universität wien

DISSERTATION

Titel der Dissertation

"Linear Time Variant Systems and Gabor Riesz Bases"

verfasst von

Dipl. Ing. Nina Engelputzeder

angestrebter akademischer Grad

Doktorin der Naturwissenschaften (Dr.rer.nat)

Wien, 2015

Studienkennzahl lt. Studienblatt:	A 091 405
Dissertationsgebiet lt. Studienblatt:	Mathematik
Betreuer:	ao. Univ.-Prof. Dr. Hans Georg Feichtinger

Contents

Acknowledgments

Firstly, I would like to express my sincere gratitude to my advisor Prof. Dr. Hans Georg Feichtinger for the continuous support of my Ph.D study and related research, for his patience, motivation, immense knowledge and insight. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my Ph.D study and the map of function spaces will stay forever in my mind.

Besides my advisor, I would like to particularly thank Dr. Peter Balazs for his continuing support, inspiring discussions and for giving me the opportunity to spend a wonderful time joining his team and contributing to the outstanding research work and scientific dialogue of the Acoustic Research institute. This strong support and motivation in combination with Prof. Feichtingers patience and vision were finally the decisive factors making it possible for me to finish this work.

Then I also want to thank Dr. Gerald Matz, the third representative of my thesis committee and all the broader Nuhag and ARI family for their insightful comments and encouragement, but also for the hard question which incented me to widen my research from various perspectives. Also, I thank my fellow IK colleagues in for the stimulating discussions and the good time we spent together.

I want to thank the two most important persons, both entering my life during the period of this work - my spouse Peter and my beloved daughter Malina Victoria who always supported me mentally in good and not so good times and by stepping back when required in terms of time and attention.

Then, I want to thank to my friends for support and friendship as well as for always listening and keeping me going.

Last but not the least, I would like to thank my parents and my brother for supporting me spiritually throughout writing this thesis and my my life in general.

Abstract

This thesis can be seen as a contribution to *application oriented harmonic analysis*, as it was specifically promoted by H.G.Feichtinger since the 80th. [48]. Good overviews of theory and applications can be found e.g. in [60], [59] and [65]. Once the developed tools at hand the underlying ideas of simultaneously considering time and frequency information, have expanded rapidly to wide spread applications from musics [41], acoustics [92], mobile communication [112], engineering [75] , [95], neuroscience and abstract mathematics [6].

Embedded in this context, the first part of the work is devoted to linear time variant systems (LTV). Such systems appear in most applied fields from geophysics over electrical engineering, acoustics to economics. Applications may concern the design of digital filters for acoustical denoising problems or or a potential divider in electrical engineering . In economics one may think of describing an economies internal structure to project a future outcome as the economic response to a tax increase or the reaction of financial markets to interest rate cuts.

The correct mathematical framework to analyze LTV systems is operator theory. For functions and distributions time-frequency expansions are well established and already in use in many applications. The time frequency representation of operators is the next logical step, which is under investigation [45] [68] but has still many challenging questions. As a particularly important class of operators (LTV systems) the structure of Gabor multipliers is studied in this thesis. Gabor multipliers are operators which are generated through decomposition of a function through a Gabor frame, pointwise multiplication with a weight function (mask) and reconstruction via the dual frame. Besides constituting an intuitive way of time variant filtering, they are of special importance as most operators coming up in engineering applications can be approximated well with Gabor multipliers [42] [69]. There are two key research question motivated by numerical filtering applications in acoustics which are treated in the first part of this thesis.

First, it currently seems to be common practice in the acoustics research community to use Gabor multipliers for the purpose of near time implementations of linear time invariant (LTI) filters. It will be shown that from an analytical point of view this operators are not exactly identical even though in most cases results seem very similar. The error will be analysed and error estimates will be given. Moreover suited parameter combinations will be deduced and put at hand to practitioners in order to guarantee sufficient approximation quality for applications. In this manner heuristic practice is put on a thorough analytical basis and it is shown in which context the usage of Gabor multipliers for the approximation of LTI filters is appropriate and when approximation quality is insufficient.

Second, the interconnection between the finite discrete Gabor multiplier as implemented in applications and the continuous infinite dimensional Short-Time-Fourier-Transform multiplier - for which abstract mathematical results are valid - is investigated. First abstract results have been reached in [13], results on the sampling error in the finite setting can be found in [108]. In this work, as link in between, the concrete quantitative and qualitative behavior of the sampling and discretization error is investigated. Besides time frequency methods, tools from the field of numerics and in particular approximation theory will be used. This is possible due to the fact that by representing the operator via the Kohn Nierenberg symbol the problem is equivalent to approximation in shift invariant spaces [51].

The second large part of the thesis project deals with the design of Gabor Riesz bases for mobile communication applications. The underlying theoretical problem is motivated in a wider setting by the practical problem of signal transmission in wireless communication. In particular the part of signal modulation is considered.

In applications, when information in form of data symbols is transmitted through a channel, these bits of information are carried by a structured system of pulses (for classical transmission in frequency domain these pulses are just rectangular functions). Mathematically speaking the signal is decomposed with respect to a Gabor Riesz basis, then the operator is applied followed by a resynthesis mapping. An important question is how to choose this Gabor Riesz basis in order to satisfy the aim of practitioners for spectral efficiency as well as the aim for (almost) perfect identification after the channel. In particular two designs, a special structure of a multi pulse Gabor Riesz basis and the Hermite functions as basis for the transmission as suggested in [70], are compared with respect to spectral efficiency, numerical efficiency and stability of results. Such designs can be used for Cyclic prefix orthogonal frequency-division multiplexing (CP-OFDM) [97] [106] [118] [29], which is a multi-carrier scheme being used or proposed for numerous wireless standards like WLAN (IEEE 802.11a/g/n, HIPERLAN/2), broadband wireless access (IEEE 802.16), wireless personal area networks (IEEE 802.15), and digital audio and video broadcasting (DAB, DRM, DVB-T).

The work on this thesis was partly supported by the Austrian Science Fund (FWF) START-project FLAME ('Frames and Linear Operators for Acoustical Modeling and Parameter Estimation'; Y 551-N13) and the WWTF project MULAC ('Frame Multipliers: Theory and Application in Acoustics', MA07-025).

All MATLAB code developed and used for this thesis can be found on the enclosed CD.

Zusammenfassung

Die vorliegende Arbeit bettet sich in das Forschungsgebiet der numerischen harmonischen Analyse ein. Das zentrale Thema der Arbeit stellen dabei zeitvariante Systeme und Gabor Riesz Basen dar.

Der erste Teil der Arbeit beschäftigt sich mit Gabor Multiplikatoren als besonders wichtige Klasse der linearen zeitvarianten Systeme. Dabei werden im Speziellen zwei Forschungsfragen näher beleuchtet. Zum einen, scheint es gängige Praxis zu sein, zeitinvariante Filter in Echtzeit mit Hilfe von Gabor Multiplikatoren zu implementieren. In diesem Zusammenhang wird gezeigt, dass trotz der sehr guten Resultate in den meisten Anwendungen, analytisch gesehen die Operatoren bis auf den trivialen Fall nicht ident sind. Der Fehler wird analysiert und konkrete Fehlerabschätzungen berechnet. Geeignete Parameterkombinationen in Abhängigkeit der vom Anwender gewünschten Genauigkeit der Näherung werden abgeleitet. Die zweite Frage beschäftigt sich mit dem Zusammenhang zwischen endlichen, diskreten Gabor Multiplikatoren und stetigen, unendlich dimensionalen Kurzzeitfouriertransformations-Multiplikatoren. Dazu gibt es bereits abstrakte Ergebnisse in [13] und Ergebnisse ber den Abtastfehler finden sich in [108]. In der vorliegenden Arbeit, wird ergänzend dazu das konkrete quantitative und qualitative Verhalten des Abtast- und Diskretisierungsfehlers untersucht.

Der zweite Teil der Arbeit beschäftigt sich mit Gabor Riesz Basen. Solche Systeme linear unabhängiger Elemente werden unter anderem häufig im Bereich der w-lan Übertragung eingesetzt. Wenn ein Signal als Linearkombination aus Elementen ein Gabor Riesz Basis dargestellt wird - wobei die Koeffizienten die zu übertragende Information beinhalten - und diese Elemente gute approximative Eigenvektoren eines zeitvarianten Übertragungskanals sind - ist es möglich die Information übermittelte Information mittels des biorthogonalen Systems, welches wiederum als Gabor System angenommen werden kann, zu identifizieren. Da in der Konstruktion einer derartigen Gabor Riesz Basis eine Vielzahl freier Parametern beinhaltet, spielt die Frage wie diese zu wählen sind, um dem Anspruch nach einer möglichst hohen Übertragungsrate zu genügen ohne dabei die Stabilität und damit das Gesamtergebnis der Übertragung außer Acht zu lassen.

Im Speziellen werden dazu zwei mögliche Konstruktionen näher untersucht: eine spezielle Struktur einer Multi Puls Gabor Riesz Basis und die Hermite Basis als Element des Übertragungs-Schemas (vgl [70]). Vergleiche an Hand von Übertragungseffizienz, numerischer Effizienz und Stabilität der Resultate werden angestellt.

Alle MATLAB Programme, die im Rahmen der vorliegenden Arbeit entwickelt und verwendet wurden, finden sich auf der beigelegten CD.

Die Arbeit an dieser Dissertation wurde zum Teil unterstützt vom Austrian Science Fund (FWF) START- Projekt FLAME ('Frames and Linear Operators for Acoustical Modeling and Parameter Estimation'; Y 551-N13) und dem WWTF Projekt MULAC ('Frame Multipliers: Theory and Application in Acoustics', MA07-025).

CHAPTER 1

Preliminaries

This chapter is intended to give a short overview over the basic principles and definitions needed in the following work. It covers the fundamental principles and summarizes the used notations in a very short manner. More thorough introductions of the various topics can be found in the quoted literature.

1.1. Fourier Analysis

This section is intended as a short reminder and summary of the theorems used from classic Fourier analysis. For a general overview, more details and proofs, see [65]. Classical Fourier Analysis dates back to the 18^{th} century and has its roots in analyzing the frequency spectrum of periodic functions: functions for that $f(x) = f(x+k)$ for all $k \in \mathbb{Z}^d$ by the means of so called Fourier series. All the following statements and proofs are standard and can be found in the respective literature, for example [81] or [109].With $L^2(\mathbb{R}^d)$ following the standard definition of the space of square integrable functions i.e. all measurable functions f for which $\int_{x\in\mathbb{R}^d} |f(x)|^2 dx < \infty$, such a function is uniquely determined by its restriction to the cube i.e. $[0,1)^d$

Proposition 1.1. *Let $f \in L^2(\mathbb{R}^d)$ be periodic and let*

$$\hat{f}(n) = \int_{[0,1]^d} f(x) e^{-2\pi i n x} dx$$

be the n-th Fourier coefficient. Then f can be expanded into the Fourier series $f = \sum_{n\in\mathbb{Z}^d} \hat{f}(n) e^{2\pi i n x}$ with convergence as an orthonormal expansion and we have

$$\int_{[0,1]^d} |f(x)|^2 dx = \sum_{n\in\mathbb{Z}^d} |\hat{f}(n)|^2. \tag{1.1}$$

Futher this concept has also been extended to non periodic functions by defining the Fourier by the integral over $\mathbb{R}^d$.

Definition 1.2. The Fourier transform of a function $f \in L^1(\mathbb{R}^d)$ is defined as $\hat{f}(\omega) = \int_{\mathbb{R}^d} f(x) e^{-2\pi i x \omega} dx$. We also write $\mathcal{F}f$ instead of $\hat{f}$ for referring to the linear operator.

Throughout the literature different normalizations can be found; in this work we stick to this one as it has the advantage of the factor 2π neither occurring in Plancherel's theorem, nor in Poisson summation formula. It can be seen immediately by $||\hat{f}||_\infty \leq ||f||_1$ and taking the absolute values in Definition (1.2) that the Fourier transform is defined for $f \in L^1(\mathbb{R}^d)$. This observation is refined by the *Riemann-Lebesgue* Lemma indicating that if $f \in L^1(\mathbb{R}^d)$, then $\hat{f}$ is uniformly continuous and $\lim_{\omega\to\infty} |\hat{f}(\omega)| = 0$. When giving up the requirement of a pointwise definition, the Fourier tranform can also be extended to other function spaces.

Theorem 1.3. *(Plancherel) If $f \in L^1 \cap L^2(\mathbb{R}^d)$ then $||f||_2 = ||\hat{f}||_2$. As a consequence the Fourier tranform extends to a unitary operator on $L^2(\mathbb{R}^d)$ and satisfies Parseval's formula $\langle f, g\rangle = \langle \hat{f}, \hat{g}\rangle$ for all $f, g \in L^2(\mathbb{R}^d)$.*

Theorem 1.4. *(Hausdorff-Young) Let $1 \leq p \leq 2$ and let p' be such that $\frac{1}{p} + \frac{1}{p'} = 1$. Then $\mathcal{F}\colon L^p(\mathbb{R}^d) \to L^{p'}(\mathbb{R}^d)$ and $||\hat{f}||_{p'} \leq ||f||_p$.*

The relation between the Fourier transform and Fourier series is given by the Poisson summation formula.

Proposition 1.5. *(Poisson summation) Suppose that for some $\varepsilon > 0$ and $C > 0$ we have $|f(x)| \leq C(1 + |x|)^{-d-\varepsilon}$ and $|\hat{f}(\omega)| \leq C(1 + |\omega|)^{-d-\varepsilon}$. Then*

$$\sum_{n\in\mathbb{Z}^d} f(x+n) = \sum_{n\in\mathbb{Z}^d} \hat{f}(n)e^{2\pi i n x} \tag{1.2}$$

The identity holds pointwise for all $x \in \mathbb{R}^d$, and both sums converge absolutely for all $x \in \mathbb{R}^d$.

The decay conditions are only necessary for the absolute convergence of the sums but Poisson summation formula also holds under weaker conditions. [65]. The main operations in the context of Fourier analysis are inversion and convolution as defined in the following.

Theorem 1.6. *(Inversion) For $f \in L^1(\mathbb{R}^d)$ and $f \in L^1(\mathbb{R}^d)$ the inversion of the Fourier transform is given by $f(x) = \int_{\mathbb{R}^d} \hat{f}(\omega)e^{2\pi i x\omega} d\omega$.*

Definition 1.7. The *convolution* of two functions $f, g \in L^1(\mathbb{R}^d)$ is the function $(f * g)$ defined by $(f * g)(x) = \int_{\mathbb{R}^d} f(y)g(x-y)dy$. It satisfies $||f * g||_1 \leq ||f||_1||g||_1$ and $\widehat{(f * g)} = \hat{f} \cdot \hat{g}$.

By following Young's theorem the definition of convolution can be extended to other function spaces in an analogue way to the Fourier transform. We state here two versions and give the proof for the latter one.

Theorem 1.8. *(Young inequality) If $f \in L^p(\mathbb{R}^d)$ and $g \in L^q(\mathbb{R}^d)$ and $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}$, with $1 \leq p, q, r \leq \infty$, then $\|f * g\|_r \leq \|f\|_p\|g\|_q$*

To finish the short section on Fourier analysis, Young's Theorem together with a sketch of the proof is stated in order to give a first rough understanding how the concept of smoothing by convolution works.

Theorem 1.9. *Let $1 \leq p, q \leq \infty$ with $\frac{1}{p} + \frac{1}{q} = 1$, $f \in L^p(\mathbb{R}^d)$, $g \in L^q(\mathbb{R}^d)$. Then $f * g \in C^\infty(\mathbb{R}^d)$. If $f \in L^1(\mathbb{R}^d)$, $g \in L^\infty(\mathbb{R}^d)$, then $(f * g)$ is uniformly continuous.*

Proof: By the Hölder inequality, $(f * g)(x)$ is defined for all x and depends continuously on $f \in L^p(X)$ and $g \in L^q(X)$. Moreover,

$$\begin{aligned}(f * g)(x_1) - (f * g)(x_2) &= \int (f(x_1 - y) - f(x_2 - y))g(y)dy \\ &\leq (\int |f(x_1 - y) - f(x_2 - y)|^p dy)^{\frac{1}{p}} \|g\|_q \\ &\leq \omega_{p,f}(x_1 - x_2)\|g\|_q.\end{aligned}$$

Hence $(f * g)$ is uniformly continuous. For $f \in C_c(X)$ obviously $(f * g) \in C_c(X)$. If $p, q < \infty$, then $C_c(X)$is dense in $L^p(X)$, $L^q(X)$. Hence for such p, q, $(f * g)$ belongs to the closure of $C_c(X)$ in $L^\infty(X)$, which is $C^\infty(X)$. □

1.2. Functional analysis

For convenience we state here main definitions and theorems from functional analysis, which are used in this work. As they are standard and can be found in any textbook on Functional analysis, there won't be any further details added.

Definition 1.10. (Banach Algebra) [34] Let A be an algebra. An algebra norm on A is a map $\|\cdot\| : A \mapsto \mathbb{R}$ such that $(A, \|\cdot\|)$ is a normed space, and, further:

$$\|ab\| \leq \|a\|\|b\| \qquad a, b \in A \tag{1.3}$$

The normed algebra $(A, \|\cdot\|)$ is a Banach algebra if $\|\cdot\|$ is a complete norm.

A prominent example for a Banach Algebra is $C_0(X)$, the space of (complex-valued) continuous functions on a locally compact space that vanish at infinity.

Theorem 1.11. *[100] Let X, Y be Banach spaces and $T : D(T) \to Y$ a closed linear operator with domain $D(T) \in X$. If $D(T)$ is closed in X, then T is bounded.*

Theorem 1.12. *(Hölder inequality) [65] If $f \in L^p(\mathbb{R}^d)$ and $g \in L^q(\mathbb{R}^d)$ and $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}$, with $1 \leq p, q, r \leq \infty$, then $\|fg\|_1 \leq \|f\|_p\|g\|_q$.*

Hölders inequality is necessary to establish $L^q(\mathbb{R}^d)$ as the dual space to $L^p(\mathbb{R}^d)$ and the duality is defined by the standard inner product

$$\langle f, g \rangle = \int_{\mathbb{R}} f(x)g(x)dx. \tag{1.4}$$

Definition 1.13. A bounded subset $H \in C_0(\mathbb{R}^d)$ is called *tight* if for every $\varepsilon > 0$ there exists a $k \in C_c(\mathbb{R}^d)$ such that $\|h - k \cdot h\|_\infty < \varepsilon$ for all $h \in H$.

1.3. Time Frequency Shifts and Lattice

Time-Frequency shift operators are the main building blocks in Time-Frequency analysis and defined as follows.

Definition 1.14. An operator $\pi(x,\omega)$ of the form $\pi(x,\omega)f = f(t-x)e^{2\pi i x\omega}$ is called a *Time-Frequency Shift operator.*

Time Frequency shift operators are the projective representation of the so called Weyl Heisenberg group $\mathbf{H} = \{(x,\omega,\phi) \in \mathbb{R}\times\mathbb{R}\times[0,1]\}$ with the group multiplication $(x,\omega,\phi)(x',\omega',\phi') = (x+x', \omega+\omega', \phi+\phi'-\omega' x))$ [65]. Usually we restrict ourselves to the irreducible unitary representation denoted by π^0 defined by

$$\pi^0(x,\omega,\phi) = e^{2\pi i\phi} M_\omega T_x \tag{1.5}$$
$$\pi(x,\omega) = \pi^0(x,\omega,0) \tag{1.6}$$

Definition 1.15. A lattice $\Lambda \subseteq \mathbb{R}^d$ is a discrete subgroup of $\mathbb{R}^d$ of the form $\Lambda = A\mathbb{Z}^d$, where A is an invertible $d\times d$ matrix over $\mathbb{R}$.

In the following we will usually use separable lattice of the form $\Lambda = \alpha\mathbb{Z}\times\beta\mathbb{Z}$

Definition 1.16. The dual lattice of Λ with respect to the standard symplectic form on $\mathbb{R}^d \times \hat{\mathbb{R}}^d$ is the *adjoint lattice* $\Lambda^0 \subset \mathbb{R}^d\times\hat{\mathbb{R}}^d$ of Λ. It is defined by $\lambda' \in \Lambda^0 \Leftrightarrow [\lambda',\lambda] \in \mathbb{Z} \quad \forall\lambda\in\Lambda$ [65].

In particular, this means if $\Lambda = \alpha\mathbb{Z}\times\beta\mathbb{Z}$, then $\Lambda^0 = \frac{1}{\beta}\mathbb{Z}\times\frac{1}{\alpha}\mathbb{Z}$ [55] [31].

1.4. Frames and Riesz Bases

Especially in the context of signal processing it turned out that although its nice properties the concept of orthogonal basis is not always sufficient, but the introduction of a certain redundancy within the information can be beneficial. This observation led to the introduction of frames by Duffin and Schaefer [46]. Frames were made popular by I. Daubechies [36] and are now one of the basic concepts in harmonic analysis [65] and in signal processing [17]. This section is intended to sketch the underlying concepts and interconnections between orthogonal bases, frames and Gabor Riesz bases. A frame is a particular generating system and defined as follows.

Definition 1.17. [65] [10] A set of vectors (g_n) in a Hilbert Space $\mathcal{H}$ is a frame for $\mathcal{H}$ if

$$A||x||^2 \leq \sum_k |\langle x, g_k\rangle|^2 \leq B||x||^2, \tag{1.7}$$

with $0 < A \leq B < \infty$. A and B are called frame bounds. If $A = B$ the frame is called *tight.* A sequence is called *Bessel sequence* if the right inequality above is fulfilled.

Definition 1.18. [65] The frame operator S for the frame (g_k) is defined as

$$Sx = \sum_k \langle x, g_k \rangle g_k = \int_x \kappa(S)(t,s)x(s)ds. \tag{1.8}$$

The kernel $\kappa(S)$ of the frame operator is given by $\kappa(S)(t,s) = \sum_k g_k(t)g_k^*(s)$

S is a self adjoint positive semidefinite operator. The frame bounds A and B of $(g)_k$ are given by the infimum and the supremum, respectively, of the eigenvalues of the frame operator S. The concept of frames can also be extended to general Banach spaces [23]. Another important concept in time frequency analysis is the following:

Definition 1.19. [125], page 32 A sequence of vectors (g_n) in $\mathcal{H}$ is a Riesz basic sequence if there exist constants $0 < c < C < \infty$ such that

$$c(\sum |a_n|^2) \leq || \sum a_n x_n ||^2 \leq C(\sum |a_n|^2) \tag{1.9}$$

for all $a_n \in l^2$ and which is a basis for its closed linear span.

A Riesz basis is a Riesz basic sequence spanning $\mathcal{H}$. Equivalently a Riesz basis can be defined as the image of an orthonormal basis [72].

Lemma 1.20. *[125] A family (g_n) of vectors in $\mathcal{H}$ is a Riesz basis (or bounded unconditional basis) if there is a bounded invertible operator $U : \mathcal{H} \to \mathcal{H}$ and an orthonormal basis e_n for $\mathcal{H}$ such that $f_n = Ue_n$.*

One key property of a Riesz basis is the existence of a unique biorthogonal basis.

Definition 1.21. Two sequences (g_n) and $(\tilde{g}_n)$ in $\mathcal{H}$ are called biorthogonal if $\langle g_n, \tilde{g_m} \rangle = \delta_{nm}$ with δ_{nm} denoting the Kronecker δ function.

Definition 1.22. [125], p.29 (f_n) is minimal $\Leftrightarrow$ for all m, $f_m \notin \overline{\text{span}}(f_n)_{n \neq m}$

Definition 1.23. [110], Def (2.1) (f_n) is called complete in $\boldsymbol{\mathcal{HS}}$ if $\overline{\text{span}}\,\{(f_n)\} = \boldsymbol{\mathcal{HS}}$.

Lemma 1.24. *[125], p.29 g_n has a biorthogonal sequence if and only if it is minimal. Also, if a biorthogonal sequence exists, it is unique if and only if (g_n) is complete.*

The existence of a biorthogonal basis is very important as it ensures that every element $f \in \mathcal{H}$ can be written as a superposition of elements of the form $f = \sum_{n \in N} \langle f, \tilde{g_n} \rangle g_n$ with unique coefficients. The connection between frames and Riesz basis can be stated as follows

Theorem 1.25. *([27] Theorem 3.4.4, Prop 3.4.5) For a sequence $(e_k)_{k=1}^{\infty}$ in a Hilbert space $\boldsymbol{\mathcal{HS}}$, the following conditions are equivalent*

- *(e_k) is a Riesz basis for $\boldsymbol{\mathcal{HS}}$.*
- *(e_k) is complete in $\boldsymbol{\mathcal{HS}}$ and there exist constants $A, B > 0$ such that for every finite scalar sequence one has*

$$A \sum |c_k|^2 \leq || \sum c_k e_k ||^2 \leq B \sum |c_k|^2 \tag{1.10}$$

- *(e_k) is complete and its Gram matrix $G := (\langle e_k, e_j\rangle)_{k,j=1}^{\infty}$ defines a bounded, invertible operator on $\ell^2(\mathbb{N})$. In this case the optimal Riesz bounds are $A = \frac{1}{||G^{-1}||}$ and $B = ||G||$*

Theorem 1.26. *[26] [90] Let (g_n) be a frame. Then the following are equivalent*

- *(g_n) is a Riesz basis.*
- *(g_n) is minimal.*
- *(g_n) has a unique biorthogonal sequence.*
- *$(S^{-1}g_n)$ is a biorthogonal sequence.*
- *If $\sum a_n g_n = 0$ in $\mathcal{H}$ for a sequence of scalars $(a_n) \in \ell^2$, then $a_n = 0 \quad \forall n$.*

.

Theorem 1.27. *[55] [99] (g, Λ) is a Riesz basis for its closed linear span in $L^2(\mathbb{R}^d)$ if and only if (g, Λ^0) is a frame for $L^2(\mathbb{R}^d)$.*

1.5. Special functions and operators

In particular the definition of weigth functions will be needed by the following definiton of some families of weighted function spaces.

Definition 1.28. [67] In general a *weight* function m on $\mathbb{R}^d$ is simply a non-negative function. Usually it is assumed without loss of generality that the weight is continuous.

In the course of this work, we will need the following properties of weight functions. All these and more details can be found in [67]

Definition 1.29. Let v and m be non-negative functions on $\mathbb{R}^d$ or on $\mathbb{Z}^d$.

(1) A weight v is called submultiplicative, if

$$v(x+y)v(x)v(y) \qquad \forall x, y \in \mathbb{R}^d \tag{1.11}$$

.

(2) Given a submultiplicative weight v, a non-negative function m is called a v-moderate weight, if there exists a constant $C > 0$, such that

$$m(x+y)Cv(x)m(y) \qquad \forall x, y \in \mathbb{R}^d. \tag{1.12}$$

If m is v-moderate with respect to some v, then we simply call m moderate.

Definition 1.30. Let St_ρ denote the dilation operator and be defined as $\mathrm{St}_\rho f(z) = \frac{1}{\rho} f(\rho z)$, $\rho > 0, z \in \mathbb{R}^d$.

Definition 1.31. Let $\chi_\Omega : \mathbb{R} \to \mathbb{R}$ denote the *indicator function* defined through

$$\chi(x) = \begin{cases} 1 & x \in \Omega \\ 0 & \text{else} \end{cases} \tag{1.13}$$

Definition 1.32. A BUPU, a so-called **b**ounded **u**niform **p**artition of **u**nity in some Banach algebra $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ of continuous functions on the group $\mathscr{G}$ is a family $(\psi_i)_{i\in I}$ of non-negative functions on $\mathscr{G}$, if the following set of conditions is satisfied:

(1) There exists some neighborhood U of the identity element of the group $\mathscr{G}$ that for each $i \in I$ there exists $x_i \in \mathscr{G}$ such that $\operatorname{supp}(\psi_i) \subseteq x_i + U$ for all $i \in I$;
(2) The family Ψ is bounded in $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$, i.e. $sup_{i\in I}\|\psi\|_{\boldsymbol{A}} \leq C_A < \infty$;
(3) The family of supports $(x_i+U)_{i\in I}$ is it relatively separated, i.e. for each $i \in I$ the number of intersecting neighbors is uniformly bounded in the following sense
$$\#\{j \,|\, x_i + U \cap x_j + U \neq \varnothing\} \leq C_0;$$
(4) $\sum_{i\in I} \psi_i(x) \equiv 1$.

For more information on BUPUs and example of other usage compare for example [37] [58]

1.6. Function Spaces & Norms

We use the usual definition of L^p *spaces* with the norm $||f||_p = \left(\int_{\mathbb{R}} |f(x)|^p dx\right)^{\frac{1}{p}}$. The discrete analogue in form of ℓ^p spaces is defined for sequences $(x_n)_{n\in\mathbb{N}}$ through the norm $||\mathbf{x}||_p = \left(\sum_{n\in\mathbb{N}} |x_n|^p\right)^{\frac{1}{p}}$.

Definition 1.33. [71] The *weighted L^p-space*, L^p_w, is the set of all complex-valued functions on $\mathbb{R}$ for which $fw \in L^p$, with norm

$$\|f\|_{L^p_w} = \|fw\|_{L^p} = \left(\int_{\mathbb{R}} |f(x)w(x)|^p dx\right)^{(1/p)}, \qquad 1 \leq p < \infty, \tag{1.14}$$

or

$$\|f\|_{L^\infty_w} = \|fw\|_{L^\infty} = \operatorname*{ess\,sup}_{x\in\mathbb{R}} |f(x)|w(x), \qquad p = \infty. \tag{1.15}$$

If $w \equiv 1$ we simply write L^p. It can be shown that L^p_w is a Banach space for $1 \leq p \leq \infty$. The dual index to p is $p^* = \frac{p}{(p-1)}$, i.e., $\frac{1}{p} + \frac{1}{p'} = 1$. We have $(L^p_w)^* = L^{p'}_{1/w}$ for $1 \leq p < \infty$, where the star denotes the Banach space dual and the duality is defined by the standard inner product.

The following definitions and statements as well as proofs and more details can be found in [62]

Definition 1.34. (Test Functions) Let $\Omega \subseteq \mathbb{R}^d$ be an open non empty subset. Then $C_c^\infty(\Omega) = \mathcal{D}(\Omega) := \{f \in C_c^\infty(\Omega) |\operatorname{supp}(f) \text{ is compact}\}$ is the space of test functions on Ω.

Definition 1.35. (Distribution) A distribution g on Ω is a linear functional on $\mathcal{D}(\Omega)$, i.e. $g : \mathcal{D}(\Omega) \to \mathbb{C}$ linear, with the following continuity property:

$$f_n \to f \quad \text{in} \quad \mathcal{D}(\Omega) \Rightarrow g(f_n) \to g(f) \quad \text{in} \quad \mathbb{C} \tag{1.16}$$

The complex vector space of distributions on Ω is denoted by $\mathcal{D}'(\Omega)$. We often write $\langle g, f \rangle$ instead of $g(f)$

Definition 1.36. (Schwartz space) Let $f \in C^\infty(\mathbb{R}^d)$. The function f is said to be rapidly decreasing, if it satisfies the following semi-norm condition

$$\forall \alpha, \beta \in \mathbb{N}^d : \qquad q_{\alpha,\beta}(f) := \sup_{x \in \mathbb{R}^d} |x^\alpha D^\beta f(x)| < \infty. \tag{1.17}$$

The vector space of all rapidly decreasing functions on $\mathbb{R}^d$ is denoted by $\mathcal{S}(\mathbb{R}^d)$.

Lemma 1.37. *$\mathcal{F}\big(\mathcal{S}(\mathbb{R}^d)\big) \subseteq \mathcal{S}(\mathbb{R}^d)$ and $f \to \hat{f}$ is continuous $\mathcal{S}(\mathbb{R}^d) \to \mathcal{S}(\mathbb{R}^d)$.*

Definition 1.38. (Tempered Distribution) A tempered (also temperate) distribution on $\mathbb{R}^d$ is a continuous linear functional $g : \mathcal{S}(\mathbb{R}^d) \to \mathbb{C}$, i.e. $f_k \to 0$ in $\mathcal{S} \Rightarrow \langle g, f_k \rangle \to 0$ in $\mathbb{C}$. The space of tempered distributions on $\mathbb{R}^d$ is denoted by $\mathcal{S}'(\mathbb{R}^d)$.

Further let $C_c(\mathbb{R}^d) := \{f : \mathbb{R}^d \to C$, continuous and with compact support$\}$ be the space of *continuous functions with compact support*, which is dense in $L^1(\mathbb{R}^d)$ [102] (Lemma 14.7).

Definition 1.39. [39] For $m \in \mathbb{N}$ and $1 \leq p \leq \infty$, then the *Sobolev space* denoted by $W^{m,p}(\mathbb{R}^d)$ consists of the functions in $L^p(\mathbb{R}^d)$ whose partial derivatives up to order m, in the sense of distributions, can be identified with functions in $L^p(\mathbb{R}^d)$. For these derivatives, we set $\alpha = (\alpha_1, ..., \alpha_n)$ and $|\alpha| = \sum_{i=1}^n \alpha_i$. Moreover, we use the notation

$$D^\alpha u = \frac{\partial^{|\alpha|} u}{\partial \alpha_1 x_1 ... \partial \alpha_n x_n}. \tag{1.18}$$

The denition above can now be written as

$$W^{m,p}(\mathbb{R}^d) = \left\{u \in L^p(\mathbb{R}^d) | \forall \alpha \in \mathbb{N}, |\alpha| \leq m D^\alpha u \in L^p(\mathbb{R}^d)\right\}. \tag{1.19}$$

Equivalently Sobolev spaces can be defined via the Fourier transform [120], which will be one of the basic concepts we make use of when describing the convergence in Chapter (3).

$$W^m(\mathbb{R}^d) = \left\{f \in L^2(\mathbb{R}^d) : (1 + |\zeta|^2)^{\frac{m}{2}} \mathcal{F} f \in L^2(\mathbb{R}^d)\right\} \tag{1.20}$$

Definition 1.40. The Hilbert Schmidt norm of an operator $H : H_1 \to H_2$ between two Hilbert spaces is defined as usual

$$||H||_{\mathcal{HS}} := \left(\sum_{i \in I} ||He_i||_{H_2}^2\right)^{\frac{1}{2}} \tag{1.21}$$

with $(e_i)_{i \in I}$ being an orthonormal basis of H_1. The Hilbert Schmidt norm is induced by the scalar product $\langle T, U \rangle = \text{trace}(T^* U)$, with T^* denoting the adjoint of T.

Moreover we use the norm of total variation:

Definition 1.41. The norm of total variation of a function f on the interval $[a, b]$ is defined as $||f||_{BV([a,b])} = |f(a)| + \sup\limits_{P} \sum_{i=1}^{N} |f(x_i) - f(x_i - 1)|$ with $P = \{a = x_0 < x_1 < ... < x_N = b\}$.

Besides these standard spaces from functional analysis we want to shortly introduce two classes of spaces in particular useful for discrete and continuous time frequency analysis. They are the Wiener Amalgam spaces and the Modulation spaces.

1.6.1. Wiener Amalgam Spaces. Wiener Amalgam spaces combine local properties with a global behavior. This idea was first mentioned by Norbert Wiener in [121], [123] and [122]. For a good introduction and a structured analysis of theses space we refer to the work of Feichtinger [49], [50]. For weighted amalgam spaces we also refer to [71].

Definition 1.42. [71] Let $1 \le p, q \le \infty$ and a weight w, according to Definition (1.28) and following, be given. Fix a compact $Q \subset \mathbb{R}^d$ with nonempty interior. Then the amalgam space $W(L^p, L^q_w)$ consists of all functions $f : \mathbb{R}^d \to \boldsymbol{C}^d$ such that $f \cdot \chi_K \in L^p$ for each compact $K \subset \mathbb{R}^d$, and for which the control function

$$F_f^Q(x) = \|f \cdot \chi_{Q+x}\|_{L^p} = \|f \cdot T_x \chi_Q\|_{L^p} \quad x \in \mathbb{R}^d \tag{1.22}$$

lies in L^q. The norm is given by

$$\begin{aligned}
\|f\|_{W(L^p, L^q_w)} &= = \|F_f\|_{L^q_w} = \|\|f \cdot \chi_{Q+x}\|_{L^p}\|_{L^q_w} \\
&= \left(\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} |f(t)|^p \chi_Q(t-x) dt \right)^{q/p} w(x)^q dx \right)^{1/q} \\
&= \left(\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} |f(t)|^p \chi_Q(t-x) w(x)^p dt dx \right)^{q/p} \right)^{1/q}
\end{aligned}$$

with the usual adjustments if p or q is infinity.

The main properties of Wiener Amalgam spaces, which will be used in the following are the following. Proofs can be found in [71].

Proposition 1.43. *If w is a v moderate weight then $W(L^p, L^q_w)$ is a Banach space. The definition of the space is independent of the choice of Q. Different Q_s define the same space with equivalent norms.*

Proposition 1.44. *If w is a v-moderate weight, the spaces $W(L^p, L^q_w)$ are translation invariant. For $w \equiv 1$, translation is an isometry on $W(L^p, L^q_w)$.*

Proposition 1.45. *If w is a moderate weight, then $W(L^p, L^p_w) = L^p_w$.*

Theorem 1.46. *Let w be a moderate weight, and assume that $1 \le p, q < \infty$. Then $W(L^p, L^q_w)' = W(L^{p'}, L^{q'}_{1/w})$.*

Theorem 1.47. *Assume the indices p_i, q_i and the moderate weights w_i are such $L^{p_1} * L^{p_2} \subset L^{p^3}$ and $L^{q_1}_{w_1} * L^{q_2}_{w_2} \subset L^{q_3}_{w_3}$ then*

$$W(L^{p_1}, L^{q_1}_{w_1}) * W(L^{p_2}, L^{q_2}_{w_2}) \subset W(L^{p_3}, L^{q_3}_{w_3}). \tag{1.23}$$

1.6.2. Modulation Spaces. Modulation spaces measure the time frequency concentration of a signal. In that sense, they are the appropriate function spaces for time frequency analysis. The classical modulation spaces are defined via the Short Time Fourier transform. We will state immediately the weighted version, as the unweighted modulation spaces are just special cases for $w \equiv 1$.

Definition 1.48. [65] Assume w is a v moderate weight function and given a window function $g \in S(\mathbb{R}^d)$, then the modulation space $M^w_{p,q}(\mathbb{R}^d)$ is the space of all distributions $f \in S'(\mathbb{R}^d)$ for which the following norm is finite

$$\|f\|_{M^w_{p,q}} = \left(\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} |V_g f(x,y)|^p w(x,y)^p dx \right)^{q/p} dy \right)^{1/q} \tag{1.24}$$

with usual modification for p or q is equal to infinity.

Proposition 1.49. *[65] $M^w_{p,q}$ is a Banach space whose definition is independent of the choice of the window g. Different windows define the same space with equivalent norms.*

Modulation spaces are a very powerful concept and a lot of other families of function spaces can be seen as special cases, for example

- The space $M^1_{2,2}(\mathbb{R}^d)$ is equal to the space $L^2(\mathbb{R}^d)$.
- For $w(x,y) = (1+|x|)^s$ weighted L^2 spaces are given by
$$L^2_s(\mathbb{R}^d) = M^w_{2,2}(\mathbb{R}^d) \tag{1.25}$$
- Classical Sobolev spaces can be found in the theory of modulation spaces for $w(x,y) = (1+|y|^2)^s$
$$H^s(\mathbb{R}^d) = M^{2,2}_w(\mathbb{R}^d). \tag{1.26}$$
- A special attention should also be devoted to the spaces $M^{p,p}_r(\mathbb{R}^d)$ with radial symmetric weight function $r(x,y) = (1+|x|^2+|y|^2)$. They are also called Sobolev Shubin spaces and have besides others the property to be invariant under the Fourier transform. For further details on Sobolev Shubin spaces we refer for example to [16].

Lemma 1.50. *[65] If $g_1 \in M^1_{v_p}(\mathbb{R}^d)$ and $g_2 \in M^2_{v_p}(\mathbb{R}^d)$ with $v_p(z) = (1+|z|)^p$, then $V_{g_1} g_2 \in L^2_{v_p}(\mathbb{R}^{2d})$.*

For convenience and in particular as it will be further used later on in this work, also the proof is stated here. **Proof**: Let $\gamma_1, \gamma_2 \in S_0(\mathbb{R}^d)$. We use that

$$|V_{g_1} g_2(x,\omega)| \leq |\langle \gamma_1, \gamma_2 \rangle|^{-1} (|V_{\gamma_1} g_1| * |V_{g_2} \gamma_2|)(x,\omega). \tag{1.27}$$

(compare [65] Lemma 11.3.3) and apply the convolution relation (Proposition 11.1.3 from [65]) $||V_{g_1}g_2||_{L^2_v} \leq C||V_{\gamma_1}g_1||_{L^1_v}||V_{\gamma_2}g_2||_{L^2_v}$ The last expression is equivalent to $V_{g_1}g_2 \in L^2_v$. (compare [65] Lemma 11.3.3) and apply the convolution relation (Proposition 11.1.3 from [65]) $||V_{g_1}g_2||_{L^2_v} \leq C||V_{\gamma_1}g_1||_{L^1_v}||V_{\gamma_2}g_2||_{L^2_v}$ The last expression is equivalent to $V_{g_1}g_2 \in L^2_v$. □

A good overview over modulation spaces can be found in [65].

1.6.3. The space S_0. The space $S_0(\mathbb{R}^d)$ fits in the theory of modulation spaces as $M_0^{1,1}(\mathbb{R}^d)$. It is devoted an own section as it was historically the first of the modulation spaces to be found and has some exceptionally nice properties.

Definition 1.51. [61] For $\phi \in S(\mathbb{R}^d)$, $S_0(\mathbb{R}^d) = \left\{ f \in L^2(\mathbb{R}^d) : \|f\|_{S_0} := ||V_\phi f||_{L^1(\mathbb{R}^{2d})<\infty} \right\}$. Its main properties are as follows:

- $S_0(\mathbb{R}^d)$ is a Banach space.
- $S_0(\mathbb{R}^d)$ is dense in $L^2(\mathbb{R}^d)$.
- The definition is independent of the choice of the window.

Moreover $S_0(\mathbb{R}^d)$ as well as its dual $S_0'(\mathbb{R}^d) = M_0^{\infty,\infty}(\mathbb{R}^d)$ is invariant under the Fourier transform. The Banach spaces $\boldsymbol{S}_0$ and $\boldsymbol{S}_0{}'$ may also be interpreted as Wiener amalgam spaces in the following way: $\boldsymbol{S}_0 = W(\mathcal{F}L^1, L^1)$ and $\boldsymbol{S}_0{}' = W(\mathcal{F}L^\infty, L^\infty)$ with $\mathcal{F}L^1$ denoting the Fourier image of absolutely integrable functions, compare [61] Sect. 3.2.2.

1.7. Operator representation

This section is devoted to the representation theory of linear operators as foundation for the further investigations. Different representations of operators give the ability to shed light on different aspects of an operator. They provide the possibility to choose the most convenient one for the further calculations, system identification and system design.

Symbolic calculus, the representation of an operator in terms of coefficients corresponding to appropriately chosen building blocks, is the key for the solution of many problems in the field of operator theory [32]. The 3 symbols primarily used in this work are the operator kernel, the spreading symbol and the Kohn-Nirenberg symbol, also called Zadeh's time varying transfer function.

1.7.1. Integral kernel.

Theorem 1.52. *[65], Theorem 14.3.4., [104], [74] For K being a continuous linear operator mapping from $\mathscr{S}(\mathbb{R}^d)$ into $\mathscr{S}'(\mathbb{R}^d)$ (with weak * topology) there exists a tempered distribution $\kappa \in \mathscr{S}(\mathbb{R}^d)$ such that*

$$\langle Kf, g \rangle = \langle k, f \otimes g \rangle \tag{1.28}$$

$$Kf(t) = \int \kappa(K)(t,y)f(y)dy. \tag{1.29}$$

This can be seen as a generalization of the matrix representation in the finite dimensional case. Therefore it is probably the best known representation of an operator.

1.7.2. Kohn-Nirenberg Symbol / Zadeh Transfer Function.

Definition 1.53. [65] Section 14.1 Let σ be a measurable function or a tempered distribution on $\mathbb{R}^{2d}$. Then the operator

$$Kf = \int_{\mathbb{R}^d} \sigma(x,\omega)\widehat{f}(\omega)e^{2\pi i x\omega}d\omega. \tag{1.30}$$

is called the *pseudodifferential operator* with symbol σ. The mapping $\sigma \to K$ is usually referred to as the Kohn Nirenberg correnspondence and σ is called the *Kohn Nirenberg symbol.*

The Kohn-Nirenberg symbol is very similar to the representation as Zadeh's time varying transfer function used in the engeneering community [94]. The term transfer function comes from the fact that the operator representation of (1.30) can also be written in the form

$$Kf(t) = \int_{\mathbb{R}^d} h(t, t-x)f(x)dx \tag{1.31}$$

Equation 1.31 can now be interpreted as the transfer function like in the LTI case in equation (1.45) with an additional second time argument.

The Konn Nirenberg symbol as defined above has some important properties, which simplify the calculus. One of them is the shift covariance property:

$$[\sigma(\pi_2(\lambda)K)] = T_\lambda[\sigma(K)]. \tag{1.32}$$

The second important property is the trace formula [59] p.326 (10.2.9):

$$tr(K) = \int_{\mathbb{R}^d} \kappa(K)(x,x)dx = \int_{\mathbb{R}^d\times\widehat{\mathbb{R}}^d} \sigma(K)(\lambda)d\lambda \tag{1.33}$$

An equivalent to the Kohn-Nirenberg symbol often used in physics is the Weyl symbol [98] [82] given by

$$w(K)(x,\zeta) = \int_{\mathbb{R}^d} \kappa(K)(x+\frac{1}{2}t, x-\frac{1}{2}t)e^{-2\pi i\zeta t}dt \tag{1.34}$$

Through its symmetry the Weyl symbol possesses the additional nice properties that $w(H^*) = \bar{w}(H)$ and it is covariant under transformations by metaplectic operators.

An integrative study on symbols of the kind of the KNS and the Weyl symbol has been done by Werner Kozek [82]. In this work the generalized Weyl symbol combining all these representations has been introduced.

One inconvenience of the Kohn Nirenberg symbol as well as the Weyl symbol is that the operator product usually does not equal the product of the symbols [38], Chapter 10, Theorem 213.

1.7.3. Spreading function. An operator K can be expressed as a superposition of time frequency shifts as follows [43] [65]

$$K = \int_{\mathbb{R}^d} \int_{\widehat{\mathbb{R}}^d} \eta(K)(x,\omega) M_\omega T_x d\omega dx \tag{1.35}$$

The spreading symbol is related to the kernel of the operator by the following transform

$$\eta(K)(t,\nu) = \int_{\mathbb{R}^d} \kappa(K)(x, x-t) e^{-2\pi i \nu x} dx \tag{1.36}$$

In addition, the spreading symbol is related to the KNS via the symplectic Fourier Transform:

$$\eta(K)(t,\nu) = \mathcal{F}_s[\sigma(K)](t,\nu) = \int_{\mathbb{R}^d} \int_{\widehat{\mathbb{R}}^d} \sigma(K)(x,\zeta) e^{2\pi i (x\nu - \zeta t)} dx d\zeta \tag{1.37}$$

For engineering applications, the spreading function provides an intuitive way of describing the channel operator, i.e. the operator mapping the transmitted signal to the received signal for example in a wireless transmission scheme. An engineer usually thinks of the channel acting on signal through various reflectors. The action of a reflector consists of a time delay (time shift) and a Doppler effect (frequency shift). Thus the operator is somehow a sum of time shifts and frequency shifts with different amplitudes. These amplitudes are exactly the values of the spreading function.

Because of the practical interpretation of its coefficients the spreading function is commonly used in statistical modeling. Basis expansion models with respect to the basis of complex exponentials as e.g. described in [89] are spreading representations with a limited number of coefficients, which are chosen due to higher order output statistics.

To get a better intuition for the spreading function it shall be noted that the spreading symbol of a time frequency shift $\pi(\lambda)$ is equal to the Delta distribution δ_λ. The support of the spreading function of a time invariant system is limited to the time axis. Moreover the value of the spreading function at the origin equals the operator's trace and its magnitude is covariant to metaplectic transformations [94].

Lemma 1.54. *If $g_1, g_2 \in M^2_{v_p}(\mathbb{R}^d)$ and $Sf = \int \langle f, \pi(\lambda) g_1 \rangle \pi(\lambda) g_2$ is Hilbert Schmidt, then $\eta(S) \in L^2_{v_p}(\mathbb{R}^{2d})$ and $\sigma(S) \in W^p_2(\mathbb{R}^{2d})$.*

Proof: $S \in \boldsymbol{\mathcal{HS}}$ is equivalent to $m \in L^2$ and by Plancherel's theorem then also $\mathscr{M} \in L^2$. Moreover, if $g_1, g_2 \in M^2_{v_p}$ this means that $V_{g_1} g_2 \in L^2(v_p)$ To show this we use that

$$|V_{g_1} g_2(x,\omega)| \leq |\langle \gamma_1, \gamma_2 \rangle|^{-1} (|V_{\gamma_1} g_1| * |V_{g_2} \gamma_2|)(x,\omega). \tag{1.38}$$

(compare [65] Lemma 11.3.3) and apply the convolution relation (Proposition 11.1.3 from [65])

$$||V_{g_1}g_2||_{L^2_v} \leq C||V_{\gamma_1}g_1||_{L^1_v}||V_{\gamma_2}g_2||_{L^2_v} \tag{1.39}$$

then $V_{g_1}g_2 \in L^2_v$. This is not trivial and can be found in [65]. The spreading function of the STFT multiplier is given by $\eta(S) = \mathscr{M} \cdot V_{g_1}g_2$ as already used extensively in this work and therefore $\eta(S) \in L^2_v$. □

1.7.4. Transition formulas. The mappings between integration kernel, Kohn Nirenberg symbol and Spreading function are Gelfand triple isometries, meaning all of the introduced symbols preserve inner products and Hilbert Schmidt norms of operators [86].

$$||K||_{\boldsymbol{\mathcal{HS}}} = ||\kappa(K)||_2 = ||\sigma(K)||_2 = ||\eta(K)||_2 \tag{1.40}$$

Explicit transition formulas can be calculated and a table which overviews clearly arranged the relation between the presented operator symbols can be found in Table (1.7.4).

	Integral Kernel $k(x,y)=$	TV Impulse Response $h(x,y)=$	Kohn Nirenberg Symbol = Zadeh's TV Transfer Function $\sigma(x,w)=$	Spreading Function $\eta(v,y)=$	Operator $Af=$
IK	$\cdot$	$k(x,x-y)$	$\int k(x,x-y)e^{-2\pi iyw}dy$	$\int k(x,x-y)e^{-2\pi ixv}dx$	$\int k(x,y)f(y)dy$
TVIR	$h(x,x-y)$	$\cdot$	$\int h(x,y)e^{-2\pi iyw}dy$	$\int h(x,y)e^{-2\pi ixv}dx$	$\int h(x,y)f(x-y)dy$
KNS	$\int \sigma(x,w)e^{2\pi iw(x-y)}dw$	$\int \sigma(x,w)e^{2\pi iyw}dw$	$\cdot$	$\int\int \sigma(x,w)e^{2\pi i(xv-yw)}dxdw$	$\int \sigma(x,w)\hat{f}(w)e^{2\pi ixw}dw$
SF	$\int \eta(v,x-y)e^{2\pi ixv}dv$	$\int \eta(v,y)e^{2\pi ixv}dv$	$\int\int \eta(v,y)e^{2\pi i(vx-wy)}dvdy$	$\cdot$	$\int\int \eta(v,y)f(x-y)e^{2\pi ivx}dvdy$

TABLE 1. Transition Formulas between the main Operator symbols.

1.8. Banach Gelfand Triples

A good over the use of Banach Gelfand triples in time frequency analysis also including the following statements can be found in [14].

Definition 1.55. A triple, consisting of a Banach space $\boldsymbol{B}$, which is dense in some Hilbert space $\mathcal{H}$, which in turn is contained in $\boldsymbol{B}'$ is called a Banach Gelfand triple.

Definition 1.56. If $(\boldsymbol{B}_1, \mathcal{H}_1, \boldsymbol{B}_1')$ and $(\boldsymbol{B}_2, \mathcal{H}_2, \boldsymbol{B}_2')$ are Gelfand triples then a linear operator T is called a [unitary] Gelfand triple isomorphism if

(1) A is an isomorphism between $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$.
(2) A is [a unitary operator resp.] an isomorphism between $\mathcal{H}_1$ and $\mathcal{H}_2$.
(3) A extends to a weak* isomorphism as well as a norm-to-norm continuous isomorphism between $\boldsymbol{B}_1'$ and $\boldsymbol{B}_2'$.

In this work we use the Banach Gelfand triple $(\boldsymbol{S}_0, L^2, \boldsymbol{S}_0{}')$. The spreading function is an isometric Banach Gelfand Triple isomorphism [see [65], Chap. 9]

Theorem 1.57. *Let $H \in (\boldsymbol{B}, \boldsymbol{H}, \boldsymbol{B}')$; then there exists a spreading function $\eta(H)$ in the triple $(\boldsymbol{S}_0(\mathbb{R}^2), L^2(\mathbb{R}^2), \boldsymbol{S}_0{}'(\mathbb{R}^2))$ such that*

$$H = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \eta H(b,\nu)\pi(b,\nu)dbd\nu. \tag{1.41}$$

For $H \in \boldsymbol{H}$, the correspondence $H \leftrightarrow \eta(H)$ is isometric, i.e.$\|H\|_{\boldsymbol{H}} = \|\eta H\|_{L^2(P)}$.

For $H \in \boldsymbol{B}$, the decomposition is absolutely convergent, whereas, for$H \in \boldsymbol{B}'$ it holds in the weak sense of bilinear forms on $\boldsymbol{S}_0$. When $\eta(H) \in L^2(P)$, H is a Hilbert-Schmidt operator, and the above integral is dened as a Bochner integral.

1.9. Linear Time Invariant and slowly time variant systems

1.9.1. LTI systems. Before we go on to time varying systems, we want to recall the most important facts about linear time invariant systems as they are essential for the later treatment of time varying systems. A time invariant system S is characterized by the fact that it responds to a delayed input signal x with the same delay in the output signal y i.e.

$$y(t) = Sx(t) \Rightarrow y(t-t_0) = Sx(t-t_0). \tag{1.42}$$

This means that the kernel of a time invariant system is just a convolution kernel $\kappa(t,s) = g(t-s)$. For a communication channel for example this would mean a wired or a fixed wireless (location of sender and receiver is fixed) data transmission.

Lemma 1.58. *[52] A matrix A represents a translation invariant linear mapping on C^d if and only if it is circulant i. e. if it is constant along side diagonals.*

Proof: We denote the ith column of the matrix A with a_i. Therefore we have $a_i = Ae_i \quad i = 1, ..., d$. But the ith unit vector is nothing else then the shifted version of the first unit vector $e_i = T_i e_1$. If we now in addition apply the definition of a time invariant system, we get:

$$a_i = Ae_i = AT_i e_1 = T_i a_1 \quad i = 1, ..., d \tag{1.43}$$

This is the definition of a matrix A to be circulant. On the other hand if we take a look on the action of $T\_1 \circ A \circ T_1$ we get

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & \cdots & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ 1 & 0 & 0 & \cdots & 0 \end{pmatrix} = \begin{pmatrix} a_{nn} & a_{n1} & \cdots & a_{n(n-1)} \\ a_{1n} & a_{11} & \cdots & a_{1(n-1)} \\ \vdots & & \ddots & \vdots \\ a_{(n-1)n} & & \cdots & a_{(n-1)(n-1)} \end{pmatrix}$$

This means, if the matrix A is constant a long the side-diagonals nothing will change in the action of A. $\square$

Linear time invariant systems are those operators which are situated in the span of all time shift operators. It has been shown for example in [28] that only with time shifts it is not possible to span the whole space of $N \times N$ matrices . Circulant matrices form a closed commutative subalgebra of the Banach algebra $GL(d)$. Such circulant matrices are also called convolution matrices, because the action performed is the discrete convolution. They can be diagonalized by the discrete Fourier transform, which is the reason for their computational efficiency and predominant role in practice.

