Optimal Adaptive Computations in the Jaffard Algebra and Localized Frames

Stephan Dahlke, Massimo Fornasier, Karlheinz Gröchenig*

May 29, 2008

Abstract

We study the efficient numerical solution of infinite matrix equations $\mathbf{Au} = \mathbf{f}$ for a matrix $\mathbf{A}$ in the Jaffard algebra. These matrices appear naturally via frame discretizations in many applications such as Gabor analysis, sampling theory, and quasi-diagonalization of pseudo-differential operators in the weighted Sjöstrand class. The proposed algorithm has two main features: firstly, it converges to the solution with quasi-optimal order and complexity with respect to classes of localized vectors; secondly, in addition to ℓ^2-convergence, the algorithm converges automatically in some stronger norms of weighted ℓ^p-spaces. As an application we approximate the canonical dual frame of a localized frame and show that this approximation is again a frame, and even an atomic decomposition for a class of associated Banach spaces. The main tools are taken from adaptive algorithms, from the theory of localized frames, and the special Banach algebra properties of the Jaffard algebra.

AMS subject classification: 41A25, 42C15, 46E35, 65F10, 65F20, 65F50, 65N12
Key Words: Adaptive scheme, Jaffard algebra, frames in Banach spaces, best approximation, localization of frames, sparse matrix.

1 Introduction

Fast matrix computations use either structure or sparsity. Structure, as used for the FFT or Toeplitz solvers, is more rigid and works only in very specific applications.

---

*The first two authors acknowledge the financial support provided by the European Union's Human Potential Programme, under contract HPRN–CT–2002–00285 (HASSIP). Both want to thank the Numerical Harmonic Analysis Group at the University of Vienna for the hospitality and support. Stephan Dahlke is supported by the DFG Grant Da 360/4–3. Massimo Fornasier acknowledges the financial support provided by the European Union's Human Potential Programme under contract MOIF-CT-2006-039438. Karlheinz Gröchenig is supported by the Marie-Curie Excellence Grant MEXT-CT-2005-517154.

Sparsity is more flexible and arises often in the discretization of operator equations with respect to a suitable basis. Roughly speaking, a matrix is sparse, if each row and each column contain only few non-zero entries (or few large entries). Likewise, a vector is sparse, if it has only few non-zero (or large) coefficients. The resulting matrix-vector multiplication is cheap because the operation count is determined by the number of large entries of the matrix and the vector. This observation is the key to the recent development of adaptive algorithms for the solution of infinite matrix equations by Cohen, Dahmen, and DeVore [8, 9].

Such adaptive numerical methods have been applied successfully for the solution of operator equations, in particular to PDE and integral equations. The existing stable and efficient implementations are mostly based on the discretization with wavelet bases. The innovation brought by the use of wavelet bases was the rigorous analysis of the complexity of the algorithms. One of the main results guarantees that the adaptive algorithm of [8] converges with the optimal order and optimal numerical complexity. However, for PDE on bounded domains or on a closed manifold the adaptive wavelet method faces a serious limitation: the construction of a suitable wavelet basis on domains is rather intricate, and the known constructions either have stability problems or lack sufficient smoothness.

These difficulties have motivated the use of (wavelet) frames instead of bases in adaptive schemes [11, 12, 13, 25]. Frames provide stable and redundant (non-orthogonal) expansions in a Hilbert space. In general, a wavelet frame on a domain is much easier to construct than a wavelet basis. However, by using frames, a new problem arises: the resulting stiffness matrix may be singular, and at first glance one has to solve a singular equation. This problem was settled in [12, 25], where it was shown that the adaptive strategies developed in [8, 9] can be generalized to the frame case and maintain their advantages.

In this paper we pursue further the investigation of adaptive numerical strategies for the solution of operator equations via frame discretization. Again, we deal with a form of sparse operator equations. The first innovation is the chosen measure of sparsity. Whereas the adaptive wavelet schemes work with matrices in the Lemarié algebra [12, 24], in this paper we investigate the analogous situation for the Jaffard algebra [23]. The sparsity of a matrix in the Jaffard algebra is given by the rate of its (polynomial) off-diagonal decay. This setting arises quite natural in many applications, notably in time-frequency analysis [18], sampling theory [1], and in the discretization and almost-diagonalization of pseudo-differential operators in the weighted Sjštrand class [20]. Such operators play a fundamental role in modelling wireless transmission channels in mobile communication [26]. In this paper, we show that the principal subroutines of adaptive algorithms also work for the Jaffard algebra. We derive suitable schemes to approximate an infinite vector by a finite one, and we provide an implementable algorithm for the approximation of an infinite matrix-vector product. The resulting numerical scheme for the solution of infinite matrix equations is then guaranteed to converge with quasi-optimal order and operation count. The algorithm is again *adaptive*, in the sense that it automatically refines the approximation space and enlarges the number of degrees of freedom according to the current prescribed accuracy.

Our second innovation is the application of the theory of localized frames as developed in [15, 19]. This theory provides a powerful tool for the analysis of the dual frame, for series expansions in associated Banach spaces, and for the extension of frames to Banach spaces. A frame is *localized* (more precisely, self-localized), if its Gramian matrix is in the Jaffard algebra. To our knowledge, the combination of localized frames and adaptive algorithms is new.

Our analysis of adaptive algorithms with localized frames differs in several aspects from the wavelet case. Although we could not prove that the algorithm is fully optimal – we establish only its quasi-optimality – the special structure of the Jaffard algebra allows us to prove some surprising results:

- The approximation of infinite vectors can be performed by a nearest neighborhood approximation. As a consequence, no sorting routines or binary binning strategies are needed;
- The approximation of a stiffness matrix in the Jaffard algebra is given by a banded matrix and thus much simpler than the approximation by a compressible matrix;
- The optimality of the adaptive scheme requires that a certain orthogonal projection is also bounded on weighted ℓ^p-spaces or on weak ℓ^p-spaces. This property is automatically satisfied when working with localized frames and the Jaffard algebra, whereas it has to be postulated as an additional assumption in the case of wavelet frames and the Lemarié algebra, see [25, Thm. 3.12].
- The adaptive algorithm converges not only in the underlying Hilbert space, but also in a scale of stronger Banach space norms. In concrete examples these stronger norms imply the convergence of derivatives and convergence in weighted L^p-spaces. The automatic convergence of the adaptive algorithm in finer norms (Theorem 3.13) is surprising, and to our knowledge, is the first result of this type. It is unclear whether this stronger form of convergence also holds in the case of wavelet frames and the Lemarié algebra.

Let us emphasize that at the heart of our results is a special Banach algebra property of the Jaffard algebra. The key is that the Jaffard algebra is closed under taking inverses [23], whereas the Lemarié algebra lacks this property.

As an important application of the new adaptive strategies, we investigate the computation of the canonical dual frame. The canonical dual is necessary for computing the coefficients of a frame expansion. Each vector of the canonical dual is defined implicitly by an operator equation involving the frame operator. Therefore the properties of the dual frame are often hard to check and usually no explicit formulas are available. Computational issues about the dual frame are investigated in [3, 4] and in [7] (by means of the *finite section method* and localization properties of the frame), see also [22] for a quantitative study on the performances of this method.

We apply the adaptive algorithm to the discretization of the frame operator, where the discretization may be with respect to a different frame. For a suitably localized frame, the corresponding stiffness matrix is in the Jaffard algebra, and thus the adaptive algorithm yields an efficient approximation of each element of the dual frame. Our main result (Theorem 5.2) asserts that the approximation of the dual frame is again a frame and that this approximation works in much finer norms (involving decay and smoothness conditions). These results are far from obvious and require the entire machinery of adaptive methods and localized frame theory.

This paper is organized as follows. In Section 2, we discuss the frame setting as far as it is needed for our purposes. Special emphasis is layed on Banach frames and localization properties. Section 3 is concerned with matrix computations in the Jaffard algebra. First we derive a subroutine to approximate infinite vectors by finite ones. Then we describe an algorithm to compute finite vectors that approximate infinite matrix-vector products up to a given precision. The combination of these subroutines yields an adaptive algorithm **SOLVE** for the numerical solution of infinite matrix equations in the Jaffard algebra. We carry out a detailed analysis of the convergence and complexity of this algorithm in Thms. 3.11 and 3.13. In Section 4, we deal with the efficient computation of the canonical dual frame by means of **SOLVE**. Finally, in Section 5, we present the error estimates that guarantee that the approximated elements of the canonical dual again form a frame.

Throughout this paper '$a \lesssim b$' means that there exists a positive constant C such that $a \leq Cb$. If $a \lesssim b$ and $b \lesssim a$ then we will write $a \asymp b$. We determine the constants explicitly only if their value is crucial for further analysis. The expression $\mathcal{C}(A)$ stands for the number of algebraic operations needed to compute the quantity A. Often we will use C, κ with subscripts or superscripts to indicate some positive relevant constants. By $L(\mathcal{B})$ we denote the space of bounded linear operators on a Banach space $\mathcal{B}$.

2 Intrinsically Localized Frames in Banach Spaces

2.1 Frames in Hilbert and Banach Spaces

In this section we recall the concept of frames. Frames provide stable and redundant nonorthogonal expansions in Hilbert spaces, and they can be used to define certain associated Banach spaces and to obtain stable decompositions in these Banach spaces. The *canonical dual* frame is used to compute the coefficients of such expansions and plays a pivotal role in the theory of Banach spaces associated to frames and in many concrete applications.

In the following we assume that the index set is $\mathcal{N} = \mathbb{Z}^d$. This is no loss of generality, because we can map any relatively separated set of $\mathbb{R}^d$ into $\mathbb{Z}^d$ by a trick in [2]. A subset $\mathcal{G} = \{g_n : n \in \mathcal{N}\}$ of a separable Hilbert space $\mathcal{H}$ is called a *frame* for $\mathcal{H}$, if

$$A_{\mathcal{G}}\|f\|^2 \leq \sum_{n\in\mathcal{N}} |\langle f, g_n\rangle|^2 \leq B_{\mathcal{G}}\|f\|^2, \quad \text{for all} \quad f \in \mathcal{H}, \tag{1}$$

for some constants $0 < A_{\mathcal{G}} \leq B_{\mathcal{G}} < \infty$. Associated to a frame are the following bounded operators, namely the analysis operator F and the synthesis operator F^* defined by

$$F : \mathcal{H} \to \ell^2(\mathcal{N}), \quad f \mapsto \big(\langle f, g_n\rangle\big)_{n\in\mathcal{N}}, \tag{2}$$

$$F^* : \ell^2(\mathcal{N}) \to \mathcal{H}, \quad \mathbf{c} \mapsto \sum_{n\in\mathcal{N}} c_n g_n. \tag{3}$$

The composition $S := F^*F$ is a boundedly invertible, positive operator on $\mathcal{H}$ called the *frame operator*. The set $\{\tilde{g}_n := S^{-1}g_n : n \in \mathcal{N}\}$ is again a frame for $\mathcal{H}$, the *canonical dual frame*, with corresponding analysis and synthesis operators

$$\tilde{F} = F(F^*F)^{-1}, \quad \tilde{F}^* = (F^*F)^{-1}F^*. \tag{4}$$

In particular, one has the following orthogonal decomposition of $\ell^2(\mathcal{N})$

$$\ell^2(\mathcal{N}) = \operatorname{ran}(F) \bigoplus \ker(F^*), \tag{5}$$

and

$$\mathbf{P} := F(F^*F)^{-1}F^* : \ell^2(\mathcal{N}) \to \operatorname{ran}(F), \tag{6}$$

is the orthogonal projection onto $\operatorname{ran}(F)$. In general $\operatorname{ran}(F) \neq \ell^2(\mathcal{N})$, and $\operatorname{ran}(F) = \ell^2(\mathcal{N})$ if and only if $\mathcal{G}$ is a Riesz basis. From the invertibility of S one has also the following reproducing formulas

$$f = \sum_{n\in\mathcal{N}} \langle f, \tilde{g}_n\rangle g_n = \sum_{n\in\mathcal{N}} \langle f, g_n\rangle \tilde{g}_n, \quad \text{for all } f \in \mathcal{H}. \tag{7}$$

More information on frames can be found in the book [5].

The concept of frame can be extended to Banach spaces as follows.

Definition 1 ([17])**.** A *Banach frame* for a separable Banach space $\mathcal{B}$ is a sequence $\mathcal{G} = \{g_n : n \in \mathcal{N}\}$ in $\mathcal{B}'$ with an associated sequence space $\mathcal{B}_d$ such that the following properties hold.

(a) Norm equivalence:

$$\|f\|_{\mathcal{B}} \asymp \|\langle f, g_n\rangle_{n\in\mathcal{N}}\|_{\mathcal{B}_d}, \text{ for all } f \in \mathcal{B}.$$

(b) There exists a bounded operator R from $\mathcal{B}_d$ onto $\mathcal{B}$, a so-called *synthesis or reconstruction operator*, such that

$$R\left(\langle f, g_n\rangle_{n\in\mathcal{N}}\right) = f, \text{ for all } f \in \mathcal{B}.$$

A dual concept and a different extension of Hilbert frames to Banach spaces is given by the notion of atomic decomposition.

Definition 2. An *atomic decomposition* for a separable Banach space $\mathcal{B}$ consists of a pair of sets $\mathcal{G} = \{g_n : n \in \mathcal{N}\}$ in $\mathcal{B}$ and $\tilde{\mathcal{G}} = \{\tilde{g}_n : n \in \mathcal{N}\}$ in $\mathcal{B}'$ and an associated sequence space $\mathcal{B}_d$ such that the following properties hold.

(a) Norm equivalence:

$$\|f\|_{\mathcal{B}} \asymp \|\langle f, \tilde{g}_n\rangle_{n\in\mathcal{N}}\|_{\mathcal{B}_d}, \text{ for all } f \in \mathcal{B}.$$

(b) The series expansion for the reconstruction of f,

$$f = \sum_{n\in\mathcal{N}} \langle f, \tilde{g}_n\rangle g_n\,,$$

converges unconditionally for all $f \in \mathcal{B}$.

2.2 Discretization of Operator Equations by Frames

In this subsection, we explain how frames can be used for the numerical treatment of operator equations

$$\mathcal{L}u = f, \tag{8}$$

where $\mathcal{L}$ denotes a boundedly invertible linear operator on $\mathcal{H}$. We want to solve (8) approximately with the aid of a suitable numerical scheme based on frames. A natural idea is the use of a Galerkin scheme. There one chooses a finite subset of frame elements, considers their span $V \subset \mathcal{H}$, and searches for $u_V \in V$ such that

$$\langle \mathcal{L}u_V, v\rangle = \langle f, v\rangle, \qquad \text{for all} \quad v \in V. \tag{9}$$

However, this standard approach may face serious problems, because the stiffness matrix corresponding to (8) may be singular in the frame case. Nevertheless, it is possible to transform (8) into an equivalent bi-infinite matrix equation on $\ell^2(\mathcal{N})$ and to derive a series representation for the solution as we shall explain now. The following lemma has been proved in [12], see also [11, 13, 25].

Lemma 2.1. *If $\mathcal{L}$ is a self-adjoint invertible operator on $\mathcal{H}$, then the operator*

$$\mathbf{A} := F\mathcal{L}F^* \tag{10}$$

is a bounded operator from $\ell^2(\mathcal{N})$ to $\ell^2(\mathcal{N})$. Moreover $\mathbf{A} = \mathbf{A}^*$ and it is boundedly invertible on its range $\operatorname{ran}(\mathbf{A}) = \operatorname{ran}(F)$.

In principle, under suitable assumptions, a linear system $\mathbf{A}\mathbf{u} = \mathbf{f}$ can be solved by a simple Richardson-Landweber iteration, as we show in the following.

Theorem 2.2. *Let $\mathcal{L}$ be a boundedly invertible, positive operator on $\mathcal{H}$. Let $\mathbf{A}$ be as in* (10) *and denote*

$$\mathbf{f} := Ff\,. \tag{11}$$

Then the solution u of (8) *can be computed by*

$$u = F^*\mathbf{u} \tag{12}$$

where $\mathbf{u}$ is given by

$$\mathbf{u} = \left(\alpha \sum_{n=0}^{\infty} (\mathrm{id} - \alpha\mathbf{A})^n\right)\mathbf{f}, \tag{13}$$

with $0 < \alpha < 2/\lambda_{\max}$ and $\lambda_{\max} = \|\mathbf{A}\|$.

Observe that (13) is simply a damped Richardson iteration

$$\begin{aligned} \mathbf{u}^{(n+1)} &= \mathbf{u}^{(n)} - \alpha(\mathbf{A}\mathbf{u}^{(n)} - \mathbf{f}), \quad n \geq 1, \\ \mathbf{u}^{(0)} &= 0, \\ \mathbf{u} &= \lim_{n\to+\infty} \mathbf{u}^{(n)}. \end{aligned} \tag{14}$$

Clearly (14) cannot be implemented directly since it involves infinite vectors and bi-infinite matrices. Nevertheless, an implementable numerical scheme can be derived by approximating the bi-infinite matrices and vectors in (13) by finite ones. This issue will be discussed in Section 3.

REMARK: According to Theorem 2.2 we have to compute (13) on the range of $\mathbf{A}$. However, if we perturb (13) by approximating the bi-infinite matrices and the infinite vectors, then the resulting vectors will have components in the kernel of $\mathbf{A}$. Nevertheless, since $\ker(\mathbf{A}) = \ker(F^*)$ by (10), the iteration will still converge if the *projected* error onto $\mathrm{ran}(\mathbf{A})$ tends to zero.

2.3 Intrinsic Localized Frames and Associated Banach Spaces

The concept of *localized frames* has been recently introduced and investigated in [2, 10, 14, 15, 16, 19, 21] as a tool for extending a frame for a Hilbert space to a Banach frame (or an atomic decomposition) for a family of *associated Banach spaces*. The localization is a measure of the *sparseness* of a frame and is defined by the off-diagonal decay of the Gramian matrix of the frame. We first recall the concept of mutual localization of two frames and then the necessary results from Banach algebra theory.

In this paper we work with the *Jaffard algebra* [23] which is defined as the class of matrices $\mathbf{A} = (a_{kl}), k, l \in \mathcal{N}$, such that

$$|a_{kl}| \lesssim (1 + |k - l|)^{-\gamma} \qquad \text{for all} \quad k, l \in \mathcal{N}, \quad \gamma > d.$$

We denote the Jaffard algebra by $\mathcal{A} := \mathcal{A}_\gamma$ and endow it with the norm

$$\|\mathbf{A}\|_{\mathcal{A}_\gamma} := \sup_{k,l\in\mathcal{N}} |a_{kl}|(1+|k-l|)^\gamma.$$

One can show [19, 23] the following properties:

(A0) If $\gamma > d$, then $\mathcal{A} \subseteq L(\ell^2(\mathcal{N}))$, i.e., each $\mathbf{A} \in \mathcal{A}$ defines a bounded operator on $\ell^2(\mathcal{N})$.

