Proceedings of the Project Review, Geo-Mathematical Imaging Group (Purdue University, West Lafayette IN), Vol. 1 (2013) pp. 17-35.

EXACT AND APPROXIMATE EXPANSIONS WITH PURE GAUSSIAN WAVEPACKETS

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

Abstract. We consider wavepackets generated by dilation, rotation and translation of (a finite sum of) Gaussians following parabolic scaling, and give asymptotics on the analogue of Daubechies' frame criterion. We give the characterization of the wavefront set with this frame, also using an approximate dual frame.

1. Introduction. We consider wavepackets generated by dilation, rotation and translation of a function, $\varphi : \mathbb{R}^2 \to \mathbb{C}$, following parabolic scaling analogously to curvelets,

$$\varphi_{j,k,\lambda}(x) := 2^{3/2j}\varphi(A_{j,k}x - \lambda), \qquad (j \geq 1, 0 \leq k \leq 2^j - 1, \lambda \in \Lambda), \tag{1.1}$$

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 fast decaying, the packet $\varphi_{j,k,\lambda}$ is localized near $A_{j,k}^{-1}\lambda$ and aligned at an angle of $\frac{-2\pi k}{2^j}$ radians. Like other parabolic systems, such as curvelets and shearlets, these wavepackets are particularly useful in geophysics (see for example [7]).

In this paper, we consider expansions and approximate expansions of $L^2(\mathbb{R}^2)$ functions in terms of these packets using Gaussian windows. As with Morlet's wavelets, one can circumvent the problem of the lack of vanishing moments by considering linear combinations of modulated Gaussians. As a fundamental example we consider windows of the form,

$$\widehat{\varphi}(\xi_1, \xi_2) := \sum_{k=1}^{N} a_k \mathrm{e}^{-\delta_0|\xi_1 - t_k|^2 + \delta_k|\xi_2|^2}, \qquad \xi = (\xi_1, \xi_2) \in \mathbb{R}^2, \tag{1.2}$$

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

$$\widehat{\varphi}(0) = \partial_{\xi_1}\widehat{\varphi}(0) = \dots = \partial_{\xi_1}^N \widehat{\varphi}(0) = 0. \tag{1.3}$$

For such a choice of φ, the wavepackets in (1.1) 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 [2]). The use of wavepackets that consist of pure Gaussians has several technical advantages because the relevant computations can be done explicitly [1]. These packets are also important to represent initial or boundary data, to generate a multi-scale Gaussian-beam-like solution to the wave equation [14].

The formula in (1.2) allows for a number of rather different windows. One possibility is to let $\widehat{\varphi(\xi)}$ be essentially a translated Gaussian $\mathrm{e}^{-|\xi-(0,t_1)|^2}$, 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 so large as 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 (1.2) leads to an expansion in terms of Gaussians.

*Department of Mathematics, Purdue University, 150 N. University Street, West Lafayette IN 47907, USA., mdehoop@math.purdue.edu

†Faculty of Mathematics, University of Vienna, Nordbergstrasse 15,A-1090 Wien, Austria, karlheinz.groechenig@univie.ac.at

‡Faculty of Mathematics, University of Vienna, Nordbergstrasse 15,A-1090 Wien, Austria, jose.luis.romero@univie.ac.at

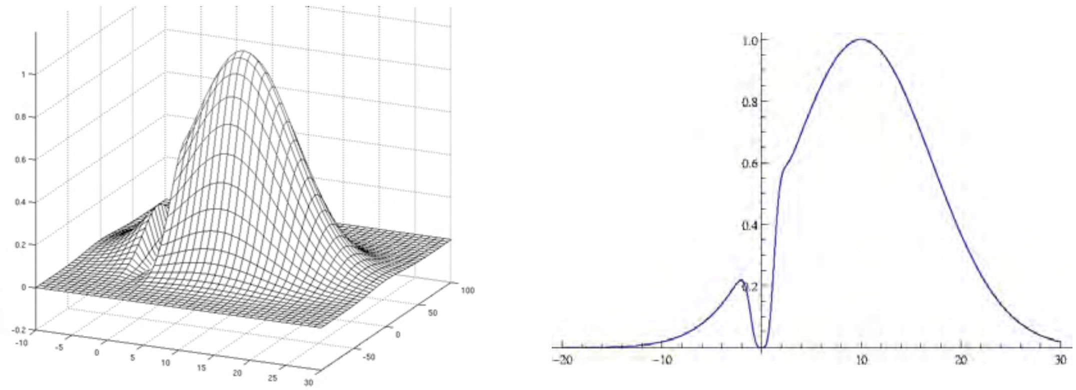

Fig. 1: A possible choice of $\widehat{\varphi}$ with vanishing moments of order 3. On the left, a plot of $\widehat{\varphi}$, and on the right a plot of its restriction to the ξ_1 axis.

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

In order to handle the coarse scales we use translates of a second window $\phi_\lambda := \phi(\cdot - \lambda)$. We typically let ϕ be a non-eccentric Gaussian centered at 0. The complete collection of wavepackets is then

$$\mathcal{C}_\Lambda := \left\{ \phi_\lambda \,\middle|\, \lambda \in \Lambda \right\} \cup \left\{ \varphi_{j,k,\lambda} \,\middle|\, \lambda \in \Lambda, j \geq 1, 0 \leq k \leq 2^j - 1 \right\}. \tag{1.4}$$

Here, we give asymptotics on the lattice parameters required for these packets to give a frame. This is done by examining the quantities that appear in (the analogue of) Daubechies' criterion for wavelets, and bounding the correlation between rotations and dilations of the window. A similar approach was recently carried out for the so-called cone-adapted shearlets in [12]. 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. This complements the results in [1] that treat the effective approximation of bounded compactly supported functions by this kind of packets, but left open the question of a full expansion in $L^2(\mathbb{R}^2)$.

Daubechies' criterion consists in comparing the frame operator to a certain Fourier multiplier. This provides an approximate dual frame, that can be easily computed by simply applying a frequency filter. The error introduced by such approximation depends on the density at which the translation parameter is sampled. In the applications to geophysics that motivated us, the data has a limited amount of numerical accuracy [7]. Therefore, having an explicit expansion in terms of inner products with dilation-rotation packets at the expense of numerical precision is a good trade-off. We consider such an approximate expansion and show that the coefficients decay fast in the scale parameter j, away from the wave-front set of the *original data*.

The paper is organized as follows. We first derive certain estimates on rotation-dilation overlaps. These are used to validate a Daubechies-like frame criterion. We then study the decay of the coefficients $\langle f, \varphi_{j,k,\lambda} \rangle$ as $j \to +\infty$ and the parameters $k = k_j, \lambda = \lambda_j$ vary with the scale so as to track a given phase-space point. We show that if that 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}. Finally we show that similar estimates hold for the coefficients in the approximate expansion related to Daubechies criterion.

