MULTISCALE MODEL. SIMUL.
Vol. 8, No. 5, pp. 1803–1837

FAST MULTISCALE GAUSSIAN WAVEPACKET TRANSFORMS AND MULTISCALE GAUSSIAN BEAMS FOR THE WAVE EQUATION*

JIANLIANG QIAN† AND LEXING YING‡

Abstract. We introduce a new multiscale Gaussian beam method for the numerical solution of the wave equation with smooth variable coefficients. The first computational question addressed in this paper is how to generate a Gaussian beam representation from general initial conditions for the wave equation. We propose fast multiscale Gaussian wavepacket transforms and introduce a highly efficient algorithm for generating the multiscale beam representation for a general initial condition. Starting from this multiscale decomposition of initial data, we propose the multiscale Gaussian beam method for solving the wave equation. The second question is how to perform long time propagation. Based on this new initialization algorithm, we utilize a simple reinitialization procedure that regenerates the beam representation when the beams become too wide. Numerical results in one, two, and three dimensions illustrate the properties of the proposed algorithm. The methodology can be readily generalized to treat other wave propagation problems.

Key words. fast multiscale Gaussian wavepacket transforms, multiscale Gaussian beams, wave equations

AMS subject classifications. 65N30, 65M60

DOI. 10.1137/100787313

1. Introduction.

This paper is concerned with the numerical solution of the linear wave equations

$$\begin{aligned} U_{tt} - V^2(x)\Delta U = 0, \quad x \in \mathbb{R}^d, \ \ t > 0, \\ U|_{t=0} = f_1(x), \\ U_t|_{t=0} = f_2(x), \end{aligned}$$

where the velocity function V is smooth, positive, and bounded away from zero; the functions $f_1(x)$ and $f_2(x)$ are compactly supported L^2-functions, presumably highly oscillatory.

When the initial condition contains oscillations of a small wavelength, the wave equation propagates these oscillations in space and time. Resolving such small oscillations by direct numerical methods, such as finite-difference, finite-element, or even spectral methods, requires an enormous computational grid and is very costly in practice. Therefore, methods based on geometrical optics are sought as an alternative numerical approximation for capturing such highly oscillatory phenomena.

One of such geometrical-optics methods is the so-called Gaussian beam method. The traditional geometrical-optics method starts from the Wentzel–Kramers–Brillouin–Jeffreys (WKBJ) ansatz consisting of a real phase function that solves an

*Received by the editors March 2, 2010; accepted for publication (in revised form) July 19, 2010; published electronically October 14, 2010.

http://www.siam.org/journals/mms/8-5/78731.html

†Department of Mathematics, Michigan State University, East Lansing, MI 48824 (qian@math.msu.edu). This author is partially supported by the National Science Foundation (grants CCF-0830161 and DMS-0810104) and by AFOSR grant FA9550-04-1-0143.

‡Department of Mathematics and ICES, University of Texas at Austin, Austin, TX 78712 (lexing@math.utexas.edu). This author is partially supported by a Sloan Research Fellowship and NSF CAREER Award DMS-0846501.

eikonal equation and a real amplitude function that solves a transport equation. The WKBJ ansatz breaks down at the caustics, in the neighborhood of which the phase function is multivalued and the amplitude function blows up. The Gaussian beam method, however, constructs a global *complex* phase and a global *complex* amplitude that satisfy the eikonal equation and the transport equation, respectively, *up to certain specified orders near a specified ray path*. Away from the ray path, the quadratic imaginary part of the phase function provides a rapidly decaying Gaussian profile. Consequently, this gives rise to a single Gaussian-like asymptotic solution, which is accurate near the underlying ray path. Since the wave equation is linear, a superposition of such single-beam solutions yields a global asymptotic solution.

In order to apply the Gaussian beam methods efficiently and accurately, one faces two major computational problems. The first one is how to generate a beam decomposition for a general initial condition. There have been some recent advances to overcome this obstacle in the settings of the wave equation [20, 12, 18, 19] and the Schrödinger equation [11]. Most recently in [15], we introduced the single-scale Gaussian wavepacket transforms and developed on top of them a highly efficient initialization algorithm for the Schrödinger equation. In this paper, we extend this line of research to the wave equation. Since the Hamiltonian of the wave equation is homogeneous of degree one, which essentially constrains the resulting Hamiltonian flow to the cosphere bundle, a Gaussian beam of the wave equation should satisfy the parabolic scaling principle [17, 3] at any given time. Motivated by this principle, we introduce a new set of multiscale Gaussian wavepacket transforms that enables us to decompose arbitrary initial conditions of the wave equation and polarize mixed high-frequency initial data into different wave modes at multiscale resolutions. Finally, based on this multiscale decomposition, we propose the multiscale Gaussian beam method that solves the wave equation with general initial conditions.

The second computational problem addressed in this paper is how to carry out long-term beam propagation for the wave equation. A simple analysis reveals that the width of a Gaussian beam in the case of the wave equation is related to the derivatives of the underlying velocity, and a beam width may grow exponentially during the evolution process. This implies that the beam loses its localized significance, leading to deteriorating accuracy in the Taylor expansion for the phase function and high cost in the beam summation. Following [15], we propose to reinitialize the beam propagation using the fast multiscale Gaussian wavepacket transforms when the beams become too wide. This controls the widths of the Gaussian beams and enables one to propagate the wave efficiently and accurately for a long period of time.

Furthermore, we show that our new multiscale Gaussian beam method yields a global asymptotic solution for the wave equation.

1.1. Related work. The idea underlying Gaussian beams is simply to build asymptotic solutions to PDEs concentrated on a single curve through the domain; this single curve is nothing but a ray as shown in [16]. The existence of such solutions has been known to the pure mathematics community since sometime in the 1960s [1], and these solutions have been used to obtain results on propagation of singularities in hyperbolic PDEs [10, 16]. An integral superposition of these solutions can be used to define a more general solution that is not necessarily concentrated on a single curve. Gaussian beams can be used to treat pseudodifferential equations in a natural way, including Helmholtz [12] and Schrödinger equations [11].

Gaussian beam superpositions have been used in geophysical applications for seismic wave modeling [6] and migration [9]. The numerical implementations in these

areas are based on ray-centered coordinates which prove to be computationally inefficient [6, 9]. More recently, based on [16, 20], a purely Eulerian computational approach was proposed in [12] which overcomes some of these difficulties; it can be easily applied to both high frequency waves and semiclassical quantum mechanics [11]. In [20] Lagrangian Gaussian beams are successfully constructed to simulate mountain waves, a kind of stationary gravity wave forming over mountain peaks and interfering with aviation. See [14, 2] for other recent works and see also [12, 11] for further references.

From the work in [3, 17], it is well known that, for the wave equation with smooth coefficients, a wavepacket remains a wavepacket at a later time if it satisfies the so-called parabolic scaling principle, i.e., the wavelength of the typical oscillation of the wavepacket being equal to the square of the width of the wavepacket. Examples of such wavepackets include curvelets [5, 4] and wave atoms [7, 8]. The multiscale Gaussian wavepackets introduced here are inspired by these constructions and in fact inherit the overall architecture of the wave atom construction. However, the transforms are modified appropriately so that the wavepackets maintain a Gaussian profile.

1.2. Contents. The rest of the paper is organized as follows. Section 2 introduces the Lagrangian formulation of the Gaussian beam method of the wave equation. In section 3, we discuss the fast multiscale Gaussian wavepacket transforms in detail. Section 4 describes how to polarize the initial condition and generate a Gaussian beam representation of the initial condition based on the tools introduced in section 3. Section 5 presents the reinitialization procedure for long time propagation. Section 6 shows that the solution obtained by our algorithm is a global asymptotic solution of the wave equation with the given initial condition. Extensive numerical results are presented in section 7. Finally, we conclude with some further discussions in section 8.

2. Gaussian-beam setup for the wave equation. We consider the following wave equation:

$$U_{tt} - V^2(x)\Delta U = 0, \quad x \in \mathbb{R}^d, \ \ t > 0, \tag{1}$$

$$U|_{t=0} = f_1(x), \tag{2}$$

$$U_t|_{t=0} = f_2(x), \tag{3}$$

where the velocity function $V(x)$ is smooth, positive, and bounded away from zero. The functions $f_1(x)$ and $f_2(x)$ are compactly supported functions in $L^2(\mathbb{R}^d)$.

We are looking for asymptotic solutions of the wave equation in geometrical-optics form,

$$A(x,t)e^{\imath\omega\tau(x,t)}, \tag{4}$$

where $\tau(x,t)$ is the phase function, $A(x,t)$ the amplitude function, and $\imath = \sqrt{-1}$. In the ansatz (4), the frequency ω is a large parameter, and an asymptotic solution for the wave equation is sought in the sense that the wave equation (1) and its associated initial conditions (2) and (3) are satisfied approximately with a small error when ω is large. After substituting the ansatz (4) into the wave equation (1) and considering the leading orders in inverse powers of the large parameter ω, we end up with the following eikonal and transport equations:

$$\tau_t^2 - V^2(x)|\nabla_x \tau(x,t)|^2 = 0, \tag{5}$$

$$2A_t\tau_t - 2V^2\nabla_x A \cdot \nabla_x \tau + A(\tau_{tt} - V^2 \mathrm{trace}(\tau_{xx})) = 0. \tag{6}$$

Factorizing the eikonal equation (5) gives

$$\tau_t^{\pm} + G^{\pm}(x, \nabla_x \tau(x,t)) = 0, \tag{7}$$

where $G^{\pm}(x, \nabla_x \tau(x,t)) = \pm V(x)|\nabla_x \tau(x,t)|$ correspond to two polarized wave modes in the second-order wave equation. Accordingly, we define the Hamiltonians,

$$G^{\pm}(x,p) = \pm V(x)|p|,$$

where $G^{\pm}(x,p)$ is clearly homogeneous of degree one in the momentum variable p.

To construct asymptotic solutions for the wave equation, we are going to use Gaussian beams [16, 13, 20]. Because the two polarized wave modes will be treated essentially in the same way, we consider the following generic situation for the eikonal equation:

$$\tau_t + G(x, \nabla_x \tau(x,t)) = 0, \tag{8}$$

where G can be taken to be either G^+ or G^- and τ to be either τ^+ or τ^-. According to the Gaussian beam theory [16, 13, 20], a single Gaussian beam is an asymptotic solution to the wave equation, and it is concentrated near a ray path which is the x-projection of a certain bicharacteristic. To construct a bicharacteristic, we apply the method of characteristics to the eikonal equation (8) to obtain the following Hamiltonian system:

$$\begin{aligned} \dot{x} &= \frac{dx}{dt} = G_p, \quad x|_{t=0} = x_0, \\ \dot{p} &= \frac{dp}{dt} = -G_x, \quad p|_{t=0} = p_0, \end{aligned} \tag{9}$$

where t is time parameterizing bicharacteristics. Solving this system yields the bicharacteristic

$$\{(x(t), p(t)) : t \geq 0\},$$

which emanates from the initial point (x_0, p_0) in phase space at $t = 0$. The corresponding ray path is $\gamma = \{(x(t), t) : t \geq 0\}$, which is defined in the (x,t)-space. Notice that along the ray path $\gamma = \{(x(t), t) : t \geq 0\}$, we have by construction $p(t) = \tau_x(x(t), t)$ due to the method of characteristics. Furthermore, the phase function $\tau(x(t), t)$ along the ray path satisfies

$$\frac{d\tau(x(t),t)}{dt} = \tau_t(x(t),t) + p(t) \cdot G_p(x(t), p(t)) = \tau_t(x(t),t) + G(x(t), \tau_x(x(t),t)) = 0,$$

which implies that the phase function $\tau(x(t), t)$ does not change along γ because the Hamiltonian G is homogeneous of degree one; we will take $\tau(x(t), t) = 0$.

So far we have computed the phase function τ and its first-order derivative $p(t) = \tau_x(x(t), t)$ along the ray path $\gamma = \{(x(t), t) : t \geq 0\}$. To construct a second-order Taylor expansion for the phase function along the ray path, one needs to compute the Hessian of the phase along the ray. Following [16, 20], we differentiate the eikonal equation (8) with respect to t and x near the ray path γ:

$$\tau_{t,x}(x,t) + G_x(x, \tau_x(x,t)) + \tau_{xx}(x,t) G_p(x, \tau_x(x,t)) = 0, \tag{10}$$

$$\tau_{t,t}(x,t) + G_p(x, \tau_x(x,t)) \cdot \tau_{x,t}(x,t) = 0. \tag{11}$$

Differentiating equation (10) further with respect to x yields

$$\tau_{t,xx}(x,t) + G_{xx}(x,\tau_x(x,t)) + \tau_{xx}G_{xp} + G_{xp}^T\tau_{xx} + \tau_{xx}G_{pp}\tau_{xx} + \tau_{xxx}G_p = 0. \tag{12}$$

Let $M(t) = \tau_{xx}(x(t),t)$. Since (10), (11), and (12) are valid along the ray path, combining the first term and the last term in (12) yields the following Riccati equation for $M(t)$:

$$\frac{dM(t)}{dt} + G_{xx} + M(t)G_{xp} + G_{xp}^T M(t) + M(t)G_{pp}M(t) = 0, \tag{13}$$

which is appended with an initial condition $M|_{t=0} = M_0 = \imath\epsilon I$, where ϵ is a positive number of order $O(1)$. In addition, once $M(t)$ is available, the mixed derivative $\tau_{t,x}$ along the ray path γ can be computed by

$$\tau_{t,x}(x(t),t) = -G_x(x(t),p(t)) - M(t)G_p(x(t),p(t)); \tag{14}$$

consequently, the second-order time derivative $\tau_{t,t}$ in (11) along the ray path γ can be computed by

$$\tau_{t,t}(x(t),t) = G_p(x(t),p(t)) \cdot G_x(x(t),p(t)) + G_p^T(x(t),p(t))M(t)G_p(x(t),p(t)). \tag{15}$$

Equations (13), (14), and (15) allow us to compute the second-order derivative (the Hessian) of the phase function τ along γ once a suitable initial condition is assigned to (13).

Although the Riccati equation (13) does not admit a global smooth solution in general, it turns out that complexifying the equation by specifying a complex initial value will guarantee that a global smooth solution exists because of the underlying symplectic structure associated with the related Hamiltonian system; see [16, 13, 20] for theoretical justification. Generally, there is a high probability for so-called transmission caustics to occur in inhomogeneous media [21]; therefore, it is critical for a numerical method to be capable of treating caustics automatically. Moreover, we have the following lemma [16, 13, 20].

LEMMA 2.1. *If the Hamiltonian G is smooth enough, then the Hessian $M(t)$ along the ray path γ has a positive-definite imaginary part, provided that it initially does.*

As we will see, Lemma 2.1 is significant for the Gaussian beam construction because it guarantees that an initial Gaussian profile with small variances will propagate along the ray path if such an initial Gaussian profile is chosen properly to peak at the initial point of the ray path.

Now with the Hessian of the phase function at our disposal, we may solve the transport equation (6) for the amplitude $A(t) = A(x(t),t)$ along the ray path γ. Since $\tau_t(x(t),t) = -G(x(t),p(t))$ along the ray path, the transport equation (6) is reduced to the following:

$$A_t(x(t),t) + \frac{V^2(x(t))A_x \cdot p(t)}{G} + \frac{A}{2G}\left(V^2(x(t))\mathrm{trace}(M(t)) - \tau_{tt}\right) = 0. \tag{16}$$

Since $\frac{dx}{dt} = \frac{V^2(x(t))}{G(x(t),p(t))}p(t)$, we have

$$\frac{dA}{dt} + \frac{A(x(t),t)}{2G}\left(V^2(x(t))\mathrm{trace}(M(t)) - G_x \cdot G_p - G_p^T M(t)G_p\right) = 0, \tag{17}$$

which is appended with a suitable initial condition $A|_{t=0} = A_0$.

At this stage, we are ready to construct a single Gaussian beam along the ray path γ by defining the following two global, smooth approximate functions for the phase and amplitude:

$$\tau(x,t) \equiv p(t)\cdot(x-x(t)) + \frac{1}{2}(x-x(t))^T M(t)(x-x(t)), \tag{18}$$

$$A(x,t) \equiv A(x(t),t) = A(t), \tag{19}$$

which are accurate near the ray path $\gamma = \{(x(t),t) : t \geq 0\}$. These two functions allow us to construct a single-beam asymptotic solution

$$\Phi(x,t) = A(x,t)\exp(\imath\omega\tau(x,t)). \tag{20}$$

This beam solution is concentrated on a single smooth curve $\gamma = \{(x(t),t) : t \geq 0\}$, which is the x-projection of the bicharacteristic $\{(x(t),p(t)) : t \geq 0\}$ emanating from (x_0,p_0) at $t = 0$. Because the phase $\tau(x,t)$ has an imaginary part, $\mathrm{Im}(\tau(x,t)) = \frac{1}{2}(x-x(t))^T\mathrm{Im}(M(t))(x-x(t))$, $\Phi(x,t)$ has a Gaussian profile of the form

$$\exp\left(-\frac{\omega}{2}(x-x(t))^T\mathrm{Im}(M(t))(x-x(t))\right),$$

which is concentrated on the smooth ray path γ.

