g(x_1, x_2, \ldots, x_m) = g_1(x_1) \cdot g_2(x_2) \cdot \ldots \cdot g_m(x_m).$$

In this case, following [14], the frame matrix of g is the Kronecker product (cf. [12]) of the frame matrices of g_i, i.e. $S_g = S_{g_1} \otimes S_{g_2} \otimes \ldots \otimes S_{g_m}$. Thus everything can be reduced to a Gabor expansion in a one-dimensional space.

III. Properties of the Janssen and non-zero block matrix

In this section the non-zero block matrix and the Janssen matrix are investigated in more detail. We start with a result, that shows, that these matrices are connected:

Corollary 11: For a given matrix the following properties are equivalent

1) having a Walnut representation, as in Theorem 7,
2) commuting with all M_{ka} and T_{lb} and
3) being represented by a Janssen matrix, as in Definition 5.

The proof is straightforward [3].

Even more, the Janssen matrix and the non-zero block matrix are connected by the following algorithm:

Theorem 12: Let $B_{g,\gamma,a,b}$ be the $b \times a$ associated non-zero block matrix of g, γ, a, b, and $J_{g,\gamma,a,b}$ be the corresponding Janssen-matrix. Then $F_a \cdot B^T_{g,\gamma,a,b} = a \cdot J_{g,\gamma,a,b}$.

For a proof see appendix A Theorem 21.

A. Norms

For many algorithms for the inversion of the frame matrix, it is important to have a good estimate for the upper frame bound B or equivalently for $\|S\|_{Op}$, which can be calculated efficiently. An example for such an algorithms is presented e.g. in [19]. We have seen two types of "small" matrices (with $b \cdot a$ elements) which characterize Gabor matrices. We can now define a norm for each type, which gives us an estimation of $\|S\|_{Op}$. We will use these norms for the numerical experiments in Section VI to get a numerical efficient way of measuring the quality of an approximate inverse .

Definition 13: Let S be a Gabor-type matrix, B be its non-zero block matrix and J be its Janssen matrix. Then, we define

$$\|S\|_{Wal} = \left\|B^T\right\|_{\infty,1} \text{ and } \|S\|_{Jan} = \|J\|_{1,1},$$

the so called *Walnut* respectively *Janssen* norm of S.

As the matrices B and J are smaller than the Gabor matrix S, the computation of the norms above is relatively simple. More precisely, both can be calculated from the block matrix (see Theorem 12).

It can be shown (see appendix A) that the norms above are bounds for the operator norm, and that in the Gabor frame case they can be ordered as follows:

$$\|S\|_{Op} \leq \|S\|_{Wal} \leq \|S\|_{Jan} \leq \|S\|_{fro} \tag{6}$$

This means that the Walnut norm is the best approximation of the operator norm, and therefore it can be used as an efficient way to find a (close) upper bound for it.

On the other hand, the Janssen matrix and norm give us some insight on the behavior in the time-frequency plane. For example in the case of approximation it tells us where in the time-frequency plane the coefficients of the difference between S and the identity are big, see Section VI-D for an example. In the numerical part of the paper, i.e. Section VI, all the algorithms use the block structure of the frame matrix. In that section, the Walnut and Janssen norms are very convenient as they can be calculated directly from the block matrix.

Regarding the Frobenius norm or equivalently the Frobenius inner product, $M_{m,n}$ forms a Hilbert space. Although it is not a very close approximation for the operator norm, as can be seen in VI-B, the Hilbert space property is very useful from an analytic point of view.

In summary, each one of the norms introduced above has its usefulness. As we will work with finite-dimensional spaces, all norms have to be equivalent, see Appendix A.

IV. Single Preconditioning Of The Gabor Frame Operator

We propose two preconditioning methods. In the first method we consider the best approximation of S with diagonal matrices, and approximate S^{-1} by inverting its diagonal approximation. The second method is based on the same idea but considering circulant matrices.

A. Diagonal Matrices

As a preconditioning matrix the inverse of the diagonal part of the frame operator is used. For every square matrix A we can find a diagonal matrix just by cutting out the diagonal part of A:

Definition 14: Let $A = (a_{i,j})_{n,n}$ be a square $n \times n$ matrix, then let $D(A) = (d_{i,j})_{n,n}$ with

$$d_{i,j} = \begin{cases} a_{i,i} & i = j \\ 0 & otherwise \end{cases}$$

the *diagonal part of A*.

The set of all diagonal $n \times n$ matrices is spanned by the matrices E_k with $E_k = D(\delta_k)$. They clearly form an orthonormal basis (ONB) (with the Frobenius inner product) and therefore $D : A \mapsto D(A)$ is an orthogonal projection. This means that the best approximation of A in $\|.\|_{HS}$ by diagonal matrices is exactly $D(A)$.

The diagonal part of a Gabor-type matrix is clearly block-circulant, and therefore also of Gabor-type. This allows us to use the efficient block-matrix algorithms from [26].

If the window g is compactly supported on an interval with length smaller than $\tilde{b}$ then $S_{g,g}$ is a diagonal matrix, see [12]. In this case the inverse matrix is very easy to calculate, by just taking the reciprocal value of the diagonal entries, which are always non-zero for a Gabor frame matrix [14]. Even in the case where the window g is not compactly supported, but S is strictly diagonal dominant, then S^{-1} is well approximated by D^{-1}. It is known [17] that, if the matrix A is *strictly diagonal dominant*, i.e. $\max\limits_{i=0,\ldots,n-1} \sum\limits_{k=0,k\neq i}^{n-1} \frac{|a_{ik}|}{|a_{ii}|} < 1$, then the *Jacobi algorithm*, $x_m = D^{-1}(D - A)x_{m-1} + D^{-1}b$, converges for every starting vector x_0 to $A^{-1}b$. The efficiency of the Jacobi algorithm follows from the fact that it is easy to find the diagonal part of a matrix and to invert it. As can be seen from the above formula the Jacobi algorithm is equivalent to preconditioning with $D(S)^{-1}$.

$$\boxed{P = D(S)^{-1}}$$

Fig. 2. The diagonal preconditioning matrix

The use of block-matrices leads to very efficient algorithms. Motivated by this fact, we would like to find criteria for the convergence of the Jacobi algorithm for non-zero block matrices, which means that by just using the diagonal preconditioning matrix and an iterative scheme we will get the inverse matrix respectively the canonical dual window.

