Constructive realization of dual systems for generators of multi-window spline-type spaces

Hans G. Feichtinger and Darian M. Onchis

March 8, 2010

Abstract

Multi-window spline-type spaces arise naturally in many areas. Among others they have been used as model spaces in the theory of irregular sampling. This class of shift-invariant spaces is characterized by possessing a Riesz basis which consists of a set of translates along some lattice Λ of a finite family of atoms. Part of their usefulness relies on the explicit knowledge of the structure of the projection operator on such a space using the existence of a finite family of dual atoms. The main goal of this paper is to address the problems arising from the discrepancy between a constructive description and an implementable approximate realization of such concepts. Using function space concepts (e.g. Wiener amalgam spaces) we describe how approximate dual atoms can be computed for any given degree of precision.

As an application of our result we describe the best approximation of Hilbert-Schmidt operators by generalized Gabor multipliers, using smooth analysis and synthesis windows. The Kohn-Nirenberg symbols of the rank-one operators formed from analysis and synthesis windows satisfy our general assumptions. Applications to irregular sampling are given elsewhere.

1 Introduction

We are considering in this paper multi-window spline-type spaces. Elsewhere these spaces are called (finitely generated) shift-invariant spaces (e.g. [1, 11, 13, 14, 33, 41]), but we prefer to the terminology used in the context of sampling theory and sometimes wavelet theory. It also reminds of the fact that (e.g. cubic) splines are a prototypical example of the situation, where a set of functions and their translates along some lattice in \( \mathbb{R}^d \) form a Riesz basis for their closed linear span in \( (L^2(\mathbb{R}^d), \| \cdot \|_2) \).

Spline type spaces are a widely used instrument in signal analysis ([2, 3, 5, 44]), wavelet theory or approximation theory or in the theory of irregular sampling. There is a substantial body of results in the literature (using various terminologies) on the subject. Often one assumes compact support, but in other settings one cannot rely on this assumption, e.g. in the context of minimal norm interpolation in Sobolev spaces for the case that Sobolev’s embedding is valid, i.e. for \( s > d/2 \) (cf. [30]). These spaces are a well established tool in (numerical) harmonic analysis.
([1, 8]) and (true) spline functions are popular tools among numerical analysts. For this reason it is also important to have numerically efficient and theoretically justified ways to compute e.g. the projection onto a spline-type space in a constructive way, with a controlled small error. It will not come as a surprise to the reader that the method is based on a constructive (approximate) computation of the family of dual generators, with a control of the error in appropriate function spaces. It is this aspect in which the present note differs from more theoretical investigations in the literature (see [11–14, 34, 39]).

For the description of the problem the Wiener amalgam space \( W(\mathbb{R}^d) := W(C_0, \ell^1)(\mathbb{R}^d) \) also known as Wiener’s algebra ([18, 42]) and \( S_0(\mathbb{R}^d) \), also known as Feichtinger’s algebra play a crucial role ([19, 36, 42]). The first because it is a large subspace of \( L^1(\mathbb{R}^d) \), controlling the norm of the analysis and synthesis operator on general \( L^p \)-spaces, the second because of its Fourier invariance and because cross-correlation functions of pairs of functions from \( W(\mathbb{R}^d) \) are in \( S_0(\mathbb{R}^d) \).

The structure of the paper is as follows. In the second section we introduce the necessary concepts from mathematical signal analysis, in the setting of the Euclidean spaces \( \mathbb{R}^d \), with a focus on spline-type spaces. The third section is devoted to the theoretical and practical considerations of computing the dual system in a multi-window spline-type space. In section 4, we give a numerical approach of the direct method, followed in section 5 by the problem reformulation in a continuous setting in order to effectively construct a good approximation to the exact biorthogonal system. The algorithmic realization for the finite case follows in section 6. The application of our result to the best approximation of a Hilbert-Schmidt operator by generalized Gabor multipliers, using smooth analysis and synthesis windows is introduced in Section 7. In section 8 the numerical realization of approximate dual systems is treated in detail and corresponding error estimates are given. In the final section we address several practical aspects, like typical applications in Gabor analysis or a numerical table showing the errors caused by the reduction of dimensionality.

2 Mathematical preliminaries

We will work with \( L^p \)-spaces over \( \mathbb{R}^d \), \( 1 \leq p \leq \infty \), with respect to the Lebesgue measure, with special emphasize on the Hilbert space \( L^2(\mathbb{R}^d) \). These spaces are isometrically invariant under translation operators \( T_x \), \( x \in \mathbb{R}^d \), defined by \( T_x \varphi(z) = \varphi(z-x) \), and the modulation operators \( M_s \), given as \( M_s \varphi(z) = e^{2\pi is \cdot z} \varphi(z) \), for \( z,s \in \mathbb{R}^d \), i.e., modulation is just multiplication of a function by the character \( \chi_s(t) := e^{2\pi is \cdot t} \), where we interpret \( s \cdot t \) as the scalar product between the vectors \( t,s \in \mathbb{R}^d \). When combining them we prefer to think of time-frequency shift operators \( \pi(\lambda) = M_s T_t \), where \( \lambda = (t,s) \) is a general point in the time-frequency plane \( \mathbb{R}^d \times \mathbb{R}^d \).

The Short-time Fourier transform (STFT) of an \( L^2 \)-function \( \varphi \) with respect to some window function \( g \in L^2(\mathbb{R}^d) \) is defined as a function over the TF-plane by \( \mathcal{V}_g(\varphi)(t,s) = \langle \varphi, M_s T_t g \rangle \). For fixed \( t \) it is – as a function of \( s \) – the Fourier transform of \( \varphi \cdot T_t g \), describing the local frequency content of \( \varphi \) near \( t \) (usually it is displayed in the \( y \)-direction).
Since we are going to work with the Fourier transform (FT) normalized as

$$\hat{\varphi}(s) = \int_{\mathbb{R}^d} \varphi(x) e^{-2\pi i x \cdot s} dx,$$

Plancherel’s theorem states that $\mathcal{F}: \varphi \mapsto \hat{\varphi}$ is a unitary mapping, and its inverse is obtained by replacing the exponential term in the definition of the FT by its conjugate. Moreover Nyquist sampling for band-limited functions with spectrum in $[-1/2, 1/2]^d$ can take place over the integer lattice $\mathbb{Z}^d$, and for any lattice $\Lambda = a\mathbb{Z}^d \lhd \mathbb{R}^d$ the orthogonal lattice $\Lambda^\perp$ is of the form $1/a \cdot \mathbb{Z}^d$ (cf. [15, 42]). As another consequence the Short-time Fourier transform satisfies

$$\|\mathcal{V}_g(\varphi)\|_2 = \|\varphi\|_2 \cdot \|g\|_2$$

for $\varphi, g \in L^2(\mathbb{R}^d)$. Hence for $\|g\|_2 = 1$ the mapping $\varphi \mapsto \mathcal{V}_g(\varphi)$ is an isometric map from $L^2(\mathbb{R}^d)$ into $L^2(\mathbb{R}^2d)$, so that its adjoint mapping inverts the STFT on its range. This gives the standard reconstruction formula for the STFT for normalized $g$:

$$\varphi = \int_{\mathbb{R}^d \times \hat{\mathbb{R}}^d} \mathcal{V}_g(\varphi)(t, s) M_s T_t g \, dt \, ds. \quad (2)$$

This representation is to be understood in the weak, vector-valued sense, but for good functions and atoms $\varphi, g$ it can even be interpreted as the limit of Riemannian sums, convergent in the norm sense ([45]). Furthermore it is essentially a consequence of the Cauchy-Schwarz inequality that $\mathcal{V}_g \varphi$ is indeed a bounded and continuous function on $\mathbb{R}^d \times \hat{\mathbb{R}}^d$, with $\|\mathcal{V}_g \varphi\|_\infty \leq \|\varphi\|_2 \|g\|_2$.

Whenever $g$ is not just $L^2$, but for example in the Schwartz-space of rapidly decreasing functions, one can take the STFT for any tempered distribution and find that it is a continuous function over $\mathbb{R}^d \times \hat{\mathbb{R}}^d$ with at most polynomial growth. It is therefore possible and indeed very useful to define a family of function spaces (usually called modulation spaces over $\mathbb{R}^d$) which are characterized by the (global) behaviour of the STFT (cf. Gröchenig [36] for a detailed discussion). Indeed, according to the theory of coorbit spaces ([22, 23]), for any weighted $L^p$-space (or even more generally, any translation-invariant, solid Banach space of functions over $\mathbb{R}^d \times \hat{\mathbb{R}}^d$) the space of tempered distributions with $\mathcal{V}_g(\varphi)$ in that space is a well-defined Banach space with respect to its natural norm (i.e., $\|\varphi\| := \|\mathcal{V}_g(\varphi)\|_p$), and that different non-zero “windows” $g$ define the same space with equivalent norms. Therefore it is convenient to choose $g = g_0$, with $g_0(t) = e^{-\pi t^2}$, the Fourier-invariant Gauss function, as the window in the definition of such spaces.