Applying the above construction to the two polarized modes with $G = G^{\pm}$ results in two sets of solutions $x^{\pm}(t)$, $p^{\pm}(t)$, $M^{\pm}(t)$, $A^{\pm}(t)$, $\tau^{\pm}(x,t)$, $A^{\pm}(x,t)$, and $\Phi^{\pm}(x,t)$. These functions are uniquely determined by the initial data x_0, p_0, M_0, and A_0. We denote these initial data collectively by a tuple $\alpha = (x_0,p_0,M_0,A_0)$. In the rest of this paper, in order to emphasize the dependence on α, the solutions are denoted, respectively, by $x^{\pm}_{\alpha}(t)$, $p^{\pm}_{\alpha}(t)$, $M^{\pm}_{\alpha}(t)$, $A^{\pm}_{\alpha}(t)$, $\tau^{\pm}_{\alpha}(x,t)$, $A^{\pm}_{\alpha}(x,t)$, and $\Phi^{\pm}_{\alpha}(x,t)$.

For a given tuple $\alpha = (x_0,p_0,M_0,A_0)$, the Gaussian beams $\Phi^{\pm}_{\alpha}(x,t)$ have a simple Gaussian envelope. For a general initial condition $(U(x,0),U_t(x,0))$, one needs to find two sets I^+ and I^- of tuples such that at time $t = 0$

$$U(x,0) \approx \sum_{\alpha\in I^+}\Phi^+_{\alpha}(x,0) + \sum_{\alpha\in I^-}\Phi^-_{\alpha}(x,0),$$
$$U_t(x,0) \approx \sum_{\alpha\in I^+}\Phi^+_{\alpha,t}(x,0) + \sum_{\alpha\in I^-}\Phi^-_{\alpha,t}(x,0).$$

Once this initial decomposition is given, the linearity of the wave equation gives the Gaussian beam solution

$$U(x,t) \approx \sum_{\alpha\in I^+}\Phi^+_{\alpha}(x,t) + \sum_{\alpha\in I^-}\Phi^-_{\alpha}(x,t).$$

To justify that the beam solution constructed this way is a valid asymptotic solution for the wave equation (1) with initial conditions (2) and (3), we have to take into account the initial conditions in the beam construction as well. However, this depends on how the initial conditions are decomposed into Gaussian profiles and how the beam propagation is initialized; see [11, 15] for several different approaches in the case of the Schrödinger equation. In particular, we have proposed in [15] an approach based on single-scale fast Gaussian wavepacket transforms to initialize beam propagation for the Schrödinger equation. In the Schrödinger case, the Hamiltonian is not homogeneous, and we cannot restrict the Hamiltonian flow to the cosphere

bundle; consequently, we are only able to use the single-scale Gaussian wavepacket transform to initialize the beam propagation for the Schrödinger equation. For the wave equation, since the Hamiltonian $G^{\pm}(x,p)$ is homogeneous of degree one, the initialization requires new *multiscale* transforms with basis functions satisfying the parabolic scaling principle.

3. Multiscale Gaussian wavepacket transforms. From the discussion of section 2, at time $t=0$ each Gaussian beam of the wave equation takes the form

$$A_0 \exp\left(\imath\omega\left(p_0\cdot(x-x_0)+\frac{1}{2}(x-x_0)^T M_0(x-x_0)\right)\right),$$

where $p_0 = O(1)$ and the Hessian M_0 is purely imaginary and of order $O(1)$. It is clear that this is a modulated Gaussian function that oscillates at a wavelength of order $O(1/\omega)$ and has an effective support of width $O(1/\sqrt{\omega})$ in space. Therefore, the initialization step is equivalent to decomposing the functions $U(x,0)$ and $U_t(x,0)$ into a linear combination of such Gaussian functions.

The transforms to be detailed in the rest of this section follow the architecture of the wave atoms proposed in [7, 8]. However, since the wave atoms do not have a Gaussian-like profile, one needs to adapt the transform appropriately so that the basis functions have the Gaussian profile.

3.1. Continuous transforms. We start by partitioning the Fourier domain $\mathbb{R}^d$ into Cartesian coronae $\{C_\ell\}$ for $\ell \geq 1$ as follows:

$$\begin{aligned} C_1 &= [-4,4]^d, \\ C_\ell &= \{\xi=(\xi_1,\xi_2,\ldots,\xi_d) : \max_{1\leq s\leq d}|\xi_s| \in [4^{\ell-1},4^\ell]\}, \quad \ell\geq 2. \end{aligned}$$

See Figures 1 and 2 for illustration. It is clear that $\xi\in C_\ell$ implies that $|\xi| = O(4^\ell)$. Each corona C_ℓ is further partitioned into boxes

$$B_{\ell,i} = \prod_{s=1}^{d}[2^\ell\cdot i_s, 2^\ell\cdot(i_s+1)],$$

where the integer multiindex $i=(i_1,i_2,\ldots,i_d)$ ranges over all possible choices that satisfy $B_{\ell,i}\subset C_\ell$. All boxes in a fixed C_ℓ have the same length $W^\ell = 2^\ell$ in each dimension, and the center of the box $B_{\ell,i}$ is denoted by $\xi_{\ell,i} = (\xi_{\ell,i,1},\xi_{\ell,i,2},\ldots,\xi_{\ell,i,d})$. To each box $B_{\ell,i}$, we associate a smooth function $g_{\ell,i}(\xi)$, which is compactly supported in a box centered at $\xi_{\ell,i}$ with size $L_\ell = 2W_\ell$ in each dimension (i.e., $\prod_{s=1}^{d}[\xi_{\ell,i,s} - W_\ell, \xi_{\ell,i,s}+W_\ell]$). $g_{\ell,i}(\xi)$ is also required to approximate a Gaussian profile

$$g_{\ell,i}(\xi)\approx e^{-\left(\frac{|\xi-\xi_{\ell,i}|}{\sigma_\ell}\right)^2}$$

with $\sigma_\ell = W_\ell/2$.

One can easily choose $g_{\ell,i}(\xi)$ to satisfy the following three *admissible* conditions:

1. $0\leq g_{\ell,i}(\xi)\leq 1$ for any ξ, i, and ℓ.
2. There exists $C_N>0$ such that for any ξ, $|\{(\ell,i) : g_{\ell,i}(\xi)>0\}|$ is bounded by C_N, where $|\{(\ell,i) : g_{\ell,i}(\xi)>0\}|$ indicates the cardinality of the set $\{(\ell,i) : g_{\ell,i}(\xi)>0\}$.
3. There exists $C_V>0$ such that for any ξ, there exists (ℓ,i) such that $g_{\ell,i}(\xi) > C_V$.

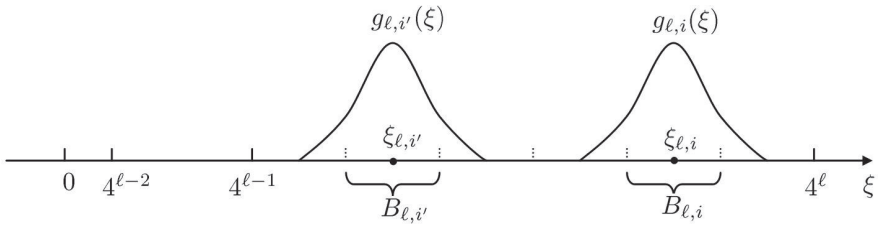

FIG. 1. *Partitioning of the one-dimensional frequency domain.*

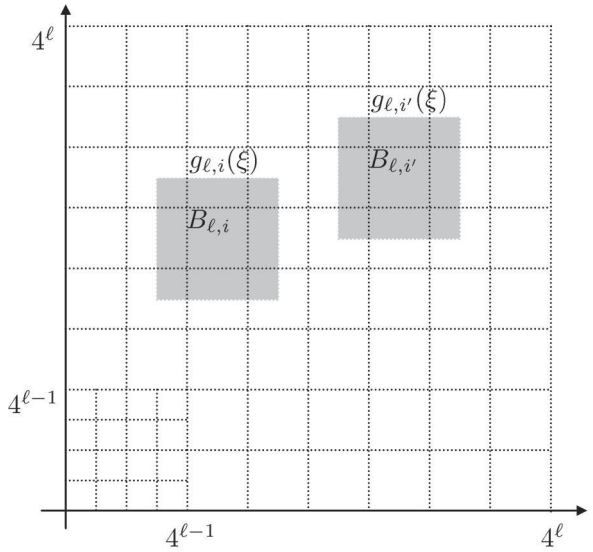

FIG. 2. *Partitioning of the two-dimensional frequency domain.*

Based on $g_{\ell,i}(\xi)$, we introduce for each $B_{\ell,i}$ the conjugate filter $h_{\ell,i}(\xi)$:

$$h_{\ell,i}(\xi) = \frac{g_{\ell,i}(\xi)}{\sum_{\ell,i}(g_{\ell,i}(\xi))^2}.$$

From the smoothness of $g_{\ell,i}(\xi)$ and the three admissible conditions, it is easy to conclude that $h_{\ell,i}(\xi)$ is also a smooth function with the same support. By construction, the products of $g_{\ell,i}(\xi)$ and $h_{\ell,i}(\xi)$ form a partition of unity: $\sum_{\ell,i} g_{\ell,i}(\xi)h_{\ell,i}(\xi) = 1$.

We introduce two sets of functions $\{\varphi_{\ell,i,k}(x)\}$ and $\{\psi_{\ell,i,k}(x)\}$, which are defined in the Fourier domain by

$$\hat{\varphi}_{\ell,i,k}(\xi) = \frac{1}{L_\ell^{d/2}} e^{-2\pi\imath\frac{k\cdot\xi}{L_\ell}} g_{\ell,i}(\xi) \quad \forall k \in \mathbb{Z}^d, \tag{21}$$

$$\hat{\psi}_{\ell,i,k}(\xi) = \frac{1}{L_\ell^{d/2}} e^{-2\pi\imath\frac{k\cdot\xi}{L_\ell}} h_{\ell,i}(\xi) \quad \forall k \in \mathbb{Z}^d. \tag{22}$$

Taking the inverse Fourier transforms gives their definitions in the spatial domain:

$$\varphi_{\ell,i,k}(x) = \frac{1}{L_\ell^{d/2}} \int_{\mathbb{R}^d} e^{2\pi\imath(x-\frac{k}{L_\ell})\cdot\xi} g_{\ell,i}(\xi)d\xi \quad \forall k \in \mathbb{Z}^d, \tag{23}$$

$$\psi_{\ell,i,k}(x) = \frac{1}{L_\ell^{d/2}} \int_{\mathbb{R}^d} e^{2\pi\imath(x-\frac{k}{L_\ell})\cdot\xi} h_{\ell,i}(\xi)d\xi \quad \forall k \in \mathbb{Z}^d. \tag{24}$$

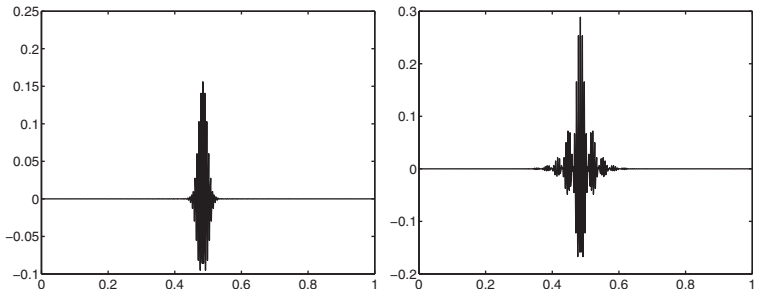

FIG. 3. *Typical profiles of $\varphi_{\ell,i,k}(x)$ and $\psi_{\ell,i,k}(x)$. Left: the real part of a $\varphi_{\ell,i,k}(x)$. Right: the real part of a $\psi_{\ell,i,k}(x)$.*

Figure 3 shows the typical profiles of these two classes of functions. Qualitatively, $\psi_{\ell,i,k}(x)$ is also a wavepacket with slightly larger support in x compared to $\varphi_{\ell,i,k}(x)$.

The definitions of $g_{\ell,i}(\xi)$ and $\varphi_{\ell,i,k}(x)$ imply that

$$\varphi_{\ell,i,k}(x) \approx \left(\sqrt{\frac{\pi}{L_\ell}}\sigma_\ell\right)^d \cdot e^{2\pi\imath(x-\frac{k}{L_\ell})\cdot\xi_{\ell,i}} \cdot e^{-\sigma_\ell^2\pi^2|x-\frac{k}{L_\ell}|^2};$$

i.e., $\varphi_{\ell,i,k}(x)$ is approximately a Gaussian function that is spatially centered at k/L_ℓ, oscillates at frequency $\xi_{\ell,i}$ with $|\xi_{\ell,i}| = O(4^\ell)$, and has an $O(\sigma_\ell) = O(W_\ell) = O(2^\ell)$ effective width in the Fourier domain and an $O(1/\sigma_\ell) = O(2^{-\ell})$ effective width in the spatial domain. The functions $\{\varphi_{\ell,i,k}(x)\}$ fit exactly into the profile of a Gaussian beam with $\omega = O(4^\ell)$.

First, we prove that $\{\varphi_{\ell,i,k}(x)\}$ and $\{\psi_{\ell,i,k}(x)\}$ are two frames of $L^2(\mathbb{R}^d)$.

LEMMA 3.1. *There exists constants C_1 and C_2 such that for any $f \in L^2(\mathbb{R}^d)$*

$$C_1\|f\|_2^2 \le \sum_{\ell,i,k} |\langle\varphi_{\ell,i,k}, f\rangle|^2 \le C_2\|f\|_2^2,$$
$$C_1\|f\|_2^2 \le \sum_{\ell,i,k} |\langle\psi_{\ell,i,k}, f\rangle|^2 \le C_2\|f\|_2^2.$$

Proof. We prove the lemma for $\{\psi_{\ell,i,k}(x)\}$. The proof for $\{\varphi_{\ell,i,k}(x)\}$ is similar:

$$\begin{aligned}\sum_{\ell,i,k} |\langle\psi_{\ell,i,k}, f\rangle|^2 &= \sum_{\ell,i}\sum_k \left|\frac{1}{L_\ell^{d/2}}\int_{\mathbb{R}^d} e^{2\pi\imath\frac{k\cdot\xi}{L_\ell}} h_{\ell,i}(\xi)\hat{f}(\xi)d\xi\right|^2 \\ &= \sum_{\ell,i}\int_{\mathbb{R}^d} |h_{\ell,i}(\xi)\hat{f}(\xi)|^2 d\xi \\ &= \int_{\mathbb{R}^d}\left(\sum_{\ell,i}|h_{\ell,i}(\xi)|^2\right)|\hat{f}(\xi)|^2 d\xi. \qquad (25)\end{aligned}$$

It is then sufficient to show that $\sum_{\ell,i}|h_{\ell,i}(\xi)|^2$ is bounded from below and from above uniformly.

From the conditions that $g_{\ell,i}(\xi)$ satisfies, we have for each ℓ, i

$$h_{\ell,i}(\xi) = \frac{g_{\ell,i}(\xi)}{\sum_{\ell,i}(g_{\ell,i}(\xi))^2} > \frac{g_{\ell,i}(\xi)}{C_N}.$$

Since one of $g_{\ell,i}(\xi)$ is at least C_V, there exists (ℓ, i) such that $h_{\ell,i}(\xi) > \frac{C_V}{C_N}$. This implies that $\sum_{\ell,i} |h_{\ell,i}(\xi)|^2$ is bounded from below by $\frac{C_V^2}{C_N^2}$.

On the other hand,

$$h_{\ell,i}(\xi) = \frac{g_{\ell,i}(\xi)}{\sum_{\ell,i}(g_{\ell,i}(\xi))^2} < \frac{g_{\ell,i}(\xi)}{C_V^2} < \frac{1}{C_V^2}.$$

Since $h_{\ell,i}(\xi) \neq 0$ only for at most C_N indices, $\sum_{\ell,i} |h_{\ell,i}(\xi)|^2$ is bounded from above by $\frac{C_N}{C_V^4}$. ∎

Next, we show that $\{\varphi_{\ell,i,k}(x)\}$ and $\{\psi_{\ell,i,k}(x)\}$ are dual frames.