(A1) If $\mathbf{A} \in \mathcal{A}$ is invertible on $\ell^2(\mathcal{N})$, then $\mathbf{A}^{-1} \in \mathcal{A}$ as well. In the language of Banach algebras, $\mathcal{A}$ is *inverse-closed* in $L(\ell^2(\mathcal{N}))$.

(A2) $\mathcal{A}$ is solid: i.e., if $\mathbf{A} \in \mathcal{A}$ and $|b_{kl}| \leq |a_{kl}|$ for all $k, l \in \mathcal{N}$, then $\mathbf{B} \in \mathcal{A}$ as well.

We refer to [21] where several other examples of algebras with properties (A0-2) are presented. Let us denote by $w_\gamma(x) = (1 + |x|)^\gamma$ the polynomially growing, submultiplicative, and radial symmetric weight function on $\mathbb{R}^d$. A weight m on $\mathbb{R}^d$ is called γ-*moderate* if $m(x + y) \leq w_\gamma(x)m(y)$. In particular, if m is γ-moderate, then m^{-1} is also γ-moderate and both $m(x)^{-1} \lesssim w_\gamma(x)$ and $m(x) \lesssim w_\gamma(x)$ for all $x \in \mathbb{R}^d$.

Definition 3. Given two frames $\mathcal{G} = \{g_n : n \in \mathcal{N}\}$ and $\mathcal{F} = \{f_x : x \in \mathcal{N}\}$ for the Hilbert space $\mathcal{H}$, the (cross-) Gramian matrix $\mathbf{A} = A(\mathcal{G}, \mathcal{F})$ of $\mathcal{G}$ with respect to $\mathcal{F}$ is the $\mathcal{N} \times \mathcal{N}$-matrix with entries

$$a_{xn} = \langle g_n, f_x\rangle.$$

The frame $\mathcal{G}$ for $\mathcal{H}$ is called $\mathcal{A}$-*localized* with respect to the frame $\mathcal{F}$ if $A(\mathcal{G}, \mathcal{F}) \in \mathcal{A}$. In this case we write $\mathcal{G} \sim_{\mathcal{A}} \mathcal{F}$. If $\mathcal{G} \sim_{\mathcal{A}} \mathcal{G}$, then $\mathcal{G}$ is called $\mathcal{A}$-*self-localized* or *intrinsically* $\mathcal{A}$-*localized.*

Intrinsic localization of frames is a very powerful concept and is essential for the following general principle which has been shown in [15, Corollary 3.7].

Theorem 2.3. *Let $\mathcal{G}$ be a frame for $\mathcal{H}$ and let $\gamma > d$. If the Gramian of $\mathcal{G}$ satisfies the condition*

$$|\langle g_k, g_l\rangle| \leq C w_\gamma(k-l)^{-1} \quad \textit{for all} \quad k, l \in \mathcal{N},$$

then the Gramian of the dual frame $\tilde{\mathcal{G}}$ also satisfies

$$|\langle \tilde{g}_k, \tilde{g}_l\rangle \leq C' w_\gamma(k-l)^{-1} \quad \textit{for all} \quad k, l \in \mathcal{N},$$

and

$$|\langle \tilde{g}_k, g_l\rangle| \leq C' w_\gamma(k-l)^{-1} \quad \textit{for all} \quad k, l \in \mathcal{N}.$$

More generally, if $\mathcal{G}$ is $\mathcal{A}$-self-localized, then $\widetilde{\mathcal{G}}$ is also $\mathcal{A}$-self-localized and $\widetilde{\mathcal{G}} \sim_{\mathcal{A}} \mathcal{G}$.

REMARK: Since the canonical dual frame $\tilde{\mathcal{G}}$ is defined implicitly by the equations

$$S\tilde{g}_n = g_n, \quad n \in \mathcal{N}\,, \tag{15}$$

it is usually difficult to check the properties of $\tilde{\mathcal{G}}$ and almost impossible to derive explicit formulas for $\tilde{g}_n$. Theorem 2.3 provides some control of the dual frame and lies at the heart of the efficient and implementable methods for the approximation of $\tilde{\mathcal{G}}$.

Next we illustrate how certain families of Banach spaces can be characterized by $\mathcal{A}$-self-localized frames. In the following we assume that $\gamma > s + d$ and m is an s-moderate weight.

Let $(\mathcal{G}, \tilde{\mathcal{G}})$ be a pair of dual $\mathcal{A}$-self-localized frames for $\mathcal{H}$, and assume $1 \le p \le \infty$:

- if $\ell^p_m(\mathcal{N}) \subset \ell^2(\mathcal{N})$, then the Banach space $\mathcal{H}^p_m(\mathcal{G}, \tilde{\mathcal{G}})$ is defined to be
$$\mathcal{H}^p_m(\mathcal{G}, \tilde{\mathcal{G}}) := \{f \in \mathcal{H} : \quad f = \sum_{n \in \mathcal{N}} \langle f, \tilde{g}_n\rangle g_n, \quad (\langle f, \tilde{g}_n\rangle)_{n\in\mathcal{N}} \in \ell^p_m(\mathcal{N})\} \tag{16}$$
with the norm
$$\|f\|_{\mathcal{H}^p_m} = \|(\langle f, \tilde{g}_n\rangle)_{n\in\mathcal{N}}\|_{\ell^p_m}\,.$$
Since $\ell^p_m(\mathcal{N}) \subset \ell^2(\mathcal{N})$, $\mathcal{H}^p_m$ is a dense subspace of $\mathcal{H}$;
- if $\ell^p_m(\mathcal{N})$ is not included in $\ell^2(\mathcal{N})$ and $1 \le p < \infty$, then we define $\mathcal{H}^p_m$ to be the completion of the subspace $\mathcal{H}_0$ of all finite linear combinations in $\mathcal{G}$ with respect to the $\mathcal{H}^p_m$-norm;
- if $p = \infty$, then we take the weak*-completion of $\mathcal{H}_0$ to define $\mathcal{H}^\infty_m$.

REMARK: The definition of $\mathcal{H}^p_m(\mathcal{G}, \tilde{\mathcal{G}})$ does not depend on the particular $\mathcal{A}$-self localized dual chosen, and any other $\mathcal{A}$-self-localized frame $\mathcal{F}$ which is localized with respect to $\mathcal{G}$ generates the same space with an equivalent norm. For more details we refer to [15], from which also the following theorem is taken.

Theorem 2.4. *Assume that $\mathcal{G}$ is an $\mathcal{A}_\gamma$-self-localized frame for $\mathcal{H}$ for some $\gamma > s+d$. Then both $\mathcal{G}$ and and its canonical dual frame $\tilde{\mathcal{G}}$ form a Banach frame for $\mathcal{H}^p_m(\mathcal{G}, \tilde{\mathcal{G}})$ for $1 \le p \le \infty$ and every s-moderate weight m. Moreover, for the same range of parameters, the pair $(\mathcal{G}, \tilde{\mathcal{G}})$ yields an atomic decomposition of $\mathcal{H}^p_m$ with sequence space ℓ^p_m.*

3 Matrix Computations in the Jaffard Algebra

In this section, we want to discuss the basic subroutines required for the approximate numerical solution of the system of equations

$$\mathbf{Au} = \mathbf{f}\,. \tag{17}$$

As already indicated in Subsection 2.2, this task requires the approximation of infinite vectors and bi-infinite matrix-vector products by finite ones. The first issue is addressed in Subsection 3.1 and is settled by the N–nearest neighborhood approximation. The second problem will be discussed in Subsection 3.3 where we derive a subroutine for the computation of a finitely supported vector $\mathbf{w}_\varepsilon$ such that

$$\|\mathbf{w}_\varepsilon - \mathbf{A}\mathbf{v}\|_{\ell^2} \leq \varepsilon. \tag{18}$$

Finally, in Subsection 3.4 we combine these building blocks and obtain a numerical scheme that converges with optimal order.

REMARK: The occurrence of the many parameters γ, r, s, t, etc. is unavoidable and requires some clarification. First, γ parametrizes the off-diagonal decay of a matrix and can be understood as a measure for sparsity. The parameter s indicates the decay of an infinite vector and serves as a measure for the localization. Usually s depends on γ, the common hypothesis is $s + d < \gamma$. The parameter r is a measure for the complexity of an algorithm and occurs in the operation count. It is always given by $r = s/d - 1/2$. Finally, for the convergence of iterative algorithms we will use weighted ℓ^p-spaces. In this context the parameter t measures the maximal growth of the admissible weight, t depends on s and γ.

3.1 Nearest Neighborhood Approximation

In this section, we introduce the sequence spaces and the approximation schemes that are needed for our purpose.

Let us start by clarifying our notion of an optimal numerical algorithm. Let $\mathbf{V} \subset \ell^2$ be a normed vector space. Assume that there exists an r such that every $\mathbf{v} \in \mathbf{V}$ possesses a finite approximation $\mathbf{v}_\varepsilon \in \mathbf{V}$ with the properties:

(a) $\|\mathbf{v} - \mathbf{v}_\varepsilon\|_{\ell^2} \leq \varepsilon$;

(b) $\#\operatorname{supp}(\mathbf{v}_\varepsilon) \lesssim \varepsilon^{-1/r}\|\mathbf{v}\|_{\mathbf{V}}^{1/r}$.

Clearly, the larger r, the smaller the support of $\mathbf{v}_\varepsilon$. We will denote the maximal exponent that works for all $\mathbf{v} \in \mathbf{V}$ by $r = r(\mathbf{V})$. Then a numerical scheme will be called *optimal*, if it produces an approximation $\overline{\mathbf{v}}_\varepsilon$ with the properties (a) and (b) and with computational costs satisfying

(c)
$$\mathcal{C}(\overline{\mathbf{v}}_\varepsilon) \lesssim \varepsilon^{-1/r}\|\mathbf{v}\|_{\mathbf{V}}^{1/r}.$$

Let us now introduce the sequence spaces $\ell^\infty_{w_s}$.

Definition 4. For $s > d$, $x \in \mathcal{N}$, and for $v \in \ell^\infty_{w_s}$ we define the norm

$$\|\mathbf{v}\|_{s,x} := \sup_{k \in \mathcal{N}} |v_k|(1 + |x - k|)^s. \tag{19}$$

Of course, for $x, y \in \mathcal{N}$, $y \neq x$ one has

$$(1 + |x - y|)^{-s} \|\mathbf{v}\|_{s,x} \leq \|\mathbf{v}\|_{s,y} \leq (1 + |x - y|)^{s} \|\mathbf{v}\|_{s,x}. \tag{20}$$

Therefore, the norms $\| \cdot \|_{s,x}$ are equivalent to $\| \cdot \|_{s,0} = \| \cdot \|_{\ell^\infty_{w_s}}$. Nevertheless, we will use this notation to indicate that a vector is "localized".

Definition 5. A vector $\mathbf{v} \in \ell^2(\mathcal{N})$ is *s-localized* at $x \in \mathcal{N}$, if $\mathbf{v} \in \ell^\infty_{w_s}(\mathcal{N})$ and

$$\|\mathbf{v}\|_{s,x} = \min_{y \in \mathcal{N}} \|\mathbf{v}\|_{\ell_{s,y}}. \tag{21}$$

In $\ell^\infty_{w_s}$, we consider the following approximation scheme.

Definition 6. Given a vector $\mathbf{v} \in \ell^\infty_{w_s}$ localized at x, we define its *N-nearest neighborhood approximation* by

$$v_k^{N-nearest} := \begin{cases} v_k & , \ |k - x| \leq N \\ 0 & , \ \text{otherwise.} \end{cases}$$

By a small computation, we have

$$\|\mathbf{v} - \mathbf{v}^{N-nearest}\|_{\ell^2} \lesssim \|\mathbf{v}\|_{s,x} N^{d/2-s}.$$

Given $\varepsilon > 0$, set $r = \frac{s}{d} - \frac{1}{2}$, $N = (\|\mathbf{v}\|_{s,x} \varepsilon^{-1})^{\frac{1}{dr}}$, and $\mathbf{v}_\varepsilon = \mathbf{v}^{N-nearest}$. Then

$$\|\mathbf{v} - \mathbf{v}_\varepsilon\|_{\ell^2} \leq \varepsilon \quad \text{and} \quad \# \operatorname{supp}(\mathbf{v}_\varepsilon) \lesssim \varepsilon^{-1/r} \|\mathbf{v}\|_{s,x}^{1/r}.$$

Moreover, since there is no need of algebraic operations to compute $\mathbf{v}_\varepsilon$ the computational cost can be assumed constant and certainly $\mathcal{C}(\mathbf{v}^\varepsilon) \lesssim \varepsilon^{-1/r} \|\mathbf{v}\|_{s,x}^{1/r}$. Consequently, the nearest neighborhood approximation gives rise to an optimal approximation for vectors in $\ell^\infty_{w_s}$, provided that we have clarified that $r = \frac{s}{d} - \frac{1}{2}$ is really the maximal choice for the rate of approximation, as we shall now explain.

As an alternative to the N-nearest neighborhood approximation we consider the *best M-term approximation* of a vector $\mathbf{v} \in \ell^2$. Let $\mathbf{v}^{M-best}$ be the vector of the M coefficients of $\mathbf{v}$ that are largest in modulus (or, equivalently, the first M coefficients of its non-increasing rearrangement $\gamma(\mathbf{v})$). If $\mathbf{v} \in \ell^\infty_{w_s}$ and $M = \#\{k : |k - x| \leq N\}$, then clearly

$$\|\mathbf{v} - \mathbf{v}^{M-best}\|_{\ell^2} \leq \|\mathbf{v} - \mathbf{v}^{N-nearest}\|_{\ell^2}, \quad \text{where} \quad \#\text{supp}\mathbf{v}^{N-nearest} = \#\text{supp}\mathbf{v}^{M-best} \asymp N^d.$$

The M-term approximation is related with the weak ℓ^τ-spaces. Let $\gamma_n(\mathbf{v})$ be the n-th term of a non-increasing rearrangement of $\mathbf{v}$ and $0 < \tau < 2$. Then the space $\ell^{\tau,w}(\mathcal{N})$ is defined by

$$\ell^{\tau,w}(\mathcal{N}) := \{\mathbf{v} \in \ell^2(\mathcal{N}) : |\mathbf{v}|_{\ell^{\tau,w}} := \sup_{n \in \mathbb{N}} n^{1/\tau} |\gamma_n(\mathbf{v})| < \infty\}. \tag{22}$$

It is easy to verify the following properties of $\ell^{\tau,w}$:

(a) $\|\mathbf{v}\|_{\ell^{\tau,w}}$ is a quasi-norm, i.e., $\|\mathbf{v}+\mathbf{w}\|_{\ell^{\tau,w}} \le C_\tau(\|\mathbf{v}\|_{\ell^{\tau,w}} + \|\mathbf{w}\|_{\ell^{\tau,w}})$ for some constant $C_\tau > 1$;

(b) $\ell^\tau \subset \ell^{\tau,w} \subset \ell^{\tau+\delta}$ for any $\delta \in (0, 2-\tau]$;

(c) if $\tau = (1/2+r)^{-1}$, then

$$|\mathbf{v}|_{\ell^{\tau,w}} \asymp \sup_{M\ge 1} M^r \|\mathbf{v} - \mathbf{v}^{M-best}\|_{\ell^2}.$$

Thus, if $\varepsilon = M^{-r}|\mathbf{v}|_{\ell^{\tau,w}}$ and $\mathbf{v}_\varepsilon = \mathbf{v}^{M-best}$, then

$$\|\mathbf{v}-\mathbf{v}_\varepsilon\|_{\ell^2} \le \varepsilon \quad \text{and} \quad \#\operatorname{supp}\mathbf{v}_\varepsilon = M = (\varepsilon^{-1}|\mathbf{v}|_{\ell^{\tau,w}})^{1/r}.$$

This implies immediately that

$$\ell^\infty_{w_s} \subset \ell^{\tau,w}, \tag{23}$$

for $\tau = (1/2+r)^{-1}$ and $r = \frac{s}{d} - \frac{1}{2}$. Moreover, there exists $\mathbf{v} \in \ell^\infty_{w_s}$, but $\mathbf{v} \notin \ell^{\tilde\tau,w}$ for $\tilde\tau < \tau = (1/2+r)^{-1}$ and $r = \frac{s}{d} - \frac{1}{2}$, for which the best M-coefficient approximation cannot be more efficient than the N-nearest neighborhood approximation. Just consider for example $\mathbf{v} = w_s^{-1}$. Thus the N-nearest neighborhood approximation is optimal.

The discussion above shows that for vectors $\mathbf{v} \in \ell^\infty_{w_s}(\mathcal{N})$ the N–nearest neighborhood approximation provides an implementable, optimal procedure, which we call **RHS** as in [8]. It has the following properties.

RHS$[\varepsilon, \mathbf{v}] \to \mathbf{v}_\varepsilon$: determines for $\mathbf{v} \in \ell^\infty_{w_s}(\mathcal{N})$ a finitely supported $\mathbf{v}_\varepsilon$ such that

(a)
$$\|\mathbf{v}-\mathbf{v}_\varepsilon\|_{\ell^2} \le \varepsilon; \tag{24}$$

(b) $\operatorname{supp}(\mathbf{v}_\varepsilon) \subseteq B(x,N)$ and $\#\operatorname{supp}(\mathbf{v}_\varepsilon) \lesssim N^d \lesssim \varepsilon^{-1/r}\|\mathbf{v}\|_{s,x}^{1/r}$;

(c) $\mathcal{C}(\mathbf{v}_\varepsilon) \lesssim \varepsilon^{-1/r}\|\mathbf{v}\|_{s,x}^{1/r}$.

3.2 Coarsening in $\ell^\infty_{w_s}$

In this section, we introduce the *coarsening* procedure and its properties. This procedure starts with some approximation vector and produces a fines approximation via a nearest neighborhood approximation. The procedure will be a crucial component of an adaptive algorithm for the numerical solution of linear systems with localized matrices. Our analysis is a natural adaptation of the one given in [8] for the coarsening of sparse/compressible vectors.

