

EXACT AND APPROXIMATE EXPANSIONS WITH PURE GAUSSIAN WAVEPACKETS

MAARTEN V. DE HOOP, KARLHEINZ GRÖCHENIG, AND JOSÉ LUIS ROMERO

ABSTRACT. We construct frames of wavepackets produced by parabolic dilation, rotation and translation of (a finite sum of) Gaussians and give asymptotics on the analogue of Daubechies frame criterion. We show that the coefficients in the corresponding approximate expansion decay fast away from the wavefront set of the original data.

1. INTRODUCTION

Wavepackets are a powerful tool for the microlocal analysis of (Fourier integral) operators and their wave front sets [13, 38, 39]. A “second dyadic decomposition” [40] with parabolic scaling and rotations endows them with the necessary directional sensitivity to resolve singularities. Related objects, such as curvelets and shearlets, incorporate similar features, in particular the parabolic dilations. Frames of these also facilitate methods to sparsify Fourier integral operators and resolve the wavefront set; the emphasis is often on the applications to image processing and numerical aspects. See [10, 11, 23, 32, 6] for representative references.

We consider wavepackets produced by parabolic dilation, rotation and translation of a function $\varphi : \mathbb{R}^2 \rightarrow \mathbb{C}$, given by

$$(1) \quad \varphi_{j,k,\lambda}(x) := 8^{j/2} \varphi(A_{j,k}x - \lambda), \quad (j \geq 1, 0 \leq k \leq 2^j - 1, \lambda \in \Lambda),$$

where

$$A_{j,k} := \begin{bmatrix} 4^j & 0 \\ 0 & 2^j \end{bmatrix} \begin{bmatrix} \cos(\frac{2k\pi}{2^j}) & -\sin(\frac{2k\pi}{2^j}) \\ \sin(\frac{2k\pi}{2^j}) & \cos(\frac{2k\pi}{2^j}) \end{bmatrix},$$

and $\Lambda \subseteq \mathbb{R}^2$ is a lattice. If φ is rapidly decaying, the packet $\varphi_{j,k,\lambda}$ is localized near $A_{j,k}^{-1}\lambda$ and aligned to an angle of $\frac{-2\pi k}{2^j}$ radians. For the coarse scales we use translates of a second window $\varphi_{0,\lambda} := \varphi_0(\cdot - \lambda)$, which typically is a symmetric Gaussian centered at 0. The complete collection of such wavepackets is then the set

$$(2) \quad \mathcal{C}_\Lambda := \{ \varphi_{0,\lambda} \mid \lambda \in \Lambda \} \cup \{ \varphi_{j,k,\lambda} \mid \lambda \in \Lambda, j \geq 1, 0 \leq k \leq 2^j - 1 \}.$$

2010 *Mathematics Subject Classification.* 41A25, 42B05, 42C40.

Key words and phrases. Gaussian wavepackets, wavefront, parabolic scaling.

K.G. and J.L.R. gratefully acknowledge support from the Austrian Science Fund (FWF) [P26273-N25] and [M1586-N25]. M. dH. was supported in part by National Science Foundation grant CMG DMS-1025318.

Our contribution is twofold: First we consider exact expansions and approximate expansions of $L^2(\mathbb{R}^2)$ functions in terms of these packets using Gaussian windows. Second we study the wavefront set with respect to a frame of wavepackets.

The motivation for using purely Gaussian windows comes from highly efficient numerical algorithms [4], with applications in reflection and global seismology. The Gaussian is the window of choice, because it has the best localization in phase-space and yields the best resolution. The use of wavepackets that consist of pure Gaussians has several technical advantages because a number of computations can be done explicitly.

Applications include data compression, denoising, and wavefield recovery from finite data [2, 3, 41], and wave propagation and inverse scattering described by Fourier integral operators associated with canonical graphs. For the latter application, see [17]. Purely Gaussian wave packets are effective in the parametrix construction of the wave equation. Here the the initial data are decomposed with respect to a frame of such wave packets, because a Gaussian wave packet naturally matches the initial conditions for a multi-scale Gaussian-beam-like solution [35]. Gaussian beams are designed to account for the formation of caustics. The frame property facilitates the derivation of error estimates for the parametrix which decay under scale refinement. One can also generate multi-scale Gaussian-beam-like solutions from the decomposition of boundary values, which play a role in the development of inverse scattering algorithms via wavefield cross correlation that are based on efficient reverse-time continuation [33].

We refer the reader to [4] for a comparison of curvelets and Gaussian wavepackets and [7] for the use of Gaussians as wavelets for the continuous wavelet transform. Whereas the recent results on curvelets and shearlets emphasize the construction of tight frames with compactly supported windows or bandlimited windows [1, 12, 23, 14, 36, 30, 28], we understand much less about the spanning properties and the frame properties of wavepackets with Gaussian windows. For instance, Andersson, Carlsson, and Tenorio [4] treat the effective *approximation* of bounded compactly-supported functions by this kind of packets, but they left open the question whether a full expansion by a set \mathcal{C}_Λ of Gaussian wavepackets represents every function in $L^2(\mathbb{R}^2)$. Likewise, most of the work on Gaussian beams [3, 34, 35, 37] emphasizes the algorithmic aspects and contents itself with good numerical approximations. The question of *exact* representations by a wavepacket expansion is omitted.

Our work complements the numerical results in [4, 35]. We will show in Theorem 3.3 that the wavepacket system \mathcal{C}_Λ constitutes a frame for a large class of Gaussian-like windows and suitable lattices Λ .

Using a Gaussian window adds numerous (technical) difficulties to the rigorous analysis of wavepacket expansions. On the one hand, the Gaussian does not have vanishing moments, which is an absolute must in this analysis. On the other hand, the Gaussian is neither compactly supported nor bandlimited; this requires delicate estimates to control the overlap of dilates and rotations of the Gaussian (see Proposition 2.3).

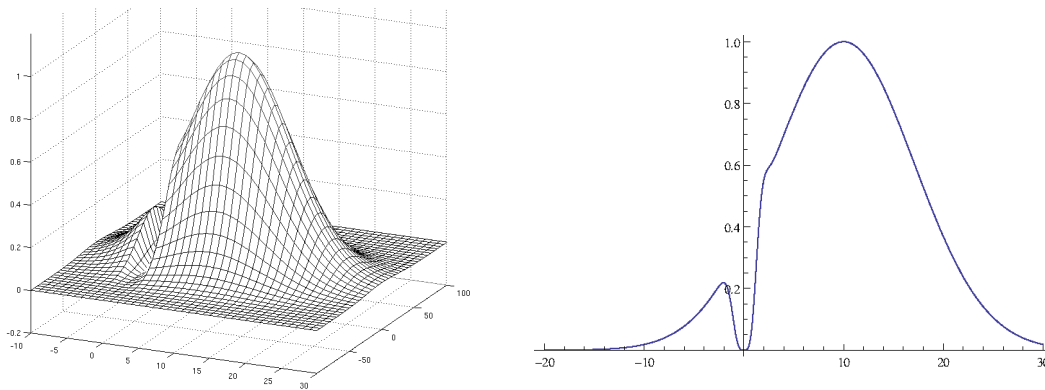


FIGURE 1. A possible choice of $\widehat{\varphi}$ with vanishing moments of order 3. On the left, a plot of $\widehat{\varphi}$, and on the right a plot of its restriction to the ξ_1 axis.

$$\begin{aligned}\widehat{\varphi}(\xi) &= g_1(\xi_1)g_2(\xi_2) \\ g_1(t) &\approx e^{-\frac{(t-10)^2}{100}} + 0.578e^{-(t-1)^2} - 3.45205e^{-(t-0.5)^2} + 4.66167e^{-(t-0.25)^2} - 2.27129e^{-t^2} \\ g_2(t) &= e^{-\frac{t^2}{1100}}.\end{aligned}$$

The problems due to the lack of vanishing moments can be circumvented, as with Morlet's wavelets, by considering linear combinations of modulated Gaussians. As a fundamental example we consider windows of the form,

$$(3) \quad \widehat{\varphi}(\xi_1, \xi_2) := \sum_{k=1}^N a_k e^{-(\delta_0|\xi_1 - t_k|^2 + \delta_k|\xi_2|^2)}, \quad \xi = (\xi_1, \xi_2) \in \mathbb{R}^2,$$

where the points t_1, \dots, t_N , the dilation parameters $\delta_0, \dots, \delta_N$ and the coefficients a_1, \dots, a_N are chosen so that

$$(4) \quad \widehat{\varphi}(0) = \partial_{\xi_1} \widehat{\varphi}(0) = \dots = \partial_{\xi_1}^N \widehat{\varphi}(0) = 0.$$

For such a choice of φ , the wavepackets in (1) are finite linear combinations of Gaussians with different eccentricities, scales and centers. It should be noted that this is different from other Gaussian wavepackets that enforce the vanishing moments through a scale-dependent frequency cut-off (see for example [5]).

The formula in (3) allows for a number of rather different windows. One possibility is to let $\widehat{\varphi}(\xi)$ be essentially a translated Gaussian $e^{-|\xi - (0, t_1)|^2}$, where $t_1 \gg 0$. This window is then corrected with other Gaussians centered near the origin, so as to obtain several vanishing moments. If $t_1 \gg 0$, the coefficients corresponding to these correcting terms can be numerically neglected. The case where t_1 is not large enough to make the correcting terms negligible is also of practical interest. In this case the window φ does not look like a pure Gaussian (see Figure 1). Nevertheless, an expansion in terms of the packets in (3) leads to an expansion in terms of Gaussians.

Throughout the article we make the following assumptions on the windows $\varphi_0, \varphi \in L^2(\mathbb{R}^2)$.

$$(5) \quad 0 < A \leq |\widehat{\varphi}_0(\xi)|^2 + \sum_{j \geq 1} \sum_{k=0}^{2^j-1} |\widehat{\varphi}(B_{j,k}\xi)|^2 \leq B < +\infty, \text{ where } B_{j,k} := (A_{j,k}^*)^{-1},$$

$$(6) \quad |\widehat{\varphi}_0(\xi)| \lesssim e^{-\delta|\xi|^2},$$

$$(7) \quad |\widehat{\varphi}(\xi)| \lesssim \min(1, |\xi_1|^\varsigma) e^{-\delta|\xi|^2}, \text{ for some } \delta > 0 \text{ and } \varsigma > 2.$$

The assumption in (5) means that the windows are wide enough so as to cover the whole frequency plane (see figure 2). The Gaussian wavepackets in (3) clearly satisfy the decay conditions in (6) and (7) while the vanishing moments in the direction ξ_1 are granted by (4).

Our main result shows that \mathcal{C}_Λ is a frame for $L^2(\mathbb{R}^2)$ for sufficiently fine lattices Λ . In Theorem 3.6 we give asymptotics on the lattice parameters required for these packets to generate a frame. The proof uses the method that was developed by Daubechies for wavelets [15]. We will examine the quantities that appear in (the analogue of) Daubechies' criterion for wavelets, and then bound the correlation between rotations and dilations of the window. A similar approach (involving fewer technicalities) was recently carried out in [28] for the so-called cone-adapted shearlets, in order to obtain compactly supported atoms. The fact that Daubechies' criterion can be satisfied at all proves that every $L^2(\mathbb{R}^2)$ function can be expanded in terms of pure Gaussian packets. Thus our main result provides theoretical support for the accuracy of numerical wavepacket methods in [4].

As a by-product of using Daubechies' criterion, we obtain a simple approximate dual frame of the wavepacket system \mathcal{C}_Λ . The main term of the associated frame operator can be written as a Fourier multiplier. Applying the inverse of this Fourier multiplier to the wavepackets we obtain an approximate dual frame that can be easily computed with a simple frequency filter. The error estimate given in Theorem 3.12 may be used to justify when to use the simple approximate expansion (without knowledge of a precise dual frame). In the applications to geophysics that motivated us, the data are noisy and possess a limited numerical accuracy [18]. Therefore, the explicit, simple, but only approximate expansion of the wavefield will be sufficient for many purposes in this context.

Our second contribution is the investigation of the wavefront set with Gaussian wavepackets. We will study the behavior of the coefficients $\langle f, \varphi_{j,k,\lambda} \rangle$ as $j \rightarrow +\infty$. Since \mathcal{C}_Λ is parametrized by a discrete set and is not invariant under translations and rotations, we need to select the parameters $k = k_j, \lambda = \lambda_j$ as a function of the scale in order to track a given phase-space point. Thus the discrete description is more subtle than the descriptions with a continuous transform in [19, 11, 21, 29, 25]. In Theorem 4.1 we will show that if a point does not belong to the wavefront set of f with respect to a Sobolev space H^s , then the corresponding coefficients decay like 4^{-js} . We also consider the approximate wavepacket expansion of Section 3.4 and show that the coefficients decay fast in the scale parameter j , away from the wavefront set of the *original data*. It is worth pointing out that wavefront sets can

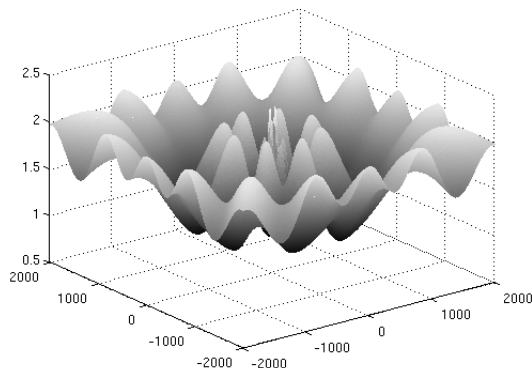


FIGURE 2. A plot of the quantity in (5) for the function φ from Fig. 1 and $\varphi_0(\xi) = e^{-\frac{|\xi|^2}{10000}}$, showing $A \approx 0.9503$ and $B \approx 2.2483$.

