Local and Global Aspects of Time-Frequency Analysis with Applications to Sound Analysis

Monika Dörfler

Habilitationsschrift

Wien, January 10, 2013

Universität Wien
Fakultät für Mathematik
2010 Mathematics Subject Classification: 33C45, 42A45, 42A65, 42B15, 42B35
42C15, 42C30 42C40, 47A70 47B38, 47G30, 65F08 , 65F20, 65F35, 94A12

Key words and phrases:
Time-frequency analysis; Gabor frame; Adaptive representation; Uncertainty principle; Frame bounds; Time-frequency localization; Short-time Fourier transform; Modulation space; Localization operator; Sparse regression
Preface

This habilitation thesis consists of 12 papers which were published in renowned journals (iii, vi, vli, viii, vii, x, xii), refereed conference proceedings (ix, xi) or have been submitted during the last months and are under review in renowned journals (i, ii, iv). The included papers are complemented by an introductory Section 1 which briefly introduces the historical and technical context of time-frequency analysis and gives a short summary of the content and most important statements of each of the papers. While all the included papers are thematically connected, they are ordered according to the three topic areas Adaptive representations, cf. Section 1.3, Localization Operators and Gabor multipliers, cf. Section 1.4 and Applications of time-frequency analysis in sound analysis, cf. Section 1.5.

Acknowledgments

I gratefully acknowledge financial support provided by the Austrian Science Fund (FWF) under project LOCATIF(T384-N13) and by the WWTF within project Audio-Miner (MA09-024).

I would like to thank all my co-authors and students for their inspiring and constructive cooperation. In particular, I thank those colleagues who co-authored work presented in this thesis: Daniel Abreu, Thomas Grill, Nicki Holighaus, Karlheinz Gröchenig, Ewa Matusiak, José Luis Romero, Kai Siedenburg, Bruno Torrésani and Gino Velasco.

I am deeply grateful to Hans Georg Feichtinger for his constant support and encouragement.

Without the love of my family and their enthusiasm for life, I would not be able to work in science. My warmest thanks go to all my family, very particularly to Michael, Milena and Simon, and to my parents Maria and Willi.
List of Included Papers

Adaptive representations


Localization Operators and Gabor multipliers


Applications of time-frequency analysis in sound analysis


Contents

1 Introduction and Summary of Included Articles 9

1.1 Harmonic analysis - History and Impact ................................. 9

1.1.1 The need for localization in time and frequency ...................... 10

1.2 Signal representations and norm equivalence .......................... 11

1.2.1 The Heisenberg group, the short-time Fourier transform and
Gabor frames ............................................................................. 13

1.3 Adaptive Representations ...................................................... 14

1.3.1 Nonstationary Gabor Frames - Existence and Construction
- [i] ....................................................................................... 17

1.3.2 Nonstationary Gabor Frames - Approximately Dual Frames
and Reconstruction Errors - [ii] ................................................. 19

1.3.3 Quilted Gabor frames - A new concept for adaptive time-
frequency representation - [iii] ................................................... 21

1.3.4 Frames adapted to a phase-space cover - [iv] ......................... 25

1.4 Localization operators and Gabor multipliers .......................... 28

1.4.1 An inverse problem for localization operators - [v] ................. 30

1.4.2 Time-Frequency partitions and characterizations of modu-
lations spaces with localization operators - [vi] ......................... 32

1.4.3 Representation of operators in the time-frequency domain
and generalized Gabor multipliers - [vii] ................................. 34

1.4.4 Representation of operators by sampling in the time-frequency
domain - [viii] ........................................................................ 36

1.5 Applications of time-frequency analysis in sound analysis .......... 38

1.5.1 Constructing an invertible constant-Q transform with non-
stationary Gabor frames - [ix] .................................................. 41
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.5.2</td>
<td>A framework for invertible, real-time constant-Q transforms</td>
<td>43</td>
</tr>
<tr>
<td>1.5.3</td>
<td>Structured sparsity for audio signals</td>
<td>44</td>
</tr>
<tr>
<td>1.5.4</td>
<td>Persistent Time-Frequency Shrinkage for Audio Denoising</td>
<td>46</td>
</tr>
</tbody>
</table>

Reference [ii] 74
Reference [iii] 94
Reference [iv] 114
Reference [v] 135
Reference [vi] 151
Reference [vii] 173
Reference [viii] 206
Reference [ix] 226
Reference [x] 233
Reference [xi] 244
Reference [xii] 248
Chapter 1

Introduction and Summary of Included Articles

 [...] The fundamental things apply
As time goes by.
(Herman Hupfeld)

1.1 Harmonic analysis - History and Impact

From its very beginning, which is actually not easy to date, harmonic analysis has developed in close connection to and inspired by problems from applications in astronomy, physics, communication theory, to name but a few. While the idea to render a comprehensive review of the utterly fascinating history of harmonic analysis is way beyond the scope of this short introductory section, I would like to recall a few sign-posts which should help to understand the current research in a very lively and highly active field, which still carries the characteristics that have shaped harmonic analysis: inter-disciplinarity.

When Fourier, in 1822, published his Analytical Theory of Heat, the idea to represent a function as a sum of infinitely many harmonics, that is, sinusoids, was not entirely new, but had been proposed by Bernoulli, d’Alembert, Euler and Gauss

1It was discovered in 1977 [Gol77], that Gauss, when working on the motion of planets, extended previous work by Euler and Lagrange and, while computing, by using discrete Fourier coefficients, the correct orbit of the "lost" minor Planet Ceres (previously discovered by G. Piazzi) even developed an algorithm that was later re-invented by Cooley and Tukey and is now known
Fourier, however, started a more rigorous mathematical analysis of previous work and developed the computation of Fourier coefficients for periodic functions. Thus, he set the beginning of the interaction between application and theoretical work that is typical in harmonic analysis.

1.1.1 The need for localization in time and frequency

Fourier analysis provides information about the frequency content of a signal, or class of signals, of interest. Applying a Fourier transform to a signal always assumes some form of stationarity in the signal, since changes of the frequency content may not be captured in this manner. However, many meaningful signal classes, foremost speech, music and images, have highly time-variant frequency content. Thus, in the analysis of these natural signals of utmost importance for humans, the desire for time-varying Fourier analysis arose quite naturally and led to the birth of time-frequency analysis. The immediate idea to cut a given signal into pieces and consider the Fourier transform of each time-localized piece separately is, while still quite commonly used in some applications, not appealing from a mathematical point of view; this method is equivalent to the multiplication with translated box-functions and thus, if Fourier transformed, to the convolution of the signal’s spectrum with a sinc-function. The sinc-function, however, features slow decay and, therefore, frequency localization is here sacrificed for the sake of localization in time. This observation is a first glimpse of a very central problem in time-frequency analysis, namely the trade-off between good time- and good frequency resolution. Since, as made precise by Heisenberg’s uncertainty principle, no function may be arbitrarily well concentrated in both time and frequency, the multiplication with a function (window) for the sake of time-localization always must be considered together with the resulting frequency resolution. What we measure is determined by how we measure.

On the other hand, simply cutting a signal into pieces bears the advantage that the familiar concept of orthonormal bases is still largely sufficient to understand the outcome of our analysis endeavor. Using smoother windows which allow for better concentration in both time and frequency, requires more flexible concepts for the preservation of signal energy, beyond the usual Parseval or Plancherel equations.

as the Fast Fourier Transform (FFT).
Once this perspective is adopted, a new and exciting realm can be entered.

Dennis Gabor was a Hungarian-British electrical engineer and physicist, who received the Nobel Prize in Physics for inventing holography. He introduced, in the context of communication theory, the idea to represent a given signal by means of time-frequency shifted Gaussian windows, thus localizing in both time and frequency, [Gab46]. However, his suggested model to "cover" the time-frequency plane with Gaussian windows at critical density, that is, without redundancy, fails to provide stable expansions. Still, his idea turned out to be seminal for the development of signal representations, cf. [FS08] for a historical review of Gabor analysis and references.

1.2 Signal representations and norm equivalence

We will now formulate a core idea which is at the heart of harmonic analysis and the scientific work presented in this thesis, namely, the concept to represent, or synthesize, a function, or signal, \( f \) by means of a sum of - possibly infinitely many - weighted basic atoms. In the case of Fourier series, the basic functions are complex sinusoids and form an orthonormal basis for the space of periodic functions. In the situation introduced by Gabor, the basic functions are modulated Gaussians and, in the originally proposed form, they are complete but do not even form a Schauder basis. For a Hilbert space \( \mathcal{H} \) and some index set \( I \), we consider a sequence \( \Phi = (\varphi_k)_{k \in I} \) of functions \( \varphi_k \in \mathcal{H} \), often called dictionary in applications without further specifying its mathematical properties.

**Definition 1** (Synthesis Operator). Let a dictionary \( \Phi \) and a sequence \( \mathbf{c} = (c_k)_{k \in I} \) of complex numbers be given. Then, the synthesis operator \( D_{\Phi} \) corresponding to \( \Phi \) is formally defined by

\[
D_{\Phi}(\mathbf{c}) = \sum_{k \in I} c_k \varphi_k
\]

While, in applications, important questions regarding \( \Phi \) and \( \mathbf{c} \) usually depend heavily on the concrete task one is interested in, be it denoising, source separation, compression, or many others, some immediate mathematical questions arise from the sloppy formulation of Definition 1: what are the properties of the sequence \( \mathbf{c} \), e.g. (square) summability, and what are resulting properties of \( D_{\Phi} \); can any...
function \( f \in \mathcal{H} \) be represented in the dictionary, i.e., is \( \Phi \) complete in \( \mathcal{H} \) or a subspace thereof, what are the convergence properties of \((1.1)\) and, given \( f \in \mathcal{H} \), how can appropriate coefficients be computed, what are their properties and what happens, if known coefficients that lead to synthesis of \( f \) in \((1.1)\) are modified? Some of these questions have been successfully and extensively treated for some important and widely used dictionaries. We will give some answers and references in due course. Let us, however, briefly comment on the last-mentioned question, namely the modification of synthesis-coefficients in a given representation. In many applications, this modification is the very reason for the usage of a certain dictionary in the first place. For example, using a sinusoidal representation, e.g. Fourier series for periodic signals, allows us to damp or amplify certain frequencies, which corresponds to a filtering process. In compression, one is interested in finding a dictionary which allows us to express the significant part of the signal of interest with few non-zero coefficients, that is, less important coefficients are set to zero. Further examples will be found later on - the idea is always similar: by choosing a dictionary in a judicious way, a particular application is more easily tackled, and modification of coefficients is almost always involved.

Let us, for now, assume that the coefficient sequences are from the Hilbert space \( \ell^2 \); then it can immediately be seen that the adjoint operator of \( D_\Phi(c) \) is given by the following analysis operator with respect to \( \Phi \).

**Definition 2 (Analysis Operator).** For a dictionary \( \Phi \) in \( \mathcal{H} \) and a function \( f \in \mathcal{H} \), the synthesis operator \( C_\Phi \) corresponding to \( \Phi \) is formally defined by

\[
(C_\Phi(f))(k) = \langle f, \varphi_k \rangle_{\mathcal{H}}
\]  

(1.2)

From a mathematical point of view, keeping the above-mentioned questions as well as properties of more familiar concepts such as (orthonormal) basis in mind, we may wish for our dictionary to be at least complete, to maybe provide us with a stable possibility to reconstruct our original function \( f \) from the analysis coefficients \((1.2)\), and to give a bounded (continuous) relation between the function and its analysis (or synthesis) coefficients. In order to give all these wishes a concise mathematical formulation, one considers the composition of analysis and synthesis operator, that is, the frame operator \( S_\Phi \):
Definition 3 (Frame Operator and Frame). For a dictionary $\Phi$ in $\mathcal{H}$, the frame operator $S_\Phi$ is formally defined by

$$S_\Phi f = D_\Phi C_\Phi f = \sum_{k \in I} \langle f, \varphi_k \rangle \varphi_k.$$  

(1.3)

If $S$ is bounded and invertible, $\Phi$ is called a frame for $\mathcal{H}$. Equivalently, $\Phi$ is a frame if there exist constants $0 < A, B < \infty$ such that

$$A\|f\|^2 \leq \sum_{k \in I} |\langle f, \varphi_k \rangle|^2 \leq B\|f\|^2.$$  

(1.4)

The inequality (1.4) may be interpreted as a relaxation of the Parseval’s equality and is called frame inequality. The definition of frames has been extended to Banach spaces [Grö91] and is closely related to characterization of certain function spaces by coefficients given by $C_\Phi$. In particular, modulation spaces, cf. [Grö01], are related to the decay of the Gabor coefficients of a given function.

1.2.1 The Heisenberg group, the short-time Fourier transform and Gabor frames

While general frame theory [Chr03] does not rely on any special structure of the frame elements, time-frequency analysis is closely related to Gabor frames, which, in turn, are connected with the Heisenberg group and the Schrödinger representation. In fact, the beautiful and complex, non-commutative structure of the Heisenberg group arises from the non-commutativity of time-frequency shifts.

Definition 4 (Time-frequency shift). Given $\varphi \in L^2(\mathbb{R})$, let $T_x \varphi(t) := \varphi(t-x)$ and $M_w \varphi(t) := e^{2\pi i wt} \varphi(t)$.

$T_x$ is a translation operator or time shift, $M_w$ is a modulation operator, or frequency shift, and operators of the form

$$T_x M_w = e^{-2\pi i wx} M_w T_x$$  

(1.5)

are called time-frequency shifts.

Now, the Schrödinger representation is a (unitary) representation of the Heisenberg group by time-frequency shifts, cf. [Grö01, Chapter 9] for details and, up to a phase factor, its representation coefficients coincide with the short-time Fourier transform (STFT), ubiquitous in signal analysis, which we define next.
Definition 5 (STFT). The short-time Fourier transform (STFT) of a distribution \( f \in \mathcal{S}'(\mathbb{R}^d) \) with respect to \( \varphi \in \mathcal{S}(\mathbb{R}^d) \) is a function on \( \mathbb{R}^d \times \mathbb{R}^d \) defined by

\[
V_\varphi f(x,w) = \int_{\mathbb{R}^d} f(t) \overline{\varphi(t - x)} e^{-2\pi i \xi t} dt = \langle f, T_x M_w \varphi \rangle, \quad (x, w \in \mathbb{R}^d \times \mathbb{R}^d).
\]

Note that this definition is reminiscent of Gabor’s idea presented in [Gab46]. However, in the STFT, the value \( V_\varphi f(x, w) \), representing the energy of \( f \) near time \( x \) and frequency \( \xi \), is measured continuously. In this situation, reconstruction is always possible for non-zero \( \varphi \) by

\[
f(t) = \frac{1}{\| \varphi \|^2} \int_{\mathbb{R}^d \times \mathbb{R}^d} V_\varphi f(x, w) M_w T_x \varphi(t) dx dw. \tag{1.6}
\]

Gabor’s approach, now, corresponds to a discretization, or sampling, of the STFT on the lattice \( \mathbb{Z} \times \mathbb{Z} \subset \mathbb{R} \times \mathbb{R} \). Using more general lattices \( a\mathbb{Z} \times b\mathbb{Z} \subset \mathbb{R} \times \mathbb{R} \), for some real, positive \( a, b \), then the following definition is obtained.

Definition 6 (Gabor frame). Given a non-zero function \( \varphi \in L^2(\mathbb{R}) \), let \( \varphi_{k,l}(t) = M_{bl} T_{ak} \varphi(t) \). The set \( G(\varphi, a, b) = \{ \varphi_{k,l} : k, l \in \mathbb{Z} \} \) is called a Gabor system for any real, positive \( a, b \) and Gabor frame, if the frame condition (1.4) is satisfied.

An important theorem in Gabor analysis is the Balian-Low theorem, which states that Gabor frames defined on the critical lattice \( \mathbb{Z} \times \mathbb{Z} \subset \mathbb{R} \times \mathbb{R} \) as originally suggested by Gabor, can not be obtained with a window function (or Gabor atom) \( \varphi \) which is simultaneously well-localized in both time and frequency, cf. [Grö01, Section 8.4] for the precise statement. The implication of this fact is the insight that in time-frequency analysis with Gabor frames, redundancy is a necessary requirement in order to obtain good time-frequency representations.

1.3 Adaptive Representations

As pointed out in the previous section, Gabor frames provide more flexibility than Gabor (Riesz) bases, in particular concerning the joint time-frequency concentration of the window \( \varphi \). However, requirements arising in applications may go beyond the flexibility obtained within the scheme of regular Gabor frames and may lead to a situation in which the underlying group structure - of the Heisenberg group in the case of Gabor frames - has to be abandoned.
As illustrated in Figure 1.1, natural signals may, and usually will, have various components with distinct time-frequency localization properties. Just as a pure sinusoid is perfectly represented in the Fourier basis, in the sense of requiring just one non-zero coefficient, while strictly time-localized signals, ideally close to the Dirac delta function, lead to badly decaying coefficients, different Gabor frames are better or worse adapted to certain signal properties, cf. [Koz97, DP02]. As an illustrating example, Figure 1.1, taken from [BJ+11], shows two versions of Gabor coefficients obtained by analysis with a narrow (short time duration) and a wide (long time duration) window. It becomes immediately apparent that signal components, which are localized in time, mathematically closer to a Dirac delta function, are better represented by the short window than the more sinusoidal components which appear much sharper in the second plot, corresponding to the long window. This example now hints at a fundamental problem one faces when forced to choose one particular size and shape of $\phi$ for the representation of the entire signal: the trade-off between good time- and good frequency-resolution. Due to this observation, the desire to change, and thus adapt, the time-frequency resolution over time or frequency, or time and frequency, arose and has led to various approaches enabling more flexible resolution of the time-frequency plane, cf. [HRVK93, TV96, Shl97, XM01] among others.
In our work, we have considered, on the one hand, adaptive time- or frequency-resolution by investigating a variant of the classical (regular) Gabor frames, denoted nonstationary Gabor frames (NSG) and, to account for more flexible tilings, frame that allow for adaptation in time and frequency. The latter require more refined mathematical techniques and two approaches have been proposed: Quilted Gabor frames [iii] and frames derived from eigenfunctions of time-frequency localization operators, [iv].

An example of the representation coefficients corresponding to the previously presented Glockenspiel signal, this time obtained by means of NSG frames with adaptivity in time, is shown in the right plot of in Figure 1.2 in comparison to a regular Gabor representation with an analysis window whose width lies in between the windows used for the plots in Figure 1.1. For the adaptive representation, the onsets, where short-lived and thus time-concentrated signal components are expected, were automatically extracted from the signal by a standard peak-tracking algorithm, and, subsequently, the windows were adapted such that narrow windows are used near the transient signal components. In [BDJ+11] this algorithm, based on painless NSG frames, was introduced, and is described in detail. It is visible and was evaluated in the above-mentioned reference, that the usage of adapted
frames leads to a more sparse representation of the signal of interest. The construction in [BDJ+11], and similar approaches suggested in [WRP10], is based on the assumption that the analysis windows have either compact support or compact bandwidth. This is in parallel to the development of an understanding of existence of classical Gabor frames, for which the painless case, with compactly supported window, was also the first situation to be fully understood, cf. [DGM86], while more general existence results followed later [Wal92, RS97]. The straightforward analysis and implementation of painless non-orthogonal expansions is due to the fact that, if the generating window is compactly supported, the frame operator $S_{G,(\varphi,a,b)}$ is diagonal under suitable conditions on the sampling parameters. The generalization of this fact to painless NSG frames is rather straightforward, cf. [BDJ+11]. Painless NSG frames with adaptivity in frequency have been exploited in order to construct invertible constant-Q transform, further described in Section 1.5.2 and Section 1.5.1.

The work summarized in the next two sections addresses a more general situation that does not require the rigid support conditions on the windows. In particular, the first paper [i] addresses existence and construction of general NSG frames, while in [ii] approximate reconstruction with these frames is investigated. The two subsequent sections then summarize the paper on quilted frames [iii] and more recent work frames derived from eigenfunctions of time-frequency localization operators, [iv], which is closely connected to the work, presented in Section 1.4, on localization operators.

1.3.1 Nonstationary Gabor Frames - Existence and Construction - [i]

Nonstationary Gabor systems are a generalization of classical Gabor systems of regular time-frequency shifts of a single window function.

Definition 7. Let $g = \{g_k \in W(L^\infty, \ell_1) : k \in \mathbb{Z}\}$ be a set of window functions and let $b = \{b_k : k \in \mathbb{Z}\}$ be a corresponding sequence of frequency-shift parameters. Set $g_{k,l} = M_{b_k}g_k$. Then, the set

$$\mathcal{G}(g, b) = \{g_{k,l} : k, l \in \mathbb{Z}\}$$

(1.7)

is called a nonstationary Gabor (NSG) system.
A class of functions which is commonly used in time-frequency analysis on $L^2(\mathbb{R})$ is the Wiener space. We give its definition next for convenience of the reader and as reference for subsequent use.

**Definition 8.** A function $g \in L^\infty(\mathbb{R})$ belongs to the Wiener space $W(L^\infty, \ell^1)$ if

$$
\|g\|_{W(L^\infty, \ell^1)} := \sum_{k \in \mathbb{Z}} \text{ess sup}_{t \in Q} |g(t + k)| < \infty, \quad Q = [0, 1].
$$

In [i], by assuming mild joint decay conditions on all the windows, we derive an existence result for nonstationary Gabor frames, which is based on the following, new Walnut-like representation of the frames operator of $G(g, b)$.

**Proposition 1.3.1.** The frame operator $S_{g, \gamma}$ of the NSG system $G(g, b)$ admits a Walnut representation for $f \in L^2(\mathbb{R})$:

$$
S_{g, \gamma} f = \sum_{k, l \in \mathbb{Z}} G_{k, l}^{g, \gamma} \cdot T_{b_k^{-1}, f}, \quad \text{where} \quad G_{k, l}^{g, \gamma}(t) = b_k^{-1} g_k(t - b_k^{-1} l) \gamma_k(t). \quad (1.8)
$$

A bound for the operator norm of $S_{g, \gamma}$ is also derived. The existence result is the formulated as follows.

**Theorem 1.3.2.** Let $g = \{g_k \in W(L^\infty, \ell^1) : k \in \mathbb{Z}\}$ be a set of windows such that

i) for some positive constants $A_0, B_0$

$$
0 < A_0 \leq \sum_{k \in \mathbb{Z}} |g_k(t)|^2 \leq B_0 < \infty \quad \text{a.e.}; \quad (1.9)
$$

ii) for all $k \in \mathbb{Z}$, the windows decay polynomially around a $\delta$-separated set $\{a_k : k \in \mathbb{Z}\}$ of time-sampling points $a_k$

$$
|g_k(t)| \leq C_k (1 + |t - a_k|)^{-p_k}, \quad (1.10)
$$

where $p_k \in [p_L, p_U] \subset \mathbb{R}$, $p_L > 2$ and $C_k \in [C_L, C_U]$.

Then there exists a sequence $\{b^0_k\}_{k \in \mathbb{Z}}$, such that for $b_k \leq b^0_k$, $k \in \mathbb{Z}$, the nonstationary Gabor system $G(g, b)$ forms a frame for $L^2(\mathbb{R})$. 
Its proof relies on irregular sampling estimates for polynomially decaying functions. An analog statement is formulated for frequency-adaptive nonstationary frames.

Further results in the paper include the concrete construction of nonstationary Gabor frames as perturbation of a known, in particular, of a painless nonstationary Gabor frame, stated in Proposition 3.7. This is a situation of practical relevance, since sometimes it may be beneficial to use windows that violate the compact support condition, but decay very fast. In this situation, the frame property of the system obtained from the non-compactly supported windows may be derived from a perturbation result with respect to a painless frame obtained from appropriate restriction of the windows to a bounded interval, which, in turn, determines the frequency sampling parameters. This result is given in Corollary 3.8 of the paper. The paper concludes with two examples which describe the construction of nonstationary Gabor frame. In Example 1, the windows considered are derived from Hannin windows of various widths by band-limiting. This construction is of particular interest for obtaining frequency-adaptive frames with windows that are compactly supported in time, which allows for real-time implementation with finite impulse response filters, cp. [EDM12]. In the second example, dilated Gaussian windows are investigated, for which the corresponding painless NSG frame is obtained by applying a cut-off to the Gaussian windows by multiplication with a characteristic function corresponding to the intervals $I_k$. The size of these intervals subsequently determine the frequency-sampling constant $b_k$.

1.3.2 Nonstationary Gabor Frames - Approximately Dual Frames and Reconstruction Errors - [ii]

Perfect reconstruction from coefficients obtained by applying nonstationary Gabor frames which are constructed with non-compactly supported window, as proposed in the previous section, may be hard and, in particular, computationally extremely inefficient, to obtain. In [ii], it is shown that for nonstationary Gabor frames that are related to some known frames, for which dual frames can be computed, good approximate reconstruction is achieved by resorting to approximately dual frames. In particular, we give constructive examples for so-called almost painless nonstationary frames, that is, frames that are closely related to nonstationary frames.
with compactly supported windows. As an application example, we mention the
construction of nonuniform filter banks via nonstationary Gabor frames, in which
case the restriction to bandlimited windows forbids finite impulse response filters;
the latter are, however, imperative for real-time processing applications.

In [ii], we therefore go beyond the results presented in [BDJ+11] and consider
nonstationary Gabor frames with windows showing fast decay but unbounded sup-
port. The existence of this kind of frames was shown in [i]. Here we are concerned
with methods for approximate reconstruction for these adaptive systems.

For the error estimates given for reconstruction with approximately dual frames,
we start from the Walnut representation for NSG frames (1.8) and introduce the
the following correlation functions of a pair of Gabor systems:

\[ G_{l}^{g,\gamma}(t) = \sum_{k \in \mathbb{Z}} b_{k}^{-1} |g_{k}(t - lb_{k}^{-1})||\gamma_{k}(t)|, \text{ for } l \in \mathbb{Z}. \]  

(1.11)

and from which we derive various bounds for the frame operator, e.g.:

\[ \|S_{g,\gamma}\|^{2} \leq \text{ess sup}_{l \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} G_{l}^{g,\gamma} \cdot \text{ess sup}_{l \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} G_{l}^{\gamma,g}. \]  

(1.12)

Denoting the off-diagonal entries of the frame operator by

\[ R_{g,\gamma} = \text{ess sup}_{l \in \mathbb{Z} \setminus \{0\}} \sum_{k \in \mathbb{Z}} b_{k}^{-1} |g_{k}(\cdot - lb_{k}^{-1})||\gamma_{k}(\cdot)|, \]  

(1.13)

we first derive conditions on NSG frames to form Bessel sequences (Proposition 3.1)
and then state the following estimate for the frame operator of two NSG Bessel
sequences (Lemma 3.2):

Let \( G(g, b) \) and \( G(\gamma, b) \) be two NSG Bessel sequences. Then

\[ \|I - S_{g,\gamma}\| \leq \left\| 1 - \sum_{k \in \mathbb{Z}} b_{k}^{-1}\overline{g_{k}}\gamma_{k} \right\|_{\infty} + \sqrt{R_{g,\gamma} \cdot R_{\gamma,g}}. \]  

(1.14)

In Section 4, the notion of approximately dual frames, introduced in [CL10],
is applied to NSG frames. First, in Proposition 4.1, the existence result obtained
in [i] is reformulated in the context of approximately dual frames by giving an error
estimate for the reconstruction with a single preconditioning dual frame, obtained
by multiplication with the inverse of the diagonal of the frame operator: \( \gamma_{k} = (G_{0}^{g,g})^{-1}g_{k}, k \in \mathbb{Z} \). Then, in Proposition 4.3, alternative approximate dual frames
are proposed in a situation where the NSG frame of interest can be interpreted as
a perturbation of another NSG frames which is easier to handle. Typically, the latter is given by a painless NSG frame, that is, one that has compactly supported analysis windows. This special case is treated in Section 4.0.1, where painless NSG frames are defined and more explicit error estimates are given in Corollary 4.3.

The results on almost painless NSG frames are then used to construct a concrete example, namely, in Example 1, which investigates an NSG frame whose analysis windows are dilated Gaussian windows. The windows’ fast decay is exploited to derive error estimate by considering the resulting frame as a perturbation of a painless frame that is obtained by multiplying the dilated Gaussian windows with characteristic functions of appropriate bounded intervals $I_k$.

A second, numerical, example treats the case of a frequency-adaptive NSG frames constructed from Hanning windows which are compactly supported in time. Again, the reconstruction quality obtained with the three proposed approximate dual frames is computed and the windows are shown in comparison to each other.

1.3.3 Quilted Gabor frames - A new concept for adaptive time-frequency representation - [iii]

The idea behind the concept of quilted Gabor frames is to allow for the flexible design of frames with adaptivity in both time and frequency, in contrast to NSG frames, which allow adaptation only in one domain. Quilted frames have been referred to in applied work [JT07, LBR11] and their existence has been derived for polynomially decaying windows within an abstract framework in [Rom11]. This shows that quilted frames also bear theoretical interest in themselves and should be compared to constructions such as fusion frames [CKL08] and the frames proposed in [ACM04]. Quilted Gabor frames are introduced as follows:

**Definition 9** (Quilted Gabor frames). Let Gabor frames $G(g^j, \Lambda^j)_{j \in J}$ for $L^2(\mathbb{R}^d)$ and an (admissible) covering $\Omega_r \subseteq (B_{R_r}(x_r))_{r \in I}$ of $\mathbb{R}^{2d}$ be given. Define the local index sets $X^r = \Omega_r \cap \Lambda^{m(r)}$, where $m : I \mapsto J$ is a mapping assigning a frame from the given Gabor frames to each member of the covering. Then the set

$$\bigcup_{r \in I} G(g^{m(r)}, X^r)$$

is called a quilted Gabor frame for $L^2(\mathbb{R}^d)$, if there exist constants $0 < A, B < \infty$,.
such that
\[
A \|f\|_2^2 \leq \sum_{r \in \mathcal{I}} \sum_{\lambda \in \Lambda^r} |\langle f, \pi(\lambda)g^{m(r)} \rangle|^2 \leq B \|f\|_2^2
\] (1.16)
holds for all \( f \in L^2(\mathbb{R}^d) \).

Figure 1.3: Partition in time-frequency and resulting quilted lattice.

Figure 1.3 from [iii] illustrates the concept schematically.

The first main result of the paper, Theorem 1, gives, in generalization of earlier work in [LW03] on the Bessel property of irregular time-frequency shifts of a single atom, the existence of an upper frame bound (or Bessel bound) for quilted frames. Then, a criterion for compactly supported windows to form a quilted frame with varying lattices and windows in time, is presented in Proposition 1. This is different from the construction of NSG frames, since, here, the construction starts from global regular Gabor frames which are then used in assigned time-strips.
In Section 6 of the paper, quilted Gabor frames are constructed by replacing atoms in a given global frame by atoms from a different frame in a certain bounded region of the time-frequency plane. A sample statement is as follows.

**Proposition 1.3.3.** Assume that tight Gabor frames $\mathcal{G}(g, \Lambda^1)$ and $\mathcal{G}(h, \Lambda^2)$ with $g, h \in S_0(\mathbb{R}^d)$ are given. Assume further that in a compact region $\Omega \subset \mathbb{R}^{2d}$ the time-frequency shifted atoms $\pi(\lambda)g$, $\lambda \in F_1 = \Omega \cap \Lambda^1$ are to be replaced by a finite set of time-frequency shifted atoms $\pi(\mu)h$, $\mu \in F_2 \subset \Lambda^2$. Then, a compact set $\Omega^*$ in $\mathbb{R}^{2d}$ can be chosen such that for $F_2 = \Omega^* \cap \Lambda^2$, the union

$$\mathcal{G} = \{\pi(\lambda)g : \lambda \in \Lambda^1 \setminus F_1\} \cup \{\pi(\mu)h : \mu \in F_2\}$$

(1.17)

is a (quilted Gabor) frame.

![Image](https://via.placeholder.com/150)

**Figure 1.4:** Replacing atoms with and without overlap, low redundancy.

In Section 7 the paper presents simulations, in particular concerning reconstruction from coefficients obtained from the quilted frames investigated in the
paper. The influence of various parameters such as the density of the sampling lattices $\Lambda^j$ and the overlap between the local patches $\Omega^r$ is investigated, in terms of the condition number of the resulting quilted frame and the convergence of the corresponding (iterative) frame algorithm. As an example, we reprint Figure 7 from [iii] in Figure 1.4, where three amounts of overlap are shown in the upper display. The condition numbers of the resulting systems and the convergence behavior of the corresponding frame algorithm are shown in the lower display. It can be observed that some overlap is essential in order to obtain fast convergence in iterative reconstruction while increasing overlap beyond a certain amount, does not dramatically improve the condition numbers.

Figure 1.5: Four different rectangular masks in time-frequency domain and the first eigenfunctions of the corresponding localization operators. Middle plots show the absolute value squared of the STFT and right plots show the real part.
1.3.4 Frames adapted to a phase-space cover - [iv]

When representing a signal (or image) in a time-frequency dictionary, for example an NSG frame or quilted frame, the choice of atoms usually relies on an underlying lattice or a family of underlying lattices. The choice of the lattice together with the characteristics (shape, width) of the basic window or family of windows determines the ability of the representation to localize certain signal components and, further, the possibility to separate them. Various approaches have been taken to circumvent the restrictions possibly imposed by a rigid application of lattice structure, for example methods using reassignment [AF95, CMDAF97, FG04] or the recently introduced synchrosqueezed transform [WFD11, DLW11]. A different option is to consider eigenfunctions of time-frequency localization operators, whose concentration is no more restricted to be located at lattice points, and projections onto corresponding appropriate eigenspaces. By definition, the eigenfunctions corresponding to high eigenvalues of the localization operators are maxi-
Figure 1.7: Chirp-shaped mask, absolute value squared of the STFT of projection of random noise onto most localized resulting eigenfunctions, real part of the projection.

mally localized within a (weighted) subfamily of the time-frequency shifted atoms; thus, they provide potentially better localization in a certain time-frequency region than the time-frequency atoms themselves. In [iv] is is shown, that new frames can be constructed by choosing a finite number of eigenfunctions of each localization operator corresponding to a partition of the time-frequency plane. Given a possibly irregular cover or partition of the time-frequency plane, we allow to vary the trade-off between time and frequency resolution along the time-frequency plane by constructing a frame with atoms whose time-frequency concentration follows the shape of the cover members. The adapted frames are constructed by selecting a family of functions maximizing their concentration in the region of the time-frequency plane corresponding to member of a given cover/partition. These adapted functions can be obtained as eigenfunctions of time-frequency localization operators, cf. [Dau88, Dau90].

**Definition 10.** Given a compact set $\Omega \subseteq \mathbb{R}^{2d}$ in the time-frequency plane, the
time-frequency localization operator $H_{\Omega}$ is defined as

$$H_{\Omega}f(t) = \int_{\Omega} \mathcal{V}_\varphi f(x, w) M_w T_x \varphi(t) dx dw.$$  \hfill (1.18)

$H_{\Omega}$ is self-adjoint and trace-class and thus has a spectral decomposition

$$H_{\Omega} f = \sum_{k=1}^{\infty} \lambda_k \langle f, \phi^\Omega_k \rangle \phi^\Omega_k.$$  

The first $N$ eigenfunctions of $H_{\Omega}$ form the orthonormal set in $L^2(\mathbb{R}^d)$ that maximizes the quantity $\sum_{j=1}^{N} \int_{\Omega} |V_\varphi \phi^\Omega_j(z)|^2 \, dz$ among all orthonormal sets of $N$ functions in $L^2(\mathbb{R}^d)$. In this sense, their time-frequency profile is optimally adapted to $\Omega$. Figure 1.5 illustrates this principle by showing some time-frequency boxes $\Omega$ along with the STFT and real part of the corresponding localization operator’s first eigenfunctions, while Figure 1.6 and Figure 1.7 show that more general and exotic shapes may also be chosen, for example corresponding to chirped components. In the construction of the new frames, we allow for covers that are arbitrary in shape as long as they satisfy a mild admissibility condition, $B_r(\gamma) \subseteq \Omega \gamma \subseteq B_R(\gamma)$ with $\Gamma$ a lattice and $R >> r > 0$. The following theorem is a sample of the results proved in [iv].

**Theorem 1.3.4.** Let $\{\Omega_\gamma : \gamma \in \Gamma\}$ be an admissible cover of $\mathbb{R}^{2d}$. Then, there exists a constant $C > 0$ such that for every choice of $N_\gamma$, $C|\Omega_\gamma| \leq N_\gamma \leq N < \infty$, the family of functions $\{\phi_\gamma^\Omega_k : \gamma \in \Gamma, 1 \leq k \leq N_\gamma\}$ is a frame of $L^2(\mathbb{R}^d)$.

For proving the above theorem and analogue versions in the context of time-scale analysis (Theorem 8) and discrete time-frequency representations (Theorem 9), we work with an abstract model for phase space that allows for a variety of settings.

The proof of the main result in this paper is based on the following observations. The norm equivalence

$$\|f\|_2^2 \approx \sum_{\gamma} \|H_{\Omega_\gamma} f\|_2^2,$$  \hfill (1.19)

holds, provided that $\{\Omega_\gamma : \gamma \in \Gamma\}$ is an admissible cover of $\mathbb{R}^{2d}$. In the paper, we consider more general multipliers $\eta_\gamma$ than just characteristic functions. The inequalities (1.19) were first proved in [DFG06] for symbols of the form $\eta_\gamma(z) =$
$h(z - \gamma)$ and $\Gamma = \mathbb{Z}^{2d}$, then for a general lattice in \textit{[vi]}, and finally for fully irregular symbols, satisfying $\sum_\gamma \eta_\gamma(z) \approx 1$, in \textit{[Rom12]}. As we will explain in Section \textit{1.4.2} the proofs in \textit{[vi]} are based on the observation that under the condition $\sum_\gamma \eta_\gamma(z) \approx 1$, the norm-equivalence \textit{(1.19)} is equivalent to the fact that finitely many eigenfunctions of the operator $H_h$ generate a multi-window Gabor frame over the lattice $\Gamma$. On the other hand, the proof in \textit{[Rom12]} does not rely on spectral properties of the operators $H_{\eta_\gamma}$ nor on tools specific to Gabor frames on lattices. Thus the question arose whether the irregular case allows for the construction of a frame consisting of eigenfunctions of $H_{\eta_\gamma}$ and was given a positive answer in this paper by observing that \textit{(1.19)} remains valid when the operators $H_{\eta_\gamma}$ are replaced by finite rank approximations $H_{\eta_\gamma}^\varepsilon$ obtained by thresholding their eigenvalues. This finite rank approximation is in turn achieved by showing the following extension of \textit{(1.19)}

$$\|f\|_2^2 \approx \sum_\gamma \|H_{\eta_\gamma}f\|_2^2 \approx \sum_\gamma \|(H_{\eta_\gamma})^2 f\|_2^2.$$  \textit{(1.20)}

As a by-product (Lemma 1), it is shown that the operators $H_{\eta_\gamma}$ have infinite rank (even if $\eta_\gamma$ is the characteristic function of a compact set), hence $\|H_{\eta_\gamma}f\|_2 \not\approx \|(H_{\eta_\gamma})^2 f\|_2$ and the global properties of the family $\{(H_{\eta_\gamma})^2 : \gamma \in \Gamma\}$ are crucial to prove \textit{(1.19)}. The transition from the operators $H_{\eta_\gamma}$ to $(H_{\eta_\gamma})^2$ is made possible by the fact that, while the latter are not time-frequency localizations operators themselves, they preserve the time-frequency localizing behavior of $H_{\eta_\gamma}$. This property is made precise by the notion of a family of operators that is \textit{well-spread in the time-frequency plane} and exploited by using tools from \textit{[Rom12]} are valid for these operator families.

\section{1.4 Localization operators and Gabor multipliers}

Time-frequency localization operators, as defined in \textit{(1.18)}, transform a signal into one that is localized in a given region by reducing the signal energy outside that region to a negligible amount. In \textit{[iv]}, these operators were introduced in order to serve as a tool to derive frames with certain time-frequency localization properties; in this section, we will first take a deeper look in the analytical properties of
localization operators and summarize paper [vi] in Section 1.4.1, which contains the converse of a classical result by Daubechies on the spectral properties of localization operators, cf. [Dau88], which states that the eigenfunctions of time-frequency localization operators with the Gaussian as analysis window and concentric multiplier functions $\eta$ are given, independently of the precise shape and radius of $\eta$, by the Hermite functions.

The first approach to time-frequency localization, introduced in 1961, consists in separately selecting time- and frequency-content by projecting onto bounded intervals, and is described in a famous series of papers known as the “Bell labs papers”, [Sle83]. In this case, the eigenfunctions are the prolate spheroidal functions [SP61].

The classical time-frequency localization operators are then defined, in generalization of (1.18), by using a symbol or multiplier $\eta \in L^1(\mathbb{R}^d)$, as

$$H_\eta f = V_\omega^* \eta V_\omega f. \quad (1.21)$$

While, as a consequence of the uncertainty principle, no signal can possibly have compact support in the time-frequency domain, the properties of localization operators motivate a different point of view on time-frequency localization, compatible with the uncertainty principle in a certain sense: if $\eta$ is assumed to have compact support or, at least, to be well concentrated in $\Omega$, e.g. $\eta = \chi_\Omega$, where $\chi_\Omega$ is the characteristic function on $\Omega$, then $H_\eta f$ can be interpreted as the part of $f$ that ”essentially” lives in $\Omega$. Following the introduction by Daubechies in [Dau88], localization operators have been studied by Wong [Won99], Ramanathan and Topiwala [RT93, RT94]. More recently, the properties of localization operators using modulation spaces as classes for symbols and windows have been investigated for their boundedness and spectral properties, cf. [CG03, CG07, Pan09] and references therein.

The intuitive understanding of the localization operators suggests a partitioning of a signal in the time-frequency domain, similar to a partition in time- or frequency domain. Although, as opposed to the latter, no orthogonal partitions can be obtained by partitioning in the time-frequency plane, a characterization of functions can be obtained in a manner similar to the characterization of $L^p$-spaces in the Littlewood-Paley sense. Work on this problem is found in [DFG06, Rom12] and the paper [vi], which is included in this thesis and summarized in Section 1.4.2.
While the theoretical results introduced so far, start from time-frequency localization operators defined via the continuous STFT, time-frequency multipliers that occur in practice usually rely on a discretization of the time-frequency domain and are thus related to Gabor frames. The resulting operators are called Gabor multipliers and were first discussed in [FN03, Dör02]. Similar to the definition of continuous time-frequency localization operators, Gabor multipliers are defined by modification (weighting) of the discrete time-frequency coefficients obtained via Gabor frames.

**Definition 11 (Gabor Multiplier).** Let $C_{\varphi,\Lambda}$ denote the analysis operator corresponding to a Gabor frame with window $\varphi$ and lattice $\Lambda$ and consider the synthesis operator $C_{\phi,\Lambda}^*$ with a possibly different window $\phi$. Further, $m \cdot C_{g,\Lambda} f$ denotes the pointwise multiplication of $C_{\varphi,\Lambda} f$ with the symbol $m \in \ell^\infty(\Lambda)$. Then, a Gabor multiplier is defined as

$$G_m : f \in \mathcal{H} \mapsto G_m f = C_{\phi,\Lambda}^* (m \cdot C_{\varphi,\Lambda} f).$$

(1.22)

In signal processing, in particular speech and audio processing, the implicit usage of Gabor multipliers is common practice, consider [LW09, BLED10, OKMT10, AFK+07] for some recent work. The papers [vii, viii] included in this thesis extend and complement work on Gabor multipliers in the spirit of [BP06, FHK04, Grö11].

### 1.4.1 An inverse problem for localization operators - [v]

In [Dau88], Daubechies considered time-frequency localization operators with the Gaussian window $g(t) = \varphi(t) = 2^{1/2} e^{-\pi t^2}$, and indicator function $\sigma(z) = \chi_\Omega(z)$ of a set $\Omega \subset \mathbb{R}^2$, and investigated the eigenvalue problem

$$H_\Omega f := H_{\chi_\Omega,\varphi} f = \lambda f.$$

(1.23)

If $\Omega$ is a disc centered at zero, she concluded, that the eigenfunctions of $H_{\chi_\Omega,\varphi}$ are the Hermite functions. Since its solutions are the functions with best concentration in $\Omega$, problem (1.23) is important in time-frequency analysis. Here, as before, we consider the time-frequency concentration of a function $f$ in $\Omega \subset \mathbb{R}^2$ defined as

$$C_{\Omega}(f) = \frac{\int_\Omega |V_{\varphi} f(z)|^2 dz}{\|f\|_{2}^2}.$$  

(1.24)
In [V], we consider the inverse situation of the one considered by Daubechies and are led to the following question:

*Given a localization operator with unknown localization domain $\Omega$, can we recover the shape of $\Omega$ from information about its eigenfunctions?*

We call this new type of inverse problem the “inverse problem of time-frequency localization” and solve it in the case where explicit computations can be made, in analogy with the set-up of [Dau88]. Our main contribution is the following.

**Theorem 1.4.1.** Let $\Omega \subset \mathbb{R}^2$ be simply connected. If one of the eigenfunctions of the localization operator $H_\Omega$ is a Hermite function, then $\Omega$ must be a disk centered at 0.

This result and its variant for time-scale analysis, given in Theorem 2 in the paper, have interesting consequences, both theoretical and in applications. We mention applications in UWB communication, where Hermite functions are commonly suggested as modulation pulses, as well as system identification.

Some theoretical consequences of Theorem 1 are formulated in Corollary 2, concerning the connection between Gaussian growth of eigenfunctions of a localization operator and the shape of the localization domain. Corollary 3 then gives the connection between double orthogonality of orthonormal bases in phase-space and the possible solutions of (1.23).

Theorem 1 is also extended to more general localization domains by means of Corollary 1. A consequence for the sub-optimality of the usage of Gaussian windows in Gabor frames with rectangular lattices is formulated in Remark 1 and it is conjectured that hexagonal lattices are optimal sampling sets for Gabor frames based on Gaussian windows, cp. [SB03].

The proof of the main statements in Theorem 1 and Theorem 2 is based on the observation that both problems can be understood as special cases of a more general formulation with general radial measures on complex domains. A corresponding problem, stated in Proposition 1 is solved by weighted monomials, which, in turn, form a doubly orthogonal set in the complex plane with respect to radial measures. It had been observed before, and was adapted and generalized to our needs in Proposition 2, that this double orthogonality of the monomials force the localization domain to be a disk, if it is assumed to be simply connected. More
general domains are treated in Corollary 1. Then, following a classical approach in
time-frequency analysis, the Bargmann transform is used to transform the corre-
sponding problem for time-frequency localization to the Bargmann-Fock space of
analytic functions. The statement of Theorem 1 then follows from Proposition 2.
Equivalently, while slightly more involved, a similar statement on wavelet local-
ization operators is obtained by applying the Bergman transform. The role of the
Hermite functions is taken over by special functions, whose Fourier transforms are
the Laguerre functions and the corresponding localization domains are proved to
be pseudohyperbolic discs centered at $i \in \mathbb{C}$.

1.4.2 Time-Frequency partitions and characterizations of
modulations spaces with localization operators - [VI]

In this paper we do not investigate just a single localization operator $H_\eta$, but the
behavior of an entire collection of localization operators; given a lattice $\Lambda \subseteq \mathbb{R}^{2d}$ of
the time-frequency plane, we consider the collection of operators $\{H_{T_\lambda \eta} : \lambda \in \Lambda\}$
and the mapping $f \rightarrow \{H_{T_\lambda \eta}f\}$. The intuitive idea that, if the supports of $T_{\lambda \eta}$
cover $\mathbb{R}^{2d}$, then $\{H_{T_\lambda \eta}f, \lambda \in \Lambda\}$ should contain enough information to recover
$f$ from its local components, is made precise by deriving a new characterization
of modulation spaces from sequence space norms applied to $L^2$-norm of the lo-
calized pieces. Modulation spaces are smoothness spaces, for which smoothness
is measured by means of time-frequency distribution rather than differences and
derivatives, as in the case of Besov spaces.

Definition 12 (Modulation space). Fix a non-zero function $\varphi$ in the Schwartz
space $S(\mathbb{R}^d)$ and a weight function $m$ on $\mathbb{R}^{2d}$ that satisfies $m(z_1 + z_2) \leq C(1 +
|z_1|)^Nm(z_2)$ for some constants $C, N \geq 0$ and all $z_1, z_2 \in \mathbb{R}^{2d}$. Then a tempered
distribution $f$ is in $M^p_m(\mathbb{R}^d)$, if

$$
\|f\|_{M^p_m} = \left( \int_{\mathbb{R}^{2d}} |\mathcal{V}_\varphi f(z)|^p m(z) \, dz \right)^{1/p} < \infty. \quad (1.25)
$$

As a special case of the main theorem of citedogr11 is a follows.

Theorem 1.4.2. A tempered distribution $f$ satisfies

$$
\left( \sum_{\lambda \in \Lambda} \|H_{T_\lambda \eta}f\|^p m(\lambda)^p \right)^{1/p} < \infty \quad (1.26)
$$
if and only if $f \in M^p_m(\mathbb{R}^d)$.

In other words, the expression in (1.26) (using the time-frequency localized components of $f$) is an equivalent norm on $M^p_m(\mathbb{R}^d)$.

In [vi], we prove the norm equivalence of the above theorem for a large class of modulation spaces and arbitrary time-frequency lattices. As mentioned before, an analogous result was derived in [DFG06] for lattices with integer oversampling, with the main argument based on Zak transform methods. For a general lattice, these methods are no longer available, and a completely new approach was developed; as a by-product of the new techniques we have found several results of independent interest.

- Several structural results and characterizations of Gabor frames for multi-window Gabor frames in generalization of the results from [Grö07]

- A finite intersection property for time-frequency invariant subspaces of the distribution space $M^\infty(\mathbb{R}^d)$, resembling the finite intersection property that characterizes compact sets.

- A new, independent proof for the existence of multi-window Gabor frames with well-localized windows with additional insight how the windows can be chosen.

- Precise estimates for the localization of the eigenfunctions of a localization operator.

It should be noted, that the statement of Theorem 1.4.2 is closely related to the work presented in [iv, vi] by the fact that the main tool in its proof relies on the properties of the spectral representation of the operators $H_{T,\chi,\eta}$ and the fact, proved in Lemma 9 in paper [vi], that finitely many of their eigenfunctions generate a multi-window Gabor frame over the corresponding lattice $\Lambda$. As mentioned before, using methods from [Rom12], where the statement of Theorem 1.4.2 is generalized to sampling sets without a lattice structure, this finite generation property is then also generalized to entirely arbitrary covers of phase-space in [vi].
1.4.3 Representation of operators in the time-frequency domain and generalized Gabor multipliers - [vii]

The work in this paper is inspired by a general operator representation in the time-frequency domain via a twisted convolution. It turns out, that this representation, respecting the underlying structure of the Heisenberg group, has an interesting connection to the so-called spreading function representation of operators.

Given a linear operator $T$ on $L^2(\mathbb{R})$ with integral kernel $\kappa_T$, its spreading function $\eta_T$ is defined as

$$\eta_T(b, \nu) = \int_{\mathbb{R}} \kappa_T(t, t-b) e^{-2\pi i \nu t} dt,$$

An operator’s spreading function comprises the amount of time-shifts and modulations, i.e. of time-frequency-shifts, effected by the operator. Its investigation is hence decisive in the study of time-frequency multipliers and their generalizations. Although no direct discretization of the continuous representation by an operator’s spreading function is possible, the twisted convolution turns out to play an important role in the generalizations of time-frequency multipliers. In the main section of this article, we introduce a general model for multiple Gabor multipliers (MGM), which uses several synthesis windows simultaneously. By adapting the masks corresponding to each of the synthesis windows, more general operators than by regular Gabor multipliers may be well-represented.

For the sake of generality, most statements in [vii] are given in a Gelfand-triple $GT = (S_0, L^2, S'_0)$, compare [DFG06, CFL08, Fei09], rather than a pure Hilbert space setting. This approach is motivated by the fact that many important operators and signals may not be described in a Hilbert-space setting, starting from simple operators as the identity. Furthermore, by using distributions, continuous and discrete concepts may be considered in a unified framework.

In the context of Gabor multipliers, one is interested in the following rank one operators (oblique projections), which can be considered as the basic building blocks of Gabor multipliers: Let $g, h \in S_0$ be such that $\langle g, h \rangle = 1$. Let $\lambda \in \Lambda$, and consider $P_{\varphi, \phi, \lambda}$ defined by

$$P_{\varphi, \phi, \lambda} f = \varphi^*_\lambda \otimes \phi_\lambda f = \langle f, \varphi_\lambda \rangle \phi_\lambda, \quad f \in (S_0, L^2, S'_0).$$

(1.28)
. Using these operators, (1.22) is rephrased as
\[ G_m = \sum_{\lambda \in \Lambda} m(\lambda) P_{\varphi,\phi,\lambda}. \]  
(1.29)

If several synthesis windows \( \phi_j, j \in \mathcal{I} \), are used, we define multiple Gabor multipliers (MGM) by
\[ M_m = \sum_{j \in \mathcal{I}} \sum_{\lambda \in \Lambda} m^j(\lambda) P_{\varphi,\phi_j,\lambda}. \]  
(1.30)

The main contributions of this paper may be summarized as follows:

• It is shown that the STFT of any operator with spreading function in \( GT \) may be represented by a twisted convolution in phase-space (Proposition 10).

• The spreading function of an STFT multiplier as well as a Gabor multiplier is derived and shown to be the product of the (continuous or discrete) symplectic Fourier transform of the multiplier’s symbol with the STFT of the synthesis window with respect to the analysis window (Lemma 14).

• A method to obtain the multiplier \( m \) for the best approximation of a given operator (whose spreading function is known) by a Gabor multiplier (Theorem 19) or a multiple Gabor multiplier (Proposition 29), as well as corresponding error estimates for the Gabor multiplier case\(^2\) (Corollary 20) are derived. Note that these methods can be directly used to implement the (computationally efficient) approximation algorithms.

• Previous work [BP06] gave a characterization of the Riesz basis property of the projections operators \( P_{\varphi,\phi,\lambda}, \lambda \in \Lambda \). Proposition 28 extends this characterization to several windows \( \phi_j, j \in \mathcal{I} \), that is, gives a characterization of the Riesz basis property of the family of projection operators \( \{P_{\varphi,\phi_j,\lambda}, j \in \mathcal{I}, \lambda \in \Lambda \} \).

• Theorem 39 shows that, when specifying to a separable mask in the modification of time-frequency coefficients within MGM, as well as a specific sampling lattice for the synthesis windows, it turns out, that the MGM reduces to one or the sum of a finite number of regular Gabor multipliers with adapted synthesis windows.

The paper is concluded by a few concrete examples.

\(^2\)Error estimates for the generalized Gabor multiplier approximation are derived in [viii].
1.4.4 Representation of operators by sampling in the time-frequency domain - [viii]

This paper continues the work presented in [vii], in particular, error estimates for the best approximation of general operators by means of Gabor multipliers and multiple Gabor multipliers (MGM, cf. (1.30)) are derived and refined. While, in signal processing, the manipulation of given signals in the time-frequency domain is common practice, consider [LW09, BLED10, OKMT10] for some recent work, time-frequency multipliers have rather rarely been studied with a focus on the influence of the various involved parameters on the outcome of methods based on time-frequency analysis of signals. This paper gives some important guidelines in this direction.

The spreading representation of an operator, introduced in the previous section, provides a novel kind of insight in certain operators’ behavior since it reflects the operator’s action in the time-frequency domain. In [viii], we developed an approach that uses the spreading representation of time-frequency multipliers in order to optimize the parameters involved. More specifically, in the one-dimensional, continuous-time case, given a linear operator \( T \) on \( L^2(\mathbb{R}) \) with integral kernel \( \kappa_T \) and spreading function \( \eta_T \) as defined in (1.27), we aim at modeling the operator by its action on the sampled short-time Fourier transform (STFT) or Gabor coefficients.

Purely multiplicative modification of the Gabor coefficients \( C_{g,\Lambda} f(\lambda) \) leads to the definition of classical Gabor multipliers, for which the linear operator applied to the coefficients is diagonal. We also consider generalizations of the classical Gabor multipliers by relaxing the restriction to diagonality in order to achieve good approximation for a wider class of operators. Also, in certain approximation tasks, the drawback resulting from coarse sub-sampling, i.e. large \( b_0 \) and/or \( \nu_0 \), can, to a certain extent, be compensated by using two or three instead of just one synthesis window. This is the case of MGMs defined in (1.30).

In [viii], we derive error estimates for the approximation of operators by multiple and generalized Gabor multipliers, based on the operator’s spreading function. The results give some insight in the choice of the parameters involved in approximation, in particular, of the windows and the lattice constants \( b_0, \nu_0 \).
More in detail, after a brief introduction to Gabor multipliers in Section 2, Section 3 presents, in Proposition 1, a general error estimate for the approximation of Hilbert-Schmidt operators by MGM, as well as several special cases. As a noteworthy special case, the approximation of short-time Fourier multipliers by Gabor multipliers is considered. A simplified version of this result is re-stated here as an example.

Denote by $S_a$ a short time Fourier multiplier, defined by $S_a : f \mapsto S_a f = V'_h(a \cdot V_g f)$, where $V_g$ and $V_h$ denote short time Fourier transforms with respect to windows $g$ and $h$ respectively, and the symbol, denoted by $a$, is a function on the whole time-frequency space.

The spreading function of an STFT-multiplier is given by $\eta_{S_a} = \tilde{a} \cdot V_g h$. Here, $\tilde{a}$ is the continuous symplectic Fourier transform of $a$.

Furthermore, for a given lattice $\Lambda$, we denote by $\Lambda^\circ$ the adjoint lattice, by $\Omega^\circ$ the corresponding fundamental domain, and by $\Pi^\circ$ the periodization operator with respect to $\Lambda^\circ$. For example, for a product lattice of the form $\Lambda = \nu_0 \mathbb{Z} \times b_0 \mathbb{Z}$, we have $\Lambda^\circ = t_0 \mathbb{Z} \times \xi_0 \mathbb{Z}$ with $t_0 = 1/\nu_0$, $\xi_0 = 1/b_0$, and $\Pi^\circ f(\zeta) = \sum_{\lambda^\circ \in \Lambda^\circ} f(\zeta + \lambda^\circ)$, $\zeta \in \Omega^\circ$.

**Proposition 1.4.3.** Let $T = S_a$ be a STFT multiplier with spreading function $\eta_T = \tilde{a} \cdot V_g h$, and denote by $T' = G_m$ its best Gabor multiplier approximation with the same windows, and lattice $\Lambda$, as defined in (5) of Theorem 1 in [viii].

1. The approximation error is given by

$$
\|T - T'\|_{HS}^2 = \int_{\Omega^\circ} \left[ \Pi^\circ((\tilde{a} \cdot V_g h)^2)(\zeta) - \frac{\Pi^\circ(\tilde{a} \cdot |V_g h|^2)(\zeta)^2}{\Pi^\circ(|V_g h(\zeta)|^2)} \right] d\zeta \quad (1.31)
$$

2. Furthermore, the difference between the best approximation $T'$ and the Gabor multiplier $T''$ obtained from the $\Lambda$-subsampled version $a|\Lambda$ of the symbol $a$ is given by

$$
\|T'' - T'\|_{HS}^2 = \int_{\Omega^\circ} \left| \Pi^\circ(\tilde{a})(\zeta)^2 \Pi^\circ(|V_g h|^2)(\zeta) - \frac{\Pi^\circ(\tilde{a} \cdot |V_g h|^2)(\zeta)}{\Pi^\circ(\tilde{a})(\zeta) \Pi^\circ(|V_g h|^2)(\zeta)} \right|^2 d\zeta . \quad (1.32)
$$

From the error estimates given in Section 3 and the corresponding operator descriptions, guidelines for the choice of good parameters are discussed in Section 4. In particular, it is observed, that the sampling lattice $\Lambda$ used in the
operator’s approximation, may be adapted to the eccentricity of the spreading function according to the obtained error expressions. For generalized Gabor multipliers, allowing for several synthesis windows, it is remarked that choosing a few eigenfunctions of an STFT-multiplier corresponding to the fundamental domain \( \Omega^o \) of \( \Lambda^o \) can significantly reduce aliasing terms in the approximation of under-spread operators by generalized Gabor multipliers. This idea was already noted in [HMKK00] in a slightly different context. Corollary 1 in [viii] now provides an explanation why these particular functions, by their property of being maximally concentrated inside \( \Omega^o \), represent a good choice for the synthesis windows of multiple Gabor multipliers. The connection to our work presented in [vi] and [iv] should be noted. Finally, Section 5 gives various insightful numerical experiments on the best approximation of various operators by both classical and generalized Gabor multipliers. Thereby, the influence of the various parameters, in particular the number of synthesis windows and the lattice parameters, is empirically investigated. The experiments show that the choice of these parameters has considerable influence on the performance of approximation by (generalized) Gabor multipliers.

1.5 Applications of time-frequency analysis in sound analysis

Sound signals play a central role in human life and the manner sound is perceived is highly sophisticated, complex and context-dependent. Time-frequency analysis has implicitly played an important role in (digital) audio processing since its beginning. The amount of sound data that are stored, searched and processed, grows dramatically, thus, there is also a growing need for sophisticated processing methods, which require an understanding of the inherent structures of sound and their implications for human listeners.

Ideally, an analysis tool should be able to render a representation that allows for visual display reflecting a user’s acoustical impression. In the subsequent sections we will describe several newly designed analysis tools, that render more sophisticated representations of sound signals than classical representations such as the short-time (or sliding window ) Fourier transform.

In particular, in [ix], Section 1.5.1 and [x], Section we show how adaptiv-
ity in the transformation parameters can sharpen the visual display while assuring a perfect connection between signal and representation in the sense of invertibility.

Then, in [x1], Section 1.5.3 and [x2], Section 1.5.4, we show how various Bayesian coefficient priors enable us to highlight particular structures by means of informed analysis.

In the sense of reproducible research, all software involved in the production of the simulation examples is available, along with many additional examples and sound files, on the webpages [http://www.univie.ac.at/nonstatgab/](http://www.univie.ac.at/nonstatgab/) and [http://homepage.univie.ac.at/monika.doerfler/StrucAudio.html](http://homepage.univie.ac.at/monika.doerfler/StrucAudio.html).

**Adaptivity in time-frequency transforms**

Adaptive signal transformations have been introduced in Section 1.3 from a rather theoretical point of view. In applications, new aspects, such as computational efficiency in analysis and reconstruction must be taken into account. Section 1.5.1 and Section 1.5.2 summarize recent work on the application of nonstationary Gabor frames to the development of a new algorithm of considerable importance for the audio processing community, namely the implementation of an invertible, and an invertible, real-time CQ-transform.

**Sparsity in Signal Processing**

A wide range of in signal processing applications have benefited from sparsity, which has become an increasingly popular topic since the mid 90’s. Introduced by Chen, Donoho and Saunders [CDS98], the idea behind sparsity is the efficient representation of a signal as linear combination of elementary atoms chosen from an appropriate dictionary, where efficiently is understood in the sense that only few atoms are needed to reconstruct the signal. The same idea appeared in the machine learning community [Tib96], where often only few variables are relevant in inference tasks based on observations living in very high dimensional spaces. Sparsity of coefficients may be enforced by \(\ell^1\)-regression, the solution is given by the Lasso [Tib96], which, given a noisy observation \(y = s + e\) in \(\mathbb{C}^L\) finds

\[
\hat{c} = \arg \min_{c \in \mathbb{C}^L} \frac{1}{2} \|y - \Phi c\|_2^2 + \lambda \Psi(c) \tag{1.33}
\]
with penalty term \(\Psi(\cdot) = \|\cdot\|_1\) and \(\lambda > 0\). Since, for Gabor frames, the sequence \(c_{k,l}\) is ordered along two dimensions, the \(\ell^1\)-prior \(\Psi\) in (1.33) may be replaced by a two-dimensional mixed norm \(\ell^{p,q}\) which acts differently on groups (indexed by \(g\) in the sequel, may be either time or frequency) and their members (indexed by \(m\)):

\[
\Psi(c) = \|c\|_{p,q} = \left(\sum_g \left(\sum_m |c_{g,m}|^p\right)^{q/p}\right)^{1/q}
\]  

(1.34)

Of particular interest are the cases \(p = 2, q = 1\) and \(p = 1, q = 2\). The former is known as Group-Lasso \[YL06\] (promoting sparsity in groups and diversity in members) and the latter was termed Elitist-Lasso in \[TK08\]: the \(\ell^{1,2}\) constraint promotes sparsity in members, only the “strongest” members (relative to an average) of each group are retained.

Landweber iterations, which solve (1.33) in the \(\ell^1\)-case, \[DDDM04\], also yield a solution of the novel problem if \(S_\lambda\) is replaced by a generalized thresholding operator

\[
S_{\lambda,\xi}(z_{k,j}) = z_{k,j}(1 - \xi(z_{k,j}))^+\text{, where }\xi = \xi_{(k,j),\lambda}\text{ is a non-negative function dependent on the index } (k,j) \text{ and } \lambda.
\]

The solution to (1.33) is then given by the iterative Landweber algorithm: choosing arbitrary \(c^0\), set

\[
c^{n+1} = S_{\lambda,\xi}(c^n - \Phi^*(y - \Phi c^n))
\]  

(1.35)

It was shown in \[Kow09\], that the solution of

\[
\hat{c} = \arg\min_{c \in \mathbb{C}^p} \frac{1}{2}\|y - \Phi c\|_2^2 + \lambda\|c\|_{p,q}
\]  

(1.36)

is given by the limit of the iterative sequence (1.35) with threshold operators

\[
S_\lambda(z) = z_\gamma(1 - \xi_\lambda(z))^+
\]

defined via \(\xi\) as follows:

\[
p = 1, q = 1 : \xi^L(c_{g,m}) = \frac{\lambda}{|c_{g,m}|} \quad \text{(Lasso)}
\]  

(1.37)

\[
p = 2, q = 1 : \xi^{GL}(c_{g,m}) = \frac{\lambda}{\left(\sum_m |c_{g,m}|^2\right)^{1/2}} \quad \text{(Group-Lasso)}
\]  

(1.38)

\[
p = 1, q = 2 : \xi^{EL}(c_{g,m}) = \frac{\lambda}{1 + M_g \lambda |c_{g,m}|} \quad \text{(Elitist-Lasso)}
\]  

(1.39)

where \(c_g = (c_{g,1}, \ldots, c_{g,M_g})\) and \(c'_g\) denotes for each group \(g\) the sequence of scalars \(|c_{g,m}|\) in descendant order. \(M_g\) denotes some natural number depending on the magnitudes of coefficients in the group \((c_{g,1}, \ldots, c_{g,M})\).
Our contributions, described in Section 1.5.3 and Section 1.5.4, extend the known algorithms by introducing neighborhood-based thresholding operators, which allows the inherent structures of natural signals to be taken into account.

1.5.1 Constructing an invertible constant-Q transform with nonstationary Gabor frames - [ix]

While most widely used signal transforms, in particular, Fourier transform based methods such as the STFT, lead to a frequency resolution that does not depend on frequency, the constant-Q transform (CQT), introduced by Brown [Bro91], transforms a time signal into the time-frequency domain with geometrically spaced center frequencies of the frequency bins. In this setting, similar to a wavelet transform, the Q-factor (the ratio of the center frequencies to the window’s bandwidth) is constant, hence the representation allows for a finer frequency resolution at lower frequencies and a finer time resolution at the higher frequencies. This kind of resolution is sometimes preferable to the fixed resolution of the standard STFT or Gabor transforms, for which the frequency bins are linearly spaced, for both perceptual and structural reasons; on the one hand, a non-linear frequency-resolution is closer to the resolution of the human auditory system, on the other hand, the parameters of the CQT can be set to coincide with the temperament, e.g. semitone or quarter tone, used in Western music.

However, the originally introduced constant-Q transform is neither computationally efficient, nor does it provide a satisfactory way for inversion, that is, the original signal or its components can not be reliably reconstructed from the CQT-coefficients. The lack of invertibility makes the convenient modification of CQT-coefficients with subsequent re-synthesis required in complex music processing tasks such as masking or transposition impossible. This may be the main reason why the CQT has not been widely used in applications to date.

A computationally efficient CQT was proposed in [SK10], but inversion introduced in this method is approximate and leads to reconstruction errors that may be unacceptable, in particular for moderate redundancy of around 4.

In this paper, we construct an invertible nonstationary Gabor transform by resorting to NSG frames with adaptive resolution in frequency. Figure 2 in [ix] illustrates the time-frequency sampling grid of the set-up, the center frequencies
being geometrically spaced and time sampling points regularly spaced in each frequency band.

It should be mentioned, that this paper has been highly acclaimed in the audio processing community; in particular, it received the Golden Best student paper award at DAFx-11 [http://dafx11.ircam.fr/wordpress] and, since its appearance has already been cited around 15 times.

The central artifice of the painless NSG frames to allow for fast, FFT-based implementation, is the design of the involved transform parameters in such a way that the frame operator is diagonal with the diagonal bounded away from zero. In frequency-adaptive NSG frames, the frame operator is diagonal in the frequency domain by careful design of the window supports and the corresponding time-shift parameters, as detailed in Section 3 of [ix] for the particular setting of a CQT.

The principal idea is to consider the family of windows $\varphi_k, k \in \mathbb{Z}$, band-limited, with Fourier transforms $\psi_k = \hat{\varphi}_k$ centered around irregularly spaced frequency points $\xi_k$ and to choose frequency dependent time-shift parameters $a_k$ such that, if the support of $\hat{\varphi}_k$ is contained in an interval of length $|I_k|$, we have $a_k \leq \frac{1}{|I_k|}$ for all $k$, i.e., the time-sampling points must be chosen dense enough. Then, having obtained the frame members by setting $\varphi_{n,k} = T_{na_k}\varphi_k$, the frame operator is diagonal in the Fourier domain:

$$Sf = \mathcal{F}^{-1}(\sum_k \frac{1}{a_k}|\hat{\varphi}_k|^2\hat{f}).$$

(1.40)

This corresponds to investigating the operator $\mathcal{F}S\mathcal{F}'$, which acts on the Fourier transform of a signal of interest.

The support conditions on $\hat{\varphi}_k$ immediately imply

$$0 < A \leq \sum_k \frac{1}{a_k}|\hat{\varphi}_k|^2 \leq B < \infty$$

(1.41)

almost everywhere and thus the invertibility of the frame operator. Then, the dual frame, used in the reconstruction process, is given by the elements

$$\gamma_{n,k} = T_{na_k} \left[ \mathcal{F}^{-1}(\hat{\varphi}_k/\sum_l \frac{1}{a_l}|\hat{\varphi}_l|^2) \right].$$

The given framework allows for fast realization of the CQT by considering the Fourier transform of the input signal. The CQ- coefficients are then obtained as

$$c_{n,k} = \langle f, T_{na_k}\varphi_k \rangle = \langle \hat{f}, M_{-na_k}\hat{\varphi}_k \rangle,$$
and can be calculated, for each $k$, with an FFT of length determined by the support of $\hat{\varphi}_k$. Reconstruction is similarly realized by applying the dual windows in a simple overlap-add process.

The paper gives several examples which are complemented by more extensive ones, including the sound files, on the web page [http://www.univie.ac.at/nonstatgab/](http://www.univie.ac.at/nonstatgab/), where all corresponding MATLAB and Python code may also be downloaded.

Finally, we note that the computation time of the NSG-based CQ-transform is better than the most recent fast CQT implementation [SK10], cf. the tables presented in the paper. However, the advantage of the NSG method in terms of computation efficiency decreases with increasing signal length. Since efficient and possibly real-time processing is desired, the next natural step is a method that processes the incoming samples in a piecewise manner. This method was investigated in [x], as summarized in the next section.

1.5.2 A framework for invertible, real-time constant-Q transforms - [x]

The work in this paper starts from the technique introduced in [ix] for the implementation of an invertible constant-Q transform based on NSG frames. The additional new ingredient is the piecewise processing of the incoming signal samples, which leads to bounded delay in, and thus linear, processing time. Therefore, the sliCQ (sliced constant-Q transform)-algorithm lends itself to real-time processing.

The article first introduces the concept of NSG systems, their properties and the conditions for these systems to constitute painless frames, as described in the previous section, are recalled. The CQ-transform by NSG frames (CQ-NSGT) with adaptivity in the frequency domain then is the starting point for the sliCQ transform. The general idea of sliCQ transform is the processing of incoming signal samples by applying CQ-NSGT in a blockwise manner, i.e. to (fixed length) slices of the input signal. This slicing process leads to two challenges:

On the one hand, the windows used for cutting the signal must be smooth and zero-padding has to be applied to suppress time aliasing and blocking artifacts, which may arise if the analysis coefficient are modified before reconstruction.
On the other hand, the coefficients obtained from the new transform should be equivalent to the CQ-coefficients obtained from a full-length CQ-NSGT. This is achieved to high precision by careful choice of both the slicing windows and the analysis windows used in the CQ-NSGT. The interpretation of the sliCQ-coefficients in relation to the full-length transform is investigated in Proposition 4.

The algorithms involved in sliCQ processing are described in detail and a pseudo code is presented for the 4 involved steps: Algorithm 1: NSG analysis, Algorithm 2: NSG synthesis, Algorithm 3: sliCQ analysis and Algorithm 4: sliCQ synthesis. The following proposition (Proposition 3 in the paper) contains the central statement of the paper: the sliCQ framework provides perfect reconstruction from its analysis coefficients.

**Proposition 1.5.1.** Let $G(g, a)$ and $\tilde{G}(\tilde{g}, a)$ be dual NSG systems for $\mathbb{C}^{2N}$. If $s$ is the output of sliCQ (Algorithm 3), then the output $\tilde{f}$ of isliCQ (Algorithm 4) equals $f$, i.e., $\tilde{f} = f$.

Furthermore, the transform’s numerical properties, in particular computation time and complexity, are analyzed and the approximation quality of the CQ-NSGT coefficients obtained by the sliCQ is investigated numerically.

Some examples for analysis and processing of real-life signals are shown in Section VI of the paper. In particular, a single component of the Glockenspiel signal, shown before, is extracted, transposed, and re-inserted in the signal, by using the sliCQ transform. Various transform settings, in particular different slice lengths, are used and compared. Corresponding sound files and complementary material, in particular for the windows that occur in the sliCQ transform, are found, together with the MATLAB code, on the accompanying web site [http://www.univie.ac.at/nonstatgab/slicq/index.html](http://www.univie.ac.at/nonstatgab/slicq/index.html).

### 1.5.3 Structured sparsity for audio signals - $\xi$

The natural measure of the cardinality of a support set, and hence its sparsity, is the $\ell_0$ “norm” which counts the number of non-zero coefficients. Minimizing such a norm leads to a combinatorial problem which is usually relaxed into a $\ell_1$ norm which is convex.

However, a property of sparse modeling based on $\ell_1$-priors is the implicit assumption that coefficients are independent, while most natural signals are highly
structured; the structures of a signal correspond to the physical priors which may be exploited for improved processing results. In particular, most audio signals of importance for humans, in particular speech and music, are highly structured in time and frequency. Typically, salient signal components are sparse in time (or frequency) and persistent in frequency (or time) in the sense that sparsity in time is connected to transient events, while sparsity in frequency is observed in harmonic components. These observed sparse structure can be enhanced by the introduction of sparsity criteria which take into account the two-dimensionality of the time-frequency representations used. A first step in this direction was the introduction of mixed norms cf. (1.34), on the coefficient arrays in order to enforce sparsity in one, and persistence in the other domain. Regression with mixed-norm priors was first proposed in [TK08] and the convergence properties of corresponding algorithms were investigated in [Kow09].

In our contribution [xi], we consider a family of regression problems involving mixed $\ell^1$ and $\ell^2$ priors on the coefficients obtained from (tight) Gabor frames $\varphi_{k,j} = M_{bj}T_{ka}\varphi, j = 0, \ldots, J-1, k = 0, \ldots K-1$, and refine algorithms which exist for regression problems with mixed norm priors [TK08, Kow09], which were introduced in the introduction to this section, by using local neighborhood-weighting. More precisely the coefficient $c_{a,m}$ are thresholded according to the energy of their time-frequency neighborhood, which, in contrast to the groups of Group- and Elitist-LASSO, can be modeled flexibly, in particular, by applying different shapes, weighting and overlap between adjacent neighborhoods. The thresholding operators defined in (1.37)-(1.39) are then composed with some neighborhood weighting functional $\eta_N$, defined in Section 2.2 of [xi], to obtain the following, generalized shrinkage operators by setting

$$
\xi^{WGL} = \xi^{L} \circ \eta_N \ (\text{windowed Group-Lasso}), \\
\xi^{PEL} = \xi^{EL} \circ \eta_N \ (\text{persistent Elitist-Lasso}), \\
\xi^{PGL} = \xi^{GL} \circ \eta_N \ (\text{persistent Group-Lasso}).
$$

Since these generalized shrinkage operators are not associated to a simple convex penalty functional, cp. [TK08], the convergence properties of the associated (iterative) algorithms Landweber-iteration are more involved and have been studied in [KSD12].
In Section 3 of the paper, the performance of the resulting different operators is systematically evaluated for classical signal processing tasks like de-noising and sparse multi-layer decomposition. The obtained results are quite satisfactory in terms of both SNR and informal listening test. Different sizes and shapes of neighborhoods as well as the various suggested thresholding operators are compared. The neighborhoods’ shapes do not need to be symmetric around the origin, which can be exploited to feature different parts of a signal under observation. Shapes with varying symmetry yield different perspectives on the signal content: while the symmetric neighborhoods capture parts before and after time-points with high energy (corresponding to attacks in the signal), the asymmetric ones rather retain components before or after, depending on the orientation of the neighborhood’s asymmetry, the attacks. The orientation of the neighborhood therefore systematically promotes the preservation of different temporal segments of the signal. For example, pre-echo, an often undesired artifact in sparse regression of audio signals, can be suppressed by using the correct asymmetry of a coefficient’s neighborhood.

In the de-noising applications investigated in [xii], the windowed group lasso (1.42) appeared to perform best; since their performance was quite promising in comparison to state-of-the-art algorithms, they were further studied in [SD12] and [xii], summarized in the next section.

1.5.4 Persistent Time-Frequency Shrinkage for Audio Denoising - [xii]

In [xii] denoising is considered as a problem of structured sparse approximation. While using the formal framework introduced in [xi], where different mixed norm priors where applied, the current paper solely considers the case of weighted ℓ1-regularization, for which an accelerated algorithm (FISTA, cf. [BT09]) is presented and a variety of refined, weight-dependent thresholding operators are systematically evaluated for denoising.

Time-frequency based audio denoising algorithms can be divided in those applying diagonal and those applying non-diagonal estimation. In diagonal denoising algorithms, the attenuation factor is determined independently for each time-frequency coefficient. These algorithms, as the classical (empirical) Wiener estima-
tor, ignore the correlation between coefficients, cf. [BSM79, LO79] and thus suffer of artifacts called musical noise. Non-diagonal estimation helps to avoid these - rather severe - artifacts.

Diagonal and non-diagonal denoising techniques have been carefully reviewed in [YMB08], where the authors introduced a denoising algorithm by time-frequency block thresholding, which can now be considered as the state of the art in non-diagonal audio denoising and serves as a reference algorithm for the experiments in our work.

In Section 2.3, we introduce a further generalization to the framework used for formalizing persistent shrinkage in [xii], namely, the thresholding operator itself may be transformed, as specified in Definition 2.2 in the paper. This step allows the inclusion of classical operators such as the empirical Wiener estimator [Mal09], which is, by composition with neighborhood weighting as defined in (1.42), generalized to a persistent empirical Wiener shrinkage operator.

The generalized shrinkage operators were implemented in MATLAB and the corresponding toolbox and the audio files used in the evaluation, as well as further audio-visual examples, are available at the webpage homepage.univie.ac.at/monika.doerfler/StrucAudio.html.

For all simulations a tight Gabor frame with Hannning window of 1024 samples length at 44100 Hz audio sampling rate and redundancy 4 is used; this is a standard setting in audio processing, since it allows for perfect reconstruction with the analysis window as synthesis window. Convergence of the applied iterative shrinkage operators within FISTA, while not covered by results derived from a simple convex optimization problem, is suggested by numerical evidence, compare Figure 3 and Figure 4 in [xii].

Section 3.2 then addresses the question, which neighborhood orientation leads to optimal denoising results within certain signal classes (music played by strings, piano and percussion, respectively) whose properties are known. As an example, we reproduce Table 2 from the paper, which shows the results for different sizes of the neighborhoods (all extended in time). Depending on signal characteristics, there seem to be optimal neighborhood lengths. Table 1.1 shows the corresponding numerical results. The longest neighborhood works best for the strings excerpt which features the most temporally persistent structures. Similarly, medium and
Table 1.1: Comparison of the neighborhood’s extension w.r.t. to signal characteristics. The maximal SNRs of WGL with short, medium and long neighborhoods are averaged over noise levels of 0, 10, and 20 dB SNR.

<table>
<thead>
<tr>
<th></th>
<th>Strings</th>
<th>Piano</th>
<th>Percussion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short</td>
<td>20.2</td>
<td>19.3</td>
<td><strong>20.8</strong></td>
</tr>
<tr>
<td>Medium</td>
<td>20.5</td>
<td><strong>19.5</strong></td>
<td>20.4</td>
</tr>
<tr>
<td>Long</td>
<td><strong>20.6</strong></td>
<td>19.4</td>
<td>20.5</td>
</tr>
</tbody>
</table>

short extensions optimize the SNR measure for piano and percussion phrases.

Finally, in Section 3.3, the performance of the new algorithms introduced in this paper are compared, in particular to the block-thresholding denoising algorithm proposed in [YMB08] and shown to perform or even better, but with significantly reduced computation time.
Bibliography


[SD12] K. Siedenburg and Monika Dörfler. Audio denoising by generalized
time-frequency thresholding. *Proceedings of the AES 45th Confer-
ence on Applications of Time-Frequency Processing*, Helsinki, Fin-

[Shl97] Seymour Shlien. The modulated lapped transform, its time-varying
forms, and its applications to audio coding standards. *Speech and

[SK10] Christian Schörkhuber and Anssi Klapuri. Constant-Q toolbox for
music processing. In *Proceedings of the 7th Sound and Music Com-

[Sle83] D. Slepian. Some comments on Fourier Analysis and Uncertainty

[SP61] D. Slepian and H. O. Pollak. Prolate Spheroidal Wave Functions,
63, 1961.

[Tib96] Robert Tibshirani. Regression shrinkage and selection via the lasso.*

[TK08] Bruno Torrésani and M. Kowalski. Sparsity and Persistence: mixed
norms provide simple signal models with dependent coefficients. *Sig-

[TV96] Christoph M. Thiele and Lars Vilmoees. A fast algorithm for


[WFD11] Hau-Tieng Wu, Patrick Flandrin, and Ingrid Daubechies. One or
two frequencies? The synchrosqueezing answers. *Adv. Adapt. Data


Nonstationary Gabor Frames - Existence and Construction

Monika Dörfler, Ewa Matusiak*

Department of Mathematics, NuHAG, University of Vienna, Austria

Abstract

Nonstationary Gabor frames were recently introduced in adaptive signal analysis. They represent a natural generalization of classical Gabor frames by allowing for adaptivity of windows and lattice in either time or frequency. In this paper we show a general existence result for this family of frames. We then give a perturbation result for nonstationary Gabor frames and construct nonstationary Gabor frames with non-compactly supported windows from a related painless nonorthogonal expansion. Finally, the theoretical results are illustrated by two examples of practical relevance.

Keywords: adaptive representations, nonorthogonal expansions, irregular Gabor frames, existence

1. Introduction

The principal idea of Gabor frames was introduced in [14] with the aim to represent signals in a time-frequency localized manner. Since the work of Gabor himself, a lot of research has been done on the topic of atomic time-frequency representation. While it turned out that the original model proposed by Gabor does not yield stable representations in the sense of frames [4, 7, 10], the existence of Gabor frames was first established in the so called painless case, [7], which requires the use of compactly supported analysis windows. The existence of Gabor frames in more general situations was proved later [19, 23] and the proof often uses an argument invoking the invertibility of diagonally dominant matrices.

Various irregular and adaptive versions of Gabor frames have been introduced over the years, cf. [1, 5, 12, 21]. In these approaches, the irregularity usually concerns either the sampling set, which is allowed to deviate from a lattice, or the window, which is allowed to be modified. In [1], varying windows as well as irregular sampling points are allowed, however, existence of a local frame is assumed, from which a global frame is constructed. Nonstationary Gabor frames give up the strict regularity of the classical Gabor setting, but, as opposed to irregular frames, maintain enough structure to guarantee efficient implementation and, possibly approximate, efficient reconstruction. In analogy to the classical, regular case [7], painless nonstationary Gabor frames were introduced in [2], where the principal idea is described and illustrated in detail. The construction of painless nonstationary Gabor
frames is similar to, but more flexible than the construction of windowed modified cosine transforms and other lapped transforms [17, 24] that allow for adaptivity of the window length. An efficient and perfectly invertible constant-Q transform was recently introduced using nonstationary Gabor transforms [22]. In this and similar situations, redundancy of the transform is crucial, since non-redundant versions of the constant-Q transform lead to dyadic wavelet transforms, which are often inappropriate for audio signal processing. Redundancy allows for good localization of both the analysis and synthesis windows, and their respective Fourier transforms and often promote sparse representations in adaptive processing.

Painless non-orthogonal expansions can only be devised if the involved analysis windows are either compactly supported or band limited. This requirement may sometimes be too restrictive. For instance, one may be interested in designing frequency-adaptive nonstationary Gabor frames with windows that are compactly supported in time, i.e. can be implemented as FIR filters, cf. [11] and lend themselves to real-time implementation, cp. [9].

The present contribution addresses the case of nonstationary Gabor frames with more general windows than used in the painless case. In Theorem 3.4, we derive the existence of nonstationary Gabor frames directly from a generalized Walnut representation: under mild uniform decay conditions on all windows involved, we show an existence result of nonstationary Gabor frames in parallel to the result given in [23] for regular Gabor frames, also compare [15, Theorem 6.5.1]. Note that the existence of a different class of nonstationary Gabor frames, the *quilted Gabor frames*, [8], was recently proved in the general context of spline type spaces in the remarkable paper [18].

This paper is organized as follows. In the next section, we introduce notation and state some auxiliary results. In Section 3, we first define nonstationary Gabor frames and recall known results for the painless case. In Section 3.2 a Walnut representation and a corresponding bound of the frame operator in the general setting is derived and Section 3.3 provides the existence of nonstationary Gabor frames. In Section 3.4 we pursue two basic approaches for the construction of nonstationary Gabor frames. Using tools from the theory of perturbation of frames, we construct nonstationary frames from an existing frame in Proposition 3.7. In Corollary 3.8 we design nonstationary Gabor frames by exploiting knowledge about a related painless frame, to obtain "almost painless nonstationary Gabor frames". In Section 4 we provide examples based on the two introduced construction principles.

2. Notation and Preliminaries

Given a non-zero function $g \in L^2(\mathbb{R})$, let $g_{k,l}(t) = M_{bl}T_{ak}g(t) := e^{2\pi i blt}g(t-ak)$. $M_{bl}$ is a modulation operator, or frequency shift, and $T_{ak}$ is a time shift.

The set $\mathcal{G}(g, a, b) = \{g_{k,l} : k, l \in \mathbb{Z}\}$ is called a Gabor system for any real, positive $a, b$. $\mathcal{G}(g, a, b)$ is a Gabor frame for $L^2(\mathbb{R})$, if there exist frame bounds $0 < A \leq B < \infty$ such that for every $f \in L^2(\mathbb{R})$ we have

$$A\|f\|_2^2 \leq \sum_{k,l \in \mathbb{Z}} |\langle f, g_{k,l} \rangle|^2 \leq B\|f\|_2^2.$$  

(1)

To every Gabor system, we associate the analysis operator $C_g$ given by $(C_g f)_{k,l} = \langle f, g_{k,l} \rangle$, and the synthesis operator $U_\gamma = C^*_\gamma$, given by $U_\gamma c = \sum_{k,l \in \mathbb{Z}} c_{k,l}^* \gamma_{k,l}$ for $c \in \ell^2$. The operator

$$U_\gamma c$$

is the inverse of the operator $C_g$. For a more detailed analysis of the Gabor system and its properties, see [22].
$S_{g,\gamma}$ associated to $\mathcal{G}(g, a, b)$ and $\mathcal{G}(\gamma, a', b')$, where $S_{g,\gamma} = U_{\gamma}C_g$ reads

$$S_{g,\gamma} f = \sum_{k,l \in \mathbb{Z}} \langle f, g_{k,l} \rangle \gamma_{k,l}.$$  

The inequality (1) is equivalent to the invertibility and boundedness of the frame operator $S_{g,g}$ of $\mathcal{G}(g, a, b)$.

The analysis operator is the sampled short-time Fourier transform (STFT). For a fixed window $g \in L^2(\mathbb{R})$, the STFT of $f \in L^2(\mathbb{R})$ is

$$V_g f(x, \omega) = \int_{\mathbb{R}} f(t) e^{-2\pi i \omega t} g(t - x) dt = \langle f, M_{\omega} T_x g \rangle.$$  

Setting $(x, \omega) = (ak, bl)$, leads to $V_g f(ak, bl) = (C_g f)_{k,l}$.

When working with irregular grids, we assume that the sampling points form a separated set: a set of sampling points $\{a_k : k \in \mathbb{Z}\}$ is called $\delta$-separated, if $|a_k - a_m| > \delta$ for $a_k, a_m$, whenever $k \neq m$. $\chi_I$ will denote the characteristic function of the interval $I$.