Lemma 3.1. *Let* $s > d$ *and* $r = \frac{s}{d} - \frac{1}{2}$. *Suppose that* $\mathbf{v} \in \ell^\infty_{w_s}(\mathcal{N})$ *and* $\mathbf{w} \in \ell^2(\mathcal{N})$ *satisfy*

$$\|\mathbf{v} - \mathbf{w}\|_{\ell^2} \leq \varepsilon \tag{25}$$

for $\varepsilon > 0$. *If* $M = (\varepsilon^{-1}\|\mathbf{v}\|_{s,x})^{\frac{1}{rd}}$, *then*

$$\|\mathbf{w} - \mathbf{w}^{M-nearest}\|_{\ell^2} \leq 3\varepsilon, \tag{26}$$

where $\mathbf{w}^{M-nearest}$ *is the best* M*-nearest approximation as introduced in Definition6.*

Proof. The result follows by an application of the triangle inequality,

$$\|\mathbf{w}-\mathbf{w}^{M-nearest}\|_{\ell^2} \leq \|\mathbf{w}-\mathbf{v}\|_{\ell^2}+\|\mathbf{v}-\mathbf{v}^{M-nearest}\|_{\ell^2}+\|\mathbf{v}^{M-nearest}-\mathbf{w}^{M-nearest}\|_{\ell^2} \leq 3\varepsilon.$$

■

Lemma 3.2. *Let* $s > d$ *and* $r = \frac{s}{d} - \frac{1}{2}$. *Suppose that* $\mathbf{v} \in \ell^\infty_{w_s}(\mathcal{N})$ *and* $\mathbf{w} \in \ell^2(\mathcal{N})$ *satisfy*

$$\|\mathbf{v} - \mathbf{w}\|_{\ell^2} \leq \varepsilon \tag{27}$$

for some $\varepsilon > 0$. *If* N *is the minimal integer such that*

$$\|\mathbf{w} - \mathbf{w}^{N-nearest}\|_{\ell^2} \leq 3\varepsilon, \tag{28}$$

then

$$N^d \leq M^d \asymp (\varepsilon^{-1}\|\mathbf{v}\|_{s,x})^{\frac{1}{r}}, \tag{29}$$

and

$$\|\mathbf{v} - \mathbf{w}^{N-nearest}\|_{\ell^2} \leq 4\varepsilon. \tag{30}$$

Proof. By Lemma 3.1 $M \asymp (\varepsilon^{-1}\|\mathbf{v}\|_{s,x})^{\frac{1}{rd}}$ has the property $\|\mathbf{w} - \mathbf{w}^{M-nearest}\|_{\ell^2} \leq 3\varepsilon$. Since N is the minimal integer with the same property, $N \leq M$ necessarily holds. The estimate (30) follows by the triangle inequality. ■

From now on we denote the ball of radius R centered at x by $B(x, R) = \{y : |y - x| \leq R\}$ and note that $\#(B(x, R) \cap \mathcal{N}) \asymp R^d$.

Lemma 3.3. *Let* $s > d$. *Suppose that* $\mathbf{v} \in \ell^\infty_{w_s}(\mathcal{N})$ *and* $\mathbf{z} \in \ell^2(\mathcal{N})$ *is a finite vector such that* $\operatorname{supp} \mathbf{z} \subset B(x, M)$. *Then*

$$\|\mathbf{z}\|_{s,x} \lesssim M^s\|\mathbf{z} - \mathbf{v}\|_{\ell^2} + \|\mathbf{v}\|_{s,x}. \tag{31}$$

Proof. We have

$$\begin{aligned}
\|\mathbf{z}\|_{s,x} &\leq \|\mathbf{z} - \mathbf{v}^{M-nearest}\|_{s,x} + \|\mathbf{v}^{M-nearest}\|_{s,x} \\
&= \sup_{|\ell-x|\leq M} |z_\ell - v_\ell^{M-nearest}|(1 + |\ell - x|)^s + \|\mathbf{v}^{M-nearest}\|_{s,x} \\
&\lesssim M^s\|\mathbf{z} - \mathbf{v}^{M-nearest}\|_{\ell_2} + \|\mathbf{v}\|_{s,x}.
\end{aligned}$$

Observe now that we trivially have $\|\mathbf{z} - \mathbf{v}^{M-nearest}\|_{\ell_2} \leq 2\|\mathbf{z} - \mathbf{v}\|_{\ell_2} \lesssim \|\mathbf{z} - \mathbf{v}\|_{\ell_2}$, and this concludes the proof. ■

Lemma 3.4. *Let $s > d$ and $r = \frac{s}{d} - \frac{1}{2}$. Suppose that $\mathbf{v} \in \ell^\infty_{w_s}(\mathcal{N})$ and that $\mathbf{w} \in \ell^2(\mathcal{N})$ is a finite vector which satisfies*

$$\|\mathbf{v} - \mathbf{w}\|_{\ell^2} \leq \varepsilon \tag{32}$$

for some small $\varepsilon > 0$. If we choose N as the minimal integer such that

$$\|\mathbf{w} - \mathbf{w}^{N-nearest}\|_{\ell^2} \leq 3\varepsilon, \tag{33}$$

then

$$\|\mathbf{w}^{N-nearest}\|_{s,x} \leq C(s) \max\left\{\|\mathbf{v}\|^{\sigma}_{s,x}\varepsilon^{1-\sigma}, \|\mathbf{v}\|_{s,x}\right\}. \tag{34}$$

where $\sigma = \frac{s}{dr} = \frac{1}{1-\frac{d}{2s}} > 1$ and $C(s) > 0$ is a suitable uniform constant.

Proof. For the second estimate we combine Lemma 3.2 with Lemma 3.3. We get

$$\begin{aligned}
\|\mathbf{w}^{N-nearest}\|_{s,x} &\lesssim (\|\mathbf{v}\|^{1/r}_{s,x}\varepsilon^{-1/r})^{s/d}\|\mathbf{w}^{N-nearest} - \mathbf{v}\|_{\ell_2} + \|\mathbf{v}\|_{s,x} \\
&= 4\|\mathbf{v}\|^{s/(dr)}_{s,x}\varepsilon^{1-s/(dr)} + \|\mathbf{v}\|_{s,x} \\
&\lesssim \max\left\{\|\mathbf{v}\|^{\sigma}_{s,x}\varepsilon^{1-\sigma}, \|\mathbf{v}\|_{s,x}\right\}.
\end{aligned}$$

The latter estimate is due to $\varepsilon > 0$ small. ■

By combining the results of this section we define the following procedure:

Assume $\varepsilon > 0$, $s > d$, $r = s/d - 1/2$, and $0 < \theta < 1/4$.

COARSE$[(1-\theta)\varepsilon, \mathbf{w}] \to \mathbf{w}_\varepsilon$: determines, for a finitely supported vector $\mathbf{w}$ (localized at x), a finitely supported $\mathbf{w}_\varepsilon = \mathbf{w}^{M-nearest}$, where M is the minimal integer for which $\|\mathbf{w} - \mathbf{w}^{M-nearest}\|_{\ell^2} \leq (1-\theta)\varepsilon$, with the following property.
If for $\mathbf{v} \in \ell^\infty_{w_s}(\mathcal{N})$ we have the approximation $\|\mathbf{w} - \mathbf{v}\|_{\ell_2} \leq \theta\varepsilon$ then

(a) $\|\mathbf{v} - \mathbf{w}_\varepsilon\|_{\ell^2} \leq \varepsilon$;

(b) $\mathrm{supp}(\mathbf{w}_\varepsilon) \subset B(x, M)$ and $M^d \lesssim \#\,\mathrm{supp}(\mathbf{w}_\varepsilon) \lesssim \varepsilon^{-1/r}\|\mathbf{v}\|^{1/r}_{s,x}$;

(c) $\mathcal{C}(\mathbf{w}_\varepsilon) \lesssim \#\,\mathrm{supp}(\mathbf{w})$;

(d) $\|\mathbf{w}_\varepsilon\|_{s,x} \lesssim \max\left\{\|\mathbf{v}\|^{\sigma}_{s,x}\varepsilon^{1-\sigma}, \|\mathbf{v}\|_{s,x}\right\}$.

3.3 Matrix Computations

The aim of this section is to establish the second fundamental subroutine, namely a fast algorithm for the computation of a finite vector $\mathbf{w}_\varepsilon$, possibly with small (or minimal) support such that

$$\|\mathbf{w}_\varepsilon - \mathbf{A}\mathbf{v}\|_{\ell^2} \leq \varepsilon.$$

We first study the approximation of an arbitrary matrix $\mathbf{A}$ on $\mathcal{N}$ by a banded matrix. For $N \in \mathbb{N}$, we define the matrix $\mathbf{B}^N$ by

$$b^N_{hk} := \begin{cases} a_{hk} & , \ |h-k| \le N \\ 0 & , \ \text{otherwise.} \end{cases}$$

Clearly, a matrix with fast off-diagonal decay will be approximated well by banded matrices. The following lemma studies the error $\mathbf{A} - \mathbf{B}^N$ for the Jaffard class on various sequence spaces.

Lemma 3.5. *Assume that* $\mathbf{A} \in \mathcal{A}_\gamma$.
(a) If $\gamma > d$, *then we have, in the operator norm on* $\ell^2(\mathcal{N})$,

$$\|\mathbf{A} - \mathbf{B}^N\| := \|\mathbf{A} - \mathbf{B}^N\|_{\ell^2 \to \ell^2} \lesssim N^{d-\gamma} \, . \tag{35}$$

(b) If $s + d < \gamma$, $1 \le p \le \infty$, *and* m *is an* s*-moderate weight, then, in the operator norm on* ℓ^p_m, *we have*

$$\|\mathbf{A} - \mathbf{B}^N\|_{\ell^p_m \to \ell^p_m} \lesssim N^{d+s-\gamma}. \tag{36}$$

Moreover, $\mathbf{A}$ *defines a bounded operator from* $\ell^p_m(\mathcal{N})$ *to itself.*
(c) In particular, if $s + d < \gamma$, *then*

$$\|\mathbf{A} - \mathbf{B}^N\|_{s,x} := \|\mathbf{A} - \mathbf{B}^N\|_{\ell^\infty_{w_s} \to \ell^\infty_{w_s}} \lesssim N^{d+s-\gamma} \, , \tag{37}$$

where the constant does not depend on x. *Here the notation* $\|\mathbf{C}\|_{s,x}$ *indicates the operator norm of* $\mathbf{C}$ *acting on* $\ell^\infty_{w_s}$ *endowed with the norm* $\|\cdot\|_{s,x}$.

Proof. (a) We use the Schur test to estimate the operator norm on ℓ^2 .

$$\begin{aligned} \sup_{h \in \mathcal{N}} \sum_{k \in \mathcal{N}} |a_{hk} - b^N_{hk}| &= \sup_{h \in \mathcal{N}} \sum_{k \in \mathcal{N}: \ |k-h| > N} |a_{hk}| \\ &\le \|\mathbf{A}\|_{\mathcal{A}_\gamma} \sup_{h \in \mathcal{N}} \sum_{k \in \mathcal{N}: \ |k-h| > N} |(1 + |k-h|)^{-\gamma} \\ &\lesssim \|\mathbf{A}\|_{\mathcal{A}_\gamma} N^{d-\gamma}, \end{aligned}$$

and likewise with k and h interchanged. Thus (35) follows.

(b) We fix the weight m and prove (36) first for $p = 1$ and $p = \infty$. Since $m(k) \lesssim (1 + |k-l|)^s \, m(l)$, we obtain

$$\begin{aligned} \|(\mathbf{A} - \mathbf{B}^N)\mathbf{c}\|_{\ell^1_m} &= \sum_k \Big| \sum_{l: |l-k| > N} a_{kl} c_l \Big| m(k) \\ &\le \|\mathbf{A}\|_{\mathcal{A}_\gamma} \sum_k \sum_{l: |l-k| > N} (1 + |k-l|)^{-\gamma} |c_l| m(k) \\ &\lesssim \|\mathbf{A}\|_{\mathcal{A}_\gamma} \sum_l \Big(\sup_l \sum_{k: |k-l| > N} (1 + |k-l|)^{-\gamma+s} \Big) |c_l| m(l) \\ &\lesssim \|\mathbf{A}\|_{\mathcal{A}_\gamma} N^{-\gamma+s+d} \|\mathbf{c}\|_{\ell^1_m} \, . \end{aligned}$$

Similarly, for $p = \infty$ we obtain

$$\begin{aligned} \|(\mathbf{A}-\mathbf{B}^N)\mathbf{c}\|_{\ell^\infty_m} &= \sup_k \Big| \sum_{l:|l-k|>N} a_{kl} c_l \Big| m(k) \\ &\leq \|\mathbf{A}\|_{\mathcal{A}_\gamma} \sup_k \sum_{l:|l-k|>N} (1+|k-l|)^{-\gamma+s} |c_l| m(l) \\ &\lesssim \|\mathbf{A}\|_{\mathcal{A}_\gamma} \|\mathbf{c}\|_{\ell^\infty_m} \sup_k \sum_{l:|k-l|>N} (1+|k-l|)^{-\gamma+s} \qquad (38) \\ &\lesssim \|\mathbf{A}\|_{\mathcal{A}_\gamma} \|\mathbf{c}\|_{\ell^\infty_m} N^{-\gamma+s+d} . \end{aligned}$$

For $1 < p < \infty$, (36) now follows by interpolation of ℓ^p_m-spaces. Finally, observe that $\mathbf{B}^0$ is clearly a bounded operator on ℓ^p_m-spaces, since it is a diagonal matrix with bounded diagonal elements. Hence, $\mathbf{A}$ itself is also bounded on ℓ^p_m-spaces.

(c) Since $m(k) = (1+|k-x|)^s$ is s-moderate, the uniform estimate in the (s,x)-norm is a special case of (38). ■

Lemma 3.5 gives an error estimate for the approximation of a matrix in the Jaffard class by a banded matrix. This should be distinguished from the more general concept of approximation by sparse matrices of [8]. According to [8], a matrix $\mathbf{A} : \ell^2(\mathcal{N}) \to \ell^2(\mathcal{N})$ is called r^**–compressible* for $r^* > 0$ if for each $j \in \mathbb{N}$ there exist constants α_j, C_j, and a matrix $\mathbf{A}^j$ having at most $\alpha_j 2^j$ non–zero entries in each column, such that

$$\|\mathbf{A}-\mathbf{A}^j\| \leq C_j, \tag{39}$$

where $(\alpha_j)_{j\in\mathbb{N}}$ is summable, and for any $0 < r < r^*$, $(C_j 2^{rj})_{j\in\mathbb{N}}$ is summable.

Lemma 3.6. *If* $\mathbf{A} \in \mathcal{A}_\gamma$ *for* $\gamma > d + 1/2$*, then* $\mathbf{A}$ *is at least* $(\gamma-d)/d$*–compressible.*

Proof. For all $j \in \mathbb{N}$ denote $\mathbf{A}^j := \mathbf{B}^{2^{j/d}/j^2}$. Then $\mathbf{A}^j$ has at most $2^j/j^{2d}$ entries in each column, and $\sum_j \alpha_j = \sum_j j^{-2d} < \infty$. By Lemma 3.5, one has

$$\|\mathbf{A}-\mathbf{A}^j\| \lesssim \Big(\frac{2^{j/d}}{j^2}\Big)^{d-\gamma} =: C_j, \quad \text{and} \quad \sum_j C_j 2^{j(\gamma-d)/d} = \sum_j j^{-2(\gamma-d)} < \infty ,$$

and the last sum converges because $\gamma > d + 1/2$. ■

We now turn to the fast matrix-vector multiplication for matrices in the Jaffard class and vectors in $\ell^\infty_{w_s}$. The results are very much inspired by [8] and [25], and the proofs partially follow their lines. Nevertheless, there is a substantial difference. Our algorithm exploits decay conditions instead of sparsity, and it does not require any sorting routines or binary binning strategies.

We now introduce the following numerical procedure for approximating $\mathbf{A}\mathbf{v}$.

APPLY$[\varepsilon, \mathbf{A}, \mathbf{v}] \to \mathbf{w}_\varepsilon$:

(i) With

$$C_k = \mathcal{N} \cap ([-2^{k/d-1}, 2^{k/d-1}]^d + x), \quad C_0 = \emptyset$$

define the dyadic coronas

$$V_k = C_k \backslash C_{k-1}.$$

(ii) Set $\mathbf{v}^{[k]} := \mathbf{v}\chi_{V_k}$*. Then* $\# \operatorname{supp} \mathbf{v}^{[k]} = \# V_k \asymp 2^{k-1}$*. Choose* k^* *such that*

$$\|\mathbf{A}\| \left\| \mathbf{v} - \sum_{k=0}^{k^*} \mathbf{v}^{[k]} \right\|_{\ell^2} \leq \frac{\varepsilon}{2}. \tag{40}$$

(iii) Compute the smallest $j \geq k^* + 1$ *such that*

$$\sum_{k=0}^{k^*} C_{j-k} \|\mathbf{v}^{[k]}\|_{\ell^2} \leq \frac{\varepsilon}{2}, \tag{41}$$

where C_j *is as in the proof of Lemma 3.6.*

(iv) Compute

$$\mathbf{w}_\varepsilon := \sum_{k=0}^{k^*} \mathbf{A}^{j-k} \mathbf{v}^{[k]}. \tag{42}$$

Theorem 3.7. *Let* $\gamma > \max(s + d, d + 1/2)$*,* $r = \frac{s}{d} - \frac{1}{2}$*, and* $\varepsilon > 0$*. Assume that* $\mathbf{A} \in \mathcal{A}_\gamma$ *and* $\mathbf{v} \in \ell^\infty_{w_s}$ *is a vector with finite support. Then the algorithm* **APPLY** *produces a vector* $\mathbf{w}_\varepsilon$ *with the following properties:*

(a) $\|\mathbf{w}_\varepsilon - \mathbf{A}\mathbf{v}\|_{\ell^2} \leq \varepsilon$*;*

(b) $\operatorname{supp}(\mathbf{w}_\varepsilon) \subseteq B(x, N)$ *and* $\# \operatorname{supp}(\mathbf{w}_\varepsilon) \lesssim N^d \lesssim \varepsilon^{-1/r} \|\mathbf{v}\|_{s,x}^{1/r}$*;*

(c) $\mathcal{C}(\mathbf{w}_\varepsilon) \lesssim \varepsilon^{-1/r} \|\mathbf{v}\|_{s,x}^{1/r} + \# \operatorname{supp}(\mathbf{v})$*;*

(d) the procedure is bounded in the sense that the estimate

$$\|\mathbf{w}_\varepsilon\|_{s,x} \lesssim \|\mathbf{v}\|_{s,x}, \tag{43}$$

holds with a constant independent of ε *and* x*.*

Thus **APPLY** *is optimal.*

Proof. **Step 1.** The error estimates (a) and (d)
Since $\mathbf{v} = \sum_{k=0}^\infty \mathbf{v}^{[k]}$, we may write

$$\mathbf{w}_\varepsilon - \mathbf{A}\mathbf{v} = \sum_{k=0}^{k^*} (\mathbf{A}^{j-k} - \mathbf{A}) \mathbf{v}^{[k]} + \mathbf{A} \Big(\sum_{k=0}^{k^*} \mathbf{v}^{[k]} - \mathbf{v} \Big).$$

Taking first the ℓ^2-norm, we estimate, with formulas (40) and (41),

$$\begin{aligned}\|\mathbf{w}_\varepsilon - \mathbf{A}\mathbf{v}\|_{\ell^2} &\leq \sum_{k=0}^{k^*} \|\mathbf{A}^{j-k} - \mathbf{A}\| \, \|\mathbf{v}^{[k]}\|_{\ell^2} + \|\mathbf{A}\| \left\| \sum_{k=0}^{k^*} \mathbf{v}^{[k]} - \mathbf{v} \right\|_{\ell^2} \\ &\leq \sum_{k=0}^{k^*} C_{j-k} \|\mathbf{v}^{[k]}\|_{\ell^2} + \frac{\varepsilon}{2} < \varepsilon \,,\end{aligned}$$

and so (a) is proved.