The first column of such a convolution matrix is often called impulse response. The circulant structure of the matrix is also the reason why electrical engineers often say that a LTI system can be 'completely characterized by a single function' [5], the so called impulse response h. A LTI system can also be written as

$$y(t) = x(t) * h(t). \tag{1.44}$$

As convolution in time domain corresponds to pointwise multiplication in frequency domain one can also write

$$Y(f) = X(f) \cdot H(f) \tag{1.45}$$

with $Y(f) = \mathcal{F}y(t)$. $H(f)$ is often called transfer function. In applications, the aim is now to estimate this function $h(t)$ or $H(f)$ respectively.
Another convenient property of time invariant system is the fact that the $e^{2i\pi ft}$ are generalized eigenfunctions of any time invariant operator with eigenvalues $H(f)$. This means $Ax = H(f)e^{2\pi ift}$.

The study of linear time invariant systems is also of special interest for the estimation of stationary processes, i.e. a stochastic process for which the joint probability distribution is constant over time. When considering stochastic analysis LTI system concepts are applied to the estimation and description of stationary processes. The definition of a weak stationary process are constant first and second moments, which means the the covariance matrix has a circulant structure. Thus the process can be predicted with constant parameters for all time points t. Time shifts, however, do not span the whole space of $n \times n$ matrices, but form a real subspace. To be able to characterize all operators, an additional component - either frequency shifts in addition to time shifts or dilation and scale as in wavelet analysis is necessary.

1.9.2. Slowly time varying systems and the underspread condition. For linear time variant (LTV) systems, the action of the operator is no longer independent of the time instant t. As a contrary to convolution operators, general linear operators do not have to be normal and usually do not commute. General linear operator theory applies in the continuous setting, the theory of general regular matrices $A \in \mathbb{C}^{2d}$ in the discrete setting.

Most systems in applications do not change arbitrarily fast from one time instant to another but only vary slowly in time. This corresponds to a finite support in frequency in the spreading domain or to a correlation of the operator kernel along the diagonal. The underspread condition imposes in addition to the slowly time varying condition also a restriction to the length of the memory of the operator. To put it in another way, it restricts the number of time shifts (delays) included. An underspread system is not necessarily slowly time varying. A slowly time varying system is narrow with respect to ω.

1.9.2.1. *The strict underspread condition.* The extension of the spreading function of an operator gives information about the operator's length of memory as well as about the fastness of the operators fluctuations. The spreading function of most operators which are relevant in practice is concentrated around the origin. Now the idea is to measure the extension of the spreading function or how much the spreading function is concentrated around the origin. If the support of $\eta(A)$ is compactly supported, such a measure can be the support of $\eta(A)$ or the size of the smallest rectangle containing the support respectively.

Definition 1.59. [83] Let $|\operatorname{supp}(f)|$ denote the area of the support of a function f. Then, if

$$|\operatorname{supp} \eta(A)| < 1, \tag{1.46}$$

the operator A is called strictly underspread.

As contrary to the time invariant case, general linear time varying operators are not uniquely determined by a single observation. Therefore it is even more interesting that according to the paper [86] the strict underspread condition is sufficient and necessary for the identifiability of a linear time varying operator. This means exactly

Theorem 1.60. *[86] Theorem 3.1 The set $\mathbb{H}_{Q_{a,b}}$ of operators with spreading functions of finite support contained in the rectangel of size a times b is identifiable e.g. there is $f \in \boldsymbol{S}_0{}'(\mathbb{R})$ such that $\Psi_f : \mathbb{H}_{a,b} \to L^2(\mathbb{R})$ is bounded and stable where $H_{Q_{a,b}}$ is equipped with the Hilbert Schmidt norm if and only if $ab < 1$.*

1.9.2.2. *Weaker underspread condition.* In applications most operators are almost underspread in the sense that most of the spreading functions energy is concentrated around the origin, but hardly ever meet the strict condition of the last section. There have been different concepts developed to describe the concentration or decay behavior of the spreading function.

As the spreading function is the symplectic Fourier transform of the Kohn-Nirenberg symbol, all these kind of conditions correspond to a certain smoothness in the Kohn -Nirenberg symbol. Matz and Hlawatsch [94] extended the underspread concept by the introduction of weighted integrals of the spreading function of the form

$$m_A := \frac{1}{||\eta(A)||_1} \int_x \int_\omega v(x,\omega)|\eta(A)(x,\omega)|dxd\omega \tag{1.47}$$

$$M_A := \frac{1}{||\eta(A)||_2} \int_x \int_\omega v^2(x,\omega)|\eta(A)(x,\omega)|^2 dxd\omega \tag{1.48}$$

with $\phi(x,\omega)$ being a non negative weight function, which satisfies $\phi(x,\omega) \geq \phi(0,0) = 0$.

This is the normalized weighted L^2 norm of the spreading function. The expressions are finite if and only if the spreading function is in the weighted L^2 space L^2_v.If the spreading function has compact support on a region Ω around the origin, both integrals are bounded by $B = \max_{\lambda\in\Omega} \phi(\lambda)$. The spreading function of the operator being in a weighted L^2_v space corresponds to the Kohn-Nirenberg Symbol σ being in the Sobolev space W^2_v. Besides the weighted L^2 spaces also other function spaces have been proposed. In this context Thomas Strohmer [112] showed that most real world communication channels meet the following conditions:

$$\eta(x,y) = 0 \quad \text{for} \quad |x| > \omega_{\max} \quad \text{and} \tag{1.49}$$

$$|\eta(x,y)| \leq ce^{-a|y|}. \tag{1.50}$$

He also showed that if this conditions are met bay the spreading function, the KNS belongs to $\mathcal{F}L^1_w(\mathbb{R}^2)$ with weight function $w(\eta,u) = e^{|\eta|^b+|u|^b}$ and $0 < b < 1$.

The KNS then also belongs to the weighed Sjörstrand class $M^{\infty,1}_w$. This symbol class is especially attiring, because operators with KNS in $M^{\infty,1}_w$ form a Wiener type Banach algebra [66]. This special structure makes numerically efficient algorithms e.g. for channel equalization possible [112]. Another possibility would be to think of Wiener Amalgam spaces which combine a local boundedness with a global decay condition such as $W(L^\infty,\ell^2)$ or $W(L^\infty,\ell^1)$.

1.10. Gabor Multipliers

1.10.1. Definition.

Definition 1.61. A *Gabor Multiplier* G_m from the Hilbert space $L^2(\mathbb{R}^d)$ to $L^2(\mathbb{R}^d)$ with symbol $m \in L^2(\mathbb{R}^{2d})$ and lattice Λ is given by

$$G_m(f) = \sum_{\lambda\in\Lambda} m(\lambda)\langle f,\pi(\lambda)g_1\rangle\pi(\lambda)g_2. \tag{1.51}$$

It is possible to extend the definition also to general frames, Bessel sequences, continuous frames as shown in [10] [11] [12] [13].

If not otherwise stated, Λ will in the following denote a separable lattice of the form $\Lambda_k = a\mathbb{Z}^d \times b\mathbb{Z}^d$. The functions g_1 and g_2 are analysis and synthesis window respectively. If we choose $g_1, g_2 \in S_0$ and $m \in l^\infty$, then the convergence of the infinite sum is guaranteed. In case other function spaces are chosen this will be explicitely stated. The continuous counterpart of a Gabor Multiplier is the Short Time Fourier Transform multiplier, also called Localization Operator or Anti-Wick operator [13].

Definition 1.62. A *Short Time Fourier Transform Multiplier* (STFT multiplier) H_m form the Hilbert space $L^2(\mathbb{R}^d)$ to $L^2(\mathbb{R}^d)$ with symbol $m \in L^2(\mathbb{R}^{2d})$ is given by

$$H_m f = \int_{\mathbb{R}^{2d}} \mu(\lambda) V_{g_1} f(\lambda) \pi(\lambda) g_2 d\lambda = \int_{\mathbb{R}^{2d}} \mu(\lambda) \big\langle f, \pi(\lambda) g_1 \big\rangle \pi(\lambda) g_2 d\lambda. \tag{1.52}$$

From [57], we know that for windows in $\boldsymbol{S}_0$ the dependence of the Gabor multiplier on its building blocks in particular the symbol, analysis and synthesis window as well as the lattice constants is continuous.

1.10.2. Properties.

1.10.2.1. *Symbols of the Gabor Multiplier.* In order to reveal the special structure of these operators it is useful to have a look at the symbols of the different operator representations [40]. Moreover we will need them in the further calculations. The kernel of a Gabor multiplier is given by

$$\kappa(G_m)(x, \zeta) = \sum_{\lambda \in \Lambda} m(\lambda) g_\lambda(x) g_\lambda(\zeta). \tag{1.53}$$

The Kohn Nirenberg Symbol of a Gabor Multiplier is given by the discrete convolution between the upper symbol $\mathbf{m}$ and the KNS of $P_0 = (g_1 \otimes g_2^*)$ over the lattice Λ. This means

$$\sigma(G_m) = m *_\Lambda \sigma(P_0). \tag{1.54}$$

The Kohn Nirenberg Symbol hence links the investigation of Gabor Multipliers to the setting of a translation invariant spaces [51]. This will be explained further in the next chapter. In addition, the KNS of the rank one operator P_0 is given by

$$\sigma(P_0)(x, \zeta) = \int_{\mathbb{R}^d} g_1(x) \overline{g_2(x-t)} e^{-2\pi i \zeta t} dt = e^{-2\pi i \zeta x} f(x) \overline{\widehat{g}(\zeta)}. \tag{1.55}$$

The spreading function is given by the symplectic Fourier transform of the Kohn Nirenberg symbol. For Gabor multipliers the following equation for the spreading function reveals a lot of structure [42]

$$\eta(G_m)(b, \nu) = \mathscr{M}(b, \nu) V_g g(b, \nu), \tag{1.56}$$

with $\mathscr{M}$ denoting the symplectic Fourier transform of the upper symbol m. As m is sampled on the lattice Λ, the symplectic Fourier transform is a Λ^0 periodic function.

The lattice Λ^0 denotes the adjoint to the lattice Λ. For details on the interconnection of STFTs and spreading funcion, [45] is a good reference.

In the special context of Gabor multipliers there are two more symbols not mentioned before: the lower symbol and the upper symbol of the operator.

Definition 1.63. [40] The *lower symbol* of an operator G with respect to a window $g \in \boldsymbol{S}_0$ and a lattice Λ is defined as:

$$\sigma_L(G)(\lambda) = \langle G, g_\lambda \otimes g_\lambda^* \rangle = \langle G, P_\lambda \rangle \tag{1.57}$$

The idea behind a Gabor multiplier is to suppress some time frequency localized components of a signal. Consequently one would expect the perfectly localized Gauss function and also roughly $\boldsymbol{S}_0-$atoms should behave almost as eigenfunctions. Thus the main part of the operator should be captured by lower symbol. We get the following estimate for the operator norm of $G : ||\sigma_L(G)||_\infty \leq ||G||$ [40].

Definition 1.64. [40] The multiplier sequence $m(\lambda)$ is also called *upper symbol* of the Gabor multiplier G_m. This name is inspired from the estimate of the operator norm of $G : ||G_m|| \leq C||m||_\infty$, if m is bounded with $C \leq C_\Lambda ||g_1||_{S_0} ||g_2||_{S_0}$.

Lemma 1.65. *[40] The upper symbol is uniquely determined if and only if $(P_\lambda)_{\lambda \in \Lambda}$ forms a Riesz Basis within the Hilbert Space of all Hilbert Schmidt operators.*

Moreover, we also know from [15] that P_λ is either a Riesz basis or not a frame. From [10] we also know that for Riesz bases the mapping from the upper symbol to the Bessel multiplier is injective.

The operator properties are closely related to the properties of the upper symbol. We want to give a short summery on this relations. Let g_1 and $g_2 \in \boldsymbol{S}_0$. We know that (g, Λ) forms a Gabor frames for an arbitrary lattice Λ. Then the following relations between the Gabor multiplier G_m and the upper symbol m are established:

- G_m is bounded, if $m \in \boldsymbol{\ell}^\infty(\Lambda)$.
- G_m is Hilbert-Schmidt if $m \in \boldsymbol{\ell}^2(\Lambda)$.
- G_m is a trace class operator on $L^2(\mathbb{R}^d)$ mapping $\boldsymbol{S}_0'(\mathbb{R}^d)$ to $\boldsymbol{S}_0'(\mathbb{R}^d)$ if $m \in \boldsymbol{\ell}^1(\Lambda)$.
- G_m is compact if $m \in \boldsymbol{C}_0$.
- G_m is self adjoint if m is real valued (and $g_1 = g_2$).

For more general results also on Bessel multipliers see again [10]. Proofs for finite dimensional Gabor analysis see [56] [57].

1.11. Finite discrete setting

In the discrete setting, we will always identify $\boldsymbol{C}^N$ with periodic vectors or equivalently functions on the cyclic group $\mathbb{Z}_N$. Therefore time shifts are taken modulo N.

Besides, in this work, we are always considering rectangular lattices Λ of the form $\Lambda = a\mathbb{Z}^A \times b\mathbb{Z}^B$ with $A := \frac{N}{a}$ and $B := \frac{N}{b}$.

Definition 1.66. A linear time invariant filter or convolution operator H is given by

$$Hf(m) = (h * f)(m) = \sum_{l=1}^{N} h(m-l)f(l), \quad m = 1, ..., N \tag{1.58}$$

where $h \in \boldsymbol{C}^N$ is called the *impulse response*. The convolution on the lattice Λ will be denoted with $*_\Lambda$, i.e. for $u = 1, ..., a$ and $\nu = 1, ..., b$

$$(m *_\Lambda k)(u, \nu) = abN \sum_{k=0}^{A-1} \sum_{l=0}^{B-1} m(ak, bl)k(u - ak, \nu - bl). \tag{1.59}$$

Corollary 1.67. *The kernel matrix representation $Hf(t) = K_H(t,\tau)f(\tau)$ of a linear time invariant filter $H : \boldsymbol{C}^N \to \boldsymbol{C}^N$ is given by*

$$K_H(t, \tau) = h(t - \tau) \tag{1.60}$$

Definition 1.68. A *time frequency shift* $\pi(k, l) \in \boldsymbol{C}^{N\times N}$is defined as

$$\pi(k, l)f(m) = T_k M_l f(m) = f(m-k)e^{\frac{2\pi ilm}{N}} \quad k, l \in \mathbb{Z} \quad m = 1, ..., N \tag{1.61}$$

with T_k denoting the *time shift operator*

$$T_k f(m) = f(m - k) \tag{1.62}$$

and M_l being the *modulation operator*

$$M_l f(m) = f(m)e^{\frac{2\pi iml}{N}}. \tag{1.63}$$

Definition 1.69. The discrete Spreading Function $\eta_H \in \boldsymbol{C}^{N\times N}$ of an operator with matrix representation $H \in \boldsymbol{C}^{N\times N}$ in its explicit form is given by

$$\eta_H(k, l) = \sum_{m=0}^{N-1} H(m, m-k)e^{\frac{-2\pi iml}{N}} \tag{1.64}$$

Proposition 1.70. *[8] The time frequency shifts form an orthonormal basis in $\boldsymbol{C}^{N\times N}$. This means every matrix H can be written in a unique way as*

$$H = \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} \eta_K(k, l)\pi(k, l). \tag{1.65}$$

η_H is called the spreading function. Therefore the spreading representation $H \leftrightarrow \eta_H$ is an isometric isomorphism from $\boldsymbol{C}^{N\times N}$ to $\boldsymbol{C}^{N\times N}$.

Definition 1.71. The discrete Gabor transform of a signal $f \in \boldsymbol{C}^N$ with respect to a window function $g \in \boldsymbol{C}^N$ is an operator from $\boldsymbol{C}^N$ to $\boldsymbol{C}^{B\times A}$ given by

$$V_g f(k, l) = \langle f, \pi(ak, bl)g \rangle = \sum_{m=0}^{N-1} f(m)g(m - ak)e^{\frac{2\pi ilm}{B}} \tag{1.66}$$

Definition 1.72. Let (g_1, Λ) be a frame for $\boldsymbol{C}^N$, then $g_2 \in \boldsymbol{C}^N$ is called a *dual window* for (g_1, λ) if

$$f = \sum_{k=0}^{A-1}\sum_{l=0}^{B-1} \langle f, \pi(ak, bl)g_1\rangle \pi(ak, bl)g_2 = \sum_{k=0}^{A-1}\sum_{l=0}^{B-1} \langle f, \pi(ak, bl)g_2\rangle \pi(ak, bl)g_1 \quad \forall f \in \boldsymbol{C}^N \tag{1.67}$$

Definition 1.73. The *canonical dual window* is defined as the dual window having minimal norm. It can be found by several methods [111].

Definition 1.74. Let m be in $\boldsymbol{C}^{N\times N}$ and $g_1, g_2 \in \boldsymbol{C}^N$. Then, a Gabor Multiplier G from the Hilbert Space $\boldsymbol{C}^N$ to $\boldsymbol{C}^N$ is defined as

$$Gf = \sum_{k=0}^{A-1}\sum_{l=0}^{B-1} m(ak, bl)\langle f, \pi(ak, bl)g_1\rangle \pi(ak, bl)g_2. \tag{1.68}$$

In the following we call m the *symbol* of the Gabor Multiplier [8] and g_1, g_2 are the *analysis* and *synthesis* window, respectively.

It is easy to see that the Gabor multipliers with fixed window functions g_1, g_2 and lattice parameters a, b form a subspace of the space of $N \times N$ matrices which we will denote with $\boldsymbol{G}(g_1, g_2, a, b)$.

Theorem 1.75. *The matrix representation $Gf(t) := K_G(t, \tau)f(\tau)$ of a Gabor multiplier $G_m^{g_1,g_2}$ with symbol $m \in \mathbb{R}^{N\times N}$ and window functions $g_1, g_2 \in \boldsymbol{C}^N$ for $f \in \mathbb{R}^N$ is given by*

$$K_G(t, \tau) = \sum_{k=0}^{A-1}\sum_{l=0}^{B-1} m(ak, bl)\overline{g_1(\tau - ak)}g_2(t - ak)e^{\frac{2\pi i bl(t-\tau)}{N}} \tag{1.69}$$

We will in the following identify the operator G with its matrix representation K_G.

Proof: See [12] [9] Let $f = [f(1)...f(N)]$ denote the signal vector and $\boldsymbol{G}_i \in \boldsymbol{C}^{AB\times N}$ be

$$\boldsymbol{G}_i := \begin{pmatrix} g_i \\ T_a M_0 g_i \\ \vdots \\ T_{(A-1)a} M_{(B-1)b} g_i \end{pmatrix} \qquad \text{for} \qquad i = 1, 2 \tag{1.70}$$

the matrix of row vectors containing the shifted and modulated versions of $g_i = [g_i(1)...g_i(N)]$ on the lattice $\Lambda = (ak, bl)$ and $M_{a,b} \in \mathbb{R}^{AB\times AB}$ be defined as $M_{a,b} := diag(m(ak, bl))$ with $a = 0, ..., (A-1)$ and $b = 0, ..., (B-1)$. Then, equation (1.68) can be written as $Gf = \langle\langle f, \boldsymbol{G}_1\rangle M, \boldsymbol{G}_2^*\rangle = \langle f, \boldsymbol{G}_1 M \boldsymbol{G}_2^*\rangle = \langle \boldsymbol{G}_2 \overline{M} \boldsymbol{G}_1^*, f\rangle$ with $\langle \cdot, \cdot\rangle$ denoting the standard complex scalar product. Now, for f and m real, we have $K_G := \boldsymbol{G}_1^* M \boldsymbol{G}_2$ gives us exactly the expression in (1.69). $\square$

Definition 1.76. Let a and b two vectors with $a, b \in \boldsymbol{C}^N$, then we define the *tensor product* $a \otimes b$ as

$$a \otimes b = \begin{pmatrix} a_1 b_1 & a_1 b_2 & \cdots & a_1 b_N \\ a_2 b_1 & a_2 b_2 & \cdots & a_2 b_N \\ & & \ddots & \\ a_N b_1 & a_N b_2 & \cdots & a_N b_N \end{pmatrix}. \tag{1.71}$$

Definition 1.77. The symplectic Fourier Transform $\mathcal{F}^s_\Lambda(m)$ of a matrix $m \in \boldsymbol{C}^{N\times N}$ on the lattice Λ is defined as

$$\mathcal{F}^s(m)(u,v) = \frac{1}{N} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} m(k,l) e^{2\pi i(ku - lv)}. \tag{1.72}$$

The symplectic Fourier transform is a self inverse mapping from $\boldsymbol{C}^{N\times N}$ to $\boldsymbol{C}^{N\times N}$ with $||m||_2 = ||\,\mathcal{F}^s(m)||_2$.

In the following we will denote with $\mathscr{M} := \mathcal{F}^s(m^T)$ the symplectic Fourier transform of the transpose of the lower symbol m of the Gabor multiplier G_m.

Theorem 1.78. *The spreading function $\eta_G \in \boldsymbol{C}^{N\times N}$ of a Gabor multiplier $G_m^{g_1,g_2} \in \boldsymbol{C}^{N\times N}$ on the lattice Λ is given through*

$$\eta_G = B \sum_{l=0}^{a-1} \sum_{k=0}^{b-1} \mathscr{M}(\tau + kB, \nu + lA) V_{g_1} g_2(\tau, \nu) \tag{1.73}$$

Proof: This proof is based on the same idea as for the L^2 and the ℓ^2 case as shown in [44]. Combining the kernel representation of the Gabor multiplier from Proposition (1.75) with the definition (1.69) of the discrete spreading function, we get

$$\begin{aligned}
\eta_G(\tau,\nu) &= \sum_{t=0}^{N-1} K_G(t, t-\tau) e^{\frac{-2\pi i t\nu}{N}} \\
&= \sum_{t=0}^{N-1} \sum_{k=0}^{A-1} \sum_{l=0}^{B-1} m(ak, bl) \overline{g_1(t-\tau-ak)} g_2(t-ak) e^{\frac{2\pi i bl(t-(t-\tau))}{N}} e^{\frac{-2\pi i t\nu}{N}} \\
&= \sum_{k=0}^{A-1} \sum_{l=0}^{B-1} m(ak, bl) e^{\frac{2\pi i bl\tau}{N}} \sum_{t=0}^{N-1} \overline{g_1(t-\tau-ak)} g_2(t-ak) e^{\frac{-2\pi i t\nu}{N}} \\
&= \sum_{k=0}^{A-1} \sum_{l=0}^{B-1} m(ak, bl) e^{\frac{2\pi i bl\tau}{N}} \sum_{t=0}^{N-1} g_2(t) \overline{g_1(t-\tau)} e^{\frac{-2\pi i (t+ak)\nu}{N}} \\
&= \sum_{k=0}^{A-1} \sum_{l=0}^{B-1} m(ak, bl) e^{\frac{2\pi i (bl\tau - ak\nu)}{N}} \sum_{t=0}^{N-1} g_2(t) \overline{g_1(t-\tau)} e^{\frac{-2\pi i t\nu}{N}} \\
&= \sum_{k=0}^{A-1} \sum_{l=0}^{B-1} m(ak, bl) e^{\frac{2\pi i (bl\tau - ak\nu)}{N}} V_{g_1} g_2(\tau,\nu)
\end{aligned}$$

$$= \frac{N}{b}\sum_{k=0}^{b-1}\sum_{l=0}^{a-1}\mathscr{M}(\tau+kB,\nu+lA)V_{g_1}g_2(\tau,\nu)$$

The last step is due to the definition of $\mathscr{M}$ and Poisson summation formula as stated in Proposition(1.5). □

In the following we denote the periodization of $\mathscr{M}$ with $\mathscr{M}_{\mathbf{P}} := \sum_{k=0}^{b-1}\sum_{l=0}^{a-1}\mathscr{M}(\tau + kB, \nu + lA)$, for simplicity of notation.

Definition 1.79. The discrete Kohn Nirenberg symbol (KNS) $\sigma_H \in \boldsymbol{C}^{N\times N}$ of an operator with kernel matrix representation $H(k,n) \in \boldsymbol{C}^{N\times N}$ in its explicit form is given by

$$\sigma_H(k,l) = \sum_{n=0}^{N-1} H(k,k-n)e^{\frac{-2\pi inl}{N}} \tag{1.74}$$

Theorem 1.80. *The Kohn Nirenberg symbol of a linear time invariant filter H is given by $\sigma_H(k,l) = \hat{h}(l)$, $\forall k$.*

Proof: Using the matrix representation of a LTI filter from Theorem (1.67) together with definition (1.79) we get

$$\sigma_H(k,l) = \sum_{n=0}^{N-1} K_H(k,k-n)e^{\frac{-2\pi inl}{N}} = \sum_{n=0}^{N-1} h(k-(k-n))e^{\frac{-2\pi inl}{N}} = \sum_{n=0}^{N-1} h(n)e^{\frac{-2\pi inl}{N}} = \hat{h}(l). \tag{1.75}$$

□

The following theorem and proof is also given in (compare e.g. [103] Prop (2.6). It shall be repeated here as it will be key for further calculations in Chapter (2).

Theorem 1.81. *The Kohn Nirenberg symbol of a Gabor Multiplier $G^m_{g_1,g_2}$ with symbol m, window functions $g_1, g_2 \in \boldsymbol{C}^N$ and lattice parameters a, b is given by*

$$\sigma_G(u,\nu) = m *_\Lambda \sigma(P_0) = abN\sum_{k=0}^{A-1}\sum_{l=0}^{B-1} m(ak,bl)g_1(u-ak)\check{g}_2(\nu-bl)e^{\frac{-2\pi i(u-ak)(\nu-bl)}{N}} \tag{1.76}$$

$$= abN\sum_{k=0}^{A-1}\sum_{l=0}^{B-1} m(u-ak,v-bl)g_1(ak)\check{g}_2(bl)e^{\frac{-2\pi iakbl}{N}} \tag{1.77}$$

Proof: The representation of a Gabor Multiplier in the Kohn Nirenberg domain corresponds to $\sigma_G = m *_\Lambda \sigma_{P_0}$ denoting the Kohn Nirenberg Symbol of $P_0 := (g_1 \otimes g_2)$ with σ_{P_0} being of the form

$$\sigma_{P_0}(x,k) = \sum_{y=0}^{N-1} P_0(x,x-y)e^{\frac{-2\pi iky}{N}} = \sum_{y=0}^{N-1} g_1(x)g_2(x-y)e^{\frac{-2\pi iky}{N}}$$

$$
\begin{aligned}
&= g_1(x) \sum_{y=0}^{N-1} g_2(x-y) e^{\frac{-2\pi i k y}{N}} = g_1(x) \sum_{z=x}^{N-1+x} g_2(z) e^{\frac{-2\pi i k(x-z)}{N}} \\
&= g_1(x) e^{\frac{-2\pi i x k}{N}} \sum_{z=0}^{N-1} g_2(z) e^{\frac{2\pi i k z}{N}} = N g_1(x) \check{g}_2(k) e^{\frac{-2\pi i x k}{N}}
\end{aligned}
$$

Now we have

$$
\begin{aligned}
\sigma_G(u,v) &= m *_\Lambda \sigma_{P_0} = ab \sum_{k=0}^{A-1} \sum_{l=0}^{B-1} m(ak,bl) \sigma_{P_0}(u-ak, v-bl) \\
&= abN \sum_{k=0}^{A-1} \sum_{l=0}^{B-1} m(ak,bl) g_1(u-ak) \check{g}_2(v-bl) e^{\frac{-2\pi i(u-ak)(v-bl)}{N}} \\
&= abN \sum_{k=0}^{A-1} \sum_{l=0}^{B-1} m(u-ak,v-bl) g_1(ak) \check{g}_2(bl) e^{\frac{-2\pi i akbl}{N}}
\end{aligned}
$$

□

Remark 1.82. Note the interconnection between the Kohn Nirenberg Symbol and the Spreading function by symplectice Fourier transform.

$$
\begin{aligned}
\eta_G &= B \mathscr{M}_{\mathbf{P}}(\tau,\nu) V_{g_1} g_2(\tau,\nu) = B \mathscr{M}_{\mathbf{P}}(\tau,\nu) \eta_{P_0}(\tau,\nu) && (1.78) \\
&= \mathcal{F}^s \left\{ \mathcal{F}^s(\mathscr{M}_{\mathbf{P}}) * \mathcal{F}^s(V_{g_1} g_2) \right\} = \mathcal{F}^s \left\{ m_\Lambda * \sigma_{P_0} \right\} = \mathcal{F}^s(\sigma_G) && (1.79)
\end{aligned}
$$

with m_Λ denoting the sampled lower symbol m on the lattice Λ given through

$$
m_\Lambda(k,l) = \begin{cases} ab\, m(k,l) & \text{for} \quad (k,l) \in \Lambda \\ 0 & \text{else} \end{cases} \tag{1.80}
$$

Definition 1.83. The *Frobenius norm* of a matrix A is given by

$$
\|A\|_{\mathrm{Fro}} = \sqrt{\sum_{k=1}^{N} \sum_{l=1}^{N} |a_{ij}|^2}. \tag{1.81}
$$

This is equal to $\sqrt{\mathrm{trace}(A^H A)}$, with A^H being the adjoint to A. The Frobenius norm is the equivalent to the Hilbert Schmidt norm in the finite discrete setting.

Theorem 1.84. *The symbol of the best approximation of an operator H by a Gabor Multiplier $G^m_{g_1,g_2}$ is given through*

$$
\mathscr{M}^B_{\mathbf{P}}(\tau,\nu) = \frac{\sum_{k=0}^{b-1} \sum_{l=0}^{a-1} \eta(\tau+kB, \nu+lA) \overline{V_{g_1} g_2(\tau+kB,\nu+lA)}}{\sum_{k=0}^{b-1} \sum_{l=0}^{a-1} |V_{g_1} g_2(\tau+kB,\nu+lA)|^2}. \tag{1.82}
$$

Proof: (Compare [44]) In order to find the best approximation in Hilbert Schmidt sense we want to find m for which $||H - G_m||_2 \to \min$. Looking at the problem in the spreading domain, we get

$$
\min_{\mathscr{M}_{\mathbf{P}} \in \boldsymbol{C}(\Omega)} ||\eta_H(u,v) - \mathscr{M}_{\mathbf{P}}(u,v) V_{g_1} g_2(u,v)||_2,
$$

with $\boldsymbol{C}(\Omega)$ denoting all matrices $C \in \boldsymbol{C}^{N\times N}$ which are (B,A)-periodic. We calculate

$$||\eta_H - \eta_G||_{\boldsymbol{\mathcal{HS}}} = \sum_{u=0}^{N-1}\sum_{v=0}^{N-1} |\eta(u,v) - \mathscr{M}_{\mathbf{P}}(u,v)V_{g_1}g_2(u,v)|^2$$

If we now think of vectors in $\boldsymbol{C}^{2N}$ instead of matrices in $\boldsymbol{C}^{N\times N}$, we can calculate the least squares solution which gives

$$\mathscr{M}_{\mathbf{P}}^B(\tau,\nu) = \frac{\sum_{k=0}^{b-1}\sum_{l=0}^{a-1} \eta(\tau+kB,\nu+lA)\overline{V_{g_1}g_2(\tau+kB,\nu+lA)}}{\sum_{k=0}^{b-1}\sum_{l=0}^{a-1} |V_{g_1}g_2(\tau+kB,\nu+lA)|^2}.$$

By the Hilbert projection theorem this is the optimal solution. For an explicit formulation of the matrix vector representation in case H is a LTI filter compare also Theorem (2.13) □

Definition 1.85. Let the *Kronecker Delta* δ function be defined through

$$\delta(v) = \begin{cases} 1 & v = 0 \\ 0 & \text{else} \end{cases} \tag{1.83}$$

Theorem 1.86. *The spreading function of a linear time invariant filter H with $Hf = (h * f)$ is given through*

$$\eta_H(u,v) = h(u)\delta(v) \tag{1.84}$$

Proof: From the explicit definition of the spreading function (1.64) we get

$$\eta_H(l,r) = \sum_{n=0}^{N-1} H(n,n-l)e^{-2\pi irn/N}. \tag{1.85}$$

As the kernel of a linear time invariant filter is circulant we get by Poisson formula (compare [65])

$$\sum_{n=0}^{N-1} H(n,n-l)e^{-2\pi irn/N} = \sum_{n=0}^{N-1} h(l)e^{-2\pi irn/N} = h(l)\sum_{n=0}^{N-1} e^{-2\pi irn/N} = h(l)\delta(r). \tag{1.86}$$

□

Theorem 1.87. *Assume in particular H to be a linear time invariant filter. Then, the symbol of the best approximation of H by a Gabor Multiplier G_{g_1,g_2}^m is given through*

$$\mathscr{M}_{\mathbf{P}}^B(\tau,\nu) = \frac{\sum_{k=0}^{b-1}\sum_{l=0}^{a-1} h(\tau+kB)\delta(\nu+lA)\overline{V_{g_1}g_2(\tau+kB,\nu+lA)}}{\sum_{k=0}^{b-1}\sum_{l=0}^{a-1} |V_{g_1}g_2(\tau+kB,\nu+lA)|^2}. \tag{1.87}$$

Proof: According to Theorem (1.86), for a linear time invariant filter the spreading function is of the particular form $\eta_H(u,v) = h(u)\delta(v)$. If we now insert this in formula (1.82), we get exactly (1.87). □

CHAPTER 2

The interconnection between LTI filters and Gabor multipliers

2.1. Introduction

In many real time software applications as e.g. Stx and LEA [7], linear filtering is performed by using Gabor Multipliers instead of performing a convolution in time domain for several reasons. First, multipliers are easy to implement. Then, quasi real time processing with finite support windows is possible. Moreover, in order to implement a filter in a linear setting it is straight forward to do this by masking unwanted frequency components. In many chosen settings the results also closely match the expectations (see Example 3).

The mathematical question is to determine under which conditions a Gabor multiplier is equivalent to a linear time invariant ilter. If exact equivalence is not possible, the interest for a best approximation appears naturally. In this case it is interesting to know how the symbol of the Gabor Multiplier looks like. We also derive error estimates as a function of lattice parameters, and the synthesis window used. Given the transfer function it is then possible to guarantee a certain approximation quality under the condition that the windows and the lattice are chosen appropriately and vice versa.

As a last step we will even loosen the condition of best approximation and consider a more intuitive approximation approach often used in practice, where the mask is chosen as constant in time and equal to the frequency filter in the frequency domain. Also for this so called simple estimation approach as described in our first example we will give an upper bound on the difference between the LTI filter and the Gabor multiplier and study the effectively implemented operator by this approach.

We will mainly study the finite dimensional discrete setting, meaning we are considering signals $f \in \boldsymbol{C}^N$. We chose this setting for a more detailed analysis as it is the one where practitioners work and therefore the setting giving rise to the initial question. The problems under investigation are similar to the ones for the continuous setting, the tools at hand however are sometimes different. Where necessary for a better understanding of the problem structure we will however additionally move to signals in ℓ^2 and L^2 and the suited function spaces.

The focus of the subsequent sections will cover the following aspects of the interconnection between LTI filters (also called Convolution operators or Fourier multipliers) and Gabor multipliers: First we see under which conditions a linear time invariant filter can be represented exactly as a Gabor Multiplier. Then we will slightly loosen the condition of equivalence and consider the best approximation of a LTI filter by a Gabor multiplier in Hilbert Schmidt sense. After an analysis of the respective matrix representation of the projection space of Gabor Multipliers and the shape of the best approximating symbol we move on to give an upper bound for the error between a LTI filter and its best approximation by a Gabor Multiplier.

There has been previous work by Ferenc Weisz on this topic indicating that L^p boundedness of the Gabor multiplier with constant symbol in time implies boundedness of the respective Fourier multiplier [119]. More specific, there has been shown that if the window functions are in the Wiener algebra $W(C, L^1)(\mathbb{R}^d) \supset M_1(\mathbb{R}^d)$ and the symbol $m(x, \cdot)$ is a multiplier for L^p uniformly in x then it is also a STFT multiplier for $W(L^p, L^q)(\mathbb{R}^d)$ independent of q. Then in this context also the work of Dörfler and Torresani [44] on best approximation by Gabor multipliers has to mentionned. Then there is also the work by Schnass on Gabor multipliers in general also pointing out the connection of the KNS to spline type spaces [103] and experimental work of Kirian Döpfler on Gabor multiplier approximation of bandlimited functions. Besides, no work on this particular topic is known to the author.

In the following when we consider the continuous setting, unless otherwise stated, we will choose analysis and synthesis window functions g_i from S_0. Then for every lattice Λ, $(g_i, \lambda)_{\lambda\in\Lambda}$ builds Bessel sequence. This means that the upper frame bound is finite automatically, and therefore the sequence only needs to be independent in order to get a frame. Moreover the condition is rather mild as S_0 is dense in L^2 and has become almost standard in Gabor analysis (compare [42]).

2.2. First example

The first example, we want to give here, is that of a low pass filter as it is often implemented in practice. We choose the frequency response $\hat{h} \in \boldsymbol{C}^N$ equal to the characteristic function, which is 1 on $[-R, R]$ and zero else. The resulting convolution operator is compared to the filter generated by a Gabor Multiplier with symbol $m = \hat{h} \otimes 1$. As analysis and synthesis window for the Gabor Multiplier we choose a Hanning window and the appropriate dual. Both operations are applied to a low frequency random signal f_0 e.g. a random signal with Fourier transform $\hat{f}_0$ supported in an interval $[-c, c]$.

Then, we also do the best approximation by a Gabor Multiplier to the convolution kernel H in Hilbert Schmidt sense with the cut off impulse response $\tilde{h} = \mathcal{F}^{-1}(\hat{h})$ being the finite discrete inverse Fourier Transform of $\hat{h}$. We compare the result to the result of the exact LTI filter as well as to the first approximation as described above. A graphical comparison of the approximation error for the different approaches is shown in figure (2.2)

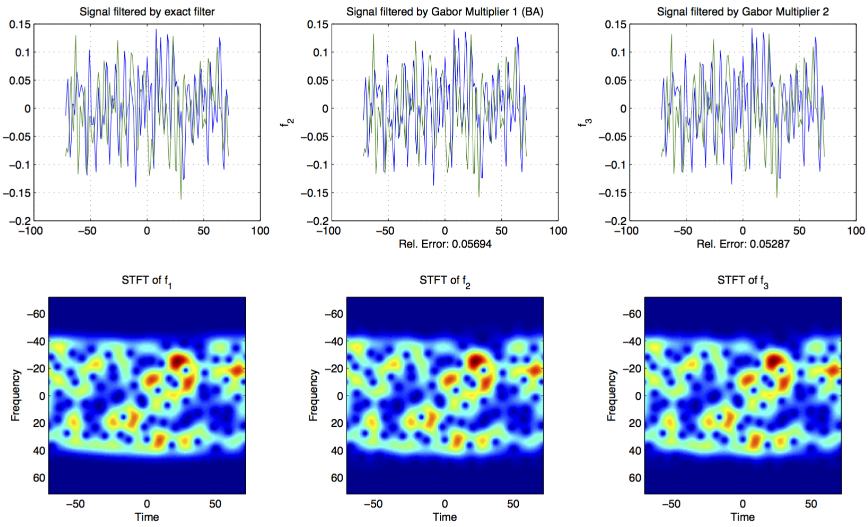

FIGURE 1. Short time Fourier transform of a random signal filtered by a low pass filter with cut off frequency $R = 80$ samples. The filter is once implemented by point wise multiplication in the frequency domain, then with a convolution kernel, its best approximation by a Gabor Multiplier and by a Gabor multiplier with mask $m = \hat{h} \otimes 1$. The relative error is calculated in Frobenius Norm between the Gabor Multiplier and the finite convolution kernel.

2.3. Representation as Gabor Multiplier

Using intuition and visual comparison as an indication like in the first example, the implementation of a LTI filter by a Gabor multiplier seems to work quite well. Given this as a starting point, we are now going to analyze the equivalence of both operators also analytically. In our first theorem,we will see immediately that in particular the most interesting class of perfect filters with characteristic function as frequency response does not qualify as suited candidates for a representation as Gabor multiplier.

Theorem 2.1. *Let $H : L^2 \to L^2$ be a LTI filter with frequency response $\mathcal{F}(h) = \chi_\Omega$. Then H can never be represented exactly as STFT multiplier and*

$$\|H - G\|_{Op} \geq \frac{1}{2} \qquad \forall G(m, g_1, g_2) \quad \textit{with} \quad m \in L^\infty, \quad g_1, g_2 \in L^2 \tag{2.1}$$

Proof: The spreading function is a unitary Banach Gelfand triple isomorphism between $(\boldsymbol{B}, \boldsymbol{\mathcal{HS}}, \boldsymbol{B}')$ and (S_0, L^2, S_0') compare section (1.8). Therefore two operators are identical in $(\boldsymbol{B}, \boldsymbol{\mathcal{HS}}, \boldsymbol{B}')$ if and only if their spreading functions are identical. The spreading function $\eta_H \in S_0'(\mathbb{R}^{2d})$ of a convolution operator $H \in \boldsymbol{B}'(L^2(\mathbb{R}^d), L^2(\mathbb{R}^d))$ with $H = (h * f)$ is calculated as

$$\eta_A(\tau,\nu) = \int_{x\in\mathbb{R}} \hat{h}(x, x-\tau)e^{-2\pi i x\nu} = \int_{x\in\mathbb{R}} \hat{h}(\tau)e^{-2\pi i x\nu} = h(\tau)\delta(\nu) \tag{2.2}$$

with the integral defined in the weak sense. The spreading function of a Gabor Multiplier G is calculated as $\eta_G(\tau,\nu) = \mathscr{M}(\tau,\nu)\cdot V_{g_1}g_2(\tau,\nu)$, where $\mathscr{M}$ is the $(\frac{1}{\beta},\frac{1}{\alpha})$-periodic symplectic Fourier Transform of the symbol m (compare [44]). Therefore a Gabor Multiplier is equivalent to a convolution operator if and only if

$$h(\tau)\delta(\nu) = \mathscr{M}(\tau,\nu)V_{g_1}g_2(\tau,\nu) \qquad \forall(\tau,\nu)\in\mathbb{R}^2 \tag{2.3}$$

As we assume the multiplier symbol m to be time invariant, we have $\mathscr{M}(\tau,\nu) = \mathcal{F}^{-1}(m_2)\otimes\delta$. This then gives then

$$h(\tau) = \mathcal{F}^{-1}(m_2)V_{g_1}g_2(\tau,0) \quad\Leftrightarrow\quad \mathcal{F}h = m_2 * \mathcal{F}\{V_{g_1}g_2(\tau,0)\}\,. \tag{2.4}$$

Further, we can calculate with $\mathcal{P}f(x) := f(-x)$

$$\begin{aligned}\mathcal{F}\{V_{g_1}g_2\} &= \int\int g_2(x)\bar{g}_1(x-\tau)e^{-2\pi i\tau\omega}d\tau dx = \int\int g_2(x)\bar{g}_1(y)e^{-2\pi i(x-y)}dxdy & (2.5)\\ &= \mathcal{F}(g_2)\,\mathcal{F}^{-1}(\bar{g}_1) = \mathcal{F}(g_2)(\mathcal{F}\circ\mathcal{P})(\bar{g}_1) & (2.6)\end{aligned}$$

Therefore we get in total

$$\mathcal{F}(h) = m_2 * \mathcal{F}(g_2)(\mathcal{F}\circ\mathcal{P})(\bar{g}_1) \tag{2.7}$$

By assumption $g_1, g_2 \in L^2$ and therefore also $\mathcal{F}(g_1)$ and $\mathcal{F}^{-1}(\bar{g}_2)$ are in L^2. This gives us that $\mathcal{F}(g_2)(\mathcal{F}\circ\mathcal{P})(\bar{g}_1)$ is in L^1. Further, we have $m \in L^\infty$ and according to Theorem (1.9), the convolution of a function in $f \in L^1$ and a function $g \in L^\infty$ is uniformly continuous. Therefore the expression on the right hand side of equation (2.7) is uniformly bounded and continuous. Let us now denote $\tilde{m}_2(x) := m_2 * \mathcal{F}(g_2)(\mathcal{F}\circ\mathcal{P})(\bar{g}_1)(x)$. Due to the continuity, this means that if $\tilde{m}_2(x_1) \geq \frac{1}{2}$ for an x_1 inside the interval Ω and $\tilde{m}_2(x_2) \leq \frac{1}{2}$ for an x_2 outside the interval Ω, then $\exists x \in \mathbb{R}$ such that $\tilde{m}_2 = \frac{1}{2}$. Further $\mathcal{F}(h)(x)$ is always equal to 0 or 1 and therefore $\exists x \in \mathbb{R}$ such that $|\,\mathcal{F}(h)(x) - \tilde{m}_2(x)| \geq \frac{1}{2}$. On the other hand if $\tilde{m}_2(x) < \frac{1}{2} \quad \forall x \in \Omega$ then $|\,\mathcal{F}(h)(x) - m_2 * \mathcal{F}(g_2)(\mathcal{F}\circ\mathcal{P})(\bar{g}_1)(x)| \geq \frac{1}{2}$ for all $x \in \Omega$. The same is true for all x in $\mathbb{R}\backslash\Omega$ if $\tilde{m}_2(x) > \frac{1}{2} \quad \forall x \in \mathbb{R}\backslash\Omega$. Therefore in any case $\exists x \in \mathbb{R}$ such that

$$|\,\mathcal{F}(h)(x) - m_2 * \mathcal{F}(g_2)(\mathcal{F}\circ\mathcal{P})(\bar{g}_1)(x)| \geq \frac{1}{2} \Rightarrow \sup_x |\,\mathcal{F}(h)(x) - m_2 * \mathcal{F}(g_2)(\mathcal{F}\circ\mathcal{P})(\bar{g}_1)(x)| \geq \frac{1}{2} \tag{2.8}$$

As H and G are multiplication operators in the frequency domain, the operator norm is given by $\|H - G\|_{Op} = \sup_x |\,\mathcal{F}(h)(x) - m_2 * \mathcal{F}(g_2)(\mathcal{F}\circ\mathcal{P})(\bar{g}_1)(x)| \geq \frac{1}{2}$. □

Note that due to the Banach Gelfand Triple isometry, it is also possible to choose m_2 in $\boldsymbol{S}_0{}'$ if the window functions g_1 and g_2 are taken from $\boldsymbol{S}_0$. If $g_1, g_2 \in \boldsymbol{S}_0$, then also $g_0 := \mathcal{F}(g_2)\,\mathcal{F}^{-1}(\bar{g}_1)$ is in $\boldsymbol{S}_0$. Further we have then $h = m_2 * g_0 \in \boldsymbol{S}_0{}'(\mathbb{R}) * \boldsymbol{S}_0 = W(\mathcal{F}L^\infty, \ell^\infty) * W(\mathcal{F}L^1, \ell^1) = W(\mathcal{F}L^1, \ell^\infty) \subseteq W(C^0, \ell^\infty) = C^b$, which means that the implemented frequency response of the LTI filter is always continuous.

In the following theorem we will further show that if we assume the frequency mask m_2 of a STFT multiplier with symbol $m = 1 \otimes m_2$ constant in time to be a reasonable function and we want G to be a bounded linear operator, then m_2 has to be in L^∞. This result is new, even though there are similar results by Cordero and Gröchnig in this direction [33] for window functions $g_1, g_2 \in \mathrm{Sp}(\mathbb{R}^d)$ and $m \in M^{p,\infty}(\mathbb{R}^{2d})$.

Theorem 2.2. *Let $G(m, g_1, g_2)$ be a STFT multiplier with symbol $m = (1 \otimes m_2)$ and window functions $g_1, g_2 \in L^2(\mathbb{R}^d)$. Further assume $m \in (L^1 + L^\infty)(\mathbb{R}^d)$. Then G is a bounded operator $L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ if and only if $m_2 \in L^\infty(\mathbb{R}^d)$.*

Lemma 2.3. *For all $g_0 \in L^1(\mathbb{R}^d)$, $\exists g_1, g_2 \in L^2(\mathbb{R}^d)$ such that $g_0 = \hat{g}_2 \bar{\check{g}}_1$.*

Proof: Every $g_0 \in L^1(\mathbb{R}^d)$ can be written as $g_0 = \sqrt{|g_0|}\sqrt{|g_0|}\frac{g_0}{|g_0|}$. Then we can set $\hat{g}_2 := \sqrt{|g_0|}$. As by assumption g_0 is in $L^1(\mathbb{R}^d)$ also $|g_0| \in L^1\mathbb{R}^d$. The square root is bounded and continuous on $\mathbb{R}^d_+$ and therefore also $\sqrt{|g_0|} \in L^1(\mathbb{R}^d)$. Further we get from $g_0 \in L^1(\mathbb{R}^d)$ that $\int_{\mathbb{R}} \hat{g}_2^2 = \int_{\mathbb{R}^d} \hat{g}_2 \bar{\hat{g}}_2 = \int_{\mathbb{R}^d} \sqrt{|g_0(t)|}\overline{\sqrt{|g_0(t)|}}dt = \int_{\mathbb{R}^d} \sqrt{|g_0(t)|}\sqrt{|g_0(t)|}dt = \int_{\mathbb{R}^d} |g_0(t)|dt \leq C < \infty$ and therefore $\hat{g}_2 \in L^1 \cap L^2$. By Plancherel's Therem it follows that $g_2 \in L^2(\mathbb{R}^d)$. The same can be now done for $\check{\bar{g}}_1 := \sqrt{|g_0|}\frac{g_0}{|g_0|}$. Again $\int_{\mathbb{R}^d} \check{\bar{g}}_1(t)\bar{\check{\bar{g}}}_1(t)dt = \int_{\mathbb{R}^d} \sqrt{|g_0|}\frac{g_0}{|g_0|}\overline{\sqrt{|g_0|}}\frac{\bar{g}_0}{|g_0|}dt$. This is further equal to $\int_{\mathbb{R}^d} \frac{|g_0| g_0 \bar{g}_0}{|g_0|^2}dt = \int_{\mathbb{R}^d} \frac{g_0 \bar{g}_0}{|g_0|}dt = \int_{\mathbb{R}^d} |g_0|dt \leq C < \infty$. This means that $\check{\bar{g}}_1 \in L^1 \cap L^2$ and by applying again Plancherel's Theorem $\bar{g}_1 \in L^2(\mathbb{R}^d)$ As we have $\int_{\mathbb{R}^d} g_1(t)\bar{g}_1(t)dt = \int_{\mathbb{R}^d} \bar{g}_1(t)g_1(t)dt$, this is equivalent to $g_1 \in L^2(\mathbb{R}^d)$. □

Lemma 2.4. *The STFT multiplier $G(m, g_1, g_2)$ with $m = 1 \otimes m_2$ is a bounded operator in $\mathcal{L}(\boldsymbol{B}, \boldsymbol{\mathcal{H}S}, \boldsymbol{B}')$ if and only if $(m_2 * g_0) \in L^\infty(\mathbb{R}^d)$.*

Proof: Following Equation (2.7) the Gabor multiplier $G(m, g_1, g_2)$ is equivalent to a LTI filter with frequency response $\mathcal{F}(h) = m_2 * g_0$. This is equivalent to the mutliplication operator $H : f \to \int_{\mathbb{R}^d}(m_2 * g_0)(t)f(t)dt$, which is bounded if and only if $\exists C : \quad \sup_{||f|| \leq 1} ||Hf|| \leq C \Leftrightarrow ||m_2 * g_0||_\infty = \sup_{||f|| \leq 1} || \int_{\mathbb{R}^d} (m_2 * g_0)(t) f(t)dt|| \leq C$. On the other hand if $||(m_2 * g_0)||_\infty \leq C$ it follows immediately that the multiplication operator is bounded. □

Lemma 2.5. *For $m_2 \in (L^1 + L^\infty)(\mathbb{R}^d)$, $L : g_0 \mapsto m_2 * g_0$ is a bounded linear mapping from $L^1(\mathbb{R}^d)$ to $L^\infty(\mathbb{R}^d)$ with $||m_2 * g_0||_\infty \leq C_{m_2}||g_0||_1$.*

Proof: This can be shown by the closed graph theorem (1.11): If $L : X \to Y$ is a linear operator between two normed spaces then L is closed if and only if whenever a sequence $x_n \to x$ with $x_n \in X$ and $L(x_n) \to y \in Y$ then $x \in X$ and $L(x) = y$. We know that both spaces $L^1(\mathbb{R}^d)$ and $L^\infty(\mathbb{R}^d)$ are Banach spaces and $(m_2 * g_0) \in L^\infty(\mathbb{R}^d)$ by Young's inequality as $m_2 \in (L^\infty + L^1)(\mathbb{R}^d)$ and $g_0 \in L^1(\mathbb{R}^d)$. It remains to show that whenever a sequence $(g_n) \in L^1(\mathbb{R}^d) \to g_0$ and $(m_2 * g_n) \to u \in L^\infty(\mathbb{R}^d)$ then $u = (m_2 * g_0)$. Now $||m_2 * g_n - m_2 * g_0||_{L^1} = ||m_2 * (g_n - g_0)||_{L^1} \leq ||m_2||_{L^1}||g_n - g_0||_{L^1}$ for $m_2 \in L^1(\mathbb{R}^d)$ and $||m_2 * (g_n - g_0)||_{L^\infty} \leq ||m_2||_\infty||(g_n - g_0)||_{L^1}$ again by Young's inequality. Thus we can conclude that $m_2 * g = u$ almost everywhere and by the closed graph theorem, $g_0 \mapsto m_2 * g_2$ describes a bounded linear mapping $L^1(\mathbb{R}^d)$ to $L^\infty(\mathbb{R}^d)$. $\square$

Proof: *[of Theorem (2.2)]*
From Equation (2.7) we know that the Gabor multiplier $G(m, g_1, g_2)$ is equivalent to the multiplication operator $G(h) := \int_{\mathbb{R}^d} \left(m_2 * \hat{g}_2\overline{\hat{g}}_1\right)(t)h(t)dt$. Now, we set $g_0 := \hat{g}_2\overline{\hat{g}}_1$ and consider all $G(h) := \int_{\mathbb{R}^d} (m_2 * g_0)(t)h(t)dt$ with $g_0 \in L^1(\mathbb{R}^d)$. From Lemma (2.3) we know that this is no restriction of generality as every function in $L^1(\mathbb{R}^d)$ can be written in that form as a superposition of $g_1, g_2 \in L^2(\mathbb{R}^d)$. From Lemma (2.4), we know that G is a bounded operator in $\boldsymbol{\mathcal{L}}(\boldsymbol{B}, \boldsymbol{\mathcal{HS}}, \boldsymbol{B}')$ if and only if $(m_2 * g_0) \in L^\infty(\mathbb{R}^d)$. Further we can write

$$\begin{aligned} ||m_2||_{L^\infty} &= \sup_{||f||\leq 1} \left\| \int_{\mathbb{R}^d} f(t)m_2(t)dt \right\| = \sup_{||f||\leq 1} ||m_2(f(t))||, \qquad f \in C_c(\mathbb{R}^d) \quad (2.9) \\ &= \sup_{\rho>0} \sup_{||f||\leq 1} ||m_2(\mathrm{St}_\rho\, g_0 * f)|| \quad (2.10) \end{aligned}$$

for $f \in C_c(\mathbb{R}^d)$ which is dense in $L^1(\mathbb{R}^d)$. Let St_ρ denote the dilation operator. Then we have $||\,\mathrm{St}_\rho\, g_0||_{L^1} \leq |||g_0||_{L^1} \quad \forall \rho > 0$. By applying Fubini's theorem we can switch the order $m_2(\mathrm{St}_\rho\, g_0 * f) = m_2(\mathrm{St}_\rho^\wedge\, g_0 f) = m_2 * (\mathrm{St}_\rho\, g_0(f))$. Therefore we get further

$$||m||_{L^\infty} = \sup_{||f||\leq 1} \sup_{\rho>0} |(m * \mathrm{St}_\rho\, g_0(f)| \quad (2.11)$$

Now we have $||m_2 * \mathrm{St}_\rho\, g_0(f)||_{L^\infty} \leq ||m_2 * \mathrm{St}_\rho\, g_0||_{L^\infty}||f||_{L^1} \leq ||m * \mathrm{St}_\rho\, g_0||_{L^\infty}$ and therefore $||m||_{L^\infty} \leq \sup_{\rho>0} ||m_2 * \mathrm{St}_\rho\, g_0||_{L^\infty}$. Further we can make use of Lemma (2.5) saying that $||m||_{L^\infty} \leq \sup_{\rho>0} C_{m_2}||\,\mathrm{St}_\rho\, g_0||_{L^1} \leq C_{m_2}||g_0||_{L^1} \leq C < \infty$ as $\mathrm{St}_\rho\, g_0 \to g_0$ in $L^1(\mathbb{R}^d)$. On the other hand it is easy to see in the same way that if $(m_2 * g_0) \in L^\infty(\mathbb{R}^d)$, the resulting mutliplication operator is bounded. This direction was also proofed in [91]. $\square$

In the finite discrete case the problem presents itself in a similar way. The following theorem states the necessary conditions on the window functions in order to get perfect equivalence. It can be seen in Theorem (2.8) then, that without subsampling perfect equivalence would in theory be always possible if $\mathrm{supp}(h) \subseteq \mathrm{supp}(g_1 * g_2)$. Numerically, we observe however the same behavior as we have in the continuous case; the error blowing up for h having slower decay than $(g_1 * \bar{g}_2)$.

