Eidgenössische Technische Hochschule Zürich
*Ecole polytechnique fédérale de Zurich*
*Politecnico federale di Zurigo*
*Swiss Federal Institute of Technology Zurich*

# Intrinsic localization of anisotropic frames

P. Grohs

Research Report No. 2012-04
March 2012

Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

# Intrinsic Localization of Anisotropic Frames

P. Grohs

March 9, 2012

**Abstract**

The present article studies off-diagonal decay properties of Moore-Penrose pseudoinverses of (bi-infinite) matrices satisfying an analogous condition. Off-diagonal decay in our paper is considered with respect to specific index distance functions which incorporates those usually used for the study of localization properties for wavelet frames but also more general systems such as curvelets or shearlets. Our main result is that if a matrix satisfies an off-diagonal decay condition, then its Moore-Penrose pseudoinverse satisfies a similar condition. Applied to the study of frames this means that, if a wavelet, curvelet or shearlet frame is intrinsically localized, then its canonical dual is, too.

**Keywords:** Frame Localization, Curvelets, Shearlets, nonlinear Approximation

**2000 Mathematics Subject Classification.** Primary 41AXX, Secondary 41A25, 53B, 22E.

## 1 Introduction

This paper is concerned with results on off-diagonal decay properties of the Moore-Penrose pseudoinverse $\mathbf{A}^+ = (a^+_{\lambda,\lambda'})_{\lambda,\lambda'\in\Lambda}$ of a symmetric (bi-infinite) matrix $\mathbf{A} = (a_{\lambda,\lambda'})_{\lambda,\lambda'\in\Lambda}$ satisfying an off-diagonal decay condition. This is in general a difficult problem, especially if no additional structural properties for $\mathbf{A}$ are assumed. Our main result is that if $\mathbf{A}$ satisfies a condition of the form

$$|a_{\lambda,\lambda'}| \le C_0\omega(\lambda,\lambda')^{-N}$$

for a specific type of *index distance function* $\omega : \Lambda \times \Lambda \to \mathbb{R}_+$, then, if $\mathbf{A}$ is also well-conditioned as an operator from $l_2(\Lambda)$ to itself, the entries of the Moore-Penrose pseudoinverse $\mathbf{A}^+$ satisfy the analogous estimate

$$|a^+_{\lambda,\lambda'}| \le C'\omega(\lambda,\lambda')^{-N^+}$$

for some $N^+ > 0$ which depends on $N, C_0$ and the spectrum of $\mathbf{A}$ and which goes to infinity with $N$. The precise dependence of $N^+$ on the various parameters will be made explicit. The index distance function $\omega$ has to be of a specific kind. Index functions which fall into our framework for instance arise naturally in the study of the off-diagonal decay properties of the Gramian matrix of wavelet, curvelet, or shearlet frames. See [11, 3, 29] and the references therein for more information regarding these systems.

In this context, the study of off-diagonal decay properties is of fundamental importance: applications such as compression, approximation and the study of function spaces crucially depend on them. In connection with frames, the study of off-diagonal properties of the Gramian matrix is usually referred to as *localization*, see [18, 15, 8, 24] for related results on Gabor- or wavelet frames. In this lingo, our main result is that if a wavelet, curvelet or shearlet frame is *intrinsically localized*, then its dual frame is, too. More precisely, given a frame $\Psi = (\psi_\lambda)_{\lambda\in\Lambda}$ for a Hilbert space $\mathcal{H}$ with canonical dual frame $\tilde{\Psi} =( \tilde{\psi}_\lambda)_{\lambda\in\Lambda}$ such that

$$f = \sum_{\lambda\in\Lambda}(f,\psi_\lambda)_{\mathcal{H}}\tilde{\psi}_\lambda = \sum_{\lambda\in\Lambda}(f,\tilde{\psi}_\lambda)_{\mathcal{H}}\psi_\lambda \quad \text{for all } f \in \mathcal{H}, \tag{1}$$

a condition of the form

$$|(\psi_\lambda,\psi_{\lambda'})_{\mathcal{H}}| \le C_0\omega(\lambda,\lambda')^{-N} \tag{2}$$

implies that

$$|(\tilde{\psi}_\lambda,\tilde{\psi}_{\lambda'})_{\mathcal{H}}| \le C^+\omega(\lambda,\lambda')^{-N^+}$$

and also

$$|(\tilde{\psi}_\lambda, \psi_{\lambda'})_{\mathcal{H}}| \leq C^+ \omega(\lambda, \lambda')^{-N^+}$$

for some $N^+$ which depends on $\omega, C_0, N$ and the frame bounds of $\Psi$. There exists a variety of applications where results of this kind are crucial. For instance, localization properties imply that the primal frame $\Psi$ and the dual frame $\tilde{\Psi}$ generate the same function spaces. In particular their approximation properties are equivalent.

Another important application concerns the compression of operators $F : \mathcal{H} \to \mathcal{H}$. Given $f \in \mathcal{H}$, by (1) we can represent it via the frame coefficient sequence $((f, \psi_\lambda)_{\mathcal{H}})_{\lambda \in \Lambda}$. Discretization of the operator $F$ entails the computation of the mapping

$$((f, \psi_\lambda)_{\mathcal{H}})_{\lambda \in \Lambda} \mapsto ((Ff, \psi_\lambda)_{\mathcal{H}})_{\lambda \in \Lambda}$$

which can be represented in matrix form as $\mathbf{F} := \left(F\tilde{\psi}_\lambda, \psi_{\lambda'}\right)_{\lambda,\lambda' \in \Lambda}$. It turns out that localization properties of the matrix $\mathbf{F}$ are often crucial to enable an efficient (i.e., linear complexity) matrix multiplication

$$\mathbf{c} \mapsto \mathbf{Fc}$$

for $\mathbf{c} = ((f, \psi_\lambda)_{\mathcal{H}})_{\lambda \in \Lambda}$ and $f$ possessing a sparse representation in $\tilde{\Psi}$ [1, 5]. However, often the matrix entries of $\mathbf{F}$ cannot be computed or estimated directly, the reason being that the dual frame $\tilde{\Psi}$ is not accessible explicitly. What usually *can* be computed are the entries of the matrix

$$\tilde{\mathbf{F}} := ((F\psi_\lambda, \psi_{\lambda'})_{\mathcal{H}})_{\lambda,\lambda' \in \Lambda}$$

and often it can be deduced that this matrix is indeed well-localized. Our results on localization imply that in this case also the matrix $\mathbf{F}$ is well-localized, thus enabling efficient discretization of $F$. As a final application we also show that one can use our results to prove optimal complexity of adaptive frame methods in the spirit of [32].

**Caveat Emptor.** We want to remark at this point that, as it is common for results on exponential localization, our results are mainly of a qualitative nature, meaning that if $N$ becomes large, so does $N^+$. The precise value of $N^+$ which is guaranteed by our proof can (and will) be much smaller than $N$. Most annoyingly, it depends on the constant $C_0$ from (2). This remark also applies to previous results on frame localization which – in contrast to our results – are confined to the study of wavelet frames, see e.g. [30, 9].

## 1.1 Previous Work and Novelties

Our main result can be seen as a generalization of [24, 30] beyond the setting of wavelet bases. It fits best into the theory of *exponential localization* which has been studied in an abstract setting in [18, 15, 8, 12, 24] and which is mainly confined to the study of wavelet- and Gabor frames. There exist quite deep studies of localization properties of the inverses of matrices [19, 20] but all of them crucially require very strong structural properties which are not valid for the problems which motivate us. In particular, properties such as inverse closedness of the Banach algebra of off-diagonal decay matrices do in general not hold for affine systems such as wavelets.

Our motivation for this work is to analyze the localization properties of more general frames than wavelet- and Gabor frames, with a particular focus on shearlets and curvelets. These novel systems have had a dramatic impact on both pure and applied mathematics in recent years; the former for their ability to diagonalize Fourier Integral Operators [31, 2], the latter for their ability to sparsely represent (and hence efficiently compress) multivariate functions with singularities along curved hypersurfaces [22, 3, 28].

The main contribution of this paper is an extension of known results on exponential localization to these systems. This turns out to be quite nontrivial: Previous results heavily rely on a certain algebraic structure of the index set $\Lambda$, which is however not valid for anisotropic systems like shearlets and curvelets.

Nevertheless we demonstrate that analogous results can be established under much weaker requirements on $\omega$ and $\Lambda$, making our approach much more robust. Specifically, we can apply our machinery to obtain localization properties for more general frames than wavelet frames, most notably shearlets and curvelets . We consider such results to be especially important, since, in contrast to wavelet- or Gabor frames, for anisotropic frame systems it is with current technology impossible to get a good grip on the structure of the dual frame. For instance [25] recently gave the first construction of shearlet frames which are compactly supported. These frames are not tight and it is certainly of interest to understand the properties of the associated dual frames. Our results provide a first step in that direction.

We also would like to stress that our results are of much wider generality, in that they hold for arbitrary matrices, which are not necessarily Gramians of frame systems.

### 1.2 A Brief Outline

In the following Section 2 we start by introducing the fundamental concept of localization and collecting the assumptions which we will require later. Section 3 contains our main result, namely that the Moore-Penrose pseudoinverse of a localized matrix is localized, too. Note that thus far, all results hold for general matrices with no reference to frame theory. We close with some applications to the study of wavelet- shearlet- and curvelet frames in Section 4.

### 1.3 Notation

We use boldface letters for (bi-infinite) matrices $\mathbf{A}$. The symbol $l_2(\Lambda)$ denotes the usual Lebesgue space with $\Lambda$ some (discrete) measure space. For an operator $\mathbf{A} : l_2(\Lambda) \to l_2(\Lambda)$ we use the symbol $\sigma(\mathbf{A})_2$ for its spectrum. In general, for an operator $S : \mathcal{H} \to \mathcal{H}$ with $\mathcal{H}$ a Hilbert space we denote its spectrum by $\sigma(S)$. The inner product of $\mathcal{H}$ is denoted by $(\cdot,\cdot)_{\mathcal{H}}$. We will also use the notation $A \lesssim B$ to describe that $A$ is bounded by a uniform constant times $B$.

## 2 Basic Assumptions

The present section collects the fundamental assumptions which will be necessary in our results to follow. We are concerned with matrix operators indexed by an index set $\Lambda$ which possesses a multiscale structure:

**Definition 2.1.** *The mapping $s : \Lambda \to \mathbb{N}$ maps every element $\lambda \in \Lambda$ to its* scale $s_\lambda \in \mathbb{N}$.

Furthermore, we equip our index set $\Lambda$ with a distance function.

**Definition 2.2.** *Let $\mathrm{d} : \lambda \to \mathbb{R}_+$ be such that*

$$\mathrm{d}(\lambda, \lambda) = 0 \quad \text{for all } \lambda \in \Lambda.$$

*Then we define the* index distance function

$$\omega(\lambda, \lambda') := 2^{|s_\lambda - s_{\lambda'}|} \left(1 + 2^{\min(s_\lambda, s_{\lambda'})}\mathrm{d}(\lambda, \lambda')\right).$$

We impose the following assumptions on d and $\Lambda$:

**Assumption 2.3** (Pseudo-symmetry and pseudo-triangle inequality)**.** *There exist constants $c_T, c_S > 0$ such that*

$$\mathrm{d}(\lambda, \lambda') \leq c_T \left(\mathrm{d}(\lambda, \lambda'') + \mathrm{d}(\lambda'', \lambda')\right) \tag{3}$$

*and*

$$\mathrm{d}(\lambda, \lambda') \leq c_S \mathrm{d}(\lambda', \lambda). \tag{4}$$

In addition we assume the following admissibility condition, reminiscent of the Schur condition required in [24]:

**Assumption 2.4** (Admissibility)**.** *We make the following $(M, K)$-admissibility assumption that for all $k > K$ and $\mu \in \Lambda$ we have*

$$\sum_{\lambda \in \Lambda_j} (1 + 2^q \mathrm{d}(\mu, \lambda))^{-k} \leq C_A 2^{M(j-q)_+}, \tag{5}$$

*where*

$$\Lambda_j := \{\lambda \in \Lambda : \ s_\lambda = j\}$$

*and $C_A > 0$.*

Aside from the $(M,K)$-admissibility assumption we shall require that the index set $\Lambda$ is well-separated in a certain sense.

