Journal of Computational Physics 229 (2010) 7848–7873

Contents lists available at ScienceDirect

Journal of Computational Physics

journal homepage: www.elsevier.com/locate/jcp

Fast Gaussian wavepacket transforms and Gaussian beams for the Schrödinger equation

Jianliang Qian [a,*], Lexing Ying [b]

[a] Department of Mathematics, Michigan State University, East Lansing, MI 48824, United States
[b] Department of Mathematics and ICES, University of Texas at Austin, Austin, TX 78712, United States

ARTICLE INFO

Article history:
Received 1 March 2010
Received in revised form 27 May 2010
Accepted 28 June 2010
Available online 21 July 2010

Keywords:
Fast Gaussian wavepacket transforms
Gaussian beams
Schrödinger equation

ABSTRACT

This paper introduces a wavepacket-transform-based Gaussian beam method for solving the Schrödinger equation. We focus on addressing two computational issues of the Gaussian beam method: how to generate a Gaussian beam representation for general initial conditions and how to perform long time propagation for any finite period of time. To address the first question, we introduce fast Gaussian wavepacket transforms and develop on top of them an efficient initialization algorithm for general initial conditions. Based on this new initialization algorithm, we address the second question by reinitializing the beam representation when the beams become too wide. Numerical examples in one, two, and three dimensions demonstrate the efficiency and accuracy of the proposed algorithms. The methodology can be readily generalized to deal with other semi-classical quantum mechanical problems.

1. Introduction

We consider the following linear evolution equation:

$$\begin{aligned} &-\imath\hbar U_t + H(x, -\imath\hbar\partial_x)U = 0, \quad x \in \mathbb{R}^d,\ t > 0,\\ &U(x,0) = U_0(x), \end{aligned}$$

where $\hbar$ is a small positive parameter, $\imath = \sqrt{-1}$, U_0 is a compactly supported L^2-function, and $H(x,p)$ is a real principal symbol. The example that we will focus on is the Schrödinger equation with $H(x,p) = V(x) + \frac{|p|^2}{2}$. As this equation propagates oscillations of wavelength $\hbar$ in space and time, resolving such oscillations by direct finite difference methods requires computational grid of size $O(1/\hbar)$ in each direction, which is costly in practice.

Classical mechanics provides a deterministic description of the motion of a classical particle by the Newton laws with well-defined position and momentum. On the other hand, quantum mechanics gives an indeterministic description of the motion of a quantum particle by the Schrödinger equation with uncertainty in its position and momentum. However, when a quantum particle is appropriately localized both in space and in momentum in a balanced wave, its mean location and mean momentum follow the Newton laws, up to quantum correction. This results in the field of semi-classical quantum mechanics, which provides a short-wavelength link between classical mechanics and quantum mechanics and maintains contact with the structure imposed by classical mechanics without sacrificing quantum mechanical accuracy. On the computational side, numerical methods based on the semi-classical approaches are also sought as efficient alternatives to capture the highly oscillatory phenomenon arising from quantum mechanics.

* Corresponding author. Tel.: +1 517 353 6334; fax: +1 517 432 1562.
E-mail addresses: qian@math.msu.edu (J. Qian), lexing@math.utexas.edu (L. Ying).

0021-9991/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jcp.2010.06.043

One of such semi-classical methods is the so-called Gaussian beam method. The idea underlying the Gaussian beam method is to construct an asymptotic solution along the trajectory of a classical particle. Similar to the WKBJ expansion, a Gaussian beam has an amplitude and phase decomposition. However, an important difference is that the quadratic term of the phase function is constructed so as to have a positive imaginary part away from the ray trajectory, thus providing the beam with a Gaussian profile at any time. The evolution of the Gaussian beam over time is described by a set of ordinary differential equations that are independent of the small parameter $\hbar$.

In order to apply the Gaussian beam methods, one faces two major questions. A single Gaussian beam has a Gaussian profile and it is certainly incapable of describing a general initial condition $U_0(x)$. The first question concerns the initialization of the Gaussian beam representation, *i.e.*, how to *decompose a general $U_0(x)$ into a collection of Gaussian beams efficiently*. Several initialization algorithms have been proposed in the literature. A commonly used approach is the asymptotic decomposition (AD) method [25,23,13], which assumes the initial condition $U_0(x)$ to have a WKBJ form, *i.e.*, $U_0(x) = A(x)\exp(\imath\tau(x)/\hbar)$. However, this method is clearly not general. Another approach is the FBI transform (Fourier–Bros–Iagolinitzer) method [13], which first computes the FBI transform of $U_0(x)$ and then discretizes this phase-space distribution to obtain the beam representation. This method is quite general; however, sometimes the number of resulting beams can be quite large. Recently in [24], Tanushev et al. proposed a method that places beams one by one via greedy optimization. This method sometimes can result in very few beams; however, its computational cost is rather high. In this paper, we introduce fast Gaussian wavepacket transforms and build on top of them a highly efficiently algorithm for constructing beam representation for arbitrary initial condition. The cost of this fast wavepacket transform is comparable to that of a single fast Fourier transform.

The second question concerns *long time propagation* of Gaussian beams. A simple analysis shows that the size of a Gaussian beam can grow exponentially with time. A overly-extended Gaussian beam not only leads to large approximation error but also makes the final summation step extremely expensive. One needs to address this question in order to carry out long time propagation of Gaussian beams efficiently and accurately. Our solution is to monitor the beam support and reinitialize the beam representation whenever necessary. This simple reinitialization process is rather difficult to implement under existing initialization algorithms, as they are either too special or too costly. However, when combined with our initialization algorithm based on fast Gaussian wavepacket transforms, it works very well and significantly improves the accuracy and efficiency of the Gaussian beam methods.

Furthermore, we show that the new wavepacket-transform-based Gaussian beam method yields an asymptotic solution to the Schrödinger equation.

1.1. Related work

The idea underlying Gaussian beams is simply to build asymptotic solutions to partial differential equations concentrated on a single curve through the domain; this single curve is nothing but a ray as shown in [22]. 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 [10,22]. 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 pseudo-differential equations in a natural way, including Helmholtz [14] and Schrödinger equations [13].

In geophysical applications, Gaussian beam superpositions have been used for seismic wave modeling [5] and seismic wave migration [9]. The numerical implementations in these areas are based on ray-centered coordinates that prove to be computationally inefficient [5,9]. More recently, based on Ralston and co-workers [22,25] a purely Eulerian computational approach was proposed in [14] to overcome some of these difficulties; it can be easily applied to both high frequency waves and semi-classical quantum mechanics. In [25], 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 [23,24,15,20,3] for recent works on Gaussian beams related to wave equations and other applications.

In quantum mechanics, some variants of Gaussian beams, such as frozen Gaussian beams and Gaussian wavepackets, have been used to construct approximate solutions to Schrödinger equations in the semi-classical regime [12,7,8]. More recently, Leung and Qian [13] presented an Eulerian formulation of the Gaussian beam methods for the Schrödinger equation following their earlier work in [14]. A related work is outlined by Jin et al. in [11].

We would like to mention that our Gaussian wavepacket transforms are closely related to the FBI transform [17], which has been used in [13] to initialize the beam representation. The FBI transform computes the phase space density of a function and, through a reproducing formula, writes the function as a sum of Gaussian wavepackets. However, as the number of dimensions is doubled by moving into the phase space, the number of Gaussian wavepackets required for a prescribed accuracy can be quite large, depending on the initial wave function. The Gaussian wavepacket transforms introduced here can be viewed as a discrete version of the FBI transform in the sense that it effectively samples the FBI transform on a (product) Cartesian grid with an $O(\hbar^{1/2})$ spacing in space and $O(\hbar^{-1/2})$ spacing in frequency. Through this careful decimation, the Gaussian wavepacket transforms provide much more efficient representations without sacrificing the reproducing property.

1.2. Contents

The rest of the paper is organized as follows. Section 2 describes the Lagrangian formulation of the Gaussian beam method for the Schrödinger equation. In Section 3, we present the fast Gaussian wavepacket transforms for continuous and

discrete data, and describe how to use them for initial decomposition. Section 4 discusses the reinitialization algorithm for long time propagation. In Section 5, we show that the solution obtained through wavepacket transforms is a global asymptotic solution of the Schrödinger equation. Section 6 presents numerical examples in one-, two-, and three-dimensional cases. Finally, conclusions and discussion on future work are given in Section 7.

2. Gaussian beam methods for the Schrödinger equation

2.1. General setup

We consider the following evolution equation:

$$-\imath\hbar U_t + H(x, -\imath\hbar\partial_x)U = 0, \quad x \in \mathbb{R}^d,\ t > 0, \tag{1}$$

$$U(x,0) = U_0(x), \tag{2}$$

where $\hbar$ is a small positive parameter, U_0 is a compactly supported L^2-function, and $H(x,p)$ is a real principal symbol that satisfies certain smoothness conditions. In addition, we assume that $H(x,-\imath\hbar\partial_x)$ is a self-adjoint (or symmetric) differential operator.

To construct Gaussian beams for the above equation, we follow [22,25,18] and start with the WKBJ ansatz,

$$A(x,t)\exp\left(\frac{\imath\tau(x,t)}{\hbar}\right). \tag{3}$$

The functions $A(x,t)$ and $\tau(x,t)$ are all assumed to be smooth, and these requirements are feasible because the beam solution is constructed to be concentrated on a single curve γ; this is the essential difference between traditional WKBJ asymptotic solutions and Gaussian beam solutions. As a result, the requirements on the phase function $\tau(x,t)$ are slightly different from those of traditional WKBJ asymptotics. We require that $\tau(x,t)$ be real valued on γ, but away from this curve γ, $\tau(x,t)$ can be complex valued with the restriction that the imaginary part of the second-order derivative $\tau_{xx}(x,t)$ is positive definite. This will make $U(x,t)$ look like a Gaussian with variance $O(\hbar)$ on planes perpendicular to γ for any fixed time t.

Substituting the ansatz (3) into (1) and considering the leading-order singular terms corresponding to $\hbar$, we have the following eikonal and transport equations (see [19]):

$$\tau_t(x,t) + H(x, \tau_x(x,t)) = 0, \tag{4}$$

$$A_t(x,t) + H_p(x,\tau_x)\cdot A_x(x,t) + \frac{A(x,t)}{2}\left(\operatorname{trace}(H_{pp}(x,\tau_x)\tau_{xx}(x,t) + H_{xp}(x,\tau_x))\right) = 0. \tag{5}$$

Notice that if the differential operator $H(x,-\imath\hbar\partial_x)$ is not symmetric, then the transport equation will take a slightly different form; see [19]. In order to simplify the presentation, we assume that the operator $H(x,-\imath\hbar\partial_x)$ is symmetric.

We are going to build Gaussian beams around central rays $\gamma = \{(x(t),t):\ t \geqslant 0\}$, which are the x-projections of bicharacteristics for the Hamilton–Jacobi equation (4), according to the Gaussian beam theory [22]. Setting $p(t) = \tau_x(x(t),t)$, the bicharacteristic $\{(x(t),p(t)):\ t \geqslant 0\}$ satisfies the following Hamiltonian system:

$$\begin{aligned} \dot{x} &= \frac{dx}{dt} = H_p(x(t),p(t)), \quad x|_{t=0} = x_0, \\ \dot{p} &= \frac{dp}{dt} = -H_x(x(t),p(t)), \quad p|_{t=0} = p_0, \end{aligned} \tag{6}$$

where t is a parameter parameterizing the bicharacteristic emanating from the initial point (x_0,p_0). Along the ray $\gamma = \{(x(t),t):\ t \geqslant 0\}$ (the x-projection of the bicharacteristic $\{(x(t),p(t)):\ t \geqslant 0\}$), the phase function $\tau(t) \equiv \tau(x(t),t)$ satisfies

$$\dot{\tau} = \frac{d\tau}{dt} = p(t)\cdot H_p(x(t),p(t)) - H(x(t),p(t)), \quad \tau|_{t=0} = \tau_0, \tag{7}$$

where τ_0 is an initial value of the phase function τ for the bicharacteristic emanating from the initial point (x_0,p_0).

Since $\tau(t) = \tau(x(t),t)$ and $p(t) = \tau_x(x(t),t)$ are constructed along the ray $\{(x(t),t):\ t \geqslant 0\}$ by solving Eqs. (6) and (7), we have available the zero- and first-order derivative information of the phase function along the ray. Next we determine the second-order derivative (the Hessian), $M(t) = \tau_{xx}(x(t),t)$, along the ray. To achieve this, we differentiate the Hamilton–Jacobi equation (4) twice with respect to x along the ray, yielding

$$\tau_{t,xx} + \tau_{xxx}H_p + H_{xp}^T\tau_{xx} + \tau_{xx}H_{px} + \tau_{xx}H_{pp}\tau_{xx} + H_{xx} = 0. \tag{8}$$

By using the first component of the Hamiltonian system (6) along the ray to reorganize the first two terms in the above equation, we have

$$\dot{M} + H_{xp}^T M + MH_{px} + MH_{pp}M + H_{xx} = 0, \tag{9}$$

which is appended with a certain initial condition $M|_{t=0} = M_0$. This is a Riccati equation which as a nonlinear ODE does not have a unique global smooth solution in general. However, since this Riccati equation for the symmetric matrix M is related

to a Hamiltonian system, complexifying the equation by specifying an appropriate complex initial condition for M can lead to a unique global smooth solution for the Riccati equation; in [22,18,25] this has been shown to be true by complexifying the underlying symplectic structure for the Hamiltonian system. The essential idea of such a proof is introducing the following variational system for $d \times d$ matrix-valued functions $B(t)$ and $C(t)$ along the bicharacteristic starting from (x_0, p_0):

$$\begin{aligned} \dot{B} &= -H_{xp}^T B - H_{xx} C, \quad B|_{t=0} = B_0 + \imath\epsilon I, \\ \dot{C} &= H_{pp} B + H_{px} C, \quad C|_{t=0} = I, \end{aligned} \tag{10}$$

where ϵ is a positive constant, I is the $d \times d$ identity matrix, and matrix B_0 is chosen to take into account the initial phase function. Here $B = B(t; x_0, p_0)$ 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 $x_0 = \alpha$,

$$B(t; x_0, p_0) = \frac{\partial p}{\partial \alpha}, \quad C(t; x_0, p_0) = \frac{\partial x}{\partial \alpha}.$$

By choosing the initial data for B to be of the form $B_0 + \imath\epsilon I$, we effectively complexify the variational system (10); consequently, a solution to the above equations exists on any interval $t \in [0, T]$. Moreover, we have the following lemma on the bound of the solution; its proof can be found in [22,18,25].

Lemma 2.1. *Under the above assumptions, $C(t)$ is non-singular for any t, and $\mathrm{Im}(B(t)C^{-1}(t))$ is positive definite.*

On the other hand, one may also manipulate the system (10) to conclude that $B(t)C^{-1}(t)$ satisfies the Riccati equation (9) with the initial condition $B(0)C^{-1}(0) = B_0 + \imath\epsilon I$. Therefore, if we choose the initial data for M such that $M(0) = B_0 + \imath\epsilon I = B(0)C^{-1}(0)$, then by uniqueness we have $M(t) = B(t)C^{-1}(t)$, and $\mathrm{Im}(M(t)) = \mathrm{Im}(B(t)C^{-1}(t))$ is positive definite along the ray $\{(x(t), t)\colon t \geqslant 0\}$ by Lemma 2.1.

Now that we have available the zero-, first-, and second-order derivative information of the phase function along the ray, we can use the following second-order Taylor expansion centered at the ray $\{(x(t), t)\colon t \geqslant 0\}$ to define a smooth global approximate phase function:

$$\tau(x, t) \equiv \tau(t) + p(t) \cdot (x - x(t)) + \frac{1}{2}(x - x(t))^T M(t)(x - x(t)), \tag{11}$$

which is valid near the ray $\{(x(t), t)\colon t \geqslant 0\}$.