Theorem 2.6. *A convolution matrix $H \in \boldsymbol{C}^{N\times N}$, with $Hf = (h * f)$ can be represented exactly as a Gabor Multiplier $G_m^{g_1,g_2}$ with lattice constants $a, b \geq 1$, symbol $m \in \boldsymbol{C}^{N\times N}$ and window functions $g_1, g_2 \in \boldsymbol{C}^N$ if and only if $\forall \tau \in \mathrm{supp}(h)$:*

1) $V_{g_1}g_2(\tau, 0) = (g_1 * g_2) \neq 0$
2) $V_{g_1}g_2(\tau + mB, lA) = 0, \quad \forall m \in 0, ..., (b-1), \quad \forall l \in 1, ..., (a-1)$
3) $V_{g_1}g_2(\tau + mB, 0) = 0, \quad \forall m \in 1, ..., (b-1) \quad s.t. \quad (\tau + mB) \notin \mathrm{supp}(h)$
4) $V_{g_1}g_2(\tau + mB, 0) = \frac{h(\tau+mB)}{h(\tau)} V_{g_1}g_2(\tau, 0), \quad \forall m \in 1, ..., (b-1) \quad s.t. \quad (\tau + mB) \in \mathrm{supp}(h)$

Proof: As we already know that the spreading representation is an isometric isomorphism, two operators are identical if and only if their spreading functions are identical. Together with equation (1.73) and equation (1.86) this means $H = G$ if and only if

$$h(\tau)\delta(\nu) = B\mathscr{M}(\tau,\nu)V_{g_1}g_2(\tau,\nu), \quad \forall(\tau,\nu) \tag{2.12}$$

This, in turn is equivalent to

$$\begin{aligned} h(\tau) &= B\mathscr{M}(\tau,0)V_{g_1}g_2(\tau,0), \quad \forall\tau & (2.13)\\ 0 &= \mathscr{M}(\tau,\nu)V_{g_1}g_2(\tau,\nu), \quad \forall\tau, \quad \forall\nu > 0 & (2.14)\end{aligned}$$

From equation (2.13) it follows immediately that

$$V_{g_1}g_2(\tau,0) \neq 0 \quad \forall\tau \in \operatorname{supp}(h). \tag{2.15}$$

which is condition 1). From the periodicity of $\mathscr{M}$ in the frequency domain together with equation (2.14) it follows condition 2). Note that for $\mathscr{M}(\tau,0) = 0$ we have from equation (2.12) $h(\tau) = 0$ and therefore $\tau \notin \operatorname{supp}(h)$. The periodicity of $\mathscr{M}$ in the time domain together with equation (2.13) gives condition 3). Finally through

$$\mathscr{M}(\tau,0) = \frac{h(\tau)}{BV_{g_1}g_2(\tau,0)} = \frac{h(\tau+lB)}{BV_{g_1}g_2(\tau+lb,0)} = \mathscr{M}(\tau+lB) \qquad \forall l = 0,...,b \tag{2.16}$$

we get condition 4). On the other hand we see that if conditions 1) to 4) are fulfilled then we can set

$$\mathscr{M}(\tau,0) = \frac{Bh(\tau)}{V_{g_1}g_2(\tau,0)} \quad \forall\tau \in \operatorname{supp}(h) \cap [0,B) \tag{2.17}$$

and the free parameters of $\mathscr{M}$ to zero, then equation (2.12) is fulfilled. Moreover the symbol m is then of the particular form $m = (\widehat{\chi_{[0,B-1]}a}) \otimes 1$. □

Remark 2.7. Theorem (2.6) can be seen as a special result on the reproducing property, compare [76] or [114] equation (4). If $\operatorname{supp}(h) \subseteq (-B,B)$ we would get perfect reproduction for $V_{g_1}g_2(\tau,0) = 1$ on $\tau \in \operatorname{supp}(h)$ and $V_{g_1}g_2(\tau,\nu) = 0$ outside the fundamental region of the adjoint lattice for $(\tau,\nu) \notin \Omega := \{(-B,B) \times (-A,A)\}$. If $\operatorname{supp}(h) \subset (-B,B)$ is a real subset, the region X with $\operatorname{supp}(h) \subset X \subset (-B,B)$ introduces the freedom to choose $(g_1 * g_2)(\tau)$ having smooth decay on X. As for an LTI filter $\eta(\tau,\nu) = 0 \quad \forall \quad \nu \neq 0$ we have this freedom in the frequency domain for $Y := \{(x,y) : 0 < |y| < A\}$ irrespective of the choice of h.

The conditions given in Theorem (2.6) mirror the central structure of the spreading domain problem. Therefore a visual outline of them is shown in Figure (2.3). The next theorem can be seen as a special case of the last result, having no subsampling e.g. $a = b = 1$. Then the necessary condition reduces to h having faster decay than $(g_1 * g_2)$. It also gives the explicit formula for mask of the STFT multiplier perfectly representing the LTI filter with impulse response h.

Theorem 2.8. *Assume two window functions g_1 and g_2 in $\boldsymbol{C}^N$ with $(g_1 * \bar{g}_2)(\tau) \neq 0, \quad \forall\tau \in \operatorname{supp} h$. Then the convolution matrix $H \in \boldsymbol{C}^{N\times N}$ with $Hf = (h * f)$ can be*

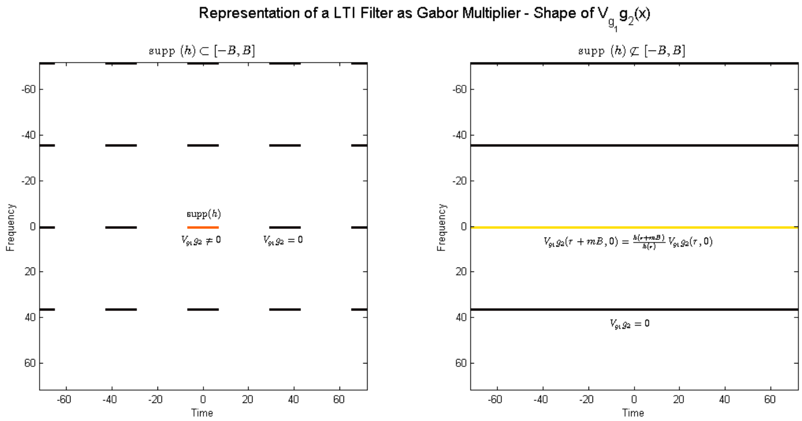

FIGURE 2. This figure gives a visual outline of Theorem (2.6) on the representation of a LTI filter by a Gabor Multiplier. The conditions on the support of $V_{g_1} g_2$ are shown once for $\mathrm{supp}(h) \subset [-B, B]$ and once for $h \not\subset [-B, B]$. Black Lines indicate the regions where the the support of $V_{g_1} g_2$ has to be zero.

represented as STFT multiplier with $a = b = 1$ and lower symbol

$$m = \frac{1}{B} \left\{ \mathbf{1} \otimes \mathcal{F}\left(\frac{h}{g_1 * \bar{g}_2} \right) \right\}. \tag{2.18}$$

Proof: As in the last proof we start from the representation in the spreading domain

$$h(\tau)\delta(\nu) = B\mathscr{M}(\tau, \nu) V_{g_1} g_2(\tau, \nu) \quad \forall(\tau, \nu) \tag{2.19}$$

Now if $m(k, l) = \mathcal{F}\left(\frac{h(\tau)}{BV_{g_1} g_2(\tau, 0)} \right)(k), \forall l$ this means $\mathscr{M}(\tau, \nu) = \frac{h(\tau)}{BV_{g_1} g_2(\tau, 0)} \otimes \delta(\nu)$ and therefore we get

$$\begin{aligned} h(\tau)\delta(\nu) &= B\frac{h(\tau)}{BV_{g_1} g_2(\tau, 0)} \delta(\nu) V_{g_1} g_2(\tau, \nu) \qquad \forall(\tau, \nu) \\ h(\tau) &= \frac{h(\tau)}{V_{g_1} g_2(\tau, 0)} V_{g_1} g_2(\tau, 0) \qquad \forall \tau = 1, ..., N \end{aligned}$$

his is always true for $V_{g_1} g_2(\tau, 0) = (g_1 * \bar{g}_2)(\tau) \neq 0$ on $\mathrm{supp}\, h$. □

This means, given window functions, for which the convolution is bounded away from zero on the support of the impulse response, a LTI filter can always be represented exactly as STFT multiplier. The error between the LTI filter and the Gabor multiplier is the error introduced through subsampling of the mask m. Even though theoretically possible in the finite case for nonzero windows, the numerical behavior reflects the problem of the continuous setting poor calculation efficiency and bad numerical behavior for $(g_1 * \bar{g}_2)$ close to zero. The representation only works well with a bandlimited $\hat{h}$.

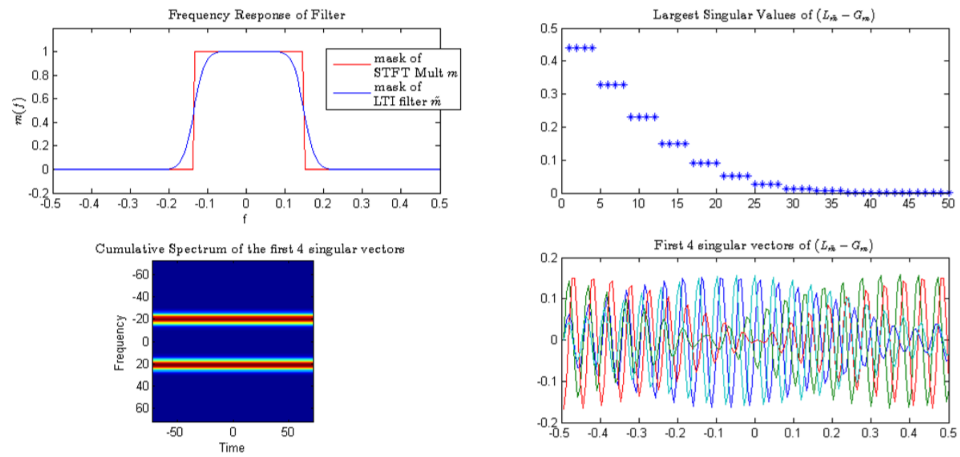

FIGURE 3. The figure shows the effect of implementing a STFT multiplier with mask $m = \chi_\Omega \otimes 1$. The resulting operator is still a LTI operator as long as no subsampling is performed, but the effectively performed masking effect is different. The fist image in the upper left shows the frequency profile as implemented in the STFT multiplier in red vs the frequency mask of the equivalent LTI filter. It can be easily seen that the implementation as STFT multiplier leads to a smoothing effect in the edges of the frequency filter $\tilde{h} = \chi_\Omega * \{\mathcal{F}(g_2)\,\mathcal{F}^{-1}(\bar{g}_1)\}$. The upper right plot shows the largest singular values of the difference between the implemented STFT multplier and perfect low pass filter $G(m, g, 1, 1) - H_{\chi_\Omega}$. In the two plots below the first four singular vectors corresponding to the four largest singular values and their cumulative spectrum are plotted. It can be seen that for signals with energy concentrated at the cut off frequency the difference becomes largest. Reducing the class of signals under consideration to bandlimited functions with bandwiths only conained in Ω would of course reduce the error, but does not make sens in practice as then filtering would not be necessary.

Knowing from Theorem (2.8) that every LTI filter with bandlimited impulse response can be represented as STFT multiplier, we are now turning the focus to the opposite direction, asking whether it is clear that a Gabor multiplier having a mask constant in time is equivalent to a LTI filter. Read ing equation (2.18) the other way round, we see implicitly that a STFT multiplier with time invariant symbol is a convolution operator. The frequency response of this convolution operator, however, is not exactly equal to the frequency mask of the STFT multiplier but smoothed by a convolution with the Fourier transform of the window functions $\mathcal{F}(h) = m_2 * [\mathcal{F}^{-1}(\bar{g}_1)\,\mathcal{F}(g_2)]$. For g_1 and g_2 being real valued, symmetric and tight this reduces to $\hat{h} = m_2 * \hat{g}^2$. In figure (2.3) a visual representation can be found. Smooth window functions have the advantage of preserving the edges of the frequency mask rather well at the cost of a longer time delay needed in return.Theorem (2.9) formalizes this fact.

Theorem 2.9. *If the lower symbol m of a Gabor multiplier $G_m^{g_1,g_2}$ with no time subsampling ($a = 1$) and symmetric window functions $g_1, g_2 \in \boldsymbol{C}^N$ is of the form*

$$m = \left\{ \mathbf{1} \otimes \hat{h} \right\}, \qquad \hat{h} \in \boldsymbol{C}^N \tag{2.20}$$

then its matrix representation is a LTI filter. It's impulse is given by $\tilde{h}(\tau) = \sum_{k=0}^{b-1} h(\tau + kB)(g_1 * g_2)(\tau)$

Proof: We start from the kernel representation of the Gabor multiplier (1.75)

$$\begin{aligned} K_G(t,\tau) &= \sum_{k=0}^{A-1} \sum_{l=0}^{B-1} m(ak,bl) \overline{g_1(\tau - ak)} g_2(t-ak) e^{\frac{2\pi i bl(t-\tau)}{N}} \\ &= \sum_{k=0}^{A-1} \sum_{l=0}^{B-1} \hat{h}(bl) \overline{g_1(\tau - ak)} g_2(t-ak) e^{\frac{2\pi i bl(t-\tau)}{N}} \\ &= \sum_{l=0}^{B-1} \hat{h}(l) e^{\frac{2\pi i bl(t-\tau)}{N}} \sum_{k=0}^{A-1} \overline{g_1(\tau - ak)} g_2(t-ak) \end{aligned} \tag{2.21}$$

Now we see that for the STFT case we have $a = 1$ and therefore we can rewrite the second sum as $\sum_{k=0}^{A-1} g_1(\tau-k) g_2(t-k) = \sum_{k_1=\tau}^{\tau+A-1} \overline{g_1(k_1)} g_2(t-\tau+k_1) = \sum_{k_1=0}^{A-1} \overline{g_1(k_1)} g_2(t-\tau+k_1)$ and for symmetric windows this gives further $\sum_{k_1=0}^{A-1} \overline{g_1(k_1)} g_2(t-\tau-k_1) = (g_1 * \bar{g}_2)(t-\tau)$.

$$K_G(t,\tau) = B \sum_{l=0}^{b-1} h((t-\tau) + lB)(g_1 * \bar{g}_2)(t-\tau), \tag{2.22}$$

which is only dependent of the difference $(t - \tau)$. □

More general results for the continuous setting have been shown in [119]:

Theorem 2.10. *If g_1, g_2 are in the Wiener algebra $W(C, L^1)(\mathbb{R}^d) \supset M_1(\mathbb{R}^d)$ and if $m(\tau, .)$ is a multiplier for $L^p(\mathbb{R}^d)$ uniformly in τ, then it is a STFT multiplier for $W(L^p, L^q)(\mathbb{R}^d)$ independent of q.*

For $g_1, g_2 \in W(L^\infty, L^1)(\mathbb{R}^d) \supset W(C, L^1)(\mathbb{R}^d)$ the STFT multiplier has been defined as a limit in $W(L^p, L^q)(\mathbb{R}^d)$ and the same results have been verified there.

It is important to note that the LTI property is only valid in case of no time subsampling. In case of a common Gabor multiplier with $a > 1$, in contrast, the second sum in equation (2.21) depends on t and is $a-$periodic. Therefore as soon as we have time domain subsampling of the signal, the LTI property of the operator is lost even though the mask being constant in time.

Theorem 2.11. *Let G be the Gabor multiplier with mask $m = 1 \otimes m_2$ and window functions $g_{1,2} \in \boldsymbol{C}^n$ and time subsampling $a > 0$, then G is a LTI operator if g_2 is bandlimited with $V_{g_1} g_2(\tau, \nu) = 0 \ \forall \nu : |\nu| \geq A$.*

Proof: Transferring the problem again to spreading domain, an operator is linear time invariant if and only if the spreading function is of the form $\eta(H) = \hat{h}(\tau)\delta(\nu)$; compare Theorem (1.86). The spreading function of the Gabor multiplier is of the form $\eta(G) = \mathscr{M}(\tau,\nu)V_{g_1}g_2(\tau, nu)$. For $a > 1$, $\mathscr{M}(\tau + kB, lA) = \sum_{j=0}^{b-1} \hat{m}_2(\tau + jB)$ for $k = 0, .., b-1$ and $l = 0, ..., a-1$. Now it can be seen immediately if $V_{g_1}g_2(\tau,\nu) = 0$ $\forall|\nu| \geq A$, then $\eta(G)(\tau,\nu) = \sum_{j=0}^{b-1} \hat{m}_2(\tau + jB)\delta(\nu)$ and therefore time invariant. $\square$

For usual window functions this condition is also necessary, but due to the sparse structure of $\mathscr{M}$ it is only necessary that $V_{g_1}g_2(\tau + kB, lA) = 0$ $\forall \tau \in \mathrm{supp}(m_2)$ and $\forall k = 0, ..., b-1$ $\forall l = 0, ...a-1$. Compare also the conditions of Theorem 2.6. Note further that G then is not really an operator with impulse response evolving in time, but rather with a cyclical behavior in the sense that the impulse response becomes the same every a samples. An example of the size of the impact of this effect can be seen in Figure (2.3).

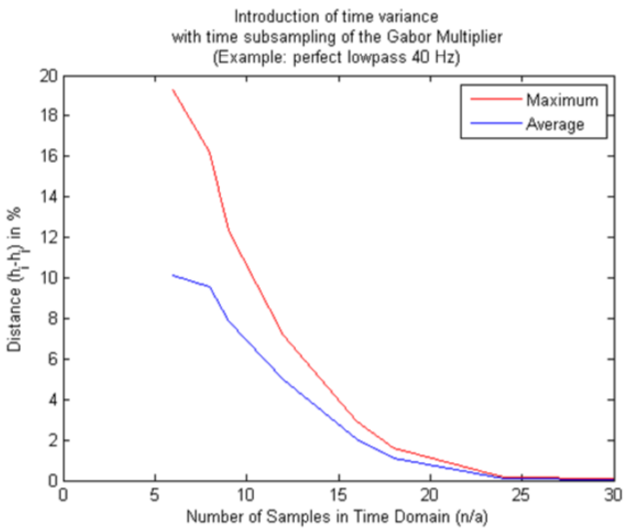

FIGURE 4. The graph shows the effect of time subsampling of the signal on the time invariance property of the Gabor Multiplier. Therefore a Gabor Multiplier with mask $m = (\chi_\Omega \otimes 1)$ and different subsampling rates in time has been implemented. The signal length was chosen to be $n = 144$. Then the distance between the impulse responses $|h(\tau_1, t) - h(\tau_2, t)|$ for the operator was measures and expressed as a percentage value on the y-axis. Mean and maximum values are plotted. For this graph the frequency sampling parameter b was set to 4, but the results are independent of frequency subsampling. Again, differences are mainly located in the frequency cut off region.

An LTI filter can be considered as a special case of a STFT multiplier with degenerated window functions $g_1 = g_2 = 1$. We want to put emphasis also the interconnection between sharp frequency cut off of the filter and smoothness of the window functions corresponding to a time delay in filtering, which we will further highlight in the subsequent chapters. Condition 1) of Theorem 2.6 requires the impulse response h to have a faster decay than $(g_1 * g_2)$. This means that in case we want to have a sharp cut off in the frequency filter $\hat{h}$, which corresponds to a slow decay in h, we have to choose a smooth window function which corresponds to a large time lag.

Having now established a rather clear picture on the interconnection between STFT multipliers and LTI filters during this section, as well as worked out the trade off between time subsampling and the time invariance property, the last theorem will now consider the case for $a, b > 1$. We state the necessary for the conditions of Theorem (2.6) to be fullfilled .

Theorem 2.12. *A convolution matrix $H \in \boldsymbol{C}^{N\times N}$, with $Hf = (h * f)$ can be represented exactly as a Gabor Multiplier $G_m^{g_1,g_2}$ with lattice constants $a, b > 1$, symbol $m \in \boldsymbol{C}^{N\times N}$ and window functions $g_1, g_2 \in \boldsymbol{C}^N$ only if*

$$\sharp\,\mathrm{supp}(h) \leq \frac{n}{ab} \tag{2.23}$$

Proof: We start from condition 1) - 4) of Proposition (2.6). Now let $\mathrm{supp}^B(h)$ be defined as

$$\mathrm{supp}^B(h) := \{\tau : \tau \in [0, B) \wedge \exists k \in [0, b) \quad s.t \quad h(\tau + kB) \neq 0\} \tag{2.24}$$

and $\quad s_h^B = \sharp\,\mathrm{supp}^B(h)$. Then as $V_{g_1}g_2(\tau + kB, lA) = \langle g_2, \pi(\tau + kB, lA)g_1\rangle$, we can write the condtions as a linear system of equations

$$\underbrace{\begin{pmatrix} T_\tau\big(g_1[1] & \cdots & g_1[n]\big) \\ \hline T_{\tau+kB}\big(g_1[1] & \cdots & g_1[n]\big) \\ \hline T_{\tau+uB}M_{lA}\big(g_1[1] & \cdots & g_1[n]\big) \end{pmatrix}}_{G_1} \begin{pmatrix} \overline{g_2}[1] \\ \vdots \\ \overline{g_2}[n] \end{pmatrix} = \underbrace{\begin{pmatrix} c_\tau \\ \hline \frac{h(\tau+kB)}{h(\tau)} c_\tau \\ \hline 0 \end{pmatrix}}_{c_1} \qquad \begin{array}{l} \forall \tau \in \mathrm{supp}^B(h) \\ \forall k = 1, ..., b-1 \\ \forall l = 0, ..., a-1 \\ \forall u = 0, ..., b-1 \end{array} \tag{2.25}$$

As by the definition of a Gabor multiplier (g_1, a, b) is a frame, the Ron-Shen duality principle gives us that (g_1, A, B) is a Riesz basis for its closed linear span. The definition of a Riesz basis gives

$$c \sum_{\lambda\in\Lambda} |a_\lambda|^2 \leq || \sum_{\lambda\in\Lambda} a_\lambda \Pi_\lambda T_\tau g_1||^2 \leq C \sum_{\lambda\in\Lambda} |a_\lambda|^2 \tag{2.26}$$

with $\Lambda = \{(Bl, Ak) : k = 0, ..., a-1, l = 0, ...b-1\}$ and as $|| \sum_{\lambda\in\Lambda} a_\lambda \Pi_\lambda T_\tau g_1||^2 = || \sum_{\lambda\in\Lambda} a_\lambda T_\tau \Pi_\lambda g_1||^2 = ||T_\tau \sum_{\lambda\in\Lambda} a_\lambda \Pi_\lambda g_1||^2 = || \sum_{\lambda\in\Lambda} a_\lambda \Pi_\lambda g_1||^2$ also $(T_\tau g_1, A, B)$ is a Gabor Riesz Basis for its closed linear span. This in turn means that $(T_{\tau+kB})_{k=0,...,b-1}$ is linear independent for all $\tau \in \mathrm{supp}^B(h)$. Considering this, the first two horizontal blocks of system (2.25) give that there can only be a solution for g_2 if

$$h(\tau + kB) = 0 \qquad \forall \tau \in \mathrm{supp}^B(h), \quad \forall k = 1, ..., b-1 \tag{2.27}$$

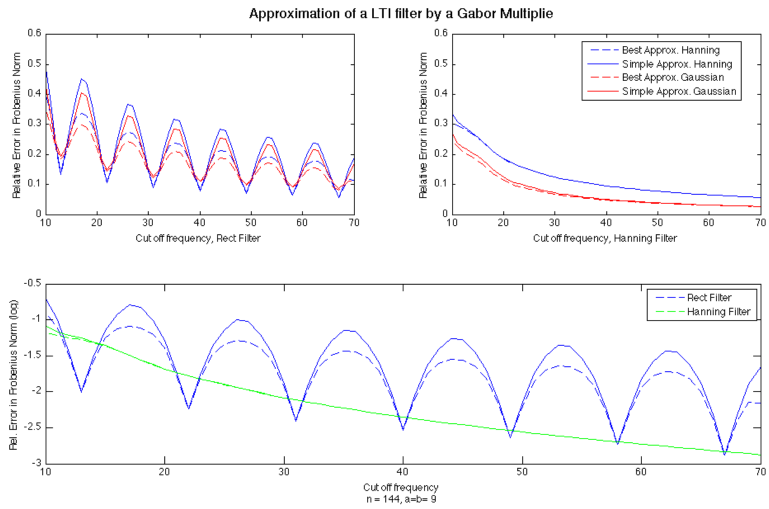

FIGURE 5. Approximation errors of an LTI filter with rectangular frequency response (top left) and Hanning window as frequency response (top right) of different pass regions. As analysis window once a Gaussian and once a Hanning window is used. The synthesis mapping is done with the respective duals.
The parameters are chosen as $N = 144, a = 9, b = 9$. The lower graph confronts the error using different filtering windows. The dashed lines show the error for the best approximation; the solid line for the simple approach. Here a Hanning window is used as analysis window and its dual as synthesis window.

Then the subsystem composed of the two horzional blocks in the bottom of the matrix G_1 give a homogenos system of linear equations with rank min $\left\{N, (ab-1)s_h^B\right\}$. Therefore there can only be a nontrivial solution $g_2 \neq 0$ if $(ab-1)s_h^B < N$. This means $s_h^b \leq$ red. $\square$

As in practice the redundancy is usually chosen close to one, this would mean that h could only be chosen as a multiple of the delta function which corresponds to a constant filter. Therefore we can say that it is not possible to represent a suitable LTI filter for common filtering applications exactly as a Gabor multiplier. Numerically we still have the possibility to make this effect vanish as we will see in the next section.

2.4. Approximations

2.4.1. Numerical result. Figure (5) shows the approximation error of a LTI filter by a Gabor multiplier for different pass frequencies. A rectangular filter is opposed to a Hanning filtering window. It can be seen that if the frequency response is chosen as rectangular window, the typical cyclic behaviour can be observed. This effect produces maximal errors exactly at the cut off frequency being a mutiple of b.

Comparing, the Best Approximation by a Gabor Multiplier to the simple approximation approach - as often used in practice and as will be described in more details in the according section - for a rectangular filter the best approximation seems to be superior to the simple approach. This effect can in particular be observed at the peaks whereas when using a Hanning filtering window the difference of the approaches becomes very small. Compare Figure (5).

2.4.2. Best Approximation. Having Theorem (2.12) in section (2.3) saying that conditions for equivalence of a LTI filter and a Gabor multiplier are hardly ever met in practice, the analysis is extended to approximations. There are different criteria how to find a suited approximation, which are easy construction, numerical efficiency or minimal error. In this section we first go for the minimal error and analyze the best approximation of a LTI filter by a Gabor multiplier in terms of construction, structure and size of the error.

Throughout this work, best approximation in the continuous setting always refers to Hilbert Schmidt norm and to Frobenius norm in the finite discrete setting as defined in Definition (1.83). This choice is motivated by the fact of providing a Hilbert space setting. We will investigate the problem in the different domains of operator representation, in particular Spreading representation, KNS representation and matrix representation to get a clearer picture of the underlying structure. In the infinite dimensional setting a linear time invariant filter is clearly never a Hilbert Schmidt operator. However, as the spreading function is a a Banach Gelfand Triple isometry $(\boldsymbol{B}, \boldsymbol{\mathcal{HS}}, \boldsymbol{B}') \leftrightarrow (S_0, L^2, S_0')$, we can extend the algorithm of best approximation to operators in $\boldsymbol{B}'$ (Compare [57]). We will call this approximation in the following multiplier approximation. For Hilbert Schmidt operators the multiplier approximation will always give us the best approximation, for multipliers in $\boldsymbol{B}'$ we know that it is well defined.

Apart from the section on the matirx structure of the finite dimensional projection space, most of the subsequent theorems can be translated as they are from the finite discrete setting to the continuous setting and vice versa. They will not be written for both settings in order to avoid not necessary redundancy. The setting will however be sometimes interchanged to give an impression how to switch between the settings.

2.4.2.1. *Projection Space.* The term best approximation refers to the approximation of the LTI filter by a linear combination of vectors of a given subspace (the space of Gabor multipliers). In this first section we want to illustrate the vector space on which the orthogonal projection is performed. In finite dimensions this can be done in terms of matrix vector notation.

Theorem 2.13. *Let $\eta_H = [\eta(0,0)\eta(0,1)...\eta(0,N-1)\eta(1,0)\eta(1,1)...\eta(1,N-1)...\eta(N-1,N-1)]^T$ be the vector representation of the spreading function of the operator H. Then, the Best Approximation of H by a Gabor Multiplier is equivalent to the orthogonal projection of η_H on the space spanned by AB vectors of size N^2. The matrix $X \in \boldsymbol{C}^{N^2 \times AB}$ spanning the projection space is of block diagonal form each of the blocks*

$M_{i,j}$ being of size $N \times A$ with the following structure.

$$X := \begin{bmatrix} M_0 & 0 & \dots & 0 \\ 0 & M_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & M_{B-1} \\ M_B & 0 & \dots & 0 \\ 0 & M_{B+1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & M_{2B-1} \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ M_{(b-1)B} & 0 & \dots & 0 \\ 0 & M_{(b-1)B+1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & M_{N-1} \end{bmatrix} \quad M_j := \begin{pmatrix} V_{g_1}g_2(j,0) & 0 & \dots & 0 \\ 0 & V_{g_1}g_2(j,1) & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \dots & 0 & V_{g_1}g_2(j,A-1) \\ V_{g_1}g_2(j,A) & 0 & \dots & 0 \\ 0 & V_{g_1}g_2(j,A+1) & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \dots & 0 & V_{g_1}g_2(j,2A-1) \\ \vdots & \vdots & \vdots & \vdots \\ V_{g_1}g_2(j,(a-1)A) & 0 & \dots & 0 \\ 0 & V_{g_1}g_2(j,(a-1)A+1) & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \dots & 0 & V_{g_1}g_2(j,N-1) \end{pmatrix} \tag{2.28}$$

Proof: The Gramian Matrix $(X'X)$ of size $AB \times AB$ is given by the diagonal matrix

$$(X'X) = \begin{pmatrix} \sum_{l=0}^{b-1} M_{lB}^2 & & 0 & \cdots & 0 \\ 0 & \sum_{l=0}^{b-1} M_{lB+1}^2 & & & \vdots \\ \vdots & & & \ddots & \vdots \\ \vdots & \ddots & \ddots & & \vdots \\ \vdots & & & & 0 \\ 0 & \cdots & 0 & & \sum_{l=0}^{b-1} M_{(l+1)B-1}^2 \end{pmatrix}$$

with the building blocks

$$M_{lB+i}^2 = \operatorname{diag}\left(\sum_{k=0}^{a-1} |V_{g_1}g_2(lB+i, kA+j)|^2 \right)_{j=0,\dots,A-1}$$

Now, further considering the projection matrix $P = (X(X'X)^{-1}X')$ we get

$$\begin{pmatrix} \tilde{M}_0^2 & 0 & \dots & 0 & \dots & \tilde{M}_0\tilde{M}_{(b-1)B} & 0 & \dots & 0 \\ 0 & \tilde{M}_1{}^2 & \ddots & \vdots & \ddots & 0 & \tilde{M}_1\tilde{M}_{(b-1)B+1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 & & \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \tilde{M}_{B-1}^2 & & 0 & \dots & 0 & \tilde{M}_{B-1}\tilde{M}_{N-1} \\ \vdots & & & & \ddots & & & & \\ \tilde{M}_{(b-1)B}\tilde{M}_0 & 0 & \dots & 0 & & \tilde{M}_{(b-1)B}^2 & 0 & \dots & 0 \\ 0 & \tilde{M}_{(b-1)B+1}\tilde{M}_1 & \ddots & \vdots & \ddots & 0 & \tilde{M}_{(b-1)B+1}^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 & & \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \tilde{M}_{N-1}\tilde{M}_{B-1} & \dots & 0 & \dots & 0 & \tilde{M}_{N-1}^2 \end{pmatrix}$$

with the blocks $\tilde{M}_i \tilde{M}_j \in \boldsymbol{C}^{N \times N}$ being of the form

$$
\begin{bmatrix}
\frac{V(i,0)\overline{V(j,0)}}{|V(i,0)|_p^2} & 0 & \cdots & 0 & \cdots & \frac{V(i,0)\overline{V(j,(a-1)A)}}{|V(i,0)|_p^2} & 0 & \cdots & 0 \\
0 & \frac{V(i,1)\overline{V(j,1)}}{|V(i,1)|_p^2} & \ddots & \vdots & \ddots & 0 & \frac{V(i,1)\overline{V(j,(a-1)A+1)}}{|V(i,1)|_p^2} & \ddots & \vdots \\
\vdots & \ddots & \ddots & 0 & & \vdots & \ddots & \ddots & 0 \\
0 & \cdots & 0 & \frac{V(i,A-1)\overline{V(j,A-1)}}{|V(i,A-1)|_p^2} & & 0 & \cdots & 0 & \frac{V(i,A-1)\overline{V(j,N-1)}}{|V(i,A-1)|_p^2} \\
\vdots & & & & \ddots & & & & \\
\frac{V(i,(a-1)A)\overline{V(j,0)}}{|V(i,(a-1)A)|_p^2} & 0 & \cdots & 0 & & \frac{V(i,(a-1)A)\overline{V(j,(a-1)A)}}{|V(i,(a-1)A)|_p^2} & 0 & \cdots & 0 \\
0 & \frac{V(i,(a-1)A+1)\overline{V(j,1)}}{|V(i,(a-1)A+1)|_p^2} & \ddots & \vdots & \ddots & 0 & \frac{V(i,(a-1)A+1)\overline{V(j,(a-1)A+1)}}{|V(i,(a-1)A+1)|_p^2} & \ddots & \vdots \\
\vdots & \ddots & \ddots & 0 & & \vdots & \ddots & \ddots & 0 \\
0 & \cdots & 0 & \frac{V(i,N-1)\overline{V(j,A-1)}}{|V(i,N-1)|_p^2} & \cdots & 0 & \cdots & 0 & \frac{V(i,N-1)\overline{V(j,N-1)}}{|V(i,N-1)|_p^2}
\end{bmatrix}
$$

with

$$
|V_{g_1} g_2(\tau, \nu)|_p^2 = \sum_{k=0}^{a-1} \sum_{l=0}^{b-1} |V_{g_1} g_2(\tau + lB, \nu + kA)|^2 \tag{2.29}
$$

denoting the periodization of $V_{g_1} g_2$. Multiplication by η_H gives

$$
G_{g_1, g_2}^{\Lambda}(\tau, \nu) = \frac{\sum_{k=0}^{a-a} \sum_{l=0}^{b-1} \eta(\tau + lB, \nu + kA) \overline{V_{g_1} g_2(\tau + lB, \nu + kA)} V_{g_1} g_2(\tau, \nu)}{\sum_{k=0}^{a-1} \sum_{l=0}^{b-1} |V_{g_1} g_2(\tau + Bl, \nu + Ak)|^2} \tag{2.30}
$$

which is equivalent to the expression for the best approximation by a Gabor Multiplier as derived in [44]. $\square$

Theorem 2.14. *Let η_H be the spreading function a LTI filter H. Then, the Best Approximation of H by a Gabor Multiplier is equivalent to the orthogonal projection of η_H on the space spanned by B matrices of size $N \times N$ being of the form*

$$
R_k := \begin{bmatrix}
\cdots & V_{g_1} g_2(k, 0) & 0 & \cdots & 0 & V_{g_1} g_2(B + k, 0) & 0 & \cdots & 0 & V_{g_1} g_2((b-1)B + k, 0) & \cdots \\
 & 0 & & & & & & & & 0 & \\
 & \vdots & & & & 0 & & & & \vdots & \\
 & 0 & & & & & & & & 0 & \\
\cdots & V_{g_1} g_2(k, A) & 0 & \cdots & 0 & V_{g_1} g_2(B + k, A) & 0 & \cdots & 0 & V_{g_1} g_2((b-1)B + k, A) & \cdots \\
 & 0 & & & & & & & & 0 & \\
 & \vdots & & & & 0 & & & & \vdots & \\
 & 0 & & & & & & & & 0 & \\
\cdots & V_{g_1} g_2(k, (a-1)A) & 0 & \cdots & 0 & V_{g_1} g_2(B + k, (a-1)A) & 0 & \cdots & 0 & V_{g_1} g_2((b-1)B + k, (a-1)A) & \cdots \\
 & 0 & & & & & & & & 0 & \\
 & \vdots & & & & & & & & \vdots & \\
 & 0 & & & & 0 & & & & 0 &
\end{bmatrix} \tag{2.31}
$$

The projection error $\varepsilon_P = \|(I - P)\eta_H\|_{\text{Fro}}$ is given by

$$
\varepsilon_P^2 = \sum_{\tau=0}^{N-1} \left(h(\tau) - \frac{\sum_{l=0}^{b-1} h(\tau + lB) \overline{V(\tau, +lB, 0)} V(\tau, 0)}{|V(\tau, 0)|_P^2} \right)^2 + \sum_{k=1}^{a-1} \sum_{\tau=0}^{N-1} \left(\frac{\sum_{l=0}^{b-1} h(\tau + lB) \overline{V(\tau + lB, 0)} V(\tau, kA)}{|V(\tau, 0)|_P^2} \right)^2 . \tag{2.32}
$$

with $V(\tau,\nu) = V_{g_1}g_2(\tau,\nu)$ *and* $|V(\tau,\nu)|_p^2 = \sum_{k=0}^{a-1}\sum_{l=0}^{b-1} |V_{g_1}g_2(\tau+lB,\nu+kA)|^2$.

Proof: We start from Theorem (2.13) and use the fact that for a LTI filter the spreading function is of the particular form

$$\eta_H = [h(0)\underbrace{0\dots0}_{N-1}h(1)\underbrace{0\dots0}_{N-1}h(2)\dots h(N-1)\underbrace{0\dots0}_{N-1}]^T. \tag{2.33}$$

Therefore we can reduce the blocks M_iM_j to $M_j(k,l) = 0 \quad \forall k > 1, \forall l$. In case the filter is bandlimited and its support is contained in $[0, B-1]$ also all colums after the first NB columns would reduce to 0. We can now calculate the norm of the projection error $\varepsilon_P^2 = \|(I-P)\eta_H\|_{\text{Fro}}^2$ giving

$$\sum_{\tau=0}^{N-1} \underbrace{\left(h(\tau) - \frac{\sum_{l=0}^{b-1} h(\tau+lB)\overline{V(\tau,+lB,0)}V(\tau,0)}{|V(\tau,0)|_P^2}\right)^2}_{\text{diagonal entries}} +$$

$$\sum_{k=1}^{a-1}\sum_{\tau=0}^{N-1} \underbrace{\left(\frac{\sum_{l=0}^{b-1} h(\tau+lB)\overline{V(\tau+lB,0)}V(\tau,kA)}{|V(\tau,0)|_P^2}\right)^2}_{\text{off diagonal entries}}.$$

□

Finally we can identify the vectors in C^{N^2} spanning the projection space with the matrices in $C^{N\times N}$ and get the representation in (2.31).

2.4.2.2. *Lower Symbol.* Having worked out the structure of the subspace of Gabor Multipliers on which the projection is performed, in section (2.4.2.1) we are now going to have a look at the coefficients, also called lower symbol or mask of the Gabor multiplier.

Proposition 2.15. *The lower symbol of the best approximation of a LTI operator by a Gabor Multiplier is given by*

$$m^B(u,v) = \hat{h} * \left(\frac{V_{g_1}g_2(\tau,0)}{\sum_{k=0}^{b-1}\sum_{l=0}^{a-1} |V_{g_1}g_2(\tau+kB,lA)|^2}\right)^{\wedge}(u) \tag{2.34}$$

Proof: We know from Theorem (1.87) that the periodization of the symplectic Fourier transform of the best approximating symbol is given through

$$\mathscr{M}_{\mathbf{P}}^B(\tau,\nu) = \frac{\sum_{k=0}^{b-1}\sum_{l=0}^{a-1} h(\tau+kB)\delta(\nu+lA)\overline{V_{g_1}g_2(\tau+kB,\nu+lA)}}{\sum_{k=0}^{b-1}\sum_{l=0}^{a-1} |V_{g_1}g_2(\tau+kB,\nu+lA)|^2}.$$

Therefore we have

$$\mathscr{M}^B(\tau,\nu) = \frac{h(\tau)\delta(\nu)\overline{V_{g_1}g_2(\tau,\nu)}}{\sum_{k=0}^{b-1}\sum_{l=0}^{a-1} |V_{g_1}g_2(\tau+kB,\nu+lA)|^2} = \frac{h(\tau)\overline{V_{g_1}g_2(\tau,0)}\delta(\nu)}{\sum_{k=0}^{b-1}\sum_{l=0}^{a-1} |V_{g_1}g_2(\tau+kB,lA)|^2}.$$

The symplectic Fourier is self inverse, which leads to

$$\begin{aligned}
m^B(l,u) &= \mathcal{F}^s(\mathscr{M}^B) = \frac{1}{N}\sum_{k=0}^{N-1}\sum_{l=0}^{N-1}\mathscr{M}^B(\tau,\nu)e^{2\pi i(u\tau-l\nu)}\\
&= \frac{1}{N}\sum_{k=0}^{N-1}\sum_{l=0}^{N-1}\frac{h(\tau)\overline{V_{g_1}g_2(\tau,0)}}{\sum_{k=0}^{b-1}\sum_{l=0}^{a-1}|V_{g_1}g_2(\tau+kB,lA)|^2}\delta(\nu)e^{2\pi i(u\tau-l\nu)}\\
&= \left(\frac{h(\tau)V_{g_1}g_2(\tau,0)}{\sum_{k=0}^{b-1}\sum_{l=0}^{a-1}|V_{g_1}g_2(\tau+kB,lA)|^2}\right)^{\wedge}(u)\\
&= \hat{h}*\left(\frac{V_{g_1}g_2(\tau,0)}{\sum_{k=0}^{b-1}\sum_{l=0}^{a-1}|V_{g_1}g_2(\tau+kB,lA)|^2}\right)^{\wedge}(u)
\end{aligned}$$

□

Having a look at formula (2.34) given in Proposition (2.15), we see two things: first, the best approximating symbol is time independent and second, along the frequency domain it is given by the convolution of the filters frequency response with a function adapting it to the Gabor system. This smoothing function is dependent of the window functions g_1, g_2 and the lattice parameters. In the degenerated case of constant window functions, it becomes equal to the Kronecker delta and convolution is then equal to the frequency response itself. In the general case, deviations from the Konecker delta lead to the typical effects on the edges of the filter' s frequency response. A visual representation of this can be found in figure (2.4.2.2).

Remark 2.16. Note that for the STFT case formula (2.34) of proposition (2.15) reduces to $m^B(u,v) = \hat{h}*\widehat{\left(\frac{1}{|V_{g_1}g_2(\tau,0)|}\right)}$. For $\mathrm{supp}(g_1 * g_2) \subseteq \mathrm{supp}(h)$ this gives exactly the representation of Theorem (2.8).

2.4.2.3. *Spreading Representation.* Having seen the structure of the projection space and the best approximating mask in the last two sections, we want to analyze the difference between the LTI filter and its best approximation by a Gabor Multiplier further in this section. The aim is to find an estimate for the upper bound of the difference in Hilbert Schmidt norm as well as to identify the main drivers for the size of the error. The problem can be transferred one to one in the different domains of operator representation kernel representation, KNS representation and spreading representation.

We start our analysis with the spreading representation. For LTI filters this representation suits very well one due to the fact that the support of a LTI filter in the spreading domain is given by the size of the support of the filter's impulse response only. All entries of the spreading matrix are situated on one single row. Given this special sparse form, the structure of the differences and periodization effects becomes more easy to handle. Equation (2.35) of Theorem (2.17) first structures the difference

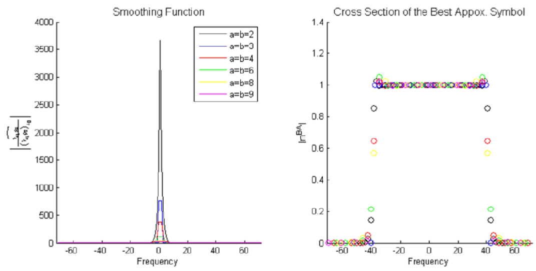

FIGURE 6. Outline of the shape of the lower symbol of the best approximation of a LTI filter by a Gabor multiplier. The cross section of the symbol m^{BA} is plotted for different sampling distances a and b on the right side of the figure. The dimension n is set to 144. The symbol m^{BA} is the convolution of the filter's frequency response $\hat{h}$ with a smoothing function $\left(\frac{\widehat{V_{g_1}g_2(\tau,0)}}{\sum_{k=0}^{b-1}\sum_{l=0}^{a-1}|V_{g_1}g_2(\tau+kB,lA)|^2}\right)$ plotted on the left side of the figure. It can be seen that this function is smooth for a large sampling grids and apporaches the δ function for a denser grid.

in a nice way making apparent the dependence on the periodization effects in the second part of the estimate.

Theorem 2.17. *The error $||H-G||_F^2$ for the best approximation of a convolution operator $H : \boldsymbol{C}^N \mapsto \boldsymbol{C}^N$, $Hf = (h * f)$ by a Gabor Multiplier $G_m^{g_1,g_2}$ with lattice constants a and b is given by*

$$N\sum_{\tau=0}^{B-1}\sum_{m=0}^{b-1}|h(\tau+mB)|^2\left|1-\frac{|\sum_{k=0}^{b-1}h(\tau+kB))\overline{V_{g_1}g_2(\tau+kB,0)}|^2}{\sum_{k=0}^{b-1}|h(\tau+kB)|^2\sum_{k=0}^{b-1}\sum_{l=0}^{a-1}|V_{g_1}g_2(\tau+kB,lA)|^2}\right| \tag{2.35}$$

for $\sum_{k=0}^{b-1}h(\tau+kB)\neq 0\forall\tau\in[0,E)$ with $E=\max_{\tau\in\text{supp}(h)}\tau$

Proof: Theorem and proof are deduced from Corollary 1 in [44] for the L^2 case. Let $\|\cdot\|_F$ denote the Frobenius norm. Then, we are again considering the spreading domain.

$$\varepsilon=\|H-G\|_F^2=N\|\eta_H-\eta_G\|_F^2=\sum_{\tau=0}^{N-1}\sum_{\nu=0}^{N-1}|h(\tau)\delta(\nu)-B\mathscr{M}^B(\tau,\nu)\cdot V_{g_1}g_2(\tau,\nu)|^2 \tag{2.36}$$

with $\mathscr{M}^B$ denoting the symplectic Fourier Transform of the best approximating symbol of the according Gabor Multiplier G. Then we look at the periodization of equation (2.36)

$$\varepsilon=N\sum_{\tau=0}^{B-1}\sum_{\nu=0}^{A-1}\left(\sum_{m=0}^{b}\sum_{l=0}^{a}|h(\tau+mB)\delta(\nu+lA)-B\mathscr{M}^B\cdot V_{g_1}g_2(\tau+mB,\nu+lA)|^2\right) \tag{2.37}$$

As G is the orthogonal projection of H on the closed subspace of Gabor Multipliers, we have $\langle H,G\rangle=\langle\eta_H,\eta_G\rangle=||\eta_G||^2$ and therefore $||\eta_H-\eta_G||^2=||\eta_K||^2-||\eta_G||^2$. We get further

$$
\begin{aligned}
\varepsilon &= N\sum_{\tau=0}^{B-1}\sum_{\nu=0}^{A-1}\left(\sum_{m=0}^{b-1}\sum_{n=0}^{a-1}|h(\tau+mB)\delta(\nu+nA)|^2 - B|\mathscr{M}^B|^2\sum_{k=0}^{b-1}\sum_{l=0}^{a-1}|V_{g_1}g_2(\tau+kB,\nu+lA)|^2\right)\\
&= N\sum_{\tau=0}^{B-1}\sum_{\nu=0}^{A-1}\left(\sum_{m=0}^{b-1}\sum_{n=0}^{a-1}|h(\tau+mB)\delta(\nu+nA)|^2\right.\\
&\qquad\left.\left|1-\frac{|\sum_{k=0}^{b-1}\sum_{l=0}^{a-1}h(\tau+kB)\delta(\nu+lA)\overline{V_{g_1}g_2(\tau+kB,\nu+lA)}|^2}{\sum_{k=0}^{b-1}\sum_{l=0}^{a-1}|h(\tau+kB)\delta(\nu+lA)|^2\sum_{k=0}^{b-1}\sum_{l=0}^{a-1}|V_{g_1}g_2(\tau+kB,\nu+lA)|^2}\right|\right) \qquad (2.38)\\
&= N\sum_{\tau=0}^{B-1}\sum_{\nu=0}^{A-1}\left(\sum_{m=0}^{b-1}|h(\tau+mB)|^2\left|1-\frac{|\sum_{k=0}^{b-1}h(\tau+kB)\overline{V_{g_1}g_2(\tau+kB,0)}|^2}{\sum_{k=0}^{b-1}|h(\tau+kB)|^2\sum_{k=0}^{b-1}\sum_{l=0}^{a-1}|V_{g_1}g_2(\tau+kB,\nu+lA)|^2}\right|\right)\\
&= N\sum_{\tau=0}^{B-1}\sum_{m=0}^{b-1}|h(\tau+mB)|^2\left|1-\frac{|\sum_{k=0}^{b-1}h(\tau+kB)\overline{V_{g_1}g_2(\tau+kB,0)}|^2}{\sum_{k=0}^{b-1}|h(\tau+kB)|^2\sum_{k=0}^{b-1}\sum_{l=0}^{a-1}|V_{g_1}g_2(\tau+kB,lA)|^2}\right|
\end{aligned}
$$

□

Starting from the spreading representation of a Gabor multiplier, it becomes transparent that the approximation error is due to the interplay of two phenomaenon

- The first part introduced through cut off of the filters impulse response avoiding periodization effects due to sampling. We will denote this by ε_h.
- The second part caused by non perfect reconstruction of the spreading function through $V_{g_1}g_2$. If $V_{g_1}g_2$ would be compactly supported with $\mathrm{supp}(\eta_K) \subseteq \mathrm{supp}\, V_{g_1}g_2 \subseteq \Omega$, we would have a perfect representation for underspread operators with $\mathrm{supp}(\eta_k) \subseteq \Omega$ by Gabor multipliers. Due to the uncertainty principle, however, indicating that $V_{g_1}g_2$ can never be simultaneously of finite support in the time and in the frequency domain, this can never be fullfilled exactly for $a, b > 1$. We will denote the part of the error induced by this ε_{g_1,g_2}.

Theorems (2.18) to (2.20) work out estimates for the two components separately and Theorem (th:estTotal) puts them together as main result of this section.

Theorem 2.18. *The error for the best approximation of a convolution operator $H \in \boldsymbol{C}^{N\times N}$, $Hf = (h * f)$ by a Gabor Multiplier $G_m^{g_1,g_2}$ with lattice constants a and b and window functions g_1, g_2 such that*

$$V_{g_1}g_2(\tau,\nu) = 0, \quad \forall(\tau,\nu) \notin \{([0,A)\times[0,B)\} \quad \textit{and} \quad V_{g_1}g_2(\tau,0) > 0 \quad \forall\tau \in [0,B) \tag{2.39}$$

is given by

$$\frac{||H-G||_F}{||H||_F} = \frac{||h||_{\tau\geq B}}{\|h\|} \tag{2.40}$$

Proof: Again we consider the problem in spreading domain. Together with the finite support of $V_{g_1}g_2$, we get

$$\varepsilon^2 \;=\; N\sum_{m=0}^{b-1}\sum_{\tau=0}^{B-1}|h(\tau+mB)|^2 - \frac{|h(\tau)\overline{V_{g_1}g_2(\tau,0)}|^2}{|V_{g_1}g_2(\tau,0)|^2} = N\sum_{\tau=0}^{B-1}\sum_{k=0}^{b-1}|h(\tau+kB)|^2 - |h(\tau)|^2$$

$$= \quad N \sum_{\tau=0}^{B-1} \sum_{k=1}^{b-1} |h(\tau+kB)|^2 = N \sum_{\tau=B}^{N-1} |h(\tau)|^2$$

Dividing by $||H||^2 = N||h||^2$ and taking the square root gives equation (2.40) $\square$

For a compactly supported Short Time Fourier transform of the window functions and lattice parameters such that the support of $V_{g_1}g_2$ matches the fundamental region of the adjoint lattice, the error only depends on the energy of the part of h not contained in the support of $V_{g_1}g_2$. In order to get an estimate dependent of b analogue to Theorem (2.21) , it is of course possible to assume h in an L^p space and proceed with the estimate as in Theorem (2.21).