Taking next the (s,x)-norm and using Lemma 3.5(c), one has the following estimates:

$$\begin{aligned}\|\mathbf{w}_\varepsilon - \mathbf{A}\mathbf{v}\|_{s,x} &\leq \sum_{k=0}^{k^*} \|\mathbf{A}^{j-k} - \mathbf{A}\|_{s,x} \, \|\mathbf{v}^{[k]}\|_{s,x} + \|\mathbf{A}\|_{s,x} \left\| \sum_{k=k^*+1}^{\infty} \mathbf{v}^{[k]} \right\|_{s,x} \\ &\lesssim \sum_{k=0}^{k^*} \left(\frac{2^{(j-k)/d}}{(j-k)^2} \right)^{d+s-\gamma} \|\mathbf{v}\|_{s,x} + \|\mathbf{A}\|_{s,x} \|\mathbf{v}\|_{s,x} \lesssim \|\mathbf{v}\|_{s,x} \,.\end{aligned}$$

Consequently,

$$\|\mathbf{w}_\varepsilon\|_{s,x} \lesssim \|\mathbf{w}_\varepsilon - \mathbf{A}\mathbf{v}\|_{s,x} + \|\mathbf{A}\mathbf{v}\|_{s,x} \lesssim \|\mathbf{v}\|_{s,x} + \|\mathbf{A}\|_{s,x} \, \|\mathbf{v}\|_{s,x} \lesssim \|\mathbf{v}\|_{s,x} \,,$$

and (d) is proved.

Step 2. Support and operation count for $\mathbf{w}_\varepsilon$.
First of all, observe that the construction of $\mathbf{v}^{[k]}$, $k = 0, \ldots, k^*$, has a cost not exceeding $\#\operatorname{supp}(\mathbf{v})$. If $\mathbf{B}$ is a banded matrix with $\mathbf{B}_{hk} = 0$ for $|h-k| > N$ and $\mathbf{v}$ is a localized vector with $v_k = 0$ for $|k-x| > M$, then $(\mathbf{B}\mathbf{v})(h) = 0$ for $|h-x| > M+N$ and so $\operatorname{supp} \mathbf{B}\mathbf{v} \subseteq B(x, M+N)$ and $\#\operatorname{supp} \mathbf{B}\mathbf{v} \lesssim (M+N)^d$. The computation of each entry of $\mathbf{B}\mathbf{v}$ requires N^d (the number of non-zero entries in a row or column of $\mathbf{B}$) or $M^d \lesssim \#\operatorname{supp} \mathbf{v}$ multiplications, whichever number is smaller. This means that the computation of $\mathbf{B}\mathbf{v}$ requires $\lesssim (M+N)^d \, \min(N^d, M^d) \leq 2^d M^d N^d$ operations.

As a consequence $\mathbf{w}_\varepsilon = \sum_{k=0}^{k^*} \mathbf{A}^{j-k} \mathbf{v}^{[k]}$ is supported on the set $\bigcup_{k=0}^{k^*} B(x, \alpha_{j-k}^{1/d} 2^{(j-k)/d} + 2^{k/d}) \subseteq B(x, 2^{j/d})$ and $\#\operatorname{supp} \mathbf{w}_\varepsilon \lesssim 2^j$.

For the operation count we have, according to the conventions of Lemma 3.6, that

$$\mathcal{C}(\mathbf{w}_\varepsilon) \lesssim \#\operatorname{supp}(\mathbf{v}) + \sum_{k=0}^{k^*} \alpha_{j-k} 2^{j-k} \#\operatorname{supp} \mathbf{v}^{[k]} \lesssim \#\operatorname{supp}(\mathbf{v}) + \sum_{k=0}^{k^*} \alpha_{j-k} 2^{j-k} 2^k \lesssim \#\operatorname{supp}(\mathbf{v}) + 2^j \,.$$

Step 3. To conclude the proof, it suffices to show that $2^j \lesssim \varepsilon^{-1/r} \|\mathbf{v}\|_{s,x}^{1/r}$ for j as defined in (41). Let us estimate the norm of $\mathbf{v}^{[k]}$. Since $\operatorname{supp} \mathbf{v}^{[k]} \subseteq V_k \subseteq \{l : |l-x| \geq 2^{(k-1)/d}\}$, the norm is bounded by

$$\begin{aligned}\|\mathbf{v}^{[k]}\|_{\ell^2} &\leq \|\mathbf{v}\|_{s,x} \Big(\sum_{\{l : |l-x| \geq 2^{(k-1)/d}\}} (1 + |l-x|)^{-2s} \Big)^{1/2} \qquad (44) \\ &\lesssim \|\mathbf{v}\|_{s,x} \left(2^{(k-1)/d} \right)^{d/2-s} \lesssim \|\mathbf{v}\|_{s,x} \, 2^{-rk} \,.\end{aligned}$$

Since j is the *smallest* integer satisfying (41), we have

$$\begin{aligned}\frac{\varepsilon}{2} &\leq \sum_{k=0}^{k^*} C_{j-1-k} \|\mathbf{v}^{[k]}\| \\ &\lesssim \|\mathbf{v}\|_{s,x} \sum_{k=0}^{k^*} C_{j-1-k} 2^{-rk} \\ &= 2^{-r(j-1)} \|\mathbf{v}\|_{s,x} \sum_{k=0}^{k^*} C_{j-1-k} 2^{(j-1-k)r}\,.\end{aligned}$$

The hypothesis $\gamma > s + d$ implies that $r = \frac{s}{d} - \frac{1}{2} < \frac{\gamma-d}{d}$. Therefore

$$2^{rj} \varepsilon \|\mathbf{v}\|_{s,x}^{-1} \lesssim \sum_{k=0}^{k^*} C_{j-1-k} 2^{\frac{\gamma-d}{d}(j-1-k)} < \infty\,,$$

and so $2^j \lesssim \left(\varepsilon^{-1} \|\mathbf{v}\|_{s,x}\right)^{1/r}$. As observed above, this concludes the proof. ∎

3.4 Numerical Solution of Bi-Infinite Systems of Linear Equations

We come back to the numerical treatment of operator equations (8). As already outlined in Remark 2.2, the discretization of (8) leads to a bi-infinite system

$$\mathbf{A}\mathbf{u} = \mathbf{f}, \tag{45}$$

which can be treated by means of the damped Richardson iteration (14). We shall focus on matrices $\mathbf{A} \in \mathcal{A}_\gamma$ and $\mathbf{u}, \mathbf{f} \in \ell^\infty_{w_s}(\mathcal{N})$.

The iterations (14) cannot be implemented numerically since in general they act on infinite sequences. To turn the abstract iteration (14) into a realizable algorithm, we substitute the infinite sequence $\mathbf{f}$ and the infinite exact matrix-vector multiplication by the finite approximations $\mathbf{RHS}[\varepsilon, \mathbf{f}]$ and $\mathbf{APPLY}[\varepsilon, \mathbf{A}, \mathbf{v}]$ as introduced in the Subsections 3.1 and 3.3. Furthermore, we allow the accuracy to depend on the iteration by choosing a suitable sequence ε_n converging to 0. We now try the following iteration scheme:

$$\mathbf{v}^{(n+1)} = \mathbf{v}^{(n)} - \alpha(\mathbf{APPLY}[\varepsilon_n, \mathbf{A}, \mathbf{v}^{(n)}] - \mathbf{RHS}[\varepsilon_n, \mathbf{f}]), \quad \mathbf{v}^{(0)} = \mathbf{0}, \qquad n = 0, 1, \ldots. \tag{46}$$

The precise algorithm must also include a stopping criterion and reads as follows, where again we follow the lines of [9, 12, 25].

Algorithm 1. $\mathbf{SOLVE}[\varepsilon, \mathbf{A}, \mathbf{f}] \to \mathbf{u}_\varepsilon$:

Let $\theta < 1/4$ and $(K_i)_{i\in\mathbb{N}} \in \mathbb{N}$ be fixed such that $3\rho^{K_i} < \theta$, for all $i \in \mathbb{N}$, $\rho := \|(\mathrm{id} - \alpha\mathbf{A})|_{\mathrm{ran}(\mathbf{A})}\|_{\ell^2(\mathcal{N})}$.

```
$i := 0$, $\mathbf{u}^{(0)} := 0$, $\varepsilon_0 := \|\mathbf{A}^{-1}_{|\,\mathrm{ran}(\mathbf{A})}\|\|\mathbf{f}\|_{\ell_2(\mathcal{N})}$
While $\varepsilon_i > \varepsilon$ do
    $i := i + 1$
    $\varepsilon_i := 3\rho^{K_i}\varepsilon_{i-1}/\theta$
    $\mathbf{f}^{(i)} := \mathbf{RHS}[\frac{\theta\varepsilon_i}{6\alpha K_i}, \mathbf{f}]$
    $\mathbf{v}^{(i,0)} := \mathbf{u}^{(i-1)}$
    For $j = 1, ..., K_i$ do
        $\mathbf{v}^{(i,j)} := \mathbf{v}^{(i,j-1)} - \alpha(\mathbf{APPLY}[\frac{\theta\varepsilon_i}{6\alpha K_i}, \mathbf{A}, \mathbf{v}^{(i,j-1)}] - \mathbf{f}^{(i)})$
    enddo
    $\mathbf{u}^{(i)} := \mathbf{COARSE}[(1-\theta)\varepsilon_i, \mathbf{v}^{(i,K_i)}]$
enddo
$\mathbf{u}_\varepsilon := \mathbf{u}^{(i)}$.
```

Algorithm 1 converges to the solution of the bi-infinite system of linear equations (45) as shown by the following proposition.

Proposition 3.8. *In the situation of Theorem 2.2, let* $\mathbf{u} \in \ell_2(\mathcal{N})$ *be a solution of* (45)*. Then* $\mathbf{SOLVE}[\varepsilon, \mathbf{A}, \mathbf{f}]$ *produces finitely supported vectors* $\mathbf{v}^{(i,K_i)}, \mathbf{u}^{(i)}$ *such that*

$$\left\|\mathbf{P}(\mathbf{u} - \mathbf{u}^{(i)})\right\|_{\ell_2(\mathcal{N})} \leq \varepsilon_i, \quad i \geq 0. \tag{47}$$

Moreover, it holds that

$$\left\|\mathbf{P}\mathbf{u} + (\mathrm{id} - \mathbf{P})\mathbf{u}^{(i-1)} - \mathbf{v}^{(i,K_i)}\right\|_{\ell_2(\mathcal{N})} \leq \frac{2\theta\varepsilon_i}{3}, \quad i \geq 1. \tag{48}$$

The proof of this proposition is omitted since it is essentially identical to a similar one in [12, 25]. The only modification is that the number of inner iterations K_i may vary, but this does not change the proof.

REMARK: Since the components of $\mathbf{v}^{(n)}$ in $\ker(\mathbf{A})$ are not reduced in the iteration, we only get an error estimate for the *projected* error $\mathbf{P}(\mathbf{u}_\varepsilon - \mathbf{u})$. As already outlined in Subsection 2.2, this does not effect the overall convergence of the scheme. It remains to investigate the optimality of Algorithm 1 with respect to the support of the approximation $\mathbf{u}_\varepsilon$ and the operation count. The assumptions $\mathbf{A} \in \mathcal{A}_\gamma$ and $\mathbf{u}, \mathbf{f} \in \ell^\infty_{w_s}(\mathcal{N})$ are again crucial.

Theorem 3.9. *Let* $\mathcal{M}$ *be a closed subspace of* $\ell^2(\mathcal{N})$ *with orthogonal projection* $\mathbf{P}$ *onto* $\mathcal{M}$*. Assume that* $\mathbf{A} = \mathbf{A}^* \in \mathcal{A}_\gamma$*,* $\ker\mathbf{A} = \mathcal{M}^\perp$ *and that* $\mathbf{A} : \mathcal{M} \longrightarrow \mathcal{M}$ *is invertible. Then the pseudoinverse* $\mathbf{A}^\dagger$*, i.e., the unique element in* $L(\ell^2)$ *satisfying* $\mathbf{A}^\dagger\mathbf{A} = \mathbf{A}\mathbf{A}^\dagger = \mathbf{P}$ *and* $\ker\mathbf{A}^\dagger = \mathcal{M}^\perp$*, is an element of* $\mathcal{A}_\gamma$*. In particular* $\mathbf{P} \in \mathcal{A}_\gamma$*.*

Proof. See [15, Theorem 3.4]. ■

Corollary 3.10. *Let $\mathcal{M}$ be a closed subspace of $\ell^2(\mathcal{N})$ with orthogonal projection $\mathbf{P}$ onto $\mathcal{M}$. Assume that $\mathbf{A} = \mathbf{A}^* \in L(\ell^2)$, $\ker \mathbf{A} = \mathcal{M}^\perp$ and that $\mathbf{A} : \mathcal{M} \longrightarrow \mathcal{M}$ is invertible. If $\mathbf{f} \in \mathcal{M}$ then there exists a unique solution $\mathbf{u} \in \mathcal{M}$ of equation* (45). *Moreover if one assumes $\mathbf{A} \in \mathcal{A}_\gamma$ and if $d + s < \gamma$ and $\mathbf{f} \in \mathcal{M} \cap \ell^\infty_{w_s}(\mathcal{N})$, then there exists a unique $\mathbf{u} \in \mathcal{M} \cap \ell^\infty_{w_s}(\mathcal{N})$ such that $\mathbf{A}\mathbf{u} = \mathbf{f}$.*

Proof. Since $\mathbf{A} : \mathcal{M} \longrightarrow \mathcal{M}$ is invertible and $\mathbf{f} \in \mathcal{M}$ then there exists a unique $\mathbf{u} \in \mathcal{M}$ solution of (45). Moreover, the inverse of $\mathbf{A}$ on $\mathcal{M}$ coincides with $\mathbf{A}^\dagger$. Since $\mathbf{A} \in \mathcal{A}_\gamma$ then by Theorem 3.9 $\mathbf{A}^\dagger \in \mathcal{A}_\gamma$ and $\mathbf{u} = \mathbf{A}^\dagger \mathbf{f}$. If $\mathbf{f} \in \ell^\infty_{w_s}(\mathcal{N})$, then $\mathbf{u} \in \ell^\infty_{w_s}(\mathcal{N})$ by Lemma 3.5(b). ■

Theorem 3.11. *Let $\mathcal{M}$ be a closed subspace of $\ell^2(\mathcal{N})$ with orthogonal projection $\mathbf{P}$ onto $\mathcal{M}$. Assume that $\mathbf{A} = \mathbf{A}^* \in \mathcal{A}_\gamma$ is a positive operator, $\ker \mathbf{A} = \mathcal{M}^\perp$ and that $\mathbf{A} : \mathcal{M} \longrightarrow \mathcal{M}$ is invertible. Suppose that $\mathbf{f} \in \mathcal{M} \cap \ell^\infty_{w_s}$ is s-localized at x and assume $\tilde{s} > s > d$ and $\gamma > \tilde{s} + d$.*

Set $r = \frac{s}{d} - \frac{1}{2}$, $\tilde{r} = \frac{\tilde{s}}{d} - \frac{1}{2}$, $\tilde{\tilde{s}} = s\left(\frac{\tilde{s}-s}{s-d/2} + 1\right)$, and $\tilde{\sigma} = \frac{\tilde{s}}{\tilde{r}d} = \frac{\tilde{s}}{\tilde{s}-d/2}$.
Fix $0 < \eta < 1$ and $0 < c < 1$ and choose $(K_i)_{i \in \mathbb{N}}$ in Algorithm 1, such that

$$c\eta \leq \max \left\{ C(\tilde{s}) \| \operatorname{id} - \mathbf{P} \| \left(\frac{3\rho^{K_i}}{\theta} \right)^{\tilde{\sigma} \frac{\tilde{\tilde{s}}}{s} - 1}, \right.$$

$$\left. 2^{\tilde{\sigma}-1} C(\tilde{s}) \left(\frac{3\rho^{K_i}}{\theta} \right)^{\tilde{\sigma}(\frac{\tilde{\tilde{s}}}{s} - 1)} \varepsilon_{i-1}^{1-\tilde{\sigma}} \| \mathbf{u}^{(i-1)} \|_{\tilde{s},x}^{\tilde{\sigma}-1} \| \operatorname{id} - \mathbf{P} \|^{\tilde{\sigma}} \right\} := \eta_i \leq \eta \,. \tag{49}$$

Then for all $0 < \varepsilon \leq \varepsilon_0$ the vector $\mathbf{u}_\varepsilon = \mathbf{SOLVE}[\varepsilon, \mathbf{A}, \mathbf{f}]$ satisfies the following properties:

(a) $\|\mathbf{P}(\mathbf{u}_\varepsilon - \mathbf{u})\|_{\ell^2} \leq \varepsilon$, *where $\mathbf{u} \in \mathcal{M}$ is the unique solution of $\mathbf{A}\mathbf{u} = \mathbf{f}$ in $\mathcal{M}$;*

(b) $\operatorname{supp}(\mathbf{u}^\varepsilon) \subseteq B(x, M)$ *and* $\# \operatorname{supp}(\mathbf{u}^\varepsilon) \lesssim M^d \lesssim \varepsilon^{-\tilde{\sigma}/r} \max\{\|\mathbf{u}\|_{s,x}^{1/\tilde{\sigma}}, \|\mathbf{u}\|_{s,x}\}^{\tilde{\sigma}/r}$;

(c) $\mathcal{C}(\mathbf{u}_\varepsilon) \lesssim \varepsilon^{-\left(\frac{\tilde{\sigma}}{r} + \xi(\tilde{\sigma}-1)\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s} + \delta\right)} \max\{\|\mathbf{u}\|_{s,x}^{1/\tilde{\sigma}}, \|\mathbf{u}\|_{s,x}\}^{\frac{\tilde{\sigma}}{r} + \xi(\tilde{\sigma}-1)\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s} + \delta}$, *for $\delta > 0$ arbitrarily small and some constants $\xi > 0$;*

(d) The procedure is quasi-bounded *in the sense that*

$$\varepsilon^{\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s} - 1} \|\mathbf{u}_\varepsilon\|_{\tilde{s},x} \lesssim \max\{\|\mathbf{u}\|_{s,x}^{1/\tilde{\sigma}}, \|\mathbf{u}\|_{s,x}\}^{\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}}, \tag{50}$$

uniformly with respect to $\varepsilon \to 0$ and $x \in \mathcal{N}$.