LEMMA 3.2. *For any* $f \in L^2(\mathbb{R}^d)$,

$$f(x) = \sum_{\ell,i,k} \langle \psi_{\ell,i,k}, f\rangle \varphi_{\ell,i,k}(x).$$

Proof. In the Fourier domain, we have

$$\begin{aligned}
\sum_{\ell,i,k} \langle \psi_{\ell,i,k}, f\rangle \hat{\varphi}_{\ell,i,k}(\xi) &= \sum_{\ell,i,k} \left(\frac{1}{L_\ell^{d/2}} \int_{\mathbb{R}^d} e^{2\pi\imath \frac{k\cdot\eta}{L_\ell}} h_{\ell,i}(\eta) \hat{f}(\eta) d\eta \right) \left(\frac{1}{L_\ell^{d/2}} e^{-2\pi\imath \frac{k\cdot\xi}{L_\ell}} g_{\ell,i}(\xi) \right) \\
&= \sum_{\ell,i} \left(\sum_k \left(\frac{1}{L_\ell^d} \int_{\mathbb{R}^d} e^{2\pi\imath \frac{k\cdot\eta}{L_\ell}} h_{\ell,i}(\eta) \hat{f}(\eta) d\eta \right) e^{-2\pi\imath \frac{k\cdot\xi}{L_\ell}} \right) g_{\ell,i}(\xi) \\
&= \sum_{\ell,i} h_{\ell,i}(\xi) \hat{f}(\xi) g_{\ell,i}(\xi) \\
&= \hat{f}(\xi). \qquad (26)
\end{aligned}$$

Here we use $\sum_{\ell,i} g_{\ell,i}(\xi) h_{\ell,i}(\xi) = 1$ and the fact that the function $h_{\ell,i}(\eta)\hat{f}(\eta)$ is supported in an interval of size L_ℓ in each direction and hence can be considered as a periodic function. ∎

Lemma 3.2 offers a way to decompose any function $f \in L^2(\mathbb{R}^d)$ into a sum of Gaussian-like functions. For a given function f, the *forward multiscale Gaussian wavepacket transform* computes the coefficients $\{c_{\ell,i,k}\}$ defined by

$$c_{\ell,i,k} = \langle \psi_{\ell,i,k}, f\rangle = \langle \hat{\psi}_{\ell,i,k}, \hat{f}\rangle = \frac{1}{L_\ell^{d/2}} \int_{\mathbb{R}^d} e^{2\pi\imath \frac{k\cdot\xi}{L_\ell}} h_{\ell,i}(\xi) \hat{f}(\xi) d\xi, \qquad (27)$$

where $\langle , \rangle$ is the usual $L^2(\mathbb{R}^d)$ inner product and $\hat{f}$ denotes the Fourier transform of f. We remark that by Lemma 3.1 each coefficient $c_{\ell,i,k}$ is at most of order $O(1)$ and many coefficients are negligible. Given a set of coefficients $\{c_{\ell,i,k}\}$, the *inverse multiscale Gaussian wavepacket transform* synthesizes a function $u(x)$ defined by

$$u(x) = \sum_{\ell,i,k} c_{\ell,i,k} \varphi_{\ell,i,k}(x). \qquad (28)$$

3.2. Transforms of discrete signals. Let us now consider the corresponding discrete transforms. For simplicity, assume that we work with the periodized d-dimensional cube $[0,1]^d$. The spatial grid and Fourier grid are defined, respectively, by

$$X = \{(n_1/N, n_2/N, \ldots, n_d/N) : 0 \le n_1, n_2, \ldots, n_d < N, n_1, n_2, \ldots, n_d \in \mathbb{Z}\}, \qquad (29)$$

$$\Omega = \left\{ (k_1, k_2, \ldots, k_d) : -\frac{N}{2} \le k_1, k_2, \ldots, k_d < \frac{N}{2}, k_1, k_2, \ldots, k_d \in \mathbb{Z} \right\}. \qquad (30)$$

Without loss of generality, we assume that N is an integer power of 2. The regions C_ℓ and the intervals $B_{\ell,i}$ are defined in the same way as in the continuous setting. For a given signal f defined on the spatial grid X, we define the discrete Fourier transforms as follows:

$$\hat{f}(\xi) = \frac{1}{N^{d/2}} \sum_{x \in X} e^{-2\pi \imath x \cdot \xi} f(x) \quad \forall \xi \in \Omega, \tag{31}$$

$$f(x) = \frac{1}{N^{d/2}} \sum_{\xi \in \Omega} e^{2\pi \imath x \cdot \xi} \hat{f}(\xi) \quad \forall x \in X. \tag{32}$$

The functions $g_{\ell,i}(\xi)$ and $h_{\ell,i}(\xi)$ are defined in the same way as the continuous case, except that now the addition and subtraction in the ξ variable are understood modulus N. Mimicking the continuous version, we define the discrete versions of $\{\varphi_{\ell,i,k}(x)\}$ and $\{\psi_{\ell,i,k}(x)\}$ in the Fourier domain:

$$\hat{\varphi}^D_{\ell,i,k}(\xi) = \frac{1}{L_\ell^{d/2}} e^{-2\pi \imath \frac{k \cdot \xi}{L_\ell}} g_{\ell,i}(\xi), \quad k \in \{0, 1, \ldots, L_\ell - 1\}^d, \tag{33}$$

$$\hat{\psi}^D_{\ell,i,k}(\xi) = \frac{1}{L_\ell^{d/2}} e^{-2\pi \imath \frac{k \cdot \xi}{L_\ell}} h_{\ell,i}(\xi), \quad k \in \{0, 1, \ldots, L_\ell - 1\}^d. \tag{34}$$

In the spatial domain,

$$\varphi^D_{\ell,i,k}(x) = \frac{1}{(NL_\ell)^{d/2}} \sum_{\xi \in \Omega} e^{2\pi \imath (x - \frac{k}{L_\ell}) \cdot \xi} g_{\ell,i}(\xi), \quad k \in \{0, 1, \ldots, L_\ell - 1\}^d, \tag{35}$$

$$\psi^D_{\ell,i,k}(x) = \frac{1}{(NL_\ell)^{d/2}} \sum_{\xi \in \Omega} e^{2\pi \imath (x - \frac{k}{L_\ell}) \cdot \xi} h_{\ell,i}(\xi), \quad k \in \{0, 1, \ldots, L_\ell - 1\}^d. \tag{36}$$

Similar to the continuous case, $\varphi^D_{\ell,i,k}(x)$ approximates a Gaussian function, but with slightly different scaling due to the definition of the discrete Fourier transform:

$$\varphi^D_{\ell,i,k}(x) \approx \left(\sqrt{\frac{\pi}{NL_\ell}} \sigma_\ell \right)^d \cdot e^{2\pi \imath (x - \frac{k}{L_\ell}) \cdot \xi_{\ell,i}} \cdot e^{-\sigma_\ell^2 \pi^2 |x - \frac{k}{L_\ell}|^2} \quad \forall x \in X.$$

Here the subtraction in the spatial domain is understood modulus the periodic domain $[0, 1]^d$.

Based on the above setup, the discrete version of the forward multiscale Gaussian wavepacket transform for an input signal f on X is given by

$$c^D_{\ell,i,k} = \langle \psi^D_{\ell,i,k}, f \rangle = \langle \hat{\psi}^D_{\ell,i,k}, \hat{f} \rangle = \sum_\xi \frac{1}{L_\ell^{d/2}} e^{2\pi \imath \frac{k \cdot \xi}{L_\ell}} h_{\ell,i}(\xi) \hat{f}(\xi). \tag{37}$$

As $h_{\ell,i}(\xi)\hat{f}(\xi)$ is supported in a subgrid of size L_ℓ in each dimension, the summation (37) is a d-dimensional inverse Fourier transform of size L_ℓ in each dimension by following the wrapping idea proposed in [4, 7, 15]. The algorithm for computing the coefficients $\{c^D_{\ell,i,k}\}$ is given as follows.

ALGORITHM 3.3 (Discrete forward multiscale Gaussian wavepacket transform). *Given a signal f defined at $x \in X$, compute the coefficients $\{c^D_{\ell,i,k}\}$.*

1. Compute $\hat{f}(\xi)$ for $\xi \in \Omega$ using a d-dimensional forward FFT of size N in each dimension.

2. For each level ℓ and each box $B_{\ell,i}$, form $h_{\ell,i}(\xi)\hat{f}(\xi)$ at the support of $h_{\ell,i}(\xi)$, wrap the result modulus L_ℓ to the domain $[-L_\ell/2, L_\ell/2)^d$, and apply a d-dimensional inverse FFT of size L_ℓ in each dimension to the wrapped result to obtain $c^D_{\ell,i,k}$ for all k.

The cost of Algorithm 3.3 is $O(N^d \log N)$.

On the other hand, given a set of coefficients $\{c^D_{\ell,i,k}\}$, the discrete version of the inverse multiscale Gaussian wavepacket transform is defined by

$$f(x) = \sum_{\ell,i,k} c^D_{\ell,i,k} \varphi^D_{\ell,i,k}(x)$$

or equivalently

$$\hat{f}(\xi) = \sum_{\ell,i,k} c^D_{\ell,i,k} \hat{\varphi}^D_{\ell,i,k}(\xi) = \sum_{\ell,i} \left(\sum_k \frac{1}{L^{d/2}} e^{-2\pi \imath \frac{k\cdot\xi}{L_\ell}} c^D_{\ell,i,k} \right) g_{\ell,i}(\xi).$$

The inner summation in k of the last formula is a forward Fourier transform of size L_ℓ.

ALGORITHM 3.4 (Discrete inverse multiscale Gaussian wavepacket transform). *Given coefficients $\{c^D_{\ell,i,k}\}$, reconstruct the function $f(x)$ for $x \in X$.*

1. For each level ℓ and each box $B_{\ell,i}$, apply a d-dimensional forward FFT of size L in each dimension to the coefficients $c^D_{\ell,i,k}$, unwrap the result modulus L_ℓ to the support of $g_{\ell,i}(\xi)$, multiply the unwrapped data with $g_{\ell,i}(\xi)$, and add the product to $\hat{f}(\xi)$.
2. Compute $f(x)$ for $x \in X$ using a d-dimensional inverse FFT of size N in each dimension.

The cost of Algorithm 3.4 is $O(N^d \log N)$.

4. Multiscale Gaussian beams for the wave equation. Equipped with the fast transforms presented above, we discuss how to initialize the Gaussian beam representation of a general initial condition $(U(x,0), U_t(x,0))$. To simplify the discussion, we assume that the domain is the unit torus $[0,1]^d$, the functions $U(x,0)$ and $U_t(x,0)$ are periodic, and $U(x,0)$ and $U_t(x,0)$ are sampled with the uniform Cartesian grid X. As we mentioned earlier, the number of samples N in each direction should be proportional to the highest frequency in $U(x,0)$ and $U_t(x,0)$, as one uses a finite number of samples per unit wavelength.

The accuracy of the Gaussian beam ansatz gets better as the large parameter ω increases. For a Gaussian wave packet $\varphi_{\ell,i,k}(x)$, ω is on the order of $O(|\xi_{\ell,i}|)$, where $\xi_{\ell,i}$ is the center of frequency of the packet. This implies that we should not use the Gaussian beam representation for the low frequency part of the initial data, as it will introduce an error of $O(1)$.

We first introduce the low frequency band pass filter $w^L(\xi)$ and the high frequency band pass filter $w^H(\xi)$ such that $w^L(\xi) + w^H(\xi) = 1$ for all ξ. Decompose the initial data $U(x,0)$ and $U_t(x,0)$ into the low frequency part $U^L(x,0)$ and $U^L_t(x,0)$ defined by

$$\hat{U}^L(\xi,0) = \hat{U}(\xi,0) w^L(\xi), \quad \hat{U_t}^L(\xi,0) = \hat{U}_t(\xi,0) w^L(\xi),$$

and the high frequency part $U^H(x,0)$ and $U^H_t(x,0)$ defined by

$$\hat{U}^H(\xi,0) = \hat{U}(\xi,0) w^H(\xi), \quad \hat{U_t}^H(\xi,0) = \hat{U}_t(\xi,0) w^H(\xi).$$

For the low frequency part $U^L(x,0)$ and $U_t^L(x,0)$, we solve the wave equation using standard finite difference, finite element, or spectral methods. For the high frequency part $U^H(x,0)$ and $U_t^H(x,0)$, we first apply polarization to get polarized components $U^+(x,0)$ and $U^-(x,0)$ which will be detailed in the following.

4.1. Two polarized wave modes. Let us begin by assuming that the initial conditions (2) and (3) belong to the high frequency part. Applying the multiscale wavepacket transform to the initial conditions yields the following decompositions:

$$U|_{t=0} = f_1(x) = \sum_{\ell,i,k} a_{\ell,i,k}\varphi_{\ell,i,k}(x), \tag{38}$$

$$U_t|_{t=0} = f_2(x) = \sum_{\ell,i,k} b_{\ell,i,k}\varphi_{\ell,i,k}(x). \tag{39}$$

Since the wave equation is linear and each $\varphi_{\ell,i,k}(x)$ in the above behaves like a Gaussian wavepacket, this motivates us to consider the wave equation with the following special initial conditions in order to obtain correct initial conditions for the two polarized wave modes:

$$U_{tt} - V^2(x)\Delta U = 0, \quad x \in \mathbb{R}^d, \ \ t > 0, \tag{40}$$

$$U|_{t=0} = a_{\ell,i,k}\varphi_{\ell,i,k}(x), \tag{41}$$

$$U_t|_{t=0} = b_{\ell,i,k}\varphi_{\ell,i,k}(x). \tag{42}$$

Motivated by the approximation

$$\varphi_{\ell,i,k}(x) \approx \left(\sqrt{\frac{\pi}{L_\ell}}\sigma_\ell\right)^d \cdot e^{\imath|\xi_{\ell,i}|(x-\frac{k}{L_\ell})\cdot\frac{2\pi\xi_{\ell,i}}{|\xi_{\ell,i}|}} \cdot e^{-|\xi_{\ell,i}|\left(\frac{\sigma_\ell^2\pi^2}{|\xi_{\ell,i}|}|x-\frac{k}{L_\ell}|^2\right)},$$

where $|\xi_{\ell,i}|$ behaves like a large parameter ω, we can construct one Gaussian beam for each wave mode, respectively, by solving

$$\begin{aligned}
&\dot{x} = G_p, \quad x|_{t=0} = \frac{k}{L_\ell},\\
&\dot{p} = -G_x, \quad p|_{t=0} = 2\pi\frac{\xi_{\ell,i}}{|\xi_{\ell,i}|},\\
&\dot{M} = -(G_{xp})^T M - MG_{px} - MG_{pp}M - G_{xx}, \quad M|_{t=0} = \imath\cdot(2\pi^2\sigma_\ell^2/|\xi_{\ell,i}|)I,\\
&\dot{A} = -\frac{A}{2G}\left(V^2\mathrm{trace}(M) - G_x\cdot G_p - G_p^T M G_p\right), \quad A|_{t=0} = \left(\sqrt{\frac{\pi}{L_\ell}}\sigma_\ell\right)^d,
\end{aligned} \tag{43}$$

where we take $G = G^+$ to obtain the "+" wave mode and $G = G^-$ to obtain the "−" wave mode, respectively. Denote the solutions by $x_{\ell,i,k}^{\pm}(t)$, $p_{\ell,i,k}^{\pm}(t)$, $M_{\ell,i,k}^{\pm}(t)$, and $A_{\ell,i,k}^{\pm}(t)$.

We assume that $\tau_{\ell,i,k}^{\pm}(x,t)$ and $A_{\ell,i,k}^{\pm}(x,t)$ are available as defined by Taylor expansions in (18) and (19), which are rewritten as

$$\tau_{\ell,i,k}^{\pm}(x,t) = p_{\ell,i,k}^{\pm}(t)\cdot(x - x_{\ell,i,k}^{\pm}(t)) + \frac{1}{2}(x - x_{\ell,i,k}^{\pm}(t))^T M_{\ell,i,k}^{\pm}(t)(x - x_{\ell,i,k}^{\pm}(t)), \tag{44}$$

$$A_{\ell,i,k}^{\pm}(x,t) = A_{\ell,i,k}^{\pm}(t). \tag{45}$$

The corresponding Gaussian beams are given by

$$\Phi_{\ell,i,k}^{\pm}(x,t) = A_{\ell,i,k}^{\pm}(x,t)\exp\left(\imath|\xi_{\ell,i}|\cdot\tau_{\ell,i,k}^{\pm}(x,t)\right).$$

Given these two Gaussian beams, we propose a global asymptotic solution to the wave equation (40) in the following form:

$$U(x,t) \approx c^+_{\ell,i,k}\Phi^+_{\ell,i,k}(x,t) + c^-_{\ell,i,k}\Phi^-_{\ell,i,k}(x,t), \tag{46}$$

where the coefficients $c^\pm_{\ell,i,k}$ are to be determined by matching the beam asymptotic solution with the initial conditions given in (41) and (42).