Theorem 2.19. *Assume a rectangular lattice Λ with lattice constant $a, b \geq 1$, an analysis window $g_1 \in \boldsymbol{C}^N$, the canonical dual as synthesis window $g_2 := S^{-1}g_1$ and a bandlimited filter function such that h is compactly supported and* $\mathrm{supp}(h) \subset [0, B)$. *Choose the variable $0 \leq c_g(a,b) \leq 1$ such that*

$$1 - c_g(a,b) \leq \frac{|V_{g_1}g_2(\tau,0)|^2}{\sum_{k=0}^{b-1}\sum_{l=0}^{a-1} |V_{g_1}g_2(\tau+kB,lA)|^2} \leq 1 \qquad \forall \tau \in \mathrm{supp}(h) \tag{2.41}$$

Then, the relative error between the LTI filter $H : \boldsymbol{C}^N \mapsto \boldsymbol{C}^N$, with $Hf = (h * f)$ and its Best Approximation by a Gabor multiplier $G := G^{\Lambda}_{g_1,g_2}$ can be estimated by

$$\frac{\|H - G\|_{\text{Fro}}}{\|H\|_{\text{Fro}}} \leq \sqrt{c_g(a,b)} \tag{2.42}$$

Proof: We start from Theorem (2.17) disregarding the periodization effects due to the finite support.

$$\|H - G\|_F^2 \leq N \sum_{\tau=0}^{B-1} |h(\tau)|^2 \left| 1 - \frac{|V_{g_1}g_2(\tau,0)|^2}{\sum_{k=0}^{b-1}\sum_{l=0}^{a-1} |V_{g_1}g_2(\tau+kB,lA)|^2} \right|$$

As $\mathrm{supp}(h) = [0, E)$ with $E \leq B$ by assumption, we have

$$\begin{aligned}
\|H - G\|_F^2 &\leq N \sum_{\tau=0}^{E-1} |h(\tau)|^2 \max_{\tau \leq E} \left| 1 - \frac{|V_{g_1}g_2(\tau,0)|^2}{\sum_{k=0}^{b-1}\sum_{l=0}^{a-1} |V_{g_1}g_2(\tau+kB,lA)|^2} \right| \\
&\leq N\|h\|^2 \max_{\tau \leq E} \left| 1 - \frac{|V_{g_1}g_2(\tau,0)|^2}{\sum_{k=0}^{b-1}\sum_{l=0}^{a-1} |V_{g_1}g_2(\tau+kB,lA)|^2} \right| \qquad (2.43) \\
&\leq N\|h\|^2 |1 - (1 - c_g(a,b)| \leq N\|h\|^2 c_g(a,b)
\end{aligned}$$

Division by $||H||^2 = N||h||^2$ and taking the square root gives the result. $\square$

Figure (2.4.2.3) shows the function $\mathcal{V}(g_1, g_2, a, b) := \frac{|V_{g_1}g_2(\tau,0)|^2}{\sum_{k=0}^{a-1}\sum_{l=0}^{b-1} |V_{g_1}g_2(\tau+kB,lA)|^2}$. It can be seen that the the decay of the correlation of the window functions becomes smoother inside the relevant interval $(-B, B)$ through the division by the periodization. The choice of the canonical dual as synthesis window or using tight windows guarantees optimal adaption to the lattice. It becomes obvious that the approximation works well as long as E is well contained in the interval $(-B/2, B/2)$. It has also been confirmed by experiments that this error component stays small as long as the essential support of h is contained in $(-B/2, B/2)$ and window and lattice fit acceptably well to each other (e.g not extremely asymetric lattices combined with the standard Gaussian window or windows assymetric in time and frequency domain combined with a regular lattice) also for low redundancies.

In case $V_{g_1}g_2$ has only very moderate decay on E the above estimate works very well. On the other hand, if $V_{g_1}g_2$ decays considerably on E because for example the windows were not chosen dual to each other it doesn't give appropriate results. In that cast it is better to apply the supremum on the filter function and the integral on the function $\mathcal{V}(g_1, g_2, a, b)$ as in Theorem (2.4.2.3)

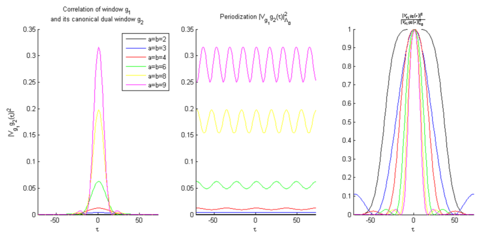

FIGURE 7. The figure shows the squared absolute correlation $|V_{g_1}g_2(\tau, 0)|^2$ of a Hanning window g_1 of size $N = 144$ with its canonical dual window for different lattice parameters a and b on the left side. In the middle the periodization $|V_{g_1}g_2|^2_{\Lambda_0}(\tau, 0) = \sum_{k=0}^{a-1}\sum_{l=0}^{b-1}|V_{g_1}g_2(\tau + kB, lA)|^2$ is plotted. The right plot shows the function $\frac{|V_{g_1}g_2(\tau,0)|^2}{\sum_{k=0}^{a-1}\sum_{l=0}^{b-1}|V_{g_1}g_2(\tau+kB,lA)|^2}$.

Theorem 2.20. *The error between a convolution operator* $H \in \boldsymbol{B}'(L^2, L^2)$, $\quad Hf = (h * f)$ *with* $\operatorname{supp}(h) \subseteq (0, \frac{1}{\beta}]$ *and its multiplier approximation by* $G_m^{g_1,g_2}$ *can be estimated by*

$$||H - G||_{\boldsymbol{H}} \leq \frac{\sup_\tau |h(\tau)|}{||g_1|| ||g_2||} \int_{|x| \geq \frac{1}{\beta}} \int_{|y| \geq \frac{1}{\alpha}} |V_{g_1}g_2(x, y)|^2 \tag{2.44}$$

Proof: Using the already described Banach Gelfand triple isometry of the spreading representation gives $||H - G||^2_{\mathcal{HS}} = ||\eta(H) - \eta(G)||_{L^2(\mathbb{R}^{2d})}$

$$\int_{|\tau| \leq \frac{1}{\beta}} \sum_{k \in \mathbb{Z}} |h(\tau + \frac{k}{\beta})|^2 \cdot \left| 1 - \frac{|\sum_{k\in\mathbb{Z}} h(\tau + \frac{k}{\beta}) V_{g_1}g_2(\tau + \frac{k}{\beta}, 0)|^2}{\sum_{k\in\mathbb{Z}} |h(\tau + \frac{k}{\beta})|^2 \sum_{k\in\mathbb{Z}}\sum_{l\in\mathbb{Z}} |V_{g_1}g_2(\tau + \frac{k}{\beta}, \frac{l}{\alpha})|^2} \right|^2$$

This is also a special case of Theorem (19) in [45] and more details can also be found there. Then we can calculate further

$$||H - G||^2_{\mathcal{HS}} \leq \sup_\tau |\sum_{k\in\mathbb{Z}} h(\tau + \frac{k}{\beta})|^2 \cdot \left\| 1 - \frac{|\sum_{k\in\mathbb{Z}} h(\tau + \frac{k}{\beta}) V_{g_1}g_2(\tau + \frac{k}{\beta}, 0)|^2}{\sum_{k\in\mathbb{Z}} |h(\tau + \frac{k}{\beta})|^2 \sum_{k\in\mathbb{Z}}\sum_{l\in\mathbb{Z}} |V_{g_1}g_2(\tau + \frac{k}{\beta}, \frac{l}{\alpha})|^2} \right\|^2$$

Due to the finite support assumption, the expression reduces to

$$||H-G||^2 \leq \sup_\tau |h(\tau)|^2 \cdot \left\| 1 - \frac{|V_{g_1}g_2(\tau,0)|^2}{\sum_{k\in\mathbb{Z}}\sum_{l\in\mathbb{Z}} |V_{g_1}g_2(\tau+\frac{k}{\beta},\frac{l}{\alpha})|^2} \right\|^2 \tag{2.45}$$

$$\leq \sup_\tau |h(\tau)|^2 \cdot \left\| \frac{\sum_{k\in\mathbb{Z}\backslash 0}\sum_{l\in\mathbb{Z}\backslash 0} |V_{g_1}g_2(\tau+\frac{k}{\beta},\frac{l}{\alpha})|^2}{\sum_{k\in\mathbb{Z}}\sum_{l\in\mathbb{Z}} |V_{g_1}g_2(\tau+\frac{k}{\beta},\frac{l}{\alpha})|^2} \right\|^2 \tag{2.46}$$

Therefore we get in total

$$||H-G|| \leq \frac{\sup_\tau |h(\tau)|}{||g_1||||g_2||} \int_{|x|\geq\frac{1}{\beta}} \int_{|y|\geq\frac{1}{\alpha}} |V_{g_1}g_2(x,y)|^2 \tag{2.47}$$

□

Theorem 2.21. *The error between a convolution operator $H \in \boldsymbol{B}'(L^2,L^2)$, $Hf = (h*f)$ with $\operatorname{supp}(h) \subseteq [0,\frac{1}{\beta})$ and a Gabor multiplier $G_m^{g_1,g_2}$ with $g_1 \in M^1_{v_p}$ and $g_2 \in M^2_{v_p}$ can be estimated by*

$$||H-G||_{\mathcal{HS}} \leq C \cdot \sup_\tau |h(\tau)| \max(\alpha,\beta)^p \frac{||g_1||_{M^1_{v_p}}||g_2||_{M^2_{v_p}}}{||g_1||||g_2||} \tag{2.48}$$

Proof: First we use the estimate of Theorem (2.20), then apply Lemma (1.54) and use the fact that $\|Vg_1g_2\| \in L^2_v$ is by its definition equivalent to $\int_{z\in\mathbb{R}^2} Vg_1g_2(1+|z|)^p dz < C$. Therefore we have

$$\int_{|x|\geq\frac{1}{\beta}} \int_{|y|\geq\frac{1}{\alpha}} V_{g_1}g_2 dz \leq \int_{|z|>\min(\frac{1}{\beta},\frac{1}{\alpha})} \frac{V_{g_1}g_2(1+|z|)^p}{(1+|z|)^p}$$
$$\leq \frac{\|V_{g_1}g_2\|_{L^2_v}}{\min(\frac{1}{\beta},\frac{1}{\alpha})^p} \leq \max(\alpha,\beta)^p \|V_{g_1}g_2\|_{L^2_v}$$

□

It is possible to refine the estimate of Theorem (2.21) if there is further information on the problem available e.g. by adjusting the weight to the window function and not choose it in a radial symmetric manner. Even though the support condition of Theorem (2.19) and the following mostly not being met exactly, the numerical reality is not too far away in the sense that the energy of the transfer function is numerically equal to zero outside a small region of essential support. After having put the restriction on the support of h, we will consider the the second error component and put the support condition on the time and frequency support of the window function in the next Theorem (2.18). Again, theory tells us that due to the uncertainty principle this condition will never be met exactly. On the other hand, starting again from a numerical point of view, the entries of $|V_{g_1}g_2|$ are so close to zero that it gets not distinguishable whether it is fulfilled or not for a number of common applications and therefore worth considering.

Knowing the results of the special cases with finite support out of the last two theorems, we are now going to consider the general setting. Therefore, we loosen the finite support condition on h moving to the notion of the essential support E. At the same time, we describe the concentration of $V_{g_1}g_2$ by putting it in relation to its periodization on the adjoint lattice. It is worth noting that the split in the essential support and the outside region is necessary for most applications. Usually E, however, is not the limiting parameter as for common filtering functions it can be chosen rather small. In the following, we propose suited values for the size of essential support of common filter functions in order to get useful error estimates. These results are stated independent of the signal dimension N to be of utmost utility to the reader.

Assume therefore, we have a signal of length N samples, which needs to be filtered with a maximum pass frequency of P. Then the percentage value e_P of the number of samples N will be given. Therefore E as used in the preceding theorem, can be obtained by $E = e_P N$. On the other hand, as already indicated by the usage of P as a subscript index, the choice of the essential support also depends on the size of the pass region. We will use here a pass region of $P = 100Hz$, but these results can be easily transferred by $e_{P_2} = \frac{P_1}{P_2} e_{P_1}$.

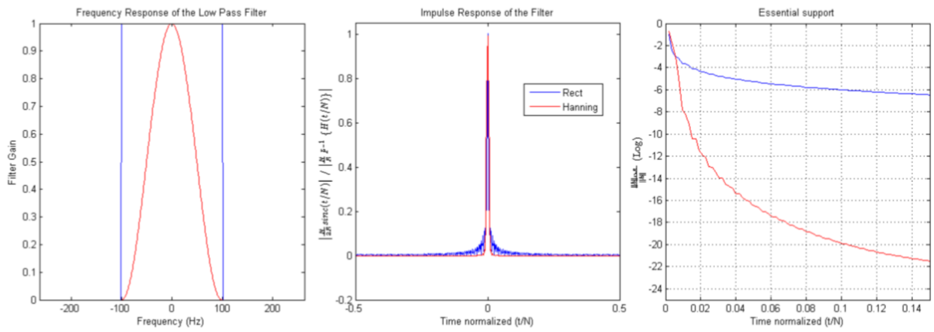

FIGURE 8. The figure shows the essential support of the impulse response for a low pass filter with rectangular filtering window and Hanning filtering window. Results are normalized by the number of samples and therefore independent of the dimension N of the signal space. The example uses a maximum pass frequency of $100Hz$ but results can be easily transfered to other filter regions by $e_{P_2} = \frac{P_1}{P_2} e_{P_1}$.

Theorem 2.22. *Assume a rectangular lattice Λ with lattice constants $a, b > 1$, an analysis window $g_1 \in \boldsymbol{C}^N$, the canonical dual as synthesis window $g_2 = S^{-1} g_1$ and a filter function h such that for $E \geq 0$*

$$\frac{\sum_{\tau=E}^{N-1} |h(\tau)|^2}{\sum_{\tau=0}^{N-1} |h(\tau)|^2} \leq c_E^h \tag{2.49}$$

Further choose $0 \leq c_g(a,b) \leq 1$ such that

$$1 - c_g(a,b) \leq \frac{|V_{g_1} g_2(\tau, 0)|^2}{\sum_{k=0}^{b-1} \sum_{l=0}^{a-1} |V_{g_1} g_2(\tau + kB, lA)|^2} \leq 1 \qquad \forall \tau \in \operatorname{supp}(h), \tau < E \tag{2.50}$$

Then the error between H and its best approximation by a Gabor multiplier can be estimated by

$$\frac{\|H-G\|_{\mathrm{Fro}}}{\|H\|_{\mathrm{Fro}}} \leq \sqrt{c_g^E} + c_h^E \tag{2.51}$$

Proof: We decompose the spreading function $\eta(H) = \eta^E(H)(\tau,\nu) + \eta^R(H)(\tau,\nu) = h\Omega_E \times \delta(\nu) + (1-h\Omega_E) \times \delta(\nu)$ in the part covering the essential support of h and the complement. Then we can calculate the error as $\|\eta(H) - P_{\boldsymbol{G}}\eta(H)\|_{\mathrm{Fro}} = \|\eta(H) - P_{\boldsymbol{G}}(\eta^E(H) + \eta^R(H))\|_{\mathrm{Fro}} = \|\eta(H) - P_{\boldsymbol{G}}\eta^E(H) - P_{\boldsymbol{G}}\eta^R(H)\|_{\mathrm{Fro}} \leq \|\eta(H) - P_{\boldsymbol{G}}\eta^E(H)\|_{\mathrm{Fro}} + \|P_{\boldsymbol{G}}\eta^R(H)\|_{\mathrm{Fro}} \leq \|\eta(H) - P_{\boldsymbol{G}}\eta^E(H)\|_{\mathrm{Fro}} + \|\eta^R(H)\|_{\mathrm{Fro}}$ with $P_{\boldsymbol{G}}$ being the projection on the space of Gabor multipliers. Then $\|\eta^R(H)\|_{\mathrm{Fro}}^2 = \sum_{\tau>E} |h(\tau)|^2 = c_h||h||_2^2$. For $P_G G\eta^E(H)$ we apply theorem (2.19), which gives in total equation (2.51). □

While $c_h(E)$ is a usual measure for the concentration of h on its essential support and its values can be found in figure (2.4.2.3) accordingly. The last theorem gives a special result for tight Gabor windows.

Theorem 2.23. *Assume a rectangular lattice Λ with lattice constant $a, b \geq 1$, a Gaussian tight window function g with $\|g\| = 1$ and an impulse response h with essential support $0 < E < B$. Choose the variables $c_h(E), c_g(a,b) \geq 0$ such that*

$$\begin{array}{llcll} \text{1)} & 1 - c_g(a,b) & \leq & \frac{A|V_{g_1}g_2(\tau,0)|^2}{\sum_{k=0}^{b-1}\sum_{l=0}^{a-1}|V_{g_1}g_2(\tau+kB,lA)|^2} \leq 1 & \forall \tau \in [0,E) \\ \text{2)} & & & \frac{\sum_{\tau=E}^{N-1}|h(t)|^2}{\sum_{\tau=0}^{N-1}|h(t)|^2} \leq c_h(E)\ . & \end{array}$$

Then, the error between a convolution matrix $H \in \boldsymbol{C}^{N\times N}$, with $Hf = (h * f)$ and its Best Approximation by a Gabor multiplier $G := G_g^\Lambda$ can be estimated by

$$\frac{\|H-G\|_{\mathrm{Fro}}^2}{\|H\|^2} \leq 1 - (1-c_h(E))\,(1-c_g(a,b)) \tag{2.52}$$

Proof: We make use of the fact that as $\{T_{ak}M_{bl}g\}$ is a frame, then by duality $\{T_{Bk}M_{Al}g\}$ is a Gabor Riesz basis. Therefore, we can write the numerator as

$$\begin{aligned} \left|\sum_{k=0}^{b-1} h(\tau+kB)\overline{V_g g(\tau+kB,0)}\right|^2 &= \left|\big\langle \sum_{k=0}^{b-1} h(\tau+kB)\big\langle T_{-\tau}gT_{kB}M_{lA}g\big\rangle\right|^2 \\ &= \frac{1}{\|g\|^2}\left\|\sum_{k=0}^{b-1} h(\tau+kB)\big\langle T_{-\tau}g, T_{kB}g\big\rangle T_{kB}g\right\|^2 \\ &\geq \frac{c_1}{\|g\|^2}\sum_{k=0}^{b-1}\left|h(\tau+kB)\big\langle T_{-\tau}g, T_{kB}g\big\rangle\right|^2 \end{aligned}$$

with c_1 denoting the lower Riesz bound. As we have by assumption a tight frame with g Gaussian and $\|g\| = 1$, we also have $c_1 = 1$ and get further

$$\begin{aligned}
\|H - G\|_{\text{Fro}}^2 &\le N \sum_{\tau=0}^{B-1} \left(\sum_{m=0}^{b-1} |h(\tau + mB)|^2 - \left| \frac{\sum_{k=0}^{b-1} |h(\tau + kB)\langle T_{-\tau}g, T_{kB}g\rangle|^2}{\sum_{k=0}^{b-1}\sum_{l=0}^{a-1} |\langle T_{-\tau}g, T_{kB}M_{lA}g\rangle|^2} \right| \right) \\
&\le N \sum_{\tau=0}^{B-1}\sum_{m=0}^{b-1} |h(\tau + mB)|^2 - N \sum_{\tau=0}^{B-1} \left| \frac{c_1 |h(\tau)|^2 |\langle T_{-\tau}g, g\rangle|^2}{\sum_{k=0}^{b-1}\sum_{l=0}^{a-1} |\langle T_{-\tau}g, T_{kB}M_{lA}g\rangle|^2} \right| \\
&\le N\|h\|^2 - N \sum_{\tau=0}^{B-1} \frac{|h(\tau)|^2 |\langle T_{-\tau}g, g\rangle|^2}{\sum_{k=0}^{b-1}\sum_{l=0}^{a-1} |\langle T_{-\tau}g, T_{kB}M_{lA}g\rangle|^2} \\
&\le N\|h\|^2 - N \sum_{\tau=0}^{E-1} \frac{|h(\tau)|^2 |\langle T_{-\tau}g, g\rangle|^2}{\sum_{k=0}^{b-1}\sum_{l=0}^{a-1} |\langle T_{-\tau}g, T_{kB}M_{lA}g\rangle|^2} \\
&\le N\|h\|^2 - N(1 - c_g(a,b)) \sum_{\tau=0}^{E-1} |h(\tau)|^2 \\
&\le N\|h\|^2 - N(1 - c_h(E))\|h\|^2(1 - c_g(a,b)) \\
&\le N\|h\|^2 \left[1 - (1 - c_h(E))(1 - c_g(a,b))\right]
\end{aligned}$$

□

we are now going to investigate the second variable $c_g(a, b)$ in more detail. As indicated it is dependent on the size of the essential support, the sampling parameters a and b, but also the choice of the window function. For commonly used window functions g_1 as Hanning or Hamming windows, we have a very fast decay in the frequency domain of $|V_{g_1}g_2(k, l)|$ opposed to a just moderate decay in the time domain. In the example stated above with $n = 144, a = b = 9$ and $R = 80$ we get $|V_{g_1}g_2(k, Al)| \le c10^{-4} \quad \forall k, \forall l > 0$ whereas $\max_{k \ge B} |V_{g_1}g_2(k, 0)| \approxeq 0.9796$.

Lemma 2.24. *Let g_1 be a window function in $\boldsymbol{C}^N$ and g_2 its dual with respect to a rectangular lattice Λ with lattice constants a and b and $E \ge 0$ then $\forall \tau : |\tau| \le E$*

$$\frac{\sum_{k=0}^{b-1}\sum_{l=0}^{a-1} |V_{g_1}g_2(\tau + kB, lA)|^2}{|V_{g_1}g_2(\tau, 0)|^2} \le 1 + \frac{(ab-1)NE^2F_u^2 ||g_1||^2 \, \|\Delta g_1\|_{\max}^2}{\left|1 - \sqrt{N}EF_u||g_1|| \, \|\Delta g_1\|_{\max}\right|^2} \tag{2.53}$$

with F_u denoting the upper frame bound of the Gabor frame (g_1, a, b).

Proof: As g_2 is the dual window to g_1, we can write the error as the deviation of the exact reproducing property:

$$\begin{aligned}
V_{g_1}g_2(\tau + kB, lA) &= \langle T_{-kB}M_{-lA}g_2, T_\tau g_1\rangle = \langle T_{-kB}M_{-lA}g_2, g_1 + (T_\tau g_1 - g_1)\rangle \\
&= \langle g_2, T_{kB}M_{lA}g_1\rangle + \langle T_{-kB}M_{-lA}g_2, \Delta_\tau g_1\rangle \\
&= \delta(k, l) + \langle T_{kB}M_{lA}g_2, \Delta_\tau g_1\rangle
\end{aligned}$$

by Cauchy Schwartz we get further

$$|V_{g_1}g_2(\tau+kB,lA)|^2 \;\; \leq \;\; \|\delta(k,l)+\Delta_\tau g_1\|\| \, ||g_2||^2 \leq \left|\delta(k,l)+\sqrt{N}\tau||g_2|| \, \|\Delta g_1\|_{\max}\right|^2,$$

which leads us to

$$\begin{aligned}
\left|\frac{\sum_{k=0}^{b-1}\sum_{l=0}^{a-1}|V_{g_1}g_2(\tau+kB,lA)|^2}{|V_{g_1}g_2(\tau,0)|^2}\right| \;\; &\leq \;\; \left|1+\frac{\sum\sum_{(k,l)\neq(0,0)}|V_{g_1}g_2(\tau+kB,lA)|^2}{\left|1-\sqrt{N}\tau\|g_2\| \, \|\Delta g_1\|_{\max}\right|^2}\right| \\
&\leq \;\; \left|1+\frac{(ab-1)N\tau^2\|g_2\|^2 \, \|\Delta g_1\|_{\max}^2}{\left|1-\sqrt{N}\tau\|g_2\| \, \|\Delta g_1\|_{\max}\right|^2}\right| \\
&\leq \;\; \left|1+\frac{(ab-1)N\tau^2||S||_{Op}^2||g_1||^2 \, \|\Delta g_1\|_{\max}^2}{\left|1-\sqrt{N}\tau||S||_{Op}||g_1|| \, \|\Delta g_1\|_{\max}\right|^2}\right| \\
&\leq \;\; \left|1+\frac{(ab-1)NE^2F_u^2||g_1||^2 \, \|\Delta g_1\|_{\max}^2}{\left|1-\sqrt{N}EF_u||g_1|| \, \|\Delta g_1\|_{\max}\right|^2}\right|
\end{aligned}$$

for all $\tau \in E$ with F_u being the upper frame bound of the Gabor frame (g_1, a, b).
□

Based on this result, we are now going to investigate the problem from the other direction. Therefore, we choose a fixed upper bound for the error constant c_g to guarantee a certain approximation quality and deduce the condition on the window function g_1:

$$\begin{aligned}
\frac{\sqrt{(ab-1)}\sqrt{N}EF_u||g_1|| \, \|\Delta g_1\|_{\max}}{\left|1-\sqrt{N}EF_u||g_1|| \, \|\Delta g_1\|_{\max}\right|} &\leq c_g \\
\sqrt{(ab-1)}\sqrt{N}EF_u||g_1|| \, \|\Delta g_1\|_{\max} + c_g\sqrt{N}EF_u||g_1|| \, \|\Delta g_1\|_{\max} &\leq c_g \\
\sqrt{N}||g_1|| \, \|\Delta g_1\|_{\max} &\leq \frac{c_g}{EF_u\sqrt{ab-1}}
\end{aligned}$$

This means that the larger we choose the essential support and the larger the upper frame bound of the Gabor system, the smoother the window function g_1 has to be chosen in order to ascertain a given approximation quality.

In essence, the preceding theorems indicate that in order to guarantee a good approximation quality the lattice constant b has to be chosen in a way such that the essential support of the impulse response of the frequency filter is contained in the interval $[0, B)$. The analysis and synthesis window of the Gabor Multiplier have to be chosen such that $V_{g_1}g_2$ has good decay properties outside Ω and $(g_1 * g_2)$ stays away from zero on supp(h). In case of h, g_1, g_2 all having at minimum decay of order q the error also reduces of order q with the lattice parameters.

2.4.2.4. *Kohn Nirenberg Symbol.* As like the spreading representation also the Kohn-Nirenberg Correspondance is an isometry for the Hilbert Schmidt norm, it is possible to consider the error in this domain in an analogue way. To be more specific, the relation between the spreading function and the Kohn-Nirenberg symbol is given by the symplectic Fourier transform. Therefore if we consider the Frobenius norm of the Kohn Nirenberg Symbols we get the analoge problem in the Fourier domain.

The Kohn Nirenberg Symbol of the LTI filter is constant in time domain and equal to the filter's frequency response in the frequency domain $\sigma(H) = \left(\hat{h} \otimes 1\right)$. On the other hand the Kohn Nirenberg symbol of the Gabor multiplier is given by the convolution of the sampled lower symbol and the Kohn Nirenberg symbol of P_0 meaning $\sigma(G) = m \cdot ⧢_\Lambda * \sigma(P_0)$ with $P_0 = g_1 \otimes g_2$. For lattice parameters $a, b > 1$ it is therefore not slightly time dependent with variations in time depending on the interplay between the smoothness of $\sigma(P_0)$ and the lattice parameters. We have

$$\begin{aligned}
\|\sigma_H(u,v) - \sigma_G(u,v)\|_F &= \|\hat{h} \otimes 1 - m *_\Lambda \sigma(P_0)\|_F = \|\hat{h} \otimes 1 - \{m \cdot ⧢_\Lambda\} * \sigma(P_0)\|_F \\
&= \|\hat{h} \otimes 1 - \mathcal{F}^s \{\mathcal{F}^s (m \cdot ⧢_\Lambda) \cdot \mathcal{F}^s (\sigma(P_0))\}\|_F \\
&= \|\hat{h} \otimes 1 - \mathcal{F}^s \{(\mathcal{F}^s(m) * ⧢_{\Lambda_0}) \cdot \mathcal{F}^s (\sigma(P_0))\}\|_F \\
&= \|\hat{h} \otimes 1 - \mathcal{F}^s \{\mathscr{M}_{\Lambda_0} \cdot \mathcal{F}^s (\sigma(P_0)')\}\|_F \\
&= \|\mathcal{F}^s \{h \otimes \delta - \mathscr{M}_{\Lambda_0} \cdot \mathcal{F}^s (\sigma(P_0)')\}\|_F
\end{aligned}$$

and as we have

$$\begin{aligned}
\mathcal{F}^s \{\sigma(P_0)'\}(u,v) &= \frac{1}{N} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} g_1(l) \check{g}_2(k) e^{\frac{-2\pi ikl}{N}} e^{\frac{2\pi i(ku-lv)}{N}} \\
&= \frac{1}{N} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} g_1(l) \check{g}_2(k) e^{\frac{-2\pi ik(l-u)}{N}} e^{\frac{-2\pi ilv}{N}} \\
&= \frac{1}{N} \sum_{k=0}^{N-1} g_1(l) g_2(l-u) e^{\frac{-2\pi ilv}{N}} = V_{g_2} g_1(u,v)
\end{aligned}$$

this is exactly the Fourier transform of the problem considered in spreading domain (compare Theorem (2.17), equation (2.36)). Using the formula for the best approximating symbol from proposition (2.34) with the notation

$$g_0^{-1}(a,b) = \mathcal{F}\left(\frac{V_{g_1} g_2}{\sum_{k=0}^{b-1} \sum_{l=0}^{a-1} |V_{g_1} g_2(\tau + kB, lA|)^2}\right),$$

we get

$$\|\sigma_H(u,v) - \sigma_G(u,v)\|_{\text{Fro}} = \|\hat{h} \otimes 1 - \left(\hat{h} * g_0^{-1}(a,b) \otimes 1\right) *_\Lambda \sigma(P_0)\|. \tag{2.54}$$

2.4.2.5. *Kernel Domain.* As the kernel domain in finite dimensions is the representation the most well know, we also want to sketch shortly how the phenomena discribed above translate to this domain of operator representation. The matrix representation of a time invariant filter is given by $K(x,y) = h(x-y)$, whereas the kernel representation of the best approximation of a LTI filter by a Gabor multiplier is given by

$$\begin{aligned}
K_G(t,\tau) &= \sum_{k=0}^{A-1}\sum_{l=0}^{B-1} m(ak,bl)\overline{g_1(\tau-ak)}g_2(t-ak)e^{\frac{2\pi ibl(t-\tau)}{N}} \\
&= \sum_{k=0}^{A-1}\sum_{l=0}^{B-1} \left\{ \hat{h} * \frac{\widehat{V_{g_1}g_2(\tau,0)}}{|V_{g_1}g_2(\tau,0)|^2_{\Lambda_0}} \right\}(ak)\overline{g_1(\tau-ak)}g_2(t-ak)e^{\frac{2\pi ibl(t-\tau)}{N}} \\
&= \begin{cases} \sum_{k=0}^{A-1} \left\{ \hat{h} * \frac{\widehat{V_{g_1}g_2(\tau,0)}}{|V_{g_1}g_2(\tau,0)|^2_{\Lambda_0}} \right\}(ak)\overline{g_1(\tau-ak)}g_2(t-ak) & \text{for} \quad |t-\tau| \mod B = 0 \\ 0 & \text{else} \end{cases}
\end{aligned}$$

We see the - also for Gabor frame operators typical - a periodization effects along all side diagonals as well as B periodization effects for fixed t. A visualization of this effects is shown in figure (2.4.2.5).

2.4.3. Simple Approximation Approach. In the following we will consider a simple approximation of a LTI filter by a Gabor multiplier as is often used in practice for near time applications. Therefore the mask of the Gabor multiplier is chosen constant in time domain and as the LTI filter's frequency response in the frequency domain. We will now give a formal definition of this Gabor multiplier approximation as well as the different operator representations. Figure (2.4.3) shows a visual outline of the idea.

We will always assume here that g_2 is chosen either as the canonical dual window of g_1 or as tight Gabor window with respect to the lattice Λ. It is of course also possible to chose g_2 in a different way but even when choosing normalization constants accordingly, empirical results during this work never showed a superior approximation quality for other windows.

Definition 2.25. The simple approximation $G^{SA_h}_{g_1,g_2,\Lambda}$ of the LTI filter H with impulse response h by a Gabor multiplier is defined as the Gabor multiplier with mask $m^{SA_h} = (\mathbf{1} \otimes \hat{h})$ and window functions g_1, g_2 on the lattice Λ.

Proposition 2.26. *The operator kernel of the Gabor Multiplier G^{SA_h} is given by*

$$G^{SA_h}(t,\tau) = B\sum_{l=0}^{b-1} h(t-\tau+lB)\sum_{k=0}^{A-1} g_2(t-ak)\overline{g_1(\tau-ak)} \tag{2.55}$$

Proof: We get the above statement by the definition $m^{SA_h} = (\mathbf{1} \otimes \hat{h})$ and the same calculations as in Theorem (2.9) up to equation (2.21). $\square$

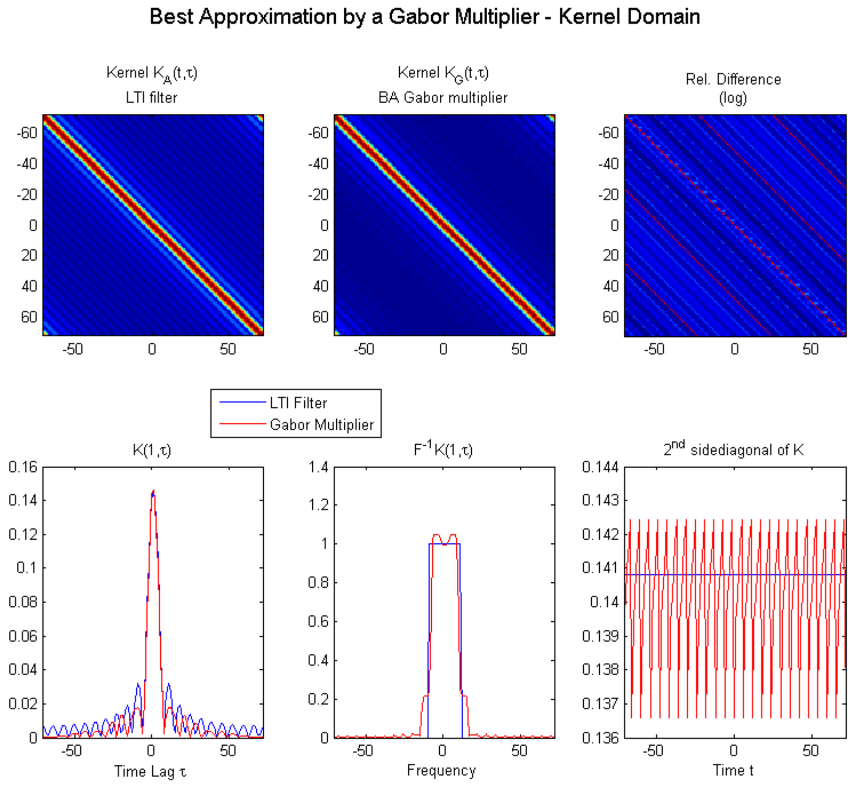

FIGURE 9. Overview over the Best Approximation of a LTI filter by a Gabor Multiplier in Kernel Domain. The periodization effects produced by the Gabor Multiplier Approximation in both time domains (t and τ) can be seen.

Proposition 2.27. *The spreading function of the Gabor Multiplier G^{SA_h} is given by*

$$\eta^{SA_h}(\tau,\nu) = AB\sum_{k=0}^{b-1}\sum_{l=0}^{a-1} h(\tau+kB)\delta(\nu+lA)V_{g_1}g_2(\tau,\nu) \tag{2.56}$$

Proof: Following Theorem (1.78), the spreading function of G^{SA_h} is given by $\eta^{SA_h}(\tau,\nu) = AB\mathscr{M}_{\mathbf{P}}^{SA_h}(\tau,\nu)V_{g_1}g_2(\tau,\nu)$ Further we have

$$\begin{aligned}
\mathscr{M}^{SA_h}(\tau,\nu) &= \frac{1}{N}\sum_{k=0}^{N-1}\sum_{l=0}^{N-1} m^{SA_h}(l,k)e^{2\pi i(l\tau-k\nu)} = \frac{1}{N}\sum_{k=0}^{N-1}\sum_{l=0}^{N-1}\hat{h}(l)e^{2\pi i(l\tau-k\nu)}\\
&= \frac{1}{N}\sum_{k=0}^{N-1}\hat{h}(l)e^{2\pi i l\tau}\sum_{l=0}^{N-1}e^{-2\pi i k\nu} = h(\tau)\sum_{k=0}^{N-1}e^{-2\pi i k\nu} = h(\tau)\delta(\nu).
\end{aligned}$$

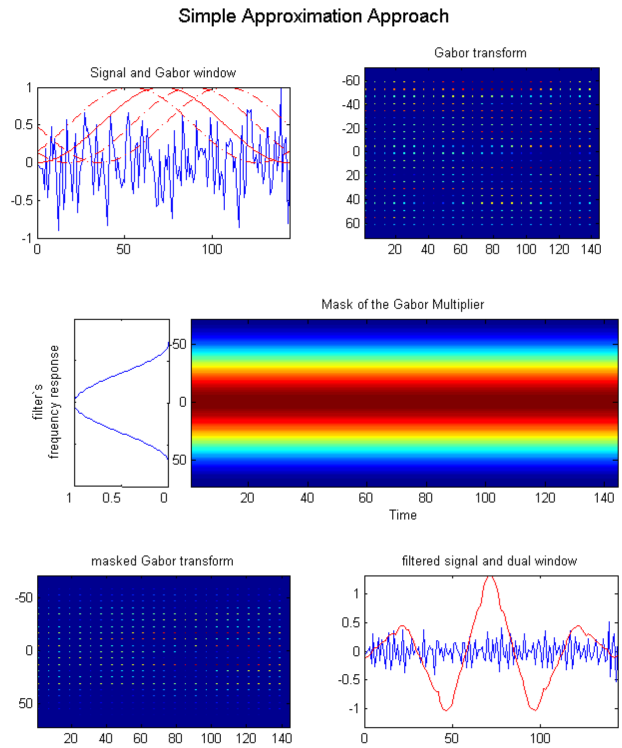

FIGURE 10. Overview over the idea of the simple approximation approach of a LTI filter by a Gabor multiplier. $N = 144$.

Therefore the periodization is $\mathscr{M}_{\mathbf{P}}^{SA_h}(\tau,\nu) = \sum_{k=0}^{b-1}\sum_{l=0}^{a-1} h(\tau + kB)\delta(\nu + lA)$ and the spreading function is given by

$$\eta^{SA_h}(\tau,\nu) = AB\sum_{k=0}^{b-1}\sum_{l=0}^{a-1} h(\tau + kB)\delta(\nu + lA)V_{g_1}g_2(\tau,\nu)$$

$\square$

Proposition 2.28. *The Kohn Nirenberg symbol of the Gabor Multiplier G^{SA_h} is given by*

$$\sigma^{SA_h}(u,v) = \sum_{l=0}^{B-1}\hat{h}(bl)\sum_{k=0}^{A-1}\Pi(ak,bl)\sigma(P_0) \quad \textit{with} \quad \sigma(P_0) = g_1(u)\check{g}_2(v)e^{-2\pi iuv} \tag{2.57}$$

Proof: From Theorem (1.81) together with definition (2.25), we get

$$\sigma^{SA_h}(u,v) \;=\; m^{SA_h} *_\Lambda \sigma(P_0) = \sum_{k=0}^{A-1}\sum_{l=0}^{B-1} m(ak,bl)\Pi(ak,bl)\sigma(P_0)$$

$$= \sum_{l=0}^{B-1} \hat{h}(bl) \sum_{k=0}^{A-1} \Pi(ak, bl)\sigma(P_0) \quad \text{with} \quad \sigma(P_0) = g_1(u)\check{g}_2(v)e^{-2\pi iuv}$$

□

2.4.3.1. *Difference to STFT Multipliers.* As a first indication for the approximation behaviour of the simple approach and as a lower bound of the error, in this section we want to have a closer look on STFT case when $a = b = 1$. Equation (2.55) then becomes $G^{SA_h}(t,\tau) = Nh(t-\tau)\,(g_1 * g_2)\,(t-\tau)$ meaning the simple approach equals the convolution operator with impulse response $\tilde{h} := Nh \cdot (g_1 * g_2)$. Apart from the degenerated cases $h = \delta$ or $g_1 = \text{const}, g_2 = \delta$ or $g_1 = \delta, g_2 = \text{const}$, the simple approach does never represent the convolution operator with impulse response h exactly. For the difference we will now give estimates.

Proposition 2.29. *Assume g being a real valued, symmetric and tight Gabor window with good decay properties and m_2 smooth such that $\|m_2 - m_2 * \hat{g}^2\|_\infty \le \delta$, then the error between the STFT multiplier G with symbol $m = (1 \otimes m_2)$ and the LTI filter H_{m_2} with frequency mask m_2 can be estimated by $\|H_{m_2} - G(m,g)\|_{Op} \le \delta$.*

Proof: From Equation (2.7), we know that the impulse response of the LTI filter corresponding to the STFT multiplier $G(m, g_1, g_2)$ is given through $\mathcal{F}(h) = m_2 * \{\mathcal{F}(g_2)\,\mathcal{F}^{-1}(\bar{g}_1)\}$, which reduces to $\hat{h} = m_2 * \hat{g}^2$ for g real valued, symmetric and tight. The operator norm is given through $\sup_{\|f\|=1}\|(H-G)f\|_{L^2(\mathbb{R}^d)} = \sup_{\|\mathcal{F}(f)\|=1}\|(H-G)f\|_{L^2(\mathbb{R}^d)}$ which is further equal to $\sup_{\|\mathcal{F}(f)\|=1}\|\mathcal{F}((H-G)f)\|_{L^2(\mathbb{R}^d)}$. As H and G are both convolution kernels, also $(H-G)$ is a convolution kernel with impulse response $(h - \mathcal{F}^{-1}(m_2 * \hat{g}^2))$. Therefore in the Fourier domain, we have the multiplication operator $\mathcal{F}\{(H-G)f\}\,(\omega) = (h - \mathcal{F}^{-1}\{m_2 * \hat{g}^2)\}\,(\omega)\,\mathcal{F}(f)(\omega)$. This in turn gives $\sup_{\|\mathcal{F}(f)\|=1}\|\mathcal{F}((H-G)f)\|_{L^2(\mathbb{R}^d)} \le \|m_2 - m_2 * \hat{g}^2\|_{L^\infty(\mathbb{R}^d)} \le \delta$. □

Theorem 2.30. *Let $h(\tau)$ be concentrated on its essential support E such that*

$$\frac{\sum_{\tau=E}^{N-1}|h(\tau)|^2}{\sum_{\tau=0}^{N-1}|h(\tau)|^2} \le (c_E^h)^2. \tag{2.58}$$

Further, let g_1, g_2 normed and such that the convolution is bounded away from zero on $[0,E)$:

$$1 - c_E^g \le N\,(g_1 * g_2(\tau)) \le 1 \quad \forall \tau \in [0,E) \qquad \textit{with} \qquad 0 \le c_E^g < 1 \tag{2.59}$$

then the error between the convolution operator H and the Gabor multiplier G^{SA_h} can be estimated as

$$\frac{\|H - G^{SA_h}\|_F}{\|H\|_F} \le (c_E^g + c_E^h) \tag{2.60}$$

Proof: We are going to calculate the error directly in the matrix representation of the operators. From proposition (2.26), we see that the kernel of the STFT multiplier is given by $G^{SA_h}(t,\tau) = Nh(\tau - t)\,(g_1 * g_2)\,(\tau - t)$. The kernel of H is by definition given through $H(t,\tau) = h(\tau - t)$ and therefore we get

$$\begin{aligned}
\|K(t,\tau) - G(t,\tau)\|_{\text{Fro}}^2 &= \|K(t,\tau+t) - G(t,\tau+t)\|_{\text{Fro}}^2 \\
&= N \sum_{\tau=0}^{N-1} |h(\tau) - Nh(\tau)(g_1 * g_2)(\tau)|^2 \\
&= N \sum_{\tau=0}^{N-1} |h(\tau)\left(1 - N(g_1 * g_2)(\tau)\right)|^2 \qquad (2.61)
\end{aligned}$$

Now, we split the error in two parts: The first part is covering the essential support of h and the second part the region outside the essential support.

$$\begin{aligned}
\|H - G^{SA_h}\|^2 &= N \sum_{\tau=0}^{E-1} |h(\tau)\left(1 - N(g_1 * g_2)\right)(\tau)|^2 + N \sum_{\tau=E}^{N-1} |h(\tau)\left(1 - N(g_1 * g_2)\right)(\tau)|^2 \\
&\leq N \sum_{\tau=0}^{E-1} |h(\tau)|^2 |\left(1 - N(g_1 * g_2)\right)(\tau)|^2 + N \sum_{\tau=E}^{N-1} |h(\tau)|^2 |\left(1 - N(g_1 * g_2)\right)(\tau)|^2 \\
&\leq N \sum_{\tau=0}^{E-1} |h(\tau)|^2 \max_{\tau\in[0,E)} |\left(1 - N(g_1 * g_2)\right)(\tau)|^2 + N \sum_{\tau=E}^{N-1} |h(\tau)|^2 \qquad (2.62)
\end{aligned}$$

By assumption (2.59) we have now $\max_{\tau\in[0,E)} |1 - N(g_1 * g_2)(\tau)| \leq c_E^g$

$$\begin{aligned}
&\leq N \sum_{\tau=0}^{E-1} |h(\tau)|^2 (c_E^g)^2 + N \sum_{\tau=E}^{N-1} |h(\tau)|^2 \leq N (c_E^g)^2 \sum_{\tau=0}^{N-1} |h(\tau)|^2 + N (c_E^h)^2 \sum_{\tau=0}^{N-1} |h(\tau)|^2 \\
&\leq N \|h\|_2^2 \left((c_E^g)^2 + (c_E^h)^2\right)
\end{aligned}$$

as both parts are positive $\sqrt{\varepsilon^2} \leq \sqrt{a^2 + b^2} \leq \sqrt{(a+b)^2} = a + b$ and as $\|g_1\|\|g_2\| \leq \frac{1}{N}$, $\max |g_1 * g_2| \leq \frac{1}{N}$ and we can take the square root. Finally, by $\|H\|_{\text{Fro}} = \sqrt{N}\|h\|_2$ the statement for the relative error follows. □

Note that the estimate c_g^E is not the same as in the last section, but is larger due to the fact that we are here considering the decay of $|V_{g_1} g_2(\tau, 0)|$ and no longer the one of the normalized function $\frac{|V_{g_1} g_2(\tau,0)|2}{\sum_k \sum_l |V_{g_1} g_2(\tau+kB,lA)|^2}$. Figure (2.4.2.3) shows the effect of the normalization. Figure (2.4.3.1) shows the numerical result for the error bound c_g for different window functions.

Theorem 2.31. *Let H be a convolution matrix with $Hf = (h * f)$. Assume $\exists E$ such that the impulse response h has a decay of order p on its tails:*

$$|h(\tau)| \leq \frac{\alpha}{|\tau|^p} \qquad \forall \tau \geq E \qquad (2.63)$$

and g_1, g_2 are chosen such that $(g_1 * g_2)$ is bounded away from zero on $[0, E)$:

$$N\left(g_1 * g_2(\tau)\right) \geq 1 - c_E^g \quad \forall \tau \in [0, E) \qquad \textit{with} \qquad 0 \leq c_E^g < 1 \qquad (2.64)$$

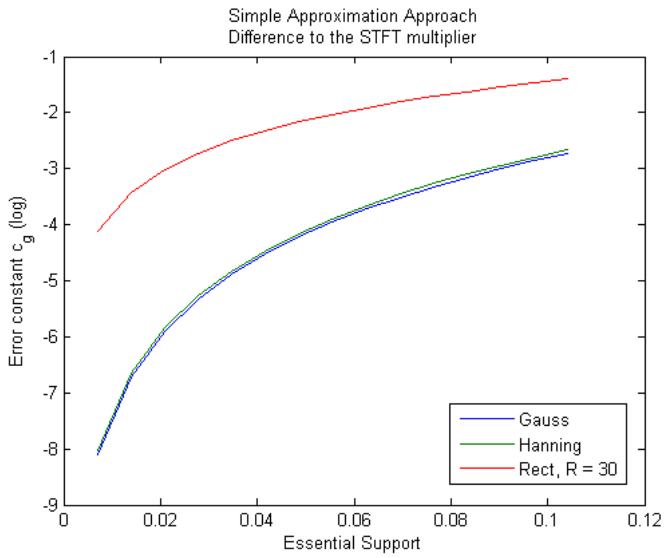

FIGURE 11. Numerical result for the error constant of the estimate inTheorem (2.30). The sample size n has been chosen to be 144 in this experiment. The error bound indicates one part of the error between the Simple approximation approach and the exact STFT multiplier. The second part of the estimate is given by the energy of the impulse response outside the essential support.

Then, the relative error between G^{SA_h} and H can be estimated by

$$\frac{\|H - G^{SA_h}\|}{\|H\|} \leq c_E^g + \frac{\alpha}{\|h\|} \left(\mathbf{H}_{N,p} - \mathbf{H}_{T,p}\right) \tag{2.65}$$

Proof: We start from Theorem (2.30) equation (2.62). Taking the square root and dividing by $\sqrt{N}\|h\|$ gives

$$\frac{||H - G^{SA_h}||}{\|H\|} \leq \max_{\tau<E} |1 - N(g_1 * g_2)(\tau)| + \frac{\|h\|_{\tau \geq E}}{\|h\|}$$

Now applying assumption (2.64) to the first part of the estimate and (2.63) to the second part, we get further

$$\begin{aligned} \frac{\|H - G^{SA_h}\|}{\|H\|} &\leq c_E^g + \frac{1}{\|h\|} \sum_{\tau=E}^{N-1} \frac{\alpha}{|\tau|^p} \leq c_E^g + \frac{\alpha}{\|h\|} \sum_{\tau=E}^{N-1} \frac{1}{\tau^p} \\ &\leq c_E^g + \frac{\alpha}{\|h\|} \left(\mathbf{H}_{N,p} - \mathbf{H}_{E,p}\right) \end{aligned}$$

with $\mathbf{H}_{N,p}$ denoting the N^{th} harmonic number of order p given through $\mathbf{H}_{N,p} := \sum_{k=1}^{N} \frac{1}{k^p}$. $\square$

Analysing the above theorems, we see that the error using the simple approach depends on the decay of $(g_1 * g_2)$ as well as the decay of the impulse response. As for the best approximation, we see again that the relation of the subsampling rate in the frequency domain B and the decay of the impulse response h is essential.

The second important observation is that if g_2 is chosen dual to g_1,the decay of $(g_1 * g_2)$ in turn depends on the chosen sampling distance as is visually outlined in figure (2.4.3.1). The larger the sampling distance is chosen, the faster the decay in $(g_1 * g_2)$ and the more the frequency response of the original filter becomes smoothened by the simple approximation approach. If there is the need for a sharp cut off in frequency this corresponds to the necessity of a dense sampling grid in the time domain.

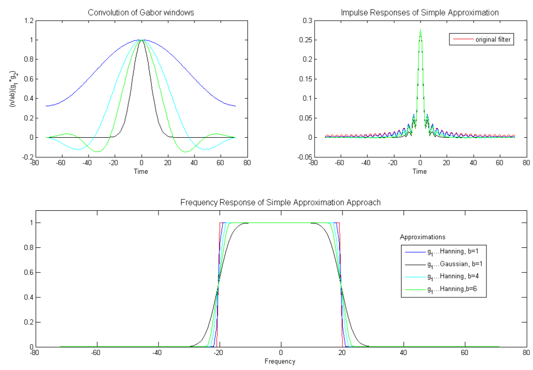

FIGURE 12. Simple Approximation approach using different window functions for the STFT multiplier: In the STFT case the simple approach equals the convolution operator with impulse response $\tilde{h} := Nh \cdot (g_1 * g_2)$. The figure shows $(g_1 * g_2)$ in the top left, the impulse response $\tilde{h}$ in the top right and the frequency response $\hat{\tilde{h}}$ in the bottom window. The synthesis window g_2 is chosen to be the canonical dual window of the analysis window g_1 with respect to the square lattice with sampling distance b. In can be seen that a fast decay in g_1 or using the dual window with respect to a coarser sampling grid leads to smoothing of $\hat{\tilde{h}}$

Third, the smoothness $(g_1 * g_2)$ on E is not only related to the sampling distance but also to the chosen window length of the Gabor multiplier. Assume in particular g_1 and g_2 are compactly supported on an interval I, then $(g_1 * g_2)$ is also compactly supported on $2I$. In general, this means if we have poor decay properties in h, we have to chose long Gabor windows in the time domain in order to get a good approximation quality.

Figure (2.4.3.1) gives an impression of the size of the two error components of Theorem (2.31). It becomes immediately obvious that it is important to adjust the the Gabor window functions to the impulse response of the filter. This can be seen by the fact that for example - contrary to the intuition that a Gassian window for the multiplier is always well suited - the combination of a Gaussion window for the

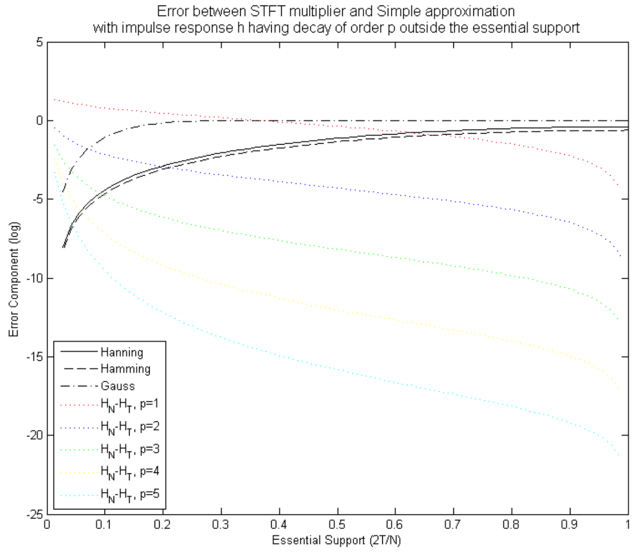

FIGURE 13. Comparison of the two parts of the error estimate in Theorem (2.31) for different window functions g_1 of the STFT multiplier and the corresponding duals. Calculations were don e for $N = 144$.

STFT multiplier and a decay of order $p = 1$ in the impulse response gives especially poor results.

2.4.3.2. *Mapping to the subspace of Gabor Multipliers.* In an analogue way as we can represent the best approximation by a Gabor multiplier as the projection of the spreading function on the AB dimensional subspace of Gabor Multipliers of $\boldsymbol{C}^{N^2}$, it is also possible to represent the the simple approximation approach as a mapping of $\eta_H \to \eta_{G^{SA_h}}$ in this subspace even though it is not an orthogonal projection any more.

Theorem 2.32. *Let η_H be the spreading representation of a LTI filter H. Then the mapping $\eta_H \to \eta_{G^{SA_h}}$ can be represented by a matrix S of the same structure as the projection matrix given in equation (2.28) but with the blocks being of the form $\tilde{M}_i \tilde{M}_j =$*

$$\begin{bmatrix}
V_{g_1}g_2(i,0) & 0 & \dots & 0 & \dots & V_{g_1}g_2(i,0) & 0 & \dots & 0 \\
0 & V_{g_1}g_2(i,1) & \ddots & \vdots & \ddots & 0 & V_{g_1}g_2(i,1) & \ddots & \vdots \\
\vdots & \ddots & \ddots & 0 & & \vdots & \ddots & \ddots & 0 \\
0 & \dots & 0 & V_{g_1}g_2(i,A-1) & & 0 & \dots & 0 & V_{g_1}g_2(i,A-1) \\
\vdots & & & & \ddots & & & & \\
V_{g_1}g_2(i,(a-1)A) & 0 & \dots & 0 & & V_{g_1}g_2(i,(a-1)A) & 0 & \dots & 0 \\
0 & V_{g_1}g_2(i,(a-1)A+1) & \ddots & \vdots & \ddots & 0 & V_{g_1}g_2(i,(a-1)A+1) & \ddots & \vdots \\
\vdots & \ddots & \ddots & 0 & & \vdots & \ddots & \ddots & 0 \\
0 & \dots & 0 & V_{g_1}g_2(i,N-1) & \dots & 0 & \dots & 0 & V_{g_1}g_2(i,N-1)
\end{bmatrix}$$

Proof: Considering the vector matrix multiplication $S\eta_H$ with $\eta_H = [\eta(0,0)\eta(0,1)...\eta(0,N-1)\eta(1,0)\eta(1,1)...\eta(1,N-1)...\eta(N-1,N-1)]^T$ we get exactly the spreading representation of the simple approach as given in proposition (2.27). $\square$

REMARK 2.33. Analogue to the considerations in Theorem (2.14) also here we restrict ourselfs to the space of LTI filters and therefore the entries in the columns 2 to N of the blocks $\widetilde{M}_i\widetilde{M}_j$ do not matter and can be set to one for $i = j$ and zero else. Therefore the representation reduces to a matrix of dimensions $\boldsymbol{C}^{N^2 \times N}$.