Corollary 15: Let S be a Gabor-type matrix and B be the associated non-zero block matrix. Then the following conditions are sufficient for the Jacobi-algorithm to converge

1) $\max\limits_{i=0,\ldots,a-1} \left\{ \sum\limits_{k=1}^{b-1} \frac{\left|B_{k,(i-k\cdot\tilde{b}) \bmod a}\right|}{B_{0,i}} \right\} < 1$

2) $\max\limits_{k=0,\ldots,a-1} \left\{ \sum\limits_{i=1}^{b-1} \frac{\left|B_{i,k}\right|}{B_{0,(k+i\cdot\tilde{b}) \bmod a}} \right\} < 1$

3) $\sum\limits_{i=0}^{a-1} \sum\limits_{k=1}^{b-1} \left(\frac{B_{k,(i+k\cdot\tilde{b}) \bmod a}}{B_{0,i}} \right)^2 < \frac{a}{n}$

4) $\sum_{k=0}^{a-1}\sum_{i=1}^{b-1}\left(\frac{B_{i,k}}{B_{0,(k+i\cdot\tilde{b})\bmod a}}\right)^2 < \frac{a}{n}$

Notice that the first column of B is always positive, as the diagonal of the Gabor frame operator has this property for frames [15]. This property can be shown by translating the result for the full matrix to the block matrix and by properly manipulating the indices, using the property $j = \left\lfloor \frac{j}{a} \right\rfloor \cdot a + (j \bmod a)$, refer to [3].

B. Excursus: A criteria for Gabor frames

The question, whether a given set of parameters (g,a,b) generate a Gabor frame, is not directly connected to our main question of finding an approximate dual window. But Corollary 15 provides an immediate criterion for Gabor frames.

Corollary 16: Sufficient conditions for a Gabor triple (g,a,b) to generate a Gabor frame are:

1) $\left|\sum_{k=0}^{\tilde{a}-1}\overline{g}(i-ak)\,g\left(i-j\tilde{b}-ak\right)\right| < \frac{1}{b-1}\sum_{k=0}^{\tilde{a}-1}\left|g(i-ak)\right|$ for all $i=0,\ldots,a-1$ and $j=1,\ldots,b-1$.

2) $\left|\sum_{k=0}^{\tilde{a}-1}\overline{g}\left(j-i\tilde{b}-ak\right)g(j-ak)\right| < \frac{1}{b-1}\sum_{k=0}^{\tilde{a}-1}\left|g\left(j+i\tilde{b}-ak\right)\right|$ for all $j=0,\ldots,a-1$ and $i=1,\ldots,b-1$.

The result above can be easily shown by combining Equation 3 with Corollary 15, cf. [3]. A similar result was stated in a corollary in [14], which is emended and expanded here.

This result, as a direct and immediate consequence of main results of this paper, gives sufficient conditions on the window g and the lattice parameters a,b to constitute a Gabor frame. Moreover, note that this corollary provides a highly efficient numerical method to verify the frame condition.

C. Circulant Matrices

Instead of considering diagonal matrices we can approximate S by projecting on the algebra of circulant matrices. This means we are using the mean value of the side-diagonals to define $C(S)$ as follows:

Definition 17: Let $C(S) = (c_{i,j})_{i,j}$ with

$$c_{i,j} = \frac{1}{n}\sum_{k=0}^{n-1} S_{k+(j-i),k}.$$

The two classes of matrices we have investigated so far are connected as follows [3]:

Theorem 18: For a circulant matrix M the matrix $\hat{M}$ is diagonal and vice versa.

Due to properties of the Matrix Fourier Transform [3] the result above means that $S \mapsto C(S)$ is again a projection, and it can be calculated by using

$$C(S) = F_n \cdot [D(F_n \cdot S \cdot F_n^*)] \cdot F_n^*$$

which implies that

$$C(S)^{-1} = F_n^* \cdot [D(F_n \cdot S \cdot F_n^*)]^{-1} \cdot F_n$$

Therefore the computation of $C(S)^{-1}$ can be done in a very efficient way by using the FFT-algorithm. Analogue to Section IV-A this can be used as preconditioning matrix.

$$P = C(S)^{-1}$$

Fig. 3. The circulant preconditioning matrix

V. Double Preconditioning Of The Gabor Frame Operator

The main result of this work is the double-preconditioning method. In a rather natural way, we will combine the two single preconditioning methods introduced above. More precisely, after an approximation with diagonal matrices and inversion we do an approximation with circulant matrices. The double preconditioning algorithm can be implemented very efficiently using the block multiplication algorithm of [14], since, if S is a Gabor-type matrix, then $C(S)$ and $D(S)$ are also Gabor-type matrices and hence can also be represented by $b \times a$ block matrices.

$$P = C\left(D(S)^{-1} \cdot S\right)^{-1} D(S)^{-1}$$

Fig. 4. The double preconditioning matrix

For a basic description of the algorithm see figure 5. In this figure the subscript *'block'* indicates a calculation on the block matrix level, which makes this algorithm very efficient. The expressions $diag_{block}(M)$, $circ_{block}(M)$, $inv_{block}(M)$ and $block(g,a,b)$ stand for the calculation of the block matrix of $D(M)$, $C(M)$, M^{-1} and $S_{g,a,b}$ respectively. The matrix multiplication on block matrix level is signified by $\bullet_{block}$.

- Parameter:
  the window g, the lattice parameters a,b
- Initialization:
  $$B = block(g,a,b)$$
- Preconditioning :
  - (first preconditioning)
    $$P_1 = inv_{block}\left(diag_{block}(B)\right)$$
    $$S_1 = P_1 \bullet_{block} B$$
  - (second preconditioning)
    $$P_2 = inv_{block}\left(circ_{block}(S_1)\right)$$
    $$S_2 = P_2 \bullet_{block} S_1$$

Fig. 5. The double preconditioning algorithm

In the following two sections we will look at two special properties of our algorithm, namely how to do the second preconditioning step and in which order to multiply the matrices. We will give reasons, why we have chosen this particular setting.

A. Choice of method

Roughly speaking, the double preconditioning method consists in two single preconditioning steps. To this end, there are two possibilities, namely, to use the original matrix S for every step or to use the result of the first step in the second one. More precisely:

(Method 1) $C\left(D(S)^{-1}S\right)^{-1}D(S)^{-1}$, or more naively
(Method 2) $C(S)^{-1}D(S)^{-1}$

The first method seems to be more sensible, as each single preconditioning step uses projections. Even more, it also enjoys the following property: if S is diagonal, after the first step we will reach identity and this will stay identity in the second step (up to the machine precision). Also, if S is circulant, after the first step we still have a circulant matrix as the multiplication of an arbitrary matrix A and a diagonal matrix D is $D \cdot A = (d_{i,i} \cdot a_{i,j})_{i,j}$. So for the circulant Matrix $C = (c(i-j))_{i,j}$ we get $\left(D^{-1}(C) \cdot C\right)_{i,j} = \left(c(0)^{-1} \cdot c(i-j)\right)_{i,j}$. Hence the second step leads to identity again. Note that the Gabor-type structure is also preserved with this method.