Next we need to determine the amplitude function A. By reorganizing the first two terms in the transport equation (5) according to the beam theory, the amplitude function $A(t) \equiv A(x(t), t)$ along the ray $\{(x(t), t)\colon t \geqslant 0\}$ satisfies the following equation:

$$\dot{A} = -\frac{A}{2}(\mathrm{trace}(H_{pp} M + H_{xp})), \quad A|_{t=0} = A_0. \tag{12}$$

To obtain a smooth global approximate amplitude function, we use the following extension:

$$A(x, t) \equiv A(x(t), t) = A(t), \tag{13}$$

which is valid near the ray $\{(x(t), t)\colon t \geqslant 0\}$.

Inserting (11) and (13) into the WKBJ ansatz yields an asymptotically valid solution:

$$\Phi(x, t) = A(x, t) \exp\left(\imath \frac{\tau(x, t)}{\hbar}\right). \tag{14}$$

This beam solution is concentrated on a single smooth curve $\gamma = \{(x(t), t)\colon t \geqslant 0\}$ which is the x-projection of the bicharacteristic $\{(x(t), p(t))\colon t \geqslant 0\}$ emanating from (x_0, p_0) at $t = 0$.

The functions $x(t)$, $p(t)$, $\tau(t)$, $M(t)$, $A(t)$, $\tau(x, t)$, $A(x, t)$, and $\Phi(x, t)$ introduced above are uniquely determined once the initial data x_0, p_0, τ_0, M_0, and A_0 are specified. We call x_0, p_0, τ_0, M_0, and A_0 the parameters of the Gaussian beam $\Phi(x, t)$ and denote them collectively as a tuple $\alpha = (x_0, p_0, \tau_0, M_0, A_0)$. In order to emphasize the dependency of the Gaussian beam on α, these functions are denoted, respectively, by $x_\alpha(t)$, $p_\alpha(t)$, $\tau_\alpha(t)$, $M_\alpha(t)$, $A_\alpha(t)$, $\tau_\alpha(x, t)$, $A_\alpha(x, t)$, and $\Phi_\alpha(x, t)$.

2.2. Gaussian beam method for the Schrödinger equation

In the case of the Schrödinger equation for a particle with unit mass, we have

$$-\imath\hbar U_t + V(x)U - \frac{\hbar^2}{2}\Delta U = 0, \quad x \in \mathbb{R}^d, \ t > 0, \tag{15}$$

$$U(x, 0) = U_0(x), \tag{16}$$

where the potential $V(x)$ is smooth, $\hbar$ is reduced Planck's constant. In this case, the Hamiltonian is $H(x, p) = V(x) + \frac{|p|^2}{2}$.

Putting the definition of $H(x,p)$ into (4) and (5) gives

$$\tau_t + V(x) + \frac{1}{2}\tau_x^2 = 0,$$
$$A_t + \tau_x \cdot A_x + \frac{1}{2}\Delta\tau A = 0.$$

Simplifying the equations of $x(t)$, $p(t)$, $\tau(t)$, $M(t)$, and $A(t)$ with $H(x,p) = V(x) + \frac{|p|^2}{2}$ gives the following set of equations:

$$\begin{aligned}
&\dot{x} = p, \quad x|_{t=0} = x_0,\\
&\dot{p} = -V_x, \quad p|_{t=0} = p_0,\\
&\dot{\tau} = \frac{|p|^2}{2} - V(x), \quad \tau|_{t=0} = \tau_0,\\
&\dot{M} = -M^2 - V_{xx}, \quad M|_{t=0} = M_0,\\
&\dot{A} = -\frac{A}{2}(\mathrm{trace}(M)), \quad A|_{t=0} = A_0.
\end{aligned} \tag{17}$$

Once the values of $x(t)$, $p(t)$, $\tau(t)$, $M(t)$, and $A(t)$ are ready, the Gaussian beam at time t is given by (11) and (14).

For a given tuple of parameters $\alpha = (x_0, p_0, \tau_0, M_0, A_0)$, the associated Gaussian beam $\Phi_\alpha(x,t)$ takes a fixed Gaussian profile at time $t = 0$. For a general $U_0(x)$, one needs to find a set I of tuples such that at time $t = 0$ the sum of the Gaussian beams associated with the tuples in I reproduces the initial condition $U_0(x)$, *i.e.*,

$$U_0(x) \approx \sum_{\alpha\in I} \Phi_\alpha(x,0).$$

Once this is done, the linearity of (15) and (16) gives the following Gaussian beam solution

$$U(x,t) \approx \sum_{\alpha\in I} \Phi_\alpha(x,t).$$

3. Gaussian wavepacket transforms and initialization

The discussion in Section 2 shows that at time $t = 0$ a Gaussian beam associated to the initial data $\alpha = (x_0, p_0, \tau_0, M_0, A_0)$ takes the form

$$A_0 \exp\left(\frac{\imath}{\hbar}\left(\tau_0 + p_0\cdot(x-x_0) + \frac{1}{2}(x-x_0)^T M_0 (x-x_0)\right)\right),$$

where the Hessian M_0 is $O(1)$ and purely imaginary. It is clear from the formula that this is a modulated Gaussian with an $O(\hbar^{1/2})$ effective width in space and an $O(\hbar^{-1/2})$ effective width in frequency. Therefore, the task of initial decomposition is equivalent to representing the function $U_0(x)$ as a linear combination of such modulated Gaussian functions. The tool that we will use for this task is essentially the windowed Fourier transform [16]; however, one needs to adapt it for the correct scaling and profile.

3.1. Gaussian wavepacket transforms

Let W be a constant of order $O(\hbar^{-1/2})$. The frequency (Fourier) domain $\mathbb{R}^d$ is partitioned into d-dimensional boxes of size W in each dimension. See Fig. 1 for the one-dimensional case and Fig. 2 for the two-dimensional case. Denote these boxes by B_i where $i = (i_1, i_2, \ldots, i_d)$ is a multiindex with integer components and the center of each B_i by $\xi_i = (\xi_{i,1}, \xi_{i,2}, \ldots, \xi_{i,d}) \in \mathbb{R}^d$.

Each B_i is associated with a smooth window function $g_i(\xi)$ that is compactly supported in a box centered at ξ_i with width $L = 2W$ (*i.e.*, $\prod_{s=1}^d [\xi_{i,s} - W, \xi_{i,s} + W]$). We further require $g_i(\xi)$ to approximate a Gaussian profile

$$g_i(\xi) \approx e^{-\left(\frac{|\xi-\xi_i|}{\sigma}\right)^2}$$

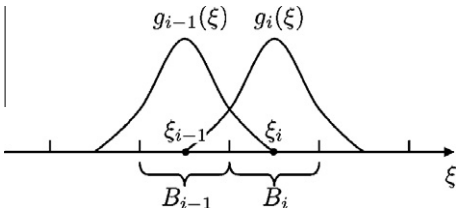

Fig. 1. Partitioning of the frequency domain in 1D.

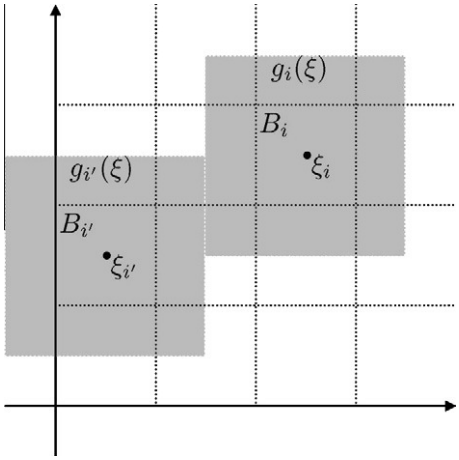

Fig. 2. Partitioning of the frequency domain in 2D. Each gray region stands for the support of a window function $g_i(\xi)$.

with $\sigma = W/2$. This choice of σ ensures that the tail of the Gaussian outside the support of $g_i(\xi)$ is sufficiently small. One can easily choose $g_i(\xi)$ to satisfy the following three *admissible* conditions:

1. $0 \leqslant g_i(\xi) \leqslant 1$ for any ξ and any i.
2. There exists $C_N > 0$ such that, for any ξ, $|\{i: g_i(\xi) > 0\}| \leqslant C_N$, where $|\cdot|$ stands for the cardinality of a set.
3. There exists $C_V > 0$ such that, for any ξ, there exists i such that $g_i(\xi) \geqslant C_V$.

Based on $g_i(\xi)$, we next introduce a dual window function $h_i(x)$ for each B_i by

$$h_i(\xi) = \frac{g_i(\xi)}{\sum_i (g_i(\xi))^2}.$$

From the properties of $g_i(\xi)$, it is not difficult to see that $h_i(\xi)$ is also a smooth function with the same support as $g_i(\xi)$. Moreover, by construction the products of $g_i(\xi)$ and $h_i(\xi)$ form a partition of unity: $\sum_i h_i(\xi) g_i(\xi) = 1$.

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

$$\hat{\phi}_{i,k}(\xi) = \frac{1}{L^{d/2}} e^{-2\pi\imath \frac{k\cdot\xi}{L}} g_i(\xi), \quad \forall k \in \mathbb{Z}^d, \tag{18}$$

$$\hat{\psi}_{i,k}(\xi) = \frac{1}{L^{d/2}} e^{-2\pi\imath \frac{k\cdot\xi}{L}} h_i(\xi), \quad \forall k \in \mathbb{Z}^d. \tag{19}$$

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

$$\phi_{i,k}(x) = \frac{1}{L^{d/2}} \int_{\mathbb{R}^d} e^{2\pi\imath \left(x - \frac{k}{L}\right)\cdot\xi} g_i(\xi) d\xi, \quad \forall k \in \mathbb{Z}^d, \tag{20}$$

$$\psi_{i,k}(x) = \frac{1}{L^{d/2}} \int_{\mathbb{R}^d} e^{2\pi\imath \left(x - \frac{k}{L}\right)\cdot\xi} h_i(\xi) d\xi, \quad \forall k \in \mathbb{Z}^d. \tag{21}$$

The definitions of $g_i(\xi)$ and $\phi_{i,k}(x)$ imply that

$$\phi_{i,k}(x) \approx \left(\sqrt{\frac{\pi}{L}}\sigma\right)^d \cdot e^{2\pi\imath \left(x - \frac{k}{L}\right)\cdot\xi_i} \cdot e^{-\sigma^2\pi^2 \left|x - \frac{k}{L}\right|^2}.$$

$\phi_{i,k}(x)$ is approximately a Gaussian function that centers at k/L, oscillates at frequency ξ_i, and has an $O(\sigma) = O(\hbar^{-1/2})$ effective width in the Fourier domain and an $O(1/\sigma) = O(\hbar^{1/2})$ effective width in the spatial domain. The functions $\{\phi_{i,k}(x)\}$ fit exactly into the requirement of a Gaussian beam and we will decompose the initial condition $U_0(x)$ into a linear combination of them.

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

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

$$C_1 \|f\|_2^2 \leqslant \sum_{i,k} |\langle \phi_{i,k}, f\rangle|^2 \leqslant C_2 \|f\|_2^2,$$

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

Proof. We prove the lemma for $\{\psi_{i,k}(x)\}$. The proof for $\{\phi_{i,k}(x)\}$ is similar and easier. We compute

$$\sum_{i,k}|\langle\psi_{i,k},f\rangle|^2=\sum_i\sum_k\left|\frac{1}{L^{d/2}}\int_{\mathbb{R}^d}e^{2\pi\mathrm{i}\frac{k\cdot\xi}{L}}h_i(\xi)\hat{f}(\xi)d\xi\right|^2=\sum_i\int_{\mathbb{R}^d}|h_i(\xi)\hat{f}(\xi)|^2d\xi=\int_{\mathbb{R}^d}\left(\sum_i|h_i(\xi)|^2\right)|\hat{f}(\xi)|^2d\xi. \tag{22}$$

It is then sufficient to show that $\sum_i|h_i(\xi)|^2$ is bounded from below and from above uniformly. From the conditions that $g_i(\xi)$ satisfies, we have for each i,

$$h_i(\xi)=\frac{g_i(\xi)}{\sum_i(g_i(\xi))^2}>\frac{g_i(\xi)}{C_N}.$$

Since one of $g_i(\xi)$ is at least C_V, there exists i such that $h_i(\xi)>\frac{C_V}{C_N}$. This implies that $\sum_i|h_i(\xi)|^2$ is bounded from below by $\frac{C_V^2}{C_N^2}$. On the other hand,

$$h_i(\xi)=\frac{g_i(\xi)}{\sum_i(g_i(\xi))^2}<\frac{g_i(\xi)}{C_V^2}<\frac{1}{C_V^2}.$$

Since $h_i(\xi)\neq 0$ only for at most C_N indices, $\sum_i|h_i(\xi)|^2$ is bounded from above by $\frac{C_N}{C_V^4}$. □

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

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

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

Proof. In the Fourier domain, we have

$$\begin{aligned}\sum_{i,k}\langle\psi_{i,k},f\rangle\hat{\phi}_{i,k}(\xi)&=\sum_{i,k}\left(\frac{1}{L^{d/2}}\int_{\mathbb{R}^d}e^{2\pi\mathrm{i}\frac{k\eta}{L}}h_i(\eta)\hat{f}(\eta)d\eta\right)\left(\frac{1}{L^{d/2}}e^{-2\pi\mathrm{i}\frac{k\cdot\xi}{L}}g_i(\xi)\right)\\&=\sum_i\left(\sum_k\left(\frac{1}{L^d}\int_{\mathbb{R}^d}e^{2\pi\mathrm{i}\frac{k\eta}{L}}h_i(\eta)\hat{f}(\eta)d\eta\right)e^{-2\pi\mathrm{i}\frac{k\cdot\xi}{L}}\right)g_i(\xi)=\sum_i h_i(\xi)\hat{f}(\xi)g_i(\xi)=\hat{f}(\xi).\end{aligned} \tag{23}$$

Here we use the fact that the function $h_i(\eta)\hat{f}(\eta)$ is supported in an interval of size L in each dimension and hence can be considered as a periodic function with period L. □

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

$$c_{i,k}=\langle\psi_{i,k},f\rangle=\langle\hat{\psi}_{i,k},\hat{f}\rangle=\frac{1}{L^{d/2}}\int_{\mathbb{R}^d}e^{2\pi\mathrm{i}\frac{k\cdot\xi}{L}}h_i(\xi)\hat{f}(\xi)d\xi. \tag{24}$$

We remark that according to Lemma 3.1 each coefficient $c_{i,k}$ is at most of order $O(1)$ and many coefficients are negligible. Given a set of coefficients $\{c_{i,k}\}$, the *inverse Gaussian wavepacket transform* synthesizes a function $u(x)$ defined by

$$u(x)=\sum_{i,k}c_{i,k}\phi_{i,k}(x). \tag{25}$$

3.2. Transforms of discrete signals

Let us now discuss 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

$$X=\{(n_1/N,n_2/N,\ldots,n_d/N):0\leqslant n_1,n_2,\ldots,n_d<N,n_1,n_2,\ldots,n_d\in\mathbb{Z}\}, \tag{26}$$

$$\Omega=\left\{(k_1,k_2,\ldots,k_d):-\frac{N}{2}\leqslant k_1,k_2,\ldots,k_d<\frac{N}{2},k_1,k_2,\ldots,k_d\in\mathbb{Z}\right\}. \tag{27}$$

For our purpose, N is assumed to be on the order $O(h^{-1})$, which corresponds to have a finite number of samples per wavelength for the field $U(x,t)$. Furthermore, we assume that N is an integer multiple of W and L is an even integer. The latter can be satisfied easily by replacing the definition of $L=2W$ by $L=2\lceil W\rceil$.

For a given function f defined on the Cartesian grid X, the discrete Fourier transforms are defined by:

$$\hat{f}(\xi)=\frac{1}{N^{d/2}}\sum_{x\in X}e^{-2\pi\mathrm{i}x\cdot\xi}f(x),\quad\forall\xi\in\Omega, \tag{28}$$

$$f(x) = \frac{1}{N^{d/2}} \sum_{\xi \in \Omega} e^{2\pi i x\cdot\xi} \hat{f}(\xi), \quad \forall x \in X. \tag{29}$$

We would like to point out that there are different conventions about the constant in front of the discrete Fourier transforms. The constant $1/N^{d/2}$ in the above definitions ensures that the transforms are discrete isometries.