2.4.4. Difference between the two approaches. In this section, we want to point out further the difference between the two estimates. For applications this reflects the room for error reduction by choosing the best approximation instead of the simple approach. It turns out that in practice this component gets very small for smooth analysis and synthesis windows in combination with a smooth frequency response of the filter.

Due to the projection theorem for Hilbert spaces, we have $\left\langle H - G^{BA}, G \right\rangle = 0 \quad \forall G \in \boldsymbol{G}(g_1, g_2, a, b)$ and therefore also $\left\langle H - G^{BA}, G^{BA} - G^{SA_h} \right\rangle = 0$. Thus, the error between the simple approximation approach and the LTI filter can be calculated as the sum

$$\|H - G^{SA_h}\|_F = \|H - G^{BA}\|_F + \|G^{BA} - G^{SA_h}\|_F \tag{2.66}$$

The first part of this error we were considering in section (2.4.2). The following section will analyse the second part further.

2.4.4.1. *Numerical Result.* Even though, the two approaches are not equivalent and the simple approach adds a small additional error on top of the Best Approximation, for most applications the main part of the error stems out of the sampling of the mask. For a smooth filtering window (e.g. Hanning) and the essential support of its Fourier transform contained in the interval $(0, B]$ experiments confirm these theoretical findings and show very similar results for both approaches. The aliasing effects are very small. Figure (14) opposes the results for best approximation (BA) and simple approximation (SA) once for a Hanning filtering window and once for a rectangular filtering window for different redundancies. It can be seen, that the error becomes smaller using a smooth filtering window as well as the supremacy using the best approximation reduces.

2.4.4.2. *Lower Symbol.* Figure (2.4.4.2) shows a comparison of the frequency cross section of the lower symbol of the simple approximation (equal to $\hat{h}$) and the best approximation.

Theorem 2.34. *Let $m_1 \in \boldsymbol{C}^{N \times N}$ denote the mask of the simple approximation of the LTI filter $H : \boldsymbol{C}^N \mapsto \boldsymbol{C}^N$ by a Gabor multiplier in $\boldsymbol{G}(g_1, g_2, a, b)$ and m_2 be the mask of the best approximation of a LTI filter by a Gabor multiplier in $\boldsymbol{G}(g_1, g_2, a, b)$. Then $m_1 = m_2$ $\forall h \in \boldsymbol{C}^N$ if and only if $V_{g_1}g_2(x, y) = 0$ $\forall (x, y) \in (N \times N)\backslash\Omega$ and $V_{g_1}g_2 = c$ $\forall x = 1, ..., N$ $c \neq 0$.*

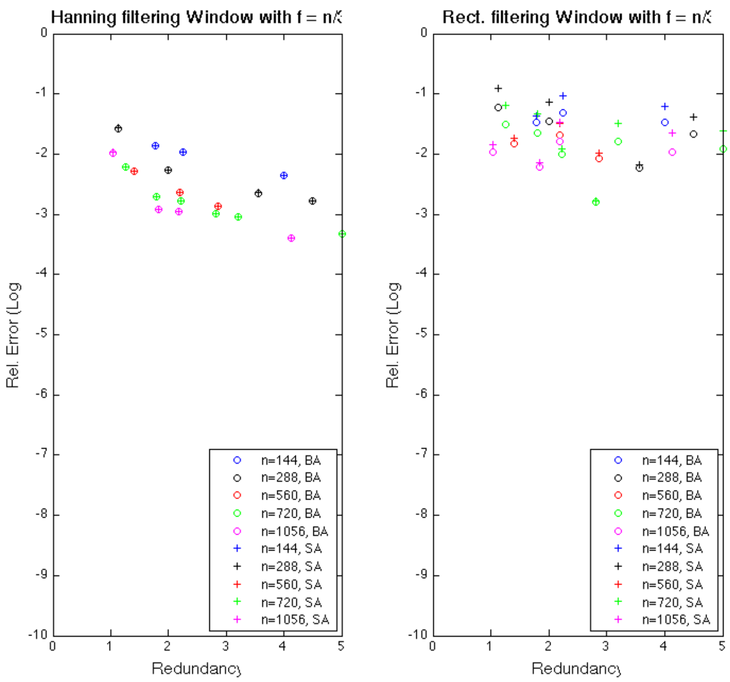

FIGURE 14. Approximation errors of an LTI filter with rectangular frequency response (right) and Hanning window as frequency response (left) for different dimensions n. The logarithmic relative errors are plotted against the different redundancies. The pass region is always chosen as $R = N/3$. As analysis window a Hanning window is used. The synthesis mapping is done with the respective duals.For the smooth filtering window both approaches produce very similar results and the error in general reduces.

Proof: By writing the best approximating mask as

$$\begin{aligned} m_2 = \mathcal{F}^s(\mathscr{M}) &= \mathcal{F}^s\left(\frac{\sum_{m=0}^{b-1}\sum_{l=0}^{a-1} h(u+mB)\delta(v+lA)\overline{V_{g_1}g_2(u+mB,v+lA)}}{\sum_{m=0}^{b-1}\sum_{l=0}^{a-1}|V_{g_1}g_2(u+mB,v+lA)|^2}\right) \\ &= (\hat{h}\otimes 1) *_\Lambda \mathcal{F}^s\left(\frac{V_{g_1}g_2(u,v)}{\sum_{m=0}^{b-1}\sum_{l=0}^{a-1}|V_{g_1}g_2(u+mB,v+lA)|^2}\right) \end{aligned} \tag{2.67}$$

and calculating the difference

$$m_1 - m_2 = (\hat{h}\otimes 1) *_\Lambda \left(\delta(u)\delta(v) - \mathcal{F}^s\left(\frac{V_{g_1}g_2(u,v)}{\sum_{m=0}^{b-1}\sum_{l=0}^{a-1}|V_{g_1}g_2(u+mB,v+lA)|^2}\right)\right) \tag{2.68}$$

we see that both approaches are equivalent if $Vg_1g_2(x,y) = 0 \quad \forall(x,y) \in \{(N\times N)\setminus\Omega\}$ and $V_{g_1}g_2 = c \quad \forall x = 1,..N \quad c \neq 0$. The other direction is easy to show by contradiction with 2 counterexamples. Choose $h(x) = 1 \ \forall x = 1,...,N$ and $V_{g_1}g2 = \chi_R$

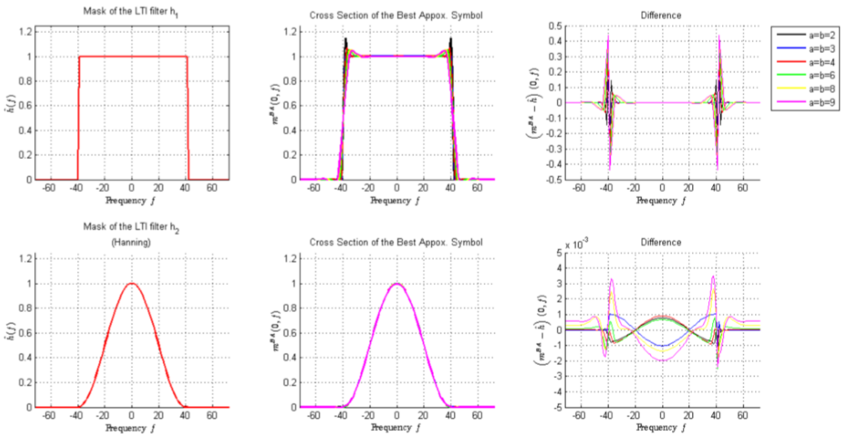

FIGURE 15. Difference in the lower symbol between the simple approximation and the best approximation approach. The difference is calculated once for a rectangular filtering mask and once for a Hannng window. The window length n was chosen to be 144 in this experiment, cut off frequency $R = 40$. As window for the Gabor Multiplier a Hanning window with the respective dual was chosen. Results are stated for different time and frequency sampling distances a and b. It can be seen that the differences are situated at the discontinuities of the mask and become smaller for smooth masks.

with $R \subset \Omega$, then the symplectic Fourier transform is the sinc function and the equation not fulllled or choose $V_{g_1}g_2(x,0) = c \ \forall x \in [0, B] \ c \neq 0$ and again $h(x) = 1$ $\forall x = 1, ..., N$. □

Due to the uncertainty principle this condition can never be fulfilled exactly for $a, b > 1$. We give a schematic overview over the difference between the two approaches.

2.4.4.3. *Spreading Domain.*

Theorem 2.35. *The difference between the simple approximation of a LTI filter H with $Hf = h * f$ and its best approximation by a Gabor multiplier in $\boldsymbol{G}(g_1, g_2, a, b)$ is given by*

$$\frac{1}{N}\|G^{SA} - G^{BA}\|_F^2 = \sum_{\tau=0}^{B-1}\sum_{m=0}^{b-1}\sum_{n=0}^{a-1} |V_{g_1}g_2(\tau + mB, nA)|^2$$
$$\left|\left(\sum_{k=0}^{b-1} h(\tau + kB) - \frac{\sum_{k=0}^{b-1} h(\tau + kB)\overline{V_{g_1}g_2(\tau + kB, 0)}}{\sum_{k=0}^{b-1}\sum_{l=0}^{a-1} |V_{g_1}g_2(\tau + kB, lA)|^2}\right)\right|^2 \tag{2.69}$$

Proof:

$$\frac{1}{N}\|\varepsilon_2\|_F^2 = \|\eta(G^{SA}) - \eta(G^{BA})\|_F^2 =$$

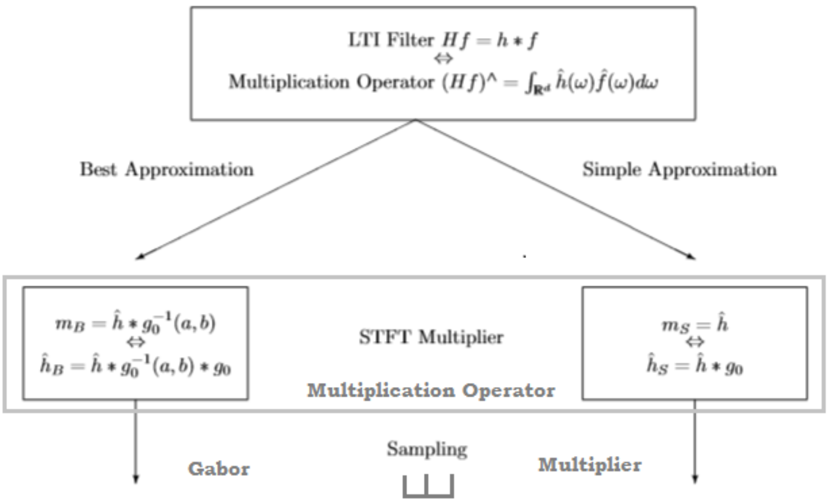

FIGURE 16. Difference between the two approximation approaches. For both STFT multipliers the frequency response is smoothed by a convolution with the superposition of the window functions $g_0(x) := \check{\bar{g}}_1(x)\hat{g}_2(x)$. The best approximation however has an additional convolution with the correction factor $g_0^{-1}(a,b)(x) = \left(\frac{\overline{V_{g_1}g_2}(x,0)}{\sum_{k=0}^{b-1}\sum_{l=0}^{a-1}|V_{g_1}g_2(x+kB,lA)|}\right)^\wedge$

$$\|\sum_{k=0}^{b-1}\sum_{l=0}^{a-1} h(\tau+kB)\delta(\nu+lA)V_{g_1}g_2(\tau,\nu) - \sum_{k=0}^{b-1}\sum_{l=0}^{a-1}\frac{h(\tau+kB)\overline{V_{g_1}g_2(\tau+kB)}\delta(\nu+lA)}{\sum_{k=0}^{b-1}\sum_{l=0}^{a-1}|V_{g_1}g_2(\tau+kB,\nu+lA)|^2}V_{g_1}g_2(\tau,\nu)\|_F^2$$

$$\sum_{\tau=0}^{N-1}\sum_{\nu=0}^{N-1}\left|\sum_{k=0}^{b-1}\sum_{l=0}^{a-1} h(\tau+kB)\delta(\nu+lA)V_{g_1}g_2(\tau,\nu) - \sum_{k=0}^{b-1}\sum_{l=0}^{a-1}\frac{h(\tau+kB)\overline{V_{g_1}g_2(\tau+kB,0)}\delta(\nu+lA)}{\sum_{k=0}^{b-1}\sum_{l=0}^{a-1}|V_{g_1}g_2(\tau+kB,\nu+lA)|^2}V_{g_1}g_2(\tau,\nu)\right|^2$$

$$\sum_{\tau=0}^{N-1}\sum_{\nu=0}^{N-1}\left|\left(\sum_{k=0}^{b-1}\sum_{l=0}^{a-1} h(\tau+kB)\delta(\nu+lA) - \sum_{k=0}^{b-1}\sum_{l=0}^{a-1}\frac{h(\tau+kB)\overline{V_{g_1}g_2(\tau+kB,0)}\delta(\nu+lA)}{\sum_{k=0}^{b-1}\sum_{l=0}^{a-1}|V_{g_1}g_2(\tau+kB,\nu+lA)|^2}\right)V_{g_1}g_2(\tau,\nu)\right|^2$$

$$\sum_{\tau=0}^{B-1}\sum_{\nu=0}^{A-1}\sum_{m=0}^{b-1}\sum_{n=0}^{a-1}|V_{g_1}g_2(\tau+mB,\nu+nA)$$

$$\left(\sum_{k=0}^{b-1}\sum_{l=0}^{a-1} h(\tau+kB)\delta(\nu+lA) - \sum_{k=0}^{b-1}\sum_{l=0}^{a-1}\frac{h(\tau+kB)\overline{V_{g_1}g_2(\tau+kB,0)}\delta(\nu+lA)}{\sum_{k=0}^{b-1}\sum_{l=0}^{a-1}|V_{g_1}g_2(\tau+kB,\nu+lA)|^2}\right)\Bigg|^2$$

$$\sum_{\tau=0}^{B-1}\sum_{\nu=0}^{A-1}\sum_{m=0}^{b-1}\sum_{n=0}^{a-1}|V_{g_1}g_2(\tau+mB,\nu+nA)|^2$$

$$\left|\left(\sum_{k=0}^{b-1}\sum_{l=0}^{a-1} h(\tau+kB)\delta(\nu+lA) - \frac{\sum_{k=0}^{b-1}\sum_{l=0}^{a-1} h(\tau+kB)\overline{V_{g_1}g_2(\tau+kB,0)}\delta(\nu+lA)}{\sum_{k=0}^{b-1}\sum_{l=0}^{a-1}|V_{g_1}g_2(\tau+kB,lA)|^2}\right)\right|^2$$

$$\sum_{\tau=0}^{B-1}\sum_{m=0}^{b-1}\sum_{n=0}^{a-1}|V_{g_1}g_2(\tau+mB,nA)|^2\left|\left(\sum_{k=0}^{b-1} h(\tau+kB) - \frac{\sum_{k=0}^{b-1} h(\tau+kB)\overline{V_{g_1}g_2(\tau+kB,0)}}{\sum_{k=0}^{b-1}\sum_{l=0}^{a-1}|V_{g_1}g_2(\tau+kB,lA)|^2}\right)\right|^2$$

$\square$

Hence, the difference between the two approaches depends on the smoothness of $V_{g_1}g_2(u,0) = (g_1 * g_2)(u) \quad \forall u \in \text{supp}(h)$. This means that if a smooth Gabor window is chosen, if the impulse response has good decay properties and the sampling density in the time domain is chosen appropriately both approaches give very similar results.

Theorem 2.36. *Let $H : \boldsymbol{C}^N \to \boldsymbol{C}^N, Hf = h * f$ be a LTI filter with impulse response h. Assume h is bandlimited e.g. $\exists 0 \leq E \leq B$ such that $|h(\tau)| = 0$ for all $\tau > B$ and further assume there is a constant $c_g^E > 0$*

$$1 - c_g^E \leq \frac{|V_{g_1} g_2|}{\sum_{k=0}^{b-1} \sum_{l=0}^{a-1} |V_{g_1} g_2(\tau + kB, lA)|^2} \leq 1 + c_g^E \tag{2.70}$$

then the difference between the best approximation by a Gabor multiplier with lattice parameters a, b, analysis window g_1 and the canonical dual g_2 as synthesis window, can be estimated as

$$\frac{1}{N} \|G^{SA} - G^{BA}\|_F^2 \leq \frac{(c_g^E)^2}{(1 - c_g^E)} \sum_{\tau=0}^{E} |h(\tau)|^2 |V_{g_1} g_2(\tau, 0)| \tag{2.71}$$

Proof: Starting from Theorem (2.35), we get together with the finite support condition of the impulse response

$$\begin{aligned}
\frac{1}{N} \|G^{SA} - G^{BA}\|_F^2 &= \sum_{\tau=0}^{B-1} \sum_{m=0}^{b-1} \sum_{n=0}^{a-1} |V_{g_1} g_2(\tau + mB, nA)|^2 \\
&\quad \left| \left(h(\tau) - \frac{h(\tau) \overline{V_{g_1} g_2(\tau, 0)}}{\sum_{k=0}^{b-1} \sum_{l=0}^{a-1} |V_{g_1} g_2(\tau + kB, lA)|^2} \right) \right|^2 \\
&\leq \sum_{\tau=0}^{B-1} \frac{|V_{g_1} g_2(\tau, 0)|}{(1 - c_g^E)} \left| h(\tau) \left(1 - \frac{\overline{V_{g_1} g_2(\tau, 0)}}{\sum_{k=0}^{b-1} \sum_{l=0}^{a-1} |V_{g_1} g_2(\tau + kB, lA)|^2} \right) \right|^2 \\
&\leq \sum_{\tau=0}^{B-1} \frac{|V_{g_1} g_2(\tau, 0)|}{(1 - c_g^E)} |h(\tau)|^2 \left| \left(1 - \frac{\overline{V_{g_1} g_2(\tau, 0)}}{\sum_{k=0}^{b-1} \sum_{l=0}^{a-1} |V_{g_1} g_2(\tau + kB, lA)|^2} \right) \right|^2 \\
&\leq \sum_{\tau=0}^{B-1} \frac{(c_g^E)^2}{(1 - c_g^E)} |V_{g_1} g_2(\tau, 0)| |h(\tau)|^2
\end{aligned}$$

□

2.4.4.4. *The problem in KNS Domain.*

Theorem 2.37. *Let $H : \boldsymbol{C}^N \to \boldsymbol{C}^N, Hf = h * f$ be a LTI filter with impulse response h. Assume $||\hat{h} - \hat{h} * g_0||_1 \leq \delta$ with $g_0 = \left(\frac{\overline{V_{g_1} g_2(x,0)}}{\sum_{k=0}^{b-1} \sum_{l=0}^{a-1} |V_{g_1} g_2(x+kB,lA)|^2} \right)^\wedge$, then the difference between the best approximation by a Gabor multiplier with lattice parameters a, b, analysis window g_1 and the canonical dual g_2 as synthesis window, can be estimated as*

$$\|G^{SA}(u, v) - G^{BA}(u, v)\|_F \leq \frac{ab}{F_L} \delta ||g_1||^2 \tag{2.72}$$

with F_L denoting the lower frame bound of (g_1, a, b).

Proof:

$$\begin{aligned}
\|\sigma_{G^{SA}}(u,v) - \sigma_{G^{BA}}(u,v)\|_2 &= \|m_S *_{\Lambda_0} \sigma(P_0) - m_B *_{\Lambda_0} \sigma(P_0)\|_2 \\
&= \|(m_S - m_B) *_{\Lambda_0} \sigma(P_0)\|_F \\
&\leq \frac{1}{AB}\|m^S - m^B\|_1 \|g_1\|_2 \|g_2\|_2 \\
&\leq \frac{ab}{N}\|m_2^S - m_2^B\|_1 \|g_1\|_2 \|g_2\|_2 \\
&\leq \frac{ab}{N}\|\hat{h} - \hat{h} * g_0\|_1 \|g_1\|_2 \|g_2\|_2 \\
&\leq \frac{ab}{NF_L}\|\hat{h} - \hat{h} * g_0\|_1 \|g_1\|_2^2
\end{aligned}$$

The last step is due to the fact that we assume g_2 to be the canonical dual window to g_1 and therefore $g_2 = S^{-1}g_2$ with S denoting the frame operator and therefore $\|g_1\|_2\|g_2\|_2 \leq \|g_1\|_2\|S^{-1}\|_{Op}\|g_1\|_2 \leq \frac{\|g_1\|_2^2}{F_L}$ with F_L being the lower frame bound. □

2.4.5. Comparison to Butterworth filter. In applications Butterworth filters are often used in order to circumvent the problem of having an impulse response k of infinite support, in case we want the transfer function $\hat{k}$ to be the rectangle function. Butterworth lowpass filters are characterized by the fact that $\hat{k}$ is maximal flat for $|\omega| \leq |\omega_c|$, with ω_c denoting the cut off frequency. Maximally flat means that for a filter of order n, the first $(n-1)$ derivatives of $|\hat{h}(0)|$ are equal to zero. The transfer function $\hat{k}(s)$ of a Butterworth filter is given by

$$\hat{k}(s) = \frac{1}{\pi_{l=1}^{n}(s - s_l)/\omega_c}, \qquad s_l = \omega_c e^{\frac{i\pi(2l+n-1)}{2n}} \quad l = 1, ..., n \tag{2.73}$$

The denomination of $\hat{k}(s)$ is also called Butterworth polynomial. The gain of a Butterworth filter of order n is given by

$$|\hat{k}(\omega)|^2 = \frac{1}{1 + (\frac{\omega}{\omega_c})^{2n}} \tag{2.74}$$

The higher we choose the parameter n the sharper is the cut off in the frequency domain. For $n \to \infty$, $|\hat{k}(\omega)|^2$ converges to the rectangle function. For further details on Butterwoth filter we refer to [21]. As we saw in the preceeding part of the chapter for the STFT multiplier with sharp cut off in the mask, smoothing in the LTI comes by nature. We are now going to compare the approximation errors between a STFT multiplier and a convolution operator to those obtained by a Butterworth filter of order n. We use the same setting as in example one with $N = 144$, $R = 80$ and $a = b = 9$. The results are shown in figure (2.4.5)

2.4.6. More Examples. To make the results more visible, we are now going to give a numerical example. Therefore we chose a system with dimension $n = 144$ and a

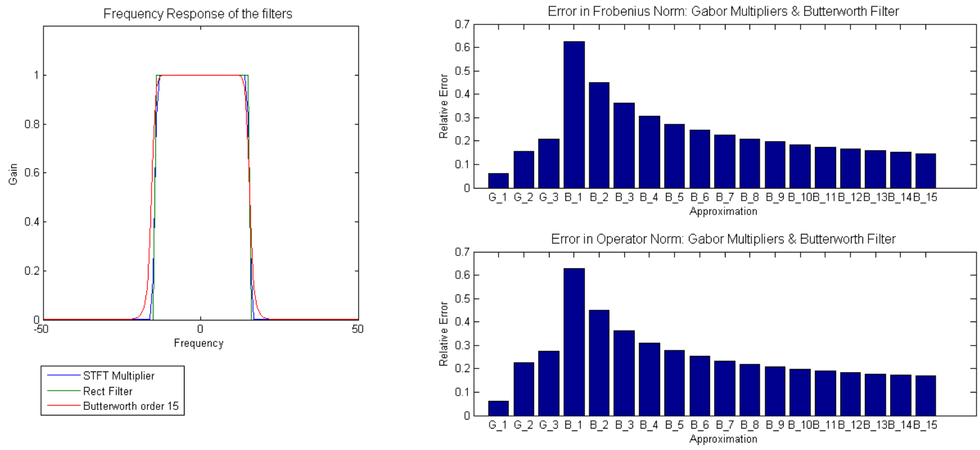

FIGURE 17. Approximation error of a low pass filter by a Gabor Multiplier compared to the one of Butterworth filters of orders $n = 1, ..., 15$. The parameters are chosen as $N = 144$ and cut off frequency is $R = 30$. As analysis and synthesis windows of the Gabor multiplier a Hanning window and the appropriate dual were taken. Relative errors are given in Operator norm as well as in Frobenius norm. G_1 shows the error of the STFT multiplier by the simple approach. For G_2 and G_3 a subsampling rate of $a = b = 4$ (red$=9$) and $a = b = 9$ (red $= 1.78$) respectively was introduced. In can be seen that the Gabor mutliplier with redundancy of 1.78 ha about the same relative error as a butterworth filter of order 8 in Frobenius norm and 5 in operator norm. The relative error of the STFT multiplier is lower than the one of a butterworth filter of reasonable order.

symmetric low pass filter with a Hanning window and maximum frequency $f_{\max} = 25$. As analysis as well as synthesis window, we choose the sampled periodized Gaussian. Then we calculate the according Gabor multiplier as well as the kernel of the Linear Time invariant filter with the same filtering window. For the lattice constants we choose $a = b = 9$, which gives a redundancy of 1.8. Figure (18) shows the results for the kernels of the two operators.

We see clearly the occurring aliasing effects. This aliasing effect stem from the sampling procedure of the symbol of the Gabor multiplier. This becomes even more obvious looking at the spreading function, given through $\eta_G = \mathscr{M} \cdot V_{g_1} g_2$. Figure (19) shows the symbol, the symplectic Fourier transform of the sampled symbol as well as the Short time Fourier transform of the analysis and synthesis window.

2.4.7. Outlook. For slowly time variant systems analogue results can be obtained. Estimates and proofs can be done in an analogue way, but formulas do not simplify in the same way as in the time invariant case as the entries in $\eta_H(x, y)$ do

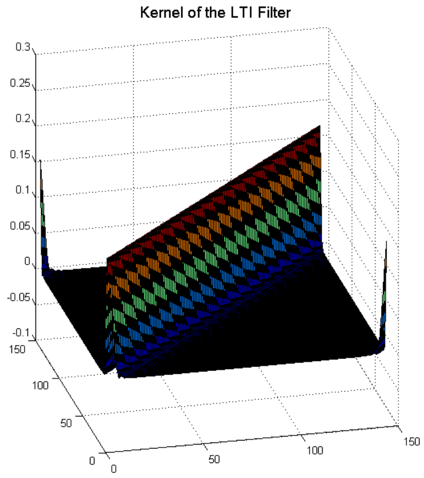

FIGURE 18. The figure shows filtering with a Hanning window with maximal pass freuency of $R = 25$. This is once implemented by a convolution kernel and once with a Gabor multiplier. Both kernels of the resulting operators are displayed. The parameters are chosen as $N = 144, a = 9, b = 9$.

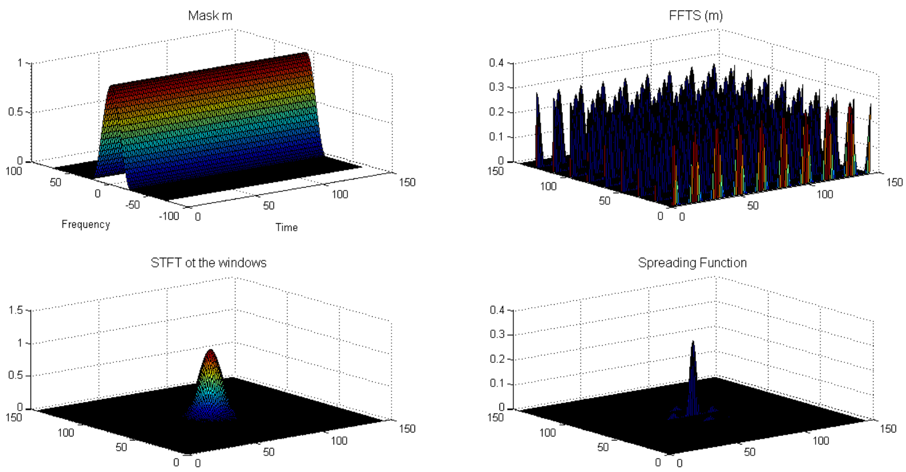

FIGURE 19. The figure shows filtering with a Hanning window with maximal pass freuency of $R = 25$. The plot in the top left shows the mask of the used Gabor Multiplier; in the top right the symplectic Fourier Transform of the symbol is shown. In the lower left image the short time Fourier transform of the analysis and synthesis windows (both chosen to be sampled periodized Gaussians) is shown. Finally the lower right plot displays the spreading function of the operator given through multiplication of $\mathscr{M} \cdot V_g g$. The parameters are chosen as $N = 144, a = 9, b = 9$.

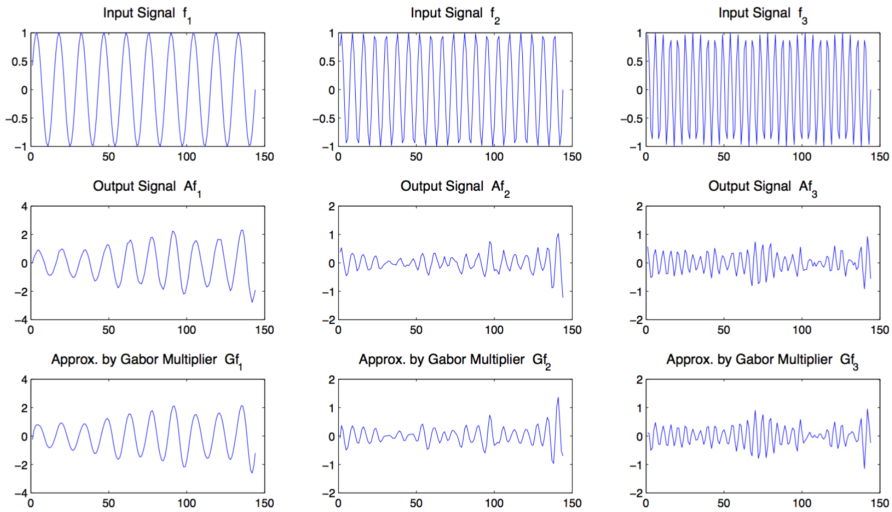

FIGURE 20. Best approximation of a slowly time variant filter A by a Gabor Multiplier G. Both operators are applied to three different signals f_i. The input signals are plotted in the first row. The parameters are chosen as $N = 144, a = 2, b = 4$.

not have to be zero for $y \neq 0$. Of course it would be possible to impose other conditions on the spreading function as e.g. to be of finite support which would lead to underspread operators or to have certain decay properties.

We want to give here a numerical result for the approximation of slowly time variant systems. We therefore construct a STV-system. Three different algorithms for the generation of STV-systems can be found in (reference code). Then we perform the best approximation by a Gabor multiplier. The results are shown in figure (20) to figure (23) for $N = 144$, $a = 2$, $b = 4$ which gives a redundancy of 18. As analysis and synthesis window a tight Gaussian window is used.

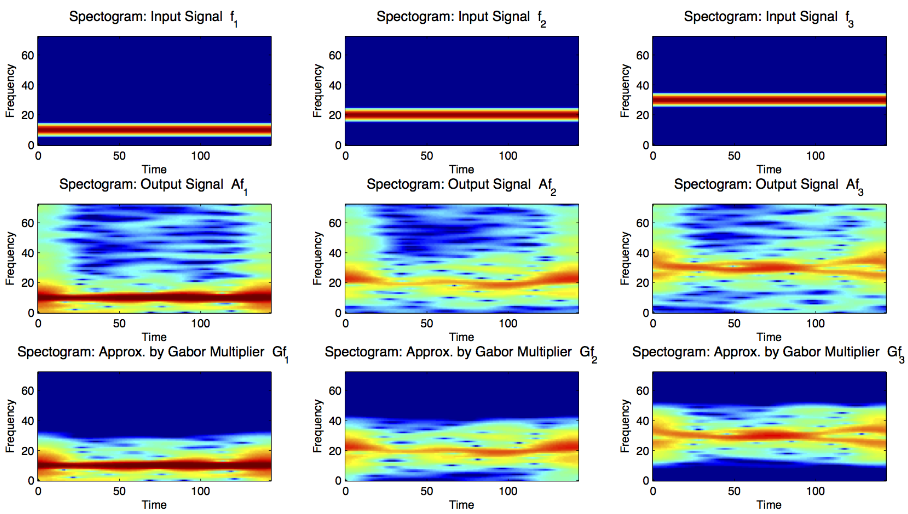

FIGURE 21. Best approximation of a slowly time variant filter A by a Gabor Multiplier G. Both operators are applied to three different signals. The figure shows the spectograms of the input signals f_i $i = 1, ..3$ as well as the output signals Af_i and Gf_i. The parameters are chosen as $N = 144, a = 2, b = 4$.

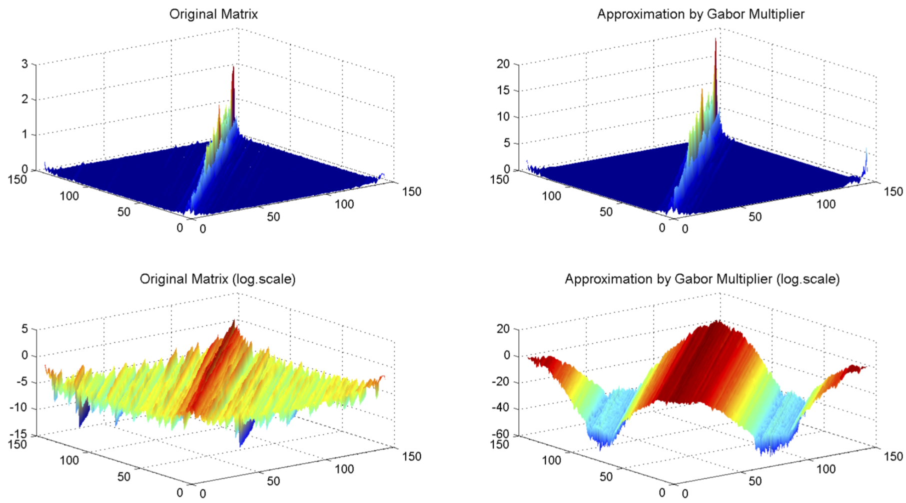

FIGURE 22. Best approximation of a slowly time variant filter A by a Gabor Multiplier G. The figure shows the both matrices in linear as well as in log scale. The parameters are chosen as $N = 144, a = 2, b = 4$.The relative error in Frobenius norm is 0.14127

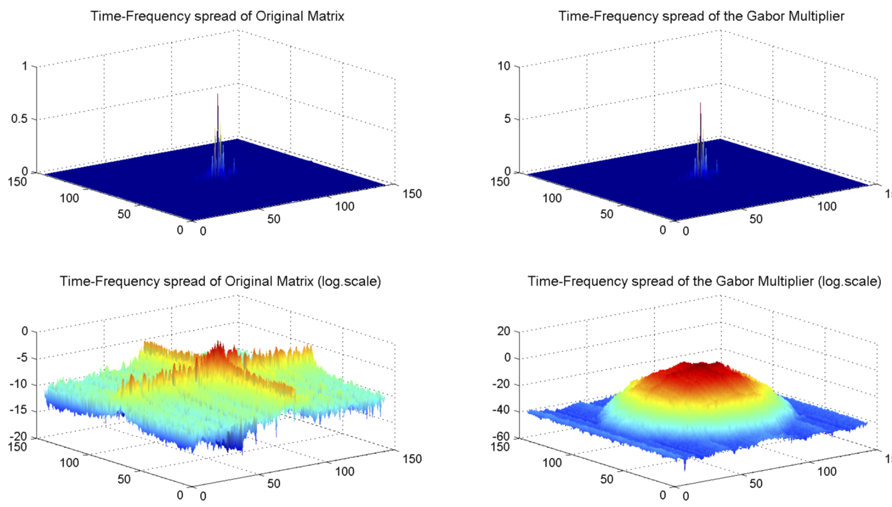

FIGURE 23. Best approximation of a slowly time variant filter A by a Gabor Multiplier G. The figure shows the spreading function of both matrices in linear as well as in log scale. The parameters are chosen as $N = 144, a = 2, b = 4$.

CHAPTER 3

From the STFT multiplier to the Gabor Multiplier

3.1. Convergence Results

Theorem 3.1. *Assume $m \in C^0(\mathbb{R}^{2d})$ and $g_1, g_2 \in \boldsymbol{S}_0(\mathbb{R}^d)$. Further, let $G(m, g, \alpha) : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ denote the Gabor multiplier with symbol m sampled on the rectangular lattice $\Lambda = \alpha\mathbb{Z} \times \alpha\mathbb{Z}$ and $S(m, g) : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ be the continuous STFT multiplier, then $G(m, g, \alpha) \to S(m, g)$ in weak operator topology for $\alpha \to 0$.*

Proof: For the weak operator topology we need to show that $\langle G^\alpha_m f, g\rangle \to \langle S_m f, g\rangle$ $\forall f, g \in L^2(\mathbb{R}^d)$. This means that $\forall \varepsilon > 0 \quad \exists \alpha_0$ such that $\langle G^\alpha_m f, g\rangle - \langle S_m f, g\rangle < \varepsilon \quad \forall \alpha \geq \alpha_0$. The mapping from the integral kernel to the KNS is linear and continuous and therefore bounded. The problem can thus be transferred one to one to the KNS domain. As stated in section (1.10.2) Equation (1.54) the KNS of a Gabor multiplier can be written as the discrete convolution of the upper symbol and $\sigma(P_0)$. The KNS of the STFT multiplier is given by the continuous convolution $\sigma(S) = m * \sigma(P_0)$. Now we write $⧢'_\Lambda := \alpha^2 ⧢_\Lambda$ for the scaled Shah distribution. Via KNS representation the convergence in weak operator topology therefore means $\langle \sigma_{G_m}, u_{f,g}\rangle - \langle \sigma_{S_m}, u_{f,g}\rangle = \langle m⧢'_\Lambda * \sigma(P_0), u_{f,g}\rangle - \langle m * \sigma(P_0), u_{f,g}\rangle < \varepsilon \quad \forall \alpha \geq \alpha_0$ with $u_{f,g}$ being the Rihaczek distribution $u_{f,g} = e^{-2\pi i x\omega}\hat{f}(\omega)g(x)$. As f and g are both in $\boldsymbol{S}_0(\mathbb{R}^d)$, $u_{f,g} \in \boldsymbol{S}_0(\mathbb{R}^{2d})$. Now, we can write

$$\begin{aligned}
\langle m⧢'_\Lambda * \sigma(P_0), u_{f,g}\rangle - \langle m * \sigma(P_0), u_{f,g}\rangle &= \langle (m⧢'_\Lambda - m) * \sigma(P_0), u_{f,g}\rangle \\
&= \langle m⧢'_\Lambda - m, \underbrace{\underbrace{\sigma(P_0)}_{\in S_0(\mathbb{R}^{2d})} * \underbrace{u_{f,g}}_{\in S_0(\mathbb{R}^{2d})}}_{\in S_0(\mathbb{R}^{2d})} \rangle.
\end{aligned}$$

Further $⧢'_\Lambda \to 1 = d\lambda$ for $\alpha \to 0$ in $w*$-topology, which finalizes the proof. □

Theorem 3.2. *Assume m to be in $C^0(\mathbb{R}^{2d})$ and $g_1, g_2 \in L^2(\mathbb{R}^d)$. Further, let $G(m, g, \alpha) : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ denote the Gabor multiplier with symbol m sampled on the rectangular lattice $\Lambda = \alpha\mathbb{Z} \times \alpha\mathbb{Z}$ and $S(m, g) : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ be the continuous STFT multiplier, then $\|G(m, g, \alpha)f - S(m, g)f\|_{L^2} \to 0$ for $\alpha \to 0$,* uniformly *$\forall f \in L^2(\mathbb{R}^d)$.*

Proof: As m is in $C^0(\mathbb{R}^{2d})$, m is tight and according to the definition of tightness, we can find for every $\varepsilon > 0 : \quad m_1 \in C^c(\mathbb{R}^{2d})$ such that $\|m_1 - m\|_\infty < \varepsilon$. Further, let ψ^0 be a BUPU according to Definiton (1.32) such that we can write $\sigma(P_0) = \sum_{j\in\mathbb{Z}} \sigma(P_0)\psi^0_j$. Then let us define $\tilde{\sigma} := \sum_{j\leq J} \sigma(P_0)\psi^0_j$ as a finite sum of building blocks and write the error as $||\sigma(P_0) - \tilde{\sigma}|| \leq \varepsilon(J)$. Therefore we can find for every $\varepsilon > 0: \quad J(\varepsilon)$ such that $||\sigma(P_0) - \tilde{\sigma}|| \leq \varepsilon, \quad \forall J > J(\varepsilon)$ and $\tilde{\sigma}$ is compactly supported.

The norm error between the STFT mutlipliers can be written as $||S(m,g) - \tilde{S}(m,g)||_\infty \leq c||\sigma(S(m,g)) - \sigma(\tilde{S}(m_1,g))||_\infty \leq c||m * \sigma(P_0) - m * \tilde{\sigma}||_\infty \leq ||(m - m_1) * \sigma(P_0) + m_1 * (\sigma(P_0) - \tilde{\sigma})|| \leq ||m_1 - m||_\infty ||\sigma(P_0)||_1 + ||m||_\infty ||\sigma(P_0) - \tilde{\sigma}||_1$. Therefore we can find for every $\varepsilon > 0$ compactly supported m_1 and $\tilde{\sigma}$ such that $||\sigma(S(m,g)) - \sigma(\tilde{S}(m,g))||_\infty \leq \varepsilon\,(||\sigma(P_0)||_1 + ||m||_\infty)$. Then we can find a compact interval K such that

$$\left(m_1 ⧢'_\Lambda * \tilde{\sigma} - m_1 * \tilde{\sigma}\right)(x) = 0 \quad \forall x \notin K \tag{3.1}$$

with $⧢'_\Lambda$ being again the scaled Shah distribution as in proof of Theorem (3.1). For $x \in K$, we can write the convolution

$$\begin{aligned}
\left(m_1 ⧢'_\Lambda - m_1\right) * \tilde{\sigma}(x) &= \int_{\mathbb{R}^{2d}} \left(m_1 ⧢'_\Lambda - m_1\right)(y) \tilde{\sigma}(x-y) dy \\
&= \int_{\mathbb{R}^{2d}} \left(m_1 ⧢'_\Lambda - m_1\right)(y) T_x \tilde{\sigma}^\vee(y) dy.
\end{aligned}$$

We know that $m_1 ⧢'_\Lambda \to m_1$ in w^* and therefore $(m_1 ⧢'_\Lambda - m_1) * \tilde{\sigma}(x) \to 0 \quad \forall x \in K$. Now remains to show that all functions are uniformly continuous.

$$\begin{aligned}
\|m_1 ⧢'_\Lambda * \tilde{\sigma} - T_x\left(m_1 ⧢'_\Lambda * \tilde{\sigma}\right)\|_\infty &\leq \|m_1 ⧢'_\Lambda * \left(\tilde{\sigma} - T_x \tilde{\sigma}\right)\|_\infty \\
&\leq \|m_1 ⧢'_\Lambda\|_{W'} \|\tilde{\sigma} - T_x \tilde{\sigma}\|_W \\
&\leq \|m_1\|_\infty \|⧢'_\Lambda\|_{W'} \|\tilde{\sigma} - T_x \tilde{\sigma}\|_W \\
&\leq \|m\|_\infty \|⧢'_\Lambda\|_{W'} \|\tilde{\sigma} - T_x \tilde{\sigma}\|_W
\end{aligned}$$

Now $\|⧢'_\Lambda\|_{W'} \to 0$ for $\alpha \to 0$ and for the last part of the estimate, we have

$$\sup_{x \in K} \left| \int_{\mathbb{R}^{2d}} \left(T_{z+x}\tilde{\sigma} - T_z \tilde{\sigma}\right) df(z) \right| \leq \|\Delta_z * \tilde{\sigma} - T_z \tilde{\sigma}\|_\infty \leq 1 \sup_{|x| \leq \operatorname{supp} \Delta} \|T_x \tilde{\sigma} - \tilde{\sigma}\|_1$$

with Δ_z being the triangle function of size one around z. $\square$

REMARK 3.3. For $m \in C^{ub}$, also the symbol m can be represented with respect to a family of BUPUs (ψ_k^1) such that $\operatorname{Sp}_{\psi^1} m = \sum_{k \in \mathbb{Z}} m(x_k) \psi_k^1$. Then $m = m \sum_{k \in \mathbb{Z}} \psi_k^1$ and the norm can be estimated

$$\begin{aligned}
\|(\operatorname{Sp}_{\psi^1} m - m)(x)\|_\infty &\leq \sup_x \left| \sum_i \left(m(x_i)\psi_i(x) - m(x)\psi_i(x)\right) \right| \\
&\leq \sup_{|\delta| \leq \alpha} |(m(x) - m(x+\delta)| \sup_x \left| \sum_i \psi_i(x) \right| \\
&\leq \sup_{|\delta| \leq \alpha} |(m(x) - m(x+\delta)| \leq \|m\|_{BV(\alpha)}
\end{aligned}$$

, with $BV(\alpha)$ as defined in Definition (1.41).

3.2. Shift invariant spaces and approximation order

In this chapter we want to establish the connection of Gabor Multipliers and approximation orders of shift invariant spaces. There is a vast literature on shift invariant spaces and it is well known e.g. by [61] that the KNS representation of a Gabor multiplier gives a shift invariant space.

Definition 3.4. Given a lattice $\Lambda \lhd \mathbb{Z}^d$, a closed space S is called *shift invariant* if $\forall f \in S, f \in S \Rightarrow T_\lambda f \in S \;\forall \lambda \in \Lambda$ with T denoting the shift operator. The shift invariant space generated by $\psi \subset S$ is the smallest shift invariant space containing ψ.

Definition 3.5. A shift invariant space generated by a single element is called *principle shift invariant space (PSI space).* A shift invariant space generated by finitely many elements is called a *finitely generated shift invariant space* (PSI space).

Definition 3.6. We denote with $\boldsymbol{G}^\alpha_{g_1,g_2} = \{m *_\Lambda \sigma(g_1 \otimes \bar{g}_2) : m \in \ell^2, \Lambda = \alpha\mathbb{Z}^2\}$ the space of Gabor multipliers with analysis window $g_1 \in S_0$, synthesis window $g_2 \in S_0$ and lattice $\Lambda_\alpha = \alpha\mathbb{Z}^2$.

Lemma 3.7. *The Kohn-Nirenberg symbol, i.e. the mapping from operators to their Kohn-Nirenberg symbol allows to identify for any given lattice $\alpha\mathbb{Z}^d$ the space of Gabor multipliers $\boldsymbol{G}^\alpha_{g_1,g_2}$ with the principal shift invariant space generated by $\sigma(g_1 \otimes \bar{g}_2)$.*

Proof: The Kohn Nirenberg symbol of a Gabor multiplier G is given through the discrete convolution operation $\sigma(G_m) = m *_\Lambda \sigma(P_0)$ with P_0 being the Rihaczek distribution of analysis and synthesis window $P_0 = (g_1 \otimes \bar{g}_2)$. This is exactly the shift invariant space generated by $\sigma(P_0)$. Compare for example [51] for a more detailed analysis of this fact. $\square$

Definition 3.8. We say that the family of spaces $(\boldsymbol{G}^\alpha_{g_1,g_2})_{\alpha>0}$ provides *approximation order k* if for every STFT multiplier $S \in \boldsymbol{\mathcal{HS}}$ with analysis window $g_1 \in S_0$ and synthesis window $g_2 \in S_0$ the approximation error can be estimated by

$$E(S, \boldsymbol{G}^\alpha) = \inf_{G \in \boldsymbol{G}^\alpha} ||S - G||^2_{\boldsymbol{\mathcal{HS}}} \leq C\alpha^k ||\sigma(S)||^2_{W_2^k}, \tag{3.2}$$

with W_2^k denoting the Sobolev space of smoothness k.

Remark 3.9. In the following we will also see that for any $S \in \boldsymbol{\mathcal{HS}}(L^2(\mathbb{R}^d), L^2(\mathbb{R}^d))$, $\sigma(S) \in W_2^k$ is fulfilled naturally by the imposed smoothness conditions on the window functions.

Lemma 3.10. *The Kohn Nirenberg symbol σ_A of an operator A being in the Sobolev space $W_2^k(\mathbb{R}^d)$ is equivalent to the spreading function of the operator $\eta_A \in L^2_v(\mathbb{R}^d)$ with weight function $v = (1 + |z|)^k$.*

Proof: This is due to the already described isometry of the operator symbols and duality of the spaces. $\square$

Lemma 3.11. *If $g_1, g_2 \in M^2_{v_p}(\mathbb{R}^d)$ and $Sf = \int_{\mathbb{R}^{2d}} \langle f, \pi(\lambda)g_1\rangle \pi(\lambda)g_2 d\lambda$ is Hilbert Schmidt, then $\eta(S) \in L^2_{v_p}(\mathbb{R}^{2d})$ and $\sigma(S) \in W_2^p(\mathbb{R}^{2d})$.*

Proof: $S \in \boldsymbol{\mathcal{HS}}$ is equivalent to $m \in L^2(\mathbb{R}^{2d})$ and by Plancherel's Theorem then also $\mathscr{M} \in L^2(\mathbb{R}^{2d})$. Moreover, if $g_1, g_2 \in M^2_{v_p}(\mathbb{R}^d)$ this means that $V_{g_1}g_2 \in L^2_{v_p}(\mathbb{R}^{2d})$. To show this, we use $|V_{g_1}g_2(x,\omega)| \leq |\langle \gamma_1, \gamma_2 \rangle|^{-1}(|V_{\gamma_1}g_1| * |V_{g_2}\gamma_2|)(x,\omega)$ as shown in [65] Lemma 11.3.3 and apply the convolution relation (Proposition 11.1.3 from [65]): $||V_{g_1}g_2||_{L^2_v} \leq C||V_{\gamma_1}g_1||_{L^1_v}||V_{\gamma_2}||_{L^2_v}$ then $V_{g_1}g_2 \in L^2_v$. This is not trivial and can also be found in [65]. The spreading function of the STFT multiplier is given by $\eta(S) = \mathscr{M} \cdot V_{g_1}g_2$ as already used extensively in this work and therefore $\eta(S) \in L^2_v(\mathbb{R}^{2d})$. $\square$

Theorem 3.12. *For $g_1, g_2 \in S_0(\mathbb{R}^d)$, the family of spaces $(\boldsymbol{G}^\alpha_{g_1,g_2})_{\alpha>0}$ provides approximation order $k \geq p$ if and only if $g_1, g_2 \in M^1_{v_p}(\mathbb{R}^d)$ with the weight $v_p(z) = (1+|z|)^p$.*

Proof: Let us first assume that $g_1, g_2 \in M^1_{v_p}(\mathbb{R}^d)$. In order to get an estimate for the approximation order, we have to estimate the error

$$E(S, \boldsymbol{G}^\alpha) = \inf_{G \in \boldsymbol{G}^\alpha} ||S - G||^2_{\boldsymbol{\mathcal{HS}}} = \inf_{G \in \boldsymbol{G}^\alpha} ||\eta(S) - \eta(G)||^2_{L_2} \tag{3.3}$$

Let Ω now be the fundamental domain of the adjoint lattice $\Omega := [0, \frac{1}{\alpha}[\times[0, \frac{1}{\alpha}[$.. Then we define the projection operator P_Ω in the spreading domain as $P_\Omega S := \eta^{-1}\{\chi_\Omega \cdot \eta(S)\}$ with χ_Ω denoting the characteristic function which is equal to one on the region Ω ad zero else. Then we can estimate further using the norm triangle inequality

$$E(S, \boldsymbol{G}^\alpha) \leq \underbrace{E(P_\Omega S, \boldsymbol{G}^\alpha)}_{A} + \underbrace{E((1-P_\Omega)S, \boldsymbol{G}^\alpha)}_{B} \tag{3.4}$$

and for a particular $G \in \boldsymbol{G}^\alpha$ we even have the equality

$$E(S, G) = E(P_\Omega S, G) + E((1-P_\Omega)S, G) - ||G||^2, \tag{3.5}$$

as

$$\begin{aligned} E(S,G) &= \langle S-G, S-G \rangle = \langle P_\Omega S + (1-P_\Omega)S - G, P_\Omega S + (1-P_\Omega)S - G \rangle \\ &= \langle P_\Omega S - G, P_\Omega S - G \rangle + \langle (1-P_\Omega)S, (1-P_\Omega)S - G \rangle - \langle G, (1-P_\Omega)S \rangle, \end{aligned}$$

because of orthogonality and then this is further equal to

$$E(S,G) = ||P_\Omega S - G||^2 + ||(1-P_\Omega)S - G||^2 - \langle G, G \rangle.$$

Now we are going to give an estimate for the two expressions A and B from equation (3.4). For expression B, we first use the simple observation that

$$E((1-P_\Omega)S, \boldsymbol{G}^\alpha) \leq ||(1-P_\Omega)S||_{\boldsymbol{\mathcal{HS}}} = ||(1-\chi_\Omega)\eta(S)||_{L_2}. \tag{3.6}$$

As S is a STFT multiplier $\eta(S)$ is given by $\eta(S) = \mathscr{M} \cdot V_{g_1}g_2$ according to [42]. Therefore the error is calculated as

$$\begin{aligned} E((1-P_\Omega)S, \boldsymbol{G}^\alpha) &\leq \int_{|z|\geq\frac{1}{\alpha}} |\mathscr{M}(z)|^2 |V_{g_1}g_2(z)|^2 d(z) && (3.7) \\ &\leq \int_{|z|\geq\frac{1}{\alpha}} |\mathscr{M}(z)|^2 dz \int_{|z|\geq\frac{1}{\alpha}} \frac{|V_{g_1}g_2(z)|^2 v_p(z)}{v_p(z)} && (3.8) \\ &\leq \alpha^k ||V_{g_1}g_2||_{L^2_v} \int_{|z|\geq\frac{1}{\alpha}} |\mathscr{M}(z)|^2 dz && (3.9) \end{aligned}$$

The $L_{v_p^2}$ norm in the last expression is finite due to Lemma (3.11). As we consider Hilbert Schmidt operators, we know that $||m||_2$, and by Plancherel also $||\mathscr{M}||_2$ is finite. We have therefore that $\int_{|z|\geq\frac{1}{\alpha}} |\mathscr{M}(z)|^2 dz \leq C < \infty$.

