Communications in Partial Differential Equations, 38: 92–134, 2013

ISSN 0360-5302 print/1532-4133 online
DOI: 10.1080/03605302.2012.727130

A Convergent Multiscale Gaussian-Beam Parametrix for the Wave Equation

GANG BAO[1,2], JIANLIANG QIAN[2], LEXING YING[3], AND HAI ZHANG[2]

[1]Department of Mathematics, Zhejiang University, Hangzhou, China

[2]Department of Mathematics, Michigan State University, East Lansing, Michigan, USA

[3]Department of Mathematics and ICES, The University of Texas at Austin, Austin, Texas, USA

The Gaussian beam method is an asymptotic method for wave equations with highly oscillatory data. In a recent published paper by two of the authors, a multiscale Gaussian beam method was first proposed for wave equations by utilizing the parabolic scaling principle and multiscale Gaussian wavepacket transforms, and numerical examples there demonstrated excellent performance of the multiscale Gaussian beam method. This article is concerned with the important convergence properties of this multiscale method. Specifically, the following results are established. If the Cauchy data are in the form of non-truncated multiscale Gaussian wavepackets, the multiscale Gaussian beam method provides a convergent parametrix for the wave equation with highly oscillatory data, and the convergence rate is $\frac{1}{\sqrt{\lambda}}$, where λ is the smallest frequency contained in the highly oscillatory data. If the highly oscillatory Cauchy data are in the form of truncated multiscale Gaussian wavepackets, the multiscale Gaussian beam method converges with a rate controlled by $\frac{1}{\sqrt{\lambda}} + \epsilon$, where ϵ is the error from initializing the Gaussian beam method by multiscale Gaussian wavepacket transforms. To prove these convergence results, it is essential to characterize multiscale properties of wavepacket interaction and beam decaying by carrying out some highly-oscillatory integrals of Fourier-integral-operator type, so that those multiscale properties lead to precise convergence orders for the multiscale Gaussian beam method.

Keywords Multiscale Gaussian beams; Multiscale Gaussian wave packets; Phase space transform; Wave equations.

Mathematics Subject Classification Primary 35S05; Secondary 35A27, 35S30, 47G30.

Received February 10, 2012; Accepted August 14, 2012
Address correspondence to Jianliang Qian, Department of Mathematics, Michigan State University, East Lansing, MI 48824, USA; E-mail: qian@math.msu.edu

1. Introduction

The Gaussian beam method is a high-frequency asymptotic method which can be used to construct parametrices for wave equations with highly oscillatory data. The idea of Gaussian beams dates back to 1960s; see [2]. Since then, Gaussian beams have been used for many different applications, such as propagation of singularities [24] and seismic modeling and imaging [7, 13, 14, 19, 21, 30]. However, all these Gaussian beam methods cited above only deal with single frequency data. In a recent work [23], Qian and Ying proposed a multiscale computational Gaussian beam method for wave equations with highly oscillatory data by making use of the parabolic scaling principle and multiscale Gaussian wavepacket transforms, and the multiscale method is able to handle the Cauchy data with a broad band of frequencies simultaneously. The aim of the current paper is to analyze the convergence properties of this multiscale Gaussian beam method.

In order to apply the Gaussian beam methods efficiently and accurately, one has to generate a beam decomposition for a general initial condition. There have been some recent advances in beam decomposition for the wave equation [19, 26, 27] and the Schrödinger equation [17, 18, 22]. In particular, [22] introduced single-scale Gaussian wavepacket transforms and developed on top of them a highly efficient initialization algorithm for the Schrödinger equation. For the wave equation, since the Hamiltonian 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 [4, 25] at any given time. Motivated by this principle, [23] introduced a set of multiscale Gaussian wavepacket transforms that enable to decompose arbitrary initial conditions of the wave equation and polarize mixed high-frequency initial data into different wave modes at multiscale resolutions. Based on this multiscale decomposition, [23] further proposed a multiscale Gaussian beam method for the wave equation with general initial conditions.

From the work in [4, 25], 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 [1, 5, 6, 9, 25] and wave atoms [10, 11]. The multiscale Gaussian wavepackets proposed in [23] were inspired by these constructions and in fact inherited the overall architecture of the wave atom construction; however, the transforms were modified appropriately so that the wavepackets maintain a Gaussian profile. Such multiscale Gaussian wavepackets are also similar to the wavepackets of Cordoba-Fefferman [8], which are sufficiently localized in phase space. From a more general perspective, such multiscale wavepackets are also related to phase space transforms which were recently used to construct parametrices for wave operators with rough coefficients in [12, 28].

It was further shown in [23] that the multiscale Gaussian beam method yields an asymptotic solution for the wave equation. In this paper, we carry out the convergence study of the multiscale method. Specifically, we prove the following results. If the Cauchy data are in the form of non-truncated multiscale Gaussian wavepackets, the multiscale Gaussian beam method provides a convergent parametrix for wave equations with highly oscillatory data, and the convergence rate is $\frac{1}{\sqrt{\lambda}}$, where λ is the smallest frequency contained in the highly oscillatory data. If the highly oscillatory Cauchy data are in the form of truncated multiscale Gaussian wavepackets, the multiscale Gaussian beam method converges with a rate controlled by $\frac{1}{\sqrt{\lambda}} + \epsilon$, where ϵ is the error due to initializing the Gaussian beam method with multiscale Gaussian wavepacket transforms.

At this point, we mention that in [3] some convergence results for single-scale Gaussian beams are given for wave equations; our convergence analysis aims at multiscale Gaussian beams for wave equations so that our starting framework and analysis tools are different from [3].

The accuracy of parametrices for this problem is tied to the regularity of the coefficients of the wave equation. In [29] Waters constructed a parametrix for general second-order wave equations assuming the minimal regularity as in Smith [25], and following [25] she also proved that it was sufficiently accurate to allow correction to an exact solution by a Volterra integral equation. She used a frame of (modulated) Gaussians and also observed that in the presence of more coefficient regularity this frame could be propagated as Gaussian beams. At essentially the same time Qian and Ying [23] constructed a parametrix for the wave equation using Gaussian beams initialized by a frame closer to the one in [25]. In the present paper we prove the accuracy of that parametrix. Naturally at quite a few places the arguments are close to those in [29].

The rest of the paper is organized as follows. Section 2 summarizes the construction of a multiscale Gaussian beam parametrix, including Gaussian beam setup and multiscale Gaussian wavepacket transforms. Section 3 details some properties of Hamiltonian flows and wavepacket interactions and presents convergence analysis of the multiscale Gaussian beam parametrix. Section 4 proves some technical lemmas that are needed in the analysis.

The following notations are needed:

1. Let x be a vector in $\mathbb{R}^d$, and let its usual Euclidean norm be denoted by $|x|$.
2. Let A be a symmetric matrix and c be a real number. The expression $A > c$ means that the matrix $A - cI$ is positive definite and symmetric. Here I denotes the identity matrix.
3. Let A be a matrix, and denote by $\|A\|$ the matrix norm defined by $\|A\| = \sup_{x\neq 0} \frac{|Ax|}{|x|}$.
4. Let $f \in C^k(\mathbb{R}^d)$, where $k \geq 0$ is an integer. Write $\|f\|_{C^k} = \sup_{x\in\mathbb{R}^d}\{|\frac{\partial^\alpha f}{\partial x^\alpha}|; |\alpha| \leq k\}$.
5. Let f be a function defined in $\mathbb{R}^d$, x_0 a point in $\mathbb{R}^d$, and k an integer. The expression $f = O(|x - x_0|^k)$ means that $\frac{\partial^\alpha f}{\partial x^\alpha}(x_0) = 0$ for all $|\alpha| \leq k - 1$.
6. Let a, b, and c be three positive numbers. The expression $a \lesssim_c b$ means that there exists a constant C depending on c such that $a \leq Cb$.

2. Multiscale Gaussian Beam Parametrix Construction

2.1. *Gaussian Beams*

Consider the following wave equation,

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

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

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

where the velocity function $V(x)$ is smooth, positive, and bounded away from zero; the functions $f_1(x)$ and $f_2(x)$ belong to $H^1(\mathbb{R}^d)$ and $L^2(\mathbb{R}^d)$, respectively, and they are assumed to be highly oscillatory.

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

$$A(x,t)e^{\sqrt{-1}\omega\tau(x,t)}, \tag{4}$$

where $\tau(x,t)$ is the phase function and $A(x,t)$ the amplitude function. 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 for large ω. After substituting the ansatz (4) into the wave equation (1) and considering the leading orders in inverse powers of the large parameter ω, we arrive at 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

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

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

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

where the square root is defined in the complex plane except the non-positive axis by analytical continuation. Here $G^{\pm}(x,p)$ is clearly homogeneous of degree one in the momentum variable p.

To construct asymptotic solutions for the wave equation, we employ Gaussian beams [20, 24, 27]. Because the two polarized wave modes may 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 [20, 24, 27], 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 obtain Gaussian beam ingredients, we solve the following system of equations,

$$\dot{x} = \frac{dx}{dt} = G_p, \quad x|_{t=0} = x_0, \tag{9}$$

$$\dot{p} = \frac{dp}{dt} = -G_x, \quad p|_{t=0} = p_0, \tag{10}$$

$$\dot{M} = \frac{dM(t)}{dt} = -(G_{xx} + M(t)G_{xp} + G_{xp}^T M(t) + M(t)G_{pp}M(t)), \quad M|_{t=0} = \sqrt{-1}\epsilon I, \tag{11}$$

$$\dot{A} = \frac{dA}{dt} = \left(\frac{A(x(t),t)}{2G}\left(G_x \cdot G_p + G_p^T M(t) G_p - V^2(x(t))\text{trace}(M(t))\right)\right), \quad A|_{t=0} = A_0, \tag{12}$$

where t is time parameterizing bicharacteristics. See [20, 23, 24, 27] for detailed derivations.

Since the corresponding ray path is $\gamma = \{(x(t), t) : t \geq 0\}$, by construction, we have $p(t) = \tau_x(x(t), t)$, $M(t) = \tau_{xx}(x(t), p(t))$, and $A(t) = A(x(t), t)$ along γ. In addition, by homogeneity of the Hamiltonian G, $\tau(t) = \tau(x(t), t)$ can be taken to be zero along γ. It follows from the symplectic structure of the Hamiltonian system that the Hessian $M(t)$ along γ has a positive-definite imaginary part provided that it initially does; see [20, 24, 27].

We are now 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:

$$\tau(x, t) \equiv p(t) \cdot (x - x(t)) + \frac{1}{2}(x - x(t))^T M(t)(x - x(t)), \tag{13}$$

$$A(x, t) \equiv A(x(t), t) = A(t), \tag{14}$$

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,

$$\Phi(x, t) = A(x, t) \exp\left(\sqrt{-1}\omega\tau(x, t)\right). \tag{15}$$

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 $t = 0$. Since the phase $\tau(x, t)$ has an imaginary part, $\text{Im}(\tau(x, t)) = \frac{1}{2}(x - x(t))^T \text{Im}(M(t))(x - x(t))$, the function $\Phi(x, t)$ has a Gaussian profile of the form,

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

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 $x_\alpha^{\pm}(t)$, $p_\alpha^{\pm}(t)$, $M_\alpha^{\pm}(t)$, $A_\alpha^{\pm}(t)$, $\tau_\alpha^{\pm}(x, t)$, $A_\alpha^{\pm}(x, t)$, and $\Phi_\alpha^{\pm}(x, t)$.

For a given tuple $\alpha = (x_0, p_0, M_0, A_0)$, the Gaussian beams $\Phi_\alpha^{\pm}(x, t)$ have a simple Gaussian profile in the spatial variable x. 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), one must 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 [17–19, 22] for several different approaches for the Helmholtz and the Schrödinger equations. In the Schrödinger case, the Hamiltonian is not homogeneous, hence one cannot restrict the Hamiltonian flow to the cosphere bundle; consequently, the single-scale Gaussian wavepacket transform is used 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 *multiscale* transforms with basis functions satisfying the parabolic scaling principle. In [23], Qian and Ying designed such multiscale transforms, called multiscale Gaussian wavepacket transforms, to carry out the needed multiscale decomposition. These multiscale transforms follow the architecture of the wave atoms proposed in [10, 11].

In the next subsection, we summarize the formulation of these multiscale Gaussian wavepacket transforms and prove some approximation properties of the resulting frames.

2.2. *Multiscale Gaussian Wavepacket Transforms*

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

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

It is clear that $\xi \in C_\ell$ implies that $|\xi| = O(4^\ell)$. Each corona C_ℓ is further partitioned into boxes

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

where the integer multi-index $i = (i_1, i_2, \ldots, i_d)$ 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 Gaussian function $\tilde{g}_{\ell,i}(\xi)$ by

$$\tilde{g}_{\ell,i}(\xi) = e^{-\left(\frac{|\xi-\xi_{\ell,i}|}{\sigma_\ell}\right)^2},$$

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

To construct a Gaussian-beam parametrix, we need to decompose the Cauchy data into Gaussian wavepackets; however, computationally it is highly nontrivial since a Gaussian function is not compactly supported in the Fourier domain. Consequently, we introduce truncated Gaussian bumps. To do that, we first choose a cut-off function $\chi \in C_0^\infty(\mathbb{R}^n)$ with $0 \le \chi \le 1$ such that $\chi(\xi) = 1$ for $\xi \in \{\xi : \max_{1\le s\le d} |\xi_s| \le 1\}$ and $\chi(\xi) = 0$ for $\xi \in \{\xi : \max_{1\le s\le d} |\xi_s| \ge 2\}$. Since $\sigma_\ell = \frac{W_\ell}{2} = 2^{\ell-1}$, we define a truncation function $\chi_{\ell,i}$ for each box $B_{\ell,i}$ by

$$\chi_{\ell,i}(\xi) = \chi\left(\frac{\xi - \xi_{\ell,i}}{\sigma_\ell}\right).$$

The truncated Gaussian bump $g_{\ell,i}$ in the frequency domain is given by

$$g_{\ell,i}(\xi) = \chi_{\ell,i}(\xi)\tilde{g}_{\ell,i}(\xi),$$

where $g_{\ell,i}(\xi)$ is compactly supported in a box centered at $\xi_{\ell,i}$ of length $L_\ell = 2W_\ell$.

Now, we define a conjugate filter $h_{\ell,i}$ for the truncated Gaussian bump $g_{\ell,i}$ by setting

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

We notice that this conjugate filter is slightly different from the one used in [23].

We have following properties about the conjugate filter $h_{\ell,i}$.

Lemma 2.1.

1. *For any given ξ, there exist at most 3^d indices (ℓ, i) such that $\chi_{\ell,i}(\xi) \neq 0$.*
2. *The denominator of $h_{\ell,i}$ is always positive. Moreover, there exist positive constants C_1 and C_2 such that*

$$C_1 \le h_{\ell,i}(\xi) \le C_2 \quad \textit{for all } \xi \in B_{\ell,i}.$$

Proof. Let (ℓ, i) be an index such that $\chi_{\ell,i}(\xi) \neq 0$. By the definition of $\chi_{\ell,i}$, we have

$$\max_{1\le s\le d} |\xi_s - \xi_{\ell,i,s}| < 2\sigma_\ell,$$

or $|\xi_s - \xi_{\ell,i,s}| < 2\sigma_\ell$ for all $1 \le s \le d$. Since for each s there exist at most three real numbers of $\xi_{\ell,i,s}$'s such that $|\xi_s - \xi_{\ell,i,s}| < 2\sigma_\ell$, it follows that there are at most 3^d numbers of $\xi_{\ell,i}$'s such that $\max_{1\le s\le d} |\xi_s - \xi_{\ell,i,s}| < 2\sigma_\ell$. The first part of the lemma is proved.

Next we show the second part. Assume that $\xi \in B_{\ell,i}$, which implies that

$$\max_{1\le s\le d} |\xi_s - \xi_{\ell,i,s}| \le \sigma_\ell.$$

Thus $\chi_{\ell,i}(\xi) = 1$ and

$$g_{\ell,i}(\xi) = \chi_{\ell,i}(\xi)\tilde{g}_{\ell,i}(\xi) \geq e^{-\left(\frac{|\xi-\xi_{\ell,i}|}{\sigma_\ell}\right)^2} \geq e^{-d}.$$

Therefore

$$\sum_{\ell,i} \chi_{\ell,i}(\xi) g_{\ell,i}(\xi) \geq e^{-d}.$$

On the other hand,

$$\sum_{\ell,i} \chi_{\ell,i}(\xi) g_{\ell,i}(\xi) \leq \sum_{\ell,i} \chi_{\ell,i}(\xi)\tilde{g}_{\ell,i}(\xi) \leq \sum_{\ell,i} \chi_{\ell,i}(\xi) \leq 3^d,$$

where we have used the result in the first part of the lemma.

Consequently, we have proved that

$$e^{-d} \leq \sum_{\ell,i} \chi_{\ell,i}(\xi) g_{\ell,i}(\xi) \leq 3^d. \tag{16}$$

The second part of the lemma follows. □

By construction, the products of $g_{\ell,i}(\xi)$ and $h_{\ell,i}(\xi)$ form a partition of unity:

$$\sum_{\ell,i} g_{\ell,i}(\xi) h_{\ell,i}(\xi) = 1,$$

and $h_{\ell,i}(\xi)$ is a smooth function compactly supported in a box centered at $\xi_{\ell,i}$ with size $L_\ell = 2W_\ell$ in each dimension (*i.e.*, $\prod_{s=1}^{d}[\xi_{\ell,i,s} - W_\ell, \xi_{\ell,i,s} + W_\ell]$).

