

## REGULARITY IN, AND MULTI-SCALE DISCRETIZATION OF THE SOLUTION CONSTRUCTION OF HYPERBOLIC EVOLUTION EQUATIONS OF LIMITED SMOOTHNESS \*

MAARTEN V. DE HOOP<sup>†</sup>, SEAN F. HOLMAN<sup>‡</sup>, HART F. SMITH<sup>§</sup>, AND GUNTHER UHLMANN<sup>¶</sup>

**Abstract.** We present a multi-scale solution scheme for hyperbolic evolution equations with curvelets. We assume, essentially, that the second-order derivatives of the symbol of the evolution operator are uniformly Lipschitz. The scheme is based on a solution construction introduced by Smith [19] and is composed of generating an approximate solution following a paradifferential decomposition of the mentioned symbol, here, with a second-order correction reminiscent of geometrical asymptotics involving a Hamilton-Jacobi system of equations and, subsequently, solving a particular Volterra equation. We analyze the regularity of the associated Volterra kernel, and then determine the optimal quadrature in the evolution parameter. Moreover, we provide an estimate for the spreading of (finite) sets of curvelets, enabling the multi-scale numerical computation with controlled error.

**Key words.** dyadic parabolic decomposition, curvelet transform, computational harmonic analysis, propagator

**1. Introduction.** We study the regularity in the construction of solutions of a general class of evolution equations with limited smoothness. We have applications to wave propagation in non-smooth media in mind. The construction makes use of a frame of curvelets [12, 10, 11], generates the weak solution on the one hand but reveals the geometrical properties reminiscent of the propagation of singularities in the case of smooth media on the other hand.

Let  $p(z, x, \xi)$  be a real-valued function defined on  $[0, Z] \times \mathbb{R}_x^n \times (\mathbb{R}_\xi^n \setminus \{0\})$  that is smooth and positively homogeneous of degree 1 with respect to  $\xi$ . If  $z_0 \in [0, Z]$ , we then consider the initial value problem

$$(1.1) \quad (\partial_z - ip(z, x, D_x))u(z, x) = 0 \quad \text{and} \quad u(z_0, x) = u_0(x).$$

Here  $p(z, x, D_x)$  is a pseudodifferential operator ( $\Psi$ DO) whose symbol,  $p$ , may be rough in the  $z$  and  $x$  variables. We will require here that  $p \in C^{m,1}S_{cl}^1$  with  $m = 1$  or  $2$ . This means that the  $m$ th derivatives of  $p$  with respect to  $z$  and  $x$  exist everywhere, are uniformly Lipschitz, and for every multi-index  $\alpha$  there is a constant  $C_\alpha$  such that for all  $\xi$  sufficiently large

$$\|\partial_\xi^\alpha p(\cdot, \xi)\|_{C^{m,1}} \leq C_\alpha(1 + \|\xi\|)^{1-|\alpha|}.$$

The notation of the left-hand side above indicates the  $C^{m,1}$  norm in  $(z, x)$ .

The technique used here for construction of solutions to (1.1) was introduced by Smith [19]. More recently, properties of these solutions were studied from the point of view of concentration of curvelets motivated by the propagation of singularities for the case of smooth symbols [3]. The solution construction is initiated by the construction of an approximate solution following the smoothing, that is, paradifferential decomposition of the symbol  $p$ , and is completed by solving a Volterra equation of the second kind which corrects for the symbol smoothing and essentially accounts for scattering between curvelets. The approximate solution is constructed using geometrical asymptotics and involves solving the Hamilton and Hamilton-Jacobi systems generated by the smoothed symbols. The Volterra equation can be solved by a Neumann series – as in the computation of certain multiple scattering series – revealing a curvelet-curvelet interaction. The main goal of this work is to develop regularity estimates in the evolution coordinate  $z$  for the Volterra kernel and

---

\*This research was supported in part under NSF ARRA grant DMS 0908274, and by the members of the Geo-Mathematical Imaging Group.

<sup>†</sup> Center for Computational and Applied Mathematics, and Geo-Mathematical Imaging Group, Purdue University, West Lafayette, IN 47907, USA.

<sup>‡</sup> Center for Computational and Applied Mathematics, and Geo-Mathematical Imaging Group, Purdue University, West Lafayette, IN 47907, USA (sfholman@purdue.edu).

<sup>§</sup> Department of Mathematics, University of Washington, Seattle, WA 98195-4350 USA.

<sup>¶</sup> Department of Mathematics, University of Washington, Seattle, WA 98195-4350 USA.

solution. These estimates govern the choice of quadrature used when solving the Volterra equation, and subsequently the initial value problem, numerically.

Our main result uses an adapted underlying approximate solution operator (para-metrix) for (1.1) with second-order correction. With this parametrix we provide scale-independent regularity estimates of the associated Volterra kernel in  $H^s$ , and likewise estimates for the regularity of the solution  $g(z, x)$  of the Volterra equation in the  $z$  variable as a map into  $H^s$ , when  $m = 2$  (or larger) and  $-1 \leq s < 2$ . Specifically, we obtain a Hölder estimate of order  $1/2$ . Thus a natural choice of quadrature when considering the numerical solution of the Volterra equation becomes the trapezoidal rule [6]. The approximate solution construction to second order is obtained from results pertaining to expansions of Fourier integral operators generated by canonical transformations [13, 9]. This second order parametrix improves on first order parametrices in at least two ways. First, the Hölder regularity mentioned above is required to prove that a discretization of the Volterra equation in the evolution parameter  $z$  converges as the discretization step size goes to zero. Second, the Volterra kernel associated with the second order parametrix is actually compact acting on  $H^s$  (in fact it maps into  $H^{s+1/2}$ ) and so exhibits better behavior when iterated.

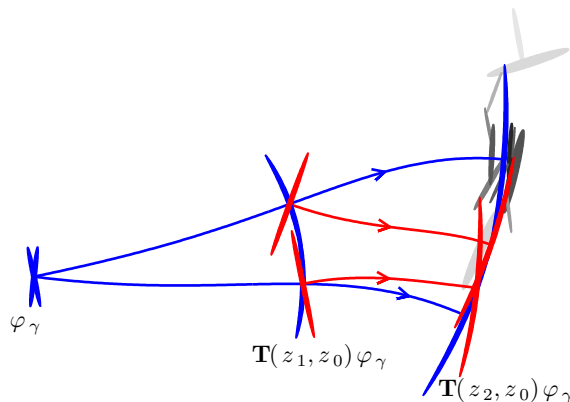


FIG. 1. Diagram describing a numeric solution including a single Volterra iteration. The parametrix is based on the second-order (geometric) approximation and depicted in blue. Some “scattered” curvelets produced by the Volterra iteration are depicted in red. The decay of curvelet frame coefficients for one of the wave packets propagated by the parametrix is illustrated by grayscale.

The results obtained here can be extended directly to apply to solving the second-order wave equation and associated Cauchy initial value problem.

A key aspect of developing an efficient computational algorithm will rely on available sparse decompositions of  $u_0$  (that is, the initial data at  $z_0$ ), and of the Volterra operator applied to the current solution by the Neumann series expansion (that is, the residual force at values of  $z$  dictated by the chosen quadrature). We have developed first steps towards an approach based on nonlinear approximation [1, 2], motivated by the work of Beylkin and Monzón [4, 5]. Here, we provide an estimate of the spreading of the set of curvelet coefficients under propagation as a function of scale. Following the decomposition of  $u_0$  into wave packets, a natural solution strategy – tracing the convergence of the Neumann series expansion – starts at the finest available scale and progresses to the coarser scales. The Volterra equation can be solved with a step-by-step method reminiscent of the semi-group property. The numerical analysis of curvelet-like transforms can be found in [8, 15]; this analysis plays a role in developing a fast algorithm for the above mentioned approximate solution. We note that the regularity and spreading estimates obtained here imply error estimates of corresponding numerical schemes, which will be published elsewhere.

The results obtained in this paper have direct applications, for example, in seismic imaging. Indeed most imaging procedures can be expressed in terms of evolution equations [16]. We mention “reverse-time migration” based imaging [14] and “downward continuation (reverse depth)” based

imaging [20, 21]. Furthermore, curvelet based data regularization dove tails perfectly with these imaging techniques.

**2. Solution of the evolution equation.** We find solutions for (1.1) in two steps. We first construct an approximate solution operator, which we will refer to as a parametrix, and then we use this parametrix to transform (1.1) into an equivalent Volterra equation of the second kind for a function with values in a Sobolev space. To be more precise, we first construct a family of operators  $\mathbf{T}(z, z') : H^s(\mathbb{R}^n) \rightarrow H^s(\mathbb{R}^n)$  for some  $s \in \mathbb{R}$  parametrized by  $(z, z') \in [0, Z] \times [0, Z]$  which satisfy the following two properties

$$(2.1) \quad \mathbf{T}(z', z') = \text{Id} \quad \text{for all } z' \in [0, Z], \text{ and}$$

$$(2.2) \quad (\partial_z - ip(z, x, D_x)) \mathbf{T}(z, z') : H^s \rightarrow H^s \quad \text{uniformly for } (z, z') \in [0, Z] \times [0, Z].$$

Because  $(\partial_z - ip(z, x, D_x))$  is a (possibly rough)  $\Psi$ DO of order 1, the mapping property (2.2) is better than would be expected and so it is this property that makes  $\mathbf{T}(z, z')$  an approximate solution operator. Here “uniformly” means that there is a single modulus of continuity that holds for all  $(z, z')$ . We think of  $\mathbf{T}(z, z_0)u_0$  as giving the approximate solution of (1.1). Any family of operators satisfying these properties will be called a *parametrix*.

Once we have a parametrix we look for an exact solution for (1.1) in the form

$$(2.3) \quad u(z, x) = [\mathbf{T}(z, z_0)u_0](x) + \int_{z_0}^z [\mathbf{T}(z, z')g(z', \cdot)](x) dz'.$$

Intuitively, we are setting  $u(z, x)$  to be the approximate solution plus an error term that we expect can be found or at least estimated. Now the function to be determined is  $g(z, x)$  which we refer to as the *residual*. A calculation making use of (2.1) shows that  $u(z, x)$  is a solution of (1.1) if and only if

$$g(z, x) = -[(\partial_z - ip(z, x, D_x))\mathbf{T}(z, z_0)u_0](x) - \int_{z_0}^z [(\partial_z - ip(z, x, D_x))\mathbf{T}(z, z')g(z', \cdot)](x) dz'.$$

Motivated by this fact we introduce the *Volterra kernel*

$$(2.4) \quad \mathbf{K}(z, z') = -(\partial_z - ip(z, x, D_x))\mathbf{T}(z, z')$$

so that the equation for  $g(z, x)$  becomes

$$(2.5) \quad g(z, x) = [\mathbf{K}(z, z_0)u_0](x) + \int_{z_0}^z [\mathbf{K}(z, z')g(z', \cdot)](x) dz'.$$

This is a linear Volterra equation of the second kind where the function to be determined ( $g(z, \cdot)$ ) takes values in the Sobolev space  $H^s$ . For a review of the classical theory of this type of equation see [6]. Although the theory there only explicitly deals with real and complex valued functions many of the results still hold in the case of functions valued in general Banach spaces with the same proofs.

The solution of (2.5) may be obtained via a Neumann series. Indeed, let us define

$$\mathbf{K}^1(z) = \mathbf{K}(z, z_0), \quad \text{and for } n > 1 \quad \mathbf{K}^n(z) = \int_{z_0}^z \mathbf{K}(z, z')\mathbf{K}^{n-1}(z') dz'.$$

Note that by (2.2)  $\mathbf{K}(z, z') : H^s \rightarrow H^s$  uniformly, and thus the composition used in the iterative definition of  $\mathbf{K}^n(z)$  is still an operator on  $H^s$ . Furthermore, if  $\|\mathbf{K}(z, z')\|_{(H^s, H^s)} \leq C(Z)$  for all  $z$  and  $z'$ , then for all  $n$  and  $z$  it follows from the definition that

$$(2.6) \quad \|\mathbf{K}^n(z)\|_{(H^s, H^s)} \leq \frac{Z^{n-1}}{(n-1)!} C(Z)^n.$$

The solution of (2.5) is then

$$g(z, x) = \sum_{n=1}^{\infty} [\mathbf{K}^n(z)u_0](x) =: [\mathbf{R}(z)u_0](x).$$

By (2.6) this sum converges absolutely in  $H^s$  for every  $z \in [0, Z]$ , and in fact

$$(2.7) \quad \|\mathbf{R}(z)\|_{(H^s, H^s)} \leq C(Z)e^{ZC(Z)}.$$

We refer to  $\mathbf{R}(z)$  as the *resolvent* corresponding to the parametrix  $\mathbf{T}$ .

This method of solution was first introduced for the half wave equation in [19], and has been used previously to analyze the equation (1.1) in [3]. In both of these works the parametrix  $\mathbf{T}$  is constructed by decomposing  $u_0$  in the curvelet frame, and then applying a rigid motion to each individual curvelet. We refer to this ‘‘rigid motion’’ parametrix as  $\mathbf{T}_1$ . In the current work we will introduce a new parametrix,  $\mathbf{T}_2$ , which still uses a curvelet decomposition of  $u_0$ , but also incorporates spreading into the evolution of each individual curvelet. As we will see, when  $\mathbf{T}_2$  is used as the parametrix the corresponding Volterra kernel  $\mathbf{K}_2(z, z')$  will have additional regularity properties in the  $z$  variables.

**3. Construction of the parametrices.** In this section we describe two possible ways to construct a parametrix satisfying the requirements (2.1) and (2.2). Both methods are based upon a curvelet decomposition. The first uses only a rigid motion of the curvelets, while the second also incorporates spreading. The treatment of caustics in the second method needs special attention, which we do not elaborate on here. The first method does not provide strong enough estimates to guarantee that numerical solutions of the Volterra equation will converge. The proofs in this section and section 4 make use of the results in both of the appendices, and in particular the rules for manipulating families of curvelet like functions (FCLFs) developed in appendix A. When dealing with a FCLF  $\mathcal{F}$  we sometimes use the notation  $f_\gamma \in \pi_S(\mathcal{F})$  for a function in the family corresponding to the curvelet index  $\gamma$ .

The first step for both parametrices is to smooth the rough symbol  $p$  of (1.1) in the  $x$  variable according to scale. In this way we obtain a sequence of smooth (in  $x$ ) symbols  $p_k$  which approximate  $p$ . Indeed, let  $\psi \in C_c^\infty(\mathbb{R}^n)$  be an even function such that  $\psi(\xi) = 1$  for  $|\xi| \leq 1$  and  $\psi(\xi) = 0$  for  $|\xi| \geq 2$ . We also assume  $0 \leq \psi \leq 1$  everywhere. Then define

$$(3.1) \quad p_k(z, x, \xi) = [\psi(2^{-k/2}D_x)p(z, \cdot, \xi)](x)$$

for all  $k \in \mathbb{N}$ . Thus for each  $k$  we low pass filter  $p$  in the  $x$  variable around the frequency  $2^{k/2}$  to obtain  $p_k \in C^\infty$ . This sequence of approximations to  $p$  satisfies the following estimates. For  $j + |\beta| \leq m + 1$  (when  $j + |\beta| = m + 1$  estimate (3.2) holds everywhere the left hand side is defined)

$$(3.2) \quad |\partial_z^j \partial_x^\beta \partial_\xi^\alpha (p - p_k)(z, x, \xi)| \lesssim 2^{-k(m-|\beta|-j+1)/2} \|\partial_\xi^\alpha p(\cdot, \xi)\|_{C^{m,1}}$$

and

$$(3.3) \quad |\partial_z^j \partial_x^\beta \partial_\xi^\alpha p_k(z, x, \xi)| \lesssim \|\partial_\xi^\alpha p(\cdot, \xi)\|_{C^{m,1}}.$$

Also, if  $j \leq m$  and  $|\beta| \geq m + 1 - j$  then

$$(3.4) \quad |\partial_z^j \partial_x^\beta \partial_\xi^\alpha p_k(z, x, \xi)| \lesssim 2^{k(|\beta|+j-m-1)/2} \|\partial_\xi^\alpha p(\cdot, \xi)\|_{C^{m,1}}.$$

Here and in the remainder of this work the notation  $A_k \lesssim B_k$  means that there exists a constant  $C > 0$  independent of the scale  $k$ , or more generally the index  $\gamma = (x, \nu, k)$ , such that  $A_k \leq CB_k$ . In the following parametrix constructions  $p_k(z, x, D_x)$  will be used to approximate the action of  $p(z, x, D_x)$  on the curvelets at scale  $k$ .