**Assumption 2.5** (Separability)**.** *There exists a constant* $c_\Lambda > 1$ *with*

$$\omega(\lambda,\lambda') \geq c_\Lambda \quad \text{for all } \lambda \neq \lambda' \in \Lambda. \tag{6}$$

The following definition introduces the crucial notion of localization for matrices.

**Definition 2.6.** *A matrix*

$$\mathbf{A} = (a_{\lambda,\lambda'})_{\lambda,\lambda'\in\Lambda}$$

*is called* $N$-localized *if*

$$|a_{\lambda,\lambda'}| \leq C_0 \omega(\lambda,\lambda')^{-N} \tag{7}$$

*for a constant* $C_0 > 0$.

We will assume that the matrix $\mathbf{A}$ is symmetric and possesses a spectral gap:

**Definition 2.7.** *The matrix* $\mathbf{A}$ *viewed as an operator from* $l_2(\Lambda)$ *to itself possesses a* spectral gap *if there exist numbers* $0 < A \leq B < \infty$ *such that*

$$\sigma(\mathbf{A})_2 \subset \{0\} \cup [A,B]. \tag{8}$$

Note that the point 0 must necessarily be an eigenvalue of $\mathbf{A}$ since it is isolated.

**Assumption 2.8** (Symmetry and spectral gap)**.** *The operator* $\mathbf{A}$ *is symmetric and possesses a spectral gap.*

**Remark 2.9.** *Let us pause for a moment to discuss Assumptions 2.4, 2.8, 2.3, 2.5 and their relation to the assumptions of previous works. We have already mentioned that a stronger form of the Admissibility Assumption 2.4 can already be found as a Schur condition in the classical work [24]. In [30], a stronger form of Assumption 2.3 is used for establishing various results on exponential localization. The Spectral Assumption 2.8 is also used in all previous works on exponential localization that we are aware of, such as [30, 12], although mostly the eigenvalue zero is not permitted.*

*Our goal is to extend the results on exponential localization of wavelet Riesz bases to the case of curvelet and shearlet frames. As it turns out, in this endeavor it is necessary to work with our weaker assumptions, due to the structure of the associated index distance function to be defined in Section 4. In our opinion, the main insight of the present article is that actually our weaker assumptions suffice to prove localization results.*

*Note that by symmetry of* $\mathbf{A}$ *we actually have that*

$$|a_{\lambda,\lambda'}| \leq C_0 \max\left(\omega(\lambda,\lambda'),\omega(\lambda',\lambda)\right)^{-N}$$

*The assumption on pseudosymmetry could easily be strengthened to full symmetry, e.g.* $c_S = 1$ *by symmetrizing the index distance function without qualitative changes in the analysis. The triangle-inequality which only needs to be satisfied up to a constant* $c_T$ *on the other hand indeed complicates the analysis, cf. the proof of Proposition 3.2.*

We remark that Assumption 2.8 implies that $\mathbf{A}$ has a closed range [26], and thus, by the closed range theorem, we can write

$$l_2(\Lambda) = \ker(\mathbf{A}) \oplus \operatorname{ran}(\mathbf{A}). \tag{9}$$

We are concerned with properties of the Moore-Penrose pseudoinverse $\mathbf{A}^+$ of the symmetric matrix $\mathbf{A}$, which satisfies the normal equations

$$\mathbf{A}^2\mathbf{A}^+ = \mathbf{A}. \tag{10}$$

**Lemma 2.10.** *The matrix* $\mathbf{A}^+$ *can be computed via a Landweber-type iteration by the formula*

$$\mathbf{A}^+ = \beta \sum_{k\in\mathbb{N}} \left(I - \beta\mathbf{A}^2\right)^k \mathbf{A}, \tag{11}$$

*where* $\beta = \frac{2}{A^2+B^2}$. *We have*

$$\mathbf{A}\mathbf{A}^+ = \mathbf{A}^+\mathbf{A} = \mathrm{Proj}_{\mathrm{Ran}(\mathbf{A})}, \tag{12}$$

*and*

$$\left(\mathbf{A}^+\right)^2 \mathbf{A} = \mathbf{A}^+. \tag{13}$$

*Proof.* We first show that (11) is consistent with (10). This can be seen by writing the infinite sum (11) as an iteration scheme

$$\mathbf{A}^{(k+1)} = \beta\mathbf{A} + \left(I - \beta\mathbf{A}^2\right)\mathbf{A}^{(k)}.$$

This shows that the limit of the sequence $\mathbf{A}^{(k)}$ satisfies (10). Existence of this limit follows by observing that the spectrum of the matrices $\left(I-\beta\mathbf{A}^2\right)^k\mathbf{A}$ is contained in the interval $[-Br^k, Ar^k]$ with $r := \frac{B^2-A^2}{A^2+B^2} < 1$:

$$\sigma\left(\left(I-\beta\mathbf{A}^2\right)^k\mathbf{A}\right)_2 = \left\{(1-\beta\nu^2)^k\nu : \nu\in\{0\}\cup[A,B]\right\} \subset [-Br^k, Ar^k]. \tag{14}$$

We come to the proof of (12). The operator $\mathbf{A}^2\big|_{\mathrm{Ran}(\mathbf{A})}$ is bijective with spectrum contained in $[A^2, B^2]$ and its inverse is given by the Neumann series

$$\left(\mathbf{A}^2\big|_{\mathrm{Ran}(\mathbf{A})}\right)^{-1} = \beta\sum_{k\in\mathbb{N}}\left(I - \beta\mathbf{A}\big|^2_{\mathrm{Ran}(\mathbf{A})}\right)^k.$$

It follows that

$$\mathbf{A}\mathbf{A}^+\big|_{\mathrm{Ran}(\mathbf{A})} = \mathbf{A}\big|_{\mathrm{Ran}(\mathbf{A})}\left(\mathbf{A}^2\big|_{\mathrm{Ran}(\mathbf{A})}\right)^{-1}\mathbf{A}\big|_{\mathrm{Ran}(\mathbf{A})} = \mathbf{Id}\big|_{\mathrm{Ran}(\mathbf{A})}.$$

On the other hand, we have

$$\mathbf{A}\mathbf{A}^+\big|_{\mathrm{Ker}(\mathbf{A})} = \mathbf{0},$$

which establishes that

$$\mathbf{A}\mathbf{A}^+ = \mathrm{Proj}_{\mathrm{Ran}(\mathbf{A})}.$$

Noting that $\mathbf{A}\mathbf{A}^+ = \mathbf{A}^+\mathbf{A}$ finally yields (12). It remains to show (13). To this end we use (12) which reduces the statement to showing that

$$\mathbf{A}^+\mathrm{Proj}_{\mathrm{Ran}(\mathbf{A})} = \mathbf{A}^+$$

and this is easy to see from the definition and (9). □

# 3 Main result

In this section we show our main result. In our proofs we will for simplicity assume that $c_S = 1$. By Remark 2.9 this presents no loss in generality.

Consider a symmetric matrix $\mathbf{A}$ which possesses a spectral gap. Our main result is the following theorem.

**Theorem 3.1.** *For any* $\mathbf{A}$ *which is* $N+L$*-localized, e.g.,*

$$|a_{\lambda,\lambda'}| \le C_0\omega(\lambda,\lambda')^{-N-L},$$

*with* $L$ *satisfying (26), and which has a spectral gap, e.g.,*

$$\sigma(\mathbf{A})_2 \subset \{0\}\cup[A,B],$$

*the Moore-Penrose pseudoinverse* $\mathbf{A}^+$ *is* $N^+$*-localized with*

$$N^+ = N\left(1 - \frac{\log\left(C_0C\left(1+\beta C_0C\right)\right)}{\log(r)}\right)^{-1} \tag{15}$$

*with*

$$r := \frac{B^2-A^2}{B^2+A^2} < 1,$$

*and* $C$ *given by (24).*

*Proof.* We will make use of the Landweber-type expansion (11). Therefore the starting point is to look at the matrices

$$\mathbf{A}^{(k)} = \left(a^{(k)}_{\lambda,\lambda'}\right)_{\lambda,\lambda'\in\Lambda} = \left(I - \beta\mathbf{A}^2\right)^k \mathbf{A}$$

**Claim 1:** There exists a constant $C$, given by (24), *independent of* $N$ (and only depending on $\Lambda$ and $\omega$) such that

$$\left|a^{(k)}_{\lambda,\lambda'}\right| \le C_0^{k+1}\left(1+\beta CC_0\right)^k C^k \omega(\lambda,\lambda')^{-N}. \tag{16}$$

To see this, we iteratively utilize Proposition 3.2 for the matrices $I - \beta\mathbf{A}^2$ and $\mathbf{A}$ and use induction in $k$. By symmetry of all matrices which occur in our computations, we may restrict ourselves to index pairs $\lambda, \lambda'$ with $s_\lambda \le s_{\lambda'}$. Clearly the claim holds for $k = 0$. Now assume that

$$\left|a^{(k-1)}_{\lambda,\lambda'}\right| \le C_0^{k}\left(1+\beta CC_0\right)^{k-1} C^{k-1} \omega(\lambda,\lambda')^{-N}. \tag{17}$$

For the calculation below we shall denote the entries of a matrix $\mathbf{B}$ corresponding to indices $\lambda, \lambda'$ by $\mathbf{B}_{\lambda,\lambda'}$.

Since $\mathbf{A}^{(k)} = \left(\mathbf{I} - \beta\mathbf{A}^2\right)\mathbf{A}^{(k-1)}$, we can invoke Proposition 3.2 together with (17) and deduce that

$$|\mathbf{A}\mathbf{A}^{(k-1)}_{\lambda,\lambda'}| \le C_0 C C_0^{k}\left(1+\beta CC_0\right)^{k-1} C^{k-1} \omega(\lambda,\lambda')^{-N}$$

and by another application of Proposition 3.2, that

$$|\beta\mathbf{A}^2\mathbf{A}^{(k-1)}_{\lambda,\lambda'}| \le \beta C_0 C C_0^{k+1}\left(1+\beta CC_0\right)^{k-1} C^{k} \omega(\lambda,\lambda')^{-N}$$

Using again the induction hypothesis (17) and the triangle inequality

$$\left|\left(\mathbf{I} - \beta\mathbf{A}^2\right)\mathbf{A}^{(k-1)}_{\lambda,\lambda'}\right| \le \left|\mathbf{A}^{(k-1)}_{\lambda,\lambda'}\right| + \left|\beta\mathbf{A}^2\mathbf{A}^{(k-1)}_{\lambda,\lambda'}\right|$$

establishes the claim (16).

**Claim 2:** We have the general estimate

$$\left|a^{(k)}_{\lambda,\lambda'}\right| \lesssim r^k. \tag{18}$$

This claim is a simple consequence of the Cauchy-Schwartz inequality and the spectral estimate (14): Denoting $\mathbf{e}_\lambda$ the unit vector in $l_2(\Lambda)$ with all zeros except at the index $\lambda \in \Lambda$, we have

$$\left|a^{(k)}_{\lambda,\lambda'}\right| = \left|\left(\mathbf{A}^{(k)}\mathbf{e}_\lambda, \mathbf{e}_{\lambda'}\right)_{l_2(\Lambda)}\right| \le \left\|\mathbf{A}^{(k)}\mathbf{e}_\lambda\right\|_{l_2(\Lambda)} \|\mathbf{e}_{\lambda'}\|_{l_2(\Lambda)} \le Br^k$$

and this is Claim 2.