The proof is rather technical and it is deferred to the Appendix. Let us now discuss some crucial elements of the previous results.
REMARKS:

(i) Observe that for highly localized solutions, i.e., for $\tilde{s} \approx s > 0$ very large (consequently $\gamma > 0$ must be also very large), we have $\tilde{\sigma} \approx 1 \approx \frac{\tilde{\tilde{s}}}{s}$. In this case the estimates of Theorem 3.11 (b) and (c) ensure the *quasi-optimality* of the algorithm. The complexity approaches $\mathcal{C}(\mathbf{u}_\varepsilon) \lesssim \varepsilon^{-\left(\frac{1}{r}+\delta\right)} \|\mathbf{u}\|_{s,x}^{\frac{1}{r}+\delta}$ for $\delta > 0$ arbitrarily small, as compared to the optimal complexity $\epsilon^{-1/r}\|u\|_{s,x}^{1/r}$. It seems difficult to produce an optimal **SOLVE** procedure based only on M-nearest neighborhood approximations. However, despite the sub-optimal asymptotic behavior, we expect that the simplification due to the use of M-nearest neighborhood approximations instead of best N-term approximations will improve the performance. Actually, the computational costs of **COARSE** and **APPLY** are already optimal, and expected to be smaller than the procedures designed for approximations in $\ell^{\tau,w}$ [8, 9, 25].

(ii) For the complexity estimates stated in Theorem 3.11 it is crucial that the projection $\mathbf{P}$ and the pseudo-inverse $\mathbf{A}^\dagger$ are both in $\mathcal{A}_\gamma$ and therefore bounded on $\ell^\infty_{w_s}$. The analogous statement for wavelet-based adaptive algorithms is open: it is not clear whether the projection $\mathbf{P}$ is bounded in $\ell^{\tau,w}$, because one has to deal with the Lemarié algebra which is not inverse-closed, see [12]. Although the corresponding algorithm does perform optimally in practice, as demonstrated in [11, 13], its optimality on $\ell^{\tau,w}$ has been shown rigorously only under the additional hypothesis that $\mathbf{P}$ is bounded on $\ell^{\tau,w}$, see [25, Thm. 3.12].

Next we show that the **SOLVE** routine converges not only in the ℓ^2-norm, but also for a much larger class of norms. First we state a technical lemma.

Lemma 3.12. *Assume that* $\mathbf{u} \in \ell^\infty_{w_s}(\mathcal{N})$ *is localized at* x. *If* $\mathbf{v}$ *is a finitely supported vector such that*

$$\operatorname{supp}(\mathbf{v}) \subset B(x, M)$$

and

$$\|\mathbf{u} - \mathbf{v}\|_{\ell^2} \le \varepsilon,$$

then

$$\|\mathbf{u} - \mathbf{v}\|_{\ell^p_m} \lesssim m(x) \left(M^{t + d\max(0, \frac{1}{p} - \frac{1}{2})} \varepsilon + M^{t - s + \frac{d}{p}} \|\mathbf{u}\|_{s,x} \right)$$

where $1 \le p \le \infty$ *and* m *is* t*-moderate for some* $t < s - d/p$.

Proof. Set $S = \operatorname{supp}(\mathbf{v}) \subset B(x, M)$ and note that $\#S \lesssim M^d$. Restricting $\mathbf{u}$ to S and to S^c, respectively, we may write

$$\mathbf{u} - \mathbf{v} = \mathbf{u}|_S - \mathbf{v} + \mathbf{u}|_{S^c}.$$

For the estimate of $\mathbf{u}$ outside S, we may use the embedding $\ell^\infty_{w_s} \subset \ell^p_m$ and obtain

$$\begin{aligned}
\|\mathbf{u}|_{S^c}\|^p_{\ell^p_m} &= \sum_{|k-x|>M} |u_k|^p m(k)^p \\
&\lesssim \|\mathbf{u}\|^p_{s,x} \sum_{|k-x|>M} (1+|k-x|)^{-ps} m(k-x+x)^p \\
&\lesssim \|\mathbf{u}\|^p_{s,x} m(x)^p \sum_{|k|>M} (1+|k|)^{-ps}(1+|k|)^{tp} \\
&\lesssim \|\mathbf{u}\|^p_{s,x} m(x)^p M^{(t-s)p+d}.
\end{aligned}$$

On S, we extract the weight and find

$$\begin{aligned}
\|\mathbf{u}|_S - \mathbf{v}\|_{\ell^p_m} &\lesssim \max_{|k-x|\le M} m(k)\|\mathbf{u}|_S - \mathbf{v}\|_{\ell^p} \\
&\lesssim m(x) \max_{|k|\le M} (1+|k|)^t \|\mathbf{u}|_S - \mathbf{v}\|_{\ell^p} \\
&\lesssim m(x) M^t \|\mathbf{u}|_S - \mathbf{v}\|_{\ell^p}.
\end{aligned}$$

If $p \ge 2$, the embedding $\ell^2 \subset \ell^p$ yields

$$\|\mathbf{u}|_S - \mathbf{v}\|_{\ell^p} \le \|\mathbf{u}|_S - \mathbf{v}\|_{\ell^2} \le \|\mathbf{u} - \mathbf{v}\|_{\ell^2} \le \varepsilon.$$

If $p < 2$, then Hölders's inequality with exponents $q = \frac{2}{p}$ and $q' = \frac{2}{2-p}$ yields

$$\|\mathbf{u}|_S - \mathbf{v}\|_{\ell^p} \le \|\mathbf{u}|_S - \mathbf{v}\|_{\ell^2}\ (\#S)^{\frac{1}{p}-\frac{1}{2}} \lesssim \varepsilon M^{d(\frac{1}{p}-\frac{1}{2})}.$$

By combining these estimates, we obtain

$$\|\mathbf{u} - \mathbf{v}\|_{\ell^p_m} \lesssim m(x) \left(M^{t+d\max(0,\frac{1}{p}-\frac{1}{2})}\varepsilon + M^{t-s+\frac{d}{p}} \|\mathbf{u}\|_{s,x} \right).$$

■

Theorem 3.13. *If* $\mathbf{f} \in \ell^\infty_{w_s}$ *is localized at* x*, then under the assumptions and notations of Theorem 3.11* **SOLVE** *converges in* ℓ^p_m *for* $1 \le p \le \infty$ *and every* t*-moderate weight with* $t < \frac{s^2}{\tilde\sigma\tilde{\tilde s}} - \frac{ds}{2\sigma\tilde{\tilde s}} - d\max\left\{\frac{1}{p} - \frac{1}{2}, 0\right\}$*. The error can be estimated by*

$$\|\mathbf{P}(\mathbf{u} - \mathbf{u}_\varepsilon)\|_{\ell^p_m} \lesssim \varepsilon^{1-\frac{\tilde\sigma\tilde{\tilde s}}{s}\left(\frac{t}{rd}+\frac{1}{r}\max(0,\frac{1}{p}-\frac{1}{2})\right)} + \varepsilon^{\frac{\tilde\sigma\tilde{\tilde s}}{s}\left(\frac{s-t}{rd}-\frac{1}{rp}\right)+\left(1-\frac{\tilde\sigma\tilde{\tilde s}}{s}\right)}.$$

Proof. By Corollary 3.10, the solution $\mathbf{u} = \mathbf{P}\mathbf{u}$ of (45) belongs to $\ell^\infty_{w_s}$. Let $\mathbf{u}_\varepsilon = \mathbf{SOLVE}(\varepsilon, \mathbf{A}, \mathbf{f})$ be the output of **Algorithm 1**, and set

$$\mathbf{v} = \mathbf{RHS}(\varepsilon, \mathbf{P}\mathbf{u}_\varepsilon).$$

By Theorem 3.11 $\mathbf{u}_\varepsilon$ is localized at x, and by Theorem 3.9 $\mathbf{Pu}_\varepsilon$ is also localized at x. We write

$$\mathbf{P}(\mathbf{u}-\mathbf{u}_\varepsilon) = \mathbf{Pu} - \mathbf{v} + \mathbf{v} - \mathbf{Pu}_\varepsilon.$$

Then, by the properties of **RHS**,

$$\|\mathbf{v} - \mathbf{Pu}_\varepsilon\|_{\ell^2} < \varepsilon,$$

and likewise

$$\|\mathbf{Pu} - \mathbf{v}\|_{\ell^2} \leq \|\mathbf{Pu} - \mathbf{Pu}_\varepsilon\|_{\ell^2} + \|\mathbf{Pu}_\varepsilon - \mathbf{v}\|_{\ell^2} < 2\varepsilon.$$

Here we have used the fact that $\mathbf{u}_\varepsilon$ is the outcome of the **SOLVE** algorithm and Theorem 3.11 (a). Furthermore $\operatorname{supp}\mathbf{v} \subset \{k \ : \ |k-x| \leq M\}$, where, by the properties of **RHS**

$$\#\operatorname{supp}\mathbf{v} \lesssim M^d \lesssim \varepsilon^{-\frac{1}{r}}\|\mathbf{Pu}_\varepsilon\|_{s,x}^{\frac{1}{r}}.$$

Since, by Theorem 3.11 (d),

$$\varepsilon^{\frac{\tilde\sigma\tilde{\tilde{s}}}{s}-1}\|\mathbf{Pu}_\varepsilon\|_{s,x} \lesssim \varepsilon^{\frac{\tilde\sigma\tilde{\tilde{s}}}{s}-1}\|\mathbf{Pu}_\varepsilon\|_{\tilde{s},x} \lesssim \varepsilon^{\frac{\tilde\sigma\tilde{\tilde{s}}}{s}-1}\|\mathbf{u}_\varepsilon\|_{\tilde{s},x} \lesssim \max\{\|\mathbf{u}\|_{s,x}^{1/\tilde\sigma}, \|\mathbf{u}\|_{s,x}\}^{\frac{\tilde\sigma\tilde{\tilde{s}}}{s}},$$

or, equivalently,

$$\|\mathbf{Pu}_\varepsilon\|_{s,x} \lesssim (\varepsilon^{-1}\max\{\|\mathbf{u}\|_{s,x}^{1/\tilde\sigma}, \|\mathbf{u}\|_{s,x}\})^{\frac{\tilde\sigma\tilde{\tilde{s}}}{s}-1}\max\{\|\mathbf{u}\|_{s,x}^{1/\tilde\sigma}, \|\mathbf{u}\|_{s,x}\},$$

the estimate for M becomes

$$M \lesssim \varepsilon^{-\frac{1}{rd}\frac{\tilde\sigma\tilde{\tilde{s}}}{s}}\max\{\|\mathbf{u}\|_{s,x}^{1/\tilde\sigma}, \|\mathbf{u}\|_{s,x}\}_{s,x}^{\frac{1}{rd}\frac{\tilde\sigma\tilde{\tilde{s}}}{s}}.$$

Both $\mathbf{Pu}$ and $\mathbf{Pu}_\varepsilon$ satisfy the assumptions of Lemma 3.12, hence the conclusion for ℓ^p_m is

$$\begin{aligned}
\|\mathbf{P}(\mathbf{u}-\mathbf{u}_\varepsilon)\|_{\ell^p_m} &\leq \|\mathbf{Pu}-\mathbf{v}\|_{\ell^p_m} + \|\mathbf{v}-\mathbf{Pu}_\varepsilon\|_{\ell^p_m} \\
&\lesssim m(x)\Big((\max\{\|\mathbf{u}\|_{s,x}^{1/\tilde\sigma}, \|\mathbf{u}\|_{s,x}\}\varepsilon^{-1})^{\frac{1}{rd}\frac{\tilde\sigma\tilde{\tilde{s}}}{s}(t+d\max(0,\frac{1}{p}-\frac{1}{2}))}\varepsilon \\
&\qquad +(\varepsilon^{-1}\max\{\|\mathbf{u}\|_{s,x}^{1/\tilde\sigma}, \|\mathbf{u}\|_{s,x}\})^{\frac{\tilde\sigma\tilde{\tilde{s}}}{s}\left(\frac{t-s}{rd}+\frac{1}{rp}\right)+\left(\frac{\tilde\sigma\tilde{\tilde{s}}}{s}-1\right)}\max\{\|\mathbf{u}\|_{s,x}^{1/\tilde\sigma}, \|\mathbf{u}\|_{s,x}\}\Big) \\
&\lesssim \varepsilon^{1-\frac{\tilde\sigma\tilde{\tilde{s}}}{s}\left(\frac{t}{rd}+\frac{1}{r}\max(\frac{1}{p}-\frac{1}{2},0)\right)} + \varepsilon^{\frac{\tilde\sigma\tilde{\tilde{s}}}{s}\left(\frac{s-t}{rd}-\frac{1}{rp}\right)+\left(1-\frac{\tilde\sigma\tilde{\tilde{s}}}{s}\right)}.
\end{aligned}$$

This expression tends to 0, precisely when $t < \frac{s^2}{\tilde\sigma\tilde{\tilde{s}}} - \frac{ds}{2\sigma\tilde{\tilde{s}}} - d\max\left\{\frac{1}{p}-\frac{1}{2}, 0\right\}$. ■

REMARKS: 1. For highly localized vectors, i.e., for $\tilde{s} \approx s > 0$ very large, we have $\tilde\sigma \approx 1 \approx \frac{\tilde{\tilde{s}}}{s}$ and the approximation $t < \frac{s^2}{\tilde\sigma\tilde{\tilde{s}}} - \frac{ds}{2\sigma\tilde{\tilde{s}}} - d\max\left\{\frac{1}{p}-\frac{1}{2}, 0\right\} \approx s - \frac{d}{p}$.

2. Once again, Theorem 3.13 relies on the specific structure of the algebra $\mathcal{A}_\gamma$, and it illuminates another important difference to adaptive schemes on $\ell^{\tau,w}$. For these schemes, the convergence is only guaranteed in ℓ^2, and to our knowledge no result is known about the convergence in stronger norms.

4 Convergence of the Frame Algorithm

In this section we apply **Algorithm 1** to the frame operator and the efficient approximation of the canonical dual frame. The hypotheses of **Algorithm 1** are a perfect match for the class of (intrinsically) localized frames. The key point is that this algorithm produces an approximation in all associated Banach spaces $\mathcal{H}^p_m(\mathcal{G}, \tilde{\mathcal{G}})$. This removes a serious restriction of the previous contribution [12, Section 7], where the approximation worked only for the underlying Hilbert space $\mathcal{H}$.

4.1 The Discretized Frame Algorithm

To find the canonical dual frame $\tilde{f}_n = S^{-1} f_n, n \in \mathcal{N}$, we have to solve the equation

$$Su = f_n \qquad \text{for all} \quad n \in \mathcal{N}. \tag{51}$$

Since the frame operator is positive and boundedly invertible on $\mathcal{H}$, we are exactly in the setting of the Subsections 2.2 and 3.4, so that the entire machinery developed so far can be applied. Consequently, by discretizing (51) by means of a second frame $\mathcal{G} = \{g_n\}_{n\in\mathcal{N}}$, we end up with the bi-infinite matrix equation

$$\mathbf{A}\mathbf{u} = \mathbf{f}_n, \tag{52}$$

where

$$\mathbf{A} = FSF^* = (\langle Sg_n, g_m \rangle)_{n,m}. \tag{53}$$

However, to fully exploit the theory of Subsection 3.4, we need to impose further conditions on $\mathbf{A}$. The following lemma establishes the link between the localization of a frame and the almost diagonalization of $\mathbf{A}$, see also [15, 19].

Lemma 4.1. *Assume that $\mathcal{F} \sim_{\mathcal{A}_\gamma} \mathcal{G}$ for some $\gamma > d$. Then $\mathbf{A} \in \mathcal{A}_\gamma$.*

Proof. By hypothesis, the (cross) Gramian $\mathbf{C} = A(\mathcal{G}, \mathcal{F})$ of $\mathcal{G}$ and $\mathcal{F}$ with entries $\mathbf{C}_{l,n} = \langle g_n, f_l \rangle$ is contained in $\mathcal{A}_\gamma$. Rewriting $\mathbf{A}$ as

$$\mathbf{A}_{m,n} = \langle Sg_n, g_m \rangle = \sum_{l\in\mathcal{N}} \langle g_n, f_l \rangle \langle f_l, g_m \rangle = (\mathbf{C}^*\mathbf{C})_{m,n},$$

we see that $\mathbf{A} = \mathbf{C}^*\mathbf{C}$. Since $\mathcal{A}_\gamma$ is a Banach $*$-algebra, we obtain $\mathbf{A} \in \mathcal{A}_\gamma$. ∎

As a first consequence of Lemma 4.1, we show that the standard frame algorithm converges in many norms beside $\mathcal{H}$.

Theorem 4.2. *(a) If $\mathcal{F}, \mathcal{G}$ are two frames for $\mathcal{H}$, then the canonical dual of $\mathcal{F}$ can be computed by*

$$\tilde{f}_n = F^* \tilde{\mathbf{f}}_n = \sum_{l\in\mathcal{N}} \left(\tilde{\mathbf{f}}_n\right)_l g_l, \quad \tilde{\mathbf{f}}_n = \left(\alpha \sum_{n=0}^{\infty} (\mathrm{id} - \alpha\mathbf{A})^n\right) \mathbf{f}_n, \tag{54}$$

for $0 < \alpha < \frac{2}{\|\mathbf{A}\|}$.

(b) If $\mathcal{F}, \mathcal{G}$ are both intrinsically $\mathcal{A}_\gamma$-localized and $\mathcal{F} \sim_{\mathcal{A}_\gamma} \mathcal{G}$ for some $\gamma > d$, then the series in (54) *converges in the ℓ^p_m-norm on* $\operatorname{ran}_{\ell^p_m}(F) = \Big\{\mathbf{c} \in \ell^p_m : \exists f \in \mathcal{H}^p_m(\mathcal{G}, \tilde{\mathcal{G}}), (\langle f, g_n\rangle)_n = \mathbf{c}\Big\}$*, for all $1 \le p \le \infty$ and every s-moderate weight m with $s < \gamma - d$. Consequently, $\tilde{\mathbf{f}}_n \in \ell^p_m(\mathcal{N})$ and $\tilde{f}_n \in \mathcal{H}^p_m(\mathcal{G}, \tilde{\mathcal{G}})$ for all $n \in \mathcal{N}$.*

Proof. (a) is a consequence of Theorem 2.2. It remains to show (b). By Theorem 3.9, the orthogonal projector $\mathbf{P}$ onto $\operatorname{ran}_{\ell^2}(F) = \operatorname{ran}(\mathbf{A})$ is contained in $\mathcal{A}_\gamma$, hence $\mathbf{P}$ is bounded on ℓ^p_m for every s-moderate weight m.