**3.1. Rigid motion parametrix.** We first review the construction from [3] of a parametrix, referred to here as  $\mathbf{T}_1(z, z')$ , based only on the rigid motion of curvelets. The purpose of this review is twofold. First, some of the techniques involved will be used again in the construction of the new parametrix in section 3.2, and second we eventually wish to compare some results for this parametrix and associated Volterra kernel to those for the new parametrix. In this interest we will also prove regularity estimates for  $\mathbf{T}_1(z, z')$  in the  $z$  and  $z'$  variables. We will always assume that  $m = 1$  when we are considering  $\mathbf{T}_1$ .

We begin by considering the system

$$(3.5) \quad \frac{dy_k}{dz}(z, z') = -\partial_\xi p_k(z, y_k, \nu_k), \quad \frac{d\nu_k}{dz}(z, z') = \partial_x p_k(z, y_k, \nu_k) - \left\langle \nu_k, \partial_x p_k(z, y_k, \nu_k) \right\rangle \nu_k,$$

and

$$(3.6) \quad \frac{d\Theta_k}{dz}(z, z') = \Theta_k \left[ \nu_k \otimes \partial_x p_k(z, y_k, \nu_k) - \partial_x p_k(z, y_k, \nu_k) \otimes \nu_k \right].$$

which gives the projected Hamiltonian flow associated to  $p_k$ . We write

$$y_k(z, z', x, \nu), \quad \nu_k(z, z', x, \nu), \quad \text{and} \quad \Theta_k(z, z', x, \nu)$$

for the solution of (3.5) and (3.6) with initial data

$$y_k(z', z', x, \nu) = x, \quad \nu_k(z', z', x, \nu) = \nu, \quad \text{and} \quad \Theta_k(z', z', x, \nu) = \text{Id},$$

and refer to the map  $(x, \nu) \mapsto (y_k(z, z', x, \nu), \nu_k(z, z', x, \nu))$  as  $\Psi_{z, z'}^k$ . We also consider the system (3.5) and (3.6) with  $p_k$  replaced by  $p$ , and introduce a corresponding map  $\Psi_{z, z'}$  defined in the analogous manner (note that since  $p \in C^{1,1}$  the problem is well-posed). In [19] it is shown that

$$(3.7) \quad d(\Psi_{z, z'}(x, \nu), \Psi_{z, z'}^k(x, \nu)) \lesssim 2^{-k},$$

where  $d$  is the pseudodistance defined in appendix A.

If  $\gamma = (x, \nu, k)$  is a curvelet index, then the *flow out* of the individual curvelet  $\varphi_\gamma$  is given by

$$\varphi_{1, \gamma}(z, z', y) = \varphi_\gamma(\Theta_k(z, z', x, \nu)(y - y_k(z, z', x, \nu)) + x).$$

If  $u \in L^2(\mathbb{R}^n)$ , then the parametrix  $\mathbf{T}_1(z, z')$  is defined by

$$(3.8) \quad [\mathbf{T}_1(z, z')u](y) = u_0 \varphi_0(y) + \sum_{\gamma} u_\gamma \varphi_{1, \gamma}(z, z', y)$$

where the  $u_\gamma$  are the coefficients of  $u$  given by the curvelet co-frame. Since it will be useful below, we also define operators  $\mathbf{T}_1^{k'}(z, z')$  which only give the contributions of curvelets at scale  $k'$ :

$$(3.9) \quad [\mathbf{T}_1^{k'}(z, z')u](y) = \sum_{\{\gamma=(x, \nu, k): k=k'\}} u_\gamma \varphi_{1, \gamma}(z, z', y).$$

It is proven in [3] that  $\mathbf{T}_1(z, z')$  is a parametrix as defined in section 2 for  $-1 \leq s < 2$ . To finish this section we prove the following regularity result for  $\mathbf{T}_1(z, z')$ .

**LEMMA 3.1.** *The operator  $\mathbf{T}_1(z, z')$  is uniformly Lipschitz in both of its arguments as a map from  $H^s$  to  $H^{s-1}$  for any  $s$  and on any fixed domain  $[0, Z] \times [0, Z]$ . That is*

$$(3.10) \quad \|\mathbf{T}_1(z, z') - \mathbf{T}_1(\bar{z}, \bar{z}')\|_{(H^s, H^{s-1})} \leq C_1(Z)|z - \bar{z}|$$

for all  $z, \bar{z}$ , and  $z' \in [0, Z]$ , and the same estimate holds when  $z'$  is varied rather than  $z$ . Furthermore,

$$(3.11) \quad \|\mathbf{T}_1^{k'}(z, z') - \mathbf{T}_1^{k'}(\bar{z}, z')\|_{(L^2, L^2)} \leq C_1(Z) 2^{k'} |z - \bar{z}|$$

and the same holds when  $z'$  is varied instead of  $z$ .

*Proof.* For  $\gamma = (x_0, \nu^0, k)$  the index of a curvelet, we define the change of variables

$$y \mapsto \Phi_{1,\gamma}(z, z', y) := \Theta_k(z, z', x_0, \nu^0)(y - y_k(z, z', x_0, \nu^0)) + x_0.$$

This family of maps satisfies the hypotheses of lemma A.7 relative to the FCLF given by the curvelet frame. Therefore

$$\left\{ 2^{-k} \partial_z \varphi_{1,\gamma}(z, z', y), \Phi_{1,\gamma}^*(\gamma) \right\}$$

is a FCLF, and by lemma A.2 we have for every  $\delta > 0$  a constant  $C_\delta$  such that

$$(3.12) \quad \left| \langle \varphi_{\tilde{\gamma}}, \partial_z \varphi_{1,\gamma}(z, z', y) \rangle \right| \leq C_\delta 2^k \mu_\delta(\tilde{\gamma}, \Phi_{1,\gamma}^*(\gamma)),$$

where  $\mu$  is the weight function introduced in appendix A. Now let us consider the infinite matrix

$$(3.13) \quad c_{1,\tilde{\gamma}\gamma}(z, \bar{z}, z') = \int_{\mathbb{R}^n} \overline{\varphi_{\tilde{\gamma}}(y)} (\varphi_{1,\gamma}(z, z', y) - \varphi_{1,\gamma}(\bar{z}, z', y)) dy$$

which provides a representation of  $\mathbf{T}_1(z, z') - \mathbf{T}_1(\bar{z}, z')$  with respect to the curvelet frame. Also, we have representations of  $\mathbf{T}_1^{k'}(z, z') - \mathbf{T}_1^{k'}(\bar{z}, z')$  for any  $k'$  given by  $c_{1,\tilde{\gamma}\gamma}^{k'}(z, \bar{z}, z') = c_{1,\tilde{\gamma}\gamma}(z, \bar{z}, z')$  when  $k' = k$ , and  $c_{1,\tilde{\gamma}\gamma}^{k'}(z, \bar{z}, z') = 0$  otherwise. By results in [19], to prove (3.10) it is sufficient to show that

$$(3.14) \quad |c_{1,\tilde{\gamma}\gamma}(z, \bar{z}, z')| \leq C_\delta |z - \bar{z}| 2^k u_\delta(\tilde{\gamma}, \Psi_{z,z'}(\gamma))$$

for every  $\delta > 0$ . Here  $\Psi_{z,z'}((x, \nu, k)) = (\Psi_{z,z'}(x, \nu), k)$ . First, the estimate

$$(3.15) \quad |\partial_z c_{1,\tilde{\gamma}\gamma}(z, \bar{z}, z')| \leq C_\delta 2^k \mu_\delta(\tilde{\gamma}, \Psi_{z,z'}^k(\gamma))$$

is equivalent to (3.12). Second, using (3.7) we may replace  $\Psi_{z,z'}^k$  by  $\Psi_{z,z'}$  in (3.15). Finally, because  $c_{1,\tilde{\gamma}\gamma}(\bar{z}, \bar{z}, z') = 0$  the mean value theorem implies (3.14), which completes the proof that  $\mathbf{T}_1(z, z')$  is Lipschitz in the  $z$  variable. Further, it follows from (3.14) that

$$|c_{1,\tilde{\gamma}\gamma}^{k'}(z, \bar{z}, z')| \leq C_\delta |z - \bar{z}| 2^{k'} \mu_\delta(\tilde{\gamma}, \Psi_{z,z'}(\gamma))$$

which implies the estimate for  $\mathbf{T}_1^{k'}$ .

The result for the  $z'$  variable follows by the same proof if we begin by obtaining (3.12) where the differentiation is with respect to  $z'$  instead of  $z$ .  $\square$

**REMARK 3.2.** We observe that both  $\mathbf{T}_1^k(z, z')$  and  $\mathbf{T}_1^k(z, z') - \mathbf{T}_1^k(\bar{z}, z')$  are families of operators satisfying the hypotheses required for  $F_k$  in lemma B.1 with, respectively,  $r = 0$  and  $r = 1$  and in the latter case with  $C = C_1(Z) |z - \bar{z}|$ .

At this point we note that lemma 3.1 together with the fact that  $\mathbf{T}_1$  is a parametrix allow us to prove a (already known, see e.g. [22]) regularity result for the solution of (1.1). Indeed, from (2.3) we see that if  $m = 1$ , then for  $-1 \leq s < 2$  and initial data  $u_0 \in H^s$  the solution  $u(z, x)$  of (1.1) is in  $C^{0,1}([0, Z]; H^{s-1})$ .

**3.2. Parametrix with second order correction.** In this section we will construct a parametrix,  $\mathbf{T}_2(z, z')$ , that takes into account the spreading of curvelets. The action of this parametrix will be specified in the same way as in section 3.1 by defining an action on each curvelet individually. The underlying motivation for the parametrix comes from an approximation to a Fourier integral operator, via a phase expansion as discussed in [13], with phase function defining the propagation of singularities for (1.1). We should note that this parametrix construction only works in the absence of caustics (this restriction will be made more precise below). However, if there is a global minimum time before any caustics develop, then it is possible to repeatedly apply the construction stepping forward in sufficiently small time steps. Thus, though we do not formulate the precise statements here, these results can also apply past caustics. When we consider  $\mathbf{T}_2(z, z')$  we will always assume that  $m = 2$ .

We begin the construction by introducing the Hamiltonian system that gives the propagation of singularities for (1.1). In contrast to (3.5) and (3.6) the integral curves here are not projected onto the unit co-sphere. For every  $(x, \eta) \in \mathbb{R}_x^n \times (\mathbb{R}_\eta^n \setminus \{0\})$  we consider the flow given by

$$(3.16) \quad \frac{dy_k}{dz}(z, z', (x, \eta)) = -\partial_\xi p_k(z, y_k, \nu_k), \quad \text{and} \quad \frac{d\nu_k}{dz}(z, z', (x, \eta)) = \partial_x p_k(z, y_k, \nu_k)$$

with initial data  $y_k(z', z', (x, \eta)) = x$ , and  $\nu_k(z', z', (x, \eta)) = \eta$ . The curves

$$(y_k(z, z', (x, \eta)), \nu_k(z, z', (x, \eta)))$$

are the integral curves of the  $z$ -dependent Hamiltonian vector field given by  $p_k$  with initial data  $(x, \eta)$ . We consider the following system of equations

$$(3.17) \quad (y, \nu) = (y_k(z, z', (x, \eta)), \nu_k(z, z', (x, \eta))).$$

For every  $k$  these define a mapping from  $(z, z', (x, \eta))$  to  $(y, \nu)$  which is the canonical relation of the solution operator for (1.1) if  $p$  is replaced by  $p_k$ . Using them to define implicit relations amongst the various variables amounts to parametrizing this canonical relation by different subsets of the variables.

Now we supplement the flow (3.16) with another system that gives the dependence of  $(y, \nu)$  on perturbations of  $(x, \eta)$ . This system, the Hamilton-Jacobi system associated to  $p_k$ , is

$$(3.18) \quad \frac{dW_k}{dz}(z, z', (x, \eta)) = \begin{pmatrix} -\partial_{\xi x}^2 p_k(z, y_k, \nu_k) & -\partial_{\xi \xi}^2 p_k(z, y_k, \nu_k) \\ \partial_{x x}^2 p_k(z, y_k, \nu_k) & \partial_{x \xi}^2 p_k(z, y_k, \nu_k) \end{pmatrix} W_k(z, z', (x, \eta))$$

where  $W_k(z, z', (x, \eta))$  is a  $2n \times 2n$  matrix with initial data  $W_k(z', z', (x, \eta)) = \text{Id}_{2n}$ . We split  $W_k(z, z', (x, \eta))$  up into four  $n \times n$  matrices

$$W_k(z, z', (x, \eta)) = \begin{pmatrix} {}_k W_1(z, z', (x, \eta)) & {}_k W_2(z, z', (x, \eta)) \\ {}_k W_3(z, z', (x, \eta)) & {}_k W_4(z, z', (x, \eta)) \end{pmatrix},$$

and then we have

$$\frac{\partial y_k}{\partial x}(z, z', (x, \eta)) = {}_k W_1(z, z', (x, \eta)).$$

We will assume that  ${}_k W_1(z, z', (x, \eta))$  is always invertible, and so by the implicit function theorem the equations (3.17) can be solved for  $x$  and  $\nu$  as a function of  $(z, z', y, \eta)$  at least locally. Since these functions depend on  $k$  we will label them as  $\tilde{x}_k$  and  $\tilde{\nu}_k$ . We can then introduce a defining function  $S_k(z, z', y, \eta)$  for the canonical relation defined by (3.17) given by

$$S_k(z, z', y, \eta) = \langle \tilde{x}_k(z, z', y, \eta), \eta \rangle.$$

We will always assume that this map  $\tilde{x}_k$  exists globally for  $z$  and  $z' \in [0, Z]$ . This is the assumption that there are no caustics. We can also find formulas for the derivatives of  $\tilde{x}_k(z, z', y, \eta)$

and  $\tilde{\nu}_k(z, z', y, \eta)$ . In the following the matrices  ${}_k W_i$  are understood to be evaluated at the point  $(z, z', \tilde{x}_k(z, z', y, \eta), \eta)$ .

$$(3.19) \quad \frac{\partial \tilde{x}_k}{\partial y}(z, z', y, \eta) = {}_k W_1^{-1},$$

$$(3.20) \quad \frac{\partial \tilde{x}_k}{\partial \eta}(z, z', y, \eta) = -{}_k W_1^{-1} {}_k W_2,$$

$$(3.21) \quad \frac{\partial \tilde{\nu}_k}{\partial y}(z, z', y, \eta) = {}_k W_3 {}_k W_1^{-1},$$

and

$$(3.22) \quad \frac{\partial \tilde{\nu}_k}{\partial \eta}(z, z', y, \eta) = {}_k W_4 - {}_k W_3 {}_k W_1^{-1} {}_k W_2.$$

Using the homogeneity of  $p_k$  we can also prove the two following properties

$$(3.23) \quad \eta = {}_k W_1^T \tilde{\nu}_k(z, z', y, \eta),$$

and

$$(3.24) \quad {}_k W_2^T {}_k W_1^{-T} \eta = 0.$$

Here  ${}_k W_1^{-T}$  refers to the inverse of the transpose of  ${}_k W_1$ . Finally, since  $S_k(z, z', y, \eta)$  is a smooth function, using the above properties and the equality of the mixed partials of  $S_k$  we have

$$(3.25) \quad {}_k W_1^{-T} = \left( \frac{\partial x_k}{\partial y} \right)^T = \left( \frac{\partial^2 S_k}{\partial y \partial \eta} \right)^T = \frac{\partial^2 S_k}{\partial \eta \partial y} = \frac{\partial \nu_k}{\partial \eta} = {}_k W_4 - {}_k W_3 {}_k W_1^{-1} {}_k W_2.$$

Thus (3.22) becomes

$$(3.22') \quad \frac{\partial \tilde{\nu}_k}{\partial \eta}(z, z', y, \eta) = {}_k W_1^{-T}.$$

Finally, we can check using some of the above identities that  ${}_k W_3 {}_k W_1^{-1}$  and  ${}_k W_1^{-1} {}_k W_2$  are always symmetric matrices.