We will now combine Claims 1 and 2 in order to obtain the desired result. By (11) we need to get a grip on

$$\sum_{k=0}^{\infty}\left|a^{(k)}_{\lambda,\lambda'}\right| = I + II$$

with

$$I := \sum_{k=0}^{k_0}\left|a^{(k)}_{\lambda,\lambda'}\right|$$

and

$$II := \sum_{k=k_0+1}^{\infty}\left|a^{(k)}_{\lambda,\lambda'}\right|,$$

and $k_0 \in \mathbb{N}$ arbitrary. Before choosing $k_0$ we first estimate $I$ using (16) and $II$ using (18) as follows:

$$|I| \lesssim \sum_{k=0}^{k_0} C_0^k\left(1+\beta CC_0\right)^k C^k \omega(\lambda,\lambda')^{-N} \lesssim C_0^{k_0}\left(1+\beta CC_0\right)^{k_0} C^{k_0} \omega(\lambda,\lambda')^{-N}. \tag{19}$$

On the other hand we can estimate

$$|II| \lesssim \sum_{k=k_0+1}^{\infty} r^k \lesssim r^{k_0}. \tag{20}$$

In summary we have, for *any* $k_0 \in \mathbb{N}$ the estimate

$$\left|a^+_{\lambda,\lambda'}\right| \lesssim C_0^{k_0} \left(1+\beta C C_0\right)^{k_0} C^{k_0} \omega(\lambda,\lambda')^{-N} + r^{k_0}. \tag{21}$$

Put $D := C_0(1+\beta C C_0)C$. We need to balance the two estimates in (21) by finding $k_0$, $N_+$ such that both can be estimated by a constant times $\omega(\lambda,\lambda')^{-N^+}$. Starting with $r^{k_0}$ this yields

$$k_0 = -N^+ \frac{\log(\omega(\lambda,\lambda'))}{\log(r)}. \tag{22}$$

For the first term in (21) we get

$$k_0 = (N-N^+)\frac{\log(\omega(\lambda,\lambda'))}{\log(D)}. \tag{23}$$

Setting equal (22) and (23) we get

$$N^+ = N\left(1 - \frac{\log(D)}{\log(r)}\right)^{-1}$$

which proves the result. □

Now comes the hard part. The following proposition forms the key technical result needed in the proof of Theorem 3.1.

**Proposition 3.2.** *With the constant*

$$C := \frac{4C_A}{1-2^{M-L/2}} + \frac{2C_A}{1-2^{-N}} + 1 \tag{24}$$

*we have*

$$\sum_{\lambda''} \omega(\lambda,\lambda'')^{-N-L}\omega(\lambda'',\lambda')^{-N} \leq C\omega(\lambda,\lambda')^{-N} \tag{25}$$

*for all*

$$L > \max\left(\frac{2\log(2c_T)}{\log(c_\Lambda)}N, 2M\right) \tag{26}$$

*and* $s_\lambda \leq s_{\lambda'}$.

*Proof.* We start by giving a lower bound for a term

$$I_{\lambda''} := \left(1 + 2^{\min(s_\lambda, s_{\lambda''})}\mathrm{d}(\lambda,\lambda'')\right)\left(1 + 2^{\min(s_{\lambda''}, s_{\lambda'})}\mathrm{d}(\lambda'',\lambda')\right). \tag{27}$$

We have

$$\begin{aligned}
I_{\lambda''} &= 1 + 2^{\min(s_\lambda,s_{\lambda''})}\mathrm{d}(\lambda,\lambda'') + 2^{\min(s_{\lambda''},s_{\lambda'})}\mathrm{d}(\lambda'',\lambda') + 2^{\min(s_\lambda,s_{\lambda''})}\mathrm{d}(\lambda,\lambda'')2^{\min(s_{\lambda''},s_{\lambda'})}\mathrm{d}(\lambda'',\lambda') \\
&\geq 1 + 2^{\min(s_\lambda,s_{\lambda'},s_{\lambda''})}\left(\mathrm{d}(\lambda,\lambda'') + \mathrm{d}(\lambda'',\lambda')\right) + 2^{\min(s_\lambda,s_{\lambda''})}\mathrm{d}(\lambda,\lambda'')2^{\min(s_{\lambda''},s_{\lambda'})}\mathrm{d}(\lambda'',\lambda') \\
&\geq 1 + 2^{\min(s_\lambda,s_{\lambda'},s_{\lambda''})}\frac{1}{c_T}\mathrm{d}(\lambda,\lambda') + 2^{\min(s_\lambda,s_{\lambda''})}\mathrm{d}(\lambda,\lambda'')2^{\min(s_{\lambda''},s_{\lambda'})}\mathrm{d}(\lambda'',\lambda') \\
&\geq \frac{1}{2c_T}\left(1 + 2^{\min(s_\lambda,s_{\lambda'},s_{\lambda''})}\mathrm{d}(\lambda,\lambda')\right) \times \dots \\
&\qquad \dots \times \left(1 + 2^{\max(\min(s_\lambda,s_{\lambda''}),\min(s_{\lambda''},s_{\lambda'}))}\min\left(\mathrm{d}(\lambda,\lambda''),\mathrm{d}(\lambda'',\lambda')\right)\right).
\end{aligned}$$

Hence, we arrive at the inequality

$$I_{\lambda''}^{-1} \leq 2c_T\left(1 + 2^{\min(s_\lambda,s_{\lambda'},s_{\lambda''})}\mathrm{d}(\lambda,\lambda')\right)^{-1} J_{\lambda''}^{-1},$$

with

$$J_{\lambda''} := \left(1 + 2^{\max(\min(s_\lambda, s_{\lambda''}), \min(s_{\lambda''}, s_{\lambda'}))} \min\left(\mathrm{d}(\lambda, \lambda''), \mathrm{d}(\lambda'', \lambda')\right)\right).$$

Since

$$J_{\lambda''} \quad \geq \quad \min\left(1 + 2^{\min(s_\lambda, s_{\lambda''})}\mathrm{d}(\lambda, \lambda''), 1 + 2^{\min(s''_\lambda, s_{\lambda'})}\mathrm{d}(\lambda'', \lambda')\right),$$

it follows that

$$J_{\lambda''}^{-N} \leq \max\left(\left(1 + 2^{\min(s_\lambda, s_{\lambda''})}\mathrm{d}(\lambda, \lambda'')\right)^{-N}, \left(1 + 2^{\min(s_{\lambda''}, s_{\lambda'})}\mathrm{d}(\lambda'', \lambda')\right)^{-N}\right) \tag{28}$$

It remains to estimate the sum

$$\sum_{\lambda'' \in \Lambda} \omega(\lambda, \lambda'')^{-N-L} \omega(\lambda'', \lambda')^{-N} = \sum_{\lambda'' \in \Lambda} 2^{-N(|s_\lambda - s_{\lambda''}| + |s_{\lambda''} - s_{\lambda'}|)} \omega(\lambda, \lambda'')^{-L} I_{\lambda''}^{-N}. \tag{29}$$

First note, that it is sufficient to get the desired bound for the sum

$$\sum_{\lambda'' \in \Lambda,\ \lambda'' \neq \lambda} 2^{-N(|s_\lambda - s_{\lambda''}| + |s_{\lambda''} - s_{\lambda'}|)} \omega(\lambda, \lambda'')^{-L} I_{\lambda''}^{-N}, \tag{30}$$

since the estimate for $\lambda'' = \lambda$ is trivial (it accounts for the "+1"-term in (24)). We split this sum into three terms as follows:

$$\sum_{\lambda'' \in \Lambda,\ \lambda'' \neq \lambda} 2^{-N(|s_\lambda - s_{\lambda''}| + |s_{\lambda''} - s_{\lambda'}|)} \omega(\lambda, \lambda'')^{-L} I_{\lambda''}^{-N} = I + II + III,$$

where

$$I := \sum_{s_{\lambda''} \geq s_{\lambda'}},\ II := \sum_{s_{\lambda''} \leq s_\lambda},\ III := \sum_{s_\lambda \leq s_{\lambda''} \leq s_{\lambda'}}.$$

We now study these three cases separately.
$\mathbf{s}_{\lambda''} \geq \mathbf{s}_{\lambda'}$. The estimate for (30) becomes

$$\begin{aligned}
I &\leq (2c_T)^N \sum_{j \geq s_{\lambda'}} 2^{-N(2s_{\lambda''} - s_\lambda - s_{\lambda'})} \sum_{s_{\lambda''} = j} (1 + 2^{s_\lambda}\mathrm{d}(\lambda, \lambda'))^{-N} \times \dots \\
&\quad \dots \times \omega(\lambda, \lambda'')^{-L} \left((1 + 2^{s_\lambda}\mathrm{d}(\lambda, \lambda''))^{-N} + (1 + 2^{s_{\lambda'}}\mathrm{d}(\lambda'', \lambda'))^{-N}\right) \\
&\leq (2c_T)^N c_\Lambda^{-L/2} \sum_{j \geq s_{\lambda'}} 2^{-N(2j - s_\lambda - s_{\lambda'})} \sum_{s_{\lambda''} = j} (1 + 2^{s_\lambda}\mathrm{d}(\lambda, \lambda'))^{-N} \times \dots \\
&\quad \dots \times 2^{-L/2(s''_\lambda - s_\lambda)} \left((1 + 2^{s_\lambda}\mathrm{d}(\lambda, \lambda''))^{-N} + (1 + 2^{s_{\lambda'}}\mathrm{d}(\lambda'', \lambda'))^{-N}\right) \\
&\leq C_A (2c_T)^N c_\Lambda^{-L/2} \sum_{j \geq s_{\lambda'}} 2^{-N(2j - s_\lambda - s_{\lambda'})} (1 + 2^{s_\lambda}\mathrm{d}(\lambda, \lambda'))^{-N} \times \dots \\
&\quad \dots \times 2^{-L/2(j - s_\lambda)} \left(2^{M(j - s_\lambda)} + 2^{M(j - s_{\lambda'})}\right) \\
&\leq \frac{2C_A}{1 - 2^{-2N - L/2 + M}} 2^{-2Ns_{\lambda'} - L/2 s_{\lambda'} + Ms_{\lambda'} + Ns_\lambda + Ns_{\lambda'} + L/2 s_\lambda - Ms_\lambda} (1 + 2^{s_\lambda}\mathrm{d}(\lambda, \lambda'))^{-N} \\
&\leq \frac{2C_A}{1 - 2^{-L/2 + M}} 2^{-N(s_{\lambda'} - s_\lambda)} (1 + 2^{s_\lambda}\mathrm{d}(\lambda, \lambda'))^{-N} 2^{(M - L/2)(s_{\lambda'} - s_\lambda)} \\
&\leq \frac{2C_A}{1 - 2^{-L/2 + M}} \omega(\lambda, \lambda')^{-N}.
\end{aligned}$$

The first inequality is (28), the second one follows from Assumption (6) and $\lambda'' \neq \lambda$, the third one holds by Assumption (5), and the rest by choosing

$$L > \max\left(N \frac{2\log(2c_T)}{\log(c_\Lambda)}, 2M\right) \tag{31}$$

and a geometric summation.