A convenient class of window functions for time-frequency analysis on $L^2(\mathbb{R})$ is the Wiener space.

**Definition 2.1.** A function $g \in L^\infty(\mathbb{R})$ belongs to the Wiener space $W(L^\infty, \ell^1)$ if

$$\|g\|_{W(L^\infty, \ell^1)} := \sum_{k \in \mathbb{Z}} \text{ess sup}_{t \in Q}|g(t + k)| < \infty, \quad Q = [0, 1].$$

For $g \in W(L^\infty, \ell^1)$ and $\delta > 0$ we have [15]

$$\text{ess sup}_{t \in \mathbb{R}} \sum_{k \in \mathbb{Z}} |g(t - \delta k)| \leq (1 + \delta^{-1})\|g\|_{W(L^\infty, \ell^1)}. \quad (2)$$

In dealing with polynomially decaying windows, we will repeatedly use the following lemma.

**Lemma 2.2.** For $p > 1$ the following estimates hold:

(a) Let $\delta > 0$, then

$$\sum_{k=1}^{\infty} (1 + \delta k)^{-p} \leq (1 + \delta)^{-p}(\delta^{-1} + p)(p - 1)^{-1}.$$  

(b) Let $\{a_k : k \in \mathbb{Z}\} \subset \mathbb{R}$ be a $\delta$-separated set. Then

$$\text{ess sup}_{t \in \mathbb{R}} \sum_{k \in \mathbb{Z}} (1 + |t - a_k|)^{-p} \leq 2 \left(1 + (1 + \delta)^{-p}(\delta^{-1} + p)(p - 1)^{-1}\right).$$

**Proof.** To show (a) we write

$$\sum_{k=1}^{\infty} (1 + \delta k)^{-p} = (1 + \delta)^{-p} + \sum_{k=2}^{\infty} (1 + \delta k)^{-p} = (1 + \delta)^{-p} + \sum_{k=2}^{\infty} \int_{[0,1]+k} (1 + \delta k)^{-p} dt.$$
for $t \in [k, k+1]$, we have $\delta t \leq \delta (k + 1)$ which implies that $1 + \delta (t - 1) \leq 1 + \delta k$. Therefore,

$$\sum_{k=2}^{\infty} \int_{[0,1]+k} (1 + \delta k)^{-p} \, dt \leq \sum_{k=2}^{\infty} \int_{[0,1]+k} (1 + \delta (t - 1))^{-p} \, dt = \int_{2}^{\infty} (1 + \delta (t - 1))^{-p} \, dt = (1 + \delta)^{-p+1}\delta^{-1}(p-1)^{-1},$$

and the estimate follows.

To prove (b), fix $t \in \mathbb{R}$. Since $|a_k - a_l| > \delta$ for $k \neq l$, each interval of length $\delta$ contains at most one point $t - a_k, k \in \mathbb{Z}$. Therefore we may write $t - a_k = \delta n_k + x_k$ for unique $n_k \in \mathbb{Z}$ and $x_k \in [0, \delta)$, and by the choice of $\delta$, we have $n_k \neq n_l$ for $k \neq l$. We assume, without loss of generality, that $n_k = 0$ for $k = 0$, and we find that

$$\sum_{k \in \mathbb{Z}} (1 + |t - a_k|)^{-p} = \sum_{k \in \mathbb{Z}} (1 + |\delta n_k + x_k|)^{-p} \leq 1 + \sum_{k \in \mathbb{Z}; n_k > 0} (1 + \delta n_k + x_k)^{-p} + \sum_{k \in \mathbb{Z}; n_k > 0} (1 + \delta n_k - x_k)^{-p} \leq 1 + \sum_{k \in \mathbb{Z}; n_k > 0} (1 + \delta n_k)^{-p} + \sum_{k \in \mathbb{Z}; n_k > 0} (1 + \delta n_k - \delta)^{-p} \leq 1 + \sum_{k=1}^{\infty} (1 + \delta k)^{-p} + \sum_{k=1}^{\infty} (1 + \delta (k - 1))^{-p} = 2 \left(1 + \sum_{k=1}^{\infty} (1 + \delta)^{-p}\right) \leq 2 \left(1 + (1 + \delta)^{-p}(\delta^{-1} + p)(p-1)^{-1}\right).$$

The last expression is independent of $t$, and the claim follows. \hfill \square

**Remark 1.** When the set $A = \{a_k : k \in \mathbb{Z}\} \subset \mathbb{R}$ is relatively $\delta$–separated, meaning

$$\text{rel}(A) := \max_{t \in \mathbb{R}} \#\{A \cap ([0, \delta] + t)\} < \infty,$$

then the estimate (b) in Lemma 2.2 becomes

$$\text{ess sup}_{t \in \mathbb{R}} \sum_{k \in \mathbb{Z}} (1 + |t - a_k|)^{-p} \leq 2 \text{rel}(A) \left(1 + (1 + \delta)^{-p}(\delta^{-1} + p)(p-1)^{-1}\right).$$

Notice, that for a separated set, $\text{rel}(A) = 1$.

### 3. Nonstationary Gabor frames

Nonstationary Gabor systems provide a generalization of the classical Gabor systems of time-frequency-shifted versions of a single window function.

**Definition 3.1.** Let $g = \{g_k \in L^2(\mathbb{R}) : k \in \mathbb{Z}\}$ be a set of window functions and let $b = \{b_k : k \in \mathbb{Z}\}$ be a corresponding sequence of frequency-shift parameters. Set $g_{k,l} = M_{b_k}g_k$. Then, the set

$$\mathcal{G}(g, b) = \{g_{k,l} : k, l \in \mathbb{Z}\}$$

is called a **nonstationary Gabor system**.
Note that, conceptually, we assume that the windows \( g_k \) are centered at points \( \{a_k : k \in \mathbb{Z}\} \), in direct generalization of the regular case, where \( g_k(t) = g(t - ak) \) for some time-shift parameter \( a \). In this sense, we have a two-fold generalization: the sampling points can be irregular and the windows can change for every sampling point. We are interested in conditions under which a nonstationary Gabor system forms a frame. We first recall the case of nonstationary Gabor frames with compactly supported windows, see [2] and http://www.univie.ac.at/nonstatgab/ for further information.

3.1. Compactly supported windows: the painless case

Based on the support length of the windows \( g_k \), we can easily determine frequency-shifts parameters \( b_k \), for which we obtain a frame. The following result is the nonstationary version of the result given in [7].

**Proposition 3.2** ([2]). Let \( g = \{g_k \in L^2(\mathbb{R}) : k \in \mathbb{Z}\} \) be a collection of compactly supported functions with \( |\text{supp } g_k| \leq 1/b_k \). Then \( \mathcal{G}(g, b) \) is a frame for \( L^2(\mathbb{R}) \) if there exist constants \( A > 0 \) and \( B < \infty \) such that

\[
A \leq G_0(t) = \sum_{k \in \mathbb{Z}} b_k^{-1} |g_k(t)|^2 \leq B \quad \text{a.e.}.
\]

The dual atoms are then \( \gamma_{k,l}(t) = M_{lb_k} G_0^{-1}(t) g_k(t) \).

**Remark 2.** An analogous theorem holds for bandlimited functions \( g_k \).

The above theorem follows from the fact that the frame operator associated to the collection of atoms described in the theorem can be written as

\[
S_{g,g} f(t) = \sum_{k \in \mathbb{Z}} b_k^{-1} |g_k(t)|^2 f(t) \quad \text{a.e.}.
\]

The diagonality of the frame operator in the painless case is derived from a generalized Walnut representation for the frame operator \( S_{g,g} \) of nonstationary Gabor frames. In the next section we will see that this representation immediately implies diagonality of \( S_{g,g} \) under the assumptions of Proposition 3.2.

3.2. A Walnut representation for nonstationary Gabor Frames

Let us now consider nonstationary Gabor systems \( \mathcal{G}(g, b) \) and \( \mathcal{G}(\gamma, b) \), with all windows \( g_k \) and \( \gamma_k \) in \( W(L^\infty, \ell^1) \). The operator associated to \( \mathcal{G}(g, b) \) and \( \mathcal{G}(\gamma, b) \) reads

\[
S_{g,\gamma} f(t) = \sum_{k,l \in \mathbb{Z}} \langle f, M_{lb_k} g_k \rangle M_{lb_k} \gamma_k.
\] (3)

**Proposition 3.3.** The operator \( S_{g,\gamma} \) in (3) admits a Walnut representation

\[
S_{g,\gamma} f = \sum_{k,l \in \mathbb{Z}} G_{k,l}^{g,\gamma} \cdot T_{lb_k^{-1}} f, \quad \text{where} \quad G_{k,l}^{g,\gamma}(t) = b_k^{-1} g_k(t - lb_k^{-1}) \gamma_k(t),
\] (4)
for \( f \in L^2(\mathbb{R}) \). Moreover, its operator norm can be bounded by

\[
|\langle S_{g,\gamma} f, h \rangle| \leq \left( \sup_{k \in \mathbb{Z}} (1 + b_k^{-1}) \|g_k\|_{W(L^\infty, L^1)} \right)^{1/2} \left( \text{ess sup}_{t \in \mathbb{R}} \sum_{k \in \mathbb{Z}} |g_k(t)| \right)^{1/2} \\
\cdot \left( \sup_{k \in \mathbb{Z}} (1 + b_k^{-1}) \|g_k\|_{W(L^\infty, L^1)} \right)^{1/2} \left( \text{ess sup}_{t \in \mathbb{R}} \sum_{k \in \mathbb{Z}} |\gamma_k(t)| \right)^{1/2} \|f\|_2 \|h\|_2 
\]

for all \( f, h \in L^2(\mathbb{R}) \).

Proof. First assume that \( f, h \in L^2(\mathbb{R}) \) are compactly supported. Since \( \langle f, M_{lb_k} g_k \rangle = \langle f \hat{g_k}(lb_k) \rangle \) we can write \( S_{g,\gamma} \) as

\[
S_{g,\gamma} f(t) = \sum_{k, l \in \mathbb{Z}} \langle f \hat{g_k}(lb_k) \rangle M_{lb_k} \gamma_k(t) = \sum_{k \in \mathbb{Z}} m_k(t) \gamma_k(t),
\]

where \( m_k(t) = \sum_{l \in \mathbb{Z}} \langle f \hat{g_k}(lb_k) \rangle e^{2\pi i lb_k t} \), for every \( k \in \mathbb{Z} \). The functions \( m_k \) are \( b_k^{-1} \) periodic and by the Poisson formula can be written as

\[
m_k(t) = b_k^{-1} \sum_{l \in \mathbb{Z}} \langle f \hat{g_k}(t - lb_k^{-1}) \rangle.
\]

Therefore, substituting (7) in (6) yields the Walnut representation.

We next prove the boundedness (5). In the following chain of inequalities, we will use Cauchy-Schwartz inequality for sums and integrals and, since all summands have absolute value, Fubini’s theorem to justify changing the order of summation and integral. We thus find

\[
|\langle S_{g,\gamma} f, h \rangle| = \left| \left\langle \sum_{k, l \in \mathbb{Z}} b_k^{-1} g_k(\cdot - lb_k^{-1}) \gamma_k(\cdot) f(\cdot - lb_k^{-1}), h \right\rangle \right|
\]

\[
\leq \sum_{k, l \in \mathbb{Z}} b_k^{-1} \int_{\mathbb{R}} |g_k(t - lb_k^{-1})| |\gamma_k(t)||f(t - lb_k^{-1})||h(t)| \, dt
\]

\[
\leq \sum_{k, l \in \mathbb{Z}} b_k^{-1} \left[ \int_{\mathbb{R}} |g_k(t - lb_k^{-1})| |\gamma_k(t)||f(t - lb_k^{-1})|^2 \, dt \right]^{1/2} \left[ \int_{\mathbb{R}} |g_k(t - lb_k^{-1})| |\gamma_k(t)||h(t)|^2 \, dt \right]^{1/2}
\]

\[
\leq \left[ \sum_{k, l \in \mathbb{Z}} b_k^{-1} \int_{\mathbb{R}} |g_k(t)| |\gamma_k(t + lb_k^{-1})||f(t)|^2 \, dt \right]^{1/2} \left[ \sum_{k, l \in \mathbb{Z}} b_k^{-1} \int_{\mathbb{R}} |\gamma_k(t)||g_k(t - lb_k^{-1})||h(t)|^2 \, dt \right]^{1/2}
\]

\[
= \left[ \int_{\mathbb{R}} |f(t)|^2 \sum_{k, l \in \mathbb{Z}} b_k^{-1} |g_k(t)||\gamma_k(t - lb_k^{-1})| \, dt \right]^{1/2} \left[ \int_{\mathbb{R}} |h(t)|^2 \sum_{k, l \in \mathbb{Z}} b_k^{-1} |\gamma_k(t)||g_k(t - lb_k^{-1})| \, dt \right]^{1/2}.
\]
The first term in the last expression can be bounded as follows

\[
\int_{\mathbb{R}} |f(t)|^2 \sum_{k,l \in \mathbb{Z}} b_k^{-1} |g_k(t)\gamma_k(t - lb_k)| dt = \sum_{k \in \mathbb{Z}} b_k^{-1} \int_{\mathbb{R}} \sum_{l \in \mathbb{Z}} |\gamma_k(t - lb_k)| |f(t)|^2 |g_k(t)| dt
\]

\[
\leq \sum_{k \in \mathbb{Z}} \left( b_k^{-1} \sup_{t \in \mathbb{R}} \sum_{l \in \mathbb{Z}} |\gamma_k(t - lb_k)| \right) \int_{\mathbb{R}} |f(t)|^2 |g_k(t)| dt
\]

\[
\leq \sup_{k \in \mathbb{Z}} \left( b_k^{-1} \sup_{t \in \mathbb{R}} \sum_{l \in \mathbb{Z}} |\gamma_k(t - lb_k)| \right) \sum_{k \in \mathbb{Z}} \int_{\mathbb{R}} |f(t)|^2 |g_k(t)| dt
\]

\[
\leq \sup_{k \in \mathbb{Z}} \left( b_k^{-1} \sup_{t \in \mathbb{R}} \sum_{l \in \mathbb{Z}} |\gamma_k(t - lb_k)| \right) \left( \sup_{t \in \mathbb{R}} \sum_{k \in \mathbb{Z}} |g_k(t)| \right) \|f\|_2^2 . \quad (9)
\]

Using relation (2) for \( \sup_{t \in \mathbb{R}} \sum_{l \in \mathbb{Z}} |\gamma_k(t - lb_k)| \) and substituting (9) into (8) yields (5).

The estimate for the second term in (8) is obtained analogously. By the density of compactly supported functions in \( L^2(\mathbb{R}) \), the estimate holds for all of \( L^2(\mathbb{R}) \). \( \square \)

**Remark 3.** From (8) in the proof of Proposition 3.3, it follows that the operator \( S_{g,\gamma} \) is also bounded by

\[
|\langle S_{g,\gamma} f, h \rangle| \leq \left( \sup_{t \in \mathbb{R}} \sum_{k,l \in \mathbb{Z}} b_k^{-1} |g_k(t - lb_k)| |\gamma_k(t)| \right)^{1/2} \cdot \left( \sup_{t \in \mathbb{R}} \sum_{k,l \in \mathbb{Z}} b_k^{-1} |\gamma_k(t - lb_k)||g_k(t)| \right)^{1/2} \|f\|_2 \|h\|_2 .
\]

**Remark 4.** In the case of a frame operator \( S_{g,g} \), the Walnut representation (4) becomes

\[
S_{g,g} f(t) = \sum_{k,l \in \mathbb{Z}} b_k^{-1} g_k(t - lb_k) g_k(t) f(t - lb_k) \quad \text{a.e.}, \quad (10)
\]

and the above bounds reduce to

\[
|\langle S_{g,g} f, h \rangle| \leq \left( \sup_{k \in \mathbb{Z}} (1 + b_k^{-1}) \|g_k\|_{W(L^\infty,\ell^1)} \right) \left( \sup_{k \in \mathbb{Z}} \sum_{l \in \mathbb{Z}} |g_k(t)| \right) \|f\|_2 \|h\|_2
\]

\[
|\langle S_{g,g} f, h \rangle| \leq \left( \sup_{t \in \mathbb{R}} \sum_{k,l \in \mathbb{Z}} b_k^{-1} |g_k(t - lb_k)| |g_k(t)| \right) \|f\|_2 \|h\|_2 .
\]

**Remark 5.** Note that in the painless case of Theorem 3.2, the frame operator reduces to the multiplication operator \( S_{g,f} = \sum_{k \in \mathbb{Z}} G_{k,0} \cdot f = G_0 \cdot f \).

**Remark 6.** In the standard setting of Gabor frames, i.e. \( g_k(t) = g(t - ak) \) for fixed \( a > 0 \), and \( b_k = b \) for all \( k \in \mathbb{Z} \), the above bound reduces to the well-know bound

\[
|\langle S_{g,f} f, h \rangle| \leq (1 + a^{-1})(1 + b^{-1}) \|g\|_{W(L^\infty,\ell^1)}^2 \|f\|_2 \|h\|_2
\]
3.3. Existence of nonstationary Gabor frames

For windows $g_k$ that are neither compactly supported nor bandlimited, we are interested in the existence of frames of the form $\mathcal{G}(g, b)$ and in the construction of the involved parameters. The following theorem derives a sufficient condition for the existence of nonstationary Gabor frames and shows that this condition can be satisfied.

In this and the subsequent sections, $[b_L, b_U]$, $[p_L, p_U]$, $[C_L, C_U]$ are compact intervals of positive real numbers.

**Theorem 3.4.** Let $g = \{g_k \in W(L^\infty, \ell^1) : k \in \mathbb{Z}\}$ be a set of windows such that

i) for some positive constants $A_0, B_0$

$$0 < A_0 \leq \sum_{k \in \mathbb{Z}} |g_k(t)|^2 \leq B_0 < \infty \text{ a.e.;} \quad (11)$$

ii) for all $k \in \mathbb{Z}$, the windows decay polynomially around a $\delta$-separated set $\{a_k : k \in \mathbb{Z}\}$ of time-sampling points $a_k$

$$|g_k(t)| \leq C_k (1 + |t - a_k|^{-p_k}), \quad (12)$$

where $p_k \in [p_L, p_U] \subset \mathbb{R}$, $p_L > 2$ and $C_k \in [C_L, C_U]$.

Then there exists a sequence $\{b_k^0\}_{k \in \mathbb{Z}}$, such that for $b_k \leq b_k^0$, $k \in \mathbb{Z}$, the nonstationary Gabor system $\mathcal{G}(g, b)$ forms a frame for $L^2(\mathbb{R})$.

**Proof.** Let $f \in L^2(\mathbb{R})$. Applying (10), we write

$$\langle S_{g,g} f, f \rangle = \int_{\mathbb{R}} \sum_{k \in \mathbb{Z}} b_k^{-1} |g_k(t)|^2 |f(t)|^2 \, dt + \int_{\mathbb{R}} \sum_{l \in \mathbb{Z} \setminus \{0\}} \sum_{k \in \mathbb{Z}} b_k^{-1} g_k(t) \overline{g_k(t - lb_k^{-1})} f(t - lb_k^{-1}) \overline{f(t)} \, dt$$

Using similar arguments as in the derivation of (8), we obtain

$$\left| \int_{\mathbb{R}} \sum_{l \in \mathbb{Z} \setminus \{0\}} \sum_{k \in \mathbb{Z}} b_k^{-1} g_k(t) \overline{g_k(t - lb_k^{-1})} f(t - lb_k^{-1}) \overline{f(t)} \, dt \right| \leq$$

$$\leq \left( \text{ess sup}_{t \in \mathbb{R}} \sum_{l \in \mathbb{Z} \setminus \{0\}} \sum_{k \in \mathbb{Z}} b_k^{-1} |g_k(t)||g_k(t - lb_k^{-1})| \right) \|f\|^2_2$$

$$\leq \max_{k \in \mathbb{Z}} \{b_k^{-1}\} \sum_{l \in \mathbb{Z} \setminus \{0\}} \left( \text{ess sup}_{t \in \mathbb{R}} \sum_{k \in \mathbb{Z}} |g_k(t)||g_k(t - lb_k^{-1})| \right) \|f\|^2_2.$$

Therefore, lower and upper frame bounds are obtained from

$$\langle S_{g,g} f, f \rangle \|f\|^2_2 \geq \min_{k \in \mathbb{Z}} \{b_k^{-1}\} \left( \text{ess inf}_{l \in \mathbb{R}} \sum_{k \in \mathbb{Z}} |g_k(t)|^2 - \max_{k \in \mathbb{Z}} \{b_k^{-1}\} \frac{R}{\min_{k \in \mathbb{Z}} \{b_k^{-1}\}} \right) \quad (13)$$

$$\langle S_{g,g} f, f \rangle \|f\|^2_2 \leq \max_{k \in \mathbb{Z}} \{b_k^{-1}\} \left( \text{ess sup}_{l \in \mathbb{R}} \sum_{k \in \mathbb{Z}} |g_k(t)|^2 + R \right),$$

8
We need to construct a sequence of $b_k$, $k \in \mathbb{Z}$, such that for all $f \in L^2(\mathbb{R})$, (13) is bounded away from zero.

Let $\epsilon < C_L$ and consider the sequence of frequency shifts $b_k = (\frac{\epsilon}{\epsilon_k})^{1/p_k}$. Then $\min_{k \in \mathbb{Z}} \{ b_k^{-1} \} \geq (C_L \epsilon^{-1})^{1/p_2}$, and $\max_{k \in \mathbb{Z}} \{ b_k^{-1} \} \leq (C_L \epsilon^{-1})^{1/p_1}$ and

$$\frac{\max_{k \in \mathbb{Z}} \{ b_k^{-1} \}}{\min_{k \in \mathbb{Z}} \{ b_k^{-1} \}} \leq C_{U}^{1/p} C_{L}^{-1/p_u} \epsilon^{1/p_u - 1/p_L}. \quad (14)$$

Since $(1 + |x + y|)^{-p} \leq (1 + |x|)^p (1 + |y|)^{-p}$ for $x, y \in \mathbb{R}$ and $p \geq 0$, using (12), we have, for some $\mu$ with $p_L - 2 > \mu > 0$:

$$|g_k(t)||g_k(t - lb_k^{-1})| \leq C^{2}_{k}(1 + |t - a_k|)^{-p_k} (1 + |t - a_k - lb_k^{-1}|)^{-p_k + (1+\mu)} \leq C^{2}_{k}(1 + |t - a_k|)^{(1+\mu)}(1 + |l|b_k^{-1})^{-p_k + (1+\mu)} \leq C^{2}_{k}(1 + |t - a_k|)^{(1+\mu)} |l|^{-p_k + (1+\mu)} b_k^{p_k - (1+\mu)} = \frac{C^{2}_{k}(1 + |t - a_k|)^{(1+\mu)} |l|^{-p_k + (1+\mu)}}{E_k} \epsilon^{1-(1+\mu)/p_k}. \quad (15)$$

Hence,

$$\sum_{k \in \mathbb{Z}} |g_k(t)||g_k(t - lb_k^{-1})| \leq \sum_{k \in \mathbb{Z}} E_k (1 + |t - a_k|)^{-1(\mu)} |l|^{-p_k + (1+\mu)} \epsilon^{1-(1+\mu)/p_k} \leq \max_{k \in \mathbb{Z}} E_k |l|^{-p_k + (1+\mu)} \epsilon^{1-(1+\mu)/p_L} \sum_{k \in \mathbb{Z}} (1 + |t - a_k|)^{-1(\mu)} \leq \max_{k \in \mathbb{Z}} E_k |l|^{-p_k + (1+\mu)} \epsilon^{1-(1+\mu)/p_L} 2 (1 + (1 + \delta)^{-1(\mu)}(\delta^{-1} + 1 + \mu)^{-1}) \quad (15)$$

where the last estimate follows from Lemma 2.2 (b). Summing the expression (15) over $l \in \mathbb{Z} \setminus \{0\}$ using Lemma 2.2 (a), we see that $R$, as a function of $\epsilon$, behaves like $\epsilon^{1-(1+\mu)/p_L}$, i.e. $R \approx \epsilon^{1-(1+\mu)/p_L}$, and $R$ tends to 0 for $\epsilon \to 0$. Moreover,

$$\frac{\max_{k \in \mathbb{Z}} \{ b_k^{-1} \}}{\min_{k \in \mathbb{Z}} \{ b_k^{-1} \}} R \approx \epsilon^{1-(2+\mu)/p_L + 1/p_U}$$

can be made arbitrarily small by choosing $\epsilon$ small since $1 - (2 + \mu)/p_L + 1/p_U > 0$. Therefore, if $\epsilon_0$ is such that for $b_k^0 := \left(\frac{\epsilon_0}{\epsilon_k}\right)^{1/p_k}$, $\max_{k \in \mathbb{Z}} \{ b_k^{-1} \} R < A_0$, then $\{ M_{b_k} g_k \}_{k \in \mathbb{Z}}$ is a frame for all $b_k \leq b_k^0$.

For completeness, we state the equivalent result for analysis windows $g_k$ with polynomial decay in the frequency domain.

**Corollary 3.5.** Let $g = \{ g_k \in L^2(\mathbb{R}) : \hat{g}_k \in W(L^\infty, \ell^1), k \in \mathbb{Z} \}$ be a set of windows such that

i) for some positive constants $A_0, B_0$

$$0 < A_0 \leq \sum_{k \in \mathbb{Z}} |\hat{g}_k(t)|^2 \leq B_0 \leq \infty \ a.e. \quad (16)$$

10
ii) for all $k \in \mathbb{Z}$, the windows decay polynomially around a $\delta$-separated set $\{b_k : k \in \mathbb{Z}\}$ of frequency-sampling points $b_k$:

$$|\hat{g}_k(t)| \leq C_k(1 + |t - b_k|)^{-p_k},$$

(17)

where $p_k$ and $C_k$ are chosen as in Theorem 3.4.

Then there exists a sequence $\{a_k^0\}_{k \in \mathbb{Z}}$, such that for $a_k \leq a_k^0$, $k \in \mathbb{Z}$, the nonstationary Gabor system $\{T_{ia_k}g_k : k, l \in \mathbb{Z}\}$ forms a frame for $L^2(\mathbb{R})$.

3.3.1. Nonstationary Gabor frames on modulation spaces

Modulation spaces, cf. [13, 15], are considered as the appropriate function spaces for time-frequency analysis and in particular, for the study of Gabor frames. By their definition, modulation spaces require decay in both time and frequency. Under additional assumptions on the windows $g_k$, the collection $\mathcal{G}(g, b)$ is a frame for all modulation spaces $M^p$, $1 \leq p \leq \infty$.

**Proposition 3.6.** Let $\mathcal{G}(g, b)$ be a frame for $L^2(\mathbb{R})$ satisfying the uniform estimate

$$|V_\phi g_k(x, \omega)| \leq C(1 + |x - a_k|)^{-r-2}(1 + |\omega|)^{-r-2}, \quad r > 2$$

(18)

where $\phi$ is a Gaussian window. Then the frame operator $S$ is invertible simultaneously on all modulation spaces $M^p$ for $1 \leq p \leq \infty$.

**Proof.** Notice, that

$$|V_\phi g_{k,l}(x, \omega)| \leq C(1 + |(x - a_k, \omega - lb_k)|)^{-r-2},$$

since $(1 + |x| + |\omega|)^{-r} \geq (1 + |x|)^{-r}(1 + |\omega|)^{-r}$ and the weights $(1 + |(x, \omega)|)^r$ and $(1 + |x| + |\omega|)^r$ are equivalent.

A result on Gabor molecules [16] states that, if an $L^2$-frame $\{g_z : z = (z_1, z_2) \in \mathcal{Z} \subseteq \mathbb{R}^2\}$, where $\mathcal{Z}$ is separable, satisfies the uniform estimate $|V_\phi g_z(x, \omega)| \leq C(1 + |(x - z_1, \omega - z_2)|)^{-r-2}$, then the frame operator $Sf = \sum_{z \in \mathcal{Z}}(f, g_z)g_z$ is invertible simultaneously on all $M^p$ for each $1 \leq p \leq \infty$. The result hence follows from condition (18). \qed

3.4. Constructing nonstationary Gabor frames

Theorem 3.4 shows that for windows with sufficient uniform decay, nonstationary Gabor frames can always be constructed by choosing sufficient density in the frequency samples. In the present section we assume the existence of a certain nonstationary Gabor frame and explicitly construct a new frame by exploiting the prior information about the original one. This is a situation of practical relevance, since we may often be interested in using windows that decay fast and are negligible outside a support of interest. In particular, we will use the fact that painless nonstationary Gabor frames are easily constructed and deduce new frames from painless frames. The new frames thus obtained will be called almost painless nonstationary Gabor frames.

We will subsequently assume $b_k \in [b_L, b_U]$ for frequency-shift parameters and we let $C_k \in [C_L, C_U]$ and $p_k \in [p_L, p_U]$ with $p_L > 1$ for the constants involved in the decay assumptions.
for the analysis windows. We then work with the following constants that depend on the separation of the sampling points, the decay of the windows and the frequency-sampling parameters:

\[
E_1 = 1 + \frac{\delta^{-1} + p_L}{(1 + \delta)^p_L(p_L - 1)} \quad \text{and} \quad E_2 = 1 + \frac{b_U + p_U}{(1 + b_U^{-1})p_L(p_L - 1)}
\]  

(19)

We first consider nonstationary Gabor frames obtained by perturbation of a known frame. This result is in the spirit of similar results for regular Gabor frames [3, 6].

**Proposition 3.7.** Let \(\{a_k : k \in \mathbb{Z}\}\) be a \(\delta\)-separated set and \(G(h, b)\) a nonstationary Gabor frame with frame bounds \(A_h, B_h\) and frame operator \(S_{h,h}\). Let \(g_k \in L^2(\mathbb{R})\) be a set of windows such that for all \(k \in \mathbb{Z}\) and for almost all \(t \in \mathbb{R}\)

\[
|g_k(t) - h_k(t)| \leq C_k(1 + |t - a_k|)^{-p_k}.
\]  

(20)

If \(C_U < \sqrt{A_h \lambda^{-1}}\) for

\[
\lambda = 4b_L^{-1} \cdot E_1 \cdot E_2,
\]  

(21)

then \(G(g, b)\) is a frame for \(L^2(\mathbb{R})\) with frame bounds \(A = A_h(1 - \sqrt{C_U^2 \lambda A_h^{-1}})^2\) and \(B = B_h(1 + \sqrt{C_U^2 \lambda B_h^{-1}})^2\).

**Proof.** By applying Cauchy-Schwartz inequality, it is easy to see that, for \(G(g, b)\) to be a frame for \(L^2(\mathbb{R})\), it suffices to show that \(\sum_{k,l}\|f, g_{k,l} - h_{k,l}\|^2 \leq R\|f\|^2\) for some \(R < A_h\), also cf. [6, Proposition 4.1.]. Then, frame bounds of \(G(g, b)\) can be taken as \(A_h(1 - \sqrt{R/A_h})^2\) and \(B_h(1 + \sqrt{R/B_h})^2\).

To obtain the required error bound, we let \(\psi_k(t) := g_k - h_k\) and use the estimate given in (9):

\[
\sum_{k,l \in \mathbb{Z}} |\langle f, \psi_{k,l} \rangle|^2 \leq \sup_{k \in \mathbb{Z}} \left( b_k^{-1} \text{ess sup}_{t \in \mathbb{R}} \sum_{l \in \mathbb{Z}} |\psi_k(t - lb_k^{-1})| \right) \left( \text{ess sup}_{t \in \mathbb{R}} \sum_{k \in \mathbb{Z}} |\psi_k(t)| \right) \|f\|^2.
\]  

(22)

The first term of the above in the above inequality is bounded by assumption (20) and Lemma 2.2 (b):

\[
\sum_{k \in \mathbb{Z}} |\psi_k(t)| \leq C_U \sum_{k \in \mathbb{Z}} (1 + |t - a_k|)^{-p_k} \leq C_U \sum_{k \in \mathbb{Z}} (1 + |t - a_k|)^{-p_1} \leq 2CUE_1.
\]

To bound the second term, note that \(\sum_{l \in \mathbb{Z}} (1 + |t - a_k - lb_k^{-1}|)^{-p_k}\) is \(b_k^{-1}\)-periodic, therefore we can simplify

\[
\text{ess sup}_{t \in \mathbb{R}} \sum_{l \in \mathbb{Z}} |\psi_k(t - lb_k^{-1})| \leq C_k \text{ess sup}_{t \in [0,b_k^{-1}]} \sum_{l \in \mathbb{Z}} (1 + |t - lb_k^{-1}|)^{-p_k}.
\]
Hence, by Lemma 2.2, we obtain for $t \in [0, b^{-1}_k]$:
\[
\sum_{l \in \mathbb{Z}} (1 + |t - lb^{-1}_k|)^{-pk} \leq 1 + \sum_{l=1}^{\infty} (1 + |t - lb^{-1}_k|)^{-pk} + \sum_{l=1}^{\infty} (1 + l + lb^{-1}_k)^{-pk} (1 + (l - 1)b^{-1}_k)^{-pk}
\]
\[
= 2 \sum_{l=0}^{\infty} (1 + lb^{-1}_k)^{-pk} = 2 \left(1 + \sum_{l=1}^{\infty} (1 + lb^{-1}_k)^{-pk}\right) \leq 2E_2.
\]

Gathering all the estimates, we obtain
\[
\sum_{k,l \in \mathbb{Z}} |\langle f, \psi_{k,l} \rangle|^2 \leq C_U^2 4b^{-1}_1 E_1 \cdot E_2 \|f\|_2^2 = C_U^2 \lambda \|f\|_2^2.
\]

By assumption $C_U^2 \lambda < A_h$, and the proof is complete. \hfill \Box

Using Proposition 3.7 we next construct a special class of nonstationary Gabor frames by relying on knowledge of a painless nonstationary Gabor frame. We construct new windows which are no more compactly supported, but coincide with the known, compact windows on their support. We call the resulting new systems almost painless nonstationary Gabor frames.

**Corollary 3.8.** Let $g = \{g_k \in W(L^\infty, \ell^1) : k \in \mathbb{Z}\}$ be a set of windows, and let $I_k$ be the intervals $I_k = [a_k - (2b_k)^{-1}, a_k + (2b_k)^{-1}]$ where $\{a_k : k \in \mathbb{Z}\}$ forms a $\delta$-separated set. Assume that $G(h, b)$, where $h_k = g_k \chi_{I_k}$, is a Gabor frame with lower frame bound $A_h$, and that for $\psi_k = g_k - h_k$, all $k \in \mathbb{Z}$ and almost all $t \in \mathbb{R}$

\[
|\psi_k(t)| \leq \begin{cases} C_k (1 + t - a_k - (2b_k)^{-1})^{-p_k}, & t > a_k + (2b_k)^{-1}; \\ 0, & t \in I_k; \\ C_k (1 - t + a_k - (2b_k)^{-1})^{-p_k}, & t < a_k - (2b_k)^{-1}. \end{cases}
\]

If $C_U < \sqrt{A_h \lambda^{-1}}$ for $\lambda = 4b^{-2}_L \cdot \delta^{-1} \cdot E_1 \cdot E_2$, then $G(g, b)$ forms a nonstationary Gabor frame for $L^2(\mathbb{R})$.

**Proof.** The proof follows the steps of the proof of Proposition 3.7 with small changes on how to approximate the terms in (22). First, observe, that for any $t \in \mathbb{R}$ and all $k$, we have

\[
|\psi_k(t)| \leq C_k \left[(1 + |t - a_k - (2b_k)^{-1}|)^{-p_k} + (1 + |t - a_k + (2b_k)^{-1}|)^{-p_k}\right]
\]

Since the frequency shifts $b_k$ are taken from a finite interval and the set $\{a_k : k \in \mathbb{Z}\}$ is $\delta-$separated, the sets $\Gamma^+ = \{a_k + (2b_k)^{-1} : k \in \mathbb{Z}\}$ and $\Gamma^- = \{a_k - (2b_k)^{-1} : k \in \mathbb{Z}\}$ are relatively $\delta-$separated with $\text{rel}(\Gamma) = \text{rel}(\Gamma^+) = \text{rel}(\Gamma^-) = [(2b_L \delta)^{-1}]$. Therefore, by Remark 1, it follows that

\[
\sum_{k \in \mathbb{Z}} |\psi_k(t)| \leq C_U \sum_{k \in \mathbb{Z}} (1 + |t - a_k - (2b_k)^{-1}|)^{-p_L} + C_U \sum_{k \in \mathbb{Z}} (1 + |t - a_k + (2b_k)^{-1}|)^{-p_L}
\]
\[
\leq 4 C_U \text{rel}(\Gamma) (1 + (1 + \delta)^{-p_L}(\delta^{-1} + p_L)(p_L - 1)^{-1}).
\]

12
Now, the expression \( \sum_{l \in \mathbb{Z}} |\psi_k(t - lb_k^{-1})| \) is \( b_k^{-1} \)-periodic. Let \( t \in I_k \), then, by (23)

\[
\sum_{l \in \mathbb{Z}} |\psi_k(t - lb_k^{-1})| \leq C_k \left[ \sum_{l > 0} (1 + a_k - (2b_k)^{-1} - t + lb_k^{-1})^{-pk} \right.
\]

\[
+ \sum_{l < 0} (1 - a_k - (2b_k)^{-1} + t - lb_k^{-1})^{-pk} \right]
\]

\[
\leq C_k 2 \sum_{l=0}^{\infty} (1 + lb_k^{-1})^{-pk} \leq 2C_k (1 + (1 + b_k^{-1})^{-pk} (b_k + p_k)(p_k - 1)^{-1}) ,
\]

where the last estimate follows from Lemma 2.2(a).

4. Examples

We illustrate our theory with two examples. In both examples, we consider a basic window and dilations by 2 and \( 1/2 \), respectively. Since the dilation parameters take only three different values, there are three kinds of windows, with support size \( 1/2, 1 \) and 2, respectively. Note that, while theoretically possible, sudden changes in the shape and width of adjacent windows turn out to be undesirable for applications, hence we only allow for stepwise change in dilation parameters.

In the first example we consider a frame that arises as perturbation of a painless nonstationary Gabor frame. The perturbation consists in the application of a bandpass filter in order to obtain windows with compact support in the frequency domain.

**Example 4.1.** Let \( h \) be a Hann or raised cosine window, i.e. \( h(t) = 0.5 + 0.5 \cos(2\pi t) \) for \( t \in [-1/2, 1/2] \), and zero otherwise. We construct a painless nonstationary Gabor frame by dilating \( h \) by 2 or \( 1/2 \), respectively: Let \( \{s_k\}_{k \in \mathbb{Z}} \) be a sequence with values from the set \( \{-1, 0, 1\} \) with the restriction that \( |s_k - s_{k-1}| \in \{0,1\} \) to avoid sudden changes between adjacent windows. We then define corresponding shift-parameters by setting \( a_0 = 0 \) and

\[
a_{k+1} = a_k + 2^{-s_k} \cdot \frac{5}{6} \quad \text{if} \quad s_k > s_{k+1},
\]

\[
a_{k+1} = a_k + 2^{-s_k+1} \cdot \frac{1}{3} \quad \text{if} \quad s_k = s_{k+1},
\]

\[
a_{k+1} = a_k + 2^{-s_k+1} \cdot \frac{5}{6} \quad \text{if} \quad s_k < s_{k+1}.
\]

The points \( a_k, \ k \in \mathbb{Z} \), form a separated set with minimum separation \( \delta = 1/3 \). Setting \( b_k = 2^s \) and \( h_k(t) = T_{a_k} \sqrt{2^s} h(2^s t) \), the system \( \{M_{lb_k}h_k : k, l \in \mathbb{Z}\} \) forms a painless nonstationary Gabor frame with lower frame bound \( A_k = 0.5 \).

Let \( \Omega = 0.02 \) and \( \phi \) be a bandlimited filter given by

\[
\hat{\phi}(\omega) = 0.5 + 0.5 \cos(2\pi \Omega^{-1} \omega) \text{ on support } [-\Omega/2, \Omega/2] .
\]

We build new windows \( g_k \) by convolving \( \phi \) with \( h_k \). The windows \( g_k := \phi * h_k \) are no more compactly supported. Since \( |\phi(t)| \leq \Omega(1 + |t|)^{-3} \), we rely on [20, Theorem 9.9] to deduce
the following bound, with \( C' = \|h_k\|_\infty \frac{\Omega}{2} \), \( I_k = [a_k - 2^{-s'_k}, a_k + 2^{-s'_k}] \) and \( s'_k = s_k + 1 \):

\[
|g_k(t) - h_k(t)| \leq C' \begin{cases}
(1 + (t - a_k) - 2^{-s'_k})^{-2} - (1 + (t - a_k) + 2^{-s'_k})^{-2} & t > a_k + 2^{-s'_k} \\
2 - (1 + (t - a_k) + 2^{-s'_k})^{-2} - (1 - (t - a_k) + 2^{-s'_k})^{-2} & t \in I_k \\
(1 - (t - a_k) - 2^{-s'_k})^{-2} - (1 - (t - a_k) + 2^{-s'_k})^{-2} & t < a_k - 2^{-s'_k}
\end{cases}
\]

We obtain the joint bound, \( |g_k(t) - h_k(t)| \leq C_U (1 + |t - a_k|)^{-2} \) by setting \( C_k = C' \cdot (1 + 2^{-s'_k})^2 \) for all \( k \in \mathbb{Z} \) and \( C_U = \max_{k \in \mathbb{Z}} C_k = 0.0282 < \sqrt{A_k} \lambda^{-1} = 0.0768 \), with \( \lambda \) as defined in (21). Thus, by Proposition 3.7, \( \{M_{I_{lb_k}} g_k : k, l \in \mathbb{Z}\} \) is a nonstationary Gabor frame with a lower frame bound \( A = 0.2 \).

**Remark 7.** Note that the construction presented in the Example 4.1 is of particular interest for constructing frequency-adaptive frames with windows that are compactly supported in time. This is a situation of considerable interest in applications, since it allows for real-time implementation with finite impulse response filters, cp. [11].

In the second example we construct a nonstationary Gabor frame by applying Corollary 3.8. The windows of the new frame coincide with the windows of a painless frame on their support. The windows in this example are constructed in analogy to the windows used in scale frames, introduced in [2] to automatically improve the resolution of transients in audio signals.

**Example 4.2.** As in the previous example, let \( s_k \in \{-1, 0, 1\} \) with \( |s_k - s_{k-1}| \in \{0, 1\} \) for all \( k \in \mathbb{Z} \). We consider a sequence of windows \( g_k \) that are translated and dilated versions of the Gaussian window \( g(t) = e^{-\pi (2.5)^2} \cdot g_k(t) = T_{a_k} \sqrt{2^{s_k}} g(2^{s_k} t) \) with \( a_0 = 0 \) and

\[
a_{k+1} = a_k + 2^{-s_{k+1}} - 1 & \quad \text{if} \quad s_k = s_{k+1} \\
a_{k+1} = a_k + \frac{1}{3} \cdot 2^{-s_{k+1}} & \quad \text{if} \quad s_k > s_{k+1} \\
a_{k+1} = a_k + \frac{1}{3} \cdot 2^{-s_k} & \quad \text{if} \quad s_k < s_{k+1}
\]

Here, the \( \{a_k : k \in \mathbb{Z}\} \) are separated with minimum distance \( \delta = 1/4 \). We arrange the windows as follows: after each change of window size, no change is allowed in the next step; in other words, each window has at least one neighbor of the same size. An example of the arrangement is shown in Figure 1.

Let \( I_k = [a_k - 2^{-s_k-1}, a_k + 2^{-s_k-1}] \) and define a new set of windows by \( h_k(t) = g_k(t) \chi_{I_k} \). Then \( \{M_{I_{lb_k}} h_k : k, l \in \mathbb{Z}\} \) with \( b_k = 2^{s_k} \) is a painless nonstationary Gabor frame. By numerical calculations, its lower frame bound is \( A_h = 0.1609 \) and \( \psi_k(t) = g_k(t) - h_k(t) \) can be bounded by

\[
|\psi_k(t)| \leq \begin{cases}
\sqrt{2^{s_k}} g(1/2)(1 + t - a_k - 2^{-s_k-1})^{-19} & t > a_k + 2^{-s_k-1} \\
0 & t \in I_k \\
\sqrt{2^{s_k}} g(1/2)(1 + a_k - 2^{-s_k-1} - t)^{-19} & t < a_k - 2^{-s_k-1}
\end{cases}
\]

From Proposition 3.8 it follows that \( 4(\delta b^2) (1 + 2^{-s_k-1}) \cdot C_U^2 \cdot E_1 \cdot E_2 = 0.0071 < A_h \), and \( \{M_{I_{lb_k}} T_{a_k} g_k : k, l \in \mathbb{Z}\} \) is a nonstationary Gabor frame with a lower frame bound \( A = 0.1538 \).
5. Acknowledgement

This work was supported by the WWTF project Audiominer (MA09-24) and the Austrian Science Fund (FWF):[T384-N13] Locatif.


Nonstationary Gabor Frames - Approximately Dual Frames and Reconstruction Errors

Monika Dörfler, Ewa Matusiak ∗†

January 9, 2013

Abstract

Nonstationary Gabor frames, recently introduced in adaptive signal analysis, represent a natural generalization of classical Gabor frames by allowing for adaptivity of windows and lattice in either time or frequency. Due to the lack of a complete lattice structure, perfect reconstruction is in general not feasible from coefficients obtained from nonstationary Gabor frames. In this paper it is shown that for nonstationary Gabor frames that are related to some known frames for which dual frames can be computed, good approximate reconstruction can be achieved by resorting to approximately dual frames. In particular, we give constructive examples for so-called almost painless nonstationary frames, that is, frames that are closely related to nonstationary frames with compactly supported windows. The theoretical results are illustrated by concrete computational and numerical examples.

Keywords: adaptive representations, nonorthogonal expansions, irregular Gabor frames, reconstruction, approximately dual frame

1 Introduction

Adapted and adaptive signal representation have received increasing interest over the past few years. As opposed to classical approaches such as the short-time Fourier

∗This work was supported by the WWTF project Audiominer (MA09-24)
†The authors are with the Department of Mathematics, NuHAG, University of Vienna, Nordbergstrasse 15, 1090 Wien, Austria (e-mail: monika.doerfler@univie.ac.at, ewa.matusiak@univie.ac.at)
transform (STFT) or wavelet transform, adaptive representations allow for a variation of parameters such as window width or sampling density over time, frequency or both. Changing parameters in the frequency domain leads, for example, to non-uniform filter banks while adapting window width and sampling density in time is reminiscent of the approach suggested in the construction of nonuniform lapped transforms. Transforms featuring simultaneous adaptivity in time and frequency are notoriously difficult to construct and implement, cp. [15, 7, 12]; however they have shown to be useful in some applications, cf. [14]. On the other hand, fast and efficient implementations exist for representations with adaptivity in only time or frequency. One recent method to obtain this kind of representations is represented by nonstationary Gabor frames, first suggested in [11] and further developed in [1, 18, 10]. All the known implementations rely on compactness of the used analysis window in either time of frequency. This assumption allows for usage of tools developed for painless non-orthogonal expansions [6]. While a priori very convenient, the restriction to using compactly supported windows in the domain for which one wishes a flexible representation can be undesirable. As an example, we mention the construction of nonuniform filter banks via nonstationary Gabor frames, in which case this restriction forbids finite impulse response (FIR) filters; the latter are, however, imperative for real-time processing applications.

In the current contribution, we therefore go beyond the results presented in the references above and consider nonstationary Gabor frames with fast decay but unbounded support. The existence of this kind of frames was shown in [8]. Here we are concerned with methods for approximate reconstruction for these adaptive systems.

2 Notation and Preliminaries

Given a non-zero function $g \in L^2(\mathbb{R})$, a modulation, or frequency shift, operator $M_{bl}$ is defined by $M_{bl}g(t) := e^{2\pi iblt}g(t)$, and time shift operator $T_{ak}$ by $T_{ak}g(t) := g(t-ak)$. A composition, $g_{k,l} = M_{bl}T_{ak}g(t) := e^{2\pi iblt}g(t-ak)$ is a time-frequency shift operator.

The set $\mathcal{G}(g, a, b) = \{g_{k,l} : k, l \in \mathbb{Z}\}$ is called a Gabor system for any real, positive $a, b$. $\mathcal{G}(g, a, b)$ is a Gabor frame for $L^2(\mathbb{R})$, if there exist frame bounds $0 < A \leq B < \infty$ such that for every $f \in L^2(\mathbb{R})$ we have

$$A\|f\|_2^2 \leq \sum_{k,l \in \mathbb{Z}} |\langle f, g_{k,l} \rangle|^2 \leq B\|f\|_2^2.$$  (1)
When working with irregular grids, we assume that the sampling points form a separated set: a set of sampling points \( \{a_k : k \in \mathbb{Z}\} \) is called \( \delta \)-separated, if \( |a_k - a_m| > \delta \) for \( a_k, a_m \), whenever \( k \neq m \). \( \chi_I \) will denote the characteristic function of the interval \( I \).

A convenient class of window functions for time-frequency analysis on \( L^2(\mathbb{R}) \) is the Wiener space.

**Definition 1.** A function \( g \in L^\infty(\mathbb{R}) \) belongs to the Wiener space \( W(L^\infty, \ell^1) \) if

\[
\|g\|_{W(L^\infty, \ell^1)} := \sum_{k \in \mathbb{Z}} \text{ess sup}_{t \in Q} |g(t + k)| < \infty, \quad Q = [0, 1].
\]

For \( g \in W(L^\infty, \ell^1) \) and \( \delta > 0 \) we have [9]

\[
\text{ess sup}_{t \in \mathbb{R}} \sum_{k \in \mathbb{Z}} |g(t - \delta k)| \leq (1 + \delta^{-1}) \|g\|_{W(L^\infty, \ell^1)}.
\]  

For \( f \in L^2(\mathbb{R}) \) we use the following normalization of the Fourier transform, which we denote by \( \mathcal{F} \):

\[
\mathcal{F} f(\omega) = \hat{f}(\omega) = \int_{\mathbb{R}} f(t) e^{-2\pi i \omega t} \, dt.
\]

### 3 Nonstationary Gabor Frames

Nonstationary Gabor systems are a generalization of classical Gabor systems of regular time-frequency shifts of a single window function.

**Definition 2.** Let \( g = \{g_k \in W(L^\infty, \ell^1) : k \in \mathbb{Z}\} \) be a set of window functions and let \( b = \{b_k : k \in \mathbb{Z}\} \) be a corresponding sequence of frequency-shift parameters. Set \( g_{k,l} = M_{b_k,l} g_k \). Then, the set

\[
\mathcal{G}(g, b) = \{g_{k,l} : k, l \in \mathbb{Z}\}
\]

is called a nonstationary Gabor (NSG) system.

In generalization of regular Gabor frames, for which \( g_k = T_{a_k} g \), we will usually assume that the windows \( g_k \) are localized around points \( a_k \) in a separated set of time-sampling points \( \{a_k : k \in \mathbb{Z}\} \). Further, we will always make the assumption that the frequency sampling parameters \( b_k \) are positive numbers contained in a closed interval, i.e. \( b_k \in [b_L, b_U] \subset \mathbb{R}^+ \) for all \( k \in \mathbb{Z} \).
To every collection (3) we associate the analysis operator $C_g$ given by $(C_g f)_{k,l} = \langle f, g_{k,l} \rangle$, and synthesis operator $U_g$, where $U_g c = \sum_{k,l \in \mathbb{Z}} c_{k,l} g_{k,l}$ and $c \in \ell^2$. For two Gabor systems $\mathcal{G}(g, b)$ and $\mathcal{G}(\gamma, b)$ the composition $S_{g,\gamma} = U_\gamma C_g$,

$$\begin{align*}
S_{g,\gamma} f &= \sum_{k,l \in \mathbb{Z}} \langle f, M_{lb} g_k \rangle M_{lb} \gamma_k, \\
\end{align*}
$$

admits a Walnut representation for all $f \in L^2(\mathbb{R})$, [8]:

$$\begin{align*}
S_{g,\gamma} f(t) &= \sum_{k,l \in \mathbb{Z}} b_k^{-1} g_k(t - lb_k^{-1}) \gamma_k(t) f(t - lb_k^{-1}).
\end{align*}
$$

We will frequently use the following correlation functions of a pair of Gabor systems:

$$\begin{align*}
G^{g,\gamma}_l(t) &= \sum_{k \in \mathbb{Z}} b_k^{-1} |g_k(t - lb_k^{-1})| |\gamma_k(t)|, \text{ for } l \in \mathbb{Z}.
\end{align*}
$$

Note that this definition is asymmetric with respect to $g$ and $\gamma$. Using this notation, we obtain the following bounds for the frame operator (4):

$$\begin{align*}
\|S_{g,\gamma}\|^2 &\leq \operatorname{ess sup} \sum_{l \in \mathbb{Z}} G^{g,\gamma}_l \cdot \operatorname{ess sup} \sum_{l \in \mathbb{Z}} G^{\gamma,g}_l.
\end{align*}
$$

By inspection of (5), we note that the summands corresponding to $l \neq 0$ may be seen as the off-diagonal entries of the frame operator. We thus isolate the diagonal part

$$\begin{align*}
G^{g,\gamma}_0(t) &= \sum_{k \in \mathbb{Z}} b_k^{-1} |g_k(t)| |\gamma_k(t)|
\end{align*}
$$

and denote the off-diagonal entries as follows:

$$\begin{align*}
R_{g,\gamma} &= \operatorname{ess sup} \sum_{l \in \mathbb{Z} \setminus \{0\}} \sum_{k \in \mathbb{Z}} b_k^{-1} |g_k(\cdot - lb_k^{-1})| |\gamma_k(\cdot)|.
\end{align*}
$$

Note that, if $g = \gamma$, then the diagonal part of the frame operator $S_{g,g}$ is equal to $G^{g,g}_0$. Using this notation, we obtain the following additional bound:

$$\begin{align*}
\langle S_{g,\gamma} f, f \rangle \|f\|_2^{-2} &\leq \operatorname{ess sup} G^{g,\gamma}_0 + \sqrt{R_{g,\gamma} \cdot R_{\gamma,g}}.
\end{align*}
$$

Bessel sequences are of particular importance in the theory of frames and Riesz bases,[5, 19]. In the regular Gabor case, where $g_k(t) = g(t - ak)$ for some $a > 0$, it is sufficient to assume $g \in W(L^\infty, \ell^1)$ to obtain a Bessel sequence. We next provide a generalization of this property to NSG frames.
Proposition 3.1. Let $G(g, b)$ be a NSG system. If $g_k \in W(L^\infty, \ell^1)$ for all $k \in \mathbb{Z}$ with $\sup_{k \in \mathbb{Z}} \|g_k\|_{W(L^\infty, \ell^1)}$ bounded, and $\sum_{k \in \mathbb{Z}} |g_k(t)| \leq B$ almost everywhere for some $B < \infty$, then the sequence $g_{k,l}$ is a Bessel sequence.

Proof. Let $f \in L^2(\mathbb{R})$. Then by the assumption on the windows $g_k$ and estimate (7)

$$\sum_{k,l \in \mathbb{Z}} |\langle f, g_{k,l} \rangle|^2 = \langle S_g g, f \rangle \leq \|f\|_2^2 \cdot \text{ess sup} \sum_{l \in \mathbb{Z}} G_l^g g$$

$$= \|f\|_2^2 \cdot \text{ess sup} \sum_{k \in \mathbb{Z}} b_k^{-1} |g_k(\cdot)| \sum_{l \in \mathbb{Z}} |g_k(\cdot - lb_k^{-1})|$$

$$\leq \|f\|_2^2 \cdot \text{ess sup} \sum_{k \in \mathbb{Z}} b_k^{-1} |g_k(\cdot)| (1 + b_k^{-1}) \|g_k\|_{W(L^\infty, \ell^1)}$$

$$\leq \|f\|_2^2 \cdot B \cdot \sup_{k \in \mathbb{Z}} [(1 + b_k) \|g_k\|_{W(L^\infty, \ell^1)}].$$

Given a frame, it is well known that there exists at least one dual frame $G(\gamma, b)$ such that

$$f = \sum_{k,l \in \mathbb{Z}} \langle f, \gamma_{k,l} \rangle g_{k,l}, \quad \text{for all } f \in L^2(\mathbb{R}).$$

(11)

The canonical dual frame is given by $\gamma_{k,l} = S^{-1} g_{k,l}$. In the regular Gabor case, the dual frames are again Gabor frames, i.e., they consist of time-frequency shifted versions of one dual window. This is due to the fact, that the frame operator $S$ commutes with time-frequency shifts, hence $\gamma_{k,l} = S^{-1} g_{k,l} = S^{-1} M_{lk} T_{ak} g = M_{lk} T_{ak} S^{-1} g = M_{lk} T_{ak} \gamma$. In general, we cannot expect, that the dual frame of a NSG frame is again a NSG frame. However, even in the case of regular Gabor frames, it is often difficult to calculate a dual frame explicitly. For that reason, alternative possibilities to obtain perfect or approximate reconstruction have been proposed, [4, 2]. The following lemma quantifies the reconstruction error using general pairs of Bessel sequences.

Lemma 3.2. Let $G(g, b)$ and $G(\gamma, b)$ be two Bessel sequences. Then

$$\|I - S_{g,\gamma}\| \leq \left\|1 - \sum_{k \in \mathbb{Z}} b_k^{-1} \overline{g_k} \gamma_k\right\|_{\infty} + \sqrt{R_{g,\gamma} \cdot R_{\gamma,g}}.$$  

(12)

Proof. Starting from the Walnut representation of $S_{g,\gamma}$, we estimate using Cauchy-Schwartz inequality for sums and integrals and, since all summands have absolute value, Fubini’s theorem to justify changing the order of summation and integral:
\[ |\langle f - S_{g, \gamma} f, f \rangle| = |\langle f - \sum_{k \in \mathbb{Z}} b_k^{-1} g_k \gamma_k f, f \rangle - \left( \sum_{l \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} b_k^{-1} \gamma_k (\cdot - lb_k^{-1}) f (\cdot - lb_k^{-1}), f \right) | \]
\[ \leq \left\| 1 - \sum_{k \in \mathbb{Z}} b_k^{-1} g_k \gamma_k \right\|_\infty \|f\|_2^2 + \sqrt{R_{g, \gamma} \cdot R_{\gamma, g}} \|f\|_2^2, \tag{13} \]

since
\[ \left| \left\langle \sum_{k \in \mathbb{Z}} b_k^{-1} \gamma_k (\cdot - lb_k^{-1}) f (\cdot - lb_k^{-1}), f \right\rangle \right| \leq \sum_{k \in \mathbb{Z}} b_k^{-1} \int_\mathbb{R} |\gamma_k(t)||g_k(t - lb_k^{-1})||f(t - lb_k^{-1})||f(t)| \, dt \]
\[ \leq \sum_{k \in \mathbb{Z}} b_k^{-1} \left[ \int_\mathbb{R} |g_k(t - lb_k^{-1})||\gamma_k(t)||f(t - lb_k^{-1})|^2 \, dt \right]^{1/2} \left[ \int_\mathbb{R} |\gamma_k(t)||f(t)|^2 \, dt \right]^{1/2} \]
\[ \leq \left[ \int_\mathbb{R} |f(t)|^2 \sum_{k \in \mathbb{Z}} b_k^{-1} |g_k(t)||\gamma_k(t - lb_k^{-1})| \, dt \right]^{1/2} \left[ \int_\mathbb{R} |f(t)|^2 \sum_{k \in \mathbb{Z}} b_k^{-1} |\gamma_k(t)||g_k(t - lb_k^{-1})| \, dt \right]^{1/2}. \]

A special class of NSG systems are collections of compactly supported windows. They were first addressed in [1]. The collection \( G(g, b) \) with windows \( g_k \) being compactly supported with \( |\text{supp } g_k| \leq \frac{1}{b_k} \) for all \( k \) is a frame for \( L^2(\mathbb{R}) \) if there exist constants \( A > 0 \) and \( B < \infty \) such that
\[ A \leq G_0^{g, b}(t) \leq B \text{ \ a.e..} \tag{15} \]

In this situation, \( G(g, b) \) is called **painless NSG frame**. The canonical dual atoms are given by \( \gamma_{k,l} = M_{lb_k} (G_0^{g, b})^{-1} g_k \). Note again that, in general, we may have \( \gamma_{k,l} = S^{-1}(M_{lb_k} g_k) \neq M_{lb_k} (S^{-1} g_k) \). If \( b_k = b \) for all \( k \), then the frame operator commutes with the frequency-shifts and the dual frame is an NSG frame.

The existence of NSG frames with not necessarily compactly supported windows was established in [8]. For these frames, finding canonical dual frames requires the inversion of the frame operator. This computation is expensive since the operator has considerably less structure than the frame operator in the classical, regular Gabor frame case, for which fast algorithms now exist, [17, 13, 16]. To circumvent the problem, we suggest the use of windows other than canonical duals to obtain sufficiently good approximate reconstruction.
4 Approximately dual atoms

The notion of approximately dual pairs was discussed in [4]. For NSG Bessel sequences we adapt their definition as follows.

**Definition 3.** Two Bessel sequences $\mathcal{G}(g, b)$ and $\mathcal{G}(\gamma, b)$ are said to be approximately dual frames if $\|I - S_{g, \gamma}\| < 1$ or $\|I - S_{\gamma, g}\| < 1$.

Note that the two conditions given in the definition are equivalent since

$$
\|I - S_{g, \gamma}\| = \|I - C_g U_{\gamma}\| = \|I - U_{\gamma}^* C_g^*\| = \|I - C_{\gamma} U_g\| = \|I - S_{\gamma, g}\|.
$$

In Definition 3 it is implicitly stated that, if two Bessel sequences are approximately dual frames, then each of them is a frame. This result was proved in [4] for general frames and it will be useful to reformulate the conditions for the existence of NSG frame, given in [8], in the context of approximately dual frames.

**Proposition 4.1.** Let $\mathcal{G}(g, b)$ be a Bessel sequence with Bessel bound $B$ and $0 < A_1 \leq \sum_{k \in \mathbb{Z}} |g_k(t)|^2 \leq A_2 < \infty$ a.e. for some positive constants $A_1, A_2$.

i) The multiplication operator $G_{0}^{g^*g}$ is invertible a.e. and, for $\gamma_k = (G_{0}^{g^*g})^{-1} g_k$,

$$\|I - S_{g, \gamma}\| \leq \frac{R_{g,g}}{\text{ess inf } G_{0}^{g^*g}}. \tag{16}$$

ii) If

$$R_{g,g} < \text{ess inf } G_{0}^{g^*g}, \tag{17}$$

then $\mathcal{G}(g, b)$ and $\mathcal{G}(\gamma, b)$ are approximately dual frames for $L^2(\mathbb{R})$.

iii) Assume, additionally, for some $\delta$-separated set of time-sampling points $\{a_k : k \in \mathbb{Z}\}$ and constants $0 < p_U, C_L, C_U < \infty$ such that for $p_k \in ]2, p_U[ \subset \mathbb{R}, C_k \in [C_L, C_U]$ we have

$$|g_k(t)| \leq C_k (1 + |t - a_k|)^{-p_k} \text{ for all } k \in \mathbb{Z}. \tag{18}$$

Then there exists a sequence $\{b_k^0\}_{k \in \mathbb{Z}}$, such that for all sequence $b_k \leq b_k^0$, $k \in \mathbb{Z}$, (17) holds.

**Remark 1.** If (17) holds, $\mathcal{G}(\gamma, b)$ is called a single preconditioning dual system for $\mathcal{G}(g, b)$. 

7
Proof. Since all frequency modulation parameters $b_k$ are taken from a closed interval in $\mathbb{R}^+$, the invertibility of $G_{0,g}^{g,g}$ is straightforward and the windows $\gamma_k$ are well defined. Moreover, since $G(g,b)$ is a Bessel sequence, so is $G(\gamma,b)$, with Bessel bound $(\text{ess inf } G_{0,g}^{g,g})^{-2}B$. Substituting $\gamma_k = (G_{0,g}^{g,g})^{-1}g_k$ for $\gamma_k$ in the proof of Lemma 3.2, the first term in (13) vanishes and we obtain (i):

$$|\langle f - S_{g,\gamma} f, f \rangle| \leq \left( (G_{0,g}^{g,g})^{-1} \sum_{l \in \mathbb{Z} \setminus \{0\}} \sum_{k \in \mathbb{Z}} b_k^{-1} g_k(\cdot) g_k(\cdot - lb_k^{-1}) f(\cdot - lb_k^{-1}), f \right)$$

$$\leq (\text{ess inf } G_{0,g}^{g,g})^{-1} \left( \sum_{l \in \mathbb{Z} \setminus \{0\}} \sum_{k \in \mathbb{Z}} b_k^{-1} |g_k(\cdot)| g_k(\cdot - lb_k^{-1}) ||f(\cdot - lb_k^{-1})||, |f| \right)$$

$$\leq (\text{ess inf } G_{0,g}^{g,g})^{-1} R_{g,g} \|f\|_2^2. \quad (19)$$

(ii) follows directly from Definition 3. Finally, (iii) follows from [8, Theorem 3.4], where it is shown that the assumptions (18) on the windows $g_k$ guarantee the existence of a sequence $b_k^0$ such that

$$\langle S_{g,g} f, f \rangle \|f\|_2^{-2} \geq \text{ess inf } G_{0,g}^{g,g} - R_{g,g} > 0. \quad (20)$$

Single preconditioning dual windows are a good choice for reconstruction, whenever the frame operator is close to diagonal. This is the case, if the original windows $g_k$ decay fast and frequency sampling is fast.

If the frame of interest is close to some other frame, which, ideally, is better understood, other approximate dual windows may be derived from this frame. The prototypical situation is a NSG frame which is close to a painless NSG frame in the sense of a small perturbation. Approximately dual frames in the context of perturbation theory were recently studied in [4]. In the following proposition we give error estimates for the reconstruction with approximately dual frames in such a situation. This provides different reconstruction methods apart from single preconditioning which was addressed in Proposition 4.1.

**Proposition 4.2.** Assume that $G(g,b)$ is a Bessel sequence with bound $B$ and that $G(h,b)$ is a NSG frame with lower and upper frame bound $A_h$ and $B_h$, respectively. We set $\psi_k = h_k - g_k$ and define the following windows:

(a) $\gamma_{k,l}^1 = S_{h,h}^{-1} h_{k,l}$ (canonical dual of $h_{k,l}$) \hspace{1cm} (21)

(b) $\gamma_{k,l}^2 = S_{h,h}^{-1} g_{k,l}$ \hspace{1cm} (22)

Then the following hold:
\[(i)\]
\[\|I - S_{g,\gamma^1}\| \leq A_h^{-1/2}\|C_\psi\|.
\] (23)
If \(\text{ess sup} \sum_{l \in \mathbb{Z}} G_{l}^{\psi,\psi} < A_h\), then \(G(\gamma^1, b)\) and \(G(g, b)\) are approximately dual frames.

\[(ii)\]
\[\|I - S_{g,\gamma^2}\| \leq A_h^{-1}(\sqrt{B_h} + \sqrt{B})\|C_\psi\|.
\] (24)
If \(\text{ess sup} \sum_{l \in \mathbb{Z}} G_{l}^{\psi,\psi} < \frac{A_h^2}{(\sqrt{B_h} + \sqrt{B})^2}\), then \(G(\gamma^2, b)\) and \(G(g, b)\) are approximately dual frames.

**Remark 2.** The first statement of Proposition 4.2 is contained in [4].

According to [8], the assumption that \(\text{ess sup} \sum_{l \in \mathbb{Z}} G_{l}^{\psi,\psi} < A_h\) can be satisfied if the functions \(\psi_k\) decay polynomially, i.e., 
\[|\psi_k(t)| \leq C_k(1 + |t|)^{-p_k}\]
with appropriate constants \(C_k\) and decay rates \(p_k > 1\).

**Proof.** From (7) it follows that 
\[\|C_\psi\|^2 \leq \text{ess sup} \sum_{l \in \mathbb{Z}} G_{l}^{\psi,\psi} .\]
The same estimate holds for \(\|U_\psi\|^2\).

Since \(G(h, b)\) is a frame with canonical dual frame \(G(\gamma^1, b)\), an upper frame bound of \(G(\gamma^1, b)\) is given by \(A_h^{-1}\). We thus obtain (i) as follows:

\[\|I - S_{g,\gamma^1}\| = \|U_{\gamma^1}C_h - U_{\gamma^1}C_g\| \leq \|U_{\gamma^1}\||\|C_\psi\|| \leq A_h^{-1/2}\|C_\psi\| .
\] (25)

If \(\text{ess sup} \sum_{l \in \mathbb{Z}} G_{l}^{\psi,\psi} < A_h\), then \(\|I - S_{g,\gamma^1}\| < 1\) and \(G(\gamma^1, b)\) and \(G(g, b)\) are approximately dual frames as claimed.

To show (ii), we note that \(U_{\gamma^2} = S_{h,h}^{-1}U_g\) and thus

\[\|I - S_{g,\gamma^2}\| = \|S_{h,h}^{-1}S_{h,h} - S_{h,h}^{-1}S_{h,g}\| = \|S_{h,h}^{-1}(U_hC_h - U_gC_g)\|\]
\[\leq \|S_{h,h}^{-1}\||U_hC_h - U_gC_h + U_gC_h - U_gC_g\| = A_h^{-1}\|U_\psiC_h - U_gC_\psi\|\]
\[\leq A_h^{-1}\|C_\psi\|(\sqrt{B_h} + \sqrt{B})\] (28)

and (24) follows. \(\Box\)

### 4.1 Perturbation of painless nonstationary Gabor frames

If a NSG system can be derived as a perturbation of a painless NSG frame, the approximately dual windows given in Proposition 4.2 are particularly simple to compute. In this situation, the frame \(G(h, b)\) is the painless frame \(G(g^o, b)\) and the frame operator
$S_{h,h} = S_{g^\gamma,g^\gamma}$ is the multiplication operator $G_0^{\gamma\gamma}$. Moreover, in this particular case, the approximately dual frames, given by $\gamma_{k,l}^1 = (G_0^{\gamma\gamma})^{-1} g_{k,l}^0$ and $\gamma_{k,l}^2 = (G_0^{\gamma\gamma})^{-1} g_{k,l}$ are NSG frames. This is a very important asset, since for NSG frames fast algorithms for analysis and reconstruction using FFT exist.

In [8] we constructed a special class of NSG frames, arising from painless NSG frames, which we introduce next.

**Definition 4** (Almost painless NSG frames). Let $\mathcal{G}(g,b)$ be a NSG system, assume that the windows $g_k$ are essentially bounded away from zero on the intervals $I_k = [a_k - (2b_k)^{-1}, a_k + (2b_k)^{-1}]$ and set $g_k^0 = g_k \chi_{I_k}$. If $\mathcal{G}(g^\gamma,b)$ is a (painless) frame for $L^2(\mathbb{R})$, then we call the system $\mathcal{G}(g,b)$ an almost painless NSG system (or frame).

For almost painless NSG systems, the estimates given in Proposition 4.2 can be written more explicitly.

**Corollary 4.3.** Assume that $\mathcal{G}(g,b)$ is an almost painless NSG system and let $A_0 = \text{ess inf } G_0^{\gamma\gamma}$ to be the lower frame bound of the painless frame $\mathcal{G}(g^\gamma,b)$ and $g_k^0 = g_k - g_k \chi_{I_k}$. Then the following hold:

(i) for $\gamma_{k}^1 = (G_0^{\gamma\gamma})^{-1} g_k^0$,

$$\|I - S_{g^\gamma^1}\| \leq A_0^{-1} \sqrt{R_{g^\gamma^\gamma} \cdot R_{g^\gamma^\gamma}}$$

(ii) for $\gamma_{k}^2 = (G_0^{\gamma\gamma})^{-1} g_k^0$,

$$\|I - S_{g^\gamma^1}\| \leq A_0^{-1} \left( R_{g^\gamma^\gamma} + R_{g^\gamma^\gamma} + \text{ess sup} \sum_{l \in \mathbb{Z}} G_l^{\gamma\gamma} \right)$$

**Proof.** The estimates follow from Lemma 3.2. First, substituting $\gamma_{k}^1 = (G_0^{\gamma\gamma})^{-1} g_k^0$ for $\gamma_k$ in (12), the first term vanishes, since $\sum_{k \in \mathbb{Z}} b_k^{-1} g_k^0 = G_0^{\gamma\gamma}$, and we obtain

$$\|I - S_{g^\gamma^1}\| \leq \sqrt{R_{g^\gamma^1} \cdot R_{g^\gamma^1}} \leq (\text{ess inf } G_0^{\gamma\gamma})^{-1} \sqrt{R_{g^\gamma^\gamma} \cdot R_{g^\gamma^\gamma}},$$

since

$$R_{g^\gamma^1} = \text{ess sup} \sum_{l \in \mathbb{Z} \setminus \{0\}} \sum_{k \in \mathbb{Z}} b_k^{-1} |g_k(\cdot - lb_k^{-1})|| (G_0^{\gamma\gamma})^{-1} g_k^0(\cdot) |$$

$$\leq (\text{ess inf } G_0^{\gamma\gamma})^{-1} \cdot \text{ess sup} \sum_{l \in \mathbb{Z} \setminus \{0\}} \sum_{k \in \mathbb{Z}} b_k^{-1} |g_k(\cdot - lb_k^{-1})||g_k^0(\cdot) |$$

$$= (\text{ess inf } G_0^{\gamma\gamma})^{-1} R_{g^\gamma^\gamma},$$

10
similarly for $R_{\gamma, g}$.

For (ii) we substitute $\gamma_2^k = (G_{0}g_0^g)^{-1}g_k$ for $\gamma$ in the proof of Lemma 3.2, $g_k^g + g_k^r$ for $g_k$ and use the fact that $g_k^g$ and $g_k^r$ have disjoint supports. Then, since $g_k^g(\cdot - lb_k^{-1})g_k^g(\cdot)$ is zero for $l \neq 0$, using (14) we obtain that

$$|\langle f - S_{g, \gamma} f, f \rangle| \leq \left| \left\langle f - \sum_{k \in \mathbb{Z}} \sum_{l \in \mathbb{Z}} g_k(\cdot - lb_k^{-1})(G_{0}g_0^g)^{-1}g_k(\cdot - lb_k^{-1}), f \right\rangle \right|$$

$$+ (\text{ess inf}(G_{0}g_0^g)^{-1} (R_{g^r, g^g} + R_{g^r, g^r} + \|S_{g^r, g^r}\|) \|f\|_2^2 ).$$

The first term vanishes and by (7), $\|S_{g^r, g^r}\| \leq \text{ess sup} \sum_{l \in \mathbb{Z}} G_{l}g^r_g^r$. 

\hfill \Box

\section{Examples}

We present two examples to illustrate our theory. In the first example we deal with almost painless frames using Gaussian windows. We consider three different approximately dual frames and check their performance in terms of reconstruction.

In both examples, we consider a basic window and dilations by 2 and $\frac{1}{2}$, respectively. Since the dilation parameters take only three different values, there are three kinds of windows, with support size $1/2$, 1 and 2, respectively. Note that, while theoretically possible, sudden changes in the shape and width of adjacent windows turn out to be undesirable for applications, hence we only allow for stepwise change in dilation parameters.

\begin{example}
Let $s_k \in \{-1, 0, 1\}$ with $|s_k - s_{k-1}| \in \{0, 1\}$ for all $k \in \mathbb{Z}$. We consider a sequence of windows $g_k$ that are translated and dilated versions of the Gaussian window $g(t) = e^{-\pi |\sigma| t^2}$: $g_k(t) = T_{a_k} \sqrt{b_k} g(b_k(t - a_k))$, with $\sigma = 2.5$, $b_k = 2^k$, $a_0 = 0$ and for all $k \in \mathbb{Z}$

$$a_{k+1} = a_k + (2b_k)^{-1} \quad \text{if} \quad s_k = s_{k+1},$$

$$a_{k+1} = a_k + (3b_{k+1})^{-1} \quad \text{if} \quad s_k > s_{k+1},$$

$$a_{k+1} = a_k + (3b_{k})^{-1} \quad \text{if} \quad s_k < s_{k+1}.$$  

Here, $b_L = 1/2$, $b_U = 2$ and the $\{a_k : k \in \mathbb{Z}\}$ are separated with minimum distance $\delta = 1/4$. We arrange the windows as follows: after each change of window size, no change is allowed in the next step; in other words, each window has at least one neighbor of the same size.

\end{example}
Let $I_k = [a_k - (2b_k)^{-1}, a_k + (2b_k)^{-1}]$ and define a new set of windows by $g_k^o(t) = g_k(t)\chi_{I_k}$. Then $\{M_{b_k}g_k^o : k, l \in \mathbb{Z}\}$ is a painless nonstationary Gabor frame with lower frame bound $A_0 = 0.1718$.

The system $\mathcal{G}(g, b)$ arises from the painless frame $\mathcal{G}(g^o, b)$, and therefore we are interested in the approximate dual windows proposed in Corollary 4.3. We first consider $\gamma_k^1 = (G_{0,b}^{g^o})^{-1}g_k^o$, the canonical dual frame of $\mathcal{G}(g^o, b)$. According to (29), we need to calculate $R_{g^o,g^o}$ and $R_{g^o,g^r}$ in order to obtain an estimate of the reconstruction error.

**Claim 1:** $R_{g^o,g^o} \leq 0.00827 + O(10^{-6}).$ 
For fixed $k \in \mathbb{Z}$,

$$b_k^{-1}|g_k^o(t)| \sum_{l \in \mathbb{Z}\{0\}} |g_k^r(t - lb_k^{-1})| = b_k^{-1/2}|g_k^o(t)| \sum_{l \in \mathbb{Z}} b_k^{-1/2}|g_k^r(t - lb_k^{-1})|$$

since $g_k^o$ and $g_k^r$ have disjoint support. Then, due to $b_k^{-1}$-periodicity of $\sum_{l \in \mathbb{Z}} b_k^{-1/2}|g_k^o(t - lb_k^{-1})|$, we have

$$R_{g^o,g^o} = \text{ess sup} \sum_{k \in \mathbb{Z}} b_k^{-1/2}|g_k^o(\cdot)| \sum_{l \in \mathbb{Z}} b_k^{-1/2}|g_k^r(\cdot - lb_k^{-1})|$$

$$\leq \text{ess sup} \sum_{k \in \mathbb{Z}} b_k^{-1/2}|g_k^o(\cdot)| \cdot \text{ess sup}_{t \in I_k} \sum_{l \in \mathbb{Z}} b_k^{-1/2}|g_k^r(t - lb_k^{-1})|$$

(32)

In order to obtain more accurate estimates, we split $I_k$ by setting $I_k^+ = [a_k, a_k + (2b_k)^{-1}]$ and $I_k^- = [a_k - (2b_k)^{-1}, a_k]$ and estimate expression (32) by its values at the end points of $I_k^+$ or $I_k^-$, respectively. We observe that, for $t \in I_k^-:

$$\sum_{l \in \mathbb{Z}} b_k^{-1/2}|g_k^r(t - lb_k^{-1})| = \sum_{l=1}^{\infty} \left[ e^{-\pi|\sigma b_k(t - a_k - lb_k^{-1})|^2} + e^{-\pi|\sigma b_k(t - a_k + lb_k^{-1})|^2} \right]$$

$$\leq \sum_{l=1}^{\infty} e^{-\pi|\sigma|^2} + \sum_{l=1}^{\infty} e^{-\pi|\sigma(2l - 1)/2|^2} \leq 0.00738 + O(10^{-6}).$$

By the symmetry of $g_k^o$ with respect to $a_k$, an analogous estimate holds for $t \in I_k^+$. Furthermore, due to the arrangement of the windows $g_k$, $\text{ess sup} \sum_{k \in \mathbb{Z}} b_k^{-1/2}|g_k^o(\cdot)| \leq 1.1206$ and Claim 1 follows.

**Claim 2:** $R_{g^o,g^r} \leq 0.0157 + O(10^{-6}).$

Observe that $\sum_{l \in \mathbb{Z}\{0\}} b_k^{-1/2}|g_k^o(t - lb_k^{-1})| \leq \|b_k^{-1/2}g_k^o\|_\infty = 1$. Then, with $I_m^+$ and
We bound $C^+$ and $C^-$ by their maximal values on $I_m^+$, respectively $I_m^-$. The set \( \{a_k : k \in \mathbb{Z}\} \) is \( \delta \)-separated, hence \( |a_k - a_m| \geq |k - m|\delta \). By the arrangement of the windows, there are at most two windows \( g_k^r \) which assume their maximum \( e^{-\pi|\sigma/2|^2} \) in the interval $I_m^+$ or in $I_m^-$. Without loss of generality we assume that the two maximal values occur in $I_m^+$, corresponding to the windows $g_{m-1}^r$ and $g_{m+2}^r$. Then, $g_{m+1}^r$, $g_m^r$.
are zero in $I^+_m$, cf. Figure 2 for an example situation. Therefore,

$$C^+ \leq \sum_{k>m+2} e^{-\pi|\sigma b_k(a_m+(2b_m)^{-1}-a_k)|^2} + 2e^{-\pi|\sigma b_k(2b_m)^{-1}|^2} + \sum_{k<m-1} e^{-\pi|\sigma b_k(a_m-a_k)|^2}$$

$$\leq 2 \left( e^{-\pi|\sigma/2|^2} + \sum_{k>0; kb_k>2} e^{-\pi|\sigma b_k|^2} \right).$$

Since we assumed that two windows $g_k^r$ reached their maximum in $I^+_m$, $C^- \leq C^+$. As before, the bound from Claim 2 follows by numerical calculations.

In summary, due to (29), Claim 1 and 2 and the lower frame bound $A_0 = 0.1718$, the reconstruction error using the approximate duals $\gamma_k^1$ is bounded by

$$\|I - S_{g,\gamma^1}\| \leq 0.0663 + \mathcal{O}(10^{-6}). \quad (33)$$

Another choice of approximate dual system are the windows $\gamma_k^2 = (G_{g,g}^r)^{-1}g_k$. In this setting we use (30) to derive

$$\|I - S_{g,\gamma^2}\| \leq A_0^{-1} \left( R_{g^r,g^r} + R_{g^r,g^o} + \text{ess sup} \sum_{l \in \mathbb{Z}} G^r_l \right) \leq 0.1402 + \mathcal{O}(10^{-6}),$$

since, by previous calculations,

$$\text{ess sup} \sum_{l \in \mathbb{Z}} G^r_l = \text{ess sup} \sum_{k \in \mathbb{Z}} b_k^{-1/2} |g_k^r(t)| \sum_{l \in \mathbb{Z}} b_k^{-1/2} |g_k^r(t-lb_k^{-1})| \quad (34)$$

$$\leq \text{ess sup} \sum_{k \in \mathbb{Z}} b_k^{-1/2} |g_k^r(t)| \cdot \text{ess sup} \sum_{l \in \mathbb{Z}} b_k^{-1/2} |g_k^r(t-lb_k^{-1})|$$

$$\leq 0.0001158 + \mathcal{O}(10^{-6}).$$

As a third choice of approximate dual windows we consider single preconditioning windows $\gamma_k = (G_{0}^{g,g})^{-1}g_k$ introduced in Proposition 4.1. It can easily be seen that

$$R_{g,g} \leq R_{g^r,g^r} + R_{g^r,g^o} + R_{g^o,g^r} \leq R_{g^r,g^r} + R_{g^r,g^o} + \text{ess sup} \sum_{l \in \mathbb{Z}} G^r_l \quad (35)$$

and, by previous calculations, it follows

$$R_{g,g} \leq 0.0241 + \mathcal{O}(10^{-6}) < A_0 \leq \text{ess inf} G_{0}^{g,g}.$$ 

Therefore, by Proposition 4.1 (ii), $G(g, b)$ and $G(\gamma, b)$ are approximately dual frames. Moreover

$$\|I - S_{g,\gamma}\| \leq \frac{R_{g,g}}{\text{ess inf} G_{0}^{g,g}} \leq \frac{R_{g,g}}{A_0} \leq 0.1402 + \mathcal{O}(10^{-6}). \quad (37)$$
Remark 3. Observe that from each of the approximately dual frame estimates given in the above example, the frame property of $G(g,b)$ follows.

Note that, from (34), the frame property of $G(g,b)$ may also be derived by applying results from perturbation theory cf. [3]. Indeed, since

$$\sum_{k,l \in \mathbb{Z}} |\langle f, g_{k,l} - g_{o,k,l} \rangle|^2 \leq \text{ess sup} \sum_{l \in \mathbb{Z}} G_{g_{r_1}^* f} \| f \|_2^2 < A_0 \| f \|_2^2,$$

it follows that $G(g,b)$ is a frame with a lower frame bound $A = 0.1630$.

Example 5.2. In our second example, we turn to the situation mentioned in the introduction, namely, the construction of non-uniform filter banks with compactly supported, that is, FIR filters, via NSG frames. In this situation, the frame operator $S$ does not have a Walnut-like structure as given in (5) on the time side. However, $S$ may be considered on the frequency side by applying a Fourier transform. Then, we encounter the same structure as before and may exploit the developed techniques to deduce the frame property and to construct approximate dual frames for reconstruction. The situation is schematically depicted in Figure 3.

![Figure 3: An example for the arrangement of dilated windows in Example 2](image)

It is obvious that, while, in the time domain, various windows $g_k$ with different pass-bands, are shifted to obtain the overall system, applying a Fourier transform yields the known situation: various windows $\hat{g}_k$ are modulated to create an NSG system.

On the other hand, we are interested in using FIR filters, that is, the windows $g_k$
are all compactly supported, hence not bandlimited. Similar to the construction in Example 5.1, we can cut the windows \( \hat{g}_k \) to obtain a painless NSG reference frame.

More precisely, we consider the family of windows \( g_k \) and a vector of corresponding time-shift parameters \( a_k \). Then, we set \( g_{k,l} = T_{a_k} g_k \) and are interested in the frame property of the set of functions \( \{ g_{k,l} : k, l \in \mathbb{Z} \} \). Considering, due to the lack of structure on the time side as mentioned above, the corresponding frame operator on the frequency side corresponds to investigating the operator \( \mathcal{F} S \mathcal{F}^* \), which acts on the Fourier transform of a signal of interest. In other words, we are now dealing with the set of functions \( \{ \mathcal{F}(g_{k,l}) = M_{a_k} \hat{g}_k : k, l \in \mathbb{Z} \} \).

For the current example, we consider Hanning windows \( h_k \) and, as in the previous example, apply dilations by \( 2^{-1} \) and \( 2 \), respectively, to obtain various time- and frequency resolutions. The time-shift parameters \( a_k \) are chosen in parallel to the choice of the frequency-shift parameters \( b_k \) in Example 5.1.

Given the explicit knowledge of the spectral properties of the Hanning windows, explicit error estimates can be derived in a similar manner as in the previous example. Here, we also numerically calculate the errors resulting from reconstruction by means of the three different proposed approximate dual systems. We use the same nomenclature as before, that is, \( \gamma_{k,1} \), \( \gamma_{k,2} \) and \( \gamma_{k,3} \) denote the canonical duals of the painless frame, the set of windows \( (G_0^{\sigma_1,\sigma_0})^{-1} g_k \) and the single preconditioning windows, respectively. Then

1. \( \| I - S_{g,\gamma_1} \| = 0.0210 \)
2. \( \| I - S_{g,\gamma_2} \| = 0.0407 \)
3. \( \| I - S_{g,\gamma_3} \| = 0.0407 \)

As before, the canonical duals of the painless frame provide the best approximate reconstruction.

The set of windows \( \hat{g}_k \) used in this example, together with their true, single preconditioning duals \( \gamma_{k,3} \) and the canonical duals of the corresponding painless frame, \( \gamma_{k,1} \), are depicted in Figure 4. On the right plots, zoom-ins are shown to better compare the detailed behavior.

A comparison between a basic window from the true dual frame and the preconditioning duals \( \gamma_{k,3} \) and the true dual and the approximate dual \( \gamma_{k,1} \) is depicted in Figures 5 and 6, respectively, in both the frequency domain (upper plot) and the time domain (lower plot). It should be noted that, while compactly supported in frequency, \( \gamma_{k,1} \)
still have better decay in the time domain than both the true dual windows and the other approximate dual windows.

References


Figure 4: The original windows and the windows from various (approximate) dual frames.
Figure 5: Comparison between true dual and single preconditioning dual $\gamma^3$.

Figure 6: Comparison between true dual and painless approximate dual $\gamma^1$. 
Quilted Gabor frames – A new concept for adaptive time-frequency representation

Monika Dörfler

Institut für Mathematik, Universität Wien, Alserbachstrasse 23, A-1090 Wien, Austria

**A R T I C L E   I N F O**

**Article history:**
Received 18 January 2010
Revised 23 February 2011
Accepted 24 February 2011
Available online 1 April 2011

**MSC:**
42A65
42C15
42C40

**Keywords:**
Time-frequency analysis
Adaptive representation
Uncertainty principle
Frame bounds
Frame algorithm

**A B S T R A C T**

Certain signal classes such as audio signals call for signal representations with the ability to adapt to the signal’s properties. In this article we introduce the new concept of quilted frames, which aim at adaptivity in time-frequency representations. As opposed to Gabor or wavelet frames, this new class of frames allows for the adaptation of the signal analysis to the local requirements of signals under consideration. Quilted frames are constructed directly in the time-frequency domain in a signal-adaptive manner. Validity of the frame property guarantees the possibility to reconstruct the original signal. The frame property is shown for specific situations and the Bessel property is proved for the general setting. Strategies for reconstruction from coefficients obtained with quilted Gabor frames and numerical simulations are provided as well.