We then introduce three sets of functions $\{\phi_{\ell,i,k}(x)\}$, $\{\tilde{\phi}_{\ell,i,k}(x)\}$, and $\{\psi_{\ell,i,k}(x)\}$, defined in the Fourier domain by

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

$$\hat{\tilde{\phi}}_{\ell,i,k}(\xi) = \frac{1}{L_\ell^{d/2}} e^{-2\pi\sqrt{-1}\frac{k\cdot\xi}{L_\ell}} \tilde{g}_{\ell,i}(\xi), \quad \forall k \in \mathbb{Z}^d,$$

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

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

$$\phi_{\ell,i,k}(x) = \frac{1}{L_\ell^{d/2}} \int_{\mathbb{R}^d} e^{2\pi\sqrt{-1}(x-\frac{k}{L_\ell})\cdot\xi} g_{\ell,i}(\xi) d\xi, \quad \forall k \in \mathbb{Z}^d, \tag{17}$$

$$\tilde{\phi}_{\ell,i,k}(x) = \frac{1}{L_\ell^{d/2}} \int_{\mathbb{R}^d} e^{2\pi\sqrt{-1}(x-\frac{k}{L_\ell})\cdot\xi} \tilde{g}_{\ell,i}(\xi) d\xi, \quad \forall k \in \mathbb{Z}^d, \tag{18}$$

$$\psi_{\ell,i,k}(x) = \frac{1}{L_\ell^{d/2}} \int_{\mathbb{R}^d} e^{2\pi\sqrt{-1}(x-\frac{k}{L_\ell})\cdot\xi} h_{\ell,i}(\xi) d\xi, \quad \forall k \in \mathbb{Z}^d. \tag{19}$$

As shown in [23], the definition of $\tilde{g}_{\ell,i}(\xi)$ implies that

$$\tilde{\phi}_{\ell,i,k}(x) = \left(\sqrt{\frac{\pi}{L_\ell}}\sigma_\ell\right)^d \cdot e^{2\pi\sqrt{-1}(x-\frac{k}{L_\ell})\cdot\xi_{\ell,i}} \cdot e^{-\sigma_\ell^2\pi^2|x-\frac{k}{L_\ell}|^2}; \tag{20}$$

i.e., $\tilde{\phi}_{\ell,i,k}(x)$ is 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.

For ease of notation, following [25] we introduce the triple $\gamma = (\ell, i, k) \in \mathbb{N} \times \mathbb{Z}^d \times \mathbb{Z}^d$. Then $\tilde{\phi}_{\ell,i,k}(x) = \tilde{\phi}_\gamma(x)$, $\phi_{\ell,i,k}(x) = \phi_\gamma(x)$, and $\psi_{\ell,i,k}(x) = \psi_\gamma(x)$. We also write $\lambda_\gamma = |\xi_\gamma| = |\xi_{\ell,i}|$, $\xi_\gamma = \xi_{\ell,i}$, and $B_\gamma = B_{\ell,i}$. Furthermore, we define $x_\gamma = \frac{k}{L_\ell}$ and $D_\gamma = \prod_{s=1}^d [x_{\gamma,s} - \frac{1}{2L_\ell}, x_{\gamma,s} + \frac{1}{2L_\ell}]$. Note that the products of the boxes $D_\gamma \times B_\gamma$ form a tiling of the phase space $\mathbb{R}^{2d}$.

We next present some properties about the three sets of functions $\{\phi_\gamma(x)\}$, $\{\tilde{\phi}_\gamma(x)\}$ and $\{\psi_\gamma(x)\}$. First functions $\{\psi_\gamma(x)\}$ and $\{\phi_\gamma(x)\}$ are dual frames for the space $L^2(\mathbb{R}^d)$ as shown in [23].

Lemma 2.2. [23] *For any $f \in L^2(\mathbb{R}^d)$,*

$$f(x) = \sum_\gamma \langle \psi_\gamma, f\rangle \phi_\gamma(x). \tag{21}$$

Proof. See [23] for the proof. □

We then give the following stability estimate for the co-frame $\{\psi_\gamma(x)\}$.

Lemma 2.3. *There exist positive constants K_1 and K_2 such that the following hold:*

$$K_1\|f\|_2^2 \le \sum_\gamma |\langle \psi_\gamma, f\rangle|^2 \le K_2\|f\|_2^2, \tag{22}$$

$$K_1\|f\|_{H^1}^2 \le \sum_\gamma \lambda_\gamma^2 |\langle \psi_\gamma, f\rangle|^2 \le K_2\|f\|_{H^1}^2. \tag{23}$$

Proof. Note that

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

$$\sum_\gamma \lambda_\gamma^2|\langle \psi_\gamma, f\rangle|^2 = \sum_{\ell,i}|\xi_{\ell,i}|^2\sum_k \left|\frac{1}{L_\ell^{d/2}}\int_{\mathbb{R}^d} e^{2\pi\sqrt{-1}\cdot\frac{k\cdot\xi}{L_\ell}} h_{\ell,i}(\xi)\hat{f}(\xi)d\xi\right|^2$$

$$= \sum_{\ell,i} \int_{\mathbb{R}^d} |\xi_{\ell,i}|^2 |h_{\ell,i}(\xi)\hat{f}(\xi)|^2 d\xi$$

$$= \int_{\mathbb{R}^d} \left(\sum_{\ell,i} |\xi_{\ell,i}|^2 |h_{\ell,i}(\xi)|^2 \right) |\hat{f}(\xi)|^2 d\xi.$$

Using properties of functions $h_{\ell,i}$'s in Lemma 2.1, one can check that there exist positive numbers K_1 and K_2 such that

$$K_1 \leq \sum_{\ell,i} |h_{\ell,i}(\xi)|^2 \leq K_2,$$

$$K_1(|\xi|^2 + 1) \leq \sum_{\ell,i} |\xi_{\ell,i}|^2 |h_{\ell,i}(\xi)|^2 \leq K_2(|\xi|^2 + 1),$$

which yield (22) and (23). □

Since the proof of Lemma 2.2 relies on $g_{\ell,i}$ being compactly supported in an essential way, the representation $f(x) = \sum_\gamma \langle \psi_\gamma, f \rangle \phi_\gamma(x)$ holds only for the set of (truncated) Gaussian wavepackets $\{\phi_\gamma(x)\}$, not for the set of (non-truncated) Gaussian wavepackets $\{\tilde{\phi}_\gamma(x)\}$; namely, $f(x) \neq \sum_\gamma \langle \psi_\gamma, f \rangle \tilde{\phi}_\gamma(x)$. However, when initializing multiscale Gaussian beam propagation as illustrated in [23], we need the Cauchy data to be in the form of (non-truncated) Gaussian wavepackets $\tilde{\phi}_\gamma$. Therefore, a natural question is: what is the difference between $\{\phi_\gamma(x)\}$ and $\{\tilde{\phi}_\gamma\}$? To illustrate this point, we estimate $\|\hat{\phi}_\gamma - \hat{\tilde{\phi}}_\gamma\|_{L^2(\mathbb{R}^d)}$. By direct calculation,

$$\begin{aligned} \|\hat{\phi}_\gamma - \hat{\tilde{\phi}}_\gamma\|^2_{L^2(\mathbb{R}^d)} &= \int_{\mathbb{R}^d} \frac{1}{L_\ell^d} \cdot |\tilde{g}_{\ell,i}(\xi)|^2 \cdot (1 - \chi_{\ell,i}(\xi))^2 d\xi \\ &= \frac{1}{L_\ell^d} \int_{\mathbb{R}^d} e^{-\frac{2|\xi - \xi_{\ell,i}|^2}{\sigma_\ell^2}} \left(1 - \chi\left(\frac{\xi - \xi_{\ell,i}}{\sigma_\ell}\right)\right)^2 d\xi \\ &= \frac{1}{4^d} \int_{\mathbb{R}^d} e^{-2|\xi|^2} \cdot (1 - \chi(\xi))^2 d\xi \\ &< \frac{1}{4^d} \int_{|\xi|>1} e^{-2|\xi|^2} d\xi \\ &< (0.0072)^d, \end{aligned}$$

which is a small number. Moreover, the above approximation error $(0.0072)^d$ can be made arbitrarily small by increasing the support of the truncation function χ.

Consequently, we may substitute $\{\phi_\gamma(x)\}$ with $\{\tilde{\phi}_\gamma(x)\}$ in the expansion of f. Moreover, we have the following two results.

Lemma 2.4. *Let $f \in L^2(\mathbb{R}^d)$ be such that*

$$f(x) = \sum_\gamma \langle \psi_\gamma, f \rangle \phi_\gamma(x) = \sum_\gamma c_\gamma \phi_\gamma(x).$$

Define

$$\tilde{f}(x) = \sum_\gamma c_\gamma \tilde{\phi}_\gamma(x).$$

Then there is a number $\varepsilon > 0$ *which is independent of* f *such that* $\|f - \tilde{f}\|_{L^2(\mathbb{R}^d)} \leq \varepsilon \|f\|_{L^2(\mathbb{R}^d)}$.

Proof. See Subsection 4.2. □

Similarly, we have

Lemma 2.5. *Let* $f \in H^1(\mathbb{R}^d)$ *satisfy*

$$f(x) = \sum_\gamma \langle \psi_\gamma, f \rangle \phi_\gamma(x) = \sum_\gamma c_\gamma \phi_\gamma(x).$$

Define

$$\tilde{f}(x) = \sum_\gamma c_\gamma \tilde{\phi}_\gamma(x).$$

Then there is a number $\varepsilon > 0$ *which is independent of* f *such that* $\|f - \tilde{f}\|_{H^1(\mathbb{R}^d)} \leq \varepsilon \|f\|_{H^1(\mathbb{R}^d)}$.

2.3. *Multiscale Gaussian Beam Parametrix*

We construct the multiscale Gaussian beam parametrix for the wave equation (1)–(3) in this section. We begin by assuming that the initial conditions (2) and (3) are highly oscillatory. By Lemma 2.2, we have the following decompositions for the initial conditions:

$$f_1(x) = \textstyle\sum_\gamma a_\gamma \phi_\gamma(x), \tag{24}$$

$$f_2(x) = \textstyle\sum_\gamma b_\gamma \phi_\gamma(x) \;, \tag{25}$$

where $a_\gamma = \langle f_1, \psi_\gamma \rangle$ and $b_\gamma = \langle f_2, \psi_\gamma \rangle$.

To construct the Gaussian beam parametrix, we need to decompose initial conditions (2) and (3) into non-truncated Gaussian wavepackets. Since $\tilde{\phi}_\gamma(x)$ is a good approximation for $\phi_\gamma(x)$, we can get approximate decompositions by substituting $\phi_\gamma(x)$ with $\tilde{\phi}_\gamma(x)$ in (24) and (25). Although this introduces an extra error in the approximation of initial data, it is reasonable as long as the error is small.

Therefore, to simplify subsequent discussions, we consider the wave equation with the following approximate initial data instead of the exact data (24) and (25):

$$U_{tt} - V^2(x)\Delta U = 0, \quad x \in \mathbb{R}^d, \quad t > 0, \tag{26}$$

$$U|_{t=0} = \textstyle\sum_\gamma a_\gamma \tilde{\phi}_\gamma(x), \tag{27}$$

$$U_t|_{t=0} = \textstyle\sum_\gamma b_\gamma \tilde{\phi}_\gamma(x). \tag{28}$$

Motivated by the approximation

$$\tilde{\phi}_\gamma(x) = \left(\sqrt{\frac{\pi}{L_\ell}}\sigma_\ell\right)^d \cdot e^{\sqrt{-1}\cdot\lambda_\gamma(x-\frac{k}{L_\ell})\cdot\frac{2\pi\xi_\gamma}{\lambda_\gamma}} \cdot e^{-\lambda_\gamma\left(\frac{\sigma_\ell^2\pi^2}{\lambda_\gamma}|x-\frac{k}{L_\ell}|^2\right)},$$

where $\lambda_\gamma = |\xi_{\ell,i}|$ behaves like a large parameter ω, as shown in [23] we can construct one Gaussian beam for each wave mode, respectively, by solving

$$\dot{x} = G_p, \quad x|_{t=0} = \frac{k}{L_\ell}, \tag{29}$$

$$\dot{p} = -G_x, \quad p|_{t=0} = 2\pi\frac{\xi_\gamma}{\lambda_\gamma}, \tag{30}$$

$$\dot{M} = -G_{xp}^T M - MG_{px} - MG_{pp}M - G_{xx}, \quad M|_{t=0} = \sqrt{-1}\cdot\frac{2\pi^2\sigma_\ell^2}{\lambda_\gamma}I, \tag{31}$$

$$\dot{A} = -\frac{A}{2G}\left(V^2\text{trace}(M) - G_x\cdot G_p - G_p^T M G_p\right), \quad A|_{t=0} = \left(\sqrt{\frac{\pi}{L_\ell}}\sigma_\ell\right)^d, \tag{32}$$

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

By the homogeneity of the Hamiltonian function $G(x, p)$ we can easily verify that $\xi_\gamma(t) = \lambda_\gamma \cdot p_\gamma(t)$ and $x_\gamma(t)$ also satisfy differential equations (30) and (29), respectively.

We are now ready to construct the multiscale Gaussian beam parametrix for the wave equation (26). For each wave mode, we define the corresponding phase function $\tau_\gamma^\pm(x, t)$, amplitude function $A_\gamma^\pm(x, t)$, and Gaussian beam $\Phi_\gamma^\pm(x, t)$ by

$$\tau_\gamma^\pm(x,t) = p_\gamma^\pm(t)\cdot(x - x_\gamma^\pm(t)) + \frac{1}{2}(x - x_\gamma^\pm(t))^T M_\gamma^\pm(t)(x - x_\gamma^\pm(t)), \tag{33}$$

$$A_\gamma^\pm(x,t) = A_\gamma^\pm(t), \tag{34}$$

$$\Phi_\gamma^\pm(x,t) = A_\gamma^\pm(x,t)\exp\left(\sqrt{-1}\cdot\lambda_\gamma\cdot\tau_\gamma^\pm(x,t)\right). \tag{35}$$

The global multiscale Gaussian beam parametrix to the wave equation (26) takes the following form:

$$\widetilde{U}(x,t) = \sum_\gamma c_\gamma^+\Phi_\gamma^+(x,t) + \sum_\gamma c_\gamma^-\Phi_\gamma^-(x,t), \tag{36}$$

where the coefficients $c_\gamma^\pm$ are determined by matching the beam asymptotic solution with initial conditions (27) and (28). As derived in [23], these coefficients are given by

$$c_\gamma^+ = \frac{1}{2}\left(a_\gamma - \frac{b_\gamma}{\sqrt{-1}\cdot G^+\left(\frac{k}{L_\ell}, 2\pi\xi_\gamma\right)}\right), \tag{37}$$

$$c_\gamma^- = \frac{1}{2}\left(a_\gamma + \frac{b_\gamma}{\sqrt{-1}\cdot G^+\left(\frac{k}{L_\ell}, 2\pi\xi_\gamma\right)}\right). \tag{38}$$

This finishes our construction of the multiscale Gaussian beam parametrix.

3. Analysis of Multiscale Gaussian Beam Parametrix

In this section, we prove the convergence of the multiscale Gaussian beam parametrix (36) for the linear wave equation (26)-(28). To do that, we need some technical lemmas.

3.1. *Properties of Hamiltonian Flows and Phase Functions*

First we present some properties of solutions to the system (29)–(32) which are crucial to our analysis.

Lemma 3.1. *For $T > 0$ fixed, assume that $\|M_\gamma(t)\| \le C$ for all $0 \le t \le T$. There exist positive constants C_1, C_2, C_3, and C_4 depending on $\|V\|_{C^1}$, T and C such that*

$$C_1 \le |p_\gamma(t)| \le C_2, \tag{39}$$

$$C_3\lambda_\gamma^{\frac{d}{4}} \le |A_\gamma(t)| \le C_4\lambda_\gamma^{\frac{d}{4}}. \tag{40}$$

Proof. For simplicity, we suppress the index γ and take $G(x, p) = V(x)|p|$. To prove (39), we note that $\dot{p}(t) = -V_x(x(t))|p(t)|$. By using the Gronwall inequality, we have

$$\begin{aligned}\frac{d}{dt}|p(t)|^2 &= -2p(t)^T V_x(x(t))|p(t)| \le 2|p(t)|^2\|V\|_{C^1}\\ &\Rightarrow |p(t)|^2 \le |p(0)|^2 e^{2T\|V\|_{C^1}}.\end{aligned}$$

Similarly,

$$\begin{aligned}\frac{d}{dt}|p(t)|^2 &= -2p(t)^T V_x(x(t))|p(t)| \ge -2|p(t)|^2\|V\|_{C^1}\\ &\Rightarrow |p(t)|^2 \ge |p(0)|^2 e^{-2T\|V\|_{C^1}}.\end{aligned}$$

Since $|p(0)| = 2\pi$, we may choose $C_1' = e^{-T\|V\|_{C^1}}$ and $C_2' = e^{T\|V\|_{C^1}}$, so that (39) holds for $C_1 = 2\pi C_1'$ and $C_2 = 2\pi C_2'$.

