Proceedings of the Project Review, Geo-Mathematical Imaging Group (Purdue University, West Lafayette IN), Vol. 1 (2009) pp. ??-??.

DISCRETE GRT COMPRESSION USING WAVE PACKETS

HERWIG WENDT[*], FREDRIK ANDERSSON[†], AND MAARTEN V. DE HOOP[‡]

Abstract. Recently, approximate expressions of the action of GRT operators on a single wave packet have been obtained. In this note, the potential use of this result in practically useable numerical schemes is investigated, and numerical key elements for efficient practical implementations are proposed. Notably, a novel procedure for obtaining the tensor product representation of a complex exponential, involved in such approximations, is obtained. It is based on orthogonal polynomial expansion and does not involve any computationally costly singular value decomposition step. The procedure is illustrated on a synthetic CO imaging problem for homogeneous medium. Finally, the practical use of approximations per single wave packet for evaluating the approximate action of GRT operators on functions composed of wave packets is discussed, and a potential numerical scheme for this problem is outlined.

Key words. Wave Packets, Curvelets, Dyadic Parabolic Decomposition, Singular Value Decomposition, Generalized Radon Transform, Fourier Integral Operator

1. Introduction.

A large class of seismic imaging and inverse scattering techniques can be formulated in terms of a Generalized Radon Transform (GRT) and its extension (cf. e.g. [12] and references therein). The GRT (scattering or imaging) operator F can microlocally be brought in the form of a Fourier integral operator (FIO)[1]:

$$(Fu)(y) = \int a(y,\xi)\exp(iS(y,\xi))\hat{u}(\xi)d\xi, \tag{1.1}$$

where the amplitude function $a(y,\xi)$ and the generating function $S(y,\xi)$ are determined by the ray-geometry of the background medium, and where the latter describe the propagation of singularities by the operator[2] [4]. Under appropriate conditions, the operator F has sparse representation in the curvelet frame [11]. In the remainder of this report, we refer to curvelets (e.g. [2] and references therein) by their original name "wave packets".

Recently, Hoop et al. [4] proposed a simple approximation of the action of the operator F on a single wave packet, with error $2^{-k/2}$ for a wave packet at frequency scale 2^k (cf. (2.5) below). The result is summarized in Section 2 for convenience. The work presented in this report elaborates on this result and proposes a numerical procedure for evaluating it in practice: First, a numerical procedure for obtaining the tensor product representation of the complex exponential involved in approximation (2.5) is proposed (cf. Section 3). Then, this procedure is used to numerically illustrate the effectiveness of the approximation (2.5) (cf. Section 4). Finally, in Section 5 a potential strategy for obtaining a efficient parallel numerical scheme for approximating the action of the operator F on input functions composed of wave packets is outlined.

2. GRT Approximation.

For convenience, this section briefly recalls the result in [4] that will be elaborated on throughout the remainder of this report. It approximates the action of the operator F on a single wave packet $\varphi_\gamma(x)$, $\gamma = (j,\nu,k)$, with central position x_j and orientation ν at scale k:

$$(F\varphi_\gamma)(y) = \rho_k^{-1/2}\int a(y,\xi)\exp[i(S(y,\xi) - \langle\xi,x_j\rangle)]\hat{\chi}_{\nu,k}(\xi)d\xi, \tag{2.1}$$

where $\hat{\varphi}_\gamma(\xi) = \rho_k^{-1/2}\hat{\chi}_{\nu,k}(\xi)\exp[-i\langle\xi,x_j\rangle]$ is the Fourier transform of φ_γ. The strategy of [4] for obtaining an approximation of $(F\varphi_\gamma)(y)$ up to error of order $2^{-k/2}$ is to replace S by a sufficient

[*] CCAM, Department of Mathematics, Purdue University, West Lafayette, IN, USA (hwendt@math.purdue.edu)

[†] Mathematics LTH, Centre for Mathematical Sciences, Lund Institute of Technology, Lund University, SE (fa@maths.lth.se)

[‡] CCAM, Department of Mathematics, Purdue University, West Lafayette, IN, USA (mdehoop@math.purdue.edu)

[1] We denote the Fourier transform of a function by $\hat{\cdot}$, and by ξ the Fourier (frequency) variables.