Let $\sigma_{\operatorname{ran}_{\ell^p_m}(F)}(\mathbf{A})$ be the spectrum of $\mathbf{A}$ acting on $\operatorname{ran}_{\ell^p_m}(F)$ and $r_{\operatorname{ran}_{\ell^p_m}(F)}(\mathbf{A}) := \max\{|\lambda| : \lambda \in \sigma_{\operatorname{ran}_{\ell^p_m}(F)}\}$ the spectral radius. If $\lambda \notin \sigma_{\operatorname{ran}(F)}(\mathbf{A})$ then $\mathbf{A} - \lambda\mathbf{P}$ is invertible on $\operatorname{ran}(F)$ and by [15, Theorem 3.4] there exists $(\mathbf{A} - \lambda\mathbf{P})^\dagger \in \mathcal{A}_\gamma$ such that

$$(\mathbf{A} - \lambda\mathbf{P})^\dagger(\mathbf{A} - \lambda\mathbf{P}) = (\mathbf{A} - \lambda\mathbf{P})(\mathbf{A} - \lambda\mathbf{P})^\dagger = \mathbf{P}. \tag{55}$$

Since $\mathcal{A}_\gamma \subset L(\ell^p_m)$, (55) also holds as an identity of operators on ℓ^p_m for $1 \le p \le \infty$ and all s-moderate weights. Restricting (55) to the invariant subspace $\operatorname{ran}_{\ell^p_m}(F)$, we see that

$$\lambda \notin \sigma_{\operatorname{ran}_{\ell^p_m}(F)}(\mathbf{A}) \qquad \text{and so} \qquad \sigma_{\operatorname{ran}_{\ell^p_m}(F)}(\mathbf{A}) \subseteq \sigma_{\operatorname{ran}(F)}(\mathbf{A}). \tag{56}$$

Applying (56) to $\mathbf{P} - \alpha\mathbf{A}$, we find

$$r_{\operatorname{ran}_{\ell^p_m}(F)}(\mathbf{P} - \alpha\mathbf{A}) \le r_{\operatorname{ran}(F)}(\mathbf{P} - \alpha\mathbf{A}) < 1$$

(by our choice of $\alpha < 2/\|\mathbf{A}\|_{\ell^2 \to \ell^2}$). Consequently, the geometric series $\sum_{n=0}^\infty (\mathbf{P} - \alpha\mathbf{A})^n$ converges on $\operatorname{ran}_{\ell^p_m}(F)$.

If $\mathcal{F} \sim_{\mathcal{A}_\gamma} \mathcal{G}$, then $\mathbf{f}_n = (\langle f_n, g_m\rangle)_{m \in \mathcal{N}} \in \operatorname{ran}_{\ell^p_m}(F)$, and thus $\tilde{\mathbf{f}}_n \in \operatorname{ran}_{\ell^p_m}(F)$ or equivalently $\tilde{f}_n \in \mathcal{H}^p_m(\mathcal{G}, \tilde{\mathcal{G}})$. ■

Next we show that the adaptive numerical schemes discussed in Subsection 3.4 can again be applied to approximate the infinite series in (54) up to a given precision.

Theorem 4.3. *Assume $\tilde{s} > s > d$, $\tilde{s} + d < \gamma$, $r = \frac{s}{d} - \frac{1}{2}$, $\tilde{\sigma} = \frac{\tilde{s}}{\tilde{r}d}$, and $\tilde{\tilde{s}} = s\left(\frac{\tilde{s} - s}{s - d/2} + 1\right)$. Let $\mathcal{F}, \mathcal{G}$ be intrinsically $\mathcal{A}_\gamma$-localized frames, $\mathcal{F} \sim_{\mathcal{A}_\gamma} \mathcal{G}$ and $\varepsilon > 0$.*

(A) Assume that $\mathbf{f}_n$ is localized at n. Then the finite vector

$$\tilde{\mathbf{f}}_{n,\varepsilon} = \mathbf{SOLVE}[\varepsilon, \mathbf{A}, \mathbf{f}_n], \tag{57}$$

has the following properties:

(a) $\|\mathbf{P}(\tilde{\mathbf{f}}_n - \tilde{\mathbf{f}}_{n,\varepsilon})\|_{\ell^2} \le \varepsilon$, where $\tilde{\mathbf{f}}_n$ is the solution to (52);

(b) $\operatorname{supp} \tilde{\mathbf{f}}_{n,\varepsilon} \subseteq B(n, N)$ and $\# \operatorname{supp} \tilde{\mathbf{f}}_{n,\varepsilon} \lesssim N^d \lesssim \varepsilon^{-\tilde{\sigma}/r} \max\{\|\tilde{\mathbf{f}}_n\|_{s,n}^{1/\tilde{\sigma}}, \|\tilde{\mathbf{f}}_n\|_{s,n}\}^{\tilde{\sigma}/r}$;

(c) $\mathcal{C}(\tilde{\mathbf{f}}_{n,\varepsilon}) \lesssim \varepsilon^{-(\frac{\tilde{\sigma}}{r}+\xi(\tilde{\sigma}-1)\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}+\delta)} \max\{\|\tilde{\mathbf{f}}_n\|^{1/\tilde{\sigma}}\|\tilde{\mathbf{f}}_n\|\}^{\frac{\tilde{\sigma}}{r}+\xi(\tilde{\sigma}-1)\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}+\delta}$, *for* $\delta > 0$ *arbitrarily small, and some constant* $\xi > 0$.

(d) $\varepsilon^{\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}-1}\|\tilde{\mathbf{f}}_{n,\varepsilon}\|_{\tilde{s},x} \lesssim \max\{\|\tilde{\mathbf{f}}_n\|_{s,x}^{1/\tilde{\sigma}}, \|\tilde{\mathbf{f}}_n\|_{s,x}\}^{\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}}$.

Therefore, one has the following approximation of the canonical dual

$$\left\|\tilde{f}_n - \sum_m \left(\tilde{\mathbf{f}}_{n,\varepsilon}\right)_m g_m\right\|_{\mathcal{H}} \leq B_{\mathcal{G}}^{\frac{1}{2}}\varepsilon. \tag{58}$$

(B) **SOLVE** *converges in* ℓ_m^p *for* $1 \leq p \leq \infty$ *and every* t*-moderate weight with* $t < \frac{s^2}{\tilde{\sigma}\tilde{\tilde{s}}} - \frac{ds}{2\sigma\tilde{\tilde{s}}} - d\max\left\{\frac{1}{p}-\frac{1}{2}, 0\right\}$. *The error can be estimated by*

$$\left\|\tilde{f}_n - \sum_m \left(\tilde{\mathbf{f}}_{n,\varepsilon}\right)_m g_m\right\|_{\mathcal{H}_m^p} \lesssim \|\mathbf{P}(\tilde{\mathbf{f}}_n - \tilde{\mathbf{f}}_{n,\varepsilon})\|_{\ell_m^p} \lesssim \varepsilon^{1-\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}\left(\frac{t}{rd}+\frac{1}{r}\max(0,\frac{1}{p}-\frac{1}{2})\right)} + \varepsilon^{\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}\left(\frac{s-t}{rd}-\frac{1}{rp}\right)+\left(1-\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}\right)}. \tag{59}$$

Proof. (A) and (B) follow immediately by applying Theorem 3.11 and Theorem 3.13, respectively.

■

REMARKS:

(i) This last theorem not only ensures the convergence of the procedure to the canonical dual in the $\mathcal{H}$ norm, but also in the norm of $\mathcal{H}_m^p$ for $1 \leq p \leq \infty$ and for certain t-moderate weights m.

(ii) Let us remark again that, although the error estimate (59) is stated for the projected error, this does not destroy the convergence of the scheme. Moreover, it is *not* necessary to have an explicit knowledge of $\mathbf{P}$.

Example 1 (Gabor frames)**.** Let $z = (x, \omega) \in \mathbb{R}^{2d}$ and

$$\pi(z)f(t) = e^{2\pi i\omega\cdot t}f(t-x) \quad t, x, \omega \in \mathbb{R}^d$$

be the time-frequency shift of the function f by $z \in \mathbb{R}^{2d}$. Assume that $\mathcal{X}$ is a relatively separated subset of $\mathbb{R}^{2d}$ and that $g \in \mathcal{S}(\mathbb{R}^d) \setminus \{0\}$ (or that g possesses sufficient time-frequency localization). If the set $\mathcal{G} = \mathcal{G}(g, \mathcal{X}) := \{\pi(z)g : z \in \mathcal{X}\}$ generates a frame (called a Gabor frame), then $\mathcal{G}$ is intrinsically $\mathcal{A}_\gamma$-localized for any $\gamma > d$ and thus it possesses an intrinsically $\mathcal{A}_\gamma$-localized canonical dual $\tilde{\mathcal{G}} = \{\tilde{g}_z : z \in \mathcal{X}\}$ by the results in [15, 19]. Therefore, an approximation of the canonical dual frame can be computed by using Theorem 4.3.

In this case, both $\mathcal{G} = \mathcal{G}(g, \mathcal{X}) := \{\pi(z)g : z \in \mathcal{X}\}$ and $\tilde{\mathcal{G}}$ form a Banach frame for the class of modulation spaces $M_m^{p,q}$ for any s-moderate weight with $s + d <$

γ [18, Chpt. 13]. If $p = q$, then $M_m^{p,p}$ coincides with the abstract Banach space $\mathcal{H}_m^p(\mathcal{G}, \tilde{\mathcal{G}})$ [15]. Since for a suitable weight m, $M_m^{2,2}(\mathbb{R}^d)$ coincides with weighted L^2-spaces and also with $H^t(\mathbb{R}^d)$, the L^2-Sobolev space of Sobolev smoothness t [18, Thm. 11.3.1], the approximation in (59) ensures the convergence of derivatives and convergence in weighted L^2-spaces.

Another possible application of Algorithm 1 is the fast approximate reconstruction of functions in shift-invariant spaces, because the theory of localized frames and hence our main theorems are applicable. For details about sampling theory see [1, 16, 19].

5 Error Estimates

So far we have shown how to approximate a single vector of the dual frame by applying the **SOLVE**-routine. More precisely, if $\mathbf{A}$ is the matrix of the frame operator with respect to a frame $\mathcal{G}$ as defined in (53) and $\mathbf{f}_n = (\langle f_n, g_m \rangle)_m$, we compute a sequence of finitely supported vectors $\boldsymbol{\gamma}_n = (\boldsymbol{\gamma}_{nm})_m$ by

$$\boldsymbol{\gamma}_n = \boldsymbol{\gamma}_n^{\varepsilon} = \mathbf{SOLVE}[\varepsilon, \mathbf{A}, \mathbf{f}_n] \,. \tag{60}$$

In this section we assume two additional technical conditions:

(a1) $\mathcal{G}$ is a Riesz basis, hence the projection $\mathbf{P}$ is just the identity operator. This condition seems unavoidable and the following results could not be shown for $\mathcal{G}$ being an arbitrary frame;

(a2) $\sup_{n \in \mathcal{N}} \|\mathbf{f}_n\|_{s,n} < \infty$. This condition just normalizes the precision ε we use to compute individual elements of the canonical dual. However, this condition is also natural in certain situations as for Gabor frames, where the elements of the frame have a uniform "localization envelope".

Under these conditions and the assumptions of Theorem 3.11, we have

$$\begin{aligned}
\operatorname{supp} \boldsymbol{\gamma}_n &\subseteq \{m : |m - n| \le N\} && (61)\\
\#\operatorname{supp} \boldsymbol{\gamma}_n &\asymp N^d \lesssim \varepsilon^{-\tilde{\sigma}/r} \max\{\|\tilde{\mathbf{f}}_n\|_{s,n}^{1/\tilde{\sigma}} \|\tilde{\mathbf{f}}_n\|_{s,n}\}^{\tilde{\sigma}/r} \lesssim \varepsilon^{-\tilde{\sigma}/r} && (62)\\
\|\tilde{\mathbf{f}}_n - \boldsymbol{\gamma}_n\|_{\ell^2} &\le \varepsilon \,. && (63)
\end{aligned}$$

Note that the projection $\mathbf{P}$ appears no longer in condition (63), since we assume that $\mathcal{G}$ is a Riesz basis. Setting $\tilde{f}_n^{\varepsilon} = \sum_m \boldsymbol{\gamma}_{nm} g_m$ for the approximate dual of $\tilde{f}_n = \sum_m \langle \tilde{f}_n, \tilde{g}_m \rangle g_m$, we then have the individual error estimate $\|\tilde{f}_n - \tilde{f}_n^{\varepsilon}\|_{\mathcal{H}} \lesssim \varepsilon$ for each n. For the solution of a single operator equation $\mathbf{Au} = \mathbf{f}_n$ such an estimate is good enough. However, when approximating the dual frame of $\mathcal{F}$, we need to know much more about the collection $\mathcal{F}^{\varepsilon} = \{\tilde{f}_n^{\varepsilon} : n \in \mathcal{N}\}$. In particular, we need to compare the exact frame expansion $f = \sum_n \langle f, f_n \rangle \tilde{f}_n$ with the approximate expansion

$$f^{\varepsilon} = \sum_{n \in \mathcal{N}} \langle f, f_n \rangle \tilde{f}_n^{\varepsilon}, \tag{64}$$

and derive error estimates, if possible. A priori, it is not at all clear whether $\mathcal{F}^\varepsilon$ is again a frame or a Banach frame or a set for an atomic decomposition. In general, an estimate $\|\tilde{f}_n - \tilde{f}_n^\varepsilon\| \lesssim \varepsilon$ for all n does not guarantee that $\mathcal{F}^\varepsilon$ is again a frame. Once again the crucial property is a localization property, this time in the form (61)-(63).

We first prove a small lemma, then derive an error estimate for $\|f - f^\varepsilon\|$, and finally apply the perturbation theory of (Banach) frames [5, 6] to show that $\mathcal{F}^\varepsilon$ is also a frame.

Lemma 5.1. *Assume that* $\mathbf{A}$ *is a banded matrix, such that* $a_{kl} = 0$ *for* $|k-l| > N$ *and* $|a_{kl}| \le \varepsilon$ *for* $|k-l| \le N$. *If* m *is a* t*-moderate weight, then the operator norm of* $\mathbf{A}$ *on* ℓ^p_m *is majorized by*

$$\|\mathbf{A}\|_{\ell^p_m \to \ell^p_m} \lesssim \varepsilon N^{t+d}\,. \tag{65}$$

Proof. Using a naive estimate and Hölder's inequality, we find that

$$\begin{aligned} |(\mathbf{A}\mathbf{c})(k)| &= \Big| \sum_{l:|l-k|\le N} a_{kl} c_l \Big| \\ &\le \varepsilon \sum_{l:|l-k|\le N} |c_l| \\ &\lesssim \varepsilon \Big(\sum_{l:|l-k|\le N} |c_l|^p \Big)^{1/p} N^{d/p'}\,. \end{aligned}$$

So the ℓ^p_m-norm of $\mathbf{A}\mathbf{c}$ is bounded by

$$\begin{aligned} \|\mathbf{A}\mathbf{c}\|^p_{\ell^p_m} &= \sum_k |\mathbf{A}\mathbf{c}(k)|^p\, m(k)^p \\ &\lesssim \varepsilon^p\, N^{dp/p'} \sum_k \big(\sum_{l:|l-k|\le N} |c_l|^p \big)\, m(k-l+l)^p \\ &\lesssim \varepsilon^p\, N^{dp/p'} \sum_l |c_l|^p m(l)^p \big(\sum_{k:|k-l|\le N} (1+|k-l|)^{tp} \big) \\ &\lesssim \varepsilon^p\, N^{dp/p'}\, N^{tp+d}\, \|\mathbf{c}\|^p_{\ell^p_m}\,. \end{aligned}$$

Taking the p-th root, we obtain

$$\|\mathbf{A}\|_{\ell^p_m \to \ell^p_m} \lesssim \varepsilon N^{d/p' + t + d/p} = \varepsilon N^{t+d}\,.$$

■

Theorem 5.2. *Assume that* $\mathcal{G}$ *and* $\mathcal{F}$ *are* $\mathcal{A}_\gamma$*-intrinsically localized frames and* $\mathcal{F} \sim_{\mathcal{A}_\gamma} \mathcal{G}$ *for* $\gamma > s+d$. *Also assume that* $\mathcal{G}$ *is a Riesz basis. Let* $1 \le p \le \infty$ *and* m *be a* t*-moderate weight for* $t < \frac{s-d/2}{\tilde{\sigma}} - d$.

If $\mathcal{F}^\varepsilon$ *is a set satisfying conditions* (61)-(63), *then*

$$\|f - f^\varepsilon\|_{\mathcal{H}^p_m} \lesssim \Big(\varepsilon^{(\gamma-t-d)\tilde{\sigma}/(rd)} + \varepsilon^{1-\frac{\tilde{\sigma}(t+d)}{rd}} \Big) \|f\|_{\mathcal{H}^p_m} \qquad \textit{for all } f \in \mathcal{H}^p_m\,. \tag{66}$$

Proof. We first look at the difference $f - f^\varepsilon$ in detail. We expand both $\tilde{f}_n$ and $\tilde{f}_n^\varepsilon$ with respect to the frame $\mathcal{G}$ and obtain

$$\begin{aligned} f - f^\varepsilon &= \sum_n \langle f, f_n \rangle (\tilde{f}_n - \tilde{f}_n^\varepsilon) \\ &= \sum_n \langle f, f_n \rangle \Big(\sum_m (\langle \tilde{f}_n, \tilde{g}_m \rangle - \gamma_{nm}) g_m \Big) \\ &= \sum_m \Big(\sum_n \langle f, f_n \rangle (\langle \tilde{f}_n, \tilde{g}_m \rangle - \gamma_{nm}) \Big) g_m \,. \end{aligned}$$

The computation of the coefficients of g_m involves the cross-Gramian $\mathbf{C} = A(\tilde{\mathcal{F}}, \tilde{\mathcal{G}})$ with entries $\mathbf{C}_{mn} = \langle \tilde{f}_n, \tilde{g}_m \rangle$ and the banded matrix $\mathbf{\Gamma}$ with entries γ_{mn}. Thus we can write the error as

$$f - f^\varepsilon = \sum_m \Big((\mathbf{C} - \mathbf{\Gamma}) F_{\mathcal{F}} f \Big)(m) \, g_m = F_{\mathcal{G}}^* (\mathbf{C} - \mathbf{\Gamma}) F_{\mathcal{F}} f \,. \tag{67}$$

The initial estimates now follow from the fact that $\mathcal{G}$ is a Banach frame for $\mathcal{H}_m^p$ and the assumption $\mathcal{F} \sim_{\mathcal{A}_\gamma} \mathcal{G}$. On the one hand we know that $\|F_{\mathcal{G}}^* \mathbf{c}\|_{\mathcal{H}_m^p} \lesssim \|\mathbf{c}\|_{\ell_m^p}$ for any s-moderate weight function m with $s < \gamma - d$ by [19, Proposition 8 (b)]. On the other hand, we have the norm equivalence $\|f\|_{\mathcal{H}_m^p} := \|F_{\mathcal{G}} f\|_{\ell_m^p} \asymp \|F_{\mathcal{F}} f\|_{\ell_m^p}$ by [15, Proposition 2.4]. In addition, we know that the cross-Gramian $\mathbf{C} = \mathbf{A}(\tilde{\mathcal{F}}, \tilde{\mathcal{G}})$ is in $\mathcal{A}_\gamma$, whence follows the boundedness of $\mathbf{C}$ on ℓ_m^p for the same class of weights and $1 \leq p \leq \infty$ by Lemma 3.5(b). Likewise the banded matrix $\mathbf{\Gamma}$ is bounded on any ℓ_m^p by Lemma 5.1. Thus all steps of the following estimate are well defined.

$$\begin{aligned} \|f - f^\varepsilon\|_{\mathcal{H}_m^p} &\lesssim \|(\mathbf{C} - \mathbf{\Gamma}) F_{\mathcal{F}} f\|_{\ell_m^p} \\ &\lesssim \|\mathbf{C} - \mathbf{\Gamma}\|_{\ell_m^p \to \ell_m^p} \|F_{\mathcal{F}} f\|_{\ell_m^p} \\ &\lesssim \|\mathbf{C} - \mathbf{\Gamma}\|_{\ell_m^p \to \ell_m^p} \|f\|_{\mathcal{H}_m^p} \,. \end{aligned}$$

This estimate reveals the key issue arising in the error analysis. We need a good bound on the operator norm of $\mathbf{C} - \mathbf{\Gamma}$. The necessary preparations have already been accomplished in Lemma 5.1 and 3.5(b). As in Lemma 3.5 we approximate $\mathbf{C}$ by a banded matrix $\mathbf{B}^N$ with entries $\mathbf{B}_{kl}^N = \mathbf{C}_{kl}$ for $|k - l| \leq N$ and $\mathbf{B}_{kl}^N = 0$ for $|k - l| > N$. Then $\mathbf{C} - \mathbf{\Gamma} = \mathbf{C} - \mathbf{B}^N + \mathbf{B}^N - \mathbf{\Gamma}$.