We next show (40). Rewrite equation (32) as

$$\frac{d\log A(t)}{dt} = -\frac{1}{2G}\left(V^2\text{trace}(M) - G_x \cdot G_p - G_p^T M G_p\right).$$

Recall that $G(x, p) = V(x)|p|$, $G_p(x, p) = V(x)\frac{p}{|p|}$, and $G_x(x, p) = V_x(x)|p|$, we have

$$\begin{aligned}\frac{d\log A(t)}{dt} = -\frac{1}{2|p(t)|}\Bigg(&V(x(t))\text{trace}(M(t))\\ &- V_x(x(t)) \cdot p(t) - V(x(t))\left(\frac{p(t)}{|p(t)|}\right)^T M(t)\left(\frac{p(t)}{|p(t)|}\right)\Bigg).\end{aligned}$$

Denote the right hand side of the above equation by $f(t)$. Note that f is a complex valued function. Since $|\text{trace}(M(t))| \le d \cdot \|M(t)\| \le d \cdot C$ and $C_2 \ge |p(t)| \ge C_1$, we see that $|f(t)| \le \frac{\|V\|_{C^1}(dC+C+C_2)}{C_1}$. We can check that

$$A(t) = A(0)e^{\frac{1}{2}\int_0^t f(s)\,ds}$$

is the solution to the equation (32). It follows that

$$e^{-T\frac{\|V\|_{C^1}(dC+C+C_2)}{2C_1}} \leq \frac{|A(t)|}{|A(0)|} \leq e^{T\frac{\|V\|_{C^1}(dC+C+C_2)}{2C_1}}.$$

Now, since $A(0) = \left(\sqrt{\frac{\pi}{L_\ell}}\sigma_\ell\right)^d = \left(\sqrt{\frac{\pi}{2^\ell}}2^{\ell-1}\right)^d = (\frac{\pi}{8})^{\frac{d}{2}} \cdot (4^\ell)^{\frac{d}{4}} = O(\lambda^{\frac{d}{4}})$, (40) follows for properly chosen constants C_3 and C_4. □

We also have the following two lemmas about the Hamiltonian flow.

Lemma 3.2. *Let $T > 0$ be fixed. Then there exist positive constants C_5 and C_6, depending on $\|V\|_{C^2}$ and T such that the following holds for all $t \in [0, T]$:*

$$\begin{aligned} &C_5(\lambda_\gamma\lambda_{\gamma'}|x_\gamma(t) - x_{\gamma'}(t)|^2 + |\xi_\gamma(t) - \xi_{\gamma'}(t)|^2) \\ &\quad \leq \lambda_\gamma\lambda_{\gamma'}|x_\gamma(0) - x_{\gamma'}(0)|^2 + |\xi_\gamma(0) - \xi_{\gamma'}(0)|^2 \\ &\quad \leq C_6(\lambda_\gamma\lambda_{\gamma'}|x_\gamma(t) - x_{\gamma'}(t)|^2 + |\xi_\gamma(t) - \xi_{\gamma'}(t)|^2). \end{aligned} \tag{41}$$

Proof. Denote $v(t) = x_\gamma(t) - x_{\gamma'}(t)$, $w(t) = \xi_\gamma(t) - \xi_{\gamma'}(t) = \lambda_\gamma p_\gamma(t) - \lambda_{\gamma'} p_{\gamma'}(t)$. We have

$$\begin{aligned} \left|\frac{dv(t)}{dt}\right| &= \left|V(x_\gamma(t))\frac{\xi_\gamma(t)}{|\xi_\gamma(t)|} - V(x_{\gamma'}(t))\frac{\xi_{\gamma'}(t)}{|\xi_{\gamma'}(t)|}\right| \\ &\leq |(V(x_\gamma(t)) - V(x_{\gamma'}(t)))|\left|\frac{\xi_\gamma(t)}{|\xi_\gamma(t)|}\right| + V(x_{\gamma'}(t))\left|\frac{\xi_\gamma(t)}{|\xi_\gamma(t)|} - \frac{\xi_{\gamma'}(t)}{|\xi_{\gamma'}(t)|}\right| \\ &\leq \|V\|_{C^1}|v(t)| + \|V\|_{C^0}\frac{2|w(t)|}{|\xi_\gamma(t)|} \\ &\leq \|V\|_{C^1}|v(t)| + \frac{2\|V\|_{C^0}}{\lambda_\gamma C_1}|w(t)|, \end{aligned}$$

where we have used the inequality $|\frac{x}{|x|} - \frac{y}{|y|}| = |\frac{x|y|-y|x|}{|x|\cdot|y|}| = |\frac{x(|y|-|x|)+|x|(x-y)}{|x|\cdot|y|}| \leq \frac{2|x-y|}{|y|}$ for all $x, y \neq 0$. By the symmetry between γ and γ', we also have

$$\left|\frac{dv(t)}{dt}\right| \leq \|V\|_{C^1}|v(t)| + \frac{2\|V\|_{C^0}}{\lambda_{\gamma'} C_1}|w(t)|.$$

Thus

$$\begin{aligned} \left|\frac{dv(t)}{dt}\right| &\leq \|V\|_{C^1}|v(t)| + 2\|V\|_{C^0}|w(t)| \min\left\{\frac{1}{\lambda_\gamma C_1}, \frac{1}{\lambda_{\gamma'} C_1}\right\} \\ &\leq \|V\|_{C^1}|v(t)| + \frac{2\|V\|_{C^0}}{C_1}\frac{1}{\sqrt{\lambda_\gamma\lambda_{\gamma'}}}|w(t)|. \end{aligned}$$

It follows that

$$\left|\frac{d(\sqrt{\lambda_\gamma\lambda_{\gamma'}}v(t))}{dt}\right| \leq \|V\|_{C^1}|\sqrt{\lambda_\gamma\lambda_{\gamma'}}v(t)| + \frac{2\|V\|_{C^0}}{C_1}|w(t)|.$$

Using the fact that $\frac{d|f(t)|^2}{dt} = 2\frac{df(t)}{dt} \cdot f(t) \leq 2|\frac{df(t)}{dt}| \cdot |f(t)|$ and the Cauchy-Schwartz inequality, we further get

$$\frac{d(\lambda_\gamma \lambda_{\gamma'} |v(t)|^2)}{dt} \leq \left(2\|V\|_{C^1} + \frac{4\|V\|_{C^0}^2}{C_1^2}\right) \lambda_\gamma \lambda_{\gamma'} |v(t)|^2 + |w(t)|^2. \tag{42}$$

We next estimate $|\frac{dw(t)}{dt}|$,

$$\begin{aligned} \left|\frac{dw(t)}{dt}\right| &= \big|V_x(x_\gamma(t))|\xi_\gamma(t)| - V_x(x_{\gamma'}(t))|\xi_{\gamma'}(t)|\big| \\ &\leq |V_x(x_\gamma(t)) - V_x(x_{\gamma'}(t))| \cdot |\xi_\gamma(t)| + |V_x(x_{\gamma'}(t))(|\xi_\gamma(t)| - |\xi_{\gamma'}(t)|)| \\ &\leq \|V\|_{C^2}|v(t)| \cdot \lambda_\gamma C_2 + \|V\|_{C^1}|w(t)| \\ &= C_2\|V\|_{C^2}|\lambda_\gamma v(t)| + \|V\|_{C^1}|w(t)|. \end{aligned}$$

The symmetry between γ, γ' yields

$$\left|\frac{dw(t)}{dt}\right| \leq C_2\|V\|_{C^2}|\lambda_{\gamma'} v(t)| + \|V\|_{C^1}|w(t)|.$$

Thus

$$\begin{aligned} \left|\frac{dw(t)}{dt}\right| &\leq C_2\|V\|_{C^2} \cdot \min\{\lambda_\gamma, \lambda_{\gamma'}\} \cdot |v(t)| + \|V\|_{C^1}|w(t)| \\ &\leq C_2\|V\|_{C^2}\sqrt{\lambda_\gamma \lambda_{\gamma'}}|v(t)| + \|V\|_{C^1}|w(t)|. \end{aligned}$$

It follows that

$$\frac{d|w(t)|^2}{dt} \leq (2\|V\|_{C^1} + \|V\|_{C^2}^2 C_2^2) \cdot |w(t)|^2 + \lambda_\gamma \lambda_{\gamma'} |v(t)|^2. \tag{43}$$

Define $E(t) = \lambda_\gamma \lambda_{\gamma'} |v(t)|^2 + |w(t)|^2$. Then (42) and (43) yield

$$\frac{dE}{dt} \leq \left(1 + 2\|V\|_{C^1} + \|V\|_{C^2}^2 C_2^2 + \frac{4\|V\|_{C^0}^2}{C_1^2}\right) E(t).$$

A direct application of Gronwall's inequality gives the second inequality in (41). A similar consideration for $\tilde{v}(t) = v(T-t)$ and $\tilde{w}(t) = w(T-t)$ yields the first inequality in (41). □

We remark that the inequality in Lemma 3.2 stated in a different form has also been proved in [29] by a different argument.

Lemma 3.3. *For a fixed constant $T > 0$, the following estimate holds,*

$$|x_\gamma(t) - x_{\gamma'}(t)|^2 \geq e^{-T(1+2\|V\|_{C^1})} \cdot |x_\gamma(0) - x_{\gamma'}(0)|^2 - 4T \cdot \|V\|_{C^0} \tag{44}$$

Proof. Denote $v(t) = x_\gamma(t) - x_{\gamma'}(t)$. By direct calculation

$$\begin{aligned}\left|\frac{dv(t)}{dt}\right| &= \left|V(x_\gamma(t))\frac{p_\gamma(t)}{|p_\gamma(t)|} - V(x_{\gamma'}(t))\frac{p_{\gamma'}(t)}{|p_{\gamma'}(t)|}\right| \\ &\leq |V(x_\gamma(t)) - V(x_{\gamma'}(t))|\left|\frac{p_\gamma(t)}{|p_\gamma(t)|}\right| + V(x_{\gamma'}(t))\left|\frac{p_\gamma(t)}{|p_\gamma(t)|} - \frac{p_{\gamma'}(t)}{|p_{\gamma'}(t)|}\right| \\ &\leq \|V\|_{C^1}|v(t)| + 2\|V\|_{C^0}.\end{aligned}$$

Using the fact that $\frac{d|v(t)|^2}{dt} = 2\frac{dv}{dt}\cdot v(t) \geq -2|\frac{dv(t)}{dt}|\cdot|v(t)|$ and the Cauchy-Schwartz inequality, we further get

$$\frac{d|v(t)|^2}{dt} \geq -(1 + 2\|V\|_{C^1})\cdot|v(t)|^2 - 4\|V\|_{C^0}^2.$$

By Gronwall's inequality, (44) follows. □

The following result is implied in a theorem shown in [15, page 101].

Lemma 3.4 [15]. *For any given $T > 0$, the solution $M_\gamma^\pm(t)$ for the Riccati equation* (31) *is well defined for $0 \leq t \leq T$. Furthermore, there exist positive constants c_0, c_1, and c_2 which depend on T and the velocity field $V(x)$, such that $c_0 \leq Im\{M_\gamma^\pm(t)\} \leq c_1$ and $\|Re\{M_\gamma^\pm(t)\}\| \leq c_2$ uniformly for all $0 \leq t \leq T$ and γ.*

Next we present two lemmas about the phase function τ_γ and the single beam Φ_γ.

Lemma 3.5. *Assume that Lemma* 3.4 *holds.*

$$\tau_{\gamma,t}(t,x) = -G(x, \tau_{\gamma,x}(t,x)) + O(|x - x_\gamma(t)|^3).$$

Proof. We suppress the index γ. Direct calculation yields

$$\begin{aligned}\tau_t(t,x) &= \frac{dp}{dt}\cdot(x - x(t)) + p(t)\cdot\left(-\frac{dx}{dt}\right) \\ &\quad + \frac{1}{2}(x - x(t))^T\frac{dM}{dt}(x - x(t)) - (x - x(t))^T M(t)\frac{dx}{dt} \\ &= -G_x(x(t), p(t))\cdot(x - x(t)) - p(t)\cdot G_p(x(t), p(t)) \\ &\quad + \frac{1}{2}(x - x(t))^T\cdot(-G_{xp}^T\cdot M^T - M\cdot G_{px} - M^T G_{pp}M - G_{xx})\cdot(x - x(t)), \\ \tau_x(t,x) &= p(t) + M(t)(x - x(t)).\end{aligned} \tag{45}$$

On the other hand, by using the Taylor expansion for G around $x = x(t)$ up to the third order term, we have

$$\begin{aligned}G(x, \tau_x(t,x)) &= G(x, p(t) + M(t)\cdot(x - x(t)) \\ &= G(x(t) + x - x(t), p(t) + M(t)(x - x(t)))\end{aligned}$$

$$
\begin{aligned}
&= G(x(t), p(t)) + G_x \cdot (x - x(t)) + G_p^T M(t)(x - x(t)) \\
&\quad + \frac{1}{2}(x - x(t))^T G_{xx}(x - x(t)) \\
&\quad + \frac{1}{2}(x - x(t))^T M(t)^T G_{pp} M(t)(x - x(t)) \\
&\quad + (x - x(t))^T G_{xp} M(t)(x - x(t)) \\
&\quad + \sum_{|\alpha|+|\beta|=3} \frac{1}{6}\int_0^1 \frac{\partial^3 G}{\partial^\alpha x \partial^\beta p}(x(t) + s(x - x(t)), p(t) + sM(t)(x - x(t))) \\
&\quad \times (x - x(t))^\alpha (M(t)(x - x(t)))^\beta ds. \qquad (46)
\end{aligned}
$$

We may write the last term as $\sum_{|\alpha|=3}(x - x(t))^\alpha F_\alpha(t, x)$. It is clear that

$$\|F_\alpha(t, \cdot)\|_{C^m(\mathbb{R}^d)} \lesssim \|V\|_{C^{m+3}(\mathbb{R}^d)},$$

provided that the norm of the matrix $M(t)$ is bounded.

Summing up (45) and (46), we have

$$
\begin{aligned}
\tau_t(t, x) + G(x, \tau_x(t, x)) &= -p(t)^T \cdot G_p(x(t), p(t)) + G(x(t), p(t)) \\
&\quad + \sum_{|\alpha|=3} (x - x(t))^\alpha F_\alpha(t, x).
\end{aligned}
$$

By the homogeneity of G, $p^T \cdot G_p(x, p) = G(x(t), p(t))$. It follows that

$$\tau_t + G(x, \tau_x(t, x) = \sum_{|\alpha|=3} (x - x(t))^\alpha F_\alpha(t, x). \qquad \square$$

Lemma 3.6.

$$
\begin{aligned}
\Phi_{\gamma,t}(t, x) &= -\Phi_\gamma(t, x) \cdot \lambda_\gamma \cdot G(x_\gamma(t), p_\gamma(t)) + \Phi_\gamma(t, x) \cdot D_\gamma(t) \\
&\quad - \sqrt{-1} \cdot \Phi_\gamma(t, x) \cdot \lambda_\gamma \cdot O(|x - x_\gamma(t)|) + \sqrt{-1} \cdot \lambda_\gamma \Phi_\gamma(t, x) \cdot O(|x - x_\gamma(t)|^3),
\end{aligned}
$$

where $D_\gamma(t)$ is defined as

$$
\begin{aligned}
D_\gamma(t) &= \frac{\dot{A}_\gamma(t)}{A_\gamma(t)} \\
&= -\frac{\begin{array}{l} V^2(x_\gamma(t))\text{trace}(M_\gamma(t)) - G_p(x_\gamma(t), p_\gamma(t)) \cdot (G_x(x_\gamma(t), p_\gamma(t)) \\ \quad + M_\gamma(t)\ G_p(x_\gamma(t), p_\gamma(t))) \end{array}}{2G(x_\gamma(t), p_\gamma(t))}.
\end{aligned}
$$

Proof. Suppressing the index γ, we have

$$
\begin{aligned}
\Phi_t(t, x) &= \Phi(t, x)\left(\frac{\dot{A}(t)}{A(t)} + \sqrt{-1} \cdot \lambda \cdot \tau_t(t, x)\right) \\
&= \Phi(t, x)\left(D(t) + \sqrt{-1} \cdot \lambda \cdot (-G(x, \tau_x(t, x)) + O(|x - x(t)|^3))\right)
\end{aligned}
$$

$$= \Phi(t,x)\Big(D(t) + \sqrt{-1}\cdot\lambda\cdot(-G(x,p(t)+M(t)\cdot(x-x(t)))$$
$$+ O(|x-x(t)|^3))\Big).$$

A Taylor expansion of the function $G(x, p(t)+M(t)\cdot(x-x(t)))$ around $x = x(t)$ yields

$$G(x, p(t)+M(t)(x-x(t))) = G(x(t), p(t)) + O(|x-x(t)|).$$

Since $\lambda \cdot G(x(t), p(t)) = G(x(t), \xi(t))$, we have

$$\lambda\cdot G(x, p(t)+M(t)\cdot(x-x(t))) = G(x(t),\xi(t)) + \lambda\cdot O(|x-x(t)|).$$

The lemma follows. □

3.2. *Wavepacket Interaction and Beam Decaying*

The following lemma describes interaction between Gaussian wavepackets which plays an important role in the proof of Lemmas 3.8 and 3.9.