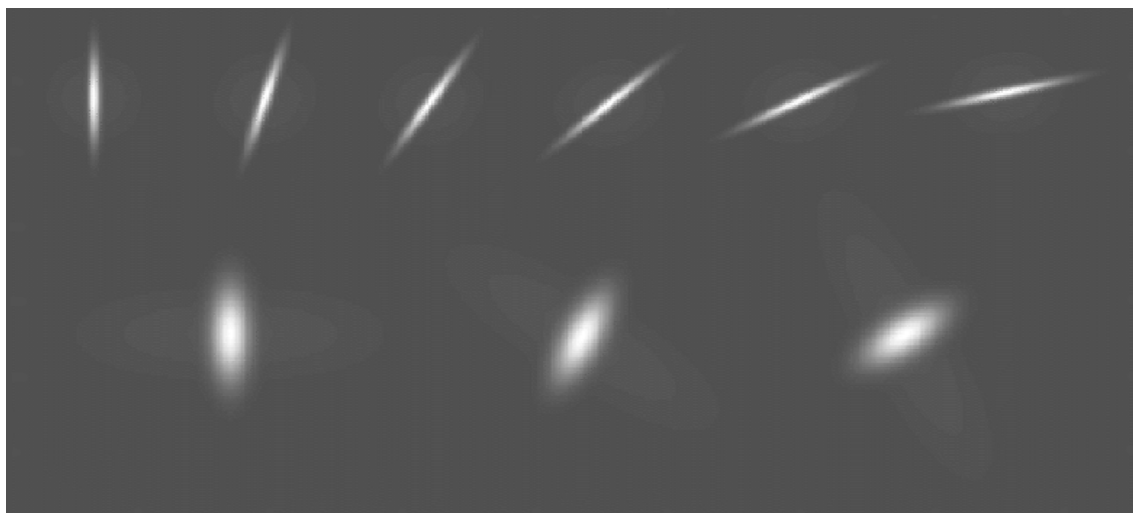


FIGURE 3. A time-domain display of the wavepackets associated with the window depicted in Figure 1 (absolute value).

be characterized with other types of anisotropies besides the parabolic one, see for example [19, Theorem 3.29]. The parabolic scaling, however, is essential for the efficient representation of functions with singularities along smooth curves.

1.1. Related work. The Gaussian wavepackets that we treat in this article share the geometric features of other parabolic wavepackets including curvelets and shearlets. Specifically, the atoms are concentrated on elongated regions obeying the relation: $width \approx length^2$ and oscillate across their ridge. The packets in (1) are produced exactly by applying a parabolic dilation, rotation and translation to a single window function. In contrast, the classic Gaussian wavepackets from [13, 39, 5] use an additional frequency cut-off to enforce vanishing moments. The curvelets from [10] are produced by parabolic dilation and rotation of a slightly different

window at each scale. In the case of shearlets, rotations are replaced by the shear maps $S_s(x, y) = (x + sy, y)$. Wave packets based on tilings of frequency space are discussed in [1, 30, 8, 31].

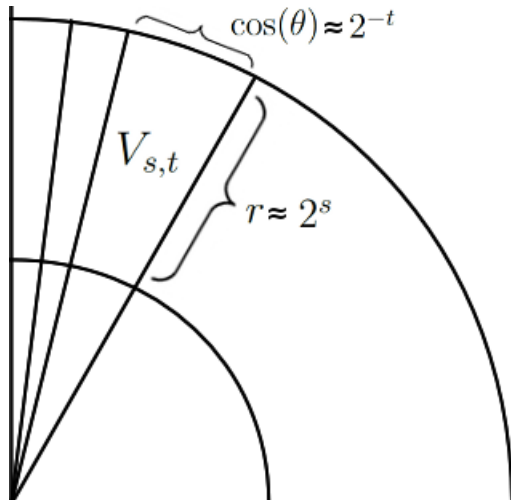
While the differences between the various parabolic wavepackets are very relevant in practice, their asymptotic properties are similar. The Gaussian wavepackets in (1) are an example of *curvelet molecules* from [9] and therefore they provide the same sparse approximation properties as curvelets for functions with discontinuities along smooth edges [10]. The notion of curvelet molecules and the related notion of *shearlet molecules* in [24] are in turn particular cases of the notion of *parabolic molecules*, as recently introduced in [22]. In that article it is shown that all expansions with adequate systems of parabolic molecules share similar asymptotic properties. As a consequence, our results on the decay of frame coefficients away from the wavefront set can be transferred to other parabolic systems.

Closest to our work is the construction of curvelet-type expansions in [36]. Whereas we construct frames by using an analogue of Daubechies criterion for wavelets, the curvelet-type frames for $L^2(\mathbb{R}^2)$ in [36] are obtained by using a perturbation argument. The main focus of that work is the production of compactly-supported atoms, but the results also apply to sums of modulated Gaussians. As mentioned in [36, Section 6], the perturbation argument yields a frame where each element is a linear combination of dilated, rotated and translated Gaussians. However, the coefficients in that linear combination depend on the indices of the particular frame element, but the number of terms is proved to be uniformly bounded. By contrast, the packets in (1) consist exactly of parabolic dilations, rotations and translations of a *single* function, which may be taken to be a fixed linear combination of Gaussians.

The Gaussian wavepackets with parabolic scaling in (1) provide an optimal representation of functions with singularities along smooth curves [10], sparsify Fourier wave propagators [9], and are particularly useful in geophysics [18]. For other kinds of applications, other geometries are more adequate. For example, in [16] it is shown that a dictionary of wave-atoms with the scaling “*wavelength* \approx *diameter*²” is optimal for the representation of oscillatory patterns (texture).

1.2. Organization. The paper is organized as follows. In Section 2 we derive certain estimates on rotation-dilation overlaps. This is perhaps the main technical contribution when working with Gaussian wavepackets. In Section 3 we develop a Daubechies-like frame criterion and derive conditions for when the wavepacket system \mathcal{C}_Λ is a frame. We also study the approximate expansions. In Section 4 we study the complement of the wavefront set and describe it by the decay of the coefficients $\langle f, \varphi_{j,k,\lambda} \rangle$ as $j \rightarrow +\infty$. Finally we show that similar estimates hold for the coefficients in the approximate expansion related to Daubechies’ criterion.

Notation: In the sequel we use the following notation. We let 1_E denote the characteristic function of a set $E \subseteq \mathbb{R}^2$ and $|x|$ denote the Euclidean norm of a vector $x \in \mathbb{R}^2$. The Fourier transform is normalized as $\mathcal{F}f(\xi) = \hat{f}(\xi) := \int_{\mathbb{R}^2} f(x)e^{-2\pi i\xi x} dx$. Translations and dilations of function $f : \mathbb{R}^2 \rightarrow \mathbb{C}$ are defined


 FIGURE 4. The set $V_{s,t}$ defined in (9).

as

$$\begin{aligned} \text{Dil}_A f(x) &:= |\det(A)|^{1/2} f(Ax), \quad A \in \mathbb{R}^{2 \times 2}, \\ T_y f(x) &:= f(x - y). \end{aligned}$$

With this notation, $\varphi_{j,k,\lambda} = \text{Dil}_{A_{j,k}} T_\lambda \varphi$. Note also that with this notation $\text{Dil}_{A \cdot B} f = \text{Dil}_B \text{Dil}_A f$.

Let us further denote the parabolic dilations and the rotations by

$$D_j = \begin{bmatrix} 4^j & 0 \\ 0 & 2^j \end{bmatrix}, \quad R_\theta = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}.$$

Hence $A_{j,k} = D_j R_{\frac{2k\pi}{2^j}}$ and $B_{j,k} := (A_{j,k}^*)^{-1} = D_j^{-1} R_{\frac{2k\pi}{2^j}}$.

If the lattice Λ is given by $\Lambda = P\mathbb{Z}^2$, with $P \in \mathbb{R}^{2 \times 2}$ invertible, then we denote its *volume* by $|\Lambda| = |\det P|$ and its *dual lattice* by $\Gamma = (P^{-1})^* \mathbb{Z}^2$.

2. ESTIMATES FOR ROTATION-DILATION OVERLAPS

We now prove certain technical estimates on rotation-dilation overlaps that will be needed throughout the article. For $f : \mathbb{R}^2 \rightarrow \mathbb{C}$ let the *star-norm* of f be defined by

$$(8) \quad \|f\|_* := \text{esssup}_{\xi \in \mathbb{R}^2} \sum_{j \geq 1} \sum_{k=0}^{2^j-1} |f(B_{j,k}\xi)|.$$

Let us define the circular sectors,

$$(9) \quad V_{s,t} := \left\{ (r \cos(\theta), r \sin(\theta)) : \begin{array}{l} 2^s \leq r \leq 2^{s+1}, 0 \leq \theta \leq \pi/2, \\ 2^{-(t+1)} \leq \cos(\theta) \leq 2^{-t} \end{array} \right\}, \quad (s \in \mathbb{Z}, t \geq 0).$$

The next lemma estimates the overlaps of the orbit of each of these sets under

parabolic dilations and rotations.

Lemma 2.1. *Let $V_{s,t}$ be defined by (9). Then $\|1_{V_{s,t}}\|_* \lesssim 4^t$.*

Proof. The effect of the parabolic dilation D_j can be described (on the first quadrant) in the following way.

$$(10) \quad D_j(r \cos(\theta), r \sin(\theta)) = (r' \cos(\theta'), r \sin(\theta')),$$

$$(11) \quad r' = \rho(\cos(\theta), j)r,$$

$$(12) \quad \tan(\theta') = 2^{-j} \tan(\theta),$$

where $0 \leq \theta \leq \pi/2$, $r > 0$, and $\rho(\alpha, j) \geq 0$ is given by

$$(13) \quad \rho(\alpha, j)^2 = \alpha^2 4^{2j} + (1 - \alpha^2) 2^{2j}, \quad (\alpha \in [0, 1], j \in \mathbb{N}).$$

We note that $\rho(\alpha, j)$ satisfies the growth estimates,

$$(14) \quad \rho(\alpha, j+1) \geq 2\rho(\alpha, j),$$

$$(15) \quad \rho(\alpha/2, j+2) \geq 2\rho(\alpha, j), \quad (\alpha \in [0, 1], j \in \mathbb{N}).$$

Note also that for fixed $j \in \mathbb{N}$, $\rho(\alpha, j)$ is increasing in α .

With this description, we estimate the image of $V_{s,t}$ under the parabolic dilation D_j . For $s \in \mathbb{Z}$, $t \geq 0$ and $j \in \mathbb{N}$, let us define

$$W_{j,s,t} := \left\{ (r \cos(\theta), r \sin(\theta)) : \begin{array}{l} 2^s \rho(2^{-(t+1)}, j) \leq r \leq 2^{s+1} \rho(2^{-t}, j) \\ 0 \leq \theta \leq \pi/2 \min\{1, 2^{-j} 4^{t+1}\} \end{array} \right\}.$$

We note that

$$(16) \quad D_j V_{s,t} \subseteq W_{j,s,t}.$$

Indeed, if $\xi = (r \cos(\theta), r \sin(\theta)) \in V_{s,t}$ and $D_j \xi = (r' \cos(\theta'), r' \sin(\theta'))$, then by (10),

$$r' = \rho(\cos(\theta), j)r \in r[\rho(2^{-(t+1)}, j), \rho(2^{-t}, j)] \subseteq [2^s \rho(2^{-(t+1)}, j), 2^{s+1} \rho(2^{-t}, j)].$$

Since ξ belongs to the first quadrant, so does $D_j \xi$. Hence, $\theta' \in [0, \pi/2)$. In addition, using (12) and the mean value theorem we estimate,

$$\theta' \leq \tan(\theta') = 2^{-j} \tan(\theta) = 2^{-j} \theta (\cos(\tau))^{-2},$$

for some $0 \leq \tau \leq \theta$. Hence, $\cos^2(\tau) \geq \cos^2(\theta) \geq 4^{-t-1}$ and, $\theta' \leq 2^{-j} 4^{t+1} \pi/2$. This shows that (16) holds true.