The functions $g_i(\xi)$ and $h_i(\xi)$ 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 $\{\phi_{i,k}(x)\}$ and $\{\psi_{i,k}(x)\}$ as follows in the Fourier domain: for $\xi \in \Omega$,

$$\hat{\phi}^D_{i,k}(\xi) = \frac{1}{L^{d/2}} e^{-2\pi i \frac{k\cdot\xi}{L}} g_i(\xi), \quad k \in \{0,1,\ldots,L-1\}^d, \tag{30}$$

$$\hat{\psi}^D_{i,k}(\xi) = \frac{1}{L^{d/2}} e^{-2\pi i \frac{k\cdot\xi}{L}} h_i(\xi), \quad k \in \{0,1,\ldots,L-1\}^d. \tag{31}$$

In terms of the spatial variable $x \in X$,

$$\phi^D_{i,k}(x) = \frac{1}{(NL)^{d/2}} \sum_{\xi \in \Omega} e^{2\pi i \left(x-\frac{k}{L}\right)\cdot\xi} g_i(\xi), \quad k \in \{0,1,\ldots,L-1\}^d, \tag{32}$$

$$\psi^D_{i,k}(x) = \frac{1}{(NL)^{d/2}} \sum_{\xi \in \Omega} e^{2\pi i \left(x-\frac{k}{L}\right)\cdot\xi} h_i(\xi), \quad k \in \{0,1,\ldots,L-1\}^d. \tag{33}$$

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

$$\phi^D_{i,k}(x) \approx \left(\sqrt{\frac{\pi}{NL}}\sigma\right)^d \cdot e^{2\pi i \left(x-\frac{k}{L}\right)\cdot\xi_i} \cdot e^{-\sigma^2\pi^2\left|x-\frac{k}{L}\right|^2}, \quad \forall x \in X.$$

Here the subtraction in the spatial domain is understood modulus the periodic unit interval.

Based on the above setup, the discrete version of the forward Gaussian wavepacket transform for an input signal $\{f(x), x \in X\}$ is given by

$$c^D_{i,k} = \left\langle \psi^D_{i,k}, f \right\rangle = \left\langle \hat{\psi}^D_{i,k}, \hat{f} \right\rangle = \sum_{\xi \in \Omega} \frac{1}{L^{d/2}} e^{2\pi i \frac{k\cdot\xi}{L}} h_i(\xi) \hat{f}(\xi). \tag{34}$$

As $h_i(\xi)\hat{f}(\xi)$ is supported in a subgrid of size L in each dimension, the summation (34) is in fact a d-dimensional inverse Fourier transform of size L in each dimension by following the "wrapping" idea [4,6]. More precisely, we introduce a new function $t(\eta)$ for $\eta \in \{-L/2,\ldots,L/2-1\}^d$ by defining

$$t(\eta) = h_i(\xi)\hat{f}(\xi)$$

with ξ in the support of $h_i(\xi)\hat{f}(\xi)$ and $\xi \equiv \eta \bmod L$. Computationally, the definition of $t(\eta)$ is equivalent to "wrapping" the function $h_i(\xi)\hat{f}(\xi)$ with period L towards the center of the Fourier domain. Using the definition of $t(\eta)$, (34) becomes

$$c^D_{i,k} = \sum_{\eta} \frac{1}{L^{d/2}} e^{2\pi i \frac{k\cdot\eta}{L}} t(\eta),$$

which is clearly a d-dimensional Fourier transform of size L in each dimension. Based on this observation, the algorithm for computing the forward transform consists of the following steps:

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

1. Compute $\hat{f}(\xi)$ for $\xi \in \Omega$ using a d-dimensional forward FFT of size N in each dimension.
2. For each box B_i, form $h_i(\xi)\hat{f}(\xi)$ at the support of $h_i(\xi)$, wrap the result modulus L to the domain $[-L/2,L/2)^d$ to get $t(\eta)$, and apply a d-dimensional inverse FFT of size L in each dimension to $t(\eta)$ to obtain $c^D_{i,k}$ for all k.

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

Given a set of coefficients $\{c^D_{i,k}\}$, the discrete version of the inverse Gaussian wavepacket transform is defined by:

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

or equivalently in the Fourier domain,

$$\hat{f}(\xi) = \sum_{i,k} c^D_{i,k} \hat{\phi}^D_{i,k}(\xi) = \sum_i \left(\sum_k \frac{1}{L^{d/2}} e^{-2\pi i \frac{k\cdot\xi}{L}} c^D_{i,k} \right) g_i(\xi).$$

The inner summation in k of the last formula is a forward Fourier transform of size L from k to ξ, followed by "unwrapping" the result to the support of $g_i(\xi)$.

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

1. For each box B_i, apply a d-dimensional forward FFT of size L in each dimension to the coefficients $c^D_{i,k}$, unwrap the result modulus L to the support of $g_i(\xi)$, multiply the unwrapped data with $g_i(\xi)$, 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 $O(N^d \log N)$.

3.3. Initialization based on Gaussian wavepacket transforms

With the Gaussian wavepacket transforms at our disposal, we discuss how to initialize the Gaussian beam representation for a general initial condition $U_0(x)$.

We first apply the forward Gaussian wavepacket transform to the initial condition $U_0(x)$ to compute the coefficients $\{c_{i,k}\}$ defined by $c_{i,k} = \langle \psi_{i,k}, U_0 \rangle$. Then from the set of $\{c_{i,k}\}$, the inverse Gaussian wavepacket transform synthesizes the function defined by

$$U_0(x) = \sum_{i,k} c_{i,k} \phi_{i,k}(x). \tag{35}$$

Recall the approximation

$$\phi_{i,k}(x) \approx \left(\sqrt{\frac{\pi}{L}}\sigma\right)^d \cdot e^{2\pi\imath\left(x-\frac{k}{L}\right)\cdot\xi_i} \cdot e^{-\sigma^2\pi^2\left|x-\frac{k}{L}\right|^2}. \tag{36}$$

For each (i,k), we introduce a Gaussian beam with an initial profile equal to the right hand side of (36):

$$\begin{aligned}
&\dot{x} = p, \quad x|_{t=0} = \frac{k}{L},\\
&\dot{p} = -V_x, \quad p|_{t=0} = 2\pi\hbar\xi_i,\\
&\dot{\tau} = \frac{|p|^2}{2} - V(x), \quad \tau|_{t=0} = 0,\\
&\dot{M} = -M^2 - V_{xx}, \quad M|_{t=0} = \imath\cdot(2\hbar\pi^2\sigma^2)I,\\
&\dot{A} = -\frac{A}{2}\mathrm{trace}(M), \quad A|_{t=0} = \left(\sqrt{\frac{\pi}{L}}\sigma\right)^d.
\end{aligned} \tag{37}$$

Denote the solutions by $x_{i,k}(t)$, $p_{i,k}(t)$, $\tau_{i,k}(t)$, $M_{i,k}(t)$, and $A_{i,k}(t)$. The Gaussian beam associated with (i,k) at time t is given by

$$\Phi_{i,k}(x,t) = A_{i,k}(x,t)\exp\left(\frac{\imath\tau_{i,k}(x,t)}{\hbar}\right),$$

where

$$\tau_{i,k}(x,t) = \tau_{i,k}(t) + p_{i,k}(t)\cdot(x - x_{i,k}(t)) + \frac{1}{2}(x - x_{i,k}(t))^T M_{i,k}(t)(x - x_{i,k}(t)), \tag{38}$$

$$A_{i,k}(x,t) = A_{i,k}(t). \tag{39}$$

We have by construction $\Phi_{i,k}(x,0) \approx \phi_{i,k}(x)$. Combining this with Eq. (35) yields

$$U_0(x) \approx \sum_{i,k} c_{i,k}\Phi_{i,k}(x,0),$$

which is a decomposition of $U_0(x)$ into a set of Gaussian beams. Therefore, the global asymptotic solution at time t is given by the summation,

$$U(x,t) \approx \sum_{i,k} c_{i,k}\Phi_{i,k}(x,t). \tag{40}$$

We notice that in the wavepacket-transform-based Gaussian beam framework (37), the initial values for bicharacteristics in the phase space define a mesh with the x-resolution $\Delta x \approx O(\sqrt{\hbar})$ and the p-resolution $\Delta p \approx O(\sqrt{\hbar})$ so that $\Delta x \Delta p \approx O(\hbar)$, which satisfies the Heisenberg uncertainty principle. Since each beam is an asymptotic solution, the summation (40) is also an asymptotic solution. Moreover, since the summation (40) satisfies the initial condition with a small error, the summation (40) is a global asymptotic solution to the initial value problem (1) and (2).

For a typical initial function $U_0(x)$, most of the coefficients $\{c_{i,k}\}$ have small norms. Therefore, in light of computational efficiency, one only needs to keep the coefficients $\{c_{i,k}\}$ such that $|c_{i,k}|$ is greater than a certain prescribed threshold η. For the rest of the coefficients, we simply set the value to zero. More specifically, for arbitrary $\eta > 0$ define the index set

$$G_\eta = \{(i,k) : |c_{i,k}| = |\langle \psi_{i,k}, U_0\rangle| > \eta\}.$$

Then the solution $U(x,t)$ can be approximated by

$$U_\eta(x,t) \equiv \sum_{(i,k)\in G_\eta} c_{i,k}\Phi_{i,k}(x,t).$$

Equipped with the fast transforms presented above, we now discuss how to implement the Gaussian beam propagation of a general initial condition $U_0(x)$ in the discrete framework. To simplify the discussion, we assume that the domain is the unit torus $[0,1]^d$, $U_0(x)$ is periodic, and $U_0(x)$ is sampled with the uniform Cartesian grid X. The number of samples N in each direction is proportional to $\frac{1}{\hbar}$ as one uses a finite number of samples per unit wavelength.

Algorithm 3.5 (*Gaussian beam propagation*).

1. Apply the discrete forward transform (Algorithm 3.3) to $U_0(x)$ to compute the coefficients $\left\{c^D_{i,k}\right\}$. Define the index set of non-negligible coefficients G_η by

$$G_\eta = \left\{(i,k) : \left|c^D_{i,k}\right| > \eta\right\}.$$

2. From the above discussion, we know that

$$\phi^D_{i,k}(x) \approx \left(\sqrt{\frac{\pi}{LN}}\sigma\right)^d \cdot e^{2\pi\imath\left(x-\frac{k}{L}\right)\cdot\xi_i} \cdot e^{-\sigma^2\pi^2\left|x-\frac{k}{L}\right|^2}, \quad \forall x \in X.$$

Therefore, for each $(i,k) \in G_\eta$, set up the following Gaussian beam equations

$$\begin{aligned}
&\dot{x} = p, \quad x|_{t=0} = \frac{k}{L},\\
&\dot{p} = -V_x, \quad p|_{t=0} = 2\pi\hbar\xi_i,\\
&\dot{\tau} = \frac{|p|^2}{2} - V(x), \quad \tau|_{t=0} = 0,\\
&\dot{M} = -M^2 - V_{xx}, \quad M|_{t=0} = \imath\cdot(2\hbar\pi^2\sigma^2)I,\\
&\dot{A} = -\frac{A}{2}\mathrm{trace}(M), \quad A|_{t=0} = \left(\sqrt{\frac{\pi}{LN}}\sigma\right)^d,
\end{aligned} \tag{41}$$

and trace it until the final time T. Denote the solutions by $x_{i,k}(t)$, $p_{i,k}(t)$, $\tau_{i,k}(t)$, $M_{i,k}(t)$, and $A_{i,k}(t)$. The Gaussian beam associated with $c^D_{i,k}$ at time T is given by

$$\Phi_{i,k}(x,T) = A_{i,k}(T)\exp\left(\frac{\imath\tau_{i,k}(x,T)}{\hbar}\right)$$

with $\tau_{i,k}(x,T) = \tau_{i,k}(T) + p_{i,k}(T)\cdot(x - x_{i,k}(T)) + \frac{1}{2}(x - x_{i,k}(T))^T M_{i,k}(T)(x - x_{i,k}(T))$. Notice that, since the domain is assumed to be the unit torus, the computation of $x_{i,k}(t)$ is performed modulus one.

3. Perform the final summation. For each $x \in X$, set

$$U_\eta(x,T) = \sum_{(i,k)\in G_\eta} c^D_{i,k}\Phi_{i,k}(x,T).$$

Besides the periodic case mentioned above, another important setting is the free space problem with compactly supported initial data. Without loss of generality, assume that the domain of interest is $[0,1]^d$ (*i.e.*, if a Gaussian beam exits from this domain, it does not come back) and that the initial condition $U_0(x)$ is compactly supported in $[O(\hbar^{1/2}), 1 - O(\hbar^{1/2})]^d$. The above algorithm can be easily adapted to address this case with two modifications. First, we periodize $U_0(x)$ to be a periodic function with fundamental domain $[0,1]^d$ and apply the (discrete) forward Gaussian wavepacket transform to it. Since $U_0(x)$ is supported $[O(\hbar^{1/2}), 1 - O(\hbar^{1/2})]^d$ and each Gaussian wavepacket has an $O(\hbar^{1/2})$ essential support, the artifacts from periodization is negligible. Second, in the solution of the Gaussian beam equations (41), whenever the location $x_{i,k}(t)$ of the (i,k)th beam is about $O(\hbar^{1/2})$ distance away from the fundamental domain $[0,1]^d$, we simply discard this beam and ignore it in the final summation.

Let us estimate the computational cost of this algorithm. The first step of Algorithm 3.5 clearly takes $O(N^d \log N) = O(\hbar^{-d}\log(\hbar^{-1}))$ steps. In the second step, since the Gaussian beam equations themselves do not contain the small parameter $\hbar$, integrating each beam over a finite time period takes $O(1)$ steps. Therefore, the cost of step 2 is proportional to the

cardinality of G_η. In the third step, the support of each Gaussian beam is of size $O(\hbar^{1/2})$ on average in each dimension. As a result, each beam at time T covers $O(N\cdot\hbar^{1/2}) = O(\hbar^{-1/2})$ points in each direction. Therefore, the overall cost of step 3 is $O(|G_\eta|\cdot\hbar^{-d/2})$. Summing these estimates together shows that the overall cost of Algorithm 3.5 is $O(\hbar^{-d}\log(\hbar^{-1}) + |G_\eta|\cdot\hbar^{-d/2})$. Compared with the $O(\hbar^{-(d+1)})$ of the standard finite difference algorithm, it is clear that Algorithm 3.5 is more efficient when the number of beams $|G_\eta|$ is small.

We remark that in a recent work [3] on Gaussian beams for the wave equation, Bougacha et al. have assumed that the FBI transforms of the initial data are infinitely small on the complement of some ring so that asymptotically Gaussian beams emanating from the complement of this ring are negligible. Moreover, the FBI transform of an L^2-function is uniformly locally infinitely small outside its frequency set as $\hbar$ going to 0 (see [17, p. 98]); consequently, asymptotically Gaussian beams which emanate from the complement of the frequency set of the initial data are negligible as well. Therefore, to apply the methodology developed here, we may assume that the frequency set of the initial data for the Schrödinger equation is compact so that we may further justify when the number of the Gaussian beams $|G_\eta|$ is small. We will report on this in a future paper.

4. Long time propagation

4.1. Gaussian beam width and the Hessian of a potential

We consider the Schrödinger equation to illustrate how the beam width is related to the Hessian of a potential function. As constructed above, a single Gaussian beam centered at a ray trajectory $\{(x(t),t): t \geqslant 0\}$ propagates according to a quadratic approximation to the potential function defined by a second-order Taylor expansion centered at $x = x(t)$ for $t \geqslant 0$:

$$V(x) = V(x(t)) + V_x(x(t))(x-x(t)) + \frac{1}{2}(x-x(t))^T V_{xx}(x(t))(x-x(t)).$$

The properties of $V_{xx}(x(t))$ control how the beam width grows.

To simplify the presentation, we assume that the potential function is given as a quadratic function:

$$V(x) = V(x_0) + V_x(x_0)(x-x_0) + \frac{1}{2}(x-x_0)^t V_{xx}(x_0)(x-x_0),$$

where x_0 is a given point. Then for a beam that starts at x_0 with momentum p_0, the Riccati equation for the Hessian of the phase function can be written as:

$$\dot{M} = -M^2 - V_{xx}(x_0), \quad M|_{t=0} = \imath\epsilon I, \tag{42}$$

where ϵ is a given positive number, and I is the $d\times d$ identity matrix.