Lemma 3.5(b) and (62) imply that

$$\|\mathbf{C} - \mathbf{B}^N\|_{\ell_m^p \to \ell_m^p} \lesssim N^{d+t-\gamma} \lesssim \varepsilon^{(\gamma - t - d)\tilde{\sigma}/rd} \,.$$

For the banded part $\mathbf{B}^N - \mathbf{\Gamma}$ we note that $|\langle \tilde{f}_n, \tilde{g}_m \rangle - \gamma_{mn}| \leq \|\tilde{\mathbf{f}}_n - \boldsymbol{\gamma}_n\|_{\ell^2} \leq \varepsilon$ by construction (63). Thus all non-zero entries of $\mathbf{B}^N - \mathbf{\Gamma}$ are bounded by ε. Consequently Lemma 5.1 implies that

$$\|\mathbf{B}^N - \mathbf{\Gamma}\|_{\ell_m^p \to \ell_m^p} \lesssim \varepsilon N^{t+d} \lesssim \varepsilon^{1 - \frac{\tilde{\sigma}(t+d)}{rd}} \,,$$

and we have $\|\mathbf{C}-\mathbf{\Gamma}\|_{\ell^p_m\to\ell^p_m}\lesssim\varepsilon^{(\gamma-t-d)\tilde\sigma/rd}+\varepsilon^{1-\frac{\tilde\sigma(t+d)}{rd}}$. For convergence, as $\varepsilon\to 0$, we need that the exponents are positive. Since $\frac{rd}{\tilde\sigma}=\frac{s-d/2}{\tilde\sigma}>t+d$ by assumption, we have $1-\frac{\tilde\sigma(t+d)}{rd}>0$, and obviously $\gamma-t-d>0$. Thus the statement is proved. ■

REMARK: Note that in the proof of Theorem 5.2 we have used only established estimates for localized frames and the properties (61)–(63) for the approximate dual frame. We have not used any special features of the **SOLVE**-algorithm. Therefore the error analysis is valid for any approximation of the dual frame satisfying (61)–(63). The virtue of **SOLVE** is to provide a practical numerical method for the approximation of the dual frame.

Corollary 5.3. *For $\varepsilon>0$ small enough $\mathcal{F}^\varepsilon$ provides an atomic decomposition for $\mathcal{H}^p_m$.*

Proof. Set $A_\varepsilon f=f^\varepsilon=\sum_n\langle f,f_n\rangle\tilde f^\varepsilon_n$. Theorem 5.2 implies that $\|\operatorname{id}-A_\varepsilon\|_{\mathcal{H}^p_m\to\mathcal{H}^p_m}<1$ for ε small enough. Then A_ε is invertible on $\mathcal{H}^p_m$. The factorization $f=A_\varepsilon A_\varepsilon^{-1}f=\sum_n\langle A_\varepsilon^{-1}f,f_n\rangle\tilde f^\varepsilon_n$ and the unconditional convergence of this sum together imply that $\mathcal{F}^\varepsilon$ provides an atomic decomposition for $\mathcal{H}^p_m$. See [6, Theorem 2.3] and its proof for more details. ■

6 Appendix

In this appendix we carry out the proof of Theorem 3.11.

Proof of Theorem 3.11

The estimate for the projected error (a) is stated in Proposition 3.8.

Proof of the quasi-boundedness (claim (d)). First of all, observe that for any vector $\mathbf{v}\in\ell^\infty_{w_{\tilde s}}(\mathcal{N})$ such that $\operatorname{supp}(\mathbf{v})\subset B(x,M)$ we have

$$\|\mathbf{v}\|_{\tilde s,x}\lesssim M^{\tilde s-s}\|\mathbf{v}\|_{s,x}\,. \tag{68}$$

Let us now choose M_i such that

$$\|\mathbf{Pu}-(\mathbf{Pu})^{M_i-nearest}\|_{\ell^2}\le\frac{\theta\varepsilon_i}{3}. \tag{69}$$

In particular, we have

$$M_i^d\lesssim\varepsilon_i^{-1/r}\|\mathbf{Pu}\|_{s,x}^{1/r}\lesssim\varepsilon_i^{-1/r}\|\mathbf{u}\|_{s,x}^{1/r}.$$

So, by (68) and $\|(\mathbf{Pu})^{M_i-nearest}\|_{\tilde s,x}\le\|\mathbf{Pu}\|_{\tilde s,x}$ we have

$$\varepsilon_i^{\frac{\tilde s}{s}-1}\|(\mathbf{Pu})^{M_i-nearest}\|_{\tilde s,x}\ \lesssim\ \varepsilon_i^{\frac{\tilde s}{s}-1}\left(\varepsilon_i^{-1/r}\|\mathbf{u}\|_{s,x}^{1/r}\right)^{\frac{\tilde s-s}{d}}\|\mathbf{Pu}\|_{s,x}\lesssim\|\mathbf{u}\|_{s,x}^{\frac{\tilde s}{s}}, \tag{70}$$

where we have used that $\frac{\tilde{s}-s}{d}\frac{1}{r} = \frac{\tilde{s}-s}{s-d/2} = \frac{\tilde{\tilde{s}}}{s} - 1$. Using $\|\mathbf{P}\mathbf{u} + (\mathrm{id} - \mathbf{P})\mathbf{u}^{(i-1)} - \mathbf{v}^{(i,K_i)}\|_{\ell^2} \leq 2\theta\varepsilon_i/3$ from Proposition 3.8, and formula (69), we have

$$\|(\mathbf{P}\mathbf{u})^{M_i-nearest} + (\mathrm{id} - \mathbf{P})\mathbf{u}^{(i-1)} - \mathbf{v}^{(i,K_i)}\|_{\ell^2} \leq \theta\varepsilon_i.$$

An application of Lemma 3.4, i.e., property (d) of **COARSE**, gives

$$\begin{aligned} &\|\mathbf{u}^{(i)}\|_{\tilde{s},x} \leq C(\tilde{s}) \max\big\{\|(\mathbf{P}\mathbf{u})^{M_i-nearest} + \\ &\qquad\qquad (\mathrm{id} - \mathbf{P})\mathbf{u}^{(i-1)}\|_{\tilde{s},x}, \varepsilon_i^{1-\tilde{\sigma}} \|(\mathbf{P}\mathbf{u})^{M_i-nearest} + (\mathrm{id} - \mathbf{P})\mathbf{u}^{(i-1)}\|_{\tilde{s},x}^{\tilde{\sigma}}\big\} \\ \leq\ & C(\tilde{s}) \max\big\{\|(\mathbf{P}\mathbf{u})^{M_i-nearest}\|_{\tilde{s},x} + \|(\mathrm{id} - \mathbf{P})\|\|\mathbf{u}^{(i-1)}\|_{\tilde{s},x}, \\ &\qquad\qquad \varepsilon_i^{1-\tilde{\sigma}} \left(\|(\mathbf{P}\mathbf{u})^{M_i-nearest}\|_{\tilde{s},x} + \|\mathrm{id} - \mathbf{P}\|\|\mathbf{u}^{(i-1)}\|_{\tilde{s},x}\right)^{\tilde{\sigma}}\Big\} \\ \leq\ & C(\tilde{s}) \max\big\{\|(\mathbf{P}\mathbf{u})^{M_i-nearest}\|_{\tilde{s},x} + \|(\mathrm{id} - \mathbf{P})\|\|\mathbf{u}^{(i-1)}\|_{\tilde{s},x}, \\ &\qquad\qquad 2^{\tilde{\sigma}-1}\varepsilon_i^{1-\tilde{\sigma}} \left(\|(\mathbf{P}\mathbf{u})^{M_i-nearest}\|_{\tilde{s},x}^{\tilde{\sigma}} + \|\mathrm{id} - \mathbf{P}\|^{\tilde{\sigma}}\|\mathbf{u}^{(i-1)}\|_{\tilde{s},x}^{\tilde{\sigma}}\right)\big\}. \end{aligned}$$

In the last inequality we used the estimate $(a+b)^{\tilde{\sigma}} \leq 2^{\tilde{\sigma}-1}(a^{\tilde{\sigma}} + b^{\tilde{\sigma}})$ for $a, b > 0$. Next we multiply the last chain of inequalities by $\varepsilon_i^{\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}-1}$. Since $\varepsilon_i < 1$ and thus $\varepsilon_i^{\sigma} \leq \varepsilon_i$, (70) yields that

$$\varepsilon_i^{\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}-1}\|(\mathbf{P}\mathbf{u})^{M_i-nearest}\|_{\tilde{s},x} \leq \varepsilon_i^{\frac{\tilde{\tilde{s}}}{s}-1}\|(\mathbf{P}\mathbf{u})^{M_i-nearest}\|_{\tilde{s},x} \lesssim \|\mathbf{u}\|_{s,x}^{\frac{\tilde{\tilde{s}}}{s}}.$$

We obtain that

$$\begin{aligned} &a_i := \varepsilon_i^{\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}-1}\|\mathbf{u}^{(i)}\|_{\tilde{s},x} \leq C(\tilde{s}) \max\Bigg\{\varepsilon_i^{\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}-1}\|(\mathbf{P}\mathbf{u})^{M_i-nearest}\|_{\tilde{s},x} + \varepsilon_i^{\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}-1}\|(\mathrm{id} - \mathbf{P})\|\|\mathbf{u}^{(i-1)}\|_{\tilde{s},x}, \\ &\qquad 2^{\tilde{\sigma}-1}\varepsilon_i^{1-\tilde{\sigma}}\varepsilon_i^{\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}-1}\|(\mathbf{P}\mathbf{u})^{M_i-nearest}\|_{\tilde{s},x}^{\tilde{\sigma}} + 2^{\tilde{\sigma}-1}\varepsilon_i^{1-\tilde{\sigma}}\varepsilon_i^{\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}-1}\|\mathrm{id} - \mathbf{P}\|^{\tilde{\sigma}}\|\mathbf{u}^{(i-1)}\|_{\tilde{s},x}^{\tilde{\sigma}}\Bigg\} \\ \leq\ & C_0 \max\{\|\mathbf{u}\|_{s,x}^{1/\tilde{\sigma}}, \|\mathbf{u}\|_{s,x}\}^{\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}} \\ &\underbrace{+ \max\left\{C(\tilde{s})\|\mathrm{id} - \mathbf{P}\|\left(\frac{3\rho^{K_i}}{\theta}\right)^{\tilde{\sigma}\frac{\tilde{\tilde{s}}}{s}-1}, 2^{\tilde{\sigma}-1}C(\tilde{s})\left(\frac{3\rho^{K_i}}{\theta}\right)^{\tilde{\sigma}(\frac{\tilde{\tilde{s}}}{s}-1)}\varepsilon_{i-1}^{1-\tilde{\sigma}}\|\mathbf{u}^{(i-1)}\|_{\tilde{s},x}^{\tilde{\sigma}-1}\|\mathrm{id} - \mathbf{P}\|^{\tilde{\sigma}}\right\}}_{=\eta_i} \\ &\qquad\qquad\qquad\qquad\qquad\qquad \times \left(\varepsilon_{i-1}^{\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}-1}\|\mathbf{u}^{(i-1)}\|_{\tilde{s},x}\right) \\ \leq\ & \underbrace{C_0 \max\{\|\mathbf{u}\|_{s,x}^{1/\tilde{\sigma}}, \|\mathbf{u}\|_{s,x}\}^{\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}}}_{:=b} + \eta_i \underbrace{\left(\varepsilon_{i-1}^{\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}-1}\|\mathbf{u}^{(i-1)}\|_{\tilde{s},x}\right)}_{:=a_{i-1}}. \end{aligned}$$

By definition, the number of iterations K_i of the inner loop is chosen precisely to make $\eta_i \leq \eta$. We may therefore rewrite this recursion in the abstract form

$$\begin{aligned} a_i &\leq b + \eta a_{i-1} \\ &\leq \left(\sum_{n=0}^{i-1} \eta^n\right) b + \eta^i a_0. \end{aligned}$$

Since $\mathbf{u}^{(0)} = 0$, we have $a_0 = 0$, and thus we obtain the quasi-boundedness

$$\varepsilon_i^{\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}-1}\|\mathbf{u}^{(i)}\|_{\tilde{s},x} \lesssim \max\{\|\mathbf{u}\|_{s,x}^{1/\tilde{\sigma}}, \|\mathbf{u}\|_{s,x}\}^{\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}}. \tag{71}$$

Control of the support of $\mathbf{u}_\varepsilon$ (Claim (b)). By property (b) of **COARSE** and by observing that $\frac{\tilde{r}}{r} = \frac{\tilde{\tilde{s}}}{s}$ we have

$$\begin{aligned}
\#\operatorname{supp}(\mathbf{u}^{(i)}) &\lesssim \varepsilon_i^{-\frac{1}{\tilde{r}}} \left(\|(\mathbf{Pu})^{M_i-nearest}\|_{\tilde{s},x} + \|\operatorname{id} - \mathbf{P}\|\|\mathbf{u}^{(i-1)}\|_{\tilde{s},x}\right)^{\frac{1}{\tilde{r}}} \\
&= \varepsilon_i^{-\frac{1}{r}} \left(\varepsilon_i^{\frac{\tilde{\tilde{s}}}{s}-1}\|(\mathbf{Pu})^{M_i-nearest}\|_{\tilde{s},x} + \varepsilon_i^{\frac{\tilde{\tilde{s}}}{s}-1}\|\operatorname{id} - \mathbf{P}\|\|\mathbf{u}^{(i-1)}\|_{\tilde{s},x}\right)^{\frac{1}{\tilde{r}}} \\
&\leq \varepsilon_i^{-\frac{\tilde{\sigma}}{r}} \left(\varepsilon_i^{\frac{\tilde{\tilde{s}}}{s}-1}\|(\mathbf{Pu})^{M_i-nearest}\|_{\tilde{s},x} + \varepsilon_i^{\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}-1}\|\operatorname{id} - \mathbf{P}\|\|\mathbf{u}^{(i-1)}\|_{\tilde{s},x}\right)^{\frac{1}{\tilde{r}}}.
\end{aligned}$$

In the latter inequality we used once more $\varepsilon_i^{\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}-1} \leq \varepsilon_i^{\frac{\tilde{\tilde{s}}}{s}-1}$. By (70) and (71) we obtain the estimate

$$\#\operatorname{supp}(\mathbf{u}^{(i)}) \lesssim \varepsilon_i^{-\frac{\tilde{\sigma}}{r}} \max\{\|\mathbf{u}\|_{s,x}^{1/\tilde{\sigma}}, \|\mathbf{u}\|_{s,x}\}^{\frac{\tilde{\sigma}}{r}}. \tag{72}$$

Estimate of the operation count (Claim (c)). For this we need to assess the computational cost of the K_i iterations of each inner loop. Therefore, we first estimate how K_i grows with i. Condition (49) can be re-written as

$$c_1 \leq \rho'^{K_i} \max\{1, \kappa_0 \varepsilon_{i-1}^{1-\tilde{\sigma}} \|\mathbf{u}^{(i-1)}\|_{\tilde{s},x}^{\tilde{\sigma}-1}\} \leq c_2$$

for suitable constants $\kappa_0 > 0$, $0 < c_1, c_2 < 1$, and $0 < \rho' = \rho^{\tilde{\sigma}\frac{\tilde{\tilde{s}}}{s}-1} < 1$. This implies that

$$0 < c_1' \leq K_i + \frac{1}{\log \rho'} \log\Big(\max\{1, \kappa_0 \varepsilon_{i-1}^{1-\tilde{\sigma}} \|\mathbf{u}^{(i-1)}\|_{\tilde{s},x}^{\tilde{\sigma}-1}\}\Big) \leq c_2'.$$

For the lower bound we may take $K_i \geq K$ such that $\frac{3\rho^K}{\theta} < 1$. This choice is necessary to increase the precision in each iteration of the outer loop.

For the upper bound, we may therefore assume that $\kappa_0 \varepsilon_{i-1}^{1-\tilde{\sigma}} \|\mathbf{u}^{(i-1)}\|_{\tilde{s},x}^{\tilde{\sigma}-1} \geq 1$ and proceed as follows. From (71) we have

$$\begin{aligned}
\varepsilon_{i-1}^{(\tilde{\sigma}-1)(\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}-1)} \|\mathbf{u}^{(i-1)}\|_{\tilde{s},x}^{\tilde{\sigma}-1} &\lesssim \max\{\|\mathbf{u}\|_{s,x}^{1/\tilde{\sigma}}, \|\mathbf{u}\|_{s,x}\}^{(\tilde{\sigma}-1)\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}}, \text{ and, thus,} \\
\max\{1, \kappa_0 \varepsilon_{i-1}^{1-\tilde{\sigma}} \|\mathbf{u}^{(i-1)}\|_{\tilde{s},x}^{\tilde{\sigma}-1}\} &\lesssim \varepsilon_{i-1}^{(1-\tilde{\sigma})\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}} \max\{\|\mathbf{u}\|_{s,x}^{1/\tilde{\sigma}}, \|\mathbf{u}\|_{s,x}\}^{(\tilde{\sigma}-1)\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}}.
\end{aligned}$$

Thus, we can estimate K_i by

$$K_i \lesssim c_2' + \frac{(\tilde{\sigma}-1)\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}}{\log \frac{1}{\rho'}} \log(\varepsilon_{i-1}^{-1} \max\{\|\mathbf{u}\|_{s,x}^{1/\tilde{\sigma}}, \|\mathbf{u}\|_{s,x}\}). \tag{73}$$

So far we have established how K_i depends on the current precision ε_{i-1}. Next we assess how ε_i decays with i, i.e.,

$$\varepsilon_i = \frac{3\rho^{K_i}}{\theta}\varepsilon_{i-1} = \left(\frac{3}{\theta}\right)^i \rho^{\sum_{\ell=0}^{i} K_\ell}\varepsilon_0. \tag{74}$$

With these estimates we can now count the number of operations needed to compute $\mathbf{u}_\varepsilon$.