$\mathbf{s}_{\lambda''} \leq \mathbf{s}_{\lambda}$. We need to estimate

$$\begin{aligned}
II &\leq (2c_T)^N \sum_{j\leq s_\lambda} 2^{-N(-2s_{\lambda''}+s_\lambda+s_{\lambda'})} \sum_{s_{\lambda''}=j} (1+2^{s_{\lambda''}}\mathrm{d}(\lambda,\lambda'))^{-N} \times \dots \\
&\quad \dots \times \omega(\lambda,\lambda'')^{-L}\left((1+2^{s_{\lambda''}}\mathrm{d}(\lambda,\lambda''))^{-N} + (1+2^{s_{\lambda''}}\mathrm{d}(\lambda'',\lambda'))^{-N}\right) \\
&\leq (2c_T)^N \sum_{j\leq s_\lambda} 2^{-N(-2s_{\lambda''}+s_\lambda+s_{\lambda'})} \sum_{s_{\lambda''}=j} 2^{N(s_\lambda - s_{\lambda''})}(1+2^{s_\lambda}\mathrm{d}(\lambda,\lambda'))^{-N} \times \dots \\
&\quad \dots \times \omega(\lambda,\lambda'')^{-L}\left((1+2^{s_{\lambda''}}\mathrm{d}(\lambda,\lambda''))^{-N} + (1+2^{s_{\lambda''}}\mathrm{d}(\lambda'',\lambda'))^{-N}\right) \\
&\leq 2C_A(2c_T)^N \sum_{j\leq s_\lambda} 2^{-N(-2j+s_\lambda+s_{\lambda'})}2^{N(s_\lambda-j)}(1+2^{s_\lambda}\mathrm{d}(\lambda,\lambda'))^{-N} c_\Lambda^{-L} \\
&\leq 2C_A \sum_{j\geq 0} 2^{-Nj}\omega(\lambda,\lambda')^{-N}.
\end{aligned}$$

The first estimate is again (28), for the second estimate we utilize the simple fact that

$$2^{s_{\lambda''}-s_\lambda}(1+2^{s_\lambda}\mathrm{d}(\lambda,\lambda'')) \leq 1+2^{s''_\lambda}\mathrm{d}(\lambda,\lambda''),$$

the third inequality uses (5) and the fourth one holds for $L$ according to (31).

$\mathbf{s}_\lambda \leq \mathbf{s}_{\lambda''} \leq \mathbf{s}_{\lambda'}$. Here we estimate

$$\begin{aligned}
III &\leq (2c_T)^N \sum_{s_\lambda\leq j\leq s_{\lambda'}} 2^{-N(s_{\lambda'}-s_\lambda)}(1+2^{s_\lambda}\mathrm{d}(\lambda,\lambda'))^{-N} \times \dots \\
&\quad \dots \times \sum_{s_{\lambda''}=j} \omega(\lambda,\lambda'')^{-L}\left((1+2^{s_\lambda}\mathrm{d}(\lambda,\lambda''))^{-N} + (1+2^{s_{\lambda''}}\mathrm{d}(\lambda'',\lambda'))^{-N}\right) \\
&\leq (2c_T)^N c_\Lambda^{L/2}\omega(\lambda,\lambda')^{-N} \sum_{s_\lambda\leq s_{\lambda''}\leq s_{\lambda'}} 2^{-L/2(s_{\lambda''}-s_\lambda)} \times \dots \\
&\quad \dots \times \left((1+2^{s_\lambda}\mathrm{d}(\lambda,\lambda''))^{-N} + (1+2^{s_{\lambda''}}\mathrm{d}(\lambda'',\lambda'))^{-N}\right) \\
&\leq 2C_A(2c_T)^N c_\Lambda^{L/2}\omega(\lambda,\lambda')^{-N} \sum_{s_\lambda\leq j\leq s_{\lambda''}} 2^{-(L/2-M)(j-s_\lambda)} \\
&\leq 2C_A \sum_{j\geq 0} 2^{-(L/2-M)j}\omega(\lambda,\lambda')^{-N}.
\end{aligned}$$

The first estimate is (28), the second one follows from Assumption (6) and $\lambda'' \neq \lambda$, the third one holds by Assumption (5), and the fourth by choosing $L$ such that (31) holds. We finally conclude that the desired statement holds true with $L$ satisfying (31). Summing up the estimates for $I$, $II$ and $III$ finally yields the desired bound. This proves (26). □

**Remark 3.3.** *In the special case $c_S = c_T = 1$ a different route leads to a similar result. Putting $\rho$ the Poincaré metric defined by*

$$\theta(\lambda,\lambda') := \left(\frac{\mathrm{d}(\lambda,\lambda')^2 + |2^{s_\lambda} - 2^{s_{\lambda'}}|^2}{\mathrm{d}(\lambda,\lambda')^2 + |2^{s_\lambda} + 2^{s_{\lambda'}}|^2}\right)^{1/2}$$

*and*

$$\rho(\lambda,\lambda') := \log\left(\frac{1+\theta(\lambda,\lambda')}{1-\theta(\lambda,\lambda')}\right)^{1/2}$$

*one can show that*

$$\log R\omega \sim \rho$$

*for a constant $R$. After renormalizing $\omega := R\omega$ it can be shown that $\log(\omega)$ satisfies a triangle inequality which is what is needed to apply the machinery developed in [30, 9] on exponential localization. It turns out that the resulting localization estimate for the pseudoinverse (valid only for the case $c_S = c_T = 1$ which excludes the curvelet index distance to be defined below) is not qualitatively different from ours which is why we chose to omit the details.*

**Remark 3.4.** *We already mentioned in the introduction that a somewhat unsatisfactory feature of our result is that the quantity $N^+$ depends on $C_0$. We think that it ought to be possible to remove this dependency, maybe using some clever application of the tensor-power trick (see* `http://www.tricki.org/article/The_tensor_power_trick`*), although we have not succeeded in doing so. Once again we note that known results from wavelet theory also suffer from this caveat. Our result is as strong as known results for wavelet Riesz bases but much more general in that it can be applied to much more general scenarios such as shearlet frames.*

# 4 Applications for Frames

## 4.1 Basic Definitions

Our original motivation comes from the field of applied harmonic analysis where one studies frames, which are a redundant generalization of orthonormal bases, see [4] for more information.

**Definition 4.1.** *Let $\mathcal{H}$ be a Hilbert space and $\Psi := (\psi_\lambda)_{\lambda\in\Lambda}$ a collection of elements $\psi_\lambda \in \mathcal{H}$. The system $\Psi$ is called a* frame *for $\mathcal{H}$ if there exist constants $0 < A \leq B < \infty$ such that*

$$A^2\|f\|_{\mathcal{H}}^2 \leq \sum_{\lambda\in\Lambda} |(f,\psi_\lambda)_{\mathcal{H}}|^2 \leq B^2\|f\|_{\mathcal{H}}^2. \tag{32}$$

*The* frame operator $S : \mathcal{H} \to \mathcal{H}$*, defined as*

$$Sf := \sum_{\lambda\in\Lambda} (f,\psi_\lambda)_{\mathcal{H}}\,\psi_\lambda$$

*is self-adjoint and satisfies*

$$\sigma(S) \subset [A,B]. \tag{33}$$

*The* canonical dual frame $\tilde{\Psi} = \left(\tilde{\psi}_\lambda\right)_{\lambda\in\Lambda}$ *is defined as*

$$\tilde{\psi}_\lambda := S^{-1}\psi_\lambda$$

*and we have the representation formula*

$$f = \sum_{\lambda\in\Lambda} (f,\psi_\lambda)_{\mathcal{H}}\,\tilde{\psi}_\lambda = \sum_{\lambda\in\Lambda}(f,\tilde{\psi}_\lambda)_{\mathcal{H}}\psi_\lambda, \quad \textit{for all } f\in\mathcal{H}. \tag{34}$$

**Remark 4.2.** *For later convenience we use the following notation for sequences $\mathbf{c} = (c_\lambda)_{\lambda\in\Lambda} \in \mathbb{C}^\Lambda$, elements $f \in \mathcal{H}$ and systems $\Psi = (\psi_\lambda)_{\lambda\in\Lambda} \in \mathcal{H}^\Lambda$, $\Phi = (\phi_\mu)_{\mu\in M} \in \mathcal{H}^M$:*

$$(\Psi,\Phi)_{\mathcal{H}} := ((\psi_\lambda,\phi_\mu)_{\mathcal{H}})_{\lambda\in\Lambda,\mu\in M} \in \mathbb{C}^{\Lambda\times M}, \quad \mathbf{c}^\top\Psi := \sum_{\lambda\in\Lambda} c_\lambda\psi_\lambda, \quad \textit{and} \quad (f,\Psi)_{\mathcal{H}} := ((f,\psi_\lambda)_{\mathcal{H}})_{\lambda\in\Lambda} \in \mathbb{C}^\Lambda.$$

Given a frame $\Psi$ we now wish to consider the *Gramian*

$$\mathbf{A} := (\Psi,\Psi)_{\mathcal{H}} := ((\psi_\lambda,\psi_{\lambda'}))_{\lambda,\lambda'\in\Lambda},$$

e.g.,

$$\mathbf{Ac} = (\mathbf{c}^\top\Psi,\Psi)_{\mathcal{H}} := \left((\sum_{\lambda'\in\Lambda} c_{\lambda'}\psi_{\lambda'},\psi_\lambda)_{\mathcal{H}}\right)_{\lambda\in\Lambda}, \ \mathbf{c}\in l_2(\Lambda).$$

**Lemma 4.3.** *We have*

$$\sigma(\mathbf{A})_2 \subset \{0\}\cup[A,B], \tag{35}$$

*and, denoting the Moore-Penrose pseudoinverse of $\mathbf{A}$ by $\mathbf{A}^+$, that*

$$(\tilde{\Psi},\tilde{\Psi})_{\mathcal{H}} = \mathbf{A}^+. \tag{36}$$

*Proof.* Equation (35) follows from [15, 7]. In view of (36), in [15, Lemma 3.1] it is shown that

$$(\tilde{\Psi}, \tilde{\Psi})_{\mathcal{H}} = (\mathbf{A}^{+})^2 \mathbf{A},$$

which, by (13), is equal to $\mathbf{A}^{+}$. □

We can now consider an index distance function $\omega$ satisfying all of the assumptions of Section 2 and introduce the notion of an *$N$-localized frame* $\Psi$, meaning that its Gramian $\mathbf{A}$ is $N$-localized. In this context the main result Theorem 3.1 reads as follows:

**Theorem 4.4.** *Assume that $\Psi$ is a frame of $\mathcal{H}$ with frame constants $A, B$. Assume further that $\Psi$ is indexed with $\Lambda$ and that $\omega$ is an index distance function on $\Lambda$ satisfying the assumptions from Section 2. Then, if $\Psi$ is $N$-localized, its dual frame $\tilde{\Psi}$ is $N^{+}$-localized with $N^{+}$ given by (15).*

## 4.2 Function Spaces and $N$-term Approximation

A result of the above type is of great interest since it provides information on the properties of dual frames. We might for instance consider the spaces

$$\mathcal{B}^{\alpha}_{p,q}(\mathcal{H}) := \left\{ f \in \mathcal{H} : \ \|(f, \Psi)\|_{\dot{b}^{\alpha}_{p,q}} < \infty \right\},$$

with

$$\|\mathbf{c}\|_{\dot{b}^{\alpha}_{p,q}} := \left\| 2^{\alpha j} \, \|\mathbf{c}_j\|_{l_p(\Lambda_j)} \right\|_{l_q(\mathbb{N})},$$

where $\mathbf{c}_j := (c_\lambda)_{\lambda \in \Lambda_j}$ and $\Lambda_j := \{\lambda \in \Lambda : s_\lambda = j\}$. A natural question to ask is under which circumstances do we have

$$\tilde{\mathcal{B}}^{\alpha}_{p,q}(\mathcal{H}) = \mathcal{B}^{\alpha}_{p,q}(\mathcal{H}) \tag{37}$$

with

$$\tilde{\mathcal{B}}^{\alpha}_{p,q}(\mathcal{H}) = \left\{ f \in \mathcal{H} : \ \|(f, \tilde{\Psi})\|_{\dot{b}^{\alpha}_{p,q}} < \infty \right\}?$$