[2] In this report, we will not discuss how S and a is obtained in practice, and consider it to be given.

number of terms of its Taylor expansion in ξ on the frequency support of the wave packet φ_γ. Using the first order Taylor term only, and neglecting higher order terms, (2.1) becomes:

$$(F\varphi_\gamma)(y) = a(y,\nu)\varphi_\gamma(T_{\nu,k}(y)) \; + \; O(0), \tag{2.2}$$

where the function $T_{\nu,k}$ is defined by:

$$y \longrightarrow T_{\nu,k}(y) = \frac{\partial S}{\partial \xi}(y,\nu) \tag{2.3}$$

and corresponds to a coordinate transform. Essentially, approximation (2.2) describes the propagation of the wave packet φ_γ along a ray according to geometrical optics, without altering the support or shape of the wave packet. Refining the approximation by including the leading term of the second order Taylor expansion of S on the support of φ_γ, and neglecting higher order terms and terms leading to smoother errors, gives the desired result, with error of order $2^{-k/2}$ (Theorem 4.1 in [4]):

Theorem 2.1 (Theorem 4.1 in [4]). *With functions $T_{\nu,k}(y)$ defined by (2.3), $\alpha^1_{r;\nu,k}(y)$ and $\hat{\alpha}^2_{r;\nu,k}(\xi)$ defined by:*

$$a(y,\nu)\exp\left(i\frac{1}{2}\frac{\xi_2^2}{\xi_1}\frac{\partial^2 S}{\partial \xi_2^2}(y,\nu)\right) \approx \sum_{r=1}^{R} \alpha^1_{r;\nu,k}(y)\hat{\alpha}^2_{r;\nu,k}(\xi), \tag{2.4}$$

one may express:

$$(F\varphi_\gamma)(y) = \sum_{r=1}^{R} \alpha^1_{r;\nu,k}(y)\left(\alpha^2_{r;\nu,k} * \varphi_\gamma\right)(T_{\nu,k}(y)) \; + \; 2^{-k/2} f_\gamma, \tag{2.5}$$

where $R \sim k/\log(k)$, and f_γ is a curvelet-like function centered at $\chi(\gamma)$.

Theorem 2.1 states that $(F\varphi_\gamma)(y)$ can be approximated, to order $O(2^{-k/2})$, by a sum over R modified wave packets $\tilde{\varphi}_{r;\gamma}(x) = (\alpha^2_{r;\nu,k} * \varphi_\gamma)(x)$ with amplitude corrections $\alpha^1_{r;\nu,k}(y)$, followed by a coordinate transform $T_{\nu,k}(y)$. The functions $T_{\nu,k}(y)$, $\alpha^1_{r;\nu,k}(y)$ and $\hat{\alpha}^2_{r;\nu,k}(\xi)$ do not depend on the position x_j of the wave packet φ_γ, but only on the scale k and orientation ν, and are hence the same for all wave packets with a certain orientation ν at a given scale k.

Discrete Implementation. The goal of this work is to propose an efficient numerical scheme for the evaluation of (2.5) at discrete frequency and output points ξ_n and y_m, respectively[3]. Eventually, this result will be used for the computation of the approximate action of F on a general function. The convolution in (2.5) can be accounted for in the ξ domain, and hence:

$$(F\varphi_\gamma)(y) = \sum_{r=1}^{R} \alpha^1_{r;\nu,k}(y) \sum_{\xi} e^{i\langle T_{\nu,k}(y),\xi\rangle}\hat{\alpha}^2_{r;\nu,k}(\xi)\hat{\varphi}_\gamma(\xi) \; + \; O(2^{-k/2}). \tag{2.6}$$

Then, evaluating (2.6) involves addressing the following two key issues:

1. **Tensor product representation of a complex exponential.** The functions $\alpha^1_{r;\nu,k}(y)$ and $\hat{\alpha}^2_{r;\nu,k}(\xi)$ in the separated representation (2.4) need to be found for each orientation ν and scale k. Obviously, the representation should have small separation rank R, and should be obtainable with low computational complexity. A fast numerical procedure for addressing this issue is proposed in the following Section 3.
2. **Evaluation of coordinate transform.** The sum over ξ in (2.6) corresponds to a non-equispaced (inverse) Fourier transform from the discrete set ξ to space points which are, for