On the other hand, the second method does not enjoy aforementioned property in the case of circulant matrices. For example take $n=6$, $a=1$,$b=6$ and $g=(1,2,3,4,5,6)$. Then S is a circulant matrix, but the double preconditioning deteriorate the approximation, as $\left\|C(S)^{-1}D(S)^{-1}S - I\right\|_{Wal} = 0.994505$. This is a big disadvantage, since, for these simple matrices, the method should give satisfactory results.

So we always use the first method. In order to simplify the notation we will use $\mathcal{C}(S)$ to denote $C(D(S)^{-1}S)$.

B. Order of preconditioning matrices

If the preconditioning matrix is diagonal, it makes no difference if it multiplies S from the left or from the right. This is because, as S is self-adjoint (see II-B), $D(S)^{-1}$ is too, and therefore, $\left(D(S)^{-1}\cdot S\right)^{*} = S \cdot D(S)^{-1}$ and $\left(D(S)^{-1}\cdot S - I\right)^{*} = S \cdot D(S)^{-1} - I$. Finally

$$\left\|D(S)^{-1}\cdot S - I\right\| = \left\|S \cdot D(S)^{-1} - I\right\|$$

So the norm of the difference to the identity is equal for

1) $D(S)^{-1}S$ or
2) $SD(S)^{-1}$

The same property holds for single preconditioning with circulant matrices.

In the case of double-preconditioning, the influence of the order in the multiplication has still to be investigated. Numerical experiments (see VI-B) suggests that also for the double preconditioning method the order is not of relevant importance. In this paper, unless specified otherwise, the order $\mathcal{C}(S)^{-1}D(S)^{-1}S$ will always be used.

C. Algorithm for an approximate dual

The double preconditioning method has two applications:

1) It can be used to speed up the convergence of an iterative scheme, here the Neumann algorithm, using S_2 in Figure 5 to get the canonical dual (up to a certain, predetermined error).
2) In order to get a real fast algorithm for the calculation of an approximate dual we propose the following method: The double preconditioning matrix itself, P_2 in Figure 5, is used as an approximation of the inverse Gabor frame operator. Then the approximated dual can be calculated as
$$\widetilde{g}^{(ap)} = \left(\mathcal{C}(S)^{-1}D(S)^{-1}\right)g \ .$$
This can be used for example for adaptive Gabor frames in real time, where the computation of the canonical dual window needs to be done repeatedly.

VI. Numerical Results

A. The shapes of the approximated duals

In this first, introductory example we will use the double preconditioning matrix to get an approximate dual, as mentioned in Section V-C (2), to see

1) that the different single preconditioning steps can capture certain properties of the dual window but fail to do so for others
2) the double preconditioning lead to a good approximation of the dual.

This experiment was done with a Gaussian window with $n=640$, $a=20$ and $b=20$. In this case it is interesting to see the difference between the diagonal and the circulant 'dual' windows. We will use the names *diagonal dual*, *circulant dual* and *double dual* for the window we get when we apply the preconditioning matrix to the original window. Note that this should not imply that these windows are true duals. See Figure 6.

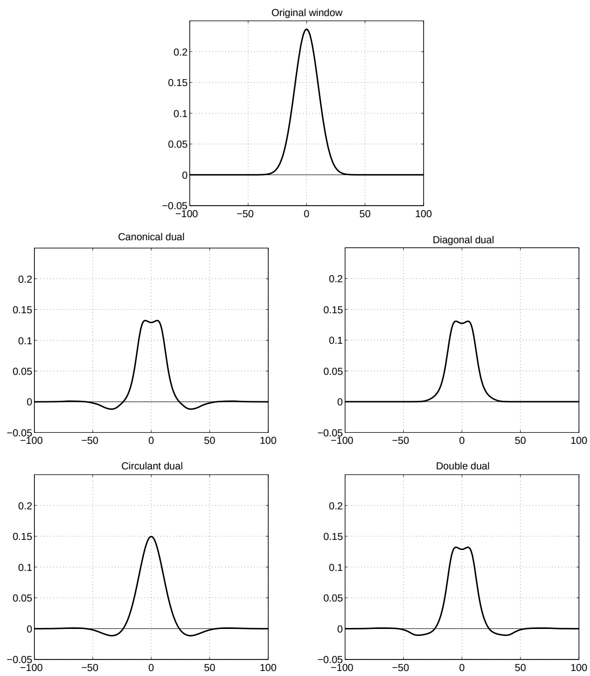

Fig. 6. *Windows*: Top: the full original window; Mid left: the true canonical dual window, Mid right: 'diagonal dual', bottom left: 'circulant dual', bottom right: 'double dual'.

The diagonal dual seems similar to the canonical dual 'away from the center' but not near the center, while the circulant dual just has the opposite property. Opposed to these 'single duals' the 'double dual' seems to combine these properties to become very similar to the true dual everywhere.

B. Order

We are now investigating if the order has any influence. In this case we use a Gaussian window, $n = 144$, $a = 6$ and $b = 9$ and we look at the norms of the difference to identity:

method \ norm	Operator	Walnut	Janssen	Frobenius
$D^{-1}S$	0.1226	0.1232	0.1234	1.0397
SD^{-1}	0.1226	0.1232	0.1234	1.0397
$C^{-1}S$	0.0038	0.0045	0.0046	0.0324
SC^{-1}	0.0038	0.0045	0.0046	0.0324
$\mathcal{C}^{-1}D^{-1}S$	0.0006	0.0007	0.0008	0.0048
$\mathcal{D}^{-1}C^{-1}S$	0.0006	0.0007	0.0009	0.0048
$\mathcal{C}^{-1}SD^{-1}$	0.0006	0.0007	0.0009	0.0048
$\mathcal{D}^{-1}SC^{-1}$	0.0006	0.0008	0.0009	0.0048
$SC^{-1}\mathcal{D}^{-1}$	0.0006	0.0007	0.0009	0.0048
$SD^{-1}\mathcal{C}^{-1}$	0.0006	0.0007	0.0008	0.0048

We see in this case that the order is irrelevant. Also other experiments lead the authors to believe, that the order is not relevant. This has to be investigated further. In this experiment we also see a nice example of the norm inequality (6).

C. Iteration

Instead of using the preconditioning matrix as approximation of the inverse, we can iterate this scheme using the Neumann algorithm.