**Definition 1.** A family $\Psi = (\psi_k)_{k \in \mathbb{Z}^d} = (T_k \psi_0)_{k \in \mathbb{Z}^d}$ is called a regular, uniform, and bounded partition of unity in the Fourier algebra on $\mathbb{R}^d$ (for short: a BUPU in $\mathcal{F}L^1(\mathbb{R}^d)$) if it satisfies:

1. $\psi_0$ is compactly supported and $\hat{\psi}_0 \in L^1(\mathbb{R}^d)$;
2. $\sum_{k \in \mathbb{Z}^d} \psi_k(x) = \sum_{k \in \mathbb{Z}} \psi_0(x - k) \equiv 1$ on $\mathbb{R}^d$.

For the justification of some key steps in our argument and the description of error estimates we will need Wiener amalgam spaces. They can be defined over locally compact groups, using
general partitions of unity, but for our setting it is enough to make use of regular BUPUs in the Fourier algebra, generated by any suitable function \( \psi_0 \) using its translates along \( \mathbb{Z}^d \).

Regularity refers to the fact that the BUPU is obtained using shifts of a single function along a lattice, uniformity refers to the uniform size of the supports of the family, here \( \text{supp}(\psi_k) = k + \text{supp}(\psi_0) \), and boundedness is understood with respect to the norm of the Fourier algebra \( (\mathcal{F}L^1, \| \cdot \|_{\mathcal{F}L^1}) \), which is obvious from the fact that \( \| T_x \psi \|_{\mathcal{F}L^1} = \| \psi \|_{\mathcal{F}L^1} \) for all \( x \in \mathbb{R}^d \).

The norm of \( f \) in some Wiener amalgam spaces measures the global behaviour of local norms, which can be taken in some quite general function space \( (X, \| \cdot \|_X) \), usually called the local component of the amalgam space. All we need here is that \( \mathcal{F}L^1 \) acts boundedly by pointwise multiplications, i.e. that \( \| h \cdot f \|_X \leq \| h \|_{\mathcal{F}L^1} \| f \|_X, f \in X, h \in \mathcal{F}L^1. \) Then we define for \( q \geq 1 \) the Wiener amalgam space as follows (see [20]):

\[
W(X, \ell^q) = \{ f \in X_{\text{loc}} \mid \| f \|_{W(X, \ell^q)} := \left( \sum_k \| f \psi_k \|_X^q \right)^{1/q} < \infty \}. 
\] (3)

Natural modifications occur for the case \( q = \infty \), or in case one wants to have more general global components, such as weighted version \( \ell^q_w \) instead of simply \( \ell^q \). With their natural norm \( W(X, \ell^q) \) is a Banach space, and different partitions of unity define the same space and equivalent norms. Extensions to general LCA groups are easy to obtain using minor modifications.

One of the important properties of these spaces are the natural inclusion relations, such as \( W(C_0, \ell^1) \subseteq W(L^2, \ell^1) \subseteq L^1(\mathbb{R}^d) \) and the convolution relations, such as \( L^1(\mathbb{R}^d) * W(C_0, \ell^1) \subseteq W(C_0, \ell^1) \) and \( W(C_0, \ell^1) * W(C_0, \ell^1) \subseteq W(\mathcal{F}L^1, \ell^1) \), and corresponding norm estimates ([20,37]).

The most important amalgam space for us will be \( S_0(\mathbb{R}^d) = W(\mathcal{F}L^1, \ell^1) \), which can also be defined by means of the short-time Fourier transform (it is called \( M^1 \) in [36]):

\[
S_0(\mathbb{R}^d) = \{ \varphi \in L^2(\mathbb{R}^d) \mid V_{g_0}(\varphi) \in L^1(\mathbb{R}^d \times \hat{\mathbb{R}}^d) \},
\] (4)

with natural alternative norm \( \| \varphi \|_{S_0} := \| V_{g_0} \varphi \|_{L^1} \). In the sequel we need a couple of basic results concerning this space (cf. [27,31], for further details see also [24]):

**Theorem 1.** (1) \( (S_0(\mathbb{R}^d), \| \cdot \|_{S_0}) \) is a Banach space, continuously embedded into \( (L^p(\mathbb{R}^d), \| \cdot \|_p) \), for \( 1 \leq p \leq \infty \), and TF-shifts act isometrically on \( S_0(\mathbb{R}^d) \), i.e.,

\[
\| \pi(\lambda) \varphi \|_{S_0} = \| \varphi \|_{S_0} \forall \lambda \in \mathbb{R}^d \times \hat{\mathbb{R}}^d, \forall \varphi \in S_0(\mathbb{R}^d);
\]

(2) \( S_0(\mathbb{R}^d) \) is invariant under the Fourier transform as well as under automorphisms of \( \mathbb{R}^d \).

(3) For any lattice \( \Lambda \vartriangleleft \mathbb{R}^d \) there exists a constant \( C_\Lambda \) such that

\[
\sum_{\lambda \in \Lambda} | \varphi(\lambda) | \leq C_\Lambda^1 \| \varphi \|_{S_0} \forall \varphi \in S_0(\mathbb{R}^d).
\]

Hence \( S_0(\mathbb{R}^d) \) is a natural domain for the Poisson summation formula (PSF): For \( C_\Lambda > 0 \)

\[
\sum_{\lambda \in \Lambda} \varphi(\lambda) = C_\Lambda \sum_{\lambda^+ \in \Lambda^+} \varphi(\lambda^+) \forall \varphi \in S_0(\mathbb{R}^d);
\] (5)
(4) For \( \varphi \in S_0(\mathbb{R}^d) \) and for any lattice \( \Lambda \), the \( \Lambda \) -periodization of \( \varphi \) (obtained by the absolutely convergent sum \( \varphi^\Lambda(x) := \sum_{\lambda \in \Lambda} \varphi(x + \lambda) \)) belongs to the space of periodic functions having absolutely convergent Fourier series expansion, which we denote by \( \mathcal{A}(\mathbb{R}^d/\Lambda) \).

**Proof.** The first three statements are covered by the references indicated above. The last statement is essentially a combination of the first two ones, since the Fourier coefficients of \( \varphi^\Lambda \) are exactly the sampling values of \( \hat{\varphi} \in S_0(\mathbb{R}^d) \) over the orthogonal lattice \( \Lambda^\perp \), and the fact that the periodization converges uniformly over compact sets.

\[ \square \]

**Remark 1.** Another consequence of the convolution relations for Wiener amalgam spaces is the fact that the cross-correlation \( \varphi \ast \psi^* \) of two functions in \( W(L^2, \ell^1) \) (and hence in particular for those in \( W(C_0, \ell^1) \)) belong to \( S_0(\mathbb{R}^d) \), because we have \( W(L^2, \ell^1) \ast W(L^2, \ell^1) \subseteq W(\mathcal{F}L^1, \ell^1) = S_0 \), hence (by (3) above) the samples of this autocorrelation function belong to \( \ell^1(\Lambda) \), using here of course the simple fact that \( W(L^2, \ell^1) \) is invariant under the involution \( \varphi^*(z) = \bar{\varphi}(-z) \).

Next, we want to state various results concerning spline-type spaces. Let us first give the definition of spline-type spaces (sometimes also called principal shift-invariant spaces ( [14]), when they are generated by a single function and its translates). Before doing so we remind the reader that a Banach space of sequences \( B_d \) is called solid if for any sequence \( c \in B_d \) and \( d \) with \( |d(\lambda)| \leq |c(\lambda)| \) (i.e. in the pointwise sense), \( d \in B_d \) and \( \|d\|_{B_d} \leq \|c\|_{B_d} \).

**Definition 2.** Let \( \Lambda \) be any discrete subgroup of \( \mathbb{R}^d \), and \( \varphi \) be any element in some translation-invariant Banach space \( (B, \| \cdot \|_B) \) of functions or distributions on \( \mathbb{R}^d \). Then the closed linear span of the family \( (T_\lambda g)_{\lambda \in \Lambda} \) is a closed, \( \Lambda \)-invariant subspace of \( (B, \| \cdot \|_B) \). We denote this space by \( V^B_{\varphi, \Lambda} \), and call it a **spline-type space** (in \( B \)) if the family \( (T_\lambda g)_{\lambda \in \Lambda} \) is a Riesz projection basis for \( V^B_{\varphi, \Lambda} \), i.e., if the following two properties are valid:

1. There is an associated solid Banach space \( (B_d, \| \cdot \|_{B_d}) \) of sequences on \( \Lambda \) (associated with \( (B, \| \cdot \|_B) \) such that the synthesis mapping (representation operator \( R \)) described by \( R : c \mapsto \sum_{\lambda \in \Lambda} c_\lambda T_\lambda \varphi \) is well-defined and defines a continuous bijection between \( B_d \) and \( V^B_{\varphi, \Lambda} \).

2. There is a bounded linear mapping \( C \) (coefficient mapping) defined on \( B \) with \( C \circ R = Id_{\mathcal{V}} \).