Next we will introduce the phase function used to construct the action of our parametriz on curvelets at scale  $k$ . Let  $\gamma = (x_0, \nu^0, k)$  be the index of a curvelet. Then, for  $z, z' \in [0, Z]$ ,  $y \in \mathbb{R}^n$ , and  $\eta \in \mathbb{R}^n$ , define (motivated in part by an expansion of  $S_k(z, z', y, \eta)$ )

$$\begin{aligned} \tilde{S}_\gamma(z, z', y, \eta) &= \langle \tilde{x}_k(z, z', y, \nu^0), \eta \rangle - \frac{1}{2 \langle \nu^0, \eta \rangle} \langle {}_k W_1^{-1} {}_k W_2 \eta, \eta \rangle \\ &\quad - \frac{1}{2} \int_{z'}^z \text{tr} \left( {}_k W_3 {}_k W_1^{-1} \partial_{\xi\xi}^2 p_k(z, y_k, \nu_k) \right) dz \end{aligned}$$

where the  ${}_k W_i$ ,  $y_k$ , and  $\nu_k$  are the functions defined above all evaluated at the point  $(z, z', x_0, \nu^0)$ . This will be the convention for the remainder of this work when  ${}_k W_i$ ,  $y_k$ , or  $\nu_k$  are written without any argument. Note that the last term in the definition only depends on  $z, z'$  and the curvelet index  $\gamma$ . Because of this we introduce the notation

$$U_\gamma(z, z') = \frac{1}{2} \int_{z'}^z \text{tr} \left( {}_k W_3 {}_k W_1^{-1} \partial_{\xi\xi}^2 p_k(z, y_k, \nu_k) \right) dz.$$



Now we define the action of an operator on the curvelet  $\varphi_\gamma$  as

$$\varphi_{2,\gamma}(z, z', y) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{i\tilde{S}_\gamma(z, z', y, \eta)} \widehat{\varphi}_\gamma(\eta) d\eta.$$

Note that  $\varphi_{1,\gamma}$  could be written using the same formula if  $\tilde{S}_\gamma$  were replaced by a simpler phase function. Now  $\mathbf{T}_2(z, z')$  and  $\mathbf{T}_2^{k'}(z, z')$  are defined respectively by (3.8) and (3.9) with  $\varphi_{1,\gamma}$  replaced by  $\varphi_{2,\gamma}$ . In theorem 4.2 we will prove that  $\mathbf{T}_2$  is a parametrrix for (1.1), but for now we prove only the following analog of lemma 3.1.

**LEMMA 3.3.** *The results of lemma 3.1 hold with  $\mathbf{T}_1$  replaced by  $\mathbf{T}_2$  assuming that no caustics develop in the interval  $[0, Z]$ .*

*Proof.* As one might suspect, the proof is similar to that of lemma 3.1. We first note that

$$(3.26) \quad \partial_z \varphi_{2,\gamma}(z, z', y) = \frac{i}{(2\pi)^n} \int e^{i\tilde{S}_\gamma(z, z', y, \eta)} \partial_z \tilde{S}_\gamma(z, z', y, \eta) \widehat{\varphi}_\gamma(\eta) d\eta.$$

Now if we define new functions by

$$(3.27) \quad \widehat{f}_\gamma(z, z', \eta) = e^{\frac{-i}{2\langle \nu^0, \eta \rangle}} \langle {}_k W_1^{-1} {}_k W_2 \eta, \eta \rangle \widehat{\varphi}_\gamma(\eta),$$

then using (3.24) we see that  $\{f_\gamma(z, z', \cdot), \gamma\}_{\gamma \in \Gamma_0}$ , where  $\Gamma_0$  is the grid of indices corresponding to the curvelet frame, is a FCLF. Further, if we write  $\Phi_{2,\gamma}(z, z', y) = \tilde{x}_k(z, z', y, \nu^0)$ , then from (3.26) we have

$$(3.28) \quad \partial_z \varphi_{2,\gamma}(z, z', y) = i e^{-iU_\gamma(z, z')} \Phi_{2,\gamma}^*(z, z', \cdot) \left[ \partial_z \tilde{S}_\gamma(z, z', \Phi_{2,\gamma}^{-1}(z, z', \cdot), D) f_\gamma(z, z', \cdot) \right] (y)$$

where  $\Phi_{2,\gamma}^{-1}(z, z', \cdot)$  is the inverse of  $y \mapsto \Phi_{2,\gamma}(z, z', y)$ . From the definition of  $\tilde{S}_\gamma$  as well as identities (3.19) and (3.25), we may calculate

$$(3.29) \quad \begin{aligned} \partial_z \tilde{S}_\gamma(z, z', \Phi_{2,\gamma}^{-1}(z, z', x), \eta) = \\ \langle \partial_\xi p_k(z, y_k(z, z', x, \nu^0), \nu_k(z, z', x, \nu^0)), {}_k W_1^{-T}(z, z', x, \nu^0) \eta \rangle \\ + \frac{1}{2\langle \nu^0, \eta \rangle} \langle \partial_{\xi\xi}^2 p_k(z, y_k, \nu_k) {}_k W_1^{-T} \eta, {}_k W_1^{-T} \eta \rangle - \frac{1}{2} \text{tr} \left( {}_k W_3 {}_k W_1^{-1} \partial_{\xi\xi}^2 p_k(z, y_k, \nu_k) \right) \end{aligned}$$

From this formula and using (3.4) we see that  $\partial_z \tilde{S}_\gamma(z, z', \Phi_{2,\gamma}^{-1}(z, z', x), \eta)$  satisfies the hypotheses of lemma A.6 where  $z$  and  $z'$  are considered as extraneous parameters. By that lemma and lemma A.7,

$$(3.30) \quad \left\{ 2^{-k} \partial_z \varphi_{2,\gamma}(z, z', y), \Phi_{2,\gamma}^*(z, z', \cdot)(\gamma) \right\}_{\gamma \in \Gamma_0}$$

is a FCLF.

Now let us introduce the matrix coefficients  $c_{2,\tilde{\gamma}\gamma}(z, \bar{z}, z')$  defined by (3.13) with  $\varphi_{1,\gamma}$  replaced by  $\varphi_{2,\gamma}$ . Just as in the proof of Lemma 3.1 from (3.30) it follows that

$$|c_{2,\tilde{\gamma}\gamma}(z, \bar{z}, z')| \leq C_\delta |z - \bar{z}| 2^k \mu_\delta(\tilde{\gamma}, \Phi_{2,\gamma}^*(z, z', \cdot)(\gamma))$$

for any  $\delta > 0$ . Now we can calculate using (3.19) and (3.23)

$$\Phi_{2,\gamma}^*(z, z', \cdot)(\gamma) = \left( y_k(z, z', (x_0, \nu^0)), \frac{\nu_k(z, z', (x_0, \nu^0))}{|\nu_k(z, z', (x_0, \nu^0))|}, k + \log_2(|\nu_k(z, z', (x_0, \nu^0))|) \right).$$

The only difference between this and  $\Phi_{1,\gamma}^*(z, z', \cdot)(\gamma)$  is the potential shift in the scale  $k$  by  $\log_2(|\nu_k(z, z', (x_0, \nu^0))|)$ . However the size of this shift can be bounded uniformly by a constant times  $Z \sup_{\omega \in \mathbb{S}^{n-1}} \|p(\cdot, \omega)\|_{C^{m,1}}$ ,

and so as before we may replace  $\Phi_{2,\gamma}^*(z, z', \cdot)(\gamma)$  by  $\Psi_{z,z'}(\gamma)$ . The results now follow as in the proof of lemma 3.1.  $\square$

**REMARK 3.4.** *From the proof of lemma 3.3, using lemma A.5, we may conclude that  $\mathbf{T}_2^k(z, z')$  and  $\mathbf{T}_2^k(z, z') - \mathbf{T}_2^k(\bar{z}, z')$  satisfy the hypotheses for  $F_k$  in lemma B.1 with respectively  $r = 0$  and  $r = 1$ , and in the latter case  $C = C_2(Z)|z - \bar{z}|$ . The constant  $A$  is related to the change in scale  $\log_2(|\nu_k(z, z', (x_0, \nu^0))|)$ , giving the frequency localization.*

**4. Properties of the Volterra kernels and solutions.** In this section we will prove a number of properties of the Volterra kernels  $\mathbf{K}_1$  and  $\mathbf{K}_2$  associated, by (2.4), respectively to the parametrices  $\mathbf{T}_1$  and  $\mathbf{T}_2$  introduced in the previous section.

**4.1. Regularity estimates for the Volterra kernels.** We will prove two theorems which give respectively Lipschitz and Hölder regularity estimates for  $\mathbf{K}_1$  and  $\mathbf{K}_2$ . The key distinction is that  $\mathbf{K}_2$  has Hölder regularity as a map from  $H^s$  to  $H^s$  for certain values of  $s$ , while  $\mathbf{K}_1$  only has this type of regularity as a map from  $H^s$  to  $H^{s-\epsilon}$  for some positive epsilon. Actually we will just prove a Lipschitz estimate for  $\mathbf{K}_1$  as a map from  $H^s$  to  $H^{s-1}$ , but using interpolation such Hölder estimates could be found.

The first result concerns  $\mathbf{K}_1$ .

**THEOREM 4.1.** *For  $-1 \leq s < 1$*

$$(4.1) \quad \|\mathbf{K}_1(z, z') - \mathbf{K}_1(\bar{z}, z')\|_{(H^{s+1}, H^s)} \leq C_1^H(Z) |z - \bar{z}|,$$

*uniformly in  $z' \in [0, Z]$ . An equivalent estimate holds when  $z'$  is varied instead of  $z$ .*

*Proof.* To begin we make a decomposition of  $(\partial_z - i p(z, y, D_y))\mathbf{T}_1(z, z') - (\partial_{\bar{z}} - i p(\bar{z}, y, D_y))\mathbf{T}_1(\bar{z}, z')$  into the following three pieces

$$(4.2) \quad \sum_k (\partial_z - i p_k(z, y, D_y))\mathbf{T}_1^k(z, z') - (\partial_{\bar{z}} - i p_k(\bar{z}, y, D_y))\mathbf{T}_1^k(\bar{z}, z'),$$

$$(4.3) \quad i \sum_k (p_k(z, y, D_y) - p(z, y, D_y))(\mathbf{T}_1^k(z, z') - \mathbf{T}_1^k(\bar{z}, z')),$$

and

$$(4.4) \quad i \sum_k ((p_k(z, y, D_y) - p_k(\bar{z}, y, D_y)) - (p(z, y, D_y) - p(\bar{z}, y, D_y)))\mathbf{T}_1^k(\bar{z}, z').$$

The proof that the kernel is Lipschitz in the  $z$  variable will be complete if we can estimate the norm of each of the previous three operators by  $|\bar{z} - z|$ . To estimate (4.3) we note, referring to remark 3.2, that  $\mathbf{T}_1^k(z, z') - \mathbf{T}_1^k(\bar{z}, z')$  satisfies the requirements of  $F_k$  in lemma B.1 with  $r = 1$  and  $C = C_1(Z)|z - \bar{z}|$ . Also  $p(z, y, \xi)$  takes the place of  $a(y, \xi)$  with  $m = 1$  for fixed  $z$ , and so lemma B.1 implies the required estimate for (4.3). Similarly, lemma B.1 implies the result for (4.4) taking this time  $F_k = \mathbf{T}_1^k(\bar{z}, z')$ ,  $r = 0$ ,  $m = 0$ , and  $a(y, \xi) = p(z, y, \xi) - p(\bar{z}, y, \xi)$ . We now continue to analyze (4.2).

We will use the same notation as in the proof of lemma 3.1. First define

$$(4.5) \quad \tilde{\varphi}_{1,\gamma}(z, z', y) = (\partial_z - i p_k(z, y, D_y))\varphi_{1,\gamma}(z, z', y)$$

and consider

$$(4.6) \quad \partial_z \tilde{\varphi}_{1,\gamma}(z, z', y) = (\partial_z - i p_k(z, y, D_y))\partial_z \varphi_{1,\gamma}(z, z', y) - i \partial_z p_k(z, y, D_y)\varphi_{1,\gamma}(z, z', y).$$

Applying several of the lemmas from appendix A to this formula we see that

$$\left\{ 2^{-k} \partial_z \tilde{\varphi}_{1,\gamma}(z, z', y), \Phi_{1,\gamma}^*(z, z', \cdot)(\gamma) \right\}_{\gamma \in \Gamma_0}$$

is a FCLF. Here we have omitted some calculations that show the cancellation of certain terms, but these calculations are essentially the same as some which can be found in [3], and a more sophisticated version is given in the proof of theorem 4.2. The result for (4.2) now follows as in the proof of lemma 3.1.

To prove the final statement about regularity in  $z'$  we write

$$\begin{aligned} \mathbf{K}_1(z, z') - \mathbf{K}_1(z, \bar{z}) &= \sum_k (\partial_z - i p_k(z, y, D_y)) (\mathbf{T}_1^k(z, z') - \mathbf{T}_1^k(z, \bar{z})) \\ &\quad - i \sum_k (p_k(z, y, D_y) - p(z, y, D_y)) (\mathbf{T}_1^k(z, z') - \mathbf{T}_1^k(z, \bar{z})). \end{aligned}$$

The required estimate for the first term in the sum above follows just as the estimate for (4.2), while the second term is estimated in the same way as (4.3).  $\square$

The following theorem regarding regularity of  $\mathbf{K}_2$  is the main technical result of this paper.

**THEOREM 4.2.** *For  $-3/2 \leq s < 3$ ,  $\mathbf{K}_2(z, z') : H^{s-1/2} \rightarrow H^s$  continuously. If  $-1 \leq s < 2$  then*

$$(4.7) \quad \|\mathbf{K}_2(z, z') - \mathbf{K}_2(\bar{z}, z')\|_{(H^s, H^s)} \leq C_2^H(Z) |z - \bar{z}|^{1/2},$$

*uniformly in  $z' \in [0, Z]$ . An equivalent estimate holds if we vary  $z'$  instead of  $z$ .*

**REMARK 4.3.** *Note that the first statement of the theorem shows that  $\mathbf{K}_2$  is a parametrix for  $-3/2 \leq s < 3$ .*

*Proof.* First assume that  $-3/2 \leq s < 3$ . We begin as before by splitting  $\mathbf{K}_2$  into smooth and rough parts:

$$\sum_k (\partial_z - i p_k(z, y, D_y)) \mathbf{T}_2^k + i \sum_k (p_k(z, y, D_y) - p(z, y, D_y)) \mathbf{T}_2^k.$$

For the rough part (the second summand above) we use the fact that  $\mathbf{T}_2^k$  satisfies the requirements of lemma B.1 with  $r = 0$ , and so the required estimates follow by applying the lemma with  $m = 2$ . Now we continue to analyze the smooth part given by the first summand.

We use the same notation as in the proof of lemma 3.3, and begin with formulas (3.28) and (3.29). Indeed, by (3.29) we have

$$\begin{aligned} &\partial_z \tilde{S}_\gamma(z, z', \Phi_{2,\gamma}^{-1}(z, z', x), D) f_\gamma \\ &= -i \langle \partial_\xi p_k(z, y_k(z, z', x, \nu^0), \nu_k(z, z', x, \nu^0)), {}_k W_1^{-T}(z, z', x, \nu^0) \partial_x f_\gamma \rangle \\ &\quad - \frac{1}{2} \left( \text{tr} \left( {}_k W_1^{-1} \partial_{\xi\xi}^2 p_k(z, y_k, \nu_k) {}_k W_1^{-T} \left[ \frac{1}{\langle \nu^0, D \rangle} \partial_{xx}^2 f_\gamma \right] \right) \right. \\ &\quad \left. + \text{tr} \left( {}_k W_3 {}_k W_1^{-1} \partial_{\xi\xi}^2 p_k(z, y_k, \nu_k) f_\gamma \right) \right). \end{aligned}$$

Combined with (3.28) this gives a formula for  $\partial_z \varphi_{2,\gamma}(z, z', y)$ .