We consider the one-dimensional case first and discuss the following three cases to illustrate the situation.

Case 1. $V_{xx}(x_0)\equiv 0$. Solving the resulting Riccati equation for M yields

$$M(t) = \frac{t\epsilon^2 + \imath\epsilon}{1+t^2\epsilon^2}.$$

Thus the beam decays like

$$\exp\left(-\frac{1}{2\hbar}\mathrm{Im}(M(t))(x-x(t))^2\right) = \exp\left(-\frac{1}{2\hbar}\frac{\epsilon}{1+t^2\epsilon^2}(x-x(t))^2\right),$$

and the beam width is roughly proportional to $\sqrt{\frac{\hbar(1+t^2\epsilon^2)}{\epsilon}}$, which grows almost linearly as time t increases.

Case 2. $V_{xx}(x_0) > 0$. To solve the resulting Riccati equation, we use the variational system for B and C,

$$\dot{B} = -V_{xx}(x_0)C, \quad B|_{t=0} = \imath\epsilon,$$
$$\dot{C} = B, \quad C|_{t=0} = 1.$$

This yields that

$$C(t) = c_1 e^{\imath\omega t} + c_2 e^{-\imath\omega t},$$
$$B(t) = \imath\omega(c_1 e^{\imath\omega t} - c_2 e^{-\imath\omega t}),$$

where $\omega = \sqrt{V_{xx}(x_0)}$, $c_1 = \frac{1}{2}(1+\frac{\epsilon}{\omega})$, and $c_2 = \frac{1}{2}(1-\frac{\epsilon}{\omega})$.

As a result,

$$M(t) = B(t)C^{-1}(t) = \frac{\omega(\epsilon^2-\omega^2)\sin\omega t\cos\omega t + \imath\epsilon\omega^2}{\omega^2\cos^2\omega t + \epsilon^2\sin^2\omega t}.$$

Thus the beam decays like

$$\exp\left(-\frac{1}{2\hbar}\mathrm{Im}(M(t))(x-x(t))^2\right)=\exp\left(-\frac{\epsilon\omega^2}{2\hbar(\omega^2\cos^2\omega t+\epsilon^2\sin^2\omega t)}(x-x(t))^2\right),$$

and the beam width is roughly proportional to

$$\sqrt{\frac{\hbar(\omega^2\cos^2\omega t+\epsilon^2\sin^2\omega t)}{\epsilon\omega^2}},$$

which is bounded from above by $\sqrt{\frac{\hbar}{\epsilon}}$ if $\epsilon\leqslant\omega$ or by $\sqrt{\frac{\hbar\epsilon}{\omega^2}}$ if $\epsilon\geqslant\omega$. In particular, when $\epsilon=\omega$, the beam width does not change.

Case 3. $V_{xx}(x_0)<0$. Solving the resulting variational system for B and C yields that

$$C(t)=c_1e^{\omega t}+c_2e^{-\omega t},$$
$$B(t)=\omega(c_1e^{\omega t}-c_2e^{-\omega t}),$$

where $\omega=\sqrt{-V_{xx}(x_0)}$, $c_1=\frac{1}{2}\left(1+\frac{\imath\epsilon}{\omega}\right)$, and $c_2=\frac{1}{2}\left(1-\frac{\imath\epsilon}{\omega}\right)$.

Consequently,

$$M(t)=B(t)C^{-1}(t)=\frac{\omega\left(1+\frac{\epsilon^2}{\omega^2}\right)(1-e^{-4\omega t})+\imath4\epsilon e^{-2\omega t}}{(1+e^{-2\omega t})^2+\frac{\epsilon^2}{\omega^2}(1-e^{-2\omega t})^2}.$$

Thus the beam decays very slowly like

$$\exp\left(-\frac{1}{2\hbar}\mathrm{Im}(M(t))(x-x(t))^2\right)=\exp\left(-\frac{4\epsilon e^{-2\omega t}}{2\hbar\left[(1+e^{-2\omega t})^2+\frac{\epsilon^2}{\omega^2}(1-e^{-2\omega t})^2\right]}(x-x(t))^2\right),$$

and the beam width is roughly proportional to

$$e^{\omega t}\sqrt{\frac{\hbar[\omega^2(1+e^{-2\omega t})^2+\epsilon^2(1-e^{-2\omega t})^2]}{4\epsilon\omega^2}},$$

which grows exponentially as t increases. 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. In practice, along a ray trajectory the Hessian of a potential function may change from positive definite to negative definite, so that the beam width may grow exponentially unexpectedly.

Next we consider the multi-dimensional case. Since the $d\times d$ Hessian $V_{xx}(x_0)$ is a symmetric matrix, it has d real eigenvalues λ_i $(i=1,\ldots,d)$ and d corresponding linearly independent real eigenvectors r_i $(i=1,\ldots,d)$. Denoting matrix $R=(r_1|\cdots|r_d)$, we can solve the Riccati equation to conclude that the second-order derivative $M(t)$ is similar to a diagonal matrix:

$$M(t)=RD(t,\lambda_1,\ldots,\lambda_d)R^{-1},$$

where $D(t,\lambda_1,\ldots,\lambda_d)=\mathrm{diag}(d_1(t,\lambda_1),\ldots,d_d(t,\lambda_d))$, each diagonal entry $d_i=d_i(t,\lambda_i)$ $(i=1,\ldots,d)$ being a function of only one of the eigenvalues. Letting $\omega=\sqrt{\lambda_i}$ $(i=1,\ldots,d)$, each $d_i=d_i(t,\lambda_i)$ takes one of the forms of the one-dimensional $M(t)$ appearing in Case 1, 2, or 3. The details are omitted.

Therefore, in the multi-dimensional case, the growth of the beam width depends on the Hessian of the potential in more complicated ways. However, some general features will be the same as in the one-dimensional case. For example, when the Hessian of the potential is negative definite, then the beam width will grow exponentially.

To the best of our knowledge, this is the **first time** that the relationship between beam widths and accuracy of the Gaussian beams has been revealed this way.

In a recent work [20], Motamed and Runborg showed that for the Helmholtz equation with constant speed of wave propagation the local beam width is not a good indicator of accuracy, and there is no direct relation between the error and the beam width in this case; moreover, they also stated without substantiation that this may not be the true for the Helmholtz equation with varying speed of wave propagation, where the beam width can be an important factor in the Taylor expansion error. According to our analysis above for the Schrödinger equation with varying potentials, the beam width is an important factor in the Taylor expansion error and is directly related to the accuracy of Gaussian beams.

As illustrated above, beam width is related to both the Planck constant $\hbar$ and the imaginary part of the Hessian of the phase function. To further discuss the relationship between beam width and accuracy of Gaussian beams, we consider two cases: (a) $\hbar$ being arbitrarily small and (b) $\hbar$ being fixed.

Case (a): $\hbar$ being arbitrarily small. If $\hbar$ can be chosen to be arbitrarily small, then as shown above the beam width can be made to be arbitrarily small as well, so that the beam will decay rapidly across the ray direction and the resulting local Taylor expansion will be within the beam support; because the beam support is small, the local Taylor expansion will have good accuracy. This explains why the beam accuracy can be improved to some extent if the $\hbar$ is made smaller.

Case (b): $\hbar$ being fixed. If a beam is too wide, the local Taylor expansion is not accurate, and the beam accuracy will be poor. If a cut-off function is introduced to make the beam vanish outside a small enough neighborhood around the central ray such as advocated in [25,23], then the neighborhood may be extremely small. Moreover, we know that rays may diverge even if their initial conditions are very close to each other, which implies that summing up a finite number of beams with cut-offs may leave some regions with no beams passing through; consequently, the resulting numerical global asymptotic solution is inaccurate.

4.2. Long time propagation by reinitialization

A natural solution to the issue of growing width is to monitor the widths of the beams and reinitialize the Gaussian beam representation before any one of the beams becomes too wide. Although this reinitialization idea is rather straightforward, it is difficult to combine it with previous methods of beam initialization. For the asymptotic decomposition method [25,23,13], one always assumes that the wave field has an amplitude and phase representation explicitly. Even though the initial solution $U(x,0) = U_0(x)$ has such a decomposition, the solution at later time $U(x,t)$ most likely fails to possess such a decomposition due to the development of caustics. For the method based on optimization [24], reinitialization is also impractical as the initialization algorithm is quite costly.

On the other hand, the reinitialization idea fits perfectly with the initialization algorithms proposed in this paper. The main reason is that our algorithm is highly efficient and is applicable to general initial conditions. The detailed steps are given as follows.

Algorithm 4.1 (*Gaussian beam propagation with reinitialization*).

1. Set the current time $\tilde{t} = 0$ and the current solution $\widetilde{U}(x) = U_0(x)$.
2. From $\widetilde{U}(x)$, compute the coefficients $c^D_{i,k}$ using Algorithm 3.3. Let G_η be the set of indices of significant coefficients, *i.e.*, $G_\eta = \left\{(i,k) : \left|c^D_{i,k}\right| > \eta\right\}$.
3. For each $(i,k) \in G_\eta$, solve the following equations
$$\begin{aligned}
&\dot{x} = p, \quad x|_{t=\tilde{t}} = \frac{k}{L},\\
&\dot{p} = -V_x, \quad p|_{t=\tilde{t}} = 2\pi\hbar\xi_i,\\
&\dot{\tau} = \frac{|p|^2}{2} - V(x), \quad \tau|_{t=\tilde{t}} = 0,\\
&\dot{M} = -M^2 - V_{xx}, \quad M|_{t=\tilde{t}} = \imath\cdot(2\hbar\pi^2\sigma^2)I,\\
&\dot{A} = -\frac{A}{2}\mathrm{trace}(M), \quad A|_{t=\tilde{t}} = \left(\sqrt{\frac{\pi}{LN}}\sigma\right)^d
\end{aligned} \tag{43}$$
and monitor the values of $M_{i,k}(t)$ until either t reaches the final time T or the smallest eigenvalue of any $\mathrm{Im}(M_{i,k}(t))$ drops below a prescribed threshold. Let t^* be the first time (after $\tilde{t}$) such that either one of these two conditions is satisfied.
4. Set the current time $\tilde{t} = t^*$ and the current solution
$$\widetilde{U}(x) = \sum_{(i,k)\in G} c^D_{i,k} A_{i,k}(\tilde{t}) \exp\left(\frac{\imath\tau_{i,k}(x,\tilde{t})}{\hbar}\right)$$
with $\tau_{i,k}(x,\tilde{t}) = \tau_{i,k}(\tilde{t}) + p_{i,k}(\tilde{t})\cdot(x - x_{i,k}(\tilde{t})) + \frac{1}{2}(x - x_{i,k}(\tilde{t}))^T M_{i,k}(\tilde{t})(x - x_{i,k}(\tilde{t}))$.
5. Repeat the steps 2–4 until we reach the final time T. Once the final time is reached, set $U(x,T) = \widetilde{U}(x)$.

Let us briefly comment on the cost of Algorithm 4.1. Since the ODEs in step 3 do not depend explicitly on the small parameter $\hbar$, the number of reinitialization is a constant even though it depends on the potential $V(x)$. From the discussion after Algorithm 3.5, it is clear that the cost spent between two consecutive reinitializations is proportional to $O(\hbar^{-d}\log(\hbar^{-1}) + |G_\eta|\cdot\hbar^{-d/2})$. Since the number of reinitializations is of order $O(1)$, the overall cost is proportional to $O(\hbar^{-d}\log(\hbar^{-1}) + |G_\eta|\cdot\hbar^{-d/2})$ as well.

As an example, we take the case $V_{xx} < 0$ in Section 4.1 to demonstrate the benefit of reinitialization. Suppose that $V_{xx}(x_0) = -4\pi^2$, $\omega = \sqrt{-V_{xx}(x_0)} = 2\pi$, and the final time $T = 2.0$. If we use Algorithm 3.5 to propagate beams, the beam width at the final time T is roughly
$$\frac{e^{\omega T}}{\sqrt{2\pi}}\sqrt{\hbar} = \frac{e^{4\pi}}{\sqrt{2\pi}}\sqrt{\hbar} \approx 64,164\sqrt{\hbar}.$$

If we choose $\tilde{t}$ such that $e^{\omega\tilde{t}} \approx \sqrt{2}\pi$, and we divide $[0,T]$ into $\lceil \frac{T}{\tilde{t}} \rceil$ sub-intervals, then the beam width in each subinterval is roughly $\sqrt{\hbar}$. In this particular example, $\tilde{t} \approx 0.25$, and $\left\lceil \frac{T}{\tilde{t}} \right\rceil = 8$. This simple case analysis demonstrates that it is critical to carry out reinitialization.

We remark in passing that beam reinitialization may help to offset the effect of the errors in beam construction, which will make Gaussian beam valid for 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.

5. Wavepacket-transform-based global asymptotic solutions

Following the idea in [22,25], we prove that the wavepacket-transform-based Gaussian beam solution (40) is an asymptotic solution to the evolution equation (1) in the following sense.

Theorem 5.1. *Assume that the conditions for Lemmas 5.2,* 5.3 *and* 5.4 *are fulfilled. For arbitrary $\eta > 0$, let the index set*

$$G_\eta = \{(i,k) : |c_{i,k}| = |\langle \psi_{i,k}, U_0\rangle| > \eta\}$$

be finite. The solution (40) is a global asymptotic solution to the evolution equation (1) in the following sense: in a finite time interval [0,T], for $\hbar$ small enough,

$$\|(-\imath\hbar\partial_t + H(x,-\imath\hbar\partial_x))U_\eta\|_{L^2_{x,t}} \leqslant C\sqrt{|G_\eta|}\hbar^{\frac{3}{2}}, \tag{44}$$

$$\lim_{\eta\to 0} U_\eta(x,0) \approx U_0(x), \tag{45}$$

where $|G_\eta|$ denotes the number of elements in the set G_η (the number of beams), the constant C is independent of $\hbar$, and the initial condition is satisfied up to a small error.

To prove this theorem, we need some lemmas.

Lemma 5.2. *The solution for the transport equation*

$$\frac{dA_{i,k}}{dt} = -\frac{A_{i,k}}{2}(\mathrm{trace}(H_{pp}M_{i,k} + H_{xp})), \quad A_{i,k}|_{t=0} = \left(\sqrt{\frac{\pi}{L}}\sigma\right)^d$$

is

$$A_{i,k}(t) = \frac{\left(\sqrt{\frac{\pi}{L}}\sigma\right)^d}{\sqrt{\det(C_{i,k}(t))}}, \tag{46}$$

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

$$\frac{dC_{i,k}}{dt} = H_{pp}M_{i,k}C_{i,k} + H_{px}C_{i,k}, \quad C_{i,k}|_{t=0} = I. \tag{47}$$

Proof. The transport equation can be written as

$$\frac{d\log A_{i,k}^2}{dt} = -(\mathrm{trace}(H_{pp}\tau_{xx} + H_{xp})). \tag{48}$$

By Lemma 2.1, $C_{i,k}(t)$ is non-singular. Let $C_{i,k}(t) = S(t)\Lambda(t)S^{-1}(t)$ be the Schur decomposition. Setting $q(t) = \det C_{i,k}(t) = \Pi_{j=1}^d \lambda_j$, where $\Lambda = diag(\lambda_1,\ldots,\lambda_d)$, and noticing that $\frac{dC_{i,k}}{dt} = S_t\Lambda S^{-1} + S\Lambda_t S^{-1} + S\Lambda(S^{-1})_t$, we have

$$\begin{aligned}\frac{dq}{dt} &= q\,\mathrm{trace}(\Lambda^{-1}\Lambda_t) = q\,\mathrm{trace}(S\Lambda^{-1}S^{-1}S\Lambda_t S^{-1} + (S\Lambda^{-1})S^{-1}S_t(S\Lambda^{-1})^{-1} + (S^{-1})_t S) = q\,\mathrm{trace}\left(C_{i,k}^{-1}\frac{dC_{i,k}}{dt}\right)\\ &= q\,\mathrm{trace}\left(C_{i,k}^{-1}(H_{pp}M_{i,k} + H_{px})C_{i,k}\right) = q\,\mathrm{trace}(H_{pp}M_{i,k} + H_{px}).\end{aligned} \tag{49}$$

Thus, by combining Eqs. (48) and (49), we have

$$A_{i,k}(t) = \frac{A_{i,k}(0)\sqrt{q(0)}}{\sqrt{q(t)}} = \frac{\left(\sqrt{\frac{\pi}{L}}\sigma\right)^d}{\sqrt{\det(C_{i,k}(t))}},$$

which implies that $A_{i,k}(t) = O(h^{-\frac{d}{4}})$ by the definition of σ and L. □

In Eqs. (38) and (39), we define the global approximations to the phase and amplitude functions by using Taylor expansions centered on the x-projection of the bicharacteristic $\{(x_{i,k}(t), p_{i,k}(t)): t \geqslant 0\}$. We have the following estimates.