[3] With slight abuse of notation, we will continue to write ξ and y for these discrete sets.

each k and ν, determined by the coordinate transform $T_{\nu,k}(y)$.
Currently, this is implemented in a straight forward "brute force" manner by interpolation in the input domain: First, regular input and output grids x and y are defined. Then, $T_{\nu,k}$ is replaced by x in (2.6), such that the sum over ξ can be obtained as an inverse wave packet transform using standard algorithms. Finally, by interpolating and resampling on the irregular sample points determined by $\tilde{x} = T_{\nu,k}(y)$, one obtains the image. This procedure is computationally expensive, and the interpolation scheme currently used introduces numerical artifacts (cf. Section 4).
In Section 5, a strategy for resolving this issue in the context of computing the approximate action of F on a function composed of wave packets will be proposed, which is at present not implemented.

3. Approximate Action on a Wave Packet: Tensor Product Representation.

The amplitude term $a(y, \nu)$ on the left hand side of (2.4) does not depend on ξ, can hence always be trivially accounted for in the final separation and will not further discuss it in this section.
The argument of the complex exponential in (2.4) is separated in y and ξ, and the left hand side can hence be arranged as a 2D matrix $[\Gamma]$ with row and column indices m and n pointing at the discrete sets y_m and ξ_n, respectively. Hence, a discrete tensor product representation as in (2.4) can always be obtained by truncation of the singular value decomposition (SVD) of $[\Gamma]$,

$$[\Gamma] = [U][\Lambda][V]^* = [B_1][B_2], \qquad [B_1] = [U], \quad [B_2] = [\Lambda][V]^*,$$

to the eigenvectors in $[U]$ and $[V]$ that correspond to the eigenvalues above a certain threshold that determines the precision of the truncated decomposition. This generic approach is computationally expensive, since it requires the computation of the full SVD of $[\Gamma]$, while only a small number R of eigenvector-eigenvalue-couples will have significant contribution to the final expansion: Heuristically, the highly oscillatory parts of $\exp(iS(y, \xi))$ are captured by the first order term of the Taylor expansion of S (i.e., by $T_{\nu,k}(y)$), such that the remainder (2.4) in the approximation is not oscillatory and highly structure.
In the context of FIO approximation, Candès et al. [3] proposed to exploit the structure (i.e. low rank) of a complex exponential matrix in a similar problem by an algorithm based on random subsets of its rows and columns, which in average reduces the size of the SVD problem. Their algorithm is based on work by Kapur et al. [10][4].
Here, we propose an alternative to SVD based approaches that is tailored to directly match the specific form of (2.4). The strategy consists in first constructing a computationally inexpensive, but potentially large rank, expansion of the complex exponential using orthogonal polynomials, which is then used in a fast alternative least square (ALS) algorithm with low complexity, recently introduced in the literature [1], for finding a representation with fewer terms.

Separate Representation using Hermite Polynomials. The left hand side of (2.4) is of the form $e^{f(\xi)\cdot g(y)}$, with $f(y) = \frac{\partial^2 S}{\partial \xi_2^2}(y, \nu)$ and $g(\xi) = i\frac{1}{2}\frac{\xi_2^2}{\xi_1}$. By Mehler's Hermite polynomial formula (e.g. [13]):

$$\sum_{s=0}^{\infty} \frac{H_s(x)H_s(y)}{s!}\left(\frac{w}{2}\right)^s = (1-w^2)^{-1/2} \exp\left(\frac{2xyw - (x^2+y^2)w^2}{1-w^2}\right),$$

[4] See e.g. [9] for an overview of deterministic methods for reducing the complexity of the SVD for matrices with certain specific structures

one can write:

$$e^{f(\xi)g(y)} = e^{\frac{w}{2}f(y)^2} e^{\frac{w}{2}g(\xi)^2} \sum_{s=0}^{\infty} \frac{w^s}{2^s s! \sqrt{1-w^2}} H_s\left(\zeta f(y)\right)\, H_s\left(\zeta g(\xi)\right) \tag{3.1}$$