Next let us analyze  $i p_k(z, y, D_y) \varphi_{2,\gamma}(z, z', y)$ . Since

$$\varphi_{2,\gamma}(z, z', y) = e^{-i U_\gamma(z, z')} [\Phi_{2,\gamma}^*(z, z', \cdot) f_\gamma(z, z', \cdot)](y),$$

we may begin by applying the calculus of  $\Psi$ DOs (see in particular [17, Theorem 18.1.17]) as well a generalization of [13, Lemma 3.1] to obtain the formula

$$p_k(z, y, D_y) \varphi_{2,\gamma}(z, z', y) = e^{-iU_\gamma(z, z')} \Phi_{2,\gamma}^*(z, z', \cdot) [g_\gamma(z, z', \cdot)](y)$$

where

$$\begin{aligned} g_\gamma(z, z', x) &= \left( p_k(z, \Phi_{2,\gamma}^{-1}(z, z', x), d\Phi_{2,\gamma}^T(z, z', \Phi_{2,\gamma}^{-1}(z, z', x)) D) \right. \\ &\quad \left. - \frac{1}{2} \text{tr} \left( {}_k W_3 {}_k W_1^{-1} \partial_{\xi\xi}^2 p_k(z, y_k, \nu_k) f_\gamma \right) + a_k(z, z', x, D) \right) f_\gamma \end{aligned}$$

and the  $a_k(z, z', x, \eta)$  result from the remainder terms of the  $\Psi$ DO calculus and the application of the lemma. The symbols  $a_k$  are such that both

$$\{2^{k/2} a_k(z, z', x, D) f_\gamma(z, z', \cdot), \gamma\}_{\gamma \in \Gamma_0} \quad \text{and} \quad \{2^{-k/2} \partial_z a_k(z, z', x, D) f_\gamma(z, z', \cdot), \gamma\}_{\gamma \in \Gamma_0}$$

are families of curvelet like functions. This follows from analysis of these remainders and size estimates of  ${}_k W_1$  and its derivatives. By lemmas A.8 and A.4 if

$$\begin{aligned} (4.8) \quad \tilde{g}_\gamma &= i 2^{k/2} e^{-iU_\gamma(z, z')} (\partial_z \tilde{S}_\gamma(z, z', \Phi_{2,\gamma}^{-1}(z, z', x), D) f_\gamma - g_\gamma) \\ &= i 2^{k/2} e^{-iU_\gamma(z, z')} a_k(z, z', x, D) f_\gamma \end{aligned}$$

then

$$(4.9) \quad \left\{ \Phi_{2,\gamma}^*(z, z', \cdot) \tilde{g}_\gamma(z, z', \cdot), \Phi_{2,\gamma}^*(z, z', \cdot)(\gamma) \right\}_{\gamma \in \Gamma_0}$$

is a FCLF. Combining all the previous calculations we see that  $2^{k/2} (\partial_z - i p_k(z, y, D_y)) \varphi_{2,\gamma}(z, z', y) = [\Phi_{2,\gamma}^*(z, z', \cdot) \tilde{g}_\gamma(z, z', \cdot)](y)$ , and we finally conclude that

$$(4.10) \quad \left\{ 2^{k/2} (\partial_z - i p_k(z, y, D_y)) \varphi_{2,\gamma}(z, z', y), \Phi_{2,\gamma}^*(z, z', \cdot)(\gamma) \right\}_{\gamma \in \Gamma_0}$$

is a FCLF. The first statement of the theorem now follows as in previous proofs.

To prove (4.7) we combine the result already obtained for the continuity of  $\mathbf{K}_2(z, z')$  with the following estimate which we will show holds for  $-3/2 \leq s < 2$ .

$$(4.11) \quad \|\mathbf{K}_2(z, z') - \mathbf{K}_2(\bar{z}, z')\|_{(H^{s+1/2}, H^s)} \leq \frac{(C_2^H(Z))^2}{2} |z - \bar{z}|$$

for a constant  $C_2^H(Z) > 0$ . Indeed, if we establish (4.11) then (4.7) follows by interpolation and the triangle inequality. The proof of (4.11) is the same as the proof of theorem 4.1. First we split  $\mathbf{K}_2(z, z') - \mathbf{K}_2(\bar{z}, z')$  into (4.2), (4.3), (4.4) with  $\mathbf{T}_1^k$  replaced by  $\mathbf{T}_2^k$  wherever it appears. The estimates for the two rough parts, (4.3) and (4.4), follow just as before except that now  $m = 2$  and 1 in the respective applications of lemma B.1. Finally, we analyze the part corresponding to (4.2).

The object is to show that

$$(4.12) \quad \left\{ 2^{-k/2} \partial_z (\partial_z - i p_k(z, y, D)) \varphi_{2,\gamma}(z, z', y), \Phi_{2,\gamma}^*(z, z', \cdot)(\gamma) \right\}_{\gamma \in \Gamma_0}.$$

is a FCLF from which (4.11) follows as in the previous proofs. To do this, we calculate using the results from above

$$\begin{aligned} 2^{-k/2} \partial_z (\partial_z - i p_k(z, y, D)) \varphi_{2,\gamma}(z, z', y) &= 2^{-k} \left[ (\partial_{\bar{z}} \Phi_{2,\gamma}^*(\bar{z}, z', \cdot)) \tilde{g}_\gamma(z, z', \cdot) \Big|_{\bar{z}=z} \right. \\ &\quad \left. + \Phi_{2,\gamma}^*(z, z', \cdot) (\partial_z \tilde{g}_\gamma(z, z', \cdot)) \right](y). \end{aligned}$$

The first term on the right hand side gives a FCLF by lemma A.7, and we can see that the second term also gives a FCLF by analyzing the derivative of the second line of (4.8). This completes the proof of (4.12) and also the proof of the Hölder regularity in  $z$ . To prove regularity in  $z'$  we begin with the same expression as (4.12) with the first  $\partial_z$  replaced by  $\partial_{z'}$  and apply a similar analysis.  $\square$

**4.2. Estimates of the iterated Volterra kernel and solution.** Now that we have established our central technical results in the previous two sections, we apply them to the solution of the Volterra equation, and by extension the full solution of (1.1). Assume that we have a Volterra kernel  $\mathbf{K}(z, z')$  with the following properties. There exist  $r, s, \alpha, C(Z), C^H(Z) \in \mathbb{R}$  with  $r \geq 0, 1 \geq \alpha > 0$ , and  $C(Z), C^H(Z) > 0$  such that

$$(4.13) \quad \mathbf{K}(z, z') : H^s \rightarrow H^s \quad \text{uniformly for } z, z' \in [0, Z] \text{ with constant } C(Z),$$

$$(4.14) \quad \|\mathbf{K}(z, z') - \mathbf{K}(\bar{z}, \bar{z}')\|_{(H^s, H^{s-r})} \leq C^H(Z)(|z - \bar{z}|^\alpha + |z' - \bar{z}'|^\alpha)$$

for all  $z, z', \bar{z}$ , and  $\bar{z}' \in [0, Z]$ . Note that the Volterra kernel associated to any parametrix satisfies (4.13),  $\mathbf{K}_1$  from section 3.1 satisfies (4.14) with certain values of  $s, r = 1$ , and  $\alpha = 1$ , and  $\mathbf{K}_2$  from section 3.2 satisfies (4.14) with certain values of  $s, r = 0$ , and  $\alpha = 1/2$ . Thus all the estimates of this section applied to either  $\mathbf{K}_1$  or  $\mathbf{K}_2$  may be considered as corollaries of theorems 4.1 and 4.2.

We first consider the iterated Volterra kernel  $\mathbf{K}^n$  given by (3.1). The following estimate is proven by applying (4.14) to the definition of the iterated kernel and using an inductive argument.

$$(4.15) \quad \|\mathbf{K}^n(z) - \mathbf{K}^n(\bar{z})\|_{(H^s, H^{s-r})} \leq \frac{Z^{n-1}}{(n-1)!} C(Z)^{n-1} C^H(Z) |z - \bar{z}|^\alpha.$$

We next consider the resolvent  $\mathbf{R}(z)$  defined in section 2 corresponding to  $\mathbf{K}(z, z')$ . By summing up (4.15) we obtain the following.

$$(4.16) \quad \|\mathbf{R}(z) - \mathbf{R}(\bar{z})\|_{(H^s, H^{s-r})} \leq e^{ZC(Z)} C^H(Z) |z - \bar{z}|^\alpha,$$

Now, for  $u_0 \in H^r$ , let  $g(z, x) = [\mathbf{R}(z)u_0](x)$  be the solution of the Volterra equation (2.5). A straightforward application of (4.16) then immediately implies that

$$(4.17) \quad \|g(z, \cdot) - g(\bar{z}, \cdot)\|_{H^{s-r}} \leq e^{ZC(Z)} C^H(Z) |z - \bar{z}|^\alpha \|u_0\|_{H^s}.$$

This, together with the comments above, shows that if  $0 \leq s < 2$  and  $g_1(z, x)$  is the solution of (2.5) using  $\mathbf{K}_1$  with initial data  $u_0 \in H^s$ , then  $g_1 \in C^{0,1}([0, Z]; H^{s-1})$ . If  $-1 \leq s < 2$  and  $g_2(z, x)$  is the solution of (2.5) using  $\mathbf{K}_2$  with initial data  $u_0 \in H^s$ , then  $g_2 \in C^{0,1/2}([0, Z]; H^s)$ .

**5. Approximation by semi-discretization.** In this section we discretize the Volterra equation (2.5) with respect to the  $z$  variable. To accomplish this we use the repeated trapezoid rule to approximate the integral, and it is here that the regularity estimates from the previous section will play a key role. Using these estimates we have certain error bounds for the quadrature scheme which allow us to prove in turn convergence of a resulting approximation of the solution to the Volterra equation.

**5.1. Quadrature scheme.** To produce a numeric algorithm to solve the Volterra equation (2.5) we first introduce a quadrature scheme for the integration involved there. Given that the Volterra kernel and solution have Hölder regularity and in general no better, a natural choice of scheme is the trapezoid rule defined as follows.

For every  $N \in \mathbb{N}$  we introduce a partition  $\mathcal{P}^N = \{z_0^N, z_1^N, \dots, z_N^N\}$  of the interval  $[0, Z]$ . That is, we chose the  $z_i^N$  so that  $0 = z_0^N < z_1^N < \dots < z_N^N = Z$ . We then define  $h_i = z_i^N - z_{i-1}^N$ , and the weights

$$(5.1) \quad w_{ij}^N = \begin{cases} \frac{h_j^N + h_{j+1}^N}{2} & \text{if } 0 < j < i \leq N, \\ \frac{h_j^N}{2} & \text{if } j = i > 0, \\ \frac{h_1^N}{2} & \text{if } j = 0 \text{ and } i > 0, \end{cases}$$

and otherwise  $w_{ij}^N = 0$ . Also let  $h_N = \sup_{i \in \{1, \dots, N\}} h_i^N$ . If  $\mathcal{B}$  is any Banach space and  $f : [0, Z] \rightarrow \mathcal{B}$  is continuous, then the repeated trapezoid rule is given by

$$(5.2) \quad \int_0^{z_i^N} f(z) \, dz = \sum_{j=0}^i w_{ij}^N f(z_j^N) + E_i^N(f).$$

We are thinking of the sum in the previous expression as an approximation to the integral and  $E_i^N(f)$  as an error term which should approach zero as  $N \rightarrow \infty$ . Indeed, we have in general the following estimates

$$(5.3) \quad \|E_i^N(f)\|_{\mathcal{B}} \leq z_i^N (h_N)^\alpha \mathcal{L}_{[0, z_i]}^\alpha(f)$$

where

$$\mathcal{L}_{[0, i]}^\alpha(f) = \sup_{z, \bar{z} \in [0, z_i^N]: z \neq \bar{z}} \frac{\|f(z) - f(\bar{z})\|_{\mathcal{B}}}{|z - \bar{z}|^\alpha}.$$

The cases that are relevant here are when  $f(z') = \mathbf{K}_m(z_i, z')g(z', \cdot) : [0, z_i] \rightarrow H^s$  where  $m = 1$  or  $2$  and  $s$  is in the allowed range depending on  $m$ .

**5.2. Semi-discrete Volterra equation.** We now introduce the semi-discrete Volterra equation. If  $\mathbf{K}$  is the Volterra kernel associated to a parametrix, as defined in section 2, then we will write  $\mathbf{K}_{ij}^N := \mathbf{K}(z_i^N, z_j^N)$ . Also we define

$$A_{ij}^N := w_{ij}^N \mathbf{K}_{ij}^N,$$

where  $w_{ij}$  are the weights given in (5.1). Thus  $A^N$  is a  $(N+1) \times (N+1)$  matrix with entries in the space of continuous linear operators from  $H^s$  to  $H^s$ . The *semi-discrete Volterra equation* is then

$$(5.4) \quad g_i^N = \mathbf{K}_{i0}^N u_0 + \sum_{j=0}^i A_{ij}^N g_j^N \quad \text{for all } i \in \{0, \dots, N\}.$$

We assume that  $u_0 \in H^s$ . If  $h_N < 2$  then, since  $A_{ii} = h_i \text{Id}/2$  for  $i > 0$ , and  $A_{00} = 0$ , we see that this equation has a unique solution in  $(H^s)^{N+1}$ . This method of approximating  $g$ , using (5.4), is known as *direct quadrature*. The next proposition establishes how well the solution of (5.4) approximates the solution of the Volterra equation.

**PROPOSITION 5.1.** *Suppose that the Volterra kernel  $\mathbf{K}(z, z')$  satisfies (4.13) and (4.14) for some values of the parameters, and for given  $u_0 \in H^s$  let  $g(z, x) \in L^\infty([0, Z], H^s)$  be the solution of (2.5) and  $g^N \in (H^s)^{N+1}$  be the solution of (5.4). Assume also that  $\mathbf{K}(z, z')$  extends to a uniformly continuous map from  $H^{s-r}$  to  $H^{s-r}$  with the same constant  $C(Z)$  and that  $h_N < 1$ . Then*

$$(5.5) \quad \sup_{i \in \{0, \dots, N\}} \|g(z_i, \cdot) - g_i^N(\cdot)\|_{H^{s-r}} \leq 2e^{\left(Z + \frac{2Nh_N}{2-h_N}\right)C(Z)} C^H(Z) h_N^\alpha \|u_0\|_{H^s}.$$

*Proof.* We first note that by (2.7), (4.14), and (4.17)

$$(5.6) \quad \|\mathbf{K}(z_i^N, z')g(z', \cdot) - \mathbf{K}(z_i^N, \bar{z}')g(\bar{z}', \cdot)\|_{H^{s-r}} \leq 2e^{ZC(Z)} C^H(Z) |z' - \bar{z}'|^\alpha \|u_0\|_{H^s}$$

for all  $i$ ,  $z'$ , and  $\bar{z}'$ . We will write  $d_i^N(x) = g(z_i^N, x) - g_i^N(x)$ . Then (2.5) and (5.4) imply

$$\begin{aligned}
 (5.7) \quad d_i^N(x) &= \left( \int_0^{z_i^N} [\mathbf{K}(z_i^N, z')g(z', \cdot)](x) dz' - \sum_{j=0}^i [A_{ij}^N g(z_j^N, \cdot)](x) \right) \\
 &\quad + \sum_{j=0}^i [A_{ij}^N (g(z_j^N, \cdot) - g_j^N(\cdot))](x) \\
 &= E_i^N([\mathbf{K}(z_i, z')g(z', \cdot)]) + \sum_{j=0}^i [A_{ij}^N d_j^N(\cdot)](x).
 \end{aligned}$$

Using now (5.3) with  $\mathcal{B} = H^{s-r}$  and (5.6) we have

$$\begin{aligned}
 \|d_i^N\|_{H^{s-r}} &\leq 2e^{ZC(Z)} C^H(Z) h_N^\alpha \|u_0\|_{H^s} + C(Z) \sum_{j=0}^i w_{ij}^N \|d_j^N\|_{H^{s-r}} \\
 &\leq 2 \frac{e^{ZC(Z)} C^H(Z) h_N^\alpha \|u_0\|_{H^s}}{1 - h_N/2} + \frac{C(Z) h_N}{1 - h_N/2} \sum_{j=0}^{i-1} \|d_j^N\|_{H^{s-r}}
 \end{aligned}$$

A discrete Gronwall-type inequality (see [6, section 1.5.3]) now implies that

$$\|d_i^N\|_{H^{s-r}} \leq 4e^{ZC(Z)} C^H(Z) h_N^\alpha \|u_0\|_{H^s} e^{\frac{N h_N C(Z)}{1 - h_N/2}}.$$

This completes the proof.

□

If  $h_N \approx 1/N$ , as would be the case if the partition  $\mathcal{P}$  is evenly spaced, then proposition 5.1 implies that  $\|g(z_i, \cdot) - g_i^N(\cdot)\|_{H^{r-s}} \approx (1/N)^\alpha$  as  $N \rightarrow \infty$ . The proposition also reveals the key difference between the parametrices  $\mathbf{T}_1$  and  $\mathbf{T}_2$  and corresponding Volterra kernels  $\mathbf{K}_1$  and  $\mathbf{K}_2$ . For  $\mathbf{K}_1$  we only have the Hölder estimates (4.14) in the case where  $r > 0$ , and so we can only estimate the error incurred as a result of the discretization in a norm which is rougher than that of the space where the initial data  $u_0$  lies. However, for  $\mathbf{K}_2$  we can take  $r = 0$  and obtain an error estimate with respect to the original norm.