Lemma 5.3. *Assume that the Hamiltonian is C^2 differentiable. Then the functions (38) and (39) satisfy the eikonal and transport equations in the following approximate sense, respectively:*

$$\frac{\partial \tau_{i,k}}{\partial t}(x,t) + H\left(x, \frac{\partial \tau_{i,k}}{\partial x}(x,t)\right) = O(|x - x_{i,k}(t)|^3), \tag{50}$$

$$A_{i,k,t}(x,t) + H_p \cdot A_{i,k,x} + \frac{A_{i,k}}{2}\left(\mathrm{trace}(H_{pp}\tau_{i,k,xx} + H_{xp})\right) = O\left(h^{-\frac{d}{4}}|x - x_{i,k}(t)|\right). \tag{51}$$

Proof. We first prove (50). For the sake of clarity we suppress the index (i,k) in the following. By the construction of $\tau_{i,k}(x,t) = \tau(x,t)$, we have

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

$$\frac{\partial \tau}{\partial x}(x,t) = p(t) + M(t)(x - x(t)).$$

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

$$\tau_t(x,t) + H(x,\tau_x(x,t)) = \tau_t(x,t) + H(x - x(t) + x(t), p(t)) + M(t)(x - x(t)) = O(|x - x(t)|^3). \tag{52}$$

Next we prove (51). By the analytical formula (46) and using the notation as in Lemma 5.2 we have

$$A_t(x,t) = A(0)\left(q(t)^{-\frac{1}{2}}\right)_t = -\frac{1}{2}A(0)q(t)^{-\frac{3}{2}}\frac{dq}{dt} = -\frac{1}{2}A(0)q(t)^{-\frac{1}{2}}\mathrm{trace}(H_{pp}M + H_{px}) = -\frac{A}{2}\mathrm{trace}(H_{pp}M + H_{px}), \tag{53}$$

$$A_x(x,t) = O\left(h^{-\frac{d}{4}}|x - x(t)|\right). \tag{54}$$

Thus, near the ray trajectory $x_{i,k}(t)$, the estimate (51) holds. □

We also need the following lemma which is proved in [22,25].

Lemma 5.4. *Assume that $c(x,t)$ vanishes to order $S-1$ on $\gamma = \{(x(t),t): 0 \leqslant t \leqslant T\}$ which is the x-projection of a bicharacteristic $\{(x(t),p(t)): 0 \leqslant t \leqslant T\}$, $\mathrm{supp}(c) \cap \{(x,t) : x \in \mathbb{R}^d,\ 0 \leqslant t \leqslant T\}$ is compact, and $\mathrm{Im}(\phi(x,t)) \geqslant a|x - x(t)|^2$ on $\mathrm{supp}(c) \cap \{(x,t) : x \in \mathbb{R}^d,\ 0 \leqslant t \leqslant T\}$, where a is a positive constant. Then*

$$\int_0^T \int_{\mathbb{R}^d} \left|c(x,t)e^{\imath\frac{\phi(x,t)}{\hbar}}\right|^2 dx\,dt \leqslant C\hbar^{S+\frac{d}{2}}, \tag{55}$$

where C is a constant independent of $\hbar$.

Now we are ready to prove Theorem 5.1.

Proof. It is easy to show that at $t = 0$ we recover the initial data up to a small error. Next we evaluate the following,

$$\begin{aligned}(-\imath\hbar\partial_t + H(x,-\imath\hbar\partial_x))U_\eta(x,t) &= \sum_{(i,k)\in G_\eta}(-\imath\hbar\partial_t + H(x,-\imath\hbar\partial_x))c_{i,k}\Phi_{i,k}(x,t)\\ &= \sum_{(i,k)\in G_\eta}(\tau_{i,k,t} + H(x,\tau_{i,k,x}))c_{i,k}A_{i,k}(x,t)\exp\left(\imath\frac{\tau_{i,k}(x,t)}{\hbar}\right)\\ &\quad + \sum_{(i,k)\in G_\eta}\frac{\hbar}{i}c_{i,k}\left(A_{i,k,t} + H_p\cdot A_{i,k,x} + \frac{A_{i,k}}{2}\mathrm{trace}(H_{pp}M_{i,k}(t) + H_{xp})\right)\exp\left(\imath\frac{\tau_{i,k}(x,t)}{\hbar}\right)\\ &\equiv f_1(x,t) + f_2(x,t).\end{aligned}$$

We estimate f_1 first. By using the Cauchy–Schwartz inequality we have

$$|f_1(x,t)|^2 \leqslant \left(\sum_{(i,k)\in G_\eta}|c_{i,k}|^2\right)\left(\sum_{(i,k)\in G_\eta}\left|(\tau_{i,k,t} + H(x,\tau_{i,k,x}))A_{i,k}(x,t)\exp\left(\imath\frac{\tau_{i,k}(x,t)}{\hbar}\right)\right|^2\right).$$

We observe that $\sum_{(i,k)\in G_\eta}|c_{i,k}|^2 \leqslant C_2\|U_0\|^2$ by Lemma 3.1 and $A_{i,k}$ is of the order $O\left(h^{-\frac{d}{4}}\right)$. By using the results in Lemmas 5.3 and 5.4 with $S = 3$, we have

$$\|f_1\|_{L^2_{x,t}} \leqslant C\sqrt{|G_\eta|}\hbar^{\frac{3}{2}},$$

where C is a constant independent of $\hbar$.

By using the Cauchy–Schwartz inequality, Lemmas 3.1, 5.3 and 5.4 with $S = 1$, we can estimate f_2 similarly to obtain

$$\|f_2\|_{L^2_{x,t}} \leqslant C\sqrt{|G_\eta|}\hbar^{\frac{3}{2}},$$

where C is a constant independent of $\hbar$. Finally, applying the triangle inequality yields the desired estimate. □

6. Numerical results

In this section, we show one-, two-, and three-dimensional numerical examples to demonstrate the performance of our algorithm. Because an exact wave function for the Schrödinger equation is not available in general, we solve the Schrödinger equation first using an accurate direct numerical method and use the result as the "exact" solution to calibrate the beam solution. Here we adopt a Strang splitting pseudo-spectral method [21,2] to compute the "exact" solutions. We also assume that the solution is periodic so that the Strang splitting pseudo-spectral method gives rise to an accurate "exact" solution.

In the following examples, we take the cut-off parameter $\eta = 10^{-2}$. In practice, because of redundancy in beam representation one may take even larger cut-off parameters and still obtain reasonable results.

6.1. One-dimensional cases

6.1.1. Example 1: a potential well

The background potential is given by $V(x) = \cos(2\pi x)$ for $x \in [0,1]$, and the initial wave function is given by

$$\begin{aligned} U(x,t=0) &= \exp(-25(x-0.5)^2)\exp\left(\frac{\imath\tau_0(x)}{\hbar}\right) \quad \text{for } x \in [0,1], \\ \tau_0(x) &= -\frac{1}{5}\ln(\exp(5(x-0.5)) + \exp(-5(x-0.5))). \end{aligned} \tag{56}$$

The periodic boundary condition is assumed in the interval $[0,1]$.

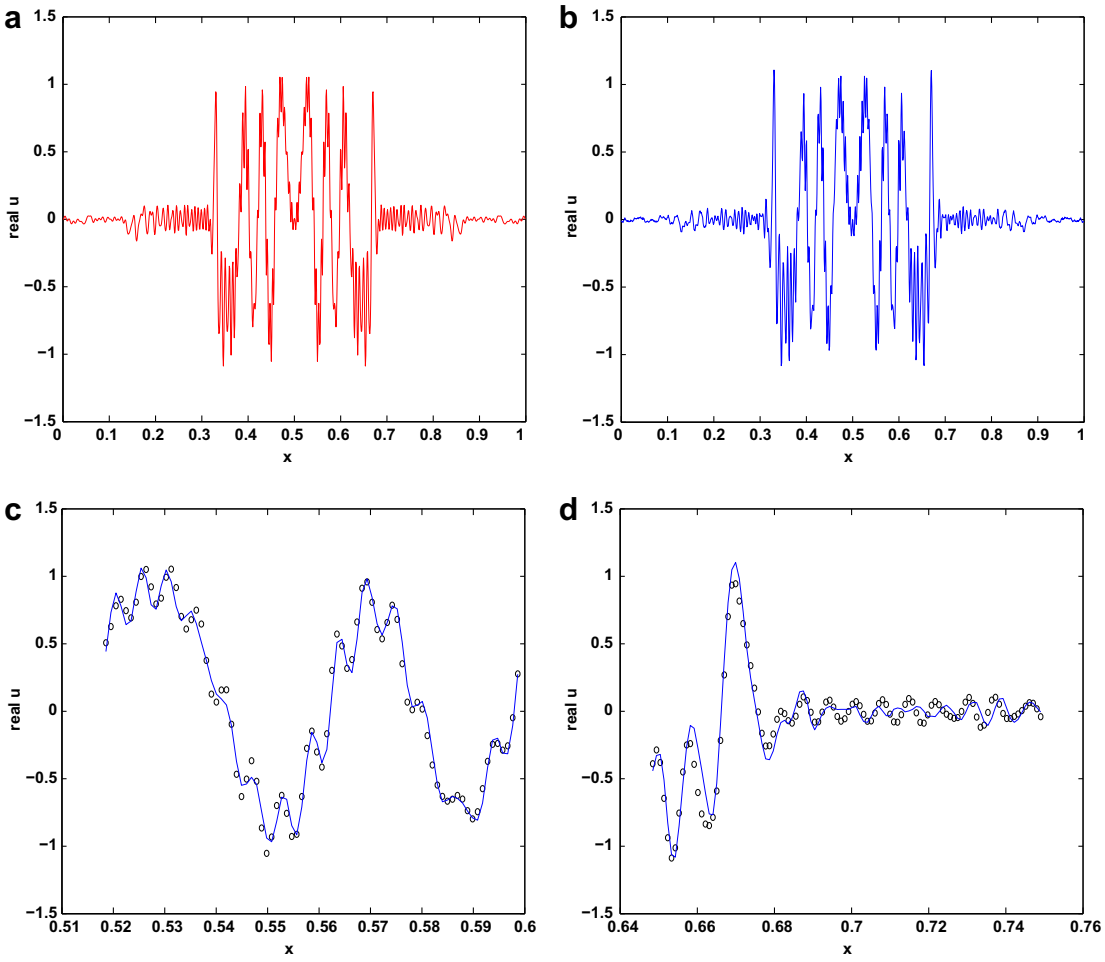

Fig. 3. Example 1: A potential well. (a) The exact solution. (b) The beam solution. (c) and (d) Comparison of the exact and beam solutions. "○": the exact solution; "-": the beam solution.

In this case, $V_{xx}(x) = -4\pi^2\cos(2\pi x)$, which is non-negative for $\frac{1}{4} \leqslant x \leqslant \frac{3}{4}$, resulting in a potential well in the interval $[0.25, 0.75]$. The initial data can be viewed as a wavepacket compactly supported in $[0.25, 0.75]$ and centered at $x = 0.5$. Therefore, according to the analysis provided in Section 4, Gaussian beam widths tend to be uniformly bounded in a given time interval, and this is observed in the numerical results.

Fig. 3 shows a case with $h = \frac{1}{256\pi}$ and the final time $T = 2.0$. We use $N = 1024$ uniformly distributed points to discretize the domain $[0, 1]$. The fast Gaussian wavepacket transforms are used to generate the initial beam decomposition and this decomposition is used to drive Gaussian beam propagation. Fig. 3(a) and (b) shows the exact solution and the beam solution at the final time $T = 2.0$, respectively, and Fig. 3(c) and (d) compares the two by plotting the solutions in different zoom-in windows. As we can see, the beam solution and the exact solution match extremely well.

6.1.2. Example 2: a potential hill

The background potential is given by $V(x) = \cos(2\pi(x + 0.5))$, and the initial wave function is the same as that in (56).

In this case, $V_{xx}(x) = -4\pi^2\cos(2\pi(x + 0.5))$, which is non-positive for $\frac{1}{4} \leqslant x \leqslant \frac{3}{4}$, resulting in a potential hill in the interval $[0.25, 0.75]$. The initial data is again a wavepacket compactly supported in $[0.25, 0.75]$ and centered at $x = 0.5$. According to the analysis provided in Section 4, Gaussian beam widths tend to grow exponentially in time, resulting in difficulties in beam summation.

Fig. 4 shows a case with $h = \frac{1}{256\pi}$ and the final time $T = 2.0$ without reinitialization of beam propagation. We again sample the initial condition with $N = 1024$ uniformly distributed points to discretize $[0, 1]$. Fig. 4(a) and (b) shows the exact solution and the beam solution at $T = 2.0$, respectively, and the two solutions differ tremendously. To see the difference more clearly, Fig. 4(c) and (d) compares the two by plotting the solutions in different zoom-in windows. As we can see, the beam solution and the exact solution do not match well.

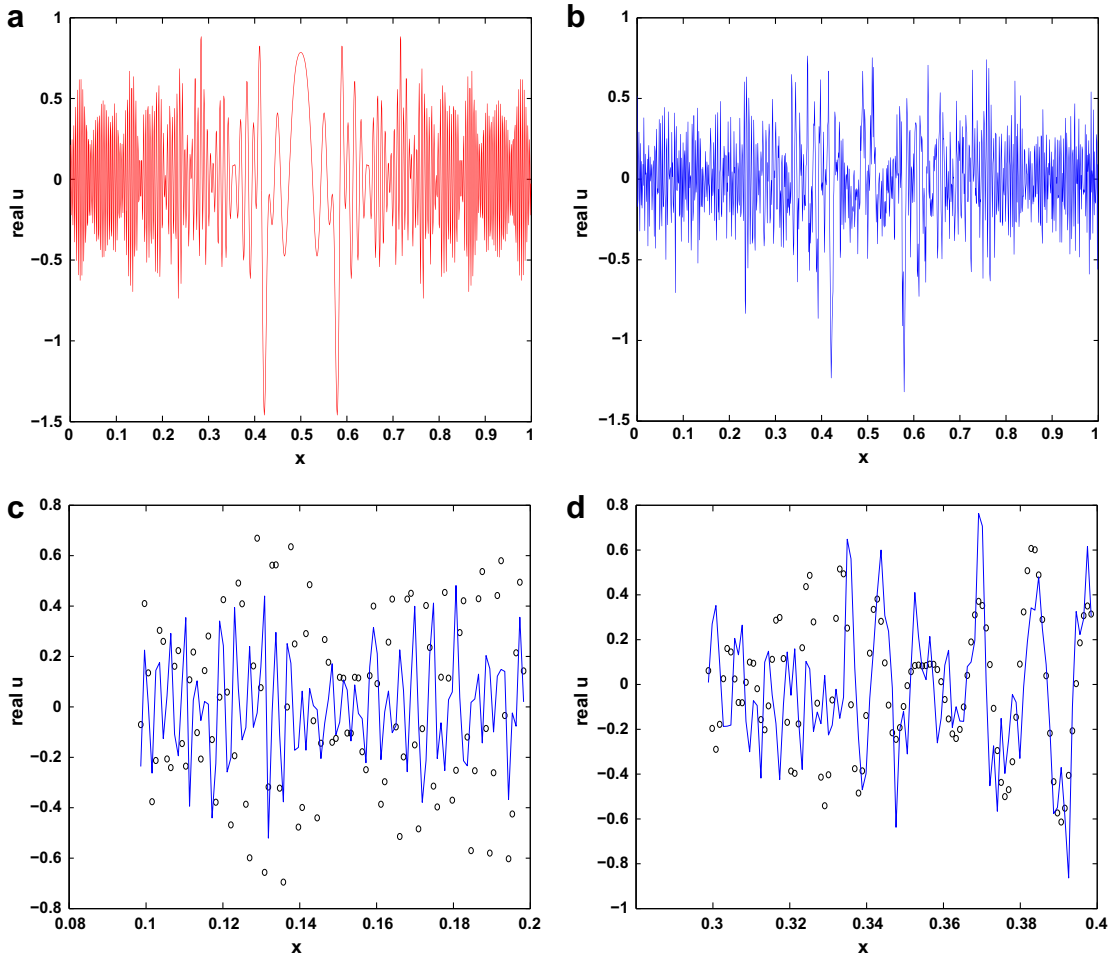

Fig. 4. Example 2: A potential hill, $h = \frac{1}{256\pi}$, without reinitialization. (a) The exact solution. (b) The beam solution. (c) and (d) Comparison of the exact and beam solutions. "∘": the exact solution; "-": the beam solution.