Letting $t = 0$ in solution (46), we obtain

$$(c^+_{\ell,i,k} + c^-_{\ell,i,k})\varphi_{\ell,i,k}(x) \approx U(x,0) = a_{\ell,i,k}\varphi_{\ell,i,k}(x). \tag{47}$$

Differentiating the right-hand side in solution (46) with respect to t, using Lemma 6.4, and letting $t \to 0$, we obtain

$$\begin{aligned} U_t|_{t=0} &\approx (c^-_{\ell,i,k} - c^+_{\ell,i,k})\left(\frac{D_0}{2G^+(x(0),p(0))} + \imath|\xi_{\ell,i}|G^+(x,p(0))\right)\varphi_{\ell,i,k}(x) \\ &\approx (c^-_{\ell,i,k} - c^+_{\ell,i,k})(\imath|\xi_{\ell,i}|G^+(x(0),p(0)))\varphi_{\ell,i,k}(x), \end{aligned} \tag{48}$$

where

$$D_0 = V^2(x(0))\mathrm{trace}(M(0)) - G_p(x(0),p(0))\cdot(G_x(x(0),p(0)) + M(0)\,G_p(x(0),p(0)))\,;$$

since the coefficient D_0 in the above is bounded and can be viewed as a lower-order quantity in comparison to $|\xi_{\ell,i}|$ and $|\xi_{\ell,i}|G^+(x,p(0)) \approx |\xi_{\ell,i}|G^+(x(0),p(0))$ if $|\xi_{\ell,i}|$ is large enough and x is in a small neighborhood of $x(0)$, the second approximation in (48) holds. Thus in the asymptotic sense of $|\xi_{\ell,i}|$ going to infinity, we have

$$(c^-_{\ell,i,k} - c^+_{\ell,i,k})(\imath\,|\xi_{\ell,i}|G^+(x(0),p(0)))\varphi_{\ell,i,k}(x) \approx U_t|_{t=0} = b_{\ell,i,k}\varphi_{\ell,i,k}(x). \tag{49}$$

Solving (47), (49) for $c^+_{\ell,i,k}$ and $c^-_{\ell,i,k}$, we have

$$c^+_{\ell,i,k} = \frac{1}{2}\left(a_{\ell,i,k} - \frac{b_{\ell,i,k}}{\imath\cdot G^+\left(\frac{k}{L_\ell}, 2\pi\xi_{\ell,i}\right)}\right), \tag{50}$$

$$c^-_{\ell,i,k} = \frac{1}{2}\left(a_{\ell,i,k} + \frac{b_{\ell,i,k}}{\imath\cdot G^+\left(\frac{k}{L_\ell}, 2\pi\xi_{\ell,i}\right)}\right). \tag{51}$$

Finally, the initial conditions (38) and (39) are polarized into the two wave modes $U(x,0) \approx U^+(x,0) + U^-(x,0)$ in the frames of the multiscale Gaussian wavepacket transforms with

$$U^\pm(x,0) \approx \sum_{\ell,i,k} c^\pm_{\ell,i,k}\Phi^\pm_{\ell,i,k}(x,0).$$

In turn, these two polarized wave modes allow us to design multiscale Gaussian beam methods to compute asymptotic solutions in the high frequency regime. The global asymptotic solution at time t is given by

$$U(x,t) \approx \sum_{\ell,i,k} c^+_{\ell,i,k}\Phi^+_{\ell,i,k}(x,t) + \sum_{\ell,i,k} c^-_{\ell,i,k}\Phi^-_{\ell,i,k}(x,t).$$

Moreover, for a typical initial condition $(U(x,0), U_t(x,0))$, most of the coefficients $\{c^{\pm}_{\ell,i,k}\}$ have small norms. Therefore, in light of computation efficiency, one needs only to keep the coefficients $\{c^{\pm}_{\ell,i,k}\}$ for which the absolute value is greater than a certain prescribed threshold η. For all other coefficients, we simply set the value to zero. Define the (index) sets of nonnegligible coefficients $S^{\pm}_{\eta}$ by

$$S^{\pm}_{\eta} = \left\{(\ell,i,k) : |c^{\pm}_{\ell,i,k}| > \eta\right\}.$$

Thus, for each $(l,i,k) \in S^{\pm}_{\eta}$ we solve (43) to obtain Gaussian beam ingredients and superpose each beam accordingly,

$$U_{\eta}(x,t) \equiv \sum_{(\ell,i,k)\in S^{+}_{\eta}} c^{+}_{\ell,i,k}\Phi^{+}_{\ell,i,k}(x,t) + \sum_{(\ell,i,k)\in S^{-}_{\eta}} c^{-}_{\ell,i,k}\Phi^{-}_{\ell,i,k}(x,t), \tag{52}$$

which yields a global asymptotic solution to the wave equation (1).

Moreover, we have the following lemma which will be needed later.

LEMMA 4.1. *The coefficients* $\{c^{\pm}_{\ell,i,k}\}$ *defined in* (50) *and* (51) *satisfy the following estimate:*

$$\sum_{(\ell,i,k)\in S^{\pm}_{\eta}} |c^{\pm}_{\ell,i,k}|^2 \le \frac{1}{2}\|f_1\|_2^2 + \frac{1}{2}\frac{C_1}{\omega^2_{\min}}\|f_2\|_2^2,$$

where $\omega_{\min} = \min\{|\xi_{\ell,i}| : (\ell,i,k) \in S^{\pm}_{\eta}\}$ *is assumed to be positive and*

$$C_1 = \frac{1}{\min_{x\in\mathbb{R}^d, p\neq 0} 4\pi^2 \left|G^{+}(x, \frac{p}{|p|})\right|^2} = \frac{1}{\min_{x\in\mathbb{R}^d} 4\pi^2 |V(x)|^2} > 0$$

by the assumption on the velocity V.

Proof. By the definition, we have

$$\begin{aligned}
\sum_{(\ell,i,k)\in S^{\pm}_{\eta}} |c^{\pm}_{\ell,i,k}|^2 &\le \frac{1}{2}\sum_{(\ell,i,k)\in S^{\pm}_{\eta}} |a_{\ell,i,k}|^2 + \frac{1}{2}\sum_{(\ell,i,k)\in S^{\pm}_{\eta}} \frac{1}{4\pi^2|\xi_{\ell,i}|^2}\frac{|b_{\ell,i,k}|^2}{|G^{+}\left(\frac{k}{L_\ell}, \frac{2\pi\xi_{\ell,i}}{2\pi|\xi_{\ell,i}|}\right)|^2} \\
&\le \frac{1}{2}\sum_{(\ell,i,k)\in S^{\pm}_{\eta}} |a_{\ell,i,k}|^2 + \frac{1}{2}\frac{C_1}{\omega^2_{\min}}\sum_{(\ell,i,k)\in S^{\pm}_{\eta}} |b_{\ell,i,k}|^2 \\
&\le \frac{1}{2}\|f_1\|_2^2 + \frac{1}{2}\frac{C_1}{\omega^2_{\min}}\|f_2\|_2^2. \qquad \square
\end{aligned}$$

4.2. Multiscale Gaussian beams: An algorithm. By polarizing the initial conditions of the wave equation in the frames of multiscale Gaussian wavepacket transforms, we have the following algorithm for carrying out multiscale Gaussian beam propagation for the wave equation in the discrete setting. We assume that the domain is $[0,1]^d$ and the functions $U(x,0)$ and $U_t(x,0)$ are discretized on the grid X defined in (29).

ALGORITHM 4.2 (Gaussian beam propagation).

1. Introduce the low frequency band pass filter $w^L(\xi)$ and the high frequency band pass filter $w^H(\xi)$ such that $w^L(\xi) + w^H(\xi) = 1$ for all ξ. Decompose the initial data $U(x,0)$ and $U_t(x,0)$ into the low frequency part $U^L(x,0)$ and $U^L_t(x,0)$ defined by

$$\hat{U}^L(\xi,0) = \hat{U}(\xi,0)w^L(\xi), \quad \hat{U_t}^L(\xi,0) = \hat{U}_t(\xi,0)w^L(\xi),$$

and the high frequency part $U^H(x,0)$ and $U_t^H(x,0)$ defined by

$$\hat{U}^H(\xi,0) = \hat{U}(\xi,0)w^H(\xi), \quad \hat{U_t}^H(\xi,0) = \hat{U}_t(\xi,0)w^H(\xi).$$

2. For the low frequency part $U^L(x,0)$ and $U_t^L(x,0)$, solve the wave problem using standard finite difference or finite element methods.
3. For the high frequency part $U^H(x,0)$ and $U_t^H(x,0)$, we first apply polarization to get the coefficients $\{c^{D,\pm}_{\ell,i,k}\}$.
4. For the positively polarized mode, define the (index) set of nonnegligible coefficients S_η^+ by

$$S_\eta^+ = \left\{(\ell,i,k) : |c^{D,+}_{\ell,i,k}| > \eta\right\}.$$

5. From the above discussion, we know that

$$\varphi_{\ell,i,k}(x) \approx \left(\sqrt{\frac{\pi}{L_\ell N}}\sigma_\ell\right)^d \cdot e^{2\pi\imath(x-\frac{k}{L_\ell})\cdot\xi_{\ell,i}} \cdot e^{-\sigma_\ell^2\pi^2|x-\frac{k}{L_\ell}|^2} \quad \forall x \in X.$$

Therefore, for each $(\ell,i,k) \in S_\eta^+$, set up the following Gaussian beam equations:

$$\dot{x} = G_p^+, \quad x|_{t=0} = \frac{k}{L_\ell},$$
$$\dot{p} = -G_x^+, \quad p|_{t=0} = 2\pi\frac{\xi_{\ell,i}}{|\xi_{\ell,i}|},$$
$$\dot{M} = -(G_{xp}^+)^T M - MG_{px}^+ - MG_{pp}^+ M - G_{xx}^+, \quad M|_{t=0} = \imath\cdot(2\pi^2\sigma_\ell^2/|\xi_{\ell,i}|)I,$$
$$\dot{A} = -\frac{A}{2G^+}\left(V^2\mathrm{trace}(M) - G_x^+\cdot G_p^+ - (G_p^+)^T M G_p^+\right), A|_{t=0} = \left(\sqrt{\frac{\pi}{L_\ell N}}\sigma_\ell\right)^d.$$

Denote the solutions by $x^+_{\ell,i,k}(t)$, $p^+_{\ell,i,k}(t)$, $M^+_{\ell,i,k}(t)$, and $A^+_{\ell,i,k}(t)$, and trace the solution until the final time T. The beam associated with $c^+_{\ell,i,k}$ at time T is then given by

$$\Phi^+_{\ell,i,k}(x,T) = A^+_{\ell,i,k}(T)\exp\left(\imath|\xi_{\ell,i}|\tau^+_{\ell,i,k}(x,T)\right)$$

with $\tau^+_{\ell,i,k}(x,T) = p^+_{\ell,i,k}(T)\cdot(x - x^+_{\ell,i,k}(T)) + \frac{1}{2}(x - x^+_{\ell,i,k}(T))^T M^+_{\ell,i,k}(T)(x - x^+_{\ell,i,k}(T))$. Summing over all beams in S_η^+ gives the time T solution of the positively polarized wave

$$U_\eta^+(x,T) \equiv \sum_{(\ell,i,k)\in S_\eta^+} c^{D,+}_{\ell,i,k}\Phi^+_{\ell,i,k}(x,T).$$

6. Repeat the above for the negatively polarized mode to compute the solution $U_\eta^-(x,T)$ by

$$U_\eta^-(x,T) \equiv \sum_{(\ell,i,k)\in S_\eta^-} c^{D,-}_{\ell,i,k}\Phi^-_{\ell,i,k}(x,T).$$

7. Finally,

$$U_\eta(x,T) = U_\eta^+(x,T) + U_\eta^-(x,T).$$

The first three steps of Algorithm 4.2 involve filtering the initial condition and performing the multiscale Gaussian wavepacket transform. The overall complexity of these steps is $O(N^d \log N)$. Since the Gaussian beam equations themselves do not contain the small parameter, integrating each beam over a finite time period takes $O(1)$ steps. Therefore, the overall cost of tracing the Gaussian beams is proportional to the cardinalities of S_η^+ and S_η^-. As the support of each Gaussian beam is of size $O(N^{1/2})$ in each dimension, each beam at time T covers about $O(N^{d/2})$ points. Therefore, the overall cost of beam summation is of order $O((|S_\eta^+| + |S_\eta^-|) \cdot N^{d/2})$. Summing these estimates together shows that the overall cost of Algorithm 4.2 is $O\left(N^d \log N + (|S_\eta^+| + |S_\eta^-|) \cdot N^{d/2}\right)$. The efficiency of the proposed algorithm evidently depends on the cardinalities of S_η^+ and S_η^-. For certain initial conditions, such as fully random initial data, the sizes of $S_\eta^\pm$ can be fairly large. However, for all usual initial conditions, such as point sources, plane waves, and curvilinear wavefronts, the multiscale Gaussian wavepackets provide theoretically sparse approximations to these initial conditions, following an argument similar to the ones in [3, 7]. Most important, such sparsity is preserved throughout the time evolution. Therefore, for those initial conditions, the sizes of $S_\eta^\pm$ would always be small. Under such situations, the $O\left(N^d \log N + (|S_\eta^+| + |S_\eta^-|) \cdot N^{d/2}\right)$ cost of Algorithm 4.2 is much more efficient compared to the $O(N^{d+1})$ cost of standard finite difference or finite element methods.

5. Long time propagation.

5.1. Exponential growth of Gaussian beam width. To illustrate the point, we consider the following one-dimensional one-way wave equation:

$$U_t + V(x)U_x = 0,$$

which is appended with highly oscillatory data. We apply the Gaussian beam method to construct an asymptotic solution for this equation.

Applying the geometrical-optics ansatz to this equation, we end up with the following Hamilton–Jacobi equation for the phase function τ:

$$\tau_t + V(x)\tau_x = 0.$$

Then we define the Hamiltonian $H(x,p) = V(x)p$. According to the general setup for Gaussian beams in section 2, we have the following Riccati equation for the second-order derivative $M(t) = \tau_{xx}(x(t),t)$:

$$\dot{M}(t) + V_{xx}(x(t))p(t) + 2M(t)V_x(x(t)) = 0, \quad M(0) = \imath\epsilon,$$

where ϵ is a positive constant; $\{(x(t),p(t)) : t > 0\}$ is the bicharacteristic initiated from (x_0,p_0).

To simplify the presentation, we further assume that the velocity $V(x)$ is a linear function, $V(x) = a + bx$, with a and b being constants. Then the Riccati equation can be solved to yield

$$M(t) = \imath\epsilon e^{-2bt}.$$

Thus the beam decays like

$$\exp\left(-\frac{\omega}{2}\mathrm{Im}(M(t))(x - x(t))^2\right) = \exp\left(-\frac{\omega\epsilon}{2}e^{-2bt}(x - x(t))^2\right),$$

and the beam width is roughly proportional to

$$\frac{1}{\sqrt{\omega\epsilon}}e^{bt}.$$

Consequently, if $b > 0$, then the beam width exponentially grows; this implies that the beam loses its localized significance, leading to deteriorating accuracy in the Taylor expansion for the phase function and high cost in beam summation as shown in numerical examples. If $b < 0$, then the beam width exponentially decreases.

We remark that since each polarized wave mode will roughly behave like one-way wave propagation and every velocity function can be approximated locally by its linearization at each point, we can conclude that exponential growth of Gaussian beam width might not be an unusual phenomenon.

5.2. Long time propagation by reinitialization. One natural solution to this issue is to monitor the widths of the Gaussian beams and reinitialize the Gaussian beam representation before any one of the beams becomes too wide. This reinitialization strategy effectively avoids the two difficulties mentioned above. Although the reinitialization idea is rather straightforward, it is rather difficult to combine with existing methods of beam initialization [20, 12, 19] for reasons related to representation and efficiency.

On the other hand, the reinitialization idea fits perfectly with the initialization algorithm proposed in this paper. The main reason is that our algorithm is highly efficient and is applicable to general initial conditions. The actual algorithm processes the positively and negatively polarized components independently. For the positively polarized component $U^+(x,0)$, the algorithm takes the following steps.

ALGORITHM 5.1 (Gaussian beam propagation with reinitialization).