**6. Concentration of sets of wave packets.** In this section we assume that the initial data,  $u_0$ , has a representation in the curvelet frame that is concentrated near a finite set of curvelet indices  $\Gamma_0$ , and then study how much the curvelet coefficients for the solution of (1.1) spread away from  $\Psi_{z,0}(\Gamma_0)$  as  $z$  increases. The motivation for this study is to apply the results to prove convergence of a numeric scheme to solve (1.1) using only a finite set of curvelets.

Following [3], we first introduce the following weighted spaces.

**DEFINITION 6.1.** *Let  $\Gamma_0$  be a finite set of curvelet indices. We define the space  $H_{\Gamma_0}^{\sigma,\alpha}$  by the norm*

$$\|f\|_{H_{\Gamma_0}^{\sigma,\alpha}}^2 = |f_0|^2 + \sum_{\gamma} |2^{k\sigma} \min_{\gamma_0 \in \Gamma_0} \left\{ (2^{\max(k,k_0)} \bar{d}(\gamma; \gamma_0))^\alpha \right\} f_\gamma|^2,$$

where  $\gamma = (x, \nu, k)$ , and  $f_\gamma$  are the coefficients of  $f$  with respect to the curvelet frame.

In this definition,  $\sigma$  corresponds to the Sobolev space regularity of  $f$  while  $\alpha$  gives the degree to which the curvelet coefficients of  $f$  are concentrated near  $\Gamma_0$ . A useful estimate is the following

$$(6.1) \quad \|f\|_{H_{\Gamma_0}^{\sigma,\alpha}}^2 \approx |f_0|^2 + \min_{\gamma_0 \in \Gamma_0} \sum_{\gamma} |2^{k\sigma} (2^{\max(k,k_0)} \bar{d}(\gamma; \gamma_0))^\alpha f_\gamma|^2.$$

The constants relating the two sides can be found based on the “radius” of the set  $\Gamma_0$  (under a proper interpretation of the radius).

Estimates in terms of the  $H_{\Gamma_0}^{\sigma, \alpha}$  norm allow us to easily estimate how well a given function is approximated by a finite set of curvelets. Indeed, given a finite set of curvelet indices  $\Gamma_0 \subset \Gamma$  let  $\Gamma_0^r$  be the set of all indices  $\gamma$  that are indices of curvelets in the curvelet frame, and such that

$$\min_{\gamma_0 \in \Gamma_0} 2^{\max(k, k_0)} \bar{d}(\gamma; \gamma_0) \leq r.$$

Then define  $\Delta_{\Gamma_0}^r$  to be the operator given

$$\Delta_{\Gamma_0}^r f = f_0 \varphi_0 + \sum_{\gamma \in \Gamma_0^r} f_\gamma \varphi_\gamma.$$

If  $f \in H_{\Gamma_0}^{\sigma, \alpha}$  it is then straight forward to check that  $f \in H^\sigma$  and

$$(6.2) \quad \|f - \Delta_{\Gamma_0}^r f\|_{H^\sigma} \leq \min(r^{-\alpha}, 1) \|f\|_{H_{\Gamma_0}^{\sigma, \alpha}}.$$

This inequality will be useful below when we estimate the error incurred by solving the Volterra equation with only a finite number of curvelets. However, to accomplish this goal we will first require the following lemma.

LEMMA 6.2. *Let  $0 \leq \alpha < \frac{m+1}{2}$ ,  $|\sigma| \leq \frac{m-1}{2}$  with  $m = 1$  ( $p \in C^{1,1}S_{\text{cl}}^1$ ) or  $2$  ( $p \in C^{2,1}S_{\text{cl}}^1$ ). It holds true that*

$$(6.3) \quad \|\mathbf{K}_m(z, z')\|_{(H_{\Psi_{z',0}(r_0)}^{\sigma, \alpha}, H_{\Psi_{z,0}(r_0)}^{\sigma, \alpha})} \leq C_K(Z)$$

uniformly in  $z, z' \in [0, Z]$ .

*Proof.* Let  $c_{m, \gamma \gamma'}(z, z')$  and  $\tilde{c}_{m, \gamma \gamma'}(z, z')$  be respectively matrices of the operators

$$\sum_k (\partial_z - i p_k(z, y, D_y)) \mathbf{T}_m^k(z, z') \quad \text{and} \quad \mathbf{T}_m(z, z')$$

with respect to the curvelet frame. Then by results in sections 3 and 4 as well as [19, Lemma 2.2] we have the estimates

$$c_{m, \gamma \gamma'}(z, z') \lesssim 2^{k(1-m)/2} \mu_\delta(\gamma, \Psi_{z, z'}(\gamma')), \quad \text{and} \quad \tilde{c}_{m, \gamma \gamma'}(z, z') \lesssim \mu_\delta(\gamma, \Psi_{z, z'}(\gamma'))$$

for any  $\delta > 0$ . Also, by [3, Theorem 5.5] (or, more accurately, using a portion of the proof of that theorem) and (6.1) we have the estimates

$$\begin{aligned} \|(p_k(z, y, D_y) - p(z, y, D_y)) \varphi_\gamma\|_{H_{\Psi_{z,0}(\Gamma_0)}^{\sigma, \alpha}} &\lesssim \|\varphi_\gamma\|_{H_{\Psi_{z,0}(\Gamma_0)}^{\sigma, \alpha}} \\ &\lesssim 2^{k\sigma} \min_{\gamma_0 \in \Psi_{z,0}(\Gamma_0)} \left\{ (2^{\max(k, k_0)} \bar{d}(\gamma; \gamma_0))^\alpha \right\} \end{aligned}$$

for  $\alpha$  and  $\sigma$  within the ranges specified in the hypotheses and where  $\varphi_\gamma$  is a curvelet at scale  $k$ . Making the same decomposition as in the proof of theorem 4.2 we have

$$\begin{aligned} \|\mathbf{K}_m(z, z') f\|_{H_{\Psi_{z,0}(\Gamma_0)}^{\sigma, \alpha}}^2 &\lesssim \\ &\left( \sum_\gamma \sum_{\gamma'} 2^{k\sigma} \min_{\gamma_0 \in \Psi_{z,0}(\Gamma_0)} \left\{ (2^{\max(k, k_0)} \bar{d}(\gamma; \gamma_0))^\alpha \right\} c_{m, \gamma \gamma'}(z, z') |f_{\gamma'}| \right. \\ &\quad \left. + \sum_\gamma \sum_{\gamma'; k'=k} 2^{k\sigma} \min_{\gamma_0 \in \Psi_{z,0}(\Gamma_0)} \left\{ (2^{\max(k, k_0)} \bar{d}(\gamma; \gamma_0))^\alpha \right\} \tilde{c}_{m, \gamma \gamma'}(z, z') |f_{\gamma'}| \right)^2. \end{aligned}$$



Now we apply the estimate

$$\begin{aligned} (2^{\max(k,k_0)}\bar{d}(\gamma; \Psi_{z,0}(\gamma_0)))^\alpha &\lesssim (2^{\max(k,k')} \bar{d}(\gamma; \Psi_{z,z'}(\gamma'))^\alpha \\ &\quad \cdot (2^{\max(k',k_0)} \bar{d}(\Psi_{z,0}(\gamma_0); \Psi_{z,z'}(\gamma'))^\alpha \end{aligned}$$

which together with the bounds from above on the matrix coefficients  $c_{m,\gamma\gamma'}$  and  $\tilde{c}_{m,\gamma\gamma'}$  gives for any  $\delta > 0$

$$\begin{aligned} \|\mathbf{K}_m(z, z')f\|_{H_{\Psi_{z,0}(\Gamma_0)}^{\sigma,\alpha}}^2 &\lesssim \\ &\min_{\gamma_0 \in \Psi_{z,0}(\Gamma_0)} \sum_{\gamma'} \left( \sum_{\gamma} \mu_\delta(\gamma_0, \Psi_{z,z'}(\gamma')) \left| 2^{k'\sigma} (2^{\max(k',k_0)} \bar{d}(\gamma_0; \Psi_{z,z'}(\gamma'))^\alpha f_{\gamma'} \right| \right)^2. \end{aligned}$$

Finally, [19, Lemmas 2.1, 2.2, 2.4] imply with the last inequality and (6.1) that

$$\|\mathbf{K}_m(z, z')f\|_{H_{\Psi_{z,0}(\Gamma_0)}^{\sigma,\alpha}}^2 \lesssim \|f\|_{H_{\Psi_{z,0}(\Gamma_0)}^{\sigma,\alpha}}^2.$$

This completes the proof.  $\square$

The lemma also yields estimates for the resolvents:

$$(6.4) \quad \|\mathbf{R}_m(z)\|_{(H_{\Gamma_0}^{\sigma,\alpha}, H_{\Psi_{z,0}(\Gamma_0)}^{\sigma,\alpha})} \leq C_{R,m}(Z)$$

uniformly for  $z \in [0, Z]$ .

With the previous result we may now prove an error estimate that relates the solution of the fully discrete Volterra equation (ie. the semi-discrete equation from the previous section truncated to a finite set of curvelets) to the true Volterra solution. We begin by modifying the semi-discrete Volterra equation to become fully discrete. Using the same notation as in section 5.2, for any given  $r > 0$  and finite set of indices  $\Gamma_0$  let

$$\tilde{A}_{ij}^N = w_{ij} \Delta_{\Psi_{z_i^N,0}(\Gamma_0)}^r \mathbf{K}_{ij}$$

The *fully discrete Volterra equation* is then (compare with (5.4))

$$(6.5) \quad \tilde{g}_i^N = \Delta_{\Psi_{z_i^N,0}(\Gamma_0)}^r \mathbf{K}_{i0} u_0 + \sum_{j=0}^i \tilde{A}_{ij} \tilde{g}_j^N.$$

Note that for every  $i$  and  $N$  the solution  $\tilde{g}_i^N$  of (6.5) is a linear combination of the curvelets corresponding to the indices  $\Psi_{z_i^N,0}(\Gamma_0)^r$ . Now we present the result, which is a sort of extension of proposition 5.1.

**COROLLARY 6.3.** *Let  $u_0$  be a linear combination of curvelets with indices in the finite set  $\Gamma_0$ . Suppose that  $g(z, x)$  is the solution of (2.5) corresponding to  $\mathbf{K}_2$ , and that  $\tilde{g}^N$  is the corresponding solution of (6.5). Then for any  $|s| \leq 1/2$  and  $0 \leq \alpha < 3/2$  we have the estimate*

$$\sup_{i \in \{0, \dots, N\}} \|g(z_i^N, \cdot) - \tilde{g}_i^N\|_{H^s} \leq C(Z) (h_N^{1/2} + \min(r^{-\alpha}, 1)) \|u_0\|_{H^s}$$

*Proof.* The proof is largely the same as the proof of proposition 5.1. The primary difference is that in (5.7)  $\tilde{d}_i^N(x) = g(z_i^N, x) - \tilde{g}_i^N(x)$  replaces  $d_i^N(x)$ ,  $\tilde{A}_{ij}^N$  replaces  $A_{ij}^N$  in the sum on the second line, and there appear the extra terms

$$\left\| \sum_{j=0}^i [A_{ij}^N - \tilde{A}_{ij}^N] g(z_j^N, \cdot) \right\|_{H^s} \leq \min(r^{-\alpha}, 1) Z C_2(Z) C_{R,2}(Z) \|u_0\|_{H^s}$$

and

$$\|\mathbf{K}_{i0} u_0 - \Delta_{\Psi_{z_i^N, 0}(\Gamma_0)}^r \mathbf{K}_{i0} u_0\|_{H^s} \leq \min(r^{-\alpha}, 1) C_2(Z) \|u_0\|_{H^s}.$$

These estimates use the result of lemma 6.2, (6.2), and (6.4) as well as the continuity of  $\mathbf{K}_2$ . Inserting these into the proof of proposition 5.1 yields the proof of the corollary.  $\square$

This last corollary establishes the possibility of approximating the solution of the Volterra equation using only curvelets that lie within a certain distance of the Hamiltonian flow corresponding to the finite number of initial curvelets. We note additionally that the estimates lend themselves well to a “step-by-step” approach to solving the fully discrete Volterra equation (6.5). Given a choice of step size in the quadrature, and a choice for  $r$ , at each time step we compute only those curvelet coefficients corresponding to indices in  $\Psi_{z_i^N, 0}(\Gamma_0)^r$ . This means, loosely speaking, that we only consider those curvelets lying within  $r$  of the original curvelets flowed forward to time  $z_i$ . Further, we point out that the approximate solution of (1.1) provided by the fully discrete Volterra equation will consist of a sum of terms each being a composition of some number of operators of the form  $\mathbf{K}_2(z_i^N, z_j^N)$  or  $\mathbf{T}_2(z_i^N, z_j^N)$ . The computation of these compositions, reserved for future work, will be accomplished via a separated representation similar to that used in the proofs of theorems 4.1 and 4.2.

### Appendix A. Curvelet like functions.

In this appendix we develop some technical machinery which we use to analyze the various operators defined in terms of the curvelet frame. In the main text, we use a curvelet frame based on parabolic scaling as defined, for example, in either [3] or [19]. Our notation for curvelets and the curvelet frame matches that of [3]. In particular, we use the notation  $\Gamma = \mathbb{R}^n \times \mathbb{S}^{n-1} \times \mathbb{R}$  and refer to  $\Gamma$  as the set of “curvelet indices.” Also  $d$  is the pseudodistance on  $\mathbb{R}^n \times \mathbb{S}^{n-1}$  introduced in [19, Definition 2.1]

$$(A.1) \quad d(x, \nu; x', \nu') = |\langle \nu, x - x' \rangle| + |\langle \nu', x - x' \rangle| + \min\{|x - x'|, |x - x'|^2\} + |\nu - \nu'|^2.$$

If  $\gamma = (x, \nu, k)$  and  $\gamma' = (x', \nu', k') \in \Gamma$ , let

$$(A.2) \quad \bar{d}(\gamma; \gamma') = 2^{-\min(k, k')} + d(x, \nu; x', \nu').$$

The weight function  $\mu_\delta(\gamma, \gamma')$  is given by

$$\mu_\delta(\gamma, \gamma') = (1 + |k' - k|^2)^{-1} 2^{-(\frac{1}{2}n + \delta)|k' - k|} 2^{-(n + \delta)\min(k', k)} \bar{d}(\gamma; \gamma')^{-(n + \delta)}.$$

This weight function is different from, but equivalent to that introduced in [18]. We also use both notations  $\hat{f}$  and  $\mathcal{F}\{f\}$  for the Fourier transform of  $f$  depending on the aesthetic demands of the individual situation.

We now begin to introduce more general classes of functions that behave in many ways like those which make up the curvelet frame. For  $k \in \mathbb{R}$  we will denote by  $C_k$  the cylinder

$$(A.3) \quad C_k = [2^{k-1}, 2^{k+1}] \times \mathbb{B}_{2^{k/2}}^{n-1} \subset \mathbb{R}^n.$$

where  $\mathbb{B}_{2^{k/2}}^{n-1}$  is the  $n - 1$  dimensional ball of radius  $2^{k/2}$  centered at the origin. The term “dyadic parabolic scaling” refers to the relative proportions of these cylinders which scale like  $2^k$  in the direction of  $e_1$ , and  $2^{k/2}$  in the perpendicular directions. Given  $\nu \in \mathbb{S}^{n-1}$  let  $\Theta_\nu \in O(n)$  represent any rotation that maps  $e_1$  into  $\nu$ , and define

$$C_{\nu, k} = \Theta_\nu C_k.$$

Naturally  $C_{\nu, k}$  is independent of the specific rotation that is chosen. Also, we write  $\rho_k = |C_k| \sim 2^{k(n+1)/2}$ . The families of functions are now defined as follows.

DEFINITION A.1. A subset  $\mathcal{F} \subset \mathcal{S}(\mathbb{R}^n) \times \Gamma$  is a **family of curvelet like functions (FCLF)** if the following conditions are met.