@2011 Elsevier Inc. All rights reserved.

1. Introduction

QUILT (verb): (a) to fill, pad, or line like a quilt
(b) to stitch (designs) through layers of cloth
(c) to fasten between two pieces of material

Natural signals usually comprise components of various different characteristics and their analysis requires judicious choice of processing tools. For audio signals time-frequency dictionaries have proved to be an adequate option. Since orthonormal bases cannot provide good time-frequency res-

---

E-mail address: monika.doerfler@univie.ac.at.

1 The author was supported by the Austrian Science Fund (FWF): [T384-N13].
olution [24], time-frequency analysis naturally leads to the use of frames. Most classes of frames commonly used in applications, be it wavelet or Gabor frames, feature a resolution following a fixed rule over the whole time-frequency or time-scale plane, respectively. The concept of quilted frames, as introduced in this contribution, gives up this uniformity and allows for different resolutions in assigned areas of the time-frequency plane.

The primary motivation for introducing this new class of frames stems from the processing of audio and in particular music signals, where the trade-off between time and frequency resolution has a strong impact on the results of analysis and synthesis, see [29,11,33,32,26]. The well-known uncertainty principle makes the choice of just one analysis window a difficult task: different resolutions might be favorable in order to achieve sparse and precise representations for the various signal components. For example, percussive elements require short analysis windows and high sampling rate in time, whereas sustained sinusoidal components are better represented with wide windows and a longer FFT, thus more sampling points in frequency.

Several approaches have been suggested to deal with the trade-off in time-frequency resolution. The notion of multi-window Gabor expansions, introduced by Zibulski and Zeevi [34], uses a finite number of windows of different shape in order to obtain a richer dictionary with the ability to better represent certain characteristics in a given signal class. Another approach is the usage of several bases in order to best describe the components of a signal with a priori known characteristics, see [10]. All these approaches, however, stick to a uniform resolution guided by the action of a certain group via a unitary representation. For quilted Gabor frames we give up this restriction and introduce systems constructed from globally defined frames by restricting these to certain, possibly compact, regions in the time-frequency or time-scale plane. The idea of realizing tilings of the time-frequency plane has been suggested in [5] and [31], however, these authors stick to the construction of orthogonal bases. In this case, every basis function corresponds to a particular tile. We will achieve a wider range of possible partitions, windows and sampling schemes by allowing for redundancy. Thus we aim at designing systems that can adapt to a class of signals considered. As a particular example of quilted frames, the notion of reduced multi-Gabor frames was first introduced in [12] and recently exploited in [26]. Note that this model allows, for example, a transform yielding constant-Q spectral resolution, which is invertible, as opposed to the original construction [6]. Reduced multi-Gabor frames were successfully applied to the task of denoising corrupted audio signals, see [33]. The processing of sound signals also yields a motivation for the next step in generalizing the idea to quilted frames, which allow arbitrary tilings of the time-frequency plane, see [27].

Quilted frames also bear theoretical interest in themselves and should be compared to constructions such as fusion frames [8] and the frames proposed in [1]. In fact, the construction of quilted frames provides constructive examples for the models presented in these contributions.

For the mathematical description of quilted frames, we start from principles of Gabor analysis [21]. The idea for the construction of quilted Gabor frames is inspired by the early work of Feichtinger and Gröbner on decomposition methods [19,17] and recent results on time-frequency partitions for the characterization of function spaces [13,14]:

Assume that a covering \( \{ \Omega_r \}_{r \in \mathbb{I}} \) of the phase space \( \mathbb{R}^{2d} \) is given. To each member of the covering a frame from a family of Gabor frames is assigned, hence, the new system locally resembles the original frames. The resulting global system will be called a quilted Gabor system. We conjecture that these systems may be shown to constitute frames under certain, rather general conditions. In this paper we will show the frame property in two special cases and proof the existence of an upper frame bound for a general setting.

The rest of this paper is organized as follows. Section 2 provides notation and gives an overview over basic results in Gabor analysis. Section 3 introduces the general concept of quilted Gabor frames. In Section 4, the existence of an upper frame bound (Bessel property) for general quilted frames is proved. In Section 5 and Section 6, a lower frame bound is constructed for two particular cases, namely, the partition of the time-frequency plane in stripes and the replacement of frame elements in a compact region of the coefficient domain. Finally, Section 7 presents numerical examples for these cases.
2. Notation and some basic facts from Gabor theory

We use the normalization \( \hat{f}(\omega) = \int_{\mathbb{R}^d} f(t) e^{-2\pi i \omega t} dt \) of the Fourier transform on \( L^2(\mathbb{R}^d) \). \( M_\omega \) and \( T_x \) denote frequency-shift by \( \omega \) and time-shift by \( x \), respectively, of a function \( g \), combined to the time-frequency shift operators \( \pi(\lambda) = M_\omega T_x g(t) = e^{2\pi i \omega t} g(t-x) \) for \( \lambda = (x, \omega) \in \mathbb{R}^{2d} \).

The short-time Fourier transform (STFT) of a function \( f \in L^2(\mathbb{R}^d) \) with respect to a window function \( g \in L^2(\mathbb{R}^d) \) is defined as

\[ \mathcal{V}_g f(\lambda) = \int_{\mathbb{R}^d} f(t) \hat{g}(t-x)e^{-2\pi i \omega t} dt = \langle f, \pi(\lambda)g \rangle. \] (1)

A lattice \( \Lambda \subset \mathbb{R}^{2d} \) is a discrete subgroup of \( \mathbb{R}^{2d} \) of the form \( \Lambda = A \mathbb{Z}^{2d} \), where \( A \) is an invertible \( 2d \times 2d \)-matrix over \( \mathbb{R} \). The special case \( \Lambda = \alpha \mathbb{Z}^{d} \times \beta \mathbb{Z}^{d} \), where \( \alpha, \beta > 0 \) are the lattice constants, is called a separable or product lattice.

A family of functions \( (g_k)_{k \in \mathbb{Z}} \) in \( (\mathbb{R}^d) \) is called a frame, if there exist lower and upper frame bounds \( A, B > 0 \), so that

\[ A \| f \|^2 \leq \sum_{k \in \mathbb{Z}} | \langle f, g_k \rangle |^2 \leq B \| f \|^2 \quad \text{for all } f \in L^2(\mathbb{R}^d). \] (2)

Assumption (2) can be understood as an “approximate Plancherel formula”. It guarantees that any signal \( f \in L^2(\mathbb{R}^d) \) can be represented as infinite series with square integrable coefficients using the elements \( g_k \). The existence of the upper bound \( B \) is called Bessel property of the sequence \( (g_k)_{k \in \mathbb{Z}} \). The frame operator \( S \), defined as

\[ Sf = \sum_{k \in \mathbb{Z}} \langle f, g_k \rangle g_k \]

allows the calculation of the canonical dual frame \( (\gamma_k)_{k \in \mathbb{Z}} = (S^{-1}g_k)_{k \in \mathbb{Z}} \), which guarantees minimal-norm coefficients in the expansion

\[ f = \sum_{k \in \mathbb{Z}} \langle f, \gamma_k \rangle g_k = \sum_{k \in \mathbb{Z}} \langle f, g_k \gamma_k \]. (3)

If \( A = B \), the frame is called tight and \( f = \frac{1}{A} A \sum_{k \in \mathbb{Z}} \langle f, g_k \rangle g_k \). We refer the interested reader to Christensen’s book [9] for more details on general frames.

In the special case of Gabor or Weyl–Heisenberg frames, the frame elements are generated by time-frequency shifts of a basic atom or window \( g \) along a lattice \( \Lambda \):

\[ g_\lambda = \pi(\lambda)g. \]

In this case we write \( S = S_g \), and \( S_g = T_g^* T_g \), where \( T_g : f \mapsto [(f, g_\lambda)]_\lambda \) is the analysis operator mapping the function \( f \in L^2(\mathbb{R}^d) \) to its coefficients \( c(f)(\lambda) = T_g f(\lambda) \). These coefficients correspond to the samples of the STFT on \( \Lambda \). Its adjoint \( T_g^* : [c_\lambda]_{\lambda \in \Lambda} \mapsto \sum_{\lambda \in \Lambda} c_\lambda \pi(\lambda)g \) is the synthesis operator. For the Gabor frame generated by time-frequency shifts of the window \( g \) along the lattice \( \Lambda \) we write \( G(g, \Lambda) \).

We next introduce the concept of partitions of unity. A family \( (\psi_r)_{r \in \mathbb{Z}} \) of non-negative functions with \( \sum_r \psi_r(x) \equiv 1 \) is called bounded admissible partition of unity (BAPU) subordinate to
(\mathcal{B}_{R_r,\omega_r})_{r \in \mathcal{I}}$, if the support $\mathcal{O}_r$ of $(\psi_r)$ is contained in $\mathcal{B}_{R_r,\omega_r}$ for $r \in \mathcal{I}$, and $(\mathcal{B}_{R_r,\omega_r})_{r \in \mathcal{I}}$ is an admissible covering in the sense of [15], i.e., $\bigcup_{r \in \mathcal{I}} \mathcal{B}_{R_r,\omega_r} = \mathbb{R}^d$ and the number of overlapping $\mathcal{B}_{R_r,\omega_r}$ is bounded above (admissibility condition). In other words, with

$$r^* := \{ s : s \in \mathcal{I}, \mathcal{B}_{R_r,\omega_r} \cap \mathcal{B}_{R_s,\omega_s} \neq \emptyset \},$$

for all $r \in \mathcal{I}$ there exists $n_0 \in \mathbb{N}$, called height of the BAPU, such that $|r^*| \leq n_0$.

For technical reasons, which do not eliminate any interesting example, we assume even more: for all $\rho < \infty$ the family $(\mathcal{B}_{\mathcal{R}_r,\omega_r+\rho})_{r \in \mathcal{I}}$ should be an admissible covering of $\mathbb{R}^d$. More precisely, we assume throughout this paper that for each $\rho > 0$ there exists $n_0 = n_0(\rho) \in \mathbb{N}$ such that the number of overlapping balls constituting the covering is uniformly controlled: $|r^*| \leq n_0$ for all $r \in \mathcal{I}$, where

$$r^* := \{ s : s \in \mathcal{I}, \mathcal{B}_{\mathcal{R}_r,\omega_r+\rho} \cap \mathcal{B}_{\mathcal{R}_s,\omega_s+\rho} \neq \emptyset \}.$$

Obviously such coverings are of uniform height.

Using the concept of BAPUs, we now turn to **Wiener amalgam spaces**, introduced by H. Feichtinger in 1980 (see [18] for an accessible publication). The definition of Wiener amalgam spaces aims at decoupling local and global properties of $L^p$-spaces. Let a BAPU $(\psi_r)_{r \in \mathcal{I}}$ for $\mathbb{R}^d$ be given. The Wiener amalgam space $W(L^p, \ell^q)$ is defined as follows:

$$W(L^p, \ell^q)(\mathbb{R}^d) = \{ f \in L^p_{\text{loc}} : \| f \|_{W(L^p, \ell^q)} = \left( \sum_{r \in \mathcal{I}} \| f \cdot \psi_r \|_{L^p} \right)^{\frac{1}{q}} < \infty \}.$$ 

We will denote by $W(C^0, \ell^p)(\mathbb{R}^d) \subseteq W(L^\infty, \ell^p)(\mathbb{R}^d)$ the subspace of continuous, locally bounded functions in $W(L^\infty, \ell^p)(\mathbb{R}^d)$. A comprehensive review of (weighted) Wiener amalgam spaces can be found in [25]. We note that in their most general form they are described as $W(B, C)$, with local component $B$ and global component $C$. Let us recall some properties which will be needed later on:

- If $B_1 \subseteq B_2$, $C_1 \subseteq C_2$, then $W(B_1, C_1) \subseteq W(B_2, C_2)$.
- If $B_1 \ast B_2 \subseteq B_3$, $C_1 \ast C_2 \subseteq C_3$, then $W(B_1, C_1) \ast W(B_2, C_2) \subseteq W(B_3, C_3)$.

A particularly important Banach space in time-frequency analysis is the Wiener amalgam space $W(F^L, \ell^1)$. This space, also known under the name Feichtinger's algebra, is better known as the modulation space $M^1_\infty$, with constant weight $m = 1$. It is often denoted by $S_0$ in the literature and we will adopt this name in the present work. For convenience, we also recall the definition of $S_0$ via the short-time Fourier transform.

**Definition 1 (S_0).** Let $g_0$ be the Gauss-function $g_0 = e^{-\pi \|x\|^2}$. The space $S_0(\mathbb{R}^d)$ is given by

$$S_0(\mathbb{R}^d) = \{ f \in L^2(\mathbb{R}^d) : \| f \|_{S_0} = \| \mathcal{F} g_0 \hat{f} \|_{L^1(\mathbb{R}^{2d})} < \infty \}.$$ 

An in-depth investigation of $S_0(\mathbb{R}^d)$ and its outstanding role in time-frequency analysis can be found in [16]. Note that $S_0(\mathbb{R}^d)$ is densely embedded in $L^2(\mathbb{R}^d)$, with $\| f \|_2 \leq \| g_0 \|_2^{-1} \| f \|_{S_0}$. Its dual space $S'_0(\mathbb{R}^d)$, the space of all linear, continuous functionals on $S_0(\mathbb{R}^d)$, contains $L^2(\mathbb{R}^d)$ and is a very convenient space of (tempered) distributions. Moreover, in the definition of $S_0(\mathbb{R}^d)$, $g_0$ can be replaced by any $g \in S_0(\mathbb{R}^d)$, see [24, Theorem 11.3.7] and different functions $g \in S_0(\mathbb{R}^d) \setminus \{0\}$ define equivalent norms on $S_0(\mathbb{R}^d)$.

One of the results of major importance in Gabor analysis states that for $g \in S_0(\mathbb{R}^d)$ the analysis mapping $T_g$ is bounded from $L^2(\mathbb{R}^d)$ to $\ell^2$ for any lattice $\Lambda$, and $T_g^*$ is then bounded by duality, see [16, Section 3.3] for details. This will be of crucial significance in our arguments.
3. Quilted Gabor frames: the general concept

For the construction of quilted frames, we start from a collection of (Gabor) frames. Usually, these frames will feature various different qualities, e.g. varying resolution quality for time and frequency. Then, a partition in time-frequency is set up according to some application-dependent criterion and a particular frame is assigned to each member of the partition. For example, in [27], the selection of the local frames is based on time-frequency sparsity criteria. Fig. 1 gives an illustration of the basic idea, for a partition assigning one out of two different Gabor frames to each of the tiles of size $64 \times 64$, where the signal length is $L = 256$ and the number of tiles thus 16. The upper displays show the lattices corresponding to the two Gabor frames, the last display shows the “quilted” lattice $\chi_1 \cup \chi_2$ resulting from concatenation. Let us emphasize at this point, that at sampling points marked with different symbols, different windows are also used.

Note that, conceptually, irregular tilings may be used just as well. However, for practical as well as theoretical reasons, tilings with some kind of structure are more convenient.

It is important to point out that the partition in different domains corresponding to various different frames actually happens in the time-frequency domain. This implies that a priori we have no knowledge about the properties of the local families, as opposed to the concept of fusion frames, as discussed in [8,7]. In particular, we are not necessarily dealing with closed subspaces which may be transformed into each other as in the approach introduced in [22]. In a very recent contribution [30], Romero proved an existence result for quilted frames.

We now give a precise definition for quilted Gabor frames.

**Definition 2 (Quilted Gabor frames).** Let Gabor frames $\mathcal{G}(g^j, \Lambda^j)_{j \in J}$ for $L^2(\mathbb{R}^d)$ and an admissible covering $\Omega_r \subseteq (B_{R_r}(x_r))_{r \in I}$ of $\mathbb{R}^{2d}$ be given. Define the local index sets $\chi^r = \Omega_r \cap \Lambda^{m(r)}$, where
\[ m : \mathcal{I} \mapsto \mathcal{J} \] is a mapping assigning a frame from the given Gabor frames to each member of the covering. Then the set
\[ \bigcup_{r \in \mathcal{I}} \mathcal{G}(g^{m(r)}, \lambda^r) \] (4)
is called a quilted Gabor frame for \( L^2(\mathbb{R}^d) \), if there exist constants \( 0 < A, B < \infty \), such that
\[ A \| f \|_2^2 \leq \sum_{r \in \mathcal{I}} \sum_{\lambda \in \lambda^r} \| f \cdot \pi(\lambda) g^{m(r)} \|_2^2 \leq B \| f \|_2^2 \] (5)
holds for all \( f \in L^2(\mathbb{R}^d) \).

The general setting of quilted frames includes, of course, various special cases. We first give some trivial examples which may however be relevant in applications.

**Example 1.** For a given Gabor frame, we may choose additional sampling points in any selected region. This may be helpful, if in some applications, finer resolution is only desirable in certain parts of the time-frequency domain. Formally, this may be rephrased as follows. We are given Gabor frames \( \mathcal{G}(g, A^j)_{j \in \mathbb{N}} \) for \( L^2(\mathbb{R}^d) \) with \( A^0 \subseteq A^j \) for \( j \in \mathbb{N} \) and \( A_0 \) the lower frame bound for \( j = 0 \). Then, for an admissible covering, the local index sets are defined by \( \lambda^r = \Omega_r \cap \Lambda^{m(r)} \), where \( m : \mathcal{I} \mapsto \mathbb{N} \) is the mapping selecting the local systems. It is then trivial to see, that the resulting quilted Gabor frame has a lower frame bound \( A_0 \). The existence of an upper frame bound is covered by Theorem 1.

**Example 2.** For a given multi-window Gabor frame \( \mathcal{G}((g_1, \ldots, g_n), A^0) \), additional sampling points for selected windows may be added in certain parts of the time-frequency domain. For a formal description, assume that an admissible covering \( \Omega_r, r \in \mathcal{I} \), is given and let \( A^0 \subseteq A^j \) for \( j = 1, \ldots, N \) as in the previous example. The mapping \( m : \mathcal{I} \rightarrow \{0, \ldots, N\}^n \) is given by \( m(r)(k) = 0, \ k = 1, \ldots, n \) whenever the original lattice is maintained for all windows in the support of \( \Omega_r \) and by \( m(r)(k) = j_0, \ j_0 \in \{1, \ldots, N\} \), if denser sampling corresponding to \( A^{j_0} \) is desired in \( \Omega_r \) for the window \( g_{j_0} \).

In the next section we will prove the Bessel property of quilted systems obtained in a rather general situation, allowing for a finite overlap between the local patches. In the construction of lower frame bounds, a certain overlap between adjacent patches is often necessary. The two subsequent sections then describe two situations, in which a lower frame bound for the resulting quilted Gabor frame can be constructed explicitly.

**4. The Bessel condition in the general case**

We prove the existence of an upper frame bound for quilted frames as defined in (4). Note that the Bessel property alone allows for interesting conclusions about operators associated with the respective sequence, compare [2]. We will deduce the Bessel property of quilted Gabor frames from a general statement on relatively separated sampling sets. This result generalizes a result given in [28] on the Bessel property of irregular time-frequency shifts of a single atom. We prove that an arbitrary function from a set of window functions satisfying a common decay condition may be chosen for every sampling point in a relatively separated sampling set to obtain a Bessel sequence.

We assume that different given Gabor systems are to be used in compact sets \( \Omega_r \) corresponding to the members of an admissible covering of \( \mathbb{R}^{2d} \). Under the assumption that the windows under consideration satisfy a common decay condition in time-frequency and that the set of lattices is compact, we claim that an upper frame bound, or Bessel bound, can be found. As before, \( g_0 \) denotes the Gaussian window.
Theorem 1. Let Gabor frames $G(g^j, \Lambda^j)_{j \in J}$ for $L^2(\mathbb{R}^d)$ and an admissible covering $\Omega_r \subseteq (B_{R_r}(x_r))_{r \in \mathcal{I}}$ of the signal domain be given. Assume further that

(i) $g^j \in H_{s,c}$ for all $j$, where

$$H_{s,c} = \left\{ g \in L^2(\mathbb{R}^d) : |\psi_{g^j}(z)| \leq C(1 + |z|^2)^{-\frac{s}{2}}, \quad s > 2d, \, C > 0 \right\}$$

(ii) The lattice constants $\alpha^j$, $\beta^j$ are chosen from a compact set in $\mathbb{R}^+ \times \mathbb{R}^+$, i.e., $\alpha^j \subseteq [\alpha_0, \alpha_1] \subset (0, \infty)$ and $\beta^j \subseteq [\beta_0, \beta_1] \subset (0, \infty)$.

(iii) The regions assigned to the different Gabor systems correspond to an admissible covering $\Omega_r$, $r \in \mathcal{I}$ with $\operatorname{supp}(\psi_r) \subseteq \Omega_r \subseteq B_{R_r}(x_r)$ for $r \in \mathcal{I}$.

Let $m : \mathcal{I} \mapsto \mathcal{J}$ be a mapping assigning a frame from $G(g^j, \Lambda^j)_{j \in J}$ to each member of the covering. Then for any $\delta < \infty$, the overall family given by

$$G_m^Q = \bigcup_{r \in \mathcal{I}} \left\{ \pi(\lambda)g^{m(r)} : \lambda \in \Lambda^r \subset \Lambda^r, \, \Lambda^r = \Lambda^m(r) \cap B_{R_r+\delta} \right\}$$

possesses an upper frame bound, i.e., is a Bessel sequence for $L^2(\mathbb{R}^d)$.

Note that the theorem states that in particular the local systems given by $\Lambda^r = \Omega_r \cap \Lambda^m(r)$ for all $r$ lead to a Bessel sequence. More generally, however, the local patches can uniformly be enlarged by a radius $\delta$.

We first prove a general statement on sampling of functions in certain Wiener amalgam spaces over relatively separate sampling sets.

Definition 3 (Relatively separated sets). A set $\mathcal{X} = \{z_i = (x_i, \xi_i), \quad i \in \mathcal{I}\}$ in $\mathbb{R}^{2d}$ is called (uniformly) $\gamma$-separated, if $\inf_{j,k \in \mathcal{I}, j \neq k} |z_j - z_k| = \gamma > 0$. A relatively separated set is a finite union of separated sets. We call $\mathcal{X}$ $(\gamma', R)$-relative separated if the number of separated sets is $R$.

Remark 1. It is easy to show that the concept of relative separation does not depend on the specific values of $\gamma$ and $R$. In other words, any $(\gamma', R)$-relative separated set is also a finite union of $\eta$-separated subsets. Of course one has to allow to compensate the smallness of $\eta$ by a larger number $R' = R'(\eta)$.

There is an equivalent point of view. A sequence is relatively separated in $\mathbb{R}^k$ if and only if for some fixed $s > 0$ the family $(B(x_k))_{k \geq 1}$ covers each point $x$ in $\mathbb{R}^k$ at most $h = h(s)$ times, uniformly with respect to $x$.

Lemma 1. Let $1 \leq p < \infty$. Given a pair $(\gamma', R)$ there exists a constant $C = C(\gamma', R)$ such that for all $(\gamma', R)$-separated sets $\mathcal{X}$ and all functions $F \in W(C^0, \ell^p)(\mathbb{R}^{2d})$ one has $F|\mathcal{X}$ in $\ell^p(\mathcal{X})$ and

$$\|F|\mathcal{X}\|_{\ell^p(\mathcal{X})} = \left( \sum_{x_i \in \mathcal{X}} |F(x_i)|^p \right)^{1/p} \leq C \|F\|_{W(C^0, \ell^p)}.$$  \hspace{1cm} (7)

Proof. Recall that $\|F\|_{W(C^0, \ell^p)} = \|a_{kn}\|_{\ell^p}$, where

$$a_{kn} = \operatorname{esssup}_{(x, \xi) \in Q} |F(x + k, \xi + n)| = \|F \cdot T_{(k,n)} \chi_Q\|_{\ell^\infty}.$$  

By assumption, $\mathcal{X}$ is the finite union of uniformly separated sets $\mathcal{X}_r, r = 1, \ldots, R$ and there exists $\gamma$, such that $\min|z_j - z_k| \geq \gamma > 0$ for any pair $z_j, z_k$ in $\mathcal{X}_r$, with $j \neq k$. Hence, for all $r = 1, \ldots, R$, we
have at most \((1 + \frac{1}{γ})^{2d}\) points \(z_i \in X^r\) in the box \((k, n) + Q, (k, n) \in \mathbb{Z}^d \times \mathbb{Z}^d\), such that the number of \(z_i \in X^r\) in this box is bounded above by \(R \cdot (1 + 1/γ)\). Clearly

\[
|F(x, ξ)| \leq \|F \cdot T_{(k, n)}X_Q\|_∞ \quad \text{for all } (x, ξ) \in (k, n) + Q.
\]

Altogether, this yields:

\[
\sum_{z_i \in X^r} |F(z_i)|^p \leq R \cdot \left(1 + \frac{1}{γ}\right)^{2d} \sum_{(k, n) \in \mathbb{Z}^d} \|F \cdot T_{(k, n)}X\|_{∞}^p = R \cdot \left(1 + \frac{1}{γ}\right)^{2d} \|F\|_{W(C^0, ℓ^p)}^p. \quad □
\]

**Remark 2.** Analogous statements are true in more general situations, especially for any weighted sequence space \(\ell^p_m\), see [20, Lemma 3.8] for example.

The condition, that \(F\) is continuous (locally in \(C^0\)), guarantees, that sampling is well defined. Of course, weaker conditions, for example, semi-continuity, are sufficient.

The upper frame bound estimate will follow from a pointwise estimate over the family of short-time Fourier transforms

\[
F(λ) = \sup_{j \in J} |V_{g^j} f(λ)|, \quad λ \in \mathbb{R}^{2d}.
\]

In the sequel, we denote by \(M\) any countable subset of \(H_{S,C}\). We will make use of the following lemma.

**Lemma 2.** Assume that \(g \in M \subseteq H_{S,C}\) for \(s \geq 2d, C > 0\). Then there exists some constant \(C > 0\) such that for all \(f \in L^2(\mathbb{R}^d)\) one has the following uniform estimate of \(V_{g^j} f\) in \(W(C^0, ℓ^2)\):

\[
\|\sup_{g \in M} |V_{g^j} f|\|_{W(C^0, ℓ^2)} \leq C_1 \|f\|_2, \quad \text{for all } f \in L^2.
\]

**Proof.** The crucial step is to invoke the convolution relation between different short-time Fourier transforms [24, Lemma 11.3.3]. For convenience in the application below let us denote the generic elements from \(M\) by \(g^j\) (instead of \(g_0\) in 11.3.3), and make the choice \(γ = g = g_0\), the normalized Gauss function \(g_0\). Then obviously \(⟨γ, g⟩ = \|g_0\|_2 = 1\) and, setting \(w_s(z) = (1 + \|z\|^2)^{-\frac{s}{2}}\), we have the following estimate

\[
|V_{g^j} f(λ)| \leq \left[|V_{g_0} f| * |V_{g^j} g_0|\right](λ).
\]

Since \(|V_{g^j} g_0(λ)| = |V_{g_0} g^j(−λ)| \leq w_s(λ)\), we may take the pointwise supremum over \(g^j \in M\) on the left side and arrive at

\[
\sup_{g^j \in M} |V_{g^j} f(λ)| \leq \left(|V_{g_0} f| * C w_s\right)(λ).
\]

Of course \(s > 2d\) implies that \(w_s \in W(C^0, ℓ^1)(\mathbb{R}^{2d})\). Using the general fact [24, Cor. 3.2.2] that \(\|V_{g_0} f\|_2 \leq C_2 \|f\|_2\) for any \(f \in L^2(\mathbb{R}^d)\) and applying the convolution relation \(L^2 \ast W(C^0, ℓ^1) \subseteq W(L^1, ℓ^2) \ast W(C^0, ℓ^1) \subseteq W(C^0, ℓ^2)\), together with the appropriate estimates, we arrive at the desired estimate. □
Remark 3. It is worthwhile to note that the above result is not just a simple compactness argument. As a matter of fact it is not difficult to construct compact sets \( M \subset S_0(\mathbb{R}^d) \) for which the above result is not valid. One may, for example, just take a null sequence of the form \((c_n T_{x_n} g)_{n \geq 1}\) for some \( g \in S_0(\mathbb{R}^d) \), and with \((c_n) \in C_0\) but \((c_n) \notin \ell^2\).

The next theorem states that for a relatively separated sampling set of time-frequency shifts we can construct a Bessel sequence by associating to each sampling point an element from \( H_{s,C} \).

**Theorem 2.** Let \( X = \{ z_i = (x_i, \xi_i), \ i \in I \} \) in \( \mathbb{R}^{2d} \) be a relatively separated set of sampling points in \( \mathbb{R}^{2d} \). Let \( m : X \to J \) be a mapping assigning a window \( g^{m(z_i)} \in M \) to each sampling point. Then the set

\[
\{ \pi(z_i) g^{m(z_i)}, \ i \in I \}
\]

(10)
is a Bessel sequence for \( L^2(\mathbb{R}^d) \).

**Proof.** We have to estimate the series

\[
\sum_{z_i \in X} \left| \langle f, \pi(z_i) g^{m(z_i)} \rangle \right|^2 = \sum_{z_i \in X} \left| \mathcal{V}_{g^{m(z_i)}} f(z_i) \right|^2 \leq \sum_{z_i \in X} \left| F(z_i) \right|^2,
\]

with \( F \) given by (9). Now, as shown in the previous lemma, \( F \) is in \( W(C^0, \ell^2) \). Hence, since \( X \) is a relatively separated set, Lemma 2 can be applied to obtain the following estimate for the Bessel bound of (10):

\[
\sum_{z_i \in X} \left| \langle f, \pi(z_i) g^{m(z_i)} \rangle \right|^2 \leq C^2 R(1 + 1/\gamma)^{2d} \| f \|_2^2. \quad \square
\]

**Lemma 3.** The union of points \( \{ \lambda = (x, \xi), \ \lambda \in \mathcal{X}^r \} \) in the discrete sets \( \mathcal{X}^r \) as chosen in Theorem 1 is relatively separated.

**Proof.** Each of the lattices \( \Lambda^r \) determining the Gabor frames used in the construction of \( G^Q_m \) is of course separated, even uniformly with respect to \( r \). The admissibility condition for \((\mathcal{O}_r)_{r \in I}\) allows only finite overlap between any pair of members in the covering, or equivalently that the family of balls of radius \( R_r \) centered at \( x_r \) forms a covering of (uniformly) bounded height. By assumption, increasing each of the balls \( B_{R_r}(x_r) \) by the finite radius \( \delta > 0 \), only the height of the covering will be increased, but not the property of (uniformly) finite height. In other words, the family of enlarged balls \( B_{R_r + \delta}(x_r) \) is still an admissible covering of \( \mathbb{R}^d \) and the union of the \( \mathcal{X}^r \) is a relatively separated set. \( \square \)

We conclude the proof of Theorem 1 by choosing \( M = \{ g^j, \ j \in J \} \) in the following corollary.

**Corollary 1.** An upper (Bessel) bound for \( G^Q_m \) as defined in (6) is given by \( C^2 n(\delta)(1 + 1/\gamma)^{2d} \).

Note that \( n(\delta) \) denotes the height of the covering, which depends on \( \delta \).

5. Reduced multi-window Gabor frames: windows with compact support or bandwidth

In this section, we show that in a specific situation, which is, however, of practical relevance, quilted Gabor frames may be constructed. In the present model, we only change the resolution in time (or frequency). This means, that the time-frequency domain is partitioned into stripes rather
than patches. Under the additional assumption that the analysis window has compact support (or bandwidth), we easily obtain a lower frame bound for the quilted system.

Assume that we are given Gabor frames \( G(g^j, \Lambda^j) \) for \( L^2(\mathbb{R}^d) \), \( j \in \mathcal{J} \), where all the windows \( g^j \) have compact support and \( \|g^j\|_{L^2} \leq C_\epsilon < \infty \forall j \). We now want to use each system for a certain time, i.e., in a restricted stripe in the time-frequency domain. The stripes are defined by means of a partition of unity: we assume that \( f = \sum_{r \in \mathcal{I}} \psi_r f \) with \( \psi_r \leq 1 \) and that the \( \psi_r \) have compact support in \([a_r, b_r] \). By means of a mapping \( m : \mathcal{I} \to \mathcal{J} \), we assign one particular frame to each of these stripes.

Now, subfamilies of the given Gabor frames may be constructed as follows. Assume, for simplicity, that \( m(0) = 0 \) and consider the task to represent \( \psi_0 f \) in terms of the given Gabor frame \( G(g^0, \Lambda^0) \):

\[
\psi_0 f = \psi_0 \left( \sum_{\lambda \in \Lambda^0} \langle f, \pi(\lambda) g^0 \rangle \pi(\lambda) g^0 \right)
= \sum_{\lambda \in \Lambda^0} \langle f, \pi(\lambda) g^0 \rangle \psi_0 \pi(\lambda) g^0.
\]

Now, there exist \( n_0^u \) and \( n_0^l \) such that for \( \lambda = (n\alpha_0, m\beta_0) \) with \( n \notin [n_0^l, n_0^u] \), we find that \( \psi_0 \pi(\lambda) g^0 \equiv 0 \), hence

\[
\psi_0 f = \sum_{\lambda \in \Lambda_0} \langle f, \pi(\lambda) g^0 \rangle \psi_0 \pi(\lambda) g^0,
\]

where \( \Lambda_0 = [n_0^l \cdot \alpha_0, n_0^u \cdot \alpha_0] \times \beta_0 \mathbb{Z} \) is the subset of \( \Lambda^0 \) corresponding to the nonzero contributions.

Analogously subsets \( \Lambda_r \subset \Lambda^{m(r)} \) are chosen for all \( r \), and we obtain:

\[
f = \sum_{r \in \mathcal{I}} \psi_r f = \sum_{r} \sum_{\lambda \in \Lambda_r} \langle f, \pi(\lambda) g^{m(r)} \rangle \psi_r \pi(\lambda) g^{m(r)}.
\]  

(11)

With this construction, we state the following proposition.

**Proposition 1.** For a family of tight Gabor frames \( G(g^j, \Lambda^j) \), \( j \in \mathcal{J} \), for \( L^2(\mathbb{R}^d) \) let \( \sup_{j \in \mathcal{J}} \|g^j\|_{L^2} = C_g < \infty \) and \( C_{\Lambda^j} = (\frac{1}{a_j} + 1) d^d (\frac{1}{b_j} + 1) d^d \leq C_{\Lambda} < \infty \) for all \( j \in \mathcal{J} \). Let a partition of unity \( (\psi_r)_{r \in \mathcal{I}} \) of compactly supported \( \psi_r \) with height \( n_0 \) be given and let a mapping \( m : \mathcal{I} \to \mathcal{J} \) assign a frame \( G(g^{m(r)}, \Lambda^{m(r)}) \) to each \( r \in \mathcal{I} \). Assume that index sets \( \Lambda_r = [n_0^l \cdot \alpha_{r,0}, n_0^u \cdot \alpha_{r,0}] \times \beta_{m(r)} \mathbb{Z} \) are chosen such that for all \( r \in \mathcal{I} \) and \( \lambda = (n\alpha_{m(r)}, m\beta_{m(r)}) \) with \( n \notin [n_0^l \cdot \alpha_{m(r)}, n_0^u \cdot \alpha_{m(r)}] \), we have that \( \psi_r \pi(\lambda) g^{m(r)} \equiv 0 \).

Then, the union of the subfamilies \( \bigcup_{r \in \mathcal{I}} (g^{m(r)}, \Lambda_r) \) is a frame for \( L^2(\mathbb{R}^d) \) with a lower frame bound given by \( 1/(n_0 C_{\Lambda} C_g^2) \).

**Proof.** First note that

\[
\|\psi_r h\|_2^2 \leq \|h\|_2^2 \quad \text{for all } h \in L^2(\mathbb{R}^d).
\]  

(12)

Now set \( h_r = \sum_{\lambda \in \Lambda_r} \langle f, \pi(\lambda) g^{m(r)} \rangle \pi(\lambda) g^{m(r)} \) and thus, with (12):

\[
\|f\|_2^2 \leq n_0 \sum_r \|\psi_r f\|_2^2 = n_0 \sum_r \left\| \sum_{\lambda \in \Lambda_r} \langle f, \pi(\lambda) g^{m(r)} \rangle \psi_r \pi(\lambda) g^{m(r)} \right\|_2^2
= n_0 \sum_r \|h_r \psi_r\|_2^2 \leq n_0 \sum_r \left\| \sum_{\lambda \in \Lambda_r} \langle f, \pi(\lambda) g^{m(r)} \rangle \pi(\lambda) g^{m(r)} \right\|_2^2
\]
\[ n_0 \sum_r (\frac{1}{\alpha_j} + 1)^d \left( \frac{1}{\beta_j} + 1 \right)^d (1^{\alpha_j} + 1) d (1^{\beta_j} + 1) \sum_{\lambda \in \mathcal{X}'} \| |f, \pi(\lambda) g^{m(r)}| |^2 \]
\[ \leq n_0 C \sum_r \sum_{\lambda \in \mathcal{X}'} \| |f, \pi(\lambda) g^{m(r)}| |^2 . \]

The last inequality is due to the boundedness of the frame-synthesis operator \( T^{*}_{g} : \ell^2(\Lambda^j) \rightarrow L^2(\mathbb{R}^d) \), whenever the window \( g^{m(r)} \) is in \( S_0(\mathbb{R}^d) \), see [24, Proposition 6.2.2]. This proves the existence of a lower frame bound as stated. The existence of an upper frame bound can be seen directly from the construction of the subfamilies, and is furthermore covered by the general case proved in Section 4. \( \square \)

**Remark 4.**

1. An analogous statement holds for general, not necessarily tight frames, for, if \( \gamma^j \) are the dual windows for each \( g^j \), then \( \| S^{-1}_j g^j \| \leq \frac{1}{\lambda_j} \| g^j \| . \)
2. The same construction may be realized in the Fourier transform domain by applying a partition of unity to \( \hat{f} \). This corresponds to the usage of different Gabor frames in different stripes of the frequency domain and hence resembles a non-orthogonal filter bank. As a particular example, a constant-Q transform may be realized [6].
3. Note that a similar yet more restrictive construction, corresponding to the classical “painless non-orthogonal expansions” was suggested in [26,3]. Very recently, another related and highly interesting construction has been suggested in [32].

6. Replacing a finite number of frame elements

We now consider the task of replacing a finite number of atoms from a given (Gabor) frame by a finite number of atoms from a different (Gabor) frame. The following theorem gives a condition valid for general frames, which will then be applied to Gabor frames. Recall, that \( T_g \) and \( T^*_{g} \) denote the analysis and synthesis mapping, respectively, for given \( g \) and \( \Lambda \). In this section, we use the notation \( T_g, \Lambda \) and \( T^*_{g, \Lambda} \) for the respective mappings corresponding to subsets of the given lattices. For example, let \( \mathcal{F}_1 \subset \Lambda \) be a finite subset of a lattice \( \Lambda \), then \( T_g, \mathcal{F}_1(f) = \langle f, \pi(\lambda) g \rangle \) for \( \lambda \in \mathcal{F}_1 \). The theorem makes use of a linear mapping \( L \) describing the replacement procedure in the coefficient domain. As long as elements from frame \( \mathcal{G}_1 \) may be replaced by elements from \( \mathcal{G}_2 \) in a controlled manner, i.e., without loosing energy, a quilted frame can be obtained.

**Theorem 3.** Assume that two frames \( \mathcal{G}_1 = \{ g_i, \ i \in \mathcal{I} \} \) and \( \mathcal{G}_2 = \{ h_j, \ j \in \mathcal{J} \} \) for \( L^2(\mathbb{R}^d) \) are given and a finite number of elements \( \{ g_i, \ i \in \mathcal{F}_1 \subset \mathcal{I} \} \) of \( \mathcal{G}_1 \) are to be replaced by a finite number of elements \( \{ h_j, \ j \in \mathcal{F}_2 \subset \mathcal{J} \} \) of \( \mathcal{G}_2 \).

Let \( \Lambda_1 \) be the lower frame bound of \( \mathcal{G}_1 \). If a bounded linear mapping \( L : \ell^2(\mathcal{F}_1) \mapsto \ell^2(\mathcal{F}_2) \) can be found such that

\[ \| T^*_{g, \mathcal{F}_1} - T^*_{h, \mathcal{F}_2} L \|_2^2 = C < \frac{\Lambda_1}{2} , \]  

then the set

\[ \{ g_i, \ i \in \mathcal{I} \setminus \mathcal{F}_1 \} \cup \{ h_j, \ j \in \mathcal{F}_2 \} \]

is a frame for \( L^2(\mathbb{R}^d) \) with a lower frame bound given by \( (\Lambda_1 - 2C)/\max(1, 2\|L^*\|_2^2) \).
Proof. First note that
\[ A_1 \| f \|_2^2 \leq \sum_{i \in I \setminus \mathcal{F}_1} |\langle f, g_i \rangle|^2 + \sum_{i \in \mathcal{F}_1} |\langle f, g_i \rangle|^2 = \| T_{g, \mathcal{F}_1} \|^2_2 + \| T_{g, \mathcal{F}_1} \|^2_2. \]

Now, we have
\[ \| T_{g, \mathcal{F}_1} \|^2_2 \leq \left( \left\| \left( T_{g, \mathcal{F}_1} - L^* T_{h, \mathcal{F}_2} \right) f \right\| + \left\| L^* T_{h, \mathcal{F}_2} f \right\| \right)^2 \]
\[ \leq 2 \cdot \left\| \left( T_{g, \mathcal{F}_1} - L^* T_{h, \mathcal{F}_2} \right) f \right\|^2_2 + 2 \cdot \left\| L^* T_{h, \mathcal{F}_2} f \right\|^2_2 \]
\[ \leq 2 \cdot C \cdot \| f \|^2_2 + 2 \cdot \| L^* T_{h, \mathcal{F}_2} f \|^2_2, \]

hence
\[ (A_1 - 2C) \| f \|^2_2 \leq \| T_{\mathcal{F}_1} f \|^2_2 + 2 \cdot \| L^* \|^2_2 \| T_{h, \mathcal{F}_2} f \|^2_2 \]
\[ \leq \max(1, 2) \| L^* \|^2_2 \left( \sum_{i \in I \setminus \mathcal{F}_1} |\langle f, g_i \rangle|^2 + \sum_{j \in \mathcal{F}_2} |\langle f, h_j \rangle|^2 \right) \]
and hence \((A_1 - 2C)/\max(1, 2)\|L^*\|^2_2\) is a lower frame bound for the system given in (14). The existence of an upper frame bound is trivial. \(\square\)

Note that the above theorem only states the existence of a lower frame bound under the given conditions, while this frame bound will usually not be optimal.

We now turn to the special case of Gabor frames.

Corollary 2. Assume that tight Gabor frames \(G(g, \Lambda^1)\) and \(G(h, \Lambda^2)\) with \(g, h \in S_0(\mathbb{R}^d)\) are given. Assume further that in a compact region \(\Omega \subset \mathbb{R}^{2d}\) the time-frequency shifted atoms \(\pi(\lambda)g, \lambda \in \mathcal{F}_1 = \Omega \cap \Lambda^1\) are to be replaced by a finite set of time-frequency shifted atoms \(\pi(\mu)h, \mu \in \mathcal{F}_2 \subset \Lambda^2\).

(a) If \(\Lambda^1 = \Lambda^2\), then
\[ G = \left\{ \pi(\lambda)g: \lambda \in \Lambda \setminus \mathcal{F}_1 \right\} \cup \left\{ \pi(\mu)h: \mu \in \mathcal{F}_2 \right\} \]
(15)
is a (quilted Gabor) frame for \(L^2(\mathbb{R}^d)\), whenever
\[ \|h - g\|^2_{S_0} = C < \frac{A_1}{2C_A}, \]
where \(C_A\) is a constant only depending on the lattice \(\Lambda = \Lambda^1 = \Lambda^2\).

(b) For general \(\Lambda^2\), a compact set \(\Omega^* \) in \(\mathbb{R}^{2d}\) can be chosen such that for \(\mathcal{F}_2 = \Omega^* \cap \Lambda^2\), the union
\[ G = \left\{ \pi(\lambda)g: \lambda \in \Lambda^1 \setminus \mathcal{F}_1 \right\} \cup \left\{ \pi(\mu)h: \mu \in \mathcal{F}_2 \right\} \]
(16)
is a (quilted Gabor) frame.

Proof. Statement (a) can easily be seen by choosing \(L = \text{Id}\) in Theorem 3:

\[ \]
\[ \| T^*_{\mathcal{F}_1} - T^*_{\mathcal{F}_2} \|_{\ell^2 \to L^2(\mathbb{R}^d)}^2 = \sup_{\|c\|_2 = 1} \left\| \sum_{\lambda \in \mathcal{F}_1} c_\lambda (\pi(\lambda)g - \pi(\lambda)h) \right\|_{L^2}^2 \]
\[ \leq C_A \cdot \|h - g\|_{S_0}^2, \]

where the last inequality follows from boundedness of the synthesis operator for windows in \( S_0 \), [24, Theorem 12.2.4].

To show (b), we first introduce the mapping \( L : \ell^2(\mathcal{F}_1) \to \ell^2(\mathcal{F}_2) \) as follows. For a finite sequence \( c = (c_\lambda)_{\lambda \in \mathcal{F}_1} \) we define \( L(c)(\mu) = (\sum_{\lambda \in \mathcal{F}_1} c_\lambda (\pi(\lambda)g, \pi(\mu)h))_{\mu \in \mathcal{F}_2} \), for which

\[ \| L \|_{\ell^2(\mathcal{F}_1) \to \ell^2(\mathcal{F}_2)}^2 = \sup_{\|c\|_2 = 1} \left\| \sum_{\lambda \in \mathcal{F}_1} c_\lambda (\pi(\lambda)g, \pi(\mu)h) \right\|_{L^2}^2 \]
\[ = \| T_{h,\mathcal{F}_2} T^*_{g,\mathcal{F}_1} \|_{L^2}^2 \leq C_A \cdot C_A^2 \cdot \|g\|_{S_0}^2 \|h\|_{S_0}^2 \]

due to the boundedness of synthesis and analysis operator, \( T_{h,\mathcal{F}_2} \) and \( T^*_{g,\mathcal{F}_1} \), respectively.

Let now \( A_1 \) be the lower frame bound of \( G(g, A^1) \) and let \( G(h, A^2) \) have lower frame bound\(^2 \) \( A_2 = 1 \) for simplicity and without restriction of generality. We then have:

\[ \| T^*_{\mathcal{F}_1} - T^*_{\mathcal{F}_2} L \|_{\ell^2 \to L^2(\mathbb{R}^d)}^2 = \sup_{\|c\|_2 = 1} \left\| \sum_{\lambda \in \mathcal{F}_1} c_\lambda (\pi(\lambda)g - \sum_{\mu \in \mathcal{F}_2} \sum_{\lambda \in \mathcal{F}_1} c_\lambda \pi(\lambda)g, \pi(\mu)h) \pi(\mu)h \right\|_{L^2}^2 \]
\[ = \sup_{\|c\|_2 = 1} \left\| \sum_{\mu \in \mathcal{F}_2} \sum_{\lambda \in \mathcal{F}_1} c_\lambda (\pi(\lambda)g, \pi(\mu)h) \pi(\mu)h \right\|_{L^2}^2 \]
\[ \leq \|h\|_2 \sup_{\|c\|_2 = 1} \left\| \sum_{\mu \in \mathcal{F}_2} \sum_{\lambda \in \mathcal{F}_1} |c_\lambda| \cdot |\nu_{\mathcal{F}_1}(\mu - \lambda)| \right\| \]
\[ \leq \|h\|_2 \sqrt{|\mathcal{F}_1|} \max_{\mu \in \mathcal{F}_2} \left\| \nu_{\mathcal{F}_1}(\mu - \lambda) \right\| \]

where \( \mathcal{F}_2 = \Lambda^2 \setminus \mathcal{F}_2 \). Now an appropriate set \( \mathcal{F}_2 \) may be constructed as follows:

Let \( G = |\nu_{\mathcal{F}_1}h| \in W(C^0, \ell^1) \). We can choose a compact set \( \Omega^* \subset \mathbb{R}^d \), such that

\[ \left\| \max_{\lambda \in \mathcal{F}_1} (T_{\lambda} G - (T_{\lambda} G) \cdot \chi_{\Omega^*}) \right\|_{W(C^0, \ell^1)} < \tilde{e} = \frac{\sqrt{A_1}}{\sqrt{2|\mathcal{F}_1|^2} \|h\|_2 C_{A^2}}, \]

where \( \chi_{\Omega^*} \) is the indicator function of the set \( \Omega \). Now set

\[ \mathcal{F}_2 = \Lambda^2 \cap \Omega^*, \]

then

\[ \sum_{\mu \in \mathcal{F}_2} \max_{\lambda \in \mathcal{F}_1} T_{\lambda} G(\mu) = \left\| \max_{\lambda \in \mathcal{F}_1} T_{\lambda} G \right\|_{\ell^1}, \]

\(^2 \) This case, i.e., tight frames with \( A = B = 1 \), is also called Parseval frame.
\[ \| \Lambda^2 \| |F^1_0| \| h \| \leq C \Lambda^2 \cdot \tilde{\varepsilon} = C \Lambda^2 \cdot \frac{\sqrt{A_1}}{\sqrt{2 |F^1_0||h|}}. \] (17)

where we used Lemma 1 in (17). Hence, \( F^2 \) can be chosen, such that

\[ \| T^{*}_{\Lambda^0_0} - T^{*}_{\bigcup_{r} \Lambda^r} L \|_{L^2(\mathbb{R}^d)} < A_1/2. \]

Of course, the existence of an upper frame bound is trivial and the resulting system (16) is a quilted Gabor frame according to Theorem 3.

Remark 5. (1) The size of the region in which atoms from \( G_{1} \) have to be replaced, influences the choice of \( \Omega^* \). In particular, \( \tilde{\varepsilon} \) is reciprocally related to \( \sqrt{|F^1_0|} \), i.e., \( \Omega^* \) grows in dependence on the perimeter rather than the area of \( \Omega \).

(2) For statement (a) in Corollary 2, if \( g, h \in L^2(\mathbb{R}^d) \), then the frame property follows whenever

\[ \| h - g \|_2^2 = C < A_1/(2 |F^1_0|). \]

Note that the size of \( \Omega \) determines the necessary similarity of the windows, whereas, for windows in \( S_0(\mathbb{R}^d) \), the good localization implied by \( S_0 \)-membership allows for a stronger statement.

(3) The construction in Corollary 2 implies explicit dependence on time and frequency of the resulting quilted frame. Similarly, Gabor atoms in finitely many compact areas can be replaced by different Gabor systems. Details on this procedure will be reported elsewhere. From an application point of view, this process corresponds to finding optimal representations for local signal components, e.g., in the sense of sparsity.

(4) While the assumption of tightness of the original Gabor frames in Corollary 2 is made for convenience and may be replaced by exploiting results on the localization of the dual windows [23], it seems unlikely that tightness can be achieved for the resulting quilted frame. However, tight frames may be obtained from quilted frames by putting weights on certain windows in the transition regions between adjacent members of the partition of unity, compare the remarks on preconditioning in Section 7. Tight frames are rather easily constructed, if adaptability is only desired in time or frequency, compare [3]. The question of existence of general tight quilted frames seems more involved and deserves further study.

(5) In Example 1, the linear mapping \( L : A^0 \rightarrow \bigcup_j \Lambda^r \) can be chosen as follows:

\[ L(c)(\lambda) = \begin{cases} c_{\lambda}, & \text{for } \lambda \in A^0, \\ 0, & \text{for } \lambda \in \bigcup_j \Lambda^r \backslash A^0. \end{cases} \]

Then

\[ \| T^*_{A^0_0,g} - T^*_{\bigcup_j \Lambda^r} L \|_{L^2(\mathbb{R}^d)} = \sup_{\| c \|_2 = 1} \left| \sum_{\lambda \in A^0} c_{\lambda} \pi(\lambda) g - \sum_{\lambda \in \bigcup_j \Lambda^r} L(c)(\lambda) \pi(\lambda) g \right|_2^2 = 0. \]

As underlined by Example 1 and Corollary 2(a), in constructing quilted Gabor frames, technically difficult situations mainly arise if the lattices and the windows vary.
7. Reconstruction and simulations

Since the frame property has been proved for systems as described in Proposition 1 and Corollary 2, reconstruction can always be performed by means of a dual frame. However, since quilted Gabor frames possess a strong local structure, alternative and numerically cheaper methods may be preferable as long as sufficient precision in the reconstruction may be guaranteed. The next two sections present numerical results for various approaches to reconstruction for which the calculation of an exact dual frame is not necessary.

7.1. Reduced multi-window Gabor frames

We first consider reduced multi-window Gabor frames. Here, Eq. (11) yields an immediate reconstruction formula by means of projections onto the members of the partition of unity. However, we are more interested in the generic reconstruction by means of dual frames. We may compare the dual frame corresponding to the quilted Gabor frame $\bigcup_{j \in I} (g_j, X_j)$ with the quilted system $\bigcup_{j \in I} (r_j, X_j)$ resulting from using the dual windows $r_j$ of the original frames $G_j$. Alternatively, we may start with tight Gabor frames.

While this approach does not result in perfect reconstruction in one step, we apply the frame algorithm (see [24, Chapter 5]) to obtain near-perfect reconstruction in a few iteration steps. In this context, the condition number of the operators involved plays an important role and depends on the amount of overlap that we introduce in the design of the system. In the following example, it turns out that, while no essential overlap is necessary to obtain a frame in the finite discrete case, the overlap as requested in the proof of Proposition 1 leads to faster convergence of the frame algorithm.

Example 3. We consider two tight Gabor frames for $\mathbb{C}^L$ and $L = 144$ and two cases of different redundancy. First, redundancy is 4.5, corresponding to the lattices $\Lambda^1$ with $a = 4$, $b = 8$ and $\Lambda^2$ with $a = 8$, $b = 4$. Second, we consider two frames with redundancy 1.125, corresponding to the lattices $\Lambda^1$ with $a = 8$, $b = 16$ and $\Lambda_2$ with $a = 16$, $b = 8$. The corresponding tight windows are shown in Fig. 2. We next generate a quilted Gabor system without overlap and a corresponding Gabor system with overlap, for both cases of redundancy. Note that in each lattice point as depicted in Fig. 3, the tight window of the original Gabor frame is used. We now look at the condition numbers of the resulting quilted systems, listed in Table 1. It is obvious, that higher redundancy leads to more stability in the process of quilting frames. On the other hand, for the system with low redundancy, overlap becomes essential in order to obtain acceptable condition numbers. These observations are consistently confirmed by more extensive numerical experiments.
In the next step, we now compare the convergence of the (iterative) frame algorithm for the 4 cases considered in this example. Table 2 gives the number of iterations necessary to attain the error threshold of $10^{-8}$ in the reconstruction of a random signal, i.e., the reconstruction is considered accurate, as soon as the relative error is below this threshold. Fig. 4 then shows the rate of convergence for the three cases with acceptable condition numbers.

We finally discuss the following, “preconditioned” version of reconstruction for the case of low redundancy with overlap. We wish to reconstruct a random signal $r \in C^L$ from its quilted Gabor coefficient. As a first guess, instead of calculating the dual frame of the quilted frame, we simply use the quilted tight frame for reconstruction: Let $T_{G_q}$ denote the analysis operator corresponding
to the quilted tight frame $G^t_q = \bigcup_{j=1}^2 (h_j, X_j)$, where $h_j$ are the tight windows. We then obtain a reconstruction $\text{rec}$ by

$$\text{rec} = T^{*}_{G^t_q} \cdot T_{G^t_q} \cdot r.$$  

Obviously, the result is not accurate and in particular in the regions of transition between the two systems, errors occur. However, we can correct a considerable amount of the deviation from the identity by simply pre-multiplying the frame operator by the inverse of its diagonal. Hence, let $D$ be the diagonal matrix with $D(n, n) = T^{*}_{G^t_q} \cdot T_{G^t_q}(n, n)$, $n = 1, \ldots, 144$.

$$\text{rec} = T^{*}_{G^t_q} \cdot T_{G^t_q} \cdot D^{-1} \cdot r.$$  

The respective results are shown in Fig. 5. The relative error, defined by $\varepsilon = \frac{||\text{rec} - r||}{||r||}$ is then 0.2239 for the uncorrected case and 0.032 for the corrected version.

### 7.2 Replacing a finite number of elements

In our next example we consider a situation similar to the one discussed in Example 3, however, this time we wish to replace elements from $G_1$ in a bounded, quadratic region of the time-frequency plane.

**Example 4.** We consider the same Gabor frames as in Example 3, and look at the high redundancy systems first. As before, we compare the condition number of the system obtained with overlap to the less redundant situation. The two situations are shown in Fig. 6. The quilted Gabor frame without overlap has condition number 1.5, while allowing for some overlap, as shown in the second display of Fig. 6 leads to condition number 1.4. Accordingly, 18 and 17 iterations are necessary for convergence of the frame operator.

We now turn to the systems with low redundancy. Here we compare three amounts of overlap as shown in the upper displays of Fig. 7. The condition numbers of the resulting systems and the convergence behavior of the corresponding frame algorithm are shown in the lower display of the
same figure. Again, it becomes obvious that for low-redundancy systems, overlap is essential in order to obtain fast convergence in iterative reconstruction. On the other hand, increasing overlap beyond a certain amount, does not dramatically improve the condition numbers.

8. Summary and outlook

We have shown the existence of a lower frame bound for two particular instances of quilted Gabor frames. Furthermore, an upper frame bound has been constructed for the general setting. We showed how to reconstruct signals from the coefficients obtained with quilted Gabor frames and numerical simulations have been provided.
Future work will mainly include the construction of lower frame bounds for more general situations. In particular, Proposition 1 will be generalized to Gabor frames with non-compactly supported windows. Furthermore, numerical simulations suggest that atoms from a given Gabor frame may be replaced by atoms from a different frame in infinitely many compact regions of the time-frequency plane under certain conditions. On the other hand, for practical applications algorithms applicable for long signals (number of sampling points $\gg 44100$) have to be developed. The results of processing with quilted Gabor frames will be assessed on the basis of real-life data. Preconditioning similar to the procedure suggested in Example 3 can be developed for the more complex situations of quilted frames, compare [4].

Acknowledgments

The author wishes to thank Hans Feichtinger for the joint development of the notion of quilted frames as well as innumerable discussions on the topic and his invaluable comments on the content of this article. She also wishes to thank Franz Luef for his comments on an earlier version of the paper, Patrick Wolfe for fruitful scientific exchange on the topic of adaptive frames from the point of view of applications and the anonymous reviewers for their helpful comments.

References


Abstract. We construct frames adapted to a given cover of the time-frequency or time-scale plane. The main feature is that we allow for quite general and possibly irregular covers. The frame members are obtained by maximizing their concentration in the respective regions of phase-space. We present applications in time-frequency, wavelet and Gabor analysis.

1. Introduction

A time-frequency representation of a distribution \( f \in \mathcal{S}'(\mathbb{R}^d) \) is a function defined on \( \mathbb{R}^d \times \mathbb{R}^d \) whose value at \( z = (x, \xi) \) represents the influence of the frequency \( \xi \) near \( x \). The short-time Fourier transform (STFT) is a standard choice for such a representation, popular in analysis and signal processing. It is defined, by means of an adequate smooth and fast-decaying window function \( \varphi \in \mathcal{S}(\mathbb{R}^d) \), as

\[
V_{\varphi} f(z) = \int_{\mathbb{R}^d} f(t) \varphi(t-x)e^{-2\pi i \xi t} dt, \quad z = (x, \xi) \in \mathbb{R}^d \times \mathbb{R}^d.
\]

The distribution \( f \) can be re-synthesized from its time-frequency content by,

\[
f(t) = \int_{\mathbb{R}^d \times \mathbb{R}^d} V_{\varphi} f(x, \xi) \varphi(t-x)e^{2\pi i \xi t} dx d\xi.
\]

This representation is extremely redundant. One of the aims of time-frequency analysis is to provide a representation of an arbitrary signal as a linear combination of elementary time-frequency atoms, which form a less redundant dictionary. The standard choice is to let these atoms be time-frequency shifts of a single window function \( \varphi \), thus providing a uniform partition of the time-frequency plane. The resulting systems of atoms are known as Gabor frames. However, in certain applications atomic decompositions adapted to a less regular pattern may be required (see for example [29, 4, 8, 13, 42]).

For example, a time-frequency partition may be derived from perceptual considerations. For audio signals, this means that low frequency bins are given a finer resolution than bins in high regions, where better time-resolution is usually desirable, cp. [44, 48]. Such a partition is schematically depicted in the left plot of Figure 1.

More irregular partitions may be desirable whenever the frequency characteristics of an analyzed signal change over time and requires adaptation in both time and frequency. For example, adaptive partitions obtained from information theoretic criteria were suggested in [36, 38]. In such a situation, partitions as irregular as shown in the right plot of Figure 1 can be appropriate.

In this article we consider the following problem. Given a - possibly irregular - cover of the time-frequency plane \( \mathbb{R}^{2d} \), we wish to construct a frame for \( L^2(\mathbb{R}^d) \) with atoms whose time-frequency concentration follows the shape of the cover members. This allows to vary the trade-off between time and frequency resolution along the time-frequency plane. The adapted frames are constructed by selecting, for each member of a given cover, a family of functions maximizing their concentration in the corresponding region of the time-frequency domain, or phase-space. These functions can be obtained as eigenfunctions of the so-called time-frequency localization operators.

Date: July 23, 2012.

2010 Mathematics Subject Classification. 42C15, 42C40, 41A30, 41A58, 40H05, 47L15.

Key words and phrases. Phase-space, localization operator, frame, short-time Fourier transform, time-frequency analysis, time-scale analysis.

Monika Dörfler was supported by the Austrian Science Fund (FWF): [T384-N13] Locatif and by the WWTF project Audiominer (MA09-24).

José Luis Romero was supported by the Austrian Science Fund (FWF): [P22746-N13]. He also gratefully acknowledges support from the Austrian Science Fund (FWF): [T384-N13].
Given a compact set $\Omega \subseteq \mathbb{R}^d$ in the time-frequency plane, the *time-frequency localization operator* $H_{\Omega}$ is defined by masking the coefficients in (1), cf. [14, 15], i.e.

$$H_{\Omega} f(t) = \int_{\Omega} V_{\varphi} f(x, \xi) \varphi(t - x) e^{2\pi i \xi t} \, dx \, dw.$$  

(2)

$H_{\Omega}$ is self-adjoint and trace-class, so we can consider its spectral decomposition

$$H_{\Omega} f = \sum_{k=1}^{\infty} \lambda_k(f, \phi_k^{\Omega}) \phi_k^{\Omega}.$$  

The first eigenfunction, $\phi_1^{\Omega}$, is optimally concentrated inside $\Omega$ in the following sense,

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

More generally, the first $N$ eigenfunctions of $H_{\Omega}$ form the orthonormal set in $L^2(\mathbb{R}^d)$ that maximizes the quantity $\sum_{j=1}^{N} \int_{\Omega} |V_{\varphi} \phi_j^{\Omega}(z)|^2 \, dz$ among all orthonormal sets of $N$ functions in $L^2(\mathbb{R}^d)$. In this sense, their time-frequency profile is optimally adapted to $\Omega$. Figure 2 illustrates this principle by showing some time-frequency boxes $\Omega$ along with the STFT and real part of the corresponding localization operator’s first eigenfunctions.

In order to construct a frame adapted to a given cover $\{\Omega_{\gamma} : \gamma \in \Gamma\}$ of $\mathbb{R}^d$, we select, for each region $\Omega_{\gamma}$, the first $N_{\gamma}$ eigenfunctions $\psi_1^{\Omega_{\gamma}}, \ldots, \psi_{N_{\gamma}}^{\Omega_{\gamma}}$ of the operator $H_{\Omega_{\gamma}}$. We will prove that, for $N_{\gamma} \approx |\Omega_{\gamma}|$, we obtain a collection of atoms that covers the whole time-frequency plane. Note that this choice of $N_{\gamma}$ agrees with the uncertainty principle which roughly says that for each time-frequency region $\Omega_{\gamma}$ there are only $\approx |\Omega_{\gamma}|$ degrees of freedom. We allow for covers that are arbitrary in shape as long as they satisfy a mild admissibility condition,

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

Under these conditions we prove the following.

**Theorem 1.** Let $\{\Omega_{\gamma} : \gamma \in \Gamma\}$ be an admissible cover of $\mathbb{R}^d$. Then, there exists a constant $C > 0$ such that for every choice of $N_{\gamma}$, $C |\Omega_{\gamma}| \leq N_{\gamma} \leq N < \infty$, the family of functions $\left\{ \phi_k^{\Omega_{\gamma}} : \gamma \in \Gamma, 1 \leq k \leq N_{\gamma} \right\}$ is a frame of $L^2(\mathbb{R}^d)$. That is, for some constants $0 < A \leq B < +\infty$, the following frame inequality
Figure 2. Four different rectangular masks in time-frequency domain and the first eigenfunctions of the corresponding localization operators. Middle plots show the absolute value squared of the STFT and right plots show the real part.

holds,

$$A \|f\|_2^2 \leq \sum_{\gamma} \sum_{k=1}^{N_\gamma} \left| \langle f, \phi_{\Omega_k}^{\gamma} \rangle \right|^2 \leq B \|f\|_2^2, \quad (f \in L^2(\mathbb{R}^d)).$$

While Theorem 1 was our main motivation, it is just a sample of our results. We work with an abstract model for phase space that allows for a variety of settings. We obtain a complete analogue of Theorem 1 in the context of time-scale analysis where the atoms we design follow a pattern prescribed in the time-scale plane. This flexibility is also further exploited in the context of time-frequency analysis, yielding a variation of Theorem 1 where continuous time-frequency representations are replaced by discrete ones.

1.1. Technical overview. We now give a technical overview, in order to highlight the main steps leading to the proof of Theorem 1 and corresponding statements in Theorem 4, Theorem 5, Theorem 8 and Theorem 9.

The proofs of the main results in this paper are based on two major observations. Firstly, the norm equivalence

$$\|f\|_2^2 \approx \sum_{\gamma} \|H_{\gamma}, f\|_2^2,$$
holds for a family of time-frequency localization operators,
\begin{equation}
H_{\gamma} f(t) = \int_{\mathbb{R}^d} \eta_{\gamma}(x, \xi) V_{\varphi}(x) f(x) \varphi(t-x)e^{2\pi i \xi t} dx dw,
\end{equation}
provided that the symbols \(\eta_{\gamma} : \mathbb{R}^{2d} \to [0, +\infty)\) satisfy
\begin{equation}
\sum_{\gamma} \eta_{\gamma}(z) \approx 1
\end{equation}
and the enveloping condition
\begin{equation}
\eta_{\gamma}(z) \leq g(z - \gamma), \quad \text{for some } g \in L^1(\mathbb{R}^{2d}) \text{ and } \gamma \in \Gamma, \text{ with } \Gamma \subseteq \mathbb{R}^{2d} \text{ a lattice}.
\end{equation}
The inequalities (4) were first proved in [17] for symbols of the form \(\eta_{\gamma}(z) = h(z - \gamma)\) and \(\Gamma = \mathbb{Z}^{2d}\), then for a general lattice in [18], and finally for fully irregular symbols satisfying (7) in [43]. It is interesting to note that the proofs in [17, 18] are based on the observation that under condition (6) the norm-equivalence (4) is equivalent to the fact that finitely many eigenfunctions of the operator \(H_{\gamma}\) generate a multi-window Gabor frame over the lattice \(\Gamma\). The proof of the general case in [43] does not explicitly involve the eigenfunctions of the operators \(H_{\gamma}\), nor does it rely on tools specific to Gabor frames on lattices. Thus the question arises whether it is also possible in the irregular case to construct a frame consisting of eigenfunctions of the operators \(H_{\gamma}\). Here, this question is given a positive answer.

Secondly, the observation that (4) remains valid when the operators \(H_{\gamma}\) are replaced by finite rank approximations \(H_{\gamma}^c\), obtained by thresholding their eigenvalues, cf. Theorem 4, is the core of the proof of our main results. This finite rank approximation is in turn achieved by proving that the operators \(H_{\gamma}\) behave “globally” like projectors. More precisely, in Proposition 6 we obtain the following extension of (4)
\begin{equation}
\|f\|_2^2 \approx \sum_{\gamma} \|H_{\gamma} f\|_2^2 \approx \sum_{\gamma} \|(H_{\gamma})^2 f\|_2^2.
\end{equation}
This will allow us to “localize in phase-space” the \(L^2\)-norm estimates relating \(H_{\gamma}^c\) and \(H_{\gamma}\).

Note that, in general, the operators \(H_{\gamma}\) have infinite rank even if \(\eta_{\gamma}\) is the characteristic function of a compact set (see Lemma 1). Consequently \(\|H_{\gamma} f\|_2 \neq \|(H_{\gamma})^2 f\|_2\) and therefore the global properties of the family \(\{(H_{\gamma})^2 : \gamma \in \Gamma\}\) are crucial to prove (4). While the squared operators \((H_{\gamma})^2\) are not time-frequency localization operators, their time-frequency localizing behavior is preserved under conditions as given by (7): in Section 3 we introduce the notion of a family of operators that is well-spread in the time-frequency plane and exploit the fact that the tools from [43] are valid for these operator families.

For clarity, we choose to accentuate the case of time-frequency analysis but all the proofs are carried out in an abstract setting that yields, for example, analogous consequences in time-scale analysis.

1.2. Organization. The article is organized as follows. Section 2 introduces the abstract model for phase-space and Section 3 presents certain key technical notions, in particular the properties required for a family of localization operators to exhibit an almost-orthogonality property. In Section 4, we first prove our results in the context of time-frequency analysis, where some technical problems of the abstract setting do not arise. In addition, in the context of time-frequency analysis, we are able to extend the result on phase-space adapted frames from \(L^2(\mathbb{R}^d)\) to an entire class of Banach spaces, the modulation spaces, by exploiting spectral invariance results for pseudo-differential operators. Theorem 5 in Section 4.1 is the general version of Theorem 1 stated above. Section 5 develops the results in the abstract setting. These are then applied to time-scale analysis in Section 5.1. Finally, Section 5.2 contains an additional application of the abstract results to time-frequency analysis, this time using Gabor multipliers, which are time-frequency masking operators related to a discrete time-frequency representation (Gabor frame). The atoms thus obtained maximize their time-frequency concentration with respect to a weight on a discrete time-frequency grid and the resulting frames are relevant in numerical applications.
For clarity, the presentation of the results highlights the case of time-frequency analysis, which was our main motivation. Most of the technicalities in Section 2 are irrelevant to that setting (although they are relevant for time-scale analysis). The reader interested mainly in time-frequency analysis is encouraged to jump directly to Section 4 and then go back to Sections 2 and 3 having a clear example in mind.

2. Abstract phase-space

We now introduce a model for phase-space that allows us to consider representations such as the STFT and the wavelet transform from a unified point of view. This approach is in the spirit of coorbit theory, cf. [23, 24]. In fact, both the STFT and the wavelet transform may be interpreted as representation coefficients of certain unitary continuous representations of a locally compact group \( \mathcal{G} \). The abstract setting goes without explicit reference to an integral transform.

2.1. Locally compact groups and function spaces. Throughout the article \( \mathcal{G} \) will be a locally compact, \( \sigma \)-compact, topological group. The left Haar measure of a set \( X \subseteq \mathcal{G} \) will be denoted by \( |X| \). Integration will always be considered with respect to the left Haar measure. For \( x \in \mathcal{G} \), we denote by \( L_x \) and \( R_x \) the operators of left and right translation, defined by \( L_x f(y) = f(x^{-1}y) \) and \( R_x f(y) = f(yx) \).

We also consider the involution \( f^\vee (x) = f(x^{-1}) \).

Given two non-negative functions \( f, g \) we write \( f \lesssim g \) if there exists a constant \( C \geq 0 \) such that \( f \leq C g \). We say that \( f \approx g \) if both \( f \lesssim g \) and \( g \lesssim f \). The characteristic function of a set \( A \) will be denoted by \( \chi_A \).

A set \( \Lambda \subseteq \mathcal{G} \) is called relatively separated if for some (or any) \( V \subseteq \mathcal{G} \), relatively compact neighborhood of the identity, the quantity - called the spreadness of \( \Lambda \) -

\[
\rho(\Gamma) = \rho_V(\Gamma) := \sup_{x \in \mathcal{G}} \#(\Gamma \cap xV)
\]

is finite, i.e. if the amount of elements of \( \Gamma \) that lie in any left translate of \( V \) is uniformly bounded.

In generalization of \( L^p \)-spaces, we will consider solid, translation invariant, Banach function (BF) spaces \( E \), i.e. Banach spaces that satisfy the following.

(i) \( E \) is continuously embedded into \( L^1_{\text{loc}}(\mathcal{G}) \), the space of complex-valued locally integrable functions on \( \mathcal{G} \).

(ii) Whenever \( f \in E \) and \( g : \mathcal{G} \to \mathbb{C} \) is a measurable function such that \( |g(x)| \leq |f(x)| \) a.e., it is true that \( g \in E \) and \( \|g\|_E \leq \|f\|_E \).

(iii) \( E \) is closed under left and right translations (i.e. \( L_x E \subseteq E \) and \( R_x E \subseteq E \), for all \( x \in \mathcal{G} \)) and the following relations hold with the corresponding norm estimates,

\[
L^1_{\text{loc}}(\mathcal{G}) \ast E \subseteq E \quad \text{and} \quad E \ast L^1_{\text{loc}} \subseteq E,
\]

where \( u(x) := \|L_x\|_{E \to E} \), \( v(x) := \Delta(x^{-1}) \|R_{x^{-1}}\|_{E \to E} \) and \( \Delta \) is the modular function of \( \mathcal{G} \).

For each space \( E \) we consider weight functions \( w : \mathcal{G} \to (0, +\infty) \) that satisfy the following admissibility conditions.

\[
w(x) = \Delta(x^{-1}) w(x^{-1}),
\]

\[
w(xy) \leq w(x) w(y) \quad \text{(submultiplicativity)},
\]

\[
w(x) \geq C_{E, w} \max \{u(x), u(x^{-1}), v(x), \Delta(x^{-1}) v(x^{-1})\},
\]

for some constant \( C_{E, w} > 0 \). Under these conditions, we say that \( w \) is admissible for \( E \). It then follows that \( w(x) \gtrsim 1 \), \( L^1_{w} \ast E \subseteq E \) and \( E \ast L^1_{w} \subseteq E \), with the corresponding norm estimates. Moreover, the constants in those estimates depend only on \( C_{E, w} \), cf. [23].

If \( E \) is a solid translation invariant BF space and \( \Gamma \subseteq \mathcal{G} \) is a relatively separated set, we construct discrete versions of \( E \) as follows. Fix \( V \), a symmetric relatively compact neighborhood of the identity and let,

\[
E_d = E_d(\Gamma) := \{ c \in \mathbb{C}^\Gamma \mid \sum_\lambda |c_\lambda| \chi_{\gamma V} \in E \},
\]
and endow it with the norm,
\[ \|(c_\gamma)_{\gamma \in \Gamma}\|_{E_d} := \sum_{\gamma} |c_\gamma| \chi_{\gamma}V \|E. \]

The definition, of course, depends on \( V \) but a different choice of \( V \) yields the same space with an equivalent norm (this is a consequence of the right invariance of \( E \), see for example [41, Lemma 2.2]). When \( E = L^\infty_\omega \), for an admissible weight \( w \), the corresponding discrete space \( E_d(\Gamma) \) is \( l^\infty_\omega(\Gamma) \), where the weight \( w \) is restricted to the set \( \Gamma \). This follows from the admissibility of \( w \) since for \( x \in \gamma V \), \( w(x) \approx w(\gamma) \).

For a solid, translation invariant BF space \( E \) we define the left Wiener amalgam space in the following way. We select again a symmetric relatively compact neighborhood of the identity \( V \). For a locally bounded function \( f : \mathcal{G} \to \mathbb{C} \) consider the left local maximum function defined by,
\[ f^\#(x) := \sup_{y \in V} |f(xy)| = \| f \cdot (L_x \chi_V) \|_\infty, \quad (x \in \mathcal{G}). \]

The Wiener amalgam space is defined as,
\[ W(L^\infty, E)(\mathcal{G}) := \{ f \in L^\infty_\infty(\mathcal{G}) \mid f^\# \in E \}, \]
and given the norm \( \|f\|_{W(L^\infty, E)} := \|f^\#\|_E \). A different choice for \( V \) yields the same space with an equivalent norm (see for example [20, Theorem 1]). The right Wiener amalgam space \( W_R(L^\infty, E) \) is defined similarly, this time using the right local maximum function,
\[ f^\#(x) := \sup_{y \in V} |f(yx)| = \| f \cdot (R_x \chi_V) \|_\infty \]
and given the norm \( \|f\|_{W_R(L^\infty, E)} := \|f^\#\|_E \). When \( E \) is a weighted \( L^p \) space, the corresponding amalgam space coincides with the classical \( L^\infty - \ell^p \) amalgam space [35, 28]. For the general theory of amalgam spaces in the broader context of possibly non-solid spaces see [20, 21]. In the present article we will be mainly interested in the spaces \( W(L^\infty, L^1_\omega) \) and \( W_R(L^\infty, L^1_\omega) \). We will need the following fact.

**Proposition 1.** The spaces \( W(L^\infty, L^1_\omega) \) and \( W_R(L^\infty, L^1_\omega) \) are convolution algebras. That is, the relations \( W(L^\infty, L^1_\omega) \ast W(L^\infty, L^1_\omega) \hookrightarrow W(L^\infty, L^1_\omega) \) and \( W_R(L^\infty, L^1_\omega) \ast W_R(L^\infty, L^1_\omega) \hookrightarrow W_R(L^\infty, L^1_\omega) \) hold together with the corresponding norm estimates.

**Proof.** The left amalgam space satisfies the (translation invariance) relation \( L^1_\omega \ast W(L^\infty, L^1_\omega) \hookrightarrow W(L^\infty, L^1_\omega) \). This is a particular case of [24, Theorem 7.1] and can also be readily deduced from the definitions. Since \( W(L^\infty, L^1_\omega) \hookrightarrow L^1_\omega \) the statement about \( W(L^\infty, L^1_\omega) \) follows. The involutions \( f^\vee(x) = f(x^{-1}) \) maps \( W_R(L^\infty, L^1_\omega) \) isometrically onto \( W(L^\infty, L^1_\omega) \) (because \( V = V^{-1} \)) and satisfies \( (f * g)^\vee = g^\vee \ast f^\vee \). Hence the statement about the right amalgam space follows from the one about the left one. \( \square \)

### 2.2. The model for phase-space

In the abstract model for phase-space we consider a solid BF space \( E \) (called the environment) and a certain distinguished subspace \( S_E \), which is the range of an idempotent operator \( P \). In the applications \( S_E \) will be taken to be the range of an integral transform, such as the STFT in the case of time-frequency analysis. The precise form of the model is taken from [43] and is designed to fit the theory in [23] and [27] (see also [39]).

We list a number of ingredients in the form of two assumptions: (A1) and (A2).

(A1)
- \( E \) is a solid, translation invariant BF space, called the environment.
- \( w \) is an admissible weight for \( E \).
- \( S_E \) is a closed complemented subspace of \( E \), called the atomic subspace.
- Each function in \( S_E \) is continuous.

The second assumption is that the retraction \( E \to S_E \) is given by an operator dominated by right convolution with a kernel in \( W(L^\infty, L^1_\omega) \cap W_R(L^\infty, L^1_\omega) \). In the case of time-frequency or time-scale analysis, the retraction operator is given by the reproducing kernel (see Sections 4 and 5.1).

(A2) We have an operator \( P \) and a non-negative function \( K \) satisfying the following.
- \( P : W(L^1, L^\infty_\omega) \to L^\infty_\omega \) is a (bounded) linear operator,
− \( P(E) = S_E \) and \( P(f) = f \), for all \( f \in S_E \),
− \( K \in W(L_\infty, L_1^w) \cap W_R(L_\infty, L_1^w) \),
− For \( f \in W(L^1, L_1^w) \),
\[
|P(f)(x)| \leq \int_{\mathcal{G}} |f(y)| K(y^{-1}x) dy, \quad (x \in \mathcal{G}).
\] (14)

When \( E = L^2(\mathcal{G}) \) we will also assume the following,
(A3) \( P : L^2(\mathcal{G}) \to S_{L^2} \) is the orthogonal projection.

For the remainder of Section 2 we assume (A1) and (A2). Under these conditions the following holds.

**Proposition 2.** [[43, Prop. 3]]

(a) \( P \) boundedly maps \( E \) into \( W(L^\infty, E) \).
(b) \( S_E \hookrightarrow W(L^\infty, E) \).
(c) If \( f \in W(L^1, L^1_{1/w}) \), then \( \|P(f)\|_{L^1_{1/w}} \lesssim \|f\|_{W(L^1, L^1_{1/w})} \|K\|_{W_R(L^\infty, L_1^w)} \).
(d) If \( f \in W(L^1, L^\infty) \), then \( \|P(f)\|_{L^\infty} \lesssim \|f\|_{W(L^1, L^\infty)} \|K\|_{W_R(L^\infty, L_1^w)} \).

**Remark 1.** Since \( w \gtrsim 1 \), \( L^\infty \hookrightarrow L^1_{1/w} \).

**Remark 2.** The estimates in Prop. 2 hold uniformly for all the spaces \( E \) with the same weight \( w \) and the same constant \( C_{E,w} \) (cf. (13)). This is relevant to the applications, where the same projection \( P \) is used with different spaces \( E \) and corresponding subspaces \( S_E = P(E) \).

### 2.3. Phase-space multipliers.

Time-frequency localization operators (2) are a particular example, be it of overwhelming practical importance, cf. [14, 16, 49, 50, 11], of the general concept of phase-space multipliers.

For \( m \in \mathcal{L}^\infty(\mathcal{G}) \), the phase-space multiplier \( M_m : S_E \to S_E \) with symbol \( m \) is defined by,
\[
M_m(f) := P(mf), \quad (f \in S_E).
\] (15)

The operator \( M_m \) is clearly bounded by Proposition 2 and the solidity of \( E \),
\[
\|M_m(f)\|_E \lesssim \|m\|_\infty \|f\|_E, \quad (f \in S_E).
\] (16)

If \( S_E \) is the range of the STFT (and \( P \) is the orthogonal projection onto it) the corresponding operators are unitarily equivalent to the time-frequency localization operators ([14, 11, 6, 12]). More generally, if the space \( S_E \) is the range of the abstract wavelet transform associated with a unitary representation of \( \mathcal{G} \), these operators are called localization operators or wavelet multipliers (see for example [50, 37]).

**Remark 3.** In this article we will mainly be concerned with multipliers with bounded symbols \( m \). However, due to the the regularizing effect of \( P \), the condition that \( m \) be bounded is by no means necessary for \( M_m : S_E \to S_E \) to be bounded. See for example [25, 11] for sharper boundedness results for time-frequency localization operators.

For future reference we note some Hilbert-space properties of phase-space multipliers (when \( E = L^2(\mathcal{G}) \)).

**Proposition 3.** Let \( E = L^2(\mathcal{G}) \) and assume (A1), (A2) and (A3). Then, the following hold.
(a) Let \( m_1, m_2 \in \mathcal{L}^\infty(\mathcal{G}) \) be real-valued. If \( m_1 \leq m_2 \) a.e., then \( M_{m_1} \leq M_{m_2} \) as operators. In particular if \( m \) is non-negative and bounded, then \( M_m \) is a positive operator.
(b) Let \( m \in \mathcal{L}^1(\mathcal{G}) \cap \mathcal{L}^\infty(\mathcal{G}) \) be non-negative. Then \( M_m : S_{L^2} \to S_{L^2} \) is trace-class and \( \text{trace}(M_m) \lesssim \|m\|_1 \).

**Proof.** To prove (a), let \( f \in S_{L^2} \) and note that by (A3),
\[
\langle M_m f, f \rangle = \langle P(m_1 f), f \rangle = \langle m_1 f, f \rangle
\]
\[
= \int_{\mathcal{G}} m_1(x) |f(x)|^2 \, dx
\]
\[
\leq \int_{\mathcal{G}} m_2(x) |f(x)|^2 \, dx = \langle M_{m_2} f, f \rangle.
\]
Let us now prove (b). For $f \in S_{L^2}$ and $x \in \mathcal{G}$, by (14)
\[
|f(x)| \leq \int_{\mathcal{G}} |f(y)| K(y^{-1} x) dy = \int_{\mathcal{G}} |f(y)| K^\vee(x^{-1} y) dy \leq \|f\|_2 \|K^\vee\|_2.
\]
This is finite because $K^\vee \in W_R(L^\infty, L^1_w) \subseteq L^2$. Hence $f(x) = \langle f, E_x \rangle$, for some function $E_x \in S_{L^2}$ with $\|E_x\|_2 \leq \|K^\vee\|_2$.

Let $\{e_k\}_k$ be an orthonormal basis of $S_{L^2}$. Since by (a) $M_m$ is a positive operator, it suffices to check that $\sum_k \langle M_m e_k, e_k \rangle \lesssim \|m\|_1$ (see for example, [45, Theorem 2.14]). To this end, note that for $x \in \mathcal{G}$,
\[
\sum_k |e_k(x)|^2 = \sum_k |\langle e_k, E_x \rangle|^2 = \|E_x\|^2 \leq \|K^\vee\|^2_2.
\]
Hence,
\[
\sum_k \langle M_m e_k, e_k \rangle = \sum_k \int_{\mathcal{G}} m(x) |e_k(x)|^2 \leq \|K^\vee\|^2_2 \|m\|_1.
\]
This completes the proof. □

3. Well-spread families of operators

One of the main technical insights of this article is the fact, that some important tools used in the investigation of families phase-space multipliers only depend on certain phase-space localization estimates. To formalize this observation, we now introduce the key concept of families of operators that are well-spread in phase-space. These families of operators are required to be dominated by product-convolution operators $f \mapsto (fg(\gamma^{-1} \cdot)) * \Theta$ centered at suitably distributed nodes $\gamma$. The advantage of working with well-spread families instead of just families of phase-space multipliers is that well-spreadness is stable under various operations, e.g. finite composition. This will be essential to the proofs of our main results.

In this section, we assume that (A1) and (A2) from Section 2.2 hold.

**Definition 1.** Let $\Gamma \subseteq \mathcal{G}$ be a relatively separated set, $\Theta \in W(L^\infty, L^1_w) \cap W_R(L^\infty, L^1_w)$ and $g \in W_R(L^\infty, L^1_w)(\mathcal{G})$ be non-negative functions. A family of operators $\{T_\gamma : S_E \to S_E : \gamma \in \Gamma\}$ is called well-spread with envelope $(\Gamma, \Theta, g)$ if the following pointwise estimate holds

\[
|T_\gamma f(x)| \leq \int_{\mathcal{G}} g(\gamma^{-1} y) |f(y)| \Theta(y^{-1} x) dy, \quad (\gamma \in \Gamma, x \in \mathcal{G}).
\]  

When we do not want to emphasize the role of the envelope, we simply say that $\{T_\gamma\}_\gamma$ is a well-spread family of operators; this implies the existence of an adequate envelope.

The canonical example of a well-spread family of operators is a family of phase-space multipliers $\{M_\eta, \eta \in \Gamma\}$ associated with an adequate family of symbols. A family of symbols $\{\eta_\gamma : \mathcal{G} \to \mathbb{C} | \gamma \in \Gamma\}$ is called well-spread if it satisfies the following.

- $\Gamma \subseteq \mathcal{G}$ is a relatively separated set.
- There is a function $g \in W_R(L^\infty, L^1_w)$ such that
\[
|\eta_\gamma(x)| \leq g(\gamma^{-1} x), \quad (x \in \mathcal{G}, \gamma \in \Gamma).
\]

The pair $(\Gamma, g)$ is called an envelope for $\{\eta_\gamma\}_\gamma$. Together with the properties of the convolution kernel $K$ dominating the projection onto $S_E$, we immediately obtain a family of well-spread operators.

**Proposition 4.** Let $\{\eta_\gamma\}_\gamma$ be a well-spread family of symbols. Then, the corresponding family of phase-space multipliers $\{M_\eta, \eta \in \Gamma\}$ is well-spread.

**Proof.** It follows readily from the definitions that if $(\Gamma, g)$ is an envelope for $\{\eta_\gamma\}_\gamma$, then $(\Gamma, K, g)$ is an envelope for $\{M_\eta, \eta \in \Gamma\}$. □
The reason why we introduce the concept of well-spread families of operators is that composition of phase-space multipliers usually fails to yield a phase-space multiplier. However, the estimate in (17) is stable under various operations. In this article we will be mainly interested in the following.

Proposition 5. Let \( \{ \eta_\gamma \}_\gamma \) be a well-spread family of symbols. Then, the family of operators \( \{ (M_\eta_\gamma)^2 : \gamma \in \Gamma \} \) is well-spread.