For expression A we consider the projection operator $\Pi : \boldsymbol{\mathcal{HS}} \to \boldsymbol{G}^\alpha$, because $E(P_\Omega S, \boldsymbol{G}^\alpha) = ||P_\Omega S - \Pi(P_\Omega S)||_{\boldsymbol{\mathcal{HS}}}$. This projection operator is well defined and non expansive due to Hilbert space setting. Π is given through

$$\eta(\Pi S) = \frac{\sum_{\lambda\in\frac{1}{\alpha}\mathbb{Z}^2} \eta(S)(\cdot+\lambda)\overline{V_{g_1}g_2}((\cdot+\lambda))}{\sum_{\lambda\in\frac{1}{\alpha}\mathbb{Z}^2} |V_{g_1}g_2((\cdot+\lambda))|^2} V_{g_1}g_2 \tag{3.10}$$

Compare for example [73] for a general result on translation invariant spaces and use the fact that $\hat{\sigma}(P_0)(n,m) = V_{g_1}g_2(-m,n)$ (See for example [65]). For Gabor Multipliers in particular the calculations can also be found in [42]. Then we can calculate further

$$\begin{aligned}
E(P_\Omega S, \boldsymbol{G}^\alpha) &= \left\| \chi_\Omega \cdot \eta(S) - \frac{\sum_{\lambda\in\frac{1}{\alpha}\mathbb{Z}^2} \{\eta(S)\cdot\chi_\Omega\}(\cdot+\lambda)\overline{V_{g_1}g_2((\cdot+\lambda))}}{\sum_{\lambda\in\frac{1}{\alpha}\mathbb{Z}^2} |V_{g_1}g_2((\cdot+\lambda))|^2} V_{g_1}g_2 \right\|_{L_2(\mathbb{R}^2)} \\
&= \left\| \eta(S) - \frac{\eta(S)|V_{g_1}g_2|^2}{\sum_{\lambda\in\frac{1}{\alpha}\mathbb{Z}^2} |V_{g_1}g_2((\cdot+\lambda))|^2} \right\|_{L_2(\Omega)} \\
&\leq \|\eta(S)\|_{L_2(\Omega)} \left\| \frac{\sum_{\lambda\in\frac{1}{\alpha}\mathbb{Z}^2\backslash 0} |V_{g_1}g_2(\cdot+\lambda)|^2}{\sum_{\lambda\in\frac{1}{\alpha}\mathbb{Z}^2} |V_{g_1}g_2(\cdot+\lambda)|^2} \right\|_{L_2(\Omega)}.
\end{aligned}$$

Again we use the fact that if $g_1, g_2 \in M^2_{v_p}$, then $V_{g_1}g_2 \in L^2_v$. Hence we can estimate

$$\begin{aligned}
\left\| \frac{\sum_{\lambda\in\frac{1}{\alpha}\mathbb{Z}^2\backslash 0} |V_{g_1}g_2(\lambda+z)|^2}{\sum_{\lambda\in\frac{1}{\alpha}\mathbb{Z}^2} |V_{g_1}g_2(\lambda+z)|^2} \right\|_{L_2(\Omega)} &= \frac{\int_{|y|\geq\frac{1}{\alpha}} |V_{g_1}g_2(y)|^2 dy}{\int_{\mathbb{R}^2} |V_{g_1}g_2(y)|^2 dy} && (3.11) \\
&= \frac{\int_{|y|\geq\frac{1}{\alpha}} \frac{|V_{g_1}g_2(y)|^2(1+|y|)^{2p}}{(1+|y|)^{2p}} dy}{C_1} && (3.12) \\
&\leq (\frac{\alpha}{1+|z|\alpha})^{2p} \frac{||V_{g_1}g_2||_{L^2_v}}{C_1} \leq C\alpha^{2p}. && (3.13)
\end{aligned}$$

In the opposite direction, we can give a proof by contradiction. Therefore assume that $(\boldsymbol{G}^\alpha_{g_1,g_2})_{\alpha>0}$ provides approximation order $k \geq p$. This means according to the definition of the approximation order that

$$E(S, \boldsymbol{G}^\alpha) = \inf_{G\in\boldsymbol{G}^\alpha} \|\eta(S) - \eta(G)\|^2_{L^2} \leq C\alpha^k ||S||_{L^2_{v_k}} \tag{3.14}$$

This in turn means

$$E(P_\Omega S, \boldsymbol{G}^\alpha) = \left\| \frac{|\eta(S)(z)|^2 \sum_{\lambda\in\frac{1}{\alpha}\mathbb{Z}^2\backslash 0} |V_{g_1}g_2(z+\lambda)|^2}{\sum_{\lambda\in\frac{1}{\alpha}\mathbb{Z}^2} |V_{g_1}g_2(z+\lambda)|^2} \right\|_{L_2(\Omega)} \tag{3.15}$$

□

Weighted modulation spaces are in that sense very well suited to describe the approximation order as they describe exactly the combination of smoothness and decay necessary for high approximation orders. We also see that exponential convergence is reached by the Gaussian window, which makes it the natural choice if no further restrictions are given to the problem.

3.3. Transition between the signal spaces

For signals the transition between the spaces L^2, ℓ^2 and C^N through sampling and periodization is well understood. We would in this section now add a short part on the possibility of interchanging the transition from the STFT multiplier to the Gabor multiplier and the sampling operation.

3.3.1. From L^2 to ℓ^2 by sampling. If we identify the the operator with its Kohn Nirenberg symbol, the problem states as follows: Given a function in $S_0(\mathbb{R}^{2d})$, what is the error if we perform a sampling operation, take then the Fourier transform, do a quasi-interpolation with a BUPU (compare [30]) in the frequency domain and then go back to S_0.

Definition 3.13. For $\alpha > 0$ and $a \in \mathbb{N}$, the sampling operators are defined as

$$S_\alpha : S_0(\mathbb{R}) \to \ell^1(\mathbb{Z}) \quad S_\alpha f(j) = \sqrt{\alpha} f(j\alpha), \quad \forall j \in \mathbb{Z}. \tag{3.16}$$

$$S_a : \ell^1(\mathbb{Z}) \to \ell^1(\mathbb{Z}) : \quad S_a f(j) = \sqrt{a} f(ja), \quad \forall j \in \mathbb{Z} \tag{3.17}$$

This corresponds to the definition given in [107]. For further details on the sampling and periodization operators we refer to the cited literature as well as [54] and [78].

Theorem 3.14. *Assume $S_m : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ is a Gabor multiplier with window functions $g_1, g_2 \in S_0(\mathbb{R})$ and symbol m bounded and continuous. Then the operator $G_m f = \sum_{\lambda \in \mathbb{Z}^2} m(\lambda) \langle f, \pi(\lambda) S_1 g_1 \rangle \pi(\lambda) S_1 g_2$ is a Gabor Multiplier from $\ell^2(\mathbb{Z})$to $\ell^2(\mathbb{Z})$ and*

$$S_1(S_m f) = G_m(S_1 f) \quad \forall f \in L^2, f\, cont. \tag{3.18}$$

Proof: For $g_1, g_2 \in S_0(\mathbb{R}^d)$ the sampled windows $g_i^s := S_1 g_i$ are in $\ell^1(\mathbb{Z})$ as S_α is a bounded operator from S_0 to $\ell^1(\mathbb{Z})$ (compare [61] Lemma 3.2.11). Moreover, if (g_1, α, β) is a frame for $L^2(\mathbb{R})$ with canonical dual window g_2, then also $(S_1 g_1, \alpha, \beta)$ is a frame for $\ell^2(\mathbb{Z}^d)$ and the frame bounds of the sampled version are controlled by the frame bounds of the continuous case and canonical dual window $S_1 g_2$. (Compare [77] for the proof and [78], [65] for the proofs that S_0 atoms fulfill the conditions). This means that the coefficient mapping $C : \ell^2(\mathbb{Z}) \to \ell^2(\Lambda)$ as well as the synthesis mapping $D : \ell^2(\Lambda) \to \ell^2(\mathbb{Z})$ are bounded and as m is bounded, the overall operator is bounded. We assume the continuity of f to be able to apply the sampling operator. One could also loosen the restrictions slightly at the cost of simplicity of the expressions (compare [77]). Then we have $\langle S_1 f, \pi(\lambda) S_1 g_1 \rangle_{\ell^2} = \langle f, \pi(\lambda) S_1 g_1 \rangle_{L^2}$, because

$$\sum_{n=-\infty}^{\infty} f(n)g(n) = \int_{\mathbb{R}} \sum_{n=-\infty}^{\infty} \delta_n(x) f(x) g(x) dx = \int_{\mathbb{R}} f(x) \sum_{n=-\infty}^{\infty} \delta_n(x) g(x) dx \qquad (3.19)$$

and with the same argument

$$\begin{aligned}
\sum_{\lambda\in\Lambda} c_\lambda \pi(\lambda) S_1 g_2(x) &= \sum_{\lambda\in\Lambda} c_\lambda \pi(\lambda) \sum_{n=-\infty}^{\infty} \delta_n(x) g_2(x) \\
&= \sum_{n=-\infty}^{\infty} \delta_n(x) \sum_{\lambda\in\Lambda} c_\lambda \pi(\lambda) g_2(x) \\
&= S_1 \left(\sum_{\lambda\in\Lambda} c_\lambda \pi(\lambda) g_2(x) \right)
\end{aligned}$$

Sampling and translation operator can be interchanged, because α is assumed to be an integer. Modulation is a continuous transform and can be interchanged with sampling. □

Theorem 3.15. *Assume $S_m : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ is a STFT multiplier with window functions $g_1, g_2 \in S_0(\mathbb{R})$ and symbol m bounded and continuous. G_m is the Gabor Multiplier with sampled symbol on the lattice $\Lambda = \alpha\mathbb{Z} \times \beta\mathbb{Z}$ with $\alpha, \frac{1}{\beta} \in \mathbb{N}$. Then for $h \in \mathbb{N}$*

$$S_h(G_m f) = G_m(S_h f) \quad \forall f \in L^2(\mathbb{R}), \quad f \text{ cont.} \qquad (3.20)$$

if and only if h is an integer divisor of α.

Proof: The proof stays the same as before. The sampling operator and the translation operator interchange if and only if h is an integer divisor of α. □

The error of sampling the signal and applying a discretized Gabor multiplier instead of a continuous Gabor multiplier is thus under the given conditions just the sampling error. This is also true for downsampling $h > 1$ as long as h is an integer divisor of α. This seems rather obvious, however confirms that it is not essential whether regionone samples the signal before applying the operator and using the sampled operator or sampling the signal after having applied the continuous operator and lets us transfer all results one to one from $L^2(\mathbb{R})$ to $\ell^2(\mathbb{Z})$. The condition on h being an integer divisor of α is naturally fullfilled for applications.

3.4. Explicite Estimate in $\boldsymbol{C}^N$

Definition 3.16. We measure the concentration of a matrix $A \in \boldsymbol{C}^{N\times N}$ on the region Ω of the finite, discrete time-frequency plane through $C_\Omega(A) = \frac{||A|_\Omega||_{Fro}}{||A||_{Fro}}$.

Theorem 3.17. *Let S_m denote the STFT Multiplier $\boldsymbol{C}^N \to \boldsymbol{C}^N$ with symbol m and window functions g_1, g_2 and G_m denote the Gabor Multiplier with symbol m sampled* *on the lattice Λ and the same window functions g_1, g_2, then the relative error in Frobenius norm can be estimated*

$$\frac{||S_m - G_m^\Lambda||_F}{||S_m||_F} \leq ||V_{g_1}g_2||_{\max}(1 - C_\Omega(\mathscr{M}))\frac{||\mathscr{M}||_F}{||\mathscr{M}V_{g_1}g_2||_F} + ||\mathscr{M}||_{\max}(1 - C_\Omega(V_{g_1}g_2))\frac{||V_{g_1}g_2||_F}{||\mathscr{M}V_{g_1}g_2||_F}$$

Proof: Let as usual $\mathscr{M}$ be the symplectic Fourier transform of the symbol m, $\mathscr{M}^{\Lambda_0}$ its periodization on the adjoint lattice, Ω the fundamental region of the adjoint lattice and Ω^C the complement of Ω, the we can estimate in the spreading domain

$$\begin{aligned}
\frac{1}{\sqrt{N}}||S_m - G_m^\Lambda||_F &= ||\eta(S_m) - \eta(G_m^\Lambda)||_F = ||\mathscr{M}V_{g_1}g_2 - \mathscr{M}^{\Lambda_0}V_{g_1}g_2||_F \\
&\leq ||\left(\mathscr{M} - \mathscr{M}^{\Lambda_0}\right)V_{g_1}g_2||_F^\Omega + ||\left(\mathscr{M} - \mathscr{M}^{\Lambda_0}\right)V_{g_1}g_2||_F^{\Omega^C} \\
&\leq ||\mathscr{M} - \mathscr{M}^{\Lambda^0}||_F^\Omega ||V_{g_1}g_2||_{\max} + ||\mathscr{M}||_{\max}||V_{g_1}g_2||_F^{\Omega^C} \\
&\leq ||V_{g_1}g_2||_{\max}||\mathscr{M}||_F^{\Omega^C} + ||\mathscr{M}||_{\max}||V_{g_1}g_2||_F^{\Omega^C} \\
&\leq ||V_{g_1}g_2||_{\max}(1 - C_\Omega(\mathscr{M}))||\mathscr{M}||_F + ||\mathscr{M}||_{\max}(1 - C_\Omega(V_{g_1}g_2))||V_{g_1}g_2||_F
\end{aligned}$$

and for the relative error this means

$$\frac{||S_m - G_m^\Lambda||_F}{||S_m||_F} \leq (1 - C_\Omega(\mathscr{M}))||V_{g_1}g_2||_{\max}\frac{||\mathscr{M}||_F}{||\mathscr{M}V_{g_1}g_2||_F} + ||\mathscr{M}||_{\max}(1 - C_\Omega(V_{g_1}g_2))\frac{||V_{g_1}g_2||_F}{||\mathscr{M}V_{g_1}g_2||_F}$$

$\square$

For finite dimensions the error in operator norm can also be calculated explicitly. For a given frame and given lattice constant α and β it can be determined by pointwise multiplication of the symplectic Fourier transform of m, given through

$$m(\lambda_1, \lambda_2) = \sum_{k \in \mathbb{Z}_N} \sum_{l \in \mathbb{Z}_N} \hat{m} e^{2\pi i (k\lambda_1 + l\lambda_2)}$$

with an error matrix. This error matrix is calculated ass

$$\begin{aligned}
||E|| &= \sup_{||f||=1} || \sum_{\lambda \in (\mathbb{Z}_N \times \mathbb{Z}_N)} m(\lambda)(g_\lambda \otimes g_\lambda)f - \alpha\beta \sum_{\lambda \in \Lambda} m(\lambda)(g_\lambda \otimes g_\lambda)f|| \\
&= \sup_{||f||=1} || \sum_{\lambda \in (\mathbb{Z}_N \times \mathbb{Z}_N)} \sum_{\mathbf{a}} \hat{m}(\mathbf{a}) e^{2\pi i \frac{\mathbf{a}\lambda}{N}} P_\lambda f - \alpha\beta \sum_{\lambda \in \Lambda} \sum_{\mathbf{a}} \hat{m}(\mathbf{a}) e^{2\pi i \frac{\mathbf{a}\lambda}{N}} P_\lambda f|| \\
&= \sup_{|f||=1} || \sum_{\mathbf{a}} \hat{m}(\mathbf{a}) \left(\sum_{\lambda \in (\mathbb{Z}_N \times \mathbb{Z}_N)} e^{2\pi i \frac{\mathbf{a}\lambda}{N}} P_\lambda - \alpha\beta \sum_{\lambda \in \Lambda} e^{2\pi i \frac{\mathbf{a}\lambda}{N}} P_\lambda \right) \chi_{\mathbf{a}} f|| \\
&\leq \sum_{\mathbf{a}} |\hat{m}(\mathbf{a})| \sup_{|f||=1} || \left(\sum_{\lambda \in (\mathbb{Z}_N \times \mathbb{Z}_N)} e^{2\pi i \frac{\mathbf{a}\lambda}{N}} P_\lambda - \alpha\beta \sum_{\lambda \in \Lambda} e^{2\pi i \frac{\mathbf{a}\lambda}{N}} P_\lambda \right) f|| \\
&\leq \sum_{\mathbf{a}} |\hat{m}(\mathbf{a})| || \left(\sum_{\lambda \in (\mathbb{Z}_N \times \mathbb{Z}_N)} e^{2\pi i \frac{\mathbf{a}\lambda}{N}} P_\lambda - \alpha\beta \sum_{\lambda \in \Lambda} e^{2\pi i \frac{\mathbf{a}\lambda}{N}} P_\lambda \right) ||_{\mathcal{L}(\boldsymbol{C}_N, \boldsymbol{C}_N)}
\end{aligned}$$

An example of such an error matrix can be found in Figure (3.4). Note that besides

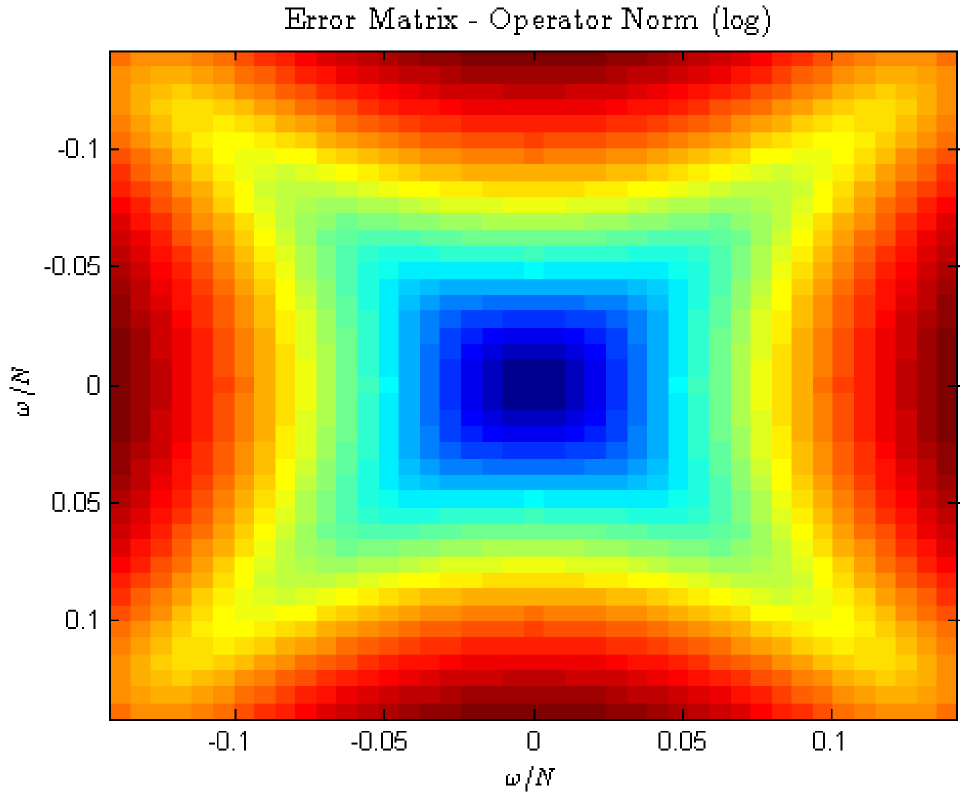

FIGURE 1. The figure gives the logarithmic error in operator norm for sampling the symbol of the Gabor Multiplier. The final error estimate is obtained by pointwhise multiplication of the absolute of the Fourier transform of the symbol by the matrix entries and building the sum. The axis indicate the frequency divided by the dimension N. It can be seen that the error blows up for higher frequencies.

the smoothness of the mask m, the concentration of the support of $V_{g_1}g_2$ is essential. Therefore results in particular for smooth masking functions can be improved by adapting the windows to the lattice. Figure (3.4) shows the comparison for the sampling error once for leaving the window functions unchanged and once for adapting the window functions to the tight window. This experiment is repeated for different masking functions, once for a STFT multiplier filtering with a Hanning window constant in time and once with a rectangular frequency filter. Based on this simulations, Figure (3.4) shows the time frequency concentration of the signals with the largest distortions. It can be seen that adapting the window function improves the result for signals concentrated within the filtering region, whereas the errors are largest in the cut off region. In contrast, when leaving the window unchanged the largest errors are observed within the pass region.

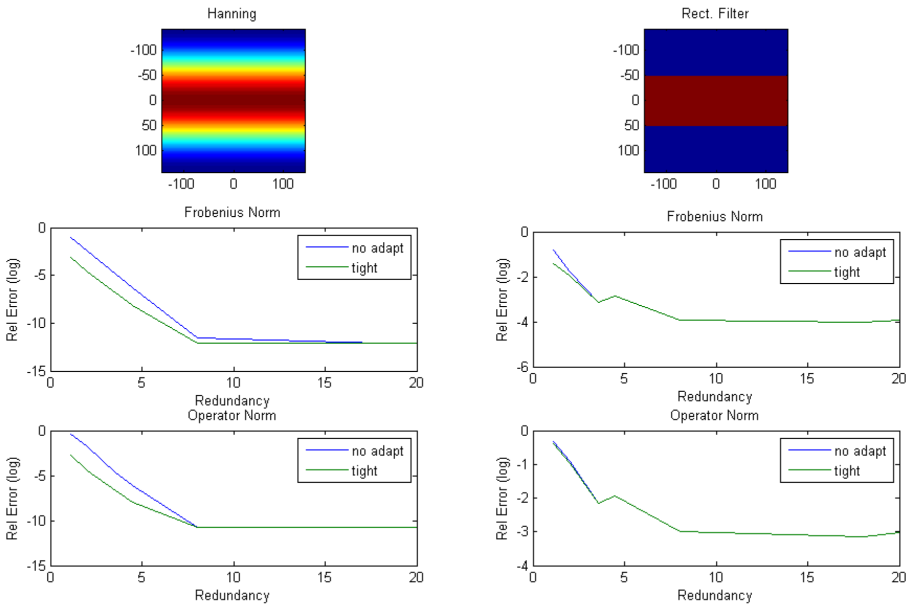

FIGURE 2. The figure gives the logarithmic error in operator norm and in Frobenius norm for sampling the symbol of the STFT Multiplier in dependency of the redundancy. The effect of adapting the window functions to the tight window is analysed for a smooth mask performing frequency filtering with a Hanning window constant in time opposed to a sharp frequency cut off filter constant in time. The two masks used are plotted at the top. It can be seen that for smooth masks the approximation can be improved by adapting the window to the lattice especially for low redundancies. For the rectangular filter the effect almost disappears.

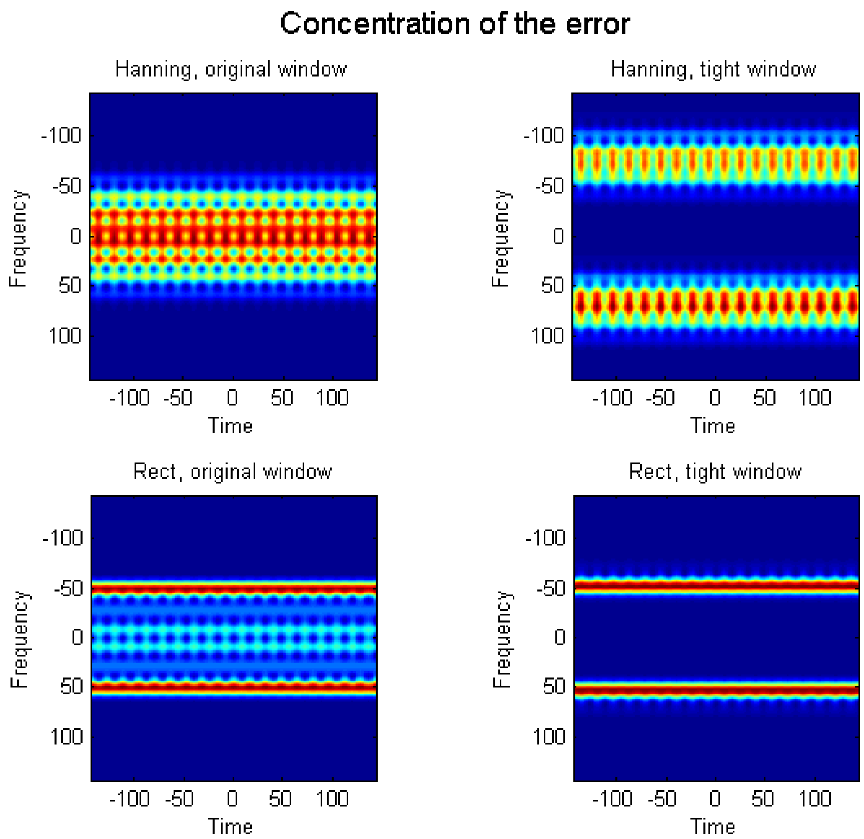

FIGURE 3. This figure shows the time frequency distribution of the subspace spanned by the 15 eigenvectors corresponding to the largest eigenvalues of the error $E = (S_m - G_m^\alpha)$. The plots on the left handside were produced without adapting the window once for a Hanning filtering window and once for a rectangular frequency filter (constant in time), whereas for the plots on the right handside the window functions were adapted to the tight window. The signal length was chosen to $n = 288$ and sampling parameters in time and in frequency were set equal $a = b$. In can be seen that the distortion of the signals within the pass region can be reduced by adapting the window to the tight one.

CHAPTER 4

Design of Gabor Riesz basis for Signal Transmission

4.1. Introduction

Systems of linear independent building blocks, e.g. Riesz basic sequences, are most suitable for mobile transmission. If a signal which is synthesized - using the coefficients carrying the information - as a linear combination of such building blocks, and if they are good joint approximate eigenvectors to any slowly varying system, one can hope to recover the coefficients from the linear combination, by means of the biorthogonal system, which in fact can be taken to be a Gabor System as well.

Since there is a lot of freedom in the design of such Gabor like Riesz basic sequences one has to solve a design problem: Given some general information about the type of slowly varying systems on wants to find Gabor Riesz sequences of relatively high density (in order to maximize the amount of information which can be transmitted by the overall system). However, it is important to also look at the robustness of the system, in order to control the overall performance of such a system.

The underlying theoretical problem is motivated in a wider setting by the practical problem of signal transmission in wireless communication. In particular we consider the part of signal modulation as used for Cyclic prefix orthogonal frequency-division multiplexing (CP-OFDM) [97] [106] [118] [29], constituting a multicarrier scheme being used or proposed for numerous wireless standards like WLAN (IEEE 802.11a/g/n, HIPERLAN/2), broadband wireless access (IEEE 802.16), wireless personal area networks (IEEE 802.15) and digital audio and video broadcasting (DAB, DRM, DVB-T).

In applications, information in form of data symbols, which are transmitted through a channel are carried by pulses, which are rotated and modulated in the time-frequency plane. For classical transmission in frequency domain these pulses are just rectangular functions. In mathematical terms, the signal is decomposed with respect to a Gabor Riesz basis, then the operator is applied followed by a resynthesis mapping. An important research question is how to choose this Gabor Riesz basis in order to satisfy the aim of practitioners for spectral efficiency as well as the aim for (almost) perfect identification of the signal after the channel. It is still common to use rectangular lattices with Gaussian pulses for the construction of Gabor Riesz bases due to their simplicity and convenient numerical properties. There have been suggestions to use multi pulse Gabor Riesz bases or Hermite functions as basis for the transmission of the data in oder to improve spectral efficiency [70], though.

The n dimensional space best concentrated on a spheroidal symmetric region around the origin is the projection space of the first n Hermite functions in the time frequency plane. This raises the possibility to transmit more data than with classical methods using the same space in the time frequency plane. In applications, however, the Hermite basis seems to have the drawback of high susceptability to distortions. Ellaborating this idea, there has been continuing work by Ewa Matusiak substituting the Hermite functions by Gaussian functions placed in a hexagonal pattern [93].

The purpose of this chapter is to analytically confirm and extend the above mentioned work. Different Gabor Riesz basis designs are compared with respect to spectral efficiency, numerical efficiency and stability of results. Special attention is paid to a composite design with a rectangular pattern for the atoms, carrying the pilot-tones, and a circular design for the data atoms. In this manner it is possible to get a better separation of the piltotones while keeping high spectral efficiency for the data packages. This in return guarantees a more stable reconstruction of the pilot-tones even for low redundancies. A more stable reconstruction of the pilot-tones leads to better estimates for the channel and a reduced overall error. A better reconstruction of the pilot-tones can be considered the leverage for the reduction of the total error keeping the redundancy constant.

One of the first questions is the optimal pattern to place the data carrying atoms. For Hermite functions the placement is somewhat natural. They are orthogonal functions in the time frequency plane covering a concentric cyclical regions starting from the center. For Gaussians, the answer to this question is a bit more difficult. Questions arise, as which size of region can be covered by a certain number of Gaussians and under which stability constraints. How big is the region that can be covered by k Gaussians compared to the region that can be covered by the first n Hermite functions or how many Gaussians are required to cover the region of the first n Hermite functions respectively.

The aim in applications is to send as much data as possible as well separated as possible in order to have minimum error and maximal efficiency. In addition, the data should be optimally separated from the pilot-tones, as distortion in the pilot-tones is even more severe than distortion of the data. This means that a quality measure for the separation of the data also appears necessary.

Along with conditioning and separation of the transmitted data, the problem of recovery has to be considered. In mathematical language this means to investigate the biorthogonal system. Main focus is put on the structure of the biorthogonal system and how it can be calculated or approximated in a computational fast manner. In order to get fast calculation, one approach is to make use of the periodic repetitions in the system and consider a local reconstruction.

An important point to emphasize here, is that the results stated in this chapter do not depend on the chosen length of the signal, which is equal to the dimension of the vector space, in a single experiment. Therefore the dimension of the signal space is chosen different for most experiments in a certain range to keep calculation time manageable. Moreover results will always be stated in relative quantities to the dimension n to make them comparable and relevant. The only advantage of choosing higher dimensions is to be able to make the calculation grid finer and produce therefore results of higher precision. The dimensions are however chosen in a manner that precision is sufficient for guiding the way to applications. We will come back to this point in a bit more detail at the end of the chapter in the subsection on channel estimation, which is the only part where results really profit from increased dimension.

4.2. Wireless signal transmission - A short overview

This section gives a short general overview over the current, state of the art, wireless data transmission in a multi-carrier communication systems shall be given. The presented methodology constitutes the basis for application technologies as OFDM, (A)DSL, DVB-T, DAB, WiMAX or LTE and therefore a considerable amount of literature can be found on the subject if further detail is needed.

In particular we will describe a setup corresponding to a pulse-shaping orthogonal frequency division multiplex system (OFDM) ([85], [95], [113]) with smooth transmit and receive pulses, which has been demonstrated ([84], [95], [115]) to be able to substantially reduce inter channel interference compared to conventional (CP) - OFDM and constitutes the basis for the following work.

4.2.1. Wireless signal transmission. The following list gives a comprehensive overview over the consecutive steps relevant for the transformation from the input signal transmitted through the channel to the receiver's output signal.

(1) The continuous signal is discretized by a sampling procedure: $S : f \mapsto Sf := \sqcup\!\sqcup_T f = f[t_n]_{n\in\mathbb{N}}$. In the following we will use brackets to refer to the discrete symbol and paranthesis describing the continuous one.
(2) The discrete signal f thereafter is mapped to a finite set of complex values A, often called the symbol alphabet, using a quantization rule. A common choice as alphabet are the roots of unity e.g. $\{1, -1, i, -i\}$. Then, values can be handled by a CPU with finite precision.
(3) The transmit signal is then produced by the modulator through

$$s[n] = \sum_{l=0}^{L-1} \sum_{k=0}^{K-1} a_{l,k} g_{l,k}[n], \tag{4.1}$$

Here we consider a multi carrier setting, which is common in broadband wireless communication systems due to its advantages over single carrier systems [95], [84]. Therefore, the available bandwidth is occupied by a set of K transmission pulses. L is the number of subcarrier symbols and K is the number of subcarriers. Moreover, a denotes the data symbol and g the Gabor atom, also called prototype pulse, with classical rectangular lattice

$$g_{l,k} = g[n - lN]e^{2i\Pi\frac{k}{K}(n-lN)}. \tag{4.2}$$

with $N \geq K$ denoting the discrete symbol duration.

(4) Finally the continuous time transmit signal, which is obtained by using an interpolation filter with impulse response $f_1(t)$:

$$s(t) = \sum_{n=1}^{N} s[n] f_1(t - nT) \tag{4.3}$$

is transmitted through the channel, which is assumed double selective and dispersive. This description often found in the literature refers to the fact that the channel operator has a spread in the time and in the frequency domain.

(5) At the receiver, the continuous signal is transformed into a discrete-time signal using an anti-aliasing filter with impulse response $f_2(t)$

$$r[n] = \int_{\mathbb{R}} r(t) f_2(nT - t)dt \tag{4.4}$$

(6) After discretization, the received signal is demodulated or resynthesized by computation of the inner products

$$x_{l,k} = \langle r, \tilde{g}_{l,k} \rangle := \sum_{n=1}^{N} r[n]\tilde{g}_{l,k}[n] \tag{4.5}$$

with $\tilde{g}_{lk} := \tilde{g}(n - lN)e^{\frac{2\pi ik(n-lN)}{K}}$ and $\tilde{g}$ denoting the receiver pulse or dual Gabor window. The design of the transmit and receive pulses is of great importance as it has to prevent symbols from interfering in time (Inter Symbol Interference) and in frequency (Inter Carrier Interference).

(7) The demodulated symbols are then equalized and quantized according to the symbol alphabet A. Equalization means, the channel is estimated and inverted by an equalizer to revert its effect. Channel estimation is a vast subject on its own and will be covered in Section (4.2.2) to the extent necessary for the current work.

(8) The symbols are then decoded through inversion of the quantization rule.

(9) The discrete signal is decompressed and transformed in a continuous signal, which finalizes the process.

For further details on the single steps, we refer to respective literature as e.g. [112].

4.2.2. The Wireless Channel. The channel is the linear operator acting on the symbol in between the sender and the receiver. As also described in more detail in Part 1, an operator can be represented in different ways. A common operator representation for the application of channel estimation, is the Zadeh time varying transfer function. The received signal $r[t]$ is then written as

$$r(t) = \int h(t,\tau)s(t-\tau)d\tau + \varepsilon(t) \tag{4.6}$$

with $h(t,\tau)$ being the time varying transfer function (impulse response) and $\varepsilon(t)$ the additive error component caused by noise. On a symbol level the channel equation (4.6) is equivalent to

$$r_{l,k} = \sum_{l'}\sum_{k'} H_{l,k,l',k'} a_{l',k'} + z_{l,k} \tag{4.7}$$

The problem of channel estimation is then equivalent to the solution of the system of linear equations for a.

$$r = Ha. \tag{4.8}$$

In an ideal environment the channel leaves the transmitted signal unchanged and $r(t) = s(t-\tau_0)$, with τ_0 denoting the time it takes the signal to reach the receiver. Under the assumption that sender and receiver are perfectly aligned τ_0 is known and the transmit and receive pulses fulfill the biorthogonality realations e.g. $\sum_n g_{k,l}[n]\tilde{g}_{k',l'}[n] = \delta(k-k')\delta(l-l')$, the demodulated symbols are equal to the transmitted symbols and the channel matrix is the identity.

In practice such ideal channels almost never occur. One physical reason for the channel matrix to differ from a simple diagonal matrix, is multipath propagation. Multipath propagation describes the phenomenon that the signal arrives at the receiver through more than one propagation path due to reflectors and scatterers in the surrounding. The resulting delay spread leads to time dispersion meaning that the transmission pulses are spread out in time. This effect can cause interference between successive data symbols and is called inter-symbol interference.

On the other hand Doppler effects on each of the multipath components occur due to relative motion of sender or receiver or one of the reflectors. This causes pure frequencies to be spread over a finite bandwidth as they pass through the channel and can lead to inter-channel interference. Inter-channel interference denotes the interference between symbols originally separated in frequency. These Doppler effects have become more severe in recent days with the use of higher frequency bands as they increase with frequency. (Compare [112]).

There are different methods and algorithms known to solve problem (4.8), which are beyond the scope of this work. Common approaches are e.g. least squares, maximum likelihood estimation or Minimum Mean Square Error (MMSE). These different methods correspond to different types of equalizers in use. Further, channel estimation can be performed either blindly, only assuming channel statistics, or pilot aided where the set $\mathcal{P} = \{a_{k,l}\}_{(k,l)\in P}$ of pilot symbols is known to the receiver. In this thesis we will only focus on pilot aided estimation schemes.

The channel matrix $H_{l,k,l',k'} = \left\langle \hat{h} g_{l,k}, g_{l',k'} \right\rangle$ is usually big and the system (4.8) is underdetermined. Therefore efficient transmission schemes aim at designing the atoms as well as the lattice in a way such that H has a simple structure with respect to this representation. Usually one would define simple in a sense to have a structure close to diagonal (having fast off diagonal decay) or being sparse meaning that it is comparable with a banded matrix. This may be achieved by choosing the transimission atoms and the structure of the transmission scheme in an appropriate manner. In particular the idea of sparcity in combination with the usage of compressed sensing methods has been recently exploited in channel estimation [117], [116], [47]. A common domain where sparsity can be assumed is the spreading domain. The discrete Delay Doppler Spreading function is given by

$$S_H[m,j] = \frac{1}{N_0} \sum_{n=0}^{N_0-1} h[n,m] e^{\frac{2\pi i n j}{N_0}} \tag{4.9}$$

Therefore we can rewrite the channel matrix H as

$$H_{l,k;l',k'} = \sum_{m=-\infty}^{\infty} \sum_{j=0}^{N_0-1} S_H[m,j] \sum_{n=-\infty}^{\infty} \tilde{g}_{l,k}[n] g_{l',k'}[n-m] e^{\frac{2\pi i n j}{N_0}} \tag{4.10}$$

and by assuming causality of the channel, one may assume that $h[n,m] = 0 \quad \forall m \notin [0, K-1]$. Hence

$$\begin{aligned} H_{l,k;l',k'} &= \sum_{m=0}^{K-1} \sum_{j=0}^{N_0-1} S_H[m,j] \sum_{n=-\infty}^{\infty} \tilde{g}_{l,k}[n] g_{l',k'}[n-m] e^{\frac{2\pi i n j}{N_0}} \\ &= \sum_{m=0}^{K-1} \sum_{j=0}^{N_0-1} S_H[m,j] V_{\tilde{g}_{k,l}} g_{k',l'}[m,j] \end{aligned}$$

Assuming now a general Gabor structure for the underlying atoms $g_{k,l} := T_{\alpha_k} M_{\beta_l} g$ and $\tilde{g}_{k,l} := T_{\alpha_k} M_{\beta_l} \tilde{g}$ with a possible irregular product grid ($\alpha_k \neq \alpha_{k'}$)

$$H_{l,k;l',k'} = \sum_{m=0}^{K-1} \sum_{j=0}^{N_0-1} S_H[m,j] V_{\tilde{g}} g[m + (\alpha_k - \alpha_{k'}), j + (\beta_l - \beta_{l'})] \tag{4.11}$$

The first part of the pointwhise product is completely determined by the dynamic of the channel. In case we would have a linear time independent channel, we would get $S_H = (\delta \otimes h)$ as we showed in Part 1. The second part, on the other hand is determined by the choice of the Gabor structure of the modulator and demodulator. Therefore we see that the design of the Gabor Riesz basis $g_{l,k} := T_{\alpha_k} M_{\beta_l} g$ has a high impact on the overall structure of the channel matrix H. The equation (4.11) would give us a perfect diagonal channel matrix in case of S_H being the Kronecker Delta and g chosen dual to $\tilde{g}$.

4.2.3. Basis Expansion Models. For the sake of completeness, we want to mention here also the existence of basis expansion models different from the Delay Doppler Spreading function given in Equation (4.9). [35], [64], [88], [18], [126].

The common underlying idea of these models is to represent the channel matrix or more precisely the channel impulse response $h[n,m]$ as a linear combination of orthonormal basis functions $\{\psi_j\}_{j=0,..,N-1}$. Then h can be written as

$$h[n,m] = \sum_{j=0}^{N-1} c_h[m,j]\psi_j[n] \quad n = 0,...,N-1 \qquad \text{with} \quad c_h[m,j] = \sum_{n=0}^{N-1} h[n,m]\tilde{\psi_j}[n] \tag{4.12}$$

For the discrete Delay-Doppler Spreading function, we have the Fourier basis functions $\psi_j = \frac{1}{N}e^{\frac{-2\pi ijn}{N}}$ i. e. the complex exponentials. Besides the complex exponentials also other basis functions such as the prolate spheroidal basis functions (the discrete analogue to the Hermite functions) have been proposed in the literature cited above.

4.2.4. Modulation schemes. The modulation scheme defines how the data bits are represented for transmission. This, in turn, directly relates to the question how much information can be carried in one data symbol $a_{k,l}$ in Equation (4.7). Usually either phase, amplitude or frequency of a base signal e.g. a sinusiod are adapted respectively to represent the data symbol.The most commonly used digital modulation schemes are PSK (phase shift keying), ASK (amplitude shift keying), FSK (frequency shift keying) and QAM (quadrature amplitude modulation), which can be seen as a combination of ASK and PSK. Especially PSK is widely used in practice due to its simplicity - for example the wireless Lan Standards IEEE 802.11b-1999, IEEE 802.15.4, IEEE 802.11g-2003 rely on different types of PSK. Those PSK schemes can be differentiated depending on the data rate which translates to the number of phases used. Most commonly binary phase shift keying and quadrature phase shift keying is applied due to their attractive relation between robustness to bit errors and spectral efficiency. The usage of more than 8 phases is not very common as bit error rates get very high compared to the gained efficiency. A visual outline over the different modulation schemes is given in Figure (4.2.4).

4.3. Gabor Riesz basis design

For the complexity of the structure of the channel matrix - cf Equation (4.11), in practical systems the choice of the modulator and demodulator is extremely important. A clever design of the underlying Gabor Riesz basis carrying the data symbols is essential. In this respect, the aim is always to end of with a matrix which is cheap and easy to invert like a complete diagonal structure of H.

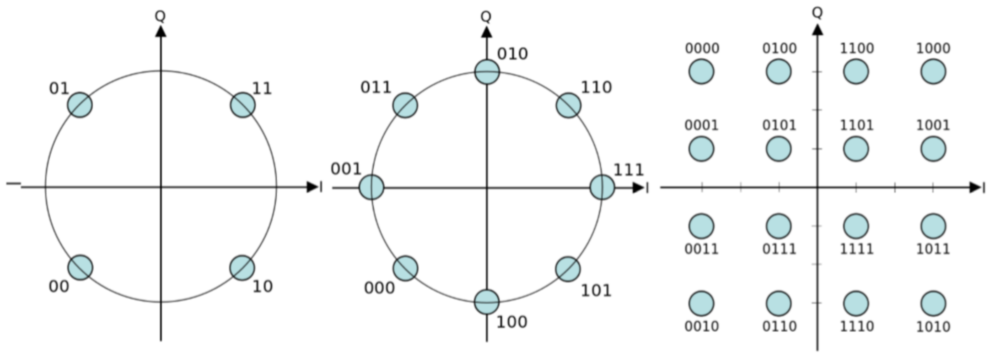

Figure 1. Overview over the most common modulation schemes showing QPSK, 8-PSK modulation and QAM modulation from left to right.

The aim in this work is to explore advantages and drawbacks of alternative designs of the Gabor Riesz basis utilized for signal transmission. The goal is to improve the off diagonal decay of the channel matrix H e.g. reduce ICI and ISI through the choice of the modulation atoms and their placement in the time frequency plane while keeping the same spectral efficiency. The equivalent problem is to improve spectral efficiency by keeping the same order of interference terms.

Special attention will be given to a new design using a rectangular scheme for the pilot-tones and a hexagonal scheme for the data packages. Through the radial symmetric data areas in the time frequency plane it is possible to reduce interference between the data and the pilot-tones which enables a more stable channel equalization while keeping the spectral efficiency. Contrary to approaches using the Hermite function as basis for data modulation the investigated scheme is still a single window approach which guarantees that the underlying structure of standard Gabor frames can be used.

Besides the the aim for low interference terms in the channel matrix, numerical efficiency and easy algorithmic implementation are important issues for applications. This is the reason why multi window approaches as the usage of the Hermite basis for data modulation have not yet become widely accepted among practitioners but CP-OFDM as described in section (4.3.2) is still a popular and widely used multicarrier modulation scheme. We will therefore also pay attention to estimate the numerical efficiency of the proposed schemes and provide algorithmic prototyping itself or reference respective literature to leverage on.

4.3.1. Measurement Dimensions. In order to give a complete picture of the quality, strengths and weaknesses of the analyzed transmission schemes, we analyse and measure the three main dimensions of interest for practitioners: Spectral Efficiency, Stability and Numerical Efficiency as well as their interrelation.

4.3.1.1. *Spectral Efficiency.* Spectral efficiency is crucial in applications as it is directly linked to cost savings due to the need for less bandwidth or an improvement in the transmission quality. We measure the spectral efficiency as the ratio of transmitted data symbols to the dimension of the signal space. This gives us an indication on how much of the space at disposal is used to transfer data.

4.3.1.2. *Stability.* In this section we investigate how distortion influences the transmission quality of the system. In applications, a certain robustness to perturbations and to noise has to be guaranteed. Spectral efficiency or transmission speed must not be too much at the cost of data transmission quality. Usually a certain imperfectness of the reconstruction is accepted or compensated by a llevel of redundancy in the information. On the other hand redundancy has the same effect as decreasing spectral efficiency. Stability will be measured predominantly in terms of the bit error rate and of special interest is here in particular the type of interrelation of spectral efficiency and stability.

4.3.1.3. *Numerical efficiency.* Numerical efficiency describes the cost and stability of the algorithmic implementation of the solution. Besides the channel estimation itself which we leave apart for the underlying work, the most critical step is the demodulation as therefore we need to get the dual system. An appropriate measure in this respect for the quality of a Gabor Riesz basis is the quotient of the Riesz bounds, which is equal to the condition number of the Gabor matrix G. The condition number itself, is a common stability measure in applications. It quantifies the change in the in coefficients when one of the basis vectors is changed or equivalently the deviation in the decomposed signal when the coefficients are perturbated. It also indicates the cost of inverting the system.

Theorem 4.1. *([27] Theorem 3.4.4, Prop 3.4.5) For a sequence $(e_k)_{k=1}^{\infty}$ in a Hilbert space $\boldsymbol{\mathcal{HS}}$, the following conditions are equivalent*

- *(e_k) is a Riesz basis for $\boldsymbol{\mathcal{HS}}$.*
- *(e_k) is complete in $\boldsymbol{\mathcal{HS}}$ and there exist constants $A, B > 0$ such that for every finite scalar sequence one has*
$$A \sum |c_k|^2 \leq || \sum c_k e_k ||^2 \leq B \sum |c_k|^2 \tag{4.13}$$
- *(e_k) is complete and its Gram matrix $G := (\langle e_k, e_j \rangle)_{k,j=1}^{\infty}$ defines a bounded, invertible operator on $\ell^2(\mathbb{N})$. The optimal Riesz bounds are $A = \frac{1}{||G^{-1}||}$ and $B = ||G||$*

Therefore the Riesz bounds can be determined through calculation of the eigenvalues of the Gram Matrix of the Gabor family G. Then the condition number is obtained through

$$c(G) = \frac{\rho_{\max}(GG')}{\rho_{\min}(GG')}. \tag{4.14}$$

with $\rho_{\min}(G)$ and $\rho_{\max}(G)$ denoting the smallest and largest singular vaule of G respectively.

Equivalent way to calculate the condition number. During the experiments it turned out and it seems to be sort of a general knowledge in parts of the research community, that there exists an equivalent way to determine the lowest condition number of the family $(g_i)_{i \in I}$ under different group automorphisms A_j.

Theorem 4.2. *Assume a Gabor family $(g_i)_{i \in I}$ and a family of group automorphisms $(A_j)_{j \in J}$ and define*

$$M_j := \frac{\max \sum_{i \in I} |V_g A_j(g_i)(z)|}{\min_{z \in \cup \operatorname{supp} V_g A_j(g_i)} \sum_{i \in I} |V_g A_j(g_i)(z)|}. \tag{4.15}$$

Then

$$j_1 = j_2 \qquad \text{with} \quad M_{j_1} = \min_{j \in J} M_j \quad c(A_{j_2}) = \min_{j \in J} c(A_j). \tag{4.16}$$

This means M reaches its minimum for the same automorphism where c reaches its minimum.

Proof: As up to our knowledge there has not been given a prove for this before, we want to give here a numerical evidence. Therefore different lattices Λ_i with the same redundancy were generated through a groupautomorphism. Then we calculate the Gabor Riesz basis (g, Λ_i), with g denoting the Gaussian. Table (4.3.1.3) shows the result for the condition numbers as well as the values of M_i for a sample of 3 different parametersets. $\square$

n	a	b	Val	Λ_1	Λ_2	Λ_3	Λ_4	Λ_5	Λ_6	Λ_7	Λ_8	Λ_9	Λ_10	Λ_{11}	Λ_{12}
256	64	16	c	1.189	1.009	1.009	1.189	2.414	2.414						
			M	646	45.16	45.16	646	170533	170533						
144	16	18	c	1.414	1.566	3.518	2.651	1.319	1.369	1.630	2.320	1.461	1.630	1.422	1.426
			M	5.745	6.297	26.55	15.02	4.094	4.905	7.268	12.23	6.103	7.268	5.654	5.967
144	24	9	c	2.077	1.873	1.873	2.077	8.162	8.162	2.085	2.436	2.436	2.085	8.162	8.162
			M	2.885	2.820	2.820	2.885	11.740	11.701	3.063	3.385	3.385	3.063	11.70	11.74

TABLE 1. The table gives the values for M and c of $(g_\lambda)_{\lambda \in \Lambda_i}$ for different groupautomorphisms Λ_i for a sample of 3 experiments. Experiments have been done for different redundancies. The respective MATLAB code in order to reproduce or extend the results, can be found in the appendix.

4.3.2. The classic transmission scheme: CP-OFDM. In order to understand the origin of the problem we want to shortly describe the classical and still popular and widely used multicarrier modulation scheme: the CP-OFDM (Cyclic Prefix Orthogonal Frequency-Devision Multiplexing) [24], [97]. In this transmission scheme a rectangular lattice in combination with rectangular windows g and $\tilde{g}$ is used. More precisely, we have

$$g[n] := \begin{cases} \frac{1}{\sqrt{K}} & n \in [-L, K-1] \\ 0 & \text{else} \end{cases} \quad \text{and} \quad \tilde{g}[n] := \begin{cases} \frac{1}{\sqrt{K}} & n \in [0, K-1] \\ 0 & \text{else} \end{cases} \tag{4.17}$$

with $L = N - K$ denoting the length of the interval for the cyclic prefix in order to avoid inter symbol interference. The transmit signal is then written as $s[n] = \sum_{l=1}^{L} s_l[n - lN]$ and

$$s_l[n] = \begin{cases} \mathcal{F}^{-1}a_l & := \frac{1}{\sqrt{K}} \sum_{k=1}^{K} a_{l,k} e^{\frac{2\pi ikn}{K}} \quad n \in [-L, K-1] \\ 0 & \text{else} \end{cases} \tag{4.18}$$

The inverse Fourier transform is K periodic and the first L entries of $s_l[n]$ are filled with the last L entries of $\mathcal{F}^{-1}a_l$ as cyclic prefix. The modulator can therefore be implemented using very low computational complexity applying the IFFT. The cyclic prefix is removed at the receiver as $\tilde{g}[n] = 0$ on $n \in [-L, 0]$. and the demodulator can be implemented using FFT. Since IDFT and DFT are of length K, the number of subcarriers is usually chosen as a power of 2 to allow fast computation. A variation of this transmission scheme is to use zero padding in the end of $s_l[n]$ instead of the cyclic prefix [63], [101].

4.3.3. Variations in Pilot Arrangements. Most classic standards use rectangular pulses and a rectangular design for the underlying lattice $\Lambda = \mathbb{Z} \times \mathbb{Z}$ as described in Section (4.3.2) out of which the pilot positioning Λ_1 is chosen according to the properties of the channel. The remaining lattice points $\Lambda_2 = \Lambda \backslash \Lambda_1$ are filled with data symbols. The most common choices for the positioning of the pilots are described subsequently.

Block type pilot arrangements. are usually used for slowly time varying channels. These are channels which stay almost constant for a block of a certain number of subsequent transmitted data symbols. Usually pilots are transmitted at all subcarriers and are equally spaced in time with a certain distance α: $\Lambda_1 = \alpha\mathbb{Z} \times \mathbb{Z}$. A block type pilote arrangement is used e.g. in WLAN standard IEEE 802.11a [1] and can be in practice combined with a decision feedback equalizer [105]

Comb type pilot arrangements. are the other extreme where pilots are inserted into each OFDM symbol whereas the set of subcarriers is subsampled with a sampling distance β. In our notation this means $\Lambda_1 = \mathbb{Z} \times \beta\mathbb{Z}$. The channel can then only be estimated at the subcarriers bearing pilots and has to be interpolated in between. Usually piece-wise constant functions, linear functions, quadratic functions or cubic splines are used with this respect [105]. WLAN standard IEEE 802.11g is for example based on such an arrangement. [2]

Hybrid pilot arrangements. are then a combination of both and used for example in multiuser WiMAX standard IEEE 802.16a [3]. They can be described as $\Lambda_1 = \cup_{i=1,2} \{(\alpha\mathbb{Z} + x_i) \times (\beta\mathbb{Z} + y_i)\}$

Frequency Domain Kronecker Delta arrangements. are used in case of large Doppler spread e.g. large ICI. This scheme is analogue to the hybrid arrangement with the difference that the piot subcarriers are not isolated but arranged in blocks. Only the center pilot is bearing information whereas the surrounding pilots have power zero and

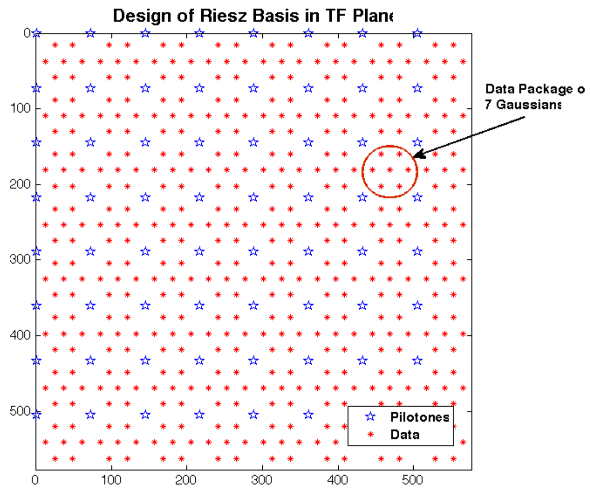

FIGURE 2. Gabor Riesz basis design in the Time Frequency plane with hexagonal pattern. The blue points show the placement of the pilot-tones, whereas the red points indicate the position of the data impulses.

only have the intentation to avoid interference between the pilot-tones and the data-symbols [79] [80]. This pilot arrangement can be found for example in applications like DVB-T [35].

4.3.4. Rectangular pilot-tones and polygonal data packages. For the design of the proposed Gabor Riesz basis, we assume a separable coarse lattice

$$\Lambda_1 = \alpha\mathbb{Z} \times \beta\mathbb{Z} \tag{4.19}$$

for the pilot-tones. In between these pilot-tones, the data packages are sent. Therefore the data carrying pulses are placed in the center an at the corners of a polygon P in the time frequency plane. This second lattice consisting of the data polygons will be denoted by

$$\Lambda_2 = \{T(\lambda)P\}_{\lambda\in\Lambda_1}. \tag{4.20}$$

Figure (4.3.4) shows this lattice design $(\Lambda_1 + \Lambda_2)$ with a hexagonal pattern. As atoms we assume to have shifted and modulated Gaussians and we denote the Gabor Riesz basis

$$G := (g_\lambda)_{\lambda\in\Lambda} \quad \text{with} \quad \Lambda = \Lambda_1 + \Lambda_2. \tag{4.21}$$

Figure(4.3.4) shows a comparison of a concrete MATLAB implementation of different Data/Pilot arrangement. Therefore 4 classical arrangements as described in section (4.3.3) as well as two different parametrisations of the new scheme togethere with the corresponding numerical efficiency are shown. These schemes will then also be used in the following sections for further experiments.