1. Set time $\tilde{t} = 0$ and current solution $\tilde{U}^+(x) = U^+(x,0)$.
2. Decompose $\tilde{U}^+(x)$ to obtain the coefficients $c^{D,+}_{\ell,i,k}$ using Algorithm 3.3. Let S^+_η be the set of indices of significant coefficients, i.e., $S^+_\eta = \{(\ell,i,k) : |c^{D,+}_{\ell,i,k}| > \eta\}$.
3. For each $(\ell,i,k) \in S^+_\eta$, set up the system

$$\begin{aligned}
&\dot{x} = G^+_p, \quad x|_{t=\tilde{t}} = \frac{k}{L_\ell},\\
&\dot{p} = -G^+_x, \quad p|_{t=\tilde{t}} = 2\pi\frac{\xi_{\ell,i}}{|\xi_{\ell,i}|},\\
&\dot{M} = -(G^+_{xp})^T M - MG^+_{px} - MG^+_{pp}M - G^+_{xx}, \quad M|_{t=\tilde{t}} = \imath\cdot(2\pi^2\sigma_\ell^2/|\xi_{\ell,i}|)I,\\
&\dot{A} = -\frac{A}{2G^+}\left(V^2\mathrm{trace}(M) - G^+_x\cdot G^+_p - (G^+_p)^T M G^+_p\right), A|_{t=0} = \left(\sqrt{\frac{\pi}{L_\ell N}}\sigma_\ell\right)^d.
\end{aligned}$$

   Denote the solutions by $x^+_{\ell,i,k}(t)$, $p^+_{\ell,i,k}(t)$, $M^+_{\ell,i,k}(t)$, and $A^+_{\ell,i,k}(t)$. Monitor the values of $M^+_{\ell,i,k}(t)$ while solving the system, and stop if either the time t reaches T or the smallest eigenvalue of any $M_{\ell,i,k}(t)$ drops below a prescribed threshold. Let t^* be the time when we stop.
4. Define $\tilde{t} = t^*$ and

$$\tilde{U}^+(x) = \sum_{(\ell,i,k)\in S^+_\eta} c^{D,+}_{\ell,i,k}A^+_{\ell,i,k}(\tilde{t})\exp\left(\imath|\xi_{\ell,i}|\tau_{\ell,i,k}(x,\tilde{t})\right)$$

with

$$\tau^+_{\ell,i,k}(x,\tilde{t}) = p^+_{\ell,i,k}(\tilde{t})\cdot(x - x^+_{\ell,i,k}(\tilde{t})) + \frac{1}{2}(x - x^+_{\ell,i,k}(\tilde{t}))^T M^+_{\ell,i,k}(\tilde{t})(x - x^+_{\ell,i,k}(\tilde{t})).$$

5. Repeat the steps 2 to 4 until we reach the final time. Once the final time is reached, set $U^+(x,T) = \tilde{U}^+(x)$.

Let us estimate the complexity of Algorithm 5.1. Since the ODEs in step 3 depend only on $V(x)$ and its derivatives, the number of reinitializations is a constant independent of the initial conditions. Combining this with the estimate of the complexity of Algorithm 4.2, it is clear that the cost of Algorithm 5.1 is also $O\left(N^d \log N + (|S^+_\eta| + |S^-_\eta|)\cdot N^{d/2}\right)$.

We remark in passing that beam reinitialization may help to offset the effect of the errors in beam construction, which will make the Gaussian beam valid for a longer time. However, due to the inherent limitation of asymptotics, beam reinitialization will not offset the effect of errors in not exactly solving the PDE, which builds up as time goes on.

6. Multiscale Gaussian beam–based global asymptotic solutions. Following the idea in [16, 20], we prove that the wavepacket-transform-based Gaussian beam solution (52) is an asymptotic solution to the evolution equation (1) in the following sense.

THEOREM 6.1. *For arbitrary $\eta > 0$, let the index sets*

$$S^\pm_\eta = \{(\ell,i,k) : |c_{\ell,i,k}| = |\langle \psi_{\ell,i,k}, U^\pm_0\rangle| > \eta\}$$

be finite. Let $\omega = \max\{|\xi_{\ell,i}| : (\ell,i,k) \in S^\pm_\eta\}$. The solution (52) *is a global asymptotic solution to the evolution equation* (1) *in the following sense: in a finite time interval $[0,T]$, for ω large enough,*

$$\|(\partial_{tt} - V^2(x)\partial_{xx})U^\eta\|_{L^2_{x,t}} \le C(|S^+_\eta| + |S^-_\eta|)\sqrt{\omega}, \tag{53}$$

$$\lim_{\eta\to 0} U^\eta(x,0) \approx f_1(x), \tag{54}$$

$$\lim_{\eta\to 0} U^\eta_t(x,0) \approx f_2(x), \tag{55}$$

where $|S^\pm_\eta|$ denotes the number of elements in the set $S^\pm_\eta$ (the number of beams) and the constant C is independent of ω. Since the initial conditions are satisfied up to a small error as η goes to 0, we use $\approx$ to indicate this approximation in (54) *and* (55).

To prove this theorem, we need some lemmas.

LEMMA 6.2. *The solution $M(t)$ for the Riccati equation* (13) *can be constructed as $M(t) = B(t)C^{-1}(t)$, where $B(t)$ and $C(t)$ satisfy the following variational system:*

$$\begin{aligned} &\dot{B} = -G^T_{xp}B - G_{xx}C, \quad B|_{t=0} = \imath\epsilon I,\\ &\dot{C} = G_{pp}B + G_{px}C, \quad C|_{t=0} = I, \end{aligned} \tag{56}$$

where ϵ is a positive constant, $B = B(t;x_0,p_0)$, and $C = C(t;x_0,p_0)$ are taken to be the variations of $p = p(t;x_0,p_0)$ and $x = x(t;x_0,p_0)$ with respect to the initial point $x_0 = \alpha$. Under the above assumptions, $C(t)$ is nonsingular for any t, and $\mathrm{Im}(B(t)C^{-1}(t))$ is positive definite.

Proof. The proof can be found in [16, 13, 20]. ∎

LEMMA 6.3. *The solution for the transport equation*

$$\frac{dA_{\ell,i,k}}{dt} = -\frac{A_{\ell,i,k}}{2G}\left(V^2\mathrm{trace}(M_{\ell,i,k}) - G_p\cdot(G_x + M_{\ell,i,k}\,G_p)\right),$$
$$A_{\ell,i,k}|_{t=0} = \left(\sqrt{\frac{\pi}{L_\ell}}\sigma_\ell\right)^d$$

is

$$A_{\ell,i,k}(t) = \frac{\left(\sqrt{\frac{\pi}{L_\ell}}\sigma_\ell\right)^d}{\sqrt{\det(C_{i,k}(t))}}\exp\left(\int_0^t \mathrm{trace}(G_{px}(x(s),p(s)))ds\right), \tag{57}$$

where $C_{\ell,i,k}(t) = \frac{\partial x_{\ell,i,k}}{\partial x_0}(t)$ *(being the variation of the current position* $x_{\ell,i,k}(t)$ *with respect to the initial value) satisfies*

$$\frac{dC_{\ell,i,k}}{dt} = G_{pp}M_{\ell,i,k}C_{\ell,i,k} + G_{px}C_{\ell,i,k}, \quad C_{\ell,i,k}|_{t=0} = I. \tag{58}$$

Proof. The transport equation can be written as

$$\frac{d\log A^2_{\ell,i,k}}{dt} = -\frac{1}{G}\left(V^2\mathrm{trace}(M_{\ell,i,k}) - G_p\cdot(G_x + M_{\ell,i,k}\,G_p)\right).$$

Using the specific form of G, the above equation can be further reduced to be

$$\frac{d\log A^2_{\ell,i,k}}{dt} = -\mathrm{trace}(G_{pp}M_{\ell,i,k} - G_{px}). \tag{59}$$

By Lemma 6.2, $C_{\ell,i,k}(t)$ is nonsingular. Let $C_{\ell,i,k}(t) = S(t)\Lambda(t)S^{-1}(t)$ be the Schur decomposition. Setting $q(t) = \det C_{\ell,i,k}(t) = \Pi_{j=1}^d\lambda_j$, where $\Lambda = \mathrm{diag}(\lambda_1,\ldots,\lambda_d)$, and noticing that

$$\frac{dC_{\ell,i,k}}{dt} = S_t\Lambda S^{-1} + S\Lambda_t S^{-1} + S\Lambda(S^{-1})_t,$$

we have

$$\begin{aligned}
\frac{dq}{dt} &= q\ \mathrm{trace}(\Lambda^{-1}\Lambda_t)\\
&= q\ \mathrm{trace}(S\Lambda^{-1}S^{-1}S\Lambda_t S^{-1} + (S\Lambda^{-1})S^{-1}S_t(S\Lambda^{-1})^{-1} + (S^{-1})_t S)\\
&= q\ \mathrm{trace}\left(C^{-1}_{\ell,i,k}\frac{dC_{\ell,i,k}}{dt}\right)\\
&= q\ \mathrm{trace}(C^{-1}_{\ell,i,k}(G_{pp}M_{\ell,i,k} + G_{px})C_{\ell,i,k})\\
&= q\ \mathrm{trace}(G_{pp}M_{\ell,i,k} + G_{px}). \qquad (60)
\end{aligned}$$

Thus, by combining (59) and (60), we have

$$A_{\ell,i,k}(t) = \frac{A_{\ell,i,k}(0)\sqrt{q(0)}}{\sqrt{q(t)}} = \frac{\left(\sqrt{\frac{\pi}{L_\ell}}\sigma_\ell\right)^d}{\sqrt{\det(C_{\ell,i,k}(t))}}\exp\left(\int_0^t \mathrm{trace}(G_{px}(x(s),p(s)))ds\right).$$

This implies that $A_{\ell,i,k}$ is of the order $O(|\xi_{\ell,i}|^{\frac{d}{4}})$ by the definitions of σ_ℓ and L_ℓ. ∎

In (44) and (45), we define the global approximations to the phase and amplitude functions by using Taylor expansions centered on the x-projection of the bicharacteristic

$$\{(x_{\ell,i,k}(t), p_{\ell,i,k}(t)) : t \geq 0\}.$$

We have the following estimates.

LEMMA 6.4. *Assume that the Hamiltonian G is C^2 differentiable. Then the functions* (44) *and* (45) *satisfy the eikonal and transport equations in the following approximate sense, respectively:*

$$\frac{\partial \tau_{\ell,i,k}}{\partial t}(x,t) + G\left(x, \frac{\partial \tau_{\ell,i,k}}{\partial x}(x,t)\right) = O(|x - x_{\ell,i,k}(t)|^3), \tag{61}$$

$$\begin{aligned} &2A_{\ell,i,k,t}\tau_{\ell,i,k,t} - 2V^2 \nabla_x A_{\ell,i,k} \cdot \nabla_x \tau_{\ell,i,k} + A_{\ell,i,k}(\tau_{\ell,i,k,tt} - V^2 \mathrm{trace}(\tau_{\ell,i,k,xx})) \\ &\qquad = O(|\xi_{\ell,i}|^{\frac{d}{4}} |x - x_{i,k}(t)|). \end{aligned} \tag{62}$$

Proof. We first prove (61). For the sake of clarity we suppress the index (ℓ, i, k) in the following. By the construction of $\tau_{i,k}(x,t) = \tau(x,t)$, we have

$$\begin{aligned} \frac{\partial \tau}{\partial t}(x,t) &= \frac{d\tau}{dt}(t) + \frac{dp}{dt} \cdot (x - x(t)) + p(t) \cdot \left(-\frac{dx}{dt}\right) \\ &\quad + \frac{1}{2}(x - x(t))^T \frac{dM}{dt}(x - x(t)) - (x - x(t))^T M(t) \frac{dx}{dt}, \\ \frac{\partial \tau}{\partial x}(x,t) &= p(t) + M(t)(x - x(t)). \end{aligned}$$

Thus, using the Taylor expansion for G around $x_{\ell,i,k}(t) = x(t)$ up to the third-order term we have

$$\begin{aligned} \tau_t(x,t) + G(x, \tau_x(x,t)) &= \tau_t(x,t) + G\left(x - x(t) + x(t), p(t)\right) + M(t)(x - x(t)) \\ &= O(|x - x(t)|^3). \end{aligned} \tag{63}$$

Next we prove (62). By the analytical formula (57) and using the notation as in Lemma 6.3 we have

$$\begin{aligned} A_t(x,t) &= A(0) \frac{d}{dt}\left(q(t)^{-\frac{1}{2}} \exp\left(\int_0^t \mathrm{trace}(G_{px}(x(s), p(s)))ds\right)\right) \\ &= -\frac{A}{2}\mathrm{trace}(G_{pp}M - G_{px}), \end{aligned} \tag{64}$$

$$A_x(x,t) = O(|\xi_{\ell,i}|^{\frac{d}{4}} |x - x(t)|). \tag{65}$$

Thus, near the ray trajectory $x_{\ell,i,k}(t)$, the estimate (62) holds. ∎

We also need the following lemma which is proved in [16, 20].

LEMMA 6.5. *Assume that $c(x,t)$ vanishes to order $S-1$ on $\gamma = \{(x(t),t) : 0 \leq t \leq T\}$ which is the x-projection of a bicharacteristic $\{(x(t),p(t)) : 0 \leq t \leq T\}$, $\mathrm{supp}(c) \cap \{(x,t) : x \in \mathbb{R}^d, 0 \leq t \leq T\}$ is compact, and $\mathrm{Im}(\varphi(x,t)) \geq a|x - x(t)|^2$ on $\mathrm{supp}(c) \cap \{(x,t) : x \in \mathbb{R}^d, 0 \leq t \leq T\}$, where a is a positive constant. Then*

$$\int_0^T \int_{\mathbb{R}^d} \left| c(x,t) e^{\imath \omega \varphi(x,t)} \right|^2 dx dt \leq C \omega^{-(S + \frac{d}{2})}, \tag{66}$$

where C is a constant independent of ω.

Now we are ready to prove Theorem 6.1.

Proof. As constructed, it is easy to check that at $t = 0$ we recover the initial conditions up to small errors as η goes to 0.

Next we evaluate the following:

$$\left(\partial_{tt} - V^2(x)\partial_{xx}\right) U^\eta(x,t) = \sum_{(\ell,i,k)\in S_\eta^+} (\partial_{tt} - V^2(x)\partial_{xx}) c^+_{\ell,i,k}\Phi^+_{\ell,i,k}(x,t) + \sum_{(\ell,i,k)\in S_\eta^-} (\partial_{tt} - V^2(x)\partial_{xx}) c^-_{\ell,i,k}\Phi^-_{\ell,i,k}(x,t).$$

Then we may just check the term associated with S_η^+; the same can be carried out for the term associated with S_η^-. Thus we have

$$\begin{aligned}
&\sum_{(\ell,i,k)\in S_\eta^+} (\partial_{tt} - V^2(x)\partial_{xx}) c^+_{\ell,i,k}\Phi^+_{\ell,i,k}(x,t)\\
&= \sum_{(\ell,i,k)\in S_\eta^+} (\partial_{tt} - V^2(x)\partial_{xx}) c^+_{\ell,i,k}A^+_{\ell,i,k}(x,t)\exp\left(\imath|\xi_{\ell,i}|\tau^+_{\ell,i,k}(x,t)\right)\\
&\approx \sum_{(\ell,i,k)\in S_\eta^+} -|\xi_{\ell,i}|^2\left((\tau^+_{\ell,i,k,t})^2 - V^2(\tau^+_{\ell,i,k,x})^2\right) c^+_{\ell,i,k}A^+_{\ell,i,k}(x,t)\exp\left(\imath|\xi_{\ell,i}|\tau^+_{\ell,i,k}(x,t)\right)\\
&+ \sum_{(\ell,i,k)\in S_\eta^+} \imath|\xi_{\ell,i}|c^+_{\ell,i,k}\times\exp\left(\imath|\xi_{\ell,i}|\tau^+_{\ell,i,k}(x,t)\right)\\
&\quad\times\left(2\tau^+_{\ell,i,k,t}A^+_{\ell,i,k,t} - 2V^2(x)\tau^+_{\ell,i,k,x}\cdot A^+_{\ell,i,k,x} + A^+_{\ell,i,k}(\tau^+_{\ell,i,k,tt} - V^2\mathrm{trace}(\tau^+_{\ell,i,k,xx}))\right)\\
&\equiv q_1(x,t) + q_2(x,t).
\end{aligned}$$

We estimate q_1 first. By using the Cauchy–Schwartz inequality we have

$$|q_1(x,t)|^2 \le \left(\sum_{(\ell,i,k)\in S_\eta^+} |c^+_{\ell,i,k}|^2\right) \times\left(\sum_{(\ell,i,k)\in S_\eta^+} \left|-|\xi_{\ell,i}|^2\left((\tau^+_{\ell,i,k,t})^2 - V^2(\tau^+_{\ell,i,k,x})^2\right)A^+_{\ell,i,k}e^{\left(\imath|\xi_{\ell,i}|\tau^+_{\ell,i,k}(x,t)\right)}\right|^2\right).$$

We observe that $\sum_{(\ell,i,k)\in S_\eta^+}|c^+_{\ell,i,k}|^2$ is uniformly bounded by Lemma 4.1 and $A^+_{\ell,i,k}$ is of order $O(|\xi_{\ell,i}|^{\frac{d}{4}})$. By using the results in Lemmas 6.4 and 6.5 with $S = 3$, we have

$$\|q_1\|_{L^2_{x,t}} \le C|S^+_\eta|\sqrt{\omega},$$

where C is a constant independent of ω as defined in the theorem.