$$= \sum_{s=0}^{\infty} \underbrace{\frac{w^s}{2^s s! \sqrt{1-w^2}} e^{\frac{w}{2}f(y)^2} H_s\left(\zeta f(y)\right)}_{\beta_s^1(y)} \underbrace{e^{\frac{w}{2}g(\xi)^2} H_s\left(\zeta g(\xi)\right)}_{\hat{\beta}_s^2(\xi)} = \sum_{s=0}^{\infty} \beta_s^1(y)\hat{\beta}_s^2(\xi), \tag{3.2}$$

where $H_s(x) = (-1)^s e^{x^2} \frac{d^s}{dx^s} e^{-x^2}$ are Hermite polynomials of order s, and $\zeta = \sqrt{\frac{1-w^2}{2w}}$. By truncating the infinite sum to within an a priori fixed numerical precision ε, one obtains the approximate separated (tensor product) representation

$$\sum_{s=0}^{\infty} \beta_s^1(y)\hat{\beta}_s^2(\xi) = \sum_{s=0}^{S} \beta_s^1(y)\hat{\beta}_s^2(\xi) + O(\varepsilon). \tag{3.3}$$

Representation (3.3) makes explicit use of the specific form of (2.4). It involves obtaining the coefficients of the polynomials H_s, which can be pre-calculated and stored, and evaluating the polynomials, which requires $\sim S$ operations per y and ξ point.

Rank Reduction. The rank S of representation (3.3) is in general too large for the practical purpose considered here[5]. Yet, it can be used efficiently for finding the desired representation with smaller separation rank R:

$$\sum_{r=0}^{R} \alpha_r^1(y)\hat{\alpha}_r^2(\xi), \qquad R \ll S. \tag{3.4}$$

This can be done using a specific ALS algorithm, originally introduced in [1] for separated representation based algorithms in high dimension. The algorithm exploits the (multi-)linear structure of (3.3) for reducing the nonlinear least squares problem:

$$|| \sum_{s=0}^{S} \beta_s^1(y)\hat{\beta}_s^2(\xi) - \sum_{r=0}^{R} \alpha_r^1(y)\hat{\alpha}_r^2(\xi) || < \varepsilon, \qquad R \ll S \tag{3.5}$$

to iterations over linear least squares problems per dimension. It has complexity $O(dR(R^2+SN)$ per iteration, where d is the dimensionality of the problem ($d=2$ in (3.5)). The reader is referred to [1] for details.

4. Numerical Example – CO Isochrone in homogeneous medium. We aim at constructing an isochrone in a common offset (CO) setup with small offset ($h/z \approx 0.05$) in homogeneous medium using the procedure proposed in the previous section. The synthetic data we consider consist of one single point at $(x=0, vt/2=2100)$, which we imagine to be obtained by a superposition of infinitesimally small line elements with slopes p_u within a certain range $[p_u - \Delta_p, p_u + \Delta_p]$. A reference image is obtained using explicit $2D$ CO map time migration expressions ([5] and references therein). The synthetic data and the reference image are shown in Fig. 1 left and center, respectively.

The data are modeled by restricting the set of wave packets obtained by decomposition of the synthetic data to those at position $(x=0, vt/2=2100)$ and with orientations (in degrees from horizontal) $[p_u - \Delta_p, p_u + \Delta_p] = [63\,^\circ, 117\,^\circ]$, at all available scales $k \le k_{max}$ (cf. Fig. 1, right). The generating function $S(y,\xi)$ and the amplitude $a(y,\xi)$ for this CO problem setup are provided in [5].

[5] For CO imaging in homogeneous medium, it is of the order of 100 for $\varepsilon = 10^{-6}$, whereas, for comparison, the representation obtained by truncated SVD involves of the order of 10 terms only for the same numerical precision.

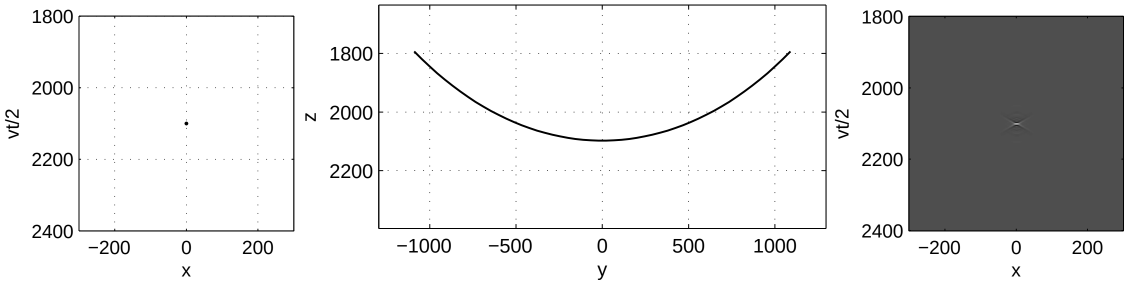

Fig. 1. *CO synthetic data (left), reference image (center), and discrete low-resolution data model obtained by decomposition of the synthetic data into wave packets (right).*