Let us look at an example with a Gaussian window, $n = 1440$, $a = 32$ and $b = 30$. See Figure 7. We look at the preconditioning steps, the frame algorithm with optimal relaxation parameter and a conjugate gradient method. The costly calculation of the frame bound for the frame algorithm was done beforehand, which has to be taken into account, when judging this algorithm.

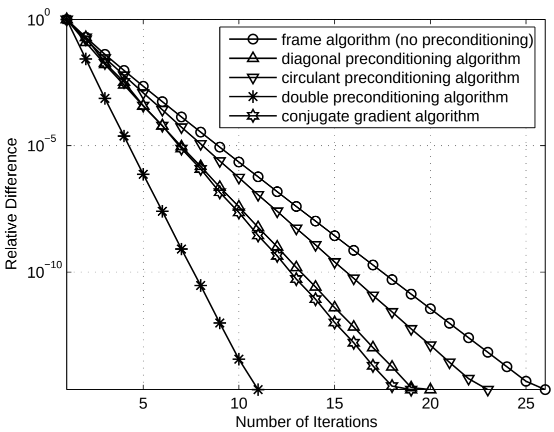

Fig. 7. *Convergence with iteration*: Relative difference of iteration steps (Gaussian window, $n = 1440$, $a = 32$ and $b = 30$.)

In this figure we see that the circulant preconditioning step is only a little bit better than the frame algorithm, iteration-wise. As the time sampling is not very small, this could be expected. Diagonal preconditioning is better, but not a lot, because the lattice parameters are quite similar to each other. The double preconditioning brings a big improvement compared to the single preconditioning methods. It can also be seen that the conjugate gradient algorithm, a method with guaranteed convergence [18], performs much worse than double preconditioning.

Generally our experiments showed that for increasing a the circulant preconditioning gets worse and for increasing b the diagonal preconditioning gets worse, as expected theoretically. Double preconditioning is not affected by these deteriorations.

D. The Janssen representation

To see what happens in the time-frequency plane let us look at the Janssen coefficients of the involved matrices. See Figure 8, where we have used a Gaussian window with $n = 144$, $a = 12$ and $b = 9$.

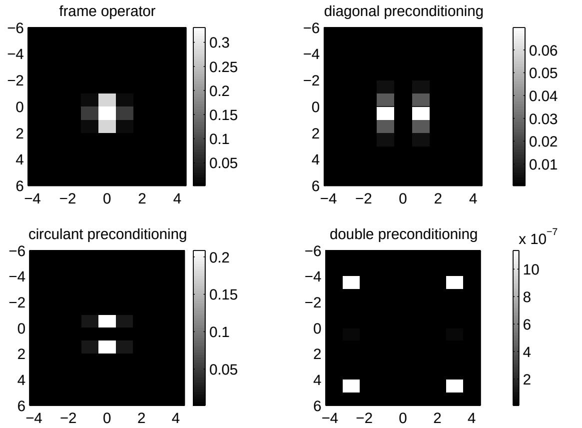

Fig. 8. Time-frequency spread of differences to identity (Centered graphs)

In the top left picture we see the time-frequency spread of the difference of identity and the original frame operator, $I - S$. It is clearly neither diagonal nor circulant, as diagonal matrices, which are linear combinations of modulations, would only have non-zero coefficients in the first row, whereas circulant matrices are non-zero only in the first column. Note that Figure 8 shows centered graphs, i.e. the first row and the first column are represented at the center of the graph.

In the top right picture we see $I - D^{-1}S$ in the Janssen representation. The first column is zero, as the diagonal part was canceled out, but some other parts remain. An analogue property is valid for the circulant preconditioning.

For the double preconditioning method we see that in this case the Janssen norm would be very small. And we see that the coefficients around the center, ‘near the diagonal and circulant case’ were approximated well. The error occurs ‘far from the center’.

So the Janssen representation gives us some insight on where in the time-frequency plane the coefficients of the difference to identity is high. As the Janssen norm just sums up the absolute value of these coefficients and is an upper bound for the operator norm, these gives us some insight on where the error of the approximation happens in the time-frequency plane.

E. Higher Dimensional Double Preconditioning

For this 2D example, see Figure 9, we use a separable window, the tensor product $g \otimes g$. We use a Gaussian 1D window g with $n = 576, a = 18, b = 24$. Here we don't get perfect reconstruction, but the reconstruction with the double dual is clearly much better than with the other two approximate duals. This can also be seen in the norm of the difference $\left\| Id - P^{-1} \cdot S \right\|_{Op} = 0.1796, 0.0914, 0.0300$. for the diagonal, circulant and double preconditioning case respectively.

For the calculation of the canonical dual with an iterative scheme $0.441s$ was needed, the 'double dual' needed only $0.060s$ on a MS Windows workstation with a Pentium III (937 MHz). For higher dimensional applications the size of the data is in general much bigger, so that numerically efficient methods, like the one presented here, become even more important.

Fig. 9. *2D Reconstruction*: Top left: the original image, top right: reconstruction with 'diagonal dual', bottom left: with 'circulant dual', bottom right: with 'double dual'.

F. Tests with Hanning window

For this experiment a zero-padded Hanning window was used as window. The length n of the signal space was chosen randomly between 1 and 1000. Out of all divisors of n the length of the Hanning window w_{supp} was chosen, as well as a and b. Because we are interested in Gabor frames, we have restricted our parameters to $a \leq w_{supp}$ and $a \cdot b \leq n$. From now on, let us use the terms *Diagonal Norm* for $\left\| D^{-1}S - Id \right\|$, *Circulant Norm* for $\left\| C^{-1}S - Id \right\|$, calling both *Single Norms*, and *Double Norm* for $\left\| D^{-1}C^{-1}S - Id \right\|$. The choice of the particular norm we use depends on the context, here in this experiment the operator norm has been used, as this experiment was intended as short and introductory, so no attention has been given to numerical efficiency.

Diagonal Norm	Circulant Norm	Double Norm	n	a	b	w_{supp}
0	1,81E-10	0	922	2	2	461
1,11E-16	1,59	1,11E-16	194	97	2	97
0,01	0,12	0,01	256	16	8	36
0,34	1,74E-06	1,74E-06	892	2	4	446
0,44	0,34	0,29	144	6	18	12
0,6	0,41	0,34	1000	40	20	100
0,65	0,56	0,52	210	5	35	10
0,94	1,29	1,36	16	4	4	8
0,96	0,97	0,94	770	14	55	22
4,39	74,52	74,65	936	36	24	312

TABLE I

COMPARISON OF THE PRECONDITIONING METHODS FOR HANNING WINDOWS WITH DIFFERENT LENGTH AND DIFFERENT LATTICE PARAMETERS

As illustration some results can be found in Table I. They are ordered by the diagonal norm. When the diagonal norm is small, the double preconditioning norm is also small. When the diagonal norm is somewhere well between 0 and 1, we see that the double norm is always smaller. Around 1 again the difference is not very big normally. Above 1, when no iterative preconditioning algorithm is converging any more, the difference might be big. This is the case, when parameters are used, such that the matrix S does not have a lot of structure and the sparsity can not be exploited anymore.