The answer lies in the localization properties of the frames $\Psi, \tilde{\Psi}$: It holds that

$$(f, \Psi) = \mathbf{A}(f, \tilde{\Psi}) \quad \text{and} \quad (f, \tilde{\Psi}) = \mathbf{A}^{+}(f, \Psi). \tag{38}$$

In order to verify (37) we thus need to establish that the operators $\mathbf{A}, \mathbf{A}^{+}$ are bounded on the space of sequences with finite $\dot{b}^{\alpha}_{p,q}$- seminorm. But this follows directly from the fact that both $\mathbf{A}, \mathbf{A}^{+}$ are $N$-localized with $N$ large enough (depending on $p, q, \alpha$). We skip the details, which are standard. A particularly important subcase is $p = q = \frac{1}{1/2+s}$ and $\alpha = 0$ which (more or less) describe the elements $f \in \mathcal{H}$ approximable by $\Psi$ with an $N$-term approximation rate $s$ [13]:

$$\tilde{\mathcal{B}}^{0}_{\frac{1}{1/2+s}, \frac{1}{1/2+s}} \subset \mathcal{A}^s(\mathcal{H}, \Psi), \tag{39}$$

where

$$\mathcal{A}^s(\mathcal{H}, \Psi) = \left\{ f \in \mathcal{H} : \ \inf \left\{ \|f - \sum_{\lambda \in \Lambda^N} c_\lambda \psi_\lambda \|_{\mathcal{H}} : \ |\Lambda^N| = N \right\} \lesssim N^{-s} \right\}.$$

We remark that the question whether the converse inclusion holds in (39) is delicate and in general unsolved for frame systems, see for example [17]. How can we verify that a given class $\mathcal{C} \subset \mathcal{H}$ possesses a certain $N$-term approximation rate? Of course, we could compute the $\tilde{\mathcal{B}}^{0}_{\frac{1}{1/2+s}, \frac{1}{1/2+s}}$-seminorm for all elements in $\mathcal{C}$ and see whether it stays bounded. This would amount to computing

$$\tilde{C} := \sup_{f \in \mathcal{C}} \left\| (f, \tilde{\Psi})_{\mathcal{H}} \right\|_{l_p(\Lambda)}, \quad p = \frac{1}{1/2+s}. \tag{40}$$

However, in general the direct computation of the quantities $(f, \tilde{\Psi})_{\mathcal{H}}$ is elusive, since the dual frame is usually not explicitly known. What we *can* do is compute

$$C := \sup_{f \in \mathcal{C}} \|(f, \Psi)_{\mathcal{H}}\|_{l_p(\Lambda)}, \quad p = \frac{1}{1/2+s}. \tag{41}$$

If this quantity is finite, we can use (38) and $N$-localization of $\mathbf{A}^+$ in order to infer finiteness of $\tilde{C}$ from the finiteness of the (computable) quantity $C$.

To make this last statement a bit more precise we give the following simple result:

**Proposition 4.5.** *Assume that $\mathbf{A}^+$ is $N$-localized with $N > p^{-1}\max(k, M)$, where $k, M$ are the constants from Assumption 2.4. Then $\mathbf{A}^+$ is bounded on $l_p(\Lambda)$. Consequently, boundedness of $C$ implies boundedness of $\tilde{C}$.*

*Proof.* Boundedness of $\mathbf{A}^+ = \left(a^+_{\lambda,\lambda'}\right)_{\lambda,\lambda'\in\Lambda}$ on $l_p(\Lambda)$ follows if

$$\max_{\lambda'\in\Lambda} \sum_{\lambda\in\Lambda} |a^+_{\lambda,\lambda'}|^p < \infty. \tag{42}$$

By the localization property, (42) can be bounded, up to a constant, by

$$\max_{\lambda'\in\Lambda} \sum_{\lambda\in\Lambda} \omega\left(\lambda, \lambda'\right)^{-Np} = \max_{\lambda'\in\Lambda} \sum_{j\in\mathbb{N}} \sum_{\lambda\in\Lambda_j} 2^{-Np|s_{\lambda'}-j|} \left(1 + 2^{\min(s_{\lambda'},j)} \mathrm{d}\left(\lambda, \lambda'\right)\right)^{-Np}.$$

By Assumption 2.4 we can bound this expression by

$$\sum_{j\in\mathbb{N}} 2^{-Np|s_{\lambda'}-j|} 2^{M|j-s'_{\lambda}|},$$

whenever $Np > k$. If additionally $Np > M$ the result follows. $\square$

A little more work yields the following result which we state without proof.

**Proposition 4.6.** *Assume that $\mathbf{A}^+$ is $N$-localized with $N$ sufficiently large. Then $\mathbf{A}^+$ is bounded on $b^\alpha_{p,q}(\Lambda)$. Consequently, in this case the equation (37) holds true.*

Another example concerns the compression of an operator $F : \mathcal{H} \to \mathcal{H}$ using a frame $\Psi$. What is typically possible is to compute the matrix $\tilde{\mathcal{F}} := (F\Psi, \Psi)$ and to show that this matrix is $N$-localized. However, in order to infer the *operator compression property*

$$F\tilde{\mathcal{B}}^0_{\frac{1}{1/2+s},\frac{1}{1/2+s}} \subset \tilde{\mathcal{B}}^0_{\frac{1}{1/2+s},\frac{1}{1/2+s}} \tag{43}$$

guaranteeing fast algorithmic evaluation of $F$ in the spirit of [1], we actually need to establish off-diagonal decay properties of the matrix

$$\mathbf{F} := (F\tilde{\Psi}, \tilde{\Psi}).$$

Again, since $\tilde{\mathbf{F}} = \mathbf{A}^+\mathbf{F}$, we can use the order $N$-localization property of $\mathbf{A}^+$ to show (43).

To sum up this discussion, our results on localization imply that under certain circumstances, approximation spaces and compression properties can be characterized both using the primal frame $\Psi$ as well as its dual $\tilde{\Psi}$, whichever is more convenient. This is the main reason for our interest in the topic of the present paper.

## 4.3 Examples

We close by giving several examples of frames falling into our setting. In all cases the Hilbert space $\mathcal{H}$ will be of the form $L_2(\mathbb{R}^d)$ for $d \in \mathbb{N}$.

### 4.3.1 Wavelets

The first example of wavelets is the most widely known one [11]. The frame $\Psi$ is generated by translation and isotropic dilation of a *mother wavelet* $\psi \in L_2(\mathbb{R}^d)$:

$$\psi_{j,k}(\cdot) := 2^{dj/2}\psi\left(2^j \cdot -k\right), \quad j \in \mathbb{N}_+,\ k \in \mathbb{Z}^d$$

and

$$\psi_{0,k}(\cdot) := \varphi(\cdot - k), \quad k \in \mathbb{Z}^d,$$

with the so-called *scaling function* $\varphi \in L_2(\mathbb{R}^d)$. The index set $\Lambda$ is thus given by $\mathbb{N} \times \mathbb{Z}^d$, where the factor $\mathbb{N}$ represents the scale of the index. The distance function $d$ is given by

$$d\left((j,k),(j',k')\right) = \left\|2^{-j}k - 2^{-j'}k'\right\|_2$$

and therefore

$$\omega\left((j,k),(j',k')\right) = 2^{|j-j'|}\left(1 + 2^{\min(j,j')}\left\|2^{-j}k - 2^{-j'}k'\right\|_2\right).$$

In this case we have $c_\Lambda = 2$ and $c_S = c_T = 1$. The approximation spaces $\mathcal{B}^\alpha_{p,q}$ in this case are the well-known *Besov spaces* [13, 16]. Our results imply that for a sufficiently well-localized wavelet frame, Besov spaces can be characterized by both the primal as well as the dual frame coefficients. A signal class $\mathcal{C}$ which is particularly well-approximable by wavelets is the class of smooth functions with point singularities. We refer to [14] for details.

The significance of our localization result to wavelet frames is limited because by now several wavelet constructions exist where the dual frame is explicitly given as another wavelet frame with explicitly specified properties [6]. In the next examples we go beyond wavelets.

#### 4.3.2 Curvelets and Shearlets

We start with briefly reviewing the idea behind curvelets and shearlets. They both represent a two-dimensional generalization of wavelets which is optimally adapted to anisotropic structures such as piecewise smooth functions with curved singularities [3, 29]. In order to achieve this, two main ingredients are necessary, namely an anisotropic dilation operation which maps a square to a rectangle obeying the *parabolic scaling relation*

$$\text{width} \sim \text{length}^2,$$

as well as a directional operation which can either be rotation (curvelets) or shearing (shearlets). We refer to [3, 29] for more information. One of the reasons for the success of curvelet and shearlets is that they are capable to optimally approximate the set $\mathcal{C}$ of bivariate, piecewise $C^2$ functions with $C^2$ singularity curves [3, 28, 22] and to optimally compress propagation operators for wave equations [2, 23].

The construction of useful curvelet and shearlet frames is not yet at a stage comparable to that of wavelet constructions. In particular, only recently the first compactly supported shearlet frames have been constructed in [25]. For these frames it is in general not known what their duals look like and what properties they have. In this respect, studying the associated localization properties is relevant.

We now show that our results can be applied to curvelets and shearlets, but first we introduce the index distance function $\omega$ which is natural for these frames, see [23, 2, 31].

**Remark 4.7.** *In order to understand the rest of this section the reader must be familiar with the constructions of curvelet and shearlet frames in [2, 23, 29, 25]. A full description of these works would be beyond the scope of this paper which is why we formulate our results directly in terms of the associated index sets and the corresponding index distance function.*

In general curvelets and shearlets form a frame $\Gamma = (\gamma_\lambda)_{\lambda\in\Lambda}$ for $L_2(\mathbb{R}^2)$ Every index $\lambda$ in the index set $\Lambda$ consists of a *scale* $s_\lambda \in \mathbb{N}$, a *direction* $\theta_\lambda \in [-\pi, \pi)$ and a *location* $x_\lambda \in \mathbb{R}^2$. Specifically for curvelets we have

$$\Lambda^C = \left\{(j,l,k) \in \mathbb{N} \times \mathbb{Z} \times \mathbb{Z}^2 : \ -2^{\lfloor j/2 \rfloor - 1} \le l < 2^{\lfloor j/2 \rfloor - 1}\right\}$$

and for $\lambda = (j,l,k) \in \Lambda$ we have

$$s_\lambda = j, \ \theta_\lambda = \pi l 2^{-\lfloor j/2 \rfloor}, \ x_\lambda = R_{\theta_\lambda}\left(k_1 2^{-j}, k_2 2^{-j/2}\right)^\top, \ k = (k_1, k_2),$$

and $R_\theta$ denotes the rotation matrix by angle $\theta \in [0, 2\pi)$. For shearlets we have

$$\Lambda^S = \left\{(j,l,k,\varepsilon) \in \mathbb{N} \times \mathbb{Z} \times \mathbb{Z}^2 \times \{0,1\} : \ -2^{\lfloor j/2 \rfloor} \le l < 2^{\lfloor j/2 \rfloor}\right\}$$

and for $\lambda = (j,l,k,\varepsilon) \in \Lambda$ we have

$$s_\lambda = j,\ \theta_\lambda = \varepsilon\pi/2 + \arctan\left(-l2^{-\lfloor j/2 \rfloor}\right),\ x_\lambda = (S_l^\varepsilon)^{-1} D_{2^{-j}}^\varepsilon k,$$

where $S_l^0$ denotes the shear matrix $\left(\begin{smallmatrix}1 & l\\ 0 & 1\end{smallmatrix}\right)$, $S_l^1 = \left(\begin{smallmatrix}1 & 0\\ l & 0\end{smallmatrix}\right)$ and $D_s^0 = \mathrm{diag}(s, s^{1/2})$, $D_s^1 = \mathrm{diag}(s^{1/2}, s)$. It is possible to construct bandlimited tight frames of curvelets or shearlets but often tightness has to be sacrificed for other desirable properties, such as compact support in space, to be fulfilled. In general this yields curvelet frame systems $\Gamma = (\gamma_\lambda)_{\lambda\in\Lambda^C}$ or shearlet frame systems $\Sigma = (\sigma_\lambda)_{\lambda\in\Lambda^S}$ which are well-localized with respect to the index distance $\omega(\lambda,\lambda') = 2^{|s_\lambda - s_{\lambda'}|}(1 + \mathrm{d}(\lambda,\lambda'))$ where d denotes the distance function as defined in the next definition.