4.1. Numerical Setup. The data consist of synthetic 256×256 real-valued 2D images, which leads to $k_{max} = 5$ available scales. The numerical precisions in the Hermite expansion and rank reduction steps are set to $\varepsilon = 10^{-6}$, leading to an overall precision of $2 \cdot 10^{-6}$ in the tensor product representation and, on average over scales and orientations, to $R = 12$. The algorithm for 2D wave packet decomposition/recomposition and for the tensor product representation presented in Section 3 have been implemented by ourselves in MATLAB and C.

4.2. Results. We compare the approximation (2.6) (i.e., coordinate transform and alteration of the wave packet) to the approximation (2.2) (i.e., coordinate transform $T_{\nu,k}(y)$ only).
Image of a single input wave packet. Fig. 2 shows one single input wave packet at scale $k = 3$ (top), the images obtained by approximation (2.2) (center row, left) and approximation (2.6) (bottom row, right), and the corresponding amplitudes along the isochrone (bottom row): The coordinate transform $T_{\nu,k}$ involved in both approximations propagates the wave packet to the correct position in the image. Yet, with approximation (2.6), the wave packet is in addition bending and spreading along the isochrone (cf. Fig. 2, center and bottom row).
Imaging the isochrone. Fig. 3 shows the final imaging result, obtained as the superposition of the images of all input wave packets: Images (top row) and amplitudes along the isochrone (bottom row) for approximation (2.2) (left column) and (2.6) (right column). The image in the left column has large gaps along the isochrone, resulting from the fact that it is obtained purely by propagation of wave packets along rays according to geometrical optics. Within approximation (2.2), the image of a singularity at large time / depth eventually breaks apart into its constituting wave packets. Similar observations have been obtained in [5]. In contrast, approximation (2.6) produces a very satisfactory image (right column): The contributing wave packets are bending to follow the isochrone and spread out such that they connect. As a result, the corresponding amplitude does not contain any significant gaps. The smaller high frequency fluctuations on top of the overall trend of the amplitude is an artifact that is entirely caused by poor interpolation in the current coordinate transform implementation (cf. Section 2, last paragraph).

5. Future Work: Approximate action on a function – discrete implementation. We finally discuss a potential strategy for the discrete approximation of the action of F on an input function composed of wave packets.
Suppose that we are given input data $u(x)$ for which we have obtained a decomposition into wave packets, i.e.:

$$u(x) = \sum_{\gamma} c_\gamma \varphi_\gamma(x) = \sum_{\gamma} \langle \psi_\gamma, u \rangle \varphi_\gamma(x) = \sum_{\xi} \left[\sum_{\nu,k} e^{i\langle x,\xi\rangle} \hat{u}(\xi) \hat{\beta}_{\nu,k}(\xi) \hat{\chi}_{\nu,k}(\xi) \right], \tag{5.1}$$

where $\{\hat{\beta}_{\nu,k}(\xi)\}$ and $\{\hat{\chi}_{\nu,k}(\xi)\}$ are the collections of window functions for defining a co-frame of wave packets (cf. [6] and references therein). Then, the action of the operator F on u can be written as