Hence, D_j maps $V_{s,t}$ into $W_{j,s,t}$, a circular sector of angle $\approx 2^{-j} 4^t$. The 2^j rotations $R_{\frac{2k\pi}{2^j}} W_{j,s,t}$, $k = 0, \dots, 2^j - 1$ have $\lesssim 4^t$ overlaps. Therefore, (16) allows us to estimate for every $j \in \mathbb{N}$ and $\xi \in \mathbb{R}^2$,

$$\begin{aligned} \sum_{k=0}^{2^j-1} 1_{V_{s,t}}(B_{j,k}\xi) &= \sum_{k=0}^{2^j-1} 1_{R_{-\frac{2k\pi}{2^j}} D_j V_{s,t}}(\xi) = \sum_{k=0}^{2^j-1} 1_{R_{\frac{2k\pi}{2^j}} D_j V_{s,t}}(\xi) \\ &\leq \sum_{k=0}^{2^j-1} 1_{R_{\frac{2k\pi}{2^j}} W_{j,s,t}}(\xi) \lesssim 4^t 1_{C_{j,s,t}}(\xi), \end{aligned}$$

where the set $C_{j,s,t}$ is the annulus

$$(17) \quad C_{j,s,t} := \{ \xi \in \mathbb{R}^2 \mid 2^s \rho(2^{-(t+1)}, j) \leq |\xi| \leq 2^{s+1} \rho(2^{-t}, j) \}.$$

Finally we note that these sets do not overlap too much as j varies. Indeed, if $C_{j,s,t} \cap C_{j',s,t} \neq \emptyset$ for $j \leq j'$, then $2^s \rho(2^{-(t+1)}, j') \leq 2^{s+1} \rho(2^{-t}, j)$. Hence, setting $\alpha := 2^{-t}$, we have that $\rho(\alpha/2, j') \leq 2\rho(\alpha, j)$. By (15) this implies that $j' \leq j + 2$. Hence,

$$\|1_{V_{s,t}}\|_* \lesssim 4^t \sup_{\xi \in \mathbb{R}^2} \sum_{j \geq 1} 1_{C_{j,s,t}}(\xi) \lesssim 4^t.$$

□

We now notice the following invariance property of the star-norm.

Lemma 2.2. *Let $f : \mathbb{R}^2 \rightarrow \mathbb{C}$ and let $S : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be one of the four symmetries $(\xi_1, \xi_2) \mapsto (\pm\xi_1, \pm\xi_2)$. Then $\|f \circ S\|_* = \|f\|_*$.*

Proof. If $S(\xi_1, \xi_2) = (\xi_1, \xi_2)$ or $S(\xi_1, \xi_2) = (-\xi_1, -\xi_2)$, then S commutes with $B_{j,k}$ and the conclusion is trivial. If $S(\xi_1, \xi_2) = (-\xi_1, \xi_2)$ or $S(\xi_1, \xi_2) = (\xi_1, -\xi_2)$, then S commutes with D_j and is related to the rotation of angle θ by: $SR_\theta = R_{-\theta}S$. Hence,

$$\begin{aligned} \|f \circ S\|_* &= \operatorname{esssup}_{\xi \in \mathbb{R}^2} \sum_{j \geq 1} \sum_{k=0}^{2^j-1} \left| f(SD_j^{-1}R_{\frac{2k\pi}{2^j}}^{-1}\xi) \right| = \operatorname{esssup}_{\xi \in \mathbb{R}^2} \sum_{j \geq 1} \sum_{k=0}^{2^j-1} \left| f(D_j^{-1}R_{-\frac{2k\pi}{2^j}}^{-1}S\xi) \right| \\ &= \operatorname{esssup}_{\xi \in \mathbb{R}^2} \sum_{j \geq 1} \sum_{k=0}^{2^j-1} \left| f(D_j^{-1}R_{-\frac{2k\pi}{2^j}}^{-1}\xi) \right| = \|f\|_*, \end{aligned}$$

where the last equality follows from the fact that $(R_{\frac{2k\pi}{2^j}})^{2^j} = I$. □

Finally we derive the main technical time-scale correlation estimate.

Proposition 2.3. *Assume that $f : \mathbb{R}^2 \rightarrow \mathbb{C}$ satisfies*

$$(18) \quad |f(\xi)| \leq |\xi_1|^{2+\varepsilon} e^{-\tau|\xi|}$$

for some $\varepsilon, \tau > 0$. Then

$$\|f\|_* = \operatorname{esssup}_{\xi \in \mathbb{R}^2} \sum_{j \geq 1} \sum_{k=0}^{2^j-1} |f(B_{j,k}\xi)| \lesssim 1.$$

Here, the implicit constant depends on τ and ε .

Remark 2.4. Since the window φ generating the wavepackets satisfies (7), it follows that $\|\widehat{\varphi}\|_* < +\infty$.

Proof of Proposition 2.3. Recall the definition of the sets $V_{s,t}$ in (9) and further define $V_{s,t}^i := S^i V_{s,t}$, $i = 0, 1, 2, 3$, where S^i are the four symmetries $(\xi_1, \xi_2) \mapsto (\pm\xi_1, \pm\xi_2)$. From Lemma 2.2, $\|1_{V_{s,t}^i}\|_* = \|1_{V_{s,t}}\|_*$ and hence, by Lemma 2.1,

$\|1_{V_{s,t}^i}\|_* \lesssim 4^t$. Since the sets $\{V_{s,t}^i : s \in \mathbb{Z}, t \in \mathbb{N}, i = 0, 1, 2, 3\}$ cover $\mathbb{R}^2 \setminus \{(\xi_1, \xi_2) : \xi_1 = 0\}$ and $f(\xi) = 0$ if $\xi_1 = 0$, we have the pointwise estimate

$$|f(\zeta)| \leq \sum_{i=0}^3 \sum_{s \in \mathbb{Z}} \sum_{t \geq 0} \sup_{\xi \in V_{s,t}^i} |f(\xi)| 1_{V_{s,t}^i}(\zeta), \quad \zeta \in \mathbb{R}^2,$$

and therefore,

$$\|f\|_* \lesssim \sum_{i=0}^3 \sum_{s \in \mathbb{Z}} \sum_{t \geq 0} 4^t \sup_{\xi \in V_{s,t}^i} |f(\xi)|.$$

Using (18) and writing $r := 2^{2+\varepsilon}$ we see that

$$(19) \quad \sup_{\xi \in V_{s,t}^i} |f(\xi)| = \sup_{\xi \in V_{s,t}} |f(\xi)| \lesssim (2^{s-t+1})^{2+\varepsilon} e^{-\tau 2^s} = (1/r)^t r^{s+1} e^{-\tau 2^s}.$$

Since $r > 4$,

$$\|f\|_* \lesssim \sum_{t \geq 0} (4/r)^t \cdot \sum_{s \in \mathbb{Z}} r^{s+1} e^{-\tau 2^s} < +\infty.$$

□

3. FRAME EXPANSIONS

3.1. Representation of the frame operator. Let S_Λ be the frame operator associated with the set of wavepackets \mathcal{C}_Λ , given by (2)

$$S_\Lambda f := \sum_{\lambda \in \Lambda} \langle f, T_\lambda \varphi_0 \rangle T_\lambda \varphi_0 + \sum_{\lambda \in \Lambda} \sum_{j=1}^{+\infty} \sum_{k=0}^{2^j-1} \langle f, \text{Dil}_{A_{j,k}} T_\lambda \varphi \rangle \text{Dil}_{A_{j,k}} T_\lambda \varphi.$$

Recall that φ and φ_0 satisfy the hypotheses (5) – (7) and that Γ denotes the dual lattice of Λ , precisely, if $\Lambda = P\mathbb{Z}^2$, with $P \in \mathbb{R}^{2 \times 2}$ invertible, then $\Gamma = (P^{-1})^* \mathbb{Z}^2$. Using Poisson's summation formula

$$(20) \quad |\Lambda|^{-1} \sum_{\gamma \in \Gamma} f(x - \gamma) = \sum_{\lambda \in \Lambda} \widehat{f}(\lambda) e^{2\pi i x \lambda}, \quad (f \in \mathcal{S}(\mathbb{R}^2)),$$

we derive a simple description of the action of S_Λ in the frequency domain that leads to an analogue of Daubechies' criterion. We now describe this explicitly.

Let us denote $\widehat{S}_\Lambda := \mathcal{F} S_\Lambda \mathcal{F}^{-1}$; i.e., $\widehat{S}_\Lambda \widehat{f} := \widehat{(S_\Lambda f)}$. Using Poisson's summation formula, \widehat{S}_Λ can be decomposed in the following way.

Lemma 3.1. *For $\gamma \in \Gamma$, consider the operator $\widehat{S}_{\Lambda,\gamma} : L^2(\mathbb{R}^2) \rightarrow L^2(\mathbb{R}^2)$ defined by*

$$(21) \quad \widehat{S}_{\Lambda,\gamma} \widehat{f}(\xi) = f(\xi - \gamma) \widehat{\varphi}_0(\xi) \overline{\widehat{\varphi}_0(\xi - \gamma)} + \sum_{j=1}^{+\infty} \sum_{k=0}^{2^j-1} f(\xi - A_{j,k}^* \gamma) \widehat{\varphi}(B_{j,k} \xi) \overline{\widehat{\varphi}(B_{j,k} \xi - \gamma)}.$$

Then the operator \widehat{S}_Λ can be written as

$$\widehat{S}_\Lambda f = |\Lambda|^{-1} \sum_{\gamma \in \Gamma} \widehat{S}_{\Lambda, \gamma} f.$$

Proof. Let us consider the operators,

$$S_0 f := \sum_{\lambda \in \Lambda} \langle f, T_\lambda \varphi_0 \rangle T_\lambda \varphi_0,$$

$$S_* f := \sum_{\lambda \in \Lambda} \langle f, T_\lambda \varphi \rangle T_\lambda \varphi,$$

$$S_{j,k} f := \sum_{\lambda \in \Lambda} \langle f, \text{Dil}_{A_{j,k}} T_\lambda \varphi \rangle \text{Dil}_{A_{j,k}} T_\lambda \varphi, \quad (j \geq 1, k = 0, \dots, 2^j - 1).$$

Hence, $S_\Lambda = S_0 + \sum_{j,k} S_{j,k}$ and $S_{j,k} = \text{Dil}_{A_{j,k}} S_* \text{Dil}_{A_{j,k}}^{-1}$.

Let us compute

$$\begin{aligned} \mathcal{F} S_0 \mathcal{F}^{-1} f(\xi) &= \sum_{\lambda \in \Lambda} \langle \mathcal{F}^{-1} f, T_\lambda \varphi_0 \rangle e^{-2\pi i \lambda \xi} \widehat{\varphi}_0(\xi) \\ &= \sum_{\lambda \in \Lambda} \langle f, \widehat{\varphi}_0 e^{-2\pi i \lambda \cdot} \rangle e^{-2\pi i \lambda \xi} \widehat{\varphi}_0(\xi) = \sum_{\lambda \in \Lambda} \mathcal{F}(f \overline{\widehat{\varphi}_0})(\lambda) e^{2\pi i \lambda \xi} \widehat{\varphi}_0(\xi). \end{aligned}$$

Using (20), this implies that

$$(22) \quad \mathcal{F} S_0 \mathcal{F}^{-1} f(\xi) = |\Lambda|^{-1} \sum_{\gamma \in \Gamma} f(\xi - \gamma) \overline{\widehat{\varphi}_0(\xi - \gamma)} \widehat{\varphi}_0(\xi).$$

Similarly,

$$\mathcal{F} S_* \mathcal{F}^{-1} f(\xi) = |\Lambda|^{-1} \sum_{\gamma \in \Gamma} f(\xi - \gamma) \overline{\widehat{\varphi}(\xi - \gamma)} \widehat{\varphi}(\xi).$$

Noting that $\mathcal{F} S_{j,k} \mathcal{F}^{-1} = \text{Dil}_{A_{j,k}^*}^{-1} \mathcal{F} S_* \mathcal{F}^{-1} \text{Dil}_{A_{j,k}^*}$, this yields

$$(23) \quad \mathcal{F} S_{j,k} \mathcal{F}^{-1} f(\xi) = |\Lambda|^{-1} \sum_{\gamma \in \Gamma} f(\xi - A_{j,k}^* \gamma) \overline{\widehat{\varphi}(B_{j,k} \xi - \gamma)} \widehat{\varphi}(B_{j,k} \xi).$$

Since $S_\Lambda = S_0 + \sum_{j,k} S_{j,k}$, the conclusion follows from (22) and (23). \square

Following the representation in Lemma 3.1, we define the time-scale correlation function,

$$(24) \quad \Theta(\zeta) := \text{esssup}_{\xi \in \mathbb{R}^2} \left(|\widehat{\varphi}_0(\xi)| |\widehat{\varphi}_0(\xi - \zeta)| + \sum_{j=1}^{+\infty} \sum_{k=0}^{2^j-1} |\widehat{\varphi}(B_{j,k} \xi)| |\widehat{\varphi}(B_{j,k} \xi - \zeta)| \right), \quad (\zeta \in \mathbb{R}^2).$$

We now note that this function bounds the operators appearing in Lemma 3.1.