For convenience we write \( V_{\varphi, \Lambda} \) for the case \( B = L^2(\mathbb{R}^d) \) and \( V^p_{\varphi, \Lambda} \) if \( B = L^p(\mathbb{R}^d) \) in the sequel.

**Remark 2.** The existence of the left inverse \( C \) to the given synthesis operator \( R \) is in fact equivalent to the assumption that there exists a bounded and linear projection operator \( P \) from \( B \) onto \( V^B_{\varphi, \Lambda} \). Indeed, if the conditions of our definition are satisfied, then the mapping \( P = R \circ C \) is such a bounded (and idempotent) linear mapping whose range is exactly \( V^B_{\varphi, \Lambda} \). Conversely, if condition (1) is satisfied and we can find such a projection, then \( C = R^{-1} \circ P \) is a well-defined left inverse to \( R \) (the inverse being understood as being the inverse on the range of \( R \), which is \( V^B_{\varphi, \Lambda} \) by assumption). Thus the definition implies that a spline-type space in \( B \) is a closed, complemented subspace of \( B \). If the finite sequences are dense in \( B_d \), then the family \( (T_\lambda g) \) is an unconditional basis (as the image of the unit vectors in \( B_d \) under \( R \)) for \( V^B_{\varphi, \Lambda} \).
The concept of Riesz projection bases has been introduced by G. Zimmermann in his thesis ([46]). Details on this concept indicating already its usefulness in the framework of Gabor analysis are found in ([31], Chap. 3 in [29]). Clearly for the case that \((B, \| \cdot \|_B)\) is a Hilbert space, such as \(L^2(G)\) (and the associated solid Banach space is \(\ell^2(\Lambda)\)), it is enough to care for assumption (1), as for any closed subspace in a Hilbert space one can use the orthogonal projection, which is of norm 1. Thus in the case of Hilbert spaces, condition (1) can be rephrased in the usual way by assuming that there exist constants \(C, D > 0\) such that for all \(c \in \ell^2(\Lambda)\) one has:

\[
C\|c\|^2 \leq \| \sum_{\lambda \in \Lambda} c_\lambda T_\lambda \varphi \|_{L^2} \leq D\|c\|^2.
\]

Indeed, such an estimate follows from the fact that \(V_{\varphi, \Lambda}\) is a closed subspace of \(L^2\), using Banach’s theorem. Conversely, it is evident that the validity of this norm equivalence not only implies that the range of \(R\) is exactly \(V_{\varphi, \Lambda}\), but also that the family \((T_\lambda g)_{\lambda \in \Lambda}\) is a Riesz basis (in the usual sense, having a suitable biorthogonal Riesz basis) for \(V_{\varphi, \Lambda}\).

Thus in short such a family is a Riesz basis for a \(V_{\varphi, \Lambda}\) \((p = 2)\) if and only if it is an \(\ell^2(\Lambda)\)-Riesz projection basis. It is also an \(\ell^1(\Lambda)\)-Riesz projection basis for \(V_{\varphi, \Lambda} \cap L^1(\mathbb{R}^d)\) (with the \(L^1\)-norm) if the corresponding synthesis and analysis windows \(\varphi\) and \(\tilde{\varphi}\) are in \(W(C_0, \ell^1)(\mathbb{R}^d)\). Here \(\tilde{\varphi}\) is the generator for the biorthogonal Riesz basis \((T_\lambda \tilde{\varphi})_{\lambda \in \Lambda}\). Following standard terminology we call \(\tilde{\varphi}\) the dual atom to \(\varphi\) (with respect to \(\Lambda\)). The stated condition ensures among others that \(f * \tilde{\varphi}\) is in \(W(C_0, \ell^1)\) for all \(f \in L^1(\mathbb{R}^d)\), its samples are in \(\ell^1(\Lambda)\), and this is one of the reasons why we have chosen this particular function space setting.

The subspace \(V_{\varphi, \Lambda}^{L^1}(\mathbb{R}^d) = V_{\varphi, \Lambda} \cap L^1(\mathbb{R}^d)\) coincides then exactly with those elements in \(V_{\varphi, \Lambda}\) which have \(\ell^1\)-coefficients in their standard expansion. More generally it follows (easily) that one has \(\ell^p\)-stability for arbitrary values of \(p \in [1, \infty]\) in this case, i.e., the \(L^p\)-norm of a function of the form \(f = \sum_{\lambda \in \Lambda} c_\lambda T_\lambda \varphi\) is equivalent to the \(\ell^p\)-norm of the sequence \((c_\lambda)_{\lambda \in \Lambda}\) (cf. also [6]).

We even have: there are constants \(C_1, C_2\) (not depending on \(p \in [1, \infty]\)), such that for any such \(p\) one has norm equivalence between the \(\ell^p\)-norm of the coefficients, and the continuous \(L^p\)-norm of the function \(f \in V_{\varphi, \Lambda}^p(\Lambda)\). Of course, one can extend such statements easily to weighted \(L^p\)-spaces (for example to weights of polynomial growth up to a certain order) whenever \(\varphi\) satisfies stronger (decay) conditions. These conditions are of course satisfied for any Schwartz functions \(\varphi\).

### 3 The role and the computation of the dual atoms

Before describing the multi-window setting let us recall the case of principal shift invariant spaces, which are generated by a single generator. This is the basis for the generalization envisaged in this paper. We discuss the structure of the biorthogonal (sometimes called dual) system and the concrete form and mapping properties of the projection operators onto the spline-type spaces.

Although the theory is valid over LCA groups, we work with general lattices \(\Lambda \lhd G = \mathbb{R}^d\), which are of the form \(\Lambda = A \mathbb{Z}^d\), for some non-singular \(d \times d\) matrix \(A\). For the applications and numerical realization we will use only \(\Lambda = \mathbb{Z}^d \lhd \mathbb{R}^d\), because by a simple transformation (automorphism of \(\mathbb{R}^d\)) of the lattice and the corresponding generator system one can reduce the
general problem to this setting, at the cost of replacing the set of generators by their transformed version (see [40]).

**Corollary 1.** For \((\varphi, \Lambda)\) as above, both families \((T_\lambda \varphi)_{\lambda \in \Lambda}\) and \((T_\lambda \tilde{\varphi})_{\lambda \in \Lambda}\) form a Riesz projection basis for all spaces \(V^p_{\varphi, \Lambda}\), defined as the closed linear span of \((T_\lambda \varphi)_{\lambda \in \Lambda}\) in \(L^p\), for \(1 \leq p < \infty\), establishing a natural isomorphism between these spaces and the corresponding \(\ell^p\)-spaces over \(\Lambda\).

The characterization of the Riesz basis (basic sequence) property of a family \((T_\lambda \varphi)_{\lambda \in \Lambda}\) (within \(L^2(G)\)) through the condition

\[
H(s) := \sum_{\lambda \in \Lambda} |\hat{\varphi}(s - \lambda^{-1})|^2 \geq \delta_0 > 0 \quad \forall s \in \hat{G}
\]  

(6)

is well-known for many different cases and works for general LCA groups just in the same way as for \(\mathbb{R}^d\) (see [5] for a survey article, or [39], [10]). In fact, it is just a translation of the bounded invertibility of the Gram-matrix on \(\ell^2(\Lambda)\) with entries \(\langle T_{\lambda_1} \varphi, T_{\lambda_2} \varphi \rangle\). In fact, this is a circulant matrix with respect to the group structure of the Abelian group \(\Lambda\). Its entries are the sampling values of the autocorrelation function \(\varphi \ast \varphi^*\) for \(\varphi\) over \(\Lambda\) (where \(\varphi^*(z) = \tilde{\varphi}(-z)\)) The fact that the biorthogonal family for the full system is obtained by applying the inverse Gram-matrix to the original system and the circulant structure of this matrix imply that the biorthogonal Riesz basis has the same structure, just with respect to a possible different set \(\tilde{\Phi}\) of dual generators.

Using them one can describe the projection on the spline-type space as follows:

**Corollary 2.** For \((\varphi, \Lambda)\) as above the orthogonal projection from \(L^2(\mathbb{R}^d)\) onto \(V_{\varphi, \Lambda}\) can be described alternatively as

\[
f \mapsto P_{\varphi, \Lambda}(f) = \sum_{\lambda \in \Lambda} [f \ast \varphi^*](\lambda) \cdot T_\lambda \tilde{\varphi} = \sum_{\lambda \in \Lambda} [f \ast \tilde{\varphi}^*](\lambda) \cdot T_\lambda \varphi.
\]  

(7)

For \(\varphi \in W(C_0, \ell^1)(\mathbb{R}^d)\), this mapping is not only bounded with respect to the \(L^2\)-norm, but also (uniformly) bounded with respect to all \(L^p\)-norms, with \(p \in [1, \infty]\).

*Proof.* One checks that \(\langle f, T_\lambda \varphi \rangle = f \ast \varphi^*(\lambda)\).