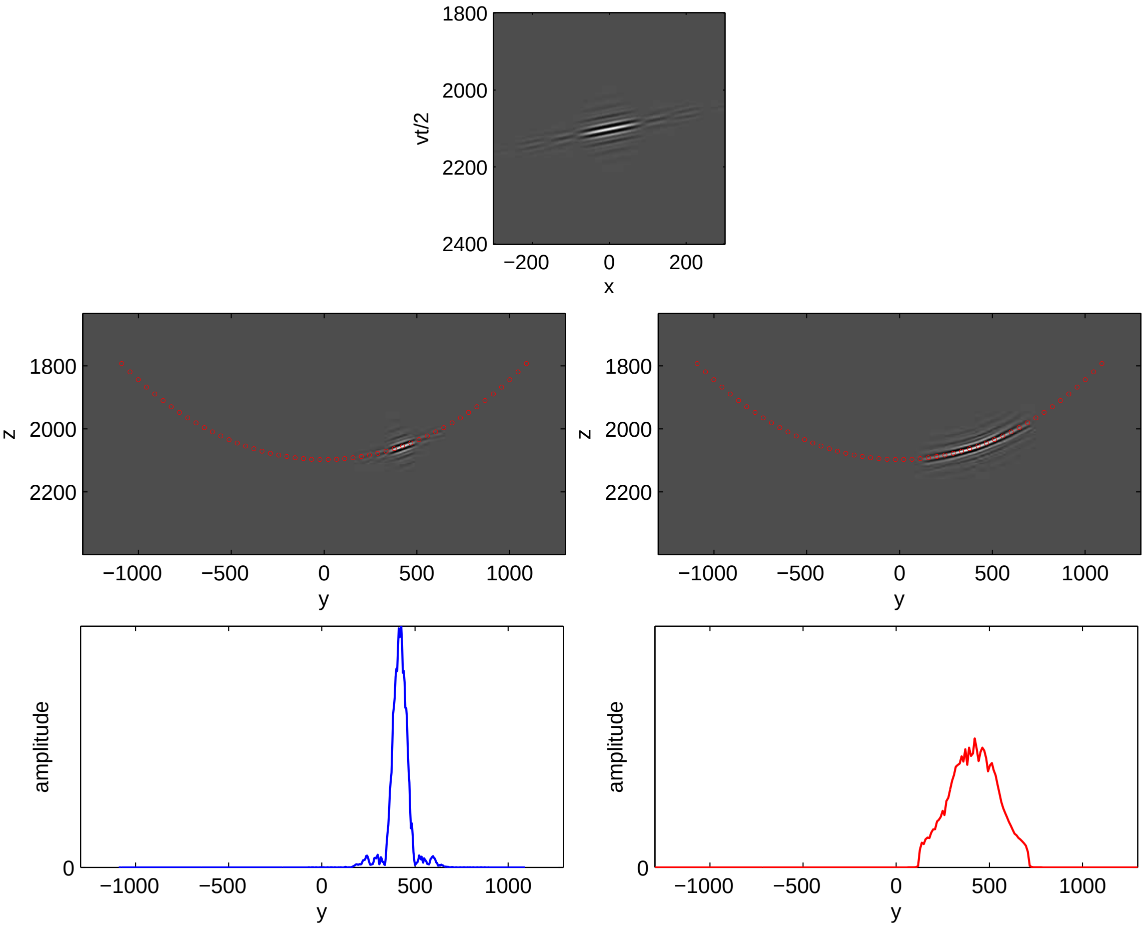

Fig. 2. *Single input wave packet at scale $k = 3$ (top), images of this wave packet (center row) and amplitudes along the isochrone (bottom row), obtained by (2.2) (left column) and (2.6) (right column).*

the sum of the action of F on each single wave packet φ_γ, weighted by the coefficients c_γ:

$$(Fu)(y) = \sum_\gamma c_\gamma (F\varphi_\gamma)(y) = \sum_{\nu,k} \sum_j c_{j,\nu,k} (F\varphi_{j,\nu,k})(y) \tag{5.2}$$

By Theorem 2.1, we have:

$$(Fu)(y) = \sum_{\nu,k} \sum_{r=1}^{R} \alpha^1_{r;\nu,k}(y) \sum_\xi \left[e^{i\langle T_{\nu,k}(y),\xi\rangle} \hat{\alpha}^2_{r;\nu,k}(\xi) \hat{u}(\xi) \hat{\beta}_{\nu,k}(\xi) \hat{\chi}_{\nu,k}(\xi) \right] + O(2^{-k_{max}/2}), \tag{5.3}$$

since neither the functions $\alpha^{\{1,2\}}_{r;\nu,k}$ nor the coordinate transform $T_{\nu,k}(y)$ depend on the index j which determines the precise position of the wave packets in the dual domain. Hence, from a computational point of view, approximating the action of F on the function $u(x)$ by wave packets induces the following differences with respect to (5.1) (i.e., the recomposition of the function $u(x)$ from its constituting wave packets):

1. An additional sum over the elements of the tensor product representation, which increases the computational complexity by a factor R. The sum over r can be taken either inside or outside of the sum $\sum_{\nu,k}$.