**Definition 4.8.** *Define the* parabolic pseudodistance *operating on* $\Lambda^C$ *or* $\Lambda^S$ *by*

$$\mathrm{d}(\lambda,\lambda') := |\theta_\lambda - \theta_{\lambda'}|^2 + |x_\lambda - x_{\lambda'}|^2 + |\langle e_\lambda, x_\lambda - x_{\lambda'}\rangle|\,, \tag{44}$$

*where* $e_\lambda := (\cos(\theta_\lambda), \sin(\theta_\lambda))^\top$.

We would like to emphasize the following important point:

*The index distance $\omega$ is the natural notion of index distance for shearlet and curvelet frames! It is the equivalent of the well-known index distance for wavelets studied in Section 4.3.1.*

The above statement has first been observed in [31] and used in [2] for curvelets and in [23] for shearlets.

This should serve as a motivation for studying localization properties with respect to $\omega$. We still need to verify that $\omega$ satisfies all the assumptions from Section 2. This is done in the following lemma.

**Lemma 4.9.** *We have*

$$c_{\Lambda^S} = \frac{5}{4},\quad c_{\Lambda^C} = 2,\quad c_S = 2,\quad c_T = 5. \tag{45}$$

*The index distance $\omega$ is $(2,2)$-admissible on both $\Lambda^C$ and $\Lambda^S$. The admissibility constant $C_A$ for the curvelet parameterization is given by (46).*

*Proof.* **Admissibility.** In order to establish the admissibility we need to consider the quantity

$$\sum_{\lambda\in\Lambda_j} (1 + 2^q \mathrm{d}(\mu,\lambda))^{-2} \le c_S^2 \sum_{\lambda\in\Lambda_j} (1 + 2^q \mathrm{d}(\lambda,\mu))^{-2}$$

for $\mu$ arbitrary. We only consider the case $j \ge q$, since for $j \le q$, the quantity to be estimated can be bounded by the case $j = q$. In the curvelet case the sum on the right-hand side can be estimated by

$$\sum_{l\in\mathbb{Z}}\sum_{k\in\mathbb{Z}^2}\left(1 + |2^{q/2}\theta_\mu - 2^{q/2}\theta_\lambda|^2 + |\langle e_\lambda, 2^{q/2}x_\mu - 2^{q/2}x_\lambda\rangle|^2 + |\langle e_\lambda^\perp, 2^{q/2}x_\mu - 2^{q/2}x_\lambda\rangle|^2 + |\langle e_\lambda, 2^q x_\mu - 2^q x_\lambda\rangle|\right)^{-2},$$

in the shearlet case, the same estimate holds with an additional sum over $\varepsilon \in \{0,1\}$. Let us start by showing the admissibility for the curvelet parameterization in which case we have to estimate the quantity

$$\sup_{\theta\in\mathbb{T},x\in\mathbb{R}^2}\sum_{l\in\mathbb{Z}}\sum_{k\in\mathbb{Z}^2}\left(1 + |\theta - 2^{q/2-\lfloor j/2\rfloor}\pi l|^2 + |x_2 - 2^{q/2-\lfloor j/2\rfloor}k_2|^2 + |x_1 - 2^{q/2-j}k_1|^2 + |x_1 - 2^{q-j}k_1|\right)^{-2}$$

which can be bounded by

$$2\sup_{\theta\in\mathbb{T},x\in\mathbb{R}^2}\sum_{l\in\mathbb{Z}}\sum_{k\in\mathbb{Z}^2}\left(1 + |\theta - 2^{(q-j)/2}l|^2 + |x_2 - 2^{(q-j)/2}k_2|^2 + |x_1 - 2^{q-j}k_1|\right)^{-2}$$

Writing this in the form

$$2^{2(j-q)+1}\sup_{\theta\in\mathbb{T},x\in\mathbb{R}^2}\sum_{l\in\mathbb{Z}}\sum_{k\in\mathbb{Z}^2} 2^{2(q-j)}\left(1 + |\theta - 2^{(q-j)/2}l|^2 + |x - 2^{(q-j)/2}k_2|^2 + |x - 2^{q-j}k_1|\right)^{-2}$$

this can be seen as a Riemann sum approximation of the integral

$$2^{2(j-q)+1}\int_{\mathbb{R}}\int_{\mathbb{R}^2} f_{\theta,x}(\xi,y)\mathrm{d}y\mathrm{d}\xi,$$

with

$$f_{\theta,x}(\xi,y) := \left(1+|\theta-\xi|^2+|x_2-y_2|^2+|x_1-y_1|\right)^{-2}.$$

By translation-invariance we only need to consider $\theta = 0$ and $x = 0$ in the evaluation of the integral. The above integral provides an upper bound for the sum

$$2^{2(j-q)+1}\sum_{l\in\mathbb{Z}}\sum_{k\in\mathbb{Z}^2} 2^{2(q-j)} \min_{(\xi,y)\in I_{l,k}} f_{\theta,x}(\xi,y)$$

with

$$I_{l,k} := 2^{(q-j)/2}[l,l+1]\times 2^{q-j}[k_1,k_1+1]\times 2^{(q-j)/2}[k_2,k_2+1].$$

Since

$$\begin{aligned}|f_{\theta,x}(2^{(q-j)/2l},2^{q-j}k_1,2^{(q-j)/2}k_2) - \min_{(\xi,y)\in I_{l,k}} f_{\theta,x}(\xi,y)| &\le \mathrm{diam}(I_{l,k})\|\nabla f_{\theta,x}\|_{L_\infty[I_{l,k}]}\\ &\le \sqrt{5}\|\nabla f_{\theta,x}\|_{L_\infty[I_{l,k}]}\end{aligned}$$

we can bound

$$\sum_{\lambda\in\Lambda_j}(1+2^q\mathrm{d}(\mu,\lambda))^{-2}$$

by

$$c_S^2 2^{2(j-q)+1}\left(\int_{\mathbb{R}}\int_{\mathbb{R}^2} f_{0,0}(\xi,y)\mathrm{d}y\mathrm{d}\xi + \sqrt{5}\sup_{(\theta,x)\in I_{0,0}}\sum_{l\in\mathbb{Z}}\sum_{k\in\mathbb{Z}^2}\|\nabla f_{\theta,x}\|_{L_\infty[I_{l,k}]}\right).$$

This shows the $(2,2)$-admissibility of $\Lambda^C$ with an admissibility constant

$$C_A = 8\left(\int_{\mathbb{R}}\int_{\mathbb{R}^2} f_{0,0}(\xi,y)\mathrm{d}y\mathrm{d}\xi + \sqrt{5}\sup_{(\theta,x)\in I_{0,0}}\sum_{l\in\mathbb{Z}}\sum_{k\in\mathbb{Z}^2}\|\nabla f_{\theta,x}\|_{L_\infty[I_{l,k}]}\right) < \infty. \tag{46}$$

We have used the fact that $c_S = 2$ which we prove below. The shearlet case can be handled in an analogous way.

**Separability.** To show that $c_{\Lambda^C} = 2$ we need to establish that

$$\omega(\lambda,\lambda') = 2^{|s_\lambda - s_{\lambda'}|}(1+2^{\min(s_\lambda,s_{\lambda'})}\mathrm{d}(\lambda,\lambda'))\ge 2, \qquad \text{for all } \lambda\neq\lambda'\in\Lambda^C.$$

Clearly, we only need to consider the case $s_\lambda = s'_\lambda$ since otherwise we would have $|s_\lambda - s_{\lambda'}|\ge 1$ which would establish the claim. Furthermore we can assume that $\theta_\lambda = \theta_{\lambda'}$, since otherwise

$$1+2^j|\theta_\lambda-\theta_{\lambda'}|^2 \ge 1+2^j\left(\pi 2^{-\lfloor j/2\rfloor}\right)^2\ge 2.$$

It remains to consider the case $s_\lambda = s_{\lambda'}$, $\theta_\lambda = \theta_{\lambda'}$ and $x_\lambda \neq x_{\lambda'}$. Since rotation by $\theta_\lambda$ is an isometry we get

$$1+2^j|x_\lambda - x_{\lambda'}|^2 + 2^j|\langle e_\lambda, x_\lambda - x_{\lambda'}\rangle| = 1+2^j2^{-2j}\delta_1^2+2^j2^{-j}\delta_2^2+2^j2^{-j}\delta_1\ge 2,$$

where $\lambda = (j,l,k)$, $\lambda' = (j,l,k')$ and $\delta = k-k'\neq 0$. Along similar lines it can be shown that also $c_{\Lambda^S} = \frac{5}{4}$. Again we may assume that $j = j'$. We first consider the case $\varepsilon = \varepsilon' = 0$ (the case $\varepsilon=\varepsilon'=1$ being the same). Then

$$|\theta_\lambda - \theta_{\lambda'}|^2 = \left|\arctan\left(-l2^{-\lfloor j/2\rfloor}\right) - \arctan\left(-l'2^{-\lfloor j/2\rfloor}\right)\right|^2 \ge \frac{1}{4}|l'-l|^2 2^{-2\lfloor j/2\rfloor},$$

since $\left|l2^{-\lfloor j/2\rfloor}\right|\le 1$ and $\frac{d}{dx}\arctan(x) = \frac{1}{1+x^2}$. Therefore we only need to consider the case $\lambda = (j,l,k,\varepsilon)$, $\lambda' = (j,l,k',\varepsilon)$. In this case we have

$$|x_\lambda - x_{\lambda'}|^2 = \left(2^{-j}\delta_1 - l2^{-\lfloor j/2\rfloor}\delta_2\right)^2 + 2^{-2\lfloor j/2\rfloor}\delta_2^2,$$

where $\delta = k - k'$. If $\delta_2 \neq 0$ the above quantity is greater or equal $2^{-j}$ which is what we need. If $\delta_2 = 0$ and $\delta_1 \neq 0$ we invoke the final term of d and estimate

$$|\langle e_\lambda, (S_l^\varepsilon)^{-1} D_{2^{-j}}^\varepsilon (\delta_1, 0)^\top \rangle| = \cos(\theta_\lambda) 2^{-j} \delta_1 \geq \frac{1}{\sqrt{2}} 2^{-j}.$$

This proves the case $\varepsilon = \varepsilon'$. For the case $\varepsilon \neq \varepsilon'$ we only need to consider the boundary case $\lambda = (j, 2^{\lfloor j/2 \rfloor}, k, 0)$ and $\lambda' = (j, -2^{\lfloor j/2 \rfloor}, k', 1)$. In this case $\theta_\lambda = \theta_{\lambda'} = \frac{\pi}{4}$. Also here one can show that

$$|x_\lambda - x_{\lambda'}|^2 + |\langle e_\lambda, x_\lambda - x_{\lambda'} \rangle| \geq \frac{1}{\sqrt{2}} 2^{-j},$$

whenever $x_\lambda \neq x_{\lambda'}$. This proves the first two statements regarding the constants $c_{\Lambda^S}$ and $c_{\Lambda^C}$.