We have collected statistical data and randomized the parameters a 1000 times. The distribution of the cases can be found in Figure 10.

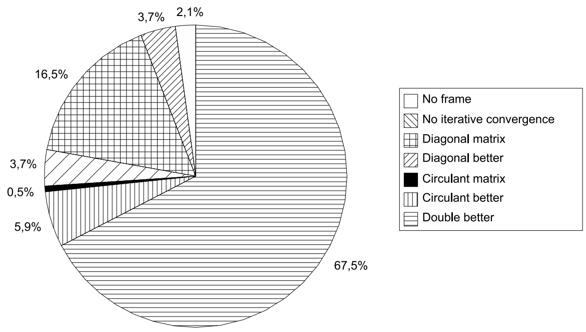

Fig. 10. *Experiment with Hanning window*: Distribution of cases.

The double preconditioning method is in 67.5% of the cases preferable to the single preconditioning methods, measured by the norm of the distance to identity of the preconditioned matrix, $\left\| I - P^{-1}S \right\|$. The remaining 32.50% are distributed as follows: The matrix is either already diagonal (16.5%) or the 'diagonal norm' is smaller (3.7%). In 0.5% the matrix was circulant and in 5.9% the 'circulant norm' was smaller. The cases remain, where the parameters do not constitute a frame (2.1%) or where all norms are larger than 1 (3.7%). As will be seen in VI-G this property does depend on the shape of the window, but even more on the lattice parameters a and b, most notably how small b is, as well as the chosen settings of the experiment. If the above mentioned quality criteria are used to measure significant difference, in this experiment only in 0.1% of the cases one of the single preconditioning methods was 'essentially' preferable, i.e. there was more than 10% difference. In these experiments we could also observe that if the norms of the both single preconditioning cases are around

1, the norm of the error of the double preconditioning method is also around 1.

G. Systematic experiments

In order to verify that the double preconditioning method is not only highly efficient for very special cases (such as the Gaussian), but for most windows typically used in Gabor analysis we have carried out systematic investigations. We have used several different windows (Gauss, Hanning, Hamming, Kaiser-Bessel, Blackman Harris, Rectangle and even noise), with various zero-padding factors and have used random signal length up to 1000 samples and also random lattice parameters a, b with $a \cdot b | n$ and the support of the window $w_{supp} \geq a$.

Here we have tried to minimize the cases, where the matrix is diagonal because of the lattice parameters (if $w_{supp} \leq n/b$). In this case it would still be possible to use the double preconditioning, we would only lose precision due to calculation and round-off errors, and the calculation is a bit slower as the double preconditioning is more complex. The conditions on the lattice parameters and the support of the window mean, of course, that we get a certain bias into our statistical investigation. For us this, nevertheless, seemed to be the most interesting situation.

The complexity of the algorithm and these tests have been further decreased by staying completely at the block matrix level, doing all calculation with the efficient block algorithms and using the Walnut norm.

We have summarized the results in Table II. For each window the experiments have been repeated 20000 times, so overall in the following table $120'000$ random parameters are used.

In the rows we see the following percentages

1) the Gabor system was no frame.
2) none of the preconditioning iteration schemes would converge, i.e. none of the norms was smaller than 1.
3) the diagonal norm was smaller than the double norm, wherein
    3') the frame matrix was already diagonal (and so both methods were essentially equal).
4) the circulant norm was smaller than the double norm, wherein
    4') the frame matrix was already circulant (and so both methods were essentially equal).
5) the double norm was larger than 1, the best single norm was smaller than 0.9.
6) the double norm was essentially larger (factor :10) than the best of the single norms.
7) the double norm was essentially smaller (factor :10) than the best of the single norms.
8) The double preconditioning method is better or essentially equal if the system is a frame. We sum up the cases, where the double preconditioning norm is smaller and the matrices are already diagonal or circulant (because then the difference is only due to calculation errors).
9) The double preconditioning method is better or essentially equal if any of the iterative scheme works.

	Han	Ham	Bla	Kai	Gau	Noi
1)	0.00 %	0.00 %	0.39 %	0.00 %	3.66 %	0.00 %
2)	36.42 %	37.36 %	29.80 %	34.05 %	22.42 %	56.00 %
3)	28.53 %	28.53 %	29.80 %	28.62 %	30.17 %	27.87 %
3')	28.50 %	28.53%	29.39 %	28.46 %	30.17 %	27.80 %
4)	13.52 %	12.30%	9.58 %	16.46 %	2.41 %	1.73 %
4')	0.00 %	0.49%	9.19 %	0.74 %	0.11 %	0.00 %
5)	0.12 %	0.50%	1.02 %	0.00 %	0.02 %	00.12 %
6)	0.00 %	0.00 %	0.00 %	0.00 %	0.14 %	0.00 %
7)	0.00 %	0.11%	0.04 %	0.08 %	7.77 %	0.00 %
8)	49.83 %	50.83%	60.45 %	50.07 %	74.34 %	49.83 %
9)	78.37 %	81.15%	86.25 %	75.92 %	96.89 %	78.37 %

TABLE II

SYSTEMATIC TESTS: (HAN)NING, (HAM)MING, (BLA)CKMAN-HARRIS, (KAI)SER-BESSEL ($\beta = 6$), (GAU)SS AND (NOI)SE

Nearly in all cases these windows form a frame. A prominent exception is the Gaussian window, which is due to the used zero-padding. About the same percentage for all windows did not allow any of the preconditioning iterative algorithm to converge, an exception being the Blackman-Harris with a somewhat low percentage, the Gaussian with a very low percentage and the noise window with a very high percentage. This leads us to the statement that the preconditioning algorithm works better for 'nice' windows.

For the windows tested it appears that the percentage of diagonal matrices is comparable, even in the case of a noise window. This is partly due to the particular properties of the chosen experiment. The percentages for circulant matrices respectively for convergence of the circulant preconditioning method seem to be quite different for different windows.