Recall that **RHS** can be implemented by M-nearest neighborhood approximation and one has $\mathbf{f}^{(i)} = \mathbf{RHS}[\frac{\theta\varepsilon_i}{6\alpha K_i}, \mathbf{f}]$ and $\operatorname{supp}(\mathbf{f}^{(i)}) \subset B(x, M)$,

$$\#\operatorname{supp}(\mathbf{f}^{(i)}) \lesssim M^d \lesssim \left(\frac{\varepsilon_i}{K_i}\right)^{-\frac{1}{r}} \|\mathbf{f}\|_{s,x}^{\frac{1}{r}} \lesssim \left(\frac{\varepsilon_i}{K_i}\right)^{-\frac{1}{r}} \|\mathbf{u}\|_{s,x}^{\frac{1}{r}}. \tag{75}$$

As in (70) we find that

$$\left(\frac{\varepsilon_i}{K_i}\right)^{\frac{\tilde{\tilde{s}}}{s}-1} \|\mathbf{f}^{(i)}\|_{\tilde{s},x} \;\; \lesssim \;\; \|\mathbf{u}\|_{s,x}^{\frac{\tilde{\tilde{s}}}{s}}, \tag{76}$$

and hence

$$\left(\frac{\varepsilon_i}{K_i}\right)^{\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}-1} \|\mathbf{f}^{(i)}\|_{\tilde{s},x} \le \left(\frac{\varepsilon_i}{K_i}\right)^{\frac{\tilde{\tilde{s}}}{s}-1} \|\mathbf{f}^{(i)}\|_{\tilde{s},x} \lesssim \|\mathbf{u}\|_{s,x}^{\frac{\tilde{\tilde{s}}}{s}} \le \max\{\|\mathbf{u}\|_{s,x}^{1/\tilde{\sigma}}, \|\mathbf{u}\|_{s,x}\}^{\tilde{\sigma}\frac{\tilde{\tilde{s}}}{s}}. \tag{77}$$

Recall that $\mathbf{v}^{(i,0)} = \mathbf{u}^{(i-1)}$; by (71) we have

$$\left(\frac{\varepsilon_i}{K_i}\right)^{\frac{\tilde{\tilde{s}}}{s}-1} \|\mathbf{v}^{(i,0)}\|_{\tilde{s},x} \le \varepsilon_i^{\frac{\tilde{\tilde{s}}}{s}-1} \|\mathbf{u}^{(i-1)}\|_{\tilde{s},x} \lesssim \max\{\|\mathbf{u}\|_{s,x}^{1/\tilde{\sigma}}, \|\mathbf{u}\|_{s,x}\}^{\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}}.$$

Assume that we have already shown that

$$\left(\frac{\varepsilon_i}{K_i}\right)^{\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}-1} \|\mathbf{v}^{(i,j-1)}\|_{\tilde{s},x} \;\; \le \;\; \kappa'\kappa_3^{j-1} \max\{\|\mathbf{u}\|_{s,x}^{1/\tilde{\sigma}}, \|\mathbf{u}\|_{s,x}\}^{\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}}, \quad j < K_i, \tag{78}$$

for $\kappa_3 > 0$ sufficiently large. Then, using

$$\mathbf{v}^{(i,j)} := \mathbf{v}^{(i,j-1)} - \alpha(\mathbf{APPLY}[\frac{\theta\varepsilon_i}{6\alpha K_i}, \mathbf{A}, \mathbf{v}^{(i,j-1)}] - \mathbf{f}^{(i)})\,,$$

combined with Theorem 3.7 (d), and (77), we derive that

$$\begin{aligned}
\left(\frac{\varepsilon_i}{K_i}\right)^{\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}-1} \|\mathbf{v}^{(i,j)}\|_{\tilde{s},x} \;\; &\le \;\; \left(\frac{\varepsilon_i}{K_i}\right)^{\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}-1} \|\mathbf{v}^{(i,j-1)}\|_{\tilde{s},x} + \alpha C_{\mathbf{APPLY}} \left(\frac{\varepsilon_i}{K_i}\right)^{\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}-1} \|\mathbf{v}^{(i,j-1)}\|_{\tilde{s},x} \\
&\qquad + \alpha \left(\frac{\varepsilon_i}{K_i}\right)^{\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}-1} \|\mathbf{f}^{(i)}\|_{\tilde{s},x} \\
&\le \;\; \kappa' \left((1 + \alpha C_{\mathbf{APPLY}})\kappa_3^{j-1} + 1\right) \max\{\|\mathbf{u}\|_{s,x}^{1/\tilde{\sigma}}, \|\mathbf{u}\|_{s,x}\}^{\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}} \\
&\le \;\; \kappa'\kappa_3^{j} \max\{\|\mathbf{u}\|_{s,x}^{1/\tilde{\sigma}}, \|\mathbf{u}\|_{s,x}\}^{\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}},
\end{aligned}$$

for $\kappa_3 > 0$ sufficiently large. In fact, we may choose $\kappa_3 = 2 + \alpha C_{\mathbf{APPLY}}$. Finally, by induction on j, we obtain

$$\left(\frac{\varepsilon_i}{K_i}\right)^{\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}-1} \|\mathbf{v}^{(i,j)}\|_{\tilde{s},x} \ \lesssim \ \kappa_3^j \max\{\|\mathbf{u}\|_{s,x}^{1/\tilde{\sigma}}, \|\mathbf{u}\|_{s,x}\}^{\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}}, \quad j = 0, \ldots, K_i, \tag{79}$$

for a suitable constant $\kappa_3 > 1$. Furthermore, let us observe that

$$\begin{aligned} \left(\frac{\varepsilon_i}{K_i}\right)^{-\frac{1}{\tilde{r}}} \|\mathbf{v}^{(i,j)}\|_{\tilde{s},x}^{\frac{1}{\tilde{r}}} &= \left(\frac{\varepsilon_i}{K_i}\right)^{-\frac{\tilde{\sigma}}{r}} \left(\left(\frac{\varepsilon_i}{K_i}\right)^{\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}-1} \|\mathbf{v}^{(i,j)}\|_{\tilde{s},x} \right)^{\frac{1}{\tilde{r}}} \\ &\lesssim \left(\frac{\varepsilon_i}{K_i}\right)^{-\frac{\tilde{\sigma}}{r}} \left(\kappa_3^j \| \max\{\|\mathbf{u}\|_{s,x}^{1/\tilde{\sigma}}, \|\mathbf{u}\|_{s,x}\}^{\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}} \right)^{\frac{1}{\tilde{r}}} \\ &= \left(\frac{\varepsilon_i}{K_i}\right)^{-\frac{\tilde{\sigma}}{r}} \max\{\|\mathbf{u}\|_{s,x}^{1/\tilde{\sigma}}, \|\mathbf{u}\|_{s,x}\}^{\frac{\tilde{\sigma}}{r}} \kappa_3^{\frac{j}{\tilde{r}}}. \end{aligned}$$

By a similar induction, using Theorem 3.7 (b), we show that

$$\# \operatorname{supp}(\mathbf{v}^{(i,j)}) \lesssim \left(\frac{\varepsilon_i}{K_i}\right)^{-\frac{\tilde{\sigma}}{r}} \max\{\|\mathbf{u}\|_{s,x}^{1/\tilde{\sigma}}, \|\mathbf{u}\|_{s,x}\}^{\frac{\tilde{\sigma}}{r}} \kappa_3^{\frac{j-1}{\tilde{r}}}.$$

By using these estimates together with Theorem 3.7 (c) and (79), the operation cost to compute $\mathbf{v}^{(i,j)}$ from $\mathbf{v}^{(i,j-1)}$ is estimated by a multiple of

$$\left(\frac{\varepsilon_i}{K_i}\right)^{-\frac{\tilde{\sigma}}{r}} \max\{\|\mathbf{u}\|_{s,x}^{1/\tilde{\sigma}}, \|\mathbf{u}\|_{s,x}\}^{\frac{\tilde{\sigma}}{r}} \kappa_3^{\frac{j}{\tilde{r}}}.$$

Therefore, the cost to compute $\mathbf{u}^{(i)}$ from $\mathbf{u}^{(i-1)}$ is estimated by

$$\left(\frac{\varepsilon_i}{K_i}\right)^{-\frac{\tilde{\sigma}}{r}} \max\{\|\mathbf{u}\|_{s,x}^{1/\tilde{\sigma}}, \|\mathbf{u}\|_{s,x}\}^{\frac{\tilde{\sigma}}{r}} \sum_{j=0}^{K_i} \kappa_3^{\frac{j}{\tilde{r}}} \lesssim \left(\frac{\varepsilon_i}{K_i}\right)^{-\frac{\tilde{\sigma}}{r}} \max\{\|\mathbf{u}\|_{s,x}^{1/\tilde{\sigma}}, \|\mathbf{u}\|_{s,x}\}^{\frac{\tilde{\sigma}}{r}} \kappa_3'^{K_i},$$

with $\kappa_3' = \kappa_3^{1/\tilde{r}}$. Let us fix $N \in \mathbb{N}$ such that $\varepsilon = \varepsilon_N = \left(\frac{3}{\theta}\right)^N \rho^{\sum_{\ell=0}^N K_\ell} \varepsilon_0$. Observe now that by (73)

$$\kappa_3'^{K_i} \lesssim \varepsilon_{i-1}^{\xi(1-\tilde{\sigma})\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}} \max\{\|\mathbf{u}\|_{s,x}^{1/\tilde{\sigma}}, \|\mathbf{u}\|_{s,x}\}^{\xi(\tilde{\sigma}-1)\frac{\tilde{\sigma}\tilde{\tilde{s}}}{s}},$$

for $\xi = \log(\kappa_3')/\log(1/\rho')$. Thus, by using the estimate (73) for K_i , and the explicit expression for ε_i above, the total cost for computing $\mathbf{u}_\varepsilon$ is proportional to

$$\begin{aligned}
&\sum_{i=0}^{N} \left(\frac{\varepsilon_i}{K_i}\right)^{-\frac{\tilde\sigma}{r}} \max\{\|\mathbf{u}\|_{s,x}^{1/\tilde\sigma}, \|\mathbf{u}\|_{s,x}\}^{\frac{\tilde\sigma}{r}} \kappa_3'^{K_i} \\
\lesssim\ & \sum_{i=0}^{N} \left(\varepsilon_i^{-1} \max\{\|\mathbf{u}\|_{s,x}^{1/\tilde\sigma}, \|\mathbf{u}\|_{s,x}\}\right)^{\beta} \log(\varepsilon_i^{-1} \max\{\|\mathbf{u}\|_{s,x}^{1/\tilde\sigma}, \|\mathbf{u}\|_{s,x}\})^{\frac{\tilde\sigma}{r}} \\
\lesssim\ & \sum_{i=0}^{N} \left(\varepsilon_i^{-1} \max\{\|\mathbf{u}\|_{s,x}^{1/\tilde\sigma}, \|\mathbf{u}\|_{s,x}\}\right)^{\beta+\delta} \\
=\ & \sum_{i=0}^{N} \left(\left(\frac{3}{\theta}\right)^{i} \rho^{\sum_{\ell=0}^{i} K_\ell} \varepsilon_0\right)^{-(\beta+\delta)} \max\{\|\mathbf{u}\|_{s,x}^{1/\tilde\sigma}, \|\mathbf{u}\|_{s,x}\}^{\beta+\delta},
\end{aligned}$$

where $\beta := (\xi(\tilde\sigma - 1)\frac{\tilde\sigma\tilde{\tilde{s}}}{s} + \frac{\tilde\sigma}{r})$, and $\delta > 0$ is arbitrarily small.

$$\begin{aligned}
&\sum_{i=0}^{N} \left(\left(\frac{3}{\theta}\right)^{i} \rho^{\sum_{\ell=0}^{i} K_\ell} \varepsilon_0\right)^{-(\beta+\delta)} \Big/ \varepsilon^{-(\beta+\delta)} \\
=\ & 1 + \sum_{i=0}^{N-1} \left(\frac{3}{\theta}\right)^{(\beta+\delta)(N-i)} \rho^{(\beta+\delta)\sum_{\ell=i+1}^{N} K_\ell} \\
\leq\ & 1 + \sum_{i=1}^{N} \left(\frac{3\rho^K}{\theta}\right)^{(\beta+\delta)i} \leq 1 + \frac{1}{1 - \frac{3\rho^K}{\theta}} < \infty.
\end{aligned}$$

The latter estimates imply that

$$\mathcal{C}(\mathbf{u}_\varepsilon) \lesssim \varepsilon^{-(\frac{\tilde\sigma}{r} + \xi(\tilde\sigma-1)\frac{\tilde\sigma\tilde{\tilde{s}}}{s} + \delta)} \max\{\|\mathbf{u}\|_{s,x}^{1/\tilde\sigma}, \|\mathbf{u}\|_{s,x}\}^{\frac{\tilde\sigma}{r} + \xi(\tilde\sigma-1)\frac{\tilde\sigma\tilde{\tilde{s}}}{s} + \delta}.$$

References

[1] A. Aldroubi and K. Gröchenig, *Nonuniform sampling and reconstruction in shift-invariant spaces*, SIAM Review **43** (2001), no. 4, 585–620.

[2] R. Balan, P. G. Casazza, C. Heil, and Z. Landau, *Density, overcompleteness, and localization of frames. I. Theory*, J. Fourier Anal. Appl. **12** (2006), no. 2, 105–143.

[3] P. G. Casazza and O. Christensen, *Approximation of the inverse frame operator and application to Gabor frames*, J. Approx. Th. **130** (2000), no. 2, 338–356.

[4] O. Christensen, *Finite–dimensional approximation of the inverse frame operator*, J. Four. Anal. Appl. **6** (2000), no. 1, 79–91.

[5] ______, *An Introduction to Frames and Riesz Bases*, Birkhäuser, 2003.

[6] O. Christensen and C. Heil, *Perturbations of Banach frames and atomic decompositions*, Math. Nachr. **186** (1997), 33–47.

[7] O. Christensen and T. Strohmer, *The finite section method and problems in frame theory.*, J. Approx. Th. **133** (2005), no. 2, 221–237.

[8] A. Cohen, W. Dahmen, and R. DeVore, *Adaptive wavelet methods for elliptic operator equations — Convergence rates*, Math. Comput. **70** (2001), 27–75.

[9] ______, *Adaptive wavelet methods II: Beyond the elliptic case*, Found. Comput. Math. **2** (2002), no. 3, 203–245.

[10] E. Cordero and K. Gröchenig, *Localization of frames. II.*, Appl. Comput. Harmon. Anal. **17** (2004), no. 1, 29–47.

[11] S. Dahlke, M. Fornasier, M. Primbs, T. Raasch, and M. Werner, *Nonlinear and adaptive frame approximation schemes for elliptic PDEs: theory and numerical experiments*, preprint (2007).

[12] S. Dahlke, M. Fornasier, and T. Raasch, *Adaptive frame methods for elliptic operator equations.*, Adv. Comput. Math. **27** (2007), no. 1, 27–63.

[13] S. Dahlke, M. Fornasier, T. Raasch, R. Stevenson, and M. Werner, *Adaptive frame methods for elliptic operator equations: The steepest descent approach*, IMA J. Numer. Anal. **27** (2007), no. 4, 717–740.

[14] M. Fornasier, *Banach frames for α-modulation spaces*, Appl. Comput. Harmon. Anal. **22** (2007), no. 2, 157–175.

[15] M. Fornasier and K. Gröchenig, *Intrinsic localization of frames*, Constr. Approx. **22** (2005), no. 3, 395–415.

[16] M. Fornasier and H. Rauhut, *Continuous frames, function spaces, and the discretization problem.*, J. Fourier Anal. Appl. **11** (2005), no. 3, 245–287.

[17] K. Gröchenig, *Describing functions: atomic decompositions versus frames*, Monatsh. Math. **112** (1991), no. 1, 1–42.

[18] ______, *Foundations of Time–Frequency Analysis*, Birkhäuser, 2000.

[19] ______, *Localization of frames, Banach frames, and the invertibility of the frame operator*, J. Fourier Anal. Appl. **10** (2004), no. 2, 105–132.

[20] K. Gröchenig, *Time-frequency analysis of Sjöstrand's class*, Rev. Mat. Iberoamericana **22** (2006), no. 2, 703–724.

[21] K. Gröchenig and M. Leinert, *Symmetry and inverse-closedness of matrix algebras and functional calculus for infinite matrices*, Trans. Amer. Math. Soc. **358** (2006), no. 6, 2695–2711 (electronic).

[22] K. Gröchenig, Z. Rzeszotnik, and T. Strohmer, *Quantitative estimates for the finite section method*, preprint (2008).

[23] S. Jaffard, *Propriétés des matrices "bien localisées" près de leur diagonale et quelques applications*, Ann. Inst. H. Poincaré Anal. Non Linéaire **7** (1990), no. 5, 461–476.

[24] P. G. Lemarié, *Bases d'ondelettes sur les groupes de lie stratifiés*, Bulletin de la Société Mathématique de France **117** (1989), no. 2, 213–232.

[25] R. Stevenson, *Adaptive solution of operator equations using wavelet frames*, SIAM J. Numer. Anal. **41** (2003), no. 3, 1074–1100.

[26] T. Strohmer, *Pseudodifferential operators and Banach algebras in mobile communications*, Appl. Comput. Harmon. Anal. **20** (2006), no. 2, 237–249.

Stephan Dahlke, FB 12 Mathematik und Informatik, Philipps-Universität Marburg, Hans-Meerwein Strasse Lahnberge, D-35032 Marburg, Germany
email: dahlke@mathematik.uni-marburg.de

Massimo Fornasier, Johann Radon Institute for Computational and Applied Mathematics, Austrian Academy of Sciences, Altenbergerstrasse 69, A-4040, Linz, Austria.
email: massimo.fornasier@oeaw.ac.at

Karlheinz Gröchenig, NuHAG, Fakultät für Mathematik, Universität Wien, Nordbergstrasse 15, A-1090, Wien, Austria.
email: karlheinz.groechenig@univie.ac.at