Proof. If \((\Gamma, g)\) is an envelope for \( \{ \eta_\gamma \}_\gamma \) then,

\[
\|(M_\eta_\gamma)^2 f(x)\| \leq \int_G g(\gamma^{-1} y) |M_\eta_\gamma f(y)| K(y^{-1} x) dy \\
\leq \int_G \int_G g(\gamma^{-1} y) g(\gamma^{-1} z) |f(z)| K(z^{-1} y) K(y^{-1} x) dy dz \\
\leq \|g\|_{\infty} \int_G g(\gamma^{-1} z) |f(z)| (K * K)(z^{-1} x) dz.
\]

Hence, if we set \( \Theta := K * K \), it follows that \((\Gamma, \Theta, \|g\|_\infty g)\) is an envelope for \( \{ (M_\eta_\gamma)^2 \}_\gamma \). Since \( K \) belongs to \( W(L^\infty, L^1_w) \cap W_R(L^\infty, L^1_w) \), by Proposition 1, so does \( \Theta \). \( \square \)

3.1. Almost-orthogonality estimates. We now introduce one of the key estimates of the article.

Theorem 2 ([43]). Let \( \{ T_\gamma : \gamma \in \Gamma \} \) be a well-spread family of operators. Suppose that the operator \( \sum_\gamma T_\gamma : S_E \to S_E \) is invertible. Then, the following norm equivalence holds,

\[
\|||T_\gamma f||_{L^2(\mathcal{G})}||_{\gamma \in \Gamma}||_{E_E} \approx ||f||_{E_E}, \quad (f \in S_E).
\]

Remark 4. The implicit constants depend only on \( \|K\|_{W(L^\infty, L^1_\delta)} \), \( \|K\|_{W_R(L^\infty, L^1_\delta)} \), \( \|\Theta\|_{W_R(L^\infty, L^1_\delta)}, \|\Theta\|_{W(L^\infty, L^1_\delta)}, \rho(\Gamma) \) (cf. (9)) and \( C_{E,w} \) (cf. (13)).

Theorem 2 is proved in [43] for families of phase-space multipliers associated with well-spread families of symbols. However, the argument works without changes for the case of general well-spread families of operators. Indeed, the definition in (17) is tailored to the requirements of the proof in [43]. For completeness, we sketch a proof of Theorem 2 in the Appendix.

We also point out that the \( L^2 \)-norm in (19) can be replaced by any solid translation invariant norm (cf. Assumption (B1) in [43]). However, the by far most important case is the one of \( L^2 \). Indeed the core of our argument in this article makes use of (19) to extrapolate certain thresholding estimates from \( L^2 \) to other Banach spaces.

We regard Theorem 2 as an almost orthogonality principle. Its main insight is that, because of the phase-space localization of the family \( \{ T_\gamma \}_\gamma \), the effect of each individual operator within the sum \( \sum_\gamma T_\gamma \) decouples.

Corollary 1. Let \( \{ T_\gamma : \gamma \in \Gamma \} \) be a well-spread family of operators. Suppose that \( E = L^2(\mathcal{G}) \) and that the operator \( \sum_\gamma T_\gamma : S_{L^2} \to S_{L^2} \) is invertible. Then, so is the operator \( \sum_\gamma T_\gamma^* T_\gamma : S_{L^2} \to S_{L^2} \).

Proof. By Theorem 2, the invertibility of \( \sum_\gamma T_\gamma \) implies that for \( f \in S_{L^2} \),

\[
\|f\|^2 \approx \sum_\gamma \|T_\gamma f\|^2 = \left( \sum_\gamma T_\gamma^* T_\gamma f, f \right).
\]

Hence \( \sum_\gamma T_\gamma^* T_\gamma \) is positive definite and therefore invertible on \( S_{L^2} \). \( \square \)

Remark 5. In the context of time-frequency analysis, Corollary 1 will be extended to other spaces \( E \) besides \( L^2(\mathcal{G}) \).

Remark 6. Corollary 1 gives further support to the interpretation of Theorem 2 as an almost orthogonality principle. Let us consider for simplicity a well-spread family of self-adjoint operators \( \{ T_\gamma \}_\gamma \). Since \( (\sum_\gamma T_\gamma)^2 = \sum_\gamma (T_\gamma)^2 + \sum_{\gamma \neq \gamma'} T_\gamma^* T_{\gamma'} \) and \( (\sum_\gamma T_\gamma)^2 \) is invertible if and only if \( \sum_\gamma T_\gamma \) is, Corollary 1 says that the invertibility of \( (\sum_\gamma T_\gamma)^2 \) implies that of its “diagonal part” \( \sum_\gamma (T_\gamma)^2 \).
4. The case of Time-Frequency analysis

In this section we consider the case $\mathcal{G} = \mathbb{R}^{2d}$, and we let the phase-space be the time-frequency (TF) plane. In time-frequency analysis, function spaces defined via their members’ properties in phase space are known as modulation spaces, whose definition we give next. Let $w : \mathbb{R}^{2d} \to (0, +\infty)$ be a submultiplicative, even weight that satisfies the GRS condition,

$$\lim_{n \to \infty} w(nx)^{1/n} = 1, \quad (x \in \mathbb{R}^{2d}).$$

Let $v : \mathbb{R}^{2d} \to (0, +\infty)$ be a $w$-moderated weight; that is

$$v(x + y) \leq C_v w(x)v(y), \quad (x, y \in \mathbb{R}^{2d}),$$

for some constant $C_v > 0$. Assume further that $v$ is moderated by a polynomial weight$^1$. Given a non-zero Schwartz-class function $\varphi \in \mathcal{S}(\mathbb{R}^d)$, the modulation space $M^p_v(\mathbb{R}^d)$, $(1 \leq p \leq +\infty)$ is defined by the set of all distributions $f \in \mathcal{S}'(\mathbb{R}^d)$ such that $V_\varphi f$ belongs to the weighted Lebesgue space $L^p_v(\mathbb{R}^{2d})$. The space $M^p_v$ is given the norm $\|f\|_{M^p_v} := \|V_\varphi f\|_{L^p_v}$. A different choice for $\varphi$ yields the same space with an equivalent norm. Moreover, not only Schwartz-class functions can be used as windows; any non-zero $\varphi \in M^1_v$ is adequate, and we will make this assumption hereafter. For convenience we also assume that $\|\varphi\|_2 = 1$. When $p = 2$ and $v \equiv 1$, $M^p_v$ is just $L^2(\mathbb{R}^d)$.

We now indicate how the abstract setting from Section 2 applies to time-frequency analysis. We let $\mathcal{G} = \mathbb{R}^{2d}$, $E = L^p_v(\mathbb{R}^{2d})$ and $S_E = S^p_v := V_\varphi(L^p_v)$. The projection $P : L^p_v(\mathbb{R}^{2d}) \to S^p_v$ is given by $P(f) := V_\varphi V^*_\varphi(f)$. If $E = L^2(\mathbb{R}^{2d})$, $P$ is in fact the orthogonal projector onto $S^2_v$. The estimate in (14) is satisfied with $K := \|V_\varphi(\varphi)\|$. By definition, $\varphi \in \mathcal{M}_w^1$ means that $V_\varphi(\varphi) \in L^1_v(\mathbb{R}^{2d})$, but it is well-known that in this case, $V_\varphi(\varphi)$ also belongs to $W(L^\infty, L^1_v)(\mathbb{R}^{2d})$ (see [30, Proposition 12.1.11]; note that this fact can also be derived from the norm equivalence in Proposition 2.) Since $\mathcal{G} = \mathbb{R}^{2d}$ is abelian, $W(L^\infty, L^1_v)(\mathbb{R}^{2d}) = W(L^\infty, L^1_v)(\mathbb{R}^{2d})$, and a family of symbols $\{\eta_\gamma : \mathbb{R}^{2d} \to \mathbb{C} \mid \gamma \in \Gamma\}$ is well-spread if, for a relatively separated set $\Gamma \subseteq \mathbb{R}^{2d}$ and $g \in W(L^\infty, L^1_v)(\mathbb{R}^{2d})$

$$|\eta_\gamma(x)| \leq g(x - \gamma), \quad (x \in \mathbb{R}^{2d}, \gamma \in \Gamma).$$

Section 3.$^2$

In the present context of time-frequency analysis, we consider operators in the signal domain, which is unitarily mapped to phase-space by $V_\varphi$. Then, in equivalence to the definition given in Section 3, a family of operators $\{H_\gamma : \mathcal{M}^1_{\mathcal{L}_w}(\mathbb{R}^2d) \to \mathcal{M}^\infty_{\mathcal{L}_w}(\mathbb{R}^2d) \mid \gamma \in \Gamma\}$ is said to be well-spread in the time-frequency plane if there exists an envelope $(\Gamma, \Theta, g)$, such that

- $\Gamma \subseteq \mathbb{R}^{2d}$ is relatively separated,
- $\Theta, g \in W(L^\infty, L^1_v)(\mathbb{R}^{2d})$ and the following estimate holds for $x \in \mathbb{R}^{2d}$ and $\gamma \in \Gamma$:

$$|V_\varphi H_\gamma f(x)| \leq \int_{\mathbb{R}^{2d}} |V_\varphi f(y)| \Theta(x - y)g(y - \gamma)dy.$$  

Note that $\{H_\gamma\}_\gamma$ is well-spread in the TF plane if and only if the family of operators

$$\{ V_\varphi H_\gamma V^*_\varphi : S^\infty_{\mathcal{L}_w} \to S^\infty_{\mathcal{L}_w} \mid \gamma \in \Gamma\}$$

is well-spread in the sense of Section 3.

Note also that, if $\{H_\gamma\}_\gamma$ is well-spread in the time-frequency plane, then, because of (21), each operator $H_\gamma$ maps $M^p_v(\mathbb{R}^d)$ into $M^p_v(\mathbb{R}^d)$ with a norm bound independent of $\gamma$,

$$\|H_\gamma f\|_{M^p_v} \leq C_v \|g\|_{\infty} \|\Theta\|_{L^1_v} \|f\|_{M^p_v}.$$ 

$^1$This assumption is only made in order to define modulation spaces as subsets of the class of tempered distributions.

For a general weight, the space $M^p_v$ can still be constructed as an abstract coorbit space.

$^2$We remark that for the results in this section the assumption $g \in W(L^\infty, L^1_v)(\mathbb{R}^{2d})$ can be relaxed to $g \in L^1(\mathbb{R}^{2d})$, but the proofs would be more technical with little practical gain (cf. [43, Section 2.4] and the Appendix).
Remark 7. In parallel to Proposition 4, if \( \{ \eta_\gamma \} \) is a well-spread family with envelope \( (\Gamma, g) \), then the corresponding family of time-frequency localization operators \( \{ H_{\eta_\gamma} \} \) is well-spread in the time-frequency plane with envelope \( (\Gamma, K, g) \).

In the case of time-frequency analysis we can strengthen Theorem 2 by means of spectral invariance results for pseudo-differential operators [46, 47, 32, 31]. Due to these results, the invertibility assumption in Theorem 2 can be replaced by assuming invertibility on \( L^2(\mathbb{R}^d) \) only.

Theorem 3. Let \( \{ H_\gamma : \gamma \in \Gamma \} \) be well-spread in the TF plane. Suppose that the operator \( \sum_\gamma H_\gamma : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d) \) is invertible. Then, also \( \sum_\gamma H_\gamma : M^p \to M^p \) is invertible and the following norm equivalence holds, for \( 1 \leq p \leq +\infty \) and \( w \)-dominated weights \( v \).

\[
\| f \|_{M^p} \approx \left( \sum_{\gamma \in \Gamma} \| H_\gamma f \|_{L^2(\mathbb{R}^d)}^p v(\gamma)^p \right)^{1/p}, \quad (f \in M^p(\mathbb{R}^d)),
\]

with the usual modification for \( p = \infty \).

Remark 8. The estimates hold uniformly for \( 1 \leq p \leq +\infty \) and any family of weights having a uniform constant \( C_0 \) (cf. (20)).

Proof of Theorem 3. Let \( T_\gamma := V_{\varphi} H_\gamma V_{\varphi}^* : S^p \to S^p_\circ \). The family \( \{ T_\gamma : \gamma \in \Gamma \} \) is well-spread (in the sense of Section 3). In order to apply Theorem 2 we need to show that \( \sum_\gamma T_\gamma : S^p \to S^p_\circ \) is invertible. This is the case if and only if \( R_{p,v} := \sum_\gamma H_\gamma : M^p \to M^p_\circ \) is invertible. By assumption we know that \( R_2 : \sum_\gamma H_\gamma : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d) \) is invertible.

Let \( \{ \pi(\lambda) h : \lambda \in \Lambda \} \) be a (Gabor) frame for \( L^2(\mathbb{R}^d) \), where \( h \in M^1_\circ \) and \( \Lambda \subseteq \mathbb{R}^{2d} \) is a lattice (for example we can take \( h(x) = e^{-\pi |x|^2} \) and \( \Lambda := \mathbb{Z}^d \times 1/2\mathbb{Z}^d \)). Using (21) and the fact that \( |V_{\varphi} \pi(\lambda) h(x)| = |V_{\varphi} h(x-\lambda)| \) we estimate,

\[
|\langle R_{p,v} \pi(\lambda) h, \pi(\mu) h \rangle| = |\langle V_{\varphi} R_{p,v} \pi(\lambda) h, V_{\varphi} \pi(\mu) h \rangle| \\
\leq \int_{\mathbb{R}^{2d}} \int_{\mathbb{R}^{2d}} |V_{\varphi} \pi(\lambda) h(y)| \Theta(x-y) \sum_\gamma g(y-\gamma) |V_{\varphi} \pi(\mu) h(x)| \, dy \, dx \\
\lesssim \| g \|_{W(L^\infty,L^1_\circ)} \int_{\mathbb{R}^{2d}} \int_{\mathbb{R}^{2d}} |V_{\varphi} h(y-\lambda)| \Theta(x-y) |V_{\varphi} h(x-\mu)| \, dy \, dx \\
\lesssim |(V_{\varphi} h) * (V_{\varphi} h) * \Theta)(\lambda - \mu)|.
\]

Since \( h \in M^1_\circ \) and \( \Theta \in W(L^\infty,L^1_\circ) \), \( |V_{\varphi} h| * |V_{\varphi} h| * \Theta \in W(L^\infty,L^1_\circ) \). Therefore, the restriction of that function to \( \Lambda \) gives an \( \ell^1 \) sequence. By [32, Theorem 3.2 and Corollary 3.3], \( R_2 \) is a pseudodifferential operator with symbol in the weighted Sjöstrand’s class \( M_{w,1}^{\infty} \), where \( \tilde{w}(z_1,z_2) = (z_2,z_1) \) and \( (z_1,z_2) \in \mathbb{R}^d \times \mathbb{R}^d \). Now the invertibility of \( R_{p,v} \) follows from the invertibility of \( R_2 \) and [32, Corollary 4.7] (see also [31]).

4.1. Frames of eigenfunctions of time-frequency localization operators. We now turn to the construction of frames comprised of eigenfunctions of localization operators. Let \( \{ \eta_\gamma \} \), be a well-spread family of non-negative functions on \( \mathbb{R}^{2d} \) and consider the corresponding family of time-frequency localization operators,

\[
H_{\eta_\gamma} f = V_{\varphi}^* (\eta_\gamma V_{\varphi} f).
\]

Since each \( \eta_\gamma \) is non-negative and belongs to \( L^1(\mathbb{R}^{2d}) \), \( H_{\eta_\gamma} \) is positive and trace class, and \( \text{trace}(H_{\eta_\gamma}) = \| \eta_\gamma \|_1 \) (see [7, 25, 50]). Hence \( H_{\eta_\gamma} : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d) \) can be diagonalized as,

\[
H_{\eta_\gamma} f = \sum_{k \geq 1} \lambda_k \langle f, \phi_k^\gamma \rangle \phi_k^\gamma, \quad (f \in L^2(\mathbb{R}^d))
\]

\(^3\text{Alternatively, use Prop. 3}\)
where \( \{ \phi_k^\gamma \mid k \in \mathbb{N} \} \) is an orthonormal subset of \( L^2(\mathbb{R}^d) \) - possibly incomplete if \( \ker(H_{\eta_i}) \neq \{0\} \) and \( (\lambda_k^\gamma)_k \) is a non-increasing sequence of non-negative real numbers satisfying,
\[
\sum_k \lambda_k^\gamma = \text{trace}(H_{\eta_i}) = \|\eta_i\|_1.
\]

The time-frequency profile of the functions \( \{ \phi_k^\gamma \mid k \in \mathbb{N} \} \) is optimally adapted to the mask \( \eta_i \) in the following sense. For each \( N \in \mathbb{N} \), the set \( \{\phi_1^\gamma, \ldots, \phi_N^\gamma\} \) is an orthonormal set maximizing the quantity,
\[
\sum_{k=1}^N \int_{\mathbb{R}^{2d}} \eta_i(z) |V_{\varphi} f_k(z)|^2 \, dz,
\]
among all orthonormal sets \( \{f_1, \ldots, f_N\} \subseteq L^2(\mathbb{R}^d) \). Moreover, the functions \( \phi_k^\gamma \) enjoy other time-frequency concentration properties. For example, since \( \eta_i \in L^1_w \) and \( \varphi \in M^1_w \), it is easy to see that \( \phi_k^\gamma \in M^1_w \) (see for example [18, Lemma 5]).

For each \( \varepsilon > 0 \), we define the operator \( H_{\eta_i}^\varepsilon \) by applying a threshold to the eigenvalues of \( H_{\eta_i} \),
\[
H_{\eta_i}^\varepsilon f := \sum_{k : \lambda_k^\gamma > \varepsilon} \lambda_k^\gamma \langle f, \phi_k^\gamma \rangle \phi_k^\gamma, \quad (f \in L^2(\mathbb{R}^d)).
\]

Hence,
\[
\|H_{\eta_i}^\varepsilon f\|_2 \leq \|H_{\eta_i} f\|_2 \leq \|H_{\eta_i}^\varepsilon f\|_2 + \varepsilon \|f\|_2, \quad (f \in L^2(\mathbb{R}^d)).
\]

We now state the main technical result.

**Theorem 4.** Let \( \{\eta_i\}_i \) be a well-spread family of non-negative symbols on \( \mathbb{R}^{2d} \) such that \( \sum_i \eta_i = 1 \). Then, there exist constants \( 0 < \varepsilon < C < +\infty \) such that for all sufficiently small \( \varepsilon > 0 \),
\[
c \|f\|_{M^p_w} \leq \left( \sum_{\gamma \in \Gamma} \|H_{\eta_i}^\varepsilon f\|^p_{L^2(\mathbb{R}^d)} v(\gamma)^p \right)^{1/p} \leq C \|f\|_{M^p_w}, \quad (f \in M^p_w(\mathbb{R}^d)),
\]
with the usual modification for \( p = \infty \).

The choice of \( \varepsilon \) and the estimates are uniform for \( 1 \leq p \leq +\infty \) and any family of weights having a uniform constant \( C_v \) (cf. (20)).

The strategy to prove Theorem 4 is to show that “globally” the operators \( \{H_{\eta_i}\}_i \) behave like projectors. Note that in general \( \|(H_{\eta_i})^2 f\|_2 \not\approx \|H_{\eta_i} f\|_2 \). Indeed, if \( \eta_i \) is the characteristic function of a set with non-empty interior, then the range operator \( H_{\eta_i} \) is infinitely dimensional (see Lemma 1) and \( \|(H_{\eta_i})^2 \phi_k^\gamma\|_2 = (\lambda_k^\gamma)^2 \not\approx \lambda_k^\gamma = \|H_{\eta_i} \phi_k^\gamma\|_2 \). However, we prove the following.

**Proposition 6.** Let \( \{\eta_i\}_i \) be a well-spread family of non-negative symbols on \( \mathbb{R}^{2d} \) such that \( \sum_i \eta_i = 1 \). Then, for \( f \in M^p_w(\mathbb{R}^d) \),
\[
\|f\|_{M^p_w} \approx \left( \sum_{\gamma \in \Gamma} \|(H_{\eta_i})^2 f\|^p_{L^2(\mathbb{R}^d)} v(\gamma)^p \right)^{1/p} \approx \left( \sum_{\gamma \in \Gamma} \|(H_{\eta_i})^2 f\|^p_{L^2(\mathbb{R}^d)} v(\gamma)^p \right)^{1/p}.
\]

**Proof.** The estimate \( \|f\|_{M^p_w} \approx \|(H_{\eta_i} f\|_2)_i \|_{L^p_w} \) is contained in [18] for the case of families of symbols consisting of lattice translates of a single function, and in [43] in the present generality. It follows readily from Theorem 3 by the following argument. Since the symbols \( \eta_i \) satisfy \( m := \sum_i \eta_i \geq A \) for some constant \( A > 0 \) it follows that,
\[
\left( \sum_{\gamma} H_{\eta_i} f, f \right) = \left( H_m f, f \right) = \left( V_{\varphi}^*(m V_{\varphi} f), f \right)
\]
\[
= \langle m V_{\varphi} f, V_{\varphi} f \rangle = \int_{\mathbb{R}^{2d}} m(z) |V_{\varphi} f(z)|^2 \, dz \geq A \|V_{\varphi} f\|_2^2 = A \|f\|_2^2.
\]
Hence, \( \sum_{\gamma} H_{\eta_i} : L^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{R}^d) \) is invertible (cf. [12, Last remark]). Now Prop. 4 and Theorem 3 yield the desired estimate.
Proof of Theorem 4. Given $\varepsilon > 0$ and $f \in \mathcal{M}_p(\mathbb{R}^d)$ we apply (25) to $H_{\eta_1}(f)$ to obtain
$$\|(H_{\eta_1})^2 f\|_2 \leq \|H_{\eta_1}^\ell H_{\eta_1} f\|_2 + \varepsilon \|H_{\eta_1} f\|_2.$$  
Since $H_{\eta_1}^\ell$ and $H_{\eta_1}$ commute, $\|H_{\eta_1}^\ell H_{\eta_1} f\|_2 = \|H_{\eta_1} H_{\eta_1}^\ell f\|_2 \lesssim \|H_{\eta_1}^\ell f\|_2$. Hence,
$$\|(H_{\eta_1})^2 f\|_2 \lesssim \|H_{\eta_1}^\ell f\|_2 + \varepsilon \|H_{\eta_1} f\|_2.$$  
Taking $\ell_p$ norm on $\gamma$ yields,
$$\|(\|H_{\eta_1}^\ell f\|_2)_{\gamma}\|_{\ell_p} \lesssim \|(\|H_{\eta_1}^\ell f\|_2)_{\gamma}\|_{\ell_p} + \varepsilon \|(\|H_{\eta_1} f\|_2)_{\gamma}\|_{\ell_p}.$$  
Using the estimates in Prop. 6 we get,
$$\|f\|_{\mathcal{M}_p} \leq C\|(\|H_{\eta_1}^\ell f\|_2)_{\gamma}\|_{\ell_p} + c\varepsilon \|f\|_{\mathcal{M}_p},$$  
for some constants $c, C$. Hence,
$$(1 - c\varepsilon)\|f\|_{\mathcal{M}_p} \leq C\|(\|H_{\eta_1}^\ell f\|_2)_{\gamma}\|_{\ell_p}.$$  
This gives the desired lower bound (for all $0 < \varepsilon < 1/c$). The upper bound follows from the first inequality in (25) and Prop. 6. \qed

Remark 9. Note that the proof of Theorem 4 only uses the estimate in (25) and the fact that $H_{\eta_1}$ and $H_{\eta_1}^\ell$ commute. Hence, more general thresholding rules besides the one in (24) can be used.

Finally, we obtain the desired result on frames of eigenfunctions.

Theorem 5. Let $\{\eta_\gamma\}_\gamma$ be a well-spread family of non-negative symbols on $\mathbb{R}^{2d}$ such that $\sum_\gamma \eta_\gamma \approx 1$ and $\nu$ a $w$-moderated weight. Then, there exists a constant $\alpha > 0$ such that, for every choice of finite subsets of eigenfunctions $\{\phi_k^\gamma : \gamma \in \Gamma, 1 \leq k \leq N_\gamma\}$ with $\alpha \|\eta_\gamma\|_1 \leq N_\gamma \leq N < +\infty$, the following frame estimates hold simultaneously for all $1 \leq p \leq +\infty$, with the usual modification for $p = +\infty$:
$$\|f\|_{\mathcal{M}_p} \approx \left(\sum_\gamma \sum_{k=1}^{N_\gamma} |\langle f, \lambda_k^\gamma \phi_k^\gamma \rangle|^p \nu(\gamma)^p\right)^{1/p}.$$  
Moreover, $\alpha$ can be chosen uniformly for any class of weights $\nu$ having a uniform constant $C_\nu$ (cf. (20)).

Remark 10. As opposed to other constructions that partition either the time or frequency domain (see e.g. [22, 4, 9]), the symbols $\eta_\gamma$ partition the time-frequency plane simultaneously.

Remark 11. Note that, if $(\Gamma, \Theta, g)$ is an envelope for $\{\eta_\gamma\}_\gamma$, since $\|\eta_\gamma\|_1 \leq \|g\|_1$, in Theorem 5 it is always possible to make a uniform choice $N_\gamma = N_0$.

Remark 12 (Characteristic functions). When $\eta_\gamma$ is the characteristic function of a set $\Omega_\gamma$, the condition $\sum_\gamma \eta_\gamma \approx 1$ means that the family of sets $\{\Omega_\gamma : \gamma \in \Gamma\}$ is a cover of $\mathbb{R}^{2d}$ having bounded overlaps. The well-spreadness condition requires the sets $\Omega_\gamma$ have bounded measure and "eccentricity." For example, the family of characteristic functions is well-spread if for some $R > 0$, $\Omega_\gamma \subseteq B_R(\gamma)$, for all $\gamma$ belonging to a well-spread set $\Gamma$ (e.g. a lattice).

In Theorem 5, we pick $\approx |\Omega_\gamma|$ eigenfunctions from each time-frequency localization operator. This agrees with the uncertainty principle that says than each region of the time-frequency plane $\Omega_\gamma$ represents $\approx |\Omega_\gamma|$ degrees of freedom in signal space.
Proof of Theorem 5. For every $\gamma \in \Gamma$ and $\varepsilon > 0$, let $I^\varepsilon_\gamma := \{ k \in \mathbb{N} / \lambda^\gamma_k > \varepsilon \}$, which is a finite set. Using Theorem 4, Proposition 6 and the orthonormality of the eigenfunctions, we can find a value of $\varepsilon > 0$ such that

$$
\| f \|_{M^p_{\varepsilon}} \approx \left( \sum_{\gamma} \left( \sum_{k \in I^\varepsilon_\gamma} |(f, \lambda^\gamma_k \phi^\gamma_k)|^2 \right)^{p/2} v(\gamma)^p \right)^{1/p},
$$

which implies that for any choice of subsets of indices $J^\varepsilon_\gamma \supseteq I^\varepsilon_\gamma$ we also have,

$$
\| f \|_{M^p_{\varepsilon}} \approx \left( \sum_{\gamma} \left( \sum_{k \in J^\varepsilon_\gamma} |(f, \lambda^\gamma_k \phi^\gamma_k)|^2 \right)^{p/2} v(\gamma)^p \right)^{1/p}.
$$

Furthermore, due to (23) we have $\# I^\varepsilon_\gamma \leq \varepsilon^{-1} \sum_k \lambda^\gamma_k \approx \varepsilon^{-1} \| \eta \|_1$. Hence, setting $\alpha := \varepsilon^{-1}$, we ensure that for $N_\gamma \geq \alpha \| \eta \|_1$, the set $J^\varepsilon_\gamma := \{ k \in \mathbb{N} | 1 \leq k \leq N_\gamma \}$ contains $I^\varepsilon_\gamma$ and therefore satisfies (27). If in addition $\# J^\varepsilon_\gamma = N_\gamma \leq N < +\infty$, then

$$
(\sum_{k \in J^\varepsilon_\gamma} |(f, \lambda^\gamma_k \phi^\gamma_k)|^2)^{1/2} \approx (\sum_{k \in J^\varepsilon_\gamma} |(f, \lambda^\gamma_k \phi^\gamma_k)|^p)^{1/p},
$$

with a constant that depends on $N$. \qed

Finally we derive a variant of Theorem 5 where, under an inner regularity assumption on the family of symbols, we renormalize the frame of eigenfunctions so that each frame element has norm 1. To this end we first prove the following lemma which may be of independent interest.

Lemma 1. Let $h \in L^2(\mathbb{R}^d)$ and let $\Omega \subseteq \mathbb{R}^d$ have non-empty interior. Then the time-frequency localization operator $H^h_{\Omega} : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$,

$$
H^h_{\Omega}f(t) = \int_{\Omega} V_h f(x, \xi) h(t-x)e^{2\pi i t \cdot \xi} dx \, d\xi, \quad (t \in \mathbb{R}^d).
$$

has infinite rank.

Proof. The proof is based on the fact that the short-time Fourier transform of Hermite functions are weighted polyanalytic functions (cf. [2, 1, 3]) and therefore cannot vanish on a ball.

Suppose, for the sake of contradiction, that $H^h_{\Omega}$ has rank $n - 1 < +\infty$. Let $h_1, \ldots, h_n \in L^2(\mathbb{R}^d)$ be multi-dimensional Hermite functions of order $\leq n$. For example, if $g_1, \ldots, g_n \subseteq L^2(\mathbb{R})$ are the first one-dimensional Hermite functions, we can let $h_k \in L^2(\mathbb{R}^d)$ be the tensor product $h_k(x_1, \ldots, x_d) := g_k(x_1)g_1(x_2) \ldots g_1(x_d)$. Let $S_n$ be the subspace of $L^2(\mathbb{R}^d)$ spanned by $h_1, \ldots, h_n$.

Since $S_n$ has dimension $n$, it follows that there exists some nonzero $f = \sum_{k=1}^n c_k h_k \in S_n$ such that $H^h_{\Omega}f = 0$. Consequently,

$$
0 = \langle H^h_{\Omega}f, f \rangle = \int_{\Omega} |V_h f(z)|^2 \, dz,
$$

and therefore $V_h f \equiv 0$ on $\Omega$. With the notation $(x, w) := (z \in \mathbb{C}^n$ and $m(z) = e^{-z \cdot w + \pi |z|^2}$, let $F(z) := m(z) V_f h(z)$. Since $V_h f(z) = \overline{V_f h(-z)}$, it follows that $F$ vanishes on $\Omega^c := -\overline{\Omega}$. We will show that $F \equiv 0$. Since $m$ never vanishes, this will imply that $f \equiv 0$, thus yielding a contradiction.

The function $F(z) := \sum_k c_k m(z) V_h h(-z)$ is a polyanalytic function of order (at most) $n$ (i.e. $(\partial^{|\alpha|}/\partial z^\alpha)F \equiv 0$ [2, 1, 3]). A polyanalytic function that vanishes on a set of non-empty interior must vanish identically. For $d = 1$ this can be proved directly by induction on $n$ or deduced from much sharper uniqueness results (see [5]). The case of general dimension $d$ reduces to $d = 1$ by fixing $d - 1$ variables of $F$ and applying the one-dimensional result. \qed
We now obtain a variant of Theorem 5.

**Theorem 6.** Let \( \{\eta_\gamma\}_\gamma \) be a well-spread family of non-negative symbols on \( \mathbb{R}^{2d} \) such that \( \sum_\gamma \eta_\gamma \approx 1 \) and \( v \) a \( w \)-moderated weight. Assume in addition that there exists a ball \( B_r \) and a constant \( c > 0 \) such that

\[
\eta_\gamma(z) \geq c \chi_{B_r}(z - \gamma) \geq 0 \quad (z \in \mathbb{R}^{2d}, \gamma \in \Gamma).
\]

Then, there exists a constant \( \alpha > 0 \) such that, for every choice of finite subsets of eigenfunctions \( \{\phi_k^\gamma | \gamma \in \Gamma, 1 \leq k \leq N_\gamma \} \) with \( \alpha \|\eta_\gamma\|_1 \leq N_\gamma \leq N < +\infty \), the following frame estimates hold simultaneously for all \( 1 \leq p \leq +\infty \), with the usual modification for \( p = \infty \):

\[
\|f\|_{M^p} \approx \left( \sum_{\gamma} \sum_{k=1}^{N_\gamma} |(f, \phi_k^\gamma)|^p v(\gamma)^p \right)^{1/p}.
\]

Moreover, \( \alpha \) can be chosen uniformly for any class of weights \( v \) having a uniform constant \( C_v \) (cf. (20)).

**Remark 13.** If \( \{\Omega_\gamma : \gamma \in \Gamma\} \) satisfies (3) and covers \( \mathbb{R}^{2d} \) with a bounded number of overlaps, then the corresponding family of characteristic functions \( \eta_\gamma \) meets the conditions of Theorem 6 (see also Remark 12). This shows that Theorem 1, in Section 1, is a particular case of Theorem 6.

**Remark 14.** The frame in Theorem 6 comprises the first \( N_\gamma \) elements of each of the orthonormal sets \( \{\phi_k^\gamma : k \geq 1\} \). These first \( n_\gamma \) functions are the ones that are best concentrated, according to the weight \( n_\gamma \). This resembles the problem studied in [42]. However, the results there require very precise information on the frames being pieced together and hence do not apply here.

**Remark 15.** In the language of [40, 10], this shows that the subspaces spanned by the finite families of eigenfunctions form a stable splitting or fusion frame. From an application point of view, it is useful to dispose of orthogonal projections onto subspaces with time-frequency concentration in a prescribed area of the time-frequency plane.

**Proof of Theorem 6.** By Theorem 5, we have that,

\[
\|f\|_{M^p} \approx \left( \sum_{\gamma} \sum_{k=1}^{N_\gamma} |(f, \lambda_k^\gamma \phi_k^\gamma)|^p v(\gamma)^p \right)^{1/p}.
\]

Hence, it suffices to show that \( \lambda_k^\gamma \approx 1 \), for \( 1 \leq k \leq N_\gamma \).

The upper bound follows from the well-spreadness condition. If \( (\Gamma, \Theta, g) \) is an envelope for \( \{H_{\psi_\gamma} : \gamma \in \Gamma\} \), then all the singular values of \( H_{\psi_\gamma} \) are bounded by \( \|H_{\psi_\gamma}\|_{2 \rightarrow 2} \leq \|g\|_\infty \), cf. (16).

By Lemma 1, the localization operator \( H_{\chi_{B_r}} \) has infinite rank. Hence, the non-zero eigenvalues of \( H_{\chi_{B_r}} \) form an infinite non-increasing sequence \( \lambda_k^R > 0, k \geq 1 \). From (28) it follows that \( H_{\eta_\gamma} \geq c H_{\chi_{B_r}} \) (cf. Prop. 3) and consequently, for \( 1 \leq k \leq N_\gamma \), \( \lambda_k^\gamma \geq c \lambda_k^R \geq c \lambda_k^R \geq 0 \).

**Remark 16.** For the results in this section, the abstract setting of Section 2 allows for the replacement of \( \ell^p \) by more general normed spaces. Indeed, the results derived in the abstract setting cover modulation spaces defined with respect to general translation-invariant solid spaces, cf. [23].

### 5. Frames of eigenfunctions: general \( L^2 \) estimates

The arguments presented in Section 4.1 make use of the general almost-orthogonality principle in Theorem 2 and carry over to the Banach-space setting by exploiting spectral-invariance results for pseudo-differential operators. In the abstract setting of Section 3 we can still carry out the proofs of Section 4.1 to obtain \( L^2 \) estimates. We briefly present these results here. The proofs are, mutatis mutandis, the same as in Section 4 and will be just sketched.
Let $E = L^2(G)$ and let us assume that (A1), (A2) and (A3) from Section 2.2 hold. Let a well-spread family $\{\eta_\gamma\}_\gamma$ of non-negative functions on $G$ be given. We consider the corresponding family of phase-space multipliers,

$$M_\gamma f = P(\eta_\gamma f), \quad (f \in S_{L^2}).$$

Since each $\eta_\gamma$ is non-negative and belongs to $L^1(G)$, according to Proposition 3, the corresponding operator $M_\gamma : S_{L^2} \to S_{L^2}$ is positive and trace-class, and $\text{trace}(M_\gamma) \lesssim \|\eta_\gamma\|_1$.

Let $M_\gamma : S_{L^2} \to S_{L^2}$ be diagonalized as $M_\gamma f = \sum_k \lambda_k^2 \langle f, \phi_k^\gamma \rangle \phi_k^\gamma$, where $\{\phi_k^\gamma \mid k \geq 1\}$ is an orthonormal subset of $S_{L^2}$ and define $M_\gamma f = \sum_{k: \lambda_k^2 \geq \varepsilon} \lambda_k^2 \langle f, \phi_k^\gamma \rangle \phi_k^\gamma$. When $M_\gamma$ has finite rank, $\lambda_k^2 = 0$, for $k >> 1$ and the choice of the corresponding eigenfunction is arbitrary.

We have the following version of the Theorems 4 and 5.

**Theorem 7.** Let $\{\eta_\gamma\}_\gamma$ be a well-spread family of non-negative symbols such that $\sum_\gamma \eta_\gamma \approx 1$. Then, there exists constants $0 < c \leq C < +\infty$ such that for all sufficiently small $\varepsilon > 0$,

$$c\|f\|_2^2 \leq \sum_\gamma \|M_\gamma f\|_2^2 \leq C\|f\|_2^2, \quad (f \in S_{L^2}).$$

Furthermore, there exists a constant $\alpha > 0$ such that, for every choice of $\alpha\|\eta_\gamma\|_1 \leq N_\gamma \leq N < +\infty$, the family

$$\{ \lambda_k^2 \phi_k^\gamma \mid \gamma \in \Gamma, 1 \leq k \leq N_\gamma \},$$

is a frame of $S_{L^2}$.

**Remark 17.** Note again that when $\eta_\gamma$ is the characteristic function of a set $\Omega_\gamma$, we are picking $\approx |\Omega_\gamma|$ eigenfunctions from each localization operator (phase-space multiplier). Here, $|\Omega_\gamma|$ is the Haar measure of $\Omega_\gamma$.

**Remark 18.** The operator $M_\gamma$ may have finite rank (for example if $G$ is a discrete group and $\eta_\gamma$ is the characteristic function of a finite set). In this case the choice of the eigenfunctions associated to the singular value is irrelevant, since in (30) these are multiplied by zero.

**Proof of Theorem 7.** This parallels the proofs in Section 4.1 and wraps up their main steps. Since $\sum_\gamma M_\gamma = M_{\eta},$ and $\sum_\gamma \eta_\gamma \geq A > 0$, it follows from Prop. 3 that $\sum_\gamma M_\gamma$ is positive definite and therefore invertible. Theorem 2 consequently yields,

$$\sum_\gamma \|M_\gamma f\|_{L^2(G)}^2 \approx \|f\|_2^2, \quad (f \in S_{L^2}).$$

In addition, Proposition 5, Corollary 1 and a second application of Theorem 2 yield,

$$\sum_\gamma \|M_\gamma^2 f\|_{L^2(G)}^2 \approx \|f\|_2^2, \quad (f \in S_{L^2}).$$

The thresholded operators $M_\gamma^\varepsilon$ satisfy,

$$\|M_\gamma^\varepsilon f\|_2 \leq \|M_\gamma f\|_2 \leq \|M_\gamma^\varepsilon f\|_2 + \varepsilon\|f\|_2, \quad (f \in S_{L^2}).$$

Applying this to $M_\gamma f$ and noting that $M_\gamma$ and $M_\gamma^\varepsilon$ commute gives,

$$\|M_\gamma^\varepsilon f\|_2 \leq \|M_\gamma^\varepsilon M_\gamma f\|_2 + \varepsilon\|M_\gamma f\|_2, \quad (f \in S_{L^2}).$$

Putting all these inequalities together gives,

$$\left(\sum_\gamma \|M_\gamma^\varepsilon f\|_2\right)^{1/2} \lesssim \|f\|_2 \lesssim \left(\sum_\gamma \|M_\gamma^\varepsilon f\|_2\right)^{1/2} + \varepsilon\|f\|_2, \quad (f \in S_{L^2}).$$

This implies that for $0 < \varepsilon << 1$,

$$\left(\sum_\gamma \|M_\gamma^\varepsilon f\|_2\right)^{1/2} \approx \|f\|_2,$$
as claimed. The fact that the system in (30) is a frame of $S_{l^2}$ now follows like in Theorem 5, this time using Proposition 3 to estimate,
\[
\# \{ \gamma_k^\varphi : \gamma_k^\varphi > \varepsilon \} \leq \varepsilon^{-1} \text{trace}(M_{\eta_\varepsilon}) \lesssim \varepsilon^{-1} \|\eta_\varepsilon\|_1.
\]
\]
\]

5.1. Application to time-scale analysis. We now show how to apply Theorem 7 to time-scale analysis. Let $\psi : \mathbb{R}^d \to \mathbb{C}$ be a Schwartz-class radial function with several vanishing moments. The wavelet transform of a function $f \in L^2(\mathbb{R}^d)$ with respect to $\psi$ is defined by,
\[
W_{\psi} f(s, x) = s^{-d/2} \int_{\mathbb{R}^d} f(t) \psi \left( \frac{t - x}{s} \right) dt, \quad (x \in \mathbb{R}^d, s > 0).
\]
If $\psi$ is properly normalized (and we assume so thereof), $W_{\psi}$ maps $L^2(\mathbb{R}^d)$ isometrically into $L^2(\mathbb{R}^d \times \mathbb{R}_+, s^{-(d+1)} dx ds)$. For a (measurable) bounded symbol $m : \mathbb{R}^d \times \mathbb{R}_+ \to \mathbb{C}$, the wavelet multiplier $W_{m \psi} : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ is defined by
\[
W_{m \psi} f(t) = \int_{0}^{+\infty} \int_{\mathbb{R}^d} m(x, s) W_{\psi} f(x, s) \pi(x, s) \psi(t) ds \frac{ds}{s^{d+1}},
\]
where $\pi(x, s) \psi(t) := s^{-d/2} \psi \left( \frac{x - t}{s} \right)$. (Here, the integral converges in the weak sense.) Note that $W_{m \psi} = W_m (W_{\psi} F)$.

In order to apply the model from Section 2.2 we consider the affine group $G = \mathbb{R}^d \times \mathbb{R}_+$, where multiplication is given by $(x, s) \cdot (x', s') = (x + sx', ss')$. We let $E := L^2(G)$ and $S_E := W_{\psi} L^2(\mathbb{R}^d)$. In complete analogy to the time-frequency analysis case, we let $P := W_{\psi}(W_{\psi})^* : L^2(G) \to W_{\psi} L^2(\mathbb{R}^d)$ be the orthogonal projection and $K := |W_{\psi}|$. We further let $w(x, s) := \max \{1, s^d\}$. The kernel $K$ belongs to $W(L^\infty, L_w^1) \cap W_R(L^\infty, L_w^1)$ if $\psi$ has sufficiently many vanishing moments (see [34, Section 4.2]).

As an example of a well-spread family of symbols we consider the characteristic functions of a cover of $G$ by irregular boxes. Let us take as centers the points
\[
\Gamma := \{ \gamma_{j,k} := (k2^j, 2^j) : j \in \mathbb{Z}, k \in \mathbb{Z}^d \},
\]
and consider a family of boxes around $(0, 1) \in \mathbb{R}^d \times \mathbb{R}_+$,
\[
V_{j,k} := [-a_{j,k}^1/2, a_{j,k}^1/2] \times \ldots \times [-a_{j,k}^d/2, a_{j,k}^d/2] \times [(b_{j,k})^{-1}, b_{j,k}],
\]
where $0 \leq a_{j,k}^l \leq a < +\infty$, $l = 1, \ldots, d$ and $0 < b^{-1} \leq b_{j,k} \leq b < +\infty$. Let us set $U_{k,j} := \gamma_{j,k} V_{j,k}$. Then the family of characteristic functions $\{ \chi_{U_{k,j}} \}_{j,k}$ is well-spread. The corresponding operators $W_{M_{k,j}} := W_{M_k} W_{U_{k,j}}$ are known as wavelet localization operators [14, 16, 15]. Let us denote again by $\{ \phi_k^\varphi \}_{k \geq 1}$ an eigenset of $W_{M_{k,j}}$ ordered according to the corresponding eigenvalues $\gamma_k^\varphi$.

Noting that $\|\chi_{k,j}\|_1 = |U_{k,j}| = |V_{k,j}| = \frac{1}{d} \prod_{l=1}^{d} a_{j,k}^l \cdot [(b_{j,k})^d - (b_{j,k})^{-d}]$, Theorem 7 yields,
\[
\text{Theorem 8. Assume that } \{ U_{j,k} \} \text{ covers } \mathbb{R}^d \times \mathbb{R}_+. \text{ Then there exists a constant } \alpha > 0 \text{ such for every choice of } \alpha \text{ and } N, N_\varepsilon \leq N < +\infty, \text{ the family}
\{
\gamma_k^\varphi \phi_k^\varphi \mid \gamma \in \Gamma, 1 \leq k \leq N_\varepsilon \},
\]
is a frame of $L^2(\mathbb{R}^d)$.

5.2. Application to Gabor analysis. We now consider as $G$ discrete subgroups of $\mathbb{R}^{2d}$, more precisely, lattices of the form $\Lambda = P \mathbb{Z}^{2d}$, where $P \in \mathbb{R}^{2d \times 2d}$ is an invertible matrix. This choice leads to Gabor analysis and in particular to the setting of Gabor frames. Here, the corresponding phase-space multipliers are Gabor multipliers, [25, 19, 33]. We now show how to apply the model from Section 2. This is completely analogous to the treatment of continuous time-frequency expansions in Section 4.

Given a lattice $\Lambda \subseteq \mathbb{R}^{2d}$ and a window $\varphi \in M^1(\mathbb{R}^d)$ with $\|\varphi\|_2 = 1$ we consider the Gabor system
\[
F(\varphi, \Lambda) = \{ (\pi(\lambda) \varphi, \lambda \in \Lambda, \pi(\lambda) \varphi(t) = \}
\[
\]

\( \varphi(t - \lambda_1)e^{2\pi i \lambda_2 t} \). We assume that \( F(\varphi, \Lambda) \) is a tight frame for \( L^2(\mathbb{R}^d) \). This means that for some constant \( A > 0 \), every function \( f \in L^2(\mathbb{R}^d) \) has an expansion,

\[
 f = A \sum_{\lambda \in \Lambda} (f, \pi(\lambda) \varphi) \pi(\lambda) \varphi.
\]

The corresponding analysis operator \( \mathcal{V}_{\Lambda, \varphi} : L^2(\mathbb{R}^d) \rightarrow \ell^2(\Lambda) \) is given by \( \mathcal{V}_{\Lambda, \varphi} f(\lambda) = \sqrt{A} (f, \pi(\lambda) \varphi) \). Let \( S_{\ell^2} = \mathcal{V}_{\Lambda, \varphi}(L^2(\mathbb{R}^d)) \). The orthogonal projection \( P : \ell^2(\Lambda) \rightarrow S_{\ell^2} \) is \( P = \mathcal{V}_{\Lambda, \varphi} \mathcal{V}_{\Lambda, \varphi}^* \) and is therefore represented by the matrix \( \kappa(\mu, \lambda) = A (\pi(\mu) \varphi, \pi(\lambda) \varphi) \). Consequently,

\[
 |\kappa(\mu, \lambda)| = A |(\pi(\mu) \varphi, \pi(\lambda) \varphi)| = A |\mathcal{V}_{\varphi} \varphi(\mu - \lambda)|, \quad (\mu, \lambda \in \Lambda).
\]

Since \( \varphi \in M^1(\mathbb{R}^d) \), \( \mathcal{V}_{\Lambda, \varphi} \) maps \( M^1(\mathbb{R}^d) \) into \( \ell^2(\Lambda) \) (see for example [26]) and we conclude that \( K := |\mathcal{V}_{\varphi} \varphi| \in \ell^1(\Lambda) \). Hence, (A1), (A2) and (A3) from Section 2 are satisfied with \( \mathcal{G} = \Lambda \), \( E = \ell^2 \) and \( w \equiv 1 \).

For a bounded sequence \( m : \Lambda \rightarrow C \), the Gabor multiplier \( GM_m : L^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{R}^d) \) is defined by applying \( m \) to the frame expansion

\[
 GM_m f = A \sum_{\lambda} m(\lambda) (f, \pi(\lambda) \varphi) \pi(\lambda) \varphi.
\]

Hence, \( GM_m f = \mathcal{V}_{\varphi}^* (Am \mathcal{V}_{\varphi} f) \) and \( \mathcal{V}_{\varphi}GM_{\varphi} \mathcal{V}_{\varphi}^* : S_{\ell^2} \rightarrow S_{\ell^2} \) is a phase-space multiplier with symbol \( Am \).

We may now consider a well-spread family \( \{ \eta_\gamma \}_\gamma \) of non-negative symbols defined on \( \mathbb{R}^{2d} \) with \( \sum_\gamma \eta_\gamma \approx 1 \) and restrict each \( \eta_\gamma \) to \( \Lambda \). Then, for \( \gamma \in \Gamma \), we let \( \{ \phi_\gamma^k : k \geq 1 \} \) be the set of eigenfunctions of \( GM_{\eta_\gamma} \), in decreasing order with respect to the corresponding eigenvalues \( \lambda_\gamma^k \). Then, Theorem 7 yields the following result.

**Theorem 9.** There exists a constant \( \alpha > 0 \) such for every choice of \( N_\gamma \geq 1 \) for which

\[
 \alpha \| \eta_\gamma \|_1 \leq N_\gamma \leq N < +\infty,
\]

the family \( \{ \lambda_\gamma^k \phi_\gamma^k \mid \gamma \in \Gamma, 1 \leq k \leq N_\gamma \} \) is a frame of \( L^2(\mathbb{R}^d) \).

**Appendix: Proof of Theorem 2**

In this appendix we prove Theorem 2. The proof is essentially contained in [43], but is not explicitly stated in the required generality. We therefore show how to derive Theorem 2 from some technical lemmas in [43].

Suppose that Assumptions (A1) and (A2) from Section 2.2 hold. Let \( \{ T_\gamma \mid \gamma \in \Gamma \} \) be a well-spread family of operators with envelope \( (\Gamma, \Theta, g) \). According to the definition, \( g \in W_R^1(L^\infty, L^1_w)(\mathcal{G}) \). The article [43] considers a technical variant of this space, the weak amalgam space \( W_R^{\text{weak}}(L^\infty, L^1_w) \) (see [43, Section 2.4]), which we do not wish to introduce here. By [43, Lemma Prop. 1], \( L^1(\mathcal{G}) \) \( \hookrightarrow \) \( W_R^{\text{weak}}(L^\infty, L^1_w)(\mathcal{G}) \) \( \hookrightarrow \) \( W_R(\ell^\infty, L^1_w)(\mathcal{G}) \). We will use certain results from [43] that are proved under the thus more general hypothesis \( g \in W_R^{\text{weak}}(L^\infty, L^1_w)(\mathcal{G}) \). Nonetheless we point out that, according to [43, Lemma Prop. 1], for certain groups \( W_R^{\text{weak}}(L^\infty, L^1_w)(\mathcal{G}) \) is just \( L^1(\mathcal{G}) \). This is the case for example for \( \mathcal{G} = \mathbb{R}^d \). Hence, in Section 4, the hypothesis \( g \in W(\ell^\infty, L^1_w) \) could be relaxed to \( g \in L^1 \).

We consider an \( L^2 \)-valued version of \( E_d(\Gamma) \),

\[
 E_{d,L^2} = E_{d,L^2}(\Gamma) := \{ f : \Gamma \rightarrow L^2(\mathcal{G}) \mid \| f_\gamma \|_{L^2} \gamma \in \Gamma \in E_d(\Gamma) \},
\]

and endow it with the norm \( \| f_\gamma \|_{E_{d,L^2}} := \| f_\gamma \|_{L^2} \gamma \in \Gamma \| \in E_d \). Let \( U \subseteq \mathcal{G} \) be a relatively compact neighborhood of the identity. Consider the operators \( C_T \) and \( S_U \) formally defined by

\[
 C_T(f) := (T_\gamma(f))_\gamma \in \Gamma, \quad (f \in S_E),
\]

\[
 S_U(f_\gamma) := \sum_\gamma P(f_\gamma) \chi_{\gamma U}, \quad (f_\gamma \in L^2(\mathcal{G})),
\]

where \( \chi_{\gamma U} \) denotes the characteristic function of the set \( \gamma U \). These operators satisfy the following mapping properties.

**Proposition 7.**

(a) The analysis operator \( C_T \) maps \( S_E \) boundedly into \( E_{d,L^2}(\Gamma) \).
(b) For every relatively compact neighborhood of the identity \( U \), and every sequence \( F \equiv (f_\gamma) \in E_{d,L^2} \), the series defining \( S_U(F) \) converge absolutely in \( L^2(\mathcal{G}) \) at every point. Moreover, the operator \( S_U \) maps \( E_{d,L^2}(\Gamma) \) boundedly into \( E \) (with a bound that depends on \( U \)).

**Proof.** Part (b) is proved in [43, Prop. 4 (b)]. Part (a) is a slight variant of [43, Prop. 4 (a)]; for completeness we give a full argument.

Let \( f \in S_E \). Since \( \eta_\gamma \) is bounded, \( f\eta_\gamma \in E \). By the definition of well-spread family (cf. (17)),

\[
|T_\gamma f(x)| \leq \int_\mathcal{G} |f(y)| g(\gamma^{-1}y)\Theta(y^{-1}x)dy
= (|f| L_2) \Theta(x).
\]

By Young’s inequality \( L^1 \ast L^2 \rightarrow L^2 \) we have,

\[
\|T_\gamma f\|_2 \leq \|\Theta\|_2 \int_\mathcal{G} |f(y)| g(\gamma^{-1}y)dy \lesssim \|\Theta\|_{W(L^\infty, L^1)} \int_\mathcal{G} |f(y)| g(\gamma^{-1}y)dy.
\]

Now the solidity of \( E \) and [43, Lemma 4] (see also [23, Lemma 3.8]) yield,

\[
\|C_T(f)\|_{E_{d,L^2}} \lesssim \|f\|_{W(L^\infty, E)} \|T\|_{W(L^\infty, L^1)}.
\]

Finally, by Prop. 2, \( f \|W(L^\infty, E) \gtrsim \|f\|_E. \)

**Remark 19.** Note that in the last proof the use of the \( L^2 \) norm is somehow arbitrary; a number of other function norms could have been used instead (cf. [43, Prop. 4]).

Now we prove the key approximation result (cf. [43, Theorem 1]).

**Theorem 10.** Given \( \varepsilon > 0 \), there exists \( U_0 \), a relatively compact neighborhood of \( \varepsilon \) such that for all \( U \supseteq U_0 \),

\[
(39) \tag{39} \| \sum_\gamma T_\gamma f - S_U C_T(f) \|_E \leq \varepsilon \|f\|_E, \quad \text{if } f \in S_E.
\]

**Remark 20.** The neighborhood \( U_0 \) can be chosen uniformly for any class of spaces \( E \) having the same weight \( w \) and the same constant \( C_{E,w} \) (cf. (13)).

Concerning the parameters in Assumptions (A1) and (A2) and (17), the choice of \( U_0 \) only depends on \( \|K\|_{W(L^\infty, L^1)}, \|\Theta\|_{W(L^\infty, L^1)} \), \( \|T\|_{W(L^\infty, L^1)} \), \( \|\Theta\|_{W(L^\infty, L^1)} \), \( \|g\|_{W(L^\infty, L^1)} \), and \( \rho(\Gamma) \) (cf. (9)).

**Proof.** Let \( f \in S_E \) and let \( U \) be a relatively compact neighborhood of \( \varepsilon \). Because of the inclusion \( S_E \hookrightarrow W(L^\infty, E) \) in Prop. 2, it suffices to dominate the left-hand side of (39) by \( \varepsilon \|f\|_{W(L^\infty, E)} \).

Note that since \( T_\gamma f \in S_E, S_U C_T(f)(x) = \sum_\gamma T_\gamma f(x)(\chi_\gamma U)(x) \). Hence, using (17) let us estimate,

\[
\left| \sum_\gamma T_\gamma f(x) - S_U C_T(f)(x) \right| \leq \sum_\gamma \left| \chi_\gamma U(x) T_\gamma f(x) \right| \leq \sum_\gamma \int_\mathcal{G} |f(y)| g(\gamma^{-1}y)\Theta(y^{-1}x) \chi_\gamma U(y)dy.
\]

The rest of the proof is carried out exactly as in [43, Theorem 1]. Indeed, the proof there only depends on the estimate just derived.4 The definition of well-spread family of operators was tailored so that the proof in [43, Theorem 1] would still work. \qed

Finally we can prove Theorem 2.

**Proof of Theorem 2.** Let \( \{T_\gamma : \gamma \in \Gamma\} \) be a well-spread family of operators and suppose that the operator \( \sum_\gamma T_\gamma : S_E \rightarrow S_E \) is invertible. We have to show that for \( f \in S_E \), \( \|f\|_E \approx \|C_T(f)\|_{E_{d,L^2}(\Gamma)} \).

The estimate \( \|C_T(f)\|_{E_{d,L^2}(\Gamma)} \lesssim \|f\|_E \) is proved in Proposition 7 (a). To establish the second inequality, consider the operator \( P S_U C_T : S_E \rightarrow S_E \). Then for \( f \in S_E \),

\[
\| \sum_\gamma T_\gamma f - P S_U C_T(f) \|_E = \| P \sum_\gamma T_\gamma f - P S_U C_T(f) \|_E \lesssim \| \sum_\gamma T_\gamma f - S_U C_T(f) \|_E.
\]

4The function \( \Theta \) is called \( H \) in the proof [43, Theorem 1].
This estimate, together with Theorem 10 implies that \(\|\sum_{\gamma} T_{\gamma} - PSU_{C_T}\|_{SE \to SE} \to 0\) as \(U\) grows to \(G\). Hence, there exists \(U\) such that \(PSU_{C_T}\) is invertible on \(SE\). Consequently, for \(f \in SE\), \(\|f\|_E \approx \|PSU_{C_T}f\|_E \leq \|C_T(f)\|_{E_{\ell_2}(\Gamma)}\). Here we have used the boundedness of \(SU\) - contained in Proposition 7 (b) - and the boundedness of \(P : E \to W(L^\infty, E) \leftrightarrow E\) - contained in Proposition 2. □

References


E-mail address: monika.doerfler@univie.ac.at

E-mail address: jose.luis.romero@univie.ac.at

Faculty of Mathematics, University of Vienna, Nordbergstrasse 15,A-1090 Wien, Austria
An inverse problem for localization operators

Luís Daniel Abreu$^{1,2}$ and Monika Dörfler$^1$

$^1$ Institut für Mathematik, Universität Wien, Alserbachstrasse 23 A-1090 Wien, Austria
$^2$ CMUC, Department of Mathematics, University of Coimbra, Portugal

E-mail: daniel@mat.uc.pt and monika.doerfler@univie.ac.at

Received 5 February 2012, in final form 8 July 2012
Published 28 September 2012
Online at stacks.iop.org/IP/28/115001

Abstract
A classical result of time–frequency analysis, obtained by Daubechies in 1988, states that the eigenfunctions of a time–frequency localization operator with circular localization domain and Gaussian analysis window are the Hermite functions. In this contribution, a converse of Daubechies’ theorem is proved. More precisely, it is shown that, for simply connected localization domains, if one of the eigenfunctions of a time–frequency localization operator with Gaussian window is a Hermite function, then its localization domain is a disc. The general problem of obtaining, from some knowledge of its eigenfunctions, information about the symbol of a time–frequency localization operator is denoted as the inverse problem, and the problem studied by Daubechies as the direct problem of time–frequency analysis. Here, we also solve the corresponding problem for wavelet localization, providing the inverse problem analogue of the direct problem studied by Daubechies and Paul.

1. Introduction

Most real-life signals of interest change their frequency properties over time. Therefore, a signal description by means of time–frequency analysis is often preferable to the signal’s Fourier transform, which reliably yields frequency information, but without any localization in time. The core purpose of time–frequency analysis is to represent a given signal as a function in the time–frequency or in the time-scale plane. However, in real world applications like optics and wireless communications, one can only ‘sense’ a signal within a certain region of those planes. This means that, in practice, the part of the signal outside the region of interest is neglected and only its ‘localized’ version is observed. Localization operators turn this observation process into rigorous mathematical terms. They transform a given signal into one that is localized in a given region by reducing the signal energy outside that region to a negligible amount.

The first approach to time–frequency localization, introduced in 1961, consists in separately selecting time- and frequency-content, and is described in a famous series of papers known as the ‘Bell labs papers’. We refer to Slepian’s review [20] for an account of this beautiful body of work. In 1988, Daubechies added a new perspective by introducing
operators, that localize directly in the time–frequency plane [4] and, together with Paul [5], extended the analysis to the time-scale plane. The time–frequency plane is associated with the short-time Fourier transform and the time-scale plane is associated with the wavelet transform. We begin our presentation by defining the short-time Fourier transform, which leads to the concept of time–frequency localization operators.

The short-time Fourier (or Gabor) transform of a function or distribution \( f \) with respect to a window function \( g \in L^2(\mathbb{R}) \) is defined to be, for \( z = (x, \xi) \in \mathbb{R}^2 \),

\[
\mathcal{V}_g f (z) = \mathcal{V}_g f (x, \xi) = \int_{\mathbb{R}} f(t) \overline{g(t - x)} \, e^{-2\pi i \xi t} \, dt,
\]

where the overline denotes complex conjugation. We let \( \pi (z) g(t) = g(t - x) \, e^{2\pi i \xi t} \) and observe that \( f \) can be resynthesized from \( \mathcal{V}_g f \) as

\[
f = \frac{1}{\|g\|_2^2} \int_{\mathbb{R}^2} \mathcal{V}_g f (z) \pi (z) g \, dz.
\]

Given a symbol \( \sigma \in L^1(\mathbb{R}^2) \), time–frequency localization operators \( H_{\sigma, g} \) are defined by

\[
H_{\sigma, g} f = \int_{\mathbb{R}^2} \sigma(z) \mathcal{V}_g f (z) \pi (z) g \, dz = \mathcal{V}_{\sigma g} f.
\]

In signal processing, it is very common to modify a signal \( f \) by acting on its time–frequency coefficients \( \mathcal{V}_g f \), for example, in order to achieve noise reduction [14]; the corresponding localization operators have been the object of research in time–frequency analysis, [7, 3]. In [4], Daubechies considered the window \( g(t) = \varphi(t) = 2^j \, e^{-\pi t^2} \), the symbol \( \sigma(z) = \chi_\Omega(z) \), i.e. the indicator function of a set \( \Omega \subset \mathbb{R}^2 \), and investigated the eigenvalue problem

\[
H_\Omega f := H_{\chi_\Omega, g} f = \lambda f
\]

for the case where \( \Omega \) is a disc centered at zero. She concluded that in this situation, the eigenfunctions of \( H_{\chi_\Omega} \) are the Hermite functions. Consequently, since, \( H_{\Omega_1 \setminus \Omega_2} = H_\Omega - H_{\Omega_2} \), for two sets \( \Omega_1 \subset \Omega_2 \), the Hermite functions are also eigenfunctions with respect to domains in the form of an annulus centered at zero and for any union of annuli.

Problem (2) is important in time–frequency analysis, because its solutions are the functions with best concentration in the subregion \( \Omega \) of the time–frequency plane, where we consider the time–frequency concentration of a function \( f \) in \( \Omega \subset \mathbb{R}^2 \) defined as

\[
C_\Omega(f) = \frac{\int_{\Omega} |\mathcal{V}_g f(z)|^2 \, dz}{\|f\|_2^2}.
\]

In this paper, we will be concerned with the inverse situation of the one considered by Daubechies. This leads us to the following question.

• Given a localization operator with unknown localization domain \( \Omega \), can we recover the shape of \( \Omega \) from information about its eigenfunctions?

This is a new type of inverse problem, and we will call it the inverse problem of time–frequency localization. We solve the problem in the case where explicit computations can be made, which is the set-up of [4]. Our main contribution is the following.

**Theorem 1.** Let \( \Omega \subset \mathbb{R}^2 \) be simply connected. If one of the eigenfunctions of the localization operator \( H_\Omega \) is a Hermite function, then \( \Omega \) must be a disc centered at 0.

Let us briefly discuss some motivations for our studies and consequences of our result. Hermite functions have been proposed as modulating pulses in pulse-shape modulation for ultra-wideband (UWB) communication, mainly due to their maximal joint concentration in time and frequency, cf [19, 15, 9] and references therein. The receiver of the communication
system often applies a filter with the modulation pulses as eigenfunctions corresponding to large eigenvalues, in order to suppress random noise that accumulates during transmission. In this situation, theorem 1 shows that the filter must be designed with a circular localization domain. Furthermore, if the filter on the receiver side is known or designed to have one single Hermite function as an eigenfunction, it is possible to guarantee the location of the time–frequency plane that the filter is sensing. This is particularly important in UWB communication, where the permitted spectrum is officially prescribed, cp [6].

This last remark also hints at an additional possible application, which is system identification. Identification of linear time-variant systems is a notoriously difficult task in general, cp [13, 2, 17]. While a linear time-invariant system is straightforwardly identified by sending an impulse to the system and retrieving the impulse response, [10, section 4.2], no similar method exists for general linear time-variant systems. By our result, and for a system that is known to be a localization operator of the form \( H_\Omega \) for some time–frequency region \( \Omega \), we may send any Hermite function to the system and judge from the response whether \( \Omega \) can be a disc. In the positive case, one can then evaluate the size of the disc by subsequently sending additional Hermite functions and evaluating the resulting scaling factors. Obviously, this approach should be extended to other shapes and its feasibility will be the topic of further research.

In analogy with theorem 1, we will also consider an inverse problem for wavelet localization operators. Here, we show that the domain of localization of the localization operators investigated by Daubechies and Paul [5] is a pseudohyperbolic disc in the upper-half plane whenever one of the operator’s eigenfunctions is the Fourier transform of a Laguerre function. We will essentially use methods from complex analysis and our techniques are strongly influenced by the ideas contained in [1] and [18].

This paper is organized as follows. Section 2.1 collects some properties of the eigenfunctions of localization operators with respect to radially weighted measures and section 2.2 deduces the geometry of localization domains under the assumption of orthogonality of any single monomial to almost all monomials. The corresponding inverse problem for Gabor localization is studied in section 3, and section 4 is devoted to the investigation of the inverse problem for wavelet localization.

2. Double orthogonality and reproducing kernel Hilbert spaces

This section is devoted to the properties of complex monomials, namely their double orthogonality with respect to any radially weighted measures and the consequences of this property.

2.1. Eigenfunctions of localization operators

Let \( D_a \) denote a disc of radius \( a, 0 < a \leq \infty \), \( d\mu(z) = \mu(|z|) \, dz \) a radially weighted measure and \( dz \) a Lebesgue measure on \( \mathbb{C} \).

In the following, we will denote by \( H_a = L^2(D_a, d\mu(z)) \) the Hilbert space of analytic functions \( F \) on \( \mathbb{C} \), such that

\[
\|F\|_{H} = \int_{D_a} |F(z)|^2 \, d\mu(z)
\]

is finite.

In proposition 1 we collect the most important facts about the ‘direct problem’ studied in [4], [5] when transferred to the complex domain. This point of view is essentially contained in [18], but we have observed that both problems can be understood as special cases of a more
general formulation with general radial measures on complex domains. This viewpoint is later reflected in our derivation of the results about the inverse problems.

Proposition 1. Consider all radial measures on discs $D_R$ with radius $R$ in the complex plane, i.e. the measures constituted by the weighted measure $d\mu(z) = \mu(|z|)\,dz$, defined on $D_R$, whose weight $\mu(|z|)$ depends only on $r = |z|$. The following statements are true.

(a) The monomials are orthogonal on any disc $D_R$ centered at zero with radius $R$ in the complex plane and with respect to all concentric measures. Consequently, the monomials are also orthogonal on any annulus centered at zero.

(b) Assume $0 < c_{n,a} < \infty$ for all moments $c_{n,a}$ of $z^a \, \mu(|z|)\,dz$. Then, the normalized monomials $e_{n,a} = z^a / \sqrt{(c_{2n+1,a})}$ constitute an orthonormal basis for $\mathcal{H}_a$.

(c) If, in addition, $\sum_{n \geq 0} (c_{2n+1,a})^{-1} |z|^{2n}$ is finite for all $z \in D_a$, then $\mathcal{H}_a$ is a reproducing kernel Hilbert space with reproducing kernel

$$K(z, w) = \sum_{n \geq 0} (c_{2n+1,a})^{-1} z^n w^n.$$

(d) The functions $F(z) = e_{n,a}$ are eigenfunctions of the problem

$$\int_{D_R} F(z)K(z, w)\,d\mu(z) = \lambda F(w). \quad (4)$$

Proof.

(a) Orthogonality can directly be seen by

$$\int_{D_R} z^n z^m \, d\mu(z) = \int_0^R r^{n+m+1} \int_0^{2\pi} e^{i(n-m)\theta} \, r \, d\mu(r) \, dr = c_{2n+1,a} \delta_{n,m}, \quad (5)$$

with $c_{n,R} = 2\pi \int_0^R r^n \, d\mu(r) \, dr$.

(b) Consider a domain $D_a$, $R < a < \infty$ such that $\lim_{r \to a} d\mu(r) = 0$. Since the power series $\sum a_n z^n$ of an analytic function $F$ on $\mathbb{C}$ converges uniformly on every $D_R$, we may interchange integral and summation in the following equations: suppose that $\langle F, e_{n,a} \rangle = 0$ for all $n \in \mathbb{Z}$, then

$$0 = \frac{1}{\sqrt{c_{2m+1,a}}} \lim_{R \to a} \int_{D_R} \sum_{n \geq 0} a_n z^n \, d\mu(|z|) \, dz = \frac{1}{\sqrt{c_{2m+1,a}}} \lim_{R \to a} \sum_{n \geq 0} a_n \int_{D_R} z^n \, d\mu(|z|) \, dz = \frac{1}{\sqrt{c_{2m+1,a}}} \lim_{R \to a} a_m C_{2m+1,R}$$

which implies $a_m = 0$ for all $m$ and hence $F \equiv 0$, which proves completeness of the functions $\{e_{n,a}\}$ in $\mathcal{H}_a$.

(c) We need to show that point evaluations of $F \in \mathcal{H}_a$ are bounded. Expanding $F$ in terms of $\{e_{n,a}\}$, we observe that

$$|F(z)| = \left| \sum_{n \geq 0} \langle F, e_{n,a} \rangle \frac{z^n}{\sqrt{c_{2n+1,a}}} \right| \leq \|F\|_{\mathcal{H}_a} \left( \sum_{n \geq 0} \frac{1}{c_{2n+1,a}} |z|^{2n} \right)^{\frac{1}{2}}.$$

Thus, by the assumption on the growth of the moments, $\mathcal{H}_a$ is a reproducing kernel Hilbert space.
Write \( U \) for the operator which multiplies a function \( F \in \mathcal{H} \) by the characteristic function of the circle \( D_R \) and \( P \) for the orthogonal projection onto \( \mathcal{H}_a \), given by the reproducing kernel. Since 
\[
0 = \int_{D_R} e_{n,a} \overline{e_{n,a}} \mu(|z|) \, dz = \int_{D_R} e_{n,a} PU(e_{n,a}) \, d\mu(z)
\]
and completeness of \( e_{n,a} \) implies that \( PU e_{n,a} = e_{n,a} \). Denoting by \( K(z, w) \) the reproducing kernel of \( \mathcal{H}_a \), the functions \( F(z) = e_{n,a} \) are eigenfunctions of problem (4).

Using appropriate unitary operators (the so-called Bargmann and Bergman transform, to be defined later in this paper), the solution to the general problem just described can be shown to be equivalent to the solution of the ‘direct’ problems considered in [4] and [5]. Indeed, the \( d \mu(z) = e^{-\pi |z|^2} \, dz \) case can be translated to the Gabor localization problem studied by Daubechies and the case \( d \mu(z) = (1 - |z|^2)^{\alpha} \, dz \) to the wavelet localization studied by Daubechies and Paul. Details will be given in sections 3 and 4.

2.2. The localization domain of monomials

We now turn to the general problem, given by (4). The following, central proposition states that orthogonality of any monomial to almost all other monomials with respect to a bounded, simply connected domain \( \Omega \subset \mathbb{C} \) forces \( \Omega \) to be a disc centered at zero. We also consider more general domains as described in corollary 1. Note that we identify \( \mathbb{R}^2 \) with \( \mathbb{C} \) for the geometric description. The proof is based on an idea of Zalcman [22] and is essentially similar to the proof given in [1], but in a more general setting, namely generalizing from area measure to general concentric measures. To adapt the original argument, we rely on proposition 1.

**Proposition 2.** Let \( d \mu(z) \) be a positive, concentric measure on \( D_\alpha \subseteq \mathbb{C} \) and consider a simply connected set \( \Omega \subset D_\alpha \). Assume, for some \( m \) and \( k \geq 0 \) that
\[
\int_{\Omega} |z|^{2m} \frac{\overline{z} - w}{z - w} \, d\mu(z) = \lambda \delta_{b,0}.
\]
Then \( \Omega \) must be a disc centered at zero.

**Proof.** Since
\[
\frac{\overline{z} - w}{z - w} = \frac{\overline{z}}{1 - \frac{w}{\overline{z}}} = -\sum_{n=1}^{\infty} \frac{z^n}{w^{n-1}},
\]
we have for every \( z \in \Omega \) and \( w \) such that \( |w| > \text{sup}\{|z|; z \in \Omega\} \), the following expansion:
\[
|z|^{2m} \frac{\overline{z} - w}{z - w} = -|z|^{2m} \left( \frac{1}{w} + \sum_{n=1}^{\infty} \frac{z^n}{w^{n+1}} \right).
\]
Integrating term wise and using (6) yields
\[
\int_{\Omega} |z|^{2m} \frac{\overline{z} - w}{z - w} \, d\mu(z) = 0;
\]

hence
\[
\int_{\Omega} |z|^{2m} \frac{|z|^2 - \overline{z} w}{|z - w|^2} \, d\mu(|z|) = \int_{\Omega} |z|^{2m} \frac{\overline{z}}{z - w} \, d\mu(z)
\]
\[
= \frac{1}{w} \int_{\Omega} |z|^{2m} \frac{\overline{z} w}{z - w} \, d\mu(z) = 0.
\]
The left expression in (8) is continuous as a function of \( \Omega \) since the integrand is locally integrable in \( z \). Therefore, (9) holds on \( \overline{\Omega} \).

We next show that 0 is inside \( \Omega \). Begin by observing that, for \(|w| > \sup \{|z| : z \in \Omega\}|\), we can expand and integrate term wise so that
\[
\int_{\Omega} |z|^{2m} \frac{1}{\bar{z} - w} \, d\mu(z) = \frac{1}{w} \int_{\Omega} |z|^{2m} \frac{\bar{w}}{\bar{z} - w} \, d\mu(z) = \frac{1}{w^2} \lambda.
\]
Let \( C > \sup \{|z| : z \in \Omega\} \). We let \( d(w, \Omega) \) denote the Euclidean distance between \( w \) and \( \Omega \), i.e. \( d(w, \Omega) = \inf_{w' \in \Omega} |w - w'| \). Then the following pointwise estimate in \( \overline{\Omega} \in \Omega^c \) holds:
\[
\int_{\Omega} |z|^{2m} \frac{1}{\bar{z} - w} \, d\mu(z) \leq \frac{C^{2m}}{d(w, \Omega)}.
\]
This allows us to extend (10) by analytic continuation to \( \overline{\Omega} \).

Suppose now that 0 \( \notin \Omega^c \). Then we can find a sequence of points \( \{w_n\} \) contained in \( \Omega^c \) such that \( w_n \to 0 \). This would give
\[
\lim_{n \to \infty} \int_{\Omega} |z|^{2m} \frac{1}{\bar{z} - w_n} \, d\mu(z) = \lim_{n \to \infty} \frac{1}{w_n} \lambda_m = \infty.
\]
On the other hand, because of the continuity of the left expression in \( \overline{\Omega} \),
\[
\lim_{n \to \infty} \int_{\Omega} |z|^{2m} \frac{1}{\bar{z} - w_n} \, d\mu(z) = \int_{\Omega} |z|^{2m} \frac{1}{\bar{z}} \, d\mu(z),
\]
and the integral on the right is bounded for every \( m \geq 0 \), since we are assuming that 0 \( \notin \Omega \).

This is a contradiction and we must have 0 \( \in \Omega \).

Finally, we can consider \( D_R \), the largest disc centered at zero and contained in \( \Omega \). Using proposition 1(a), we can repeat the steps leading to (10) with \( D_R \) instead of \( \Omega \). Pick a point \( w_0 \in \partial D_R \cap \partial \Omega \). Then
\[
\int_{\Omega \setminus D_R} |z|^{2m} \frac{|z|^2 - \text{Re} \bar{z} w_0}{|z - w_0|^2} \, d\mu(z) = 0.
\]
Since for \( z \in \Omega \setminus D_R \), \(|z||w_0| \leq |z|^2\), the integrand is positive on \( \Omega \setminus D_R \). This forces \( \Omega \setminus D_R \) to be of area measure zero, which implies \( \Omega = D_R \).

For the next statement, we consider a more general situation. Let \( \gamma_j \), \( j = 1, \ldots, n \) be a family of non-intersecting Jordan curves with interiors \( I^j \) such that \( I^{j-1} \subset I^j \) for all \( j > 1 \).

If \( n \) is even, set \( K = n/2 \) and let \( \Omega_k = I^{2k+1} \setminus I^{2k-1} \) for \( k = 1, \ldots, K \).

If \( n \) is odd, set \( K = (n-1)/2 \) and let \( \Omega_1 = I^1 \) and \( \Omega_k = I^{2k-1} \setminus I^{2k-2} \) for \( k = 2, \ldots, K \). For the situation just described, we set \( \Omega = \bigcup_{k=1}^{K} \Omega_k \) and consider the corresponding localization operator. The next corollary shows that under the double orthogonality condition (6), all curves must contain 0 in their interior. Furthermore, for \( n = 2 \), if one of the two curves is a circle, then \( \Omega \) must be an annulus.

Corollary 1.
(a) Let (6) hold for \( \Omega = \bigcup_{k=1}^{K} \Omega_k \) defined by a family of nested Jordan curves as described above. Then all curves \( \gamma_j \) must contain zero.
(b) If \( n = 2 \) and \( \gamma_j \) is a circle centered at 0 for \( j = 1 \) or \( j = 2 \), then \( \Omega \) is an annulus, see figure 1.

Proof.
(a) We will show by induction that 0 must be inside all curves \( \gamma_j \), \( j = 1, \ldots, n \).
Case \( n = 1 \). Then \( \Omega \) is the interior of \( \gamma_1 \), therefore simply connected, and it follows from the proof of proposition 2, that \( 0 \in \Omega \).

Case \( n = 2 \). Then \( \Omega = I_{\gamma_2} \setminus I_{\gamma_1} \) and \( I_{\gamma_1} \) is simply connected. We apply, by assuming that \( 0 \in (I_{\gamma_2})^c \), the argument used in the first paragraph of case \( n = 1 \) to show that \( 0 \in (\Omega \cup I_{\gamma_1})^c \). Then, either \( 0 \in \Omega \) or \( 0 \in I_{\gamma_1} \). In the first case, we consider again \( D_R \), the largest disc centered at zero contained in \( \Omega \) and argue as in case \( n = 1 \) to show that \( \Omega = D_R \), which contradicts the assumption that \( n = 2 \). Therefore, \( 0 \in I_{\gamma_1} \).

Arbitrary \( n \in \mathbb{N} \). Assume that, for \( n - 1 \) curves, \( 0 \) is inside all curves. For \( n \) curves, we first show that \( 0 \in I_{\gamma_n} \); assume that \( 0 \in \Omega_k \) and use, as before, the argument from case \( n = 1 \) to show that this leads to \( n = 1 \). Consequently, \( 0 \) must be inside the remaining \( n - 1 \) curves and, by induction hypothesis, inside all curves \( \gamma_j, j = 1, \ldots, n \).

(b) First assume that \( \Omega \) is a disc, centered at zero, with a hole—in other words, that \( \gamma_2 \) is a circle. Then, \( I_{\gamma_2} \) is a disc centered at zero, such that (6) holds for \( I_{\gamma_2} \) and therefore also for \( I_{\gamma_1} \). Since the latter is simply connected, it must be a disc centered at \( 0 \).

Now let \( I_{\gamma_n} \) enclose a disc centered at \( 0 \). We then consider the largest annulus \( \Pi \) contained in \( \Omega \); it is given by \( \Pi = D_R \setminus I_{\gamma_n} \) where \( D_R \) is the largest disc centered at zero and contained in \( I_{\gamma_n} \). Due to proposition 1(a), condition (6) holds on \( \Pi \) and we obtain (10) with \( \Pi \) instead of \( \Omega \). Pick a point \( w_0 \in \partial D_R \cap \gamma_2 \). Then

\[
\int_{\Omega \setminus \Pi} |z|^{2m} \frac{|z|^2 - \text{Re} \bar{z}w_0}{|z - w_0|^2} d\mu(z) = 0.
\]

and \( |\text{Re} \bar{z}w_0| \leq |z||w_0| \leq |z|^2 \) and the integrand is positive on \( z \in \Omega \setminus \Pi \), which implies \( \Omega = \Pi \).

\[\square\]
3. An inverse problem for Gabor localization

In this section, we prove theorem 1 and derive the complete solution of the classical eigenvalue problem (2) from the assumption that any single solution is a Hermite function.

In the following, we identify \((x, \xi)\) with \(z = x + i\xi\) and we recall that \(\pi(z)\varphi(t) = \pi(x, \xi)\varphi(t) = \varphi(t - x) e^{2\pi izt}\).

3.1. Bargmann transform

In the Gabor case, the choice of the Gaussian function \(\varphi(t) = 2^{\frac{1}{4}} e^{-\pi t^2}\) allows the translation of the time–frequency localization operator \(H_{sur}\) to the complex analysis set-up via the Bargmann transform \(\mathcal{B}\). \(\mathcal{B}\) is defined for functions of a real variable as

\[
\mathcal{B} f(z) = \int_{\mathbb{R}} f(t) e^{2\pi izt - \frac{1}{4} t^2} \, dt = e^{-\pi z^2 + \pi i z \xi} \mathcal{V}_\varphi f(x, -\xi).
\]

(11)

\(\mathcal{B}\) maps \(L^2(\mathbb{R})\) unitarily onto \(\mathcal{F}^2(\mathbb{C})\), the Bargmann–Fock space of analytic functions with the inner product obtained by choosing the measure \(d\mu(z) = e^{-\pi |z|^2} \, dz\).

3.2. The Hermite functions

The normalized monomials \(e_n = (\pi^{n/2} / n!) \cdot \xi^n = \mathcal{B} h_n(z) = e^{-\pi z^2 + \pi i z \xi} \mathcal{V}_\varphi h_n(z)\) form an orthonormal basis for \(\mathcal{F}^2(\mathbb{C})\). Here, \(h_n(t) = c_n e^{\pi t^2} \left(\frac{\pi}{2}\right)^n (e^{-2\pi t^2})^n\) are the Hermite functions, which, by appropriate choice of \(c_n\), provide an orthonormal basis of \(L^2(\mathbb{R})\). As a direct consequence of the unitarity of \(\mathcal{B}\) and \(\mathcal{V}_\varphi\), the set \(\{\mathcal{V}_\varphi h_n, n \in \mathbb{N}\}\) is orthogonal over all discs \(D_R\).

3.3. Proof of theorem 1

We first deduce the equivalent formulation of the eigenvalue problem (2) in the Bargmann domain. Since the Bargmann transform is unitary, (2) is equivalent to

\[
\int_{\Omega} \mathcal{V}_\varphi f(z) \mathcal{B}(\pi(z)\varphi(w)) \, dz = \lambda \mathcal{B} f(w).
\]

Now, since \(\mathcal{B}(\pi(z)\varphi(w)) = e^{-\pi iz\xi} e^{-\pi |z|^2/2} e^{\pi w^2}\), we write the previous equation as

\[
\int_{\Omega} \mathcal{V}_\varphi f(z) e^{-\pi iz\xi} e^{-\pi |z|^2/2} e^{\pi w^2} \, dz = \lambda \mathcal{B} f(w).
\]

Thus, by (11), we have

\[
\int_{\Omega} \mathcal{B} f(z) e^{\pi |w - \pi |z|^2} \, dz = \lambda \mathcal{B} f(w).
\]

By the unitarity of the Bargmann transform, we conclude that the eigenvalue problem (2) on \(L^2(\mathbb{R})\) is equivalent to

\[
\int_{\Omega} F(z) e^{\pi |w - \pi |z|^2} \, dz = \lambda F(w),
\]

(12)

on \(\mathcal{F}^2(\mathbb{C})\). We may now expand the kernel \(e^{\pi w^2}\) in its power series which transforms the eigenvalue equation to

\[
\lambda F(w) = \sum_{n=0}^{\infty} \frac{\pi^n}{n!} w^n \int_{\Omega} F(z) z^n e^{-\pi |z|^2} \, dz.
\]

(13)
Now we use the assumption that $z^m$ solves (13) for $\lambda = \lambda_m$—in other words, that any of the solutions of (2) is a Hermite function. Setting $F(z) = z^m$ then gives

$$\lambda_m w^m = \sum_{n=0}^{\infty} \frac{\pi^n}{m!} w^n \int_{\Omega} z^m e^{-\pi |z|^2} \, dz.$$  

By the identity theorem for analytic functions, this implies

$$\int_{\Omega} z^m e^{-\pi |z|^2} \, dz = \lambda_m \frac{m!}{\pi^m} \delta_{n,m}.$$  

In particular, setting $n = m + k$,

$$\int_{\Omega} |z|^{2m+k} e^{-\pi |z|^2} \, dz = \lambda \delta_{k,0}, \text{ for all } k \geq 1.$$  

Now proposition 2 can be applied and we conclude that $\Omega$ must be the union of $\frac{n}{2}$ annuli centered at 0 for even $n$ and the union of a disc and $\frac{n-1}{2}$ annuli centered at 0 for odd $n$. In particular, for simply connected $\Omega$, we obtain a disc centered at zero.

3.4. Consequences of theorem 1

**Corollary 2.** Let $\Omega$ be simply connected. If the Gabor transform of one of the eigenfunctions of the localization operator $H_\Omega$ has Gaussian growth, $O(e^{-\pi |z|^2})$, then $\Omega$ must be a disc. The same conclusion holds if some eigenfunction has Gaussian growth in both the time and the frequency domains.

**Proof.** This is a consequence of the version of Hardy’s uncertainty principle for the Gabor transform proved by Gröchenig and Zimmermann [11]. They showed that, if $V_g f(z) = O(e^{-\pi |z|^2})$, then both $f$ and $g$ must be time–frequency shifts of a Gaussian function. Therefore, under the hypotheses of the corollary, the Gaussian (which is the first Hermite function) is an eigenfunction of the localization operator $H_\Omega$ and by theorem 1, $\Omega$ must be a disc. The second statement follows in a similar fashion from the classical Hardy uncertainty principle [12]. □

The result of theorem 1 immediately implies that the complete solution of (2) is given by the orthonormal basis of Hermite functions.

**Corollary 3.** Assume that an orthonormal basis of $L^2(\mathbb{R})$ has doubly orthogonal Gabor transform with respect to the Gaussian window $\varphi$ and some domain $\Omega$:

$$\int_{\Omega} V_{\varphi} \psi_j(z) \overline{V_{\varphi} \psi_{j'}}(z) \, dz = c_j \delta_{j,j'}.$$  

Let $\Omega$ be simply connected or of the form stated in corollary 1(b). If, for any $j_0$, $\varphi_{j_0} = h_{j_0}$ is a Hermite function, then for every $j \geq 0$,

$$\varphi_j = h_j.$$  

**Proof.** Note that an orthonormal basis of $L^2(\mathbb{R})$ satisfies (15) if and only if it consists of eigenfunctions of the localization operator $H_\Omega$. Hence, we are in the situation of theorem 1, and $\Omega$ must be disc centered at zero, the union of a disc and a finite number of annuli centered at zero or an annulus centered at zero, respectively. This, in turn, implies that all eigenfunctions are Hermite functions. □
**Remark 1.** Note the following consequence of theorem 1: if the localization domain $\Omega$ is not a disc, then the function of optimal concentration inside $\Omega$, in the sense of (3), cannot be a Gaussian window. On the other hand, it is well-known that Gaussian windows uniquely minimize the Heisenberg uncertainty relation. In this sense, disks seem to be the optimal domain for measuring time–frequency concentration.

Gaussian windows $\psi$ (and also higher order Hermite functions) are a popular choice for the basic atom in the generation of Gabor frames, whose members are given as $\pi(\lambda)\psi$, $\lambda \in \Lambda$ for some discrete subgroup $\Lambda \subset \mathbb{R}^2$. A popular choice of $\Lambda$ is $\Lambda = a\mathbb{Z} \times b\mathbb{Z}$, i.e. a rectangular lattice. In this case, the fundamental domain of $\Lambda$ in $\mathbb{R}^2$ is rectangular and thus, according to theorem 1, no Hermite function can be maximally concentrated inside the fundamental domain. This observation suggests that Gaussian or Hermitian windows are not an ideal choice for generating Gabor frames along a rectangular lattice. Although no proof for a precise statement exists, this observation has been made before and it is consistently confirmed in numerical experiments. In particular, in [21] it is mentioned that Gabor frames generated by time–frequency shifting Gaussian pulses over a hexagonal lattice have better condition number than frames obtained via a corresponding rectangular lattice. It is well-known and was shown by Gauss in 1840 that hexagonal lattices provide the densest packing of circles in the plane\(^4\). On the other hand, it is known that a Gabor frame with a Gaussian basic window is never tight [8].

Motivated by our observations we formulate the following conjecture:

Given a fixed redundancy $red > 1$, the condition number of a Gabor frame with Gaussian window $\psi$ is optimal for a hexagonal lattice.

### 3.5. Remark (due to Karlheinz Gröchenig)

Since Daubechies’ results also extend to localization operators with symbols $\sigma$ other than indicator functions, stating that any radial symbol equally leads to localization operators diagonalized by the Hermite functions, one may ask the obvious question, whether a similar inverse statement to theorem 1 can be expected for more general symbol classes than indicator functions. The following example shows that this is not true.