Lemma 3.2. *The operators in (21) satisfy the bound,*

$$\|\widehat{S}_{\Lambda, \gamma} f\|_2 \leq \max\{\Theta(\gamma), \Theta(-\gamma)\} \|f\|_2, \quad (f \in L^2(\mathbb{R}^2), \gamma \in \Gamma).$$

Proof. It is straightforward to see that $\|\widehat{S_{\Lambda,\gamma}}f\|_1 \leq \Theta(-\gamma)\|f\|_1$ and $\|\widehat{S_{\Lambda,\gamma}}f\|_\infty \leq \Theta(\gamma)\|f\|_\infty$. Hence the conclusion follows by interpolation. \square

3.2. A Daubechies-like frame criterion. We now get the following analogue of Daubechies' criterion for wavelets.

Theorem 3.3. *Let*

$$(25) \quad \Delta(\Lambda) := \sum_{\gamma \in \Gamma \setminus \{0\}} \max\{\Theta(\gamma), \Theta(-\gamma)\}.$$

If $\Delta(\Lambda) < A$, then the set of wavepackets

$$\mathcal{C}_\Lambda = \{ \varphi_{0,\lambda} \mid \lambda \in \Lambda \} \cup \{ \varphi_{j,k,\lambda} \mid \lambda \in \Lambda, j \geq 1, 0 \leq k \leq 2^j - 1 \}$$

is a frame of $L^2(\mathbb{R}^2)$ with bounds $|\Lambda|^{-1}(A - \Delta(\Lambda))$ and $|\Lambda|^{-1}(B + \Delta(\Lambda))$.

Proof. Recall the decomposition $\widehat{S_\Lambda}f = |\Lambda|^{-1} \sum_{\gamma \in \Gamma} \widehat{S_{\Lambda,\gamma}}f$ in Lemma 3.1 and note that

$$\widehat{S_{\Lambda,0}}f(\xi) = f(\xi) \left(|\widehat{\varphi}_0(\xi)|^2 + \sum_{j=1}^{+\infty} \sum_{k=0}^{2^j-1} |\widehat{\varphi}(B_{j,k}\xi)|^2 \right).$$

The covering condition in (5) implies that

$$A|f(\xi)| \leq \left| \widehat{S_{\Lambda,0}}f(\xi) \right| \leq B|f(\xi)|, \quad (f \in L^2(\mathbb{R}^2))$$

and therefore

$$(26) \quad A\|f\|_2 \leq \|\widehat{S_{\Lambda,0}}f\|_2 \leq B\|f\|_2, \quad (f \in L^2(\mathbb{R}^2)).$$

Using Lemma 3.2 we get

$$\begin{aligned} \|\Lambda| \widehat{S_\Lambda}f - \widehat{S_{\Lambda,0}}f\|_2 &\leq \sum_{\gamma \in \Gamma \setminus \{0\}} \|\widehat{S_{\Lambda,\gamma}}f\|_2 \\ &\leq \sum_{\gamma \in \Gamma \setminus \{0\}} \max\{\Theta(\gamma), \Theta(-\gamma)\} \|f\|_2 = \Delta(\Lambda)\|f\|_2. \end{aligned}$$

This together with (26) this implies that for all $f \in L^2(\mathbb{R}^2)$

$$|\Lambda|^{-1}(A - \Delta(\Lambda))\|f\|_2 \leq \|\widehat{S_\Lambda}f\|_2 \leq |\Lambda|^{-1}(B + \Delta(\Lambda))\|f\|_2.$$

Applying the last estimate to \hat{f} and using Plancherel's theorem we conclude that for all $f \in L^2(\mathbb{R}^2)$

$$|\Lambda|^{-1}(A - \Delta(\Lambda))\|f\|_2 \leq \|S_\Lambda f\|_2 \leq |\Lambda|^{-1}(B + \Delta(\Lambda))\|f\|_2,$$

as desired. \square

Remark 3.4. The criterion remains true if $\Delta(\Lambda)$ is replaced by the following smaller quantity, cf. Lemma 3.2,

$$\sum_{\gamma \neq 0} \Theta(\gamma)^{1/2} \Theta(-\gamma)^{1/2}.$$

3.3. Asymptotics for the sampling density and the error. Using the rotation-dilation estimates from Section 2 we can give asymptotics for the frame criterion of Theorem 3.3. Let the lattice Λ be given by $\Lambda = P\mathbb{Z}^2$, with P having singular values $a, b \geq 0$. This means that Λ is a unitary image of the rectangular lattice $a\mathbb{Z} \times b\mathbb{Z}$. Motivated by this, we call a, b the *lattice parameters* of Λ . We now estimate the sampling density and reconstruction error in terms of a and b .

Lemma 3.5. *Under assumptions (5), (6), (7), the quantity $\Delta(\Lambda)$ in (25) satisfies the estimate*

$$\Delta(\Lambda) \lesssim e^{-\tau/a^2} + e^{-\tau/b^2},$$

for some $0 < \tau < \delta$ and $0 < a, b \leq 1$.

Proof. Let $\xi, \zeta \in \mathbb{R}^2$. We use repeatedly the quasi triangle inequality: $|\xi + \zeta|^2 \leq 2(|\xi|^2 + |\zeta|^2)$ and use τ to denote new (smaller) decay parameters that can vary from line to line. By (6),

$$|\widehat{\varphi}_0(\xi)| |\widehat{\varphi}_0(\xi - \zeta)| \lesssim e^{-\delta(|\xi|^2 + |\xi - \zeta|^2)} \leq e^{-\tau|\zeta|^2}.$$

Likewise, by (7),

$$|\widehat{\varphi}(\xi)| |\widehat{\varphi}(\xi - \zeta)| \lesssim |\xi_1|^\varsigma e^{-\delta(|\xi|^2 + |\xi - \zeta|^2)}.$$

Since $|\xi|^2 + |\xi - \zeta|^2 \gtrsim \max\{|\xi|^2, |\zeta|^2\}$, we have that $|\xi|^2 + |\xi - \zeta|^2 \gtrsim |\xi|^2 + |\zeta|^2$. Hence,

$$|\widehat{\varphi}(\xi)| |\widehat{\varphi}(\xi - \zeta)| \lesssim |\xi_1|^\varsigma e^{-\tau|\xi|^2} e^{-\tau|\zeta|^2} = \Phi(\xi) e^{-\tau|\zeta|^2},$$

if we let $\Phi(\xi) := |\xi_1|^\varsigma e^{-\tau|\xi|^2}$. This implies that the time-scale correlation function in (24) satisfies

$$\begin{aligned} \Theta(\zeta) &\lesssim e^{-\delta|\zeta|^2} + \operatorname{esssup}_{\xi \in \mathbb{R}^2} \sum_{j=1}^{+\infty} \sum_{k=0}^{2^j-1} |\widehat{\varphi}(B_{j,k}\xi)| |\widehat{\varphi}(B_{j,k}\xi - \zeta)| \\ &\leq e^{-\tau|\zeta|^2} + e^{-\tau|\zeta|^2} \operatorname{esssup}_{\xi \in \mathbb{R}^2} \sum_{j=1}^{+\infty} \sum_{k=0}^{2^j-1} \Phi(B_{j,k}\xi) \\ &= e^{-\tau|\zeta|^2} (1 + \|\Phi\|_*). \end{aligned}$$

Using Proposition 2.3 we conclude that

$$(27) \quad \Theta(\zeta) \lesssim e^{-\tau|\zeta|^2}.$$

Recall that $\Lambda = P\mathbb{Z}^2$ and let us write $P = UD$ with $D = \operatorname{diag}(a, b)$, U unitary and $a, b \geq 0$. Then $\Gamma = UD^{-1}\mathbb{Z}^2$ and we can estimate

$$\begin{aligned} \Delta(\Lambda) &\lesssim \sum_{(k,j) \in \mathbb{Z}^2 \setminus \{0\}} e^{-\tau|UD^{-1}(k,j)|^2} = \sum_{(k,j) \in \mathbb{Z}^2 \setminus \{0\}} e^{-\tau|D^{-1}(k,j)|^2} \\ &\leq \sum_{(k,j) \in \mathbb{Z}^2 \setminus \{0\}} e^{-\tau k^2/a^2} e^{-\tau j^2/b^2} \lesssim \sum_{(k,j) \in \mathbb{N}^2 \setminus \{0\}} e^{-\tau k/a^2} e^{-\tau j/b^2}. \end{aligned}$$

Summing the geometric series gives

$$\begin{aligned}
\Delta(\Lambda) &\lesssim (1 - e^{-\tau/a^2})^{-1}(1 - e^{-\tau/b^2})^{-1} - 1 = \frac{1 - (1 - e^{-\tau/a^2})(1 - e^{-\tau/b^2})}{(1 - e^{-\tau/a^2})(1 - e^{-\tau/b^2})} \\
&\lesssim 1 - (1 - e^{-\tau/a^2})(1 - e^{-\tau/b^2}) \\
&= e^{-\tau/a^2} + e^{-\tau/b^2} - e^{-\tau/a^2}e^{-\tau/b^2} \\
&\leq e^{-\tau/a^2} + e^{-\tau/b^2}.
\end{aligned}$$

□

As an immediate consequence of Theorem 3.3 and Lemma 3.5 we have the following result.

Theorem 3.6. *Assume that (5), (6), (7) hold. Then there are constants $C, \tau > 0$ depending on δ, ς and ε such that \mathcal{C}_Λ is a frame whenever*

$$(28) \quad \kappa := e^{-\tau/a^2} + e^{-\tau/b^2} \leq CA.$$

Moreover, in this case, $|\Lambda|^{-1}(A - C^{-1}\kappa)$ and $|\Lambda|^{-1}(B + C^{-1}\kappa)$ are frame bounds for \mathcal{C}_Λ .

In particular, for every lattice Λ with sufficiently large density (i.e. $a, b \ll 1$), the system \mathcal{C}_Λ is a frame of $L^2(\mathbb{R}^2)$ and yields the following expansions

$$(29) \quad f = \sum_{\lambda \in \Lambda} \langle f, S_\Lambda^{-1} \varphi_{0,\lambda} \rangle \varphi_{0,\lambda} + \sum_{\lambda \in \Lambda} \sum_{j=1}^{+\infty} \sum_{k=0}^{2^j-1} \langle f, S_\Lambda^{-1} \varphi_{j,k,\lambda} \rangle \varphi_{j,k,\lambda}$$

$$(30) \quad = \sum_{\lambda \in \Lambda} \langle f, \varphi_{0,\lambda} \rangle S_\Lambda^{-1} \varphi_{0,\lambda} + \sum_{\lambda \in \Lambda} \sum_{j=1}^{+\infty} \sum_{k=0}^{2^j-1} \langle f, \varphi_{j,k,\lambda} \rangle S_\Lambda^{-1} \varphi_{j,k,\lambda}.$$

As the lattice parameters $a, b \rightarrow 0^+$, the ratio of the frame bounds tends to B/A .

Remark 3.7. If the window φ is given by (3), then (29) provides a formal argument for the representability of L^2 functions in terms of the Gaussian wavepackets discussed in [4], which can be then computed with the algorithms discussed there.

Remark 3.8. The system \mathcal{C}_Λ is a family of *curvelet molecules* in the sense of [9] and hence the expansion in (30) possesses the same asymptotic sparse approximation as curvelets [10].

Remark 3.9. As the proof of Proposition 2.3 shows, its conclusion holds under much weaker conditions than those in (6), (7), for example polynomial decay.

Remark 3.10. Theorem 3.6 is analogous to a recent result for shearlets [28]. The factor $|\xi_1|^{2+\varepsilon}$ in (7) should be compared with the factor $|\xi_1|^{3+\varepsilon}$ in [28].

3.4. Approximate reconstruction. Since the frame criterion in Theorem 3.3 consists in showing that the frame operator S_Λ is close to $S_{\Lambda,0}$, in practice one often uses $S_{\Lambda,0}$ to construct an approximate dual frame. This approximation is particularly convenient because $S_{\Lambda,0}$ is simply the Fourier multiplier with symbol

$$m(\xi) := |\widehat{\varphi}_0(\xi)|^2 + \sum_{j \geq 1} \sum_{k=0}^{2^j-1} |\widehat{\varphi}(B_{j,k}\xi)|^2, \quad (\xi \in \mathbb{R}^2),$$

(cf. Lemma 3.1). (Recall that the Fourier multiplier with symbol $m \in L^\infty(\mathbb{R}^2)$ is the operator $M_m : L^2(\mathbb{R}^2) \rightarrow L^2(\mathbb{R}^2)$ given by $(\widehat{M_m f})(\xi) = m(\xi)\widehat{f}(\xi)$.)

Hence, the approximate dual system is just the image of \mathcal{C}_Λ by the Fourier multiplier with symbol $1/m$. It then readily follows from the proof of Theorem 4.1 and the estimates of Section 3.3 that this gives a reconstruction error of at most $\Delta(\Lambda)/A \lesssim e^{-\tau/a^2} + e^{-\tau/b^2}$, with a, b the lattice parameters.