1. For every  $j \in \mathbb{N}$ ,  $\alpha \in \mathbb{N}^n$ , and  $N \in \mathbb{N}$ , there exists a constant  $C_{j,\alpha,N} > 0$  such that the following estimates hold for all  $(f, (x, \nu, k)) \in \mathcal{F}$

$$\rho_k^{1/2} |\langle \nu, \partial_\xi \rangle^j \partial_\xi^\alpha (e^{i\langle x, \xi \rangle} \widehat{f}(\xi))| \leq C_{j,\alpha,N} 2^{-k(j + \frac{|\alpha|}{2})} (1 + 2^{-k/2} \|\xi - C_{\nu,k}\|)^{-N}.$$

2. There exists a constant  $C \in \mathbb{R}$  (possibly less than zero) such that for all  $(f, (x, \nu, k)) \in \mathcal{F}$ ,  $k \geq C$ .

When we have a family of curvelet like functions,  $\mathcal{F}$ , we use the notation  $\pi_\Gamma : \mathcal{F} \rightarrow \Gamma$  for the map projecting  $\mathcal{F}$  onto the set of curvelet indices, and  $\pi_x, \pi_\nu, \pi_k$  for the respective projections onto components of the curvelet indices. Also,  $\pi_S : \mathcal{F} \rightarrow \mathcal{S}(\mathbb{R}^n)$  is the projection onto  $\mathcal{S}(\mathbb{R}^n)$ . When there is only one family of curvelet like functions we will usually write  $\gamma = (x, \nu, k)$  for the curvelet index of arbitrary functions in the family.

Though we will not formulate the result precisely, we also remark that families of curvelet like functions may be added together (ie. adding functions corresponding to the same index) to produce a new family of curvelet like functions. This fact, which follows immediately from the definition, will be used extensively.

Curvelet frames with parabolic scaling give families of curvelet like functions if we remove the element of the frame whose Fourier transform covers the origin (ie. the zero frequency curvelet). The motivation for considering these families is that they are more or less preserved under most of the operations that we would like to perform on curvelets. In the following series of lemmas we will show precisely what this means, and in essence establish a calculus for families of curvelet like functions.

LEMMA A.2. Suppose that  $\mathcal{F}$  and  $\mathcal{G}$  are two families of curvelet like functions. Then for every  $\delta > 0$  there exists a constant  $C_\delta$  such that for every  $f \in \mathcal{F}$  and  $g \in \mathcal{G}$

$$\left| \langle \pi_S(f), \pi_S(g) \rangle \right| \leq C_\delta \mu_\delta(\pi_\Gamma(f), \pi_\Gamma(g)).$$

*Proof.* We first prove the result for the case when one of the families is given by a curvelet frame. Then we represent the functions in each of two families with respect to this curvelet frame and apply a slight generalization of [18, lemma 2.5] to the case when the  $\gamma$  and  $\gamma_0$  need not be in the grid corresponding to the curvelet frame.  $\square$

We next study what happens when we take derivatives of curvelet like functions.

LEMMA A.3. Suppose that  $\mathcal{F}$  is a family of curvelet like functions. Then

$$\left\{ (2^{-k} \langle \nu, \partial_y \rangle f, \gamma) \right\}_{(f,\gamma) \in \mathcal{F}} \quad \text{and} \quad \left\{ (2^k \langle \nu, y - x \rangle f, \gamma) \right\}_{(f,\gamma) \in \mathcal{F}}$$

are also families of curvelet like functions. Furthermore, if we are given a map  $\nu_\perp : \mathcal{F} \rightarrow \mathbb{S}^{n-1}$  such that  $\langle \nu_\perp(f, \gamma), \nu \rangle = 0$  for every  $(f, \gamma) \in \mathcal{F}$ , then

$$\left\{ (2^{-k/2} \langle \nu_\perp(f, \gamma), \partial_y \rangle f, \gamma) \right\}_{(f,\gamma) \in \mathcal{F}}$$

and

$$\left\{ (2^{k/2} \langle \nu_\perp(f, \gamma), y - x \rangle f, \gamma) \right\}_{(f,\gamma) \in \mathcal{F}}$$

are both families of curvelet like functions.

*Proof.* For  $(f, \gamma) \in \mathcal{F}$  we have

$$2^{-k} \mathcal{F} \left\{ \langle \nu, \partial_y \rangle f \right\} (\xi) = i 2^{-k} \langle \nu, \xi \rangle \widehat{f}(\xi).$$

Combined with the inequality

$$2^{-k} |\langle \nu, \xi \rangle| \leq 2(1 + 2^{-k/2} \|\xi - C_{\nu, k}\|)$$

this gives the first assertion of the lemma.

Next we have

$$\mathcal{F} \left\{ \langle \nu, y - x \rangle f \right\} = \langle \nu, D \widehat{f} \rangle,$$

which easily implies the second assertion.

The third and fourth assertions follow in the same way if we use also the inequality

$$2^{-k/2} |\langle \nu_{\perp}(f, \gamma), \xi \rangle| \leq (1 + 2^{-k/2} \|\xi - C_{\nu, k}\|)$$

which holds for any  $\nu_{\perp}$  satisfying the hypotheses.  $\square$

We next study how curvelet like functions change under pull-back by a change of coordinates. Suppose that  $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$  is a diffeomorphism. First recall that the pull-back of a function  $f \in \mathcal{S}(\mathbb{R}^n)$  is given by the composition  $\Phi^*(f)(x) = f(\Phi(x))$ . We define the pull-back of a curvelet index  $\gamma = (x, \nu, k)$  by

$$\Phi^*(\gamma) = \left( \Phi^{-1}(x), (d\Phi^T(x)\nu)/|d\Phi^T(x)\nu|, k + \log_2(|d\Phi^T(x)\nu|) \right).$$

Note that since  $\Phi$  is a diffeomorphism, the map  $\Phi^* : \Gamma \rightarrow \Gamma$  is invertible.

LEMMA A.4. *Suppose that  $\mathcal{F}$  is a family of curvelet like functions, and that  $\{\Phi_{\gamma}\}_{\gamma \in \pi_{\Gamma}(\mathcal{F})}$  is a family of diffeomorphisms on  $\mathbb{R}^n$  satisfying*

$$\|\partial_x^{\alpha} \Phi_{\gamma}^{-1}\| \leq C_{\alpha} \quad \text{for } 0 < |\alpha| \leq 2, \quad \text{and} \quad \|\partial_x^{\alpha} \Phi_{\gamma}^{-1}\| \leq 2^{k(|\alpha|-2)/2} C_{\alpha} \quad \text{for } 2 < |\alpha|.$$

Then

$$\left\{ (\Phi_{\gamma}^*(f), \Phi_{\gamma}^*(\gamma)) \right\}_{(f, \gamma) \in \mathcal{F}}$$

is a family of curvelet like functions. Note that  $\Phi_{\gamma}^*(f)$  is the pull-back of the function  $f$ , while  $\Phi_{\gamma}^*(\gamma)$  is the pull-back of the curvelet index  $\gamma$ .

*Proof.* Let  $(f, \gamma) \in \mathcal{F}$ . By the Fourier inversion formula we have the following formula

$$\begin{aligned} \rho_{k'}^{1/2} \langle \nu', \partial_{\xi} \rangle^j \partial_{\xi}^{\alpha} \left( e^{i\langle \Phi_{\gamma}^{-1}(x), \xi \rangle} \widehat{\Phi_{\gamma}^* f}(\xi) \right) &= \frac{i^{j+|\alpha|} \rho_{k'}^{1/2}}{(2\pi)^n \rho_k^{1/2}} \iint e^{i(\langle \Phi_{\gamma}(y-x), \eta \rangle - \langle y - \Phi_{\gamma}^{-1}(x), \xi \rangle)} \\ &\quad \cdot \langle \nu', \Phi_{\gamma}^{-1}(x) - y \rangle^j \left( \Phi_{\gamma}^{-1}(x) - y \right)^{\alpha} \left( \rho_k^{1/2} e^{i\langle x, \eta \rangle} \widehat{f}(\eta) \right) d\eta dy. \end{aligned}$$

where  $k' \in \mathbb{R}$  and  $\nu' \in \mathbb{S}^{n-1}$  are the respective components of the pull-back  $\Phi_{\gamma}^*(\gamma)$ . Note that this should be interpreted as an iterated integral with the integration done first in  $\eta$  and then in  $y$ . Making the change  $\tilde{x} = \Phi_{\gamma}(y) - x$  in the second integral gives

$$\begin{aligned} &\rho_{k'}^{1/2} \langle \nu', \partial_{\xi} \rangle^j \partial_{\xi}^{\alpha} \left( e^{i\langle \Phi_{\gamma}^{-1}(x), \xi \rangle} \widehat{\Phi_{\gamma}^* f}(\xi) \right) \\ &= \frac{i^{j+|\alpha|} \rho_{k'}^{1/2}}{(2\pi)^n \rho_k^{1/2}} \iint e^{i(\langle \tilde{x}, \eta \rangle - \langle \Phi_{\gamma}^{-1}(\tilde{x}+x) - \Phi_{\gamma}^{-1}(x), \xi \rangle)} \langle \nu', \Phi_{\gamma}^{-1}(x) - \Phi_{\gamma}^{-1}(\tilde{x}+x) \rangle^j \\ &\quad \cdot \left( \Phi_{\gamma}^{-1}(x) - \Phi_{\gamma}^{-1}(\tilde{x}+x) \right)^{\alpha} \left( \rho_k^{1/2} e^{i\langle x, \eta \rangle} \widehat{f}(\eta) \right) \frac{d\eta d\tilde{x}}{|\det(d\Phi_{\gamma}(\tilde{x}+x))|}. \end{aligned}$$

By Taylor's theorem we may write

$$\Phi_\gamma^{-1}(\tilde{x} + x)^r - \Phi_\gamma^{-1}(x)^r = d\Phi_\gamma^{-1}(x)_p^r \tilde{x}^p + \tilde{\Psi}_\gamma(\tilde{x}, x)_{pq}^r \tilde{x}^q \tilde{x}^p$$

using the summation convention. Here  $\tilde{\Psi}_\gamma$  is a smooth array of functions that can all be simultaneously bounded in  $C^l$  in terms of bounds on the derivatives of  $\Phi_\gamma^{-1}$  up to order  $l + 2$ . We will also write  $\tilde{\Phi}(\tilde{x}, x)_p^r = d\Phi_\gamma^{-1}(x)_p^r + \tilde{\Psi}_\gamma(\tilde{x}, x)_{pq}^r \tilde{x}^q$ . With this notation

$$\begin{aligned} & \rho_{k'}^{1/2} \langle \nu', \partial_\xi \rangle^j \partial_\xi^\alpha \left( e^{i\langle \Phi_\gamma^{-1}(x), \xi \rangle} \widehat{\Phi_\gamma^* f}(\xi) \right) \\ &= \frac{(-i)^{j+|\alpha|} \rho_{k'}^{1/2}}{(2\pi)^n \rho_k^{1/2}} \iint e^{i(\langle \tilde{x}, \eta \rangle - \langle \tilde{\Phi}_\gamma(\tilde{x}, x), \tilde{x}, \xi \rangle)} \langle \nu', \tilde{\Phi}_\gamma(\tilde{x}, x) \tilde{x} \rangle^j \\ & \quad \cdot \left( \tilde{\Phi}_\gamma(\tilde{x}, x) \tilde{x} \right)^\alpha \left( \rho_k^{1/2} e^{i\langle x, \eta \rangle} \widehat{f}(\eta) \right) \frac{d\eta d\tilde{x}}{|\det(d\Phi_\gamma(\tilde{x} + x))|} \\ &= \frac{(-i)^j \rho_{k'}^{1/2}}{(2\pi)^n \rho_k^{1/2}} \iint e^{i(\langle \tilde{x}, \eta \rangle - \langle \tilde{\Phi}_\gamma(\tilde{x}, x), \tilde{x}, \xi \rangle)} \left( \langle \nu', d\Phi_\gamma^{-1}(x) \partial_\eta \rangle \right. \\ & \quad \left. + \left( \nu'_r \tilde{\Psi}_\gamma(\tilde{x}, x)_{pq}^r \partial_{\eta_p}^2 \right) \right)^j \left( \tilde{\Phi}_\gamma(\tilde{x}, x) \partial_\eta \right)^\alpha \left( \rho_k^{1/2} e^{i\langle x, \eta \rangle} \widehat{f}(\eta) \right) \frac{d\eta d\tilde{x}}{|\det(d\Phi_\gamma(\tilde{x} + x))|}. \end{aligned}$$

After several more rounds of integration by parts, for any  $M$  and  $\tilde{N}$  this last expression equals

$$\begin{aligned} & \frac{(-i)^j \rho_{k'}^{1/2}}{(2\pi)^n \rho_k^{1/2}} \iint e^{i(\langle \tilde{x}, \tilde{\eta} \rangle - \langle \tilde{\Phi}_\gamma(\tilde{x}, x), \tilde{x}, \xi \rangle)} \left( 1 + i 2^{-k/2} \langle \partial_{\tilde{x}}, \eta - \partial_{\tilde{x}}(\langle \tilde{\Phi}_\gamma(\tilde{x}, x), \tilde{x}, \xi \rangle) \rangle \right)^M \\ & \quad \cdot \frac{\left( \langle \nu', d\Phi_\gamma^{-1}(x) \partial_\eta \rangle + \nu'_r \tilde{\Psi}_\gamma(\tilde{x}, x)_{pq}^r \partial_{\eta_p}^2 \right)^j}{(1 + 2^{-k/2} \|\eta - \partial_{\tilde{x}}(\langle \tilde{\Phi}_\gamma(\tilde{x}, x), \tilde{x}, \xi \rangle)\|^2)^M} \left( \frac{1 - 2^k \Delta_\eta}{1 + 2^k |\tilde{x}|^2} \right)^{\tilde{N}} \left( \tilde{\Phi}_\gamma(\tilde{x}, x) \partial_\eta \right)^\alpha \\ & \quad \cdot \left( \rho_k^{1/2} e^{i\langle x, \eta \rangle} \widehat{f}(\eta) \right) \frac{d\eta d\tilde{x}}{|\det(d\Phi_\gamma(\tilde{x} + x))|}, \end{aligned}$$

which may now be interpreted as an integral over  $\mathbb{R}^{2n}$ . Using definition A.1 and the hypotheses on  $\Phi_\gamma$ , the integrand in the previous formula can be bounded for any  $N$  by an expression of the form

$$\begin{aligned} & C 2^{-k(j + \frac{|\alpha|}{2})} (1 + 2^{-k/2} \|\eta - C_{\nu, k}\|)^{-N} (1 + 2^{-k/2} \|\eta - \partial_{\tilde{x}}(\langle \tilde{\Phi}_\gamma(\tilde{x}, x), \tilde{x}, \xi \rangle)\|)^{-M} \\ & \quad \cdot (1 + 2^k |\tilde{x}|^2)^{-\tilde{N}} \end{aligned}$$

for some positive  $C$ . Therefore, if  $M > N + n$  then

$$\begin{aligned} \left| \rho_{k'}^{1/2} \langle \nu', \partial_\xi \rangle^j \partial_\xi^\alpha \left( e^{i\langle \Phi_\gamma^{-1}(x), \xi \rangle} \widehat{\Phi_\gamma^* f}(\xi) \right) \right| & \lesssim 2^{-k(j + \frac{|\alpha|}{2})} \sup_{\eta, x' \in \mathbb{R}^n} \left\{ (1 + 2^{-k/2} \|\eta - C_{\nu, k}\|)^{-N} \right. \\ & \quad \left. (1 + 2^{-k/2} \|\eta - \partial_{\tilde{x}}(\langle \tilde{\Phi}_\gamma(\tilde{x}, x), \tilde{x}, \xi \rangle)|_{\tilde{x}=2^{-k/2}x'}\|)^{-N} \right\} \\ & \lesssim 2^{-k(j + \frac{|\alpha|}{2})} (1 + 2^{-k/2} \|\xi - d\Phi_\gamma^T(x) C_{\nu, k}\|)^{-N} \\ & \lesssim 2^{-k(j + \frac{|\alpha|}{2})} (1 + 2^{-k/2} \|\xi - C_{\nu', k'}\|)^{-N}. \end{aligned}$$

This is the required estimate and completes the proof.  $\square$

The next lemma says that we may decompose curvelet like functions into two pieces one of which is compactly supported and the other which decays very quickly with the scale  $k$ .