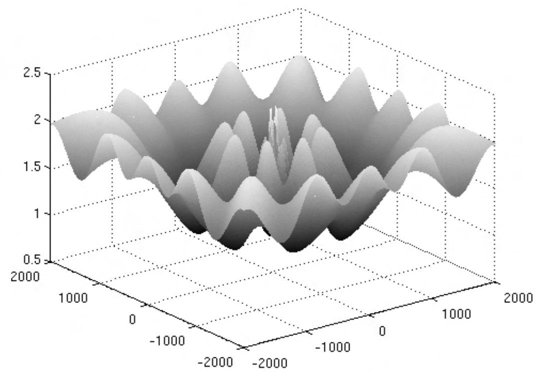

Fig. 2: A plot of the quantity in (1.5) for the function φ from Fig. 1 and $\phi(\xi) = e^{-\frac{|\xi|^2}{10000}}$, showing $A \approx 0.9503$ and $B \approx 2.2483$.

Throughout the analysis we make the following assumptions on the windows ϕ, φ.

$$0 < A \leq \left|\widehat{\phi}(\xi)\right|^2 + \sum_{j\geq 1}\sum_{k=0}^{2^j-1} |\widehat{\varphi}(B_{j,k}\xi)|^2 \leq B < +\infty, \text{ where } B_{j,k} := (A^*_{j,k})^{-1}, \tag{1.5}$$

$$\left|\widehat{\phi}(\xi)\right| \lesssim \mathrm{e}^{-\delta|\xi|^2}, \tag{1.6}$$

$$|\widehat{\varphi}(\xi)| \lesssim \min(1, |\xi_1|^{\varsigma})\mathrm{e}^{-\delta|\xi|^2}, \text{ for some } \delta > 0 \text{ and } \varsigma > 2. \tag{1.7}$$

The assumption in (1.5) means that the windows are wide enough so as to cover the whole frequency plain (see Figure 2). The Gaussian wavepackets in (1.2) clearly satisfy the decay conditions in (1.6) and (1.7) while the vanishing moments in the direction ξ_1 are granted by (1.3).

We also use the following notation. We let 1_E denote the characteristic function of a set $E \subseteq \mathbb{R}^2$. The Fourier transform is normalized as $\hat{f}(\xi) := \int_{\mathbb{R}^2} f(x)\mathrm{e}^{-2\pi \mathrm{i}\xi x}\, dx$. Translations and dilations of function $f : \mathbb{R}^2 \to \mathbb{C}$ are defined as

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

With this notation, $\varphi_{j,k,\lambda} = Dil_{A_{j,k}} T_\lambda \varphi$. Note also that with this notation $Dil_{A\cdot B} f = Dil_B Dil_A f$. We further denote the parabolic dilations and the rotations by

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

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. Here, we prove certain technical estimates on rotation-dilation overlaps that will be needed throughout the analysis. For $f : \mathbb{R}^2 \to \mathbb{C}$ let the *star-norm* of f be defined by

$$\|f\|_* := \operatorname*{ess\,sup}_{\xi\in\mathbb{R}^2} \sum_{j\geq 1}\sum_{k=0}^{2^j-1} |f(B_{j,k}\xi)|\,. \tag{2.1}$$

We define the circular sectors,

$$V_{s,t} := \left\{ (r\cos(\theta), r\sin(\theta)) : \begin{array}{c} 2^s \leq r \leq 2^{s+1}, 0 \leq \theta \leq \pi/2, \\ 2^{-(t+1)} \leq \cos(\theta) \leq 2^{-t}, \end{array} \right\}, \quad (s \in \mathbb{Z}, t \geq 0). \tag{2.2}$$

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 (2.2). 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.

$$D_j(r\cos(\theta), r\sin(\theta)) = (r'\cos(\theta'), r\sin(\theta')), \tag{2.3}$$

$$r' = \rho(\cos(\theta), j)r, \tag{2.4}$$

$$\tan(\theta') = 2^{-j}\tan(\theta), \tag{2.5}$$

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

$$\rho(\alpha, j)^2 = \alpha^2 4^{2j} + (1-\alpha^2)2^{2j}, \qquad (\alpha \in [0,1], j \in \mathbb{N}). \tag{2.6}$$

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

$$\rho(\alpha, j+1) \geq 2\rho(\alpha, j), \tag{2.7}$$

$$\rho(\alpha/2, j+2) \geq 2\rho(\alpha, j), \qquad (\alpha \in [0,1], j \in \mathbb{N}). \tag{2.8}$$

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

With this description, we estimate the image of $V_{s,t}$ by the parabolic dilation D_j. To this end, let us define for $r > 0$ and $j \in \mathbb{N}$,

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

We now note that

$$D_j V_{s,t} \subseteq W_{j,s,t}. \tag{2.9}$$

Indeed, if $\xi = (r\cos(\theta), r\sin(\theta)) \in V_{s,t}$ and $(r'\cos(\theta'), r'\sin(\theta')) = D_j\xi$, then by (2.3),

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

Since ξ belongs to the first quadrant, so does $D_j\xi$. Hence, $\theta' \in [0, \pi/2)$. In addition, using again (2.3) 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 (2.9) 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 $R_{\frac{2k\pi}{2^j}}W_{j,s,t}$, $k = 0, \ldots, 2^j - 1$ have $\lesssim 4^t$ overlaps. Hence, (2.9) allow us to estimate for each $j \in \mathbb{N}$ and $\xi \in \mathbb{R}^2$,

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

where,

$$C_{j,s,t} := \left\{ \xi \in \mathbb{R}^2 \,\middle|\, 2^s\rho(2^{-(t+1)}, j) \le |\xi| \le 2^{s+1}\rho(2^{-t}, j) \right\}. \tag{2.10}$$

Finally we note that these sets do not overlap too much as j varies. Indeed, if $C_{j,s,t} \cap C_{j',s,t} \ne \emptyset$ for $j \le j'$, then $2^s\rho(2^{-(t+1)}, j') \le 2^{s+1}\rho(2^{-t}, j)$. Hence, setting $\alpha := 2^{-t}$, we have that $\rho(\alpha/2, j') \le 2\rho(\alpha, j)$. By (2.8) this implies that $j' \le j + 2$. Hence,

$$\|1_{V_{s,t}}\|_* \lesssim 4^t \sup_{\xi\in\mathbb{R}^2} \sum_{j\ge1} 1_{C_{j,s,t}}(\xi) \lesssim 4^t.$$

∎ We have the following invariance property of the star norm.