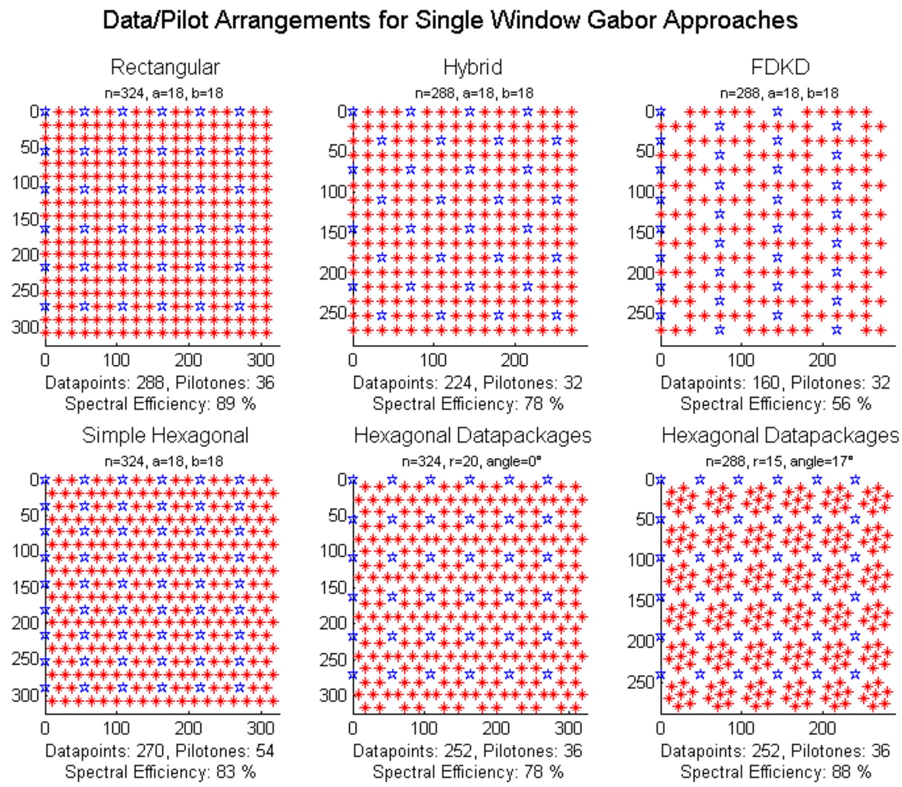

FIGURE 3. The figure shows a comparison of different Data/Pilot arrangements for single window Gabor Approaches indicating also their spectral efficiency. The first four are classical approaches whereas the last two figures in the second row give two different parametrisations of the proposed scheme.

4.3.5. Implementation. For the following numerical experiments the abstract design has been translated to a constructive algorithm in the finite discrete setting. As a constructor, an appropriate atom - as the sampled and periodized Gaussian - is chosen. Then copies of this atom are placed on the corners of the hexagon. Two different ways have been algorithmically implemented to accomplish the construction of an hexagonal Gabor Riesz basis as described above.

(1) The Gaussian atom g is shifted by the radius of the hexagon. The orthogonal basis of Hermite functions for dimension n is calculated and a coefficient mapping of g is done. Then the rotation is performed through pointwise multiplication of the coefficients. Finally a resynthesis mapping is done.
(2) The second possibility is to directly perform the shift by sin and cos of the according angle in the time frequency plane. For this method the implementation of a fractional time frequency shift is needed to avoid a discretization error.

The two methods are now described more formally and the equivalence of the two concepts is shown.

4.3.5.1. *Method 1: Fractional Fourier Transform.* The first approach uses the fractional Fourier transform $\mathcal{F}^a$. The main idea is to consider the $\mathcal{F}^a$ as a generalization of the usual Fourier transform in the sense that while the classical Fourier transform describes a rotation of a signal by $\frac{\pi}{2}$ in the time frequency plane, $\mathcal{F}^a$ corresponds to a rotation by an arbitrary angle $\alpha = \frac{a\pi}{2}$ with $a \in \mathbb{R}$. For a detailed study of the partial or fractional Fourier transform we refer to the literature, for example to [20] or [87].

The eigenfunctions of both, the classical and the fractional Fourier transform are the Hermite functions h_k . These special functions - often defined through the Hermite polynomials H_k - are given by

$$h_k(x) = \frac{2^{1/4}}{\sqrt{2^k k!}} e^{\frac{-x^2}{2}} H_k \qquad H_k(x) = (-1)^k e^{x^2} \frac{d^k}{dx} e^{-x^2} \tag{4.22}$$

As we have

$$\mathcal{F} h_k = e^{-\frac{i\pi k}{2}} h_k \tag{4.23}$$

the eigenvalues of the classical Fourier transform $\mathcal{F}$ to the eigenvectors h_k are $\lambda_k = e^{\frac{-ik\pi}{2}} = (-i)^k$. From the geometric intuition described above it seems natural to define $\mathcal{F}^a$ as

$$\mathcal{F}^a h_k = e^{-\frac{i\pi k a}{2}} h_k = \lambda_a^k h_k \tag{4.24}$$

with $\lambda_a = e^{-\frac{i\pi a}{2}} = e^{-i\alpha}$. It can be shown that

$$e^{\frac{ix\zeta}{2}} V_g f(x,\zeta) = e^{\frac{ix_a\zeta_a}{2}} V_{g_a} f_a(x_a,\zeta_a) \tag{4.25}$$

(compare [20] Theorem 5.2). As we use the Gaussian as a window function, which is invariant under Fourier transform we have further

$$e^{\frac{ix\zeta}{2}} V_g f(x,\zeta) = e^{\frac{ix_a\zeta_a}{2}} V_g f_a(x_a,\zeta_a) \tag{4.26}$$

This means

$$V_g g_j(x_a,\zeta_a) = e^{i\frac{x\zeta - x_a\zeta_a}{2}} V_g g(x,\zeta), \tag{4.27}$$

where x_a and ζ_a denote the transformed coordinates and are given by

$$x = x_a \cos\alpha + \zeta_a \sin\alpha \qquad \zeta = -x_a \sin\alpha + \zeta_a \cos\alpha. \tag{4.28}$$

In order to generate the polygon P, we first shift the Gauss function by the radius r and then we apply a partial discrete Fourier $\mathcal{F}^\alpha$ transform (compare [4], [22]) to $g(t-r)$ with an angle $\alpha = \frac{360}{k}$. In total, we get $V_g g_j(x_a,\zeta_a) =$

$$e^{i\frac{x\zeta - x_a\zeta_a}{2}} e^{-2\pi i r(-x_a\sin\alpha + \zeta_a\cos\alpha)} V_g g(x_a\cos\alpha + \zeta_a\sin\alpha - r, -x_a\sin\alpha + \zeta_a\cos\alpha) \tag{4.29}$$

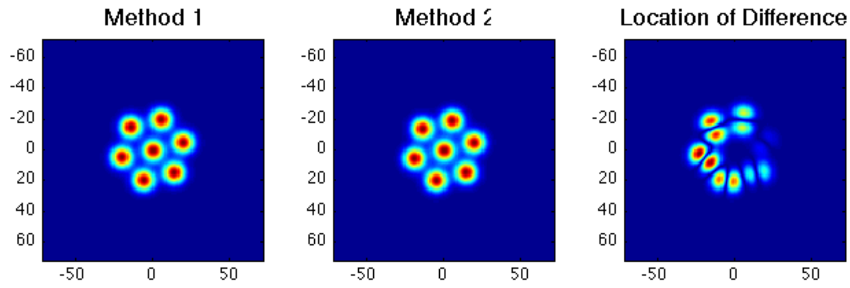

Figure 4. Location of the differences between the subspaces created by the two different methods.The visualization was implemented through projection on $(\Pi(\lambda)g)_{\lambda\in\mathbb{Z}_N}$ by the MATLAB routine 'stftsubspace'. The scaling in the third figure differs from the others by a factor of 10^{-1}.

For the implementation of the discrete fractional Fourier transform, the algorithm developed by Feichtinger as described in [96] section 4.3.3 was used.

4.3.5.2. *Method 2: Rotation and Modulation.* In the second approach, we first calculate the coordinates of the k corners of the polygon in the time frequency plane through

$$x_j = r\cos\left(\frac{2\pi j}{k}\right) \qquad \omega_j = r\left(\sin\frac{2\pi j}{k}\right). \tag{4.30}$$

Then the Gaussian g is rotated and modulated by (x_j, ω_j). Because of the covariance property of the STFT (compare [65]) we have

$$V_g g_j(t,\zeta) = e^{-2\pi i x_j\zeta} V_g g(t - x_j, \zeta - \omega_j) \qquad g_j = T_{x_j} M_{\omega_j} g. \tag{4.31}$$

This means that the absolute value of the STFT of the shifted Gabor atom ist identical to the absolute value of the shifted STFT of the atom, which is exactly what we wanted to do.

4.3.5.3. *Equivalence of the two approaches.* We compare equation (4.29) and equation (4.31). If $V_g g$ is radial symmetric e.g $V_g g(t,\zeta) = V_g g(\sqrt{t^2+\zeta^2})$ the two approaches give exactly the same result up to a phase factor as

$$\begin{aligned} & (x\cos\alpha + \zeta\sin\alpha - r)^2 + (-x\sin\alpha + \zeta\cos\alpha)^2 \\ = \;& t^2 + \zeta^2 + r^2 - 2rt = (t - r\cos\alpha)^2 + (\zeta - r\sin\alpha)^2. \end{aligned}$$

MATLAB Code for both approaches can be found in the appendix for the respective section.

Remark 4.3. In practice the results are not exactly the same due to numerical effects of the not exact rotational symmetry of the discretized Gaussians. The location of the differences between the spaces in the time frequency plane can be found in Figure (4). The higher the dimension n is chosen, the more similar the results become. The differences turned out to be of a size not of practical relevance in our simulations.

For the sake of completeness we want to mention that for the numerical experiments in the following the finite fractional Fourier transform was used.

4.3.6. Parametrization. As can also be seen in Figure (4.3.4), the parametrization of the proposed Gabor Riesz basis design is crucial for the performance of the transmission scheme with respect to spectral efficiency and stability. In the described setting, for the datapackages, there are 3 different parameters to choose:

- the number of edges of the polygon
- the radius of the polygon
- the angle of the polygon

In addition to these three main parameters one can choose the distance parameter of the pilot lattice which in turn also influences the stability of the channel estimation. First, we will assume the distance between the pilot-tones fixed and analyse the basis properties of the data packages. The target in this first optimisation cycle is to span a space, which is optimally localized in the time frequency plane while being well conditioned at the same time. This is performed via a comparison of the time frequency expansion of the space spanned by the Gaussian polygons - i.e. seven Gauss functions the STFT of which being centered in the corners and the center of a polygon in the time frequency plane - to the Hermite basis as reference for an optimally localized space. In this section we concentrate on numerical efficiency vs stability as for the numerical implementation the exact choice of parameters does not influence the result.

4.3.6.1. *Time Frequency Expansion compared to the Hermite Basis.* The aim of this section is to describe the time frequency expansion of the subspace

$$G_k = \text{span}\,\{g_i | i = 1, ..., k\}$$

spanned by k Gaussians. We will compare it to the space H_j spanned by the first j Hermite functions. The Hermite functions as special functions are perfectly orthogonal. Therefore, span $\{H_j\}$ will be the largest j dimensional rotation invariant space localized in the time frequency plane that can be covered. Figure (5) gives a visual representation of range of the projection operator the two spaces. The discrete Hermite functions have been obtained as eigenvectors of an STFT multiplier with radial symmetric weight functions (for details on this approach compare [124] section 3.3). Different measures for the overlap of the time frequency expansion of the spaces G_k and H_j are possible. The two options presented here, are the comparison of the projection operators as well as the calculation of the angle between the spaces. For both measures, we determine the optimal radius r of the polygon in G_r, which in our current analysis is equivalent to the radius covering the largest part of H_j, respectively. Finally we will explain the equivalence of both measures.

Comparison of projection operators. Let H_j denote the projection operator on the first j Hermite functions and P_k the projection operator on the space G_k of k

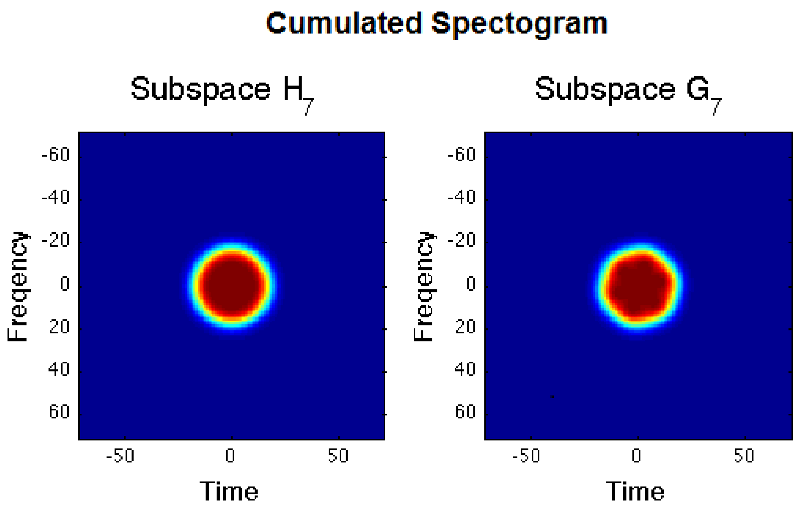

FIGURE 5. Location of the subspace of the first seven Hermite functions H_7 and the space spanned by seven Gauss functions placed in a hexagonal pattern G_7 in the time frequency plane. The visualization was implemented through projection on $(\Pi(\lambda)g)_{\lambda\in\mathbb{Z}_N}$ by the MATLAB routine 'stftsubspace'

Gaussian atoms placed as described in Section (4.3.4). We determine the size of the image of the operator $U_{jk} := (H_j - P_k \circ H_j)$ by

$$\varepsilon_{jk}^{max} = \|U_{jk}\|_{\mathrm{Op}} = ||H_j - P_k \circ H_j||_{\mathrm{Op}}, \tag{4.32}$$

which corresponds to the maximal singular value of U_{jk}. This measure gives the maximal length of a vector lying in $\{H_j \backslash P_K\}$. In case of two identical spaces, the measure gives zero as the projection operator is idempotent. On the other hand, for orthogonal spaces the measure gives 1 as for all $f \in H_j$ with $\|f\| = 1$, $H_j f = f$ and $P_k \circ H_j f = P_k f = 0$. Further, we consider the Frobenius norm or mean square error given by

$$\varepsilon_{jk}^{mse} = \frac{||H_j - P_k \circ H_j||_{\mathrm{Fro}}}{\sqrt{N}}. \tag{4.33}$$

This value can be seen as a measure for the average length of a vector being in H_j but not in G_k. Figure (6) provides an image of the results for ε_{jk}^{mse} and ε_{jk}^{max} from the MATLAB experiment for $k = 7$. The radius was normalized with respect to the dimension $r^* = \frac{\tilde{r}^*}{\sqrt{n}}$. The experiment was repeated for 10 randomly chosen n between 100 and 500 in order to exclude any numerical effect of the dimension and the difference in operator norm as well as in Frobenius norm was calculated. Restricting now to the space H_7, figure shows in the top two boxes, a cross section plot the dependence on the radius and the number of edges of the polygon in more details. It can be seen that the error reaches its minimum at around 0.2 times $\sqrt{n}$. At the minimum the error is of size around 10^{-15} and increases the steadily to one slightly above a radius of $\sqrt{n}$. If we choose the radius optimally, the spaces spanned by the

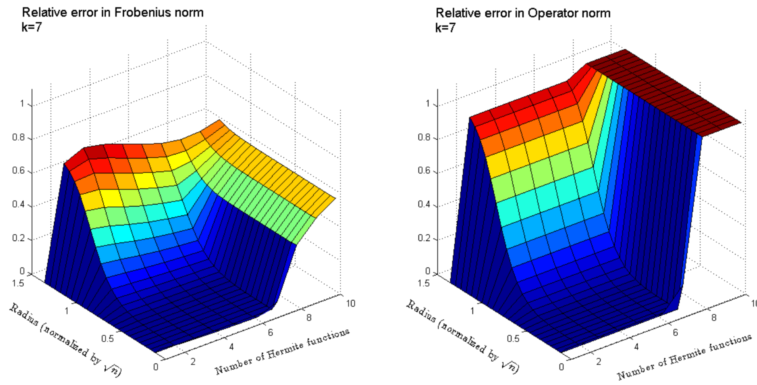

FIGURE 6. The graph shows the relative error for the difference in projections on the space of j Hermite functions and on the space of 7 Gaussians. The experiment has been performed with a dimension of $n = 288$ and repeated 10 times with randomly chosen dimensions giving similar results (Source code provided in the appendix). It can be seen that for small $r > 0$ the spaces are very similar and diverge with increasing radius starting from about 0.5 to 1 times $\sqrt{n}$. The increase in the relative error in both norms for the space of Hermite functions up to a dimension of 7 is very similar whereas it is obviously not possible to reasonably well cover a larger space with dimension 8 and larger.

first 7 Hermite functions can be considered identical up to an error of 10^{-14}. This can also be seen in the 3D plots at the bottom of the figure showing the errors between the projection spaces between different combinations of numbers of Gaussians and Hermite functions.

The plot including in the lower right surface plot reveals the behavior in case the radius in not chosen in an optimal way. There we chose the radius to be $0.5\sqrt{n}$. Then it can be seen that instead of 10^{-14} we get an error of 10^{-6}. Furthermore whereas in case the radius is optimally chosen, it does not add any value to increase the dimension of the space further by adding more edges to the polygon an improvement can be observed for the situation when the radius is not chosen in an optimal way. As we are considering a rotational symmetric space without yet adding the pilot-tones to the corners of the square yet, the angle can be chosen arbitrary in this experiment. For applications it is also of interest how large the radius has to be chosen at least in order to guarantee that $\dim\{G_7\} = 7$. Experiments show that this is the case starting already from a radius of one. In practice, this setting is extremely unstable and hence we will analyse it further in the next section.

4.3.6.2. *Stability.* A part from the spectral efficiency, for applications the central requirement is stability of the Gabor system with particular focus on the stability under the reconstruction with the dual system (compare Section 4.3.1.2).

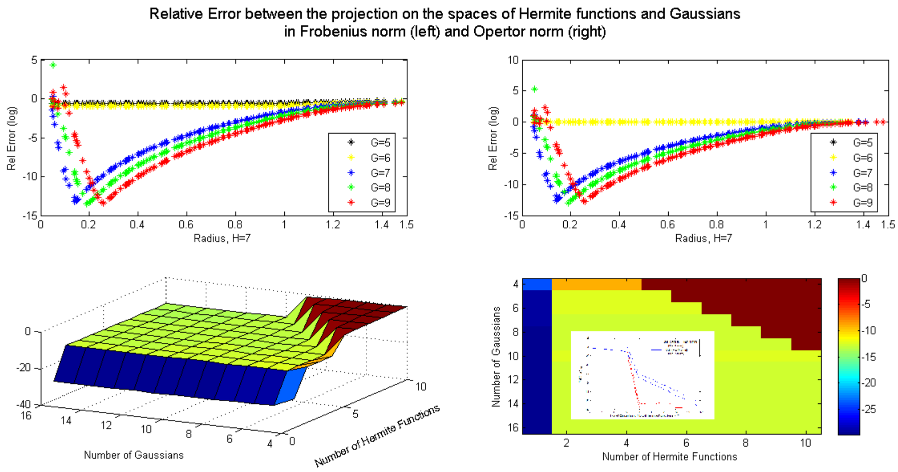

FIGURE 7. The plot shows the relative error between the projection spaces on the first j Hermite functions and i Gaussians placed in the corners and in the middle of a polygon as described in equation (4.33). Results in Frobenius norm are given on the left side and in operator norm on the right side for different radius. The experiment was done with 10 different randomly chosen dimensions N between 100 and 500. On the $y-$axis the radius normed by $\sqrt{N}$ is plotted. It can be seen that the optimal choice of the radius with respect to covering of the space H_7 is around $r = 0.2\sqrt{N}$. Moreover it can be seen that the space of j Hermite functions (j between 4 and 10) can always be covered with the same number of Gaussians placed on a polygon up to a relative error of size 10^{-14}. Whenever the radius is not chosen in an optimal way this error increases and it can be advantagous to increase the dimension of the space, which has no impact when the radius is chosen optimally.

In this respect, the condition number of the transmission matrix (i.e. Gabor matrix multiplied with the channel matrix) is a common measure for the sensitivity of the demodulated signal with respect to distortions in the received signal. Precisely, the condition number is defined as the maximum ratio of the relative error of the reconstructed signal devided by the relative error of the received signal. Using the notation of the system of linear equations at the resynthesis step in (4.5), $\mathrm{cond}(G) = \max(||\Delta\tilde{a}||/||\tilde{a}||)/(||\Delta r||/||r||)$. By rewriting the definition, the relative error in the reconstructed symbol can be estimated as $\frac{||\Delta\tilde{a}||}{||\tilde{a}||} \leq \mathrm{cond}(G)\frac{||\Delta r||}{||r||}$. In case of a theoretic perfectly deterministic channel equal to the the identity operator $r = Ga$, $\Delta r = 0$. Assuming a slightly disturbed channel, such that the received signal is $\tilde{r} = \tilde{G}a$, gives $||\Delta r|| = ||Gr - \tilde{G}r|| = ||(G - \tilde{G})r||$ and we get

$$\frac{||\Delta\tilde{a}||}{||\tilde{a}||} \leq \mathrm{cond}(G)||G - \tilde{G}||_{\mathrm{Op}}. \tag{4.34}$$

Now assume that $\tilde{G}$ is obtained by a time frequency shift operator as $\tilde{G} = \Pi G,$. By the equation

$$\frac{||\Delta\tilde{a}||}{||\tilde{a}||} \leq \operatorname{cond}(G)\|\operatorname{Id} - \Pi\|\|G\| \tag{4.35}$$

the upper bound of the error in the resynthesed symbols caused by the distortion of the channel and the condition number of the Gabor system can be estimated. In Chapter (5.2) we will look at this in more details and make the connection to the most common error measure in applications, the bit error rate.

In this section, we elaborate on the stability concept by a series of MATLAB simulations going from the simple concept to the general case. The dependency of the conditioning on parametrization of the lattice, size of distortions as well as spectral efficiency are stated in the subsequent paragraphs.

Experiment 1: 1 block, no pilot-tones. In the first experiment we want to get a general impression on the dependence of the conditioning of the problem in dependence of the radius. Therefore we chose the same setting as in the last section and create a Gabor Riesz basis of 7 Gaussian atoms, one in the center surrounded by a hexagonal pattern. As radius has to be set in relation to the dimension, it was normalized by $\sqrt{n}$. The results in Figure (8) clearly state a highly nonlinear increase of the condition number for decreasing radius.

Experiment 2: (5×5) blocks with pilot-tones. In this setting, a transmission scheme of 5×5 data blocks consisting of 7 Gaussians in a hexagonal pattern and in total 25 pilot-tones was created. Therefore the space spanned by the Gabor system has dimension $\dim(\operatorname{span}\{G\}) = 200$. For varying radius and dimension $N \geq \dim(\operatorname{span}\{G\})$ the condition numbers c of the respective systems were calculated in dependence of radius and angle. Results are given in Figure (9). In can be seen that for different spectral efficiencies the order of the condition number differs but the optimal region at around 0.35 times the grid-size of the product lattice stays stable.

Experiment 3 - Confirming the results with varying dimensions. The third experiment extends the first two by going to a higher dimension and varying system sizes Therefore additional blocks G_j where added consisting of 7 Gaussians each, placed in a hexagonal pattern. In between the hexagons, Gaussian pilot-tones are situated on a rectangular lattice with lattice constant a between 5 and 1000. To avoid numerical effects of a particular parameter combination, besides the variation of a, the experiment was done for a different number of blocks between 1 and 25 of size $(a \times a)$ each. Whereas the condition number obviously depends on the dimension and redundancy of the overall system, the choice of the optimal choice region for the radius almost only dependent on the sampling distance a of Λ_1.

Results are shown in Figure (10). It was confirmed by the experiment that the choice of the radius is crucial for the condition number of the system. Optimal results are obtained by choosing the radius in a region of 0.25 to 0.35 times the sampling

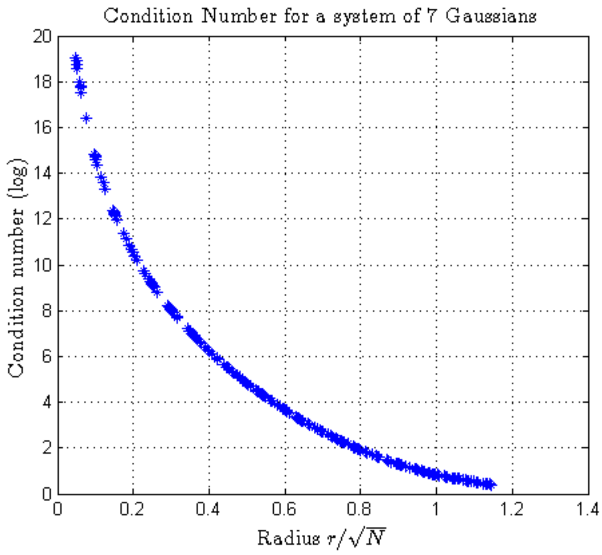

FIGURE 8. The graph shows the condition number of a Gabor Riesz basis consisting of 7 Gaussian atoms in a hexagonal design. This represents the transmission design for one data package without pilot-tones. The condition numbers are plotted on a logarithmic scale in dependence of the the radius normed by $\frac{1}{\sqrt{N}}$ and results were combine for 10 runs with different dimensions N randomly chosen between 100 and 500. It can be seen that contrary to the projection space coverage analyzed in the last section, the condition number increases sharply with decreasing radius of the hexagon.

distance of the pilot-tones, where the optimal choice for higher spectral efficiencies is rather on the upper end. These results stay stable for different choices of dimensions and redundancies. In addition the choice of the optimal angle of 17 degrees rotation also gives a slight improvement of the results. The effect is however very small and only present for high spectral efficiency. For higher subsampling, meaning lower spectral efficiency no significant impact of the choice of the angle could be found.

Figure (11) shows again a two dimensional representation of the dependence of the condition number on radius and angle for one example. In this plot the importance of the choice of the radius compared to the angle of the hexagon effect of the angle is only present for non optimal choice of the radius or a low degree of subsampling.

Experiment 4 - Simulation for optimal radius under scattered pilot-tones. To be able to estimate the effect of this random perturbations, a Monte Carlo simulation was performed. Therefore, as in the previous experiments, a Gabor Riesz basis consisting of 16 blocks of size 40 containing 7 atoms each and one pilot-tone was created. This sample design corresponds to a redundancy of 1.25. Then, the pilot-tones were perturbated by small rotations and modulations corresponding to an underspread channel with normally distributed scattering function. The sizes of rotations and

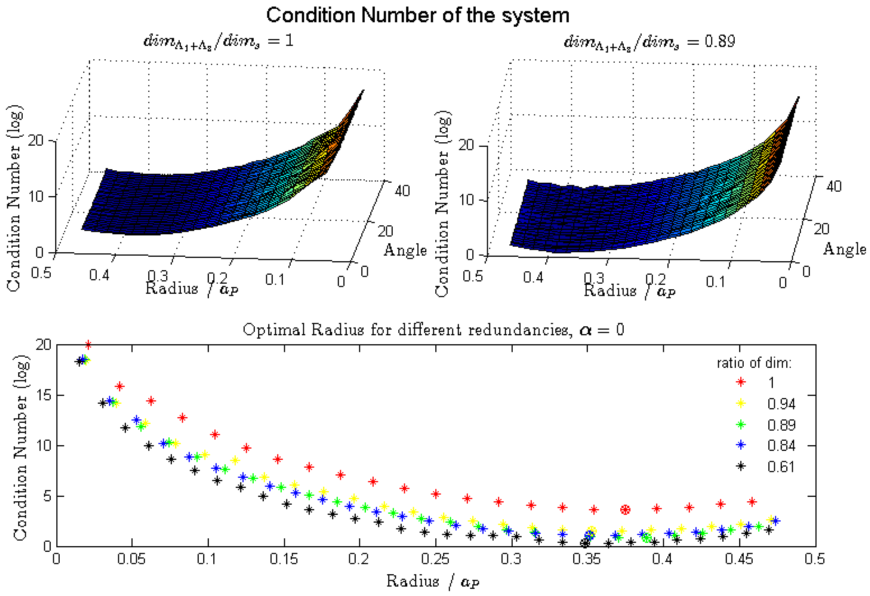

FIGURE 9. The graph shows the condition number of a Gabor Riesz basis consisting of a transmission scheme of 25 data packages consisting of 7 Gaussians each and 25 pilot-tones. This gives a 200 dimensional system of Gaussian pulses. In order not to neglect the effect of the spectral efficiency the experiment was done for different signal space dimensions $n \geq 200$ and the ratios of the dimensions are stated in the plot. The condition numbers are plotted on a logarithmic scale in dependence of the the radius divided by the grid-size of the pilot lattice a_p and the angle of the hexagon α. An $\alpha = 0$ corresponds to the hexagon having one corner exactly at 90 degree of the circle; α then denotes the rotation to the right. The upper two surface plots show the common dependence on radius and angle. It can be seen that while there is a significant dependence on the radius, the dependence on the angle is rather negligible. This holds true also for high spectral efficiencies. The plot in the bottom shows the optimal radius minimizing the condition number for different signal space dimensions. It can be seen that the radius is stable at 0.35 times the grid-size of the product lattice.

modulations were chosen according to a randomly sampled normal distribution with expectation 0 and standard deviation σ. The distribution of timeshifts and frequency shifts was chosen independently from each other. The data atoms were kept deterministic. Then the optimal radius and the corresponding minimal condition number was calculated for each run. The simulation was done with 100 runs and finally the sample statistics for radius and condition number were calculated. Figure (12) shows the results, which confirm that the optimal parametrization stays the same also under small perturbations.

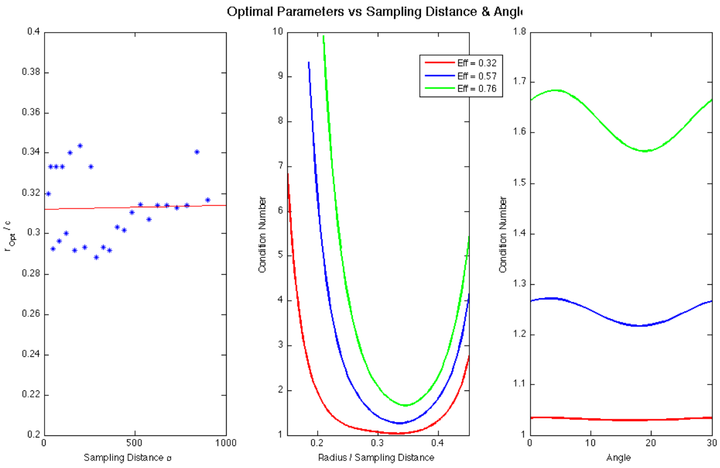

FIGURE 10. The graph on the left side shows the optimal relation of radius and sampling distance a of Λ_1 for different sampling distances. Results do not depend on the overall dimension of the system. A linear trend line was fitted to show the constancy of the obtained results. The second plot shows the relation r/a versus the condition number for a system of 4 blocks and the spectral efficiency indicated in the legend. It can be seen that the system stays stable for a choice of the radius around 0.32 times the sampling distance. For higher spectral efficiencies the system becomes more sensitive to the choice of the radius. The third plot shows that for low spectral efficiency the choice of the angle of the hexagon does not have any impact on the condition number. For higher spectral efficiency a small impact also of the angle could be found. The optimal choice of the angle would be a rotation of 17 degrees.

It becomes clear that the expectation stays stable also under large pertubations even under rather small redundancy of 1.25. The distribution for the condition number becomes wider and therefore large outliers occur more often.Those outliers would mean that in this case a small deviation in the input data symbol could lead to a large change in the reconstructed signal. From this example we also get an idea about the foremost importance of the pilot separation as they are essential for the correct reconstruction of all surrounding blocks.

Experiment 5 - Stability of the biorthogonal system. Let again G denote the system of 7 Gaussians $(g_i) := \pi(z_i)g$ arranged in a hexagonal pattern. Subsequently, the biorthogonal system is determined by the pseudoinverse $G^d := (G'G)^{-1}G'$. Let $G^p := (g_i^p) = \pi(z_i^{'})g$ denote the perturbed system. Then the entries in the Gramian are given by

$$\langle g_i^p, g_j^d \rangle = \langle \pi(z_i^{'})g^p, \pi(z_j)g^d \rangle = \langle g, \pi(z_i^{'} - z_i)g^d \rangle. \tag{4.36}$$

For the simulation in MATLAB two random vectors u and v of length 7 uniformly distributed in the interval $[-s_{\max}, s_{\max}]$ are created - one for the rotations and one

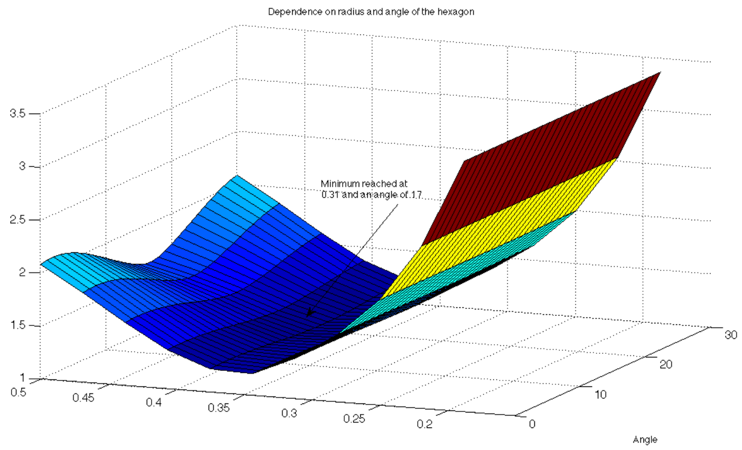

FIGURE 11. The figure shows the condition number of a Gabor Riesz basis consisting of Gaussian atoms in a hexagonal design with pilot-tones for a subsampling of 3 in dependence of the radius and angle. The results shown correspond to the experiment with 81 blocks of size 32 times 32.

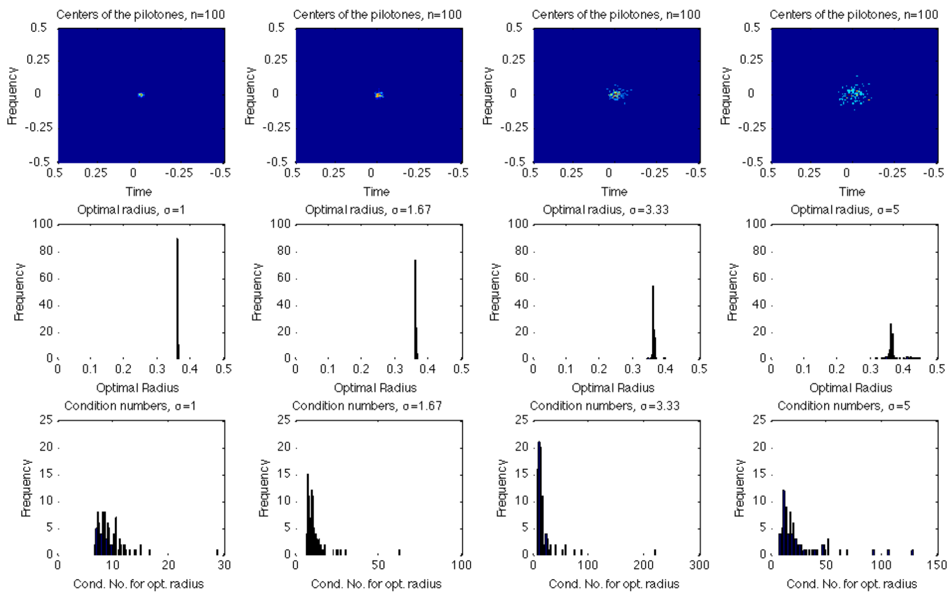

FIGURE 12. The figure shows the results of the Monte Carlo simulation for pilot-tones under normal distributed random perturbations with standard deviation σ. For each simulation run the optimal radius and the according condition number was calculated and the sample histogram is shown. The experiment was done for a setting of 16 blocks of size 40 corresponding to a redundancy of 1.25. The experiment was repeated for different values of σ between 1 and 5.

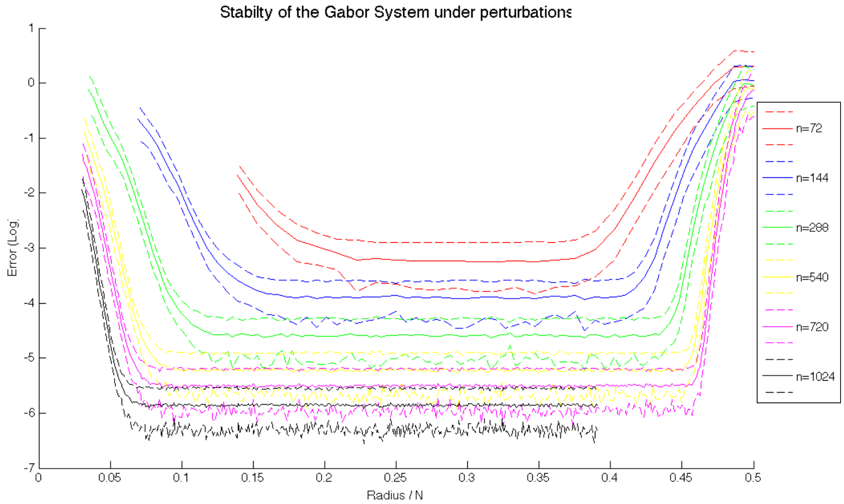

FIGURE 13. The figure shows the largest singular value of the Gramian $\left\langle g_i^p, g_j \right\rangle$, with g_i^p denoting the atoms of the perturbed Gabor system and (g_j^d) the dual system to the original Gabor system. The error is plotted with respect to the chosen radius of the hexagon for different dimensions n. Perturbations are always restricted by a maximum rotation of $1/n$ combined with a maximum modulation of $1/n$ for each atom. The perturbations are determined randomly. The plot shows the mean and an interval spanned by $\pm 2\sigma$ for a sample of 100 realisations for each dimension.

for modulations. The perturbed Gabor system $G^p := (g_i^p)$ is calculated through

$$g_i^p = \Pi(u_i, v_i) g_i \qquad i = 1, ..., 7 \tag{4.37}$$

In the following the maximal error is determined through

$$\varepsilon := \lambda_{\max}(Id_7 - A) \quad \text{with} \quad A_{ij} = \left\langle g_i^p, g_j^d \right\rangle \tag{4.38}$$

and $\lambda_{\max}$ denoting the maximal singular value of the Gramian A. Then ε describes the maximal error which can occur on the spanned 7 dimensional space. The experiment was conducted for different dimensions N and different radius r and repeated 100 times for each setting in order to exclude numerical effects. Figure (13) shows the results of the experiments. It can be seen that for proper choice of the size of the radius e.g. a region of $0.2 \leq (r/N) \leq 0.4$ the system is stable and the changes in the biorthogonal system are very small. It can also be seen from the graph that the error becomes smaller for smaller perturbation of the time frequency lattice. In the experiment the size of the perturbation has been kept constant as the dimension n increases, which is equivalent to making the perturbations smaller. We will go into more details on the relation between the size of scattering and errors in the next section. Again the according code can be found in the appendix.

Experiment 6 - Variation of the spectral efficiency. This experiment summarizes above findings and explains the variation of the choice of the optimal radius by the

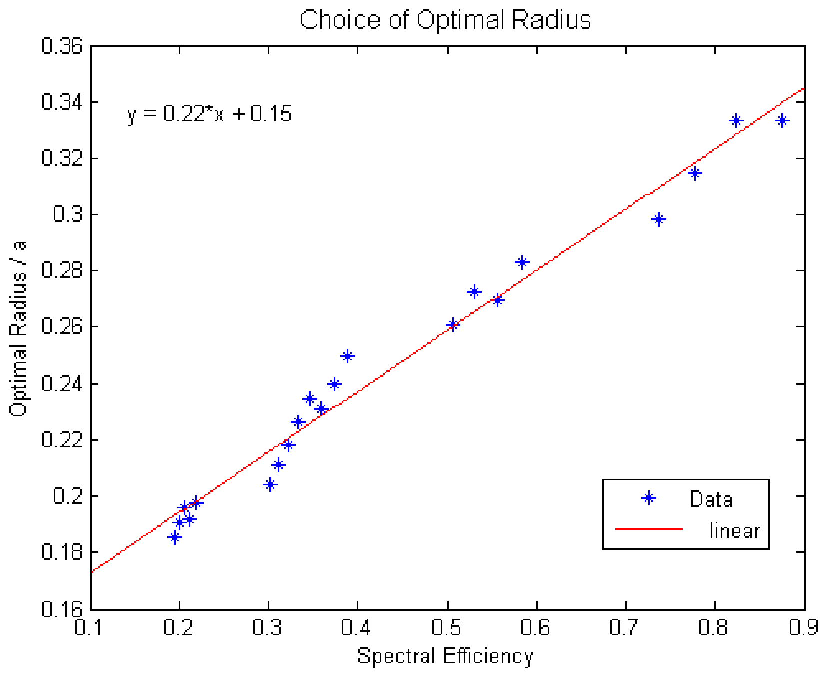

FIGURE 14. The figure shows the dependency of the optimal radius of the polygon data lattice on the spectral efficiency. Therefore, a family of Gabor Riesz bases with different spectral efficiencies following the suggested design was created. For each of them, the optimal parametrization with respect to the condition number was performed individually. The optimal radius was then normed by the gapsize of the pilot lattice and plotted against the spectral efficiency. A simple linear least squares fit was performed. It can be seen that the plot shows a linear behavior confirming also the results of the preceding experiments. The optimal ratio (r^*/a_P) starts slightly below 0.2 for very low spectral efficiency and reaches to 0.37 with a value slightly above 0.3 for the usual range.

variation of the spectral efficiency. Therefore a family of transmission schemes with different spectral efficiencies was created. For each of the schemes an optimization of the radius and angle was done individually. As no significant impact of the choice of the angle was found, the angle was then set to 15 degrees for further reiterations of the experiment and only the results for the optimization of the radius are shown here. The optimal radius was subsequently set in relation to the gapsize a_P of the pilotl attice Λ_1. Figure (14) shows the plot of the relation between (r^*/a_P) and the spectral efficiency. Results and in particular the equation of the linear fit confirm clearly the optimal ratio of slightly above 0.3 for common settings with a range between 0.2 and 0.37 for more extreme settings.

As the numerical results are very stable, we refrain from doing from a strict theoretical optimization in this setting. For more general results on the perturbation

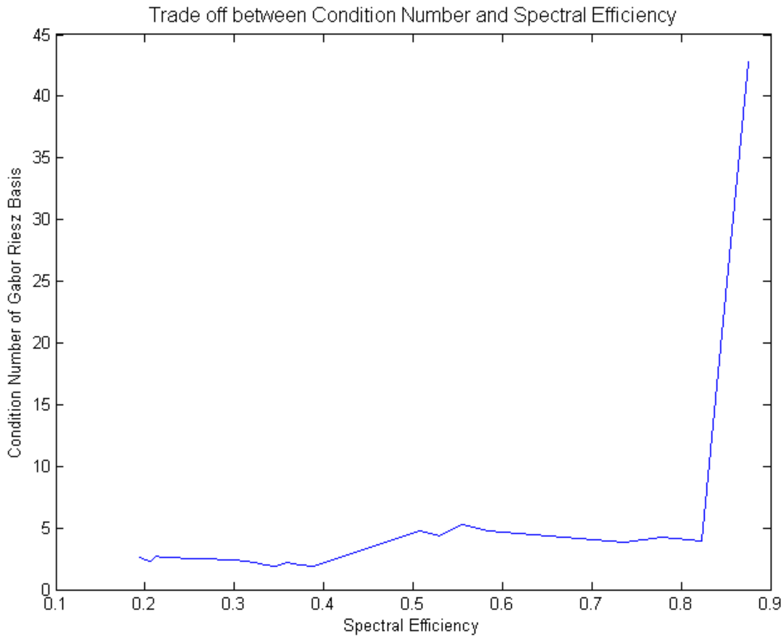

FIGURE 15. The figure shows the dependency of the condition number of the Gabor Riesz basis on the spectral efficiency. Therefore a family of Gabor Riesz bases with different spectral efficiencies following the suggested design was created. For each of them, the optimal parametrization with respect to the condition number was performed. The optimal radius was then normed by the gapsize of the pilot lattice and plotted against the spectral efficiency. Overall a very stable behavior absolute values around for even for high spectral efficiencies - except for the extreme case where the dimension of the Gabor Riesz basis (including pilot-tones) equals the dimension of the signal space - can be observed.

theory of Gabor systems we also refer to the work of Ole Christensen e.g. to [19] and [25].

4.3.6.3. *Trade off between spectral efficiency and stability.* Not only, is the conditioning of the underlying Gabor Riesz basis - and therefore the stability of the resynthesis mapping - subject to parametrisation of the data packages but also dependent on the chosen spectral efficiency of the overall transmission scheme. This has also been confirmed in Experiment (6) of the preceding section. To analyze this effect, again a family of transmission schemes with different spectral efficiencies was created. For each of the systems, the choice of parameters r and α was optimized and finally the spectral efficiencies are opposed to the condition numbers in the plot in Figure(15). Overall a stable behavior with an outlier at the boundary case was observed.

CHAPTER 5

Implementation and Error measures

5.1. Numerical Efficiency - Localization of the biorthogonal System

The main goal in applications is to recover the original information from the received data; which means to find the biorthogonal system. To calculate the overall biorthogonal system, however, is a computational challenge and cannot be performed near time due to the fact that the whole signal has to be passed before channel estimation can start. Therefore, the core question is how the global biorthogonal system relates to the local one. More specifically we want to investigate the changes within the biorthogonal system when enlarging the system to invert by adding blocks in a cyclical manner around the central block of seven Gaussian data carrying atoms.

Experiment 1. A Gabor Riesz basis with the suggested hexagonal design has been set up. We will denote one block of 7 data packages with $G = (g_j)_{j=1,\ldots,7}$. First, the local biorthogonal system $\tilde{G}^1 = (\tilde{g}_j^1)_{j=1,\ldots,7}$ of one block G and the surrounding 4 pilotones has been calculated. Then, the neighboring blocks have been added such that a system of 9 blocks was obtained in total and again the biorthogonal system $\tilde{G}^2$ was determined. In the same manner, in every subsequent iteration, an additional circle was added and the dual basis $\tilde{G}^k, k = 3, \ldots, 5$ calculated. This was done for systems of (5×5), (7×7) and (9×9) blocks . As size for one block in this experiment (90×90) was chosen and the dimension of the overall system was set to 960 .

In order to measure the error, we calculate $\varepsilon = \|Id - G'G_i\|$, which gives the maximal possible error occurring due to using the local dual system instead of the overall dual system. In addition we also measure the relative change of each dual atom by an additional iteration step $d_i = \frac{\|\tilde{g}_j^k - \tilde{g}_j^{(k-1)}\|}{\|\tilde{g}_j\|^k}$. The results of the experiments are shown in Figure (1).

From [53] we know that the calculated local dual $\tilde{g}_i^k$ will converge to the exact canonical dual $\tilde{g}_i$ for $k \to \infty$ under the assumption that the matrix is diagonal dominant. The formal proof is based on the Janssen representation and the uniqueness of the projection operator. For further details we refer to [53]. Even though no evidence of diagonal dominance in this setting could be found, the numerical results would indicate the convergence of each single dual atoms . We find the relative change in each iteration step to be of size 10^{-1}. The structure of the central local biorthogonal atom $\tilde{g}_1^k$ in the time frequency plane can be seen in Figure (2).

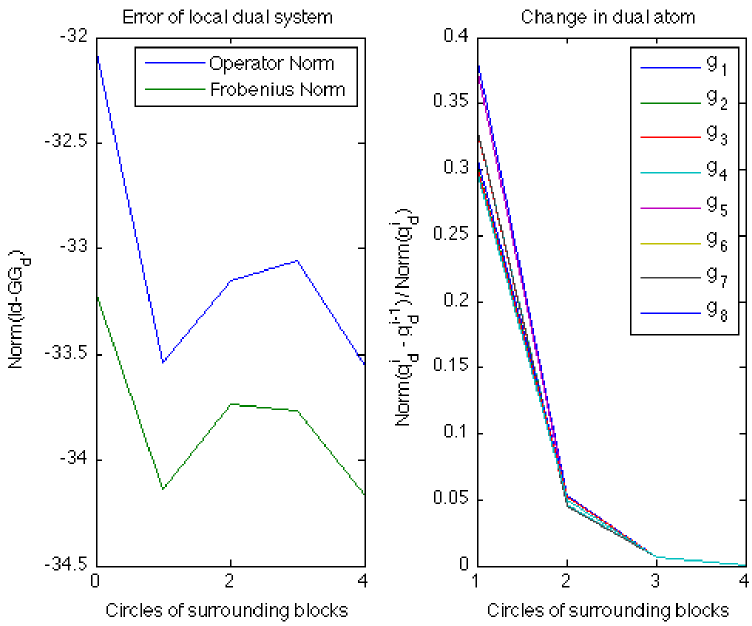

FIGURE 1. This figure shows on the left side the error in operator norm, which occurs by decomposition with the local biorthogonal system instead of the overall biorthogonal system. Therefore 4 iterations were performed by adding one additional circle of surrounding blocks in each iteration. The overall dimension in this experiment is $n = 960$. One block is of size $b = 90$. On the right hand side the relative difference of each of the 7 central data carrying atoms is shown for each iteration step. It becomes clear that whereas the relative difference in the dual atoms reduces by the size of 10^{-1} in each iteration step, the for applications relevant effect for the decomposition is already of size 10^{-32} starting from the first iteration step.

The other, probably for applications more relevant, finding is that the maximal possible error caused through decomposition with the local dual system is already without usage of any surrounding block of size 10^{-32} and does only reduce to a size of 10^{-34} by adding one additional circle of surrounding blocks. Then no considerable further improvement could be found.

Experiment 2. This experiment is intended to extend the findings of Experiment 1 and move further ahead towards applications. It is no longer limited to one block, but also takes into account interference effects between different blocks. Therefore a system of (10×10) blocks with a size of (106×106) each was set up. This gives a dimension of 1090 and a redundancy of 1.325. Then, the local pseudoinverse for each block was calculated followed by the decomposition of the system. As error measure again $\|Id - G\tilde{G}^d\|$ was used. First, the pseudoinverse of only one block of 7 data atoms was used together with the surrounding 4 pilotones. Then, an additional circle of 8 blocks was added and finally a second circle ending up with (5×5) small blocks.

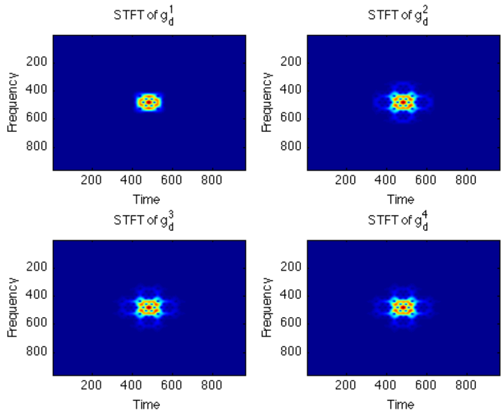

FIGURE 2. This figure shows the dual atom $\tilde{g}_1^k, k = 1, ..., 4$ in the time frequency plane. The dual atoms are obtained by calculating the local biorthogonal system and adding one surrounding circle of lattice tiles containing 7 data carrying atoms and the surrounding pilotones in each iteration step.

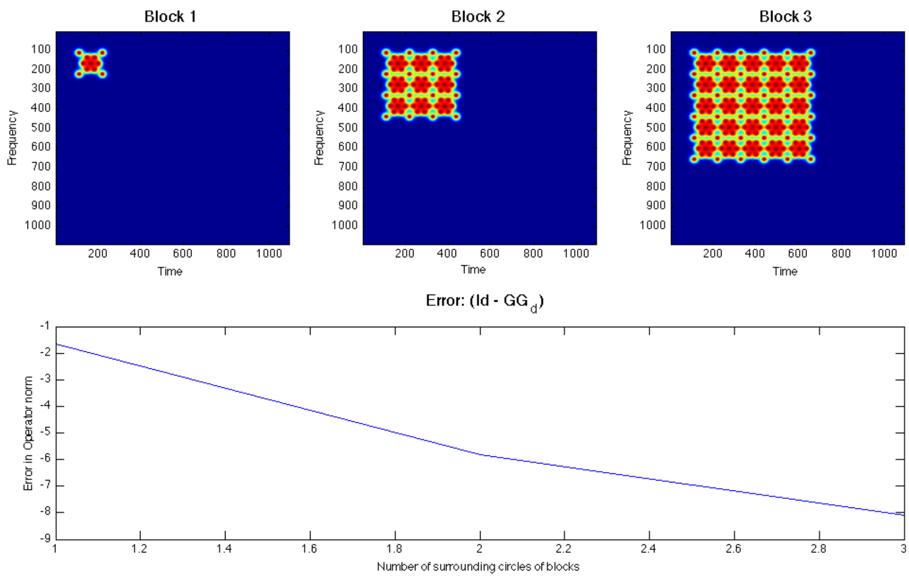

FIGURE 3. This figure shows on the top as an example the subspace spanned in the time frequency plane by the first block of atoms used for local biorthogonalization. In the bottom the error $(Id - G\tilde{G_i^d})$ is displayed in dependence of the number of the approximation i. The number of the approximation corresponds to the number of circles surrounding the atoms for which the biorthogonal atoms are calculated.

To visualize this the subspace in the time frequency plane spanned by the first block for each run is displayed in figure (3)

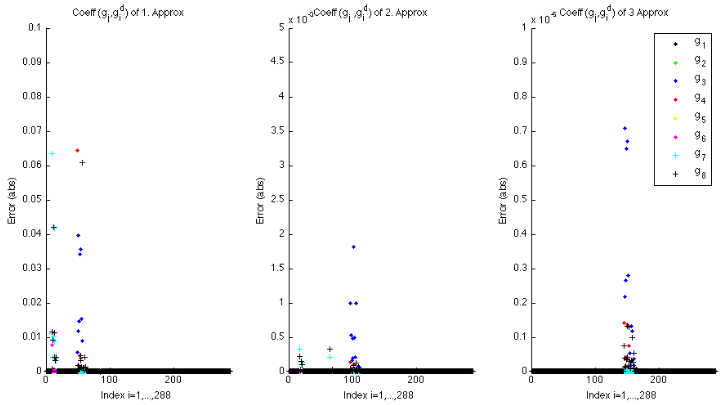

FIGURE 4. This figure shows the coefficients of the error matrix $e_{ij} = |Id_{ij} - \langle g_i, g_j \rangle|$ for $i = 1, ..., 8 \quad j = 1, ..., 288$ under local biorthogonalisation. The pilotone is g_1 and g_2 denotes the center of the hexagon. It can be seen that the error stems mostly from the atoms located at the edges of the hexagon.

The experiment shows that there are interference effects between the blocks, but the error decreases fast when using additional surrounding blocks for the calculation of the biorthogonal system. Whereas the error in the first approximation, relying only on the atoms under consideration and the 4 surrounding pilotones is about 10^{-2} it demishes to 10^{-8} in the 3^{rd} run.

A second plot shown in figure (4) displays the coefficients

$$e_{ij} = |Id_{ij} - \langle g_i, g_j \rangle| \qquad i = 1, ..., 8 \quad j = 1, ..., 288 \tag{5.1}$$

of the error matrix. Here g_i denotes the pilotone and g_2 the center of the hexagon. It can be seen that the error mainly roots from the atoms located on the edges of the hexagon.