Lemma 3.7. *Assume that M_1 and M_2 are two symmetric and positive definite matrices satisfying $0 < c_0 I < M_1, M_2 < c_1 I$, and that N_1 and N_2 are two symmetric matrices such that $\|N_1\|, \|N_2\| \le c_2$. Assume also that δx and $\delta\xi$ are two vectors in $\mathbb{R}^d$, and λ_1 and λ_2 are two positive numbers. Let $c_1^* = \frac{32 c_1^3 c_2^2}{c_0^2}$. Then*

$$\left|\int_{\mathbb{R}^d} x^\alpha\cdot(x-\delta x)^\beta\cdot e^{\sqrt{-1}\delta\xi\cdot x-\lambda_1 x^T(M_1+\sqrt{-1}\cdot N_1)x-\lambda_2(x-\delta x)^T(M_2+\sqrt{-1}\cdot N_2)(x-\delta x)}\,dx\right|$$
$$\lesssim_{c_0} \frac{1}{(\lambda_1+\lambda_2)^{\frac{d}{2}}} e^{-\frac{c_0\lambda_1\lambda_2}{4(\lambda_1+\lambda_2)}|\delta x|^2-\frac{|\delta\xi|^2}{c_1^*(\lambda_1+\lambda_2)}}\cdot\frac{1}{\lambda_1^{\frac{|\alpha|}{2}}\cdot\lambda_2^{\frac{|\beta|}{2}}}.$$

Proof. See Subsection 4.3. □

We remark that the special case of Lemma 3.7 corresponding to $\alpha = \beta = 0$, $N_1 = N_2 = 0$, and $M_1 = M_2 = I$ can be evaluated directly as done in [29].

Furthermore, the following two lemmas show that different Gaussian beams are "almost orthogonal" to some extent.

Lemma 3.8. *Assume that $\Phi_\gamma(t,x)$ is defined as in (35), and b_γ are complex numbers with $\sum_\gamma |b_\gamma|^2 < \infty$ and $b_\gamma = 0$ for all $\gamma = (\ell, i, k)$ with $\ell \le 1$. Let m be a non-negative integer. If $|\alpha(\gamma)| \ge m$ for all γ, then*

$$\left\|\sum_\gamma \lambda_\gamma^{\frac{m}{2}} b_\gamma \Phi_\gamma(t,\cdot)\cdot(\cdot - x_\gamma(t))^{\alpha(\gamma)}\right\|^2_{L^2(\mathbb{R}^d)} \lesssim \sum_\gamma |b_\gamma|^2,$$

uniformly for $0 \le t \le T$, where T is given.

Proof. See Subsection 4.4. □

In Lemma 3.8, we have characterized interactions between propagated multiscale Gaussian wavepackets along bicharacteristics in terms of various vanishing orders, and these vanishing orders in turn define the orders of the accumulated interaction. A similar result has been established in [29] (Theorem 2.1 in [29]) for the interaction between propagated Gaussian wavepackets with Gaussian frame functions, and the proof relies on some specific properties of the frame constructed there, such as the specific form of the scale factor inherent in that frame.

Lemma 3.9. *Let $\Phi_\gamma(t, x)$ be defined as in* (35); *let b_γ be complex numbers such that $\sum_\gamma |b_\gamma|^2 < \infty$ and $b_\gamma = 0$ for all $\gamma = (\ell, i, k)$ such that $\ell \leq 1$. Assume that $F_\gamma(t, x) = O(|x - x_\gamma(t)|^m)$ for each γ. Then*

$$\left\| \sum_\gamma \lambda_\gamma^{\frac{m}{2}} b_\gamma \Phi_\gamma(t, \cdot) \cdot F_\gamma(t, \cdot) \right\|^2_{L^2(\mathbb{R}^d)} \lesssim \sum_\gamma |b_\gamma|^2,$$

uniformly for $0 \leq t \leq T$, where T is given.

Proof. Taylor expanding $F_\gamma(t, x)$ about $x = x(t)$ up to the N_0-th order, where N_0 is to be determined later, we can write

$$F_\gamma(t, x) = \sum_{|\alpha|=m}^{N_0} d_{\gamma,\alpha}(x - x(t))^\alpha + \sum_{|\alpha|=N_0+1} F_{\gamma,\alpha}(t, x)(x - x(t))^\alpha \equiv F_\gamma^{(1)}(t, x) + F_\gamma^{(2)}(t, x),$$

where all $d_{\gamma,\alpha}$ are uniformly bounded complex numbers and $F_{\gamma,\alpha}(t, x) \lesssim O(1)$.

By Lemma 3.8, the following holds:

$$\left\| \sum_\gamma \lambda_\gamma^{\frac{m}{2}} b_\gamma \Phi_\gamma(t, \cdot) \cdot F_\gamma^{(1)}(t, \cdot) \right\|^2_{L^2(\mathbb{R}^d)} \lesssim \sum_\gamma |b_\gamma|^2. \tag{47}$$

Next we show that

$$\left\| \sum_\gamma \lambda_\gamma^{\frac{m}{2}} b_\gamma \Phi_\gamma(t, \cdot) \cdot F_\gamma^{(2)}(t, \cdot) \right\|^2_{L^2(\mathbb{R}^d)} \lesssim \sum_\gamma |b_\gamma|^2. \tag{48}$$

To see this, we first apply Lemma 4.6 to get

$$\begin{aligned} &| < \lambda_\gamma^{\frac{m}{2}} \Phi_\gamma F_\gamma^{(2)}(t, \cdot), \lambda_{\gamma'}^{\frac{m}{2}} \Phi_{\gamma'} F_{\gamma'}^{(2)}(t, \cdot) > | \\ &\lesssim \frac{(\lambda_\gamma \lambda_{\gamma'})^{\frac{d}{4}}}{(\lambda_\gamma + \lambda_{\gamma'})^{\frac{d}{2}}} \cdot e^{-\frac{c_0 \lambda_\gamma \lambda_{\gamma'}}{2(\lambda_\gamma + \lambda_{\gamma'})} |x_\gamma(t) - x_{\gamma'}(t)|^2} \cdot \frac{1}{\lambda_\gamma^{\frac{|\alpha(\gamma)|-m}{2}} \cdot \lambda_{\gamma'}^{\frac{|\alpha(\gamma')|-m}{2}}}. \end{aligned}$$

Since $\lambda_\gamma > 1$ for those γ with $\ell > 1$, we further get

$$\begin{aligned} | < \lambda_\gamma^{\frac{m}{2}} \Phi_\gamma F_\gamma^{(2)}(t, \cdot), \lambda_{\gamma'}^{\frac{m}{2}} \Phi_{\gamma'} F_{\gamma'}^{(2)}(t, \cdot) > | &\lesssim e^{-\frac{c_0}{4} |x_\gamma(t) - x_{\gamma'}(t)|^2} \cdot \lambda_\gamma^{\frac{-N_0-1+m}{2}} \cdot \lambda_{\gamma'}^{\frac{-N_0-1+m}{2}} \\ &\lesssim e^{-\frac{c_0}{4} |x_\gamma(t) - x_{\gamma'}(t)|^2} \cdot \lambda_\gamma^{-d-1} \cdot \lambda_{\gamma'}^{-d-1}, \end{aligned}$$

where we have taken $N_0 = 2d + 1 + m$.

Define

$$d_{\gamma,\gamma'} = e^{-\frac{c_0}{4}|x_\gamma(t)-x_{\gamma'}(t)|^2} \cdot \lambda_\gamma^{-d-1}\lambda_{\gamma'}^{-d-1}.$$

Similar to the proof of Lemma 3.8, we can show with the help of Lemma 3.3 that

$$d_{\gamma,\gamma'} \lesssim \min_{x\in D_\gamma, \xi\in B_\gamma, x'\in D_{\gamma'}, \xi'\in B_{\gamma'}} e^{-c_0'|x-x'|^2} \cdot |\xi|^{-d-1}|\xi'|^{-d-1} \tag{49}$$

for some properly chosen $c_0' > 0$, where D_γ and B_γ are defined in Subsection 2.2.

Defining function $b(x, \xi)$ by letting $b(x, \xi) = b_\gamma$ for all $x \in D_\gamma$, $\xi \in B_\gamma$, we have

$$\begin{aligned}\sum_{\gamma,\gamma'} |b_\gamma b_{\gamma'} d_{\gamma,\gamma'}| &\lesssim \int_{\mathbb{R}^{4d}} |b(x,\xi)b(x',\xi')|e^{-c_0'|x-x'|^2} \cdot |\xi|^{-d-1}|\xi'|^{-d-1} dx d\xi dx' d\xi' \\ &\le \int_{\mathbb{R}^{4d}} (|b(x,\xi)|^2 + |b(x',\xi')|^2) e^{-c_0'|x-x'|^2} \cdot |\xi|^{-d-1}|\xi'|^{-d-1} dx d\xi dx' d\xi' \\ &= 2\int_{\mathbb{R}^{4d}} |b(x,\xi)|^2 e^{-c_0'|x-x'|^2} \cdot |\xi|^{-d-1}|\xi'|^{-d-1}\, dx' d\xi' dx d\xi \\ &\lesssim \int_{\mathbb{R}^{2d}} |b(x,\xi)|^2 |\xi|^{-d-1}\, dx d\xi \\ &\le \int_{\mathbb{R}^{2d}} |b(x,\xi)|^2\, dx d\xi \\ &= \frac{1}{2^d}\sum_\gamma |b_\gamma|^2. \end{aligned} \tag{50}$$

This proves (48). Combining this with (47) yields the lemma. □

3.3. *Main Convergence Results*

With Lemmas 3.8 and 3.9 at our disposal, we are ready to estimate the error between the Gaussian beam solution and the exact solution. We first estimate the error that comes from approximating the Cauchy data.

Lemma 3.10. *Assume that* $U(0)$ *and* $U_t(0)$ *are defined as in* (27) *and* (28). *Assume also that* c_γ^+, c_γ^- *and* $\widetilde{U}(t, x)$ *are defined as in* (37), (38), *and* (36). *Then*

(a). $\sum_\gamma \lambda_\gamma^2(|c_\gamma^+|^2 + |c_\gamma^-|^2) \lesssim \sum_\gamma(\lambda_\gamma^2|a_\gamma|^2 + |b_\gamma|^2)$;
(b). $\widetilde{U}(0) = U(0)$;
(c). $\|\widetilde{U}_t(0) - U_t(0)\|_{L^2(\mathbb{R}^d)} \lesssim \frac{1}{\lambda_{\min}^{\frac{1}{2}}}\left(\sum_\gamma(\lambda_\gamma^2|a_\gamma|^2 + |b_\gamma|^2)\right)^{\frac{1}{2}}$,
where $a_\gamma, b_\gamma = 0$ *for all* γ *such that* $\lambda_\gamma \le \lambda_{\min}$ *with* $\lambda_{\min} > 4$.

Proof. We first show (a). Using (37), we have

$$\begin{aligned}\sum_\gamma \lambda_\gamma^2 |c_\gamma^+|^2 &\le \sum_\gamma \lambda_\gamma^2 \cdot \left(|a_\gamma|^2 + \frac{|b_\gamma|^2}{V^2(\frac{k}{L_\gamma})(2\pi\lambda_\gamma)^2}\right) \\ &= \sum_\gamma \lambda_\gamma^2 \cdot |a_\gamma|^2 + \sum_\gamma \frac{|b_\gamma|^2}{V^2(\frac{k}{L_\gamma})(2\pi)^2}.\end{aligned}$$

Since $V(x)$ is bounded away from zero, we get

$$\sum_\gamma \lambda_\gamma^2 \cdot |c_\gamma^+|^2 \lesssim \sum_\gamma (\lambda_\gamma^2 |a_\gamma|^2 + |b_\gamma|^2).$$

Similarly, it is easy to show that

$$\sum_\gamma \lambda_\gamma^2 \cdot |c_\gamma^-|^2 \lesssim \sum_\gamma (\lambda_\gamma^2 |a_\gamma|^2 + |b_\gamma|^2),$$

which completes the proof of (a).

The proof of (b) follows from the fact that $\Phi_\gamma^+(0, x) = \Phi_\gamma^-(0, x)$.

Next, we prove (c). By Lemma 3.6, we have

$$\begin{aligned}\widetilde{U}_t(0, x) - U_t(0, x) = &\sum_\gamma c_\gamma^+ \Phi_\gamma^+(0, x) D_\gamma^+(0) - \sqrt{-1} \cdot \sum_\gamma \lambda_\gamma c_\gamma^+ \Phi_\gamma^+(0, x) \cdot O(|x - x_\gamma(0)|) \\ &+ \sqrt{-1} \cdot \sum_\gamma \lambda_\gamma c_\gamma^+ \Phi_\gamma^+(0, x) \cdot O(|x - x_\gamma(0)|^3) \\ &+ \sum_\gamma c_\gamma^- \Phi_\gamma^-(0, x) D_\gamma^-(0) \\ &- \sqrt{-1} \cdot \sum_\gamma \lambda_\gamma c_\gamma^- \Phi_\gamma^-(0, x) \cdot O(|x - x_\gamma(0)|) \\ &+ \sqrt{-1} \cdot \sum_\gamma \lambda_\gamma c_\gamma^- \Phi_\gamma^-(0, x) \cdot O(|x - x_\gamma(0)|^3). \end{aligned} \tag{51}$$

By Lemma 3.8 and Lemma 3.9, we have

$$\left\| \sum_\gamma c_\gamma^+ \Phi_\gamma^+(0, \cdot) D_\gamma^+(0) \right\|_{L^2(\mathbb{R}^d)}^2 \lesssim \sum_\gamma |c_\gamma^+|^2 \le \frac{1}{\lambda_{\min}^2} \sum_\gamma \lambda_\gamma^2 \cdot |c_\gamma^+|^2,$$

$$\left\| \sum_\gamma \lambda_\gamma c_\gamma^+ \Phi_\gamma^+(0, \cdot) \cdot O(| \cdot - x_\gamma(0)|) \right\|_{L^2(\mathbb{R}^d)}^2 \lesssim \frac{1}{\lambda_{\min}} \sum_\gamma \lambda_\gamma^2 \cdot |c_\gamma^+|^2,$$

$$\left\| \sum_\gamma \lambda_\gamma c_\gamma^+ \Phi_\gamma^+(0, \cdot) \cdot O(| \cdot - x_\gamma(0)|^3) \right\|_{L^2(\mathbb{R}^d)}^2 \lesssim \frac{1}{\lambda_{\min}^3} \sum_\gamma \lambda_\gamma^2 \cdot |c_\gamma^+|^2.$$

The same results hold for other terms in (51) with sup-script "-". Using part (a), (c) follows. □

Now, we estimate the error that comes from beam propagation.

Lemma 3.11. *Assume that $T > 0$ is fixed and $\lambda_{\min} > 4$. Assume also that c_γ are complex numbers such that $\sum_\gamma |c_\gamma|^2 \lambda_\gamma^2 < \infty$ and $c_\gamma = 0$ for all $\lambda_\gamma < \lambda_{\min}$. Define $u(t, x) = \sum_\gamma c_\gamma \Phi_\gamma(t, x)$. Then*

$$\|\mathbf{P}u\|_{L^2(\mathbb{R}^d)}^2 \lesssim \frac{1}{\lambda_{\min}} \sum_\gamma \lambda_\gamma^2 |c_\gamma|^2, \quad \textit{uniformly for all } 0 \le t \le T,$$

where $\mathbf{P}u \equiv (\partial_{tt} - V^2(x)\Delta)u$.

Proof. By a direct calculation, we have

$$\begin{aligned}
\mathbf{P}\Phi_\gamma(t,x) &= (\partial_{tt} - V^2(x)\Delta)A_\gamma(t)e^{\sqrt{-1}\cdot\lambda_\gamma\tau_\gamma(t,x)} \\
&= -\lambda_\gamma^2(\tau_{\gamma,t}^2(t,x) - V^2(x)\tau_{\gamma,x}^2(t,x))\cdot A_\gamma(t)e^{\sqrt{-1}\cdot\lambda_\gamma\tau_\gamma(t,x)} \\
&\quad + \sqrt{-1}\cdot\lambda_\gamma\Big\{2\tau_{\gamma,t}(t,x)A_{\gamma,t}(t) \\
&\quad + A_\gamma(t)(\tau_{\gamma,tt} - V^2(x)\mathrm{trace}(\tau_{\gamma,xx}(t,x)))\Big\}e^{\sqrt{-1}\cdot\lambda_\gamma\tau_\gamma(t,x)} \\
&= -\lambda_\gamma^2\cdot\Phi_\gamma(t,x)\cdot\{\tau_{\gamma,t}^2(t,x) - V^2(x)\tau_{\gamma,x}^2(t,x)\} \\
&\quad + \sqrt{-1}\cdot\lambda_\gamma\cdot\Phi_\gamma(t,x)\Big\{2\tau_{\gamma,t}(t,x)\frac{A_{\gamma,t}(t)}{A_\gamma(t)} \\
&\quad + \tau_{\gamma,tt}(t,x) - V^2(x)\mathrm{trace}(\tau_{\gamma,xx}(t,x))\Big\}.
\end{aligned}$$

Let

$$\begin{aligned}
g_{\gamma,1}(t,x) &= \tau_{\gamma,t}^2(t,x) - V^2(x)\tau_{\gamma,x}^2(t,x), \\
g_{\gamma,2}(t,x) &= 2\tau_{\gamma,t}(t,x)\frac{A_{\gamma,t}(t)}{A_\gamma(t)} + \tau_{\gamma,tt}(t,x) - V^2(x)\mathrm{trace}(\tau_{\gamma,xx}(t,x)).
\end{aligned}$$

Then

$$\mathbf{P}\Phi_\gamma(t,x) = -\lambda_\gamma^2\cdot\Phi_\gamma(t,x)\cdot g_{\gamma,1}(t,x) + \sqrt{-1}\cdot\lambda_\gamma\cdot\Phi_\gamma(t,x)\cdot g_{\gamma,2}(t,x).$$