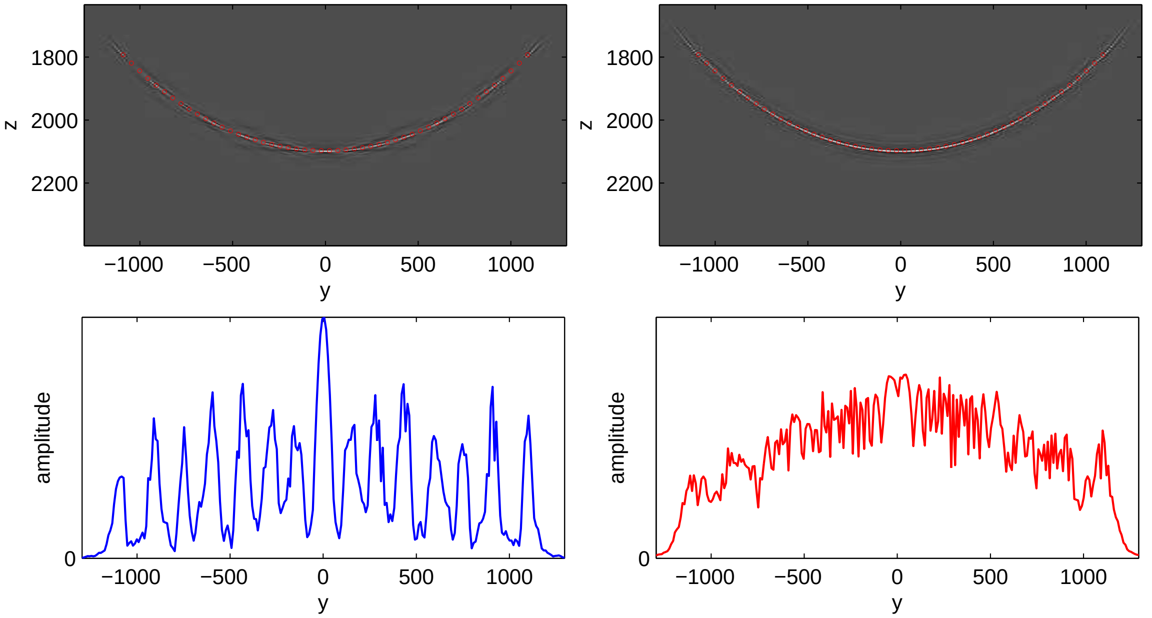

FIG. 3. *Isochrone images (top row) and amplitudes along the isochrone (bottom row), obtained by (2.2) (left column) and (2.6) (right column).*

A - Initialization	– fix numerical precision – define regular discrete grid on image domain y
B - Action per ν, k	For each orientation ν with non-zero coefficients c_γ at scale k: 1. Fix image region y corresponding to data domain x of interest 2. Evaluate $\frac{\partial S}{\partial \xi}(y,\nu)$ and $\frac{\partial^2 S}{\partial \xi_2^2}(y,\nu)$ 3. Obtain tensor product representation (2.4): a) Hermite polynomial expansion (3.3) b) rank reduction (3.5) 4. Evaluate $\sum_\xi [\,\cdot\,]$ in (5.3) for each tensor product term r 5. Evaluate $\sum_r$ in (5.3)
C - Action $(Fu)(y)$	$\sum_{\nu,k}$ in (5.3): Add up the contributions of each ν, k to the final image.

TABLE 5.1
Outline of a parallel numerical scheme for evaluating (5.3).

2. Due to the (ν,k)-dependence of the coordinate transform $T_{\nu,k}(y)$ and the amplitudes $\alpha^1_{r;\nu,k}(y)$ in (5.3), the non-equidistant Fourier transform $\sum_\xi [\,\cdot\,]$ can not be calculated over the entire ξ-domain at once, but needs to be calculated for each (ν,k) separately. The sum over ξ has to be calculated only for ξ points within the support of the wave packets at (ν,k), yet still needs to be evaluated for each y. This increases the computational complexity by a factor proportional to the number of (ν,k)-pairs, i.e., in 2D, approximately by $\sqrt{N}$.