By slightly relaxing this bound one can gain certain liberty in the design of the approximate dual system, which can be exploited to improve the function properties. If one considers a second window $\psi \in L^2(\mathbb{R}^2)$, then the frame-type operator $\widetilde{S}_\Lambda : L^2(\mathbb{R}^2) \rightarrow L^2(\mathbb{R}^2)$,

$$\widetilde{S}_\Lambda f := \sum_{\lambda \in \Lambda} \langle f, T_\lambda \varphi_0 \rangle T_\lambda \varphi_0 + \sum_{\lambda \in \Lambda} \sum_{j=1}^{+\infty} \sum_{k=0}^{2^j-1} \langle f, \text{Dil}_{A_{j,k}} T_\lambda \psi \rangle \text{Dil}_{A_{j,k}} T_\lambda \varphi$$

can be decomposed in a way completely analogous to the one in Lemma 3.1. (This can be seen from the proof of Lemma 3.1, or obtained a posteriori by polarization). The Daubechies-like criterion in this case is expressed in terms of lower and upper bounds for

$$(31) \quad \widetilde{m}(\xi) := |\widehat{\varphi}_0(\xi)|^2 + \sum_{j \geq 1} \sum_{k=0}^{2^j-1} \widehat{\varphi}(B_{j,k}\xi) \overline{\widehat{\psi}(B_{j,k}\xi)},$$

and upper bounds for

$$\widetilde{\Delta}(\Lambda) := \sum_{\gamma \in \Gamma \setminus \{0\}} \max\{\widetilde{\Theta}_1(\gamma), \widetilde{\Theta}_2(\gamma)\},$$

where

$$\begin{aligned} \widetilde{\Theta}_1(\zeta) &:= \text{esssup}_{\xi \in \mathbb{R}^2} \left(|\widehat{\varphi}_0(\xi)| |\widehat{\varphi}_0(\xi - \zeta)| + \sum_{j=1}^{+\infty} \sum_{k=0}^{2^j-1} |\widehat{\varphi}(B_{j,k}\xi)| |\widehat{\psi}(B_{j,k}\xi - \zeta)| \right), \\ \widetilde{\Theta}_2(\zeta) &:= \text{esssup}_{\xi \in \mathbb{R}^2} \left(|\widehat{\varphi}_0(\xi)| |\widehat{\varphi}_0(\xi + \zeta)| + \sum_{j=1}^{+\infty} \sum_{k=0}^{2^j-1} |\widehat{\varphi}(B_{j,k}\xi + \zeta)| |\widehat{\psi}(B_{j,k}\xi)| \right). \end{aligned}$$

In our case it makes sense to choose ψ as a small modification of φ that enforces some extra frequency cancellation. Let $\varepsilon > 0$. According to our assumptions, the set $E_\varepsilon := \{ \xi \in \mathbb{R}^2 \mid |\widehat{\varphi}(\xi)| > \varepsilon \}$ is an open bounded set at positive distance from

the origin. Let $\eta : \mathbb{R}^2 \rightarrow [0, 1]$ be a smooth function with compact support not containing the origin such that $\eta \equiv 1$ on E_ε . Let $\psi \in L^2(\mathbb{R}^2)$ be given by

$$\widehat{\psi} := \eta \widehat{\varphi}.$$

The following lemma shows that this approximation is compatible with the quantities in Daubechies' criterion.

Lemma 3.11. *The following estimates hold.*

$$\begin{aligned} |m(\xi) - \widetilde{m}(\xi)| &\leq \|\widehat{\varphi}\|_* \varepsilon \approx \varepsilon, \\ 0 &\leq \widetilde{\Delta}(\Lambda) \leq \Delta(\Lambda). \end{aligned}$$

Proof. Recall that by Remark 2.4 $\|\widehat{\varphi}\|_* < +\infty$ and note that, by definition, $(1 - \eta(\xi)) |\widehat{\varphi}(\xi)| \leq \varepsilon$. Hence, we simply estimate

$$\begin{aligned} |m(\xi) - \widetilde{m}(\xi)| &= \left| \sum_{j \geq 1} \sum_{k=0}^{2^j-1} \widehat{\varphi}(B_{j,k}\xi) (1 - \eta(B_{j,k}\xi)) \overline{\widehat{\varphi}(B_{j,k}\xi)} \right| \\ &\leq \varepsilon \sum_{j \geq 1} \sum_{k=0}^{2^j-1} |\widehat{\varphi}(B_{j,k}\xi)| = \varepsilon \|\widehat{\varphi}\|_*. \end{aligned}$$

The bound for $\widetilde{\Delta}(\Lambda)$ follows from the fact that $|\widehat{\psi}| = |\widehat{\varphi}| \eta \leq |\widehat{\varphi}|$ and consequently $\widetilde{\Theta}_1(\zeta) \leq \Theta(\zeta)$ and $\widetilde{\Theta}_2(\zeta) \leq \Theta(-\zeta)$. \square

As a consequence of these estimates, if $\varepsilon \|\widehat{\varphi}\|_* < A$, then

$$(32) \quad 0 < \widetilde{A} := A - \varepsilon \|\widehat{\varphi}\|_* \leq \widetilde{m}(\xi) \leq \widetilde{B} := B + \varepsilon \|\widehat{\varphi}\|_* < +\infty,$$

and the Fourier multiplier with symbol \widetilde{m} is invertible on $L^2(\mathbb{R}^2)$. We can then consider the following system

$$\begin{aligned} \widetilde{\mathcal{C}}_\Lambda &:= \{ \widetilde{\varphi}_{0,\lambda} \mid \lambda \in \Lambda \} \cup \{ \widetilde{\varphi}_{j,k,\lambda} \mid j \geq 1, 1 \leq k \leq 2^j - 1, \lambda \in \Lambda \}, \\ \widehat{\widetilde{\varphi}_{0,\lambda}}(\xi) &:= \frac{1}{\widetilde{m}(\xi)} \widehat{T_\lambda \varphi_0}(\xi), \\ \widehat{\widetilde{\varphi}_{j,k,\lambda}}(\xi) &:= \frac{1}{\widetilde{m}(\xi)} \widehat{\text{DiI}_{A_j,k} T_\lambda \psi}(\xi). \end{aligned}$$

This provides the following approximate reconstruction.

Theorem 3.12. *Let $\varepsilon \|\widehat{\varphi}\|_* < A$ and $\Delta(\Lambda) < A - \varepsilon \|\widehat{\varphi}\|_*$. Then every $f \in L^2(\mathbb{R}^2)$ admits the following approximate expansion.*

$$(33) \quad \widetilde{f} := |\Lambda| \left(\sum_{\lambda \in \Lambda} \langle f, \widetilde{\varphi}_{0,\lambda} \rangle \varphi_{0,\lambda} + \sum_{j \geq 1} \sum_{k=0}^{2^j-1} \sum_{\lambda \in \Lambda} \langle f, \widetilde{\varphi}_{j,k,\lambda} \rangle \varphi_{j,k,\lambda} \right),$$

with $\|f - \widetilde{f}\|_2 \leq \frac{\Delta(\Lambda)}{A - \varepsilon \|\widehat{\varphi}\|_*} \|f\|_2 \lesssim (e^{-\tau/a^2} + e^{-\tau/b^2}) \|f\|_2$.

Remark 3.13. Note that the coefficients in (33) are given by

$$\langle f, \tilde{\varphi}_{j,k,\lambda} \rangle = \langle M_{1/\tilde{m}} f, \psi_{j,k,\lambda} \rangle,$$

where $M_{1/\tilde{m}}$ is the Fourier multiplier with symbol $1/\tilde{m}$, i.e., $(\widehat{M_{1/\tilde{m}} f})(\xi) = 1/\tilde{m}(\xi)\widehat{f}(\xi)$, and the packets $\psi_{j,k,\lambda} = \text{Dil}_{A_{j,k}} T_\lambda \psi$ have rotation-dilation structure. That is why in practice we prefer the expansion in (33) to the ‘‘abstract’’ expansion in (29).

Proof of Theorem 3.12. Let us denote by $M_{\tilde{m}}$ the Fourier multiplier with symbol \tilde{m} . As in the proof of Theorem 3.3, for every $f \in L^2(\mathbb{R}^2)$

$$\| |\Lambda| \tilde{S}_\Lambda f - M_{\tilde{m}} f \|_2 \leq \tilde{\Delta}(\Lambda) \|f\|_2.$$

Applying this to $M_{\tilde{m}}^{-1} f$ and using Lemma 3.11 and (32) we obtain

$$\| |\Lambda| \tilde{S}_\Lambda M_{\tilde{m}}^{-1} f - f \|_2 \leq \Delta(\Lambda) \|M_{\tilde{m}}^{-1} f\|_2 \leq \frac{\Delta(\Lambda)}{A - \varepsilon \|\widehat{\varphi}\|_*} \|f\|_2.$$

Finally note that $\tilde{f} = |\Lambda| \tilde{S}_\Lambda M_{1/\tilde{m}} f = |\Lambda| \tilde{S}_\Lambda M_{\tilde{m}}^{-1} f$. \square

4. DECAY ESTIMATES AWAY FROM THE WAVEFRONT SET

It is well-know that the continuous transform associated with parabolic representations detects the wavefront set of a distribution, provided that the generating window has sufficiently many vanishing moments [11, 29, 21] (see also [20]). In the case of wavepacket coefficients, the situation is more technical since one needs to approximate a given phase-space point by discrete parameters (see [27] for a similar problem related to Gabor expansions).

4.1. Wavefront sets with wavepackets. Let $(x_0, \theta_0) \in \mathbb{R}^2 \times [0, 2\pi)$. For each scale $j \geq 1$, we select a wavepacket of scale j that is localized near x_0 and is approximately aligned to θ_0 . Recall that the wavepacket $\varphi_{j,k,\lambda}$ is localized near $A_{j,k}^{-1}\lambda$ and aligned to an angle of $\frac{-2\pi k}{2^j}$ radians. We choose the parameters $k_j = k_j(x_0, \theta_0) \in \{0, \dots, 2^j - 1\}$ and $\lambda_j = \lambda_j(x_0, \theta_0) \in \Lambda$ that give the best approximation of θ_0 and x_0 respectively. More precisely, for every $j \geq 1$ we choose (k_j, λ_j) such that

$$(34) \quad \frac{2\pi k_j}{2^j} \leq 2\pi - \theta_0 \leq \frac{2\pi(k_j + 1)}{2^j},$$

$$(35) \quad |A_{j,k_j} x_0 - \lambda_j| \leq L_\Lambda,$$

where L_Λ is a constant that depends on the lattice Λ (namely the diameter of its fundamental domain). For certain (x_0, θ_0) there is more than one possible choice of (k_j, λ_j) . Any of these will be adequate. The resulting sequence $\{(k_j, \lambda_j) : j \geq 1\}$ will be called the *grid parameters* related to (x_0, θ_0) .

The objective of this section is to show that when (x_0, θ_0) does not belong to the wavefront set of f , then the numbers $\langle f, \varphi_{j,k_j,\lambda_j} \rangle$ decay fast as $j \rightarrow +\infty$.

We denote by V_ε the cone of width ε aligned along the ξ_1 -axis

$$(36) \quad V_\varepsilon := \{ \xi \in \mathbb{R}^2 \mid |\xi_2| \leq \varepsilon |\xi_1| \}.$$

We say that a distribution $f \in \mathcal{S}'(\mathbb{R}^2)$ belongs to the microlocal Sobolev space $H^s(x_0, \theta_0)$ if there exist $\varepsilon > 0$ and a smooth compactly-supported function η with $\eta \equiv 1$ near x_0 and such that

$$(37) \quad \int_{R_{\theta_0} V_\varepsilon} \left| \widehat{\eta f}(\xi) \right|^2 |\xi|^{2s} d\xi < +\infty.$$

Since we work with windows with a limited number of vanishing moments, we further consider the space $H_M^s(x_0, \theta_0)$ consisting of distributions in $H^s(x_0, \theta_0)$ that satisfy, in addition to (37),

$$(38) \quad \int_{\mathbb{R}^2} \left| \widehat{\eta f}(\xi) \right|^2 (1 + |\xi|)^{-2M} d\xi < +\infty.$$

Note that $H^s(x_0, \theta_0) = \bigcup_{M>0} H_M^s(x_0, \theta_0)$.

We now state the main microlocal estimate. Although we are mainly interested in the case of L^2 -functions (i.e. $M = 0$), the estimate is valid for distributions provided that the window φ belongs to the Schwartz class.

Theorem 4.1. *Let $s, M \geq 0$. Assume that $\varphi \in \mathcal{S}(\mathbb{R}^2)$ and that $\varsigma > 2s + 2M - 1/2$.*

Let $(x_0, \theta_0) \in \mathbb{R}^2 \times [0, 2\pi)$, and let $\{(k_j, \lambda_j) : j \geq 1\}$ be the corresponding grid parameters given by (34) and (35). Let $f \in H_M^s(x_0, \theta_0)$, then

$$\sum_{j \geq 1} \left| \langle f, \varphi_{j, k_j, \lambda_j} \rangle \right|^2 4^{2js} < +\infty.$$

Since the parameters (k_j, λ_j) do not exactly capture the pair (x_0, θ_0) but only provide the best approximation at scale j , the proof of Theorem 4.1 requires us to control a certain scale-dependent approximation. For this reason we need a microlocal version of Lemma 2.1, as formulated by the following technical statement.