Let $H_\sigma$ be a time–frequency localization operator. Then, for every $N \in \mathbb{N}_0$ there exist non-negative, non-radial symbols $\sigma$, such that $H_\sigma h_N = \lambda h_N$.

To construct $\sigma$, we proceed as in section 3.3 and consider the equivalent operator on $\mathcal{F}^2(\mathbb{C})$, i.e.

$$T_\sigma F(w) = \int_\mathbb{C} \sigma(z) F(z) e^{2\pi i w \cdot z} \, dz.$$ We then claim that $T_\sigma(e_N^N) = \lambda_N e_N$ for some non-radial $\sigma$. We fix $N \in \mathbb{N}_0$ and let

$$\sigma(z) = \sigma(r e^{2\pi i r}) = \sigma_0(r) + \sigma_1(r) \cdot (e^{2\pi i (N+1) r} + e^{-2\pi i (N+1) r}),$$

where $\sigma_0(r) \geq 2|\sigma_1(r)| \forall r \geq 0$ and $\int_0^\infty \sigma_1(r) r^{3N+1} e^{-\pi r^2} \, dr = 0$. Observe that $\sigma_1$ can be chosen to be bounded, compactly supported and real-valued. Then we have $\sigma(z) \geq \sigma_0(r) - 2|\sigma_1(r)| \geq 0$. Since $\sigma_0$ is radial, we have $T_{\sigma_0}(e_N^N) = \lambda_N e_N$ with $\lambda_N > 0$. Therefore, it is enough to show that $T_{\sigma_1}(e_N^N) = 0$. However, this is easy to see by considering $\sigma_+ = \sigma_1(r) e^{2\pi i (N+1) r}$ and $\sigma_- = \sigma_1(r) e^{-2\pi i (N+1) r}$ separately and noting that, since $T_{\sigma_0}$ is

entire, the task is reduced to showing that \( \frac{\alpha}{\pi} = 0 \) for all \( \alpha \in \mathbb{N}_0 \). A straightforward calculation shows that, setting \( F(z) = z^N \) and writing \( (T_{\alpha})^{(l)} = \frac{\alpha}{\pi} T_{\alpha}(f) \), we have

\[
(T_{\alpha}(F))^{(l)}(0) = \int_{\mathbb{C}} \sigma_\pm(z) e^{\frac{\alpha}{i} |z|^2} \, dz.
\]

We finally substitute polar coordinates \( z = r e^{\pi i} \) to obtain, for \( \sigma_\pm \):

\[
(T_{\alpha}(F))^{(l)}(0) = \pi \int_0^1 \int_{\mathbb{R}} \sigma_\pm(r) r^{N+l} e^{\pi r} r \, dr \, dt \int_0^1 e^{2\pi i (N+l-t)} \, dt.
\]

The integral over \( t \) is zero for \( l \neq 2N+1 \) by orthogonality of the Fourier basis and the integral over \( r \) is zero for \( l = 2N+1 \) by assumption. The argument for \( \sigma_- \) is similar.

4. An inverse problem for wavelet localization

By replacing ‘Gabor transform’ by ‘wavelet transform’ in the formulation of the inverse problem for time–frequency localization, we may define a completely analogous inverse problem for wavelet localization. The corresponding direct problem has been treated by Daubechies and Paul in [5] and by Seip in [18]. This section is related to the previous one in the same way as the direct problem studied in [5] is related to the problem studied in [4].

It is quite remarkable that, after an appropriate reformulation of the eigenvalue problem, we may define a completely analogous inverse problem for time–frequency localization. The corresponding direct problem has been treated by Daubechies and Paul in [5] and by Seip in [18]. This section is related to the previous one in the same way as the direct problem studied in [5] is related to the problem studied in [4].

By replacing 'Gabor transform' by 'wavelet transform' in the formulation of the inverse problem for time–frequency localization, we may define a completely analogous inverse problem for wavelet localization. Since analyticity will play a fundamental role, in this section we restrict ourselves to functions defined on the half-plane, we will translate the problem to a conformally equivalent hyperbolic geometry of the upper-half plane. Since the set-up of proposition 1 is not visible in the spaces defined in the unit disc. This transformation is given by a Cayley transform as defined in section 4.2:

\[
\mathcal{B} : L^2(\mathbb{R}) \rightarrow \mathcal{F}(\mathbb{C}),
\]

we now need to further transform the images of the so-called Bergman transform (Berg \( \alpha \)) to a space defined in the unit disc. This transformation is given by a Cayley transform \( T_{\alpha} \) as defined in section 4.2:

\[
H^2(\mathbb{C}^+) \rightarrow A_{\alpha}(\mathbb{C}^+) \rightarrow A_{\alpha}(\mathbb{D}).
\]

The role of the Hermite functions is taken over by special functions, whose Fourier transforms are the Laguerre functions. This is possible, since the Laguerre functions constitute an orthogonal basis for \( L^2(0, \infty) \) and the Fourier transform provides a unitary isomorphism \( H^2(\mathbb{C}^+) \rightarrow L^2(0, \infty) \).

4.1. The wavelet transform

Since analyticity will play a fundamental role, in this section we restrict ourselves to functions in a subspace of \( L^2(\mathbb{R}) \), namely to \( f \in H^2(\mathbb{C}^+) \), the Hardy space in the upper-half plane. \( H^2(\mathbb{C}^+) \) is constituted by analytic functions \( f \) such that

\[
\sup_{0 \leq s < \infty} \int_{-\infty}^{\infty} |f(x+is)|^2 \, dx < \infty.
\]
The functions in the space $H^2(C^+)$ may be considered as being of ‘positive frequency’ since a well-known Paley–Wiener theorem says that $\mathcal{F}(H^2(C^+)) = L^2(0, \infty)$. For this reason it is common to study $H^2(C^+)$ on the ‘frequency side’, where many calculations become easier. For convenience, we will use a different normalization of the Fourier transform in this section, namely $(Ff)(\xi) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} e^{-i\xi t} f(t) \, dt$. Now consider $\mathbb{C}^+ = \{z \in \mathbb{C} : \text{Re}(z) > 0\}$. For every $x \in \mathbb{R}$ and $s \in \mathbb{R}^+$, let $z = x + is \in \mathbb{C}^+$ and define

$$\pi_t^z g(t) = s^{-2} g(s^{-1}(t - x)).$$

Fix a function $g \neq 0$ such that

$$0 < \|Fg\|_{L^2(\mathbb{R}^+,s^{-1})}^2 = C_g < \infty.$$ Such functions are called admissible and the constant $C_g$ is the admissibility constant. Then the continuous wavelet transform of a function $f$ with respect to a wavelet $g$ is defined, for every $z = x + is \in \mathbb{C}$ as

$$W_g f(z) = \langle f, \pi_z^g \rangle_{H^2(\mathbb{C}^+)}.$$ (17)

Let $d\mu^+(z)$ denote the standard normalized area measure in $\mathbb{C}^+$. The orthogonal relations for the wavelet transform

$$\int_{\mathbb{C}^+} W_{g_1} f_1(x,s) W_{g_2} f_2(x,s) s^{-2} d\mu^+(z) = \langle Fg_1, Fg_2 \rangle_{L^2(\mathbb{R}^+,s^{-1})} \langle f_1, f_2 \rangle_{H^2(\mathbb{C}^+)},$$ (18)

are valid for all $f_1, f_2 \in H^2(\mathbb{C}^+)$ and $g_1, g_2 \in H^2(\mathbb{C}^+)$ admissible. As a result, the continuous wavelet transform provides an isometric inclusion

$$W_g : H^2(\mathbb{C}^+) \to L^2(\mathbb{C}^+, s^{-2} \, dx \, ds),$$

which is an isometry for $C_g = 1$.

4.2. Bergman spaces

Let $\alpha > -1$. The Bergman space in the upper-half plane, $A_\alpha(\mathbb{C}^+)$, is constituted by the analytic functions in $\mathbb{C}^+$ such that

$$\int_{\mathbb{C}^+} |f(z)|^2 s^{\alpha} d\mu^+(z) < \infty,$$ (19)

where $d\mu^+(z)$ stands for the standard normalized area measure in $\mathbb{C}^+$.

Now consider $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$. The Bergman space in the unit disc, denoted by $A_\alpha(\mathbb{D})$, is constituted by the analytic functions in $\mathbb{D}$ such that

$$\int_{\mathbb{D}} |f(w)|^2 (1 - |w|)^\alpha \, dA(w) < \infty,$$ (20)

with $dA(w)$ being the normalized area measure in $\mathbb{D}$. The map $T_\alpha : A_\alpha(\mathbb{C}^+) \to A_\alpha(\mathbb{D})$, defined as

$$(T_\alpha f)(w) = \frac{2^{\frac{\alpha+1}{2}}}{(1 - w)^{\alpha+2}} f \left( \frac{w + 1}{i(w - 1)} \right),$$

provides a unitary isomorphism between the two spaces. The reproducing kernel of $A_\alpha(\mathbb{C}^+)$ is

$$K^\alpha_{\mathbb{C}^+}(z, w) = \frac{1}{(w - z)^{\alpha+2}}.$$ (21)

Now observe that, letting $T_\alpha$ act on the reproducing kernel of $A_\alpha(\mathbb{C}^+)$, first as a function of $w$ and then as a function of $\overline{z}$, we are led to the reproducing kernel of $A_\alpha(\mathbb{D})$,

$$K^\alpha_{\mathbb{D}}(z, w) = \frac{1}{(1 - w \overline{z})^{\alpha+2}}.$$
4.3. The Bergman transform

If we choose the window $\psi_\alpha$ as

$$\mathcal{F}\psi_\alpha(t) = \frac{1}{c_\alpha} \mathbf{1}_{[0,\infty)} t^\alpha e^{-t},$$

then we can relate the wavelet transform to Bergman spaces of analytic functions. Here,

$$c_\alpha^2 = \int_0^\infty t^{2\alpha-1} e^{-2t} \, dt = 2^{2\alpha-1} \Gamma(2\alpha),$$

where $\Gamma$ is the Gamma function. The choice of $c_\alpha$ implies $C_{\psi_\alpha} = 1$ and the corresponding wavelet transform is isometric. The Bergman transform of order $\alpha$ is the unitary map $\text{Ber}_\alpha : H(\mathbb{C}^+) \to A_\alpha(\mathbb{C}^+)$ given by

$$\text{Ber}_\alpha f(z) = s^{-\alpha-1} W_{\alpha/2} f(-s, s) = c_\alpha \int_0^\infty t^{\alpha+1} (\mathcal{F} f)(t) e^{t^2} \, dt. \quad (23)$$

4.4. The Laguerre and other related systems of functions

We define the Laguerre functions

$$l_\alpha^n(x) = \mathbf{1}_{[0,\infty)}(x) e^{-x/2} x^{\alpha/2} L_n^\alpha(x)$$

in terms of the Laguerre polynomials

$$L_n^\alpha(x) = \frac{e^x x^{-\alpha}}{n!} \frac{d^n}{dx^n} [e^{-x} x^{\alpha+n}] = \sum_{k=0}^n (-1)^k \binom{n + \alpha}{n - k} x^k. \quad (24)$$

By repeated integration by parts, one sees that the polynomials $L_n^\alpha(x)$ are orthogonal on $(0, \infty)$ with respect to the weight function $e^{-x} x^\alpha$. Thus, for $\alpha \geq 0$, the Laguerre functions $l_\alpha^n$ constitute an orthogonal basis for the space $L^2(0, \infty)$. We will use a related system of functions $\psi_\alpha^n$ defined as

$$(\mathcal{F} \psi_\alpha^n)(t) = \left( \frac{(-1)^n n!}{2^{2\alpha+2n+1} \Gamma(n+2+\alpha) \Gamma(2+\alpha)} \right)^{1/2} l_n^{\alpha+1}(2t).$$

Now consider the monomials

$$e_\alpha^n(w) = \left( \frac{\Gamma(n+2+\alpha)}{n! \Gamma(2+\alpha)} \right)^{1/2} w^n.$$

We can apply proposition 1 with $\mu(|z|) = (1 - |w|^2)^\alpha$. We conclude that $\{e_\alpha^n\}_{n=0}^\infty$ forms an orthonormal basis for $A_\alpha(\mathbb{D})$ and that they are orthogonal on every disc $D_r \subset \mathbb{D}$: for every $r > 0$,

$$\int_{D_r} e_\alpha^n(w) \overline{e_\alpha^m(w)} (1 - |w|^2)^\alpha \, d\Lambda(w) = C(r, m) \delta_{nm}. \quad (25)$$

The normalization constant $C(r, m)$ depends on $r$ and $m$ and satisfies $\lim_{r \to 1^-} C(r, m) = 1$. Now, the functions $\Psi_\alpha^n$, for every $n \geq 0$ and $\alpha > -1$,

$$\Psi_\alpha^n(z) = \frac{1}{4^{\alpha+1/2}} \left( \frac{\Gamma(n+2+\alpha)}{n! \Gamma(2+\alpha)} \right)^{1/2} \left( \frac{z - i}{z + i} \right)^n \left( \frac{1}{z + i} \right)^{\alpha+2}, \quad z \in \mathbb{C}^+, \quad (26)$$

are conveniently chosen such that

$$(T_d \Psi_\alpha^n)(w) = e_\alpha^n(w).$$
Thus, a change of variables $w = \frac{z - i}{z + i}$ in (25) leads to
\[
\int_{\varphi(z_1, z_2)} \Psi_n^\alpha(z) \overline{\Psi_n^\alpha(z)} \alpha \mu^+ \alpha = C(r, m) \delta_{nm},
\]
where $\varphi(z_1, z_2) = \left| \frac{z_1 - z_2}{z_1 + z_2} \right|$ is the pseudohyperbolic metric on $\mathbb{C}^+$. Moreover, the unitarity of the operator $T_\alpha$ translates the basis property of $\{e_n^\alpha\}_{n=0}^\infty$ in $A_\delta(\mathbb{D})$ to $A_\delta(\mathbb{C}^+)$. In other words, (26) shows that $\{\Psi_n^\alpha(z)\}_{n=0}^\infty$ is an orthogonal basis of $A_\delta(\mathbb{C}^+)$. Finally, we observe that (23) together with the special function formula
\[
\int_0^\infty x^\alpha L_n^\alpha(x) e^{-\alpha x} dx = \frac{\Gamma(\alpha + n + 1)}{n!} s^{\alpha - n - 1} (s - 1)^n
\]
gives
\[
\operatorname{Ber}_\alpha \psi_n^\alpha = \Psi_n^\alpha.
\]
For an intuitive grasp of this section, keep in mind that with the composition of transforms (16), one associates the transformations of the basis functions:
\[
\psi_n^\alpha \in H(\mathbb{C}^+) \xrightarrow{\operatorname{Ber}_\alpha} \Psi_n^\alpha \in A_\delta(\mathbb{C}^+) \xrightarrow{T_\alpha} e_n^\alpha \in A_\delta(\mathbb{D})
\]

4.5. The inverse problem

We now consider the wavelet localization operator $P_{\Delta, \alpha}$ defined as
\[
P_{\Delta, \alpha} f = \int_\Delta W_{\varphi_{\alpha+2}} f (z) \overline{\varphi_{\alpha+2}} \, d\mu^+(z)
\]
and set up the corresponding eigenvalue problem
\[
P_{\Delta, \alpha} f = \lambda f.
\]

**Theorem 2.** If one of the eigenfunctions of the localization operator $P_{\Delta, \alpha}$ belongs to the family $\{\psi_n^\alpha\}$, then $\Delta$ must be a pseudohyperbolic disc centered at $i$.

**Proof.** We first rewrite the eigenvalue problem (29). A simple change of variables on the ‘Fourier’ side of the wavelet representation gives
\[
\operatorname{Ber}_\alpha (\varphi_{\alpha+2}) (w) = m_\alpha s^{\alpha+2} \left( \frac{1}{\frac{z}{w} - 1} \right) = s^{\alpha+2} K_{\mathbb{C}^+}^\alpha (z, w),
\]
where $m_\alpha = \frac{\alpha + 1}{2}$. Now apply the Bergman transform and use (23) to rewrite (29) as
\[
\int_\Delta \operatorname{Ber}_\alpha f (z) K_{\mathbb{C}^+}^\alpha (z, w) s^\alpha \, d\mu^+(z) = \lambda \operatorname{Ber}_\alpha f (w).
\]
By the unitarity $\operatorname{Ber}_\alpha : H(\mathbb{C}^+) \to A_\delta(\mathbb{C}^+)$, we conclude that our eigenvalue problem is equivalent to
\[
\int_\Delta F(z) K_{\mathbb{C}^+}^\alpha (z, w) s^\alpha \, d\mu^+(z) = \lambda F(w),
\]
with $F \in A_\delta(\mathbb{C}^+)$. Making the change of variables
\[
z_D = \frac{z - i}{z + i}, \quad w_D = \frac{w - i}{w + i},
\]
we move the eigenvalue problem to the unit disc
\[
\int_{\Omega = T_\alpha} (T_\alpha F)(z_D) \left( \frac{1 - |z_D|^2}{1 - w_D \bar{z}_D} \right)^2 \, dA_D (z_D) = \lambda (T_\alpha F)(w_D),
\]
where \( T_\alpha F(wD) \in A_\alpha(D) \). We simplify the notation writing \( z_D = z \), \( w_D = w \). Now, using the uniformly convergent expansion of the reproducing kernel,
\[
\frac{1}{(1 - w \bar{z})^{1+\alpha}} = \sum_{n=0}^{\infty} e_n(w) e_n(z),
\]
we can transform the eigenvalue equation into
\[
\lambda(T_\alpha F)(w) = \sum_{n=0}^{\infty} \frac{\Gamma(n + 2 + \alpha)}{n! \Gamma(2 + \alpha)} u^n \int_{\Omega} (T_\alpha F)(z) \overline{e_n(1 - |z|)\alpha} dA_D(z). \tag{30}
\]
If one of the eigenfunctions of the localization operator \( P_{\Delta,u} \) belongs to the family \( \{ \psi_n^u \} \), then
\[
T_\alpha (\text{Ber}_u \psi_n^u)(z) = T_\alpha (\Psi_n^u)(z) = e_n(z)
\]
solves (30) for \( \lambda = \lambda_n \). Setting \( (T_\alpha F)(z) = e_n(z) \) gives
\[
\lambda_m u_m = \sum_{n=0}^{\infty} \frac{\Gamma(n + 2 + \alpha)}{n! \Gamma(2 + \alpha)} u^n \int_{\Omega} \overline{e_n(z)}(1 - |z|)\alpha dA_D(z).
\]
Comparison of coefficients yields
\[
\int_{\Omega} z^m \overline{e_n(z)}(1 - |z|)\alpha dA_D(z) = \lambda_m \frac{n! \Gamma(2 + \alpha)}{\Gamma(n + 2 + \alpha)} d_{n,m}
\]
and further, with \( n = m + k \), the condition of proposition 2
\[
\int_{\Omega} |z|^{2m} \overline{e_n(z)}(1 - |z|)\alpha dA_D(z) = \lambda d_{k,0}, \text{ for all } k \geq 1. \tag{31}
\]
Hence, proposition 2 can be applied and \( \Omega \) must be a disc centered at zero. Now we can go back to the upper-half plane by the change of variables
\[
u = \frac{z + 1}{1 - z} \in \mathbb{C}^+,
\]
which maps \( 0 \in \mathbb{D} \) to \( i \in \mathbb{C}^+ \) and leaves the pseudohyperbolic metric invariant. Finally, note that the condition \( |z| < r \) can be written in terms of the pseudohyperbolic metric of the disc as \( \varphi_D(z, 0) < r \). Hence, the disc \( \Omega \) centered at zero is mapped to the pseudohyperbolic disc \( \Delta = \{ \varphi_{\mathbb{C}^+}(u, i) < r \} \) centered at \( u = i \). \( \square \)

**Remark 2.** We can draw a conclusion similar to the one in remark 1 after theorem 1. Indeed, it follows from theorem 2 that, if the localization domain \( \Delta \) is not a pseudohyperbolic disc, then the function \( f \) providing optimal concentration in the sense of maximizing
\[
C_\Delta(f) = \int_{\Delta} |W_{\psi_n,f}(z)|^2 \overline{dz} \tag{32}
\]
cannot be the function \( \psi_n \) in (22)—the so-called Cauchy wavelet. On the other hand, it is known that the functions \( \psi_n \) minimize the affine uncertainty principle as first mentioned in [16]. In this sense, pseudohyperbolic discs seem to be optimal domains to measure wavelet localization.

**Acknowledgments**

We would like to thank Saptarshi Das for his precious advice on UWB technology. DA was supported by the ESF activity ‘Harmonic and Complex Analysis and its Applications’, by FCT (Portugal) project PTDC/MAT/114394/2009 and CMUC through COMPETE/FEDER. MD was supported by the Austrian Science Fund (FWF):[T384-N13] Locatif.
References

[16] Paul T 1984 Affine coherent states and the radial Schrödinger equation I: radial harmonic oscillator and hydrogen atom Technical Report CPT–84/P–1710 (Centre de Physique Théorique, Marseille)
Time-frequency partitions and characterizations of modulation spaces with localization operators

Monika Dörfler, Karlheinz Gröchenig

Institut für Mathematik, Universität Wien, Alserbachstrasse 23 A-1090 Wien, Austria

Received 7 December 2009; accepted 21 December 2010
Available online 6 January 2011
Communicated by G. Godefroy

Abstract

We study families of time-frequency localization operators and derive a new characterization of modulation spaces. This characterization relates the size of the localization operators to the global time-frequency distribution. As a by-product, we obtain a new proof for the existence of multi-window Gabor frames and extend the structure theory of Gabor frames.

Keywords: Phase-space localization; Short-time Fourier transform; Modulation space; Localization operator; Gabor frame

1. Introduction

A time-frequency representation transforms a function $f$ on $\mathbb{R}^d$ into a function on the time-frequency space $\mathbb{R}^d \times \mathbb{R}^d$. The goal is to obtain a description of $f$ that is local both in time and in frequency [5,20]. The standard time-frequency representations, such as the short-time Fourier transform and its various modifications known as Wigner distribution, radar ambiguity function, Gabor transform, all encode time-frequency information. However, the pointwise interpretation of such a time-frequency representation meets difficulties because, by the uncertainty principle, a small region in the time-frequency plane does not possess a physical meaning. Therefore...
the question arises in which sense the short-time Fourier transform describes the local properties of a function and its Fourier transform.

Following Daubechies [10], we use time-frequency localization operators to give meaning to the local time-frequency content. By investigating a whole family of localization operators and gluing together the local pieces, we are able to characterize the global time-frequency distribution of a function. In more technical terms, our main result provides a new characterization of modulation spaces.

We define the short-time Fourier transform (STFT) of a function \( f \in L^2(\mathbb{R}^d) \) with respect to a window function \( \varphi \in L^2(\mathbb{R}^d) \) as

\[
V_\varphi f(x, \omega) = \int_{\mathbb{R}^d} f(t) \bar{\varphi}(t-x) e^{-2\pi i \omega \cdot t} \, dt,
\]

for all \( z = (x, \omega) \in \mathbb{R}^{2d} \). (1)

The STFT \( V_\varphi f(z) \) is a measure of the time-frequency content near the point \( z \) in the time-frequency plane \( \mathbb{R}^{2d} \). However, the STFT cannot be supported on a set of finite measure by results in [27,29,37]. This fact complicates the interpretation of local information obtained from the STFT. In particular, it is impossible to construct a projection operator that satisfies \( V_\varphi (P_\Omega f) = \chi_\Omega \cdot V_\varphi f \). As a remedy one resorts to the following definition of localization operators.

We denote translation operators by \( T_x f(t) = f(t-x) \) and time-frequency shifts by \( \pi(z) f(t) = e^{2\pi i \omega \cdot t} f(t-x) \) for \( x, \omega, t \in \mathbb{R}^d \). Fix a non-zero function \( \varphi \in L^2(\mathbb{R}^{2d}) \) (a so-called window function) and a symbol \( \sigma \in L^1(\mathbb{R}^{2d}) \). Then the time-frequency localization operator \( H_\sigma \) acting on a function \( f \) is defined as

\[
H_\sigma f = \int_{\mathbb{R}^{2d}} \sigma(z) V_\varphi f(z) \pi(z) \varphi \, dz.
\]

The integral is defined strongly on many function spaces, in particular on \( L^2(\mathbb{R}^d) \). A useful alternative definition of \( H_\sigma \) is the weak definition

\[
\langle H_\sigma f, g \rangle_{L^2(\mathbb{R}^d)} = \sigma \langle V_\varphi f, V_\varphi g \rangle_{L^2(\mathbb{R}^{2d})}.
\]

This definition can be easily extended to distributional symbols \( \sigma \in S'(\mathbb{R}^{2d}) \). The subtleties of the definition and boundedness properties between various spaces have been investigated in many papers, see [7,36,38] for a sample of results.

If \( \sigma \) is non-negative and has compact support in \( \Omega \subseteq \mathbb{R}^d \), then \( H_\sigma f \) can be interpreted as the part of \( f \) that lives essentially on \( \Omega \) in the time-frequency plane, and so \( H_\sigma \) may be taken as a substitute for the non-existing projection onto the region \( \Omega \) in the time-frequency plane.

In this paper we investigate the behavior of an entire collection of localization operators. Namely, given a lattice \( \Lambda \subseteq \mathbb{R}^{2d} \) of the time-frequency plane, we consider the collection of operators \( \{ H_{T_\lambda \sigma} : \lambda \in \Lambda \} \) and the mapping \( f \rightarrow \{ H_{T_\lambda \sigma} f \} \). If the supports of \( T_\lambda \sigma \) cover \( \mathbb{R}^{2d} \), then \( \{ H_{T_\lambda \sigma} f, \lambda \in \Lambda \} \) should contain enough information to recover \( f \) from its local components. In particular, the set \( \{ H_{T_\lambda \sigma} f : \lambda \in \Lambda \} \) should carry the complete information about the global time-frequency properties of \( f \). We make this intuition precise and derive a new characterization of modulation spaces from it. Similar to Besov spaces, modulation spaces are smoothness spaces, but the smoothness is measured by means of time-frequency distribution rather than
differences and derivatives. Here, we establish a correspondence between the behavior of the sequence \( \| H_{T_{\lambda},\sigma} f \|_2, \lambda \in \Lambda \), and the membership of \( f \) in a modulation space.

As a special case of our main theorem we formulate the following result.

**Theorem 1.** Fix a non-zero function \( \varphi \) in the Schwartz space \( S(\mathbb{R}^d) \) and a weight function \( m \) on \( \mathbb{R}^{2d} \) that satisfies \( m(z_1 + z_2) \leq C(1 + |z_1|)^N m(z_2) \) for some constants \( C, N \geq 0 \) and all \( z_1, z_2 \in \mathbb{R}^{2d} \). Then a tempered distribution \( f \) satisfies

\[
\left( \int_{\mathbb{R}^{2d}} |\mathcal{V}_{\varphi} f(z)|^p m(z)p \, dz \right)^{1/p} < \infty, \tag{3}
\]

if and only if

\[
\left( \sum_{\lambda \in \Lambda} \| H_{T_{\lambda},\sigma} f \|_2^p m(\lambda)^p \right)^{1/p} < \infty. \tag{4}
\]

The expression in (3) is just the norm of \( f \) in the modulation space \( M^p_m(\mathbb{R}^d) \). Our main result shows that the expression in (4) (using the time-frequency components of \( f \)) is an equivalent norm on the modulation space \( M^p_m(\mathbb{R}^d) \).

In pseudodifferential calculus one often defines spaces by conditions on their time-frequency components. For instance, Bony, Chemin, and Lerner [3,4] introduced a Sobolev-type space \( H(m) \) by using Weyl operators instead of localization operators. For the (extremely simplified) case of a constant Euclidean metric on the time-frequency plane, a distribution \( f \) belongs to \( H(m) \), whenever for some test function \( \psi \) on \( \mathbb{R}^{2d} \)

\[
\| f \|_{H(m)}^2 = \int_{\mathbb{R}^{2d}} \| (T_Y \psi)^m f \|_2^2 m(Y) \, dY, \tag{5}
\]

is finite, where \( \sigma^w \) is the Weyl operator corresponding to the symbol \( \sigma \). The only difference between (5) and (4) is the use of Weyl calculus instead of time-frequency localization operators and a continuous definition instead of a discrete one. It was understood only recently that \( H(m) \) coincides with the modulation space \( M^2_m(\mathbb{R}^d) \) and that (5) is an equivalent norm on \( M^2_m(\mathbb{R}^d) \) [25]. Thus Theorem 1 can be interpreted as an extension of [3] to \( L^p \)-like spaces.

Let us also mention that in the language of [35], the operators \( \{ H_{T_{\lambda},\sigma}, \lambda \in \Lambda \} \) form a \( g \)-frame for \( L^2(\mathbb{R}^d) \). Our construction seems to be one of the few non-trivial examples of \( g \)-frames that are not frames.

In this paper we prove the norm equivalence of Theorem 1 for a large class of modulation spaces and arbitrary time-frequency lattices. For a rather restricted class of lattices, namely lattices with integer oversampling, an analogous result was derived in [12] for unweighted modulation spaces. The main arguments for the integer lattice were based on Zak transform methods and interpolation. For a general lattice, these methods are no longer available, and we have to develop a completely new approach to some of the key arguments.
As a by-product of the new techniques we have found several results of independent interest.

- We formulate several structural results and characterizations of Gabor frames for multi-window Gabor frames.
- We prove a finite intersection property for time-frequency invariant subspaces of the distribution space $M^\infty(\mathbb{R}^d)$. This property resembles the finite intersection property that characterizes compact sets.
- We give a new, independent proof for the existence of multi-window Gabor frames with well-localized windows. Previous proofs were based on coorbit theory [15] and the theory of projective modules [32]. Our proof provides additional insight how the windows can be chosen.
- We derive precise estimates for the localization of the eigenfunctions of a localization operator.

This paper is organized as follows. In Section 2 we recall necessary facts from time-frequency analysis. On the one hand, we introduce modulation spaces and explain their characterization by means of multi-window Gabor frames. On the other hand, we state and prove several properties of localization operators. In Section 3, we formulate and prove our main result (Theorem 8). In Section 3.4 we analyze some of the consequences of Theorem 8 and its proof. In Appendix A we collect and sketch the proofs of some of the structural results on Gabor frames.

2. Time-frequency analysis of functions and operators

2.1. Modulation spaces

Modulation spaces are a class of function spaces associated to the short-time Fourier transform (1). Note that for a suitable test function $\varphi$, the short-time Fourier transform can be extended to distribution spaces by duality and $V_\varphi f(z) = \langle f, \pi(z)\varphi \rangle$.

For the standard definition of modulation spaces, we fix a non-zero “window function” $g \in S(\mathbb{R}^d)$ and consider moderate weight functions $m$ of polynomial growth, i.e., $m(z_1 + z_2) \leq C(1 + |z_1|)^s m(z_2)$, $z_1, z_2 \in \mathbb{R}^{2d}$ for some $C, s \geq 0$. Given a moderate weight $m$ and $1 \leq p, q \leq \infty$, the modulation space $M_{m}^{p,q}(\mathbb{R}^d)$ is defined as the space of all tempered distributions $f \in S'(\mathbb{R}^d)$ with $V_g f \in L_{m}^{p,q}(\mathbb{R}^{2d})$, with norm

$$
\| f \|_{M_{m}^{p,q}(\mathbb{R}^d)} = \| V_g f \|_{L_{m}^{p,q}(\mathbb{R}^{2d})}.
$$

If $p = q$, we write $M_{m}^{p}(\mathbb{R}^d)$.

For weight functions of faster growth we have to resort to different spaces of test functions and distributions. Let $g(t) = e^{-\pi t^T t}$ be the Gaussian window and $\mathcal{H}_0 = \text{span}\{\pi(z)g: z \in \mathbb{R}^{2d}\}$ be the linear space of all finite linear combinations of time-frequency shifts of the Gaussian. Let $v$ be a submultiplicative even weight function on $\mathbb{R}^{2d}$ and $m$ be a $v$-moderate function; this means that $v(z_1 + z_2) \leq v(z_1)v(z_2)$, $v(z) = v(-z)$ and $m(z_1 + z_2) \leq v(z_1)m(z_2)$ for all $z, z_1, z_2 \in \mathbb{R}^{2d}$. For $1 \leq p, q < \infty$ the modulation space $M_{m}^{p,q}(\mathbb{R}^d)$ is then defined as the closure of $\mathcal{H}_0$ in the norm $\| f \|_{M_{m}^{p,q}(\mathbb{R}^d)}$ as in (6). If $p = \infty$ or $q = \infty$, we take a weak*-closure of $\mathcal{H}_0$. These general modulation spaces possess the following properties. Assume that $m$ is $v$-moderate.
and $1 \leq p, q \leq \infty$, then

$$M^1_\nu(\mathbb{R}^d) \subseteq M^{p,q}_m(\mathbb{R}^d) \subseteq M^{\infty}_1(\mathbb{R}^d) = M^1_\nu(\mathbb{R}^d)^*.$$  \hspace{1cm} (7)

Further, if $\varphi \in M^1_\nu(\mathbb{R}^d)$, then

$$\|V\varphi f\|_{L^{p,q}_m} \approx \|Vg f\|_{L^{p,q}_m} = \|f\|_{M^{p,q}_m},$$  \hspace{1cm} (8)

thus different windows in $M^1_\nu(\mathbb{R}^d)$ yield equivalent norms on $M^{p,q}_m$.

The embedding (7) says that $M^1_\nu(\mathbb{R}^d)$ may serve as a space of test functions and $M^{\infty}_1/\nu(\mathbb{R}^d)$ as a space of distributions for all modulation spaces $M^{p,q}_m$ with a $\nu$-moderate weight $m$.

If $\nu_s(z) = (1 + |z|)^s$, $s \geq 0$ and $m$ is $\nu_s$-moderate, then we have

$$S(\mathbb{R}^d) \subseteq M^1_\nu(\mathbb{R}^d) \subseteq M^{p,q}_m(\mathbb{R}^d) \subseteq M^{\infty}_1/\nu_s(\mathbb{R}^d) \subseteq S'(\mathbb{R}^d),$$

in agreement with the standard definition, but for $\nu(z) = e^{a|z|^b}$ with $a > 0$ and $0 < b \leq 1$ we have

$$M^1_\nu(\mathbb{R}^d) \subseteq S(\mathbb{R}^d) \subseteq S'(\mathbb{R}^d) \subseteq M^{\infty}_1/\nu(\mathbb{R}^d).$$

In the sequel we will start with a submultiplicative weight $\nu$ and take $M^{\infty}_1/\nu(\mathbb{R}^d)$ as the appropriate distribution space. Our results hold for arbitrary submultiplicative weights $\nu$.

For the detailed theory of modulation spaces we refer to [21, Chapters 11–13], for a discussion of weights and possible distribution spaces see [23].

**Sequence space norms.** Recall that a time-frequency lattice $\Lambda$ is a discrete subgroup of $\mathbb{R}^{2d}$ of the form $\Lambda = A\mathbb{Z}^{2d}$ for some invertible real-valued $2d \times 2d$-matrix $A$.

Given a lattice $\Lambda \subseteq \mathbb{R}^{2d}$ with relatively compact fundamental domain $Q$, the discrete space $\ell^{p,q}_m(\Lambda)$ consists of all sequences $a = (a_\lambda)_{\lambda \in \Lambda}$ for which the norm

$$\|a\|_{\ell^{p,q}_m} = \left\| \sum_{\lambda \in \Lambda} |a_\lambda| \chi_{\lambda + Q} \right\|_{L^{p,q}_m}$$  \hspace{1cm} (9)

is finite. If $\Lambda = a\mathbb{Z}^d \times b\mathbb{Z}^d$, then this definition reduces to the usual mixed-norm space $\ell^{p,q}_m(\mathbb{Z}^{2d})$ with norm

$$\|a\|_{\ell^{p,q}_m} = \left( \sum_{n \in \mathbb{Z}^d} \left( \sum_{k \in \mathbb{Z}^d} |a_{kn}|^p m(ak, bn)^p \right)^{q/p} \right)^{1/q}.$$

As a technical tool we will need amalgam spaces (in one place only). A measurable function $F$ on $\mathbb{R}^{2d}$ belongs to the (Wiener) amalgam space $W(L^{p,q}_m)$, if the sequence of local suprema $a_{kn} = \text{ess sup}_{x,w \in [0,1]^d} |F(x + k, \omega + n)| = \|F \cdot T(k,n) \chi\|_{\infty}$

belongs to $\ell^{p,q}_m(\mathbb{Z}^{2d})$. The norm on $W(L^{p,q}_m)$ is $\|F\|_{W(L^{p,q}_m)} = \|a\|_{\ell^{p,q}_m}$. See [26] for an introductory article. We need their behavior under convolution and their properties under sampling.
(a) Convolution in Wiener amalgam spaces: Let $1 \leq p, q \leq \infty$ and let $m$ be a $\nu$-moderate weight. Then
\[
\|F * G\|_{W(L^{p,q}_m)} \leq C \|F\|_{W(L^{p,q}_m)} \|G\|_{L^{1,\nu}}.
\]  
(10)

(b) Sampling in Wiener amalgam spaces: For $F \in W(L^{p,q}_m)$ the following sampling property holds:
\[
\|F|_{A}\|_{p,q} \leq C_A \|F\|_{W(L^{p,q}_m)}.
\]  
(11)

These statements are proved in [26] or [21, Proposition 11.1.4, Theorem 11.1.5].

2.2. Gabor frames

Gabor frames are closely linked to modulation spaces. They constitute “basis-like” sets for modulation spaces and are used to characterize the membership in a modulation space by the magnitude of coefficients in the corresponding series expansion.

For a given lattice $\Lambda \subseteq \mathbb{R}^{2d}$ and a window function $\varphi \in L^2(\mathbb{R}^d)$, let $G(\varphi, \Lambda)$ denote the set of functions $\{\pi(\lambda)\varphi: \lambda \in \Lambda\}$ in $L^2(\mathbb{R}^d)$. The operator
\[
S_{\varphi} f = \sum_{\lambda \in \Lambda} \langle f, \pi(\lambda)\varphi \rangle \pi(\lambda)\varphi
\]
is the frame operator corresponding to $G(\varphi, \Lambda)$. If $S_{\varphi}$ is bounded and invertible on $L^2(\mathbb{R}^d)$, then $G(\varphi, \Lambda)$ is called a Gabor frame for $L^2(\mathbb{R}^d)$. This property is equivalent to the existence of two constants $A, B > 0$ such that
\[
A \|f\|_2^2 \leq \sum_{\lambda \in \Lambda} \|\langle f, \pi(\lambda)g \rangle\|^2 = \langle S_{\varphi} f, f \rangle \leq B \|f\|_2^2 \quad \text{for all} \quad f \in L^2(\mathbb{R}^d).
\]  
(12)

Using several windows $\varphi = (\varphi_1, \ldots, \varphi_n)$, we say that the union $\bigcup_{j=1}^n G(\varphi_j, \Lambda)$ is a multi-window Gabor frame, if the associated frame operator given by
\[
S_{\varphi} f = \sum_{j=1}^n \sum_{\lambda \in \Lambda} \langle f, \pi(\lambda)\varphi_j \rangle \pi(\lambda)\varphi_j = \sum_{j=1}^n S_{\varphi_j} f
\]  
(13)
is invertible on $L^2(\mathbb{R}^d)$. The frame operator can be expressed as the composition of the analysis operator $C_{\varphi, \Lambda}$ defined by
\[
C_{\varphi, \Lambda}(f)(\lambda, j) = \langle f, \pi(\lambda)\varphi_j \rangle, \quad \lambda \in \Lambda, \quad j = 1, \ldots, n
\]
and the synthesis operator $D_{\varphi, \Lambda}$ defined by $D_{\varphi, \Lambda}(c) = \sum_{\lambda \in \Lambda} \sum_{j=1}^n c_{\lambda,j} \pi(\lambda)\varphi_j$. Then $S_{\varphi, \Lambda} = D_{\varphi, \Lambda} \circ C_{\varphi, \Lambda}$. 

2.3. Characterization of modulation spaces with Gabor frames

The following characterization of modulation spaces by means of multi-window Gabor frames is a central result in time-frequency analysis and useful in many applications. It is crucial for the proof of our main theorem (Theorem 8).

**Theorem 2.** Let \( \nu \) be a submultiplicative weight on \( \mathbb{R}^{2d} \) satisfying the condition \( \lim_{n \to \infty} \nu(nz)^{1/n} = 1 \) for all \( z \in \mathbb{R}^{2d} \) and let \( m \) be a \( \nu \)-moderate weight and \( 1 \leq p, q \leq \infty \). Assume further that \( \bigcup_{j=1}^{n} G(\varphi_j, \Lambda) \) is a multi-window Gabor frame and that \( \varphi_j \in M^p_1(\mathbb{R}^d) \) for \( j = 1, \ldots, n \).

(i) A distribution \( f \) belongs to \( M^p_m(\mathbb{R}^d) \), if and only if \( C_{\varphi_j} f, \ell^p_m \) for \( j = 1, \ldots, n \). In this case there exist constants \( A, B > 0 \), such that, for all \( f \in M^p_m(\mathbb{R}^d) \),

\[
A \| f \|_{M^p_m} \leq \left( \sum_{\lambda \in \Lambda} \left( \sum_{j=1}^{n} \left| \left( f, \pi(\lambda)\varphi_j \right) \right|^2 \right)^{p/2} m(\lambda)^p \right)^{1/p} \leq B \| f \|_{M^p_m}.
\]

(ii) Assume in addition that \( \Lambda = a\mathbb{Z}^d \times b\mathbb{Z}^d \) is a separable lattice. Then a distribution \( f \) belongs to \( M^{p,q}_m(\mathbb{R}^d) \) if and only if each sequence \( C_{\varphi_j} f, \ell^p_m \) belongs to \( \ell^{p,q}_m(\mathbb{Z}^{2d}) \). In this case there exist constants \( A \) and \( B \) depending on \( p, q, m \) such that, for all \( f \in M^{p,q}_m \)

\[
A \| f \|_{M^{p,q}_m} \leq \left( \sum_{l \in \mathbb{Z}} \left( \sum_{k \in \mathbb{Z}} \left( \sum_{j=1}^{n} \left| \left( f, \pi(ak, bl)\varphi_j \right) \right|^2 \right)^{p/2} m(ak, bl)^p \right)^{q/p} \right)^{1/q} \leq B \| f \|_{M^{p,q}_m}.
\]

(iii) Let \( \Lambda \subseteq \mathbb{R}^{2d} \) be an arbitrary lattice and \( Q \) be a relatively compact fundamental domain of \( \Lambda \). Then a distribution \( f \) belongs to \( M^{p,q}_m(\mathbb{R}^d) \), if and only if the function \( \sum_{\lambda \in A} \left( \sum_{j=1}^{n} \left| \left( f, \pi(\lambda)\varphi_j \right) \right|^2 \right)^{1/2} \chi_{\lambda+Q} \) belongs to \( L^{p,q}_m(\mathbb{R}^{2d}) \). In this case there exist constants \( A, B > 0 \), such that, for all \( f \in M^{p,q}_m(\mathbb{R}^d) \),

\[
A \| f \|_{M^{p,q}_m} \leq \left\| \sum_{\lambda \in A} \left( \sum_{j=1}^{n} \left| \left( f, \pi(\lambda)\varphi_j \right) \right|^2 \right)^{1/2} \chi_{\lambda+Q} \right\|_{L^{p,q}_m} \leq B \| f \|_{M^{p,q}_m}.
\]

Note that (ii) follows from (iii), since for \( Q = [0,a]^d \times [0,b]^d \) the norm equivalence \( \| \sum_{k,l \in \mathbb{Z}^d} a_{kl} \chi_{(ak,bl)+Q} \|_{L^{p,q}_m} \propto \| a \|_{\ell^{p,q}_m} \) holds.

Theorem 2 has a long history. It extends the basic characterizations of modulation spaces by Gabor frames to multi-window Gabor frames. For Gabor frames with a single window and lattices of the form \( \Lambda = a\mathbb{Z}^d \times b\mathbb{Z}^d \) with \( ab \in Q \) Theorem 2 was proved in [16]. For general lattices it follows from the main result in [24] and the techniques in [16]. See also the discussion in [21, Chapter 13]. The proofs for multi-window Gabor frames require only few modifications, we therefore postpone a discussion to Appendix A.
2.4. A new characterization of multi-window Gabor frames

The proof of our main statement relies on a characterization of multi-window Gabor frames without using inequalities. The following lemma is a generalization of [22] from Gabor frames to multi-window Gabor frames.

Lemma 3. Assume that \( \varphi_j \in M^1(\mathbb{R}^d) \) for \( j = 1, \ldots, n \). Then the following properties are equivalent.

(i) \( \bigcup_{j=1}^n G(\varphi_j, \Lambda) \) is a multi-window Gabor frame for \( L^2(\mathbb{R}^d) \).

(ii) The analysis operator \( C_{\varphi, \Lambda} \) is one-to-one from \( M^\infty(\mathbb{R}^d) \) to \( \ell^\infty(\Lambda, \mathbb{C}^n) \).

The idea of the proof will be given in Appendix A, where we will also list many more equivalent conditions.

2.5. Properties of localization operators

We next recall some elementary properties of the localization operators \( H_{T_\lambda \sigma} \). Time-frequency localization operators have been introduced and studied by Daubechies [11,10] and Ramanathan and Topiwala [33], and are also called STFT multipliers, time-frequency Toeplitz operators, Wick operators, time-frequency filters, etc. They are a popular tool in signal analysis for time-frequency filtering or nonstationary filtering [31,34], in quantization procedures in physics [1], or in the approximation of pseudodifferential operators [9,30]. For a detailed account of the early theory we refer to Wong’s book [38], for a study of boundedness and Schatten class properties to [7,8,18,36].

Lemma 4 (Intertwining property). If \( \sigma \in L^\infty(\mathbb{R}^2d) \), \( \varphi \in L^2(\mathbb{R}^d) \), and \( \lambda \in \Lambda \), then

\[
\pi(\lambda) H_{\sigma} \pi(\lambda)^* = H_{T_\lambda \sigma}.
\]

The proof consists of a simple calculation, see [12, Lemma 2.6].

For estimates of the STFT of \( H_\sigma f \) we introduce the formal adjoint of \( \mathcal{V}_\varphi \), namely

\[
\mathcal{V}_\varphi^* F = \int_{\mathbb{R}^{2d}} F(z) \pi(z) \varphi \, dz,
\]

which maps functions on \( \mathbb{R}^{2d} \) to functions or distributions on \( \mathbb{R}^d \). With this notation we can write the localization operator \( H_\sigma \) as

\[
H_\sigma f = \mathcal{V}_\varphi^* (\sigma \mathcal{V}_\varphi f).
\]

The STFT of \( \mathcal{V}_\varphi^* F \) satisfies a fundamental pointwise estimate [21, Proposition 11.3.2]:

\[
|\mathcal{V}_\varphi (\mathcal{V}_\varphi^* F)(z)| \leq \big(|\mathcal{V}_\varphi \varphi| * |F|\big)(z), \quad \forall z \in \mathbb{R}^{2d}.
\]

(15)
We note that for $F = \sigma \mathcal{V}_\varphi f$ this estimate becomes

$$
\left| \mathcal{V}_\varphi (H_\sigma f) (z) \right| = \left| \mathcal{V}_\varphi (\mathcal{V}_\varphi^* (\sigma \mathcal{V}_\varphi f)) (z) \right| \leq \left( |\mathcal{V}_\varphi \varphi| * (|\sigma| |\mathcal{V}_\varphi f|) \right) (z).
$$

(16)

Thus the short-time Fourier transform of $H_\sigma$ is a so-called product-convolution operator. The standard boundedness results for localization operators can be easily deduced from the well-established results for product convolution operators [6].

Estimate (16) is quite useful for the derivation of norm estimates. In the following, we fix a non-negative symbol $\sigma$ and investigate the set of operators $\{ H_{T_\lambda \sigma} : \lambda \in \Lambda \}$. To simplify notation we will write $H_\lambda$ instead of $H_{T_\lambda \sigma}$, and sometimes $H_0 = H_\sigma$ by some abuse of notation.

**Lemma 5.**

(i) Assume that $\sigma \in L^1(\mathbb{R}^{2d})$, $\sigma \geq 0$ and that $\varphi \in L^2(\mathbb{R}^d)$. Then each $H_\lambda$, $\lambda \in \Lambda$, is a positive trace-class operator.

(ii) If, in addition, $\varphi \in M^1(\mathbb{R}^d)$ and $\sigma \in L^1_{\nu}(\mathbb{R}^{2d})$, then each $H_\lambda$ is bounded from $M^\infty(\mathbb{R}^d)$ into $M^1_{1/\nu}(\mathbb{R}^d)$. In particular, all eigenfunctions $\varphi_j$ of $H_\sigma$ belong to $M^1_{1/\nu}(\mathbb{R}^d)$.

(iii) Furthermore, if $\varphi \in M^1_{\nu}(\mathbb{R}^d)$ and $\sigma \in L^1_{\nu}(\mathbb{R}^{2d})$, then each $H_\lambda$ is bounded from $M^\infty_{1/\nu}(\mathbb{R}^d)$ into $L^2(\mathbb{R}^d)$.

**Proof.** Statement (i) is well known, see, e.g., [2,17,38]. To show (ii), we use (16) to obtain, for $f \in M^\infty(\mathbb{R}^d)$,

$$
\| H_\sigma f \|_{M^1_{1/\nu}} = \| \mathcal{V}_\varphi (H_\sigma f) \|_{L^1_{1/\nu}} = \| \mathcal{V}_\varphi \mathcal{V}_\varphi^* (\sigma \mathcal{V}_\varphi f) \|_{L^1_{1/\nu}} \\
\leq \| |\mathcal{V}_\varphi \varphi| * |\sigma| |\mathcal{V}_\varphi f| \|_{L^1_{1/\nu}} \\
\leq \| \mathcal{V}_\varphi \varphi \|_{L^1_{1/\nu}} \| \sigma \mathcal{V}_\varphi f \|_{L^1_{1/\nu}},
$$

(17)

where we have used Young’s inequality. Since $\varphi \in M^1_{\nu}(\mathbb{R}^d)$ if and only if $\mathcal{V}_\varphi \varphi \in L^1_{\nu}(\mathbb{R}^{2d})$ by [21, Proposition 12.1.2], we find that

$$
\| H_\sigma f \|_{M^1_{1/\nu}} \leq \| \mathcal{V}_\varphi \varphi \|_{L^1_{1/\nu}} \| \sigma \|_{L^1_{\nu}} \| \mathcal{V}_\varphi f \|_{L^\infty} \leq C \| \sigma \|_{L^1_{\nu}} \| f \|_{M^\infty},
$$

and thus $H_\sigma$ is bounded from $M^\infty(\mathbb{R}^d)$ to $M^1_{1/\nu}(\mathbb{R}^d)$.

The proof of (iii) is similar. Again, we apply (16) to obtain for $f \in M^\infty_{1/\nu}(\mathbb{R}^d)$:

$$
\| H_\sigma f \|_{L^2} \leq \| |\mathcal{V}_\varphi \varphi| * |\sigma| \mathcal{V}_\varphi f | \|_{L^2} \leq \| \mathcal{V}_\varphi \varphi \|_{L^2} \| \sigma \mathcal{V}_\varphi f \|_{L^1_{1/\nu}}.
$$

Hence, the result follows from

$$
\| \sigma \mathcal{V}_\varphi f \|_{L^1_{1/\nu}} = \int_{\mathbb{R}^{2d}} \sigma (z) |\mathcal{V}_\varphi f (z)| \frac{1}{v(z)} \, dz \\
\leq \| \sigma \|_{L^1_{\nu}} \| f \|_{M^\infty_{1/\nu}},
$$

\(\square\)
The spectral theorem for compact self-adjoint operators provides the following spectral representation of $H_\lambda$.

**Corollary 6.** Assume $\varphi \in M^1_1(\mathbb{R}^d)$ and $\sigma \in L^1_{\nu}(\mathbb{R}^{2d})$. Then there exists a positive sequence of eigenvalues $c = (c_j) \in \ell^1$ and an orthonormal system of eigenfunctions $\varphi_j \in M^1_1(\mathbb{R}^d)$, such that

$$H_\sigma f = \sum_{j=1}^{\infty} c_j \langle f, \varphi_j \rangle \varphi_j. \quad (18)$$

It follows that

$$H_\lambda f = H T_\lambda \sigma f = \pi(\lambda) H_\sigma \pi(\lambda)^* f = \sum_{j=1}^{\infty} c_j \langle f, \pi(\lambda) \varphi_j \rangle \pi(\lambda) \varphi_j, \quad (19)$$

and $\{\pi(\lambda) \varphi_j, \ j \in \mathbb{N}\}$ is an orthonormal system of eigenfunctions of $H_\lambda$.

A priori, the spectral representation of $H_\lambda$ holds only for $f \in L^2(\mathbb{R}^d)$. The next corollary extends the spectral representation to all of $M_{1/\nu}^\infty(\mathbb{R}^d)$.

**Corollary 7.** The expansion for $H_\lambda f$ given in (19) is well defined on $M_{1/\nu}^\infty(\mathbb{R}^d)$ and converges to $H_\lambda f$ in $L^2$ for all $f \in M_{1/\nu}^\infty(\mathbb{R}^d)$.

**Proof.** Without loss of generality, we assume $\lambda = 0$ and set $H = H_\sigma$. Since $H f \in L^2(\mathbb{R}^d)$ for every $f \in M_{1/\nu}^\infty(\mathbb{R}^d)$ by Lemma 5(iii), we can expand $H f$ with respect to the orthonormal system of eigenfunctions of $H$ and obtain that

$$H f = \sum_{j=1}^{\infty} \langle H f, \varphi_j \rangle \varphi_j + r \quad (20)$$

for some $r \in L^2(\mathbb{R}^d)$ in the orthogonal complement of span$\{\varphi_j: \ j \in \mathbb{N}\}$. As $H$ is self-adjoint on $L^2(\mathbb{R}^d)$, we also have $\langle H f, \varphi_j \rangle = \langle f, H \varphi_j \rangle = c_j \langle f, \varphi_j \rangle$, and consequently

$$H f = \sum_{j=1}^{\infty} c_j \langle f, \varphi_j \rangle \varphi_j + r. \quad (21)$$

We need to show that $r = 0$. Since $r \in L^2(\mathbb{R}^d)$ is orthogonal to all eigenfunctions $\varphi_j$, we find that $\langle H f, r \rangle = \|r\|^2_2$.

To show $r = 0$, we first observe that $\langle H h, r \rangle = 0$ for all $h \in L^2(\mathbb{R}^d)$ by (18). Since $L^2(\mathbb{R}^d)$ is $w^*$-dense in $M_{1/\nu}^\infty(\mathbb{R}^d)$, we may choose an approximating sequence $f_n \in L^2(\mathbb{R}^d)$ such that $f_n \xrightarrow{w^*} f \in M_{1/\nu}^\infty(\mathbb{R}^d)$. For instance, $f_n$ may be chosen as

$$f_n = \int_{\mathbb{R}^{2d}} \chi_{B_n}(z) \mathcal{V}_{g^*} f(z) \pi(z) g \, dz,$$
where $B_n$ is the ball with radius $n$ and centered at 0. Furthermore, since $f_n \overset{w^*}{\to} f$, we obtain in particular that $\mathcal{V}_\varphi f_n$ converges to $\mathcal{V}_\varphi f$ uniformly on compact sets [13, Theorem 4.1]. Consequently

$$0 = \langle Hf_n, r \rangle = \int_{\mathbb{R}^d} \sigma(z) \mathcal{V}_\varphi f_n(z) \mathcal{V}_\varphi r(z) \, dz \to \int_{\mathbb{R}^d} \sigma(z) \mathcal{V}_\varphi f(z) \mathcal{V}_\varphi r(z) \, dz = \langle Hf, r \rangle = \|r\|_2^2.$$  

This shows that $r = 0$ and so the series (19) represents $Hf$ for all $f \in \mathcal{M}^\infty_{1/\nu}(\mathbb{R}^d)$.  

### 3. From local information to global information

We first state and prove the main result for the modulation spaces $\mathcal{M}^p_m(\mathbb{R}^d)$. The generalizations to $\mathcal{M}^p,q_m(\mathbb{R}^d)$ will be discussed later. As always, $\nu$ denotes a submultiplicative, even weight function on $\mathbb{R}^{2d}$ satisfying the condition $\lim_{n \to \infty} \nu(nz)^{1/n} = 1$ for all $z \in \mathbb{R}^{2d}$.

**Theorem 8.** Let $\sigma \in L^1_\nu(\mathbb{R}^{2d})$ be a non-negative symbol satisfying the condition

$$A \leq \sum_{\lambda \in \Lambda} T_\lambda \sigma \leq B, \quad a.e. \quad (22)$$

for two constants $A, B > 0$. Assume that $\varphi \in \mathcal{M}^1_\nu(\mathbb{R}^d)$. Then for every $\nu$-moderate weight $m$ and $1 \leq p < \infty$ the distribution $f \in \mathcal{M}^\infty_{1/\nu}(\mathbb{R}^d)$ belongs to $\mathcal{M}^p_m(\mathbb{R}^d)$, if and only if

$$\left( \sum_{\lambda \in \Lambda} \|H_\lambda f\|_2^p m(\lambda)^p \right)^{1/p} < \infty, \quad (23)$$

and the expression in (23) is an equivalent norm on $\mathcal{M}^p_m(\mathbb{R}^d)$.

Similarly, for $p = \infty$ we obtain the norm equivalence

$$\|f\|_{\mathcal{M}^\infty_m} \asymp \sup_{\lambda \in \Lambda} \|H_\lambda f\|_{2m(\lambda)}. \quad (24)$$

The norm equivalence supports the interpretation that $H_\lambda f$ carries the local time-frequency information about $f$ near $\lambda \in \mathbb{R}^{2d}$. By combining the local pieces $H_\lambda f$, one obtains the global time-frequency information as it is measured by modulation space norms.

The proof of Theorem 8 requires some preparations. We first show that finitely many eigenfunctions of $H_0 = \mathcal{V}_\varphi^* \sigma \mathcal{V}_\varphi$ generate a multi-window Gabor frame for $L^2(\mathbb{R}^d)$. With this crucial step in place, Theorem 8 can then be deduced from the characterization of modulation spaces by means of Gabor frames.

#### 3.1. Multi-window Gabor frames

**Lemma 9.** Assume that $\sigma \in L^1(\mathbb{R}^{2d})$ and $\sum_{\lambda \in \Lambda} T_\lambda \sigma \propto 1$, and that $\varphi \in \mathcal{M}^1_\nu(\mathbb{R}^d)$. Let $\{\varphi_j : j \in \mathbb{N}\}$ be the orthonormal system of eigenfunctions of $H_0$. Then there exists $n \in \mathbb{N}$, such that the finite union $\bigcup_{j=1}^n \mathcal{G}(\varphi_j, \Lambda)$ is a multi-window Gabor frame for $L^2(\mathbb{R}^d)$.
An analogous statement was proved and used in [12] for the lattice $\Lambda = \mathbb{Z}^{2d}$ and rational lattices by means of Zak transform methods. In the case of general lattices we cannot apply Zak-transform methods. As a substitute, we will use a finite intersection property for $\Lambda$-invariant subspaces of $M^\infty$. The following statement may be of interest in its own right.

**Lemma 10.** Assume that $W_n$ is a sequence of $w^*$-closed subspaces in $M^\infty(\mathbb{R}^d)$ such that

(i) $W_n \supseteq W_{n+1} \neq \{0\}$ for all $n \in \mathbb{N}$, and

(ii) $W_n$ is invariant under all operators $\pi(\lambda)$ for $\lambda \in \Lambda$.

Then $\bigcap_{n \geq 1} W_n \neq \{0\}$.

**Proof.** Let $Q$ be the closure of a relatively compact fundamental domain of $\Lambda$, for instance, if $\Lambda = A\mathbb{Z}^{2d}$, then $Q = A[0, 1]^{2d}$. We first choose a sequence $h_n \in W_n$ with $\|h_n\|_{M^\infty} = \sup_{z \in \mathbb{R}^{2d}} |\phi h_n(z)| = 1$. Then there exists a sequence of points $\lambda_n$ in $\Lambda$, such that

$$\sup_{z \in Q} |\phi(\pi(\lambda_n) h_n)(z)| = 1.$$ 

Since $W_n$ is invariant under all $\pi(\lambda)$, $\lambda \in \Lambda$, the distribution $f_n = \pi(\lambda_n)h_n$ is in $W_n$.

Next we show that the set of restrictions $\{\phi f_n|_Q\}$ is equicontinuous. We have

$$|\phi f_n(z) - \phi f_n(\xi)| = \|f_n(\pi(z) - \pi(\xi))\phi\| \leq \|f_n\|_{M^\infty} \cdot \|\phi(z - \pi(\xi))\phi\|_{M^1}. \quad (25)$$

Since $\|f_n\|_{M^\infty} = \|\pi(\lambda_n) h_n\|_{M^\infty} = 1$, the equicontinuity follows from the strong continuity of time-frequency shifts on $M^1(\mathbb{R}^d)$.

We next choose $z_n \in Q$ with $|\phi f_n(z_n)| \geq \frac{1}{2}$. Since the unit ball in $M^\infty(\mathbb{R}^d)$ is $w^*$-compact, there exists a subsequence $f_{n_k}$ that converges to some $f \in M^\infty(\mathbb{R}^d)$ in the $w^*$-sense. Furthermore, by compactness of $Q$, there also exists a subsequence $z_{\ell}$ of $z_{nk}$, such that $z_{\ell} \rightarrow z \in Q$. Hence, by equicontinuity,

$$\phi f_{\ell}(z_{\ell}) \rightarrow \phi f(z).$$

Since $|\phi f_{\ell}(z_{\ell})| \geq 1/2$, we conclude that also $|\phi f(z)| \geq 1/2$, and consequently $f \neq 0$.

By construction, $f_\ell \in W_m$ for every $\ell \geq m$, hence we obtain $f = w^* - \lim_{\ell \rightarrow \infty} f_\ell \in W_m$ for all $m$, because $W_m$ is $w^*$-closed. To summarize, we have constructed a non-zero $f \in M^\infty(\mathbb{R}^d)$ that is in $W_m$ for all $m$. □

**Proof of Lemma 9.** To prove that finitely many eigenfunctions generate a multi-window Gabor frame with respect to the lattice $\Lambda$, we assume on the contrary that $\bigcup_{j=1}^n \mathcal{G}(\varphi_j, \Lambda)$ is not a frame for every $n \in \mathbb{N}$. Using Lemma 10 and Lemma 3, we will derive a contradiction to the assumption that $A \leq \sum_{\lambda \in \Lambda} T_{\lambda\sigma} \leq B$.

We use the criterion of Lemma 3. Let $\varphi_n = (\varphi_1, \ldots, \varphi_n)$ be the vector-valued function consisting of the first $n$ eigenfunctions of $H_0$, and

$$W_n = \ker(C_{\varphi_n, \Lambda}) = \{ f \in M^\infty(\mathbb{R}^d) : \langle f, \pi(\lambda) \varphi_j \rangle = 0, \forall \lambda \in \Lambda, j = 1, \ldots, n \}$$

be the kernel of the coefficient operator $C_{\varphi_n, \Lambda}$ in $M^\infty(\mathbb{R}^d)$. 

If \( \bigcup_{n}^{\infty} G(\psi_j, \Lambda) \) is not a frame, then \( \mathcal{W}_n \) is a non-trivial subspace of \( M^{\infty}(\mathbb{R}^d) \) by Lemma 3. By construction, the \( \mathcal{W}_n \)'s form a nested sequence of \( w^* \)-closed subspaces of \( M^{\infty}(\mathbb{R}^d) \), and they are also invariant under \( \pi(\lambda), \lambda \in \Lambda \). Thus the assumptions of Lemma 10 are satisfied, and we conclude that \( \bigcap_{n}^{\infty} \mathcal{W}_n \neq \{0\} \). This means that there exists a non-zero \( f \in M^{\infty}(\mathbb{R}^d) \), such that \( \langle f, \pi(\lambda)\psi_j \rangle = 0 \) for all \( \lambda \in \Lambda \) and all \( j \in \mathbb{N} \).

We now consider \( H_{\lambda} f \). Since \( H_{\lambda} f \in M^{1}(\mathbb{R}^d) \) by Lemma 5, the bracket \( \langle H_{\lambda} f, f \rangle \) is well defined and given by

\[
\langle H_{\lambda} f, f \rangle = \int_{\mathbb{R}^{2d}} \sigma(z-\lambda) |\mathcal{V}_{\psi} f(z)|^2 \, dz. \tag{27}
\]

On the other hand, the extended spectral representation of Lemma 7 and (26) imply that

\[
H_{\lambda} f = \sum_{j=1}^{\infty} c_j \langle f, \pi(\lambda)\psi_j \rangle \pi(\lambda)\psi_j = 0. \tag{28}
\]

Consequently \( \langle H_{\lambda} f, f \rangle = 0 \) for all \( \lambda \in \Lambda \), and \( |\mathcal{V}_{\psi} f(z)|^2 \) vanishes on \( \bigcup_{\lambda \in \Lambda} \text{supp} T_{\lambda} \sigma \).

According to the crucial assumption (22) we have \( \sum_{\lambda \in \Lambda} T_{\lambda} \sigma > A > 0 \) almost everywhere, and thus \( \bigcup_{\lambda \in \Lambda} \text{supp}(T_{\lambda} \sigma) = \mathbb{R}^{2d} \). Therefore, (27) and (28) imply that \( \mathcal{V}_{\psi} f = 0 \), from which \( f = 0 \) follows. This is a contradiction to \( f \) being a non-zero element in \( \bigcap_{n}^{\infty} \mathcal{W}_n \).

This contradiction shows that there exists an \( n \in \mathbb{N} \), such that \( \bigcup_{j=1}^{n} G(\psi_j, \Lambda) \) is a multi-window Gabor frame, and we are done. \( \square \)

**Remark 1.** Note that for finite-rank operators \( H_0 \), it can be seen directly that the finite set of eigenvectors generates a multi-window Gabor frame for \( \Lambda \).

### 3.2. Proof of Theorem 8

We are now ready to prove the main theorem. We observe that for \( f \in M^{\infty}_{1/v}(\mathbb{R}^d) \), \( H_{\lambda} f \in L^{2}(\mathbb{R}^d) \) by Lemma 5(iii). Thus the terms in (23) are well defined.

First assume that \( p < \infty \) and \( f \in M^{p}_{m}(\mathbb{R}^d) \subseteq M^{\infty}_{1/v}(\mathbb{R}^d) \). Using the embedding \( M^{1}(\mathbb{R}^d) \hookrightarrow L^{2}(\mathbb{R}^d) \) and the estimate (17) with \( \nu \equiv 1 \), we majorize \( \|H_{\lambda} f\|_{2} \) as follows

\[
\|H_{\lambda} f\|_{2} \leq C_{\psi} \|H_{\lambda} f\|_{M^{1}} \leq C_{\psi} \|T_{\lambda} \sigma \cdot \mathcal{V}_{\psi} f\|_{1} \|\mathcal{V}_{\psi} \psi\|_{1} \\
= C_{\psi} C \int_{\mathbb{R}^{2d}} |\sigma(z-\lambda)| \cdot |\mathcal{V}_{\psi} f(z)| \, dz \\
= C_{\psi} C (|\mathcal{V}_{\psi} f| * \sigma^{\vee})(\lambda), \tag{29}
\]

where \( \sigma^{\vee}(z) = \sigma^{\vee}(-z) \). Thus \( \|H_{\lambda} f\|_{2} \) is majorized by a sample of \( |\mathcal{V}_{\psi} f| * \sigma^{\vee} \). To proceed further, we use the fact that \( \mathcal{V}_{\psi} f \in W(L^{p}_{m}) \) and \( \|\mathcal{V}_{\psi} f\|_{W(L^{p}_{m})} \leq C_{0} \|\psi\|_{M^{1}} \|f\|_{M^{p}_{m}} \) for \( \varphi \in M^{1}_{1}(\mathbb{R}^d) \).
and \( f \in M^p_m(\mathbb{R}^d) \) by [21, Theorem 12.2.1]. Now the convolution relation (10) and the sampling inequality (11) imply that

\[
\sum_\lambda \| H_\lambda f \|^p_{m(\lambda)} \leq C \varphi C \left\{ \left( \sigma^\vee |V_\varphi f| \right) \right\}_A \leq C \varphi C A \left\{ \sigma \right\}_{L^p} \leq C \varphi C C \left\{ \left| \sigma \right| \right\}_{L^p} \leq C \varphi C C \left\{ \left| \sigma \right| \right\}_{W(L^p)} \leq C \varphi C C \left\{ \left| \sigma \right| \right\}_{L^p} \| f \|^p_{M^p_m}. \tag{30}
\]

The same argument yields \( \sup_\lambda \| H_\lambda f \|^p_{2 m(\lambda)} \leq C \| f \|^p_{M^\infty_m} \).

Hence, for \( 1 \leq p \leq \infty \), the mapping \( f \mapsto (\| H_\lambda f \|^p_{2})_{\lambda \in A} \) is bounded from \( M^p_m(\mathbb{R}^d) \) to \( \ell^p_m(A) \).

Conversely, assume that \( p < \infty \) and

\[
\sum_\lambda \| H_\lambda f \|^p_{2 m(\lambda)} < \infty.
\]

We need to show that \( f \in M^p_m(\mathbb{R}^d) \). Since \( \| H_\lambda f \|^p_{2} = \| H_\lambda f \|^p_{m(\lambda)} \), we have the inequality

\[
\sum_\lambda \left| \langle H_\lambda f, g_\lambda \rangle \right|^p m(\lambda) \leq \sum_\lambda \| H_\lambda f \|^p_{2 m(\lambda)} < \infty.
\]

for arbitrary sequences \( g_\lambda \in L^2(\mathbb{R}^d) \) with \( \| g_\lambda \|^2 = 1 \). Applying the eigenfunction expansion of Corollary 6, we obtain

\[
\sum_\lambda \left| \sum_{j=1}^{\infty} c_j \langle f, \pi(\lambda)\varphi_j \rangle \langle \pi(\lambda)\varphi_j, g_\lambda \rangle \right|^p m(\lambda) \leq \sum_\lambda \| H_\lambda f \|^p_{2 m(\lambda)} < \infty. \tag{31}
\]

Now fix \( j_0 \in \mathbb{N} \) and set \( g_\lambda = \pi(\lambda)\varphi_{j_0} \) for \( \lambda \in A \). Since the eigenfunctions of \( H_\lambda \) are orthonormal, the sum over \( j \) collapses to a single term, and (31) becomes

\[
\sum_\lambda \left| \langle H_\lambda f, g_\lambda \rangle \right|^p m(\lambda) = \sum_\lambda \left| c_{j_0} \langle f, \pi(\lambda)\varphi_{j_0} \rangle \right|^p m(\lambda) \leq \sum_\lambda \| H_\lambda f \|^p_{2 m(\lambda)} < \infty.
\]

The last inequality holds for every \( j_0 \in \mathbb{N} \). After summing over finitely many \( j_0 \) and switching to the \( \ell^2 \)-norm on \( \mathbb{C}^n \), we obtain the inequality

\[
\sum_\lambda \left( \sum_{j=1}^{n} \left| \langle f, \pi(\lambda)\varphi_j \rangle \right|^2 \right)^{1/2} m(\lambda) \leq \sum_{j=1}^{n} \sum_\lambda \left| \langle f, \pi(\lambda)\varphi_j \rangle \right|^p m(\lambda) \leq \left( \sum_{j=1}^{n} \frac{1}{\pi_j^p} \right) \sum_\lambda \| H_\lambda f \|^p_{2 m(\lambda)} < \infty. \tag{32}
\]
We now apply Lemma 9 and choose an $n \in \mathbb{N}$, such that $\bigcup_{j=1}^{p} \mathcal{G}(\varphi_j, \Lambda)$ is a multi-window Gabor frame for $L^2(\mathbb{R}^d)$. Since all $\varphi_j$ are in $M^1(\mathbb{R}^d)$, the fundamental characterization of modulation spaces (Section 2.3) is valid. Thus Theorem 2(i) implies that $f \in M_{p,m}^p(\mathbb{R}^d)$.

If $p = \infty$ and $\sup_{\lambda \in \Lambda} \| H_\lambda f \|_{2m(\lambda)} < \infty$, then, by choosing $g_\lambda$ as before, we find

$$c_{j_0} \sup_{\lambda} \| f, \pi(\lambda) \varphi_{j_0} \|_{m(\lambda)} \leq \sup_{\lambda} \| H_\lambda f \|_{2m(\lambda)} < \infty$$

for every $j_0$.

Arguing as above, Theorem 2 says that $\| f \|_{M_\infty^{1/\nu}(\mathbb{R}^d)} \leq C \max_{j=1, \ldots, n} \sup_{\lambda} \| f, \pi(\lambda) \varphi_{j_0} \|_{m(\lambda)} \leq \left( \max_{j=1, \ldots, n} c_j \right) \sup_{\lambda} \| H_\lambda f \|_{2m(\lambda)} < \infty$.

Combining (30) and (32), we have shown that $\| f \|_{M_{p,m}^p(\mathbb{R}^d)}$ and $(\sum_{\lambda \in \Lambda} \| H_\lambda f \|_{2m(\lambda)}^p)^{1/p}$ for $1 \leq p < \infty$ (or $\sup_{\lambda \in \Lambda} \| H_\lambda f \|_{2m(\lambda)}$ for $p = \infty$) are equivalent norms on $M_{p,m}^p(\mathbb{R}^d)$.

### 3.3. Variations of Theorem 8

In order to formulate our main result for mixed-norm spaces and arbitrary lattices, we have to resort to the theory of coorbit spaces, as introduced in [13,14]. In particular, for arbitrary lattices, a sequence $(c_\lambda)_{\lambda \in \Lambda}$ is in the sequence spaces associated with $L_{p,q}^{p,q}(\mathbb{R}^d)$, if $\sum_{\lambda \in \Lambda} c_\lambda \chi_{\lambda + Q}$ is in $L_{p,q}^{p,q}(\mathbb{R}^d)$ for some fundamental domain $Q$ of $\Lambda$. With this definition, we may give the following characterization.

**Theorem 11.** Let $\Lambda$ be an arbitrary lattice in $\mathbb{R}^d$ and $Q$ be a relatively compact fundamental domain $Q$. Assume the same conditions on $\sigma$ and $\varphi$ as in Theorem 8. Then a distribution $f \in M_{1/\nu}^{1}(\mathbb{R}^d)$ belongs to $M_{p,q}^{p,q}(\mathbb{R}^d)$, $1 \leq p, q \leq \infty$, if and only if

$$\sum_{\lambda \in \Lambda} \| H_\lambda f \|_{2 \chi_{\lambda + Q}} \in L_{p,q}^{p,q}(\mathbb{R}^d),$$

and $\| \sum_{\lambda \in \Lambda} \| H_\lambda f \|_{2 \chi_{\lambda + Q}} \|_{L_{p,q}^{p,q}} \asymp \| f \|_{M_{p,q}^{p,q}}$.

**Proof.** The proof is almost identical to the proof of Theorem 8. The only modifications occur in (30), which has to be replaced by

$$\left\| \sum_{\lambda \in \Lambda} \| H_\lambda f \|_{2 \chi_{\lambda + Q}} \right\|_{L_{p,q}^{p,q}} \leq \sum_{\lambda \in \Lambda} \| \mathcal{V}_\varphi f \ast \tilde{\sigma}(\lambda) \chi_{\lambda + Q} \|_{L_{p,q}^{p,q}} \leq C \| \mathcal{V}_\varphi f \ast \tilde{\sigma} \|_{W(L_{m}^{p,q})}.$$  

Likewise, in (32) we replace the weighted $L_{p,q}^{p,q}$-norm by the general $L_{m}^{p,q}$-norm. 

For a separable lattice $\Lambda = a\mathbb{Z}^d \times b\mathbb{Z}^d$ the norm in (33) is just the $L_{m}^{p,q}$-norm on $\mathbb{Z}^2d$ with $\tilde{m}(k,n) = m(ak,bn)$. In this case, $\lambda = (ka,nb)$, $k,n \in \mathbb{Z}^d$ and we may write $H_\lambda f = H_{k,n} f$. 


Corollary 12. Let \( \Lambda = a\mathbb{Z}^d \times b\mathbb{Z}^d \) be a separable lattice and assume the same conditions on \( \sigma \) and \( \varphi \) as in Theorem 8. Then a distribution \( f \in M_{1/\nu}^\infty(\mathbb{R}^d) \) belongs to \( M_{p,q}^{m,q}(\mathbb{R}^d) \) for \( 1 \leq p, q < \infty \), if and only if

\[
\left( \sum_{n \in \mathbb{Z}^d} \left( \sum_{k \in \mathbb{Z}^d} \| H_{k,n} f \|_{2}^{p} m(ka, nb)^{p} \right)^{q/p} \right)^{1/q} < \infty, \tag{34}
\]

and (34) defines an equivalent norm on \( M_{m,q}^{p,q}(\mathbb{R}^d) \). The result holds for \( p = \infty \) or \( q = \infty \) with the usual modifications.

3.4. Existence of multi-window Gabor frames and properties of the eigenfunctions \( \varphi_j \)

We finally point out some immediate consequences of our results and methods. The intermediate results leading to Theorem 8 also imply the existence of multi-window Gabor frames for general lattices.

Theorem 13. Let \( \Lambda \) be an arbitrary lattice and \( \nu \) a submultiplicative weight on \( \mathbb{R}^{2d} \). Then there exist finitely many functions \( \varphi_j \in M_{1}^{1}(\mathbb{R}^d) \), such that \( \bigcup_{j=1}^{n} \mathcal{G}(\varphi_j, \Lambda) \) is a multi-window Gabor frame for \( L^2(\mathbb{R}^d) \).

Proof. Choose \( \sigma \in L_{1\nu}^{1}(\mathbb{R}^{2d}) \) such that \( \sum_{\lambda \in \Lambda} T_{\lambda} \sigma = 1 \) and fix a window \( \varphi \in M_{1}^{1}(\mathbb{R}^d) \). For instance, one may choose the characteristic function \( \chi_Q \) of a (relatively compact) fundamental domain of \( \Lambda \) and the Gaussian window \( \varphi(t) = e^{-\pi t \cdot t} \).

Now consider the localization operator \( H_{0} = \mathcal{V}_{\varphi}^{*} \varphi \mathcal{V}_{\varphi} \). According to Lemma 5(ii), all eigenfunctions \( \varphi_j \) of \( H_{0} \) belong to \( M_{1}^{1}(\mathbb{R}^d) \). Lemma 9 states that for some finite \( n \in \mathbb{N} \) the set \( \bigcup_{j=1}^{n} \mathcal{G}(\varphi_j, \Lambda) \) is a multi-window Gabor frame for \( L^2(\mathbb{R}^d) \). \( \square \)

The existence of multi-window Gabor frames for general lattices was known before. On the one hand, it is an immediate consequence of coorbit theory applied to the Heisenberg group. To be more precise, according to [15, Theorem 7] for every lattice \( \Lambda \) and every non-zero \( g \in M_{1}^{1}(\mathbb{R}^d) \) there exists \( n \in \mathbb{N} \), such that the set \( \mathcal{G}(g, \frac{1}{n} \Lambda) \) is a Gabor frame for \( L^2(\mathbb{R}^d) \). Using a coset decomposition \( \frac{1}{n} \Lambda = \bigcup (\mu + \Lambda) \) for suitable \( \mu \in \Lambda \), one sees that \( \mathcal{G}(g, \frac{1}{n} \Lambda) = \bigcup \mathcal{G}(\pi(\mu)g, \Lambda) \) is a multi-window Gabor frame with all windows \( \pi(\mu)g \) derived from a single window \( g \). Recently Luef [32] proved the existence of multi-window Gabor frames by exploiting a connection between Gabor analysis and non-commutative geometry. Our methods provide a third, independent proof for this interesting result.

The construction of multi-window Gabor frames in Proposition 13 yields more detailed information about the frame generators, since they are eigenfunctions of a localization operator. Intuitively the eigenfunctions corresponding to the largest eigenvalues of a localization operator concentrate their energy on the essential support of the symbol \( \sigma \) of \( H_{0} \). For the special case of compactly supported \( \sigma \), this intuition is made precise by the following result.

Proposition 14. Let the non-negative function \( \sigma \in L_{1}(\mathbb{R}^{2d}) \) be supported in a compact set \( \Omega \) in \( \mathbb{R}^{2d} \) with \( 0 \leq \sigma(z) \leq C_{\sigma} < \infty \) for \( z \in \Omega \). Consider the localization operator given by \( H_{\sigma} f = \mathcal{V}_{\varphi}^{*} \varphi \mathcal{V}_{\varphi} f \) with \( \varphi \in M_{1}^{1}(\mathbb{R}^d) \), \( \| \varphi \|_{2} = 1 \) and spectral representation as in Corollary 6.
Then the eigenfunctions $\varphi_j$ of $H_\sigma$ satisfy the following time-frequency concentration

$$\int_\Omega |\mathcal{V}_\varphi \varphi_j(z)|^2 \, dz \geq \frac{c_j}{C_\sigma}. \quad (35)$$

Equality holds, if and only if $\sigma(z)/C_\sigma = \chi_\Omega(z)$ is the characteristic function of $\Omega$.

**Proof.** Using the weak interpretation of $H_\sigma$ from (2), we obtain

$$\int_\Omega |\mathcal{V}_\varphi \varphi_j(z)|^2 \, dz \geq \frac{1}{C_\sigma} \int_\Omega \sigma(z) |\mathcal{V}_\varphi \varphi_j(z)|^2 \, dz$$

$$= \frac{1}{C_\sigma} \langle H_\sigma \varphi_j, \varphi_j \rangle = \frac{c_j}{C_\sigma} \|\varphi_j\|_2^2 = \frac{c_j}{C_\sigma}. \quad \square$$

**Appendix A. Characterizations of modulation spaces and multi-window Gabor frames**

In the appendix, we will sketch the proof of Theorem 2 and formulate a series of new characterizations of multi-window Gabor frames. These statements generalize well-known facts from Gabor analysis and the results about Gabor frames without inequalities in [22].

For the investigation of multi-window Gabor frames we need the dual concept of vector-valued Gabor systems. In this case we consider the Hilbert space $\mathcal{H} = L^2(\mathbb{R}^d, \mathbb{C}^n)$ consisting of all vector-valued functions $f(t) = (f_1(t), \ldots, f_n(t))$ with the inner product

$$\langle f, \varphi \rangle_{L^2(\mathbb{R}^d, \mathbb{C}^n)} = \sum_{j=1}^n \int f_j(t) \overline{\varphi_j(t)} \, dt = \sum_{j=1}^n \langle f_j, \varphi_j \rangle_{L^2(\mathbb{R}^d)}. \quad (A.1)$$

Time-frequency-shifts act coordinate-wise on $f$. The vector-valued Gabor system $\mathcal{G}(\varphi, \Lambda) = \{\pi(\lambda) \varphi : \lambda \in \Lambda\}$ is a Riesz sequence in $L^2(\mathbb{R}^d, \mathbb{C}^n)$, if there exist constants $0 < A, B < \infty$ such that for all finitely supported sequences $c$,

$$A\|c\|_2^2 \leq \left\| \sum_{\lambda \in \Lambda} c_\lambda \pi(\lambda) \varphi \right\|_{L^2(\mathbb{R}^d, \mathbb{C}^n)}^2 \leq B\|c\|_2^2. \quad (A.2)$$

We now proceed to the proof of Theorem 2. The crucial step is to show the invertibility of the frame operator on $M^1_\nu(\mathbb{R}^d)$. This step requires a special representation of the frame operator due to Janssen [28] and at its core uses “Wiener’s lemma for twisted convolution” [24].

For $\varphi_j, \psi_j$ in $M^1(\mathbb{R}^d), j = 1, \ldots, n$, we denote frame-type operators by

$$S_{\varphi, \varphi} f = \sum_{\lambda \in \Lambda} \sum_{j=1}^n \langle f, \pi(\lambda) \varphi_j \rangle \pi(\lambda) \psi_j = \sum_{j=1}^n S_{\varphi_j, \psi_j}. \quad \text{The frame operator of the Gabor system } \bigcup_{j=1}^n \mathcal{G}(\varphi_j, \Lambda) \text{ is } S = S_{\varphi, \varphi}. \quad \text{We usually omit the reference to the lattice } \Lambda \text{ and the windows } \varphi_j.$$
The volume \( s(\Lambda) \) of a lattice \( \Lambda = A \mathbb{Z}^{2d} \) is defined as the measure of a fundamental domain of \( \Lambda \) and is \( |\det(A)| \). The adjoint lattice of \( \Lambda \) is \( \Lambda^\circ = \{ \mu \in \mathbb{R}^{2d} : \pi(\lambda)\pi(\mu) = \pi(\mu)\pi(\lambda) \text{ for all } \lambda \in \Lambda \} \).

**Lemma 15 (Janssen’s representation).** Assume that \( \varphi_j, \psi_j \in M^1(\mathbb{R}^d) \) for all \( j = 1, \ldots, n \). Then the frame type operator associated to \( \bigcup_{j=1}^n G(\varphi_j, \Lambda) \) and \( \bigcup_{j=1}^n G(\psi_j, \Lambda) \) can be written as

\[
S_{\varphi, \psi} f = s(\Lambda)^{-1} \sum_{\mu \in \Lambda^\circ} \sum_{j=1}^n \langle \varphi_j, \pi(\mu)\psi_j \rangle \pi(\mu) f \quad (A.3)
\]

with unconditional convergence in the operator norm on \( L^2 \).

**Proof.** By Janssen’s result [28] the representation holds for a single \( S_{\varphi_j, \psi_j} \) and (A.3) follows by taking a sum. \( \square \)

The canonical dual frame is defined to be \( \gamma_{j, \lambda} = \pi(\lambda)S^{-1}\varphi_j \). Since the frame operator \( S = S_{\varphi, \varphi} \) commutes with time-frequency shifts on \( \Lambda \), we obtain the reconstruction formulas

\[
f = S^{-1}Sf = \sum_{\lambda \in \Lambda} \sum_{j=1}^n \langle f, \pi(\lambda)\varphi_j \rangle \pi(\lambda)\gamma_j
\]

\[
= SS^{-1}f = \sum_{\lambda \in \Lambda} \sum_{j=1}^n \langle f, \pi(\lambda)\gamma_j \rangle \pi(\lambda)\varphi_j
\]

\[
= D_{\varphi, \Lambda}C_{\gamma, \Lambda}f = D_{\gamma, \Lambda}C_{\varphi, \Lambda}f.
\]

As a general principle the localization of a frame is inherited by the dual frame [19]. The following statement is a generalization of [24, Theorem 9] to multi-window Gabor frames on general lattices.

**Lemma 16.** Assume that \( \nu \) is a submultiplicative, even weight on \( \mathbb{R}^{2d} \) satisfying \( \lim_{n \to \infty} \nu(nz)^{1/n} = 1 \) for all \( z \in \mathbb{R}^{2d} \). Assume further that \( \bigcup_{j=1}^n G(\varphi_j, \Lambda) \) is a frame for \( L^2(\mathbb{R}^d) \) and that \( \varphi_j \in M^1_\nu(\mathbb{R}^d) \). Then the frame operator \( S \) is invertible on \( M^1_\nu(\mathbb{R}^d) \) and \( \gamma_j = S^{-1}\varphi_j \in M^1_\nu(\mathbb{R}^d) \) for \( j = 1, \ldots, n \).

**Proof.** Janssen’s representation (A.3) implies that

\[
S = S_{\varphi, \varphi} = s(\Lambda)^{-1} \sum_{\mu \in \Lambda^\circ} c_\mu \pi(\mu), \quad (A.4)
\]

with a coefficient sequence \( c_\mu = \sum_{j=1}^n \langle \varphi_j, \pi(\mu)\varphi_j \rangle \). The hypothesis \( \varphi_j \in M^1_\nu(\mathbb{R}^d) \) guarantees that \( \sum_{\mu \in \Lambda^\circ} \langle \varphi_j, \pi(\mu)\varphi_j \rangle \nu(\mu) < \infty \) for each \( j \), see [21, Corollary 12.1.12], and therefore the coefficient sequence \( (c_\mu) \) is in \( \ell^1_\nu(\Lambda^\circ) \). Since \( \bigcup_{j=1}^n G(\varphi_j, \Lambda) \) is a frame, the frame operator \( S_{\varphi, \varphi} \) is invertible on \( L^2(\mathbb{R}^d) \). It follows from [24, Theorem 3.1] that the inverse frame operator
$S^{-1}$ is again of the form $S^{-1} = \sum_{\mu \in A^o} d_{\mu} \pi(\mu)$ with a coefficient sequence $d$ in $\ell_1(A^o)$. This representation implies that $S^{-1}$ is bounded on $M_1^1(\mathbb{R}^d)$ and that
\[
\|y_j\|_{M_1^1} = \|S^{-1} \varphi_j\|_{M_1^1} \leq C \|\varphi_j\|_{M_1^1},
\] (A.5)
Therefore the dual windows $\gamma_j$, $j = 1, \ldots, n$ are in $M_1^1(\mathbb{R}^d)$ as claimed. □

Once the invertibility of the multi-window frame operator on $M_1^1(\mathbb{R}^d)$ is established, the proof of Theorem 2 is straight-forward by using the following boundedness properties of the coefficient operator $C_{\varphi, A}$ and $D_{\varphi, A}$ from [21, Theorems 12.2.3 and 12.3.4]. If $\varphi_j \in M_1^1(\mathbb{R}^d)$ and $\gamma_j \in M_1^1(\mathbb{R}^d)$, then both $C_{\varphi, A}$ and $C_{\gamma, A}$ are bounded from $M_1^{p,q}(\mathbb{R}^d)$ into $\ell_m^{p,q}(A, \mathbb{C}^n)$ for $1 \leq p, q \leq \infty$ and for every $\nu$-moderate weight $m$. Likewise $D_{\varphi, A}$ and $D_{\gamma, A}$ are bounded from $\ell_m^{p,q}(A, \mathbb{C}^n)$ into $M_1^{p,q}(\mathbb{R}^d)$. For the $\ell_m^{p,q}(A, \mathbb{C}^n)$-norm we use the Euclidean norm on $\mathbb{C}^n$, so that $\|c\|_{\ell_m^{p,q}(A, \mathbb{C}^n)} = \|\sum_{\lambda \in A}(\sum_{j=1}^{n} |c_{\lambda,j}|^2)^{1/2} \chi_{\lambda} \circ \Omega\|_{L_m^{p,q}}$.