On the other hand, if reinitialization is used, the beam summation results improve dramatically. Fig. 5 shows the same problem with reinitialization of beam propagation. We divide the interval $[0,T]$ into eight uniform sub-intervals and reinitialize at the beginning of each sub-interval. Fig. 5(a) and (b) shows the exact solution and the beam solution at the final time $T = 2.0$, respectively, and the two solutions look very close. To compare the two solutions more closely, Fig. 5(c) and (d) compares the two by plotting the solutions in different zoom-in windows. As we can see, the beam solution and the exact solution match with each other extremely well.

To further illustrate the power of reinitialization, we test an example with $h = \frac{1}{2^{14}\pi}$ and $T = 1.0$. To well sample the initial data, we choose $N = 2^{17}$. We divide the interval $[0,T] = [0,1]$ into eight uniform sub-intervals and reinitialize at the beginning of each sub-interval. Fig. 6 shows the computational results. Fig. 6(a) and (b) plots the exact solution and the beam solution at the final time $T = 1.0$, respectively, and the two solutions look extremely similar. To compare the two solutions more closely, Fig. 6(c) shows that the two solutions match with each other extremely well in a selected window, and Fig. 6(d) shows that the modulus-squared (the position density) of the two solutions match with each other as well. In this case, the relative L^2-errors in the beam solution are 0.47% at $T = 0$ and 3.7% at $T = 2.0$, respectively. We also remark that without reinitialization the computed beam solution is simply unacceptable.

6.1.3. Example 3: a potential with hill and well

The background potential is given by $V(x) = 10 + \sin(2\pi x)$, and the initial wave function is the same as that in (56).

In this case, by the potential, $V_{xx}(x) = -4\pi^2\sin(2\pi x)$, which is non-positive for $0 \leqslant x \leqslant 0.5$ and non-negative for $0.5 \leqslant x \leqslant 1.0$, resulting in a potential hill in the interval $[0,0.5]$ and a potential well in the interval $[0.5,1.0]$. Since the initial data can be viewed as a wavepacket compactly supported in $[0.25,0.75]$ and centered at $x = 0.5$, and the resulting Gaussian beams will change their widths according to the locations of corresponding ray trajectories; in particular, some beam widths may grow exponentially as analyzed in Section 4, which leads to difficulties in beam summation as shown in numerical examples.

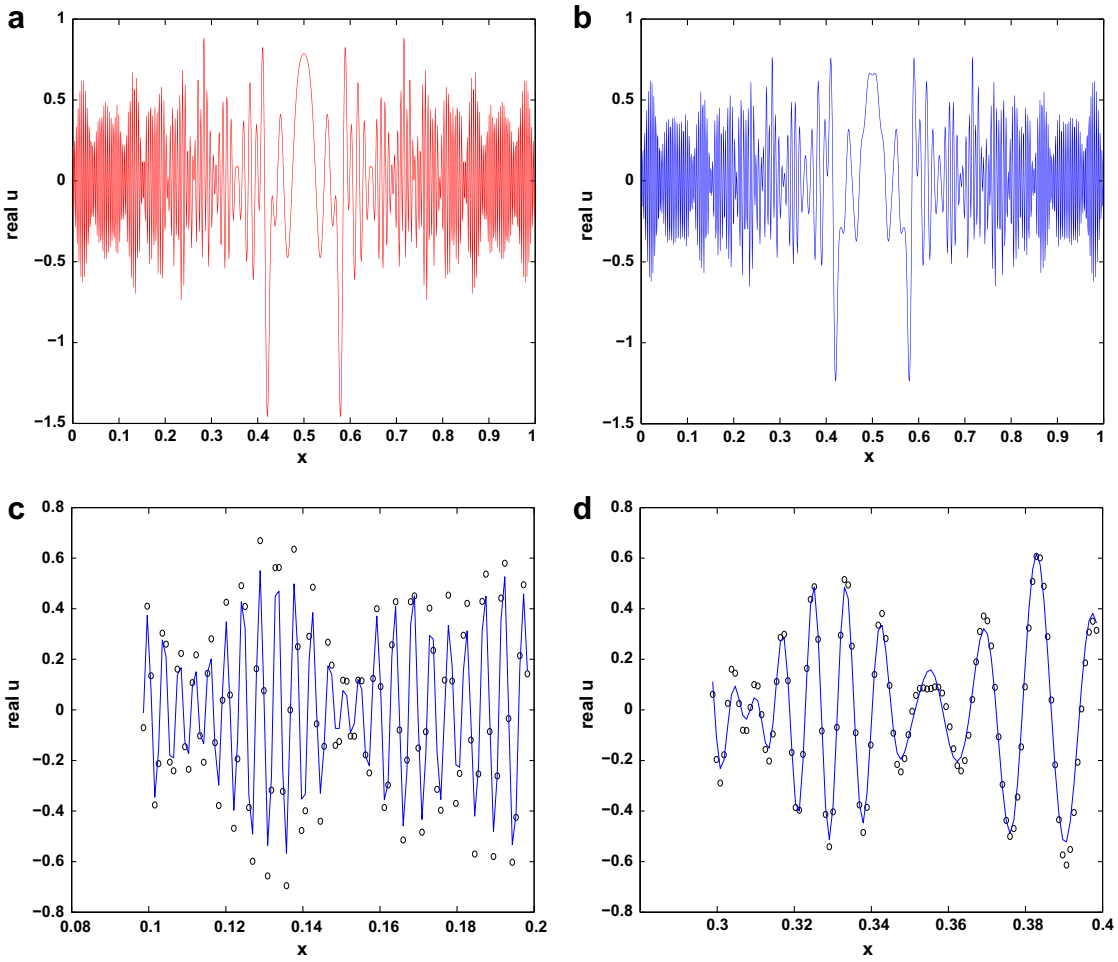

Fig. 5. Example 2. A potential hill, $h = \frac{1}{256\pi}$, with reinitialization. (a) The exact solution. (b) The beam solution. (c) and (d) Comparison of the exact and beam solutions. "∘": the exact solution; "-": the beam solution.

We set up numerical experiments by choosing $\hbar = \frac{1}{256\pi}$, $N = 1024$, and the final time $T = 2.0$. We carry out the beam propagation both with and without reinitialization. Fig. 7 shows the case without reinitialization. Fig. 7(a) and (b) shows the exact solution and the beam solution at the final time $T = 2.0$, respectively, and the two solutions differ tremendously. To see the difference more clearly, Fig. 7(c) and (d) compares the two by plotting the solutions in different zoom-in windows. As we can see, the beam solution and the exact solution do not match well.

When the reinitialization is used, the accuracy of the Gaussian beam method improves dramatically. Fig. 8 shows a case with $\hbar = \frac{1}{256\pi}$ and the final time $T = 2.0$ with reinitialization of beam propagation. We divide the interval $[0,T]$ into eight uniform sub-intervals, and reinitialize the representation at the beginning of each sub-interval. Fig. 8(a) and (b) shows the exact solution and the beam solution at the final time $T = 2.0$, respectively, and the two solutions look extremely similar. To compare the two solutions more closely, Fig. 8(c) and (d) compares the two by plotting the solutions in different zoom-in windows. As we can see, the beam solution and the exact solution match with each other extremely well.

6.2. Two-dimensional cases

6.2.1. Example 4: a two-dimensional additive potential

The background potential is given by $V(x,y) = 10 + \cos(2\pi x) + 2\cos(2\pi y)$ on a unit square, and the initial wave function is given by

$$\begin{aligned} U(x,y,t=0) &= \exp(-25(x-0.5)^2)\exp\left(\frac{i\tau_0(x)}{\hbar}\right), \quad (x,y) \in [0,1]\times[0,1], \\ \tau_0(x) &= -\frac{1}{5}\ln(\exp(5(x-0.5)) + \exp(-5(x-0.5))). \end{aligned} \tag{57}$$

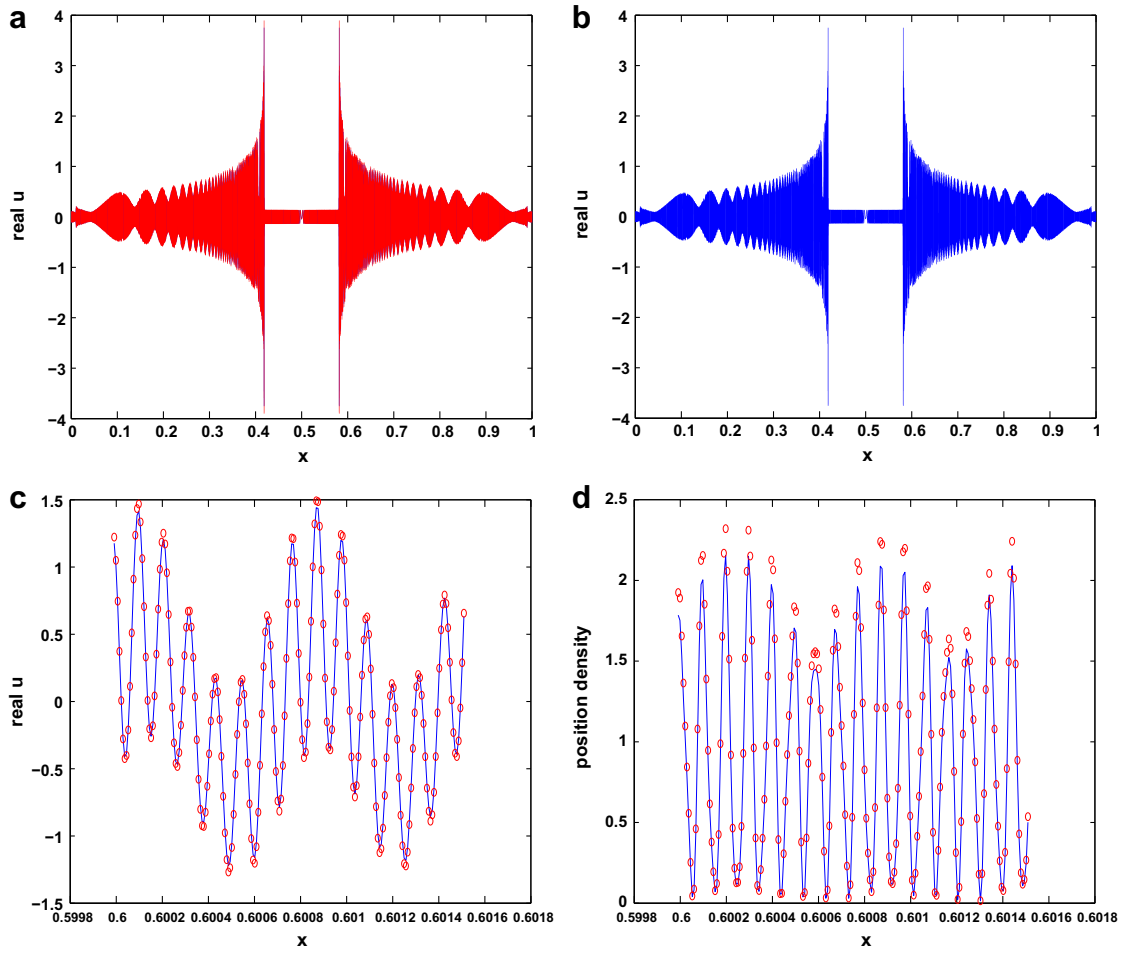

Fig. 6. Example 2. A potential hill, $\hbar = \frac{1}{2^{14}\pi}$ and $N = 2^{17}$, with reinitialization. (a) The exact solution. (b) The beam solution. (c) and (d) Comparison of the exact and beam solutions. "∘": the exact solution; "-": the beam solution.

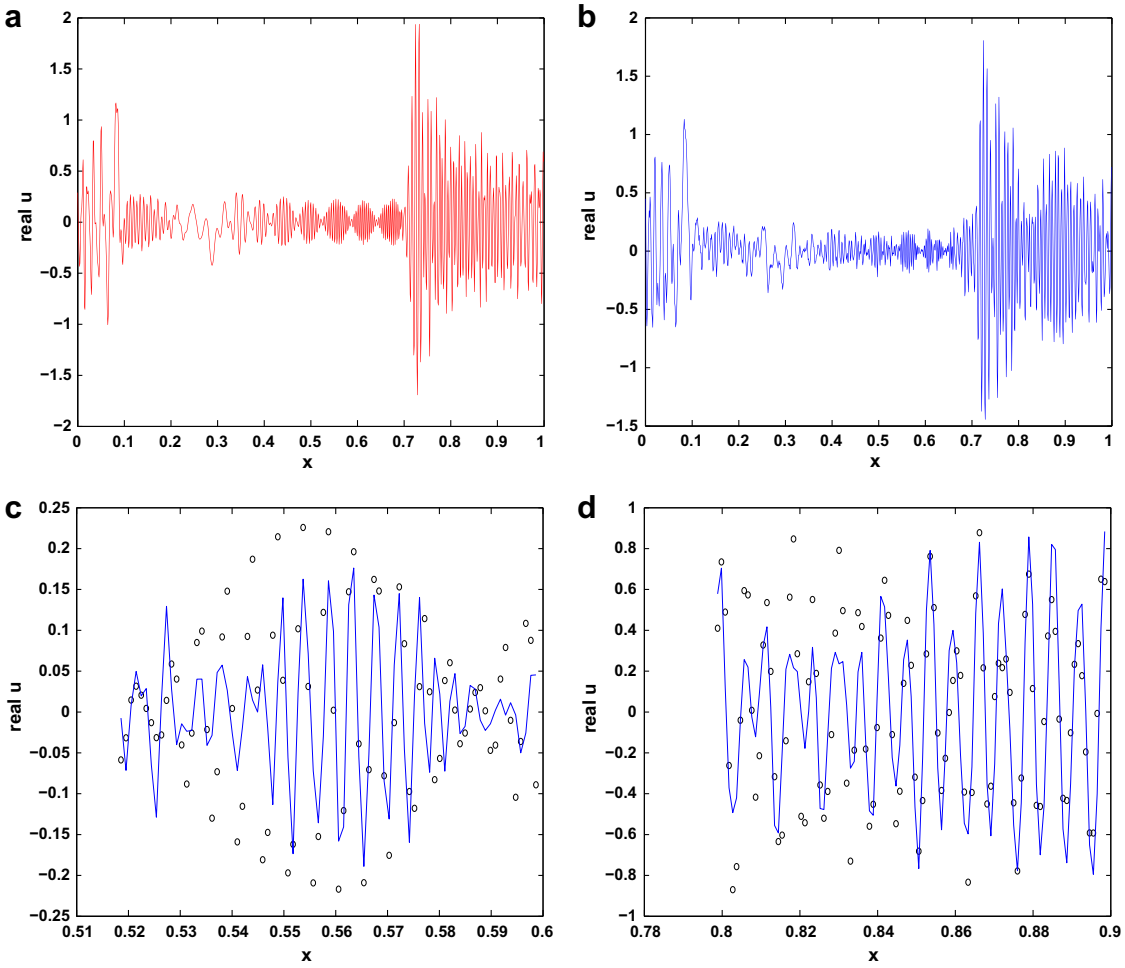

Fig. 7. Example 3. A potential with both hill and well, $\hbar = \frac{1}{256\pi}$, without reinitialization. (a) The exact solution. (b) The beam solution. (c) and (d) Comparison of the exact and beam solutions. "∘": the exact solution; "-": the beam solution.