By Lemma 3.5, we have

$$g_{\gamma,1}(t,x) = O(|x - x(t)|^3).$$

Then Lemma 3.9 yields

$$\begin{aligned}
\Big\|\sum_\gamma c_\gamma\lambda_\gamma^2\cdot\Phi_\gamma(t,\cdot)\cdot g_{\gamma,1}(t,\cdot)\Big\|_{L^2(\mathbb{R}^d)}^2 &= \Big\|\sum_\gamma\Big(c_\gamma\lambda_\gamma^{\frac{1}{2}}\Big)\cdot\lambda_\gamma^{\frac{3}{2}}\cdot\Phi_\gamma(t,\cdot)\cdot g_{\gamma,1}(t,\cdot)\Big\|_{L^2(\mathbb{R}^d)}^2 \\
&\lesssim \sum_\gamma\Big(c_\gamma\lambda_\gamma^{\frac{1}{2}}\Big)^2 \le \frac{1}{\lambda_{\min}}\sum_\gamma\lambda_\gamma^2|c_\gamma|^2. \qquad (52)
\end{aligned}$$

We can also show that

$$\Big\|\sum_\gamma c_\gamma\lambda_\gamma\cdot\Phi_\gamma(t,\cdot)\cdot g_{\gamma,2}(t,\cdot)\Big\|_{L^2(\mathbb{R}^d)}^2 \le \frac{1}{\lambda_{\min}}\sum_\gamma\lambda_\gamma^2|c_\gamma|^2. \qquad (53)$$

Indeed, using Lemma 3.9, we need only show that $g_{\gamma,2}(t,x) = O(|x - x_\gamma(t)|)$. This is done in the next lemma, namely Lemma 3.12. Thus we have proved (53). Combining this with (52) yields the theorem. □

Lemma 3.12.

$$g_{\gamma,2}(t,x) = 2\tau_{\gamma,t}(t,x)\frac{\dot{A}_\gamma(t)}{A_\gamma(t)} + \tau_{\gamma,tt}(t,x) - V^2(x)\mathrm{trace}(\tau_{\gamma,xx}(t,x)) = O(|x - x_\gamma(t)|).$$

Proof. We suppress the index γ. By Lemma 3.5, we have

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

Furthermore, direct calculation shows that

$$\begin{aligned}\tau_{tt}(t,x) &= -G_p(x, p(t) + M(t)(x - x(t)))\{\dot{p}(t) + \dot{M}(t)(x - x(t)) + M(t)(-\dot{x}(t))\}\\ &\quad + O(|x - x(t)|^2),\\ \tau_{xx}(t,x) &= M(t).\end{aligned}$$

Thus

$$\begin{aligned}g_2(t,x) &= 2\tau_t(t,x)\frac{\dot{A}(t)}{A(t)} + \tau_{tt}(t,x) - V^2(x)\mathrm{trace}(M(t))\\ &= -2\frac{\dot{A}(t)}{A(t)} \cdot G(x, \tau_x(t,x))\\ &\quad - G_p(x, p(t) + M(t)(x - x(t))) \cdot \{\dot{p}(t) + \dot{M}(t)(x - x(t)) + M(t)(-\dot{x}(t))\}\\ &\quad + O(|x - x(t)|^2) - V^2(x)\mathrm{trace}(M(t)) + O(|x - x(t)|^3).\end{aligned} \tag{54}$$

When $x = x(t)$, we have

$$g_2(t, x(t)) = 2\tau_t(t, x(t))\frac{\dot{A}(t)}{A(t)} + \tau_{tt}(t, x(t)) - V^2(x(t))\mathrm{trace}(M(t)). \tag{55}$$

Substituting the equalities

$$\begin{aligned}2\frac{\dot{A}(t)}{A(t)} &= -\frac{V^2(x(t))\mathrm{trace}M(t) - G_p \cdot G_x - G_p^T(x(t), p(t))M(t)G_p(x(t), p(t))}{G(x(t), p(t))},\\ \tau_t(t, x(t)) &= -G(x(t), p(t)),\\ \tau_{tt}(t,x) &= -G_p(x(t), p(t)) \cdot \dot{p}(t)) + G_p^T(x(t), p(t))M(t)G_p(x(t), p(t))\\ &= G_p(x(t), p(t)) \cdot G_x(x(t), p(t)) + G_p^T(x(t), p(t))M(t)G_p(x(t), p(t))\end{aligned}$$

into (55), we get

$$g_2(t, x(t)) = 0.$$

Using Taylor expansion about $x = x(t)$ for the function $g_2(t,x)$ with expression (54), the result $g_2(t,x) = O(|x - x(t)|)$ follows. □

Before we state the main result, we first recall the following estimate from [16].

Theorem 3.1 [16]. *Suppose that* $f_1 \in H^1(\mathbb{R}^d)$, $f_2 \in L^2(\mathbb{R}^d)$, *and* $f \in L^\infty(0, T; L^2(\mathbb{R}^d))$. *Then there exists an unique solution* $u \in C^1(0, T; L^2(\mathbb{R}^d)) \bigcap C(0, T; H^1(\mathbb{R}^d))$ *to the following wave equation*

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

Furthermore, the following estimate holds,

$$\|u\|_{C^1(0,T;L^2(\mathbb{R}^d))\bigcap C(0,T;H^1(\mathbb{R}^d))} \lesssim \|f_1\|_{H^1(\mathbb{R}^d)} + \|f_2\|_{L^2(\mathbb{R}^d)} + \|f\|_{L^\infty(0,T;L^2(\mathbb{R}^d))}.$$

Finally, we can state and prove our main result.

Theorem 3.2. *Consider the following wave equation*

$$\begin{aligned} U_{tt} - V^2(x)\Delta U &= 0, \quad x \in \mathbb{R}^d, \quad 0 < t < T, \\ U|_{t=0} &= \sum_\gamma a_\gamma \tilde{\phi}_\gamma(x), \\ U_t|_{t=0} &= \sum_\gamma b_\gamma \tilde{\phi}_\gamma(x). \end{aligned}$$

Assume that $\sum_\gamma(\lambda_\gamma^2|a_\gamma|^2 + |b_\gamma|^2) < \infty$, $\lambda_{\min} \gg 1$, *and* $a_\gamma = b_\gamma = 0$ *for all* γ *such that* $\lambda_\gamma \leq \lambda_{\min}$. *Let* c_γ^+ *and* c_γ^- *be defined as in* (37) *and* (38). *Define the Gaussian beam parametrix by*

$$\widetilde{U}(t, x) = \sum_\gamma c_\gamma^+ \Phi_\gamma^+(t, x) + \sum_\gamma c_\gamma^- \Phi_\gamma^-(t, x).$$

Then

$$\widetilde{U} \in C^1(0, T; L^2(\mathbb{R}^d)) \bigcap C(0, T; H^1(\mathbb{R}^d)). \tag{56}$$

Furthermore, the following error estimate holds for the Gaussian beam solution:

$$\|\widetilde{U} - U\|_{C^1(0,T;L^2(\mathbb{R}^d))\bigcap C(0,T;H^1(\mathbb{R}^d))} \lesssim \frac{1}{\lambda_{\min}^{\frac{1}{2}}} \left(\sum_\gamma \left(\lambda_\gamma^2 |a_\gamma|^2 + |b_\gamma|^2\right) \right)^{\frac{1}{2}}, \tag{57}$$

where $U \in C^1(0, T; L^2(\mathbb{R}^d)) \bigcap C(0, T; H^1(\mathbb{R}^d))$ *is the exact solution.*

Proof. Denote $u(t, x) = \sum_\gamma c_\gamma \Phi_\gamma(t, x)$, where $c_\gamma \Phi_\gamma(t, x)$ can be either $c_\gamma^+ \Phi_\gamma^+(t, x)$ or $c_\gamma^- \Phi_\gamma^-(t, x)$. It is clear that we need only show the relation (56) for $u(t, x)$. Without loss of generality, we take $c_\gamma \Phi_\gamma(t, x) = c_\gamma^+ \Phi_\gamma^+(t, x)$. By direct calculation, we have

$$\begin{aligned} u_x(t, x) &= \sum_\gamma \lambda_\gamma c_\gamma \Phi_\gamma(t, x)(p_\gamma(t) + M_\gamma(t)(x - x_\gamma(t))) \\ &= \sum_\gamma \lambda_\gamma c_\gamma \Phi_\gamma(t, x) p_\gamma(t) + \sum_\gamma \lambda_\gamma c_\gamma \Phi_\gamma(t, x) M_\gamma(t)(x - x_\gamma(t)). \end{aligned}$$

Using the fact that both $|p_\gamma(t)|$ and $\|M_\gamma(t)\|$ are uniformly bounded for $0 \le t \le T$ (Lemmas 3.1 and 3.4), we can apply Lemma 3.9 to conclude that

$$\left\| \sum_\gamma \lambda_\gamma c_\gamma \Phi_\gamma(t, \cdot) p_\gamma(t) \right\| \lesssim \sum_\gamma \lambda_\gamma^2 |c_\gamma|^2,$$

$$\left\| \sum_\gamma \lambda_\gamma c_\gamma \Phi_\gamma(t, \cdot) M_\gamma(t)(\cdot - x_\gamma(t)) \right\| \lesssim \sum_\gamma \lambda_\gamma |c_\gamma|^2.$$

Thus the series representing u_x converges uniformly for $0 \le t \le T$. It follows that $u_x \in C(0, T; L^2(\mathbb{R}^d))$ and hence $u \in C(0, T; H^1(\mathbb{R}^d))$.

The fact that $u \in C^1(0, T; L^2(\mathbb{R}^d))$ follows from

$$\begin{aligned} u_t(t, x) = & -\sum_\gamma \Phi_\gamma(t, x) \cdot \lambda_\gamma c_\gamma \cdot G(x_\gamma(t), p_\gamma(t)) + \sum_\gamma c_\gamma \Phi_\gamma(t, x) \cdot D_\gamma(t) \\ & - \sum_\gamma \sqrt{-1} \cdot \Phi_\gamma(t, x) \cdot \lambda_\gamma c_\gamma \cdot O(|x - x_\gamma(t)|) \\ & + \sum_\gamma \sqrt{-1} \cdot \lambda_\gamma c_\gamma \Phi_\gamma(t, x) \cdot O(|x - x_\gamma(t)|^3). \end{aligned}$$

Since $G(x_\gamma(t), p_\gamma(t))$ and $\|D_\gamma(t)\|$ are uniformly bounded for $0 \le t \le T$, we can apply Lemma 3.8 and Lemma 3.9 to conclude that

$$\left\| \sum_\gamma \Phi_\gamma(t, \cdot) \cdot \lambda_\gamma c_\gamma \cdot G(x_\gamma(t), p_\gamma(t)) \right\|_{L^2(\mathbb{R}^d)}^2 \lesssim \sum_\gamma \lambda_\gamma^2 |c_\gamma|^2;$$

$$\left\| \sum_\gamma c_\gamma \Phi_\gamma(t, \cdot) \cdot D_\gamma(t) \right\|_{L^2(\mathbb{R}^d)}^2 \lesssim \sum_\gamma |c_\gamma|^2;$$

$$\left\| \sum_\gamma \Phi_\gamma(t, \cdot) \cdot \lambda_\gamma c_\gamma \cdot O(|\cdot - x_\gamma(t)|) \right\|_{L^2(\mathbb{R}^d)}^2 \lesssim \sum_\gamma \lambda_\gamma |c_\gamma|^2;$$

$$\left\| \sum_\gamma \lambda_\gamma c_\gamma \Phi_\gamma(t, \cdot) \cdot O(|\cdot - x_\gamma(t)|^3) \right\|_{L^2(\mathbb{R}^d)}^2 \lesssim \sum_\gamma \lambda_\gamma^{-1} |c_\gamma|^2.$$

Thus the series representing u_t converges uniformly for $0 \le t \le T$ and hence $u_t \in C(0, T; L^2(\mathbb{R}^d))$. Besides, it is also easy to see that $u \in C(0, T; L^2(\mathbb{R}^d))$. Therefore we have $u \in C^1(0, T; L^2(\mathbb{R}^d))$. (56) is proved.

Now we show (57). Let $w(t, x) = \widetilde{U}(t, x) - U(t, x)$. It follows that

$$\begin{aligned} w_{tt} - V^2(x)\Delta w &= -\mathbf{P}U(t, x), \quad x \in \mathbb{R}^d, \quad t > 0, \\ w(t = 0) &= 0, \\ w_t(t = 0) &= \widetilde{U}_t(0) - U_t(0). \end{aligned}$$

It follows from Lemma 3.10, Lemma 3.11, and Theorem 3.1 that

$$\|w\|_{C^1(0,T;L^2(\mathbb{R}^d)) \bigcap C(0,T;H^1(\mathbb{R}^d))} \lesssim \frac{1}{\lambda_{\min}^{\frac{1}{2}}} \left(\sum_\gamma \left(\lambda_\gamma^2 |a_\gamma|^2 + |b_\gamma|^2 \right) \right)^{\frac{1}{2}},$$

which completes the proof of (57). □

Theorem 3.2 combined with Lemmas 2.4 and 2.5 yields

Theorem 3.3. *Consider the following wave equation*

$$U_{tt} - V^2(x)\Delta U = 0, \quad x \in \mathbb{R}^d, \quad 0 < t < T,$$
$$U|_{t=0} = f_1(x) = \sum_\gamma a_\gamma \phi_\gamma(x),$$
$$U_t|_{t=0} = f_2(x) = \sum_\gamma b_\gamma \phi_\gamma(x).$$

Assume that $f_1 \in H^1(\mathbb{R}^d)$, $f_2 \in L^2(\mathbb{R}^d)$, $\lambda_{\min} \gg 1$, *and* $a_\gamma = b_\gamma = 0$ *for all* γ *such that* $\lambda_\gamma \le \lambda_{\min}$. *Let* c_γ^+ *and* c_γ^- *be defined as in* (37) *and* (38). *Define the multiscale Gaussian beam solution by*

$$\widetilde{U}(t,x) = \sum_\gamma c_\gamma^+ \Phi_\gamma^+(t,x) + \sum_\gamma c_\gamma^- \Phi_\gamma^-(t,x).$$

Then

$$\widetilde{U} \in C^1(0,T; L^2(\mathbb{R}^d)) \bigcap C(0,T; H^1(\mathbb{R}^d)). \tag{58}$$

Furthermore, the following error estimate holds for the multiscale Gaussian beam solution:

$$\|\widetilde{U} - U\|_{C^1(0,T;L^2(\mathbb{R}^d))\bigcap C(0,T;H^1(\mathbb{R}^d))} \lesssim \left(\frac{1}{\lambda_{\min}^{\frac{1}{2}}} + \varepsilon\right)\left(\|f_1\|_{H^1(\mathbb{R}^d)} + \|f_2\|_{L^2(\mathbb{R}^d)}\right), \tag{59}$$

where $U \in C^1(0,T; L^2(\mathbb{R}^d)) \bigcap C(0,T; H^1(\mathbb{R}^d))$ *is the exact solution and* ε *is determined in Lemmas* 2.4 *and* 2.5.

We conclude the paper with the following remarks.

Remark 3.1. The method and results in this work can also be applied to the general second order wave equation $U_{tt}(t,x) - \sum_{i,j=1}^d a_{i,j}(t,x)U_{x_i,x_j}(t,x) = 0$ with highly oscillatory Cauchy data, where $a_{i,j}$'s are assumed to be smooth functions and the matrix $\{a_{i,j}\}_{i,j=1}^d$ formed by $a_{i,j}$'s is assumed to be symmetric and uniformly positive definite.

Remark 3.2. The method and results may also be applied to the general first order wave equation $(D_t - a(x,D_x))U(t,x) = 0$ with highly oscillatory Cauchy data, where $a(x,D_x)$ is a first-order homogeneous pseudo-differential operator.