By using the Cauchy–Schwartz inequality, Lemmas 4.1, 6.4, and 6.5 with $S = 1$, we can estimate q_2 similarly to obtain

$$\|q_2\|_{L^2_{x,t}} \le C|S^+_\eta|\sqrt{\omega},$$

where C is a constant independent of ω.

The exact same can be carried out for the term associated with S_η^-. Finally, applying the triangle inequality yields the desired estimate. □

We remark that because we decompose the initial conditions into nearly Gaussians, it is not easy to quantify the small errors incurred at the beginning of such beam decomposition. Currently, we are analyzing this issue in an ongoing work.

7. Numerical results. In this section we show one-, two-, and three-dimensional numerical examples to validate our algorithms. Because an exact wave function for the wave equation is not available in general, we solve the wave equation first with an accurate pseudospectral method to compute the "exact" solutions.

In the following examples, we have taken different cutoff parameters η for different examples. In practice, one may take larger cutoff parameters and still obtain reasonable results.

7.1. One-dimensional cases.

7.1.1. Example 1: A slow velocity with small slope. The velocity is given by $V(x,y) = 0.25 + 0.1\sin(2\pi x)$. To illustrate the effectiveness of the multiscale Gaussian method, we consider three sets of initial conditions. In the first three cases, we use $N = 4096$ uniform distributed points to discretize $[0,1]$ and compute the solution up to the final time $T = 8.0$ without reinitialization. We take $\eta = 10^{-3}$ as the threshold value for S_η^+ and S_η^-.

Case 1.

$$U|_{t=0} = 2\sin(128\pi x), \qquad U_t|_{t=0} = 0.$$

Figure 4 summarizes the results for this case. Figure 4(a) shows the velocity, which is smooth and slowly varying. Figure 4(b) plots the initial condition $U|_{t=0}$, which is highly oscillatory. Figure 4(c) plots the rays that have been traced in beam propagation, and Figure 4(d) shows the multiscale Gaussian beam solution. The exact solution and the beam solution together are plotted together shown in Figure 4(e); they are almost identical to each other. To compare these two solutions more closely, Figure 4(f) displays a windowed portion of the two solutions, from which we conclude that these two solutions match with each other very well.

Case 2.

$$U|_{t=0} = 0, \qquad U_t|_{t=0} = 6400\cos(128\pi x).$$

Figure 5 summarizes the results of this case. Figure 5(a) shows the velocity. Figure 5(b) plots the oscillatory initial condition $U_t|_{t=0}$. Figure 5(c) shows the rays that have been traced in beam propagation, and Figure 5(d) plots the multiscale Gaussian beam solution. The exact solution and the beam solution are plotted together in Figure 5(e); they are almost identical to each other. To compare these two solutions more closely, Figure 5(f) shows a windowed portion of the two solutions, from which we see that these two solutions match very well.

Case 3.

$$U|_{t=0} = 2\sin(128\pi x), \qquad U_t|_{t=0} = 64\cos(64\pi x).$$

We notice that this set of initial conditions does not have the usual geometrical-optics form, which implies that the traditional Gaussian beam is not applicable to this example.

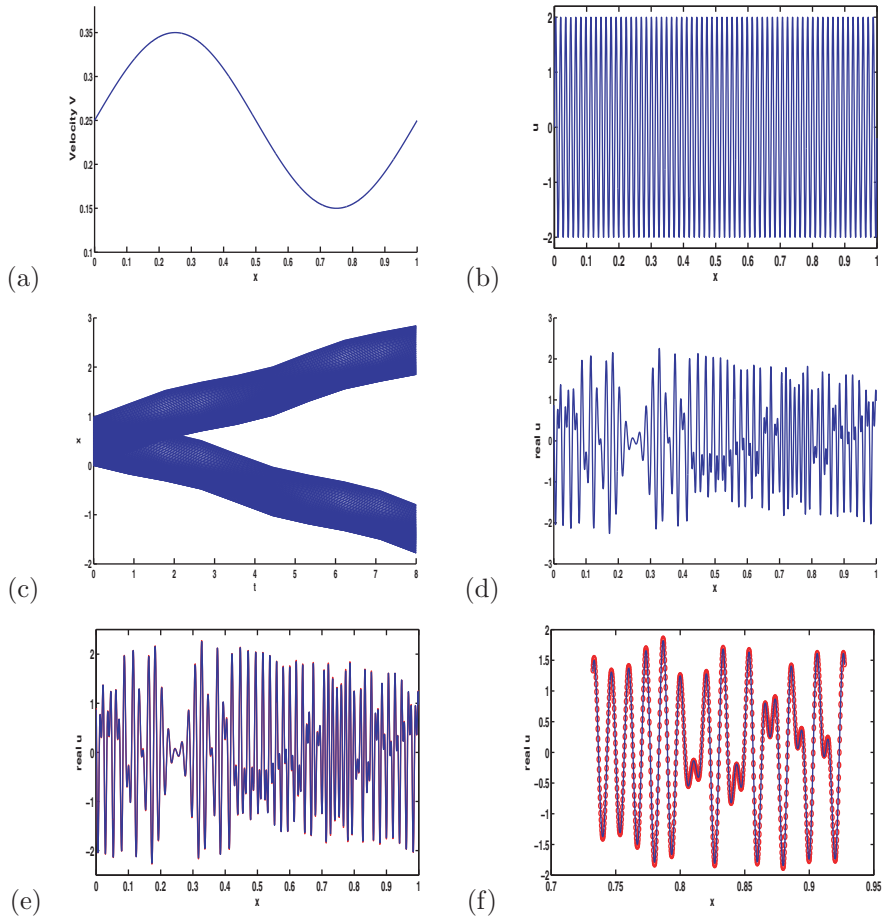

FIG. 4. *Example* 1, *Case* 1*: The final time* $T = 8.0$*:* (a) *the velocity;* (b) *the initial condition* $U|_{t=0}$*;* (c) *the rays;* (d) *the beam solution;* (e) *the exact solution overlays the beam solution;* (f) *windowed comparison: the exact solution ("o") and the beam solution ("-").*

Figure 6 summarizes the results of this case. Figure 6(a) and Figure 6(b) show, respectively, the initial condition $U|_{t=0}$ and $U_t|_{t=0}$, both of which are highly oscillatory. Figure 6(c) plots the rays that have been traced in beam propagation, and Figure 6(d) shows the multiscale Gaussian beam solution. The exact solution and the beam solution are displayed together in Figure 6(e); they are almost identical to each other. To compare these two solutions more closely, Figure 6(f) plots a windowed portion of the two solutions, from which we see that these two solutions match with each other very well.

To further demonstrate the performance of the new algorithm, we compute wave fields for the following two cases.

Case 4.

$$U|_{t=0} = 2\sin(512\pi x), \qquad U_t|_{t=0} = 256\cos(256\pi x).$$

In this case, we take $\eta = 0.1$, $N = 8192$, and $T = 8.0$. If we accept the solution by the pseudospectral method as the exact solution, the relative L^2-error in the Gaussian beam solution is 4.6% and the relative maximum error is 5.8%. Here $|S_\eta^+| = 640$ and $|S_\eta^-| = 640$.

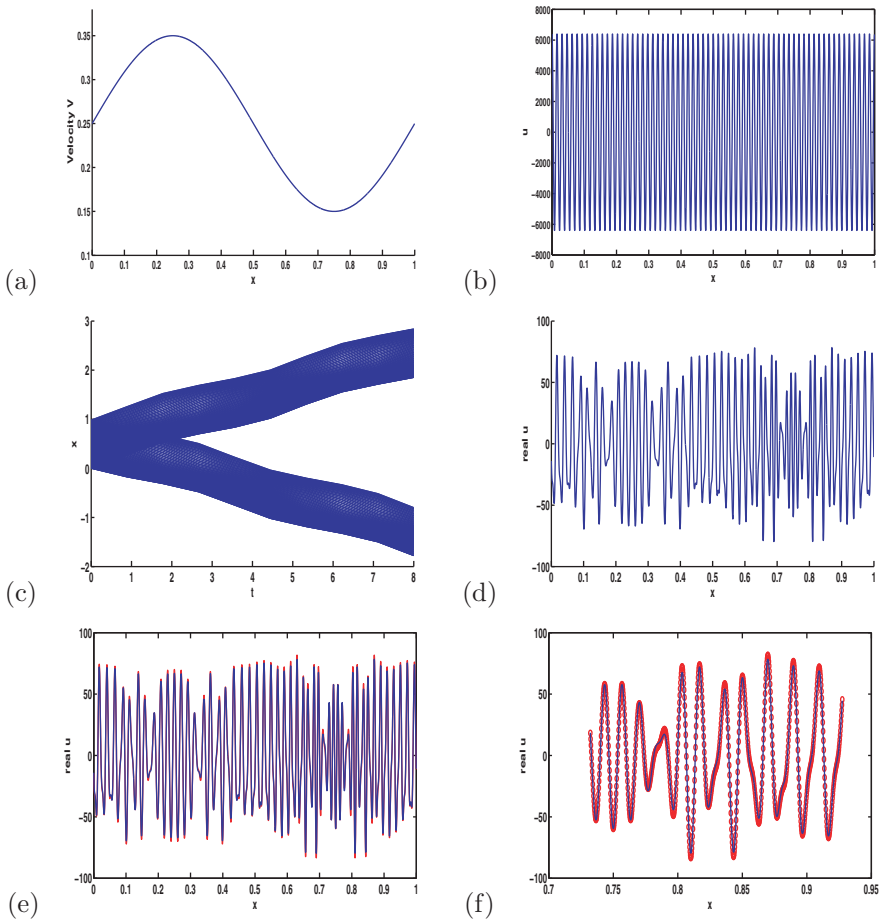

FIG. 5. *Example* 1, *Case* 2*: The final time* $T = 8.0$*;* (a) *the velocity;* (b) *the initial condition* $U_t|_{t=0}$*;* (c) *the rays;* (d) *the beam solution;* (e) *the exact solution overlays the beam solution;* (f) *windowed comparison: the exact solution ("o") and the beam solution ("-").*

Case 5.

$$U|_{t=0} = 2\sin(512\pi x), \qquad U_t|_{t=0} = 0.$$

In this case, we take $\eta = 0.1$, $N = 8192$, and $T = 8.0$. If we accept the solution by the pseudospectral method as the exact solution, the relative L^2-error in the Gaussian beam solution is 5.28% and the relative maximum error is 6.17%. Here $|S_\eta^+| = 384$ and $|S_\eta^-| = 384$.

7.1.2. Example 2: A fast velocity with large slope. The velocity is given by $V(x, y) = 1.5 + \sin(2\pi x)$. We consider the following initial conditions:

$$U|_{t=0} = 2\sin(256\pi x), \qquad U_t|_{t=0} = 128\cos(128\pi x).$$

We use $N = 8192$ uniform distributed points to discretize $[0, 1]$ and compute the solution up to the final time $T = 2.0$ without reinitialization. η is again equal to 10^{-3}. Figure 7 summarizes the numerical results of this example. Figure 7(a) and Figure 7(b) display, respectively, the initial conditions $U|_{t=0}$ and $U_t|_{t=0}$, both of

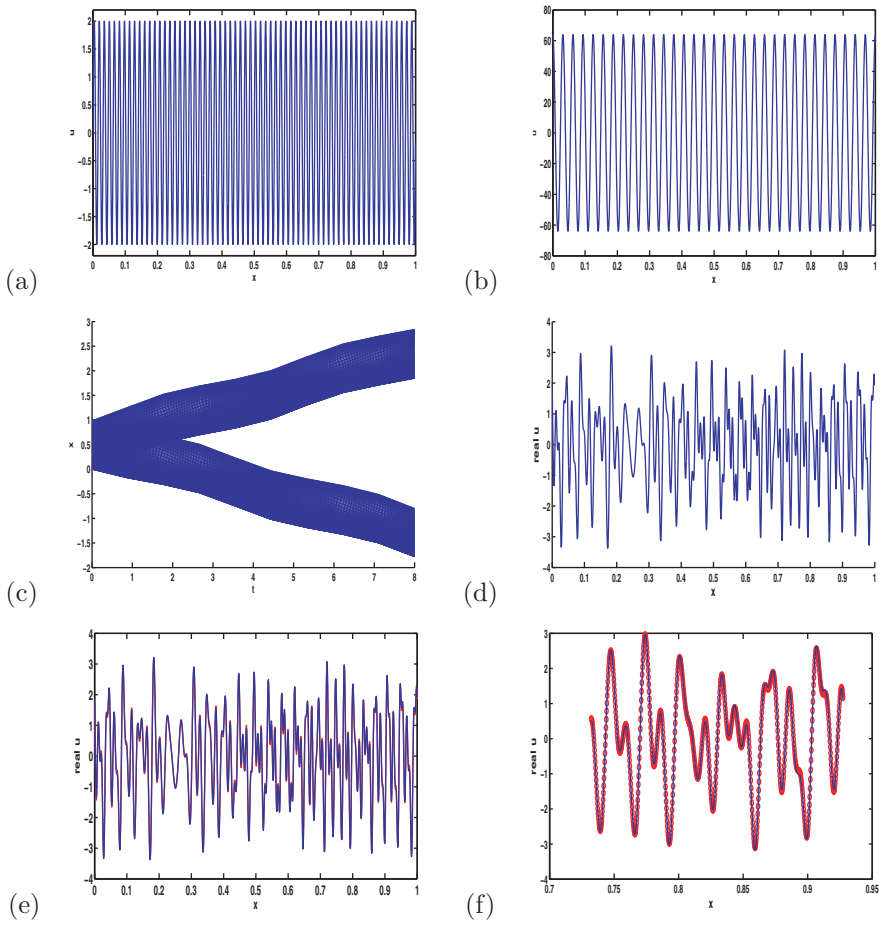

FIG. 6. *Example* 1, *Case* 3*: The final time* $T = 8.0$*;* (a) *the initial condition* $U|_{t=0}$*;* (b) *the initial condition* $U_t|_{t=0}$*;* (c) *the rays;* (d) *the beam solution;* (e) *the exact solution overlays the beam solution;* (f) *windowed comparison: the exact solution ("o") and the beam solution ("-").*

which are highly oscillatory. Figure 7(c) shows the rays that have been traced in beam propagation, and Figure 7(d) plots the multiscale Gaussian beam solution. The exact solution and the beam solution are plotted together as shown in Figure 7(e); they are almost identical to each other. To compare these two solutions more closely, Figure 7(f) shows a windowed portion of the two solutions, from which it is clear that these two solutions match with each other very well.

We remark that in Figure 7(e) the waves are compressed in the neighborhood of $x = 0.75$ and the waves are rarefied in the neighborhood of $x = 0.25$, so that resolving such highly oscillatory waves become extremely difficult; this is why a very refined mesh is used to visualize the solution at $T = 2.0$.

7.2. Two-dimensional cases.

7.2.1. Example 3: A two-dimensional velocity. The velocity is given by $V(x, y) = 1.0 + 0.1 \sin(2\pi x) \cos(2\pi y)$, and the initial conditions are given by

$$U|_{t=0} = \sin(2\pi\beta((x - 0.5) + (y - 0.5))) \exp\left(-600((x - 0.5)^2 + (y - 0.5)^2)\right),$$
$$U_t|_{t=0} = 0. \tag{67}$$

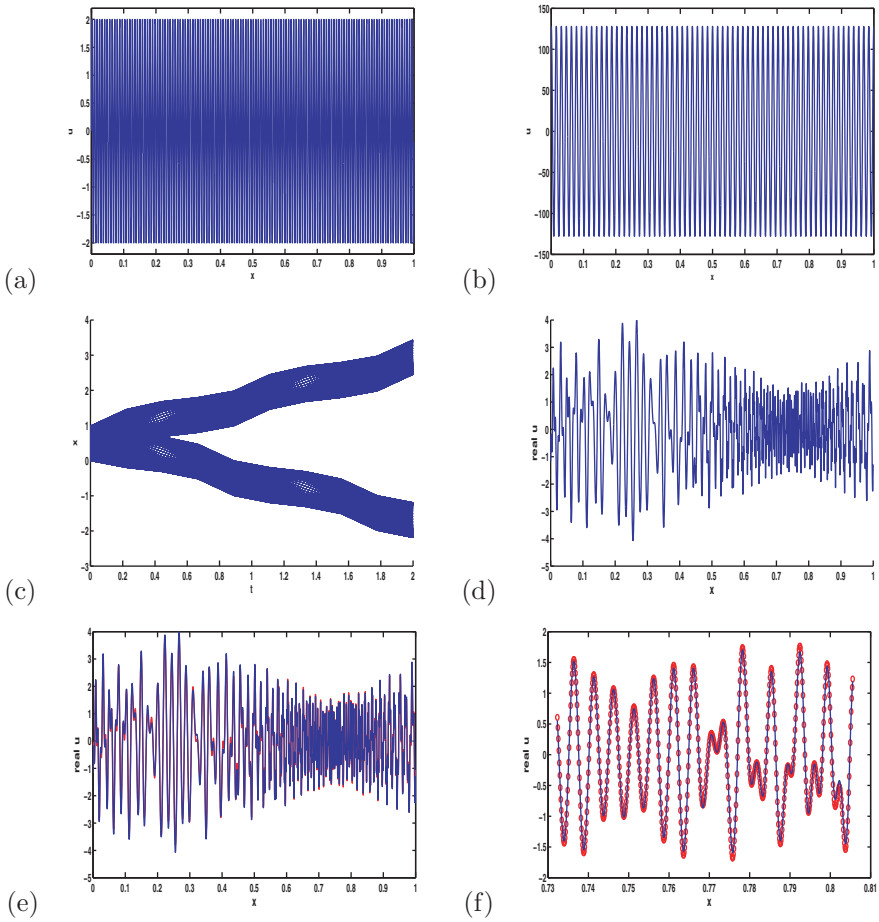

FIG. 7. *Example* 2. *The final time* $T = 2.0$*:* (a) *the initial condition* $U|_{t=0}$*;* (b) *the initial condition* $U_t|_{t=0}$*;* (c) *the rays;* (d) *the beam solution;* (e) *the exact solution overlays the beam solution;* (f) *windowed comparison: the exact solution ("o") and the beam solution ("-").*

We take $\beta = 256$ and use $N \times N = 1024^2$ uniform distributed points to discretize $[0,1]\times[0,1]$. We compute the solution up to the final time $T = 2.0$, with reinitialization set at $t = 0.5, 1.0, 1.5$. η is equal to 10^{-3}. Notice that in this case the small wavelength in the initial data is roughly equal to $\frac{1}{\beta} = \frac{1}{256}$, and we use roughly four points per wavelength to resolve the oscillation, which is optimal to some extent; consequently, this example serves the purpose to test the resolution of multiscale Gaussian beams proposed here.