Lemma 4.2. *For $r, t \in \mathbb{Z}$ consider the intervals,*

$$(39) \quad W_{r,t} := \left\{ \xi \in \mathbb{R}^2 \mid 4^{r-1} \leq |\xi_1| \leq 4^r, 2^{t-1} \leq |\xi_2| \leq 2^t \right\}.$$

For each $j \geq 1$, let $\theta_j \in [-2^{-j}2\pi, 2^{-j}2\pi]$ be arbitrary. Then

$$\sup_{r,t \in \mathbb{Z}} \sup_{\xi \in \mathbb{R}^2} \sum_{j \geq 1} 1_{R_{\theta_j} D_j W_{r,t}}(\xi) < +\infty.$$

That is, each of the families of sets $\{R_{\theta_j} D_j W_{r,t} : j \geq 1\}$ has a number of overlaps that is bounded independently of r, t .

Proof. Let $r, t \in \mathbb{Z}$ and suppose that $(R_{\theta_j} D_j W_{r,t}) \cap (R_{\theta_{j'}} D_{j'} W_{r,t}) \neq \emptyset$ for some $1 \leq j \leq j'$. Let $h := j' - j$. We want to find an absolute bound on h .

It follows that there exists some $\xi \in (R_{\theta_{j-\theta_j}} D_j W_{r,t}) \cap (D_h D_j W_{r,t})$. Since $\xi \in D_h D_j W_{r,t}$, we have

$$(40) \quad |\xi_1| \approx 4^{h+j+r} = 2^{2h+2j+2r},$$

$$(41) \quad |\xi_2| \approx 2^{h+j+t}.$$

Similarly, since $\xi \in R_{\theta_j - \theta_{j'}} D_j W_{r,t}$

$$\begin{aligned}\xi_1 &= \cos(\theta_j - \theta_{j'}) \zeta_1 + \sin(\theta_j - \theta_{j'}) \zeta_2, \\ \xi_2 &= -\sin(\theta_j - \theta_{j'}) \zeta_1 + \cos(\theta_j - \theta_{j'}) \zeta_2,\end{aligned}$$

for some $\zeta = (\zeta_1, \zeta_2)$ such that $|\zeta_1| \approx 4^{j+r} = 2^{2j+2r}$ and $|\zeta_2| \approx 2^{j+t}$. Since $|\theta_j - \theta_{j'}| \lesssim 2^{-j} + 2^{-j'} \approx 2^{-j}$, we have that $|\sin(\theta_j - \theta_{j'})| \lesssim 2^{-j}$. Using this and that $|\cos(\theta_j - \theta_{j'})| \leq 1$ we estimate

$$(42) \quad |\xi_1| \lesssim 2^{2j+2r} + 2^{-j} 2^{j+t} = 2^{2j+2r} + 2^t,$$

$$(43) \quad |\xi_2| \lesssim 2^{-j} 2^{2j+2r} + 2^{j+t} = 2^{j+2r} + 2^{j+t}.$$

Comparing these equations with (40) and (41) we get

$$\begin{aligned}2^{2h+2j+2r} &\lesssim 2^{2j+2r} + 2^t, \\ 2^{h+j+t} &\lesssim 2^{j+2r} + 2^{j+t}.\end{aligned}$$

Hence, there exists an absolute constant $C > 0$ such that

$$\begin{aligned}2h + 2j + 2r &\leq \max\{2j + 2r + C, t + C\}, \\ h + j + t &\leq \max\{j + 2r + C, j + t + C\}.\end{aligned}$$

If $h = j' - j \leq C$, we are done. If on the contrary $h > C$, then the previous estimates reduce to

$$(44) \quad 2h + 2j + 2r \leq t + C,$$

$$(45) \quad h + t \leq 2r + C.$$

Since $h, j \geq 0$ we get from (44) that $2r \leq 2h + 2j + 2r \leq t + C$. Plugging this into (45) gives $h + t \leq t + 2C$. Hence $h \leq 2C$. This completes the proof. \square

Proof of Theorem 4.1. For convenience, let us define $\theta_j := \frac{2\pi k_j}{2^j}$ and $\varphi_j(x) := \text{Dil}_{D_j} \varphi(x) = 8^{j/2} \varphi(4^j x_1, 2^j x_2)$. Hence $\varphi_{j,k_j,\lambda_j} = \text{Dil}_{R_{\theta_j}} \text{Dil}_{D_j} T_{\lambda_j} \varphi = \text{Dil}_{R_{\theta_j}} T_{D_j^{-1} \lambda_j} \varphi_j$. We now further define

$$g_j := T_{-D_j^{-1} \lambda_j} \text{Dil}_{R_{-\theta_j}} f, \quad (j \geq 1).$$

Consequently,

$$\langle f, \varphi_{j,k_j,\lambda_j} \rangle = \langle g_j, \varphi_j \rangle.$$

Since $f \in H_M^s(x_0, \theta_0)$ there exist a smooth compactly-supported function η and $\gamma > 0$ such that $\eta \equiv 1$ on the Euclidean ball $B_\gamma(x_0)$ and (37) and (38) hold. Let us consider the functions,

$$\begin{aligned}g_j^1 &:= T_{-D_j^{-1} \lambda_j} \text{Dil}_{R_{-\theta_j}} (\eta f), \quad (j \geq 1), \\ g_j^2 &:= T_{-D_j^{-1} \lambda_j} \text{Dil}_{R_{-\theta_j}} ((1 - \eta) f), \quad (j \geq 1).\end{aligned}$$

Hence,

$$\langle f, \varphi_{j,k_j,\lambda_j} \rangle = \langle g_j^1, \varphi_j \rangle + \langle g_j^2, \varphi_j \rangle.$$

The proof consists in showing that the last two terms have the desired decay as $j \rightarrow +\infty$.

Step 1. We show that $\sum_{j \geq 1} |\langle g_j^2, \varphi_j \rangle|^2 4^{2js} < +\infty$.

Since $\eta \equiv 1$ on $B_\gamma(x_0)$,

$$g_j^2 = T_{-D_j^{-1}\lambda_j} \text{Dil}_{R_{-\theta_j}}((1-\eta)f) \equiv 0 \text{ on } R_{\theta_j}B_\gamma(x_0) - D_j^{-1}\lambda_j = B_\gamma(R_{\theta_j}(x_0 - A_{j,k_j}^{-1}\lambda_j)).$$

Using (35) we estimate,

$$\begin{aligned} \|R_{\theta_j}(x_0 - A_{j,k_j}^{-1}\lambda_j)\| &= \|x_0 - A_{j,k_j}^{-1}\lambda_j\| = \|A_{j,k_j}^{-1}(A_{j,k_j}x_0 - \lambda_j)\| \\ &\leq 2^{-j}\|A_{j,k_j}x_0 - \lambda_j\| \leq 2^{-j}L_\Lambda. \end{aligned}$$

For $j \gg 0$, $2^{-j}L_\Lambda \leq \gamma/2$ and therefore $B_{\gamma/2}(0) \subseteq B_\gamma(R_{\theta_j}(x_0 - A_{j,k_j}^{-1}\lambda_j))$. Hence, for $j \gg 0$,

$$(46) \quad g_j^2 \equiv 0 \text{ on } B_{\gamma/2}(0).$$

Since $(1-\eta)f \in \mathcal{S}'$, there exist N such that for all $h \in \mathcal{S}$,

$$|\langle (1-\eta)f, h \rangle| \lesssim \sum_{|\alpha| \leq N} \|(1+|\cdot|)^N \partial^\alpha h\|_\infty.$$

Using this estimate we have that for all $j \geq 1$,

$$\begin{aligned} |\langle g_j^2, h \rangle| &= \left| \left\langle T_{-D_j^{-1}\lambda_j} \text{Dil}_{R_{-\theta_j}}(1-\eta)f, h \right\rangle \right| \\ &= \left| \left\langle (1-\eta)f, \text{Dil}_{R_{\theta_j}} T_{D_j^{-1}\lambda_j} h \right\rangle \right| \lesssim \sum_{|\alpha| \leq N} \|(1+|\cdot|)^N \partial^\alpha (\text{Dil}_{R_{\theta_j}} T_{D_j^{-1}\lambda_j} h)\|_\infty \\ &\lesssim \sum_{|\alpha| \leq N} \|(1+|\cdot|)^N \partial^\alpha T_{D_j^{-1}\lambda_j} h\|_\infty \\ &\lesssim \sum_{|\alpha| \leq N} \|(1+|\cdot|)^N \partial^\alpha h\|_\infty, \end{aligned}$$

where the last two bounds follow from the fact that $|R_{\theta_j}x| = |x|$ and that, according to (35),

$$\|D_j^{-1}\lambda_j\| = \|D_j^{-1}(\lambda_j - A_{j,k_j}x_0) + R_{\theta_j}x_0\| \leq 2^{-j}L_\Lambda + \|x_0\| \lesssim 1.$$

Taking into account the vanishing property in (46), the bound on $|\langle g_j^2, h \rangle|$ can be improved for $j \gg 0$ to

$$(47) \quad |\langle g_j^2, h \rangle| \lesssim \sum_{|\alpha| \leq N} \|(1+|\cdot|)^N \partial^\alpha h\|_{L^\infty(\mathbb{R}^2 \setminus B_{\gamma'}(0))},$$

where $\gamma' := \gamma/4$.

We finally use this to bound $\langle g_j^2, \varphi_j \rangle$. Since $\varphi \in \mathcal{S}$ we have that for all $L > 2N+2$ there is a constant $C_L > 0$ such that for all multi-indices with $|\alpha| \leq N$

$$|\partial^\alpha \varphi(x)| \leq C_L |x|^{-L}.$$

Hence, when $|x| \geq \gamma'$,

$$\begin{aligned}
 (1 + |x|)^N |\partial^\alpha \varphi_j(x)| &\leq C_{\gamma'} |x|^N |\partial^\alpha \varphi_j(x)| \\
 &\lesssim |x|^N 8^{j/2} 4^{jN} \sum_{\beta \leq \alpha} |\partial^\beta \varphi(4^j x_1, 2^j x_2)| \\
 &\lesssim C_L |x|^N 4^{j(N+1)} |(4^j x_1, 2^j x_2)|^{-L} \\
 &\leq C_L 4^{j(N+1)} 2^{-jL} |x|^{N-L} \leq C_L \gamma'^{(N-L)} 2^{j(2N+2-L)}.
 \end{aligned}$$

Plugging this into (47) shows that $|\langle g_j^2, \varphi_j \rangle| \lesssim 2^{-(2N+2-L)j}$, for all $L > 2N + 2$. This clearly implies the desired estimate.

Step 2. We show that $\sum_{j \geq 1} |\langle g_j^1, \varphi_j \rangle|^2 4^{2js} < +\infty$.

Since $g_j^1 = T_{-D_j^{-1}\lambda_j} \text{Dil}_{R_{-\theta_j}}(\eta f)$, we have

$$|\widehat{g_j^1}(\xi)| = |\widehat{\eta f}(R_{-\theta_j}\xi)|, \quad (\xi \in \mathbb{R}^2).$$

So if we let $\alpha_j := 2\pi - \theta_0 - \theta_j$ and $\Phi(\xi) := \widehat{\eta f}(R_{\theta_0}\xi)$, we see that

$$(48) \quad |\widehat{g_j^1}(\xi)| = \Phi(R_{\alpha_j}\xi), \quad (\xi \in \mathbb{R}^2).$$

Note that from (34) we get

$$(49) \quad 0 \leq \alpha_j \leq 2\pi 2^{-j}, \quad (j \geq 1).$$

Since f satisfies (38), it follows that Φ satisfies,

$$(50) \quad \int_{\mathbb{R}^2} \Phi(\xi)^2 (1 + |\xi|)^{-2M} d\xi < +\infty.$$

In addition, by (37), we have

$$(51) \quad \int_{V_\varepsilon} \Phi(\xi)^2 |\xi|^{2s} d\xi < +\infty,$$

where $\varepsilon > 0$ and the cone V_ε is defined in (36). Let $k \in \mathbb{N}$ be such that $2^{-k} < \varepsilon$. Since $\alpha_j \rightarrow 0$, there exist $j_0 \geq 1$ such that for all $j \geq j_0$,

$$(52) \quad R_{-\alpha_j}(V_{2^{-k}}) \subseteq V_\varepsilon.$$

Consider the sets $\{W_{r,t} : r, t \in \mathbb{Z}\}$ from (39). Since these sets partition \mathbb{R}^2 up to sets of measure zero, we can decompose $\widehat{\varphi}$ as

$$\begin{aligned}
 \widehat{\varphi} &= \sum_{r,t \in \mathbb{Z}} \widehat{\varphi}^{r,t}, \quad \text{a.e.} \\
 \widehat{\varphi}^{r,t} &:= \widehat{\varphi} \cdot 1_{W_{r,t}}.
 \end{aligned}$$

Let us define $K_{r,t} \geq 0$ by

$$(53) \quad K_{r,t}^2 := \int_{\mathbb{R}^2} |\widehat{\varphi}^{r,t}(\xi)|^2 d\xi = \int_{W_{r,t}} |\widehat{\varphi}(\xi)|^2 d\xi.$$

Our main assumption (7) on the window φ implies that

$$K_{r,t}^2 \lesssim 4^{r(2\varsigma+1)} e^{-2\delta 16^{r-1}} 2^t e^{-2\delta 4^{t-1}}.$$

Hence if $\rho_1, \rho_2 \in \mathbb{R}$ are such that $\rho_1 < (2\varsigma + 1)$ and $\rho_2 > -1$, then

$$(54) \quad \sum_{r,t \in \mathbb{Z}} 4^{-\rho_1 r} 2^{\rho_2 t} K_{r,t}^2 < +\infty.$$

We also use the notation $\widehat{\varphi}_j^{r,t}(\xi) := \text{Dil}_{D_{-j}} \widehat{\varphi}^{r,t}(\xi) = 8^{-j/2} \widehat{\varphi}^{r,t}(4^{-j}\xi_1, 2^{-j}\xi_2)$, so for each $j \geq 1$

$$\widehat{\varphi}_j = \sum_{r,t \in \mathbb{Z}} \widehat{\varphi}_j^{r,t}.$$

Fix $r, t \in \mathbb{Z}$ and let us bound $\langle g_j^1, \varphi_j^{r,t} \rangle$.