Lemma 2.2. *Let $f : \mathbb{R}^2 \to \mathbb{C}$ and let $S : \mathbb{R}^2 \to \mathbb{R}^2$ 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,

$$\begin{aligned}
\|f \circ S\|_* &= \operatorname*{ess\,sup}_{\xi\in\mathbb{R}^2} \sum_{j\ge1} \sum_{k=0}^{2^j-1} \left| f(SD_j^{-1}R^{-1}_{\frac{2k\pi}{2^j}}\xi) \right| = \operatorname*{ess\,sup}_{\xi\in\mathbb{R}^2} \sum_{j\ge1} \sum_{k=0}^{2^j-1} \left| f(D_j^{-1}R^{-1}_{-\frac{2k\pi}{2^j}}S\xi) \right| \\
&= \operatorname*{ess\,sup}_{\xi\in\mathbb{R}^2} \sum_{j\ge1} \sum_{k=0}^{2^j-1} \left| f(D_j^{-1}R^{-1}_{-\frac{2k\pi}{2^j}}\xi) \right| = \|f\|_*,
\end{aligned}$$

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 1. *Let $f : \mathbb{R}^2 \to \mathbb{C}$ satisfy,*

$$|f(\xi)| \le |\xi_1|^{2+\varepsilon} \, \mathrm{e}^{-\tau|\xi|}, \tag{2.11}$$

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

$$\|f\|_* = \operatorname*{ess\,sup}_{\xi\in\mathbb{R}^2} \sum_{j\ge1} \sum_{k=0}^{2^j-1} |f(B_{j,k}\xi)| \lesssim 1.$$

Here, the implicit constant depends on τ and ε.

Remark 1. *As a consequence, it follows that $\|\widehat{\varphi}\|_* < +\infty$.*

Proof. [Proof of Prop. 1] Recall the definition of the sets $V_{s,t}$ in (2.2) and further define $V^i_{s,t} := S^iV_{s,t}$, $i = 0, 1, 2, 3$, where S^i are the four symmetries $(\xi_1, \xi_2) \mapsto (\pm\xi_1, \pm\xi_2)$. From Lemma 2.2, $\|1_{V^i_{s,t}}\|_* = \|1_{V_{s,t}}\|_*$ and hence, by Lemma 2.1, $\|1_{V^i_{s,t}}\|_* \lesssim 4^t$. Since the sets $\{V^i_{s,t} : s \in \mathbb{Z}, t \in \mathbb{N}, i = 0, 1, 2, 3\}$ 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,

$$|f| \le \sum_{i=0}^{3} \sum_{s\in\mathbb{Z}} \sum_{t\ge0} \sup_{V^i_{s,t}} |f| \, 1_{V^i_{s,t}},$$

and therefore,

$$\|f\|_* \lesssim \sum_{i=0}^{3} \sum_{s\in\mathbb{Z}} \sum_{t\ge0} 4^t \sup_{V^i_{s,t}} |f| \,.$$

Using (2.11) we see that,

$$\sup_{V^i_{s,t}} |f| = \sup_{V_{s,t}} |f| \lesssim (2^{s-t+1})^{2+\varepsilon} \mathrm{e}^{-\tau 2^s} = (1/r)^t r^{s+1} \mathrm{e}^{-\tau 2^s}, \tag{2.12}$$

where $r := 2^{2+\varepsilon}$. Since $r > 4$,

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

□

3. Frame expansions.

3.1. Representation of the frame operator.

Let S_Λ be the frame operator associated with $\mathcal{C}_\Lambda$,

$$S_\Lambda f := \sum_{\lambda \in \Lambda} \langle f, T_\lambda \phi \rangle T_\lambda \phi + \sum_{\lambda \in \Lambda} \sum_{j=1}^{+\infty} \sum_{k=0}^{2^j-1} \left\langle f, Dil_{A_{j,k}} T_\lambda \varphi \right\rangle Dil_{A_{j,k}} T_\lambda \varphi.$$

Poisson's summation formula,

$$|\Lambda| \sum_{\lambda \in \Lambda} f(x-\lambda) = \sum_{\gamma \in \Gamma} \hat{f}(\gamma) \mathrm{e}^{2\pi \mathrm{i} x \gamma}, \quad f \in \mathcal{S}(\mathbb{R}^2), \tag{3.1}$$

gives 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.

We denote by $\widehat{S_\Lambda}$ the operator given by $\widehat{S_\Lambda} \hat{f} := \widehat{(S_\Lambda f)}$. Using Poisson's summation formula, $\widehat{S_\Lambda}$ can be decomposed in the following way.

Lemma 3.1. *The operator $\widehat{S_\Lambda}$ can be written as*

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

where,

$$\widehat{S_{\Lambda,\gamma}} f(\xi) = f(\xi - \gamma) \widehat{\phi_0}(\xi) \overline{\widehat{\phi_0}(\xi - \gamma)} + \sum_{j=1}^{+\infty} \sum_{k=0}^{2^j-1} f(\xi - A^*_{j,k}\gamma) \widehat{\varphi}(B_{j,k}\xi) \overline{\widehat{\varphi}(B_{j,k}\xi - \gamma)}. \tag{3.2}$$

Proof. We consider the operators,

$$\begin{aligned} S_0 f &:= \sum_{\lambda \in \Lambda} \langle f, T_\lambda \phi \rangle T_\lambda \phi, \\ S_* f &:= \sum_{\lambda \in \Lambda} \langle f, T_\lambda \varphi \rangle T_\lambda \varphi, \\ S_{j,k} f &:= \sum_{\lambda \in \Lambda} \left\langle f, Dil_{A_{j,k}} T_\lambda \varphi \right\rangle Dil_{A_{j,k}} T_\lambda \varphi, \quad (j \geq 1, k = 0, \ldots, 2^j - 1). \end{aligned}$$

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

Let $\mathcal{F}$ denote the Fourier transform; then

$$\begin{aligned} \mathcal{F} S_0 \mathcal{F}^{-1} f(\xi) &= \sum_{\lambda \in \Lambda} \left\langle \mathcal{F}^{-1} f, T_\lambda \phi \right\rangle \mathrm{e}^{2\pi \mathrm{i} \lambda \xi} \widehat{\phi_0}(\xi) \\ &= \sum_{\lambda \in \Lambda} \left\langle f, \widehat{\phi_0} \mathrm{e}^{2\pi \mathrm{i} \lambda \cdot} \right\rangle \mathrm{e}^{2\pi \mathrm{i} \lambda \xi} \widehat{\phi_0}(\xi) = \sum_{\lambda \in \Lambda} \mathcal{F}(f \overline{\widehat{\phi_0}})(\lambda) \mathrm{e}^{2\pi \mathrm{i} \lambda \xi} \widehat{\phi_0}(\xi). \end{aligned}$$

Using (3.1), this implies that,

$$\mathcal{F}S_0\mathcal{F}^{-1}f(\xi) = |\Lambda|^{-1}\sum_{\gamma\in\Gamma} f(\xi-\gamma)\overline{\widehat{\phi_0}(\xi-\gamma)}\widehat{\phi_0}(\xi). \tag{3.3}$$