As a consequence, the reconstruction formula $f = D_{\varphi, A} C_{\gamma, A} f = D_{\gamma, A} C_{\varphi, A} f$ holds for $f \in M_1^{p,q}(\mathbb{R}^d)$ with the correct norm estimates. The norm equivalence stated in Theorem 2 then follows from
\[
\|f\|_{M_1^{p,q}(\mathbb{R}^d)} = \|D_{\gamma, A} C_{\varphi, A} f\|_{M_1^{p,q}(\mathbb{R}^d)} \leq \|D_{\gamma, A}\|_{op} \|C_{\varphi, A} f\|_{\ell_m^{p,q}(A, \mathbb{C}^n)} \leq \|D_{\gamma, A}\|_{op} \|C_{\varphi, A}\|_{op} \|f\|_{M_1^{p,q}(\mathbb{R}^d)}.
\]

Next we come to the characterization of multi-window Gabor frames (Lemma 3) and extend the list of equivalent conditions. For the formulation of the dual conditions on the adjoint lattice $A^o$ we need the vector-valued versions of the analysis and synthesis operators. For $f = (f_1, \ldots, f_n) \in M_1^\infty(\mathbb{R}^d, \mathbb{C}^n)$ and $\varphi = (\varphi_1, \ldots, \varphi_n) \in M_1^1(\mathbb{R}^d, \mathbb{C}^n)$ the coefficient operator is defined to be $\tilde{C}_{\varphi, A^o}(f)(\mu) = \langle f, \pi(\mu) \varphi \rangle$, $\mu \in \Lambda^o$, and the synthesis operator is $\tilde{D}_{\varphi, A^o}(c) = \sum_{\mu \in \Lambda^o} c_{\mu} \pi(\mu) \varphi$. The Gramian operator $G_{\varphi, A^o} = \tilde{C}_{\varphi, A^o} \cdot \tilde{D}_{\varphi, A^o}$ is defined on sequences indexed by $A^o$.

Lemma 17. Assume that $\varphi_j \in M_1^1(\mathbb{R}^d)$ for $j = 1, \ldots, n$. The following are equivalent for the multi-window Gabor system $\bigcup_{j=1}^{n} G(\varphi_j, A)$:

(i) $\bigcup_{j=1}^{n} G(\varphi_j, A)$ is a frame for $L^2(\mathbb{R}^d)$.
(ii) Weixel–Raz biorthogonality: There exist $\gamma_j \in M_1^1(\mathbb{R}^d)$, $j = 1, \ldots, n$, such that
\[
\langle \varphi_j, \pi(\mu) \gamma_j \rangle = \delta_{\mu,0} \text{ for } \mu \in \Lambda^o.
\] (A.6)
(iii) Ron–Shen duality: $G(\varphi, A^o)$ is a Riesz sequence in $L^2(\mathbb{R}^d, \mathbb{C}^n)$.
(iv) $S_{\varphi, \varphi}$ is invertible on $M_1^1(\mathbb{R}^d)$.
(v) $S_{\varphi, \varphi}$ is invertible on $M_1^\infty(\mathbb{R}^d)$.
(vi) $S_{\varphi, \varphi}$ is one-to-one on $M_1^\infty(\mathbb{R}^d)$.
(vii) The analysis operator $C_{\varphi, A} : M_1^\infty(\mathbb{R}^d) \mapsto \ell^\infty(A, \mathbb{C}^n)$ is one-to-one from $M_1^\infty(\mathbb{R}^d)$ to $\ell^\infty(A, \mathbb{C}^n)$.
(viii) The synthesis operator $D_{\varphi, A}$ defined on $\ell_1(A, \mathbb{C}^n)$ has dense range in $M_1^1(\mathbb{R}^d)$.
(ix) $D_{\varphi, \Lambda}$ is surjective from $\ell^1(\Lambda, C^n)$ onto $M^1(\mathbb{R}^d)$.
(x) The synthesis operator $\tilde{D}_{\varphi, \Lambda}$ defined on $\ell^\infty(\Lambda^\circ)$ is one-to-one from $\ell^\infty(\Lambda^\circ)$ to $M^\infty(\mathbb{R}^d, C^n)$.
(xi) The analysis operator $\tilde{C}_{\varphi, \Lambda^\circ}$ defined on $M^1(\mathbb{R}^d, C^n)$ has dense range in $\ell^1(\Lambda^\circ)$.
(xii) $\tilde{C}_{\varphi, \Lambda^\circ}$ is surjective from $M^1(\mathbb{R}^d, C^n)$ onto $\ell^1(\Lambda^\circ)$.
(xiii) $G_{\varphi, \Lambda^\circ}$ is invertible on $\ell^1(\Lambda^\circ)$.
(xiv) $G_{\varphi, \Lambda^\circ}$ is invertible on $\ell^\infty(\Lambda^\circ)$.
(xv) $G_{\varphi, \Lambda^\circ}$ is one-to-one on $\ell^1(\Lambda^\circ)$.

The equivalence (i) $\iff$ (vii) is claimed in Lemma 3 and is all we need for the main results of our paper.

**Proof.** The implication (i) $\Rightarrow$ (iv) was sketched in Lemma 16.

(i) $\iff$ (ii): Time-frequency shifts on a lattice are linearly independent in the following sense: if $c = (c_\mu)_{\mu \in \Lambda^\circ} \in \ell^\infty$ and $\sum_{\mu \in \Lambda^\circ} c_\mu \pi(\mu) = 0$ (as an operator from $M^1(\mathbb{R}^d)$ to $M^\infty(\mathbb{R}^d)$), then $c_\mu = 0$ for all $\mu \in \Lambda^\circ$, see [22]. Now, if $f = S_{\varphi, \gamma} f$ for all $f \in M^1(\mathbb{R}^d)$, then by Janssen’s representation (A.3) we have

$$f = s(A)^{-1} \sum_{\mu \in \Lambda^\circ} \sum_{j=1}^n (\varphi_j, \pi(\mu) \gamma_j) \pi(\mu) f.$$ 

The linear independence of time-frequency shifts implies (A.6). The converse is obvious.

(ii) $\iff$ (iii): Assume first that $\bigcup_{j=1}^n G(\varphi_j, \Lambda)$ is a multi-window Gabor frame for $L^2(\mathbb{R}^d)$. The upper bound in (A.2) follows from the boundedness of the synthesis operator $\tilde{D}_{\varphi}$ on $L^2(\mathbb{R}^d)$. To show the existence of a lower bound, we apply the Wexler–Raz relations. Since $\bigcup_{j=1}^n G(\varphi_j, \Lambda)$ is a frame with dual $\bigcup_{j=1}^n G(\gamma_j, \Lambda)$ and $\gamma_j \in M^1(\mathbb{R}^d)$ for all $j$, we have $\langle \varphi, \pi(\mu) \gamma \rangle = \sum_{j=1}^n \langle \varphi_j, \pi(\mu) \gamma_j \rangle = s(A) \delta_{\mu,0}$, and $G(\varphi, \Lambda^\circ)$ and therefore $G(\gamma, \Lambda^\circ)$ are biorthogonal systems in $L^2(\mathbb{R}^d, C^n)$. If $f = \sum_{\mu \in \Lambda^\circ} c_\mu \pi(\mu) f$, then $c_\mu = s(A)^{-1} \langle f, \pi(\mu) \gamma \rangle_{L^2(\mathbb{R}^d, C^n)}$ and

$$c = s(A)^{-1} \tilde{C}_{\varphi, \Lambda^\circ} f,$$

from which the lower bound in (A.2) follows.

Conversely, assume that $G(\varphi, \Lambda^\circ)$ is a Riesz sequence in $L^2(\mathbb{R}^d, C^n)$. Then there exists a biorthogonal basis of the form $\{\pi(\mu) \gamma: \mu \in \Lambda^\circ\}$ contained in $K = \text{Span}(G(\varphi, \Lambda^\circ))$. It can be shown that $\gamma \in M^1(\mathbb{R}^d, C^n)$. The frame property of $G(\varphi_j, \Lambda)$ follows from the Wexler–Raz relations (A.6).

With three classical statements (A.3) and (ii), (iii) for multi-window Gabor frames the remaining equivalences follow exactly as in [22]. 

**References**

Representation of Operators in the Time-Frequency Domain and Generalized Gabor Multipliers

Monika Dörfler · Bruno Torrésani

Received: 28 August 2008 / Revised: 8 June 2009 / Published online: 4 August 2009
© Birkhäuser Boston 2009

Abstract Starting from a general operator representation in the time-frequency domain, this paper addresses the problem of approximating linear operators by operators that are diagonal or band-diagonal with respect to Gabor frames. A characterization of operators that can be realized as Gabor multipliers is given and necessary conditions for the existence of (Hilbert-Schmidt) optimal Gabor multiplier approximations are discussed and an efficient method for the calculation of an operator’s best approximation by a Gabor multiplier is derived. The spreading function of Gabor multipliers yields new error estimates for these approximations. Generalizations (multiple Gabor multipliers) are introduced for better approximation of overspread operators. The Riesz property of the projection operators involved in generalized Gabor multipliers is characterized, and a method for obtaining an operator’s best approximation by a multiple Gabor multiplier is suggested. Finally, it is shown that in certain situations, generalized Gabor multipliers reduce to a finite sum of regular Gabor multipliers with adapted windows.

Keywords Operator approximation · Generalized Gabor multipliers · Spreading function · Twisted convolution

Mathematics Subject Classification (2000) 47B38 · 47G30 · 94A12 · 65F20

Communicated by Karlheinz Gröchenig.
The first author has been supported by WWTF project MA07-025 and FWF grant T 384-N13.

M. Dörfler (✉)
Numerical Harmonic Analysis Group, Faculty of Mathematics, University of Vienna,
Alserbachstraße 23, 1090 Wien, Austria
e-mail: monika.doerfler@univie.ac.at

B. Torrésani
Laboratoire d’Analyse, Topologie et Probabilités, Centre de Mathématique et d’Informatique, 39 rue Joliot-Curie, 13453 Marseille cedex 13, France
e-mail: bruno.torensani@univ-provence.fr
1 Introduction

The goal of time-frequency analysis is to provide efficient representations for functions or distributions in terms of decompositions such as

$$f = \sum_{\lambda \in \Lambda} \langle f, g_\lambda \rangle h_\lambda.$$  

Here, $f$ is expanded as a weighted sum of atoms $h_\lambda$ well localized in both time and frequency domains. The time-frequency coefficients $\langle f, g_\lambda \rangle$ characterize the function under investigation, and a synthesis map usually allows the reconstruction of the original function $f$.

Concrete applications can be found mostly in signal analysis and processing (see [8, 9, 36] and references therein), but recent works in different areas such as numerical analysis may also be mentioned (see for example [19, 20] and references therein).

Time-frequency analysis of operators, originating in the work on communication channels of Bello [5], Kailath [30] and Zadeh [41], has enjoyed increasing interest during the last few years, [2, 35, 38, 39]. Efficient time-frequency operator representation is a challenging task, and often the intuitively appealing approach of operator approximation by modification of the time-frequency coefficients before reconstruction is the method of choice. If the modification of the coefficients is confined to be multiplicative, this approach leads to the model of time-frequency multipliers, as discussed in Sect. 2.4. The class of operators that may be well represented by time-frequency multipliers depends on the choice of the parameters involved and is restricted to operators performing only small time-shifts or modulations.

The work in this paper is inspired by a general operator representation in the time-frequency domain via a twisted convolution. It turns out, that this representation, respecting the underlying structure of the Heisenberg group, has an interesting connection to the so-called spreading function representation of operators. An operator’s spreading function comprises the amount of time-shifts and modulations, i.e. of time-frequency-shifts, effected by the operator. Its investigation is hence decisive in the study of time-frequency multipliers and their generalizations. Although no direct discretization of the continuous representation by an operator’s spreading function is possible, the twisted convolution turns out to play an important role in the generalizations of time-frequency multipliers. In the main section of this article, we introduce a general model for multiple Gabor multipliers (MGM), which uses several synthesis windows simultaneously. Thus, by jointly adapting the respective masks, more general operators may be well-represented than by regular Gabor multipliers. Specifying to a separable mask in the modification of time-frequency coefficients within MGM, as well as a specific sampling lattice for the synthesis windows, it turns out, that the MGM reduces to one or the sum of a finite number of regular Gabor multipliers with adapted synthesis windows.

For the sake of generality, most statements are given in a Gelfand-triple, rather than a pure Hilbert space setting. This choice bears several advantages. First of all, many important operators and signals may not be described in a Hilbert-space setting, starting from simple operators as the identity. Furthermore, by using distributions,
continuous and discrete concepts may be considered in a unified framework. Finally, the Gelfand-triple setting often allows for short-cut proofs of statements formulated in a general context.

This paper is organized as follows. The next section gives a review of the time-frequency plane and the corresponding continuous and discrete transforms. We then introduce the concept of Gelfand triples, which will allow us to consider operators beyond the Hilbert-Schmidt framework. The section closes with the important statement on operator-representation in the time-frequency domain via twisted convolution with an operator’s spreading function. Section 2.4 introduces time-frequency multipliers and gives a criterion for their ability to approximate linear operators. A fast method for the calculation of an operator’s best approximation by a Gabor multiplier in Hilbert-Schmidt sense is suggested. Section 3 introduces generalizations of Gabor multipliers. The operators in the construction of MGM are investigated and a criterion for their Riesz basis property in the space of Hilbert-Schmidt operators is given. We mention some connections to classical Gabor frames. A numerical example concludes the discussion of general MGM. In the final section, TST (twisted spline type) spreading functions are introduced. It is shown, that under certain conditions, a MGM reduces to a regular Gabor multiplier with an adapted window or a finite sum of regular multipliers with the same mask and adapted windows.

**2 Operators from the Time-Frequency Point of View**

Whenever one is interested in time-localized frequency information in a signal or operator, one is naturally led to the notion of the time-frequency plane, which, in turn, is closely related to the Weyl-Heisenberg group.

**2.1 Preliminaries: The Time-Frequency Plane**

The starting point of our operator analysis is the so-called spreading function operator representation. This operator representation expresses linear operators as a sum (in a sense to be specified below) of time-frequency shifts \( \pi(b, \nu) = M_\nu T_b \). Here, the translation and modulation operators are defined as

\[
T_b f(t) = f(t - b), \quad M_\nu f(t) = e^{2i\pi \nu t} f(t), \quad f \in L^2(\mathbb{R}).
\]

These (unitary) operators generate a group, called the Weyl-Heisenberg group

\[
\mathbb{H} = \{(b, \nu, \varphi) \in \mathbb{R} \times \mathbb{R} \times [0, 1]\},
\]

with group multiplication

\[
(b, \nu, \varphi)(b', \nu', \varphi') = (b + b', \nu + \nu', \varphi + \varphi' - \nu' b).
\]

The specific quotient space \( \mathbb{P} = \mathbb{H}/[0, 1] \) of the Weyl-Heisenberg group is called phase space, or time-frequency plane, which plays a central role in the subsequent analysis. Details on the Weyl-Heisenberg group and the time-frequency plane may
be found in [22, 40]. In the current article, we shall limit ourselves to the basic irreducible unitary representation of $\mathbb{H}$ on $L^2(\mathbb{R})$, denoted by $\pi^\circ$, and defined by

$$\pi^\circ(b, v, \varphi) = e^{2i\pi \varphi M_v T_b}.$$  

(3)

By $\pi(b, v) = \pi^\circ(b, v, 0)$ we denote the restriction to the phase space. We refer to [11] or [25, Chap. 9] for a more detailed analysis of this quotient operation.

The left-regular (and right-regular) representation generally plays a central role in group representation theory. By unimodularity of the Weyl-Heisenberg group, its left and right regular representations coincide. We thus focus on the left-regular one, acting on $L^2(\mathbb{H})$ and defined by

$$[L(b', v', \varphi') F](b, v, \varphi) = F(b - b', v - v', \varphi - \varphi' + b'(v - v')).$$  

(4)

Denote by $\mu$ the Haar measure. Given $F, G \in L^2(\mathbb{H}, d\mu)$, the associated (left) convolution product is the bounded function $F \ast G$, given by

$$(F \ast G)(b, v, \varphi) = \int_{\mathbb{H}} F(h)[L(h)G](b, v, \varphi) d\mu(h).$$

After quotienting out the phase term, this yields the twisted convolution on $L^2(\mathbb{P})$:

$$\langle F, G \rangle(b, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(b', v') G(b - b', v - v') e^{-2i\pi b'(v-v')} db' dv'. $$  

(5)

The twisted convolution, which admits a nice interpretation in terms of group Plancherel theory [11] is non-commutative (which reflects the non-Abelianness of $\mathbb{H}$) but associative. It satisfies the usual Young inequalities, but is in some sense nicer than the usual convolution, since $L^2(\mathbb{R}^2) \otimes L^2(\mathbb{R}^2) \subset L^2(\mathbb{R}^2)$ (see [22] for details).

As explained in [27, 28] (see also [23] for a review), the representation $\pi^\circ$ is unitarily equivalent to a subrepresentation of the left regular representation. The representation coefficient is given by a variant of the short time Fourier transform (STFT), which we define next.

**Definition 1** Let $g \in L^2(\mathbb{R})$, $g \neq 0$. The STFT of any $f \in L^2(\mathbb{R})$ is the function on the phase space $\mathbb{P}$ defined by

$$V_g f(b, v) = \langle f, \pi(b, v)g \rangle = \int_{-\infty}^{\infty} f(t) \overline{g}(t - b) e^{-2i\pi vt} dt.$$  

(6)

This STFT is obtained by quotienting out $[0, 1]$ in the group transform

$$V^\circ_g f(b, v, \varphi) = \langle f, \pi^\circ(b, v, \varphi)g \rangle.$$  

(7)

The integral transform $V^\circ_g$ intertwines $L$ and $\pi^\circ$, i.e. $L(h) V^\circ_g = V^\circ_g \pi^\circ(h)$ for all $h \in \mathbb{H}$. The latter relation still holds true (up to a phase factor) when $\pi^\circ$ and $V^\circ_g$ are replaced with $\pi$ and $V_g$ respectively.

It follows from the general theory of square-integrable representations that for any $g \in L^2(\mathbb{R})$, $g \neq 0$, the transform $V^\circ_g$ is (a multiple of) an isometry $L^2(\mathbb{R}) \to$
More precisely, given \( h \in L^2(\mathbb{R}) \) such that \( \langle g, h \rangle \neq 0 \), one has for all \( f \in L^2(\mathbb{R}) \)

\[
f = \frac{1}{\langle h, g \rangle} \int_{\mathbb{P}} \mathcal{V}_g f(b, v) \pi(b, v) h \, db \, dv. \tag{8}
\]

We refer to [8, 25] for more details on the STFT and signal processing applications.

The STFT, being a continuous transform, is not well adapted for numerical calculations, and, for practical issues, is replaced by the Gabor transform, which is a sampled version of it. To fix notation, we outline some steps of the Gabor frame theory and refer to [9, 25] for a detailed account.

**Definition 2** (Gabor transform) Given \( g \in L^2(\mathbb{R}) \) and two constants \( b_0, \nu_0 \in \mathbb{R}^+ \), the corresponding Gabor transform associates with any \( f \in L^2(\mathbb{R}) \) the sequence of Gabor coefficients

\[
\mathcal{V}_g f(mb_0, n\nu_0) = \langle f, M_{n\nu_0} T_{mb_0} g \rangle = \langle f, g_{mn} \rangle, \tag{9}
\]

where the functions \( g_{mn} = M_{n\nu_0} T_{mb_0} g \) are the Gabor atoms associated to \( g \) and the lattice constants \( b_0, \nu_0 \).

Whenever the Gabor atoms associated to \( g \) and the given lattice \( \Lambda_{b_0, \nu_0} = b_0 \mathbb{Z} \times \nu_0 \mathbb{Z} \) form a frame,\(^1\) the Gabor transform is left invertible, and there exists \( h \in L^2(\mathbb{R}) \) such that any \( f \in L^2(\mathbb{R}) \) may be expanded as

\[
f = \sum_{m,n} \mathcal{V}_g f(mb_0, n\nu_0) h_{mn}. \tag{10}
\]

In addition to the notion of a lattice we need to define the dual concept of adjoint lattices.

**Definition 3** (Adjoint lattice) For a given lattice \( \Lambda = b_0 \mathbb{Z} \times \nu_0 \mathbb{Z} \) the adjoint lattice is given by \( \Lambda^\circ = \frac{1}{\nu_0} \mathbb{Z} \times \frac{1}{b_0} \mathbb{Z} \).

Note that the adjoint lattice is the dual lattice \( \Lambda^\perp \) with respect to the symplectic character.

2.2 The Gelfand Triple \((S_0, L^2, S'_0)\)

We next set up a framework for the exact description of operators we are interested in. In fact, by their property of being compact operators, the Hilbert space of Hilbert-Schmidt operators turns out to be far too restrictive to contain most operators of

\(^1\)The operator

\[ S_g f = \sum_{m,n} \langle f, M_{mb_0} T_{n\nu_0} g \rangle M_{mb_0} T_{n\nu_0} g \]

is the frame operator corresponding to \( g \) and the lattice defined by \((b_0, \nu_0)\). If \( S_g \) is invertible on \( L^2(\mathbb{R}) \), the family of time-frequency shifted atoms \( M_{mb_0} T_{n\nu_0} g \), \( m, n \in \mathbb{Z} \), is a Gabor frame for \( L^2(\mathbb{R}) \).
practical interest, starting from the identity. Although the classical triple \((\mathcal{S}, \mathcal{L}^2, \mathcal{S}')\) might seem to be the appropriate choice of generalization, we prefer to resort to the Gelfand triple \((\mathcal{S}_0, \mathcal{L}^2, \mathcal{S}'_0)\), which has proved to be more adapted to a time-frequency environment. Additionally, the Banach space property of \(\mathcal{S}_0\) guarantees a technically less elaborate account.

**Definition 4** \((\mathcal{S}_0)\) Let \(\mathcal{S}(\mathbb{R})\) denote the Schwartz class. Fix a non-zero “window” function \(\varphi \in \mathcal{S}(\mathbb{R})\). The space \(\mathcal{S}_0(\mathbb{R})\) is given by

\[
\mathcal{S}_0(\mathbb{R}) = \{ f \in \mathcal{L}^2(\mathbb{R}) : \| f \|_{\mathcal{S}_0} := \| \mathcal{V}_\varphi f \|_{\mathcal{L}^1(\mathbb{R}^2)} < \infty \}.
\]

Note that \(\mathcal{S}_0\) is a special case of a modulation space \([25]\), namely \(M^{1,1}\), and is also called Feichtinger’s algebra in the literature. The following proposition summarizes some properties of \(\mathcal{S}_0(\mathbb{R})\) and its dual, the distribution space \(\mathcal{S}'_0(\mathbb{R})\).

**Proposition 5** \(\mathcal{S}_0(\mathbb{R})\) is a Banach space and densely embedded in \(\mathcal{L}^2(\mathbb{R})\). The definition of \(\mathcal{S}_0(\mathbb{R})\) is independent of the window \(\varphi \in \mathcal{S}(\mathbb{R})\), and different choices of \(\varphi \in \mathcal{S}(\mathbb{R})\) yield equivalent norms on \(\mathcal{S}_0(\mathbb{R})\).

By duality, \(\mathcal{L}^2(\mathbb{R})\) is densely and weak*-continuously embedded in \(\mathcal{S}'_0(\mathbb{R})\) and can also be characterized by the norm \(\| f \|_{\mathcal{S}'_0} = \| \mathcal{V}_\varphi f \|_{\mathcal{L}^\infty}\).

A triple of spaces \((\mathcal{B}, \mathcal{H}, \mathcal{B}')\) consisting of a Banach space \(\mathcal{B}\) which is densely embedded in the Hilbert space \(\mathcal{H}\), which, in turn is densely embedded in \(\mathcal{B}'\), is called (Banach) Gelfand triple. Hence, \((\mathcal{S}_0, \mathcal{L}^2, \mathcal{S}'_0)\) represents a special case of a Gelfand triple \([24]\) or Rigged Hilbert space. The prototype of a Gelfand triple is the triple of sequence spaces \((\ell^1, \ell^2, \ell^\infty)\). For a proof of Proposition 5, equivalent characterizations, and more results on \(\mathcal{S}_0\) we refer to \([13, 14, 17]\).

Via an isomorphism between integral kernels in the Banach spaces \(\mathcal{S}_0, \mathcal{S}'_0\) and the operator spaces of bounded operators \(\mathcal{S}'_0 \leftrightarrow \mathcal{S}_0\) and \(\mathcal{S}_0 \leftrightarrow \mathcal{S}'_0\), we obtain, together with the Hilbert space of Hilbert-Schmidt operators, a Gelfand triple of operator spaces, as follows. We denote by \(\mathcal{B}\) the family of operators that are bounded \(\mathcal{S}'_0 \to \mathcal{S}_0\) and by \(\mathcal{B}'\) the family of operators that are bounded \(\mathcal{S}_0 \to \mathcal{S}'_0\). We have the following correspondence between these operator classes and their integral kernels \(\kappa\):

\[
H \in (\mathcal{B}, \mathcal{H}, \mathcal{B}') \quad \longleftrightarrow \quad \kappa_H \in (\mathcal{S}_0(\mathbb{R}^2), \mathcal{L}^2(\mathbb{R}^2), \mathcal{S}'_0(\mathbb{R}^2)).
\]

We will make use of the principle of unitary Gelfand triple isomorphisms, described for the Gelfand triples just introduced in \([17]\). A linear mapping between two Gelfand triples \((\mathcal{B}_1, \mathcal{H}_1, \mathcal{B}'_1)\) and \((\mathcal{B}_2, \mathcal{H}_2, \mathcal{B}'_2)\) is called unitary Gelfand triple isomorphism if it is an isomorphism between \(\mathcal{B}_1, \mathcal{B}_2\), a unitary isomorphism on the Hilbert space level and extends to a weak*-continuous bijection from \(\mathcal{B}'_1\) to \(\mathcal{B}'_2\). In fact, it may be shown, that it suffices to verify unitarity of a given isomorphic operator on the (dense) subspace \(\mathcal{S}_0\) in order to obtain a unitary Gelfand triple isomorphism, see \([17, \text{Corollary 7.3.4]}\). The most prominent examples for a unitary Gelfand triple isomorphism are the Fourier transform and the partial Fourier transform. For all further details on the Gelfand triples just introduced, we again refer to \([17]\), only mentioning here,
that one important reason for investigating operator representations on the Gelfand triple instead of just a Hilbert space framework is the fact, that \( S'_0 \) contains distributions such as the Dirac functionals, Shah distributions or just pure frequencies and \( B' \) contains operators of great importance in signal processing, e.g. convolution, the identity or just time-frequency shifts.

Subsequently, we will usually assume that the analysis and synthesis windows \( g, h \) are in \( S_0 \). This is a rather mild condition, which has almost become the canonical choice in Gabor analysis, for many good reasons. Among others, this choice guarantees a beautiful correspondence between the \( \ell^p \)-spaces and corresponding modulation spaces [25]. In the \( \ell^2 \)-case this means, that the sequence of Gabor atoms generated from time frequency translates of an \( S_0 \) window on an arbitrary lattice \( \Lambda \) is automatically a Bessel sequence (in such a case, the window is termed “Bessel atom”), which is not true for general \( L^2 \)-windows.

The Banach spaces \( S_0 \) and \( S'_0 \) may also be interpreted as Wiener amalgam spaces [13, Sect. 3.2.2]. These time-frequency homogeneous spaces are defined as follows. Let \( \mathcal{F}L^1 \) denote the Fourier image of integrable functions and let a compactly supported function \( \phi \in \mathcal{F}L^1(\mathbb{R}) \) with \( \sum_{n \in \mathbb{Z}} \phi(x - n) \equiv 1 \) be given. Then, for \( X(\mathbb{R}) = \mathcal{F}L^1(\mathbb{R}) \) or \( X(\mathbb{R}) = C(\mathbb{R}) \), i.e. the space of continuous functions on \( \mathbb{R} \), or any of the Lebesgue spaces, we define, for \( p \in [1, \infty) \), with the usual modification for \( p = \infty \):

\[
W(X, \ell^p) = \left\{ f \in X_{\text{loc}} : \|f\|_{W(X, \ell^p)} = \left( \sum_{n \in \mathbb{Z}} \|fT_n \phi\|_X^p \right)^{1/p} < \infty \right\}.
\]

(11)

Now, \( S_0 = W(\mathcal{F}L^1, \ell^1) \) and \( S'_0 = W(\mathcal{F}L^\infty, \ell^\infty) \), see [13, Sect. 3.2.2].

2.3 The Spreading Function Representation and Its Connections to the STFT

The so-called spreading function representation, closely related to the integrated Schrödinger representation [25, Sect. 9.2], expresses operators in \( (B, \mathcal{H}, B') \) as a sum of time-frequency shifts. More precisely, one has (see [25, Chap. 9]):

**Theorem 6** Let \( H \in (B, \mathcal{H}, B') \); then there exists a spreading function \( \eta_H \) in \( (S_0(\mathbb{R}^2), L^2(\mathbb{R}^2), S'_0(\mathbb{R}^2)) \) such that

\[
H = \int_{-\infty}^\infty \int_{-\infty}^\infty \eta_H(b, v) \pi(b, v) db dv.
\]

(12)

For \( H \in \mathcal{H} \), the correspondence \( H \leftrightarrow \eta_H \) is isometric, i.e. \( \|H\|_{\mathcal{H}} = \|\eta_H\|_{L^2(\mathbb{P})} \).

**Remark 7** For \( H \in B \), the decomposition given in (12) is absolutely convergent, whereas, for \( H \in B' \), it holds in the weak sense of bilinear forms on \( S_0 \).

When \( \eta_H \in L^2(\mathbb{P}) \), \( H \) is a Hilbert-Schmidt operator, and the above integral is defined as a Bochner integral.
The spreading function is intimately related to the integral kernel $\kappa = \kappa_H$ of $H$ via

$$
\eta_H(b, \nu) = \int_{-\infty}^{\infty} \kappa_H(t, t-b) e^{-2i\pi \nu t} dt \quad \text{and}
$$

$$
\kappa_H(t, s) = \int_{-\infty}^{\infty} \eta_H(t-s, \nu) e^{2i\pi \nu t} d\nu.
$$

As a consequence, for $\kappa_H \in (S_0, L^2, S'_0)$, we also have $\eta_H \in (S_0, L^2, S'_0)$. In particular, this leads to the following expression for a weak evaluation of Gelfand triple operators:

$$
\langle K, L \rangle_{(B, H, B')} = (\langle \kappa(K), \kappa(L) \rangle_{(S_0, L^2, S'_0)} = \langle \eta(K), \eta(L) \rangle_{(S_0, L^2, S'_0)}).
$$

(13)

For $L = g \otimes f^*$, i.e. the tensor product with kernel $\kappa(s, t) = g(s) \overline{f(t)}$ and spreading function $\eta(b, \nu) = V_f g(b, \nu)$, we thus have:

$$
\langle K, g \otimes f^* \rangle_{(B, \mathcal{H}, B')} = (\langle \kappa(K), \kappa(g \otimes f^*) \rangle_{(S_0, L^2, S'_0)} = \langle \eta(K), \mathcal{V}_f g \rangle_{(S_0, L^2, S'_0)}).
$$

(14)

Let us also mention that the spreading function is related to the operator’s Kohn-Nirenberg symbol via a symplectic Fourier transform, which we define for later reference.

**Definition 8** The symplectic Fourier transform is formally defined by

$$
\mathcal{F}_s F(t, \xi) = \int \mathbb{P} F(b, \nu) e^{-2\pi i (b\xi - \nu t)} db d\nu.
$$

The symplectic Fourier transform is a self-inverse unitary automorphism of the Gelfand triple $(S_0, L^2, S'_0)$. We will make use of the following relation.

**Lemma 9** Assume that $f_1, f_2, g_1, g_2 \in L^2(\mathbb{R})$. Then

$$
\mathcal{F}_s (\mathcal{V}_{g_1} f_1 \mathcal{V}_{g_2} f_2)(x, \omega) = (\mathcal{V}_{f_2} f_1 \mathcal{V}_{g_2} g_1)(x, \omega).
$$

Proof The analogous statement for the conventional (Cartesian) Fourier transform reads $\mathcal{F}(\mathcal{V}_{g_1} f_1 \mathcal{V}_{g_2} f_2)(x, \omega) = (\mathcal{V}_{f_2} f_1 \mathcal{V}_{g_2} g_1)(-\omega, x)$ and has been shown in [26, Lemma 2.3.2]. The fact, that

$$
[\mathcal{F}_s F](x, \omega) = \hat{F}(-\omega, x)
$$

completes the argument. □

Recall that the spreading function of the product of operators corresponds to the twisted convolution of the operators’ spreading function. Assume $K_1$ in $(\mathcal{B}, \mathcal{H}, \mathcal{B}')$ and $K_2$ in $(\mathcal{B}', \mathcal{H}, \mathcal{B})$ then

$$
\eta(K_2 \cdot K_1) = \eta(K_2) \circ \eta(K_1).
$$

(15)
The spreading function representation of operators provides an interesting time-frequency implementation for operators, stated in the following proposition. It turns out to be closely connected to the tools described in the previous section, in particular twisted convolution and STFT.

**Proposition 10** Let \( H \) be in \((\mathcal{B}, \mathcal{H}, \mathcal{B}')\), and let \( \eta = \eta_H \) be its spreading function in \((S_0(\mathbb{R}^2), L^2(\mathbb{R}^2), S_0'(\mathbb{R}^2))\). Let \( g \in S_0(\mathbb{R}) \), then the STFT of \( Hf \) is given by a twisted convolution of \( \eta_H \) and \( V_g f \):

\[
V_g Hf(z) = (\eta_H \ast V_g f)(z).
\]

**Proof** By \((12)\), we may write

\[
V_g Hf(z') = \langle Hf, \pi(z')g \rangle = \int \langle \eta_H(z)\pi(z)f, \pi(z')g \rangle dz = \int \eta_H(z)\langle f, \pi(z')g \rangle dz.
\]

Note that \( S_0 \) is time-frequency shift-invariant, so \( \pi(z)g \) is in \( S_0 \) for all \( z \). Hence, the expression in \((16)\) is well-defined.

If \( f \in L^2(\mathbb{R}) \) and \( H \in \mathcal{H} \), then \( V_g f, V_g Hf \) and \( \eta_H \in L^2(\mathbb{R}^2) \), which is in accordance with the fact that \( L^2 \sharp L^2 \subseteq L^2 \).

If \( f \in S_0(\mathbb{R}) \), then \( H \) may be in \( \mathcal{B}' \), such that \( \eta_H \in S_0'(\mathbb{R}^2) \), hence \( Hf \in S_0'(\mathbb{R}) \). Hence, we have \( V_g f \in S_0'(\mathbb{R}^2) \) and \( V_g Hf \in S_0'(\mathbb{R}^2) \). This leads to the inclusion \( S_0' \sharp W(\mathcal{C}, \ell^1) \subseteq L^\infty \), which may easily be verified directly.

On the other hand, if \( f \in S_0'(\mathbb{R}) \) and \( H \in \mathcal{B} \), such that \( \eta_H \in S_0(\mathbb{R}^2) \), then \( Hf \) is in \( S_0(\mathbb{R}) \). Hence, \( V_g f \in S_0'(\mathbb{R}^2) \) and \( V_g Hf \in S_0'(\mathbb{R}^2) \). Here, this leads to the conclusion that we have, for \( f \in S_0'(\mathbb{R}) \):

\[
S_0' \sharp V_g f \subseteq S_0.
\]

Although it is known that \( V_g f \) is not only \( L^\infty(\mathbb{R}^2) \), but also in the Amalgam space \( W(\mathcal{F}L^1, \ell^\infty) \) for \( f \in S_0'(\mathbb{R}) \) and \( g \in S_0(\mathbb{R}) \), \([13]\), it is not clear, whether \((17)\) also holds for functions \( F \in W(\mathcal{F}L^1, \ell^\infty) \), which are not in the range of \( S_0'(\mathbb{R}) \) under \( V_g \). This and other interesting open questions concerning the twisted convolution of function spaces are currently under investigation.\(^2\)

Remark 11 As a consequence of the last proposition, $H$ may be realized as a twisted convolution in the time-frequency domain:

$$Hf = \frac{1}{\langle g, h \rangle} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left( \eta_H \langle V_g f \rangle (b, \nu) M_{\nu} T_b h db d\nu \right) \text{ for all } f \in (S_0, L^2, S'_0).$$

(18)

Notice that Proposition 10 implies that the range of $V_g$ is invariant under left twisted convolution. Notice also that this is no longer true if the left twisted convolution is replaced with the right twisted convolution. Indeed, in such a case, one has

$$V_g f \eta_H = V_{H^*} g f.$$

Hence, one has the following simple rule: left twisted convolution on the STFT amounts to acting on the analyzed function $f$, while right twisted convolution on the STFT amounts to acting on the analysis window $g$. It is worth noticing that in such a case, applying $V_g^*$ to $V_g f \eta_H$ yields the analyzed function $f$, up to some (possibly vanishing) constant factor.

Example 12 As an illustrative example, let $g, h \in S_0$ be such that $\langle g, h \rangle = 1$ and consider the oblique projection $P : f \mapsto \langle f, g \rangle h$. The spreading function of this operator is given by $V_g h$, and we have $V_\varphi P f (z) = \langle f, g \rangle \langle h, \pi(z) \varphi \rangle$. By virtue of the inversion formula for the STFT, which may be written as $\langle f, g \rangle h = \int \langle h, \pi(z) g \rangle \pi(z) f dz$, we obtain:

$$V_\varphi P f (z') = \int \langle h, \pi(z) g \rangle \langle \pi(z) f, \pi(z') \varphi \rangle dz = V_g h \eta V_\varphi f.$$

By completely analogous reasoning, we obtain the converse formula, if the operator is applied to the analysing window:

$$V_{P_\varphi} f = \langle f, \pi(z) P \varphi \rangle = V_\varphi f \eta V_g h.$$

Remark 13 Notice also that twisted convolution in the phase space is associated with the true translation structure. Indeed, time-frequency shifts take the form of twisted convolutions with a Dirac distribution on $\mathbb{P}$:

$$\delta_{b_0, \nu_0} V_g f = V_g M_{\nu_0} T_{b_0} f.$$

This corresponds to the usage of engineers, who “adjust the phases” after shifting STFT coefficients [10, 21].

2.4 Time-Frequency Multipliers

Section 2.3 has shown the close connection between the spreading function representation of Hilbert-Schmidt operators and the short time Fourier transform. However, the twisted convolution representation is generally of poor practical interest in the continuous case, because it does not discretize well. Even in the finite case, it relies on the full STFT on $\mathbb{C}^N$, which represents vectors with $N^2$ STFT coefficients, which
may be far too large in practice, and sub-sampling is not possible in a straightforward way.

Time-frequency (in particular Gabor) multipliers represent a valuable alternative for time-frequency operator representation (see [18, 29] and references therein for reviews). We analyze below the connections between these representations and the spreading function, and point out some limitations, before turning to generalizations.

2.4.1 Definitions and Main Properties

Let $g, h \in S_0(\mathbb{R})$ be such that $\langle g, h \rangle = 1$, let $m \in L^\infty(\mathbb{R}^2)$, and define the STFT multiplier $M_{m; g, h}$ by

$$M_{m; g, h} f = \int_{\mathbb{R}^2} m(b, \nu) \mathcal{V}_g f(b, \nu) \pi(b, \nu) h \, db \, d\nu.$$  \hspace{1cm} (19)

This defines a bounded operator on $(S_0(\mathbb{R}), L^2(\mathbb{R}), S_0'(\mathbb{R}))$.

Similarly, given lattice constants $b_0, \nu_0 \in \mathbb{R}^+$, set $\pi_{mn} = \pi(mb_0, n\nu_0) = M_{n\nu_0} T_{mb_0}$. Then, for $m \in \ell^\infty(\mathbb{Z}^2)$, the corresponding Gabor multiplier is defined as

$$M_{m; g, h}^G f = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} m(m, n) \mathcal{V}_g f(mb_0, n\nu_0) \pi_{mn} h.$$  \hspace{1cm} (20)

Note that Gabor multipliers may be interpreted as STFT multipliers with multiplier $m$ in $S_0'$. In fact, in this case, $m$ is simply a sum of weighted Dirac impulses on the sampling lattice.

The definition of time-frequency multipliers can of course be given for $g, h \in L^2(\mathbb{R})$, many nice properties only apply with additional assumptions on the windows. Abstract properties of such multipliers have been studied extensively, and we refer to [18] for a review. One may show for example that, whenever the windows $g$ and $h$ are at least in $S_0$, if $m$ belongs to $L^2(\mathbb{P})$ (or $\ell^2(\mathbb{Z}^2)$) then the corresponding multiplier is a Hilbert-Schmidt operator and maps $S_0'(\mathbb{R})$ to $L^2(\mathbb{R})$.

The spreading function of time-frequency multipliers may be computed explicitly.

**Lemma 14** The spreading function of the STFT multiplier $M_{m; g, h}$ is given by

$$\eta_{M_{m; g, h}}(b, \nu) = \mathcal{M}(b, \nu) \mathcal{V}_g h(b, \nu),$$  \hspace{1cm} (21)

where $\mathcal{M}$ is the symplectic Fourier transform of the transfer function $m$

$$\mathcal{M}(t, \xi) = \int_{\mathbb{R}^2} m(b, \nu) e^{2i\pi(vt - \xi b)} \, db \, d\nu.$$  

Specifying to the Gabor multiplier $M_{m; g, h}^G$, we see the same expression for the spreading function, however, in this case, $\mathcal{M} = \mathcal{M}^{(d)}$ is the $(\nu_0^{-1}, b_0^{-1})$-periodic symplectic Fourier transform of the discrete transfer function $m$

$$\mathcal{M}^{(d)}(t, \xi) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} m(m, n) e^{2i\pi(n\nu_0 t - mb_0 \xi)}.$$  \hspace{1cm} (22)
Proof For \( f \in S'_0 \) and \( \varphi \in S_0 \), we may write
\[
\langle M_m; g, h f, \varphi \rangle = \langle m, V_g f \cdot V_h \varphi \rangle,
\]
where the right-hand side inner product has to be interpreted as an integral or infinite sum, respectively. By Lemma 9, applying the symplectic Fourier transform, we obtain
\[
\langle M_m; g, h f, \varphi \rangle = \langle \mathcal{M}, V_f \varphi \cdot V_g h \rangle = \langle \mathcal{M} \cdot V_f \varphi, V_g h \rangle.
\]
This proves (21). By virtue of that fact that for \( g, h \in S_0 \), \( V_g h \) is certainly in \( L^1(\mathbb{P}) \) and even in the Wiener Amalgam Space \( W(C, L^1) \), hence in particular continuous, the expressions for the spreading function given in the lemma are always well-defined. The symplectic Fourier transform is a Gelfand triple isomorphism of \( (S_0, L^2, S'_0) \), i.e., \( m \in (S_0, L^2, S'_0) \iff \mathcal{M} \in (S_0, L^2, S'_0) \). Hence, \( \mathcal{M} \cdot V_f \varphi \in (S_0, L^2, S'_0) \), which is in accordance with the fact, that for \( m \) in \( (\ell^1, \ell^2, \ell^\infty) \), i.e., \( (S_0, L^2, S'_0)(\mathbb{Z}^2) \), the kernel of the resulting operator (and hence its spreading function), is in \( (S_0, L^2, S'_0) \), respectively, see [18] for details. Hence, the expression obtained in (23) is well-defined by duality.

Remark 15 All expressions derived so far are easily generalized to Gabor frames for \( \mathbb{R}^d \) associated to arbitrary lattices \( \Lambda \subset \mathbb{R}^{2d} \). In such situations, the spreading function takes a similar form, and involves some discrete symplectic Fourier transform of the transfer function \( m \), which is in that case a \( \Lambda^\circ \)-periodic function, \( \Lambda^\circ \) being the adjoint lattice of \( \Lambda \).

Notice that as a consequence of Proposition 10, one has the following “intertwining property”
\[
V_g M_h f = (\mathcal{M}_{m; g, h} V_g h) V_g f.
\]

Remark 16 It is clear from the above calculations that a general Hilbert-Schmidt operator may not be well represented by a TF-multiplier. For example, let us assume that the analysis and synthesis windows have been chosen, and let \( \eta \) be the spreading function of the operator under consideration.

- In the STFT case, if the analysis and synthesis windows are fixed, the decay of the spreading function has to be fast enough (at least as fast as the decay of \( V_g h \)) to ensure the boundedness of the quotient \( \mathcal{M} = \eta / V_g h \). Such considerations have led to the introduction of the notion of underspread operators [32] whose spreading function is compactly supported in a domain of small enough area. A more precise definition of underspread operators will be given below.
- In the Gabor case, the periodicity of \( \mathcal{M}^{(d)} \) imposes extra constraints on the spreading function \( \eta \). In particular, the shape of the support of the spreading function must influence the choice of optimal parameters for the approximation by a Gabor multiplier, i.e. the shape of the window as well as the lattice parameters. The following numerical example indicates the direction for the choice of parameters in the approximation of operators by Gabor multipliers.
Example 17 Consider two operators $OP_1$ and $OP_2$ with spreading functions as shown in Fig. 1. The values of the spreading functions are random and real, uniformly distributed in $[-0.5, 0.5]$.

Operator 1 has a spreading function with smaller support on the time-axis, which means that the corresponding operator exhibits time-shifts across smaller intervals than Operator 2, whose spreading function is, on the other hand, less extended in frequency. The effect in the opposite direction is, obviously, reverse. These characteristics are illustrated by applying the operators to a sinusoid with frequency 1 and a Dirac impulse at $-1$, respectively.

Next, we realize approximation by Gabor multipliers with two fixed pairs of lattice constants: $b_1 = 2$, $\nu_1 = 8$ and $b_2 = 8$, $\nu_2 = 2$. Furthermore, the windows are Gaussian windows varying from wide ($j = 0$) to narrow ($j = 100$). Thus, $j$ corresponds to the concentration of the window, in other words, $j$ is the reciprocal of the standard deviation. Now the approximation quality is investigated. The results are shown in the lower plots of Fig. 1, where the left subplot shows the approximation quality for operator $OP_1$ for $b_1$, $\nu_1$ (solid) and $b_2$, $\nu_2$ (dashed), while the right hand subplot gives the corresponding results for operator 2. The error is measured by \( \text{err} = \|OP - APP\|_{\mathcal{H}} / \|OP + APP\|_{\mathcal{H}} \). Here, $APP$ denotes the approximation operator and the norm is the operator norm. The results show that, as expected, the “adapted” choice of time-frequency parameters leads to more favorable approximation quality. Here, the adapted choice of $b$ and $\nu$ mimics the shape of the support of the spreading function according to formula (21) and the periodicity of $M^{(d)}$. In brief, if the operator realizes frequency-shifts in a wider range, we will need more sampling-points in frequency and vice-versa. It is also visible, that the shape of the window has considerable influence on the approximation quality.

The previous example shows, that the parameters in the approximation by Gabor multipliers must be carefully chosen. Let us point out that the approximation quality achieved in the experiment described in Example 17 is not satisfactory, especially when the time- and frequency shift parameters are not well adapted. Operators with a spreading function that is not well-concentrated around 0, i.e. “overspread operators”, don’t seem to be well-represented by a Gabor multiplier even with high redundancy (the redundancy used in the example is 8). Moreover, a realistic operator will have a spreading function with a much more complex shape. The next section will give some more details on approximation by Gabor multipliers before generalizations, which allow for approximation of more complex operators, are suggested.

2.4.2 Approximation by Gabor Multipliers

The possibility of approximating operators by Gabor multipliers in Hilbert-Schmidt sense depends on the properties of the rank one operators associated with time-frequency shifted copies of the analysis and synthesis windows.

---

3 Best approximation is realized in Hilbert-Schmidt sense, see the next section for details.
Let $g, h \in S_0(\mathbb{R})$ be such that $\langle g, h \rangle = 1$. Let $\lambda = (b_1, \nu_1) \in \mathbb{P}$, and consider the rank one operator (oblique projection) $P_\lambda$ defined by

$$P_\lambda f = (g_\lambda^* \otimes h_\lambda)f = \langle f, g_\lambda \rangle h_\lambda, \quad f \in (S_0(\mathbb{R}), L^2(\mathbb{R}), S'_0(\mathbb{R})).$$

(24)

Direct calculations show that the kernel of $P_\lambda$ is given by

$$\kappa_{P_\lambda}(t, s) = \overline{g_\lambda(s)}h_\lambda(t),$$

(25)

and its spreading function reads

$$\eta_{P_\lambda}(b, \nu) = e^{2i\pi(\nu_1 b - b_1 \nu)}V_g h(b, \nu).$$

(26)

The following result characterizes the situations for which time-frequency rank one operators form a Riesz sequence, in which case the best approximation by a Hilbert-Schmidt operator is well-defined. This result first appeared in [15]. Here, we give a slightly different version, which is obtained from the original statement by applying Poisson summation formula. This result was also given in [6] for general full-rank lattices in $\mathbb{R}^d$.

We will subsequently denote by $\Lambda_{b_0, \nu_0}^0 = \frac{1}{\nu_0} \mathbb{Z} \times \frac{1}{b_0} \mathbb{Z}$ the adjoint lattice of $\Lambda_{b_0, \nu_0}^0$, and by $\Box^0 = [\frac{1}{2\nu_0}, \frac{1}{\nu_0}] \times [\frac{1}{2b_0}, \frac{1}{b_0}]$ its fundamental domain.
Proposition 18 Let \( g, h \in L^2(\mathbb{R}) \), with \( \langle g, h \rangle \neq 0 \), let \( b_0, v_0 \in \mathbb{R}^+ \), and set

\[
U(t, \xi) = \sum_{k, \ell = -\infty}^{\infty} \left| V_g h \left( t + \frac{k}{v_0}, \xi + \frac{\ell}{b_0} \right) \right|^2.
\] (27)

The family \( \{ P_{mb_0, n\nu_0}, m, n \in \mathbb{Z} \} \) is a Riesz sequence in \( \mathcal{H} \) if and only if there exist real constants \( 0 < A \leq B < \infty \) such that

\[
0 < A \leq U(t, \xi) \leq B < \infty \quad \text{a.e. on } \Box^0.
\] (28)

We call this condition the \( U \) condition.

Note that validity of the \( U \) condition, via the fact that the subspace spanned by the corresponding Riesz basis of projection operators is closed, guarantees the existence of a unique solution to the problem of best approximation of an arbitrary Hilbert-Schmidt operator in a given Gabor multiplier setting. Accordingly it turns out, that the approximation of a given operator via a standard minimization process yields an expression, which is only well-defined if the \( U \) condition (28) holds.

Theorem 19 Assume that \( V_g h \) and \( b_0, v_0 \in \mathbb{R}^+ \) are such that the \( U \) condition (28) is fulfilled. Then the best Gabor multiplier approximation (in Hilbert-Schmidt sense) of \( H \in \mathcal{H} \) is defined by the time-frequency transfer function \( m \) whose discrete symplectic Fourier transform reads

\[
M(b, \nu) = \sum_{k, \ell = -\infty}^{\infty} \frac{V_g h(b + k/v_0, \nu + \ell/b_0) \eta_H(b + k/v_0, \nu + \ell/b_0)}{\sum_{k, \ell = -\infty}^{\infty} |V_g h(b + k/v_0, \nu + \ell/b_0)|^2}.
\] (29)

Proof Let us set \( V = V_g h \) for simplicity of notation. First, notice that if \( \eta_H \in L^2(\mathbb{P}) \), then the function \( (b, \nu) \in \Box^0 \rightarrow \sum_{k, \ell} |\eta_H(b + k/v_0, \nu + \ell/b_0)|^2 \) is in \( L^2(\Box^0) \), and is therefore well defined almost everywhere in \( \Box^0 \). Thus, by Cauchy-Schwarz inequality, the numerator in (29) is well-defined a.e.

The Hilbert-Schmidt optimization is equivalent to the problem

\[
\min_{M \in L^2(\Box^0)} \| \eta_H - M V \|^2.
\]

The latter squared norm may be written as

\[
\| \eta_H - M V \|^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\eta_H(b, \nu) - M(b, \nu) V(b, \nu)|^2 \, db \, d\nu
\]

\[
= \sum_{k, \ell = -\infty}^{\infty} \int_{\Box^0} \int_{\Box^0} |\eta_H(b + k/v_0, \nu + \ell/b_0) - M(b, \nu) V(b + k/v_0, \nu + \ell/b_0)|^2 \, db \, d\nu
\]

\[
= \int_{\Box^0} \int_{\Box^0} \left[ \sum_{k, \ell} |\eta_H(b + k/v_0, \nu + \ell/b_0)|^2 \right] \, db \, d\nu
\]
From this expression, the Euler-Lagrange equations may be obtained, which read

\[ M(b, \nu) \sum_{k, \ell} |V(b + k/\nu_0, \nu + \ell/b_0)|^2 = \sum_{k, \ell} \eta_H(b + k/\nu_0, \nu + \ell/b_0) V(b + k/\nu_0, \nu + \ell/b_0), \]

and the result follows.

We next derive an error estimate for the approximation. Let us set, for \((b, \nu) \in \Box^0\),

\[ \mathcal{E}(b, \nu) = \left| \sum_{k, \ell} \eta_H(b + k/\nu_0, \nu + \ell/b_0) V(b + k/\nu_0, \nu + \ell/b_0) \right|^2 \frac{\sum_{k, \ell} \eta_H(b + k/\nu_0, \nu + \ell/b_0)|^2}{U(b, \nu)}. \]

**Corollary 20** With the above notation, we obtain the estimate

\[ \| H - M_m \|_H^2 \leq \| \eta_H \|_H^2 \| 1 - \mathcal{E} \|_{\infty} \]

for the best approximation of \( H \) by a Gabor multiplier \( M_m \) according to (29).

**Proof** Set \( \Gamma_H(b, \nu) = \sum_{k, \ell} |\eta_H(b + k/\nu_0, \nu + \ell/b_0)|^2 \). Replacing the expression for \( \mathcal{M} \) obtained in (29) into the error term, we have

\[ \| H - M_m \|_H^2 = \int \int_{\Box^0} \left[ \Gamma_H(b, \nu) - |\mathcal{M}(b, \nu)|^2 U(b, \nu) \right] dbd\nu \]

\[ = \int \int_{\Box^0} \Gamma_H(b, \nu)[1 - \mathcal{E}(b, \nu)] dbd\nu \]

\[ \leq \| \eta_H \|_H^2 \| 1 - \mathcal{E} \|_{\infty}, \]

where we have used the fact that \( \| \Gamma_H \|_{L^2(\Box^0)} = \| \eta_H \|_{L^2(\Box)} \). Clearly, Cauchy-Schwarz inequality gives \( |\mathcal{E}(b, \nu)| \leq 1 \) on \( \Box^0 \), with equality if and only if there exists a function \( \phi \) such that \( \eta = \phi V \), i.e. if and only if \( H \) is a multiplier with the prescribed window functions. Hence we obtain

\[ \| H - M_m \|_H^2 \leq \| H \|_H^2 \left[ 1 - \text{ess inf}_{(b, \nu) \in \Box^0} \mathcal{E}(b, \nu) \right]. \]  

\( \square \)
Remark 21 \( \mathcal{E}(b, \nu) \) essentially represents the cosine of the angle between vectors 
\( \{ \eta_H(b + k/v_0, v + \ell/b_0), k, \ell \in \mathbb{Z} \} \) and \( \{ V(b + k/v_0, v + \ell/b_0), k, \ell \in \mathbb{Z} \} \). In other words, the closer to colinear these vectors, the better the approximation.

An example for operators which are poorly represented by this class of multipliers are those with a spreading function that is not “well-concentrated”. These are, in technical terms, overspread operators. The underspread/overspread terminology seems to originate from the context of time-varying multipath wave propagation channels [31]. However, different definitions exist in the literature. Here, we give the definition used in [33]. Note that underspread operators have recently found renewed interest [7, 33–35, 37, 39].

**Definition 22** Consider an operator \( K \) with compactly supported spreading function:
\[
\text{supp}(\eta_K) \subseteq Q(t_0, \xi_0), \quad \text{where } Q(t_0, \xi_0) := [-t_0, t_0] \times [-\xi_0, \xi_0].
\]
Then, \( K \) is called underspread, if \( \text{vol}(Q) < 1 \).

Most generally, operators which are not underspread, will be called overspread. It is generally known, that an operator must be underspread in order to be well-approximated by Gabor multipliers. Formula (29) enables us to make this statement more precise.

**Corollary 23** Consider an underspread operator \( H \). Choose \((b_0, \nu_0)\) such that \( 1/b_0 > 2\xi_0 \) and \( 1/\nu_0 > 2t_0 \) and consider the Gabor frame \( \{ g_{mn}, m, n \in \mathbb{Z} \} \), with lattice constants \((b_0, \nu_0)\), and dual window \( h \). Then, the symplectic Fourier transform \( \mathcal{M} \) of the time-frequency transfer function of the best Gabor multiplier takes the form
\[
\mathcal{M} = \eta_H \frac{\mathcal{V}_gh}{\mathcal{U}},
\]
and the approximation error can be bounded by
\[
\| \eta_H - \mathcal{M} \mathcal{V}_gh \| \leq \| \eta_H \| \quad \text{ess sup} \quad \frac{1 - |\mathcal{V}_gh(b, \nu)|^2}{\mathcal{U}(b, \nu)},
\]
with \( \Box_H^o = \Box^o \cap \text{supp}(\eta_H) \).

**Proof**
\[
\| \eta_H - \mathcal{M} \mathcal{V}_gh \| = \int \int_{\Box^o} \left( |\eta(b, \nu)|^2 - \frac{|\eta(b, \nu)|^2}{\mathcal{U}(b, \nu)} \right) db d\nu \leq \| \eta_H \| \quad \text{ess sup} \quad \frac{1 - |\mathcal{V}_gh(b, \nu)|^2}{\mathcal{U}(b, \nu)}. \]

The estimate in the last corollary shows, that approximation quality is a joint property of window and lattice, which is in accordance with the results of Example 17.
Remark 24 Note that, although technically only defined for Hilbert-Schmidt operators, the approximation by Gabor multipliers can formally be extended to operators from $\mathcal{B}'$, see [18, Sect. 5.8]. Also, the expression given in (29) is well-defined in $S'_0$ whenever $\eta_H$ is at least in $S'_0$. However, for non-Hilbert-Schmidt operators, it is not clear, in which sense the resulting Gabor multiplier represents the original operator. The following example shows that at least in some cases, the result is however the intuitively expected one.

Example 25 Consider the operator $\pi(\lambda)$, i.e. a time-frequency shift. Although this operator is clearly not a Hilbert-Schmidt operator, we may consider its approximation by a Gabor-multiplier according to (29). First note that the spreading function of the time-frequency shift $\pi(\lambda) = \pi(b_1, v_1)$ is given by $\eta_{\pi} = \delta(b - b_1) \cdot \delta(v - v_1)$. Then, we have

$$M(b, v) = \sum_{k,l} \frac{\mathcal{V}_g \overline{h}(b + k/v_0, v + l/b_0) \delta(b - b_1 + k/v_0, v - v_1 + l/b_0)}{\sum_{k',l'} |\mathcal{V}_g h(b + k'/v_0, v + l'/b_0)|^2}$$

$$= \frac{\mathcal{V}_g h(b_1, v_1)}{U(b_1, v_1)} \sum_{k,l} \delta(b - b_1 + k/v_0, v - v_1 + l/b_0).$$

Hence, from the inverse (discrete) symplectic Fourier transform we obtain:

$$m(m, n) = \frac{\mathcal{V}_g h(b_1, v_1)}{U(b_1, v_1)} \int_0^{1/v_0} \int_0^{1/b_0} \sum_{k,l} \delta(b - b_1 + k/v_0, v - v_1 + l/b_0)$$

$$\times e^{-2\pi i (mb_0v - n/v_0b)} dbdv$$

$$= \frac{\mathcal{V}_g h(b_1, v_1)}{U(b_1, v_1)} e^{-2\pi i (mb_0v_1 - n/v_0b_1)}.$$

As expected, the absolute value of the mask is constant and the phase depends on the displacement of $\lambda$ from the origin. This confirms the key role played by the phase of the mask of a Gabor multiplier. Specializing to $\lambda = 0$, we obtain a constant mask and thus, if $h$ is a dual window of $g$ with respect to $\Lambda = b_0\mathbb{Z} \times v_0\mathbb{Z}$, up to a constant factor, the identity.

3 Generalizations: Multiple Gabor Multipliers and TST Spreading Functions

In the last section it has become clear that most operators are not well represented as a STFT or Gabor multiplier.

Guided by the desire to extend the good approximation quality that Gabor multipliers warrant for underspread operators to the class of their overspread counterparts, we introduce generalized TF-multipliers. The basic idea is to allow for an extended
scheme in the synthesis part of the operator: instead of using just one window \( h \), we suggest the use of a set of windows \( \{h^{(j)}\} \) in order to obtain the class of **Multiple Gabor Multipliers** (MGM for short).

**Definition 26** (Multiple Gabor Multipliers) Let \( g \in S_0(\mathbb{R}) \) and a family of reconstruction windows \( h^{(j)} \in S_0(\mathbb{R}) \), \( j \in J \), as well as corresponding masks \( m_j \in \ell^\infty \) be given. Operators of the form

\[
\mathcal{M} = \sum_{j \in J} M_{M^{G}_{m_j;g,h^{(j)}}}^{G}
\]

will be called **Multiple Gabor Multipliers** (MGM for short).

**Remark 27**

(a) The most natural generalization of Gabor multipliers would allow for both several analysis and synthesis windows. However, we will see below that the generalization introduced in (31) is sufficient in order to achieve perfect representation of most operators of interest.

(b) Note, that we need to impose additional assumptions in order to obtain a well-defined operator. For example, we may assume \( \sum_j \sup_\lambda |m_j(\lambda)| = C < \infty \) and \( \max_j \|h^{(j)}\|_{S_0} = C < \infty \), which guarantees a bounded operator on \( (S_0, L^2, S'_0) \). This follows easily from the boundedness of a Gabor multiplier under the condition that \( m \) is \( \ell^\infty \). Conditions for function space membership of MGMs are easily derived in analogy to the Gabor multiplier case. For example, if \( \sum_j \sum_\lambda |m_j(\lambda)|^2 < \infty \), we obtain a Hilbert-Schmidt operator, similarly, trace-class membership follows from an analogous \( \ell^1 \)-condition.

(c) The spreading function of a MGM is (formally) given by a sum of the spreading functions corresponding to the single Gabor multipliers involved:

\[
\eta^{G}_{M_{m_j;g,h^{(j)}}}(b,\nu) = \sum_{j \in J} \mathcal{M}^{(j)}(b,\nu) \mathcal{V}_{g^{(j)}}(b,\nu),
\]

where the \( (v_0^{-1}, b_0^{-1}) \)-periodic functions \( \mathcal{M}^{(j)} \) are the symplectic Fourier transforms of the transfer functions \( m_j \).

(d) Note that in practice \(|J|\) will often be finite. Let us mention that by assuming \( m_j \in \ell^2(\mathbb{Z}^3) \), we have \( \|\eta^{G}_{M_{m_j;g,h^{(j)}}}\|_2 = C\|m\|_2 \).

It is immediately obvious that the new model gives much more freedom in generating overspread operators. However, in order to obtain structural results, we will have to impose further specifications.

Before doing so, we will state a generalization of Proposition 18 to the more general situation of the family of projection operators \( P_{\lambda}^j \) defined by

\[
P_{\lambda}^j f = (g^*_\lambda \otimes h^j_\lambda) f = \langle f, g_\lambda \rangle h^j_\lambda, \quad \text{where } \lambda \in \Lambda, \; j \in J.
\]

Note that these projection operators are the building blocks for the MGM. The following theorem characterizes their Riesz property.
Proposition 28  Let $g, h^j \in L^2(\mathbb{R})$, $j \in J$, with $\langle g, h^j \rangle \neq 0$, let $b_0, v_0 \in \mathbb{R}^+$, and let the matrix $\Gamma(b, v)$ be defined by

$$
\Gamma(b, v)_{jj'} = \sum_{k, \ell} \mathcal{V}_g h^{(j)}(b + k/v_0, v + \ell/b_0) \mathcal{V}_g h^{(j')}(b + k/v_0, v + \ell/b_0)
$$

(34)

a.e. on $\square^o = [0, v_0^{-1}] \times [0, b_0^{-1}]$. Then the family of projection operators $\{P^i_\lambda, j \in \mathbb{Z}, \lambda \in \Lambda\}$ is a Riesz sequence in $\mathcal{H}$ if and only if $\Gamma$ is invertible a.e. with uniform bounds $0 < A \leq B < \infty$ such that

$$
A \leq \Gamma(b, v) \leq B \quad \text{a.e. on } \square^o.
$$

(35)

Alternatively, the Riesz basis property is characterized by bounded invertibility of the matrix $U$ defined as

$$
U_{jj'}(t, \xi) = \sum_{k, l} \left[\mathcal{V}_g \cdot \mathcal{V}_{h^j} h^{j'}\right](kb_0, lv_0) e^{-2\pi i(lv_0t - kb_0\xi)}
$$

(36)

a.e. on the fundamental domain of $\Lambda$.

Proof  Recall that the family $\{P^i_\lambda, j \in \mathbb{Z}, \lambda \in \Lambda\}$ is a Riesz basis for its closed linear span if there exist constants $0 < A, B < \infty$ such that

$$
A \|c\|_2^2 \leq \left\| \sum_\lambda \sum_j c^j_\lambda P^j_\lambda \right\|^2_{\mathcal{H}\mathcal{S}} \leq B \|c\|_2^2
$$

(37)

for all finite sequences $c$ defined on $(\Lambda \times J)$. We have

$$
\left\| \sum_\lambda \sum_j c^j_\lambda P^j_\lambda \right\|^2_{\mathcal{H}\mathcal{S}} = \left\langle \sum_\lambda \sum_j c^j_\lambda P^j_\lambda, \sum_\mu \sum_{j'} c'^{j'}_{\mu} P'^{j'}_{\mu} \right\rangle_{\mathcal{H}\mathcal{S}} = \sum_{\lambda \mu} \sum_{jj'} c^j_\lambda \overline{c'^{j'}_{\mu}} \langle g^j_\lambda, g^j'_{\mu} \rangle \langle h^j_\lambda, h^{j'}_{\mu} \rangle.
$$

Hence, by setting $U_{jj'}(b, v) = [\mathcal{V}_g \cdot \mathcal{V}_{h^j} h^{j'}](b, v)$, we may write

$$
\left\| \sum_\lambda \sum_j c^j_\lambda P^j_\lambda \right\|^2_{\mathcal{H}\mathcal{S}} = \sum_{\lambda \mu} \sum_{jj'} c^j_\lambda \overline{c'^{j'}_{\mu}} U_{jj'}(\mu - \lambda)
$$

$$
= \sum_{\mu} \sum_{jj'} c'^{j'}_{\mu} U_{jj'}(\mu)
$$

$$
= b_0 v_0 \sum_{jj'} \langle U_{jj'}, C^j, C'^{j'} \rangle_{L^2(\square^o),}
$$

where $U_{jj'}(t, \xi) = \sum_{k,l} U_{jj'}(kb_0, lv_0) e^{-2\pi i(lv_0t - kb_0\xi)}$ is the discrete symplectic Fourier transform of $U_{jj'}$, and, analogously, $C^j$ is the discrete symplectic Fourier transform of $C^j$. \hfill \Box
transform of the sequence $c_j$, defined on $\Lambda$, for each $j$. Hence, these are $\nu_0^{-1}b_0^{-1}$-periodic functions. The last equation can be rewritten as

$$
\left\| \sum_{\lambda} \sum_j c_{j,\lambda} P_{j,\lambda} \right\|_2^2 = b_0v_0 \sum_{j,j'} \int_{\mathbb{R}^2} U^{jj'}(t,\xi) \cdot C^j(t,\xi) \overline{C'}^{jj'}(t,\xi) dt d\xi
$$

$$
= b_0v_0 \langle U \cdot C, C \rangle_{\mathcal{L}^2(\mathbb{R}^2) \times \ell^2(\mathbb{Z})},
$$

where $U$ is the matrix with entries $U^{jj'}$. Note that this proves statement (36) by positivity of the operator $U$.

In order to obtain the condition for $\Gamma$ given in (34), first note that $F_s(V_{gg} \cdot V_{hh}^{jj'})(\lambda) = (V_{gh}^{jj'} \cdot V_{gh}^{jj'})(\lambda)$ by applying Lemma 9. Furthermore, $F := V_{gh}^{jj'} \cdot V_{gh}^{jj'}$ is always in $L^1$ for $g,h \in L^2$. We may therefore look at the Fourier coefficients of its $\Lambda^\circ$-periodization, with $\lambda = (mb_0, n\nu_0)$:

$$
\mathcal{F}^{-1}_s(P_{\Lambda^\circ} F)(\lambda) = \int_{\mathbb{R}^2} \left( \sum_{k,l} F\left( b + \frac{k}{\nu_0}, v + \frac{l}{b_0} \right) \right) e^{2\pi i \left( bn\nu_0 - vmb_0 \right)} db dv
$$

$$
= \int_{\mathbb{R}^2} \left( \sum_{k,l} F\left( b + \frac{k}{\nu_0}, v + \frac{l}{b_0} \right) \right) e^{2\pi i \left( b + \frac{k}{\nu_0} \right) n\nu_0 - \left( v + \frac{l}{b_0} \right) mb_0} db dv
$$

$$
= \int_{\mathbb{R}^2} F(b, v) e^{2\pi i \left( bn\nu_0 - vmb_0 \right)} db dv = \mathcal{F}^{-1}_s(V_{gh}^{jj'} \cdot V_{gh}^{jj'})(\lambda).
$$

Hence, we may apply the Poisson summation formula, with convergence in $L^2(\mathbb{R}^2)$, to obtain:

$$
P_{\Lambda^\circ}(V_{gh}^{jj'} \cdot V_{gh}^{jj'})(b, v) = b_0v_0 \sum_{k,l} \mathcal{F}^{-1}_s(V_{gh}^{jj'} \cdot V_{gh}^{jj'})(kb_0, l\nu_0) e^{-2\pi i(\ell v_0 - \nu kb_0)}
$$

$$
= b_0v_0 \sum_{k,l} \overline{V_{gh}^{jj'}} \cdot V_{hh}^{jj'}(kb_0, l\nu_0) e^{-2\pi i(\ell v_0 - \nu kb_0)}.
$$

We conclude that

$$
\left\| \sum_{j} \sum_{\lambda \in \Lambda} c_{j,\lambda} P_{j,\lambda} \right\|_2^2 = b_0v_0 \sum_{j,j'} \int_{\mathbb{R}^2} \Gamma(b, v) j^{jj'} C^j(b, v) \overline{C'}^{jj'}(b, v) db dv,
$$

and the Riesz basis property is equivalent to the invertibility of $\Gamma$. □

In the sequel, the discrete symplectic Fourier transforms of $m_j$ will be denoted by $\mathcal{M}_j$, and the vector with $\mathcal{M}_j$ as coordinates will be denoted by $\mathcal{M}$. We then obtain an expression for the best multiplier in analogy to the Gabor multiplier case discussed in Theorem 19.
Proposition 29 Let \( g \in S_0(\mathbb{R}) \) and \( h^{(j)} \in S_0(\mathbb{R}) \), \( j \in J \) be such that for almost all \( b, v \), the matrix \( \Gamma(b, v) \) defined in (34) is invertible a.e. on \( \square^\circ \).

Let \( H \in (\mathcal{B}, \mathcal{H}, \mathcal{B}') \) be an operator with spreading function \( \eta \in (S_0, L^2, S'_0) \). Then the functions \( M_j \) yielding approximation of the form (31) may be obtained as

\[
M = \Gamma^{-1} \cdot \mathcal{B},
\]

where \( \mathcal{B} \) is the vector whose entries read

\[
\mathcal{B}_{j_0}(b, v) = \sum_{k, \ell} \eta(b + k/v_0, v + \ell/b_0) \overline{\mathcal{V}}_g h^{j_0}(b + k/v_0, v + \ell/b_0).
\]

For operators in \( \mathcal{H} \) the obtained approximation is optimal in Hilbert-Schmidt sense.

Proof The proof follows the lines of the Gabor multiplier case. The optimal approximation of the form (31) is obtained by minimizing

\[
\left\| \eta - \sum_j M_j \mathcal{V}_j \right\|^2 = \sum_{k, \ell} \int_{\square^\circ} \left| \eta(b + k/v_0, v + \ell/b_0) - \sum_j M_j(b, v) \mathcal{V}_j(b + k/v_0, v + \ell/b_0) \right|^2 dbd\nu
\]

where one has set \( \mathcal{V}_j = \mathcal{V}_g h^{(j)} \). Setting to zero the Gâteaux derivative with respect to \( M_{j_0} \), we obtain the corresponding variational equation

\[
\sum_j M_j(b, v) \sum_{k, \ell} \mathcal{V}_j(b + k/v_0, v + \ell/b_0) \overline{\mathcal{V}}_{j_0}(b + k/v_0, v + \ell/b_0) = \mathcal{B}_{j_0}(b, v),
\]

where \( \mathcal{B}_j(b, v) \) are as defined in (40). Provided that the \( \Gamma(b, v) \) matrices are invertible for almost all \( b, v \), this implies that the functions \( M_j \) for approximation of the form (31) may indeed be obtained as in (39). \( \square \)

In a next step, we are going to discern two basic approaches:

(a) \( m_j(\lambda) = m(\mu, \lambda) \), i.e. the synthesis windows are time-frequency shifted versions (on a lattice) of a single synthesis window: \( h_j = \pi(\mu_j)h, \mu_j \in \Lambda_1 \).

(b) \( m_j(\lambda) = m_1(\lambda)m_2(j) \), i.e. a separable multiplier function. If we set \( h^{(j)}(t) = \pi(b_j, v_j)h(t) \) then this approach leads to what will be called TST spreading functions in Sect. 3.1.

In both cases we will be especially interested in the situation in which the \( h^{(j)} \) are given as time-frequency shifted versions of a single synthesis window on the adjoint lattice \( \Lambda^\circ \).

We fix the synthesis windows \( h_j \) to be time-frequency translates of a fixed window function, i.e.

\[
h^{(j)}(t) = \pi(b_j, v_j)h(t) = e^{2i\pi v_j t}h(t - b_j).
\]

(41)
We may turn our attention to the projection operators associated with two fixed (Gabor) families \((g, \Lambda_1)\) and \((h, \Lambda_2)\). Note that it has been shown by Benedetto and Pfander [6] that the family of projection operators \(\{P_\lambda, \lambda \in \Lambda\}\), as discussed in Sect. 2.4 either forms a Riesz basis or not a frame (for its closed linear span). The next corollary shows that, on the other hand, if we use the extended family of projection operators \(\{P_{\lambda, \mu}, (\lambda, \mu) \in \Lambda_1 \times \Lambda_2\}\), where \(P_{\lambda, \mu} f = \langle f, \pi(\lambda)g \rangle \pi(\mu) h\), we obtain a frame of operators for the space of Hilbert-Schmidt operators, whenever \((g, \Lambda_1)\) and \((h, \Lambda_2)\) are Gabor frames. This corollary is a special case of Theorem 4.1 in [3] and Proposition 3.2 in [4].

**Corollary 30** Let two Gabor frames \((g, \Lambda_1)\) and \((h, \Lambda_2)\) be given. Then the family of projection operators \(\{P_{\lambda, \mu}, (\lambda, \mu) \in \Lambda_1 \times \Lambda_2\}\) form a frame of operators in \(\mathcal{H}\) and any Hilbert-Schmidt operator \(H\) may be expanded as

\[
H = \sum_{\lambda \in \Lambda_1, \mu \in \Lambda_2} c(\lambda, \mu) P_{\lambda, \mu}.
\]

The coefficients are given by \(c(\lambda, \mu) = \langle H, (P_{\lambda, \mu})^* \rangle = \langle H \pi(\mu) h, \pi(\lambda) g \rangle\).

Note that an analogous statement holds for Riesz sequences. Very recently, it has been shown [1], that the converse of Corollary 30 holds true for both frames and Riesz bases, i.e. the family of projection operators \(\{P_{\lambda, \mu}, (\lambda, \mu) \in \Lambda_1 \times \Lambda_2\}\) is a frame (a Riesz basis) for \(\mathcal{H}\) if and only if the two generating sequences form a frame (a Riesz basis) for \(L^2(\mathbb{R})\). In particular, this leads to the conclusion, that the characterization of Riesz sequences given in Proposition 28 also yields a characterization of frames for \(L^2(\mathbb{R})\)—it is well known, that \((g_\lambda, \lambda \in \Lambda)\) form a Gabor frame if and only if \((g_\mu, \mu \in \Lambda^\circ)\) form a Riesz sequence. We can draw two conclusions.

**Corollary 31** Let \(g \in L^2(\mathbb{R})\) and a lattice \(\Lambda = b_0 \mathbb{Z} \times v_0 \mathbb{Z}\) be given.

(a) The Gabor family \(\{g_\lambda, \lambda \in \Lambda\}\) forms a frame for \(L^2(\mathbb{R})\) if and only if the matrix

\[
\Gamma_{mn,m' n'}(b, v) = \sum_{k, \ell} \exp \left( -2i\pi [m/v_0 (v + \ell v_0 - n/b_0) - m'/v_0 (v + \ell v_0 - n'/b_0)] \right)
\]

\[
\times \mathcal{V}_g(b - m/v_0 + kb_0, v - n/b_0 + \ell v_0)
\]

\[
\times \mathcal{V}_g(b - m'/v_0 + kb_0, v - n'/b_0 + \ell v_0)
\]

is, a.e. on \(\square^\circ\), invertible on \(\ell^2\).

(b) In addition, we may state the following “Balian-Low Theorem for the tensor products of Gabor frames”:

A family of projection operators given by \(\{P_{\lambda, \mu} = g_\lambda^* \otimes g_\mu, (\lambda, \mu) \in \Lambda \times \Lambda^\circ\}\) forms a frame for the space of Hilbert-Schmidt operators on \(L^2(\mathbb{R})\) if and only if it forms a Riesz basis. Hence, in this case, \(g\) cannot be in \(S_0(\mathbb{R})\).
Proof Statement (a) is easily obtained from (34) by observing that
\[ V g \pi (mb_0, n \nu_0) g(b, \nu) = e^{-2i \pi mb_0 (\nu - n \nu_0)} V g \pi \varepsilon (b - mb_0, \nu - n \nu_0). \]
We then have that \{P_{\lambda, \mu} = g_{\lambda}^* \otimes g_{\mu}, (\lambda, \mu) \in \Lambda^\circ \times \Lambda^\circ \} forms a Riesz basis in \( \mathcal{H} \) if and only if \( \Gamma_{mn,m'n'}(b, \nu) \) is invertible. By the converse of Corollary 30, this is equivalent to the Riesz property of \( g_{\mu}, \mu \in \Lambda^\circ \), which, in turn, is equivalent to the frame property of \( \{g_{\lambda}, \lambda \in \Lambda \} \) by Ron-Shen duality, see, e.g. [25].

To see (b), note that in this case \( P_{\lambda, \mu} \) is a frame for \( \mathcal{H} \) \iff \( (g_{\lambda}/\Lambda^\circ) \) and \( (g_{\mu}/\Lambda^\circ) \) form a Riesz basis \iff \{P_{\lambda, \mu} = g_{\lambda}^* \otimes g_{\mu}, (\lambda, \mu) \in \Lambda \times \Lambda^\circ \} \) is a Riesz basis for \( \mathcal{H} \). Furthermore, by the classical Balian-Low theorem, if a Gabor system is an \( L^2 \)-frame and at the same time a Riesz sequence (hence an \( L^2 \)-Riesz basis), then the generating window \( g \) cannot be in \( S_0 \).

We may next ask, when the projection operators form a Riesz sequence, if the reconstruction windows are TF-shifted versions of a single window \( h \) on the adjoint lattice of \( \Lambda = b_0 \mathbb{Z} \times v_0 \mathbb{Z} \). In fact, in this case, the matrix \( \Gamma \) turns out to enjoy quite a simple form. To fix some notation, let
\[ A_{mn}(b, \nu) = \sum_{k, \ell} e^{2i \pi m[v - \ell/\nu_0]} V g \pi \varepsilon (b - k/\nu_0, \nu - \ell/b_0) \times V g \pi \varepsilon (b - (k - m)/\nu_0, \nu - (\ell - n)/b_0), \]
and introduce the right twisted convolution operator
\[ K^{\natural}_A(b, \nu) : \mathcal{M}(b, \nu) \rightarrow \mathcal{M}(b, \nu)^{\natural} A(b, \nu). \]

We may derive the form of \( \Gamma \) given in the following corollary by direct calculation:

**Corollary 32** Let \( g, h \in L^2(\mathbb{R}) \) as well as \( b_0, v_0 \) be given. Furthermore, let \( h^{(j)} = \pi(\pi(m/v_0, n/b_0))h \). Then the variational equations read
\[ \mathcal{M}(b, \nu)^{\natural} A(b, \nu) = B(b, \nu). \] (42)
Hence, if for all \( b, \nu \in \mathbb{R}^2 \), the discrete right twisted convolution operator \( K^{\natural}_A \) is invertible, then the family \( P_{\lambda, \mu} = g_{\lambda}^* \otimes g_{\mu}, (\lambda, \mu) \in \Lambda \times \Lambda^\circ \) forms a Riesz sequence and the best MGM approximation of an Hilbert-Schmidt operator with spreading function \( \eta \) is given by the family of transfer functions
\[ \mathcal{M}_{mn}(b, \nu) = \left[ (K^{\natural}_A(b, \nu))^{-1} B(b, \nu) \right]_{mn}, \]
where \( B \) is given in (40).

We close this section with some results of numerical experiments testing the approximation quality of MGMs for slowly time-varying systems. Note that more extensive numerical results are presented in [12].

\(^4\)More precisely, \( g \) cannot even be in the space of continuous functions in the Wiener space \( W(\mathbb{R}) \), see [25, Theorem 8.4.1].
**Example 33** We study the approximation of a (slowly) time-varying operator. The operator has been generated by perturbing a time-invariant operator. The spreading function is shown in the upper display of Fig. 2. The signal length is 32, time- and frequency-parameters are $b_0 = 4$ and $v_0 = 4$, such that the redundancy of the Gabor frame used in the MGM approximation is 2. The approximation is then realized in several steps for two different schemes. Scheme 1 adds three synthesis window corresponding to a frequency-shift by 4, a time-shift by 4 and a time-frequency-shift by $(4, 4)$. The first step 1 calculates the regular Gabor multiplier approximation. Step 2 adds one (only frequency-shift) and so on. The rank of the resulting operator families is 64, 128 for both step 2 and step 3 (adding either time- or frequency-shift) and 256 (time-shifted, frequency-shifted and time-frequency-shifted window added). The resulting approximation-errors are given by the solid line in the lower display of Fig. 2.

Scheme 2 considers synthesis windows shifted in time and frequency on the sublattice generated by $a = 8$, $b = 8$, the resulting families having rank 64, 256 and 576. Here, we only plot the results for the case corresponding to three and eight additional synthesis windows, respectively. The results are given by the dotted line.

For comparison, an approximation with a regular Gabor multiplier with redundancy 8, i.e. an approximation family of rank 256, has been performed. The approximation error for this situation is the diamond in the middle of the display. It is easy to see that, depending on the behavior of the spreading function, different schemes per-
form advantageously for a certain redundancy. Note that for scheme 1, the best MGM with the same rank as the regular Gabor multiplier performs better than the latter. In the case of the present operator, scheme 2 performs the “wrong” time-frequency shifts on the synthesis windows in order to capture important characteristics of the operator. However, in a different setting, this scheme might be favorable (e.g. if an echo with a longer delay is present).

The example shows, that the choice of an appropriate sampling scheme for the synthesis windows is extremely important in order to achieve a good and efficient approximation by MGM. An optimal sampling scheme depends on the analysis window’s STFT, the lattice used in the analysis and on the behavior of the operator’s spreading function, which reflects the amount of delay and Doppler-shift created by the operator. Additionally, structural properties of the family of projections operators used in the approximation, based on the results in this section, have to be exploited to achieve numerical efficiency. An algorithm for optimization of these parameters is currently under development.\(^5\)

3.1 Varying the Synthesis Window: TST Spreading Functions

We next turn to the special case of separable functions \(m_j(\lambda) = m_1(\lambda)m_2(j)\) for the mask in the definition (31) of MGMs. In this case the resulting operator is of the form

\[
\mathbb{M}f = \sum_\lambda m_1(\lambda)\rho(\lambda) \left( \sum_j m_2(j)(g^* \otimes h^j) \right)(f) = \sum_\lambda m_1(\lambda)\rho(\lambda)\mathbb{P}_m f,
\]

where \(\mathbb{P}_m f = \sum_j m_2(j)(f, g)h^j\) and \(\rho(\lambda)\) denotes a tensor product of time-frequency shifts:

\[
\rho(\lambda)H := \pi(\lambda)H\pi^*(\lambda).
\]

Hence, the spreading function of \(\mathbb{M}\) is given by

\[
\eta_{\mathbb{M}} = \mathcal{M} \cdot \eta_{\mathbb{P}_m},
\]

where \(\mathcal{M}\) is the discrete symplectic FT of \(m_1\). If the reconstruction windows are given by \(h^j = \pi(\mu^j)h, \mu_j = (b_j, \nu_j)\), this becomes

\[
\eta_{\mathbb{M}} = \mathcal{M} \cdot \sum_j m_2(j)V g h(\lambda - \mu_j)e^{-2\pi i(\nu - \nu_j)b_j}.
\]

Motivated by this result, we introduce the following definition.

**Definition 34** (TST spreading functions) Let \(\phi\) be a given function from the function spaces \((S_0(\mathbb{R}^2), L^2(\mathbb{R}^2), S'_0(\mathbb{R}^2))\) and let \(b_1, \nu_1\) denote positive numbers. Let \(\alpha\) be

A spreading function $\eta = \eta_H$ of $H \in (\mathcal{B}, \mathcal{H}, \mathcal{B}')$, that may be written as
\begin{equation}
\eta(b, \nu) = \sum_{k, \ell} \alpha_{k\ell} \phi(b - kb_1, \nu - \ell \nu_1) e^{-2i\pi (\nu - \ell \nu_1) kb_1} \tag{43}
\end{equation}
will be called Twisted Spline Type function (TST for short).

**Remark 35** By $\alpha$ in $\ell^1$, the series defining $\eta$ is absolutely convergent in $(S_0, L^2, S'_0)$. For $\ell^2$-sequences $\alpha$, we obtain an $L^2$-function $\eta$ for $\phi \in L^2$.

TST functions are nothing but spline type functions (following the terminology introduced in [15]), in which usual (Euclidean) translations are replaced with the natural (i.e. $\mathbb{H}$-covariant) translations on the phase space $\mathbb{P}$. In fact, by writing $\alpha(b, \nu) = \sum_{k, \ell} \alpha_{k\ell} \delta(b - kb_1, \nu - \ell \nu_1)$, the TST spreading function may be written as a twisted convolution: $\eta(b, \nu) = \alpha \hat{\otimes} \phi$. This leads to the following property of operators associated with TST spreading functions.

**Lemma 36** An operator $H \in (\mathcal{B}, \mathcal{H}, \mathcal{B}')$ possesses a TST spreading function $\eta \in (S_0, L^2, S'_0)$ as in (43) if and only if it is of the form
\begin{equation}
H_{\eta} = \sum_{k, \ell} \alpha_{k\ell} \pi(kb_1, \ell \nu_1) H_{\phi}, \tag{44}
\end{equation}
where $H_{\phi}$ is the linear operator with spreading function $\phi$.

The proof consists of a straight-forward computation which may be spared by noting that we have, by (15):
\begin{equation*}
H_{\eta} = H_{\alpha \hat{\otimes} \phi} = H_{\alpha} \cdot H_{\phi} = \sum_{k, \ell} \alpha_{k\ell} \pi(kb_1, \ell \nu_1) H_{\phi}.
\end{equation*}

As before, we are particularly interested in the situation of the synthesis windows being given by time-frequency shifted versions of a single window: $h^j = \pi(\mu_j) h$. In a next step we note, that under the condition $\pi(\lambda) \pi(\mu_j) = \pi(\mu_j) \pi(\lambda)$, i.e., $\mu_j \in \Lambda^0$, the MGM with separable multiplier results in a TST spreading function with a Gabor multiplier as basic operator $H_{\phi}$.