Since $\widehat{\varphi}^{r,t}$ is supported on $W_{r,t}$, its dilation $\widehat{\varphi}_j^{r,t}$ is supported on $D_j W_{r,t}$. Recall the numbers k, j_0 from (52). The set $D_j W_{r,t}$ is contained in the cone $V_{2^{-k}}$, if $j \geq k + t - 2r + 2$. Let us denote $j_1 = j_1(r, t) := \max\{j_0, k + t - 2r + 2\}$. For $j \geq j_1(r, t)$ we bound,

$$\begin{aligned} |\langle g_j^1, \varphi_j^{r,t} \rangle| &\leq 8^{-j/2} \int_{D_j W_{r,t}} \left| \widehat{g}_j^1(\xi) \right| \left| \widehat{\varphi}^{r,t}(4^{-j}\xi_1, 2^{-j}\xi_2) \right| d\xi \\ &\leq 8^{-j/2} 4^{-js} \int_{D_j W_{r,t}} \left| \widehat{g}_j^1(\xi) \right| |\xi_1|^s |4^{-j}\xi_1|^{-s} \left| \widehat{\varphi}^{r,t}(4^{-j}\xi_1, 2^{-j}\xi_2) \right| d\xi \\ &\leq 4^{-js} \left(\int_{D_j W_{r,t}} \left| \widehat{g}_j^1(\xi) \right|^2 |\xi_1|^{2s} d\xi \right)^{1/2} 8^{-j/2} \left(\int_{D_j W_{r,t}} |4^{-j}\xi_1|^{-2s} \left| \widehat{\varphi}^{r,t}(4^{-j}\xi_1, 2^{-j}\xi_2) \right|^2 d\xi \right)^{1/2} \\ &\leq 4^{-js} \left(\int_{D_j W_{r,t}} \left| \widehat{g}_j^1(\xi) \right|^2 |\xi_1|^{2s} d\xi \right)^{1/2} \left(\int_{W_{r,t}} |\xi_1|^{-2s} \left| \widehat{\varphi}^{r,t}(\xi) \right|^2 d\xi \right)^{1/2} \\ &\lesssim 4^{-js} 4^{-rs} \left(\int_{D_j W_{r,t}} \left| \widehat{g}_j^1(\xi) \right|^2 |\xi_1|^{2s} d\xi \right)^{1/2} \left(\int_{W_{r,t}} \left| \widehat{\varphi}^{r,t}(\xi) \right|^2 d\xi \right)^{1/2}. \end{aligned}$$

Hence, since $|\xi_1| \leq |\xi|$,

$$|\langle g_j^1, \varphi_j^{r,t} \rangle| \leq 4^{-js} 4^{-rs} K_{r,t} \left(\int_{D_j W_{r,t}} \left| \widehat{g}_j^1(\xi) \right|^2 |\xi|^{2s} d\xi \right)^{1/2}, \quad (j \geq j_1(r, t)).$$

Using (48) we get

$$|\langle g_j^1, \varphi_j^{r,t} \rangle| \leq 4^{-js} 4^{-rs} K_{r,t} \left(\int_{R_{-\alpha_j}(D_j W_{r,t})} \Phi(\xi)^2 |\xi|^{2s} d\xi \right)^{1/2}, \quad (j \geq j_1(r, t)).$$

For $j \geq j_1(r, t)$, we have that $D_j W_{r,t} \subseteq V_{2^{-k}}$ so (52) implies that $R_{-\alpha_j}(D_j W_{r,t}) \subseteq V_\varepsilon$. Combining this with (51) and the bound on the number of overlaps of the sets

$\{R_{-\alpha_j}(D_j W_{r,t}) : j \geq 1\}$ granted by Lemma 4.2, we obtain

$$(55) \quad \sum_{j \geq j_1(r,t)} |\langle g_j^1, \varphi_j^{r,t} \rangle|^2 4^{2sj} \lesssim 4^{-2rs} K_{r,t}^2.$$

Since by assumption $s < \varsigma + 1/2$, we use (54) with $\rho_1 = 2s$ and $\rho_2 = 0$ and obtain

$$(56) \quad \sum_{r,t \in \mathbb{Z}} \sum_{j \geq j_1(r,t)} |\langle g_j^1, \varphi_j^{r,t} \rangle|^2 4^{2sj} < +\infty.$$

For $1 \leq j < j_1(r,t)$ we use the bound

$$(1 + |\xi|) \leq 1 + 4^j |D_{-j}\xi| \leq 4^j (1 + |D_{-j}\xi|),$$

and carry out a similar estimate.

$$\begin{aligned} |\langle g_j^1, \varphi_j^{r,t} \rangle| &\leq 8^{-j/2} \int_{D_j W_{r,t}} \left| \widehat{g}_j^1(\xi) \right| \left| \widehat{\varphi}^{r,t}(D_{-j}\xi) \right| d\xi \\ &\leq 8^{-j/2} 4^{jM} \int_{D_j W_{r,t}} \left| \widehat{g}_j^1(\xi) \right| (1 + |\xi|)^{-M} (1 + |D_{-j}\xi|)^M \left| \widehat{\varphi}^{r,t}(D_{-j}\xi) \right| d\xi \\ &\leq 4^{jM} \left(\int_{D_j W_{r,t}} \left| \widehat{g}_j^1(\xi) \right|^2 (1 + |\xi|)^{-2M} d\xi \right)^{1/2} \left(\int_{D_j W_{r,t}} (1 + |D_{-j}\xi|)^{2M} \left| \widehat{\varphi}^{r,t}(D_{-j}\xi) \right|^2 d\xi \right)^{1/2} \\ &\leq 4^{jM} \left(\int_{D_j W_{r,t}} \left| \widehat{g}_j^1(\xi) \right|^2 (1 + |\xi|)^{-2M} d\xi \right)^{1/2} \left(\int_{W_{r,t}} (1 + |\xi|)^{2M} \left| \widehat{\varphi}^{r,t}(\xi) \right|^2 d\xi \right)^{1/2} \\ &\lesssim 4^{jM} (1 + 4^r + 2^t)^M \left(\int_{D_j W_{r,t}} \left| \widehat{g}_j^1(\xi) \right|^2 (1 + |\xi|)^{-2M} d\xi \right)^{1/2} \left(\int_{W_{r,t}} \left| \widehat{\varphi}^{r,t}(\xi) \right|^2 d\xi \right)^{1/2}. \end{aligned}$$

Therefore, for $1 \leq j < j_1(r,t)$,

$$|\langle g_j^1, \varphi_j^{r,t} \rangle| \leq 4^{jM} (1 + 4^r + 2^t)^M K_{r,t} \left(\int_{R_{-\alpha_j}(D_j W_{r,t})} \Phi(\xi)^2 (1 + |\xi|)^{-2M} d\xi \right)^{1/2}.$$

Hence, for $1 \leq j < j_1(r,t)$,

$$|\langle g_j^1, \varphi_j^{r,t} \rangle| 4^{js} \leq 4^{j_1(r,t)(M+s)} (1 + 4^r + 2^t)^M K_{r,t} \left(\int_{R_{-\alpha_j}(D_j W_{r,t})} \Phi(\xi)^2 (1 + |\xi|)^{-2M} d\xi \right)^{1/2}.$$

Using again the bound on the number of overlaps of $\{R_{-\alpha_j}(D_j W_{r,t}) : j \geq 1\}$ of Lemma 4.2, this time combined with (50), we get

$$\sum_{j=j_0}^{j_1(r,t)-1} |\langle g_j^1, \varphi_j^{r,t} \rangle|^2 4^{2sj} \lesssim 4^{2j_1(r,t)(M+s)} (1 + 4^r + 2^t)^{2M} K_{r,t}^2.$$

To bound this quantity we may assume that $j_1(r, t) = k + t - 2r + 2 > j_0$, since otherwise $j_1(r, t) = j_0$ and the sum is empty. Hence,

$$\begin{aligned} \sum_{j=j_0}^{j_1(r,t)-1} |\langle g_j^1, \varphi_j^{r,t} \rangle|^2 4^{2sj} &\lesssim 4^{2(t-2r)(M+s)} (1 + 4^r + 2^t)^{2M} K_{r,t}^2 \\ &\lesssim 2^{4(M+s)t} 4^{-4(M+s)r} (1 + 4^{2Mr} + 2^{2Mt}) K_{r,t}^2 \\ &= (4^{-(4M+4s)r} 2^{(4M+4s)t} + 4^{-(2M+4s)r} 2^{(4M+4s)t} + 4^{-(4M+4s)r} 2^{(6M+4s)t}) K_{r,t}^2. \end{aligned}$$

To bound the sum over $r, t \in \mathbb{Z}$ of the last expression, we use (54) with $\rho_1 = 4M + 4s, \rho_2 = 4M + 4s; \rho_1 = 2M + 4s, \rho_2 = 4M + 4s; \rho_1 = 4M + 4s, \rho_2 = 6M + 4s$. Note that in each case $\rho_1 \leq 4M + 4s < 2\zeta + 1$ by assumption and $\rho_2 \geq 0 > -1$. Hence the use of (54) is justified and we conclude that

$$(57) \quad \sum_{r,t \in \mathbb{Z}} \sum_{j=j_0}^{j_1(r,t)-1} |\langle g_j^1, \varphi_j^{r,t} \rangle|^2 4^{2sj} < +\infty.$$

Finally, since $\varphi_j = \sum_{r,t} \varphi_j^{r,t}$,

$$\sum_{j \geq j_0} |\langle g_j^1, \varphi_j \rangle|^2 4^{2js} \leq \sum_{r,t \in \mathbb{Z}} \sum_{j \geq j_1(r,t)} |\langle g_j^1, \varphi_j^{r,t} \rangle|^2 4^{2js} + \sum_{r,t \in \mathbb{Z}} \sum_{j=j_0}^{j_1(r,t)-1} |\langle g_j^1, \varphi_j^{r,t} \rangle|^2 4^{2js} < +\infty.$$

□

4.2. Estimates for the approximate reconstruction. Theorem 4.1 shows that the coefficients of the wavepacket expansion in (30) decay fast away from the wavefront set of f . We now show that the same is true for the coefficients in the approximate wavepacket expansion in (33). To avoid immaterial technicalities, let us assume that $\varphi_0, \varphi \in \mathcal{S}(\mathbb{R}^2)$.

Theorem 4.3. *In the setting of Theorem 3.12, let $(x_0, \theta_0) \in \mathbb{R}^2 \times [0, 2\pi)$ and let $\{(k_j, \lambda_j) : j \geq 1\}$ be the corresponding grid parameters given by (34) and (35). If $f \in L^2(\mathbb{R}^2)$ belongs to H^s microlocally at (x_0, θ_0) , then*

$$\sum_{j \geq 1} |\langle f, \tilde{\varphi}_{j,k_j,\lambda_j} \rangle|^2 4^{2js} < +\infty.$$

As mentioned in Remark 3.13, the approximate coefficients can be rewritten as

$$\langle f, \tilde{\varphi}_{j,k,\lambda} \rangle = \langle M_{1/\tilde{m}} f, \psi_{j,k,\lambda} \rangle,$$

where $\psi_{j,k,\lambda} = \text{Dil}_{A_{j,k}} T_\lambda \psi$, \tilde{m} is given by (31) and $\widehat{M}_{1/\tilde{m}}(f) := (1/\tilde{m})\widehat{f}$. Since ψ satisfies the same conditions that φ , it follows from Theorem 4.1 that the numbers $\langle f, \tilde{\varphi}_{j,k,\lambda} \rangle$ decay away from the H_s -wavefront set of $M_{1/\tilde{m}}(f)$. Hence, it suffices to show that f and $M_{1/\tilde{m}}(f)$ share the same H_s -wavefront set. We do so in the next lemma.