The periodic boundary condition is imposed. In this case, the Hessian of the potential is a diagonal indefinite matrix. Thus, depending on the initial values of beams, the beam widths may grow exponentially as time evolves.

Fig. 9 shows an example with $\hbar = \frac{1}{256\pi}$ and the final time $T = 0.5$. To well resolve the initial data, we use an $N \times N = 1024 \times 1024$ uniform grid to discretize the domain $[0,1] \times [0,1]$. Accordingly, we apply the 2-D fast Gaussian wave-packet transform to generate the beam representation from the initial data. To improve the beam propagation, we subdivide $[0,T] = [0,0.5]$ into ten uniform sub-intervals and reinitialize the beam representation at the beginning of each sub-interval. Fig. 9(a) shows the scaled image of the initial data. Fig. 9(b) and (c) displays the image of the exact solution and the beam solution at the final time $T = 0.5$, and they are consistent with each other. To compare the two solutions more clearly, we plot the sliced solutions together along $x = 0.5$ (see Fig. 9(d) and (e)) and $y = 0.5$ (see Fig. 9(f) and (g)). As we can see, the two solutions match with each other very well.

6.2.2. Example 5: a two-dimensional multiplicative potential

The background potential is given by $V(x,y) = 10 + 0.5\sin(2\pi x)\cos(2\pi y)$ on a unit square, and the initial wave function is given as in (57). The periodic boundary condition is imposed. In this case, the Hessian of the potential is an indefinite matrix function. Thus, depending on the initial values of beams, beam widths may grow exponentially as time evolves.

Fig. 10 shows an example with $\hbar = \frac{1}{512\pi}$ and the final time $T = 0.56$. To well resolve the initial data, we again use an $N \times N = 1024 \times 1024$ uniform grid to discretize $[0,1] \times [0,1]$. Fig. 10(a) and (b) shows the images of the exact solution and the beam solution at the final time $T = 0.56$. To compare the two solutions more clearly, we plot the sliced solutions along $y = 0.5$ in Fig. 10(c) and (d), respectively; although the two solutions look similar, they have large difference in some regions. In Fig. 10(e) and (f), we plot the two y-sliced solutions together for two x-windows, and we can see that the two solutions indeed are not comparable in terms of resolving some small details. In Fig. 10(g), we plot the x-sliced ($x = 0.5$)

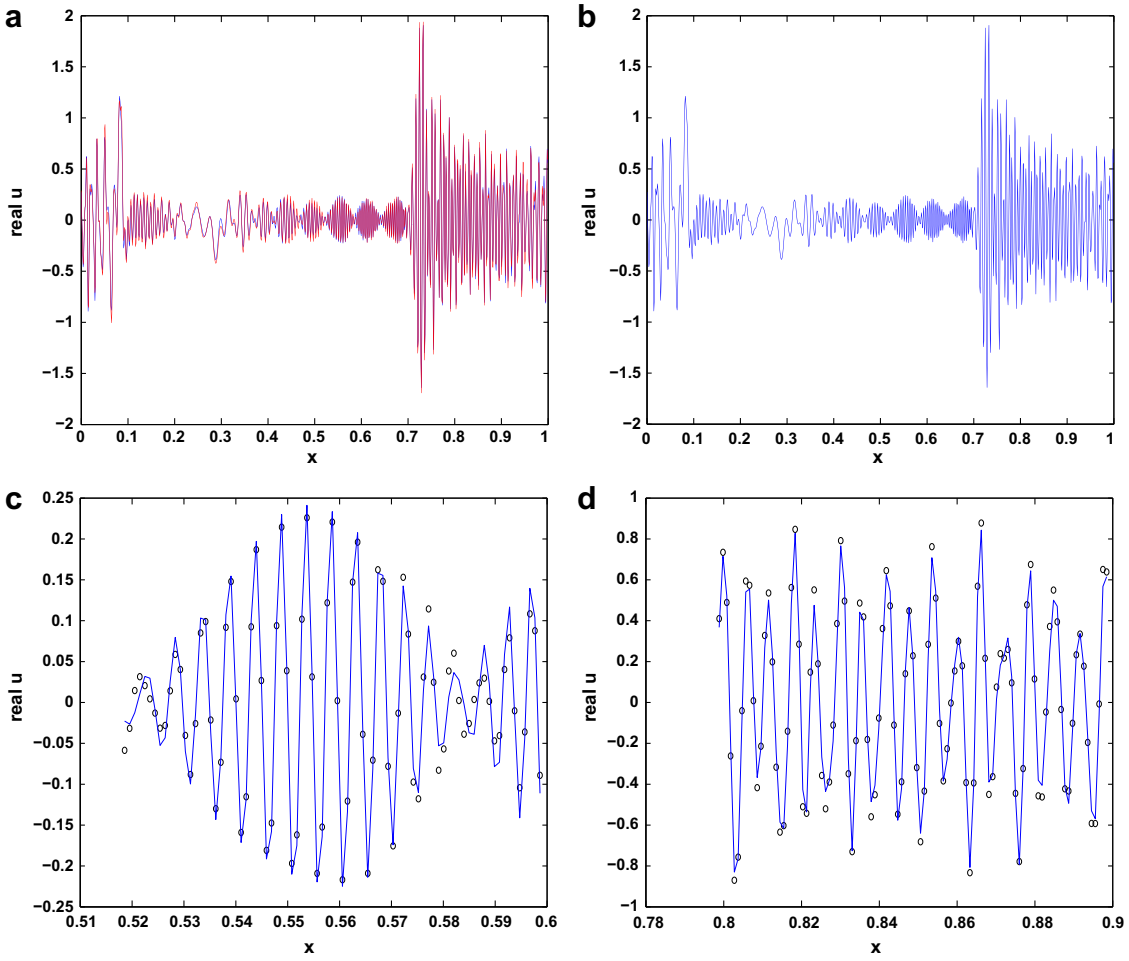

Fig. 8. Example 3. A potential with both hill and well, $\hbar = \frac{1}{256\pi}$, with reinitialization. (a) The exact solution. (b) The beam solution. (c) and (d) Comparison of the exact and beam solutions. "∘": the exact solution; "-": the beam solution.

solutions together for a y-window; in Fig. 10(h), we plot the x-sliced ($x = 0.5205$) solutions together for a y-window; these two figures illustrate that the two solutions do not match with each other well.

If we reinitialize beam propagation during time evolution, then the accuracy of beam solution improves significantly. Fig. 11 shows an example with $\hbar = \frac{1}{512\pi}$ and the final time $T = 0.56$. We subdivide $[0,T] = [0,0.56]$ into eight uniform sub-intervals and reinitialize the beam representation at the beginning of each sub-interval. Fig. 11(a) and (b) shows the images of the exact solution and the beam solution at the final time $T = 0.56$, and they are consistent with each other. To compare the two solutions more clearly, we plot the sliced solutions along $y = 0.5$ in Fig. 11(c) and (d), respectively; the two solutions look similar. In Fig. 11(e) and (f), we plot the y-sliced solutions together for two x-windows, and we can see that the two solutions match very well. In Fig. 11(g), we plot the x-sliced ($x = 0.5$) solutions together for a y-window; in Fig. 11(h), we plot the x-sliced ($x = 0.5205$) solutions together for a y-window. Overall, we can see that the two solutions match with each other very well.

6.3. Three-dimensional cases

6.3.1. Example 6: a three-dimensional multiplicative potential

The background potential is given by $V(x,y,z) = 10 + \sin(2\pi x)\cos(2\pi y)\sin(2\pi z)$ on a unit cube, and the initial wave function is given by

$$\begin{aligned} U(x,y,z,t=0) &= \exp(-25(x-0.5)^2)\exp\left(\frac{\imath\tau_0(x)}{\hbar}\right), \\ \tau_0(x) &= -\frac{1}{5}\ln(\exp(5(x-0.5)) + \exp(-5(x-0.5))). \end{aligned} \tag{58}$$

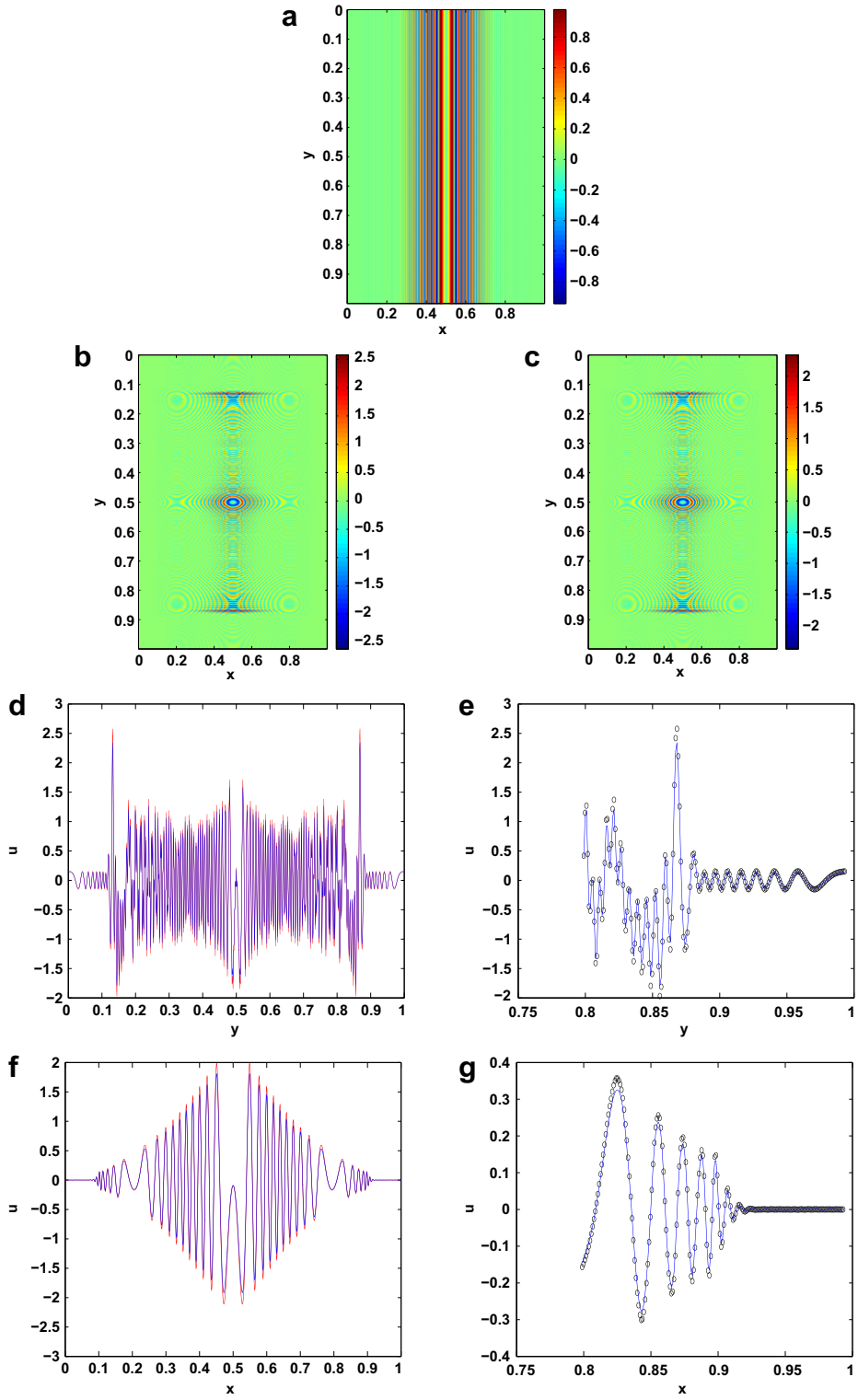

Fig. 9. Example 4. A two-dimensional additive potential, $h = \frac{1}{256\pi}$ and $T = 0.5$. (a) The initial data. (b) The exact solution. (c) The beam solution. (d) The slice along $x = 0.5$. (e) A y-window for the slice along $x = 0.5$: exact ("∘") and beam ("-"). (f) The slice along $y = 0.5$. (g) An x-window for the slice along $y = 0.5$: exact ("∘") and beam ("-").

The periodic boundary condition is imposed. Once again, the Hessian of the potential function is an indefinite matrix. Thus, depending on the initial values of beams, the beam widths may grow exponentially as time evolves.

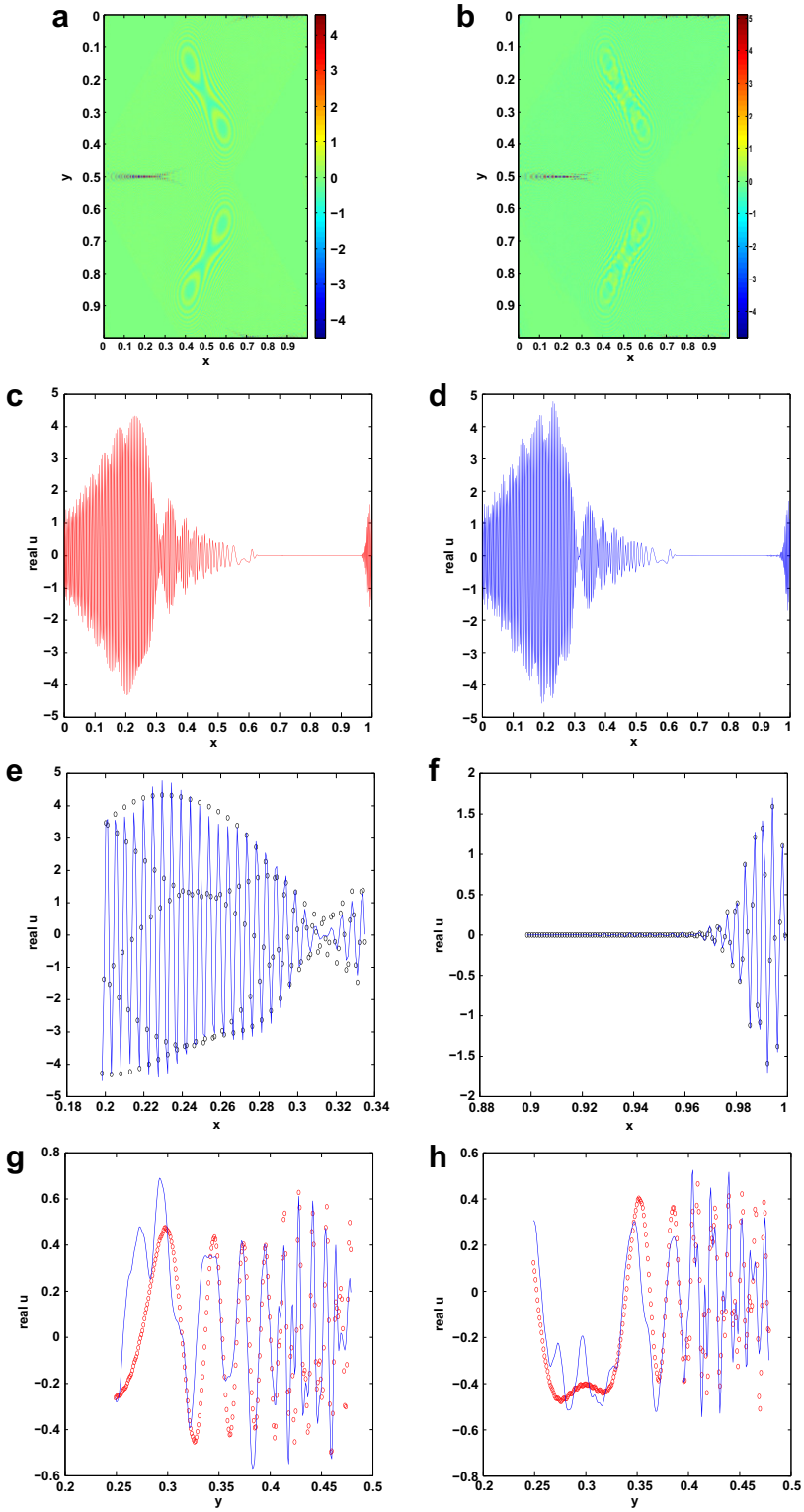

Fig. 10. Example 5. A two-dimensional multiplicative potential, $h = \frac{1}{512\pi}$ without reinitialization. The exact solution ("∘") and the beam solution ("-"). (a) The exact solution. (b) The beam solution. (c) The exact solution: the slice along $y = 0.5$. (d) The beam solution: the slice along $y = 0.5$. (e) An x-window for the slice along $y = 0.5$. (f) Another x-window for the slice along $y = 0.5$. (g) A y-window for the slice along $x = 0.5$. (h) A y-window for the slice along $x = 0.5205$.