\(\square\)

**Remark 3.** Since for \(f \in V_{\varphi, \Lambda}\) on has \(f = P_{\varphi, \Lambda} f\) the reconstruction formula (7) above resembles of course very much the classical Shannon sampling formula, where one has \(\varphi = \text{SINC}\), the cardinal sine, which has the interesting property of forming both an orthogonal family within its closed linear span (\(V_{\varphi, \Lambda}\) is just a Paley-Wiener space of band-limited functions) and being the Lagrange interpolator for this space. However, due to the non-integrability of SINC the reconstruction is limited to band-limited functions in \(L^p\), for the range \(p \in (1, \infty)\), and the corresponding bounds are deteriorating towards the critical values \(p = 1\) and \(p = \infty\) respectively.

**Remark 4.** As a consequence of the above corollary we can explicitly write the dual atom as linear combinations of the original atom with the coefficients coming from inverting the Gramian matrix, whose entries are the values of the autocorrelation function \(\varphi \ast \varphi^*\) for \(\varphi\), sampled over \(\Lambda\).

\[
\tilde{\varphi} = \sum_{\lambda \in \Lambda} b_\lambda T_\lambda \varphi.
\]  

(8)
For the multi-window spline-type case or finitely generated multi-spline type spaces, we consider a set of functions in $L^2(\mathbb{R}^d)$, namely $\Phi = \{\varphi_1, \varphi_2, ..., \varphi_k\}$. The functions $\varphi_1, \varphi_2, ..., \varphi_k$ are called a set of generators (resp. atoms) for the space $V = V_{\Phi, \Lambda} ([1, 4])$.

It is a well known fact that a indexed set of vectors $(g_\rho)_{\rho \in J}$ in a Hilbert-space (here in $\mathcal{H} = L^2(\mathbb{R}^d)$) is a Riesz basis (for its closed linear span, sometimes called a Riesz basic sequence) if and only if its Gramian, given by

$$G = (g_{\rho, \rho'}) = (\langle g_{\rho'}, g_\rho \rangle)_{\rho, \rho' \in J}$$

is boundedly invertible on $\ell^2(\Lambda)$. The inverse Gramian matrix allows to find the biorthogonal family $(\tilde{g}_\rho)_{\rho \in J}$ in the closed linear span of the family $(g_\rho)_{\rho \in J}$. We have $\tilde{g}_{\rho'} = \sum_{\rho} d_{\rho'}^\rho g_\rho$, where the coefficients are taken from the $\rho'$-th row (or column) of $G^{-1}$. Conversely, the existence of a biorthogonal family generating a bounded mapping from $L^2(\mathbb{R}^d)$ into $\ell^2(\Lambda)$ implies that $(g_\rho)_{\rho \in J}$ is a Riesz basis.

**Definition 3.** Let $\Phi = (\varphi_1, ..., \varphi_k)$ and $c = (c_1, ..., c_k)$ we can write the multi-window spline-type space - in analogy to ordinary spline-type spaces - in the form

$$V_{\Phi, \Lambda}^p = \{ f = \sum_{\lambda \in \Lambda} c_\lambda T_{\lambda} \Phi := \sum_{i=1}^k \sum_{\lambda \in \Lambda} c^i_{\lambda} T_{\lambda} \varphi^i | c^i = (c^i_{\lambda}) \in \ell^p(\Lambda), \text{ for } i = 1, ..., k \}.$$  

(9)

Under the conditions on $\Phi$ described below these spaces turn out to be closed subspaces of $L^p(\mathbb{R}^d)$ and in fact of $W(C_0, \ell^p)(\mathbb{R}^d)$, for $1 \leq p \leq \infty$.

## 4 A first numerical approach

For practical purposes, we will discuss the (algebraic part) of the problem first in the setting of finite length signals. The transition to continuous variables (resp. from long to short signals) is a separate issue to be addressed below. We write $n$ for the length of the atoms and $k$ for the number of atoms (e.g. the number of original spline windows used for analysis). We consider a finite, discrete lattice $\Lambda$ consisting of $n/a$ points, where $a$ is the translation step. Hence we have $k \leq n/a$, because otherwise we cannot have linear independence. Let us recall that we are looking for an efficient algorithm to find the biorthogonal family, assuming from the beginning that the set of all translates of the system $\Phi = \{\varphi_1, \varphi_2, ..., \varphi_k\}$ along a (common) lattice $\Lambda$ forms a Riesz basis for its closed span, which will be called a multi-windows spline-type space. The general pattern suggested by linear algebra is quite clear: one has to build the corresponding Gram-matrix $[GR_\Phi]$, calculate its inverse and then apply it to the original system in order to obtain the biorthogonal system (which by assumption is uniquely determined, belonging to the same space). The algorithm naturally follows the steps described below:
**Algorithm 1**: Dual system through direct Gramian inverse

**Input**: $\Phi = [\varphi_1, \varphi_2, ..., \varphi_k]$ the multiwindow systems, $\Lambda$ - the lattice

**Output**: $\tilde{\Phi}$ biorthogonal dual system

Establish the family $TRL = (T_\lambda \Phi)_{\lambda \in \Lambda}$ as a Riesz basis for its closed linear span $V_{\Phi, \lambda}$;

Compute the Gramian as $[GR_\Phi] = TRL' * TRL$;

(in standard MATLAB notation) Invert the Gramian $[GR_\Phi]^{-1}$, which is possible due to the Riesz basis assumption;

Getting the biorthogonal family as $\tilde{\Phi} = [GR_\Phi]^{-1} * TRL$

Although simple to describe already the storage requirements for this algorithm are enormous, for long signals or the 2D applications we have in mind/ We will show how to reduce the complexity using the translation invariance of the system. In order to store all the information about the Gramian it is enough to store these values

$$\langle T_\lambda \varphi_j, T_\lambda \varphi_k \rangle = \langle \varphi_j, T_{\lambda - \lambda} \varphi_k \rangle.$$  \hspace{1cm} (10)

Practically speaking they can be handled as a stack of $k \times k$ matrices, or equivalently, as a matrix-valued function on $\Lambda$ described by the symbol $[G_\Phi]$:

$$[G_\Phi]_{i,j} = \begin{pmatrix} [G_{\varphi_1, \varphi_1}] & [G_{\varphi_1, \varphi_2}] & \cdots & [G_{\varphi_1, \varphi_k}] \\ [G_{\varphi_2, \varphi_1}] & [G_{\varphi_2, \varphi_2}] & \cdots & [G_{\varphi_2, \varphi_k}] \\ \vdots & \vdots & \ddots & \vdots \\ [G_{\varphi_k, \varphi_1}] & [G_{\varphi_k, \varphi_2}] & \cdots & [G_{\varphi_k, \varphi_k}] \end{pmatrix}$$

where each of $[G_{\varphi_i, \varphi_j}] = (\langle \varphi_i, T_\lambda \varphi_j \rangle)_{\lambda \in \Lambda} = (\varphi_i \ast \varphi_j^*)_{\Lambda};$ $i, j \in \{1, ..., k\}$.

The most trivial but also the most inefficient way of obtaining the dual biorthogonal family is now to invert the full Gramian matrix (cf. algorithm 1). There is however an important benefit from building this huge matrix: it allows to study its structure in order to come up with a more elegant solution\(^1\) to the original problem.

In some cases (this is well known from wavelet theory) one needs an orthonormal basis for the same space. There is also a well-known method to produce such orthonormal bases using instead of the inverse the square root of the inverse of the Gramian (\([35]\)). In fact, this methods is well known since the early days of wavelet theory and realizes what is sometimes called the Loewdin orthonormalization (\([43]\)). We do not go into details on this matter here but mention here that in the case that the original multi-window Riesz basis has a good condition number the new orthonormal basis consisting of translates of a multi-window system can be shown to have similar concentration around zero as the original system.

\(^1\)This was the original approach taken by the authors, which finally led to the much more compact and general method presented below.
5 The reformulation of the problem using biorthogonality

Next we describe how to effectively construct the set $\tilde{\Phi}$ of $k$ generators of the biorthogonal system. The full biorthogonal family is then obtained by shifting this family along $\Lambda$. Obviously, there is nothing to do if they generate an orthonormal system, whereas the information about the correlation structure of the set of shifted copies has to be taken into account for the general case. So let us start from the available data:

$$a^{sr}_\lambda := \langle \varphi^*, T_\lambda \varphi^s \rangle, \quad 1 \leq r, s \leq k, \lambda \in \Lambda. \quad (11)$$

The individual elements of the biorthogonal system have to be linear combinations of the elements of the supposed Riesz basis, so each of them is of the form $h = \sum_{\lambda} \sum_{s} c^i_{\lambda} T_\lambda \varphi^i$, where $c = (c^i_{\lambda}) \in \ell^1(\Lambda)^k$. Later on the short-hand notation $f = \sum_{\lambda} c_{\lambda} T_\lambda \Phi$ will be used for convenience, in analogy to the 1D case, see [8], with the convention of writing $c_{\lambda} T_\lambda \Phi := \sum_{i} c^i_{\lambda} T_\lambda \varphi^i$.

Let us also agree to denote the coefficients for the biorthogonal family $\tilde{\Phi}$ by choosing $f = \tilde{\varphi}^j, 1 \leq j \leq k$ by $b^j_{\lambda}$. Our aim is to calculate the coefficients $(b^j_{\lambda})$ from the auto-correlation family $(a^{sr}_\lambda)$.