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 $\mathcal{F}S_{j,k}\mathcal{F}^{-1} = Dil^{-1}_{A^*_{j,k}}\mathcal{F}S_*\mathcal{F}^{-1}Dil_{A^*_{j,k}}$ this yields,

$$\mathcal{F}S_{j,k}\mathcal{F}^{-1}f(\xi) = |\Lambda|^{-1}\sum_{\gamma\in\Gamma} f(\xi-A^*_{j,k}\gamma)\overline{\widehat{\varphi}(B_{j,k}\xi-\gamma)}\widehat{\varphi}(B_{j,k}\xi). \tag{3.4}$$

Since $S_\Lambda = S_0+\sum_{k,j}S_{k,j}$, the conclusion follows from (3.3) and (3.4). □ Following the representation in Lemma 3.1, we define the time-scale correlation function,

$$\Theta(\zeta) := \operatorname*{ess\,sup}_{\xi\in\mathbb{R}^2}\left(\left|\widehat{\phi_0}(\xi)\right|\left|\widehat{\phi_0}(\xi-\zeta)\right| + \sum_{j=1}^{+\infty}\sum_{k=0}^{2^j-1}|\widehat{\varphi}(B_{j,k}\xi)|\,|\widehat{\varphi}(B_{j,k}\xi-\zeta)|\right), \quad (\zeta\in\mathbb{R}^2). \tag{3.5}$$

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

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

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

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

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

THEOREM 3.3. *Let*

$$\Delta(\Lambda) := \sum_{\gamma\in\Gamma\setminus\{0\}}\max\{\Theta(\gamma),\Theta(-\gamma)\}. \tag{3.6}$$

If $\Delta(\Lambda) < A$, then $\mathcal{C}_\Lambda$ 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

$$\widehat{S_{\Lambda,0}}f(\xi) = f(\xi)\left(\left|\widehat{\phi_0}(\xi)\right|^2 + \sum_{j=1}^{+\infty}\sum_{k=0}^{2^j-1}|\widehat{\varphi}(B_{j,k}\xi)|^2\right).$$

Hence, by the covering condition in (1.5) we have that, for $f\in L^2(\mathbb{R}^2)$,

$$A\,|f(\xi)| \le \left|\widehat{S_{\Lambda,0}}f(\xi)\right| \le B\,|f(\xi)|,$$

and therefore,

$$A\|f\|_2 \le \|\widehat{S_{\Lambda,0}}f\|_2 \le B\|f\|_2. \tag{3.7}$$

Using Lemma 3.2 and (3.8) we get

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

This together with (3.7) implies that for all $f\in L^2(\mathbb{R}^2)$

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

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

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

as desired. □

Remark 2. *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 on the sampling density and the error. Using the rotation-dilation estimates from Sec. 2 we can give asymptotics on the frame criterion of Theorem 3.3. Let the lattice Λ be given by $\Lambda=P\mathbb{Z}^2$, with P having singular values $a,b\ge 0$. This means that Λ is an 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 3.4. *The following estimate*

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

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

Proof. Let $\xi,\zeta\in\mathbb{R}^2$. We use repeatedly the quasi triangle inequality: $|\xi+\zeta|^2\le 2(|\xi|^2+|\zeta|^2)$ and use τ to denote new (smaller) decay parameters that can very from line to line. By (1.6),

$$|\phi(\xi)|\,|\phi(\xi-\zeta)|\lesssim \mathrm{e}^{-\delta(|\xi|^2+|\xi-\zeta|^2)}\le \mathrm{e}^{-\tau|\zeta|^2}.$$

Likewise, by (1.7),

$$|\varphi(\xi)|\,|\varphi(\xi-\zeta)|\lesssim|\xi_1|^t\,\mathrm{e}^{-\delta(|\xi|^2+|\xi-\zeta|^2)}.$$

Since $|\xi|^2+|\xi-\zeta|^2$ is $\gtrsim$ than both $|\xi|^2$ and $|\zeta|^2$, we have that $|\xi|^2+|\xi-\zeta|^2\gtrsim|\xi|^2+|\zeta|^2$. Hence,

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

if we let $\Phi(\xi):=|\xi_1|^t\,\mathrm{e}^{-\tau|\xi|^2}$. This implies that the time-scale correlation function in (3.5) satisfies,

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

Using Prop. 1 we conclude that,

$$\Theta(\zeta) \lesssim \mathrm{e}^{-\tau|\zeta|^2}. \tag{3.8}$$

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

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

Summing the geometric series gives,

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

□ As an immediate consequence of Theorem 3.3 and Lemma 3.4 we have the following

THEOREM 3.5. *There are constants $C,\tau > 0$ depending on δ and ε such that $\mathcal{C}_\Lambda$ is a frame whenever,*

$$\kappa := \mathrm{e}^{-\tau/a^2} + \mathrm{e}^{-\tau/b^2} \leq CA. \tag{3.9}$$

Moreover, in this case, $A - C^{-1}\kappa$ and $B + C^{-1}\kappa$ are frame bounds for $\mathcal{C}_\Lambda$. In particular, for every lattice Λ with sufficiently large density (i.e. $a, b \ll 1$), the system $\mathcal{C}_\Lambda$ is a frame of $L^2(\mathbb{R}^2)$, giving the expansions

$$f = \sum_{\lambda\in\Lambda} \left\langle f, S_\Lambda^{-1}\phi_\lambda \right\rangle \phi_\lambda + \sum_{\lambda\in\Lambda}\sum_{j=1}^{+\infty}\sum_{k=0}^{2^j-1} \left\langle f, S_\Lambda^{-1}\varphi_{j,k,\lambda} \right\rangle \varphi_{j,k,\lambda} \tag{3.10}$$

or

$$f = \sum_{\lambda\in\Lambda} \left\langle f, \phi_\lambda \right\rangle S_\Lambda^{-1}\phi_\lambda + \sum_{\lambda\in\Lambda}\sum_{j=1}^{+\infty}\sum_{k=0}^{2^j-1} \left\langle f, \varphi_{j,k,\lambda} \right\rangle S_\Lambda^{-1}\varphi_{j,k,\lambda}. \tag{3.11}$$

As the density of Λ tends to infinity, the ratio of the frame bounds tends to A/B.

REMARK 3. *If we let the window φ be given by* (1.2)*, then* (3.10) *provides a formal argument for the representability of L^2 functions in terms of the Gaussian wavepackets discussed in [1], which can be then computed with the algorithms discussed there.*

REMARK 4. *The system $\mathcal{C}_\Lambda$ is a family of* curvelet molecules *in the sense of [5] and hence the expansion in* (3.11) *shares the same asymptotic sparse approximation that curvelets do [3].*

REMARK 5. *As the proof of Prop. 1 shows, the fact that every sufficiently dense lattice yields a frame follows under conditions much milder that those in* (1.6), (1.7)*, for example polynomial decay.*

REMARK 6. *This result is analogous to the one recently obtained for shearlets in [12]. The factor $|\xi_1|^{\varsigma=2+\varepsilon}$ in* (1.7) *should be compared with the factor $|\xi_1|^{3+\varepsilon}$ in [12].*

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 this approximation to construct an approximate dual frame. This is particularly convenient because $S_{\Lambda,0}$ is simply the Fourier multiplier with symbol

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

(cf. Lemma 3.1). Hence, the approximate dual system is just the image of $\mathcal{C}_\Lambda$ by the Fourier multiplier with symbol $1/m$. It then readily follows from the proof of Theorem 4.1 and the estimates of Sec. 3.3 that this gives a reconstruction error of at most $\Delta(\Lambda)/A \lesssim \mathrm{e}^{-\tau/a^2} + \mathrm{e}^{-\tau/b^2}$, with a, b the lattice parameters.