There are very few cases, where a single preconditioning algorithm would converge, but the double preconditioning would not. A detailed investigation shows that this happens only in cases, where also the 'single norms', i.e. the norm of the deviation from identity (see Section VI-F) are high, near to one. For the cases when all relevant norms are smaller than one, we see that for the Gaussian window we have only a very small chance that the best single preconditioning method is essentially better than the double preconditioning method, but a rather high probability for the opposite. For all other windows the chance for an essential improvement using the double preconditioning method is not very big, but there is no chance for a deterioration. Note that here the double preconditioning method still has an advantage, as it can be more easily used as 'default' method than the single preconditioning methods as seen in 9) in Table II.

Overall we see that with all windows the double preconditioning algorithm works in about half of the cases, if we have a frame. And it works in about 80 percent of the cases, when any of the preconditioning would work, with the notable exception of the Gaussian window, where it works nearly always. The Hanning and Hamming window are quite similar but contrary to common belief, they are not very similar to the Gaussian window. We see that the behavior for the double and single preconditioning method depends heavily on the chosen window and so the connection of analytical properties of the windows with the efficiency of the preconditioning methods should be investigated. One can expect some connection as can be seen from the behavior of the Gaussian on one side and noise on the other side.

VII. Conclusion and perspectives

We have presented a new method for finding an approximate inverse for a typical Gabor frame operator respectively very good approximate dual windows at very low computational costs. We have introduced a fast algorithm using existing block matrix methods. The method was constructed so that, diagonal and circulant matrices are approximated perfectly (up to precision). We have shown that this method is very often preferable to other iterative schemes. For 'nice' windows and lattice parameters the first approximation, the preconditioning matrix, is already a good approximation of the inverse frame matrix.

We have shown a close connection of the non-zero block matrix and the Janssen matrix and have introduced corresponding norms. We have shown the connection between the norms and why they can be useful in different situations.

For the single preconditioning case we have found sufficient conditions for the window, when this algorithm would converge and therefore the Gabor system would form a frame. We have also found conditions on the non-zero block matrix for convergence of the Jacobi algorithm. The condition for the window is not very intuitive, but as the block matrix can be established quickly this check can be done conveniently.

An important motivation for this paper are investigations, which can not be described in detail within the given length of the paper, or are still speculative. These issues will have to be investigated more in the future. For example it would be desirable to have some simple sufficient conditions which guarantee that the Jacobi algorithm is convergent. Section VI-G gives reasons to believe that an investigation of the analytic properties of a window and the connection to its 'preconditioning behavior' is fruitful. Furthermore the idea can be extended by using preconditioning matrices using other commutative subgroups of the time-frequency plane than the translations and modulations.

We believe that these algorithms can be very useful in situations, where the calculation of the inverse frame operator or of the dual window is very expensive or cannot be done at all. For example in the situation of *quilted Gabor frames* [28] or the *Time-Frequency Jigsaw Puzzle* [5], globally we might have a frame, but certainly not a Gabor frame. So we can not find a dual Gabor window globally, but the dual frame can be approximated by the dual windows of the local Gabor frame. In these cases it might be preferable to use a good and fast approximation of the local Gabor dual windows instead of using a precise calculation of the local canonical dual, as precision will be lost at the approximation of the global dual frame.

It is also highly believed by the authors, that these algorithm will have an even bigger importance for higher dimensional data. Similar algorithms will be applicable in the cases of multi-dimensional non-separable Gabor systems, where there exists only a few efficient algorithms. This will allow a lot of applications of these or similar algorithms.

Appendix

A. Proof of the norm equivalences

For a better overview we will split the results in several lemmas and propositions.

Lemma 19:

$$\|S\|_{Op} \leq \|S\|_{Jan} \qquad \|S\|_{Wal} \leq \|S\|_{Jan}$$

$$\|S\|_{fro} \leq \sqrt{n}\,\|S\|_{Jan}$$

Proof: We know from (5) that we can represent the frame operator as sum of the time and frequency shifts and so for every norm

$$\left\|S_{g,\gamma}\right\| = \frac{n}{a\cdot b}\left\|\sum_{l=0}^{b-1}\sum_{k=0}^{a-1} c_{l,k} M_{l\tilde{a}} T_{k\tilde{b}}\right\| \leq$$

$$\leq \frac{n}{a\cdot b}\sum_{l=0}^{b-1}\sum_{k=0}^{a-1}\left|c_{l,k}\right| \|M_{l\tilde{a}} T_{k\tilde{b}}\| = \sum_{l=0}^{b-1}\sum_{k=0}^{a-1}\left|J_{l,k}\right| \|M_{l\tilde{a}} T_{k\tilde{b}}\|$$

Since $\|M_{l\tilde{a}} T_{k\tilde{b}}\|_{Op} = 1$, $\|M_{l\tilde{a}} T_{k\tilde{b}}\|_{Wal} = 1$, and $\|M_{l\tilde{a}} T_{k\tilde{b}}\|_{fro} = \sqrt{n}$ the proof is complete. ■

Lemma 20:

$$\|S\|_{Op} \leq \|S\|_{Wal}$$

Proof:

$$\|S\|_{Op} = \max_{x:\|x\|_2=1}\left\{\|Sx\|_2\right\}$$

Let b_p in $\mathbb{C}^n$ with $(b_p)_j = B_{p,j \bmod a}$. We know from (4) that

$$\|Sx\|_2 = \left\|\sum_{p=0}^{b-1} T_{-p\tilde{b}}x \cdot b_p\right\|_2 \leq$$

$$\leq \sum_{p=0}^{b-1}\left\|T_{-p\tilde{b}}x\cdot b_p\right\|_2 \leq \sum_{p=0}^{b-1}\left\|T_{-p\tilde{b}}x\right\|_2 \cdot \|b_p\|_\infty =$$

$$= \sum_{p=0}^{b-1}\|x\|_2 \cdot \max_{j=0,\ldots,a-1}\left\{B_{p,j \bmod a}\right\} =$$

$$= \|x\|_2 \cdot \underbrace{\sum_{p=0}^{b-1}\max_{j=0,\ldots,a-1}\left\{B_{p,j \bmod a}\right\}}_{\|S\|_{Wal}}$$

■

Theorem 21: Let $B_{g,\gamma,a,b}$ be the $b\times a$ associated non-zero block matrix for g,γ,a,b, and $J_{g,\gamma,a,b}$ the corresponding Janssen-matrix. Then

$$F_a \cdot B^t_{g,\gamma,a,b} = a\cdot J_{g,\gamma,a,b} \text{ and } \|B\|_{fro} = \sqrt{a}\,\|J\|_{fro}.$$