It turns out to be helpful to consider such collections of coefficients in two alternative ways: at the one hand $\lambda \rightarrow B(\lambda) := (b^j_{\lambda})_{(i,j)}$ is a matrix valued function on $\Lambda$, on the other hand one has for each fixed pair of matrix coordinates $(l,j)$ a function on $\lambda$ (in our case in $\ell^1(\Lambda)$), which could be naturally denoted by $B(i,j)$.

Using this setting the biorthogonality relation between the family $\tilde{\Phi} = [\tilde{\varphi}^1, \ldots, \tilde{\varphi}^k]$ and $\Phi = [\varphi^1, \ldots, \varphi^k]$ is equivalent to the following identity (with $\ast_{\Lambda}$ denoting convolution over $\Lambda$):

$$\delta_{i,j} \cdot \delta_{0,\lambda'} = \langle \tilde{\varphi}^j, T_\lambda \varphi^i \rangle = \langle \sum_{\lambda} \sum_{s} b^j_{\lambda} T_\lambda \varphi^i, T_\lambda \varphi^i \rangle = \sum_{\lambda} \sum_{s} b^j_{\lambda} \langle \varphi^i, T_\lambda \varphi^j \rangle =$$

$$= \sum_{\lambda} \sum_{s} a^{sj}_{\lambda} \cdot b^l_{\lambda} = \sum_{\lambda} a(i, l) \ast_{\Lambda} b(l, j).$$

Since the most efficient way to realize the convolution over $\Lambda$ is of course to apply the corresponding (discrete) Fourier transform $\mathcal{F}_\Lambda$ we can say that for any fixed set of indices $l, j, i$ one has

$$a(i, l) \ast_{\Lambda} b(l, j) = \mathcal{F}^{-1}_\Lambda[\mathcal{F}_\Lambda a(i, l) \ast \mathcal{F}_\Lambda b(l, j)] = \mathcal{F}^{-1}_\Lambda[\alpha(i, l) \ast \beta(l, j)],$$

where we write $\alpha(i, l)$ for $\mathcal{F}_\Lambda[a(i, l)]$ and $\beta(l, j)$ for $\mathcal{F}_\Lambda[b(l, j)]$, and $\ast$ stands for pointwise multiplication, following MATLAB convention. Again there are two possible views: we could consider $\alpha$ and $\beta$ as a collection of functions (in the Fourier algebra) on $\hat{\Lambda}$, or as matrix-valued functions on $\Lambda$, which is our viewpoint, i.e. we think of $\alpha(\omega)$ and $\beta(\omega)$ as functions on a compact group. In the applications $\Lambda$ will be isomorphic to $\mathbb{Z}^d$ and hence $\alpha$ and $\beta$ are matrix-valued functions on $\mathbb{T}^d$. We thus come up with the following equivalent description of the biorthogonality relation:

$$Id_k = \delta_{i,j} \cdot \mathcal{F}^{-1}_{\Lambda}(\delta_{i,j}) = \sum_{l} [\alpha(i, l)(\omega) \ast \beta(l, j)(\omega)] = \alpha(\omega) \ast \beta(\omega) \quad \forall \omega \in \Lambda,$$
where * means ordinary matrix multiplication (again following standard MATLAB conventions).

In this setting the unique solution to the biorthogonality problem is given by matrix inversion:

\[ \beta(\omega) = \alpha^{-1}(\omega), \quad \forall \omega \in \hat{\Lambda}. \]  

(12)

The collection of coefficients \( B = (b_{l,j}^i) \) to find \( \tilde{\Phi} \) is thus obtained by applying (for all pairs \((l, j)\)) the inverse Fourier transform \( F_{\Lambda}^{-1} \) to the individual coordinate functions \( \beta(l, j) \).

The considerations above show that we can recast the problem of finding the biorthogonal system for \( \Phi \) to the inversion of a continuous family of \( k \times k \)-matrices indexed by the group \( \hat{\Lambda} \). This provides us already with important information about the properties of the coefficient sequences \( b(l, j), 1 \leq l, j \leq k \). Knowing that the entries of \( \alpha(i, j) \) belong to the algebra \( A(\hat{\Lambda}) \) of absolutely convergent Fourier series on \( \hat{\Lambda} \) (being the Fourier transforms of sequences \( a(i, j) \in \ell_1(\Lambda) \)) we can invoke Wiener’s inversion theorem ([42]) in conjunction with Cramer’s rule in order to conclude that also the entries of their inverse matrices, i.e. the functions \( \omega \mapsto \beta(l, j)(\omega) \) are in \( A(\hat{\Lambda}) \). This in turn implies that each of the sequences \( b(l, j) \) belong to \( \ell^1(\Lambda) \), hence the dual atoms \( (\tilde{\varphi}_j), 1 \leq j \leq k \) are constituted by absolutely convergent series of elements from the given Riesz basis. Since translations act uniformly bounded on \( W(C_0, \ell^1) \) we have the following result:

**Theorem 2.** Assume that the set of all \( \Lambda \)-translates of a set of atoms is a Riesz basis for its closed linear span \( V_{\Phi, \Lambda} \) within \( L^2(\mathbb{R}^d) \). If in addition the family \( \Phi = (\varphi_1, \ldots, \varphi_k) \) is in \( W(C_0, \ell^1) \), then the family of dual atoms \( \tilde{\Phi} \) generating the biorthogonal Riesz basis \( (T_\lambda \tilde{\Phi})_{\lambda \in \Lambda} \) also belongs to \( W(C_0, \ell^1) \cap V_{\Phi, \Lambda} \). As a consequence the orthogonal projection operator in \( L^2(\mathbb{R}^d) \) extends to the full range of \( L^p \)-spaces for \( 1 \leq p \leq \infty \):

\[ P : f \mapsto \sum_{k \in \Lambda} \langle f, T_\lambda \tilde{\Phi} \rangle_{L^2} T_\lambda \Phi =: (f * \tilde{\Phi}^*)(\lambda)T_\lambda \Phi. \]

(13)

and defines a bounded projection from \( L^p \) onto \( V_{\Phi, \Lambda}^p \).

**Proof.** The arguments for the statement have been collected already above. As for the conclusion concerning the projections again convolution relations between Wiener amalgams can be used, see for example ([7, 20, 37]). \( \square \)

**Remark 5.** Following Theorem 2 we can explicitly write the dual generator \( \tilde{\Phi} = (\tilde{\varphi}_1, \ldots, \tilde{\varphi}_k) \) as

\[ \tilde{\Phi} = \sum_{\lambda \in \Lambda} B_{\lambda} T_{\lambda} \Phi. \]

(14)

**Remark 6.** Although this reformulation of the problem reduces the inversion of a huge (somewhat structured) Gramian matrix to the inversion of an indexed family of small matrices of size \( k \times k \) the constructive realization of the corresponding procedure is still not feasible in the literal sense. We discuss this matter later on, using the fact that all the involved sequences are in \( \ell^1(\Lambda) \), hence the coordinate functions of \( \beta \) are in the Fourier algebra \( A(\hat{\Lambda}) \) due to Wiener’s inversion theorem. The natural idea is of course to do the (small) matrix inversion only on a sublattice of \( \hat{\Lambda} \). This is plausible, since it is sufficient to know a function in \( A(\hat{\Lambda}) = S_0(\hat{\Lambda}) \) on such a lattice, because
one can use quasi-interpolation operators to reconstruct the function over all of $\hat{\Lambda}$ (see [26]). If the function is band-limited (i.e. if the corresponding sequence has finite support in $\Lambda$) one has of course perfect reconstruction, but the inverse of such a convolution operator never has finite support, so we have to rely on good approximation.

Since the finite subgroups of $\hat{\Lambda}$ are of the form $\hat{\Lambda}/\Lambda_0$, with $\Lambda_0$ being a sufficiently coarse sublattice of $\Lambda$, but it is clear that we will have to deal with $\Lambda_0$-periodic versions of our families $a$ and $b$ respectively. The point of this short discussion being only to have a pointer, indicating how to relate the general (continuous) problem to a related problem over finite Abelian groups, which can be solved using fast algorithms, to be discussed in the next section.

Theorem 2, combined with know convolution theorems for Wiener amalgam spaces allows as to draw the following important conclusion:

**Theorem 3.** Given the situation of Thm. 2 we have: For every $\varepsilon > 0$ there exists a computable method of finding an approximate version $\tilde{\varphi}_a^j$ with $\|\tilde{\varphi}_a^j - \hat{\varphi}_a^j\|_W < \varepsilon, 1 \leq j \leq k$. Consequently, we can approximate, in the operator norm, the projection operator as described in (?) by using $\tilde{\varphi}_a^j$ instead of $\hat{\varphi}_a^j$, in fact uniformly with respect to $p \in [1, \infty]$.