By slightly relaxing this bound one can gain certain flexibility 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 cross frame operator

$$\widetilde{S}_\Lambda f := \sum_{\lambda\in\Lambda} \langle f, T_\lambda \phi\rangle\, T_\lambda\phi + \sum_{\lambda\in\Lambda}\sum_{j=1}^{+\infty}\sum_{k=0}^{2^j-1} \left\langle f, Dil_{A_{j,k}} T_\lambda \psi\right\rangle Dil_{A_{j,k}} T_\lambda \varphi.$$

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

$$\widetilde{m}(\xi) := \left|\widehat{\phi}(\xi)\right|^2 + \sum_{j\geq 1}\sum_{k=0}^{2^j-1} \widehat{\varphi}(B_{j,k}\xi))\overline{\widehat{\psi}(B_{j,k}\xi))}, \tag{3.12}$$

and upper bounds for

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

where

$$\widetilde{\Theta}(\zeta) := \operatorname*{ess\,sup}_{\xi\in\mathbb{R}^2} \left(\left|\widehat{\phi_0}(\xi)\right| \left|\widehat{\phi_0}(\xi-\zeta)\right| + \sum_{j=1}^{+\infty}\sum_{k=0}^{2^j-1} \widehat{\varphi}(B_{j,k}\xi)\overline{\widehat{\psi}(B_{j,k}\xi-\zeta)} \right).$$

In our case it makes sense let ψ be a small modification of ϕ that enforces some extra frequency cancellation. Let $\varepsilon > 0$. According to our assumptions, the set $E_\varepsilon := \left\{ \xi \in \mathbb{R}^2 \;\middle|\; |\widehat{\varphi}(\xi)| > \varepsilon \right\}$ is an open bounded set at positive distance from the origin. Let $\eta : \mathbb{R}^2 \to [0,1]$ 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 3.6. *The following estimates hold true,*

$$|m(\xi) - \widetilde{m}(\xi)| \leq \|\widehat{\varphi}\|_* \varepsilon \approx \varepsilon,$$
$$\left|\Delta(\Lambda) - \widetilde{\Delta}(\Lambda)\right| \leq \|\widehat{\varphi}\|_* \varepsilon \approx \varepsilon.$$

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

$$\begin{aligned}|m(\xi) - \widetilde{m}(\xi)| &= \left|\sum_{j\ge1}\sum_{k=0}^{2^j-1} \widehat{\varphi}(B_{j,k}\xi))(1-\eta(B_{j,k}\xi))\overline{\widehat{\varphi}(B_{j,k}\xi)}\right| \\ &\le \varepsilon \sum_{j\ge1}\sum_{k=0}^{2^j-1} |\widehat{\varphi}(B_{j,k}\xi)| = \varepsilon\|\widehat{\varphi}\|_*.\end{aligned}$$

The bound for $\left|\Delta(\Lambda) - \widetilde{\Delta}(\Lambda)\right|$ follows similarly. □ As a consequence of these estimates, if $\varepsilon\|\widehat{\varphi}\|_* < A$, then

$$0 < \widetilde{A} := A - \varepsilon\|\widehat{\varphi}\|_* \le \widetilde{m}(\xi) \le \widetilde{B} := B + \varepsilon\|\widehat{\varphi}\|_* < +\infty, \tag{3.13}$$

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

$$\begin{aligned}&\mathcal{C}'_\Lambda := \left\{\, \widetilde{\phi}_\lambda \,\middle|\, \lambda\in\Lambda \,\right\} \cup \left\{\, \widetilde{\varphi}_{j,k,\lambda} \,\middle|\, j\ge1, 1\le k\le 2^j-1, \lambda\in\Lambda \,\right\}, \\ &\widehat{\widetilde{\phi}_\lambda}(\xi) := \frac{1}{\widetilde{m}(\xi)}\widehat{T_\lambda\phi}(\xi), \\ &\widehat{\widetilde{\varphi}_{j,k,\lambda}}(\xi) := \frac{1}{\widetilde{m}(\xi)}\widehat{Dil_{A_{j,k}}T_\lambda}\psi(\xi).\end{aligned}$$

This provides the following approximate reconstruction.

THEOREM 3.7. *Let $\varepsilon\|\widehat{\varphi}\|_* < A$ and $\Delta(\Lambda) < A - 2\varepsilon\|\varphi\|_*$. Then every $f \in L^2(\mathbb{R}^2)$ admits the following approximate expansion.*

$$\widetilde{f} := |\Lambda| \left(\sum_{\lambda\in\Lambda}\left\langle f, \widetilde{\phi}_\lambda\right\rangle \phi_\lambda + \sum_{j\ge1}\sum_{k=0}^{2^j-1}\sum_{\lambda\in\Lambda}\left\langle f, \widetilde{\varphi}_{j,k,\lambda}\right\rangle \varphi_{j,k,\lambda}\right), \tag{3.14}$$

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

REMARK 7. *Note that the coefficients in* (3.14) *are given by*

$$\left\langle f, \widetilde{\varphi}_{j,k,\lambda}\right\rangle = \left\langle M_{\widetilde{m}} f, \psi_{j,k,\lambda}\right\rangle,$$

where $M_{\widetilde{m}}$ is the Fourier multiplier with symbol $\widetilde{m}$ and the packets $\psi_{j,k,\lambda} = Dil_{A_{j,k}}T_\lambda\psi$ have rotation-dilation structure. That is why in practice we prefer the expansion in (3.14) *to the "abstract" expansion in* (3.10).

4. Characterization of 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 ([4, 11, 9]). 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 [10] for a similar problem related to Gabor expansions).

4.1. Wavefront set with wavepackets. Let $(x_0, \theta_0) \in \mathbb{R}^2 \times [0, 2\pi)$. For each scale $j \ge 1$, we select a wave-packet of scale j that is localized near x_0 and approximately aligned to θ_0. The wave-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. We choose the parameters $k_j = k_j(x_0,\theta_0) \in \left\{0,\dots,2^j-1\right\}$ 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,

$$\frac{2\pi k_j}{2^j} \le 2\pi - \theta_0 \le \frac{2\pi(k_j+1)}{2^j}, \tag{4.1}$$

$$\left|A_{j,k_j}x_0 - \lambda_j\right| \le L_\Lambda, \tag{4.2}$$

where L_Λ is a constant that depends on the lattice Λ (namely the reciprocal of its density). 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 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 $j \to +\infty$.