Therefore for the corresponding frame matrix S

$$\|S\|_{fro} = \sqrt{n}\,\|J\|_{fro}$$

Proof:

$$J_{k,l} = \frac{n}{ab}\left\langle\gamma, M_{k\tilde{a}} T_{l\tilde{b}} g\right\rangle = \frac{1}{a}\tilde{b}\left(\mathcal{V}_g\gamma\right)(l\tilde{b}, k\tilde{a}) =$$

$$= \frac{1}{a}\tilde{b}\widehat{(\gamma\cdot \overline{T_{l\tilde{b}} g})}(k\tilde{a})$$

Let us look at the i-th row of B $b^{(i)} \in \mathbb{C}^a$ with $b_j^{(i)} = B_{i,j}$.

$$\hat{b^{(l)}}_k \overset{(3)}{=} \tilde{b} \left(\widehat{\sum_{p=0}^{\tilde{a}-1} T_{ap} \left(T_{l\tilde{b}} \overline{g} \cdot \gamma \right)} \right)_k \overset{Poisson}{=} \tilde{b} \cdot \left(\widehat{T_{l\tilde{b}} \overline{g} \cdot \gamma} \right)_{k\tilde{a}}$$

Note that we start with a Fourier transform in $\mathbb{C}^a$ but end up in $\mathbb{C}^n$ in this equation.

$$\Longrightarrow a \cdot J_{k,l} = \hat{b^{(l)}}_k$$

$$\|B\|_{fro} = \sqrt{\sum_{i=0}^{b-1} \sum_{j=0}^{a-1} \left| B_{i,j} \right|^2} = \sqrt{\sum_{i=0}^{b-1} \left\| b^{(i)} \right\|^2} =$$

$$= \sqrt{\sum_{i=0}^{b-1} \frac{1}{a} \left\| \widehat{b^{(i)}} \right\|^2} = \sqrt{\sum_{i=0}^{b-1} \frac{1}{a} \cdot \sum_{j=0}^{a-1} a^2 \left| J_{i,j} \right|^2} = \sqrt{a} \cdot \|J\|_{fro}$$

As S consists of $\frac{n}{a}$ rotated versions of the $n \times a$ block-matrix and this larger block-matrix has the same Frobenius norm as the non-zero block matrix, clearly

$$\|S\|_{fro} = \sqrt{\frac{n}{a}} \cdot \|B\|_{fro} \tag{7}$$

and therefore

$$\|S\|_{fro} = \frac{\sqrt{n}}{\sqrt{a}} \cdot \sqrt{a} \cdot \|J\|_{fro} = \sqrt{n} \cdot \|J\|_{fro}$$

■

Lemma 22:

$$\|J\|_{fro} \leq \|J\|_{1,1} \leq \sqrt{a \cdot b}\, \|J\|_{fro}$$

Proof: This is just an analogue property to the norm equivalence for $\|.\|_2$ and $\|.\|_1$ in $\mathbb{C}^n$:

$$\|x\|_2 \leq \|x\|_1 \leq \sqrt{n}\, \|x\|_2 .$$

■

Proposition 23:

$$\sqrt{\frac{n}{a \cdot b}}\, \|S\|_{Jan} \leq \|S\|_{fro} \leq \sqrt{n} \cdot \|S\|_{Jan}$$

Proof: We know the second part from Lemma 19.

$$\|S\|_{Jan} = \|J\|_{1,1} \leq \sqrt{a \cdot b}\, \|J\|_{fro} \overset{Lemma\ 21}{=} \sqrt{\frac{a \cdot b}{n}}\, \|S\|_{fro}$$

■

If the Gabor system constitutes a frame, $red = \frac{n}{a \cdot b} \geq 1$, and so $\|S\|_{Jan} \leq \|S\|_{fro}$. Therefore the Walnut norm approximate the operator norm better.

Lemma 24:

$$\frac{1}{\sqrt{a}}\, \|B\|_{fro} \leq \|B\|_{\infty,1} \leq \sqrt{b}\, \|B\|_{fro}$$

Proof:

$$\|B\|_{\infty,1} = \sum_{i=0}^{b-1} \max_{j=0,\dots,a-1} \left\{ \left| B_{i,j} \right| \right\} = (*)$$

Clearly

$$\max_{j=0,\dots,a-1} \left\{ \left| B_{i,j} \right| \right\} \leq \sqrt{\sum_{j=0}^{a-1} \left| B_{i,j} \right|^2}$$

and

$$(*) \leq \sum_{i=0}^{b-1} \sqrt{\sum_{j=0}^{a-1} \left| B_{i,j} \right|^2} \leq \sqrt{b} \sqrt{\sum_{i=0}^{b-1} \sum_{j=0}^{a-1} \left| B_{i,j} \right|^2}$$

On the other hand

$$\sqrt{\sum_{j=0}^{a-1} \left| B_{i,j} \right|^2} \leq \sqrt{a} \cdot \max_{j=0,\dots,a-1} \left\{ \left| B_{i,j} \right| \right\} \Longrightarrow$$

$$\sqrt{\sum_{i=0}^{b-1} \sum_{j=0}^{a-1} \left| B_{i,j} \right|^2} \leq \sum_{i=0}^{b-1} \sqrt{\sum_{j=0}^{a-1} \left| B_{i,j} \right|^2} \leq$$

$$\leq \sum_{i=0}^{b-1} \sqrt{a} \cdot \max_{j=0,\dots,a-1} \left\{ \left| B_{i,j} \right| \right\} = \sqrt{a} \cdot (*)$$

■

With Theorem 21 we get immediately

Proposition 25:

$$\frac{\sqrt{n}}{a \cdot b}\, \|S\|_{Wal} \leq \|S\|_{fro} \leq \sqrt{n}\, \|S\|_{Wal} \quad .$$

Combining Lemma 19, 23, 24 and Theorem 21 we get

Proposition 26:

$$\frac{1}{\sqrt{a \cdot b}}\, \|S\|_{Jan} \leq \|S\|_{Wal} \leq \|S\|_{Jan}$$

So in combination we get (6):

Theorem 27:

$$\|S\|_{Op} \leq \|S\|_{Wal} \leq \|S\|_{Jan} \leq \|S\|_{fro}$$

ACKNOWLEDGMENTS

The authors would like to thank Tobias Werther, Damin Marelli, Piotr Majdak and the anonymous reviewers for many helpful comments and suggestions as well as Claudia Balazs for final layouting.

The first author would like to thank the hospitality of the LATP, CMI, Marseille, France, where part of this work was prepared, supported by the HASSIP-network.

REFERENCES

[1] H. Bölcskei, F. Hlawatsch, and H. G. Feichtinger, “Frame-theoretic analysis of oversampled filter banks,” *IEEE Trans. Signal Processing*, vol. 46, no. 12, pp. 3256–3268, 1998.