**Remark 7.** It is easy to generalize the above statement to the setting of weighted spaces, using (sub-multiplicative) weights $w$. Starting from functions $\varphi_j \in W(C_0, \ell^1_w)(\mathbb{R}^d)$, one finds that the matrix valued functions $(B_k)_{k \in \mathbb{Z}^d}$ have entries in $\ell^1_w(\mathbb{Z}^d)$, and thus by Wiener’s theorem for Beurling algebras (cf. [42]) also the inverse matrices. Consequently the involved functions have a good decay (for $k \to \infty$) and the convolution kernels corresponding to the Gram matrix can be well approximated by “finite” sequences (cf. below).

**Remark 8.** For the structural description of the problem we have used abstract lattices $\Lambda \triangleleft \mathbb{R}^d$ and its dual group (frequency domain). In any concrete situation (see [40] for an early paper at this generality) we can use that $\Lambda = A\mathbb{Z}^d$, where $A$ is some non-singular $n \times n$-matrix. Hence one can use some automorphism of $\mathbb{R}^d$, i.e. a transformation $\alpha$ reducing the problem to the case $\Lambda = \mathbb{Z}^d$. This can be done by replacing $\Phi$ by $\alpha^*(\Phi)$, where $\alpha^*f(x) = f(\alpha(x))$. In this way the action of $\Lambda$ can be replaced by a modified action and a modified family of translation operators, now with $(T_k)_{k \in \mathbb{Z}^d}$ instead of translation $(T_\lambda)_{\lambda \in \Lambda}$. Since both $W(C_0, \ell^1)(\mathbb{R}^d)$ and $S_0(\mathbb{R}^d)$ are invariant under automorphism also the new family $\alpha^*(\Phi)$ has the same properties. The same is of course true in the converse direction and therefore it is enough to treat the problem for the case $\Lambda = \mathbb{Z}^d$ in the sequel. Also, from the practical side, one uses generators of $\Lambda$ and therefore the natural label for a given point $\lambda = Ak \in \Lambda$ is the corresponding point $k \in \mathbb{Z}^d$.

### 6 Algorithmic realization for the finite case

For the case of finite signals Theorem 2 given in the previous section can be turned into the following implementable recipe:
Algorithm 2: Approximate biorthogonal system for 1D/2D signals

| Input: Φ = [ϕ₁, ϕ₂, ..., ϕₖ] the 1D/2D multiwindow systems, Λ = the lattice; |
| Output: ˜Φ biorthogonal dual system; |
| k - number of 1D/2D windows, n, n x n - windows length, a, (a, b) - the translation steps; |
| Stack the approximate correlations ⟨ϕᵣ, Tₗϕₛ⟩, 1 ≤ r, s ≤ k, λ ∈ Λ; |
| Downsample each correlation matrix on the lattice Λ; |
| Apply the unitary F operator to diagonalize each correlation matrix; |
| Inversion procedure for each matrix following the stack order; |
| Back from the frequency domain under the unitary F⁻¹ operator; |
| Synthesize the biorthogonal dual system from the linear combinations of stack matrices; |

Considering the cost estimation of the algorithm 2, we claim that we only have to compute now k² convolutions followed by the application of k² unitary Fourier operations with respect to the lattice Λ. The total cost of inverting the so-called "small Gramians" is reduced now to k³ x #Λ and is very efficient in our assumed case k << #Λ. From theoretical considerations regarding the redundancy, we need in general to consider k << #Λ, in order to have a good expansion and reconstruction.

The algorithm can be used to deploy our recipe both for 1D and the 2D multi-windows splines. We want to emphasize the scalability of our implementation, that allows the direct extension to 2D atoms with minimum modifications.

When comparing the computational costs of the 1D algorithm case and of the 2D case with the same number of windows we observe the following: The order of the algorithm increases only with the order of the second data index (up to a harmless logarithmic term, due to the use of suitable FFTs). In fact, we now only have to apply an FFT to a matrix of size n x n instead of a vector of size n. All the other steps are the same, as outlined in the theoretical description of the proposed method for computing the biorthogonal family. Further reduction of the actual computational costs (at least for large signal size) follow from the reduction steps (shortening of atoms via periodization) given below.

7 Application to the approximation of Hilbert Schmidt Operators by Generalized Gabor Multipliers

An interesting and non-trivial application of the 2D version of above theory is the problem of approximating an operator by a so-called generalized Gabor multiplier with respect to the Hilbert Schmidt (HS) norm. Recall that an “ordinary” Gabor multiplier is constructed from a pair of “windows” (typically the analysis window γ and the synthesis window g are taken from the space $S₀(ℝᵈ)$, a lattice Λ ⊂ ℝᵈ, and a multiplier sequence (mₗ)ₗ∈Λ (also called upper symbol), typically in $ℓ^∞(Λ)$, as follows

$$Tf = \sum_{λ∈Λ} mₗ(f, π(λ)γ)π(λ)g.$$  (15)
The connection between the problem of approximating HS-operators by Gabor multipliers using the Kohn-Nirenberg-calculus (KN) is described in [21]. The implementation of the finite dimensional variant has been the subject of [25]. A generalized Gabor multiplier is an operator which is a finite sum of operators of the above form. They are also studied by Dörfler and Toressani recently ([16,17]) who address the same question in a direct way.

A useful viewpoint on Gabor multipliers makes use of the action of $\mathbb{R}^{2d}$ on operators by $\pi \otimes \pi^*(\lambda) T = \pi(\lambda) \circ T \circ \pi(\lambda)^{-1}$, and $Q$ for the rank-one operator $f \mapsto \langle f, \gamma \rangle g$. Gabor multipliers are thus operators of the form

$$T = \sum_{\lambda \in \Lambda} m_\lambda \pi \otimes \pi^*(\lambda) Q = \sum_{\lambda \in \Lambda} m_\lambda Q_\lambda$$

if we write $Q_\lambda = \pi \otimes \pi^*(\lambda) Q$. Hence a generalized Gabor multiplier is obtained from a sequence $Q^1, \ldots, Q^k$ of rank one operators, with analysis windows $\gamma_1, \ldots, \gamma_k$ and synthesis windows $g_1, \ldots, g_k$ and thus can be compactly written in the form

$$T = \sum_{l=1}^{k} \sum_{\lambda \in \Lambda} m^l_\lambda Q^l_\lambda.$$  

The Kohn-Nirenberg (KNS) mapping $\sigma$ is a unitary isomorphism between the set of all Hilbert-Schmidt operators on $L^2(\mathbb{R}^d)$ with the standard scalar product $\langle T, S \rangle_{HS} := \text{trace}(TS^*)$ and $L^2(\mathbb{R}^{2d})$. For us it is important that $\sigma$ intertwines $\pi \otimes \pi^*$ with the ordinary translation operator, i.e. one has

$$\sigma[\pi \otimes \pi^*(\lambda) Q] = T_\lambda[\sigma(Q)], \ \lambda \in \mathbb{R}^{2d}. $$

Hence the best approximation problem for generalized Gabor multipliers in the Hilbert-Schmidt norm is translated into a best-approximation problem for multi-windows spline-spaces in $L^2(\mathbb{R}^{2d})$ over phase space. Moreover, the Kohn-Nirenberg symbol of $Q$ as above is given by

$$\sigma(Q)(x, \omega) = g(x) \gamma(\omega) exp(-2\pi i x \omega). $$

Hence the KNS-symbol $\sigma(Q)$ is in $S_0(\mathbb{R}^{2d})$ if (and only if) both $\gamma, g \in S_0(\mathbb{R}^d)$. Thus the assumption that the collection $(Q^l_\lambda)$ is a Riesz basis within the Hilbert-Schmidt operators for the closed subspace of all generalized Gabor multipliers of a certain type (with $\ell^2$-coefficients $m^l_\lambda$) implies that the problem can be translated into a spline-type space problem, showing how to approximately compute those coefficients $(m^l_\lambda)$, up to a small $\ell^2$-error. Note that the level of generality required for this task (no compact support, etc.) is well in agreement with the assumptions made in our general theory.

8 Numerical realization of the approximate dual systems

The description of the best approximation of a function by linear combinations of translates given in Thm.2 is explicit and in this sense constructive, and can be transferred as it stands to
the context of arbitrary locally compact Abelian groups. But only for the special case of finite Abelian groups it becomes a recipe that can be implemented on a computer. In fact, the outlined strategy helps to obtain an FFT-based efficient implementation. However, in practice we cannot carry out infinitely many perfect integrations (to get the correlation coefficients), nor can we invert an uncountable family (indexed by the unit square, for the case \(d = 2\)) of matrices, even if they are of small size. Therefore we have to address the errors arising from discretizations and reduction to corresponding finite dimensional approximations of the infinite dimensional problems.

Hence we will describe how to obtain for any \(\varepsilon > 0\) a suitable finite number of samples of the involved functions, and calculate the approximation to \(P_V(f)\) to the given precision, the error being expressed in the \(L^p\)-norm. In this process the role of the atoms \((\varphi_1, \ldots, \varphi_k)\) and the role of \(f\) (the function to be projected) are slightly different. By controlling the error between canonical and approximated (computable) dual elements we can a small relative error (uniformly over all the \(L^p\)-norms) for the realizable reconstruction method. In order to reach such estimates we have use \(\ell^1\)-norm error estimates at the coefficients level. The actual quality of approximation of \(P_V f\) by a computable approximation to this function depends on the smoothness of \(f\).