LEMMA A.5. *If  $\mathcal{F}$  is a family of curvelet like functions, then for every  $\epsilon > 0$  it is possible to find a family of curvelet like functions  $\mathcal{G}$  and a map,  $T_{\mathcal{G}} : \mathcal{F} \rightarrow \mathcal{G}$  such that*

1.  $\pi_{\Gamma} \circ T_{\mathcal{G}} = \pi_{\Gamma}$
2. *For every  $(g, \gamma) \in \mathcal{G}$ , if  $\xi \in \text{supp}(\widehat{g})$ , then*

$$2^k(1/2 - \epsilon) \leq \|\xi\| \leq 2^k(2 + \epsilon) \quad \text{and} \quad 2^k(1/2 - \epsilon) \leq |\langle \nu, \xi \rangle| \leq 2^k(2 + \epsilon).$$

3. *For every  $m \in \mathbb{R}$ ,*

$$\left\{ (2^{km}(f - \pi_S \circ T_{\mathcal{G}}(f, \gamma)), \gamma) \right\}_{(f, \gamma) \in \mathcal{F}}$$

*is a family of curvelet like functions.*

*Proof.* We begin by choosing a cut-off function  $\chi \in C_c^\infty$  supported within  $\epsilon$  of the set  $A_1 = \{2^{-1} \leq \|\xi\| \leq 2\} \cap \{2^{-1} \leq |\langle \nu, \xi \rangle| \leq 2\}$  and equal to 1 within  $\epsilon/2$  of this set. We construct  $\chi$  so that it is symmetric with respect to rotations that preserve  $\nu$ . Also, we set  $\chi_k(\xi) = \chi(2^{-k}\xi)$ . The first task is to show that

$$\mathcal{G} = \left\{ \chi_k(D)f, \gamma \right\}_{(f, \gamma) \in \mathcal{F}}$$

is a family of curvelet like functions which will then satisfy requirement 2. For  $(f, \gamma) \in \mathcal{F}$  we have

$$\rho_k^{1/2} \langle \nu, \partial_\xi \rangle^j \partial_\xi^\alpha e^{i\langle x, \xi \rangle} [\widehat{\chi_k(D)f}](\xi) = \langle \nu, \partial_\xi \rangle^j \partial_\xi^\alpha (\rho_k^{1/2} e^{i\langle x, \xi \rangle} \chi_k(\xi) \widehat{f}(\xi))$$

and it follows from this expression and the Liebniz rule that  $\mathcal{G}$  is a family of curvelet like functions.

It now remains to show that for any  $m$ ,

$$\mathcal{H} = \left\{ 2^{km}(1 - \chi_k(D))f, \gamma \right\}_{(f, \gamma) \in \mathcal{F}}$$

is a family of curvelet like functions. Once again for  $(f, \gamma) \in \mathcal{F}$ , we have using definition A.1 that for any  $N$

$$\begin{aligned} \left| \rho_k^{1/2} \langle \nu, \partial_\xi \rangle^j \partial_\xi^\alpha e^{i\langle x, \xi \rangle} [(1 - \widehat{\chi_k(D)})f](\xi) \right| &= \left| \langle \nu, \partial_\xi \rangle^j \partial_\xi^\alpha (\rho_k^{1/2} e^{i\langle x, \xi \rangle} (1 - \chi_k(\xi)) \widehat{f}(\xi)) \right| \\ &\lesssim 2^{-k(j + \frac{|\alpha|}{2})} (1 + 2^{-k/2} \|\xi - C_{\nu, k}\|)^{-N} \sup_{2^{-k+1}\xi \in A_1^c} (1 + 2^{-k/2} \|\xi - C_{\nu, k}\|)^{-2m} \\ &\lesssim 2^{-k(j + \frac{|\alpha|}{2})} (1 + 2^{-k/2} \|\xi - C_{\nu, k}\|)^{-N} 2^{-mk}. \end{aligned}$$

This completes the proof.  $\square$

Now we begin to examine the action of pseudodifferential operators on families of curvelet like functions.

LEMMA A.6. *Suppose that  $\mathcal{F}$  is a family of curvelet like functions, and that  $\{p_\gamma(y, \xi)\}_{\gamma \in \pi_\Gamma(\mathcal{F})}$  is a collection of smooth functions on  $\mathbb{R}^n \times (\mathbb{R}^n \setminus \{0\})$  such that for some  $m \in \mathbb{R}$ , any multi-indices  $\alpha$  and  $\beta$ , and any nonnegative integer  $j$  there is a constant  $C_{\alpha, \beta, j}$  so that*

$$|\partial_y^\beta \partial_\xi^\alpha p_\gamma(y, \xi)| \leq C_{\alpha, \beta, j} 2^{k \frac{|\beta|}{2}} (1 + \|\xi\|)^{m - |\alpha|}$$

for all  $(y, \xi)$ . Then

$$\left\{ (2^{-km} p_\gamma(y, D)f, \gamma) \right\}_{(f, \gamma) \in \mathcal{F}}$$

is a family of curvelet like functions.

*Proof.* We begin by choosing a family  $\mathcal{G}$  as in lemma A.5 with some small value of  $\epsilon > 0$ . The following calculation then applies when  $(h, \gamma)$  equals either  $T_{\mathcal{G}}(f, \gamma)$  or  $(f - \pi_S \circ T_{\mathcal{G}}(f, \gamma), \gamma)$ .

$$\begin{aligned} & \rho_k^{1/2} \langle \nu, \partial_\xi \rangle^j \partial_\xi^\alpha e^{i\langle x, \xi \rangle} \mathcal{F} \left\{ p_\gamma(y, D) h \right\}(\xi) \\ &= \frac{1}{(2\pi)^n} \iint e^{i\langle \eta - \xi, y - x \rangle} \langle \nu, \partial_\eta \rangle^j \partial_\eta^\alpha p_\gamma(y, \eta) \left( \rho_k^{1/2} e^{i\langle x, \eta \rangle} \widehat{h}(\eta) \right) d\eta dy \\ &= \frac{1}{(2\pi)^n} \iint e^{i\langle \eta - \xi, y \rangle} \langle \nu, \partial_\eta \rangle^j \partial_\eta^\alpha p_\gamma(y + x, \eta) \left( \rho_k^{1/2} e^{i\langle x, \eta \rangle} \widehat{h}(\eta) \right) d\eta dy \\ &= \frac{1}{(2\pi)^n} \iint e^{i\langle \eta - \xi, y \rangle} \left( \frac{1 - i 2^{-k/2} \langle \eta - \xi, \partial_y \rangle}{1 + 2^{-k/2} \|\eta - \xi\|^2} \right)^M \left( \frac{1 - 2^k \Delta_\eta}{1 + 2^k |y|^2} \right)^N \\ & \quad \cdot \langle \nu, \partial_\eta \rangle^j \partial_\eta^\alpha p_\gamma(y + x, \eta) \left( \rho_k^{1/2} e^{i\langle x, \eta \rangle} \widehat{h}(\eta) \right) d\eta dy. \end{aligned}$$

In the case when  $(h, \gamma) = T_{\mathcal{G}}(f, \gamma)$ , by taking  $M$  sufficiently large this integral may be bounded by the required estimate since on the support of  $h$

$$|\partial_y^\beta \partial_\eta^\alpha \langle \nu, \partial_\eta \rangle^j p_\gamma(y + x, \eta)| \lesssim 2^{k(m-j-|\alpha|+\frac{|\beta|}{2})}$$

and

$$(1 + 2^{-k/2} \|\eta - \xi\|)^{-1} (1 + 2^{-k/2} \|\eta - C_{\nu, k}\|)^{-1} \leq (1 + 2^{-k/2} \|\xi - C_{\nu, k}\|)^{-1}.$$

For the case when  $(h, \gamma) = (f - \pi_S \circ T_{\mathcal{G}}(f, \gamma), \gamma)$ , we use the fact that  $2^{k\tilde{m}}(f - \pi_S \circ T_{\mathcal{G}}(f, \gamma))$  gives a family of curvelet like functions for any  $\tilde{m}$ . Therefore the integral in this case may be bounded by a constant times

$$2^{-k\tilde{m}} (1 + 2^{-k/2} \|\xi - C_{\nu, k}\|)^{-N}$$

for any  $\tilde{m}$ . This proves the result.  $\square$

The next lemma examines the case of lemma A.4 when the diffeomorphisms depend on a parameter.

LEMMA A.7. *Suppose that  $\mathcal{F}$  is a family of curvelet like functions, and that  $\{\Phi_\gamma\}_{\gamma \in \pi_\Gamma(\mathcal{F})}$  is a smooth family of functions from  $[z_0, Z] \times \mathbb{R}^n$  to  $\mathbb{R}^n$  such that*

$$\{\Phi_\gamma(z, \cdot)\}_{\gamma \in \pi_\Gamma(\mathcal{F})}$$

*satisfies the hypotheses of lemma A.4 for every fixed  $z$  with the constants in the estimates uniform with respect to  $z$ , and*

$$|\partial_y^\alpha \partial_z \Phi_\gamma(z, y)| \leq C_\alpha 2^{k\frac{|\alpha|}{2}}.$$

*Then*

$$\left\{ (2^{-k} \partial_z [\Phi_\gamma^*(z, \cdot) f], \Phi_\gamma^*(\gamma)) \right\}_{(f, \gamma) \in \mathcal{F}}$$

*is a family of curvelet like functions.*

*Proof.* Let  $(f, \gamma) \in \mathcal{F}$ . Then, as in the proof of lemma A.4 we use the Fourier inversion formula to establish that

$$\begin{aligned} \partial_z [\Phi_\gamma^*(z, \cdot) f](y) &= \frac{i}{(2\pi)^n} \iint e^{i\langle \Phi_\gamma(z, y), \xi \rangle} \langle \partial_z \Phi_\gamma(z, y), \xi \rangle \widehat{f}(\xi) d\xi \\ &= \Phi_\gamma^*(z, \cdot) [ \langle \partial_z \Phi_\gamma(z, \Phi_\gamma^{-1}(z, \cdot)), \partial_y \rangle f ](y) \end{aligned}$$

The collection of functions  $\{\langle \partial_z \Phi_\gamma(z, \Phi_\gamma^{-1}(z, y)), \xi \rangle\}_{\gamma \in \pi_\nu(\mathcal{F})}$  satisfy the hypotheses of lemma A.6 with  $m = 1$ , and so that lemma and lemma A.4 imply the result.  $\square$

The next lemma gives an explicit expression for the leading order terms of the action of a suitable family of pseudodifferential operators with principal symbols that are homogeneous of degree 1 on a family of curvelet like functions. For every  $\nu \in \pi_\nu(\mathcal{F})$  we use the notation  $P_\nu$  for the matrix which gives orthogonal projection onto the space perpendicular to  $\nu$ .

LEMMA A.8. *Suppose that  $\mathcal{F}$  is a family of curvelet like functions, and that  $\{p_\gamma\}_{\gamma \in \pi_\Gamma(\mathcal{F})}$  is a collection of functions that satisfy the hypotheses of lemma A.6 with  $m = 1$ , and are such that every  $p_\gamma$  is positive homogeneous in  $\xi$  of degree 1 on  $\{2^{C-2} \leq |\langle \xi, \nu \rangle|\}$  where  $C$  is the constant from part 2 of definition A.1. Also, let  $\phi(t) \in C_c^\infty(\mathbb{R})$  be a function that is equal to zero when  $|t| \leq 2^{C-3}$  and equal to 1 when  $|t| > 2^{C-2}$ . If for every  $(f, \gamma) \in \mathcal{F}$  we define*

$$g = 2^{k/2} \left( p_\gamma(y, D)f - \left( \langle \partial_\xi p_\gamma(y, \nu), Df \rangle + \frac{1}{2} \text{tr} \left( \partial_\xi^2 p_\gamma(x, \nu) \frac{\phi(\langle \nu, D \rangle)}{\langle \nu, D \rangle} D^2 f \right) \right) \right)$$

then

$$\left\{ (g, \gamma) \right\}_{(f, \gamma) \in \mathcal{F}}$$

is a family of curvelet like functions.

*Proof.* First, applying both lemma A.5 and lemma A.6 we may assume without loss of generality that every  $(f, \gamma) \in \mathcal{F}$  satisfies part 2 of lemma A.5 for some small value of  $\epsilon > 0$ .

Next we make the following expansion of  $p_\gamma$ , which holds for  $\xi \in \text{supp}(\hat{f})$  and follows using the homogeneity assumption.

$$(A.4) \quad p_\gamma(y, \xi) = \langle \partial_\xi p_\gamma(y, \nu), \xi \rangle + \frac{1}{2} \sum_{q,r=1}^n \partial_\xi^2 p_\gamma(x, \nu)_{qr} \frac{\xi^q \xi^r}{\langle \xi, \nu \rangle} + R_\gamma(y, \xi)$$

where

$$\begin{aligned} R_\gamma(y, \xi) &= \frac{1}{6} \sum_{q,r,s=1}^n \left( \int_0^1 \partial_{\xi^q \xi^r \xi^s}^3 p_\gamma(y, \langle \xi, \nu \rangle \nu + t P_\nu \xi) dt \right) (P_\nu \xi)^q (P_\nu \xi)^r (P_\nu \xi)^s \\ &\quad + \frac{1}{2} \sum_{r,s=1}^{n-1} \left( \int_0^1 \partial_{y^q \xi^r \xi^s}^3 p_\gamma(x + t(y-x), \nu) \nu_r^r \nu_s^s dt \right) (x-y)^q \frac{\langle \xi, \nu_r \rangle \langle \xi, \nu_s \rangle}{\langle \xi, \nu \rangle}. \end{aligned}$$

We write  $(R_\gamma^1)_{qrs}(y, \xi)$  and  $(R_\gamma^2)_{rs}(y)$  respectively for the two arrays of functions given by the integrals in the previous formula. From the hypotheses and the fact that every  $(f, \gamma) \in \mathcal{F}$  satisfies part 2 of lemma A.5 we see that each of these functions satisfies the hypotheses of lemma A.6 with respectively  $m = -2$  and  $m = 0$  for  $\xi$  restricted to  $\text{supp}(\hat{f})$ . Therefore by A.3

$$\left\{ 2^{k/2} [R_\gamma(y, D)f], \gamma \right\}_{(f, \gamma) \in \mathcal{F}}$$

is a family of curvelet like functions. From (A.4) we observe that

$$g = 2^{k/2} [R_\gamma(y, D)f],$$

and so the proof is complete.  $\square$



**Appendix B. Lemma for paradifferential estimates.**

In this appendix we will state and prove the lemma used to deal with the “rough” parts of the Volterra kernels. The lemma is an extension of lemma 13 in [7] to a broader class than just multipliers. The idea of expanding homogeneous symbols using spherical harmonics is suggested in exercise 13.9.6 of [23].

Let  $a(x, \xi) \in C^{m,1} S_{cl}^m(\mathbb{R}^n)$  be homogeneous of order 1 in  $\xi$ , and let  $a_k$  be obtained by (3.1) applied to  $a$  instead of  $p$ . Also, let  $\beta \in C_0^\infty(\mathbb{R}^n)$  be a function such that  $0 \leq \beta \leq 1$ ,  $\text{supp}(\beta) \subset \{1/2 \leq |\xi| \leq 2\}$  for some  $l_0 \in \mathbb{Z}^+$ , and constructed so that  $\beta_0(\xi) + \sum_{k=1}^{\infty} \beta(2^{-k+1}\xi) = 1$  for another function  $\beta_0 \in C_0^\infty(\mathbb{R}^n)$  with support contained in the unit ball. For convenience we define  $\beta_k(\xi) = \beta(2^{-k+1}\xi)$  for  $k \geq 1$  (ie. so that  $\{\beta_k\}$  provides a Littlewood-Paley partition of unity). Assume that  $F_k$  is a family of operators on  $L^2(\mathbb{R}^n)$  satisfying estimates of the form

$$(B.1) \quad \|F_k u\|_{L^2(\mathbb{R}^n)} \leq C 2^{kr} \|\beta_k(D)u\|_{L^2(\mathbb{R}^n)}.$$

We will further assume that each  $F_k$  is frequency localized at the scale  $A 2^k$  for some constant  $A$  in the sense that  $(1 - \beta_k(D/A))F_k = R_k$  where  $R_k : H^{s'} \rightarrow H^s$  is continuous uniform in  $k$  for any  $s'$  and  $s$ .