**Pseudo-symmetry and pseudo-triangle inequality.** The equality for $c_S$ follows by going through the proof of [2, Proposition 2.2] and keeping track of the implicit constants. The pseudo triangle inequality has also been proven in [2] but it is fair to say that the proof is somewhat difficult to follow. Additionally, only the existence of a finite constant $c_T$ is asserted there. For the convenience of the reader we present an alternative proof which will provide us with the desired expression for the constant $c_T$. We operate in the coordinate system spanned by $e_\lambda, e_\lambda^\perp$ and use the following notation from [2]:

$x_\lambda = (0,0)^\top$, $e_\lambda = (1,0)^\top$, $x_{\lambda'} = (x_1, x_2)^\top$, $e_{\lambda'} = (\cos(\alpha), \sin(\alpha))^\top$, $x_{\lambda''} = (y_1, y_2)^\top$, $e_{\lambda''} = (\cos(\beta), \sin(\beta))^\top$.

We get

$$\mathrm{d}(\lambda, \lambda') = |x_1| + |x_1|^2 + |x_2|^2 + |\alpha|^2, \tag{47}$$

$$\mathrm{d}(\lambda, \lambda'') = |y_1| + |y_1|^2 + |y_2|^2 + |\beta|^2, \tag{48}$$

$$\mathrm{d}(\lambda'', \lambda') = |\cos(\beta)(x_1 - y_1) + \sin(\beta)(x_2 - y_2)| + |x_1 - y_1|^2 + |x_2 - y_2|^2 + |\alpha - \beta|^2. \tag{49}$$

First, we note that by the usual triangle inequality it holds that

$$|x_1|^2 + |x_2|^2 + |\alpha|^2 \leq 2 \left( |y_1|^2 + |y_2|^2 + |\beta|^2 + |x_1 - y_1|^2 + |x_2 - y_2|^2 + |\alpha - \beta|^2 \right). \tag{50}$$

It remains to estimate $|x_1|$ by $\mathrm{d}(\lambda, \lambda'') + \mathrm{d}(\lambda'', \lambda')$. If we fix any $\varepsilon > 0$ this follows easily for $|x_1 - y_1| \geq \varepsilon$ since then

$$|x_1| \leq |\, y_1| + |y_1 - x_1| \leq |y_1| + \varepsilon^{-1} |y_1 - x_2|^2 \leq \varepsilon^{-1} \left( \mathrm{d}(\lambda, \lambda'') + \mathrm{d}(\lambda'', \lambda') \right). \tag{51}$$

For $|x_1 - y_1| := \delta < \varepsilon$ and any *fixed* constant $\nu > 0$ we only need to consider the case $|\beta| \leq \nu \delta^{1/2}$ and $|x_2 - y_2| \leq \nu \delta^{1/2}$, since otherwise we would have

$$|x_1| \leq |\, y_1| + \delta \leq |y_1| + \nu^{-2} (|\beta|^2 + |x_2 - y_2|^2) \leq \nu^{-2} \left( \mathrm{d}(\lambda, \lambda'') + \mathrm{d}(\lambda'', \lambda') \right). \tag{52}$$

Now we note that

$$|\cos(\beta)(x_1 - y_1)| \geq \sqrt{1 - |\beta|^2} |y_1 - x_1| \geq \sqrt{1 - \nu^2 \varepsilon} |y_1 - x_1| \tag{53}$$

and similarly

$$|\sin(\beta)(x_2 - y_2)| \leq \nu^2 \delta \leq \nu^2 |x_1 - y_1|. \tag{54}$$

Putting together (53) and (54) we arrive at

$$|\cos(\beta)(x_1 - y_1) + \sin(\beta)(x_2 - y_2)| \geq (\sqrt{1 - \nu^2 \varepsilon} - \nu^2) |x_1 - y_1|,$$

which, using the inequality

$$\sqrt{1 - \nu^2 \varepsilon} \geq 1 - \sqrt{\varepsilon} \nu,$$

valid for $0 < \nu, \varepsilon < 1$ and putting $\tau := 1 - \sqrt{\varepsilon} \nu - \nu^2$, implies

$$|x_1| \leq \tau^{-1} \left( |y_1| + |\cos(\beta)(x_1 - y_1) + \sin(\beta)(x_2 - y_2)| \right) \leq \tau^{-1} \left( \mathrm{d}(\lambda, \lambda'') + \mathrm{d}(\lambda'', \lambda') \right). \tag{55}$$

Now we may put $\varepsilon = \frac{1}{3}$ and $\nu = \frac{1}{\sqrt{3}}$ to obtain from (51), (52) and (55) that

$$|x_1| \leq 3 \left( \mathrm{d}(\lambda, \lambda'') + \mathrm{d}(\lambda'', \lambda') \right)$$

This inequality, together with (50) and (47), (48), (49) finally yields

$$\mathrm{d}(\lambda, \lambda') \leq 5 \left( \mathrm{d}(\lambda, \lambda'') + \mathrm{d}(\lambda'', \lambda') \right),$$

which is what we wanted. □

We close with the following result which is a direct consequence of Theorem 3.1.

**Theorem 4.10.** *Assume we have a shearlet frame $\Gamma = (\gamma_\lambda)_{\lambda\in\Lambda^S}$ which is $N$-localized with respect to $\omega$, meaning that*

$$|\langle\gamma_\lambda, \gamma_{\lambda'}\rangle| \le C_0\omega(\lambda,\lambda')^{-N} \text{ for all } \lambda,\lambda' \in \Lambda^S,$$

*with*

$$\omega(\lambda,\lambda') = 2^{|s_\lambda - s_{\lambda'}|}\left(1 + 2^{\min(s_\lambda, s_{\lambda'})}\mathrm{d}(\lambda,\lambda')\right),$$

*and* d *given by (44). Then the canonical dual frame $\tilde{\Gamma}$ is $N^+$-localized with $N^+$ given by (15) and $c_S, c_T, c_{\Lambda^S}$ given by (45).*

## 4.4 Adaptive Frame Methods for Elliptic Operator Equations

We give another application concerning the adaptive solution of elliptic operator equations in the spirit of [32, 9]. Our exposition mainly utilizes the notation of [9]. Consider a *Gelfand triple*

$$\mathcal{B} \subset \mathcal{H} \subset \mathcal{B}',$$

meaning that $\mathcal{B}$ is a Banach space, $\mathcal{B}'$ its dual with respect to the pairing $(\cdot,\cdot)_\mathcal{H}$ and the embeddings are dense.

Consider further a boundedly invertible operator $F : \mathcal{B} \to \mathcal{B}'$ and, given $f \in \mathcal{B}'$ we seek $u \in \mathcal{B}$ so that

$$Fu = f. \tag{56}$$

For concreteness we can think of $\mathcal{H} = L_2(\mathbb{R}^d)$, $\mathcal{B} = H^1(\mathbb{R}^d)$, and $F = \Delta$. For a frame $\Psi$ for $\mathcal{H}$ we also need to define the *frame analysis operator*

$$G_\Psi : \left\{ \begin{array}{ccc} \mathcal{H} & \to & l_2(\Lambda) \\ g & \mapsto & (f,\Psi)_\mathcal{H} \end{array} \right.$$

and its dual the *frame reconstruction operator*

$$G_\Psi^* : \left\{ \begin{array}{ccc} l_2(\Lambda) & \to & \mathcal{H} \\ \mathbf{c} & \mapsto & \mathbf{c}^\top\Psi \end{array} \right.$$

We call a frame $\Psi = (\psi_\lambda)_{\lambda\in\Lambda}$ for $\mathcal{H}$ a *Gelfand frame* if $\Psi \subset \mathcal{B}$, $\tilde{\Psi} \subset \mathcal{B}'$, and there exists a Gelfand triple $(\mathcal{B}_d, l_2(\Lambda), \mathcal{B}_d')$ of sequence spaces such that the operators

$$G_\Psi^* : \left\{ \begin{array}{ccc} \mathcal{B}_d & \to & \mathcal{B} \\ \mathbf{c} & \mapsto & \mathbf{c}^\top\Psi \end{array} \right. \quad \text{and} \quad G_{\tilde{\Psi}} : \left\{ \begin{array}{ccc} \mathcal{B} & \to & \mathcal{B}_d \\ f & \mapsto & \left(f,\tilde{\Psi}\right)_\mathcal{H} \end{array} \right.$$

are bounded. In addition, suppose that there exists an isomorphism $D_\mathcal{B} : \mathcal{B}_d \to l_2(\Lambda)$ such that its $l_2(\Lambda)$ adjoint $D_\mathcal{B}^*$ is also an isomorphism. For $\mathcal{B} = H^1(\mathbb{R}^d)$ and $\mathcal{H} = L_2(\mathbb{R}^d)$ and $\Psi$ a wavelet (or curvelet/shearlet) frame we would typically have

$$\mathcal{B}_d = \{\mathbf{c} : (2^{s_\lambda}c_\lambda)_{\lambda\in\Lambda} \in l_2(\Lambda)\} = \dot{b}^1_{2,2}(\Lambda) \cap l_2(\lambda) \text{ and } D_\mathcal{B} = D_\mathcal{B}^* : \mathbf{c} \mapsto (2^{s_\lambda}c_\lambda)_{\lambda\in\Lambda}.$$

We briefly sketch how it can be seen that also curvelet and shearlet frames for the Gelfand triple $H^1(\mathbb{R}^2) \subset L_2(\mathbb{R}^2) \subset H^{-1}(\mathbb{R})$ can be constructed. First, observe that the tight curvelet frame constructed in [2] is indeed a Gelfand frame as desired (this follows quite straightforwardly from considering the Fourier supports of the frame elements). Then, utilize the fact that every reasonable curvelet or shearlet frame is localized with the tight curvelet frame, meaning that their cross-Gramian has fast off-diagonal decay [21]. Finally, using e.g. Theorem 4.10, we see that also its dual is localized with the tight curvelet frame. With arguments akin to those of Section 4.2 it follows that all these frames span the same function spaces, and in particular any shearlet frame which is intrinsically localized and localized with the tight curvelet frame from [2] constitutes a Gelfand frame for the triple $H^1(\mathbb{R}^2) \subset L_2(\mathbb{R}^2) \subset H^{-1}(\mathbb{R})$. The practical use of this fact is however limited by the fact that no constructions of Gelfand frames of curvelet or shearlets are known for finite domains $D \subset \mathbb{R}^2$. This is a challenging problem of ongoing interest [27] and we hope that useful constructions will appear in the near future. Now, the system (56) can be discretized to the infinite discrete system

$$\mathbf{F}\mathbf{c} = \mathbf{f} \tag{57}$$

with

$$\mathbf{F} = (D_{\mathcal{B}}^*)^{-1} G_\Psi F G_\Psi^* D_{\mathcal{B}}^{-1} \text{ and } \mathbf{f} = (D_{\mathcal{B}}^*)^{-1} G_\Psi f.$$

Under our assumptions, the operator $\mathbf{F} : l_2(\Lambda) \to l_2(\Lambda)$ is bounded and boundedly invertible on its range $\operatorname{ran}(\mathbf{F}) = \operatorname{ran}((D_{\mathcal{B}}^*)^{-1} G_\Psi)$, [9, Lemma 4.1]. Therefore, the discrete system (57) can in principle be inverted by applying an iterative solver such as a damped Richardson iteration. In practice one has to compute with finite quantities and therefore only an inexact application of the matrix $\mathbf{F}$ can be computed using a suitable matrix-vector multiplication routine [5]. In general the result of such an application lies outside of $\operatorname{ran}(\mathbf{F})$ and a possible error in the kernel of $\mathbf{F}$ will not decrease in subsequent iterations, see the discussions in [32, 9, 10]. Defining $\mathbf{Q} : l_2(\Lambda) \to \operatorname{ran}(\mathbf{F})$ the orthogonal projection onto the range of $\mathbf{F}$, it is possible to show that this does not have a negative effect on the complexity of the algorithms in these papers, *provided that* $\mathbf{Q}$ *is bounded on a certain weak* $l_p$ *space.* This condition would for instance follow from the fact that $\mathbf{Q}$ is $N$-localized with a sufficiently large $N$ [5, 32]. Establishing its validity has only been shown for certain simple cases with $\mathcal{B} = \mathcal{B}' = \mathcal{H}$ [9]. Even for the case of wavelet frames and the Laplace equation, it is an open question whether it holds true.