4. Proof of Technical Lemmas

4.1. *Some Nonstandard Inequalities*

Lemma 4.1. *Let A be a symmetric, positive-definite matrix in* $\mathbb{R}^d$, *and* c_0 *be a positive number. Then*

$$A \ge c_0 I \Leftrightarrow A^{-1} \le c_0^{-1} I.$$

Proof. Since A is symmetric and positive-definite, there exists an orthogonal basis in $\mathbb{R}^d$: $\tilde{e}_1, \ldots \tilde{e}_d$, and positive numbers $\lambda_1, \ldots \lambda_d$ such that

$$A = \sum_{j=1}^{d} \lambda_j \tilde{e}_j \tilde{e}_j^T.$$

Then

$$A \geq c_0 I \Leftrightarrow \min_{1 \leq j \leq d} \lambda_j \geq c_0 \Leftrightarrow \max_{1 \leq j \leq d} \lambda_j^{-1} \leq c_0^{-1}$$
$$\Leftrightarrow A^{-1} = \sum_{j=1}^{d} \lambda_j^{-1} \tilde{e}_j \tilde{e}_j^T \leq c_0^{-1} I. \qquad \square$$

Lemma 4.2. *Assume that $A = (a_{i,j})$ is a symmetric, positive-definite matrix in $\mathbb{R}^d$, and c_1 is a positive upper bound of A, i.e. $A \leq c_1 I$. Then $\|A\| = \sup_{x \neq 0} \frac{|Ax|}{|x|} \leq c_1$ and $|a_{i,j}| \leq c_1$.*

Proof. Since A is symmetric, we have

$$\|A\| = \sup_{x \neq 0} \frac{|Ax|}{|x|} = \sup_{x \neq 0} \frac{< Ax, x >}{< x, x >} \leq c_1,$$
$$|a_{i,j}| = |< Ae_i, e_j >| \leq |Ae_i| \leq \|A\| \leq c_1. \qquad \square$$

Lemma 4.3. *If $f(x) = e^{-\tau x^2} x^n$ with $x > 0$, then*

$$f(x) \lesssim \frac{1}{\tau^{\frac{n}{2}}}.$$

Proof. We have $f'(x) = e^{-\tau x^2} x^n (-2\tau x + \frac{n}{x})$. For $0 \leq x \leq \sqrt{\frac{n}{2\tau}}$, $f'(x) > 0$; for $x \geq \sqrt{\frac{n}{2\tau}}$, $f'(x) < 0$. Thus

$$\max_{x > 0} f(x) = f\left(\sqrt{\frac{n}{2\tau}}\right) = e^{-\frac{n}{2}} \left(\frac{n}{2}\right)^{\frac{n}{2}} \cdot \frac{1}{\tau^{\frac{n}{2}}} \lesssim \frac{1}{\tau^{\frac{n}{2}}}. \qquad \square$$

Lemma 4.4. *If α is any multi-index in $\mathbb{R}^d$, then*

$$\left| \int_{\mathbb{R}^d} e^{\sqrt{-1} \cdot \eta \cdot y - |y|^2} y^\alpha \, dy \right| \lesssim e^{-\frac{|\eta|^2}{4}}.$$

Proof. The following proof is based on direct evaluation which is analogous to the one used in [29]. Denote the integral with multi-index α by $I(\alpha)$. We prove by induction. For $\alpha = 0$, it is well known that

$$I(0) = \int_{\mathbb{R}^d} e^{\sqrt{-1} \cdot \eta \cdot y - |y|^2} \, dy = (\pi)^{\frac{d}{2}} e^{-\frac{|\eta|^2}{4}}.$$

Next we assume that the result holds for any multi-index $\alpha \leq \mu \in \mathbb{Z}^d$. Let $\mu^j \in Z^d$ be such that only the j's component is 1 and all others are 0. Consider the integral $I(\mu + \mu^j)$. We have

$$I(\mu + \mu^j) = \int_{\mathbb{R}^d} e^{\sqrt{-1}\cdot\eta\cdot y - |y|^2} y^{\mu+\mu^j}\, dy = \frac{1}{2}\int_{\mathbb{R}^d} e^{\sqrt{-1}\cdot\eta\cdot y} y^{\mu} \frac{\partial e^{-|y|^2}}{\partial y_j}\, dy$$
$$= -\frac{1}{2}\int_{\mathbb{R}^d} e^{-|y|^2} \frac{\partial}{\partial y_j}\left(e^{\sqrt{-1}\cdot\eta\cdot y} y^{\mu}\right) dy$$
$$= -\frac{1}{2}\int_{\mathbb{R}^d} e^{-|y|^2} e^{\sqrt{-1}\cdot\eta\cdot y} y^{\mu}\cdot\sqrt{-1}\cdot\eta_j\, dy - \frac{1}{2}\int_{\mathbb{R}^d} e^{-|y|^2} e^{\sqrt{-1}\cdot\eta\cdot y}\frac{\partial y^{\mu}}{\partial y_j}\, dy.$$

Thus

$$|I(\mu+\mu^j)| \le \frac{1}{2}|\eta_j|\cdot|I(\mu)| + \frac{1}{2}\left|\int_{\mathbb{R}^d} e^{-|y|^2} e^{\sqrt{-1}\cdot\eta\cdot y}\frac{\partial y^{\mu}}{\partial y_j}\, dy\right| \lesssim (|\eta|+1)e^{-\frac{|\eta|^2}{4}} \lesssim e^{-\frac{|\eta|^2}{4}}.$$

This closes our induction and the lemma is proved. □

Lemma 4.5. *For a positive number c, let* $g(\xi, \xi') = \frac{e^{-\frac{c|\xi-\xi'|^2}{|\xi|+|\xi'|}}}{|\xi'|^{\frac{d}{4}}}$ *be a function defined on* $\mathbb{R}^d \times \mathbb{R}^d$ *and* $I(\xi) = \int_{\mathbb{R}^d} g(\xi, \xi')\, d\xi'$. *Then*

$$I(\xi) \lesssim |\xi|^{\frac{d}{4}}, \quad \textit{for} \quad |\xi| \ge 1.$$

Proof. We divide the integral $I(\xi)$ into three parts: $I_1(\xi) = \int_{\frac{1}{2}|\xi|\le|\xi'|\le 2|\xi|} g(\xi,\xi')\, d\xi'$, $I_2(\xi) = \int_{|\xi'|\le\frac{1}{2}|\xi|} g(\xi,\xi')\, d\xi'$ and $I_3(\xi) = \int_{|\xi'|\ge 2|\xi|} g(\xi,\xi')\, d\xi'$. We estimate these three integrals one by one.

In the region I: $\frac{1}{2}|\xi| \le |\xi'| \le 2|\xi|$, we have $g(\xi,\xi') \le e^{-\frac{c|\xi-\xi'|^2}{|\xi|}}\cdot\frac{2^{\frac{d}{4}}}{|\xi|^{\frac{d}{4}}}$. Thus

$$I_1(\xi) \le \int_{\mathbb{R}^d} e^{-\frac{c|\xi-\xi'|^2}{|\xi|}}\cdot\frac{2^{\frac{d}{4}}}{|\xi|^{\frac{d}{4}}}\, d\xi'$$
$$= \frac{2^{\frac{d}{4}}}{|\xi|^{\frac{d}{4}}}\int_{\mathbb{R}^d} e^{-\frac{c|\xi'|^2}{|\xi|}}\, d\xi' \lesssim \frac{1}{|\xi|^{\frac{d}{4}}}\cdot\left(\frac{|\xi|}{c}\right)^{\frac{d}{2}} = \frac{1}{c^{\frac{d}{2}}}|\xi|^{\frac{d}{4}}.$$

In the region II: $|\xi'| \le \frac{1}{2}|\xi|$, we have $g(\xi,\xi') \le e^{-\frac{c|\xi|^2}{2|\xi|}}\cdot|\xi'|^{-\frac{d}{4}} = e^{-\frac{c|\xi|}{2}}\cdot|\xi'|^{-\frac{d}{4}}$. Thus

$$I_2(\xi) \le \int_{|\xi'|\le\frac{1}{2}|\xi|} e^{-\frac{c|\xi|}{2}}\cdot|\xi'|^{-\frac{d}{4}}\, d\xi' \lesssim e^{-\frac{c|\xi|}{2}}\cdot\int_0^{|\xi|} r^{-\frac{d}{4}+d-1}\, dr$$
$$\lesssim e^{-\frac{c|\xi|}{2}}\cdot|\xi|^{\frac{3d}{4}} \lesssim |\xi|^{\frac{d}{4}}.$$

In the region III: $|\xi'| \ge 2|\xi|$, we have $|\xi-\xi'| \ge |\xi'|-|\xi| > \frac{|\xi'|}{2} > \frac{1}{3}(|\xi|+|\xi'|)$. Thus $\frac{|\xi-\xi'|^2}{|\xi|+|\xi'|} \ge \frac{|\xi-\xi'|}{3}$ and consequently $g(\xi,\xi') \le e^{-\frac{c|\xi'|}{6}}\cdot(2|\xi|)^{-\frac{d}{4}}$. It follows that

$$I_3(\xi) \le \int_{|\xi'|\ge 2|\xi|} e^{-\frac{c|\xi'|}{6}}\cdot(2|\xi|)^{-\frac{d}{4}}\, d\xi' \le |\xi|^{-\frac{d}{4}}\int_{|\xi'|\ge 2|\xi|} e^{-\frac{c|\xi'|}{6}}\, d\xi'$$
$$\lesssim |\xi|^{-\frac{d}{4}}\int_{2|\xi|}^{\infty} e^{-\frac{cr}{6}}\cdot r^{d-1}\, dr \le |\xi|^{-\frac{d}{4}}\int_{2|\xi|}^{\infty} e^{-\frac{cr}{6}}\, dr$$
$$\lesssim |\xi|^{-\frac{d}{4}} e^{-\frac{c|\xi|}{3}} \lesssim |\xi|^{\frac{d}{4}}.$$

□

4.2. *Proof of Lemma 2.4*

Proof. Step 1. Denote $x_\gamma = 2\pi\frac{k}{L_\ell}$, $x_{\gamma'} = 2\pi\frac{k}{L_{\ell'}}$, $\delta\xi = \xi_{\ell,i} - \xi_{\ell',i'}$, $\delta x = x_\gamma - x_{\gamma'}$ and $a_{\gamma,\gamma'} = \langle\tilde{\phi}_\gamma - \phi_\gamma, \tilde{\phi}_{\gamma'} - \phi_{\gamma'}\rangle$. Then

$$\begin{aligned} a_{\gamma,\gamma'} &= \langle\hat{\tilde{\phi}}_\gamma - \hat{\phi}_\gamma, \hat{\tilde{\phi}}_{\gamma'} - \hat{\phi}_{\gamma'}\rangle \\ &= \frac{1}{L_\ell^{d/2}L_{\ell'}^{d/2}}\int_{\mathbb{R}^d} e^{-\sqrt{-1}\cdot\delta x\cdot\xi} e^{-\frac{|\xi-\xi_{\ell,i}|^2}{\sigma_\ell^2}-\frac{|\xi-\xi_{\ell',i'}|^2}{\sigma_{\ell'}^2}} \\ &\quad\cdot\left(1-\chi\left(\frac{\xi-\xi_{\ell,i}}{\sigma_\ell}\right)\right)\cdot\left(1-\chi\left(\frac{\xi-\xi_{\ell',i'}}{\sigma_{\ell'}}\right)\right)d\xi. \end{aligned}$$

We want to estimate $a_{\gamma,\gamma'}$. Without loss of generality, assume that $\ell \le \ell'$. By the change of variable, $\xi = \sigma_\ell z + \xi_\gamma$, we get

$$\begin{aligned} |a_{\gamma,\gamma'}| &= \frac{1}{4^d}\frac{L_\ell^{d/2}}{L_{\ell'}^{d/2}}\Bigg|\int_{\mathbb{R}^d} e^{-\sqrt{-1}\cdot\delta x\cdot\sigma_\ell z} e^{-|z|^2-|\frac{\sigma_\ell}{\sigma_{\ell'}}z-\frac{\xi_{\ell,i}-\xi_{\ell',i'}}{\sigma_{\ell'}}|^2}\cdot(1-\chi(z)) \\ &\quad\cdot\left(1-\chi\left(\frac{\sigma_\ell}{\sigma_{\ell'}}z-\frac{\xi_{\ell,i}-\xi_{\ell',i'}}{\sigma_{\ell'}}\right)\right)dz\Bigg| \\ &= \frac{\tau^{d/2}}{4^d}\left|\int_{\mathbb{R}^d} e^{-\sqrt{-1}\cdot\delta x\cdot\sigma_\ell z} e^{-|z|^2-|\tau z-\frac{\delta\xi}{\sigma_{\ell'}}|^2}\cdot(1-\chi(z))\cdot\left(1-\chi\left(\tau z-\frac{\delta\xi}{\sigma_{\ell'}}\right)\right)dz\right|, \end{aligned}$$

where $\tau = \frac{\sigma_\ell}{\sigma_{\ell'}} \le 1$ by the assumption that $\ell \le \ell'$.

Step 2. Let

$$I = \int_{\mathbb{R}^d} e^{-\sqrt{-1}\cdot\delta x\cdot\sigma_\ell z} e^{-|z|^2-|\tau z-\frac{\delta\xi}{\sigma_{\ell'}}|^2}\cdot(1-\chi(z))\cdot\left(1-\chi\left(\tau z-\frac{\delta\xi}{\sigma_{\ell'}}\right)\right)dz.$$

Define the differential operator $L = \frac{-\triangle+1}{1+\sigma_\ell^2|\delta x|^2}$ about the variable z. We have

$$Le^{-\sqrt{-1}\cdot\delta x\cdot\sigma_\ell z} = e^{-\sqrt{-1}\cdot\delta x\cdot\sigma_\ell z}.$$

Then

$$\begin{aligned} I &= \int_{\mathbb{R}^d} L^m(e^{-\sqrt{-1}\cdot\delta x\cdot\sigma_\ell z})e^{-|z|^2-|\tau z-\frac{\delta\xi}{\sigma_{\ell'}}|^2}\cdot(1-\chi(z))\cdot\left(1-\chi\left(\tau z-\frac{\delta\xi}{\sigma_{\ell'}}\right)\right)dz \\ &= \frac{1}{(1+\sigma_\ell^2|\delta x|^2)^m}\int_{\mathbb{R}^d} e^{-\sqrt{-1}\cdot\delta x\cdot\sigma_\ell z}(1-\triangle)^m\left(e^{-|z|^2-|\tau z-\frac{\delta\xi}{\sigma_{\ell'}}|^2}\cdot(1-\chi(z))\right. \\ &\quad\left.\cdot\left(1-\chi\left(\tau z-\frac{\delta\xi}{\sigma_{\ell'}}\right)\right)\right)dz \end{aligned}$$

where m is an integer to be determined later. Since $0 < \tau \le 1$, we can check that

$$\begin{aligned} &\left|(1-\triangle)^m\left(e^{-|z|^2-|\tau z-\frac{\delta\xi}{\sigma_{\ell'}}|^2}\cdot(1-\chi(z))\cdot\left(1-\chi\left(\tau z-\frac{\delta\xi}{\sigma_{\ell'}}\right)\right)\right)\right| \\ &\quad\lesssim e^{-|z|^2-|\tau z-\frac{\delta\xi}{\sigma_{\ell'}}|^2}\left(1+|z|^{2m}+\left|\frac{\delta\xi}{\sigma_{\ell'}}\right|^{2m}\right). \end{aligned}$$

Thus

$$|I| \lesssim \frac{1}{1+\sigma_\ell^{2m}|\delta x|^{2m}} \int_{|z|\geq 1} e^{-|z|^2-|\tau z-\frac{\delta\xi}{\sigma_{\ell'}}|^2}\,dz$$
$$\lesssim \frac{1}{1+\sigma_\ell^{2m}|\delta x|^{2m}} e^{-\frac{|\delta\xi|^2}{3\sigma_{\ell'}^2}} \int_{|z|\geq 1} e^{-\frac{|z|^2}{2}}\,dz$$
$$\lesssim \frac{1}{1+\sigma_\ell^{2m}|\delta x|^{2m}} e^{-\frac{|\delta\xi|^2}{3\sigma_{\ell'}^2}}.$$

Step 3. By the result in Step 2, we have

$$|a_{\gamma,\gamma'}| \lesssim \frac{\tau^{d/2}}{4^d}\frac{1}{1+\sigma_\ell^{2m}|\delta x|^{2m}} e^{-\frac{|\delta\xi|^2}{3\sigma_{\ell'}^2}} < \frac{1}{1+\sigma_\ell^{2m}|\delta x|^{2m}} e^{-\frac{|\delta\xi|^2}{3\sigma_{\ell'}^2}},$$

provided that $\ell \leq \ell'$. We now claim that

$$\sum_{\gamma'} |a_{\gamma,\gamma'}| \leq O(1)$$

uniformly for all fixed γ.

To prove this claim, we first let k_0 be the least integer such that $4^{k_0} \geq 8\sqrt{d}$. We then divide all γ''s into four regions, where regions I, II, III and IV consist of γ''s such that $\ell' \geq \ell + k_0$, $\ell' \leq \ell - k_0$, $\ell \leq \ell' < \ell + k_0$, and $\ell - k_0 < \ell' < \ell$, respectively. In what follows, we estimate $\sum_{\gamma'} |a_{\gamma,\gamma'}|$ in these four regions, respectively.