Outline for a parallel computational scheme. Tab. 5.1 summarizes the main steps of a scheme for evaluating (5.3). All of the computationally costly operations are performed in step **B**, independently for each ν, k, and can hence be performed using massive parallelization. Step **A** merely consists in setting up a discrete target grid in the image domain by fixing a reference point y_0 and target resolution dy, such that the final image can be obtained as a simple sum of the contributions from different ν, k in step **C**.

Since wave packets with different orientation ν do not propagate to the same image region, in step

B1, the image region corresponding to data wave packets at ν, k is determined in order to decrease the size of the y domain to be considered for each orientation. This could be performed either by an iterative search for y points for which $T_{\nu,k}(y)$ correspond to the borders of the data domain x, or by inversion of $T_{\nu,k}(y)$. A numerical scheme for step **B**3 has been presented in Section 3. For the evaluation of the non-equispaced Fourier transform in step **B**4, several methods with complexity close to that of a standard FFT have been proposed in the literature (cf., e.g., [7, 8]).

6. Conclusion and perspectives. In this note, numerical approaches for practically using recent theoretical work in [4] for approximating the action of GRT operators using wave packets have been proposed and discussed. Notably, a numerical procedure for obtaining the approximate low rank tensor product representation of a complex exponential has been proposed, which turns out to be one of the key steps in the approximation obtained in [4]. The proposed procedure has been used here to provide a first numerical example that illustrates the effectiveness of the result Theorem 4.1 in [4] in practice. Similar results, not presented here, have been obtained for the time-reversal of a wave fronts in the presence of multi-pathing, and for modeling the propagation of wave packets for real-world data from a physical experiment conducted by L. Pyrak-Nolte Department of Physics, Purdue University. Finally, a parallel scheme for numerically approximating the action of GRT operators on input functions composed of wave packets has been outlined. This is currently being investigated.

REFERENCES

[1] G. Beylkin and M.J. Mohlenkamp, *Algorithms for numerical analysis in high dimensions*, SIAM J. Sci. Comput., 26 (2005), pp. 2133–2159.

[2] E. Candès, L. Demanet, D. Donoho, and L. Ying, *Fast discrete curvelet transforms*, SIAM Multiscale Model. Simul., 5 (2006), pp. 861–899.

[3] E. Candès, L. Demanet, and L. Ying, *Fast computation of fourier integral operators*, SIAM J. Sci. Comput., 29 (2007), pp. 2464–2493.

[4] 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, 25 (2009), pp. 025005+.

[5] H. Douma and M.V. de Hoop, *Leading-order seismic imaging using curvelets*, Geophysics, 72 (2007), pp. 231–248.

[6] A. Duchkov, F. Andersson, and M.V. de Hoop, *Discrete almost symmetric wave packets and multi-scale representation of (seismic) waves*, preprint, (2009).

[7] A. Dutt and V. Rokhlin, *Fast fourier transforms for nonequispaced data*, SIAM J. Sci. Comput., 14 (1993), pp. 1368–1393.

[8] ———, *Fast fourier transforms for nonequispaced data II*, Appl. Comput. Harmon. Anal., 2 (1995), pp. 85–100.

[9] G.H. Golub and C.F. Van Loan, *Matrix Computations (3rd ed.)*, Series in Mathematical Science, Johns Hopkins University Press, Baltimore, MD, USA, 1994.

[10] S. Kapur and D.E. Long, *Ies3: A fast integral equation solver for efficient 3-dimensional extraction*, in Proc. IEEE/ACM Int. Conf. Computer Aided Design, Washington, DC, 1997, pp. 448–455.

[11] H. Smith, *A parametrix construction for wave equations with $c^{1,1}$ coefficients*, Ann. Inst. Fourier, Grenoble, 48 (1998), pp. 797–835.

[12] C.C. Stolk and M.V. de Hoop, *Microlocal analysis of seismic inverse scattering in anisotropic, elastic media*, Comm. Pure Appl. Math., 55 (2002), pp. 261–301.

[13] G.N. Watson, *Notes on generating functions of polynomials: (2) hermite polynomials*, J. London Math. Soc., 8 (1933), pp. 194–199.