As a remedy, in [32] it is shown that using an arbitrary bounded projector

$$\mathbf{P} : l_2(\Lambda) \to l_2(\Lambda)$$

with

$$\ker(\mathbf{P}) = \ker(G_\Psi^* D_{\mathcal{B}}^{-1}) \tag{58}$$

one can modify the algorithms from [32, 9, 10] to provably converge with optimal computational complexity *provided that the matrix* $\mathbf{P}$ *is* $N$*-localized with* $N$ *sufficiently large.* Let us consider the simple case of the Laplace equation, in which case the operator $D_{\mathcal{B}}$ is the diagonal matrix with entries $(2^{s_\lambda})_{\lambda \in \Lambda}$ (both for wavelet and curvelet/shearlet frames). Take a wavelet, curvelet or shearlet frame $\Psi$ for $\mathcal{H}$ and consider the (injective) mapping

$$Z : \begin{cases} \mathcal{B} & \to & l_2(\Lambda) \\ f & \mapsto & D_{\mathcal{B}} G_{\tilde{\Psi}} f \end{cases}$$

By the definition of a Gelfand frame, this mapping is bounded. We also have that

$$G_\Psi^* D_{\mathcal{B}}^{-1} Z f = G_\Psi^* D_{\mathcal{B}}^{-1} D_{\mathcal{B}} G_{\tilde{\Psi}} f = G_\Psi^* G_{\tilde{\Psi}} f = f \quad \text{for all } f \in \mathcal{B}. \tag{59}$$

Therefore, we can put

$$\mathbf{P} := \begin{cases} l_2(\Lambda) & \to & l_2(\Lambda) \\ \mathbf{c} & \mapsto & Z G_\Psi^* D_{\mathcal{B}}^{-1} \mathbf{c} \end{cases}$$

and see, using (59), that this mapping is indeed a projector with (58). To find the matrix representation of $\mathbf{P}$ we consider the systems

$$D_{\mathcal{B}} \tilde{\Psi} := \left(2^{s_\lambda} \tilde{\psi}_\lambda\right)_{\lambda \in \Lambda} D_{\mathcal{B}}^{-1} \Psi := \left(2^{-s_\lambda} \psi_\lambda\right)_{\lambda \in \Lambda}$$

and note that

$$\mathbf{P} = \left(D_{\mathcal{B}} \tilde{\Psi}, D_{\mathcal{B}}^{-1} \Psi\right)_{\mathcal{H}} \tag{60}$$

We claim that $\mathbf{P}$ is localized, provided that $\Psi$ is localized. Indeed, by Theorem 4.4, the cross-Gramian $(\tilde{\Psi}, \Psi)_{\mathcal{H}}$ is $N^+$-localized, provided that $\Psi$ is $N$-localized with $N$ sufficiently large. This means that

$$\left|(\tilde{\psi}_\lambda, \psi_{\lambda'})_{\mathcal{H}}\right| \lesssim 2^{-|s_\lambda - s_{\lambda'}| N^+} (1 + \mathrm{d}(\lambda, \lambda'))^{-N^+}.$$

Consequently, by (60), for the entries of $\mathbf{P} = (p_{\lambda,\lambda'})_{\lambda,\lambda' \in \Lambda}$ we have the estimate

$$|p_{\lambda,\lambda'}| = 2^{s_\lambda - s_{\lambda'}} \left|(\tilde{\psi}_\lambda, \psi_{\lambda'})_{\mathcal{H}}\right| \lesssim 2^{-|s_\lambda - s_{\lambda'}|(N^+ - 1)} (1 + \mathrm{d}(\lambda, \lambda'))^{-N^+},$$

which shows that $\mathbf{P}$ is $N^+ - 1$-localized. In summary, we can construct a projector $\mathbf{P}$ ensuring optimal solvability of the discrete operator equation (57) using the algorithm modSolve from [32] for any frame system $\Psi$ which is sufficiently localized. This includes in particular curvelet and shearlet frames, but to our knowledge our findings are also new in the more classical case of wavelet frames.

## References


[1] G. Beylkin, R. Coifman, and V. Rokhlin. Fast wavelet transforms and numerical algorithms I. *Communications on Pure and Applied Mathematics*, 44(2):141–183, 1991.

[2] E. Candes and L. Demanet. The curvelet representation of wave propagators is optimally sparse. *Communications on Pure and Applied Mathematics*, 58:1472–1528, 2004.

[3] E. Candes and D. Donoho. New tight frames of curvelets and optimal representations of objects with piecewise $C^2$ singularities. *Communications on Pure and Applied Mathematics*, 57(2):219–266, 2004.

[4] O. Christensen. *An Introduction to Frames and Riesz Bases*. Birkhäuser, 2003.

[5] A. Cohen, W. Dahmen, and R. DeVore. Adaptive wavelet methods for elliptic operator equations: convergence rates. *Mathematics of Computation*, 70(233):27–76, 2001.

[6] A. Cohen, I. Daubechies, and J. Feauveau. Biorthogonal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics*, 45(5):485–560, 1992.

[7] J. Conway. *A Course in Functional Analysis*, volume 96. Springer, 1990.

[8] E. Cordero and K. Gröchenig. Localization of frames II. *Applied and Computational Harmonic Analysis*, 17(1):29–47, 2004.

[9] S. Dahlke, M. Fornasier, and T. Raasch. Adaptive frame methods for elliptic operator equations. *Advances in Computational Mathematics*, 27(1):27–63, 2007.

[10] S. Dahlke, T. Raasch, M. Werner, M. Fornasier, and R. Stevenson. Adaptive frame methods for elliptic operator equations: The steepest descent approach. *IMA Journal of Numerical Analysis*, 27(4):717–740, 2007.

[11] I. Daubechies. *Ten Lectures on Wavelets*. SIAM, 2006.

[12] S. Demko, W. F. Moss, and P. W. Smith. Decay rates for inverses of band matrices. *Mathematics of Computation*, 43(168):491–499, 1984.

[13] R. DeVore. Nonlinear approximation. *Acta Numerica*, 7(2):51–150, 1998.

[14] D. Donoho. Unconditional bases are optimal bases for data compression and for statistical estimation. *Applied and Computational Harmonic Analysis*, 1(1):100–115, 1993.

[15] M. Fornasier and K. Gröchenig. Intrinsic localization of frames. *Constructive Approximation*, 22(3):395–415, 2005.

[16] M. Frazier and B. Jawerth. A discrete transform and decompositions of distribution spaces. *Journal of Functional Analysis*, 93(1):34–170, 1990.

[17] R. Gribonval and M. Nielsen. On a problem of Gröchenig about nonlinear approximation with localized frames. *Journal of Fourier Analysis and Applications*, 10(4):433–437, 2004.

[18] K. Gröchenig. Localization of frames, Banach frames, and the invertibility of the frame operator. *Journal of Fourier Analysis and Applications*, 10(2):105–132, 2004.

[19] K. Gröchenig and A. Klotz. Noncommutative approximation: Inverse-closed subalgebras and off-diagonal decay of matrices. *Constructive Approximation*, 32(3):429–466, 2010.

[20] K. Gröchenig and M. Leinert. Wiener's lemma for twisted convolution and Gabor frames. *Journal of the American Mathematical Society*, 17(1):1–18, 2004.

[21] P. Grohs and G. Kutyniok. Parabolic molecules. 2011. manuscript.

[22] K. Guo and D. Labate. Optimally sparse multidimensional representation using shearlets. *SIAM Journal on Mathematical Analysis*, 39(1):298–318, 2008.

[23] K. Guo and D. Labate. Representation of Fourier integral operators using shearlets. *Journal of Fourier Analysis and Applications*, 14(3):327–371, 2008.

[24] S. Jaffard. Propriétés des matrices bien localisées pres de leur diagonale et quelques applications. *Annales de l'Institut Henri Poincaré (C) Analyse Non Linéaire*, 7(5):461–476, 1990.

[25] P. Kittipoom, G. Kutyniok, and W. Lim. Construction of compactly supported shearlet frames. *Constructive Approximation*, pages 1–52, 2010.

[26] S. H. Kulkarni, M. T. Nair, and G. Ramesh. Some properties of unbounded operators with closed range. *Proc. Indian Acad. Sci. Math. Sci.*, 118(4):613–625, 2008.

[27] G. Kutyniok and W. Lim. Shearlets on bounded domains. *Arxiv preprint arXiv:1007.3039*, 2010.

[28] G. Kutyniok and W.-Q. Lim. Compactly supported shearlets are optimally sparse. *Journal of Approximation Theory*, 2011.

[29] D. Labate, W. Lim, G. Kutyniok, and G. Weiss. Sparse multidimensional representation using shearlets. In *SPIE Proc. 5914, SPIE, Bellingham, WA*, pages 254–262, 2005.

[30] P. Lemarié. Base d'Ondelettes sur les Groupes de Lie stratifiés. *Bull. Soc. Math. France*, 117(2):213–232, 1989.

[31] H. Smith. A parametrix construction for wave equations with $C^{1,1}$-coefficients. *Annales de l'Institut Fourier*, 48:797–835, 1998.

[32] R. Stevenson. Adaptive solution of operator equations using wavelet frames. *SIAM Journal on Numerical Analysis*, pages 1074–1100, 2004.

# Research Reports

| No. | Authors/Title |
|---|---|
| 12-04 | *P. Grohs*<br>Intrinsic localization of anisotropic frames |
| 12-03 | *P. Grohs*<br>Geometric multiscale decompositions of dynamic low-rank matrices |
| 12-02 | *D. Kressner and C. Tobler*<br>htucker - A Matlab toolbox for tensors in hierarchical Tucker format |
| 12-01 | *F.Y. Kuo, Ch. Schwab and I.H. Sloan*<br>Quasi-Monte Carlo methods for high dimensional integration - the standard (weighted Hilbert space) setting and beyond |
| 11-72 | *P. Arbenz, A. Hiltebrand and D. Obrist*<br>A parallel space-time finite difference solver for periodic solutions of the shallow-water equation |
| 11-71 | *M.H. Gutknecht*<br>Spectral deflation in Krylov solvers: A theory of coordinate space based methods |
| 11-70 | *S. Mishra, Ch. Schwab and J. Šukys*<br>Multi-level Monte Carlo finite volume methods for shallow water equations with uncertain topography in multi-dimensions |
| 11-69 | *Ch. Schwab and E. Süli*<br>Adaptive Galerkin approximation algorithms for partial differential equations in infinite dimensions |
| 11-68 | *A. Barth and A. Lang*<br>Multilevel Monte Carlo method with applications to stochastic partial differential equations |
| 11-67 | *C. Effenberger and D. Kressner*<br>Chebyshev interpolation for nonlinear eigenvalue problems |
| 11-66 | *R. Guberovic, Ch. Schwab and R. Stevenson*<br>Space-time variational saddle point formulations of Stokes and Navier-Stokes equations |
| 11-65 | *J. Li, H. Liu and H. Sun*<br>Enhanced approximate cloaking by SH and FSH lining |
| 11-64 | *M. Hansen and Ch. Schwab*<br>Analytic regularity and best $N$-term approximation of high dimensional parametric initial value problems |
| 11-63 | *R. Hiptmair, G. Phillips and G. Sinha*<br>Multiple point evaluation on combined tensor product supports |