Figure 8 shows the results. Figures 8(a) and 8(b) show the exact solution and the multiscale Gaussian beam solution. To compare the results, we plot the exact solution and the beam solution together by taking different slices. Figures 8(c), 8(e), 8(g), and 8(j) show slices at $x = 0.6$, $x = 0.9$, $y = 0.1$, and $y = 0.4$, respectively; Figures 8(d), 8(f), 8(i), and 8(k) show corresponding windowed portions of the slices at $x = 0.6$, $x = 0.9$, $y = 0.1$, and $y = 0.4$, respectively, with the beam solution ("-") and the exact solution ("o"). From these comparisons, we can see clearly that the beam solution ("-") matches with the exact solution ("o") very well.

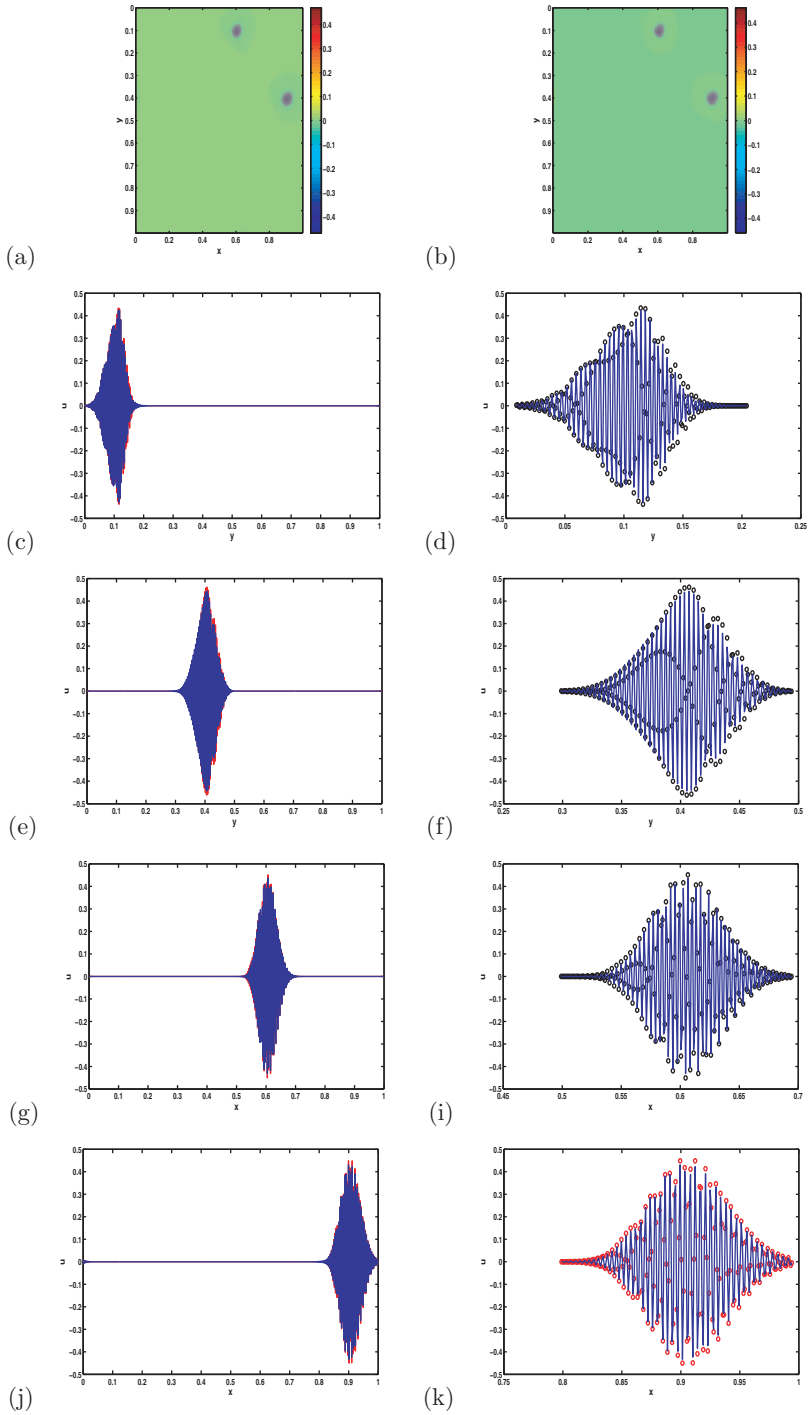

FIG. 8. *Example* 3. *(A* 2*-d velocity model.) The final time* $T = 2.0$ *and the reinitialization is set at* $t = 0.5, 1.0, 1.5$*:* (a) *exact solution;* (b) *the beam solution;* (c) *the slice at* $x = 0.6$*;* (d) *a windowed slice at* $x = 0.6$*;* (e) *the slice at* $x = 0.9$*;* (f) *a windowed slice at* $x = 0.9$*;* (g) *the slice at* $y = 0.1$*;* (i) *a windowed slice at* $y = 0.1$*;* (j) *the slice at* $y = 0.4$*;* (k) *a windowed slice at* $y = 0.4$*: the beam solution ("-") and the exact solution ("o").*

7.2.2. Example 4: A two-dimensional velocity. The velocity is given by $V(x,y) = 1.0 + 0.5\sin(2\pi x)\cos(2\pi y)$, and the initial conditions are given by

$$\begin{aligned} U|_{t=0} &= 2.0\sin(256\pi x)\cos(256\pi y), \\ U_t|_{t=0} &= 200.0\cos(256\pi x)\sin(256\pi y). \end{aligned} \tag{68}$$

We use $N \times N = 1024 \times 1024$ uniform distributed points to discretize $[0,1] \times [0,1]$. We set the cutoff threshold to be $\eta = 0.005$. To demonstrate the advantages of beam reinitialization, we compute the solution up to the final time $T = 0.60$ without reinitialization and with reinitialization, respectively.

Figure 9 shows the results without beam reinitialization. Figures 9(a) and 9(b) plot the exact solution and the multiscale Gaussian beam solution. To see the comparison of the results, we plot the exact solution and the beam solution together by taking different slices. Figures 9(c), 9(e), 9(g), and 9(i) show slices at $x = 0.50$, $x = 0.75$, $y = 0.20$, and $y = 0.50$, respectively; Figures 9(d), 9(f), 9(h), and 9(j) show corresponding windowed portions of the slices at $x = 0.50$, $x = 0.75$, $y = 0.20$, and $y = 0.50$, respectively, with the beam solution ("-") and the exact solution ("o"). From these comparisons, we can see that the beam solution ("-") differs from the exact solution ("o") significantly.

Figure 10 shows the results with beam reinitialization set at $t = 0.20$ and $t = 0.40$, respectively. Figures 10(a) and 10(b) show the exact solution and the multiscale Gaussian beam solution. To see the comparison of the results, we plot the exact solution and the beam solution together by taking different slices. Figures 10(c), 10(e), 10(g), and 10(i) show slices at $x = 0.50$, $x = 0.75$, $y = 0.20$, and $y = 0.50$, respectively; Figures 10(d), 10(f), 10(h), and 10(j) show corresponding windowed portions of the slices at $x = 0.50$, $x = 0.75$, $y = 0.20$, and $y = 0.50$, respectively, with the beam solution ("-") and the exact solution ("o"). Based on these comparisons, we can see that the beam solution ("-") matches with the exact solution ("o") very well. In particular, comparing Figure 9(f) with Figure 10(f) and also Figure 9(k) with Figure 10(k), we conclude that the Gaussian beam propagation with reinitialization yields much more accurate solutions than that without beam reinitialization. Moreover, without beam reinitialization the relative L^2-error is 69% and the relative maximum error is 60%; with beam reinitialization the relative L^2-error is 7.3% and the relative maximum error is 5.9%.

7.3. Three-dimensional cases.

7.3.1. Example 5: A three-dimensional slow velocity. The velocity is given by $V(x,y,z) = 0.50 + 0.1\sin(2\pi x)\cos(2\pi y)\sin(2\pi z)$, and the initial conditions are given by

$$\begin{aligned} U|_{t=0} &= \sin(2\pi\beta(x+y+z-1.5)) \\ &\quad \exp\left(-600((x-0.5)^2+(y-0.5)^2+(z-0.5)^2)\right), \\ U_t|_{t=0} &= 0. \end{aligned} \tag{69}$$

In Figure 11, we take $\beta = 32$, $\eta = 10^{-3}$, and use $N^3 = 128^3$ uniform distributed points to discretize the unit cube $[0,1]^3$. We compute the beam solution up to the final time $T = 0.5$ without reinitialization. Figure 11(a) shows an x-section of the three-dimensional wavefield at $x = 0.4219$, and Figure 11(b) shows slices at $x = 0.4219$ and $y = 0.3359$, in which we compare the exact solution ("o") with the beam solution

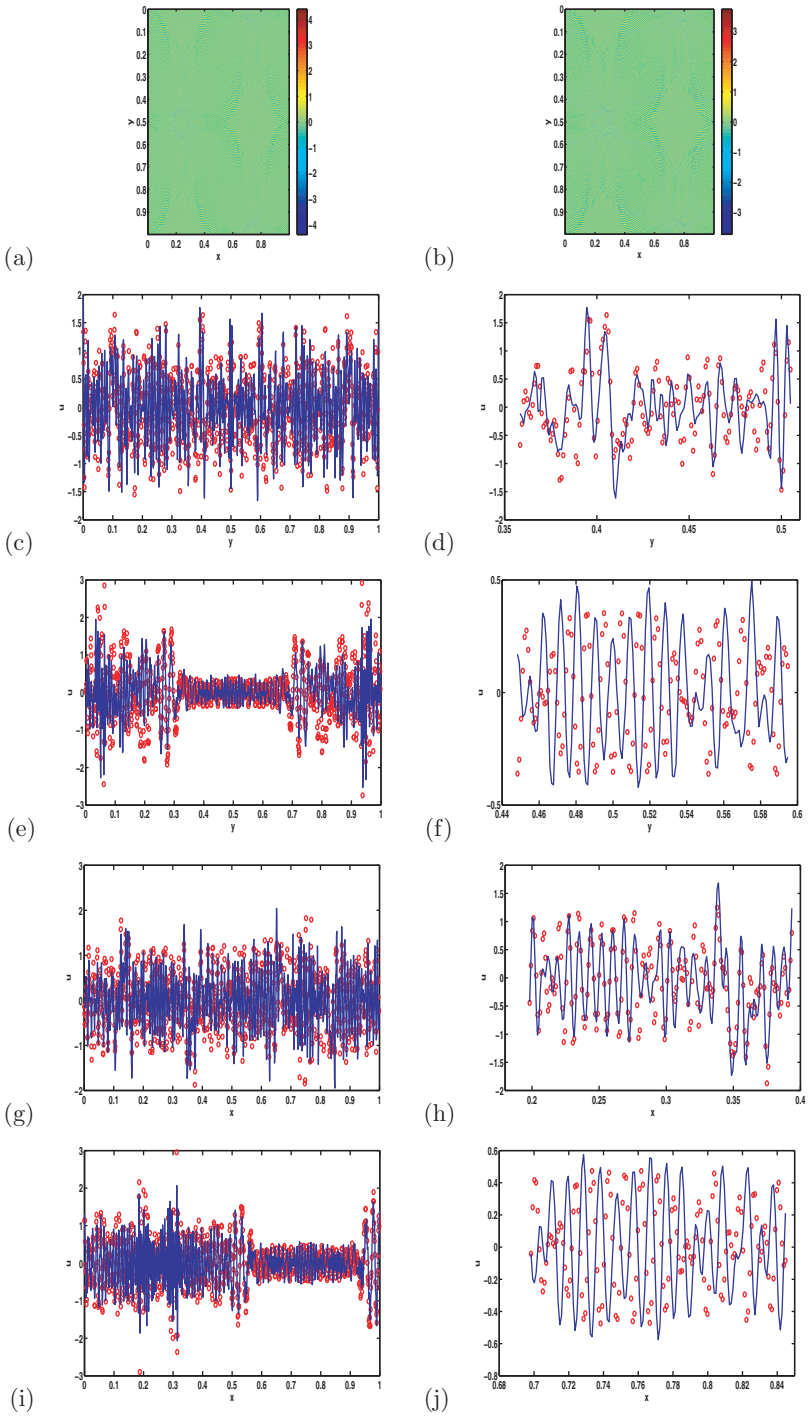

FIG. 9. *Example* 4. *No beam reinitialization at* $T = 0.6$*:* (a) *the exact solution;* (b) *the beam solution;* (c) *the slice at* $x = 0.5$*;* (d) *a windowed slice at* $x = 0.5$*;* (e) *the slice at* $x = 0.75$*;* (f) *a windowed slice at* $x = 0.75$*;* (g) *the slice at* $y = 0.2$*;* (h) *a windowed slice at* $y = 0.2$*;* (i) *the slice at* $y = 0.5$*;* (j) *a windowed slice at* $y = 0.5$*: the beam solution ("-") and the exact solution ("o").*