LEMMA B.1. *If  $s$  and  $s' \in \mathbb{R}$  are such that*

$$s < m + 1, \quad r - m \leq s' - s, \quad r - m \leq s', \quad \text{and} \quad \frac{2r - m + 1}{2} \leq s' - s,$$

*then there is an  $N \in \mathbb{N}$  such that for any  $u \in H^{s'}(\mathbb{R}^n)$*

$$\left\| \left( \sum_k (a(y, D) - a_k(y, D)) F_k \right) u \right\|_{H^s} \lesssim (C + C') \sum_{|\alpha| \leq 2N} \sup_{\omega \in \mathbb{S}^{n-1}} \|\partial_\xi^\alpha a(\cdot, \omega)\|_{C^{m,1}} \|u\|_{H^{s'}}.$$

*The constant  $C$  in this estimate is the same as the constant in (B.1), and  $C'$  is a modulus of continuity for  $R_k : H^{s'} \rightarrow H^s$  that is independent of  $k$ .*

*Proof.* For ease of notation we will write  $f_k(y, \xi) = a(y, \xi) - a_k(y, \xi)$  and now record a few properties of  $f_k$ . First, from the homogeneity of  $a$  it is still true that  $f_k$  is homogeneous of degree 1 in  $\xi$ . Second, because  $a_k$  is obtained by a low pass filter in  $y$  from  $a$ ,  $\widehat{f}_k(\eta, \xi) = \widehat{f}_j(\eta, \xi)$  for  $|\eta| \geq 2^{\max(k,j)/2}$ . Finally, the estimates

$$(B.2) \quad |\partial_y^\beta \partial_\xi^\alpha f_k(y, \xi)| \lesssim 2^{-k(m+1-|\beta|)/2} \|\partial_\xi^\alpha a(\cdot, \xi)\|_{C^{m,1}}$$

for  $|\beta| \leq m + 1$  follow from (3.2).

The first step of the proof will be to decompose  $f_k(y, D)$  in terms of a sum of multiplication and convolution operators by using spherical harmonics in the phase space. Indeed, let  $\{w_\kappa\}$  denote the set of eigenfunctions of  $\Delta_{\mathbb{S}^{n-1}}$ , with the eigenvalue of  $w_\kappa$  denoted by  $\lambda_\kappa$ , which form an orthonormal basis for  $L^2(\mathbb{S}^n)$ . By the homogeneity of  $f_k$  in  $\xi$  we have

$$f_k(y, \xi) = |\xi| \sum_\kappa f_{k\kappa}(y) w_\kappa(\xi/|\xi|)$$

where

$$f_{k\kappa}(y) = \int_{\mathbb{S}^{n-1}} f_k(y, \omega) w_\kappa(\omega) d\omega.$$

By Green's formula we have for any  $N \in \mathbb{N}$

$$\lambda_\kappa^N f_{k\kappa} = \int_{\mathbb{S}^{n-1}} f_k(y, \omega) \Delta_{\mathbb{S}^{n-1}}^N w_\kappa(\omega) d\omega = \int_{\mathbb{S}^{n-1}} (\Delta_{\mathbb{S}^{n-1}}^N f_k)(y, \omega) w_\kappa(\omega) d\omega,$$

which using (B.2) gives

$$(B.3) \quad |\partial_y^\beta f_{k\kappa}| \lesssim \lambda_\kappa^{-N} 2^{-k(m+1-|\beta|)/2} \sum_{|\alpha| \leq 2N} \sup_{\omega \in \mathbb{S}^{n-1}} \|\partial_\xi^\alpha a(\cdot, \omega)\|_{C^{m,1}}$$

for  $|\beta| \leq m+1$  and any  $N$ . Also, we can see from the definition that the family  $\{f_{k\kappa}\}$  inherits the property that  $\widehat{f_{k\kappa}}(\eta) = \widehat{f_{j\kappa}}(\eta)$  for  $|\eta| \geq 2^{\max(k,j)/2}$  from  $\{f_k\}$ . With this decomposition we have

$$\sum_k f_k(y, D) F_k = \sum_{k, \kappa} f_{k\kappa}(y) |D| w_\kappa(D/|D|) F_k$$

The operator  $|D| w_\kappa(D/|D|) F_k$  is still frequency localized at scale  $A2^k$  in the sense given above, and so this last formula in fact gives

$$\sum_k f_k(y, D) F_k = \sum_{k, \kappa} f_{k\kappa}(y) \beta_k(D/A) |D| w_\kappa(D/|D|) F_k + f_{k\kappa}(y) R_{k\kappa}$$

where  $R_{k\kappa} : H^{s'} \rightarrow H^s$  has the same continuity properties as  $R_k$  but uniform also in  $\kappa$ . To finish we apply one further decomposition to the above operator by separating frequency scales a second time to obtain

$$(B.4) \quad \sum_k f_k(y, D) F_k = \sum_{j, k, \kappa} (\beta_j(D/A) f_{k\kappa}(y) \beta_k(D/A)) |D| w_\kappa(D/|D|) F_k + \sum_{k, \kappa} f_{k\kappa}(y) R_{k\kappa}.$$

Note that by the uniform continuity of  $R_{k\kappa}$ , if we take  $N$  to be large enough then the second sum on the right gives an operator with the required properties provided that  $s < m+1$  in which case  $f_{k\kappa}$  acts as a multiplier mapping  $H^s$  to  $H^s$  with norm bounded by (B.3). Thus we have reduced the proof to estimating the first sum in (B.4).

Now we will look in more detail at the operators in parentheses from the right hand side above, which we will label as

$$\Gamma_{jk\kappa} = \beta_j(D/A) f_{k\kappa} \beta_k(D/A) |D|.$$

With this notation, for  $u \in L^2(\mathbb{R}^n)$  and  $j, k \geq 1$  we have

$$\widehat{\Gamma_{jk\kappa} u}(\xi) = \frac{\beta(2^{-j+1}\xi/A)}{(2\pi)^n} \int_{\mathbb{R}^n} \widehat{f_{k\kappa}}(\xi - \eta) \beta(2^{-k+1}\eta/A) |\eta| \widehat{u}(\eta) d\eta.$$

From this formula and the location of the support of  $\beta(\xi)$  we have that  $\widehat{\Gamma_{jk\kappa} u}(\xi)$  only depends on  $\widehat{f_{k\kappa}}(\xi - \eta)$  at points where

$$\xi \in \text{supp}(\beta(2^{-j+1} \cdot /A)) = \{\xi \in \mathbb{R}^n \mid A2^{j-2} \leq |\xi| \leq A2^j\}$$

and

$$\eta \in \text{supp}(\beta(2^{-k+1} \cdot /A)) = \{\eta \in \mathbb{R}^n \mid A2^{k-2} \leq |\eta| \leq A2^k\}.$$

We will use the notation  $l(A) = \min(0, \text{floor}(\log_2(A)))$ . Assuming that  $j-2 > k$ , by the triangle inequality

$$\begin{aligned} & \{\xi \in \mathbb{R}^n \mid A2^{j-2} \leq |\xi| \leq A2^j\} - \{\eta \in \mathbb{R}^n \mid A2^{k-2} \leq |\eta| \leq A2^k\} \\ & \subset \{\zeta \in \mathbb{R}^n \mid |\zeta| \geq A(2^{j-2} - 2^k) = A2^{j-2}(1 - 2^{k-j+2}) \geq 2^{j-3+l(A)}\}, \end{aligned}$$

and similarly assuming that  $k-2 > j$

$$\begin{aligned} & \{\xi \in \mathbb{R}^n \mid A2^{j-2+l(A)} \leq |\xi| \leq A2^j\} - \{\eta \in \mathbb{R}^n \mid A2^{k-2} \leq |\eta| \leq A2^k\} \\ & \subset \{\zeta \in \mathbb{R}^n \mid |\zeta| \geq 2^{k-3+l(A)}\}. \end{aligned}$$

Since  $f_{k\kappa}(\zeta) = f_{(2^{j-6+2l(A)})\kappa}(\zeta)$  for  $|\zeta| \geq 2^{\max(k, 2^{j-3+l(A)})/2}$ , in the case that  $j > k + 3 - l(A)$  we have

$$\Gamma_{jk\kappa} = \beta_j(D/A) f_{(2^{j-6+2l(A)})\kappa} \beta_k(D/A) |D|,$$

and in the case that  $k > j + 3 - l(A)$

$$\Gamma_{jk\kappa} = \beta_j(D/A) f_{(2^{k-6+2l(A)})\kappa} \beta_k(D/A) |D|.$$

To simplify the following formulas we now introduce the notation

$$\|a\|_N = \sum_{|\alpha| \leq 2N} \sup_{\omega \in \mathbb{S}^{n-1}} \|\partial_\xi^\alpha a(\cdot, \omega)\|_{C^{m,1}}.$$

Based on (B.3) and the fact that  $\|\beta_k(D)\|_{L^2(\mathbb{R}^n) \rightarrow L^2(\mathbb{R}^n)} \leq 1$  for any  $r$  we can see that

$$\|\Gamma_{jk\kappa}\|_{L^2 \rightarrow L^2} \lesssim \begin{cases} A 2^k \lambda_\kappa^{-N} 2^{-j(m+1)} \|a\|_N & \text{for } j > k + 3 - l(A), \\ A 2^k \lambda_\kappa^{-N} 2^{-k(m+1)/2} \|a\|_N & \text{for } k + 3 - l(A) \geq j \geq k - 3 + l(A), \\ A 2^k \lambda_\kappa^{-N} 2^{-k(m+1)} \|a\|_N & \text{for } k - 3 + l(A) > j. \end{cases}$$

With these estimates available we now return to (B.4). It remains to estimate the operator norm of the first sum on the right hand side, which we will refer to as  $S$ . As before, we may choose  $N$  sufficiently large depending on the dimension so that the sums with respect to  $\kappa$  in (B.4) converge absolutely and uniformly with respect to the other parameters. Having chosen such an  $N$  we now apply these estimates to obtain the result. For  $u \in C_0^\infty(\mathbb{R}^n)$

$$\begin{aligned} \|Su\|_{H^s}^2 &\lesssim C^2 \sum_{j,k,\kappa} 2^{2(js+k\kappa r)} \|\Gamma_{jk\kappa}\|_{L^2 \rightarrow L^2}^2 \|\beta_k(D)u\|_{L^2}^2 \\ &\lesssim C^2 \sum_{j,k,\kappa} 2^{2(js+k(r-s'))} \|\Gamma_{jk\kappa}\|_{L^2(\mathbb{R}^n)}^2 \left(2^{ks'} \|\beta_k(D)u\|_{L^2(\mathbb{R}^n)}\right)^2 \\ &\lesssim (AC)^2 \|a\|_N^2 \left( \sum_{j>k+3-l(A)} 2^{2(j(s-m-1)+k(r-s'+1))} \left(2^{ks'} \|\beta_k(D)u\|_{L^2(\mathbb{R}^n)}\right)^2 \right. \\ &\quad + \sum_{k-3+l(A) \leq j \leq k+3-l(A)} 2^{2js+k(2r-2s'-m+1)} \left(2^{ks'} \|\beta_k(D)u\|_{L^2(\mathbb{R}^n)}\right)^2 \\ &\quad \left. + \sum_{k>j+3-l(A)} 2^{2(js+k(r-s'-m))} \left(2^{ks'} \|\beta_k(D)u\|_{L^2(\mathbb{R}^n)}\right)^2 \right). \end{aligned}$$

If  $s - m - 1 < 0$  and  $r - m \leq s' - s$ , then the first sum in parentheses is bounded by  $\|u\|_{H^{s'}(\mathbb{R}^n)}^2$ . If  $(2r - m + 1)/2 \leq s' - s$ , then the second sum is also bounded by  $\|u\|_{H^{s'}(\mathbb{R}^n)}^2$ . Finally, if  $r - m \leq s'$  and  $r - m \leq s' - s$ , then the third sum is bounded by the same quantity. This completes the proof of the lemma.

□

Note that  $s - s'$  measures the extent to which the operator

$$\sum_k (a(y, D) - a_k(y, D)) F_k$$

is smoothing. The previous lemma thus states that if  $m$  is large enough, then this operator is smoothing.

## REFERENCES

- [1] FREDRIK ANDERSSON, MARCUS CARLSSON, AND MAARTEN V. DE HOOP, *Nonlinear approximation of functions in two dimensions by sums of exponential functions.*, Appl. Comput. Harmon. Anal., 29 (2010), pp. 156–181.
- [2] ———, *Nonlinear approximation of functions in two dimensions by sums of wave packets.*, Appl. Comput. Harmon. Anal., 29 (2010), pp. 198–213.
- [3] FREDRIK ANDERSSON, MAARTEN V. DE HOOP, HART F. SMITH, AND GUNTHER UHLMANN, *A multi-scale approach to hyperbolic evolution equations with limited smoothness.*, Commun. Partial Differ. Equations, 33 (2008), pp. 988–1017.
- [4] GREGORY BEYLKIN AND LUCAS MONZÓN, *On approximation of functions by exponential sums.*, Appl. Comput. Harmon. Anal., 19 (2005), pp. 17–48.
- [5] ———, *Approximation by exponential sums revisited.*, Appl. Comput. Harmon. Anal., 28 (2010), pp. 131–149.
- [6] HERMANN BRUNNER AND P.J. VAN DER HOUWEN, *The numerical solution of Volterra equations.*, CWI Monographs, 3. Centrum voor Wiskunde en Informatica, Amsterdam. Amsterdam etc.: North-Holland. XVI, 1986.
- [7] VALERIY BRYTIK, MAARTEN V. DE HOOP, AND MIKKO SALO, *Sensitivity analysis of wave-equation tomography: a multi-scale approach*, J. Fourier Anal. Appl., 16 (2010), pp. 544–589.
- [8] EMMANUEL CANDÈS, LAURENT DEMANET, DAVID DONOHO, AND LEXING YING, *Fast discrete curvelet transforms.*, Multiscale Model. Simul., 5 (2006), pp. 861–899.
- [9] EMMANUEL CANDÈS, LAURENT DEMANET, AND LEXING YING, *Fast computation of Fourier integral operators.*, Siam J. Sci. Comput., 29 (2007), pp. 2464–2493.
- [10] EMMANUEL J. CANDÈS AND DAVID L. DONOHO, *Continuous curvelet transform. I: Resolution of the wave front set.*, Appl. Comput. Harmon. Anal., 19 (2005), pp. 162–197.
- [11] ———, *Continuous curvelet transform. II: Discretization and frames.*, Appl. Comput. Harmon. Anal., 19 (2005), pp. 198–222.
- [12] EMMANUEL J. CANDÈS AND FRANCK GUO, *New multiscale transforms, minimum total variation synthesis: Applications to edge-preserving image reconstruction.*, Signal Process., 82 (2002), pp. 1519–1543.
- [13] M.V. DE HOOP, H. SMITH, G. UHLMANN, AND R.D. VAN DER HILST, *Seismic imaging with the generalized Radon transform: a curvelet transform perspective*, Inverse Problems, (2009).
- [14] M. V. DE HOOP, T. J. P. M. OP’T ROOT, AND C. C. STOLK, *Reverse time migration-inversion from single-shot data*, in SEG (Houston TX), 2009.
- [15] A. A. DUCHKOV, F. ANDERSSON, AND M. V. DE HOOP, *Discrete, almost symmetric wave packets and multiscale geometric representation of (seismic) waves*, IEEE Trans. Geosc. and Remote Sensing, 48 (2010), pp. 3408–3423.
- [16] ANTON A. DUCHKOV, MAARTEN V. DE HOOP, AND ANTÔNIO S. SÁ BARRETO, *Evolution-equation approach to seismic image, and data, continuation.*, Wave Motion, 45 (2008), pp. 952–969.
- [17] LARS HÖRMANDER, *The analysis of linear partial differential operators. III: Pseudo-differential operators. Reprint of the 1994 ed.*, Classics in Mathematics. Berlin: Springer, 2007.
- [18] HART F. SMITH, *A Hardy space for Fourier integral operators*, J. Geom. Anal., 8 (1998), pp. 629–653.
- [19] ———, *A parametrix construction for wave equations with  $C^{1,1}$  coefficients*, Ann. Inst. Fourier, 48 (1998), pp. 797–835.
- [20] C. C. STOLK AND M. V. DE HOOP, *Modelling of seismic data in the downward continuation approach*, SIAM J. Appl. Math., 65 (2005), pp. 1388–1406.
- [21] ———, *Seismic inverse scattering in the downward continuation approach*, Wave Motion, 43 (2006), pp. 579–598.
- [22] MICHAEL E. TAYLOR, *Partial differential equations. II*, vol. 116 of Applied Mathematical Sciences, Springer-Verlag, New York, 1996. Qualitative studies of linear equations.
- [23] ———, *Tools for PDE. Pseudodifferential operators, paradifferential operators, and layer potentials.*, Mathematical Surveys and Monographs. 81. Providence, RI: American Mathematical Society (AMS), 2000.