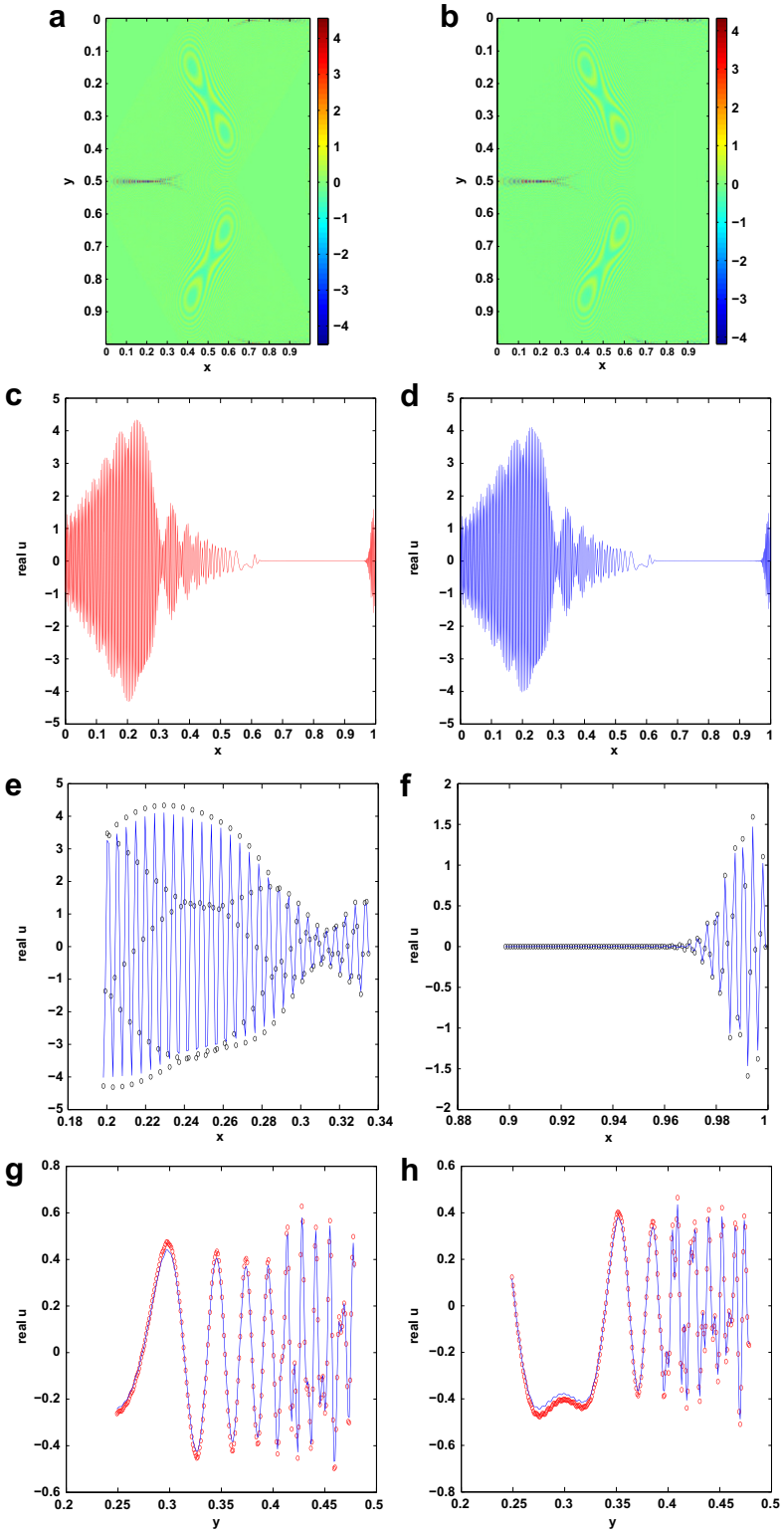

Fig. 11. Example 5. A two-dimensional multiplicative potential, $h = \frac{1}{512\pi}$ with reinitialization. The exact solution ("∘") and the beam solution ("-"). (a) The exact solution. (b) The beam solution. (c) The exact solution: the slice along $y = 0.5$. (d) The beam solution: the slice along $y = 0.5$. (e) An x-window for the slice along $y = 0.5$. (f) Another x-window for the slice along $y = 0.5$. (g) A y-window for the slice along $x = 0.5$. (h) A y-window for the slice along $x = 0.5205$.

Fig. 12 shows an example with $\hbar = \frac{1}{64\pi}$ and the final time $T = 0.56$. To well resolve the initial data, we use an $N \times N \times N = 128 \times 128 \times 128$ uniform grid to discretize $[0,1] \times [0,1] \times [0,1]$. The 3-D fast Gaussian wavepacket transform is used to generate the beam decomposition of the initial data. We subdivide $[0,T] = [0,0.56]$ into eight uniform sub-intervals and reinitialize at the beginning of each sub-interval. Fig. 12(a) shows a two-dimensional x-slice for $x = 0.5234$; Fig. 12(b) plots the exact and beam solutions together for a z-window at $x = 0.5234$ and $y = 0.5859$; Fig. 12(c) plots the exact and beam solutions together for a y-window at $x = 0.5234$ and $z = 0.5859$. Fig. 12(d) shows a two-dimensional y-slice for $y = 0.3359$; Fig. 12(e) plots the exact and beam solutions together for a z-window at $x = 0.5859$ and $y = 0.3359$; Fig. 12(f) plots the exact and beam solutions together for a x-window at $y = 0.3359$ and $z = 0.5859$. Fig. 12(g) shows a two-dimensional z-slice for $z = 0.6172$; Fig. 12(h) plots the exact and beam solutions together for a y-window at $x = 0.3906$ and $z = 0.6172$; Fig. 12(i) plots

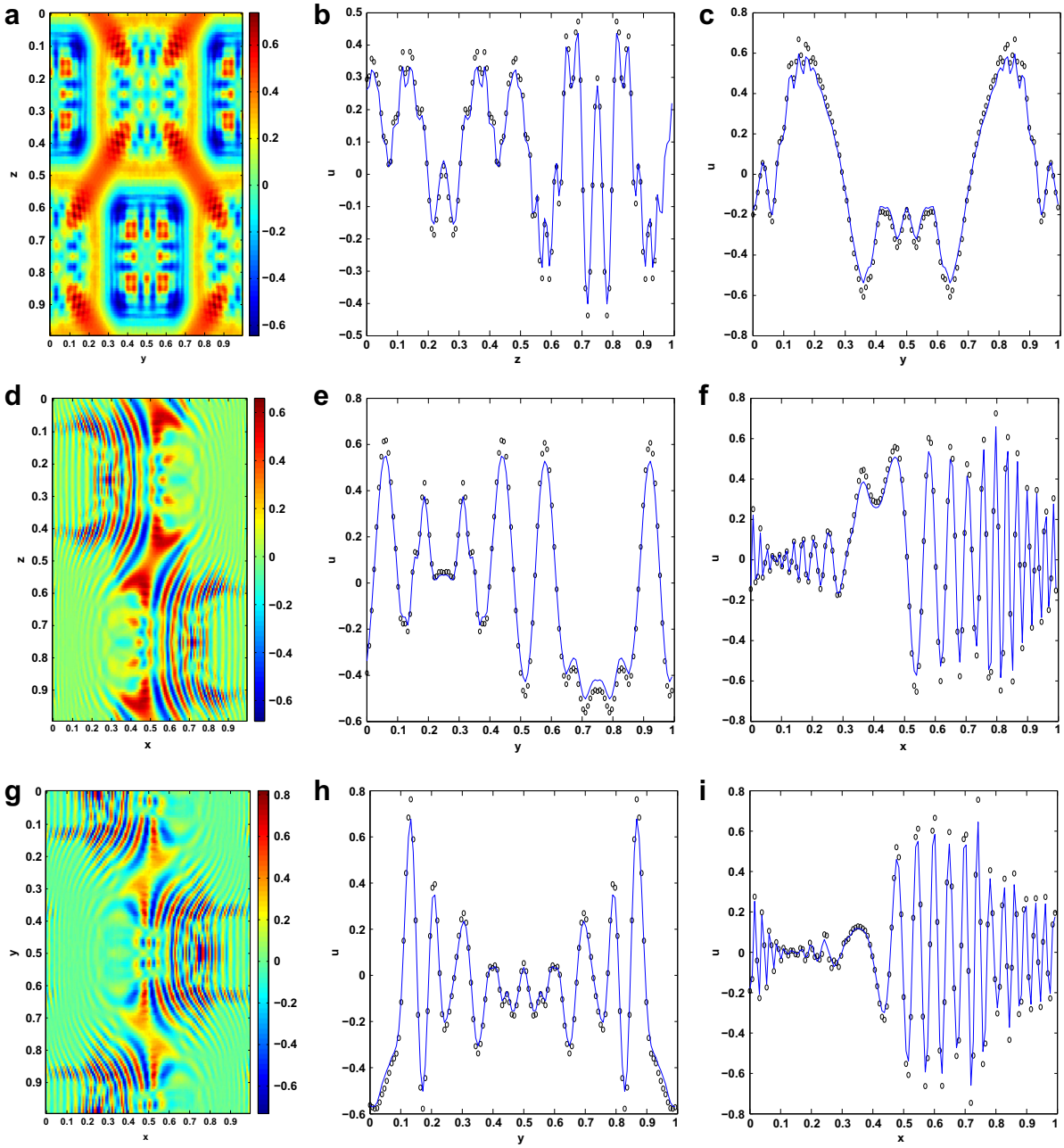

Fig. 12. Example 6. A three-dimensional multiplicative potential, $\hbar = \frac{1}{64\pi}$. The exact solution ("$\circ$") and the beam solution ("-"). (a) The x-slice for $x = 0.5234$. (b) A z-window at $x = 0.5234$ and $y = 0.5859$. (c) A y-window at $x = 0.5234$ and $z = 0.5859$. (d) The y-slice for $y = 0.3359$. (e) A z-window at $x = 0.5859$ and $y = 0.3359$. (f) An x-window at $y = 0.3359$ and $z = 0.5859$. (g) The z-slice for $z = 0.6172$. (h) A y-window at $x = 0.3906$ and $z = 0.6172$. (i) An x-window at $y = 0.3906$ and $z = 0.6172$.

the exact and beam solutions together for a x-window at $y = 0.3906$ and $z = 0.6172$. In this case, the relative L^2-error in the beam solution at the final time is 1.8%.

7. Conclusions and future works

In this paper, we addressed two critical computational problems of the Gaussian beam methods for the Schrödinger equation. We proposed fast Gaussian wavepacket transforms, which allow us to generate the beam representation for a general initial condition efficiently and accurately. Based on these efficient transforms, we addressed the long time propagation problem by introducing a reinitialization algorithm. Numerical results are presented for one-, two-, and three-dimensional examples to illustrate the properties of the proposed algorithms.

A possible direction of future work is to design more efficient initial decomposition algorithms for special initial data. For example, it is quite common to have initial data localized either in space or in frequency. In such a case, one can tailor the Gaussian wavepacket transforms to focus on certain parts of the phase space. The resulting transforms would be truly data-adaptive and potentially be more efficient. We are also working on some other related applications.

Acknowledgements

Qian would like to thank R. Burridge, S. Osher and J. Ralston for their interest and encouragement in this work. Qian is partially supported by NSF 0810104, NSF 0830161, and by AFOSR Grant #FA9550-04-1-0143. Ying would like to thank E. Candès and L. Demanet for early discussions regarding the construction of dual window functions of the transforms and B. Engquist, N. Tanushev, and R. Tsai for discussions regarding the Gaussian beam methods. Ying is partially supported by a Sloan Research Fellowship and NSF Career Award DMS-0846501.

References

[1] V.M. Babich, V.S. Buldyrev, Asymptotic Methods in Short Wave Diffraction Problems, Nauka, Moscow, 1972 (in Russian).
[2] W. Bao, S. Jin, P. Markowich, On time-splitting spectral approximations for the Schrödinger equation in the semiclassical regime, J. Comput. Phys. 175 (2002) 487–524.
[3] S. Bougacha, J. Akian, R. Alexandre, Gaussian beams summation for the wave equation in a convex domain, Commun. Math. Sci. 7 (2009) 973–1008.
[4] E. Candès, L. Demanet, D. Donoho, L. Ying, Fast discrete curvelet transforms, Multiscale Model. Simul. 5 (3) (2006) 861–899 (electronic).
[5] V. Cerveny, M. Popov, I. Psencik, Computation of wave fields in inhomogeneous media-Gaussian beam approach, Geophys. J. Roy. Astron. Soc. 70 (1982) 109–128.
[6] L. Demanet, L. Ying, Wave atoms and sparsity of oscillatory patterns, Appl. Comput. Harmon. Anal. 23 (3) (2007) 368–387.
[7] E. Heller, Cellular dynamics: a new semiclassical approach to time-dependent quantum mechanics, J. Chem. Phys. 94 (1991) 2723–2729.
[8] E. Heller, Guided Gaussian wave packets, Acc. Chem. Res. 39 (2006) 127–134.
[9] N. Hill, Gaussian beam migration, Geophysics 55 (1990) 1416–1428.
[10] L. Hörmander, On the existence and the regularity of solutions of linear pseudo-differential equations, L'Enseign. Math. 17 (1971) 99–163.
[11] S. Jin, H. Wu, X. Yang, Gaussian beam methods for the Schrödinger equation in the semi-classical regime: Lagrangian and Eulerian formulations, Commun. Math. Sci. 6 (4) (2008) 995–1020.
[12] E. Kluk, M. Herman, H. Davis, Comparison of the propagation of semiclassical frozen Gaussian wave functions with quantum propagation for a highly excited anharmonic oscillator, J. Chem. Phys. 84 (1986) 326–334.
[13] S. Leung, J. Qian, Eulerian Gaussian beam methods for Schrödinger equations in the semi-classical regime, J. Comput. Phys. 228 (2009) 2951–2977.
[14] S. Leung, J. Qian, R. Burridge, Eulerian Gaussian beams for high frequency wave propagation, Geophysics 72 (2007) SM61–SM76.
[15] S.Y. Leung, H.-K. Zhao, Gaussian beam summation for diffraction in inhomogeneous media based on the grid based particle method, Commun. Comput. Phys., in press.
[16] S. Mallat, A Wavelet Tour of Signal Processing, third ed., Elsevier/Academic Press, Amsterdam, 2009. The sparse way, With contributions from Gabriel Peyré.
[17] A. Martinez, An Introduction to Semiclassical and Microlocal Analysis, Springer, New York, 2002.
[18] V.P. Maslov, The Complex WKB Method for Nonlinear Equations I Linear Theory, Birkhauser Verlag, Basel, 1994.
[19] V.P. Maslov, M.V. Fedoriuk, Semi-classical Approximation in Quantum Mechanics, D. Reidel Publishing Company, 1981.
[20] M. Motamed, O. Runborg, Taylor expansion and discretization errors in Gaussian beam superposition, Wave Motion, in press.
[21] D. Pathria, J. Morris, Pseudo-spectral solution of nonlinear Schrödinger equations, J. Comput. Phys. 87 (1990) 108–125.
[22] J. Ralston, Gaussian beams and the propagation of singularities, in: W. Littman (Ed.), Studies in Partial Differential Equations, MAA Studies in Mathematics, vol. 23, 1983, pp. 206–248.
[23] N. Tanushev, Superpositions and higher order Gaussian beams, Commun. Math. Sci. 6 (2008) 449–475.
[24] N. Tanushev, B. Engquist, R. Tsai, Gaussian beam decomposition of high frequency wave fields, J. Comput. Phys. 228 (2009) 8856–8871.
[25] N. Tanushev, J. Qian, J. Ralston, Mountain waves and Gaussian beams, SIAM J. Multiscale Model. Simul. 6 (2007) 688–709.