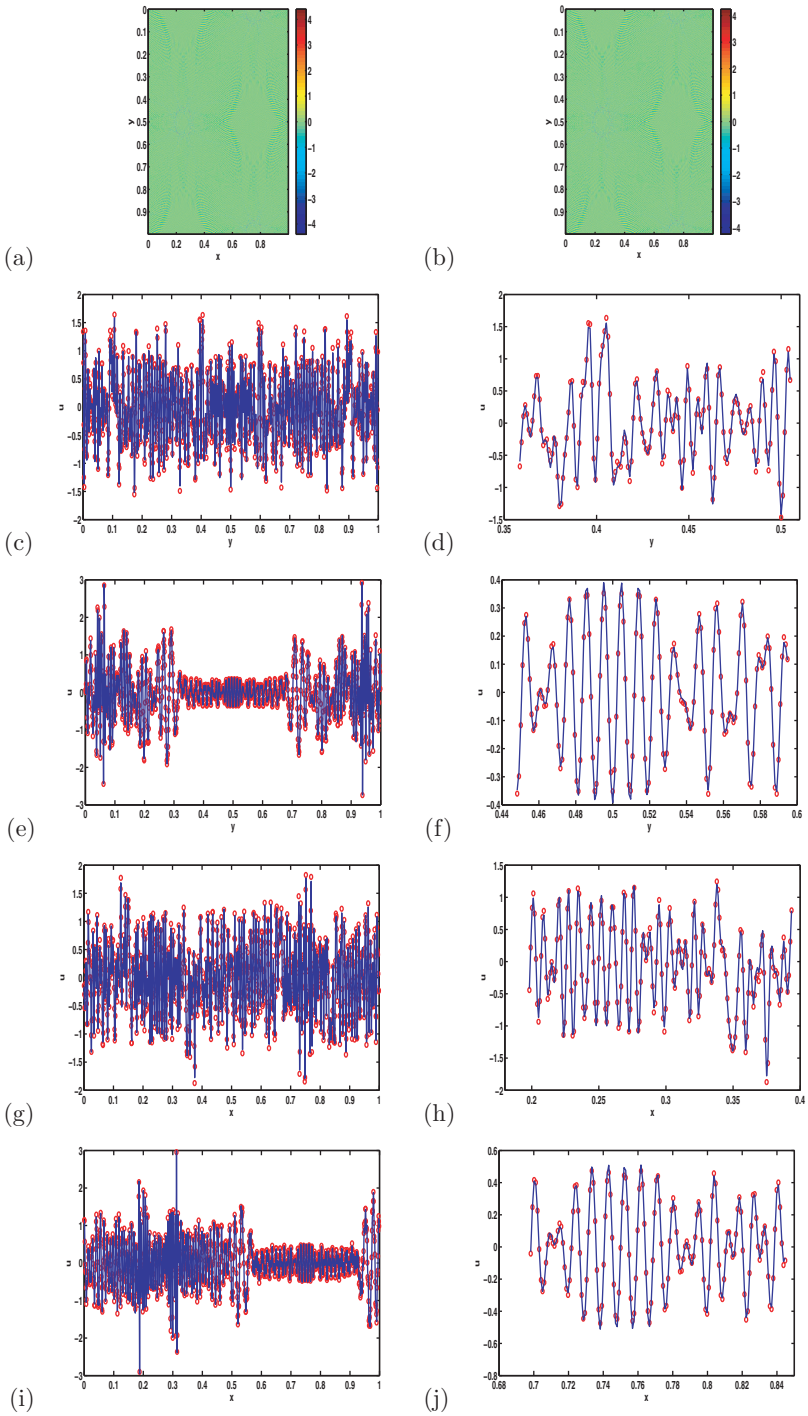

FIG. 10. *Example* 4. *With beam reinitialization at* $T = 0.6$. *The reinitialization is set at* $t = 0.2$ *and* $t = 0.40$, *respectively:* (a) *the exact solution;* (b) *the beam solution;* (c) *the slice at* $x = 0.5$*;* (d) *a windowed slice at* $x = 0.5$*;* (e) *the slice at* $x = 0.75$*;* (f) *a windowed slice at* $x = 0.75$*;* (g) *the slice at* $y = 0.2$*;* (h) *a windowed slice at* $y = 0.2$*;* (i) *the slice at* $y = 0.5$*;* (j) *a windowed slice at* $y = 0.5$*: the beam solution ("-") and the exact solution ("o").*

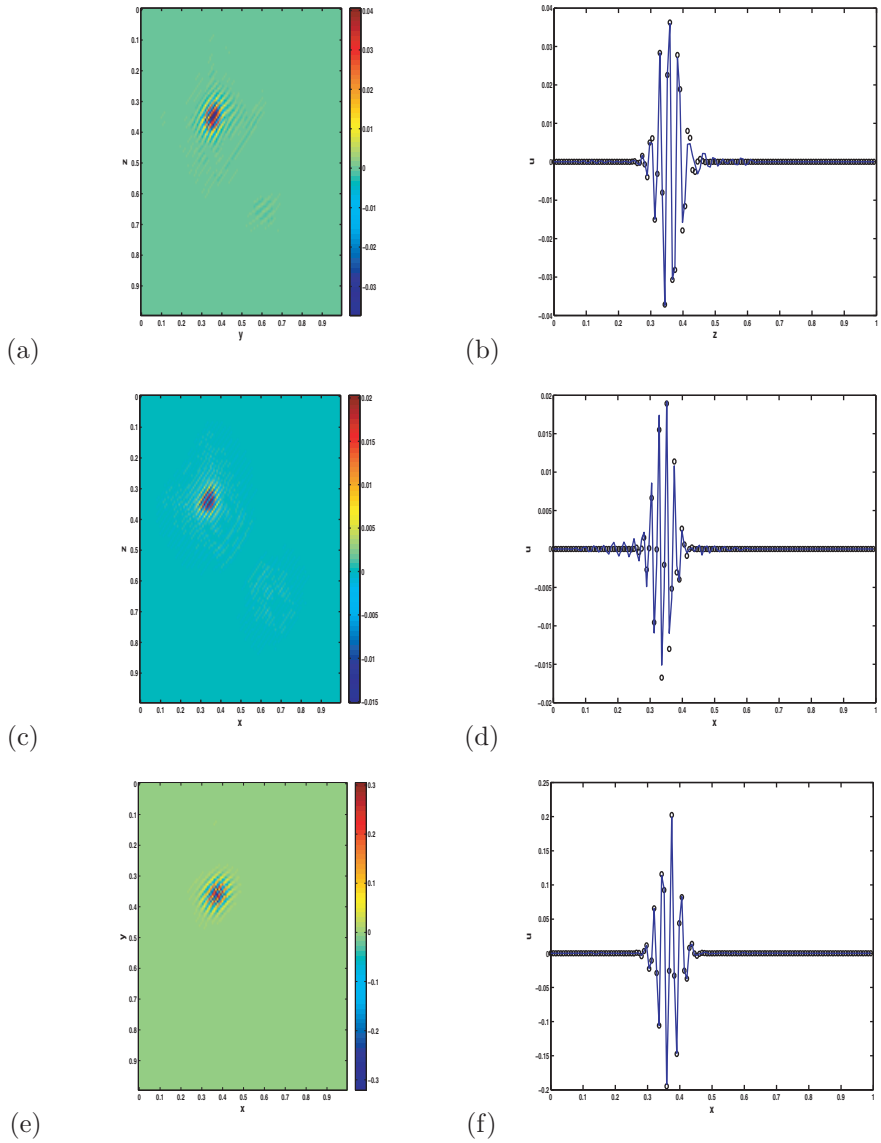

FIG. 11. *Example* 5. *(A three-dimensional slow velocity model.) The final time* $T = 0.50$ *without reinitialization:* (a) *the section of* $x = 0.4219$*;* (b) *comparison of the slices at* $x = 0.4219$ *and* $y = 0.3359$*;* (c) *the section of* $y = 0.4531$*;* (d) *comparison of the slices at* $y = 0.4531$ *and* $z = 0.3359$*;* (e) *the section of* $z = 0.3438$*;* (f) *comparison of the slices at* $z = 0.3438$ *and* $y = 0.3359$*: the beam solution ("-") and the exact solution ("o").*

("-"). Figure 11(c) plots a y-section of the three-dimensional wavefield at $y = 0.4531$, and Figure 11(d) shows slices at $y = 0.4531$ and $z = 0.3359$, in which we compare the exact solution ("o") with the beam solution ("-"). Figure 11(e) shows a z-section of the three-dimensional wavefield at $z = 0.3438$, and Figure 11(f) gives slices at $z = 0.3438$ and $y = 0.3359$, in which we compare the exact solution ("o") with the beam solution ("-"). As we can see from the comparison results, the beam solution matches well with the exact solution.

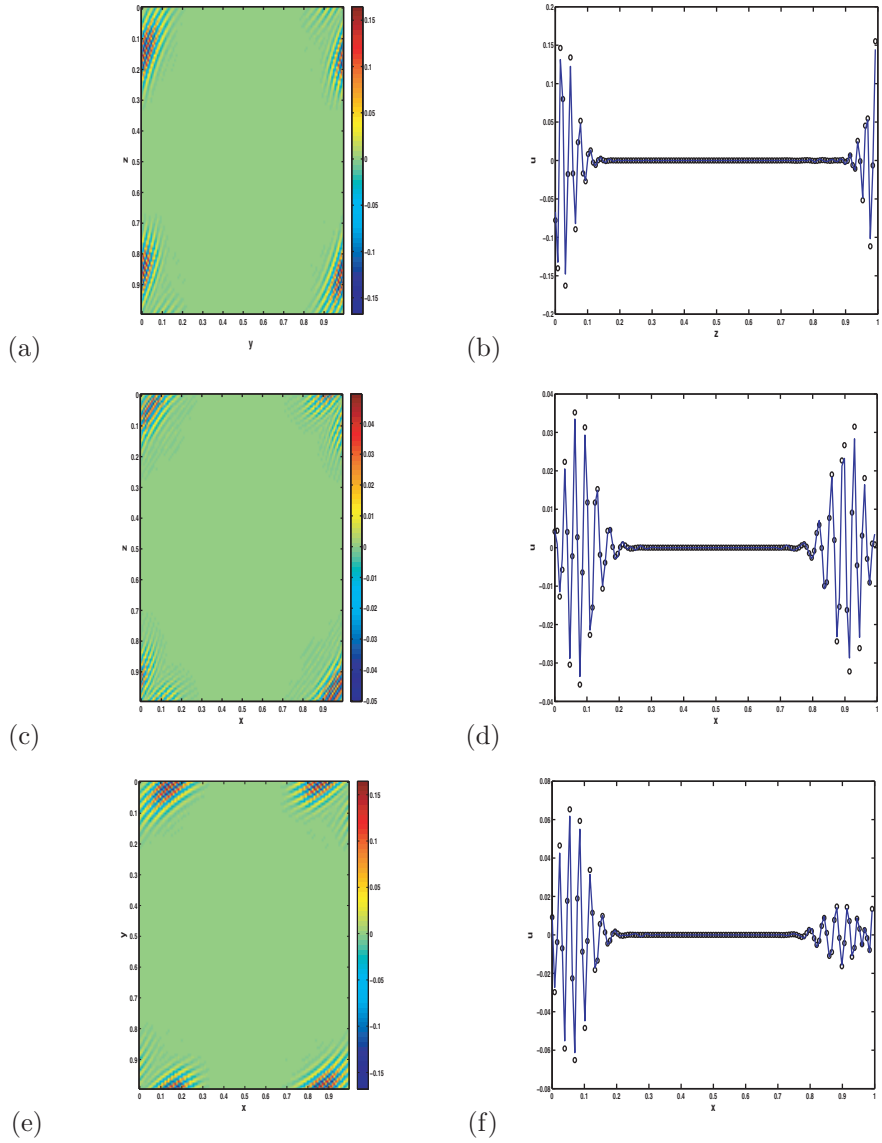

FIG. 12. *Example* 6. *(A three-dimensional fast velocity model.) The final time* $T = 0.80$ *with reinitialization at* $t = 0.2, 0.4, 0.6$*:* (a) *the section of* $x = 0.9922$*;* (b) *comparison of the slices at* $x = 0.9922$ *and* $y = 0.1406$*;* (c) *the section of* $y = 0.9922$*;* (d) *comparison of the slices at* $y = 0.9922$ *and* $z = 0.0156$*;* (e) *the section of* $z = 0.9922$*;* (f) *comparison of the slices at* $z = 0.9922$ *and* $y = 0.0391$*: the beam solution ("-") and the exact solution ("o").*

7.3.2. Example 6: A three-dimensional fast velocity. The background potential is given by $V(x, y, z) = 1.0 + 0.1 \sin(2\pi x) \cos(2\pi y) \sin(2\pi z)$, and the initial wave function is given as that in (69).

In Figure 12, we take $\beta = 32$, $\eta = 10^{-3}$, and use $N^3 = 128^3$ uniform distributed points to discretize the unit cube $[0, 1]^3$. The beam solution is computed up to the final time $T = 0.8$ with reinitialization at $t = 0.2, 0.4, 0.6$. Figure 12(a) shows an x-section of the three-dimensional wavefield at $x = 0.9922$, and Figure 12(b) plots

slices at $x = 0.9922$ and $y = 0.1406$, in which we compare the exact solution ("o") with the beam solution ("-"). Figure 12(c) shows a y-section of the three-dimensional wavefield at $y = 0.9922$, and Figure 12(d) plots slices at $y = 0.9922$ and $z = 0.0156$, in which we compare the exact solution ("o") with the beam solution ("-"). Figure 12(e) shows a z-section of the three-dimensional wavefield at $z = 0.9922$, and Figure 12(f) plots slices at $z = 0.9922$ and $y = 0.0391$, in which we compare the exact solution ("o") with the beam solution ("-"). As we can see from the comparison results, the beam solution matches well with the exact solution.

8. Conclusions. This paper addressed two main computational aspects of the Gaussian beam methods for the wave equation: beam initialization for general initial condition and long time propagation. For the beam initialization problem, we proposed fast multiscale Gaussian wavepacket transforms and developed based on them a new efficient algorithm for beam initialization for general initial data. The second problem was addressed by reinitializing the beam representation using this algorithm when the beam widths get too wide.

In [3], Candès and Demanet represented the solution operator of the wave equation in the curvelet frame and showed that this representation is optimally sparse in an asymptotic sense. One can get similar optimality results with other frames, such as wave atoms [7], that satisfy the parabolic scaling principle. However, numerical tests show that these representations often have a significant constant factor, which limits their practical application.

We can also view our Gaussian beam method as a way to represent the solution operator of the wave equation. However, this representation is quite different from the curvelet representation of the solution operator. In our Gaussian beam method, the initial data is represented as a linear combination of the multiscale Gaussian wavepackets, which are indexed by the discrete (ℓ, i, k) decimation of the phase space. However, at the final time T, the Gaussian beams are no longer multiscale Gaussian wavepackets any more, as they are in general off the grid indexed by (ℓ, i, k). Since the Gaussian beams (indexed by continuous parameters) are much more flexible and general than the multiscale Gaussian wavepackets (indexed by the discrete (ℓ, i, k) parameters), what we gain in the Gaussian beam method is a representation of the solution operator that is one-to-one and more efficient.

Acknowledgments. Qian would like to thank R. Burridge, S. Osher, and J. Ralston for their interest and encouragement in this work. Ying would like to thank E. Candès and L. Demanet for discussions regarding curvelets and wave atoms and B. Engquist, N. Tanushev, and R. Tsai for discussions regarding the Gaussian beam method.

REFERENCES

[1] V. M. Babich and V. S. Buldyrev, *Asymptotic Methods in Short Wave Diffraction Problems*, Izdat, Nauka, Moscow, 1972 (in Russian).

[2] S. Bougacha, J. Akian, and R. Alexandre, *Gaussian beams summation for the wave equation in a convex domain*, Commun. Math. Sci., 7 (2009), pp. 973–1008.

[3] E. Candès and L. Demanet, *The curvelet representation of wave propagators is optimally sparse*, Comm. Pure Appl. Math., 58 (2005), pp. 1472–1528.

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

[5] E. Candès and D. Donoho, *New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities*, Comm. Pure Appl. Math., 57 (2004), pp. 219–266.

[6] V. Cerveny, M. Popov, and I. Psencik, *Computation of wave fields in inhomogeneous media-Gaussian beam approach*, Geophys. J. R. Astr. Soc., 70 (1982), pp. 109–128.
[7] L. Demanet and L. Ying, *Wave atoms and sparsity of oscillatory patterns*, Appl. Comput. Harmon. Anal., 23 (2007), pp. 368–387.
[8] L. Demanet and L. Ying, *Wave atoms and time upscaling of wave equations*, Numer. Math., 113 (2009), pp. 1–71.
[9] N. Hill, *Gaussian beam migration*, Geophys., 55 (1990), pp. 1416–1428.
[10] L. Hörmander, *On the existence and the regularity of solutions of linear pseudo-differential equations*, Enseign. Math. (2), 17 (1971), pp. 99–163.
[11] S. Leung and J. Qian, *Eulerian Gaussian beam methods for Schrödinger equations in the semi-classical regime*, J. Comput. Phys., 228 (2009), pp. 2951–2977.
[12] S. Leung, J. Qian, and R. Burridge, *Eulerian Gaussian beams for high frequency wave propagation*, Geophys., 72 (2007), pp. SM61–SM76.
[13] V. P. Maslov, *The Complex WKB Method for Nonlinear Equations* I: *Linear Theory*, Birkhäuser-Verlag, Basel, 1994.
[14] M. Motamed and O. Runborg, *Taylor expansion and discretization errors in Gaussian beam superposition*, Wave Motion, to appear.
[15] J. Qian and L. Ying, *Fast Gaussian wavepacket transforms and Gaussian beams for the Schrödinger equation*, J. Comput. Phys., 229 (2010), pp. 7848–7873.
[16] J. Ralston, *Gaussian beams and the propagation of singularities*, in Studies in Partial Differential Equations. MAA Stud. Math. 23, Mathematical Association of America, Washington, D.C., 1983, pp. 206–248.
[17] H. Smith, *A parametrix construction for wave equations with $C^{1,1}$ coefficients*, Ann. Inst. Fourier (Grenoble), 48 (1998), pp. 797–835.
[18] N. Tanushev, *Superpositions and higher order Gaussian beams*, Commun. Math. Sci., 6 (2008), pp. 449–475.
[19] N. Tanushev, B. Engquist, and R. Tsai, *Gaussian beam decomposition of high frequency wave fields*, J. Comput. Phys., 228 (2009), pp. 8856–8871.
[20] N. Tanushev, J. Qian, and J. Ralston, *Mountain waves and Gaussian beams*, Multiscale Model. Simul., 6 (2007), pp. 688–709.
[21] B. S. White, *The stochastic caustic*, SIAM J. Appl. Math., 44 (1984), pp. 127–149.