In the region I, we have

$$|\xi_{\gamma'}| \geq 4^{\ell'-1} \geq 4^{\ell+k_0-1} \geq 2\sqrt{d}4^{\ell} \geq 2|\xi_\gamma|.$$

Thus

$$\frac{|\delta\xi|^2}{3\sigma_{\ell'}^2} = \frac{|\xi_\gamma - \xi_{\gamma'}|^2}{3\sigma_{\ell'}^2} \geq \frac{3}{4}\frac{|\xi_{\gamma'}|^2}{3\sigma_{\ell'}^2} \geq \frac{|\xi_{\gamma'}|}{4}.$$

Hence

$$|a_{\gamma,\gamma'}| \lesssim \frac{1}{1+\sigma_\ell^{2m}|\delta x|^{2m}} e^{-\frac{|\xi_{\gamma'}|}{4}}.$$

It follows that

$$\sum_{\gamma':\ell'\geq \ell+k_0} |a_{\gamma,\gamma'}| \lesssim \sum_{\ell'\geq\ell+k_0}\sum_{i'}\sum_{k'} \frac{1}{1+\sigma_\ell^{2m}|\frac{2\pi k'}{L_{\ell'}} - \frac{2\pi k}{L_\ell}|^{2m}} e^{-\frac{|\xi_{\gamma'}|}{4}}$$
$$\lesssim \sum_{\ell'\geq\ell+k_0}\sum_{i'} e^{-\frac{|\xi_{\gamma'}|}{4}} \int_{\mathbb{R}^d} \frac{1}{1+\sigma_\ell^{2m}|\frac{2\pi k'}{L_{\ell'}} - \frac{2\pi k}{L_\ell}|^{2m}}\,dk'$$
$$\lesssim \sum_{\ell'\geq\ell+k_0}\sum_{i'} e^{-\frac{|\xi_{\ell',i'}|}{4}} \left(\frac{L_{\ell'}}{\sigma_\ell}\right)^d \quad \text{where we take } m = \left[\frac{d+1}{2}\right]+1$$

$$\lesssim \left(\frac{2}{\sigma_\ell}\right)^d \sum_{\ell' \geq \ell + k_0} \sum_{i'} e^{-\frac{|\xi_{\ell',i'}|}{4}} W_{\ell'}^d$$

$$\lesssim \int_{|\xi'| \geq |4^\ell|} e^{-\frac{|\xi'|}{4}} \, d\xi' \lesssim O(1).$$

By symmetry, in the region II, we have

$$|a_{\gamma,\gamma'}| \leq \frac{1}{1+\sigma_{\ell'}^{2m}|\delta x|^{2m}} e^{-\frac{|\xi_\gamma|}{4}}.$$

Thus

$$\sum_{\gamma': \ell' \leq \ell - k_0} |a_{\gamma,\gamma'}| \lesssim \sum_{\ell' \leq \ell - k_0} \sum_{i'} \sum_{k'} \frac{1}{1+\sigma_{\ell'}^{2m}|\frac{2\pi k'}{L_{\ell'}} - \frac{2\pi k}{L_\ell}|^{2m}} e^{-\frac{|\xi_\gamma|}{4}}$$

$$\lesssim \sum_{\ell' \leq \ell - k_0} \sum_{i'} e^{-\frac{|\xi_\gamma|}{4}} \int_{\mathbb{R}^d} \frac{1}{1+\sigma_{\ell'}^{2m}|\frac{2\pi k'}{L_{\ell'}} - \frac{2\pi k}{L_\ell}|^{2m}} \, dk'$$

$$\lesssim \sum_{\ell' \leq \ell - k_0} \sum_{i'} e^{-\frac{|\xi_{\ell,i}|}{4}} \quad \text{where we take } m = \left[\frac{d+1}{2}\right] + 1$$

$$\lesssim \sum_{\ell' \leq \ell - k_0} \sum_{i'} e^{-\frac{4^{\ell-1}}{4}}$$

$$\leq \sum_{\ell'=1}^{\ell'=\ell-k_0} 2^{d\ell'} e^{-4^{\ell-2}}$$

$$\lesssim 2^{d(\ell-2)} e^{-4^{\ell-2}} \lesssim O(1).$$

In the region III, we have

$$|a_{\gamma,\gamma'}| \leq \frac{1}{1+\sigma_\ell^{2m}|\delta x|^{2m}} e^{-\frac{|\xi_\gamma - \xi_{\gamma'}|^2}{3\sigma_{\ell'}^2}} \leq \frac{1}{1+4^{-k_0 m}\sigma_{\ell'}^{2m}|\delta x|^{2m}} e^{-\frac{|\xi_\gamma - \xi_{\gamma'}|^2}{3\sigma_{\ell'}^2}},$$

$$\sum_{\gamma': \ell \leq \ell' \leq \ell + k_0 - 1} |a_{\gamma,\gamma'}| \lesssim \sum_{\ell \leq \ell' \leq \ell + k_0 - 1} \sum_{i'} \sum_{k'} \frac{1}{1+4^{-k_0 m}\sigma_{\ell'}^{2m}|\frac{2\pi k'}{L_{\ell'}} - \frac{2\pi k}{L_\ell}|^{2m}} e^{-\frac{|\xi_\gamma - \xi_{\gamma'}|^2}{3\sigma_{\ell'}^2}}$$

$$\lesssim \sum_{\ell \leq \ell' \leq \ell + k_0 - 1} \sum_{i'} e^{-\frac{|\xi_\gamma - \xi_{\gamma'}|^2}{3\sigma_{\ell'}^2}} \int_{\mathbb{R}^d} \frac{1}{1+4^{-k_0 m}\sigma_{\ell'}^{2m}|\frac{2\pi k'}{L_{\ell'}} - \frac{2\pi k}{L_\ell}|^{2m}} \, dk'$$

$$\lesssim \sum_{\ell \leq \ell' \leq \ell + k_0 - 1} \sum_{i'} e^{-\frac{|\xi_{\ell,i} - \xi_{\ell',i'}|^2}{3\sigma_{\ell'}^2}} \quad \text{where we take } m = \left[\frac{d+1}{2}\right] + 1$$

$$\lesssim \sum_{\ell \leq \ell' \leq \ell + k_0 - 1} \int_{\mathbb{R}^d} \frac{1}{W_{\ell'}^d} e^{-\frac{|\xi_{\ell,i} - \xi'|^2}{3\sigma_{\ell'}^2}} \, d\xi'$$

$$\lesssim \sum_{\ell \leq \ell' \leq \ell + k_0 - 1} O(1) = O(1).$$

In the region IV, i.e. $\ell - k_0 < \ell' < \ell$, we have

$$|a_{\gamma,\gamma'}| \leq \frac{1}{1 + \sigma_{\ell'}^{2m}|\delta x|^{2m}} e^{-\frac{|\delta\xi|^2}{3\sigma_\ell^2}}.$$

Thus

$$\begin{aligned}
\sum_{\gamma':\ell-k_0<\ell'<\ell} |a_{\gamma,\gamma'}| &\lesssim \sum_{\ell-k_0<\ell'<\ell} \sum_{i'} \sum_{k'} \frac{1}{1 + \sigma_{\ell'}^{2m}|\frac{2\pi k'}{L_{\ell'}} - \frac{2\pi k}{L_\ell}|^{2m}} e^{-\frac{|\xi_\gamma - \xi_{\gamma'}|^2}{3\sigma_\ell^2}} \\
&\lesssim \sum_{\ell-k_0<\ell'<\ell} \sum_{i'} e^{-\frac{|\xi_\gamma - \xi_{\gamma'}|^2}{3\sigma_\ell^2}} \int_{\mathbb{R}^d} \frac{1}{1 + \sigma_{\ell'}^{2m}|\frac{2\pi k'}{L_{\ell'}} - \frac{2\pi k}{L_\ell}|} dk' \\
&\lesssim \sum_{\ell-k_0<\ell'<\ell} \sum_{i'} e^{-\frac{|\xi_{\ell,i} - \xi_{\ell',i'}|^2}{3\sigma_\ell^2}} \quad \text{where we take } m = \left[\frac{d+1}{2}\right] + 1 \\
&\lesssim \sum_{\ell-k_0<\ell'<\ell} \int_{\mathbb{R}^d} \frac{1}{W_{\ell'}^d} e^{-\frac{|\xi_{\ell,i} - \xi'|^2}{3\sigma_\ell^2}} d\xi' \\
&= \sum_{\ell-k_0<\ell'<\ell} O(1) = O(1).
\end{aligned}$$

By combining the above results, we can conclude that $\sum_{\gamma'} |a_{\gamma,\gamma'}| \leq O(1)$.

Step 4. Finally, we have

$$\begin{aligned}
\|f - \tilde{f}\|_{L^2(\mathbb{R}^d)}^2 &= \left| \sum_{\gamma,\gamma'} c_\gamma c_{\gamma'} a_{\gamma,\gamma'} \right| \\
&\leq \sum_{\gamma,\gamma'} \frac{|c_\gamma|^2 + |c_{\gamma'}|^2}{2} |a_{\gamma,\gamma'}| \\
&= \sum_{\gamma,\gamma'} |c_\gamma|^2 |a_{\gamma,\gamma'}| \\
&\leq \sum_\gamma |c_\gamma|^2 \sum_{\gamma'} |a_{\gamma,\gamma'}| \\
&\lesssim \sum_\gamma |c_\gamma|^2 \lesssim \|f\|_{L^2(\mathbb{R}^d)}^2.
\end{aligned}$$

□

4.3. *Proof of Lemma* 3.7

This subsection is devoted to the proof of Lemma 3.7. We begin with two technical lemmas.

Lemma 4.6. *Let M_1 and M_2 be two symmetric and positive definite matrices such that $M_1, M_2 > c_0 I > 0$, λ_1 and λ_2 be two positive numbers, and δx be a vector in $\mathbb{R}^d$. Then the following estimate holds*:

$$\int_{\mathbb{R}^d} |x^\alpha| \cdot |(x - \delta x)^\beta| \cdot e^{-\lambda_1 x^T M_1 x - \lambda_2 (x-\delta x)^T M_2 (x - \delta x)} dx \lesssim_{c_0} \frac{1}{(\lambda_1 + \lambda_2)^{\frac{d}{2}}} e^{-\frac{c_0 \lambda_1 \lambda_2}{2(\lambda_1 + \lambda_2)}|\delta x|^2} \cdot \frac{1}{\lambda_1^{\frac{|\alpha|}{2}} \cdot \lambda_2^{\frac{|\beta|}{2}}}.$$

Proof. Denote the integral on the left hand side of the above inequality by I. Let A be the symmetric positive definite matrix such that $\lambda_1 M_1 + \lambda_2 M_2 = A^2$. By the change of variable, $x \to y = Ax$, we have

$$\begin{aligned}\lambda_1 x^T M_1 x + \lambda_2 (x-\delta x)^T M_2 (x-\delta x) &= x^T A^2 x - 2\lambda_2 \langle Ax, A^{-1} M_2 \delta x\rangle + \lambda_2 \langle \delta x, M_2 \delta x\rangle \\ &= |y - \lambda_2 A^{-1} M_2 \delta x|^2 - \lambda_2^2 \langle \delta x, M_2 A^{-2} M_2 \delta x\rangle \\ &\quad + \lambda_2 \langle \delta x, M_2 \delta x\rangle.\end{aligned}$$

Since

$$\lambda_2 M_2 - \lambda_2^2 M_2 A^{-2} M_2 = \lambda_2 M_2 A^{-2}(\lambda_1 M_1 + \lambda_2 M_2) - \lambda_2 M_2 A^{-2}(\lambda_2 M_2) = \lambda_1 \lambda_2 M_2 A^{-2} M_1,$$

we have

$$\lambda_1 x^T M_1 x + \lambda_2 (x-\delta x)^T M_2 (x-\delta x) = |y - \lambda_2 A^{-1} M_2 \delta x|^2 + \langle \delta x, \lambda_1 \lambda_2 M_2 A^{-2} M_1 \delta x\rangle.$$

Thus the integral I becomes

$$I = |\det(A^{-1})| e^{-\lambda_1\lambda_2\langle \delta x, M_2 A^{-2} M_1 \delta x\rangle} \int_{\mathbb{R}^d} |(A^{-1}y)^\alpha| \cdot |(A^{-1}y - \delta x)^\beta| \cdot e^{-|y - \lambda_2 A^{-1} M_2 \delta x|^2} dy.$$

By the change of variable, $y \to z = y - \lambda_2 A^{-1} M_2 \delta x$, and using the fact that

$$A^{-1}y - \delta x = A^{-1}z + \lambda_2 A^{-2} M_2 \delta x - \delta x = A^{-1} z - \lambda_1 A^{-2} M_1 \delta x,$$

we obtain

$$\begin{aligned}I = |\det(A^{-1})| e^{-\lambda_1\lambda_2\langle \delta x, M_2 A^{-2} M_1 \delta x\rangle} \int |(A^{-1}z + \lambda_2 A^{-2} M_2 \delta x)^\alpha| \\ \cdot |(A^{-1}z - \lambda_1 A^{-2} M_1 \delta x)^\beta| \cdot e^{-|z|^2} dz. \end{aligned} \tag{60}$$

Let $B = \lambda_1 \lambda_2 M_2 A^{-2} M_1$. Then

$$B^{-1} = \lambda_1^{-1} \lambda_2^{-1} M_1^{-1} A^2 M_2^{-1} = \lambda_2^{-1} M_2^{-1} + \lambda_1^{-1} M_1^{-1}.$$

Since $M_1, M_2 > c_0 I > 0$, Lemma 4.1 implies that $M_1^{-1}, M_2^{-1} < \frac{1}{c_0} I$. Thus $B^{-1} < (\frac{1}{\lambda_1} + \frac{1}{\lambda_2})\frac{1}{c_0} I$. It follows that $B > (\frac{1}{\lambda_1} + \frac{1}{\lambda_2})^{-1} c_0 I = c_0 \frac{\lambda_1 \lambda_2}{\lambda_1 + \lambda_2} I$. Hence

$$e^{-\lambda_1\lambda_2 <\delta x, M_2 A^{-2} M_1 \delta x>} \leq e^{-\frac{c_0 \lambda_1 \lambda_2}{(\lambda_1 + \lambda_2)} |\delta x|^2}. \tag{61}$$

In addition, since $A^2 = \lambda_1 M_1 + \lambda_2 M_2 \geq (\lambda_1 + \lambda_2) c_0 I$, Lemma 4.1 yields $|\det A^{-2}| \leq (\frac{1}{(\lambda_1+\lambda_2)c_0})^d$. It follows that

$$|\det A^{-1}| \lesssim_{c_0} \frac{1}{(\lambda_1 + \lambda_2)^{\frac{d}{2}}}. \tag{62}$$

Next, we estimate the term $|(A^{-1}z + \lambda_2 A^{-2}M_2\delta x)^\alpha| \cdot |(A^{-1}z - \lambda_1 A^{-2}M_1\delta x)^\beta|$. We have

$$\begin{aligned}
&|(A^{-1}z + \lambda_2 A^{-2}M_2\delta x)^\alpha| \cdot |(A^{-1}z - \lambda_1 A^{-2}M_1\delta x)^\beta| \\
&\quad \lesssim (|A^{-1}z|^{|\alpha|} + |\lambda_2 A^{-2}M_2\delta x|^{|\alpha|}) \cdot (|A^{-1}z|^{|\beta|} + |\lambda_1 A^{-2}M_1\delta x|^{|\beta|}) \\
&\quad \lesssim \left(\frac{|z|^{|\alpha|}}{(\lambda_1+\lambda_2)^{\frac{|\alpha|}{2}}} + \frac{\lambda_2^{|\alpha|} \cdot |\delta x|^{|\alpha|}}{(\lambda_1+\lambda_2)^{|\alpha|}} \right) \cdot \left(\frac{|z|^{|\beta|}}{(\lambda_1+\lambda_2)^{\frac{|\beta|}{2}}} + \frac{\lambda_1^{|\beta|} \cdot |\delta x|^{|\beta|}}{(\lambda_1+\lambda_2)^{|\beta|}} \right). \qquad (63)
\end{aligned}$$

Substituting inequalities (61), (62), (63) into (60) and integrating over z, we get

$$\begin{aligned}
I &\lesssim \frac{1}{(\lambda_1+\lambda_2)^{\frac{d}{2}}} e^{-\frac{c_0\lambda_1\lambda_2}{(\lambda_1+\lambda_2)}|\delta x|^2} \cdot \frac{1}{(\lambda_1+\lambda_2)^{\frac{|\alpha|+|\beta|}{2}}} \left(1 + \frac{\lambda_2^{|\alpha|} \cdot |\delta x|^{|\alpha|}}{(\lambda_1+\lambda_2)^{\frac{|\alpha|}{2}}}\right) \cdot \left(1 + \frac{\lambda_1^{|\beta|} \cdot |\delta x|^{|\beta|}}{(\lambda_1+\lambda_2)^{\frac{|\beta|}{2}}}\right) \\
&\lesssim \frac{1}{(\lambda_1+\lambda_2)^{\frac{d}{2}}} e^{-\frac{c_0\lambda_1\lambda_2}{(\lambda_1+\lambda_2)}|\delta x|^2} \cdot \frac{1}{(\lambda_1+\lambda_2)^{\frac{|\alpha|+|\beta|}{2}}} (1 + \lambda_2^{\frac{|\alpha|}{2}} \cdot |\delta x|^{|\alpha|}) \cdot (1 + \lambda_1^{\frac{|\beta|}{2}} \cdot |\delta x|^{|\beta|}).
\end{aligned}$$

Applying Lemma 4.3, we have

$$\begin{aligned}
&e^{-\frac{c_0\lambda_1\lambda_2}{2(\lambda_1+\lambda_2)}|\delta x|^2}(1 + \lambda_2^{\frac{|\alpha|}{2}} \cdot |\delta x|^{|\alpha|}) \cdot (1 + \lambda_1^{\frac{|\beta|}{2}} \cdot |\delta x|^{|\beta|}) \\
&\quad \lesssim_{c_0} \left(1 + \lambda_2^{\frac{|\alpha|}{2}} \cdot \left(\frac{\lambda_1+\lambda_2}{\lambda_1\lambda_2}\right)^{\frac{|\alpha|}{2}}\right) \cdot \left(1 + \lambda_1^{\frac{|\beta|}{2}} \cdot \left(\frac{\lambda_1+\lambda_2}{\lambda_1\lambda_2}\right)^{\frac{|\beta|}{2}}\right) \\
&\quad \le 4\frac{(\lambda_1+\lambda_2)^{\frac{|\alpha|+|\beta|}{2}}}{\lambda_1^{\frac{|\alpha|}{2}} \cdot \lambda_2^{\frac{|\beta|}{2}}}.
\end{aligned}$$

It follows that

$$I \lesssim_{c_0} \frac{1}{(\lambda_1+\lambda_2)^{\frac{d}{2}}} e^{-\frac{c_0\lambda_1\lambda_2}{2(\lambda_1+\lambda_2)}|\delta x|^2} \cdot \frac{1}{\lambda_1^{\frac{|\alpha|}{2}} \cdot \lambda_2^{\frac{|\beta|}{2}}}. \qquad \square$$