[2] M. Dolson, “The phase vocoder: a tutorial,” *Computer Musical Journal*, vol. 10, no. 4, pp. 11–27, 1986.

[3] P. Balazs, “Regular and irregular Gabor multipliers with application to psychoacoustic masking,” PhD thesis, University of Vienna, June 2005.

[4] N. Delprat, B. Escudie, P. Guillemain, R. Kronland-Martinet, P. Tchamitchian, and B. Torrésani, “Asymptotic wavelet and Gabor analysis: extraction of instantaneous frequencies,” *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 644–664, 1992.

[5] F. Jaillet and B. Torrésani, “Time-frequency jigsaw puzzle: adaptive and multilayered Gabor expansions,” preprint.

[6] P. J. Wolfe, M. Dörfler, and S. J. Godsill, “Multi-Gabor dictionaries for audio time-frequency analysis,” in *Proceedings of the IEEE Workshop of Signal Processing to Audio and Acoustics*, 43–46 2001.

[7] M. J. Bastiaans, “Gabor’s signal expansion of a signal into Gaussian elementary signals,” *Proc. IEEE*, vol. 68, no. 4, pp. 538–539, 1980.

[8] H. G. Feichtinger and T. Strohmer, *Gabor Analysis and Algorithms - Theory and Applications*. Birkhäuser Boston, 1998.

[9] T. Werther, Y.C.Eldar, and N.K.Subbanna, “Dual Gabor frames: Theory and computational aspects,” *IEEE Transactions on Signal Processing*, vol. 53, no. 11, pp. 4147–4158, 2005.

[10] Y. Z. M. Zibulski, "Frame analysis of the discrete Gabor scheme," *IEEE Transactions on Signal Processing*, vol. 42, no. 4, pp. 942–945, 1994.
[11] A. Janssen, "The Zak transform: a signal transform for sampled time-continous signals," *Philips J. Res.*, vol. 43, pp. 23–69, 1988.
[12] T. Strohmer, *Numerical Algorithms for Discrete Gabor Expansions*. Birkhäuser Boston, 1998, ch. 8, pp. 267–294.
[13] J. Wexler and S. Raz, "Discrete Gabor expansion," *Signal Processing*, vol. 21, pp. 207–220, 1990.
[14] S. Qiu and H. Feichtinger, "Discrete Gabor structures and optimal representations," *IEEE Transactions on Signal Processing*, vol. 43, no. 10, pp. 2258–2268, 1995.
[15] T. Tschurtschenthaler, "The Gabor frame operator (its structure and numerical consequences)," Master's thesis, University of Vienna, 2000.
[16] A. Meister, *Numerik linearer Gleichungssysteme*. Friedr. Vieweg & Sohn Braunschweig, 1999.
[17] J. Stoer, *Introduction To Numerical Analysis*. Springer New York, 2002.
[18] L. N. Trefethen and D. B. III, *Numerical Linear Algebra*. SIAM Philadelphia, 1997.
[19] A. Janssen, "Duality and biorthogonality for Weyl-Heisenberg frames," *J. Fourier Anal. Appl.*, vol. 2, no. 4, pp. 403–436, 1995.
[20] O. Axelson and V. Barker, *Finite Element Solution of Boundary Value Problems, Theory and Computation*. Academic Prers, Orlando, 1984.
[21] D. Luenberger, *Linear And Nonlinear Programming*. Addison-Wesley, Reading, 1984.
[22] R. J. Duffin and A. C. Schaeffer, "A class of nonharmonic Fourier series," *Trans. Amer. Math. Soc.*, vol. 72, pp. 341–366, 1952.
[23] K. Gröchenig, *Foundations of Time-Frequency Analysis*. Birkhäuser Boston, 2001.
[24] O. Christensen, *An Introduction To Frames And Riesz Bases*. Birkhäuser, 2003.
[25] D. F. Walnut, "Continuity properties of the Gabor frame operator," *J. Math. Anal. Appl.*, vol. 165, no. 2, pp. 479–504, 1992.
[26] S. Qiu and H. Feichtinger, "Gabor-type matrices and discrete huge gabor transforms," in *ICASSP-Proceeding, Detroit*, 1995, pp. 1089–1092.
[27] S. Qiu, "Gabor-type matrix algebra and fast computations of dual and tight Gabor wavelets," *Opt. Eng.*, vol. 36, no. 1, pp. 276–282, 1997.
[28] M. Dörfler and H. G. Feichtinger, "Quilted Gabor families I: reduced multi-Gabor frames," preprint.

Peter Balazs (S'02-M'05) was born in Tulln, Austria on December 11, 1970. He received the M.S. (2001) and the PhD degree (2005) (both with distinction) at the University of Vienna in mathematics. He has been a member of the Acoustics Research Institute of the Austrian Academy of Science, since 1999. He is a fellow of the HASSIP EU network, joining the LATP, CMI and LMA, CNRS Marseille from November 2003 - April 2004 and the FYMA, UCL, Louvain-La-Neuve in August 2005.

Peter Balazs is interested in Time-Frequency Analysis, Gabor Analysis, Numerics, Frame Theory, Signal Processing, Acoustics and Psychoacoustics.

Hans G. Feichtinger has been a member of the Faculty of Mathematics, University of Vienna, since 1973. He has held visiting positions in Sweden, Germany, France and USA. He is Editor-in-Chief of the Journal of Fourier Analysis and Applications, associate editor of the Journal of Approximation Theory and the new journals "Sampling Theory in Signal and Imaging Processing" and "Function Spaces and Applications". He edited two books about the field of Gabor Analysis and authored or co-authored over 150 publications in research journals and conference proceedings. He is the host scientist of the upcoming Marie Curie Excellence Grant EUCETIFA (= European Center for Time-Frequency Analysis), 2005-2009.

Dr. Feichtinger's research interests are Harmonic Analysis, Gabor Analysis, (irregular) Sampling, and Wavelets and Function spaces.

Mario Hampejs was born in Vienna, Austria on March 14, 1968. He received his M.S. degree in 1999 at the University of Vienna in mathematics. Currently he is employed by Siemens (working field: interlocking systems). He is working on the Ph.D. degree at the research group NUHAG (University of Vienna).

Mario Hampejs is interested in Gabor analysis, time-frequency analysis and hazard analysis.

Günther Kracher received the M.S. degree in 1998 from the University of Vienna in mathematics and the Ph.D in computer science in 2001 from the University of Technology, Vienna. After a PostDoc year in machine learning and object recognition he is with the research group NuHAG, faculty of mathematics, University of Vienna.

Günther Kracher is interested in time-frequency analysis, Gabor analysis, signal processing, image processing and machine learning.