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 a smooth compactly-supported function η and $\varepsilon > 0$ such that $\eta \equiv 1$ near x_0 and

$$\int_{R_\theta V_\varepsilon} \left| \widehat{\eta f}(\xi) \right|^2 |\xi|^{2s} \, d\xi < +\infty, \tag{4.3}$$

where

$$V_\varepsilon := \left\{ \xi \in \mathbb{R}^2 \,\middle|\, |\xi_2| \le \varepsilon |\xi_1| \right\}. \tag{4.4}$$

Since we work with windows with a limited number of vanishing moments, we further consider the space $H^s_M(x_0, \theta_0)$ consisting of distributions in $H^s(x_0, \theta_0)$ that satisfy, in addition to (4.3),

$$\int_{\mathbb{R}^2} \left| \widehat{\eta f}(\xi) \right|^2 (1 + |\xi|)^{-2M} \, d\xi < +\infty. \tag{4.5}$$

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

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 we assume that $\varphi \in \mathcal{S}$.

THEOREM 4.1. *Let $s, M \ge 0$. Assume that $\varphi \in \mathcal{S}(\mathbb{R}^2)$ and that $\varsigma > 2s + 2M - 1/2$. Let $(x_0, \theta_0) \in \mathbb{R}^2 \times [0, 2\pi)$, define $\{(k_j, \lambda_j) : j \ge 1\}$ by (4.1) and (4.2) and let $f \in H^s_M(x_0, \theta_0)$, then*

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

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. The following technical lemma will be used to that end.

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

$$W_{r,t} := \left\{ \xi \in \mathbb{R}^2 \,\middle|\, 4^{r-1} \le |\xi_1| \le 4^r, 2^{t-1} \le |\xi_2| \le 2^t \right\}. \tag{4.6}$$

For each $j \ge 1$, let $\theta_j \in \mathbb{R}$ be such that $|\theta_j| \le 2^{-j} 2\pi$. Then, each of the families of sets $\left\{ R_{\theta_j} D_j W_{r,t} : j \ge 1 \right\}$ has a number of overlaps that is bounded independently of r, t; that is,

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

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

It follows that there exists some $\xi \in (R_{\theta_j - \theta_{j'}} D_j W_{r,t}) \cap (D_h D_j W_{r,t})$. Because $\xi \in D_h D_j W_{r,t}$, we have

$$|\xi_1| \approx 4^{h+j+r} = 2^{2h+2j+2r}, \tag{4.7}$$

$$|\xi_2| \approx 2^{h+j+t}. \tag{4.8}$$

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

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

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 $|\cos(\theta_j - \theta_{j'})| \leq 1$ and $|\sin(\theta_j - \theta_{j'})| \lesssim 2^{-j}$. Hence, we can estimate

$$|\xi_1| \lesssim 2^{2j+2r} + 2^{-j}2^{j+t} = 2^{2j+2r} + 2^t, \tag{4.9}$$

$$|\xi_2| \lesssim 2^{-j}2^{2j+2r} + 2^{j+t} = 2^{j+2r} + 2^{j+t}. \tag{4.10}$$

Comparing the last equations with (4.7) and (4.8) we get

$$2^{2h+2j+2r} \lesssim 2^{2j+2r} + 2^t,$$
$$2^{h+j+t} \lesssim 2^{j+2r} + 2^{j+t}.$$

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

$$2h + 2j + 2r \leq \max\{2j + 2r + C, t + C\},$$
$$h + j + t \leq \max\{j + 2r + C, j + t + C\}.$$

If we further assume, as we may, that $h > C$ these estimates reduce to

$$2h + 2j + 2r \leq t + C, \tag{4.11}$$

$$h + t \leq 2r + C. \tag{4.12}$$

Since $h, j \geq 0$ we get from (4.11) that $2r \leq 2h + 2j + 2r \leq t + C$. Plugging this into (4.12) gives $h + t \leq t + 2C$. Hence $h \leq 2C$. This completes the proof. □

Proof. [Proof of Theorem 4.1] For convenience, we define $\theta_j := \frac{2\pi k_j}{2^j}$ and $\varphi_j(x) := Dil_{D_j}\varphi(x) = 2^{3/2j}\varphi(4^j x_1, 2^j x_2)$. Hence $\varphi_{j,k_j,\lambda_j} = Dil_{R_{\theta_j}} Dil_{D_j} T_{\lambda_j}\varphi = Dil_{R_{\theta_j}} T_{D_j^{-1}\lambda_j}\varphi_j$. We now further define,

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

Consequently,

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

Since $f \in H^s_M(x_0, \theta_0)$ 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 (4.3) and (4.5) hold. We further consider the functions,

$$g_j^1 := T_{-D_j^{-1}\lambda_j} Dil_{R_{-\theta_j}}(\eta f), \quad (j \geq 1),$$
$$g_j^2 := T_{-D_j^{-1}\lambda_j} Dil_{R_{-\theta_j}}((1-\eta) f), \quad (j \geq 1).$$

Hence,

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

The remainder of the proof consists in showing that the last two terms have the desired decay as $j \to +\infty$. This is done in two steps.

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

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

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

Using (4.2) we estimate,

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

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 $j \gg 0$,

$$g_j^2 \equiv 0 \text{ on } B_{\gamma/2}(0). \tag{4.13}$$

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 $j \geq 1$,

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

where the last two bounds follow from the fact that $\left|R_{\theta_j} x\right| = |x|$ and that, according to (4.2),

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

Taking (4.13) into account, the bound on $\left|\left\langle g_j^2, h\right\rangle\right|$ can be improved for $j \gg 0$ to

$$\left|\left\langle g_j^2, h\right\rangle\right| \lesssim \sum_{|\alpha|\leq N} \|(1+|\cdot|)^N \partial^\alpha h\|_{L^\infty(B_{\gamma'}(0))}, \tag{4.14}$$

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

We finally use this to bound $\left\langle g_j^2, \varphi_j\right\rangle$. Since $\varphi \in \mathcal{S}$ we have that for all $L > 2N$ 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 $|x| \geq \gamma'$,

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

Substituting this into (4.14) shows that $\left|\left\langle g_j^2, \varphi_j\right\rangle\right| \lesssim 2^{-Lj}$, for all $L > 0$. This clearly implies the desired estimate.