Let us start with the error arising in the computing of the matrix-valued autocorrelation function \(\alpha = (a_{s,r}^{\Lambda})\), describing the scalar products between the different atoms and their \(\Lambda\)-translates. We have to guarantee that it standard numerical integration methods allow to determine \(\alpha\) up to an arbitrary small \(\ell^1\)-error, if the atoms are in \(W(C_0, \ell^1)\). Since inversion is a continuous operation in the Banach convolution algebra \(\ell^1(\Lambda)\) inverting a slightly perturbed matrix results in a slightly wrong matrix, hence the inverse of the computable approximation to \(\alpha\) will give us a set of coefficient close to the true ones, still in the \(\ell^1\)-norm.

First let us recall a basic property shared by all the standard numerical integration methods: they can be considered as a sequence of bounded linear functionals on \(W(C_0, \ell^1)(\mathbb{R}^d)\). In fact, the classical numerical integration methods, such as Riemannian sums, trapezoidal method or Simpson’s rule approximate the integral by an average over sampling values of \(f\), taken in a finite number of sampling points \((x_k)_{1 \leq k \leq K}\). This means that a given function in \(W(C_0, \ell^1)\) has to be sampled fine enough and long enough (over the essential support of \(f\)) in order to compute integrals approximately. Autocorrelation coefficients can be found essentially by discrete convolution. The key argument for our error estimate only requires the following assumption applying to all the above mentioned above:

**Definition 4.** A numerical integration method is a bounded sequence (or net) of linear functionals \(N_h\) on \(W(C_0, \ell^1)(\mathbb{R}^d)\), each of which is using a finite number from samples of \(f \in W(C_0, \ell^1)(\mathbb{R}^d)\), such that for all such \(f\) one has:

\[
N_h(f) \rightarrow \int_{\mathbb{R}^d} f(x)dx \quad \text{for} \quad h \rightarrow 0. \tag{20}
\]

Since the dual space to \(W(C_0, \ell^1)(\mathbb{R}^d)\) is known to be \(W(M, \ell^\infty)(\mathbb{R}^d)\), the space of translation bounded measures, one can express it equivalently by saying that \(N_s(f) = \mu_s(f)\), where for
each fixed parameter $s$ one has $\mu_h(f) = \sum_j c_j \delta_{x_j}$, with the side condition that
\[
\sup_s \sup_x \sum_{j: x_j - x \leq 1} |c_j| \leq C_M < \infty.
\] (21)

The convergence assumption can be viewed as the assumption that these functionals are $w^*$-convergent to the linear functional $f \to \int_{\mathbb{R}^d} f(x) dx$, the (infinite) Riemann integral over $\mathbb{R}^d$ ([9]).

In most cases $\mu_h$ is just a finite discrete measure. For example, Riemannian sums correspond to $h = 1/K \mu_h = h \cdot \sum_{k=-K^2}^{K^2} \delta_{kh}$.

Thus we have to control the error between the ideal sequence $\alpha$ and the computable sequence $\alpha^{(s)}$ obtained by replacing all the scalar products defining $a^{s,r}_\lambda = \langle \varphi^r, T_\lambda \varphi^s \rangle$, for $1 \leq r, s \leq k, \lambda \in \Lambda$, which we want to describe collectively as $\alpha^{(h)}$. We will show that for $h \to 0$ one has $\|\alpha - \alpha^{(h)}\|_{\ell^1} \to 0$.

In order to prove this we need the lemma below. For the proof it will be convenient to use on $W = W(C_0, \ell^1)(\mathbb{R}^d)$ the alternative norm arising from the regular BUPU obtained by choosing as $\psi_0$ the indicator function ([32, 42]) of $Q_0 = [0, 1]^d$:
\[
\|f\|_W = \sum_{n \in \mathbb{Z}^d} \sup_{x \in Q} |f(n + x)|.
\] (22)

For this norm we can easily prove:

**Lemma 1.** For each lattice $\Lambda < \mathbb{R}^d$ there exist $C_\Lambda$ such that for $g, h \in W(L^\infty, \ell^1)$
\[
\sum_{\lambda \in \Lambda} \|T_\lambda g \cdot h\|_W \leq C_\Lambda \|g\|_W \|h\|_W.
\] (23)

**Proof.** Since $W(L^\infty, \ell^p)$ is invariant under automorphisms of $\mathbb{R}^d$ (i.e. $f \mapsto \alpha^*(f)$, with $\alpha^*(f) = f(Ax)$, for any invertible matrix $A$), we can reduce (up to equivalence of norms) the problem to the case $\Lambda = \mathbb{Z}^d$ ([36]). Then, starting from $\psi_n = T_n \psi_0$, we consequently have:
\[
\sum_{k \in \mathbb{Z}^d} \|T_k g \cdot h\|_W \leq \sum_{k \in \mathbb{Z}^d} \sum_{n \in \mathbb{Z}^d} \|(T_k g \cdot h) \psi_n\|_W.
\]

Since $\psi_n = \psi_n \cdot \psi_n$ and $T_{-k} \psi_n = \psi_{n-k}$ implies $\|T_k g \cdot \psi_n\|_\infty = \|g \cdot \psi_{n-k}\|_\infty$ we reach the estimate
\[
\sum_{k \in \mathbb{Z}^d} \sum_{n \in \mathbb{Z}^d} \|g \cdot \psi_{n-k}\|_\infty \|h \cdot \psi_n\|_\infty = \|g\|_W \|h\|_W.
\]

The lemma implies immediately that
\[
\|\alpha^{(h)}\|_{\ell^1} \leq \sum_{r,s,\lambda} |N_h(\varphi^s \cdot T_\lambda \varphi^r)| \leq C_M \sum_{r,s,\lambda} \|\varphi^s \cdot T_\lambda \varphi^r\|_W \leq C_M \cdot C_\Lambda \cdot k^2 \cdot \max_s \|\varphi^s\|_W^2.
\]
This estimate also implies that only finitely many coefficients have to be calculated approximately (which is possible by choosing $h$ small enough, according to the assumptions) in order to achieve a good $\ell^1$ approximation to $\alpha$, using a method $N_h$ with $h$ small enough.

The second problem arises when one has to determine a matrix-valued function on $\mathbb{R}^d / \mathbb{Z}^d$, which cannot be done literally in practice, because this fundamental domain is having “continuous variables”. Even a countable number of matrix-inversions would not help in the strict sense. Since the entries of these $k \times k$ matrices are in the Fourier algebra it would make sense to sample it regularly over the compact domain and apply some kind of (quasi-)interpolation to the result afterwards (cf. [26]). However, since this matrix inversion over the compact domain is followed by a Fourier transform anyway it is better to carry out the related procedure (i.e. periodization instead of sampling) on the time-side. It is about replacing the infinite discrete problem (at the coefficient level) by a corresponding finite model. In order to justify this step we will need the following proposition: The reduction from the infinite case to a finite case is done problem (at the coefficient level) by a corresponding finite model. In order to justify this step we will need the following proposition: The reduction from the infinite case to a finite case is done via periodization. Since there only a more cumbersome notation in the vector-valued case we formulate the principle in the setting of scalar-valued functions: We define for via periodization. Since there only a more cumbersome notation in the vector-valued case we will need the following proposition: The reduction from the infinite case to a finite case is done problem (at the coefficient level) by a corresponding finite model. In order to justify this step we will need the following proposition: The reduction from the infinite case to a finite case is done via periodization. Since there only a more cumbersome notation in the vector-valued case we formulate the principle in the setting of scalar-valued functions: We define for $x \in \ell^1(\mathbb{Z}^d)$ and an integer $N$ the $N-$periodization $P_x = P^N(x)$ via $P_x(k) = \sum_{l \in \mathbb{Z}^d} x_{k+l}$, $k = 0 \leq l \leq N - 1$. It describes the periodization of sequences along the subgroup $H = n \cdot \mathbb{Z}^d$. In the general notation of ([42]), the mapping $P^N$ corresponds to canonical projection mapping $T_H$ described in Theorem 3.5.4 (page 106).

**Proposition 1.** (i) The mapping $x \mapsto P^N(x)$ is an algebra homomorphism, i.e. it respects convolutions in $\ell^1(\mathbb{Z}^d)$ and $\ell^1(\mathbb{Z}^d_N)$ respectively (we use $*_N$ for convolution in $\ell^1(\mathbb{Z}^d_N)$).

(ii) Assume that $x \in \ell^1(\mathbb{Z}^d)$ defines an invertible element in the Banach algebra $\ell^1(\mathbb{Z}^d)$, with inverse $y$, i.e. such that $x * y = \delta_0$. Then $P^N(x)$ is invertible in $\ell^1(\mathbb{Z}_N)$ for every $N$ and $P^N(y) = (P^N(x))^{-1}$.