**Lemma 37** Assume that a MGM $\mathbb{M}$ with multiplier $m_j(\lambda) = m_1(\lambda) m_2(j)$ is given. If the synthesis windows $h^j$ are given by $h^j = \pi(\mu_j) h$, with $\mu_j \in \Lambda^0$, then
\begin{equation*}
\mathbb{M} = \sum_j m_2(j) \pi(\mu_j) \mathbb{M}^G_{m_1; g, h},
\end{equation*}
i.e., here, the operator $H_{\phi}$ is given by a regular Gabor multiplier with mask $m_1$ and synthesis window $h$. 

*BIRKHAUSER*
Remark 38 Comparing the expression in the previous lemma to the expression $M = \sum_\lambda m_1(\lambda)\rho(\lambda)P_m$ for the same operator, we note that in this situation, the operator may either be interpreted as a (weighted) sum of Gabor multipliers or as a Gabor multiplier with a generalized projection operator $P_m$ in the synthesis process.

In this situation, we may ask, whether the family of generalized projection operators, \(\{\rho(\lambda)P_m\}_{\lambda \in \Lambda}\) form a frame or Riesz basis for their linear span. In fact, if \(m\) is in \(\ell^1\) and \(g, h \in S_0\), this question is easily answered by generalizing the result proved in [6, Theorem 3.2]. Here, \(\{\rho(\lambda)P_m\}_{\lambda \in \Lambda}\) is either a Riesz basis or not a frame for its closed linear span. Furthermore, there exists \(r > 0\) such that \(\{\rho(\alpha \lambda)P_m\}_{\lambda \in \Lambda}\) is a Riesz basis for its closed linear span whenever \(\alpha > r\).

In generalizing the result of Lemma 37, it is a natural further step to assume that the basic function \(\phi\) entering in the composition of \(\eta\) is the spreading function of a Gabor multiplier (at least in an approximate sense). According to the discussion of Sect. 2.4, this essentially means that \(\phi\) is sufficiently well concentrated in the time-frequency domain.

(In the sequel we will write \(\pi_{mn}\) for \(\pi(mb_0, n\nu_0)\) whenever the applicable lattice constants are sufficiently clear from the context.)

Hence, we assume that a Gabor multiplier \(H_\phi\), as defined in (20) is given. We may formally compute

\[
Hf = \sum_{k, \ell} \alpha_{k\ell} \pi_{k\ell} \sum_{m, n} m(m, n) V_g f(mb_0, n\nu_0) \pi(mb_0, n\nu_0)h
\]

\[
= \sum_{m, n} m(m, n) V_g f(mb_0, n\nu_0) \sum_{k, \ell} \alpha_{k\ell} \pi(kb_1, \ell\nu_1) \pi(mb_0, n\nu_0)h.
\]

(45)

Based on this expression, one may pursue two different choices of the sampling-points \((kb_1, \ell\nu_1)\). First, in extension of the result given in Lemma 37, we assume that the sampling points are associated to the adjoint lattice \(\Lambda^o = \frac{1}{\nu_0}Z \times \frac{1}{b_0}Z\) of \(\Lambda = b_0Z \times \nu_0Z\). The second choice of sampling points on the original lattice leads to a construction as introduced in [11] as Gabor twisters and will not be further discussed in the present contribution.

The following theorem extends the result given in Lemma 37 to the case in which the sampling points in the TST expansion are chosen from a lattice containing the adjoint lattice. It turns out that the TST spreading function then leads to a representation as a sum of Gabor multipliers.

Theorem 39 Let \(b_0, \nu_0 \in \mathbb{R}^+\) generate the time-frequency lattice \(\Lambda\), and let \(\Lambda^o\) denote the adjoint lattice. Let \(g, h \in S_0(\mathbb{R})\) denote respectively Gabor analysis and synthesis windows, such that the \(U\) condition (28) is fulfilled. Let \(H\) denote the operator in \((B, H, B')\) defined by the twisted spline type spreading function \(\eta\) as in (43), with \(b_1, \nu_1 \in \mathbb{R}^+\).

1. Assume that \(b_1\) and \(\nu_1\) are multiple of the dual lattice constants. Then \(H\) is a Gabor multiplier, with analysis window \(g\), synthesis window

\[
\gamma = \sum_{k, \ell} \alpha_{k\ell} \pi(kb_1, \ell\nu_1)h,
\]

(46)
and transfer function

\[ m(m, n) = b_0 \nu_0 \int_{\Box^\circ} M(b, \nu) e^{-2i\pi (n\nu_0 b - mb_0 \nu)} \, dbd\nu, \quad (47) \]

with \( \Box^\circ \) the fundamental domain of the adjoint lattice \( \Lambda^\circ \), and

\[ M(b, \nu) = \sum_{k, \ell = -\infty}^{\infty} \left| V_g h(b + k/\nu_0, \nu + \ell/b_0) \right|^2. \quad (48) \]

2. Assume that the lattice generated by \( b_1 \) and \( \nu_1 \) contains the adjoint lattice:

\[ b_1 = \frac{1}{p \nu_0}, \quad \nu_1 = \frac{1}{q b_0}. \quad (49) \]

Then \( H \) may be written as a finite sum of Gabor multipliers

\[ Hf = \sum_{i=1}^{p} \sum_{j=1}^{q} \left( \sum_{m \equiv i \pmod{p}} \sum_{n \equiv j \pmod{q}} m(m, n) V_g f(m b_0, n \nu_0) \pi_{mn} \right) \gamma_{ij}, \quad (50) \]

with at most \( p \cdot q \) different synthesis windows \( \gamma_{ij} \) and the transfer function given in (47) and (48).

**Proof** Let us formally compute

\[
Hf = \sum_{m,n} m(m, n) V_g f(m b_0, n \nu_0) \sum_{k, \ell} \alpha_{k\ell} \pi(kb_1, \ell \nu_1) \pi_{mn} h
= \sum_{m,n} m(m, n) V_g f(m b_0, n \nu_0) \pi_{mn} \gamma_{mn},
\]

where

\[
\gamma_{mn} = \sum_{k, \ell} \alpha_{k\ell} e^{2i\pi [kn b_0 \nu_1 - \ell m \nu_0 b_1]} \pi(kb_1, \ell \nu_1) h.
\]

Now observe that if \( (b_1, \nu_1) \in \Lambda^\circ \), one obviously has

\[
\gamma_{mn} = \sum_{k, \ell} \alpha_{k\ell} \pi(kb_1, \ell \nu_1) h = \gamma_{00}, \quad \text{for } (m, n) \in \mathbb{Z}^2,
\]

i.e. the above expression for \( Hf \) involves a single synthesis window \( \gamma = \gamma_{00} \). Therefore, in this case, \( H \) takes the form of a standard Gabor multiplier, with fixed time-frequency transfer function, and a synthesis window prescribed by the coefficients in the TST expansion. This proves the first part of the theorem.

Let us now assume that the TST expansion of the spreading function is finer than the one prescribed by the lattice \( \Lambda^\circ \), but nevertheless the lattice \( \Lambda_1 = \mathbb{Z}b_1 \times \mathbb{Z} \nu_1 \) contains \( \Lambda^\circ \). In other words, there exist positive integers \( p, q \) such that (49) holds.
We then have

$$
\gamma_{mn} = \sum_{k, \ell} \alpha_{k\ell} e^{2\pi i \left(\frac{[np - \ell q]}{pq}\right) \pi (kb_1, \ell v_1)} h
$$

(51)

and it is readily seen that there are at most \(pq\) different synthesis windows \(\gamma_{ij}\),

$$
\gamma_{ij} = \gamma_{m \mod p, n \mod q}, \quad i = 1, \ldots, p; \ j = 1, \ldots, q.
$$

(52)

The operator \(H\) may hence be written as a sum of Gabor multipliers, with one prescribed time-frequency transfer function, which is sub-sampled on several sub-lattices of the lattice \(\Lambda\):

$$
\Lambda_{ij} = (pb_0 \cdot \mathbb{Z} + i \cdot b_0) \times (qv_0 \cdot \mathbb{Z} + j \cdot v_0), \quad i = 0, \ldots, p - 1; \ j = 0, \ldots, q - 1,
$$

and a single synthesis window per sub-lattice as given in (52). The resulting expression for \(H\) is hence as given in (50).

The expression for the transfer function is derived in analogy to the case discussed in Sect. 2.4.

□

Remark 40 Let us observe that in this approximation, the time-frequency transfer function \(m\) is completely characterized by the function \(\phi\) used in the TST expansion. The choice of \(\phi\) therefore imposes a fixed mask for the multipliers that come into play in (50).

Example 41 We first assume, that for a given primal lattice \(\Lambda = b_0 \mathbb{Z} \times v_0 \mathbb{Z}\), the representation of a spreading function \(\eta\) is given by 5 building blocks:

$$
\eta(b, v) = \sum_{k=-1}^{1} \alpha_{k0} \phi\left(b - \frac{k}{v_0}, v\right) + \sum_{\ell=-1}^{1} \alpha_{0\ell} \phi\left(b, v - \frac{\ell}{b_0}\right).
$$

In this case, we obtain a single Gabor multiplier with synthesis window

$$
\gamma_{00} = \sum_{k=-1}^{1} \alpha_{k0} \pi\left(\frac{k}{v_0}, 0\right) h + \sum_{\ell=-1}^{1} \alpha_{0\ell} \pi\left(0, \frac{\ell}{b_0}\right) h.
$$

If we add the windows \(\phi\left(b \pm \frac{1}{2v_0}, v \pm \frac{1}{2b_0}\right)\) to the representation of \(\eta\), we are now dealing with the finer lattice \(\Lambda = \frac{1}{2v_0} \mathbb{Z} \times \frac{1}{2b_0} \mathbb{Z}\) and we obtain the sum of 4 Gabor multipliers with the following synthesis windows:

$$
\gamma_{00} = \sum_{k=-1}^{1} \alpha_{k0} \pi\left(\frac{k}{2v_0}, 0\right) h + \sum_{\ell=-1}^{1} \alpha_{0\ell} \pi\left(0, \frac{\ell}{2b_0}\right) h,
$$

$$
\gamma_{01} = \sum_{k=-1}^{1} \alpha_{k0} e^{\pi ik} \pi\left(\frac{k}{2v_0}, 0\right) h + \sum_{\ell=-1}^{1} \alpha_{0\ell} \pi\left(0, \frac{\ell}{2b_0}\right) h,
$$

$$
\gamma_{10} = \sum_{k=-1}^{1} \alpha_{k0} \pi\left(\frac{k}{2v_0}, 0\right) h + \sum_{\ell=-1}^{1} \alpha_{0\ell} \pi\left(0, \frac{\ell}{2b_0}\right) h.
$$
\[ \gamma_{10} = \sum_{k=-1}^{1} \alpha_{k0} \pi \left( \frac{k}{2\nu_0}, 0 \right) h + \sum_{\ell=-1}^{1} e^{\pi i \ell} \alpha_{0\ell} \pi \left( 0, \frac{\ell}{2b_0} \right) h, \]

\[ \gamma_{11} = \sum_{k=-1}^{1} \alpha_{k0} e^{\pi i k} \pi \left( \frac{k}{2\nu_0}, 0 \right) h + \sum_{\ell=-1}^{1} \alpha_{0\ell} e^{-\pi i \ell} \pi \left( 0, \frac{\ell}{2b_0} \right) h, \]

and corresponding lattices: \( \Lambda_{00} = 2\mathbb{Z}b_0 \times 2\mathbb{Z}v_0 \), \( \Lambda_{01} = 2\mathbb{Z}b_0 \times (2\mathbb{Z} + 1)v_0 \), \( \Lambda_{10} = (2\mathbb{Z} + 1)b_0 \times 2\mathbb{Z}v_0 \), and \( \Lambda_{11} = (2\mathbb{Z} + 1)b_0 \times (2\mathbb{Z} + 1)v_0 \).

It is important to note, that in both cases described in Theorem 39 as well as the above example, the transfer function \( m \) can be calculated as the best approximation by a regular Gabor multiplier—a procedure which may be efficiently realized using (29). Fast algorithms for this task also exist in the literature, see [16], however, the algorithm derived from (29) is faster, since only one two-dimensional Fourier-transform, whose size depends on the fundamental domain of the adjoint lattice, is necessary. Moreover, our method may easily be adapted to the generalized case according to (39).

4 Conclusions and Perspectives

Starting from an operator representation in the continuous time-frequency domain via a twisted convolution, we have introduced generalizations of conventional time-frequency multipliers in order to overcome the restrictions of this model in the approximation of general operators. The model of multiple Gabor multipliers in principle allows the representation of any given linear operator. However, in order to achieve computational efficiency as well as insight in the operator’s characteristics, the parameters used in the model must be carefully chosen. An algorithm choosing the optimal sampling points for the family of synthesis windows, based on the spreading function, is the topic of ongoing research. On the other hand, the model of twisted spline type functions allows the approximation of a given spreading function and results in an adapted window or family of windows. By refining the sampling lattice in the TST approximation, a rather wide class of operators should be well-represented. The practicality of this approach has to be shown in the context of operators of practical relevance. All the results given in this work will also be applied in the context of estimation rather than approximation.

As a further step of generalization, frame types other than Gabor frames may be considered. Surprisingly little is known about wavelet frame multipliers, hence it will be interesting to generalize the achieved results to the affine group.

Acknowledgements We thank two anonymous reviewers for their helpful comments and suggestions.

References

Representation of Operators by Sampling in the Time-Frequency Domain

Monika Dörfler
Institut für Mathematik, Universität Wien, Alserbachstrasse 23
A-1090 Wien, Austria
Monika.Doerfler@univie.ac.at

Bruno Torrésani
Université de Provence, LATP, UMR CNRS 6632,
CMI, 39 rue Joliot-Curie, 13453 Marseille cedex 13, France.
Bruno.Torresani@cmi.univ-mrs.fr

Abstract
Gabor multipliers are well-suited for the approximation of certain time-variant systems. However, this class of systems is rather restricted. To overcome this restriction, multiple Gabor multipliers allowing for more than one synthesis windows are introduced. The influence of the choice of the various parameters involved on approximation quality is studied for both classical and multiple Gabor multipliers. Fairly simple error estimates are provided, and the study is supplemented by numerical simulations. This paper is an extended and improved version of [6].

Key words and phrases : Operator approximation, time-frequency, Gabor multiplier, error estimates

2000 AMS Mathematics Subject Classification — 47B38, 47G30, 94A12, 65F20.

1 Introduction

In signal processing, in particular speech and audio processing, the manipulation of given signals in the time-frequency domain is common practice, consider [1, 19, 20] for some recent work. However, the operators that arise from these manipulations, so called time-frequency multipliers, have rather rarely been studied with a focus on the influence of the various involved parameters on the outcome of methods based on time-frequency analysis of signals.

In a recent paper [7], the authors describe the representation of operators in the time-frequency domain by means of a twisted convolution with the operator’s spreading function. Although this description is not suitable for direct discretization, the spreading representation provides a better understanding of certain operators’ behavior: it reflects the operator’s action in the time-frequency
domain. This motivates an approach that uses the spreading representation of time-frequency multipliers, in order to optimize the parameters involved. More specifically, in the one-dimensional, continuous-time case, given a linear operator $T$ on $L^2(\mathbb{R})$ with integral kernel $\kappa_T$ and spreading function $\eta_T$:

$$\eta_T(b, \nu) = \int_{-\infty}^{\infty} \kappa_T(t, t - b) e^{-2\pi i \nu t} dt,$$

we aim at modeling the operator by its action on the sampled short-time Fourier transform (STFT) or Gabor coefficients. The STFT of $f \in L^2(\mathbb{R})$ is defined by

$$V^g_f(b, \nu) = \langle f, \pi(b, \nu) g \rangle, \quad (b, \nu) \in \mathbb{R}^2 \quad (1)$$

where $\pi(b, \nu) g(x) = M_{\nu} T_b g = g(x - b) e^{2\pi i \nu x}$ denotes the time-frequency shifted versions of a window $g \in L^2(\mathbb{R})$. Sampling the STFT on a lattice $\Lambda$ then leads to the Gabor transform, which, for the sake of clarity, is denoted by $C^g,\Lambda f$:

$$C^g,\Lambda f(\lambda) = \langle f, \pi(\lambda) g \rangle, \quad \lambda \in \Lambda \subset \mathbb{R}^2. \quad (2)$$

For the special case of a product lattice of the form $\Lambda = b_0 \mathbb{Z} \times \nu_0 \mathbb{Z}$, we obtain $C^g,\Lambda f(k, l) = \langle f, \pi(kb_0, l\nu_0) g \rangle = \langle f, M_{\nu_0} T_{kb_0} g \rangle$.\footnote{The finite-dimensional case $\mathcal{H} = \mathbb{C}^L$ is obtained similarly, replacing integrals with finite sums, and letting $k = 0, \ldots N_k - 1$, $l = 0, \ldots N_\nu - 1$, where $N_k = L/b_0$, $N_\nu = L/\nu_0$ and $b_0, \nu_0$ divide $L$.} A good overview of Gabor analysis can be found in [13].

Purely multiplicative modification of the Gabor coefficients $C^g,\Lambda f(\lambda)$ leads to the definition of classical Gabor multipliers [12]. In this case, the linear operator applied to the coefficients is diagonal. Gabor multipliers provide accurate approximation of so-called underspread operators [18]. We will consider generalizations of the classical Gabor multipliers: the restriction to diagonality is relaxed in order to achieve better approximation for a wide class of operators at low cost. Moreover, in certain approximation tasks it is efficient, e.g. in the sense of sparsity, to use several side diagonals, but a lower redundancy (coarser sampling lattice $\Lambda$) in the Gabor system used. Further, the drawback resulting from coarse sub-sampling, i.e. large $b_0$ and/or $\nu_0$, can, to a certain extent, be compensated by using two or three instead of just one synthesis window. This case will be called multiple Gabor multiplier.

The aim of this contribution is the description of error estimates for the approximation of operators by multiple and generalized Gabor multipliers. The given error estimates are based on the operator’s spreading function. The results give some insight in the choice of the parameters involved in approximation, in particular, of the windows and the lattice constants $b_0, \nu_0$.

This paper is organized as follows. Section 2 introduces Gabor multipliers and their generalizations. The basic idea of approximation in the spreading domain
is explained. In Section 3, a general error estimate for the approximation of Hilbert-Schmidt operators by Gabor multipliers is derived, and several special cases are deduced thereof. As a noteworthy special case, the approximation of short-time Fourier multipliers by Gabor multipliers is considered. From these descriptions, guidelines for the choice of good parameters will be discussed in Section 4. Finally, Section 5 gives various insightful numerical experiments.

2 Approximation in the time-frequency domain: the parameters

Throughout this paper, \( \mathcal{H} = L^2(A) \) where \( A \) is a locally compact abelian group. \( \mathcal{H} \) is therefore equipped with an action of the Heisenberg group of time-frequency shifts, and corresponding versions of short time Fourier and Gabor transform may be defined. The standard cases \( A = \mathbb{R} \) and \( A = \mathbb{Z} \) will be of special interest to us.

2.1 Gabor multipliers

Let \( g, h \in \mathcal{H} \), let \( \Lambda \) be a lattice in the time-frequency space (see [10] for details on lattices and Gabor frames), and let \( \mathcal{Y}_g \) be the short time Fourier transform associated with analysis window \( g \). To avoid confusions with the short time Fourier transform \( \mathcal{Y}_g \), we use the notation \( C_{g,\Lambda} \) for the Gabor transform with window \( g \) and lattice \( \Lambda \) (the so-called analysis operator), and denote by \( C_{g,\Lambda}^* \) the adjoint of \( C_{g,\Lambda} \) (the synthesis operator). A Gabor multiplier [12] is defined as

\[
G_m : f \in \mathcal{H} \mapsto G_m f = C_{h,\Lambda}^*(m \cdot C_{g,\Lambda} f).
\]

Here, \( m \cdot C_{g,\Lambda} f \) denotes the pointwise multiplication of \( C_{g,\Lambda} f \) with the symbol \( m \in \ell^\infty(\Lambda) \). Often \( g = h \), i.e., analysis and synthesis window may be identical. For a given lattice \( \Lambda \), we shall denote by \( \Lambda^\circ \) the adjoint lattice [10], by \( \Omega^\circ \) the corresponding fundamental domain, and by \( \Pi^\circ \) the corresponding periodization operator. In the infinite-dimensional setting \( \mathcal{H} = L^2(\mathbb{R}) \), and for a product lattice of the form \( \Lambda = b_0 \mathbb{Z} \times \nu_0 \mathbb{Z} \), we have \( \Lambda^\circ = t_0 \mathbb{Z} \times \xi_0 \mathbb{Z} \) with \( t_0 = 1/\nu_0 \), \( \xi_0 = 1/b_0 \), and \( \Pi^\circ f(\zeta) = \sum_{\lambda^\circ \in \Lambda^\circ} f(\zeta + \lambda^\circ), \zeta \in \Omega^\circ \).

In a finite-dimensional setting \( \mathcal{H} = \mathbb{C}^L \), with \( \Lambda = \mathbb{Z}_{N_b} \times \mathbb{Z}_{N_\nu} \), with \( N_b, N_\nu \) two divisors of \( L \), we have \( \Lambda^\circ = \mathbb{Z}_{L/N_\nu} \times \mathbb{Z}_{L/N_b} \), and the obvious form for the periodization operator.

In the definition of the multipliers, several parameters have to be fixed: the analysis and synthesis windows \( g \) and \( h \), the lattice \( \Lambda \), and the symbol \( m \). For practical as well as theoretical reasons, the windows should be well-localized in time and frequency. As for the lattice, it is expected that denser lattices will lead to better results in approximation, but higher computational cost. We will
see, that the eccentricity of the lattice plays an important role in approximation quality and that lattices that exceed a certain density are not suitable.

Finally, the symbol \( m \) can be optimized to best approximate a given operator. This problem was studied in [7] in the Hilbert-Schmidt setting. An explicit expression for the best Gabor multiplier approximation of an Hilbert-Schmidt operator \( T \) (in the sense that the Hilbert-Schmidt norm \( \| T - G_m \|_{HS} \) is minimized) was obtained in the spreading domain, see Theorem 1. The spreading function of a Gabor multiplier \( G_m \) takes the form \( \eta_{G_m} (\zeta) = \tilde{m} (\zeta) \cdot \mathcal{Y}_g h (\zeta), \) where \( \tilde{m} \) is the symplectic Fourier transform of the sequence \( m: \)

\[
\tilde{m}(\zeta) = \mathcal{F} \cdot m(\zeta) = \sum_{\lambda \in \Lambda} m(\lambda)e^{2\pi i \langle \lambda, \zeta \rangle}, \quad \zeta \in A \times \hat{A}.
\]

Here, \( [\cdot, \cdot] \) denotes the usual symplectic form (e.g., for the rectangular lattice \( \Lambda = b_0 \mathbb{Z} \times \nu_0 \mathbb{Z} \), we have \([kb_0, t\nu_0], (t, \xi) = t\nu_0 t - kb_0 \xi\)).

Note, that \( \tilde{m} \) is a \( \Lambda^o \)-periodic function and this periodicity has a decisive influence on the quality of operator representation by Gabor multipliers. Loosely speaking, since \( \eta_{G_m} (\zeta) = \tilde{m} (\zeta) \cdot \mathcal{Y}_g h (\zeta), \) a given spreading function can only be accurately reproduced in one fundamental domain of \( \Lambda^o \). Even for underspread operators whose spreading function’s support is contained in \( \Omega^o \), the periodicity of \( \tilde{m} \) leads to aliasing effects, since \( \mathcal{Y}_g h \) is never compactly supported (see [15] for a discussion of STFT-related uncertainty inequalities). This phenomenon is illustrated in Figure 1, where an example of an operator with symmetric spreading function is shown. It is obvious that the amount of aliasing depends on the window decay, so, as mentioned before, localized windows must be chosen.

Given the spreading function of the operator to be approximated, the above observations give immediate insight in how the parameters involved in the approximation of operators by Gabor multipliers have to be chosen. Generally speaking, good approximation by a classical Gabor multiplier is possible, if the essential support of the spreading function is contained in the fundamental domain \( \Omega^o \) of the adjoint lattice for a dense enough lattice \( \Lambda \). In this case, to reduce aliasing as much as possible, the analysis and synthesis windows must be chosen such that \( \mathcal{Y}_g h \) is small outside \( \Omega^o \) and positive on the support of the spreading function, also see Section 4.1.

### 2.2 Generalizing Gabor multipliers

In order to extend the good approximation quality of Gabor multipliers to a more general class of operators, we may allow for more sophisticated action on the time-frequency coefficients. In particular, instead of using just diagonal matrices, we may introduce several side-diagonals. This idea leads to the following definition, by considering the special case \( h^{(j)} = \pi(\mu_j)h \) in the definition below. Generally speaking, multiple Gabor multipliers are sums of Gabor multipliers with different synthesis windows and symbols.
**Definition 1** (Multiple and generalized Gabor Multiplier). Let \( g, h^{(j)} \in \mathcal{H}, \) for \( j = 1, \ldots, J \) denote an analysis and a set of synthesis window functions. Let \( \Lambda \) be a time-frequency lattice.

1. For \( \mathbf{m} = \{m_j \in \ell^\infty(\Lambda), j \in J\} \), a family of bounded functions on \( \Lambda \), the associated multiple Gabor multiplier \( G_\mathbf{m} \) is defined as follows: for all \( f \in \mathcal{H} \)

\[
G_\mathbf{m} f = \sum_{\lambda \in \Lambda} \sum_{j \in J} m_j(\lambda) \langle f, \pi(\lambda)g \rangle \pi(\lambda)h^{(j)}.
\]

2. A generalized Gabor multiplier is a multiple Gabor multiplier whose synthesis window functions are time-frequency shifted copies \( h^{(j)} = \pi(\mu_j)h \), for a finite set of time-frequency shifts \( \{\mu_j, j \in J\} \) of a unique window function \( h \).

Note that the spreading function of multiple Gabor multipliers is given by
\[
\eta_{G_\mathbf{m}} = \sum_j \tilde{m}_j \mathcal{V}_g h^{(j)}.
\]
The operator approximation with multiple Gabor multipliers involves, in addition to the choice of the parameters mentioned above, the choice of the analysis windows or the sampling points $\mu_j$. Approximation by sums of Gabor multipliers in the operator norm was treated in [16], with an application to the modeling of channel matrices in OFDM. Since the latter are assumed to be invertible, Hilbert-Schmidt norm approximation does not apply.

3 Error analysis in $L^2(\mathbb{R})$

Multiple and generalized Gabor multipliers were introduced in [7]. For the sake of completeness, let us restate the corresponding approximation result, with the notations of the present paper. Given analysis and synthesis windows $g \in L^2(\mathbb{R})$ and $h^{(j)} \in L^2(\mathbb{R})$, $j = 1, \ldots J$, introduce the $\Lambda^\circ$ periodizations:

$$U_{jj'} = \Pi^\circ(\mathcal{V}_g h^{(j)} \mathcal{V}_g h^{(j')}) , \quad B_j = \Pi^\circ(\eta_H \mathcal{V}_g h^{(j)}) , \quad j, j' = 1, \ldots J . \quad (4)$$

Denote by $U$ the matrix-valued function with matrix elements $U_{jj'}$ and by $B$ the vector with components $B_j$.

**Theorem 1.** Let $g \in \mathcal{H}$ and $h^{(j)} \in \mathcal{H}$, $j = 1, \ldots J$ be such that for almost all $\zeta \in \Omega^\circ$, the matrix $U(\zeta)$ is invertible.

Let $T$ be an Hilbert-Schmidt operator on $\mathcal{H}$, with spreading function $\eta_T$. Then the vector $\mathbf{m} = (m_1, \ldots m_J)$ of symbols of the multiple Gabor multiplier that minimizes the approximation error $\|T - G_\mathbf{m}\|_{HS}$ is obtained from the solution of the matrix equation

$$\tilde{\mathbf{m}}(\zeta) = U(\zeta)^{-1} \cdot B(\zeta) , \quad \zeta \in \Omega^\circ , \quad (5)$$

where $\tilde{\mathbf{m}} = (\tilde{m}_1, \ldots \tilde{m}_J)$ is the vector of symplectic Fourier transforms of $\mathbf{m}$.

The invertibility condition for $U(\zeta)$ (which reduces to the classical one [2, 7] in the Gabor multiplier case), is equivalent to linear independence of the system of projection operators involved, see [7] for details. Obviously, for a very dense lattice $\Lambda$ (i.e., high redundancy), $\Lambda^\circ$ is very coarse and this property is usually not fulfilled.

The case of a unique synthesis window may be immediately obtained from the above formula. Note that formula (5) allows for an efficient implementation of the otherwise expensive calculation of the best approximation by multiple Gabor multipliers, compare [9] for an algorithm that applies to the classical Gabor multiplier situation.

We may now give an expression for the error in the approximation given above, in the case $\mathcal{H} = L^2(\mathbb{R})$. We define, for the Hilbert-Schmidt operator $T$,

$$\Gamma_T = \Pi^\circ(|\eta_T|^2) .$$
Proposition 1. Let $T$ be a Hilbert-Schmidt operator on $\mathcal{H} = L^2(\mathbb{R})$, let $\tilde{m}$ denote the vector-valued function obtained as in (5) and let $m$ be its inverse symplectic Fourier transform. Then the approximation error $E = \|T - Gm\|_{HS}^2$ is given by

$$E = \int_{\Omega^o} \Gamma_T(\zeta) \left( 1 - \frac{\sum_{i,j} (U^{-1})_{ij}(\zeta) B_i(\zeta) \overline{B_j(\zeta)}}{\Gamma_T(\zeta)} \right) d\zeta \quad (6)$$

Proof. For simplicity, we set $V_j = V_gh^j$. Since the mapping $T \mapsto \eta_T$ is unitary, see [11], we may start from

$$\|H - Gm\|_{HS}^2 = \|\eta_T - \sum_j \tilde{m}_j \mathcal{V}_g h^{(j)}\|^2$$

$$= \|\eta_T\|^2 - 2\text{Re} \left( \sum_j \langle \eta_T, \tilde{m}_j \mathcal{V}_j \rangle \right) + \text{Re} \left( \sum_{j,j'} \langle \tilde{m}_j \mathcal{V}_j, \tilde{m}_{j'} \mathcal{V}_{j'} \rangle \right),$$

and find by a straight-forward calculation that

$$\sum_j \langle \eta_T, \tilde{m}_j \mathcal{V}_j \rangle = \sum_{j,j} \int_{\Omega^o} (U^{-1})_{jj'}(\zeta) B_j(\zeta) \overline{B_j}(\zeta) d\zeta = \sum_j \int_{\Omega^o} \tilde{m}_j(\zeta) \overline{B_j}(\zeta) d\zeta$$

whereas

$$\sum_{j,j'} \langle \tilde{m}_j \mathcal{V}_{j}, \tilde{m}_{j'} \mathcal{V}_{j'} \rangle = \sum_{j,j'} \int_{\Omega^o} \tilde{m}_j(\zeta) \overline{m}_{j'}(\zeta) (U^{-1})_{jj'}(\zeta) d\zeta = \sum_j \int_{\Omega^o} B_j(\zeta) \overline{m}_j(\zeta) d\zeta.$$

Hence, we have

$$\|\eta_T - \sum_j \tilde{m}_j \mathcal{V}_j\|^2 = \|\eta_T\|^2 - \int_{\Omega^o} \sum_{j,j'} (U^{-1})_{jj'}(\zeta) B_j(\zeta) \overline{B_j}(\zeta) d\zeta$$

and since $\|\eta_T\|^2 = \int_{\Omega^o} \Pi^o(|\eta_T|^2)$ we obtain the error expression as stated. \(\blacksquare\)

Notice that Proposition 1 covers the Gabor multiplier case obtained in [7]. Notice also that (6) immediately yields

$$E \leq \|\eta_T\|^2 \left| 1 - \frac{\sum_{j,j'} (U^{-1})_{jj'} B_j \overline{B_j}}{\Gamma_T} \right|_{L^\infty(\Omega^o)}$$

The finite-dimensional situation is similar, the integral over $\Omega^o$ is replaced by a finite sum over the finite fundamental domain $\{0, \ldots t_0 - 1\} \times \{0, \ldots \xi_0 - 1\}$.
Corollary 1. Let an operator $T$ with $\text{supp}(\eta_T) \subseteq \Omega^0$ be given. Then, the error of the best approximation by a generalized Gabor multiplier $G_m$ as defined in (5) is bounded by
\[
\|T - G_m\|_{HS}^2 \leq \|\eta_T\|^2 \text{ess sup} \left[ |1 - \langle U(\zeta)^{-1} \cdot V(\zeta), V(\zeta) \rangle_{C^J}| \right]
\] (7)
where the vector-valued function $V(\zeta)$ is given by $V_j(\zeta) = \gamma_j h(\zeta)$ and $\eta_l = \eta_T \cdot \chi_{\Omega^0} + \mu_l$ is the restriction of $\eta_T$ to $\Omega^0 + \mu_l$. 

Proof. Since $\text{supp}(\eta_T) \subseteq \Omega^0$, we have $B(\zeta) = \eta_T(\zeta) \cdot V(\zeta)$ and the result follows from Proposition 1. □

The last result may be interpreted as follows. For all $\zeta \in \Omega^0_T$, the matrix $U(\zeta)^{-1}$ is the Gramian matrix of the vectors $V_j(\zeta)$ defined by their entries $V_j(\lambda^0) = \gamma_j h(\zeta + \lambda^0)$ for $\lambda^0 \in \Lambda^0$. The inverse of the Gramian provides the biorthogonal system to this family of vectors, $\tilde{V}_j^\natural = [U(\zeta)^{-1} \cdot V^\natural]_j$, for every $\zeta \in \Omega^0_T$. Since, now, the support of $\eta_T$ is restricted to $\Omega^0_T$, in the present case, the vector $\eta_T^\natural$, given by $\eta_T^\natural(\lambda^0) = \eta_T(\zeta + \lambda^0)$, $\lambda^0 \in \Lambda^0$, to be approximated by the family of vectors $V_j(\zeta), j = 1, \ldots, J$, has only one non-zero entry, at $\lambda^0 = 0$. By realization of the matrix multiplication corresponding to the projection of $\eta_T^\natural$ onto the span of the $V_j^\natural$, we find that requiring equality of $\eta_T^\natural$ to its projection leads to the necessary condition that $\sum_j \tilde{V}_j^\natural(0) \cdot V_j^\natural(0) = 1$, which is equivalent to saying that $\langle U(\zeta)^{-1} \cdot V(\zeta), V(\zeta) \rangle_{C^J}$ is equal to 1, for all $\zeta \in \Omega^0_T$. Consequently, we have to find a family of windows that exhaust $\Omega^0_T$ in the sense of maximal concentration inside $\Omega^0_T$. Note that the concentration may (and will) be achieved by using functions with support outside $\Omega^0_T$ in order to cancel the aliases caused by the unavoidable energy of each $V_j$ outside this area. We will come back to this situation in an example in Section 5.2.

The next corollary generalizes the result from the previous corollary. We assume that the spreading function is supported in a finite union of translates of the fundamental domain $\Omega^0$. In other words, we assume that $\eta(\zeta) = \sum_{l=1}^L \eta_l(\zeta)$ with $\text{supp}(\eta_l) \subseteq (\Omega^0 + \mu_l)$ for $l = 1, \ldots, L$.

Corollary 2. Let an operator $T$ with $\text{supp}(\eta_T) \subseteq \bigcup_{l=1}^L (\Omega^0 + \mu_l)$ be given. Then, the error $E = \|T - G_m\|_{HS}^2$ of the best approximation by a generalized Gabor multiplier $G_m$ as defined in (5) is given by
\[
E = \int_{\Omega^0} \Gamma_T(\zeta) \left[ 1 - \sum_{l=1}^L \frac{|\eta_l(\zeta + \mu_l)|^2 \langle U(\zeta)^{-1} \cdot V(\zeta + \mu_l), V(\zeta + \mu_l) \rangle}{\sum_{l=1}^L |\eta_l(\zeta + \mu_l)|^2} \right] d\zeta ,
\] (8)
where the vector-valued function $V$ is given by $V_j = \gamma_j h^{(j)}$ and $\eta_l = \eta_T \cdot \chi_{\Omega^0 + \mu_l}$ is the restriction of $\eta_T$ to $\Omega^0 + \mu_l$. 


Proof. According to the assumption on $\eta_T$, we may write $\eta_T = \sum_{l=1}^{L} \eta_l$, with $\text{supp}(\eta_l) \subseteq (\Omega^o + \mu_l)$. Therefore, we obtain

$$B_i(\zeta)B_j(\zeta) = \sum_{k=1}^{L} (\eta_T g^{(i)}(\zeta + \mu_k)) \cdot \sum_{l=1}^{L} (\eta_T g^{(j)}(\zeta + \mu_l))$$

$$= \sum_{l=1}^{L} |\eta_l(\zeta + \mu_l)|^2 (g^{(i)} \cdot g^{(j)})(\zeta + \mu_l),$$

since the $\eta_l$ have, by definition, disjoint support. Hence, the result follows directly from (6).

Note that the last corollary shows that for sufficiently disjoint compactly supported regions $\Omega^o + \mu_l$, $l = 1, \ldots, J$, of a given spreading function, the situation is comparable to the situation observed in Corollary 1. One synthesis window or, in an improved situation, a finite set of synthesis windows, should be concentrated as much as possible in each of the $\Omega^o + \mu_l$ and close to zero outside.

On the other hand, in the case of adjacent $\Omega^o + \mu_l$, a similar strategy can be envisaged, with the virtue of mutual cancellation of aliases arising in each of the $\Omega^o + \mu_l$.

The next corollary deals with the approximation error occurring in the case of approximating an STFT multiplier by a Gabor multiplier. This is a situation of particular relevance, since in practice, the multiplier symbol $m$, originally defined on $\mathbb{R}^2$, is simply sub-sampled. Therefore, we also compare the result of this procedure to the corresponding best approximation. Let us recall that a short time Fourier multiplier is defined in a similar way as a Gabor multiplier, using pointwise multiplication in the time-frequency domain. In order to avoid confusion, we shall denote by $S_a$ a short time Fourier multiplier, defined by

$$S_a : f \in \mathcal{H} \rightarrow S_a f = \mathcal{F}_h(a \cdot \mathcal{F}_g f),$$

where $\mathcal{F}_g$ and $\mathcal{F}_h$ denote short time Fourier transforms with respect to windows $g$ and $h$ respectively, and the symbol, denoted by $a$, is now defined as a function on the whole time-frequency space (instead of a sub-lattice $\Lambda$ of it).

We recall [7], that the spreading function of an STFT-multiplier is given by $\eta_{S_a} = \tilde{a} \cdot \mathcal{F}_g h$, $\tilde{a}$ being the continuous symplectic Fourier transform of $a$.

Corollary 3. Let $T = S_a$ be a STFT multiplier with spreading function $\eta_T = \tilde{a} \cdot \mathcal{F}_g h$, and denote by $T' = G_m$ its best Gabor multiplier approximation with the same windows, and lattice $\Lambda$, as defined in (5).

1. The approximation error is given by

$$\|T - T'\|_{HS}^2 = \int_{\Omega^o} \left[ \Pi^o((\tilde{a} \cdot g^{(i)})^2)(\zeta) - \frac{||\Pi^o(\tilde{a} \cdot g^{(i)}(\zeta))||^2}{\Pi^o(|g^{(i)}(\zeta)|^2)} \right] d\zeta \quad (9)$$
2. Furthermore, the difference between the best approximation $T'$ and the Gabor multiplier $T'' = G_{a|\Lambda}$ obtained from the $\Lambda$-subsampled version $a|\Lambda$ of the symbol $a$ is given by

$$
\|T'' - T'\|_{\mathcal{HS}} = \int_{\Omega^o} \left| \Pi^o(\tilde{a})(\zeta) - \frac{\Pi^o(\tilde{a} \cdot |\mathcal{G}\tilde{a}|^2)(\zeta)}{\Pi^o(|\mathcal{G}\tilde{a}|^2)(\zeta)} \right|^2 \cdot |\mathcal{G}\tilde{a}|^2 d\zeta \cdot d\zeta .
$$

(10)

**Proof.** To prove (9), we first note that, since the spreading function of $T' = G_m$ is given by $\tilde{m} \cdot \mathcal{G} h$, with $\tilde{m}$ given in (5), we have

$$
\|T - T'\|_{\mathcal{HS}}^2 = \int_{\mathbb{R}^2} \left[ \frac{\Pi^o(\tilde{a} \cdot |\mathcal{G}\tilde{a}|^2)(\zeta)}{\Pi^o(|\mathcal{G}\tilde{a}|^2)(\zeta)} \right]^2 \cdot |\mathcal{G}\tilde{a}|^2 d\zeta.
$$

We then immediately obtain (9) since, by the usual periodization argument,

$$
\int_{\mathbb{R}^2} \frac{\Pi^o(\tilde{a} \cdot |\mathcal{G}\tilde{a}|^2)(\zeta)}{\Pi^o(|\mathcal{G}\tilde{a}|^2)(\zeta)} \cdot |\mathcal{G}\tilde{a}|^2 d\zeta = \int_{\Omega^o} \frac{\Pi^o(\tilde{a} \cdot |\mathcal{G}\tilde{a}|^2)(\zeta)}{\Pi^o(|\mathcal{G}\tilde{a}|^2)(\zeta)} d\zeta.
$$

To see (10), note that the symplectic Fourier transform of $a|\Lambda$ is just the periodization of $\tilde{a}$ on $\Lambda^o$. Hence, we have

$$
\|T'' - T'\|_{\mathcal{HS}} = \int_{\mathbb{R}^2} \left| \Pi^o(\tilde{a})(\zeta) - \frac{\Pi^o(\tilde{a} \cdot |\mathcal{G}\tilde{a}|^2)(\zeta)}{\Pi^o(|\mathcal{G}\tilde{a}|^2)(\zeta)} \right|^2 \cdot |\mathcal{G}\tilde{a}|^2 d\zeta.
$$

which proves (10).

It becomes obvious that the error between the result obtained from subsampling, is insignificant as long as the symbol’s (symplectic) Fourier transform $\tilde{a}$ is concentrated near the origin. If higher frequencies are present in $a$, simple subsampling of the multiplier may lead to undesirable aliasing effects, as shown in the following simple example.

**Example 1.** Consider the case $\mathcal{H} = L^2(\mathbb{R})$ with a separable lattice $\Lambda = b_0\mathbb{Z} \times v_0\mathbb{Z}$, and let $Sa$ be an STFT multiplier with a symbol consisting of a
low-frequency component and a higher-frequency perturbation component. For example, let $a$ be given by its symplectic Fourier transform $\tilde{a}$ with $\tilde{a}(\zeta) = 1$ for $\zeta \in [-\varepsilon, \varepsilon] \times [-\varepsilon, \varepsilon]$ and $\tilde{a}(\zeta) = -1$ for $\zeta \in \left[\frac{k_1}{\nu_0} - \varepsilon, \frac{k_1}{\nu_0} + \varepsilon\right] \times \left[\frac{k_2}{\nu_0} - \varepsilon, \frac{k_2}{\nu_0} + \varepsilon\right]$ for some integer $k_1$, $\varepsilon$ close to zero, and $\tilde{a}(\zeta) = 0$ otherwise. Then, if $\Lambda = b_0 \mathbb{Z} \times \nu_0 \mathbb{Z}$, obviously $\Pi^\circ(\tilde{a})(\zeta) \equiv 0$, such that subsampling the original multiplier $a$ results in the zero operator. On the other hand, assuming an exponential decay for $|\gamma_g h|$, e.g. $|\gamma_g h(\zeta)|^2 \leq e^{-|\zeta|^4}$, we have, with $\zeta_1 = \left(\frac{k_1}{\nu_0}, \frac{k_1}{\nu_0}\right)$: $\Pi^\circ(\tilde{a} \cdot |\gamma_g h|^2)(\zeta) \approx 1 - e^{-|\zeta_1|^4}$, for $\zeta \in \left[\frac{k}{\nu_0} - \varepsilon, \frac{k}{\nu_0} + \varepsilon\right] \times \left[\frac{k}{\nu_0} - \varepsilon, \frac{k}{\nu_0} + \varepsilon\right]$, $k \in \mathbb{Z}$, and $\Pi^\circ(\tilde{a} \cdot |\gamma_g h|^2)(\zeta) = 0$ else. So, in this case, the aliases that would be generated by subsampling are efficiently suppressed in the approximation.

The above example shows that knowing some properties of a system can help to choose appropriate parameters in the approximation of the corresponding operators. For example, if an STFT multiplier is generated by a strictly low-pass symbol, sub-sampling the symbol is a good choice. As soon as the symbol has higher-frequency components, this may result in quite different operators. For example, if an STFT multiplier is generated by a strictly low-pass symbol, sub-sampling the symbol is a good choice. As soon as the symbol has higher-frequency components, this may result in quite different operators. For example, if an STFT multiplier is generated by a strictly low-pass symbol, sub-sampling the symbol is a good choice. As soon as the symbol has higher-frequency components, this may result in quite different operators. For example, if an STFT multiplier is generated by a strictly low-pass symbol, sub-sampling the symbol is a good choice. As soon as the symbol has higher-frequency components, this may result in quite different operators. For example, if an STFT multiplier is generated by a strictly low-pass symbol, sub-sampling the symbol is a good choice. As soon as the symbol has higher-frequency components, this may result in quite different operators.

### 4 Choosing the parameters

For simplicity, we specialize the following discussion to the infinite-dimensional case $\mathcal{H} = \mathbb{L}^2(\mathbb{R})$, and rectangular lattice $\Lambda = b_0 \mathbb{Z} \times \nu_0 \mathbb{Z}$. The finite-dimensional situation is handled similarly.

#### 4.1 Gabor Multipliers

If an operator $T$ with known spreading function is to be approximated by a Gabor multiplier $G_m$, the lattice may be adapted to the eccentricity of the spreading function according to the error expression obtained in Proposition 1, which may be considerably simplified for the case of only one synthesis window, see [7]. In order to choose the eccentricity of the lattice accordingly and adapt the window to the chosen lattice as to avoid aliasing, assume, that we may find $b_0, \nu_0$, with $b_0 \cdot \nu_0 < 1$, such that $\text{supp}(\eta_T) \subseteq (\Omega^\circ + z)$, where $\Omega^\circ = [0, \frac{1}{\nu_0}] \times [0, \frac{1}{b_0}]$. In this case, the error resulting from best approximation by a Gabor multiplier with respect to the lattice $\Lambda = b_0 \mathbb{Z} \times \nu_0 \mathbb{Z}$ is bounded by $K_e : \|\eta_T\|_2^2$, with

$$K_e = 1 - \inf_{t, \xi \in \Omega_T^\circ} \sum_{k,l} |\gamma_g h(t + k/\nu_0, \xi + l/b_0)|^2,$$

with $\Omega_T = \Omega^\circ \cap \text{Supp}(\eta_T)$. Optimal results are therefore expected if $g, h$ can be chosen in such a way that $K_e$ be minimum, for a given lattice, i.e. when the
cross ambiguity function $\mathcal{V}_g h$ is “concentrated” inside the fundamental domain $\Omega^\circ$. It is worth noticing that such a concentration property, which may be seen as a lattice-constrained time-frequency uncertainty, differs from what is usually required in, e.g. radar applications, since the geometry of the adjoint lattice has to be accounted for. Heuristically as well as from numerical experiments we know, that the tight window, [13], corresponding to a given lattice and a reasonably localized window is usually a good choice to fulfill this requirement, since this window automatically adapts to the eccentricity of the lattice. On the other hand, a window that is better concentrated inside $\Omega^\circ$, may be obtained by computing the eigenfunction corresponding to the biggest eigenvalue of the time-frequency localization operator corresponding to $\Omega^\circ$. This operator is given as an STFT multiplier

$$S_{\chi_{\Omega^\circ}} = \mathcal{V}_g^* \chi_{\Omega^\circ} \mathcal{V}_g,$$

$\chi_{\Omega^\circ}$ denoting the indicator function of the fundamental domain $\Omega^\circ$. Time-frequency localization operators have been studied extensively, e.g., [8, 3, 4, 14, 5]. Of course, this approach is less straight-forward and involves considerable computational effort. However, it may be generalized, as we will explain in the next section.

### 4.2 Multiple and generalized Gabor Multipliers

The main additional task in the generalized situation is the choice of the (additional) synthesis windows, or, in case only time-frequency shifts of one synthesis window are considered (generalized Gabor Multipliers), the choice of sampling points $\mu_j$ for the (additional) synthesis windows. A good choice will again be guided by the behavior of the spreading function.

Even for operators with a spreading function with compact support inside $\Omega^\circ$, considerable aliasing effects caused by the periodization of $\tilde{a}$ have to be taken into account. It turns out that choosing a few eigenfunctions of $S_{\chi_{\Omega^\circ}}$ can significantly reduce the amount of aliasing in the approximation of underspread operators by generalized Gabor multipliers. This idea was already noted in [17] in a slightly different context. It becomes now obvious from Corollary 1, why these particular functions, by their property of being maximally concentrated inside $\Omega^\circ$, represent a good choice for the synthesis windows of multiple Gabor multipliers.

When dealing with overspread operators, the additional energy of the operator’s spreading function outside $\Omega^\circ$ poses the second source of error on top of aliasing. The relevant areas in the spreading domain should be covered as well as possible with the smallest possible “leakage” beyond $\Omega^\circ$ of the different synthesis windows’ cross-ambiguity functions. Motivated by the results from the Gabor multiplier situation, we choose a tight window with respect to the analysis lattice and look for the most appropriate sampling points for the synthesis
windows. Examples will be given in Section 5.2. Alternatively, one or several eigenfunctions of $S_{\chi_{\Omega}}$ may be time-frequency shifted to cover the areas in phase space where significant energy of the operator’s spreading function occurs.

5 Examples

We now turn to numerical experiments, in the finite case $\mathcal{H} = \mathbb{C}^L$. In the following examples, the relative approximation error for the best approximation $T' = G_m$ of $T$ is measured by the logarithmic quantity

$$E = \log_{10}\left(\frac{\|T - G_m\|}{\|T\|}\right), \quad (12)$$

where $\| \cdot \|$ represents a generic matrix norm. We display below the results obtained using the Frobenius norm, the plots obtained with the operator norm are almost identical.

5.1 Classical Gabor Multipliers

Experiment 1: We generate operators with compact support in the spreading domain, in a square of side size between 3 and 61, symmetric about 0. The values are complex, uniformly distributed random numbers, the signal length is $L = 180$. We then investigate the approximation quality for various pairs of lattice constants, with $b_0$ varying between 2 and 18 and $\nu_0$ between 2 and 10. The results, averaged over 40 realizations of the random spreading functions, are presented in Figure 2. Evidently, the error decreases for decreasing $b_0$, in each plot. Note the two distinct regimes: the error grows exponentially (linearly in the logarithmic representation of the figure) up to a certain value of the support size, depending on the lattice density, and slower thereafter. The explanation for this effect is the observation, that the error (see the bound in (11)) is comprised of an aliasing error and the inherent inaccuracy of Gabor multiplier approximation, even for very high sampling density, of overspread operators.

Experiment 2: In order to emphasize the importance of lattice adaptation to eccentricity of the spreading function’s support, we show the results for different lattice constants resulting in the same redundancy (5) in Figure 3. The solid lines show the results for $b_0 = \nu_0 = 6$, leading to far better results than the lattice constants not adapted to the (symmetric) support of the spreading function.

5.2 Generalized Gabor Multipliers

Experiment 3: We now investigate the influence of additional synthesis windows on the approximation quality. We first consider the same operators as in the previous section, but allow for additional synthesis windows. We compare
Figure 2: Logarithmic approximation error for different bandwidth of spreading function and different values of $b_0, \nu_0$

Figure 3: Logarithmic approximation error for different lattice-eccentricity
the performance of generalized Gabor multipliers and multiple Gabor multipliers. We use up to 9 synthesis windows generated as time-frequency shifted copies of the tight window corresponding to $\Lambda = 9\mathbb{Z} \times 9\mathbb{Z}$. The sampling points are chosen from the intersection of $\Omega^\circ$ with $\Lambda$. On the other hand, we choose up to 9 eigenvectors of the time-frequency localization operator $S_{\chi_{\Omega^\circ}}$, using the eigenvectors corresponding to the biggest eigenvalues, in decreasing order. The results are displayed in Figure 4, where the two upper plots show the error for growing support of the spreading function. The errors for different numbers of synthesis windows are each plotted separately, the upper left display showing the errors resulting from the usage of eigenvectors ($eV$), while $eG$ denotes the errors resulting from using Gabor atoms as synthesis windows. Obviously, additional synthesis windows improve the approximation quality. In order to better compare the performance of the two different sets of synthesis windows, the differences between the errors obtained in both cases are shown in the lower plot. The choice of eigenfunctions yields an over-all improvement of approximation quality, as expected from the analytic results.

Figure 4: *Logarithmic approximation error for growing support of spreading function, comparing eigenfunctions and Gabor atoms as synthesis windows*
Experiment 4: We investigate the following situation: an operator with two effectively disjoint components in the spreading domain is, again, approximated by a multiple Gabor multiplier with two synthesis windows. For better comparison, the two components are the component from the previous examples plus a shifted version (by 90 samples) thereof. Figure 5 shows the spreading functions of one of the operators and its best approximation with two synthesis windows, for the optimal additional window. Note the aliasing effect.

In this situation, using two appropriate synthesis windows, the obtained results are similar to those in the case of one spreading function component and one synthesis window, as discussed in the previous section. In Figure 6, we display the results for 3 symmetric pairs of lattice constants, the optimal window’s result being represented by the solid line, while the dashed lines show the results of close but suboptimal synthesis windows. As the operator was generated by a translation by 90 samples, the tight window, shifted by 90 samples itself, is expected to be the optimal additional window. This is confirmed by the experiments. Analogously, the setting in Experiment 3 can be repeated for two (or more) disjoint components in the spreading domain, leading to largely comparable results.

Experiment 5: In a last experiment, we compare operators with a compact spreading function inside $\Omega^c$ (Operator 0), to operators with compact spreading function inside $\Omega^c \cup (\Omega^c + 1/\nu_0)$ (Operator 1) and $\Omega^c \cup (\Omega^c + 2/\nu_0)$ (Operator 2), respectively. In other words, in the first case, the two fundamental domains are adjacent and in the second case they are sufficiently separated. Again, the values of the operators’ spreading function are random. We investigate the approximation quality for 1 to 6 eigenfunctions of $S_{\chi_{\Omega^c}}$ in case of Operator 0, and
these eigenfunctions are, in each case, shifted to the appropriate fundamental domain. The lattice parameters were chosen to be $b_0 = \nu_0 = 9$, so that, in this case, the support of the spreading functions lies in a (union of) fundamental domain/s of $\Lambda^\circ$. This is the situation described in Corollary 2. Figure 7 shows the results of this experiment. The errors shown result from averaging over 40 different (random) spreading functions with the prescribed support.

As expected, the two disjoint spreading function components lead to largely comparable results. On the other hand, for two adjacent components, each of the components benefits from the eigenfunctions that are essentially concentrated in the neighboring fundamental domain. Apparently, this effect grows, when eigenfunctions less concentrated around 0 (i.e. corresponding to bigger eigenvalues) are added.

We remark, that the alternative idea to directly choose eigenfunctions corresponding to the bigger region $\Omega^\circ \cup (\Omega^\circ + 1/\nu_0)$ gives significantly poorer approximation results.
Figure 7: Comparing the approximation error for one or two spreading function components and 1 to 6 synthesis windows

6 Discussion and conclusions

The examples given in the previous section show that the choice of various parameters has considerable influence on the performance of approximation by (generalized) Gabor multipliers. While the situation is rather easily understood in the case of classical Gabor multipliers (at least qualitatively, though a better theoretical control would still be desirable), the generalized case involves a more subtle choice of additional parameters. Both the analytical results provided in Section 3 and the experiments from the previous section suggest that aliasing effects resulting from coarse sampling lattice can best be suppressed by using eigenfunctions of the localization operator $S_{\chi_{\Omega}}$ as additional synthesis windows. It should be noted that, while yielding better results in the approximation, using a small number of additional synthesis windows does not dramatically increase the computational cost: in (5), going from $|J| = 1$ to larger index sets $J$ involves inverting (generally small) matrices instead of computing a point-wise ratio.

Experiment 5 in the previous section shows that components of the spreading function of operators, that are not underspread, are very efficiently covered once
the (approximate) location of these components is known. Hence, the development of an efficient method to estimate relevant components of the spreading function will be an important question in further research on the present topic.

ACKNOWLEDGEMENT

The first author was funded by the Austrian Science Fund (FWF) project LOCATIF(T384-N13). The second author acknowledges partial support from CNRS through the PEPS programme MTF&Sons and from the European Union through the EU FET Open grant UNLocX (255931).

References


CONSTRUCTING AN INVERTIBLE CONSTANT-Q TRANSFORM WITH NONSTATIONARY GABOR FRAMES

Gino Angelo Velasco†‡, Nicki Holighaus∗, Monika Dörfler∗, Thomas Grill‡

†NuHAG, Faculty of Mathematics, University of Vienna, Austria
‡Institute of Mathematics, University of the Philippines, Diliman, Quezon City, Philippines
∗Austrian Research Institute for Artificial Intelligence (OFAI), Vienna, Austria

{ginovelasco,nicki.holighaus,monika.doerfler}@univie.ac.at, thomas.grill@ofai.at

ABSTRACT

An efficient and perfectly invertible signal transform featuring a constant-Q frequency resolution is presented. The proposed approach is based on the idea of the recently introduced nonstationary Gabor frames. Exploiting the properties of the operator corresponding to a family of analysis atoms, this approach overcomes the problems of the classical implementations of constant-Q transforms, in particular, computational intensity and lack of invertibility. Perfect reconstruction is guaranteed by using an easy to calculate dual system in the synthesis step and computation time is kept low by applying FFT-based processing. The proposed method is applied to real-life signals and evaluated in comparison to a related approach, recently introduced specifically for audio signals.

1. INTRODUCTION

Many traditional signal transforms impose a regular spacing of frequency bins. In particular, Fourier transform based methods such as the short-time Fourier transform (STFT) lead to a frequency resolution that does not depend on frequency, but is constant over the whole frequency range. In contrast, the constant-Q transform (CQT), originally introduced by J. Brown [1, 2], features a frequency resolution dependent on the center frequencies of the windows used for each bin and the center frequencies of the frequency bins are not linearly, but geometrically spaced. In this sense, the principal idea of CQT is reminiscent of wavelet transforms, compare [3]: the Q-factor, i.e. the ratio of the center frequency to bandwidth is constant over all bins and thus the frequency resolution is better for low frequencies whereas time resolution improves with increasing frequency. However, the transform proposed in the original paper [1] is not invertible and does not rely on any concept of (orthonormal) bases. In fact, the number of bins used per octave is much higher than most traditional wavelet techniques would allow for. Furthermore, the computational efficiency of the original transform and its improved versions, [4], may be insufficient.

CQTs rely on perception-based considerations, which is one of the reasons for their importance in the processing of speech and music signals. In these fields, the lack of invertibility of existing CQTs has become an important issue: for important applications such as masking of certain signal components or transposition of an entire signal or, again, some isolated signal components, the unbiased reconstruction from analysis coefficients is crucial. An interesting and promising approach to music processing with CQT was recently suggested in [5], also cf. references therein.

In the present contribution, we take a different point of view and consider both the implementation and inversion of a constant-Q transform in the context of the nonstationary Gabor transform (NSGT). Classical Gabor transform [6, 7] may be understood as a sampled STFT or sliding window transform. The generalization to NSGT was introduced in [8, 9] and allows for windows with flexible, adaptive bandwidths. Figure 1 shows examples of spectrograms of the same signal obtained from the classical sampled STFT (Gabor transform) and the proposed constant-Q nonstationary Gabor transform (CQ-NSGT).

If the analysis windows are chosen appropriately, both analysis and reconstruction is realized efficiently with FFT-based methods. The original motivation for the introduction of NSGT was the desire to adapt both window size and sampling density in time, in order to resolve transient signal components more accurately. Here, we apply the same idea in frequency: we use windows with adaptive, compact bandwidth and choose the time-shift parameters dependent on the bandwidth of each window. The construction of the atoms, i.e. the shifted versions of the basic window functions used in the transform, is done directly in the frequency domain, see Sections 2.2 and 3.1. This approach allows for efficient implementation using the FFT, as explained in Section 2.3. To exploit the efficiency of FFT, the signal of interest must be transformed into the frequency domain. For long real-life signals (e.g. signals longer than 10 seconds at a sampling rate of 44100Hz), processing is therefore done on consecutive time-slices, which is a natural processing step in real-time signal analysis 1. The resolution of the proposed CQ-NSGT is identical to that of the CQT and perfect reconstruction is assured by relying on concepts from frame theory, which will be discussed next.

2. NONSTATIONARY GABOR FRAMES

Frames were first mentioned in [10], also see [11, 12]. Frames are a generalization of (orthonormal) bases and allow for redundancy and thus for much more flexibility in design of the signal representation. Thus, frames may be tailored to a specific application of the analysis coefficients have to be expected. Mathematical details and error estimates will be given elsewhere.

1If the time-slicing is done using smooth windows with a judiciously chosen amount of zero-padding, no undesired artifacts after modification of the analysis coefficients have to be expected. Mathematical details and error estimates will be given elsewhere.
or certain requirements such as a constant-Q frequency resolution. Loosely speaking, we wish to expand, or represent, a given signal of interest as a linear combination of some building blocks or atoms \( \varphi_{n,k} \), with \((n, k) \in \mathbb{Z} \times \mathbb{Z}\), which are the members of our frame:

\[
f = \sum_{n,k} c_{n,k} \varphi_{n,k}
\]

for some coefficients \( c_{n,k} \). The double indexes \((n, k)\) allude to the fact that each atom has a certain location and concentration in time and frequency, compare Figure 2. Frame theory now allows us to determine, under which conditions an expansion (1) is possible and how coefficients leading to stable, perfect reconstruction may be determined.

We introduce the concept of frames for a Hilbert space \( \mathcal{H} \). In a continuous setting, one may think of \( \mathcal{H} = L^2(\mathbb{R}) \), whereas we will choose \( \mathcal{H} = \mathbb{C}^L \), \( L \) being the signal length, for describing the implementation.

### 2.1. Frames

Consider a collection of atoms \( \varphi_{n,k} \in \mathcal{H} \) with \((n, k) \in \mathbb{Z} \times \mathbb{Z}\). Here, \( n \) may be thought of as a time index and \( k \) as an index related to frequency. We then define the frame operator \( S \) by

\[
Sf = \sum_{n,k} \langle f, \varphi_{n,k} \rangle \varphi_{n,k},
\]

for all \( f \in \mathcal{H} \). Note that, if the set of functions \( \{ \varphi_{n,k}, (n, k) \in \mathbb{Z} \times \mathbb{Z} \} \) is an orthonormal basis, then \( S \) is the identity operator. If \( S \) is invertible on \( \mathcal{H} \), then the collection \( \{ \varphi_{n,k}, (n, k) \in \mathbb{Z} \times \mathbb{Z} \} \) is a frame.

In this case, we may define a dual frame by

\[
\gamma_{n,k} = S^{-1} \varphi_{n,k}.
\]

Then, reconstruction from the coefficients \( c_{n,k} = \langle f, \varphi_{n,k} \rangle \) is possible:

\[
f = S^{-1}Sf = \sum_{n,k} \langle f, \varphi_{n,k} \rangle S^{-1} \varphi_{n,k} = \sum_{n,k} c_{n,k} \gamma_{n,k}.
\]

### 2.2. The Case of Painless Nonstationarity

In a general setting, the inversion of the operator \( S \) poses a problem in numerical realization of frame analysis. However, it was shown in [13], that under certain conditions, usually fulfilled in practical applications, \( S \) is diagonal. This situation of painless non-orthogonal expansions can now be generalized to allowing for adaptive resolution. Adaptive time-resolution was described in [8, 9], and here we turn to adaptivity in frequency in the same manner.

In the sequel, let \( T_x \) denote a time-shift by \( x \), \( M_{\omega} \) denote a frequency shift (or modulation) by \( \omega \) and \( \mathcal{F}f = \hat{f} \) the Fourier transform of \( f \). Let \( \varphi_{x,k} \in \mathbb{Z} \), be band-limited windows, well-localized in time, whose Fourier transforms \( \psi_k = \hat{\varphi}_k \) are centered around possibly irregularly (or, e.g. geometrically) spaced frequency points \( \xi_k \).

Then, we choose frequency dependent time-shift parameters (hop-sizes) \( a_k \) as follows: if the support of \( \hat{\varphi}_k \) is contained in an interval of length \( |\xi_k| \), then we choose \( a_k \) such that

\[
a_k \leq \frac{1}{|\xi_k|} \quad \text{for all } k.
\]

In other words, the time-sampling points have to be chosen dense enough to guarantee this condition. Finally, we obtain the frame members by setting

\[
\hat{\varphi}_{n,k} = T_{na_k} \hat{\varphi}_k.
\]

Under these conditions on the windows \( \varphi_x \) and the hop-sizes \( a_k \), the frame operator is diagonal in the Fourier domain: since, by unitarity of the Fourier transform [14] and the Walnut representation of the frame operator [15], we have

\[
\langle Sf, f \rangle = \sum_{n,k} |\langle f, T_{na_k} \hat{\varphi}_k \rangle|^2 = \sum_{n,k} |\langle \hat{f}, M_{-na_k} \hat{\varphi}_k \rangle|^2 = \left\langle \sum_k \frac{1}{a_k} |\hat{\varphi}_k|^2 \hat{f}, \hat{f} \right\rangle,
\]
the frame operator assumes the following form:
\[ Sf = F^{-1} \left( \sum_k \frac{1}{a_k} |\hat{\phi}_k|^2 \right) \, f \, \right]. \tag{2} \]

See [16, 13, 17] for detailed proofs of the diagonality of the frame operator in the described setting. From (2), it follows immediately that the frame operator is invertible whenever there exist real numbers \( A \) and \( B \) such that the inequalities
\[ 0 < A \leq \sum_k \frac{1}{a_k} |\hat{\phi}_k|^2 \leq B < \infty \tag{3} \]
hold almost everywhere. In this case, the dual frame is given by the elements
\[ \gamma_{n,k} = T_{n,a_k} \left[ F^{-1} \left( \frac{1}{a_k} |\hat{\phi}_k|^2 \right) \right] \].

2.3. Realization in the Frequency domain

Based on the implementation of nonstationary Gabor frames performing adaptivity in the time domain [9], the above framework permits a fast realization by considering the Fourier transform of the input signal. The transform coefficients \( c_{n,k} = \langle f, \varphi_{n,k} \rangle \) take the form
\[ c_{n,k} = \langle f, T_{n,a_k} \varphi_k \rangle = \langle \hat{f}, M_{-n,a_k} \hat{\varphi}_k \rangle \],
and can be calculated, for each \( k \), with an inverse FFT (IFFT) of length determined by the support of \( \varphi_k = \hat{\varphi}_k \). Similarly, reconstruction is realized by applying the dual windows \( \hat{\gamma}_k = \hat{\varphi}_k / \sum_k \frac{1}{a_k} |\hat{\varphi}_k|^2 \) in a simple overlap-add process:
\[ \hat{f} = \sum_{n,k} \langle f, M_{-n,a_k} \varphi_k \rangle M_{n,a_k} \hat{\gamma}_k. \tag{4} \]

3. THE CQ-NSGT PARAMETERS: WINDOWS AND LATTICES

We will now describe in detail the parameters involved in the design of a nonstationary Gabor transform with constant-Q frequency resolution.

The CQT in [1] depends on the following parameters: the window functions, the number of frequency bins per octave, the minimum and maximum frequencies. These parameters determine the Q-factor, which is, as mentioned before, the ratio of the center frequency to the bandwidth. Here, the Q-factor is desired to be constant for all the relevant bins.

Let \( B \) and \( \Omega_{\text{min}} \) denote the number of frequency bins per octave and the desired minimum frequency, respectively. For the proposed CQ-NSGT, we consider band-limited window functions \( \varphi_k \in C^L, k = 1, \ldots, K \), with center frequencies \( \xi_k \) (in Hz) satisfying \( \xi_1 = \Omega_{\text{min}} 2^{\frac{1}{Q}} \), as in the classical CQT. The maximum frequency \( \Omega_{\text{max}} \) is restricted to be less than the Nyquist frequency \( \xi_{/2} \), where \( \xi_2 \) denotes the sampling frequency. Further, we require the existence of an index \( K \) such that \( \Omega_{\text{max}} \leq \xi_K < \xi_{/2} \). We may set \( K = \left[ B \log_2 \left( \frac{\Omega_{\text{max}}}{\Omega_{\text{min}}} \right) + 1 \right] \), with \([z] \) denoting the smallest integer greater than or equal to \( z \).

Note that in the CQT, since the frequency spacing in the CQT is geometric, no 0-frequency is present and some high frequency content might not be represented. In the CQ-NSGT, however, there is freedom to use additional center frequencies, at negligible computational cost, to guarantee perfect reconstruction.

In our current implementation, tailored to (real) audio signals, we consider some symmetry in the frequency domain, and take the following values for the frequency-centers \( \xi_k \):
\[ \xi_k = \begin{cases} 0, & k = 0 \\ \frac{\xi_{\text{min}} 2^{k-1}}{2^Q - 1}, & k = 1, \ldots, K \\ \xi_{/2}, & k = K + 1 \\ \xi_{k-2} + 2^{K-k}, & k = K + 2, \ldots, 2K + 1. \end{cases} \]

The bandwidth \( \Omega_k \) (the support of the window in frequency) of \( \varphi_k \) is set to be \( \Omega_k = \xi_{k+1} - \xi_k \), for \( k = 2, \ldots, K - 1 \), which leads to a constant Q-factor \( Q = (2^3 - 2^2) \). To obtain the same Q-factor on the relevant frequency bins, \( \Omega_1 \) and \( \Omega_{K+1} \) are therefore set to be \( \xi_1 / \Omega_1 \) and \( \xi_{K+1} / \Omega_{K+1} \), respectively. Finally, we let \( \Omega_0 = 2 \xi_1 = 2 \Omega_{\text{min}} \) and \( \Omega_{K+1} = \xi_{/2} - \xi_K \). In summary, we have the following values for \( \Omega_k \):
\[ \Omega_k = \begin{cases} 2 \Omega_{\text{min}}, & k = 0 \\ \xi_k / \Omega_1, & k = 1, \ldots, K \\ \xi_K, & k = K + 1 \\ 2 \xi_{K+1} / \Omega_{K+1}, & k = K + 2, \ldots, 2K + 1. \end{cases} \]

3.1. Window Choice: Satisfying the Frame Conditions

We now give the details on the windows \( \varphi_k \) to be used such that (3) and hence the frame property is fulfilled.

We use a Hann window \( \tilde{h} \) that is zero outside \([-1/2, 1/2]\), i.e., a standard Hann window centered at 0 with support of length 1. We obtain the atoms \( \varphi_k \) by translation and dilatation of \( \hat{\varphi}_k = \tilde{h}(j \xi_k / L - \xi_k) / \Omega_k \), \( k = 1, \ldots, K, K + 2, \ldots, 2K + 1, j = 0, \ldots, L - 1 \).

For the windows corresponding to the 0 and Nyquist frequencies, we use a plateau-like function \( g, e.g., \) a Tukey window. We obtain \( \varphi_0 \) and \( \varphi_{K+1} \) by setting \( \varphi_k[j] = g(j \xi_k / L - \xi_k) / \Omega_k \), \( k = 0, K + 1 \).

Now, for the collection of time-shifts of the constructed windows \( \varphi_k \), we require \( a_k \leq \xi_k / \Omega_k \), in order to satisfy (3). The \( \varphi_{n,k} \) are then given by their Fourier transforms as:
\[ \varphi_{n,k} = M_{-n,a_k} \varphi_k, \quad n = 0, \ldots, \left[ \frac{\xi_k}{2^Q} \right] - 1. \]

Figure 2 illustrates the time-frequency sampling grid of the set-up with the sampling points taken geometrically over frequency and linearly over time. Given these parameters, the coefficients of the CQ-NSGT are of the form \( c_{n,k} = \langle f, \varphi_{n,k} \rangle = \langle \hat{f}, \varphi_{n,a_k} \rangle \).

We note that the time-shift parameters can also be fixed to have the same value \( a = \min_k \{ a_k \} \) and the coefficients obtained from the CQ-NSGT can be put in a matrix of size \( [N/2] \times (K+1) \).

From the given support condition, the system \( \{ \varphi_k \} \) has an overlap factor of around 1/2. This implies that for the case where \( a_k = \xi_k / \Omega_k \), the redundancy of the system is approximately 2.

By construction, the sum \( \sum_{m=0}^{2K+1} \frac{L}{a_k} |\hat{\varphi}_{m,k}|^2 \) is finite and bounded away from 0. From Sections 2.2 and 2.3, the frame operator is invertible and perfect reconstruction of the signal is obtained from the coefficients \( c_{n,k} \) by applying (4).
4. SIMULATIONS

We now present some experiments comparing the original CQT with the CQ-NSGT in terms of reconstruction error, computation time and (visual) representation of sound signals.

Technical framework: All simulations were done in MATLAB R2009b on a 2 Gigahertz Intel Core 2 Duo machine with 2 Gigabytes of RAM running Kubuntu 9.04. The CQTs were computed using the code published with [5], available for free download at http://www.elec.qmul.ac.uk/people/anssik/cqt/. The CQ-NSGT algorithms are available at http://univie.ac.at/nonstatgab/cqt/.

For all experiments, \( \xi_{\text{min}} \) is taken to be \( \xi_K = \xi_{\text{min}} 2^{\frac{K}{2K+1}} \), where \( K \) is the largest integer such that \( 2K < \xi \).

4.1. Reconstruction Errors

The theoretical results, stating that the CQ-NSGT allows for perfect reconstruction, are confirmed by our experiments. For five test signals and various transform parameters, the relative reconstruction error

\[
e_{\text{rec}} = \sqrt{\frac{\sum_{j=0}^{L-1} |f[j] - f_{\text{rec}}[j]|^2}{\sum_{j=0}^{L-1} |f[j]|^2}}
\]

was calculated. With \( \xi_{\text{min}} \) between 10 Hz and 130 Hz and \( B \) from 12 to 192, the largest reconstruction error of the CQ-NSGT algorithm was slightly smaller than \( 1.6 \times 10^{-15} \), perfect reconstruction up to numerical precision. For comparison, it was shown in [5] that a CQT with reasonable amounts of redundancy and bins per octave can be inverted with a relative error of \( 10^{-3} \). This might not be enough for high-quality applications.

4.2. Computation Time and Computational Complexity

The required time for construction of the transform atoms and computation of the corresponding coefficients was measured for audio signals of roughly 6 to 18 seconds length, at a sampling rate of 44.1 kHz. Each experiment was repeated 50 times, the results are listed in Table 1. We note that for all signals, the CQ-NSGT is faster than the CQT implementation proposed in [5] by a considerable factor.

Our approach is still of complexity \( O(L \log L) \), though, and the advantage over the CQT decreases for longer signals. Each frequency channel’s time samples are acquired by means of sampled IFFT from the Fourier transform of the input signal, multiplied with the corresponding window. Therefore, a preliminary full length FFT is necessary.

More explicitly, we assume \( \bar{\omega} \) to have support of length \( M_0 \) and we denote by \( N_k \) the corresponding IFFT length. Let \( N = \max_k \{N_k\} \), i.e., the maximum IFFT-length, and we have \( M_0 \leq N_k \leq N \), since we only consider the painless case. Consequently, the number of operations is as follows:

1. FFT: \( O(N \log N) \).
2. Windowing: \( M_k \) operations for the \( k \)-th window.
3. IFFT: \( O(N_k \log(N_k)) \) for the \( k \)-th window.

The number of frequency channels \( 2K + 1 \) is independent of \( L \), since it is determined directly from the transform parameters. Thus, \( M_k \) and \( N_k \) are \( L \)-dependent and the computational complexity of the discrete CQ-NSGT is \( O(L \log L) \).

In applications, the dual windows are constructed directly on the frequency side and the painless case construction involves multiplication of the window functions by the inverse of a diagonal matrix, resulting in \( O(2^{2K+2} M_k) = O(L) \) operations. Finally, the inverse CQ-NSGT has numerical complexity

\[
\begin{array}{|c|c|c|}
\hline
\text{Bins per octave } B & \text{CQT mean time} & \text{CQ-NSGT mean time} \\
& \pm \text{variance (in seconds)} & \pm \text{variance (in seconds)} \\
\hline
12 & 0.95 \pm 0.01 & 0.36 \pm 0.00 \\
24 & 1.44 \pm 0.02 & 0.44 \pm 0.00 \\
48 & 2.42 \pm 0.03 & 0.65 \pm 0.00 \\
96 & 4.50 \pm 0.23 & 1.09 \pm 0.15 \\
\hline
\end{array}
\]
We note that linear computation time may be achieved by processing the signal in a suitable piecewise manner. Some experiments on that matter have been conducted, but the details of this procedure exceed the scope of this paper and are intended to be part of a future contribution.

In a second experiment, CQT and CQ-NSGT coefficients of the shortest sample, a Glockenspiel signal, were computed for several numbers of bins per octave. We note that the complexity of the algorithm for the CQ-NSGT is linear in \( B \).

The results, listed in Table 2, illustrate that the advantage of the CQ-NSGT algorithm increases for large numbers of bins.

### Visual Representation of Sound Signals

The spectral representation provided by CQT has several desirable properties, e.g. the logarithmic frequency scale resolves musical intervals in a similar way, independent of absolute frequencies. These properties are still present in the CQ-NSGT, in fact, its visual representation is practically identical to that of classical CQT as illustrated by Figure 3 for the exemplary case of the Glockenspiel signal. Figure 4 shows the CQ-NSGT of two additional music signals, further illustrating that even highly complex signals are nicely resolved by the proposed transform, similar to CQT.

5. EXPERIMENTS ON APPLICATIONS

Our experiments show applications of the CQ-NSGT in musical contexts, where the property of a logarithmic frequency scale renders the method often superior to the traditional STFT. Corresponding sound examples can be found at http://univie.ac.at/nonstatgab/cqt/.

5.1. Transposition

A useful property of continuous constant-Q decompositions is the fact that the transposition of a harmonic structure, like a note including overtones, corresponds to a simple translation of the logarithmically scaled spectrum. Approximately, this is also the case for the finite, discrete CQ-NSGT. In this experiment, we transposed a piano chord simply by shifting the inner frequency bins accordingly. By inner frequency bins, we refer to all bins with constant Q-factor. This excludes the 0-frequency and Nyquist frequency bins. The onset portion of the signal has been damped, since inharmonic components, such as transients, produce audible artifacts when handled in this way. In Figure 5, we show spectrograms of the original and modified chords, shifted by 20 bins. This corresponds to an upwards transposition by 5 semitones.

5.2. Masking

In the masking experiment, we show that the perfect reconstruction property of CQ-NSGT can be used to cut out components from a signal by directly modifying the time-frequency coefficients. The advantage of considerably higher spectral resolution at low frequencies (with a chosen application-specific temporal resolution at higher frequencies) compared to the STFT, makes the CQ-NSGT a very powerful, novel tool for masking or isolating time-frequency components of musical signals. Our example shows in Figure 6 a mask for extracting – or inversely, suppressing – a note from the Glockenspiel signal depicted in Figure 3. The mask was created as a gray-scale bitmap using an ordinary image manipulation program and then resampled in order to conform to the irregular time-frequency grid of the CQ-NSGT. Figure 6 shows the mask spectrogram, along with the spectrograms of the synthesized, processed signal and remainder.

6. SUMMARY AND PERSPECTIVES

We presented a constant-Q transform, based on nonstationary Gabor frames, that is computationally efficient and allows for perfect reconstruction. The described framework can easily be adapted to other perceptive frequency scales (e.g. mel or Bark scale) by choosing appropriate dictionaries.

The possibility of overcoming the difficulties that stem from dependence of the proposed transform on the signal length, e.g. by piecewise processing, is currently under investigation. This will further reduce computational effort and enable the use of a single family of frame elements for signals of arbitrary length.
Figure 4: Representations of a pipe organ and piano solo recordings, respectively, using the CQ-NSGT. The transform parameters were $B = 48$ and $\xi_{\text{min}} = 50$ Hz.

Figure 5: Piano chord signal and upwards transposition by 5 semitones, corresponding to a circular shift of the inner bins by 20. The transform parameters were $B = 48$ and $\xi_{\text{min}} = 100$ Hz.
Figure 6: Note extraction from the Glockenspiel signal by masking. The CQ-NSGT coefficients of the Glockenspiel signal were weighted with the mask shown on top. The remaining signal and extracted component are depicted in the middle and bottom respectively. The transform parameters were $B = 24$ and $\xi_{\text{min}} = 50$ Hz.

7. REFERENCES


A framework for invertible, real-time constant-Q transforms

Nicki Holighaus, Monika Dörfler, Gino Angelo Velasco and Thomas Grill

Abstract—Audio signal processing frequently requires time-frequency representations and in many applications, a non-linear spacing of frequency bands is preferable. This paper introduces a framework for efficient implementation of invertible signal transforms allowing for non-uniform frequency resolution. Non-uniformity in frequency is realized by applying nonstationary Gabor frames with adaptivity in the frequency domain. The realization of a perfectly invertible constant-Q transform is described in detail. To achieve real-time processing, independent of signal length, slice-wise processing of the full input signal is proposed and referred to as slicQ transform.

By applying frame theory and FFT-based processing, the presented approach overcomes computational inefficiency and lack of invertibility of classical constant-Q transform implementations. Numerical simulations evaluate the efficiency of the proposed algorithm and the method’s applicability is illustrated by experiments on real-life audio signals.

Index Terms—Time-frequency dictionary, constant-Q, Gabor frames, real-time, audio signals

I. INTRODUCTION

Analysis, synthesis and processing of sound is commonly based on the representation of audio signals by means of time-frequency dictionaries. The short-time Fourier transform (STFT), also referred to as Gabor transform, is a widely used tool due to its straightforward interpretation and FFT-based implementation, which ensure efficiency and invertibility [15], [7]. STFT features a uniform time and frequency resolution and a linear spacing of the time frequency bins.

In contrast, the constant-Q transform (CQT), originally introduced in [22] and in music processing by J. Brown [2], provides a frequency resolution that depends on geometrically spaced center frequencies of the analysis windows. In particular, the Q-factor, i.e. the ratio of center frequency to bandwidth of each window, is constant over all frequency bins; the constant Q-factor leads to a finer frequency resolution in low frequencies whereas time resolution improves with increasing frequency. This principle makes the constant-Q transform well-suited for audio data, since it better reflects the resolution of the human auditory system than the linear frequency-spacing provided by the FFT, cf. [20] and references therein. Furthermore, musical characteristics such as overtone structures remain invariant under frequency shifts in a constant-Q transform, which is a natural feature from a perception point of view. In speech and music processing, perception-based considerations are important, which is one of the reasons why CQTs, due to their previously discussed properties, are often desirable in these fields. An example of a CQT-transform, obtained with our algorithm, is shown in Figure 1.

The principal idea of CQT is reminiscent of wavelet transforms, compare [19]. As opposed to wavelet transforms, the original CQT is not invertible and does not rely on any concept of (orthonormal) bases. On the other hand, the number of bins (frequency channels) per octave is much higher in the CQT than most traditional wavelet techniques would allow for. Partly due to this requirement, the computational efficiency of the original transform as well as its improved versions, cf. [3], may often be insufficient. Moreover, the lack of invertibility of existing CQTs has become an important issue: for some desired applications, such as extraction and modification, e.g. transposition, of distinct parts of the signal, the unbiased reconstruction from analysis coefficients is crucial. Approximate methods for reconstruction from constant-Q coefficients have been proposed before, in particular for signals which are sparse in the frequency domain [5] and by octave-wise processing in [18].

In the present contribution, we are interested in inversion in the sense of perfect reconstruction; to this end, we investigate a new approach to constant-Q signal processing. The presented framework has the following core properties:

1) Relying on concepts from frame theory, [15], we suggest the implementation of a constant-Q transform using the nonstationary Gabor transform (NSGT), which guarantees perfect invertibility. This perfectly invertible constant-Q transform is subsequently called constant-Q nonstationary Gabor transform (CQ-NSGT).
2) We introduce a preprocessing step by slicing the signal to pieces of (usually uniform) finite length. Together with FFT-based methods, this allows for bounded delay and results in linear processing time. Thus, our algorithm lends itself to real-time processing and the resulting transform is referred to as sliced constant-Q transform (slicQ).

NSGTs, introduced in [11], [1], generalize the classical sampled short-time Fourier transform or Gabor transform [15], [10]. They allow for fast, FFT-based implementation of both analysis and reconstruction under mild conditions on the analysis windows. The CQ-NSGT was first presented in [21];
the frequency-resolution of the proposed CQ-NSGT is indistinguishable from that of the CQT, cf. Figure 1 for an example.

The main drawback of the CQ-NSGT is the inherent necessity to obtain a Fourier transform of the entire signal prior to actual processing. This problem prohibits real-time implementation and is overcome by a slicing step, which preserves the perfect reconstruction property. However, blocking effects and time-aliasing may be observed if the coefficients are modified in applications such as de-noising or transposition and time-shift of certain signal components. While slicing the signal naturally introduces a trade-off between delay and finest possible frequency resolution, the parameters can be chosen to suppress blocking artifacts and to leave the constant-Q coefficient structure intact.

The rest of this paper is organized as follows. In Section II we introduce the concepts of frames as overcomplete, stable spanning sets, with a focus on nonstationary Gabor (NSG) systems and their properties. We recall the conditions for these systems to constitute so-called painless frames, a special case that allows for straightforward inversion. Section III describes the construction of the CQ-NSGT by NSG frames with adaptivity in the frequency domain. This is the starting point for the sliCQ transform, which is explored in Section IV. After giving the general idea, we describe interpretation of the sliCQ-coefficients in relation to the full-length transform in Section IV-C. Subsequently, Section V is concerned with an analysis of the transforms' numerical properties, in particular computation time and complexity, as well as the quality of approximation of the CQ-NSGT coefficients by the sliCQ, accompanied by a set of simulations. Finally, in Section VI the CQ-NSGT is applied and evaluated in the analysis and processing of real-life signals. The paper is closed by a short summary and conclusion.

II. NONSTATIONARY GABOR FRAMES

Frames, first mentioned in [8], also cf. [4], [15], generalize (orthonormal) bases and allow for redundancy and thus design flexibility in signal representations. Frames may be tailored to a specific application or certain requirements such as a constant-Q frequency resolution. Loosely speaking, we wish to represent a given signal of interest as a sum of the frame members, or atoms, $\varphi_{n,k}$, weighted by coefficients $c_{n,k}$:

$$f = \sum_{n,k} c_{n,k} \varphi_{n,k}. \quad (1)$$

The double indexes $(n,k)$ allude to the fact that each atom has a certain location and concentration in time and frequency. Frame theory establishes conditions under which an expansion of the form (1) can be obtained with coefficients leading to stable, perfect reconstruction.

For this contribution, we only consider frames for $\mathbb{C}^L$, i.e. vector spaces of finite, discrete signals, understood as functions $f, g$ on $\mathbb{C}^L$. We denote by $\langle f, g \rangle$ the inner product of $f$ and $g$, i.e. $\langle f, g \rangle = \sum_{l=0}^{L-1} f[l]^* g[l]$ and $|f|_2 = \sqrt{\langle f, f \rangle}$. The structures introduced here can easily be extended to the Hilbert space of quadratically integrable functions, $L^2(\mathbb{R})$.

A. Frames

Consider a collection of atoms $\varphi_{n,k} \in \mathbb{C}^L$ with $(n,k) \in I_N \times I_K$ for finite index sets $I_N, I_K$. We then define the frame operator $S$ by

$$S f = \sum_{n,k} \langle f, \varphi_{n,k} \rangle \varphi_{n,k}, \quad (2)$$

for all $f \in \mathbb{C}^L$. If the linear operator $S$ is invertible on $\mathbb{C}^L$, then the set of functions $\{ \varphi_{n,k} \}_{(n,k) \in I_N \times I_K}$ is a frame\(^1\). In this case, we may define a dual frame by

$$\varphi_{n,k}^* = S^{-1} \varphi_{n,k} \quad (3)$$

and reconstruction from the coefficients $c_{n,k} = \langle f, \varphi_{n,k} \rangle$ is straightforward:

$$f = S^{-1} S f = \sum_{n,k} \langle f, \varphi_{n,k} \rangle S^{-1} \varphi_{n,k} = \sum_{n,k} c_{n,k} \varphi_{n,k}^* \quad (4)$$

$$= S S^{-1} f = \sum_{n,k} \langle f, \varphi_{n,k} \rangle \varphi_{n,k} = \sum_{n,k} \langle f, \varphi_{n,k}^* \rangle \varphi_{n,k}. \quad (5)$$

We next introduce a case of particular importance, the so-called Gabor frames, for which the elements $\varphi_{n,k}$ are obtained from a single window $\varphi$ by time- and frequency-shifts along a lattice. Let $T_x$ and $M_\omega$ denote a time-shift by $x$ and a frequency shift (or modulation) by $\omega$, i.e.

$$T_x f[l] = f[l - x] \quad \text{and} \quad M_\omega f[l] = e^{2\pi i\omega l/L} f[l].$$

\(^1\)Note that, if $\{ \varphi_{n,k}, (n,k) \in I_N \times I_K \}$ is an orthonormal basis, then $S$ is the identity operator.
where $l - x$ is considered modulo $L$. In other words, this is a circular shift operation. Furthermore, we use the normalization

$$
F f[j] = \hat{f}[j] = \frac{1}{\sqrt{L}} \sum_{l=0}^{L-1} f[l] e^{-2\pi i l j / L}
$$

for the discrete Fourier transform of $f$. It follows that $F(T_x f) = M_{-x} \hat{f}$ and $F(M_{x} f) = T_x F f$.