Experiment 3. To visualize the spread in the time frequency plane induced by the locally performed biorthogonalization, a third experiment was conducted. Therefore the coefficients

$$M^j_{\lambda,\lambda'} = \langle g_\lambda, \tilde{g}^{d_j}_{\lambda'} \rangle \quad \forall \lambda, \lambda' \in \Lambda \tag{5.2}$$

were calculated, with $\tilde{g_\lambda}^{d_j}$ denoting the approximation of the dual atom $g^d_{\lambda'}$ to $g_{\lambda'}$ under the j^{th} approximation. Then the matrices M_j for $j = 1, ..., 3$ were transformed to their spreading representations $\eta(M_j)$, $j = 1, ..., 3$. Figure (5) shows the results. It displays a plot indicating all locations where the absolute value of the entry in the matrix $\eta(M_j)$ exceeds the predefined threshold of 10^{-5} to 10^{-1}. From this the periodization effects in the spreading function occurring through the localization of the biorthogonalizaion can be seen. For $\eta(M_1)$ rather strong effects for the distance of one block were observed. This effect disappeared in the second iteration. For the third approximation also the periodization effects for the distance of two blocks

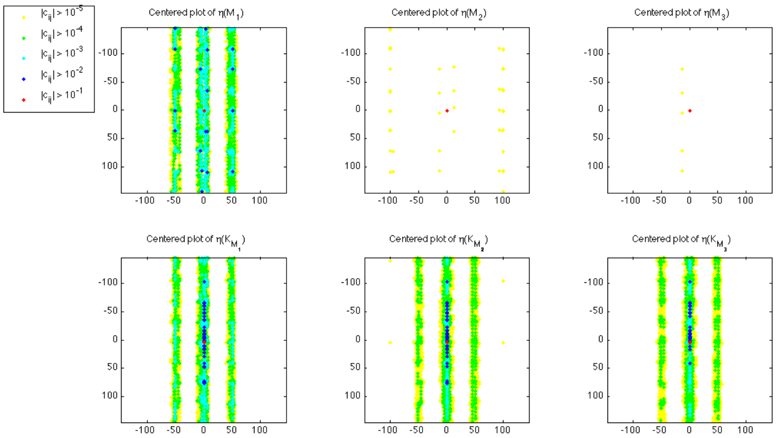

FIGURE 5. This figure shows the spread effect in the time frequency domain induced through local biorthogonalization. In the top row the effect without channel is shown, whereas in the bottom the effect after application of a channel is shown. It can be seen that the first iteration improves the results significantly, whereas the second step only brings a slight improvement. After the third iteration in the biorthogonalization the periodization effect disappears almost completely. The large spread under the channel is mostly due to the large spreading domain of the channel as a result of low dimension of the experiment. It can however be seen that from the second to the third step no remarkable improvement can be reached any more.

in the time frequency plane disappear, which are still slightly present in the second approximation.

In the bottom of Figure (5) the effect of disturbance through a channel can be seen. Therefore an underspread channel operator K was simulated. The spreading function $\eta(K)$ for the channel was constructed by generating random numbers on the predefined maximal domain of the spreading function ($\Omega := [-\tau_0, \tau_0] \times [-\nu_0, \nu_0]$) following a normal distribution. For the current experiment the domain of the spreading function was chosen of the size $(\tau_0 \nu_0) = 0.075N$, which is much larger than usually observed in typical applications. This was chosen on purpose in order to understand the effect under extreme conditions as well as because of numerically faster implementation in small dimensions. By Fourier Transform and shifting the rows of the matrix to the side diagonals, the channel operator kernel was obtained (this is performed by the MATLAB function $K = \text{spr2mat}(\eta(K))$. Then, the matrices K^{M_j} of inner products

$$K^{M_j}_{\lambda,\lambda'} := \left\langle K g_\lambda, \tilde{g}^d_{\lambda'} \right\rangle \quad \forall \lambda, \lambda' \in \Lambda \tag{5.3}$$

were calculated and the corresponding representation in the spreading domain $\eta(K^{M_j})$ investigated. Due to the large spreading domain and the low dimension the effect of the channel is predominant. Nevertheless, it cam be seen that there is a significant improvement between the first and the second approximation whereas the third approximation only yields slightly better results.

As the principle of the effect in the spreading domain already becomes clear, we will restrain at this point from further repeating the experiment in larger dimension. The code can be found in the appendix for replication of the results.

5.2. The bit error rate

This section describes the connection between the underlying work to one central measure for data transmission quality, the bit error rate. Results as outlined in Figure (10) indicate a higher tolerance to distortions caused by underspread channels of the proposed transmission scheme compare to the use of the Hermite basis.

Let $(a_\lambda)_{\lambda\in\Lambda}$ denote the set of data symbols transmitted with a QPSK scheme taken from the standard alphabet $\{1, i, -1, -i\}$ transmitting 2 bits of information each. Then the composed signal at the transmittor is of the form

$$s(t) = \sum_{\lambda\in\Lambda} a_\lambda g_\lambda(t) \qquad g_\lambda := \pi(\lambda)g. \tag{5.4}$$

The received signal after the application of the channel is then given through

$$r(t) = Hs(t) + \varepsilon(t) = H\sum_{\lambda\in\Lambda} a_\lambda g_\lambda(t) + \varepsilon(t) = \sum_{\lambda\in\Lambda} a_\lambda H g_\lambda(t) + \varepsilon(t) \tag{5.5}$$

with H denoting the channel matrix and ε the normal distributed error with expectation 0 and standard deviation σ. The symbols after demodulation are then given through

$$\tilde{a}_{\lambda'} = \Big\langle \sum_{\lambda\in\Lambda} a_\lambda H g_\lambda(t), \tilde{g}_{\lambda'} \Big\rangle \tag{5.6}$$

$$= \sum_{\lambda\in\Lambda} a_\lambda \big\langle H g_\lambda, \tilde{g}_{\lambda'} \big\rangle \tag{5.7}$$

with $\tilde{g}_{\lambda'}$ denoting the dual atom given by the pseudoinverse of the Gabor Matrix $\tilde{G} = G'(G'G)^{-1}$. In case we choose $g_1, ...g_7$ as Hermite functions $\tilde{G} = G'$. The central question is, whether the demodulated symbol $\tilde{a}_{\lambda'}$ is mapped correctly to the value of the alphabet which was sent.

The mapping is done due to proximity (the received value is mapped to the closest value of the alphabet). This means that as long as $\tilde{a}_{\lambda'}$ is in the correct quadrant we do observe correct reconstruction of both bits of information. In case $\tilde{a}_{\lambda'}$ falls in one of the neighboring quadrants, one errorneus bit out of two is received and in case it falls in the opposite quadrant all two bits of information are wrong. Figure (7) displays the result of the simulation of the possible outcomes of $\tilde{a}_\lambda$ after applying a time shift operator and demodulation with a Gaussian scheme to a single symbol a_λ. It can also be used as an illustration how the mapping is performed for a QPSK scheme.

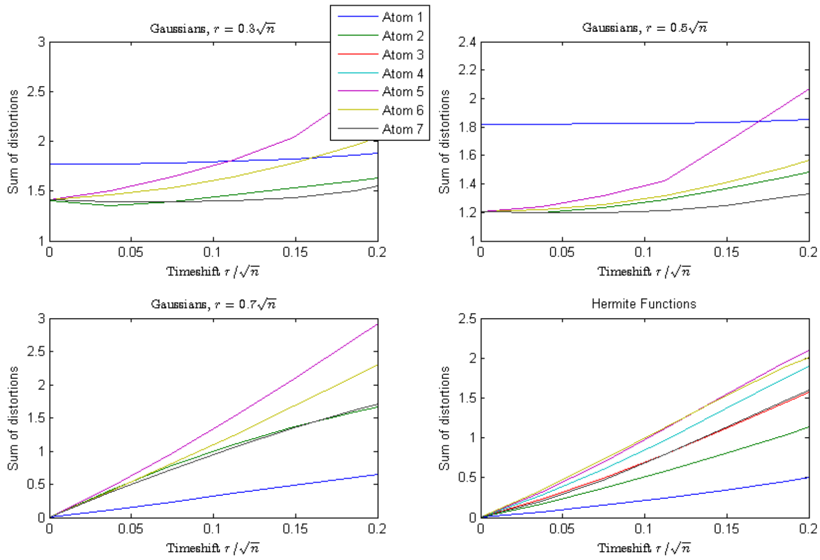

FIGURE 6. The figure shows the summed correlation effects $d_i(\tau) = |1 - c_i i(\tau)| + \sum_{i \neq j} |c_i j(\tau)|$ as measure for the distortion plotted in dependence of the timeshift τ normalized by the square root of n. It can be seen that for the Hermite basis the error increases fast for the outer Hermite functions whereas for the Gaussian setting through proper choice of parametrization the increase in the error for the single atoms is more moderate. The effect of G_1 being almost constant in the first 2 parametrizations is due to the threshold used in the calculation of the pseudoinverse.

For simplicity of calculations, we assume in the following H is a time shift operator T_τ. By time frequency rotation all results can be also translated one to one in the frequency domain due to rotation symmetry. Then $\langle H g_\lambda, \tilde{g}_{\lambda'} \rangle$ is just the correlation $(g_\lambda * \tilde{g}_{\lambda'})(\tau)$ and we calculate the autocorrelation as well as the cross correlation

$$c_{ij}(\tau) = \langle g_i, T_\tau \tilde{g}_j \rangle \tag{5.8}$$

$$c_{ii}(\tau) = \langle g_i, T_\tau \tilde{g}_i \rangle. \tag{5.9}$$

For numerical purposes, the dual basis $\tilde{G} = (\tilde{g}_i)_{i \in I}$ has always been calculated as the pseudoinverse with threshold ε. This means that singular values smaller than ε are ignored when calculating the pseudoinverse in order to avoid numerical instabilities. Next, for each symbol $(a_i)_{i \in I}$ of the transmission basis the value

$$d_i(\tau) = |1 - c_i i(\tau)| + \sum_{i \neq j} |c_i j(\tau)|$$

giving the sum of the interference effects between the atoms of one data package was calculated. Figure(6) shows an overview over the effects for the Gaussian settings with different parameters opposed to the Hermite basis.

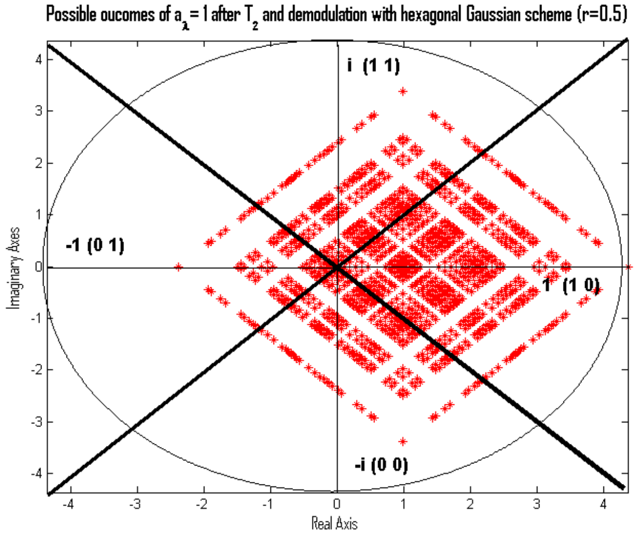

FIGURE 7. The figure shows the result of a full simulation of the possible outcomes of a transmitted symbol $a_1 = 1$ after modulation application of a time shift and demodulation with a hexagonal Gaussian scheme with radius $r = 0.5\sqrt{n}$. The dimension n was set to 720 in this experiment. The symbol was assumed to be located within a system of 7 data symbols, the others being i.i.d. with uniform distribution. It can also be seen how the mapping of the outcomes to the decoded values is performed according to the quadrant in which they fall. Finally the BER is calculated as correct outcomes / possible outcomes.

For the first simulation a data package of 7 symbols $(a_i)_{i\in I}$ with index set $I = \{1, ..., 7\}$ has been considered. The BER was then calculated by simulating independently identically distributed (i.i.d.) random symbols with a uniform distribution and dividing correct outcomes / possible outcomes for each symbol. Next, the BER of the overall data package consisting of 7 symbols (corresponding to 14 bit of information in total) has been calculated as arithmetic average of the BER of the individual symbols and plotted in dependence of the size of the time shift τ. Figure (8) shows the results. It can be seen that for the Gaussian setting there is more flexibility in the sense that the stability is highly dependent on the radius. Whereas for the Hermite functions we observe a sharp increase in the BER starting from a timeshft of $0.1\sqrt{n}$, for the Gaussian setting starting from a radius of $0.5\sqrt{n}$ the increase is already slightly more moderate. The Gaussian setting with a radius of $0.7\sqrt{n}$ gives the best stability results when it comes to distortions.

The second experiment now extends the simulation to a more realistic setting by increasing the dimension of the transmission basis and adding also pilotons to the experiment design. Therefore a $(\Pi/4)$-PSK scheme was considered. This is equivalent to a QPSK scheme, rotated by $(\Pi/4)$. It has the the alphabet $\{(1+i), (1-i), (-1-i), (-1+i)\}$ and both BER and SINR are the same as for a usual QPSK scheme. The received symbols are assumed to be i.i.d with a uniform distribution within the circle $||\Delta a|| \leq \varepsilon$

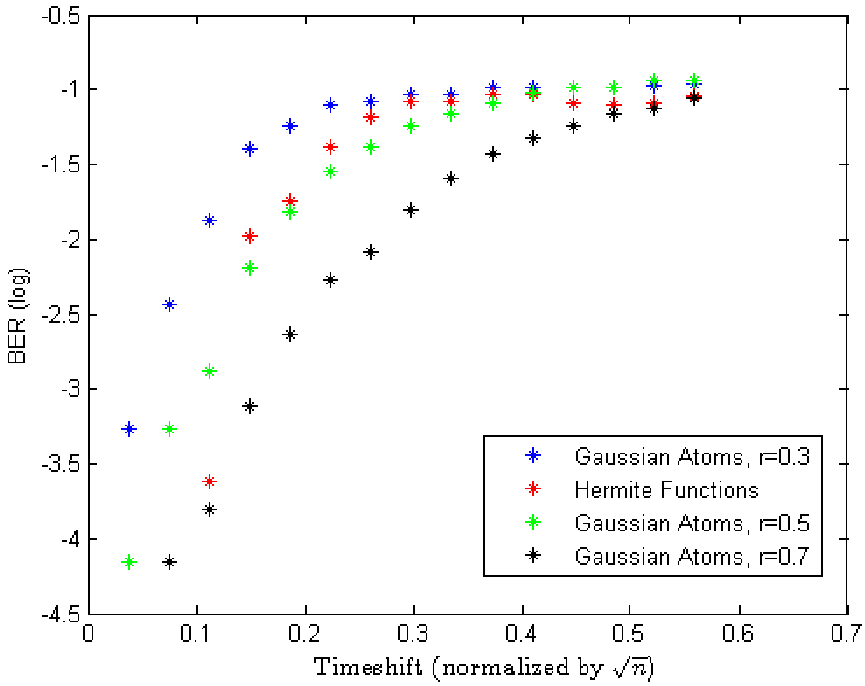

FIGURE 8. The figure shows the bit error rates for different transmission schemes of the data packages in dependence of the timeshift of the channel operator. The dimension for this experiment was set to 720 but all results have been normalized by $\sqrt{n}$. It can be seen that for the Hermite functions we observe a fast increase in the bit error rate as soon as the channel becomes distorted. This is also the case for the Gaussian settings with small radius. In this case however there is the option to increase the radius in order to get more stability in the transmission. Therefore the setting seems to be superior for distorted channels when the radius is adapted accordingly.

.

and the probability of a bit error is determined by devision of the size of the region of positive outcomes by the size of the region of possible outcomes. Therefore the probability that the received symbol $\tilde{a}$ falls in the neighboring quadrants of the correct one can be calculated through the common formula for the area of a segment of a circle Q and is given through

$$Q = ||\Delta a||^2 \arccos\left(\frac{1}{||\Delta a||}\right) - \sqrt{||\Delta a||^2 - 1} \tag{5.10}$$

Figure(9) shows the dependence of the probability of one bit error in identifying the symbol a_i on the size of the distortion of the channel $||\Delta a_i|| := ||Ha_i - a_i||$. It can be seen that the dependence is not linear and has a fast increase as soon as $||\Delta a_i||$ exceeds the lower bound of 1 for the standard alphabet. Therefore the overall transmission quality is not only dependent on the common upper bound of the error

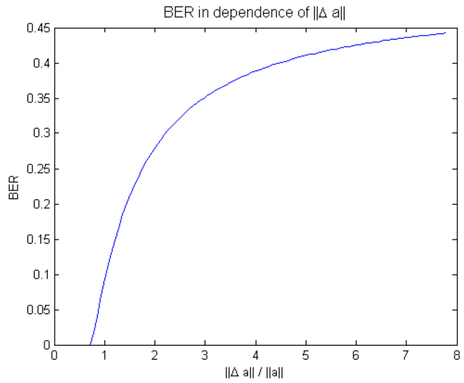

FIGURE 9. The figure shows the dependence of the probability of one bit error in identifying the symbol a_i on the size of the distortion of the channel $||\Delta a_i|| := ||Ha_i - a_i||$..

for all symbols but also on the interplay between the different symbols. Finally the expectation of the overall bit error was calculated by the arithmetic average over all the bit error porbabilities of the atoms of the transmission schemes. The results are stated in Figure (10). Therefore a transmission through a time shifted channel with increasing time shift τ - once using the Hermite basis combined with a rectangular pilote lattice and once using the suggested Gabor scheme with the same pilote lattice - was considered. In both cases, the dimension of the signal space was set to $n = 324$. Then 36 datapackages each carrying 7 datasymbols with 2 bits of information were transmitted. In addition the lattice contains 36 pilotones. The spectral efficiency of the system was therefor calculated as $0.778 = (36 * 7)/324$. The radius of the polygons was set to $0.38 * a_P = 20.52$. From the graph it can be seen that the Gabor transmission scheme shows a higher tolerance to distortions and only reacts later to timeshifts.

Overall, the simulations indicate that whereas for the overall Hermite basis conditioning is very good, in particular the higher order Hermite functions are very susceptible to distortions. This disadvantage can be coped by the suggested basis of Gaussians showing a more moderate reaction to distortions for all data carrying atoms.

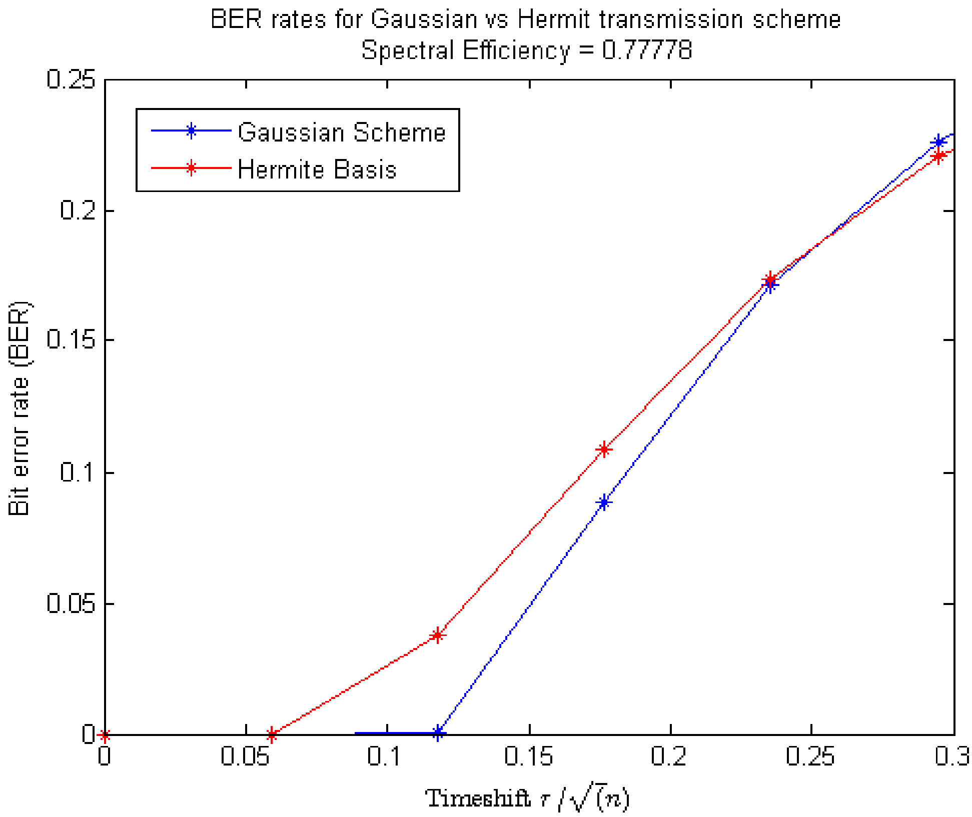

FIGURE 10. The figure shows the calculated bit error rate for a transmission through a time shifted channel once using the Hermite basis combined with a rectangular pilote lattice and once using the suggested Gabor scheme with the same pilote lattice. In both cases, the dimension of the signal space was set to $n = 324$. Then 36 datapackages each carrying 7 datasymbols with 2 bits of information were transmitted. In addition the lattice contains 36 pilotones. The spectral efficiency of the system was therefor calculated as $0.778 = (36 * 7)/324$. The radius of the polygons was set to $0.38 * a_P = 20.52$. From the graph it can be seen that the Gabor transmission scheme shows a higher tolerance to distortions and only reacts later to timeshifts.

CHAPTER 6

Appendix A

6.1. Final remark on the choice of dimension

At this point we would like to come back shortly on the question raised in the introduction concerning the applicability of these discrete results of - from a practitioner' s point of view - rather low dimensions to higher dimensions or the continuous setting. As a first point to mention we will always give results in relation to the dimension n. This means we will relate the domain of the spreading function to a certain percentage of n e.g. $(\tau_0 \nu_0) = 0.01n$, which would relate to a size of the spreading domain 0.01 in the continuous setting. We will now investigate the Cauchy property of the coefficients $H_{k,l}$ of the channel representation (4.7). Therefore, let $\eta(x, y) \in \mathbb{R}^2$ be the spreading symbol of the continuous channel. Then the coefficients of the discrete representation with dimension N obtained through sampling of the spreading function are given by

$$H^N_{k,l} = \frac{1}{N} \sum_{u=-n}^{n} \sum_{v=-n}^{n} \eta(\frac{u}{\sqrt{N}}, \frac{v}{\sqrt{N}}) \langle g_k, \pi(\frac{u}{\sqrt{N}}, \frac{v}{\sqrt{N}}) \tilde{g}^d_l \rangle \tag{6.1}$$

with $n :=$ ceil $(\frac{\sqrt{N}}{2})$ for easier notation. We only have to consider the sum from $-n$ to n as we assume the spreading function to be essentially underspread. Let, without restriction of generality, be $M \geq N$ and denote the according set of sampling points of the representations with $S^M = (s^M_i)_{i=1,\ldots,M^2}$ and $S^N = (s^N_j)_{j=1,\ldots,N^2}$. Then for every sampling point $s_i \in S^M$ we choose the sampling point $s_j \in S^N$ with the minimal distance meaning

$$\Delta^2_i := |s_i - s^*_j|^2 = \min_{j \in S^N} |s_i - s_j|^2 \tag{6.2}$$

and denote with P^j the set of points $s_i \in S^M$ assigned to the sampling point $s_j \in S^N$. Further we calculate

$$\begin{aligned} |H^N_{k,l} - H^M_{k,l}| &\leq \frac{1}{M} |\sum_{j=1}^{N} \sum_{i \in P^j} \eta(s_j) \langle g_k, \pi(s_j) g^d_l \rangle - \eta(s_j + \Delta_i) \langle g_k, \pi(s_j + \Delta_i) g^d_l \rangle| \\ &\leq \frac{1}{M} \sum_{j=1}^{N} \sum_{i \in P^j} |\eta(s_j) \langle g_k, \pi(s_j) g^d_l \rangle - \eta(s_j + \Delta_i) \langle g_k, \pi(s_j + \Delta_i) g^d_l \rangle| \end{aligned}$$

$$\leq \frac{1}{M}\sum_{j=1}^{N}\sum_{i\in P^j} |\,(\eta(s_j)-\eta(s_j+\Delta_i))\,\langle g_k,\pi(s_j)g_l^d\rangle + \eta(s_j+\Delta_i)\left(\langle g_k,(\pi(s_j)-\pi(s_j+\Delta_i))g_l^d\rangle|\right)$$

$$\leq \frac{1}{M}\sum_{j=1}^{N}\sum_{i\in P^j} |\left(\frac{d\eta(s_i)}{d(x,y)}\langle g_k,\pi(s_j)g_l^d\rangle + \frac{dV_{g_k}g_l^d}{d(x,y)}\eta(s_j+\Delta_i)\right)\Delta_i|$$

If we now assume a regular sampling, we have $\Delta_i \leq \frac{1}{2N}$. Furthermore in our experiments we assume $V_{g_k}g_l^d$ and $\eta(x,y)$ to have exponential decay with the parameters $a, b, t_1, t_2 > 0$ and get therefore

$$\begin{aligned}
|H_{k,l}^N - H_{k,l}^M| &\leq \frac{1}{M}\frac{1}{2N}\sum_{j=1}^{N}\sum_{i\in P^j} |\left(\frac{d\eta(s_j)}{d(x,y)}ae^{-t_1 s_j} + \frac{dV_{g_k}g_l^d}{d(x,y)}be^{-t_2 s_j}|\right) \\
&\leq \frac{1}{M}\frac{1}{2N}\sum_{j=1}^{N}\sum_{i\in P^j} |abt_2 e^{-t_2 s_j}e^{-t_1 s_j} + abt_1 e^{-t_2 s_j}e^{-t_1 s_j}|
\end{aligned}$$

Set t as being the minimum of the decay parameters $t := \min(t_1, t_2)$ and $p_j := \#(P^j)$ and we get

$$\begin{aligned}
|H_{k,l}^N - H_{k,l}^M| &\leq \frac{1}{M}\frac{1}{2N}\sum_{j=1}^{N} p^j |2abte^{-2ts_j}| \\
&\leq \frac{1}{N}\frac{1}{2N}\sum_{j=1}^{N} |2abte^{-2ts_j}| \\
&\leq \frac{1}{N}abt\underbrace{\sum_{j=1}^{N}\frac{1}{N}|e^{-2ts_j}|}_{\leq 1} \\
&\leq \frac{abt}{2N} =: \varepsilon(N)
\end{aligned}$$

As we consider the bit error rate as result, this can only influence the outcome in case $|H_{k,l}c_{k,l}| \leq \varepsilon(N)|$ or $|H_{k,l}c_{k,l} - i| \leq \varepsilon(N)$. We will account for that effect, by giving the interval for the bit error rate if necessary.

CHAPTER 7

Appendix B

List of MATLAB files attached:

```
adaptN.m
addPilotone.m
alldiv.m
BERfinal.m
biorth_new.m
Butterworth.fig
calcBER.m
CompSpaces.m
Comp_Projections_Herm_Gauss.m
CondOptRadAng.m
createGaborBasis.m
createHermBasis.m
CreateLattice.m
createRieszBasis.m
createSystem.m
createSystemFast.m
EffectOfWindow.m
Error.m
Exp1.m
Filter_170909.m
Filter_250809.m
gabddd.m
gaussnk.m
getCorrelations_new.m
getDataLattice.m
getLatticeHerm.m
getLatticeHex.m
getPolygon.m
gmappNE.m
hermf.m
IncreaseSpecEff.zip
MonteCarlo.m
MoreExamples.m
OptConditionNumber.m
```

```
Projections.m
Projections_Skript.m
rotmodfr.m
ScriptDiffDesigns.m
SimpleApproachError.m
Skript_BER.m
spr2mat.m
stftsubspace.m
```

Bibliography

[1] IEEE Standard 802.11a-1999(2003): Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, 2003.

[2] IEEE Standard 802.11g-2003: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, 2003.

[3] IEEE Standard 802.16-2009: Air Interface for Broadband Wireless Access Systems, 2009.

[4] T. Alieva, M. Bastiaans, and L. Stankovic. Time-frequency signal analysis based on the windowed fractional Fourier transform. *Signal Processing*, 83(11):2459 – 2468, 2003.

[5] P. Antsaklis and A. Michel. *Linear Systems.* Birkhäuser Boston Inc., 2006.

[6] G. Ascensi, H. G. Feichtinger, and N. Kaiblinger. Dilation of the Weyl symbol and Balian-Low theorem. *Trans. Amer. Math. Soc.*, pages 01–15, 2013.

[7] A. R. I. Austrian Academy of Sciences. Stx zum download. website. `http://www.kfs.oeaw.ac.at/index.php?option=com_content&view=article&id=276&Itemid=415&lang=de`;Accessed: 2014-03-14.

[8] P. Balazs. *Regular and Irregular Gabor Multiplier With Application To Psychoacoustic Masking.* PhD thesis, University of Vienna, 2005.

[9] P. Balazs. Frames and finite dimensionality: Frame transformation, classification and algorithms. 2006.

[10] P. Balazs. Basic definition and properties of Bessel multipliers. *J. Math. Anal. Appl.*, 325(1):571–585, 2007.

[11] P. Balazs. Hilbert-Schmidt operators and frames - classification, best approximation by multipliers and algorithms. *Int. J. Wavelets Multiresolut. Inf. Process.*, 6(2):315 – 330, March 2008.

[12] P. Balazs. Matrix representation of bounded linear operators by Bessel sequences, frames and Riesz sequences. In *SAMPTA'09, Marseille, May 18-22.* ARI;Frames;NHG-coop, 2009.

[13] P. Balazs, D. Bayer, and A. Rahimi. Multipliers for continuous frames in Hilbert spaces. *J. Phys. A*, Special issue: Coherent states(45):244023, 2012.

[14] S. Bannert. Banach-Gelfand Triples and Applications in Time-Frequency Analysis. Master's thesis, University of Vienna, 2010.

[15] J. J. Benedetto and G. E. Pfander. Frame expansions for Gabor multipliers. *Appl. Comput. Harmon. Anal.*, 20(1):26–40, 2006.

[16] P. Boggiatto and J. Toft. Embeddings and compactness for generalized Sobolev-Shubin spaces and modulation spaces. *Appl. Anal.*, 84(3):269–282, 2005.

[17] H. Bölcskei, F. Hlawatsch, and H. G. Feichtinger. Frame-theoretic analysis of oversampled filter banks. *IEEE Trans. Signal Process.*, 46(12):3256–3268, 1998.

[18] D. Borah and B. Hart. Frequency-selective fading channel estimation with a polynomial time-varying channel model. *IEEE Trans. Comm.*, 47:862–873, Jun. 1999.

[19] M. Bownik and O. Christensen. Characterization and perturbation of Gabor frame sequences with rational parameters. *J. Approx. Theory*, 147(1):67 – 80, 2007.

[20] A. Bultheel and H. Martinez. A shattered survey of the Fractional Fourier Transform. 2003.

[21] S. Butterworth. On the Theory of Filter Amplifiers. *Wireless Engineer*, 7:536–541, 1930.

[22] a. Candan, M. A. Kutay, and H. M. Ozaktas. The discrete fractional Fourier transform. *IEEE Trans. Signal Process.*, 48(5):1329–1337, 2000.

[23] P. G. Casazza, D. Han, and D. R. Larson. Frames for Banach spaces. In L. Baggett and D. R. Larson, editors, *The Functional and Harmonic Analysis of Wavelets and Frames (San Antonio, TX, 1999).*, volume 247 of *Contemporary Mathematics*, pages 149–182, Providence, RI, 1999. American Mathematical Society (AMS).

[24] R. Chang. Synthesis of band-limited orthogonal signals for multichannel data transmission. *Bell System Tech. J.*, 45:1775–1796, Dec. 1966.

[25] O. Christensen. Frame perturbations. *Proc. Amer. Math. Soc.*, 123:1217–1220, 1995.

[26] O. Christensen. Frames, Riesz bases, and discrete Gabor/wavelet expansions. *Bull. Amer. Math. Soc., New Ser.*, 38(3):273–291 (electronic), 2001.

[27] O. Christensen. *An Introduction to Frames and Riesz Bases.* Applied and Numerical Harmonic Analysis. Birkhäuser, Boston, 2003.

[28] O. Christensen. *Functions, spaces, and expansions. Mathematical tools in physics and engineering.* Birkhäuser, 2010.

[29] L. Cimini. Analysis and simulation of a digital mobile channel using orthogonal frequency division multiplexing. *IEEE Transactions on Communications*, 33(7):665–675, 1985.

[30] L. Condat, T. Blu, and M. Unser. Beyond Interpolation: Optimal Reconstruction by Quasi-Interpolation. pages 33–36, 2005.

[31] J. Conway and N. J. A. Sloane. *Sphere Packings, Lattices and Groups.* Volume 290. Third Edition edition, 1999.

[32] J. B. Conway. *A Course in Operator Theory.* American Mathematical Society (AMS), RI, 2000.

[33] E. Cordero and K. Gröchenig. Necessary conditions for Schatten class localization operators. *Proc. Amer. Math. Soc.*, 133(12):3573–3579, 2005.

[34] H. G. Dales, P. Aiena, J. Eschmeier, K. Laursen, and G. A. Willis. *Introduction to Banach Algebras, Operators, and Harmonic Analysis.* Cambridge University Press, Cambridge, 2003.

[35] S. Das. *Mathematical methods for wireless channel estimation and equalization.* PhD thesis, September, 2009.

[36] I. Daubechies. *Ten lectures on wavelets.*, volume 61 of *CBMS-NSF Regional Conference Series in Applied Mathematics.* SIAM, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1992.

[37] C. de Boor and R. A. DeVore. Partitions of unity and approximation. *Proc. Amer. Math. Soc.*, 93(4):705–709, 1985.

[38] M. De Gosson. *Symplectic Methods in Harmonic Analysis and in Mathematical Physics*, volume 7 of *Pseudo-Differential Operators. Theory and Applications.* Birkhäuser/Springer Basel AG, Basel, 2011.

[39] F. Demengel and G. Demengel. Sobolev spaces and embedding theorems. In *Functional Spaces for the Theory of Elliptic Partial Differential Equations*, pages 57–112. Springer, 2012.

[40] M. Dörfler. *Gabor Analysis for a Class of Signals called Music.* PhD thesis, 2002.

[41] M. Dörfler. What time-frequency analysis can do to Music Signals. In M. Emmer, editor, *Matematica e Cultura 2003.* 2003.

[42] M. Dörfler and B. Torrésani. On the time-frequency representation of operators and generalized Gabor multiplier approximations. *preprint*, submitted to Elsevier, 2007.

[43] M. Dörfler and B. Torrésani. Spreading function representation of operators and Gabor multiplier approximation. In *Proceedings of SAMPTA0́7*, Thessaloniki, June 2007.

[44] M. Dörfler and B. Torrésani. Representation of operators by sampling in the time-frequency domain. In *Proceedings of SAMPTA09.* NHG-coop;NuHAG, 2009.

[45] M. Dörfler and B. Torrésani. Representation of operators in the time-frequency domain and generalized Gabor multipliers. *J. Fourier Anal. Appl.*, 16(2):261–293, 2010.

[46] R. J. Duffin and A. C. Schaeffer. A class of nonharmonic Fourier series. *Trans. Amer. Math. Soc.*, 72:341–366, 1952.

[47] D. Eiwen. *Compressive Channel Estimation - Compressed Sensing Methods for Estimating Doubly Selective Channels in Multicarrier Systems.* PhD thesis, 2012.

[48] H. G. Feichtinger. An elementary approach to the generalized Fourier transform. In T. Rassias, editor, *Topics in Mathematical Analysis*, pages 246–272. World Sci.Pub., 1989.

[49] H. G. Feichtinger. New results on regular and irregular sampling based on Wiener amalgams. In K. Jarosz, editor, *Function Spaces, Proc Conf, Edwardsville/IL (USA) 1990*, volume 136 of *Lect. Notes Pure Appl. Math.*, pages 107–121, New York, 1992. Marcel Dekker.

[50] H. G. Feichtinger. Wiener amalgams over Euclidean spaces and some of their applications. 136:123–137, 1992.

[51] H. G. Feichtinger. Spline-type spaces in Gabor analysis. In D. X. Zhou, editor, *Wavelet Analysis: Twenty Years Developments Proceedings of the International Conference of Computational Harmonic Analysis, Hong Kong, China, June 4–8, 2001*, volume 1 of *Ser. Anal.*, pages 100–122. World Sci.Pub., 2002.

[52] H. G. Feichtinger. A Fresh Approach to Harmonic Analysis: NuHAG: Version Summer 2010. *Lecture Notes*, page 68, 2010.

[53] H. G. Feichtinger, A. Grybos, and D. Onchis. Approximate dual Gabor atoms via the adjoint lattice method. *Adv. Comput. Math.*, 40(3):651–665, 2014.

[54] H. G. Feichtinger and N. Kaiblinger. Quasi-interpolation in the Fourier algebra. *J. Approx. Theory*, 144(1):103–118, 2007.

[55] H. G. Feichtinger and W. Kozek. Quantization of TF lattice-invariant operators on elementary LCA groups. In H. G. Feichtinger and T. Strohmer, editors, *Gabor analysis and algorithms*, Appl. Numer. Harmon. Anal., pages 233–266. Birkhäuser Boston, Boston, MA, 1998.

[56] H. G. Feichtinger, W. Kozek, and F. Luef. Gabor Analysis over finite Abelian groups. *Appl. Comput. Harmon. Anal.*, 26(2):230–248, 2009.

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

[58] H. G. Feichtinger and D. Onchis. Constructive realization of dual systems for generators of multi-window spline-type spaces. *J. Comput. Appl. Math.*, 234(12):3467–3479, 2010.

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

[60] H. G. Feichtinger and T. Strohmer. *Advances in Gabor Analysis.* Birkhäuser, 2003.

[61] H. G. Feichtinger and G. Zimmermann. A Banach space of test functions for Gabor analysis. In H. G. Feichtinger and T. Strohmer, editors, *Gabor Analysis and Algorithms: Theory and Applications*, Applied and Numerical Harmonic Analysis, pages 123–170, Boston, MA, 1998. Birkhäuser Boston.

[62] F. G. Friedlander. *Introduction of the Theory of Distributions. With Additional Material by M. Joshi.* Cambridge University Press, Cambridge, 2nd ed. edition, 1998.

[63] G. Giannakis. Filterbanks for blind channel identification and equalization. *IEEE Signal Process. Letters*, pages 184–187, Jun. 1997.

[64] G. B. Giannakis and C. Tepedelenlioglu. Basis Expansion Models and Diversity Techniques for Blind Identification and Equalization of TimeVarying Channels. In *Proc. IEEE*, pages 1969–1986, 1998.

[65] K. Gröchenig. *Foundations of Time-Frequency Analysis.* Appl. Numer. Harmon. Anal. Birkhäuser Boston, 2001.

[66] K. Gröchenig. Composition and spectral invariance of pseudodifferential operators on modulation spaces. *J. Anal. Math.*, 98:65–82, 2006.

[67] K. Gröchenig. Weight functions in time-frequency analysis. In L. Rodino and et al., editors, *Pseudodifferential Operators: Partial Differential Equations and Time-Frequency Analysis*, volume 52 of *Fields Inst. Commun.*, pages 343–366. Amer. Math. Soc., Providence, RI, 2007.

[68] K. Gröchenig. Representation and approximation of pseudodifferential operators by sums of Gabor multipliers. *to appear in Appl. Anal.*, page 16, 2009.

[69] K. Gröchenig. Representation and approximation of pseudodifferential operators by sums of Gabor multipliers. *Appl. Anal.*, 90(3-4):385–401, 2011.

[70] M. M. Hartmann, G. Matz, and D. Schafhuber. Wireless Multicarrier Communications via Multipulse Gabor Riesz Bases. *EURASIP J. Appl. Signal Process.*, 2006:Article ID 23818, 15 pages, 2006.

[71] C. Heil. An introduction to weighted Wiener amalgams. In M. Krishna, R. Radha, and S. Thangavelu, editors, *Wavelets and their Applications (Chennai, January 2002)*, pages 183–216. Allied Publishers, 2003.

[72] C. Heil and D. F. Walnut. Continuous and discrete wavelet transforms. *SIAM Rev.*, 31:628–666, 1989.

[73] O. Holtz and A. Ron. Approximation orders of shift-invariant subspaces of $w_2^s(bbbr^d)$. *J. Approximation Theory*, 132(1):97–148, 2005.

[74] C. Houdré. Linear Fourier and stochastic analysis. *Probab. Theory Relat. Fields*, 87(2):167–188, 1990.

[75] T. Hrycak, S. Das, G. Matz, and H. G. Feichtinger. Low complexity equalization for doubly selective channels modeled by a basis expansion. *IEEE Trans. Signal Process.*, 58(11):5706–5719, Nov. 2010.

[76] S. Huestis. Optimum kernels for oversampled signals. *The Journal of the Acoustical Society of America*, 92:1172, 1992.

[77] A. J. E. M. Janssen. From continuous to discrete Weyl-Heisenberg frames through sampling. *J. Fourier Anal. Appl.*, 3(5):583–596, 1997.

[78] N. Kaiblinger. Approximation of the Fourier transform and the dual Gabor window. *J. Fourier Anal. Appl.*, 11(1):25–42, 2005.

[79] A. Kannu and P. Schniter. MSE-optimal training for linear time-varying channels. volume 3, pages 789–792, Mar. 2005.

[80] A. Kannu and P. Schniter. Design and analysis of MMSE pilot-aided cyclic-prefixed block trans mission for doubly selective channels. *IEEE Trans. Signal Process.*, 56:1148–1160, Mar. 2008.

[81] Y. Katznelson. *An Introduction to Harmonic Analysis*. Cambridge University Press, 2003.

[82] W. Kozek. On the generalized Weyl correspondence and its application to time-frequency analysis of linear time-varying systems. In *IEEE Int. Symp. on Time–Frequency and Time–Scale Analysis*, pages 167–170, Victoria, Canada, October 1992.

[83] W. Kozek. On the underspread/overspread classification of nonstationary random processes. In K. Kirchgaessner, . Mahrenholtz, and R. Mennicken, editors, *Proc. ICIAM 95, Hamburg, July 1995*, volume 3 of *Mathematical Research*, pages 63–66, Berlin, 1996. Akademieverlag.

[84] W. Kozek and A. Molisch. Nonorthogonal pulseshapes for multicarrier communications in doubly dispersive channels. *Selected Areas in Communications, IEEE Journal on*, 16(8):1579–1589, Oct 1998.

[85] W. Kozek, A. F. Molisch, and E. Bonek. Pulse design for robust multicarrier transmission over doubly-dispersive channels. In *Proc. International Conference on Telecommunications (ICT'98)*, pages 313–317, 1998.

[86] W. Kozek and G. E. Pfander. Identification of operators with bandlimited symbols. *SIAM J. Math. Anal.*, 37(3):867–888, 2006.

[87] M. A. Kutay, H. M. Ozaktas, and Z. Zalevsky. *The Fractional Fourier Transform, with Applications in Optics and Signal Processing*. John Wiley and Sons, 2001.

[88] G. Leus. On the estimation of rapidly varying channels. volume 4, pages 2227–2230, Sep. 2004.

[89] Y. Li, L. Cimini, and N. Sollenberger. Robust channel estimation for OFDM systems with rapid dispersive fading channels. *IEEE Trans. Comm.*, 46:902–915, Jul. 1998.

[90] J. Lim and H. Kim. New characterizations of Riesz bases. *Appl. Comput. Harmon. Anal.*, 4:222–229, 1997.

[91] Y. Liu and M. Wong. Polar wavelet transforms and localization operators. *Integr. Equ. Oper. Theory*, 58(1):99–110, 2007.

[92] D. Marelli, M. Fu, P. Balazs, and P. Majdak. An Iterative Method for Approximating LTI Systems using Subbands. In *2008 IEEE International Conference on Acoustics, Speech, and Signal Processing*. ARI;Acoustics;NuHAG;NHG-coop, 2008.

[93] E. Matusiak. *Some aspects of Gabor analysis on elementary locally compact Abelian groups*. PhD thesis, 2007.

[94] G. Matz and F. Hlawatsch. Time-frequency transfer function calculus (symbolic calculus) of linear time-varying systems (linear operators) based on a generalized underspread theory. *J. Math. Phys.*, 39(8):4041–4070, 1998.

[95] G. Matz, D. Schafhuber, K. Gröchenig, M. Hartmann, and F. Hlawatsch. Analysis, optimization, and implementation of low-interference wireless multicarrier systems. *IEEE Trans. Wireless Comm.*, 6(4):1–11, 2007.

[96] A. Missbauer. Gabor Frames and the Fractional Fourier Transform. Master's thesis, 2012.

[97] A. Peled and A. Ruiz. Frequency domain data transmission using reduced computational complexity algorithms. volume 5, pages 964–967, April 1980.

[98] J. Ramanathan and P. Topiwala. Time-frequency localization via the Weyl correspondence. *SIAM J. Math. Anal.*, 24(5):1378–1393, 1993.

[99] A. Ron and Z. Shen. Weyl-Heisenberg frames and Riesz bases in $L_2(\mathbb{R}^d)$. *Duke Math. J.*, 89(2):237–282, 1997.

[100] W. Rudin. *Functional Analysis*. McGraw-Hill Book Company, New York, 1973.

[101] A. Scaglione, G. Giannakis, and S. Barbarossa. Redundant filterbank precoders and equalizers, Parts I and II. *IEEE Trans. Signal Process.*, pages 1988006, and 2007022, Jul. 1999.

[102] R. L. Schilling. *Measures, Integrals and Martingales*. Cambridge University Press, 2005.

[103] K. Schnass. Gabor Multipliers. A self-contained survey. Master's thesis, University of Vienna, 2004.

[104] L. Schwartz. Th'eorie des noyaux. In *Proceedings of the International Congress of Mathematicians*, volume 1, pages 220–230, 1950.

[105] Y. Shen and E. Martinez. Channel estimation in OFDM systems, Feb. 2006.

[106] P. Smulders. Exploiting the 60. *IEEE Communications Magazine GH band for local wireless multimedia access: prospects and future directions*, 40(1):140–147, 2002.

[107] P. L. Sondergaard. Gabor frames by sampling and periodization. *Adv. Comput. Math.*, 2005.

[108] P. L. Sondergaard. Gabor frames by sampling and periodization. *Adv. Comput. Math.*, 27(4):355–373, 2007.

[109] E. M. Stein. *Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals*. Princeton University Press, 1993.

[110] D. T. Stoeva and P. Balazs. Riesz bases multipliers. In *Concrete Operators, Spectral Theory, Operators in Harmonic Analysis and Approximation*, volume 236, pages 475–482. Birkhuser, Heidelberg, New York, Dordrecht, 2014.

[111] T. Strohmer. Numerical algorithms for discrete Gabor expansions. In H. G. Feichtinger and T. Strohmer, editors, *Gabor Analysis and Algorithms: Theory and Applications*, pages 267–294. Birkhäuser Boston, Boston, 1998.

[112] T. Strohmer. Pseudodifferential operators and Banach algebras in mobile communications. *Appl. Comput. Harmon. Anal.*, 20(2):237–249, March 2006.

[113] T. Strohmer and S. Beaver. Optimal OFDM system design for time-frequency dispersive channels. *IEEE Trans. Comm.*, 51(7):1111–1122, July 2003.

[114] T. Strohmer and J. Tanner. Implementations of Shannon's sampling theorem, a time-frequency approach. *Sampl. Theory Signal Image Process.*, 4(1):1–17, 2005.

[115] G. Tauböck, M. Hampejs, G. Matz, F. Hlawatsch, and K. Gröchenig. LSQR-based ICI equalization for multicarrier communications in strongly dispersive and highly mobile environments. In *Proc. IEEE SPAWC 07: Signal Processing Advances in Wireless Communications, 8th Workshop on*, pages 1 – 5. MOHAWI;NuHAG, 2007.

[116] G. Tauböck and F. Hlawatsch. A compressed sensing technique for OFDM channel estimation in mobile environments: Exploiting channel sparsity for reducing pilots. *IEEE Int. Conf. on*

Acoustics, Speech, and Signal Processing (ICASSP), Las Vegas, Nevada, April 2008, pages 2885–2888, 2008.

[117] G. Tauböck, F. Hlawatsch, D. Eiwen, and H. Rauhut. Compressive estimation of doubly selective channels in multicarrier systems: leakage effects and sparsity-enhancing processing. *IEEE J. Sel. Topics Sig. Process.*, 4(2):255–271, 2010.

[118] S. B. Weinstein and P. M. Ebert. Data transmission by frequency division multiplexing using the discrete Fourier transform. *IEEE Transactions on Communications*, 19(5):628–634, 1971.

[119] F. Weisz. Multiplier theorems for the short-time Fourier transform. *Integr. Equ. Oper. Theory*, 60:133–149, 2008.

[120] D. Werner. *Funktionalanalysis. (Functional analysis) 2., überarb. Aufl.* Springer-Verlag, Berlin, 1997.

[121] N. Wiener. On the representation of functions by trigonometric integrals. *Math. Z.*, 24:575–616, 1926.

[122] N. Wiener. *The Fourier Integral and Certain of its Applications.* Dover Publications Inc, New York, 1958.

[123] N. Wiener. Tauberian theorems. *MIT Press*, pages 143–242, 1964.

[124] C. Wiesmeyr. *Construction of frames by discretization of phase space.* PhD thesis, 2014.

[125] R. M. Young. *An Introduction to Nonharmonic Fourier Series.* Academic Press, Orlando, FL, Revised 1st edition edition, 2001.

[126] T. Zemen and C. Mecklenbräuker. Time-variant channel estimation using discrete prolate spheroidal sequences. *IEEE Trans. Signal Process.*, 53:3597–3607, Sep. 2005.

Curriculum Vitæ

Curriculum Vitae

Dipl.-Ing.
Nina
Engelputzeder

Markgraf-Rüdiger-Str. 7/67
1150 Wien
+43 (0) 650 255 73 39
nina.engelputzeder@gmail.com

Mathematician / Head of global COC Rates, FX and Equity

Personal

Dipl.-Ing. Nina Engelputzeder
07.10.1982 born in Linz (Austria)
1 daughter (Malina Victoria, born 12.11.2012)

Professional

07/2015-today — UniCredit S.p.a, Milan
Group Price Control

- Head of global Competence Center Rates & FX
- Head of global Competence Center Equity & Commodity

Independent Price Verification of above mentioned asset classes
Coordination and harmonization of Market Data provisioning to all Group Front office systems for End of Day valuation.
Coordination and harmonization of Market Data provisioning to all Group Risk Systems for calculation of risk metrics.

08/2014 – 06/2015 — UniCredit Bank Austria,
Department for Market Risk and Liquidity Risk – Austria and CEE

- Head of Section Price Control
    - Independent Price Verification of BA and CEE portfolio
    - Calculation of Valuation Adjustments according to IFRS
    - Market Risk Stress Test Calculation Reporting and integration in overall stress test.
    - Market Data provision to FO systems / deriving risk factors for risk measurement, Market Data Analysis & Reporting
    - Market Conformity and Pricing Transparency Analysis
    - Project enrolements on OIS discounting methodology, new IPV methodology for OTC transactions, introduction of new internal market risk model, negative interest rate models, fund transfer pricing, interest rate volatility models
- Head of competence center CEE Bonds
    - Independent price verification for bonds (Austria & CEE)
    - SPOC for bond valuation methodology (Austria & CEE)
    - Regular and ad hoc reporting on bond portfolio (Austria & CEE)

09/2009-07/2014 — UniCredit Bank Austria, Department for Market Risk & Risk Integration
(09/2009-10/2012 full time, 05/2013-07/2014 part time, 10/2012-04/2013 maternity leave)
Market Risk Manager (Setup of reporting, financial time series analysis, derivatives price validation, project management of CEE rollout of market data platform and market conformity platform as well as respective policies)

03/2010-08/2011 — Risk Management Trainee Program

11/2013-08/2014 — Austrian Academy of Science, Acoustics Research Institute
Part time research consultant on the European START project „FLAME“ on Gabor Multipliers / Finishing Phd thesis in mathematics.

10/2008 – 10/2009 — Austrian Academy of Science, Acoustics Research Institute

	Research assistant supporting the project ‚Gabor Multipliers in Acoustics' on mathematical models in audio research.
03/2007 – 09/2009	University of Vienna, Mathematics Institute Research assistant within the team for numeric harmonic analysis Funded participant of the initiativ college: ‚Modern mathematical analysis and applications'
08/2007-05/2010	Freelance projects as statistical consultant for Caritas Austria, Fessl Gfk, Statistics Austria
02/2006-03/2007	Statistics Austria, National statistics office, Department of Economics Project responsible for the EU project ‚Flash estimates for employment' Development of econometric models for GDP calculation and quarterly projections

Education

04/2015-(06/2016)	Executive Academy of the Vienna University for Business and Economics Global Executive Management MBA (Female Leadership Scholarship)
03/2007-(09/2015)	University of Vienna, Phd in Mathematics Thesis on ‚Time variant systems and Gabor Riesz Bases ' (currently under review)
10/2000-01/2006	Technical University of Vienna Economic Mathematics (Technical Mathematics) Specializaion in Statistics and Econometrics Thesis on ‚Econometric analysis of consumption in Austria'
01/2004-07/2004	Université de téchnologie deTroyes, Troyes, France ERASMUS semester
08/2007	ECMI Modeling Week, Rouen, France
07/2003	Summer University, Liège, Belgium ‚Economics and management introduced to engeneers'
09/1992-06/2000	Gymnasium BG 3 Ramsauerstraße in Linz

Continuing Training

04/2015-07/2015	Building Leadership Skills Bank Austria & Hype Vereinsbank joint leadership program
05/2014	Pioneer Investments, Portfolio Management Advanced
04/2014	UniManagement of UniCredit Group, Operational Risk in Organisations
02/2010	Euromoney Training, London School of Banking Risk Management
05/2011-05/2012	Boccorni University, Banking Risk Academy
10/2010-06/2011	Finance Trainer, Corporate Treasury Certificate
03/2010-08/2011	UniCredit Bank Austria, Risk Management Trainee program
10/2008-06/2010	NLP Akademie, Certified NLP trainer education
05/2009 – 10/2012	WiLAK – Wiener Lebensakademie, Certified Systemic Coaching education

07/2008 – 07/2009	Bildungsforum - Society for continuing training Trainings on presentation skills, leadership and time management
04-08/2009	HR development of the University of Vienna Training on negotiation skills, training on business communication
03/2006-12/2006	SAS Institute Software GmbH Trainings: SAS basic, SAS Programming, Time Series analysis with SAS

Languages

German:	Mother tongue
English:	Fluent
French:	Very good

Compuer Skills

Operating Systems: Linux, MS Windows
Formating Languages: Html, LaTEX
Prgramming languages: C++, PHP, Fortran, VB.net
Statistical / Mathematics Software: R, Eviews, SPSS, SAS, Matlab, Maple
IT systems: SAP basics
Database Languages: MySQL, PLSQL, Access, SAS/SQL

Further skills & interest

Arts:	One year of professional dance education - Studio Halfstreet 7
Driving licence:	B
Hobbys:	Social engagement at Caritas Austria, Sports (Running, Modern Dance, Yoga, Mountain sports), Travelling

Publications

Florent Jaillet, Peter Balazs, Monika Dörfler and Nina Engelputzeder. Nonstationary Gabor Frames In SAMPTA'09, Marseille, May 18-22", 2009.

Florent Jaillet, Peter Balazs, Monika Dörfler and Nina Engelputzeder. On the structure of the phase around the Zeros of the STFT In Proceedings of DAGA09, Rotterdam, NL, 2009

Nina Engelputzeder. Flash Estimates Employment Methodology. EUROSTAT project report, Statistics Austria, 2006

Nina Engelputzeder. Eine ökonometrische Analyse des Konsums der Österreicher, Master thesis, Technical University of Vienna, 2005