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

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

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

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

$$\left|\widehat{g_j^1(\xi)}\right| = \Phi(R_{\alpha_j}\xi), \quad (\xi \in \mathbb{R}^2). \tag{4.15}$$

Note that from (4.1) we get

$$0 \le \alpha_j \le 2\pi 2^{-j}, \qquad (j \ge 1). \tag{4.16}$$

Since f satisfies (4.5), it follows that Φ satisfies,

$$\int_{\mathbb{R}^2} \Phi(\xi)^2 (1 + |\xi|)^{-2M}\, d\xi < +\infty, \tag{4.17}$$

for some $M > 0$. In addition, by (4.3), we have

$$\int_{V_\varepsilon} \Phi(\xi)^2 \left|\xi\right|^{2s}\, d\xi < +\infty, \tag{4.18}$$

where $\varepsilon > 0$ and V_ε is the cone from (4.4). Let $\varepsilon' := 2^{-k}$ with $k \in \mathbb{N}$ such that $\varepsilon' < \varepsilon$. Since $\alpha_j \to 0$, there exist $j_0 \ge 1$ such that for all $j \ge j_0$,

$$R_{-\alpha_j}(V_{\varepsilon'}) \subseteq V_\varepsilon. \tag{4.19}$$

Consider the sets $\{W_{r,t} : r, t \in \mathbb{Z}\}$ from (4.6). Since these sets partition $\mathbb{R}^2$ up to sets of null measure, we can decompose $\widehat{\varphi}$ (mod null-measure) as

$$\widehat{\varphi} = \sum_{r,t\in\mathbb{Z}} \widehat{\varphi^{r,t}},$$
$$\widehat{\varphi^{r,t}} := \widehat{\varphi} \cdot 1_{W_{r,t}}.$$

We define $K_{r,t} \ge 0$ by

$$K_{r,t}^2 := \int_{\mathbb{R}^2} \left|\widehat{\varphi^{r,t}}(\xi)\right|^2\, d\xi = \int_{W_{r,t}} \left|\widehat{\varphi}(\xi)\right|^2\, d\xi. \tag{4.20}$$

From (1.7) it follows that

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

Hence for $\rho_1, \rho_2 \in \mathbb{Z}$,

$$\sum_{r,t\in\mathbb{Z}} 4^{-\rho_1 r} 2^{\rho_2 t} K_{r,t}^2 < +\infty, \text{ if } \rho_1 < (2\varsigma + 1) \text{ and } \rho_2 > -1. \tag{4.21}$$

We also use the notation $\widehat{\varphi_j^{r,t}}(\xi) := 2^{-3/2j}\widehat{\varphi^{r,t}}(4^{-j}\xi_1, 2^{-j}\xi_2)$, so for each $j \ge 1$

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

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

Since $\widehat{\varphi^{r,t}}$ is supported on $W_{r,t}$ we have that $\widehat{\varphi_j^{r,t}}$ is supported on $D_j W_{r,t}$. This set is contained in the cone $V_{\varepsilon'=2^{-k}}$, if $j \ge k + t - 2r$. We denote $j_1 = j_1(r,t) := \max\{j_0, k + t - 2r\}$. For $j \ge j_1(r,t)$

we bound,

$$
\begin{aligned}
\left|\langle g_j^1, \varphi_j^{r,t}\rangle\right| &\le 2^{-3/2j} \int_{D_j W_{r,t}} \left|\widehat{g_j^1}(\xi)\right| \left|\widehat{\varphi^{r,t}}(4^{-j}\xi_1, 2^{-j}\xi_2)\right| \, d\xi \\
&\le 2^{-3/2j} 4^{-js} \int_{D_j W_{r,t}} \left|\widehat{g_j^1}(\xi)\right| \left|\xi_1\right|^s \left|4^{-j}\xi_1\right|^{-s} \left|\widehat{\varphi^{r,t}}(4^{-j}\xi_1, 2^{-j}\xi_2)\right| \, d\xi \\
&\le 4^{-js} \left(\int_{D_j W_{r,t}} \left|\widehat{g_j^1}(\xi)\right|^2 \left|\xi_1\right|^{2s} \, d\xi\right)^{1/2} 2^{-3/2j} \left(\int_{D_j W_{r,t}} \left|4^{-j}\xi_1\right|^{-2s} \left|\widehat{\varphi^{r,t}}(4^{-j}\xi_1, 2^{-j}\xi_2)\right|^2 \, d\xi\right)^{1/2} \\
&\le 4^{-js} \left(\int_{D_j W_{r,t}} \left|\widehat{g_j^1}(\xi)\right|^2 \left|\xi_1\right|^{2s} \, d\xi\right)^{1/2} \left(\int_{W_{r,t}} \left|\xi_1\right|^{-2s} \left|\widehat{\varphi^{r,t}}(\xi)\right|^2 \, d\xi\right)^{1/2} \\
&\lesssim 4^{-js} 4^{-rs} \left(\int_{D_j W_{r,t}} \left|\widehat{g_j^1}(\xi)\right|^2 \left|\xi_1\right|^{2s} \, d\xi\right)^{1/2} \left(\int_{W_{r,t}} \left|\widehat{\varphi^{r,t}}(\xi)\right|^2 \, d\xi\right)^{1/2}.
\end{aligned}
$$

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

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

Using (4.15) we get

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

For $j \ge j_1(r,t)$, we have that $D_j W_{r,t} \subseteq V_{\varepsilon'}$ so (4.19) implies that $R_{-\alpha_j}(D_j W_{r,t}) \subseteq V_\varepsilon$. Combining this with (4.18) and the bound on the number of overlaps of the sets $\left\{R_{-\alpha_j}(D_j W_{r,t}) : j \ge 1\right\}$ granted by Lemma (4.2), we get

$$
\sum_{j \ge j_1(r,t)} \left|\langle g_j^1, \varphi_j^{r,t}\rangle\right|^2 4^{2sj} \lesssim 4^{-2rs} K_{r,t}^2. \tag{4.22}
$$

Since $s < \varsigma + 1/2$, by (4.21), this implies that

$$
\sum_{r,t \in \mathbb{Z}} \sum_{j \ge j_1(r,t)} \left|\langle g_j^1, \varphi_j^{r,t}\rangle\right|^2 4^{2sj} < +\infty. \tag{4.23}
$$

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

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

and do a similar estimate.

$$\begin{aligned}
\left|\langle g_j^1, \varphi_j^{r,t}\rangle\right| &\leq 2^{-3/2j} \int_{D_j W_{r,t}} \left|\widehat{g_j^1}(\xi)\right| \left|\widehat{\varphi^{r,t}}(D_{-j}\xi)\right| \, d\xi \\
&\leq 2^{-3/2j} 4^{jM} \int_{D_j W_{r,t}} \left|\widehat{g_j^1}(\xi)\right| (1+|\xi|)^{-M} (1+|D_{-j}\xi|)^M \left|\widehat{\varphi^{r,t}}(D_{-j}\xi)\right| \, d\xi \\
&\leq 4^{jM} \left(\int_{D_j W_{r,t}} \left|\widehat{g_j^1}(\xi)\right|^2 (1+|\xi|)^{-2M} \, d\xi\right)^{1/2} \\
&\qquad \times 2^{-3/2j} \left(\int_{D_j W_{r,t}} (1+|D_{-j}\xi|)^{2M} \left|\widehat{\varphi^{r,t}}(D_{-j}\xi)\right|^2 \, d\xi\right)^{1/2} \\
&\leq 4^{jM} \left(\int_{D_j W_{r,t}} \left|\widehat{g_j^1}(\xi)\right|^2 (1+|\xi|)^{-2M} \, d\xi\right)^{1/2} \left(\int_{W_{r,t}} (1+|\xi|)^{2M} \left|\widehat{\varphi^{r,t}}(\xi)\right|^2 \, d\xi\right)^{1/2} \\
&\lesssim 4^{jM} (1+4^r+2^t)^M \left(\int_{D_j W_{r,t}} \left|\widehat{g_j^1}(\xi)\right|^2 (1+|\xi|)^{-2M} \, d\xi\right)^{1/2} \left(\int_{W_{r,t}} \left|\widehat{\varphi^{r,t}}(\xi)\right|^2 \, d\xi\right)^{1/2}.
\end{aligned}$$