Fixing a time-shift parameter $a$ and a frequency-shift parameter $b$, with $L/a, L/b \in \mathbb{N}$, we call the collection of atoms

$$
\mathcal{G} = \{ \varphi_{n,k} = M_{a+b} T_{na} \varphi \}_{n,k \in I_N \times I_L},
$$

with $I_N \times I_K = \mathbb{Z}_{L/a} \times \mathbb{Z}_{L/b}$, a Gabor system. If $\mathcal{G}$ is a frame, it is called a Gabor frame. For Gabor frames, the frame coefficients are given by samples of the short-time Fourier transform of $f$ with respect to the window $\varphi$:

$$
e_{n,k} = \langle f, \varphi_{n,k} \rangle = \langle f, M_{a+b} T_{na} \varphi \rangle
= \sum_{l=0}^{L-1} f[l] \varphi[l-na] e^{-2\pi i l l / L}.
$$

In a general setting, the inversion of the operator $S$ poses a problem in numerical realization of frame analysis. However, for Gabor frames, it was shown in [6], that under certain conditions, usually fulfilled in practical applications, $S$ is diagonal, and a dual frame can be calculated easily. This situation of painless non-orthogonal expansions can now be generalized to allow for adaptive resolution.

### B. Frequency-Adaptive Painless Nonstationary Gabor Frames

In classical Gabor frames, we obtain all samples of the STFT in (4) by applying the same window $\varphi$, shifted along a regular set of sampling points and taking an FFT of the same length. In order to achieve adaptivity of the resolution in either time or frequency, we relax the regularity of classical Gabor frames to derive nonstationary Gabor frames.

The original motivation for the introduction of NSGT was the desire to adapt both window size and sampling density in time, cf. [11], [1], in order to accurately resolve transient signal components. Here, we apply the same idea in frequency, i.e., adapt both the bandwidth and sampling density in frequency. From an algorithmic point of view, we apply a nonstationary Gabor system to the Fourier transform of the input signal.

The windows are constructed directly in the frequency domain by taking real-valued filters $g_k$ centered at $\omega_k$. The inverse Fourier transforms $\tilde{g}_k := F^{-1} g_k$ are the time-reverse impulse responses of the corresponding (frequency-adaptive) filters. Therefore, we let $\tilde{g}_k, k \in I_K$, denote the members of a finite collection of band-limited windows, well-localized in time, whose Fourier transforms $g_k := F \tilde{g}_k$ are centered around possibly irregularly (or, e.g., geometrically) spaced frequency points $\omega_k$.

Then, we select frequency dependent time-shift parameters (hop-sizes) $a_k$ as follows: if the support (the interval where the vector is nonzero) of $g_k$ is contained in an interval of length $L_k$, then $a_k$ is chosen such that

$$
a_k \leq \frac{L}{L_k}
$$

In other words, the time-sampling points have to be chosen dense enough to guarantee (5). If we denote by $g_{n,k}$ the modulation of $g_k$ by $-n a_k$, i.e., $g_{n,k} = M_{-n a_k} g_k$, then we obtain the frame members $\varphi_{n,k}$ by setting

$$
\varphi_{n,k} = \tilde{g}_{n,k} = F^{-1} (M_{-n a_k} g_k) = T_{n a_k} \tilde{g}_k,
$$

where $k \in I_K$ and $n = 0, \ldots , L/a_k - 1$. The system $G(g,a) := \{ \varphi_{n,k} = T_{n a_k} \tilde{g}_k \}_{n,k}$ is a painless nonstationary Gabor system, as described in [1], for $C^L$. We also define $g := \{ g_k \in C^L \}_{k \in I_K}$ and $\alpha := \{ a_k \}_{k \in I_K}$. By Parseval’s formula, we see that the frame coefficients can be written as

$$
e_{n,k} = \langle f, \varphi_{n,k} \rangle = \langle \tilde{f}, M_{n a_k} g_k \rangle.
$$

For convenience, we use the notation $c := \{ c_k \}_{k \in I_K} := \{ \{ c_{n,k} \}_{n=0}^{L/a_k-1} \}_{k \in I_K}$ to refer to the full set of coefficients and channel coefficients, respectively. By abuse of notation, we indicate by $c \in \mathbb{C}^{L/a_k \times I_K}$ that $c$ is an irregular array with $|I_K|$ columns, the $k$-th column possessing $L/a_k$ entries. The NSG coefficients can be computed using the following algorithm.

#### Algorithm 1 NSG analysis: $c = \text{CQ-NSGT}_L(f,g,a)$

1. Initialize $f, g_k$ for all $k \in I_K$
2. $f \leftarrow \text{FFT}_L(f)$
3. for $k \in I_K$, $n = 0, \ldots , L/a_k - 1$
4. $\hat{c}_k \leftarrow \sqrt{L/a_k} \cdot \text{IFFT}_{L/a_k} (f \hat{g}_k)$
5. end for

Here $(I)FFT_N$ denotes a (inverse) Fast Fourier transform of length $N$, including the necessary zero-padding preprocessing to convert the input vector to the correct length $N$. The analysis algorithm above is complemented by Algorithm 2, an equally simple synthesis algorithm that synthesizes a signal $f$ from a set of coefficients $c$.

#### Algorithm 2 NSG synthesis: $\hat{f} = \text{iCQ-NSGT}_L(c,g,a)$

1. Initialize $c_{n,k}, \hat{g}_k$ for all $n = 0, \ldots , L/a_k - 1, k \in I_K$
2. for $k \in I_K$
3. $\hat{f}_k \leftarrow \sqrt{a_k/L} \cdot \text{FFT}_{L/a_k} (c_k)$
4. end for
5. $\hat{f} \leftarrow \sum_{k \in I_K} \hat{f}_k \hat{g}_k$
6. $\hat{f} \leftarrow \text{IFFT}_L(\hat{f})$

Remark 1. The algorithms proposed in this section can also be applied for $a_k \gg \frac{L}{L_k}$. However, in this case, applying $(I)FFT_{L/a_k}$ may require periodization or periodic extension, respectively, to convert the input to length $L/a_k$ or the output to length $L_k$.

If $G(g,a)$ and $\tilde{G}(g,a)$ are a pair of dual frames, then we can reconstruct a function perfectly from its NSG analysis coefficients. For more details and a proof of the following propositions, see Appendix A.

#### Proposition 1. Let $G(g,a) = \{ \tilde{g}_{n,k} = T_{n a_k} \tilde{g}_k \}_{n,k}$ and $G(\tilde{g},a) = \{ \hat{g}_{n,k} = T_{n a_k} \hat{g}_k \}_{n,k}$ be a pair of dual frames. If $c$ is the output of CQ-NSGT$_L(f,g,a)$ (Algorithm 1), then
TABLE I
CENTER FREQUENCY AND BANDWIDTH VALUES

<table>
<thead>
<tr>
<th>$k$</th>
<th>$\xi_k$</th>
<th>$\Omega_k$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>$2\xi_{\text{min}}$</td>
</tr>
<tr>
<td>$1, \ldots, K$</td>
<td>$\xi_{\text{min}}^2\frac{2\pi}{\Gamma}$</td>
<td>$\xi_k/Q$</td>
</tr>
<tr>
<td>$K + 1$</td>
<td>$\xi_k/2$</td>
<td>$\xi_k - 2\xi_K$</td>
</tr>
<tr>
<td>$K + 2, \ldots, 2K + 1$</td>
<td>$\xi_k - \xi_{2K+2-k}$</td>
<td>$\xi_{2K+2-k}/Q$</td>
</tr>
</tbody>
</table>

the output $\tilde{f}$ of $\text{iCQ-NSGT}_L(c, \mathbf{g}, \mathbf{a})$ (Algorithm 2) equals $f$, i.e.

$$\tilde{f} = f, \quad \text{for all } f \in \mathbb{C}^L. \quad (7)$$

The remaining problem is to ascertain that $\mathcal{G}(g, a)$ is a frame and to compute the parameters leading to an NSGT with constant-Q property. The following proposition is a discrete version of an equivalent result for NG systems in $L^2(\mathbb{R})$ and achieves both, using the painless case condition (5).

**Proposition 2.** Let $\mathcal{G}(g, a)$ on NSG system satisfying (5). This system is a frame if and only if

$$0 < \sum_{k \in I_K} \frac{L}{a_k} |g_k[j]|^2 < \infty, \quad \text{for all } j = 0, \ldots, L - 1 \quad (8)$$

and the filters generating the canonical dual frame $\hat{\mathcal{G}}(\mathbf{g}, \mathbf{a})$ are given by

$$\hat{g}_k[j] = \frac{g_k[j]}{\sum_{l \in I_K} \frac{1}{\pi l} |g_l[j]|^2}. \quad (9)$$

In the next section, we construct a constant-Q NSG system satisfying (5) and (8).

**Remark 2.** The maximum and minimum of the sum in (8) give the upper and lower frame bound, respectively.

**Remark 3.** Note that NSG frames can be equivalently used to design general nonuniform filter banks [14], [16] in a similar manner.

III. THE CQ-NSGT PARAMETERS: WINDOWS AND LATTICES

The parameters of the NSGT can be designed as to implement various frequency-adaptive transforms. Here, we focus on the parameters leading to an NSGT with constant-Q frequency resolution, suitable for the analysis and processing of music signals, as discussed in the introduction. In constant-Q analysis, the functions $g_k$ are considered to be filters with support of length $L_k \leq L$ centered at frequency $\omega_k$ (in samples), such that for the bins corresponding to a certain frequency range, the respective center frequencies and lengths have (approximately) the same ratio. Using these filters, the CQ-NSGT coefficients $c_{n,k}$ are obtained via Algorithm 1, where $k$ indexes the frequency bins, and $n = 0, \ldots, L/a_k - 1$.

As detailed in [21], the construction of the filters for the CQ-NSGT depends on the following parameters: minimum and maximum frequencies $\xi_{\text{min}}$ and $\xi_{\text{max}}$ (in Hz), respectively, the sampling rate $\xi_s$, and the number of bins per octave $B$. The center frequencies $\xi_k$ satisfy $\xi_k = \xi_{\text{min}}^2\frac{2\pi}{\Gamma}$, similar to the classical CQT in [2], for $k = 1, \ldots, K$, where $K$ is an integer such that $\xi_{\text{max}} \leq \xi_K < \xi_s/2$, the Nyquist frequency. Note that the correspondence between $\xi_k$ and $\omega_k$ is the conversion ratio from Hz to samples, as detailed in the next paragraphs.

The bandwidths are set to be $\Omega_k = \xi_{k+1} - \xi_{k-1}$, for $k = 2, \ldots, K - 1$, which lead to a constant Q-factor

$$Q = \xi_k/\Omega_k = (2\hat{\xi} - 2\xi_k^{-1})^{-1},$$

where $\Omega_1$ and $\Omega_K$ are taken to be $\xi_1/Q$ and $\xi_K/Q$, respectively. Since the signals are real-valued, additional filters are considered which are positioned in a symmetric manner with respect to the Nyquist frequency. Moreover, to ensure that the union of filter supports cover the entire frequency axis, filters with center frequencies corresponding to the zero frequency and the Nyquist frequency are included. The values for $\xi_k$ and $\Omega_k$ over all frequency bins are summarized in Table I.

With these center frequencies and bandwidths, the filters $g_k$ are set to be $g_k[j] = H((j\xi_s/L - \xi_k)/\Omega_k)$, for $k = 1, \ldots, K, K + 2, \ldots, 2K + 1$, where $H$ is some continuous function centered at 0, positive inside and zero outside of $]-1/2, 1/2[$, i.e. each $g_k$ is a sampled version of a translated and dilated $H$. Meanwhile, $g_0$ and $g_{K+1}$ are taken to be plateau functions, i.e. continuous, compactly supported functions that are constant 1 on some interval, centered at the zero and the Nyquist frequencies respectively. Thus, each filter $g_k$ is centered at $\omega_k = \xi_kL/\xi_s$ and has support $L_k = \Omega_kL/\xi_s$.

With the aforementioned parameters, we compute the phase-locked CQ-NSGT coefficients as

$$c_{n,k} = \sum_{l=0}^{L-1} \overline{f(l)} |g_k[l]| e^{2\pi i ((l - \omega_k) - na_k)/L}.$$  

This phase-lock convention, while slightly different from the definition (6) above, does not affect the frame property, yet implementation is more straightforward.

It is easy to see that this choice of $\mathcal{G}(g, a)$ satisfies the conditions of Proposition 2 for any sequence $a$ with $L/a_k \geq L_k$ for all $k \in I_k = \{0, \ldots, 2K + 1\}$. Note that while $a_k$ might be rational, $L/a_k$ must be integer-valued. Consequently, perfect reconstruction of the signal is obtained from the coefficients $c_{n,k}$ by applying Algorithm 2 with a dual frame, e.g. the canonical dual given by (9).

IV. REAL-TIME PROCESSING AND THE SLICQ

The CQ-NSGT implementation introduced in the previous sections a priori relies on a Fourier transform of the entire signal. This contradicts the idea of real-time applications, which require bounded delay in processing incoming samples and linear over-all complexity. These requirements can be satisfied by applying the CQ-NSGT in a blockwise manner, i.e. to (fixed length) slices of the input signal. However, the slicing process involves two important challenges: First, the windows $h_m$ used for cutting the signal must be smooth and zero-padding has to be applied to suppress time aliasing and blocking artifacts when coefficient-modification occurs. Second, the coefficients issued from the block-wise transform should be equivalent to the CQ-coefficients obtained from a full-length CQ-NSGT. This can be achieved to high precision by careful choice of both the slicing windows $h_m$ and the analysis windows $g_k$ used in the CQ-NSGT.
A. Structure of the sliCQ transform

We now summarize the individual steps of the sliCQ algorithm and introduce the involved parameters.

I) Sliced constant-Q NSGT analysis:

1) Cut the signal $f \in \mathbb{C}^L$ into overlapping slices $f^m$ of length $2N$ by multiplication with uniform translates of a slicing window $h_0$, centered at 0.
2) For each $f^m$, obtain coefficients $c^m \in \mathbb{C}^{2N/\alpha_k \times |V_K|}$ by applying CQ-NSGT$_{2N}(f, g, a)$ (Algorithm 1).
3) Due to the overlap of the slicing windows, cf. Figure 2, each time index is related to two consecutive slices. For visualization and processing, the slice coefficients $c^m$ are re-arranged into a 2-layer array $s$, with $s := \{s^1[i] \}_{i \in [0,1]} \in \mathbb{C}^{2 \times L/\alpha_k \times |V_K|}$, cf. Figure 3.

II) Sliced constant-Q NSGT synthesis:

1) Retrieve $c^m$ by partitioning $s$.
2) Compute the dual frame $\mathcal{G}[g, a]$ for $\mathcal{G}(g, a)$ and, for all $m$, $f^m = $ iCQ-NSGT$_{2N}(c^m, g, a)$ (Algorithm 2).
3) Recover $f$ by (windowed) overlap-add.

Note that $L$ must be a multiple of $2N$; this is achieved by zero-padding, if necessary. By construction, the positions $(n, k)$ of the coefficients in $s^1$ reflect their time-frequency position with respect to the full-length signal, for $l = 0, 1$.

B. Computation of a sliced constant-Q NSGT

The sliced constant-Q NSGT (sliCQ) coefficients of $f$ with respect to $h_0$ and $\mathcal{G}(g, a)$ and slice length $2N$ are obtained according to Algorithm 3.

Note that in this and the following algorithm, negative indices are used in a circular sense, with respect to the maximum admissible index, e.g. $f[-j] := f[L-j]$ or $s^1_{l,n,k} := s^1_{l/\alpha_k+n,k}$. As the CQ-NSGT analysis before, Algorithm 3 is complemented by a synthesis algorithm with similar structure, Algorithm 4, that synthesizes a signal $\tilde{f}$ from a 2-layer coefficient array $s$.

The following proposition states that $f$ is perfectly recovered from its sliCQ coefficients by applying Algorithm 4, see Appendix B for a proof.

Proposition 3. Let $\mathcal{G}(g, a)$ and $\mathcal{G}[g, a]$ be dual NSG systems for $\mathbb{C}^{2N}$. Further let $h_0, \tilde{h}_0 \in \mathbb{C}^L$ satisfy

$$\sum_{m=0}^{L/\alpha_k-1} T_m N \left( \tilde{h}_0 h_0 \right) \equiv 1. \quad (10)$$

If $s$ is the output of sliCQ$_{L,N}(f, h_0, g, a)$ (Algorithm 3), then the output $\tilde{f}$ of iCQ-NSGT$_{L, N}(c, \tilde{h}_0, a)$ (Algorithm 4) equals $f$, i.e., $\tilde{f} = f$.

C. The relation between CQ-NSGT and sliCQ

To maintain perfect reconstruction in the final overlap-add step in Algorithm 4, we assume

$$h_m = T_m N h_0 \quad \text{with} \quad \sum_{m=0}^{L/\alpha_k-1} h_m = 1, \quad (11)$$

and use a dual window $\tilde{h}_0$ satisfying (10) in the synthesis process.

Another obvious option for the design of the slicing windows is to require $\sum_{m=0}^{L/\alpha_k-1} h_m^2 = 1$, which would allow for using

---

**Algorithm 3** sliCQ analysis: $s = $ sliCQ$_{L,N}(f, h_0, g, a)$

1. Initialize $f, h_0, g, a$ for all $k \in I_K$
2. $m \leftarrow 0$
3. for $m = 0, \ldots, L/\alpha_k - 1$ do
4. for $j = 0, \ldots, N - 1$ do
5. $f^m[j] \leftarrow f^m[j + (m-1)N]$
6. end for
7. $c^m \leftarrow $ CQ-NSGT$_{2N}(f, g, a)$
8. $l \leftarrow (m \mod 2)$
9. for $k \in I_K$, $n^k = 0, \ldots, 2N/\alpha_k - 1$ do
10. $s^1_{l,n,k} \leftarrow c^n_{l,n,k}$
11. end for
12. end for

**Algorithm 4** sliCQ synthesis: $\tilde{f} = $ iCQ-NSGT$_{L, N}(c, \tilde{h}_0, g, a)$

1. Initialize $s, \tilde{h}_0, g, a$ for all $k \in I_K$
2. $\tilde{f} \leftarrow 0$
3. $\tilde{f} \leftarrow 0$
4. for $m = 0, \ldots, L/\alpha_k - 1$ do
5. $l \leftarrow (m \mod 2)$
6. for $k \in I_K$, $n^k = 0, \ldots, 2N/\alpha_k - 1$ do
7. $c^n_{l,n,k} \leftarrow s^1_{l,n,k}$
8. end for
9. $f^m \leftarrow $ iCQ-NSGT$_{L, N}(c^m, g, a)$
10. for $j = 0, \ldots, N - 1$ do
11. $\tilde{f}[j + (m-1)N] \leftarrow \tilde{f}[j + (m-1)N] + f^m[j] \tilde{h}_0[j - N]$
12. end for
13. end for
the same windows in the final overlap-add step. However, if we want to approximate the true CQ-coefficients as obtained from a full-length transform, (11) is the more favorable condition.

In our implementation, slicing of the signal is accomplished by a uniform partition of unity constructed from a Tukey window \( h_0 \) with essential length \( N \) and transition areas of length \( M \), for some \( N, M \in \mathbb{N} \) with \( M < N \) (usually \( M < N \)). The slicing windows are symmetrically zero-padded to length \( 2N \), reducing time aliasing significantly. The uniform partition condition (11) leads to close approximation of the full-length CQ-NSGT by sliCQ. This correspondence between the sliCQ and the corresponding full-length CQ-NSGT is made explicit in the following proposition, proven in Appendix B.

**Proposition 4.** Let \( \mathcal{G}(\mathbf{g}, \mathbf{a}) \) be a nonstationary Gabor system for \( \mathbb{C}^L \). Further, let \( h_0 \in \mathbb{C}^L \) be such that (11) holds and define \( g_k \in \mathbb{C}^{2N} \), for all \( k \in I_k \) by

\[
g_k[j] = g_k[jL/(2N)].
\]

For \( f \in \mathbb{C}^L \), denote by \( \mathbf{g} \in \mathbb{C}^{L \times [n] \times |v|} \) the CQ-NSGT coefficients of \( f \) with respect to \( \mathcal{G}(\mathbf{g}, \mathbf{a}) \) and by \( \mathbf{s} \in \mathbb{C}^{2 \times L \times [n] \times |v|} \) the sliCQ coefficients of \( f \) with respect to \( h_0 \) and \( \mathcal{G}(\mathbf{g}, \mathbf{a}) \). Then

\[
\left| f_n^{(n, k) + s_n^{(n, k)} - c_n, k} \right| \\
\leq \left| f[h] \right| \left\| (1 - h_0 - h_1) \mathbf{T}_{n, a_k} g_k \right\|_2 \\
+ \left| (h_0 + h_1) \sum_{j=1}^{a_k - 1} \mathbf{T}_{n, a_k + j} g_k \right\|_2.
\]

**Remark 4.** In practice, \( g_k \) is chosen such that the translates \( \mathbf{T}_{n, a_k} g_k \) are essentially concentrated in

\[
I_{N, M} = \left[ - \frac{N - M}{2}, N + \frac{N - M}{2} \right],
\]

i.e. \( \left| \mathbf{T}_{n, a_k} g_k \right|_{\mathbb{C}^{L \times [n] \times |v|}} \ll \left| \mathbf{T}_{n, a_k} g_k \right|_2 \), for all \( n = 0, \ldots, N/a_k - 1 \). Therefore, the value of (12) is negligibly small. While more precise estimates of the error are beyond the scope of the present contribution, numerical evaluation of the approximation quality is given in Section V-C.

As a consequence of the previous definition, we propose the sliCQ spectrogram as \( \left| f^{(n, k) + s^{(n, k)}} \right|^2 \) and propose to simultaneously treat \( s_{n, k}^{(n, k)} \) and \( s_{n, k}^{(n, k)} \), corresponding to the same time-frequency position, when processing the coefficients.

**V. NUMERICAL ANALYSIS AND SIMULATIONS**

In this section we treat the computational complexity of CQ-NSGT and sliCQ and how they compare to one another. In [21] it was shown that despite superlinear complexity, CQ-NSGT outperforms state-of-the-art implementations of the classical constant-Q transform. Since sliCQ is a linear cost algorithm, it further improves the efficiency of the CQ-NSGT for sufficiently long signals. Section V-C provides experimental results confirming the good approximation of CQ-NSGT by the sliCQ coefficients, cf. Proposition 4.

The CQ-NSGT and sliCQ Toolbox (for MATLAB and Python) used in this contribution is available at http://www.univie.ac.at/nonstatgab/slicq along side extended experimental results complementing those presented in Section VI.

A. Computation Time and Approximation Complexity

We assume the number of filters \( |I_k| \) in the CQ-NSGT to be independent of the signal length \( L \) and Proposition 2 to hold, in particular \( L/a_k \geq L_k \). The support size \( L_k \) of each filter \( g_k \) depends on \( L \). Hence, the number of operations for Algorithm 1 is as follows:

\[
O\left( L \log(L) + \sum_{k \in I_k} \frac{L}{a_k} \log \left( \frac{L}{a_k} \right) + \frac{L_k}{\pi} \right).
\]

With \( L_k \) and \( L/a_k \) bounded by \( L \), this can be simplified to \( O(L \log(L)) \).

The computation of the dual frame involves inversion of the multiplication operator \( \mathbf{S} \) and applying the resulting operator \( \mathbf{S}^{-1} \) to each filter. This results in \( O(2 \sum_{k \in I_k} L_k) = O(L) \) operations, where the support of the \( g_k \) was taken into account. Complexity of Algorithm 2 can be derived to be \( O(L \log(L)) \), analogous to Algorithm 1.

For sliCQ, \( I_k \) (Algorithm 3), we assume the slice length \( 2N \) to be independent of \( L \), resulting in a computational complexity of

\[
O\left( \frac{L}{N} \cdot \left( \frac{2N \log(2N)}{\#slices} + \frac{2N}{\#slices} \right) \right) = O(L).
\]

Both the dual frame and \( \tilde{h}_0 \) can be precomputed independent of \( L \), whilst Algorithm 4 is of complexity \( O(L) \), analogous to Algorithm 3.

B. Performance Evaluation

A comparison of the CQ-NSGT algorithm with previous constant-Q implementations was given in [21]. Figure 4 re-produces and extends some of the results; it shows, for both the constant-Q implementation provided in [18] and CQ-NSGT, mean computation duration and variance for analysis followed by reconstruction, against signal length. The plot also
Fig. 5. Computation time versus signal length of the CQ transform (dotted gray), CQ-NSGT (dashed gray) and various sliCQ transforms. The sliCQ transforms were taken with slice lengths 4096 (solid gray), 16384 (dotted black), 32768 (dashed black) and 65536 (solid black) samples.

illustrates the dependence of CQ-NSGT on the prime factor decomposition of the signal length $L$.

Figure 5 illustrates the performance of sliCQ compared to the constant-Q and CQ-NSGT algorithms shown in Figure 4. Linearity of the sliCQ algorithm becomes evident, with deviations occurring due to unfavorable FFT lengths $2N/q_k$ in (i)CQ-NSGT.$_{2N}$. Performance improvements for increasing slice length can be attributed to the advanced nature of MATLAB’s internal FFT algorithm, as compared to the current implementation of the sliCQ framework.

The performance of the involved algorithms does not depend on signal content. Consequently, random signals were used in the performance experiments, although we implicitly assumed the signals to be sampled at 44.1 kHz. All the results represent transforms with 48 bins per octave, minimum frequency 50 Hz and maximum frequency 22 kHz, in Section V1 a maximum frequency of 20 kHz is used instead. For a more comprehensive comparison of the CQ-NSGT to previous constant-Q transforms, please refer to [21]. Results for other parameter values do not differ drastically and are omitted.

All computation time experiments were run in MATLAB R2011a on a 3 Gigahertz Intel Core 2 Duo machine with 2 Gigabytes of RAM running Kubuntu 10.04 using the MATLAB toolboxes available at http://www.elec.qmul.ac.uk/people/anssik/cqt/ and http://www.univie.ac.at/nonstatgab/.

C. Approximation properties

To verify the approximate equivalence of the sliCQ coefficients to those of a full-length CQ-NSGT and thus to a constant-Q transform, we computed the norm difference between $s^0 + s^1$ and $c$ as in Proposition 4, for two sets of fundamentally different signals. Set 1 contains 50 random, complex-valued signals of $2^{20}$ samples length, while Set 2 consists of 90 music samples of the same length, sampled at 44.1 kHz each, covering pop, rock, jazz and classical genres. The signals of the second set are well-structured and often well-concentrated in the time-frequency plane, characteristics that the first set lacks completely.

For discretization reasons as well as to achieve good concentration of $g^k$ in Proposition 4, sliCQ implementations must impose a lower bound on the length of $g^k$. Approximation results for various lower bounds on the filter length are summarized in Figure 6, showing the mean approximation quality over the whole set.

All errors are given in signal-to-noise ratio, scaled in dB:

$$20 \log_{10} \frac{\|c\|_2}{\|c - (s^0 + s^1)\|_2}$$

Figure 6 shows that, independent of other parameters, a minimal filter length smaller than 8 samples leads to a representation that is visibly different from, while values above 16 samples yield coefficients that are largely equivalent to those of a constant-Q transform. We can see that the slice length itself has rather small influence on the results, while the interplay of slicing window shape, specified by the ratio of transition area length to slice length, and minimal filter length is illustrated nicely; remarkably, this ratio influences the approximation quality mainly for moderately well localized filters. This is in correspondence with the characterization given in (12): the circular overspill, given by the second term of the right hand side in (12), depends on the shape and support of the sum of two adjacent slicing windows, in particular for moderately well localized filters. If the windows are very well localized, the overspill is small independent of the particular shape of the slicing area. On the other hand, very badly localized windows make the distinct influence of the slicing windows negligible. Finally, a comparison of the top and bottom graphs in Figure 6 shows that the approximation quality is largely independent of the signal class. For Set 1 the variance is generally negligible (< 0.1 dB) and was omitted. Despite some outliers in Set 2, we have found the approximation quality to depend on the minimal filter length in a stable way, cf. Figure 7. These outliers can be attributed to signals particularly sparse (smaller
Fig. 7. Coefficient approximation error (12) for all signals from Set 2 and slice and transition length of 65536, resp. 16384 samples. Line style indicates the minimal bandwidth: 8 (dotted), 16 (dashed) and 32 (solid) samples.

Fig. 8. Masks for extracting a transient (top) and sinusoidal component (bottom) of the Glockenspiel signal. The gray level plot describes the amplitude of the mask, with black and white representing 1 and 0, respectively.

error) or dense (larger error) in low frequency regions, where $q_k^C$ is least concentrated.

VI. EXPERIMENTS ON APPLICATIONS

Experiments in [21] show how the CQ-NSGT can be applied in the processing of signals taking advantage of the logarithmic frequency scaling and the perfect reconstruction property. In particular, the transposition of a harmonic structure amounted to just a translation of the spectrum along frequency bins, while the masking of the CQ-NSGT coefficients allowed for the extraction or suppression of a component of the signal. In our experiment, we show that the two procedures can be used to modify a portion of a signal.

Figure 8 shows masks for isolating a transient part and the corresponding sinusoidal part of a Glockenspiel signal, created using an ordinary image manipulation program. Therein, the layers paradigm has been used to be able to quickly switch on and off the masks in order to accurately adapt them to the CQ-NSGT representation of the audio. An “inverse mask” is also constructed for the remainder part of the signal, essentially decomposing the signal into transient, sinusoidal and background portions. The masks have been drawn in the logarithmic domain, to be able to handle the dynamics of the audio. They are linearly scaled in dB units, so that 0 in the mask corresponds to $10^{-5}$ (-100 dB) and 1 corresponds to 0 (0 dB).

While keeping the transient part, the isolated sinusoidal component of the signal is transposed upward by 2 semitones, corresponding to 8 frequency bins. The transient, the remainder, and the modified sinusoidal coefficients are then added and the inverse transform is applied to obtain the resulting processed signal. For ease of use, this process is done with a rectangular representation of the slices, obtained by choosing $L/a_k$ constant for all frequency bands which corresponds to a sinc-interpolation of the coefficients.

Figure 9 compares the CQ-NSGT spectrograms of the original and the modified signal, while Figure 10 shows the results for the same experiment using sliCQ transforms with different slice lengths. Note that the plots show the spectrogram of the synthesized signal, not the time-frequency coefficients before synthesis. Further, the exact same mask was used for CQ-NSGT and sliCQ transpositions. The sound files for this and other transposition experiments are available at http://www.univie.ac.at/nonstatgab/slicq. A script for the Python toolbox that executes the experiment, is available on the same page.

For synthesis, performed from modified coefficients, as opposed to mere reconstruction, an evaluation of the results is a highly non-trivial matter. This is due to the lack of a properly defined notion of accuracy or the existence of a target signal, not only for the algorithms presented here, but for any analysis/synthesis based signal processing framework. Thus, while the examples in this section should indicate that CQ-NSGT synthesis and sliCQ synthesis can produce results in accordance with intuition, an in-depth treatment of this subject is far beyond the scope of this article.

VII. SUMMARY AND CONCLUSION

In this contribution, we have introduced a framework for real-time implementation of an invertible constant-Q transform based on frame theory. The proposed framework allows for straight-forward generalization to other non-linear frequency scales, such as mel- or Bark scale, cp. [9]. While real-time processing is possible by means of a preprocessing step, we investigated the possible occurrence of time-aliasing. We provided a numerical evaluation of computation time and quality of approximation of the true NSGT coefficients.

In analogy to the classical phase vocoder, phase issues have to be addressed, if CQ-transformed coefficients are processed, cp. [12], [13], [17]. While preliminary experiments using the proposed framework for real-life signals were presented, undesired phasing effects, mainly due to the contribution of a signal component to several adjacent filters, will be investigated in detail in future work. Furthermore, future work
will consider the efficient realization of adaptivity both in
time and frequency by varying the length of the preprocessing
windows used for slicing.

**APPENDIX**

**A. Derivation of CQ-NSGT properties**

**Proof of Proposition 1:** By Algorithm 1, we have

\[
c_{n,k} - c_k[n] = \sqrt{L/a_k} - \frac{1}{\sqrt{L/a_k}} \sum_{m=0}^{L/a_k-1} \sum_{l=0}^{a_k-1} \left( \hat{f} \left( \frac{m - l}{a_k} \right) \right) e^{2\pi i m n a_k / L} - \sum_{m=0}^{L/a_k-1} \sum_{l=0}^{a_k-1} \left( \hat{f} \left( \frac{m - l}{a_k} \right) \right) e^{2\pi i m n a_k / L}.
\]

(13)

Since \( L/a_k \geq L_k \), only one element of the inner sum above is non-zero, for each \( m \in \{0, \ldots, L/a_k - 1\} \). It follows that

\[
c_{n,k} = \langle \hat{f}, M_{n,a_k} g_k \rangle.
\]

(14)

Inserting into Algorithm 2 yields, for all \( j \in \{0, \ldots, L - 1\} \),

\[
\hat{f}[j] = \sum_{k \in I_k} \left[ \sum_{n=0}^{L/a_k-1} c_{n,k} e^{-2\pi i m n a_k / L} \hat{g}_k[j] \right] - \sum_{k \in I_k} \sum_{n=0}^{L/a_k-1} \langle \hat{f}, M_{n,a_k} g_k \rangle M_{n,a_k} \hat{g}_k[j].
\]

**Proof of Proposition 2:** Denote by \( J_k \) an interval of

length \( L_k, L_k \) as in Section II, containing the support of \( g_k \).

By assumption

\[
0 < \sum_{k \in I_k} |g_k[j]|^2 < \infty, \quad \text{for all } j = 0, \ldots, L - 1
\]

and \( L/a_k \geq L_k = |J_k| \). Note that the frame operator (2) can

be written as follows

\[
S_f[j] = \sum_{k \in I_k} \sum_{n=0}^{L/a_k-1} \langle f, M_{n,a_k} g_k \rangle M_{n,a_k} \hat{g}_k[j]
\]

(15)

for all \( f \in \mathbb{C}^L \). Furthermore, with \( \chi_{J_k} \) the characteristic

Fig. 9. CQ-NSGT spectrograms showing an excerpt of the Glockenspiel signal before (top) and after transposition of a component (bottom).

Fig. 10. sliCQ spectrograms showing an excerpt of the Glockenspiel signal after transposition of a component. The top plot was done with a slice length of 50000 and a transition area of 20000 samples, the bottom plot with a slice length of 5000 and a transition area of 2000 samples.
function of the interval $J_k$,
\[
f_{L} = \chi_{J_k} \sum_{l=0}^{\frac{m}{a_k}} T_{l,a_k}(f_{L})
\]
and, obviously, $g_k = \chi_{J_k} g_k$. Inserting into (15) yields
\[
Sf[j] = \sum_{k \in I_k} \frac{L}{a_k}[f_{L}]g_k[j]\]
\[
= f[j] \sum_{k \in I_k} \frac{L}{a_k}[g_k]_j^k. \tag{16}
\]
With the sum bounded above and below, the inverse frame transform follows.
\[
Cf = \sum_{k \in I_k} \frac{L}{a_k}[g_k]_j^k, \quad \text{for all } f \in \mathbb{C}^L.
\tag{17}
\]
Since the elements of the canonical dual frame are given by (3), this completes the proof.

B. Derivation of sliCQ properties

Proof of Proposition 3: According to Proposition 1, $\tilde{f}^m$, the output of iCQ-NSGT in Step 9 of Algorithm 4 satisfies to $f_{m+1}^m[j] = (f \cdot T_{m,N} h_0)[j + (m - 1)N]$. Since $\sum_m T_m N (h_0 h_0) = 1$ holds,
\[
\tilde{f} = f^m \cdot T_{m,N} h_0 T_{m,N} h_0 - f \cdot \sum_m T_m N (h_0 h_0) = f
\]
follows.

Proof of Proposition 4 : Since $g_k$ is obtained by sampling $g_k^c$ with sampling period $L/2N$, the (inverse) Fourier transform $\tilde{g}_k$ of $g_k$ is given by periodicization of $g_k^c$ as follows:
\[
\tilde{g}_k[l] = \sum_{j=0}^{\frac{m}{a_k}-1} g_k^c[l + j \cdot 2N]. \tag{18}
\]
Recall from (6) that the CQ-NSGT coefficients of $f$ with respect to $G$($g_k^c$, $a$) are given by $e_{n,k} = \langle f, T_{n,a_k} g_k^c \rangle$ while the CQ-NSGT coefficients $e_n^m$ of $f^m$ are, for $m = 0, \ldots, L/N - 1, n^* = 0, \ldots, \frac{m}{a_k} - 1$ and $k \in I_k$
\[
e_{n^*,k}^m = \sum_{j=0}^{\frac{m}{a_k}-1} T_{n^*,a_k+1} g_k^c \tag{19}
\]
where the final inner product is taken over $\mathbb{C}^L$. Observe that every $n = 0, \ldots, \frac{m}{a_k} - 1$ can be written as $n = m \frac{N}{a_k} + n^*$ with $n^*$ from 0, $\ldots, \frac{N}{a_k} - 1$ and thus
\[
s_{n,k}^m = \sum_{j=0}^{\frac{m}{a_k}-1} T_{n^*,a_k+1} N g_k^c \tag{20}
\]
Hence $s_{n,k}^0 + s_{n,k}^1 - c_{n,k} = R[n]$. The result follows from Cauchy-Schwartz’ inequality, applied to the case $m = 0$, observing independence from $m$.\]

Acknowledgment

This research was supported by the WWTF project AudioMiner (MA09-024), the Austrian Science Fund (FWF):T384-N13] and the EU FET Open grant UNLocX (255931). The authors wish to thank the reviewers for their extremely helpful and constructive remarks.

References

N. Holighaus studied mathematics and theoretical computer sciences at Justus–Liebig–University, Gießen, Germany. After graduation in 2010, he took a position as research assistant at the Numerical Harmonic Analysis Group (NuHAG), Faculty of Mathematics, University of Vienna, Austria, where he currently pursues a Ph.D. degree.

His research interests include applied frame theory and time-frequency analysis, adaptive time-frequency techniques and signal processing.

M. Dörfler obtained her PhD in Mathematics from the University of Vienna and is a researcher at the Faculty of Mathematics. She studied piano at the Music University of Vienna and is working in the field of applied mathematics for audio signal processing.

She is interested in the interplay of local and global aspects of time-frequency analysis, and focuses on the benefits of theoretical results in practical applications. She is heading the interdisciplinary research project AudioMiner.

G.A. Velasco received his B.S. and M.S. degree in Mathematics from the University of the Philippines Diliman in 2001 and 2006, respectively. He is an instructor at the Institute of Mathematics, University of the Philippines Diliman and is pursuing a Ph.D. degree at the Numerical Harmonic Analysis Group (NuHAG), Faculty of Mathematics, University of Vienna.

His research interests include approximation theory, time-frequency analysis and its application to signal processing.

T. Grill obtained his PhD in Sound and Music Computing from the University of Music and Performing Arts in Graz, Austria. He holds a lectureship for sonic art at the University of Applied Arts Vienna, Austria, and a position as researcher at the Austrian Research Institute for Artificial Intelligence (OFAI).

Also being an electroacoustic composer and performer, his research interests include auditory and cross-modal perception, sound and music computing, and new interfaces for musical expression.
STRUCTURED SPARSITY FOR AUDIO SIGNALS

Kai Siedenburg\textsuperscript{\textdagger}, Monika Dörfler\textsuperscript{\dagger}

\textsuperscript{*} Department of Mathematics, Humboldt University Berlin
\textsuperscript{\dagger}NuHAG, Faculty of Mathematics, University of Vienna, Austria

kai.siedenburg@gmx.de, monika.doerfler@univie.ac.at

ABSTRACT

Regression problems with mixed-norm priors on time-frequency coefficients lead to structured, sparse representations of audio signals. In this contribution, a systematic formulation of thresholding operators that allow for weighting in the time-frequency domain is presented. The related iterative algorithms are then evaluated on synthetic and real-life audio signals in the context of denoising and multi-layer decomposition. Further, initial results on the influence of the shape of the weighting masks are presented.

1. INTRODUCTION

Most audio signals of importance for humans, in particular speech and music, are highly structured in time and frequency. Typically, salient signal components are sparse in time (or frequency) and persistent in frequency (or time). Sparsity in time is connected to transient events, while sparsity in frequency is observed in harmonic components. Processing sound signals with time-frequency dictionaries is ubiquitous. The sparse structure usually seen can be further enhanced by procedures such as basis pursuit \cite{1} or \ell^2-regression \cite{2}. In the context of time-frequency dictionaries, a natural step beyond classical sparsity approaches is the introduction of sparsity criteria which take into account the two-dimensionality of the time-frequency representations used. Mixed norms on the coefficient arrays make it possible to enforce sparsity in one domain and diversity and persistence in the other domain. Regression with mixed-norm priors was first proposed in \cite{3}. In the current contribution, we consider a family of specific regression problems with \ell^2- and \ell^4-priors on the coefficients; the algorithms derived thereof are refined by using local neighborhood-weighting. The performance of the resulting algorithms differ and are systematically evaluated for classical signal processing tasks like denoising and sparse multi-layer decomposition. Applications lead to quite satisfactory results in terms of measured (SNR) and listening. The presented results reflect a first step in the exploitation of structured shrinkage in the sense of informed analysis, i.e., using some available prior knowledge about the signal under consideration. The main contribution is the generalization and application of structured shrinkage operators \cite{3} to representations of audio signals by frames.

2. TECHNICAL TOOLS

We seek to expand a signal $s \in \mathbb{C}^L$ in the form

$$s(n) = \sum_{k,j} c_{k,j} \varphi_{k,j}(n) + r(n), \quad n = 1, \ldots, L \quad (1)$$

where the $\varphi_{k,j}$ denote the atoms of a time-frequency dictionary $\Phi$, $c_{k,j}$ are the expansion coefficients and $r$ is some residual. In order to guarantee perfect and stable reconstruction of a signal from its associated analysis coefficients $c_{k,j} = (s, \varphi_{k,j})$, we assume that the dictionary $\Phi$ forms a frame \cite{4}. We consider Gabor frames, which are exhaustively used in music processing, be it under a different name: in their simplest instantiation they correspond to a sampled sliding window or short-time Fourier transform. Gabor frames consist of a set of atoms $\varphi_{k,j} = M_J T_{k,j} \varphi$, where $T_{k,j}$ and $M_J$ denote the time- and frequency-shift-operator, resp., defined by $T_{k,j} \varphi(n) = \varphi(n-j)$, $M_J \varphi(n) = \varphi(n) e^{-2\pi i j n / L}$, and $\varphi$ is a standard window function. $a$ and $b$ are the time- and frequency sampling constants, and $j = 0, \ldots, J-1, k = 0, \ldots, K-1$, with $Ka = Jb = L$.

We will even assume more, namely tightness of the frames in use, which means that, up to a constant which may be set to 1, we have $s = \sum_{k,j} (s, \varphi_{k,j}) \varphi_{k,j},$ i.e., synthesis is done with the analysis window. Tight frames are easily calculated, see \cite{5}. In the finite case, the frame’s atoms constitute the columns of a matrix $\Phi$ which is of dimension $L \times p$; for tight frames, we have $\Phi^* \Phi \cdot s = s$. Since we are especially interested in the redundant case $L < p$, the additional degrees of freedom are used to promote sparsity of the coefficients.

2.1. Regression with mixed norms

Sparsity of coefficients may be enforced by \ell^2-regression, also known as the Lasso \cite{2}. Given a noisy observation $y = s + e$ in $\mathbb{C}^L$ it finds

$$\hat{c} = \arg \min_{c \in \mathbb{C}^L} \frac{1}{2} \| y - \Phi c \|^2_2 + \lambda \Psi(c) \quad (2)$$

with penalty term $\Psi(\cdot) = \| \cdot \|_1$, and $\lambda > 0$. Since the sequence $c_{k,j}$ is ordered along two dimensions for Gabor frames, the \ell^4-prior $\Psi$ in (2) may be replaced by a two-dimensional mixed norm $\ell^p,q$ which acts differently on groups (indexed by $g$ in the sequel, may be either time or frequency) and their members (indexed by $m$):

$$\Psi(c) = \| c \|_{p,q} = \left( \sum_g \left( \sum_m |c_{g,m}|^p \right)^{q/p} \right)^{1/q} \quad (3)$$

Subsequently, the notation $(g, m)$ will be used in reference to the group-member structure, whereas $(k, j)$ refers to the time-frequency indices of the Gabor-expansion. In terms of $\ell^p,q$, we consider the cases $p = 2, q = 1$ and $p = 1, q = 2$. The former is known as Group-Lasso (GL) \cite{6} (promoting sparsity in groups and diversity in members) and the latter was termed Elitist-Lasso (EL) in \cite{3}; the

This work was supported by the Austrian Science Fund (FWF) project LOCATIF(T384-N13) and the WWTF project Audio-Miner (MA09-024)
\(\ell^{1,2}\) constraint promotes sparsity in members, only the “strongest” members (relative to an average) of each group are retained. Landweber iterations, which solve (2) in the \(\ell^{1}\)-case, [4], also yield a solution to the generalized minimization problem induced by (3), if standard soft thresholding is replaced by a generalized thresholding operator \(S_{\lambda, \xi}(z_{g,m}) = z_{g,m} (1 - \xi(z))\). Here, \(\xi = \xi_{(g,m), \lambda}\) is a non-negative function dependent on the index \((g,m)\) and \(\lambda\). The solution to (2) is then given by the iterative Landweber algorithm: choosing arbitrary \(c^{0}\), set
\[
c^{n+1} = S_{\lambda, \xi}(c^{n} - \Phi^{*}(y - \Phi c^{n})).
\] (4)

It was shown in [7], that the use of the thresholding operators \(S_{\lambda, \xi}\), defined via \(\xi\), leads to convergence of the iterative sequence (4) to the minimizer of (2):
\[
p = 1, q = 1 : \xi^{\ell}(c_{g,m}) = \frac{\lambda}{|c_{g,m}|} \quad \text{(Lasso)}
\] (5)
\[
p = 2, q = 1 : \xi^{GL}(c_{g,m}) = \frac{\lambda}{\sum_{|c_{g,m}|^2}^2} \quad \text{(GL)}
\] (6)
\[
p = 1, q = 2 : \xi^{EL}(c_{g,m}) = \frac{\lambda}{1 + M_{g,\lambda} |c_{g,m}|} \quad \text{(EL)}
\] (7)

where \(c_{g} = (c_{g,1}, \ldots, c_{g,M})\) and \(\{c_{g,m}\}_{m}\) denotes for each group \(g\) the sequence of coefficients in the group \((c_{g,1}, \ldots, c_{g,M})\).

### 2.2. Refining the algorithms

To exploit structures in audio signals, like persistence in time or frequency, we refine the shrinking operators introduced above for application in audio analysis. The coefficient \(c_{g,m}\) (or groups of them) undergo shrinkage according to the energy of a time-frequency neighborhood. In contrast to the groups of GL and EL, the neighborhoods can be modeled flexibly, e.g., using weighting and overlap. Hence, we compose \(\xi\) with some neighborhood weighting functional \(\eta_{N}\):

To an index \(\gamma = (g, m)\) in a structured index set \(I\), we associate a (weighted) neighborhood \(N(\gamma) = \{\gamma' \in I : w(\gamma') \neq 0\}\) with weights \(w_{r}\) defined on \(I\) such that \(w(\gamma) > 0, w(\gamma') \geq 0\) for all \(\gamma' \in I\) and \(\sum_{\gamma' \in N(\gamma)} w(\gamma')^{2} = 1\). Then, with \(\eta_{N}(c_{\gamma}) = \left(\sum_{\gamma' \in N(\gamma)} w(\gamma')^{2} |c_{\gamma'}|^{2}\right)^{1/2}\), we obtain the generalized shrinking operators by setting\(^2\)
\[
\xi^{WGL} = \xi^{c} \circ \eta_{N} \quad \text{(windowed GL (WGL)),}
\xi^{PEL} = \xi^{EL} \circ \eta_{N} \quad \text{(persistent EL (PEL)),}
\xi^{PGL} = \xi^{GL} \circ \eta_{N} \quad \text{(persistent GL (PGL))}
\]
in (5)-(7). These generalized shrinking operators are not associated to a simple convex penalty functional, cp. [3]. Convergence properties of their Landweber-iterations are currently under study, and numerical experiments suggest convergence.

\(^{2}\)Cp. [7] for a more involved, but exact definition of the \(M_{g}\) in EL.

### 3. SIMULATIONS

The generalized shrinking operators were implemented in MATLAB with the following parameterization of the neighborhoods: For each time-frequency-index \((k, j)\) and a neighborhood size vector \(\sigma = (\sigma_{1}, \ldots, \sigma_{4})\), the neighborhood \(N_{\sigma}\) is defined as the set of indices \(N_{\sigma}(k, j) = \{ (k', j') : k' \in [k - \sigma_{4}, k + \sigma_{4}], j' \in [j - \sigma_{3}, j + \sigma_{3}] \}\). Neighborhoods of indices close to a border of the time-frequency plane are obtained by mirroring the index set at the respective border. Rectangular and triangular weighting of the neighborhoods was implemented, with rectangular weighting only in section 3.1 and 3.2. In the plots, an index after the operator’s abbreviation specifies the group-label as time or frequency (not needed for Lasso and WGL), e.g., PEL-t signifies that the group in the respective elitist lasso is time. For the neighborhood-smoothed operators WGL, PEL, and PGL the neighborhood-size vector \(\sigma\) is given. To test the obtained variety of shrinking operators, we used a simulated “toy”-signal consisting of a stationary, a transient and a noise part.\(^3\) The stationary part consists of four harmonics with fundamental frequency 400 Hz and decreasing amplitudes. The obtained harmonics were shaped by a linear envelope in attack and decay. The transient part was simulated by 4 equidistant impulses with similarly decaying amplitudes. Finally, Gaussian white noise with SNR about 15 and 3dB was added.

Landweber iterations are known to converge very slowly and various methods of acceleration have been proposed [8]. As it was out of the scope of this paper to elaborate on these ideas, we used the basic iteration scheme (4). The iterations presented in the following were stopped after 100 steps. Then almost all of the final relative iteration errors were below 0.3%.

### 3.1. Structured denoising

As a first experiment the standard de-noising problem with additive Gaussian white noise was considered. We use a tight Gabo-frame with Hann window of length 1024 and hop size 256 (at sampling rate 44100 Hz). We measure the operator’s performance in SNR: with the estimation’s approximate Landweber-limit \(c^{*}\) of (4) and \(\hat{s} = \Phi c^{*}\), the SNR is \(\text{SNR}(\hat{s}, s) = 10 \log_{10} \left( \frac{\|\Phi_{\hat{s}}\|^{2}}{\|\Phi_{s}\|^{2}} \right)\). For comparison, the SNR is then plotted against the number of positive coefficients. Of the variety of possible operators, Fig.

\(^{3}\)Corresponding sound files and more detailed visualizations are presented at the conference and on the webpage http://homepage.univie.ac.at/monika.doerfler/strucAudio.
3.2. Multilayer decomposition

We continue processing the toy-example by aiming to extract the signal’s tonal and transient parts at the lower noise level (15dB SNR). For estimating the tonal layer, we use a tight Gabor frame with Hann window, window-length 4096 and hop size 1024. The transients are estimated starting from the transient layer + noise (which corresponds to the unrealistic but complexity reducing assumption of perfect tonal estimation) using short windows (256 samples, hop size 64). Table 1 and 2 present the performance of a sample of operators, again of each type a “basic” and neighborhood-smoothed one. The tables show the maximum SNR value of the estimation measured w.r.t. the “true” respective layer and the corresponding percentage of retained coefficients.

Concerning the tonal estimation displayed in Table 1, WGL, with neighborhoods expanding in frequency, performs best in terms of SNR of the estimation of the transient layer. This result is not very surprising since the example’s transient layer has a simple structure which supports the performance of GL. However, GL-t should be a good choice for transient extraction in more complex signals, since it yields broadband transients without extracting many horizontal (tonal) signal parts.

The next experiment addressed the decomposition of musical audio without the presence of quasi-ground truth. For this “real-life” application, the choice of sparsity level $\lambda$ is always a difficult task. We chose the SNR-maximizing candidates from the simulations.

This yielded WGL (with rectangularly weighted neighborhoods extending 4 elements in each direction of time) as estimator of the tonal layer with sparsity level $\lambda = 0.072$, and GL (with the time-index as group label) for the transient layer with $\lambda = 0.072$. We used a 5 seconds excerpt of a Jazz-record containing piano, double-bass and drums. In the decomposition, the drums (and some percussive elements of the bass) are well separated from the harmonics of piano and bass. Using GL as transient estimator works well in this example, it captures all of the soft 16th notes drum-patterns. We observed a trade-off in the choice of the sparsity level: increasing sparsity in the tonal estimation improves the separation of both layers but leads to increased damping of higher, low-energy partials of the tonal part.

3.3. Shapes

As described at the beginning of this section, the neighborhoods’ shapes (constituted by size and weighting) were implemented and parametrized in a straight-forward fashion, so far allowing for rectangular domains with either uniform (i.e. rectangular) or triangular...
Table 1: Comparison of the performance of different operators in tonal estimation: maximum SNR values of estimation and “true” layer and corresponding number of retained coefficients in percent. * refers to neighborhoods (0, 0, 0) while † refers to (0, 0, 0, 0).

<table>
<thead>
<tr>
<th>Operator</th>
<th>Lasso</th>
<th>WGL+</th>
<th>GL-†</th>
<th>PGL+†</th>
<th>EL-†</th>
<th>PEL†</th>
</tr>
</thead>
<tbody>
<tr>
<td>max. SNR</td>
<td>28.7</td>
<td>31.2</td>
<td>30.5</td>
<td>30.7</td>
<td>26.2</td>
<td>30.6</td>
</tr>
<tr>
<td>%Coeffs</td>
<td>0.4</td>
<td>1.1</td>
<td>3.3</td>
<td>17.0</td>
<td>2.8</td>
<td>4.1</td>
</tr>
</tbody>
</table>

Table 2: Transient estimation: maximum SNR values of estimation and “true” layer and corresponding number of retained coefficients in percent. As above: * means (4, 0, 0, 0) and † means (0, 4, 0, 4).

<table>
<thead>
<tr>
<th>Operator</th>
<th>Lasso</th>
<th>WGL+</th>
<th>GL-†</th>
<th>PGL-†</th>
<th>EL-†</th>
<th>PEL†</th>
</tr>
</thead>
<tbody>
<tr>
<td>max. SNR</td>
<td>10.4</td>
<td>13.2</td>
<td>14.4</td>
<td>9.5</td>
<td>10.4</td>
<td>13.3</td>
</tr>
<tr>
<td>%Coeffs</td>
<td>1.0</td>
<td>2.9</td>
<td>2.2</td>
<td>38.9</td>
<td>1.4</td>
<td>3.7</td>
</tr>
</tbody>
</table>

(i.e. “tent”-like) weightings. These shapes do not necessarily have to be symmetric at the origin, as the energy of most audio signals is not symmetrically distributed around its peaks either. This fact can be exploited to feature different parts of a signal under observation. Consider Figure 3, where the iterated WGL-shrinkage results with four different neighborhood-shapes, each solely extending in time, are compared (based on a Gabor-frame with window length 1024 and overlap of 4). It is obvious that the shapes yield different (sparse) perspectives on the signal content. Whereas the symmetric neighborhoods naturally captures parts before and after the attacks (or rather time-points of maximum energy), the asymmetric ones rather retain components before (resp. after) the attacks. The orientation of the neighborhood therefore systematically promotes the preservation of different temporal segments of the signal.

4. SUMMARY AND PERSPECTIVES

We presented first results on structured sparsity approaches for Gabor frames to audio signals. Future work will focus on the convergence of the algorithms, both in a theoretical and computational setting. By taking into account methods as [9] the proposed algorithms should be accelerated significantly. On the contrary, evaluations of the algorithms’ perceptual qualities will be considered. Further, using various shapes for the weight, we aim at the extraction of more specific structures, in the sense of sound objects [10].

5. ACKNOWLEDGMENTS

We thank Matthieu Kowalski and the anonymous reviewers for their valuable advice.

6. REFERENCES


Persistent Time-Frequency Shrinkage for Audio Denoising

KAI SIEDENBURG, AES Student Member
(kai.siedenburg@ofai.at)
Austrian Research Institute for Artificial Intelligence, Vienna

AND MONIKA DÖRFLER
(monika.doerfler@univie.ac.at)
Numerical Harmonic Analysis Group, Faculty of Mathematics, University of Vienna

Meaningful audio signals are known to be highly structured. Incorporating knowledge about the inherent structures helps to improve denoising algorithms. In this article, audio denoising is addressed as a problem of structured sparse atomic decomposition. A generalized class of time-frequency shrinkage operators is introduced, which generalizes some well-known thresholding operators, such as the empirical Wiener filter and basis pursuit denoising. The general framework allows for the exploitation of structural properties, in particular the persistence inherent to most natural audio signals. Fast iterative shrinkage algorithms are reviewed and their convergence is numerically evaluated. The denoising performance of the proposed persistent shrinkage operators is evaluated on real-life audio signals. The novel approach shows competitive performance to the state of the art when evaluated by means of signal to noise ratio and improves existing methods in terms of preliminary perceptual evaluations.

0 Introduction

In audio processing today, signals are commonly interpreted by means of their expansion coefficients with respect to basis functions taken from so-called dictionaries. In particular, dictionaries comprised of windowed Fourier or cosine bases turn out to provide valuable representations: they are easy to interpret and reflect physical reality by expanding a signal with respect to the dimensions time and frequency. Over the years, time-frequency dictionaries also have proven to be well adapted for processing most audio signals of relevance for humans, in particular speech and music. A dictionary may be understood as being well adapted to a class of signals, if it allows for sparse representations, i.e. using few atoms from the dictionary, cp. [1, 2]. Today, sparsity in redundant dictionaries is a forceful paradigm in signal processing. While intuitively, an $\ell^0$-prior on the synthesis coefficients should yield maximum sparsity, the $\ell^1$-norm as a prior on the representation coefficients plays a central role in modeling sparsity with respect to time-frequency coefficients.

Most natural signals feature strong inherent structures, such as formants, harmonics, but also transient components. In order to exploit the a priori knowledge about these inherent structures, a growing body of research has recently focused on extending existing sparsity paradigms. The desire to take structural information into account has lead to approaches of structured sparsity, see e.g. [2] and references therein. In the audio-context, structure is often interpreted in the sense of temporal and spectral persistence. In fact, a consequence of basic acoustic laws concerning resonant systems and impact sounds is that large classes of audio components are either sparse in frequency and persistent in time or sparse in time and persistent in frequency, cp. [3].

In this contribution, denoising is considered as a problem of structured sparse approximation. Therefore, we utilize the formal framework introduced in [4]. As opposed to [4], where different mixed norm priors where applied to various applications, the current paper considers the case of weighted $\ell^1$-regularization, for which an accelerated algorithm is presented and a variety of refined, weight-dependent thresholding operators are evaluated for the denoising task in depth. The proposed framework allows to adapt the neighborhood-weighting according to global sig-
nal properties and the current work addresses the influence of the chosen neighborhood-weighting on the denoising algorithm’s performance.

The paper first gives a brief review of the state of the art, then the framework of sparse expansions using Gabor frames is developed in Section 2.1, the signal model and its link to soft thresholding algorithms is established in Section 2.2 and a generalized viewpoint on soft-thresholding is introduced in Section 2.3. In Section 3, we turn to numerical aspects of the proposed denoising scheme. In Section 3.1, convergence of the algorithms is evaluated and in Section 3.2 the properties of the neighborhood-persistent thresholding operators are investigated. A comparison to existing algorithms is provided in Section 3.3 using a variety of signals, both in terms of SNR as well as from a preliminary perceptual point of view. Perceptual aspects of the evaluation are discussed in Section 4 and the paper is concluded by a summary and some perspectives in Section 5.

1 State of the Art

In time-frequency based audio denoising, it is insightful to distinguish between diagonal and non-diagonal estimation. In diagonal denoising algorithms, the attenuation factor for each time-frequency coefficient is determined independently. Diagonal denoising procedures such as the empirical Wiener estimator and other power subtraction estimators, typically ignore the correlation between coefficients, cf. [5, 6]. Approaches of the diagonal type typically suffer of what is now known as musical noise, perceptually annoying isolated noise-residuals. Non-diagonal estimation then serves as a means to avoid these severe artifacts. Ephraim and Malah [7] were the first to apply time recursive filtering for SNR estimation, introducing some coupling between coefficients, which was equipped with different attenuation rules and was variously refined later, cf. [8, 9]. Both diagonal and non-diagonal denoising techniques have been carefully reviewed in [9]. In the same paper, the authors introduced a denoising algorithm by time-frequency block thresholding. The latter has shown to improve performance significantly, and can now be considered as the state of the art in non-diagonal audio denoising.

Since the pioneering work of Donoho and Johnstone on wavelet thresholding in statistics [10, 11], sparsity based approaches have also lead to significant contributions in signal processing and audio denoising. A particularly successful model of sparsity is based on the $\ell^1$-norm regularization (to be thought of as a convex relaxation of the $\ell^0$ norm) of the signal expansion. It was introduced by Chen et al. in signal processing [12] as basis pursuit denoising and in an equivalent approach by Tibshirani as Lasso in statistics [13]. Later, it was shown that solutions of $\ell^1$-regularized inverse problems are in many cases optimally sparse [14] and fast iterative algorithms have been proposed [15, 16, 17], following the initial theoretical work of Daubechies, Defrise and de Mol [18]. However, by treating each coefficient independently these initial sparse models as the Lasso again give rise to diagonal denoising procedures.

Aiming at a more comprehensive mathematical model for the dependencies between time-frequency coefficients, a sparse regression approach with structured priors was suggested for audio denoising in [19]. With similar motivation, Kowalski and Torresani [20] extended the $\ell^1$-regularization method by introducing mixed-norm priors and neighborhood weighting on the coefficients. Each coefficient is thresholded according to the weight of its neighborhood, hence yielding non-diagonal shrinkage operators. Solutions to the mixed-norm regularized inverse problems were shown to be obtained by generalized classes of soft-thresholding operators in [21]. This approach was revisited and refined in [4] and demonstrated promising results for a variety of classical signal processing tasks such as multi-layer-decomposition, transient estimation and denoising. In [22] it was evaluated for audio denoising in particular. The convergence properties of neighborhood system based shrinkage were investigated in [23]. Following the same idea of persistent estimation, a neighborhood-based empirical Wiener estimator was introduced and discussed in [24].

2 Denoising by Time-Frequency Shrinkage

Time-frequency methods are very common in audio signal processing, the best-known transform is the short-time Fourier transform (STFT) or sampled sliding window transform. We consider dictionaries known as Gabor frames, which, in their simplest instantiation, correspond to an STFT.

2.1 Signal Expansions using Gabor Frames

We wish to expand a signal $y \in \mathbb{R}^L$ as a linear combination of time-frequency or Gabor atoms $\phi_{k,j}$, i.e.

$$y(n) = \sum_{k,j} c_{k,j} \phi_{k,j}(n) + e(n), \quad n = 1, \ldots, L.$$ 

The atoms $\phi_{k,j}$ are generated by time-frequency shifts of a single window function, $\phi_{k,j} = M_{j,b} \phi_k$. Here, $T_x \phi(n) = \phi(n-x)$ and $M_{a,b} \phi(n) = \phi(n) e^{-2\pi i a b n}$ denote time- and frequency-shift operators and $\phi$ is a standard window function. $a$ and $b$ are the time- and frequency sampling constants, $j = 0, \ldots, J-1, k = 0, \ldots, K-1$, with $K a = J b = L$, and the complex numbers $c_{k,j}$ are the expansion coefficients. In order to guarantee perfect and stable reconstruction of a signal from its associated analysis coefficients $c_{k,j} = \langle y, \phi_{k,j} \rangle$, we always assume that the dictionary $\Phi = (\phi_{k,j})_{k,j}$ forms a frame, see [25], which, in the finite-dimensional case, means that $\Phi$, the matrix of dimension $L \times p$, with $p = JK \geq L$, whose columns consist of the atoms $\phi_{k,j}$, has full rank $L$.

We will assume even more, namely the tightness of the frames in use, which means that, up to a constant, we obtain $y = \sum_{j} \langle y, \phi_{k,j} \rangle \phi_{k,j}$, i.e., perfect reconstruction is achieved by using the same window $\phi$ for both analysis and synthesis. Tight frames are easily calculated in many situations of practical relevance, see e.g. [26]. Tight windows further
simplify the interpretation of operations like thresholding in the time-frequency domain, as they introduce a symmetric relation between analysis and synthesis.

As we are especially interested in the redundant Gabor-frame case \( L < p \), the additional degrees of freedom can be used to promote (structured) sparsity of the coefficients.

### 2.2 Sparse Approximation and Iterated Soft-Thresholding

In our basic signal model \( y = \Phi c + e \), the coefficients \( c \in \mathbb{C}^p \) are assumed to be sparsely distributed. Then, denoising the observation \( y \) means to sparsely approximate the coefficients \( c \). This implies the recovery of \( \Phi c \), the clean signal. This approach is formalized via the minimization problem

\[
\min_{c \in \mathbb{C}^p} \left\{ \frac{1}{2} \| y - \Phi c \|_2^2 + \lambda \| c \|_1 \right\}
\]

which is the well-known \( \ell^1 \) sparse minimization functional. In statistics, this problem is called Lasso [13], in signal processing basis pursuit denoising [12]. The constant \( \lambda > 0 \) is the so-called Lagrange-multiplier. It adjusts the weight given to the \( \ell^1 \) penalty term and hence adjusts the sparsity level of the solution. The higher the value of \( \lambda \), the sparser the solution will be. Note that, depending on the relation of the noise variance and \( \lambda \), the solution of (1) might not coincide with the true underlying sparse representation. In practice, finding a good sparsity level is consequently most crucial for obtaining satisfying results, cf. [24].

In case \( \Phi \) is an orthonormal basis, the solution \( c^\star \) of (1) is given by a simple soft-thresholding step of the analysis coefficients (see e.g. [18]):

\[
c^\star = S_\lambda(\Phi^\ast y) \text{ where } S_\lambda(z) = \exp(i\text{arg}(z))(|z| - \lambda)^+ \text{ and as usual, } b^\ast := \max(b, 0). \Phi^\ast \text{ denotes the analysis operator, the adjoint of the synthesis operator } \Phi, \text{ given by } \Phi^\ast s = \langle (s, \Phi_k) \rangle_{k,j}.
\]

In the general case when \( \Phi \) only forms a frame, it was shown in [18] (and extended by other means in [27]) that for arbitrary starting point \( c^0 \in \mathbb{C}^p \), the iterated soft-thresholding algorithm (ISTA) converges strongly to the solution \( c^\star \) of (1):

\[
c^\star = \lim_{n \to \infty} S_\lambda(\Phi^\ast (c^n + \Phi^\ast(y - \Phi c^n)))
\]

if only \( \| \Phi \| < \sqrt{2} \).

The IST-algorithm converges very slowly, and in general not even linearly.\(^1\) Hence, various methods of acceleration have recently been proposed. It turns out that especially the Beck-Teboulle FISTA-algorithm [16] improves convergence significantly, cf. [15]. Here, we give an adaptation of the algorithm to our setting.

#### Algorithm 2.1 (FISTA)

Set \( c^0 = b^1 = 0 \) and \( t_1 = 1 \).

Do

\[
e^t = S_\lambda(b^t + \Phi^\ast(y - \Phi b^t))
\]

\[
f_{t+1} = \frac{1}{2}(1 + \sqrt{1 + 4f_t^2})
\]

\[
b^{t+1} = e^t + \left( \frac{f_t - 1}{f_{t+1}} \right)(e^t - e^{t-1})
\]

Until convergence.

While not significantly adding computational complexity, FISTA accelerates the convergence of the objective functional value quadratically [16], i.e.

\[
\frac{1}{2} \| y - \Phi c^t \|_2^2 + \lambda \| c^t \|_1 \sim \mathcal{O}(1/n^2).
\]

### 2.3 Generalized Soft-Thresholding

Since the \( \ell^1 \)-norm acts independently on each coefficient, problem (1) does not take into account the dependencies between time-frequency coefficients which are inherent in most natural audio signals. It was shown in [21] that the \( \ell^1 \)-penalty term can be replaced by a mixed norm which acts independently on groups of coefficients and their members. The solution of the corresponding minimization functional is then obtained by generalized soft-thresholding. Furthermore, overlapping neighborhood structures can be introduced to exploit persistence properties of coefficients [20].

However, for denoising of quasi-stationary noise, the neighborhood operators derived from the \( \ell^1 \)-norm (the so-called windowed group lasso) suffice. Therefore, the formal framework introduced in this paper is restrained to the \( \ell^1 \)-case. For the more general setting encompassing mixed norms, see [4].

#### Definition 2.2.
Let \( \xi = \xi_k : \mathbb{C}^p \to \mathbb{R}^+ \) be a non-negative function, the so called threshold function and let \( \Gamma \) denote the set of time-frequency indices, i.e. \( \gamma = (j, k) \in \Gamma \) for \( j = 0, \ldots, J - 1, k = 0, \ldots, K - 1 \). Then, for \( z \in \mathbb{C}^p \) the generalized thresholding operator is defined component-wise by

\[
S_{\xi}(z) = z \mathbb{1}_{1 - \frac{\xi_k(z)}{\xi_k(z)}}
\]

and we write \( S_{\xi}(z) := (S_{\xi}(z))_{\gamma \in \Gamma} \).

#### Example 2.3.
For \( \xi_k(z) = \xi_k^L(z) := \frac{\lambda}{\xi_k^L} \), we recover the usual soft-threshold operator (Lasso) from above. For \( \xi = (\xi^L)^2 \), we obtain a so-called empirical Wiener estimator, cf. [24, 25]. Setting \( \xi = \lim_{k \to \infty}(\xi_k^L)^k \), the hard thresholding operator is approached, defined by \( S_{\lambda}(z) = z \mathbb{1}_{|z| > \lambda} \) and 0 if \( |z| \leq \lambda \).

The neighborhood-systems introduced in [20] under the name Windowed Group Lasso can be written in terms of generalized thresholding by setting \( \xi = \xi_k^L \circ \eta \), where \( \eta \) is a neighborhood smoothing function. It can be defined by

\[
\eta(c_{\gamma}) := \left( \sum_{\gamma \in \Gamma} v_\gamma(\gamma')|c_{\gamma'}|^2 \right)^{1/2}
\]

where \( v \) are appropriately chosen non-negative time-frequency neighborhood weights, that is e.g.

\[
\| v_\gamma \|_1 = 1, \text{ and } v_\gamma(\gamma') > 0 \ \forall \gamma \in \Gamma.
\]
The first condition ensures that the overall extension of the neighborhood does not interfere with the sparsity of the solution and without the second, the intuition of a neighborhood would not make sense. In this kind of persistent thresholding \( \xi = \xi^k \circ \eta \), the coefficients undergo shrinkage according to the energy of a time-frequency neighborhood which can be modeled flexibly, e.g. by using weighting and overlap.

Using the same nomenclature, the persistent empirical Wiener shrinkage (PEW) introduced in [24] simply writes as \( \xi = (\xi^L)^2 \circ \eta \).

**Example 2.4.** For the neighborhood consisting of the singleton, WGL coincides with the regular Lasso. With disjoint neighborhoods we obtain the well known Group-Lasso [28] as a special case of the WGL. Furthermore, block-thresholding estimators as introduced in [29] can be expressed in this framework by using disjoint neighborhoods and appropriately chosen threshold functions.

### 3 Numerical Evaluation

The generalized shrinkage operators were implemented in MATLAB with the following parameterization of the neighborhoods: For each time-frequency-index \( \gamma = (k, j) \) \((k\) denoting time- and \( j\) frequency-index) and a neighborhood size vector \( \sigma = (\sigma_1, \ldots, \sigma_4)\), the neighborhood \( N_\sigma \) is defined as the set of indices \( N(\gamma) = N_\sigma(k, j) = \{(k', j') : k' \in \{k - \sigma_4, k + \sigma_2\}, j' \in \{j - \sigma_3, j + \sigma_1\}\} \). That is, the vector \( \sigma \) refers to the additional extension of the neighborhood in the orientation \( \sigma = \) (north, east, south, west) from the center coefficient. For instance, the size vector \( \sigma = (1, 1, 1, 1) \) therewith refers to the rectangular neighborhood containing 9 coefficients. Neighborhoods of indices close to a border of the time-frequency plane are obtained by mirroring the index set at the respective border. Figure 1 sketches different neighborhood shapes in the time-frequency plane.

The corresponding toolbox and the audio files used in the evaluation, as well as further audio-visual examples, are available at the webpage homepage.univie.ac.at/monika.doerfler/StrucAudio.html

**Figure 1.** Sketch of different neighborhood shapes in a schematic time-frequency plane.

**Figure 2.** Clean, noisy, Lasso-denoised, and WGL-denoised time-frequency coefficients of a natural audio signal (in left to right, top to bottom order). WGL is computed with neighborhood \( \sigma = (0, 4, 0, 4) \). Using the same threshold level \( \lambda = 0.01 \), the Lasso is sparser than WGL but suffers from severe musical noise.

Figure 2 gives a first overview of the basic behavior of the denoising operators.\(^2\) It shows parts of the clean, noisy, WGL-denoised and Lasso-denoised time-frequency coefficients from a natural audio signal. While using the same threshold level, the Lasso gives obviously much sparser results. On the contrary, WGL yields a regularized solution, without apparent musical noise.

For all the following simulations we employed a tight Gabor frame with Hann-window of 1024 samples length with hop-size 256 at 44100 Hz audio sampling rate.\(^3\) In this paper we only consider the case of Gaussian white noise at base noise levels of 0, 10 and 20 dB SNR\(^4\). The denoising performance of the different operators is measured via the SNR of the denoised estimation and the original, clean signal.

### 3.1 Convergence

The proposed extended class of shrinkage operators are in general not associated to a simple convex penalty functional any more. For the case of the Windowed Group Lasso the relation between a related minimization functional and the shrinkage operator is an intricate issue, cf. [23]. However, the theoretical evidence in [23] and our practical experience suggests that convergence of WGL in the FISTA framework seems to take place and is beneficial.

\(^2\)For the sake of brevity, we leave out depictions of the (P)EW operators as they yield similar visual appearances as their respective counterparts.

\(^3\)The experiments were repeated with various alternative parameter settings, in particular, windows with longer support. Since the results were very similar, we stick to one standard setting in all evaluations.

\(^4\)The signal to noise ratio of signals \( s, \tilde{s} \in \mathbb{C}^L \) is defined by

\[
10 \log_{10} \left( \frac{\sum_{n=1}^{L} |s(n)|^2}{\sum_{n=1}^{L} |s(n) - \tilde{s}(n)|^2} \right).
\]
in terms of sparsity and denoising performance. In the following, we hence use WGL always in the iterated FISTA framework. On the contrary, the empirical Wiener estimator and the persistent empirical Wiener do not have any obvious relation to an underlying minimization functional and will simply be used as non-iterated time-frequency thresholding operators.

As an example of the convergence processes Figure 3 depicts the evolution of the SNR of the sparse approximation and the clean signal as a function of the number of iteration steps for the Lasso and the Windowed Group Lasso. It is obvious that the FISTA-algorithm accelerates the convergence significantly. Figure 4 further sheds some light on the relation of approximation precision, sparsity level and required number of iteration steps. The figure shows the number of steps required to obtain SNR-tolerances of 0.5, 0.05, and 0.005 dB w.r.t. to the SNR of the 100th iteration. Note that in this example (and also all others we have encountered so far) only less than 20 iterations are required to reach a SNR-tolerance of 0.05 dB. Furthermore, the WGL seems to regularize the convergence process as it requires much less iterations to reach its stationary point. In the following simulations, we hence used FISTA with (at least) 20 iteration steps.

3.2 Selection of Neighborhoods

In this section, we investigate the influence of neighborhood support and weights. We give some numerical illustrations of the intuition, that the neighborhood’s influence depends on local signal characteristics. In particular, we illustrate the four aspects of orientation, extension, symmetry and decay.

To evaluate the operators with respect to varying signal characteristics we chose a set of three mutually contrasting signals. The first contains a rather sustained excerpt of a string quartet, the second a Latin-style piano pattern and the third a percussive conga groove, each of 2300 ms length. The samples do not contain reverber or audible room influences and are taken from a sample library of a commercially available audio-software. Gaussian white noise was added to obtain “base” noise levels of 0, 10 and 20 dB SNR for each signal. The quantitative results presented in the following are always averaged over these three noise levels.

It is quite intuitive that the denoising performance should be optimal if the orientation of the neighborhoods in time and frequency fits to the dominating persistence properties of the signal. For instance, the string-signal contains temporally persistent and spectrally sparse orientation, hence a horizontally oriented neighborhood as \( \sigma = (0, 4, 0, 4) \) seems to be suited. The contrary of course should hold for the percussion-signal where vertically oriented neighborhoods should perform best. Table 1 presents numerical evaluations concerning three different neighborhood-orientations and the three musical test signals. It shows the respective maximal SNR (over the range of relevant sparsity levels) averaged over all three different noise levels. While the vertically oriented neighborhood performs best for the sustained string signal, the vertical neighborhood works better for the percussive signal. Contrary to a first intuition, the rectangular neighborhood \( \sigma = (1, 1, 1, 1) \) performs best for the percussive signal, while being sub-optimal for the piano signal which should have both, persistence in time and frequency.

<table>
<thead>
<tr>
<th>Orientation</th>
<th>Strings</th>
<th>Piano</th>
<th>Percussion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Horizontal</td>
<td>20.2</td>
<td>19.3</td>
<td>20.8</td>
</tr>
<tr>
<td>Vertical</td>
<td>18.1</td>
<td>17.3</td>
<td>21.3</td>
</tr>
<tr>
<td>Rectangular</td>
<td>19.1</td>
<td>18.5</td>
<td>21.5</td>
</tr>
</tbody>
</table>

Table 1. Comparison of the neighborhood’s orientation w.r.t. to signal characteristics. The maximal SNRs of WGL with vertical \( \sigma = (4, 0, 4, 0) \), horizontal \( \sigma = (0, 4, 0, 4) \) and rectangular \( \sigma = (1, 1, 1, 1) \) neighborhoods are averaged over noise levels of 0, 10, and 20 dB SNR.
Secondly, we disregard orientation and only focus on the extension of the neighborhoods for horizontally oriented ones. Again, depending on signal characteristics there are optimal neighborhood lengths. Table 2 shows the corresponding numerical results. The longest neighborhood with \(\sigma = (0, 12, 0, 12)\) works best for the strings excerpt which features the most temporally persistent structures. Similarly, medium \(\sigma = (0, 8, 0, 8)\) and short \(\sigma = (0, 4, 0, 4)\) extensions optimize the SNR measure for piano and percussion phrases.

The symmetry of the energy distribution of most audio signals varies with the instruments being played and their respective modes of excitation. It was noted in [4] that non-symmetric neighborhoods can be beneficial, especially for avoiding pre-echo artifacts in tasks like estimation of a tonal signal layer. Table 3 shows the signal-dependent denoising results of the WGL operator using symmetric \(\sigma = (4, 0, 4, 0)\), 1/3-centered \(\sigma = (2, 0, 2, 0)\) and asymmetric \(\sigma = (0, 0, 0, 8)\) neighborhoods. The SNR differences are very subtle but nonetheless point in an interesting direction. While the asymmetric neighborhood is suboptimal for all three signals, the symmetric one is best for the strings (which play rather sustained notes of continuous excitation) and the 1/3-centered neighborhood performs best for the percussive signal (with fast decay after all excitation).

A last remark concerns the decay of the neighborhood weights. It seems intuitive that rectangular weighting is sub-optimal, since it does not account for continuously varying interdependence of two points in the time-frequency plane (meaning the further away, the less important). However, it is hard to evaluate this aspect, since for instance switching from rectangular to linear weighting will have a similar impact as just changing the neighborhood’s length. In the numerical experiments directed so far, non-rectangular weighting has not improved performance significantly.

In conclusion, this numerical case study confirms the intuition that neighborhood selection should ideally be adapted to the signal. While the performance differences for the parameter of symmetry as measured in SNR seems of minor relevance (which does not imply that the same holds in terms of perception), adjusting the overall shape of the neighborhood seems to have the greatest impact.

### 3.3 Comparison with Other Algorithms

In 2008, Yu, Mallat and Bacry proposed a non-diagonal denoising procedure based on time-frequency block thresholding [9]. By minimizing the corresponding Stein estimate of the risk, the algorithm automatically adapts blocks of time-frequency coefficients on which a block-wise empirical Wiener filter is applied. The authors show that the method clearly outperforms other state of the art algorithms as power subtraction [5] and Ephraim and Malah’s MMSE-LSA [7] in terms of signal difference measured in SNR and perceptual evaluation with respect to a representative group of subjects. We employ this algorithm with the same Gabor-transform settings as above and use it as a reference to the state of the art.

Using the generalized thresholding framework, we further evaluate the simple but ubiquitously used Lasso estimate with threshold function \(\hat{\xi} = \xi L\) and its quadratic counterpart \(\hat{\xi} = (\xi L)^2\) corresponding to the empirical Wiener estimator, both classical diagonal denoising operators. In terms of neighborhoods of the WGL, we choose non-symmetric ones to avoid pre-echos and evaluate the 1/3-centered horizontal support \(\sigma = \sigma_\ell = (2, 0, 2, 0)\), as well as the rather vertically oriented \(\sigma = \sigma_v = (2, 0, 2, 1)\). For both neighborhoods rectangular weighting is used. To further enhance SNR performance, we also consider their respective non-iterated squared (empirical Wiener) counterparts. The pool of test signals was complemented by a male and a female voice (2300ms length) and a 6 seconds Jazz-quintett sample containing drums, double-bass, piano, saxophone and trumpet. All test signals were corrupted with Gaussian white noise at levels of 0, 10 and 20 dB.

Figure 5 shows the maximal (over all sparsity levels) SNR-performance of the 7 different operators, averaged over the three noise levels. Obviously, the plain Lasso is constantly worst. WGL-0206 and WGL-2021 improve the estimation, where depending on the signal type one is better than the other. For rather percussive signals as the conga-percussion part and the speech signals, the vertically oriented WGL is better, for the rather tonal signals, the horizontally oriented WGL improves performance of at least 1 dB. Most interestingly, Yu, Mallat, and Bacry’s block thresholding algorithm works best for the percussive signals, i.e. the conga and speech signals, while the non-iterated PEW outperforms block thresholding for predominantly tonal signals.

Besides only regarding the denoising performance as measured in SNR, we can compare the sizes of the algorithms’ recovered support, i.e. the model selection, in order to better understand their behavior. Figure 6 depicts the percentage of non-zero coefficients over different sparsity levels \(\lambda\) for the operators L, EW, WGL-0206, PEW-0206...
Let us finally note that the non-iterated PEW is much more computationally efficient than block thresholding (in our implementation around factor 6), where for each signal the rather sophisticated optimal partition of the time-frequency plane must be calculated.

### 4 Perceptual Aspects

The signal to noise ratio measures global energy difference, which generally does not coincide with perceptual distance. Therefore, this section briefly addresses the perceptual qualities of the proposed algorithms. As it was out of scope of this work to conduct representative listening tests, we confine ourselves to a brief subjective description of the algorithms qualities and a quantitative evaluation using a standardized audio-quality measure.

#### 4.1 Audible Differences and Artifacts

Although in terms of SNR very distinct, the difference between the WGL and PEW operators was inaudible for the authors (under the usage of appropriate audio equipment). Hence, the latter seems preferable for the denoising-task as it produces higher SNR without iteration. In terms of neighborhood shape, the vertically oriented WGL/PEW with \( \sigma = (2, 0, 2, 1) \) tends to produce "underwater-feeling"-like residuals at high noise levels because of its short time-persistence. As could be expected, it further tends to emphasize vertical time-frequency components, as e.g. the snare clicks in the jazz-quintet sample, and reduces some temporal smearing which becomes audible in the horizontal WGL and block thresholding as subtle "ghost-like" artifacts, in particular in speech. All three alternatives successfully avoid musical noise as well as pre-echo, WGL due to non-symmetric neighborhoods and block thresholding due to signal dependent block-adaptation.

A point supporting block thresholding is that its low-pass filter impact is not as strong as for the WGL-versions we tested, although the difference is of minor magnitude. A more serious artifact is that block-thresholding tends to produce a subtle but annoying background-texture of clicks. This phenomenon is audible at higher noise levels when sparser approximations are required (i.e. higher thresholds) such that no high frequency components can mask this texture. The artifact is probably due to the disjoint partition of the time-frequency plane into blocks and the corresponding rapid amplitude modulations in the time-frequency plane. The overlapping neighborhood systems of the WGL avoid this artifact.

---

5Note that for each thresholding step, both operators produce the same support.

---

6The SNR favors non-sparse solutions containing noise residuals of low energy (which are clearly audible, though) such that the high energy signal components, emphasized by the underlying \( \ell^2 \)-norm, are not shrunk too much.

7A selection of these test signals and the denoised results can be found at the above mentioned webpage.
Fig. 7. Comparison of the maximal PEAQ-scores of denoised and clean signal, averaged over noise levels 0, 10, 20 dB for 6 different signals. The operators are WGL with $\sigma = (0.2, 0.6)$, persistent empirical Wiener (PEW) with $\sigma = (0.2, 0.6)$ and $\sigma = (2, 0, 2, 1)$ and block thresholding (BT).

4.2 PEAQ-Scoring

In order to collect some quantitative data on the audio quality, we measured the distance of the estimation and the clean signal by means of ITU-R recommendation BS. 1387 perceptual evaluation of audio quality (PEAQ) [30], as implemented in [31]. The PEAQ measure attempts to model perceived audio quality differences by combining a number of psycho-acoustic features, based on a filter bank ear model. A so called objective difference grade is computed via a neural network featuring a quality scale from -4 (very annoying) to 0 (imperceptible).

Figure 7 shows the resulting PEAQ-evaluation, presented in the same form as the previous SNR evaluations. Again, the maximal PEAQ-values over all (this time pre-elected) sparsity values are respectively averaged over the three different noise levels per signal. The results clearly favor the persistent shrinkage, in particular the vertical PEW. Only for the female voice and the jazz-quintet sample, the advantage over the block thresholding algorithm becomes smaller.

The PEAQ-standard may not be accepted uncritically, cp. [31]. The assessment that the vertically oriented PEW averagely “sounds best” for all different test signals is not at all confirmed by the authors’ subjective judgements, for instance. Nonetheless, the main direction of perceptually favoring the neighborhood-smoothed WGL/PEW over block-thresholding seems plausible, mainly because of the severe click-artifacts of the latter, see Section 4.1 above.

5 Conclusion

This paper considered the denoising problem from the perspective of sparse atomic representation. A general framework of time-frequency soft-thresholding was proposed which encompasses and connects well-known shrinkage operators as special cases. In particular, it was demonstrated how neighborhood-persistent threshold operators can be efficiently approximated and how a signal adaptive choice of parameters can improve estimation and denoising quality. With respect to the denoising task it was shown that simple non-iterated operators derived from the generalized thresholding scheme perform competitively compared to cutting edge methods, as evaluated in SNR, and are computationally less complex. In terms of perceived audio quality informal description as well as standardized audio quality evaluation methods showed clear preferences for this novel method of time-frequency thresholding.

The presented $l^1$-framework is a special case of the mixed norm setting, which was already used for modeling inter-channel dependencies in multi-channel audio [20]. By enhancing the mixed norm setting with the insights gained in this paper, further improvements of multi-channel denoising algorithms can be expected.

In this paper, the issue of adapting the underlying time-frequency transform stayed untouched. In fact, using persistent thresholding in combination with non-stationary or multi-frame expansions seems to bear great potential for structured denoising. Such modified approaches, as well as more comprehensive comparisons with other denoising methods (as used in speech enhancement) will be addressed in future work.

Finally, it is to be noted that the basic idea discussed in this paper of supporting persistent structures in time-frequency representations should also bear potential for other branches of sparse representations such as dictionary learning, audio coding and source separation.

6 Acknowledgements

This work was supported by the Austrian Science Fund (FWF) : T384-N13 and the WWTF project Audio-Miner (MA09-024).

7 REFERENCES


THE AUTHORS

Kai Siedenburg studied Mathematics and Musicology, receiving his Dipl.Math (M.Sc.) degree from Humboldt University Berlin. In 2008/09 he visited the University of California, Berkeley and its Center For New Music and Audio Technologies on a Fulbright scholarship. Throughout his studies Kai was stipendiary of the German National Academic Merit Foundation. Currently, he works as researcher at the Austrian Research Institute for Artificial Intelligence.

Monika Dörfler obtained her PhD in Mathematics from the University of Vienna and is a researcher at the Faculty of Mathematics of the University of Vienna. She studied piano at the Music University of Vienna and is working in the field of applied mathematics for audio signal processing. She is interested in the interplay of local and global aspects of time-frequency analysis, and focuses on the benefits of theoretical results in practical applications. She is heading the interdisciplinary research project AudioMiner and has been a Hertha Firnberg fellow of the Austrian science fund FWF.