Written out in detail the first claim is that we have

$$P^N(x * y) = P^N(x) *_N P^N(y) \quad \text{for} \quad x, y \in \ell^1(\mathbb{Z}^d).$$

The second statement implies that $P^N(y)$ can be obtained using FFT-based calculations. It remains to ensure that one can fabricate the truly (typically infinite) inverse $y$ (up to a small $\ell^1-$error) from these periodizations. This is ensured by the next lemma. Of course it also allows to reduce the costs of inversion of huge convolution matrices to those of smaller size. Corresponding numerical results showing the usefulness of this approach are given in the last section. We do not discuss the asymptotic regime here, because it is not so relevant for the numerical realizability problem in the focus of our paper. Note also that the reduction of computational costs is more significant for higher dimensions.

Thus we need as a final justification for the reduction to the finite case a lemma that shows how to recover $y$ from $P^N(y)$:

**Lemma 2.** Write $Q^M$ for the localization operator in $\ell^1(\mathbb{Z}^d)$, which maps $x$ onto the vector $Q^M x$ which has the same coordinates $x_k$, if $|k| \leq M$, and are otherwise zero. Then one has $Q^M P^N x \to x$ in the norm of $\ell^1$, if $M \leq N/2$ and $M \to \infty$ (hence also $N \to \infty$). In particular $\lim_{M \to \infty} Q^M P^M x = x$. 

17
Proof. The key observations needed to verify the claim of the lemma are the fact that obviously
(i) $\|x - Q^M x\|_1 \to 0$ for $M \to \infty$, and that on the other hand
(ii) within the coordinates $\{k \mid |k| \leq M\}$ the tails of the $N$-periodic copies of $x$ add up to a tail
which is (altogether) in the $\ell^1$-norm smaller than $\sum_{|k| \geq N} |x_k|$. For this to be valid we use the
assumption $M \leq N/2$, so that each coordinate with $|k| \geq M$ appears at most once (in the tail
of one of the shifted copies by a multiple of $N$). This last expression tends again to zero for
$N \to \infty$. □

In the next section we will give experimental evidence of this situation.

9 Testing the algorithm and construction errors

We have done extensive testing of the algorithm both concerning the approximation of matrices by
generalized Gabor multipliers (to be reported elsewhere) and on concrete multi-window families
of atoms, arising in the context of Gabor analysis. Our approach allows among others to treat
Gaborian Riesz basic sequences where the frequencies are not coming at a regular distance and
not being of infinitely large frequency content (thus the case is not part of standard Gabor
theory). Among others we have choosen for our tests shift-invariant subfamilies of Gaborian
Riesz bases (hence the Riesz condition is clearly satisfied; Gabor theory provides a number of
test for this condition). As an example one could take a Gauss function on $\mathbb{R}$, with $\Lambda = a\mathbb{Z}$ and
as additional windows modulated Gaussians of the form $M_{kb}$, $k = +/-1, 2, \ldots, \ell$ for any choice
with $a \cdot b > 1$.

The viewpoint of considering Gabor families as multi-window spline-type families has been
already useful for theoretical purposes (cf. [38]) and allows for additional freedom (e.g. by using
different windows for different frequency channels, which is typical case not covered by the
classical approach to regular Gabor families). Also, for concrete applications one may not want
to go to arbitrary high frequencies and rather work with frequency channels up to a certain
maximal frequency, which again leads to a multi-window situation which is not within the scope
of Gabor analysis. Also slightly irregular spacing in the frequency domain could be covered in
this way, but because it does not influence the performance we stick to the regular case in our
experiments.

On one side we use algorithm 2 in order to compute the matrix contributions to the dual
system associated with a local set of lattice points and we will compare the results with the
standard version of computing the dual system for a multi-window spline-type family using
algorithm 1. To form the test family, we start from a normalized discrete Gaussian function $g$
of length $n$ (actually considered as a function on $\mathbb{Z}_n$), and take the sublattice of $\mathbb{Z}_n$ with gapsize $a$
(isomorphic to $\mathbb{Z}_{n/a}$). The multi-window spline-type system can be written in this case as
$\Phi = \{g, M_{s_2}(g), \ldots, M_{s_k}(g)\} = \{g_1, g_2, \ldots, g_k\}$.

From the overall efficiency of the number of operations of our proposed algorithm (namely
algorithm 2) we can make a direct count estimation as follows. To establish the convolution
system we have to perform $k^2/2$ FFT-operations of size $n$ and considering the order of complexity
of the Fast Fourier transform it can be estimated by the order $k^2 \times n \times \log n$. The next step of
diagonalization under the Fast Fourier transform is done for the $k^2$ windows using a small size FFT of the reduced order $\mathcal{A}$, resulting in a complexity order of $k^2 \times n \times \log n/a$ for this step. Accounting for the inversion of the "small Gramians" we have complexity $k^3 \times \mathcal{A}$ and finally the cost of obtaining the dual system from the small contributions is of the order $k^2 \times n \times \log n/a$. For typical cases the dominant term is of the form $k^2 \times n \times \log(n)$, which is already much less than $n^2$. In other words, our method is far from squaring the number of operations with the factor $n$ which is the signal length as it happens in the regular case when inverting the huge Gramian matrix (algorithm 1).

On the other side we want to provide quantitative information about the error described in Lemma 2, due to the replacement of infinite (or just very large, in order to be able to verify it numerically) convolution matrices by finite ones. As expected we observe that the error is modest if the folding factor is not too large, i.e. we save computational time without significant deterioration of the result. Details are given in the table below. We collect the errors of construction for both the approximate dual atom (in the usual Euclidean distance) and the coefficients obtained by the localized inverse of the Gramian for a signal of norm 1, by varying the folding factor named $\text{Factor}$, the signal size named $n$ and the lattice gap named $a$. In order to have comparable conditions the shape of the atom was kept fixed throughout the experiment (some discrete variant of the Gauss function).

<table>
<thead>
<tr>
<th>$\text{Factor} = 2$</th>
<th>$\text{Factor} = 4$</th>
<th>$\text{Factor} = 8$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Dual error</td>
<td>Coef. error</td>
</tr>
<tr>
<td>$n=480, a=10$</td>
<td>6.8291e-04</td>
<td>4.2745e-04</td>
</tr>
<tr>
<td>$n=480, a=12$</td>
<td>8.1635e-05</td>
<td>6.6041e-05</td>
</tr>
<tr>
<td>$n=480, a=20$</td>
<td>2.2164e-08</td>
<td>2.2045e-08</td>
</tr>
<tr>
<td>$n=960, a=10$</td>
<td>2.6511e-07</td>
<td>1.6594e-07</td>
</tr>
<tr>
<td>$n=960, a=12$</td>
<td>6.5879e-09</td>
<td>5.3295e-09</td>
</tr>
<tr>
<td>$n=960, a=20$</td>
<td>3.4499e-15</td>
<td>3.4403e-15</td>
</tr>
<tr>
<td>$n=1920, a=10$</td>
<td>4.0573e-14</td>
<td>2.6444e-14</td>
</tr>
<tr>
<td>$n=1920, a=12$</td>
<td>1.4544e-15</td>
<td>2.3839e-15</td>
</tr>
<tr>
<td>$n=1920, a=20$</td>
<td>6.2374e-16</td>
<td>5.9665e-16</td>
</tr>
</tbody>
</table>

When comparing the experimental results collected in the above table we see that the resulting dual system is directly applicable for reconstruction given the small approximation errors coming from the numerical discretization. We can observe that the quality of approximation is stable when increasing the number of samples even though the folding factor is also increasing. Also, as expected, when the lattice density is too low, the folding factor will generate small deviations on the dual approximate construction. The table also indicates that it is not the signal length that should be relevant for the computational effort, but rather the relative broadness of the dual atoms (typically expressed in multiples of the lattice constant $a$).
10 Conclusion

This paper presents a method for a constructive approximate realization of the biorthogonal family in the case of multi-window spline-type spaces with generators in $W(C_0, l^1(\mathbb{R}^d))$ or $S_0(\mathbb{R}^d)$, making use of the structure of the biorthogonal system. The method was efficiently implemented and the computational issues related to its realization were treated in three steps. The first step was a direct algorithm for computing the naive inversion of the full Gramian. In the second step the “continuous problem” is approximated by a corresponding problem over a discrete Abelian group. A key point of the paper is a detailed analysis of the errors arising in the reduction from the infinite discrete case to a corresponding finite model by periodization and folding. As a 2D application of our approach we propose a solution to the best approximation of Hilbert Schmidt operators by generalized Gabor multipliers.

11 Acknowledgement

The research was supported by the project NAHA (Numerical and Applied Harmonic Analysis) funded by the University of Vienna and by the WWTF-project MULAC (Frame Multipliers: Theory and Application in Acoustics; MA07-025) as well as the EUCETIFA Marie Curie Excellence Grant FP6-517154.

Thanks are also going to the anonymous reviewers for suggestion which have helped to improve the paper.

References


23
Authors address:
Hans G. Feichtinger
Darian Onchis-Moaca
Faculty of Mathematics
University Vienna
Nordbergstrasse 15
A-1090 Wien, AUSTRIA