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

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

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

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

Using again the bound on the number of overlaps of $\left\{R_{-\alpha_j}(D_j W_{r,t}) : j \geq 1\right\}$, this time combined with (4.17), we get

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

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

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

Using (4.21) we get

$$\sum_{r,t\in\mathbb{Z}} \sum_{j=j_0}^{j_1(r,t)-1} \left|\langle g_j^1, \varphi_j^{r,t}\rangle\right|^2 4^{2sj} < +\infty, \tag{4.24}$$

since $M, s \geq 0$ and $4(M+S) < 2\sigma+1$.

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

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

□

4.2. Estimates for the approximate reconstruction. Theorem 4.1 shows that the coefficients in (3.11) have fast decay away from the wavefront set of f. We now show that the same is true for the ones in (3.14).

THEOREM 4.3. *In the setting of Theorem 3.7, let* $(x_0, \theta_0) \in \mathbb{R}^2 \times [0, 2\pi)$*, define* $\{(k_j, \lambda_j) : j \geq 1\}$ *by* (4.1) *and* (4.2) *and let* $f \in L^2(\mathbb{R}^2)$ *belong to* H^s *microlocally at* (x_0, θ_0)*, then*

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

As mentioned in Remark 7 the approximate coefficients can be rewritten as

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

where $\psi_{j,k,\lambda} = Dil_{A_{j,k}} T_\lambda \psi$ and $\widetilde{m}$ is given by (3.12) and $\widehat{M_{\widetilde{m}}}(f) := \widetilde{m}\hat{f}$. Since ψ also satisfies the same conditions that φ, it follows from Theorem 4.1 that the numbers $\langle f, \widetilde{\varphi}_{j,k,\lambda} \rangle$ decay away from the H_s-wavefront set of $M_{\widetilde{m}}(f)$. Hence, it suffices to show that f and $M_{\widetilde{m}}(f)$ share the same H_s-wavefront set. We do so in the next lemma.

LEMMA 4.4. *Under the hypothesis of Theorem 3.7, 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* $M_{\widetilde{m}}(f)$ *does. That is,* f *and* $M_{\widetilde{m}}(f)$ *have the same* H_s*-wavefront set.*

Proof. The Fourier multiplier $M_{\widetilde{m}}$ is a pseudo-differential operator with Weyl symbol $\sigma(x,w) := \widetilde{m}(w)$. Under the hypothesis of Theorem 3.7, Equation (3.13) holds. This implies that $M_{\widetilde{m}}$ is elliptic of order 0. We will now show that σ belongs to Hörmander's symbol class $S^0_{\frac{1}{2},0}$. That is,

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

Once this is established, it will follow that $M_{\widetilde{m}}$ preserves the H_s-wavefront set. We write

$$\widetilde{m}(\xi) := |\widehat{\phi}(\xi)|^2 + \widetilde{m}'(\xi),$$

where

$$\widetilde{m}'(\xi) := \sum_{j \geq 1} \sum_{k=0}^{2^j - 1} \Phi(B_{j,k}\xi),$$

$$\Phi(\xi) := \widehat{\varphi}(\xi) \overline{\widehat{\psi}(\xi)}.$$

Since $\left| \widehat{\phi}(\xi) \right|$ is a Schwartz function we focus on $\widetilde{m}'$. Note that Φ is a compactly-supported smooth function that vanishes in a neighborhood of the origin $(-\varepsilon, \varepsilon)^2$.

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

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

Therefore,

$$\left| \partial_\xi^\alpha \widetilde{m}'(\xi) \right| \leq \sum_{j \geq 1} \sum_{k=0}^{2^j - 1} 2^{-jN} \Phi_N(B_{j,k}\xi). \tag{4.25}$$

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, if we further consider $\Phi^*_N(\xi) := |\xi|^{N/2} \Phi(\xi)$ we have that

$$|\xi|^{N/2} \left| \partial_\xi^\alpha \widetilde{m}'(\xi) \right| \leq \sum_{j \geq 1} \sum_{k=0}^{2^j - 1} \Phi^*_N(B_{j,k}\xi) \lesssim 1,$$

where the last bound follows from the fact that the support of Φ^*_N is compact and does not contain 0 (cf. Prop. 1). □

5. Acknowledgements. K.G. and J.L.R. gratefully acknowledge support from the Austrian Science Fund (FWF) [P22746-N13]. M.V. dH. was supported in part by the members of the Geo-Mathematical Imaging Group.

REFERENCES

[1] F. Andersson, M. Carlsson, and L. Tenorio. On the Representation of Functions with Gaussian Wave Packets. *J. Fourier Anal. Appl.*, 18:146–181, 2012.
[2] 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.
[3] E. J. Candès and David 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.
[4] 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.
[5] E. J. Candès and L. Demanet. The curvelet representation of wave propagators is optimally sparse. *Commun. Pure Appl. Anal.*, 58(11):1472–1528, 2005.
[6] A. Cordoba and C. Fefferman. Wave packets and Fourier integral operators. *Comm. Partial Differential Equations*, 3:979–1005, 1978.
[7] H. Douma and M. V. de Hoop, Leading-order seismic imaging using curvelets. *Geophys. J. Internat.*, 72(6):S231–S248, 2007.
[8] P. Gérard. Moyennisation et régularité deux-microlocale. *Ann. Sci. École Norm. Sup. (4)*, 23(1):89–121, 1990.
[9] P. Grohs. Continuous shearlet frames and resolution of the wavefront set. *Monatsh. Math.*, 164(4):393–426, 2011.
[10] K. Johansson, S. Pilipović, N. Teofanov, and J. Toft. Gabor pairs, and a discrete approach to wave-front sets. *Monatsh. Math.*, 166(2):181–199, 2012.
[11] G. Kutyniok and D. Labate. Resolution of the wavefront set using continuous shearlets. *Trans. Amer. Math. Soc.*, 361(5):2719–2754, 2009.
[12] P. Kittipoom, G. Kutyniok, and W.-Q. Lim. Construction of Compactly Supported Shearlet Frames. *Constr. Approx.*, 35:21–72, 2012.
[13] H. F. Smith. A Hardy space for Fourier integral operators. *J. Geom. Anal.*, 8(4):629–653, 1998.
[14] 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.