Lemma 4.4. *Under the hypothesis of Theorem 3.12, let $s \geq 0$. Then a function $f \in L^2(\mathbb{R}^2)$ belongs to the microlocal Sobolev space $H^s(x_0, \theta_0)$ if and only if $M_{1/\tilde{m}}(f)$ does. Consequently, f and $M_{1/\tilde{m}}(f)$ have the same H_s -wavefront set.*

Proof. Let $g := M_{1/\tilde{m}}(f)$. We want to show that g and $M_{\tilde{m}}g$ have the same H^s -wavefront set. The Fourier multiplier $M_{\tilde{m}}$ is a pseudo-differential operator with Weyl symbol $\sigma(x, \xi) := \tilde{m}(\xi)$. Under the hypothesis of Theorem 3.12, by (32), $0 < \tilde{A} \leq \tilde{m}(\xi) \leq \tilde{B} < +\infty$. This implies that $M_{\tilde{m}}$ is elliptic of order 0. We will show that σ belongs to Hörmander's symbol class $S_{\frac{1}{2}, 0}^0$, i.e.,

$$|\partial_\xi^\alpha \partial_x^\beta \sigma(x, \xi)| = |\partial_\xi^\alpha \tilde{m}(\xi)| \leq C_\alpha (1 + |\xi|)^{-\frac{|\alpha|}{2}}, \text{ for every multi-index } \alpha.$$

Once this is established, it will follow that $M_{\tilde{m}}$ preserves the H_s -wavefront set (see [26, Chapter 18] or [19, Chapter 2]). Let us write

$$\tilde{m}(\xi) := |\widehat{\varphi}_0(\xi)|^2 + \widetilde{m}_1(\xi),$$

where

$$\begin{aligned} \widetilde{m}_1(\xi) &:= \sum_{j \geq 1} \sum_{k=0}^{2^j-1} \Phi(B_{j,k}\xi), \\ \Phi(\xi) &:= \widehat{\varphi}(\xi) \overline{\widetilde{\psi}(\xi)}. \end{aligned}$$

Since $|\widehat{\varphi}_0(\xi)|^2$ is a Schwartz function and thus in $S_{\frac{1}{2}, 0}^0$ we focus on \widetilde{m}_1 . Note that, by the construction of ψ , Φ is a compactly supported smooth function that vanishes in some neighborhood of the origin $(-\varepsilon, \varepsilon)^2$.

Let α be a multi-index, $N = |\alpha|$ and

$$\Phi_N(\xi) := \sum_{|\beta| \leq N} |\partial_\xi^\beta \Phi(\xi)|.$$

Therefore

$$(58) \quad |\partial_\xi^\alpha \widetilde{m}_1(\xi)| \leq \sum_{j \geq 1} \sum_{k=0}^{2^j-1} 2^{-jN} \Phi_N(B_{j,k}\xi).$$

For each j, k , $|B_{j,k}\xi| \geq 4^{-j} |\xi|$ and consequently $|\xi|^{N/2} \leq 2^{jN} |B_{j,k}\xi|^{N/2}$. Hence, setting $\Phi_N^*(\xi) := |\xi|^{N/2} \Phi_N(\xi)$ we have that

$$|\xi|^{N/2} |\partial_\xi^\alpha \widetilde{m}_1(\xi)| \leq \sum_{j \geq 1} \sum_{k=0}^{2^j-1} \Phi_N^*(B_{j,k}\xi) \leq \|\Phi_N^*\|_*.$$

Since the support of Φ_N^* is compact and does not contain 0, Proposition 2.3 implies that $\|\Phi_N^*\|_* < +\infty$. Similarly,

$$|\partial_\xi^\alpha \widetilde{m}_1(\xi)| \leq \sum_{j \geq 1} \sum_{k=0}^{2^j-1} 2^{-jN} \Phi_N(B_{j,k}\xi) \leq \sum_{j \geq 1} \sum_{k=0}^{2^j-1} \Phi_N(B_{j,k}\xi) \leq \|\Phi_N\|_* < +\infty.$$

The last two estimates imply that $|\partial_\xi^\alpha \widetilde{m}_1(\xi)| \lesssim (1 + |\xi|)^{-\frac{N}{2}} = (1 + |\xi|)^{-\frac{|\alpha|}{2}}$, as desired. \square

REFERENCES

- [1] A. Aldroubi, C. Cabrelli, and U. Molter. Wavelets on Irregular Grids with Arbitrary Dilation Matrices, and Frame Atoms for $L_2(\mathbb{R}^d)$. *Appl. Comput. Harmon. Anal.*, Special Issue on Frames II.:119–140, 2004.
- [2] F. Andersson, M. Carlsson, and M. V. de Hoop. Nonlinear approximation of functions in two dimensions by sums of exponential functions. *Appl. Comput. Harmon. Anal.*, 29(2) 156–181, 2010.
- [3] F. Andersson, M. Carlsson, and M. V. de Hoop. Nonlinear approximation of functions in two dimensions by sums of wave packets. *Appl. Comput. Harmon. Anal.*, 29(2):198–213, 2010.
- [4] F. Andersson, M. Carlsson, and L. Tenorio. On the Representation of Functions with Gaussian Wave Packets. *J. Fourier Anal. Appl.*, 18:146–181, 2012.
- [5] F. Andersson, M. V. de Hoop, H. F. Smith, and G. Uhlmann. A multi-scale approach to hyperbolic evolution equations with limited smoothness. *Comm. Partial Differential Equations*, 33(4-6):988–1017, 2007.
- [6] F. Andersson, M. V. de Hoop, and H. Wendt. Multi-scale discrete approximation of Fourier Integral Operators., *Multiscale Model. Simul.*, 10(1):111–145, 2012.
- [7] P. Brault and J.-P. Antoine. A spatio-temporal Gaussian-Conical wavelet with high aperture selectivity for motion and speed analysis. *Appl. Comput. Harmon. Anal.*, 34(1):148–161, 2013.
- [8] C. Cabrelli, U. Molter, and J. L. Romero. Non-uniform painless decompositions for anisotropic Besov and Triebel-Lizorkin spaces. *Adv. Math.*, 232(1):98–120, 2013.
- [9] E. J. Candès and L. Demanet. The curvelet representation of wave propagators is optimally sparse. *Comm. Pure Appl. Math.*, 58(11):1472–1528, 2005.
- [10] E. J. Candès and D. L. Donoho. New tight frames of curvelets and optimal representations of objects with piecewise C_2 singularities. *Comm. Pure Appl. Math.*, 57(2):219–266., 2004.
- [11] E. J. Candès and D. L. Donoho. Continuous curvelet transform. I: Resolution of the wavefront set. *Appl. Comput. Harmon. Anal.*, 19(2):162–197, 2005.
- [12] E. J. Candès and D. L. Donoho. Continuous curvelet transform. II. Discretization and frames. *Appl. Comput. Harmon. Anal.*, 19(2):198–222, 2005.
- [13] A. Cordoba and C. Fefferman. Wave packets and Fourier integral operators. *Comm. Partial Differential Equations*, 3:979–1005, 1978.
- [14] S. Dahlke, G. Steidl, and G. Teschke. Shearlet coorbit spaces: Compactly supported analyzing shearlets, traces and embeddings. *J. Fourier Anal. Appl.*, 17(6):1232–1255, 2011.
- [15] I. Daubechies. The wavelet transform, time-frequency localization and signal analysis. *IEEE Trans. Inform. Theory*, 36(5):961–1005, 1990.
- [16] L. Demanet and L. Ying, Wave Atoms and Sparsity of Oscillatory Patterns. *Appl. Comput. Harmon. Anal.*, 23(3): 368–387, 2007.
- [17] A.A. Duchkov and M. V. de Hoop, Extended isochron rays in prestack depth (map) migration. *Geophysics.*, 75(4): S139-S150, 2010.
- [18] H. Douma and M. V. de Hoop, Leading-order seismic imaging using curvelets. *Geophysics.*, 72(6):S231–S248, 2007.
- [19] G. B. Folland. *Harmonic Analysis in Phase Space*. Princeton Univ. Press, Princeton, NJ, 1989.
- [20] P. Gérard. Moyennisation et régularité deux-microlocale. *Ann. Sci. École Norm. Sup. (4)*, 23(1):89–121, 1990.
- [21] P. Grohs. Continuous shearlet frames and resolution of the wavefront set. *Monatsh. Math.*, 164(4):393–426, 2011.
- [22] P. Grohs and G. Kutyniok. Parabolic Molecules. *Found. Comput. Math.*, 14(2):299–337, 2014.
- [23] K. Guo and D. Labate. Optimally sparse multidimensional representation using shearlets. *SIAM J. Math. Anal.*, 39(1):298–318, 2007.
- [24] K. Guo and D. Labate. Representation of Fourier integral operators using shearlets. *J. Fourier Anal. Appl.*, 14(3):327–371, 2008.

- [25] K. Guo and D. Labate. Characterization and analysis of edges using the continuous shearlet transform. *SIAM J. Imaging Sci.*, 2(3):959–986, 2009.
- [26] L. Hörmander. *The analysis of linear partial differential operators. III: Pseudo-differential operators*. Grundlehren der Mathematischen Wissenschaften, 274. Springer, 1985.
- [27] K. Johansson, S. Pilipović, N. Teofanov, and J. Toft. Gabor pairs, and a discrete approach to wavefront sets. *Monatsh. Math.*, 166(2):181–199, 2012.
- [28] P. Kittipoom, G. Kutyniok, and W.-Q. Lim. Construction of Compactly Supported Shearlet Frames. *Constr. Approx.*, 35:21–72, 2012.
- [29] G. Kutyniok and D. Labate. Resolution of the wavefront set using continuous shearlets. *Trans. Amer. Math. Soc.*, 361(5):2719–2754, 2009.
- [30] M. Nielsen and K. Rasmussen. Compactly Supported Frames for Decomposition Spaces. *J. Fourier Anal. Appl.*, 18(1):87–117, 2012.
- [31] M. Nielsen. Frames for decomposition spaces generated by a single function. *Collect. Math.*, 65(2):183–201, 2014.
- [32] G. Kutyniok and D. Labate, editors. *Shearlets*. Applied and Numerical Harmonic Analysis. Birkhäuser/Springer, New York, 2012. Multiscale analysis for multivariate data.
- [33] T.J.P.M. Op’t Root, C.C. Stolk and M.V. de Hoop, Linearized inverse scattering based on seismic Reverse-Time Migration. *J. Math. Pures Appl.*, 98(2): 211–238, 2012.
- [34] J. Qian and L. Ying. Fast Gaussian wavepacket transforms and Gaussian beams for the Schrödinger equation. *J. Comput. Phys.*, 229(20):7848–7873, 2010.
- [35] J. Qian and L. Ying. Fast multiscale Gaussian wavepacket transforms and multiscale Gaussian beams for the wave equation. *Multiscale Model. Simul.*, 8(5):1803–1837, 2010.
- [36] K. N. Rasmussen and M. Nielsen. Compactly Supported Curvelet-Type Systems. *J. Funct. Spaces Appl.*, 2012:1–18, Art. ID 876315, 2012. doi:10.1155/2012/876315.
- [37] A. Shlivinski and E. Heyman. Windowed Radon transform frames. *Appl. Comput. Harmon. Anal.*, 26(3):322–343, 2009.
- [38] H. F. Smith. A Hardy space for Fourier integral operators. *J. Geom. Anal.*, 8(4):629–653, 1998.
- [39] H. F. Smith. A parametrix construction for wave equations with $C^{1,1}$ coefficients. *Ann. Inst. Fourier (Grenoble)*, 48(3):797–835, 1998.
- [40] E. M. Stein. *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*. Princeton Univ. Press, Princeton, NJ, 1993. With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III.
- [41] L. Tenorio, F. Andersson, P. Ma and M. V. de Hoop, Data analysis tools for uncertainty quantification of inverse problems. *Inverse Problems*. 27, 2011. doi:10.1088/0266-5611/27/4/045001.

DEPARTMENT OF MATHEMATICS, PURDUE UNIVERSITY, 150 N. UNIVERSITY STREET, WEST LAFAYETTE IN 47907, USA.

E-mail address: mdehoop@purdue.edu

FACULTY OF MATHEMATICS, UNIVERSITY OF VIENNA, OSKAR-MORGENSTERN-PLATZ 1, A-1090 WIEN, AUSTRIA

E-mail address: karlheinz.groechenig@univie.ac.at

FACULTY OF MATHEMATICS, UNIVERSITY OF VIENNA, OSKAR-MORGENSTERN-PLATZ 1, A-1090 WIEN, AUSTRIA

E-mail address: jose.luis.romero@univie.ac.at