Lemma 4.7. *Let M_1 and M_2 be two symmetric and positive definite matrices such that $0 < c_0 I < M_1, M_2 < c_1 I$. Assume that λ_1 and λ_2 are two positive numbers; δx and $\delta\xi$ are two vectors in $\mathbb{R}^d$. Then the following estimate holds,*

$$\begin{aligned}
&\left| \int_{\mathbb{R}^d} x^\alpha \cdot (x-\delta x)^\beta \cdot e^{\sqrt{-1}\delta\xi\cdot x - \lambda_1 x^T M_1 x - \lambda_2 (x-\delta x)^T M_2 (x-\delta x)} \right| \\
&\quad \lesssim \frac{1}{(\lambda_1+\lambda_2)^{\frac{d}{2}}} e^{-\frac{c_0\lambda_1\lambda_2}{2(\lambda_1+\lambda_2)}|\delta x|^2 - \frac{|\delta\xi|^2}{4c_1(\lambda_1+\lambda_2)}} \cdot \frac{1}{\lambda_1^{\frac{|\alpha|}{2}} \cdot \lambda_2^{\frac{|\beta|}{2}}}.
\end{aligned}$$

Proof. As in the proof of Lemma 4.6, we denote the integral by I and let A be the symmetric positive-definite matrix such that $\lambda_1 M_1 + \lambda_2 M_2 = A^2$. After changing variables twice, we can transform the integral into

$$\begin{aligned}
I = &|\det(A^{-1})| e^{-\lambda_1\lambda_2 <\delta x, M_2 A^{-2} M_1 \delta x>} e^{-\sqrt{-1}\delta\xi\cdot\lambda_2 A^{-1} M_2 \delta x} \cdot \\
&\int (A^{-1}z + \lambda_2 A^{-2}M_2\delta x)^\alpha \cdot (A^{-1}z - \lambda_1 A^{-2}M_1\delta x)^\beta \cdot e^{-\sqrt{-1}A^{-1}\delta\xi\cdot z - |z|^2} dz.
\end{aligned}$$

Letting $B = (b_{i,j}) = A^{-1}$, $u = \lambda_1 A^{-1} M_1 \delta x$, and $v = \lambda_2 A^{-1} M_2 \delta x$, we first show that

$$(Bz + v)^{\alpha}(Bz - u)^{\beta} = \sum_{|\mu| \le |\alpha| + |\beta|} d_{\mu} z^{\mu},$$

where d_{μ}'s are algebraic combinations of $b_{i,j}$'s, u_j's and v_j's, and they satisfy the following estimate,

$$|d_{\mu}| \lesssim \sum_{\substack{\mu_1 + \mu_2 = \mu \\ \mu_1 \le \alpha, \mu_2 \le \beta}} \frac{\lambda_2^{|\alpha| - |\mu_1|} \lambda_1^{|\beta| - |\mu_2|}}{(\lambda_1 + \lambda_2)^{|\alpha| + |\beta| - \frac{|\mu|}{2}}} \cdot (|\delta x|)^{|\alpha| + |\beta| - |\mu|}. \tag{64}$$

Indeed, by Lemma 4.2, we have

$$|b_{i,j}| \lesssim_{c_0} \frac{1}{(\lambda_1 + \lambda_2)^{\frac{1}{2}}} \equiv b; \quad |u_j| \le |u| \lesssim_{c_0} \frac{\lambda_1}{\lambda_1 + \lambda_2} |\delta x|; \quad |v_j| \le |v| \lesssim_{c_0} \frac{\lambda_1}{\lambda_1 + \lambda_2} |\delta x|.$$

Consider the term $(Bz + v)^{\alpha}$. Direct calculation gives

$$(Bz + v)^{\alpha} = \sum_{\mu \le \alpha} d_{\mu}^{(1)} z^{\mu}$$

with $d_{\mu}^{(1)} \lesssim b^{|\mu|} |v|^{|\alpha| - |\mu|}$.

Similarly,

$$(Bz - u)^{\beta} = \sum_{\mu \le \beta} d_{\mu}^{(2)} z^{\mu}$$

with $d_{\mu}^{(2)} \lesssim b^{|\mu|} |u|^{|\beta| - |\mu|}$.

As a result, we have

$$(Bz + v)^{\alpha}(Bz - u)^{\beta} = \sum_{\mu_1 \le \alpha, \mu_2 \le \beta} d_{\mu_1}^{(1)} d_{\mu_2}^{(2)} z^{\mu_1 + \mu_2}.$$

It follows that

$$d_{\mu} = \sum_{\substack{\mu_1 + \mu_2 = \mu \\ \mu_1 \le \alpha, \mu_2 \le \beta}} d_{\mu_1}^{(1)} d_{\mu_2}^{(2)},$$

and moreover,

$$\begin{aligned}
|d_{\mu}| &\lesssim_{c_0} \sum_{\substack{\mu_1 + \mu_2 = \mu \\ \mu_1 \le \alpha, \mu_2 \le \beta}} b^{|\mu_1| + |\mu_2|} \cdot |u|^{|\alpha| - |\mu_1|} \cdot |v|^{|\beta| - |\mu_2|} \\
&\lesssim_{c_0} \sum_{\substack{\mu_1 + \mu_2 = \mu \\ \mu_1 \le \alpha, \mu_2 \le \beta}} \frac{1}{(\lambda_1 + \lambda_2)^{\frac{|\mu|}{2}}} \cdot \left(\frac{\lambda_2}{\lambda_1 + \lambda_2} \right)^{|\alpha| - |\mu_1|} \cdot \left(\frac{\lambda_1}{\lambda_1 + \lambda_2} \right)^{|\beta| - |\mu_2|} \cdot |\delta x|^{|\alpha| + |\beta| - |\mu|} \\
&= \sum_{\substack{\mu_1 + \mu_2 = \mu \\ \mu_1 \le \alpha,\ \mu_2 \le \beta}} \frac{\lambda_2^{|\alpha| - |\mu_1|} \lambda_1^{|\beta| - |\mu_2|} |\delta x|^{|\alpha| + |\beta| - |\mu|}}{(\lambda_1 + \lambda_2)^{|\alpha| + |\beta| - \frac{|\mu|}{2}}}.
\end{aligned}$$

Inequality (64) is proved.

Now, we estimate I. We have

$$|I| \lesssim_{c_0} \frac{1}{(\lambda_1+\lambda_2)^{\frac{d}{2}}} e^{-\frac{c_0\lambda_1\lambda_2}{(\lambda_1+\lambda_2)}|\delta x|^2} \sum_{|\mu|\le|\alpha|+|\beta|} |d_\mu| \cdot \left| \int_{\mathbb{R}^d} z^\mu e^{-\sqrt{-1}A^{-1}\delta\xi\cdot z-|z|^2} dz \right|.$$

Using the inequality

$$e^{-\frac{c_0\lambda_1\lambda_2}{2(\lambda_1+\lambda_2)}|\delta x|^2} \cdot |\delta x|^n \lesssim_{c_0} \left(\frac{\lambda_1+\lambda_2}{\lambda_1\lambda_2}\right)^{\frac{n}{2}},$$

Lemma 4.4, and the fact that $|A^{-1}\delta\xi|^2 = (\delta\xi)^T A^{-2}\delta\xi \ge \frac{1}{(\lambda_1+\lambda_2)c_1}|\delta\xi|^2$, we get

$$|I| \lesssim_{c_0} \frac{1}{(\lambda_1+\lambda_2)^{\frac{d}{2}}} e^{-\frac{c_0\lambda_1\lambda_2}{2(\lambda_1+\lambda_2)}|\delta x|^2-\frac{|\delta\xi|^2}{4c_1(\lambda_1+\lambda_2)}} \cdot \left(\sum_{|\mu|\le|\alpha|+|\beta|} \sum_{\substack{\mu_1+\mu_2=\mu \\ \mu_1\le\alpha,\mu_2\le\beta}} \frac{\lambda_2^{|\alpha|-|\mu_1|}\lambda_1^{|\beta|-|\mu_2|}}{(\lambda_1+\lambda_2)^{|\alpha|+|\beta|-\frac{|\mu|}{2}|}} \cdot \left(\frac{\lambda_1+\lambda_2}{\lambda_1\lambda_2}\right)^{\frac{|\alpha|+|\beta|-|\mu|}{2}} \right).$$

Note that for each $\mu_1 \le \alpha$ and $\mu_2 \le \beta$,

$$\begin{aligned}
&\frac{\lambda_2^{|\alpha|-|\mu_1|}\lambda_1^{|\beta|-|\mu_2|}}{(\lambda_1+\lambda_2)^{|\alpha|+|\beta|-\frac{|\mu|}{2}|}} \cdot \left(\frac{\lambda_1+\lambda_2}{\lambda_1\lambda_2}\right)^{\frac{|\alpha|+|\beta|-|\mu|}{2}} \\
&= \left(\frac{1}{\lambda_1+\lambda_2}\right)^{\frac{|\alpha|+|\beta|}{2}} \cdot \left(\frac{\lambda_1}{\lambda_2}\right)^{\frac{|\mu_1|-|\mu_2|}{2}} \cdot \left(\frac{\lambda_1}{\lambda_2}\right)^{\frac{|\beta|-|\alpha|}{2}} \\
&\le \left(\frac{1}{\lambda_1+\lambda_2}\right)^{\frac{|\alpha|+|\beta|}{2}} \cdot \left(\left(\frac{\lambda_1}{\lambda_2}\right)^{\frac{|\alpha|}{2}} + \left(\frac{\lambda_1}{\lambda_2}\right)^{-\frac{|\beta|}{2}}\right) \cdot \left(\frac{\lambda_1}{\lambda_2}\right)^{\frac{|\beta|-|\alpha|}{2}} \\
&\le \left(\frac{1}{\lambda_1+\lambda_2}\right)^{\frac{|\alpha|+|\beta|}{2}} \cdot \left(\left(\frac{\lambda_1}{\lambda_2}\right)^{\frac{|\beta|}{2}} + \left(\frac{\lambda_1}{\lambda_2}\right)^{-\frac{|\alpha|}{2}}\right) \\
&\le \left(\frac{1}{\lambda_1+\lambda_2}\right)^{\frac{|\alpha|+|\beta|}{2}} \cdot \frac{(\lambda_1+\lambda_2)^{\frac{|\alpha|+|\beta|}{2}}}{\lambda_1^{\frac{|\alpha|}{2}} \cdot \lambda_2^{\frac{|\beta|}{2}}} \\
&\le \frac{1}{\lambda_1^{\frac{|\alpha|}{2}} \cdot \lambda_2^{\frac{|\beta|}{2}}}.
\end{aligned}$$

Thus,

$$I \lesssim_{c_0} \frac{1}{(\lambda_1+\lambda_2)^{\frac{d}{2}}} e^{-\frac{c_0\lambda_1\lambda_2}{2(\lambda_1+\lambda_2)}|\delta x|^2-\frac{|\delta\xi|^2}{4c_1(\lambda_1+\lambda_2)}} \cdot \frac{1}{\lambda_1^{\frac{|\alpha|}{2}} \cdot \lambda_2^{\frac{|\beta|}{2}}}.$$

Now we are ready to show Lemma 3.7.

Proof of Lemma 3.7. Following the proof of Lemma 4.7, we denote the integral by I and let A be the symmetric positive-definite matrix such that $\lambda_1 M_1 + \lambda_2 M_2 = A^2$.

After changing variables twice, we get

$$
\begin{aligned}
|I| = |\det(A^{-1})| \cdot e^{-\lambda_1\lambda_2 <\delta x, M_2 A^{-2} M_1 \delta x>} \\
\bigg| \int (A^{-1}z + \lambda_2 A^{-2} M_2 \delta x)^\alpha \cdot (A^{-1}z - \lambda_1 A^{-2} M_1 \delta x)^\beta \\
\cdot e^{\sqrt{-1}\cdot\eta\cdot z - |z|^2 - \sqrt{-1}\cdot z^T A^{-1}(\lambda_1 N_1 + \lambda_2 N_2) A^{-1} z} dz \bigg|,
\end{aligned}
$$

where $\eta = A^{-1}\delta\xi - \lambda_1\lambda_2 A^{-1} N_1 A^{-2} M_2 \delta x + \lambda_1\lambda_2 A^{-1} N_1 A^{-2} M_1 \delta x$.

Since $A^{-1}(\lambda_1 N_1 + \lambda_2 N_2)A^{-1}$ is symmetric and

$$
\|A^{-1}(\lambda_1 N_1 + \lambda_2 N_2)A^{-1}\| \le \frac{\lambda_1 c_2 + \lambda_2 c_2}{c_0(\lambda_1 + \lambda_2)} = \frac{c_2}{c_0},
$$

we can find an orthogonal matrix T such that

$$
T^T A^{-1}(\lambda_1 N_1 + \lambda_2 N_2) A^{-1} T = \operatorname{diag}(\epsilon_1, \epsilon_2, \ldots, \epsilon_d),
$$

where ϵ_j's are eigenvalues of $A^{-1}(\lambda_1 N_1 + \lambda_2 N_2)A^{-1}$ and satisfy $|\epsilon_j| \le \frac{c_2}{c_0}$.

By the change of variable, $z \to T^{-1}z = w$, we have

$$
\begin{aligned}
J &= \bigg| \int (A^{-1}z + \lambda_2 A^{-2} M_2 \delta x)^\alpha \cdot (A^{-1}z - \lambda_1 A^{-2} M_1 \delta x)^\beta \\
&\quad \cdot e^{\sqrt{-1}\cdot\eta\cdot z - |z|^2 - \sqrt{-1}\cdot z^T A^{-1}(\lambda_1 N_1 + \lambda_2 N_2) A^{-1} z} dz \bigg| \\
&= \bigg| \int (A^{-1}Tw + \lambda_2 A^{-2} M_2 \delta x)^\alpha \cdot (A^{-1}Tw - \lambda_1 A^{-2} M_1 \delta x)^\beta \\
&\quad \cdot e^{\sqrt{-1}\cdot T^T\eta\cdot w - |w|^2 - \sqrt{-1}\cdot\epsilon_1 w_1^2 - \cdots - \sqrt{-1}\cdot\epsilon_d w_d^2} dw \bigg|.
\end{aligned}
$$

As in the proof of Lemma 4.7, we can show that

$$
(A^{-1}Tw + \lambda_2 A^{-2} M_2 \delta x)^\alpha (A^{-1}Tw - \lambda_1 A^{-2} M_1 \delta x)^\beta = \sum_{|\mu| \le |\alpha| + |\beta|} d_\mu z^\mu,
$$

where d_μ's satisfy the following estimate,

$$
d_\mu \lesssim \sum_{\substack{\mu_1 + \mu_2 = \mu \\ \mu_1 \le \alpha, \mu_2 \le \beta}} \frac{\lambda_2^{|\alpha| - |\mu_1|} \lambda_1^{|\beta| - |\mu_2|}}{(\lambda_1 + \lambda_2)^{|\alpha| + |\beta| - \frac{|\mu|}{2}|}} \cdot (|\delta x|)^{|\alpha| + |\beta| - |\mu|}.
$$

Let $T^T\eta = \sigma$. Then

$$
J \le \sum_{|\mu| \le |\alpha| + |\beta|} |d_\mu| \cdot \left| \int_{\mathbb{R}^d} w^\mu e^{\sqrt{-1}\cdot\sigma\cdot w - (1 + \sqrt{-1}\cdot\epsilon_1) w_1^2 - \cdots - (1 + \sqrt{-1}\cdot\epsilon_d) w_d^2} dw \right|. \tag{65}
$$

Using the well-known fact that

$$\int_{\mathbb{R}} e^{-zx^2-\sqrt{-1}\cdot\xi\cdot x}\,dx = \left(\frac{\pi}{z}\right)^{\frac{1}{2}} e^{-\frac{|\xi|^2}{4z}}$$

for any complex number z with a positive real part, we can show by a similar procedure as in the proof of Lemma 4.4 that

$$\left|\int_{\mathbb{R}} x^n e^{-zx^2-\sqrt{-1}\cdot\xi\cdot x}\,dx\right| \lesssim \frac{1}{|z|^{\frac{1}{2}}} e^{-Re\{z\}\frac{|\xi|^2}{4|z|^2}}$$

for any nonnegative integer n and any complex number z with a positive real part. Applying this fact to the right hand side of (65) we get

$$J \lesssim \sum_{|\mu|\le|\alpha|+|\beta|} |d_\mu|\cdot e^{-\frac{|\sigma|^2}{4(1+\left(\frac{c_2}{c_0}\right)^2)}} = \sum_{|\mu|\le|\alpha|+|\beta|} |d_\mu|\cdot e^{-\frac{|\eta|^2}{4(1+\left(\frac{c_2}{c_0}\right)^2)}} \le \sum_{|\mu|\le|\alpha|+|\beta|} |d_\mu|\cdot e^{-\frac{|\eta|^2}{4\tau(1+\left(\frac{c_2}{c_0}\right)^2)}},$$

where $\tau \ge 1$ will be determined later. Moreover,

$$|A^{-1}N_1A^{-2}M_2\delta x| \le \|A^{-1}\|^3\cdot\|N_1\|\cdot\|M_2\|\cdot|\delta x| \le \frac{1}{(c_0(\lambda_1+\lambda_2))^{\frac{3}{2}}}c_1c_2|\delta x|,$$

$$|A^{-1}N_2A^{-2}M_1\delta x| \le \|A^{-1}\|^3\cdot\|N_2\|\cdot\|M_1\|\cdot|\delta x| \le \frac{1}{(c_0(\lambda_1+\lambda_2))^{\frac{3}{2}}}c_1c_2|\delta x|.$$

Let $u = \lambda_1\lambda_2A^{-1}N_1A^{-2}M_2\delta x$, and $v = -\lambda_1\lambda_2A^{-1}N_2A^{-2}M_1\delta x$. Then

$$\begin{aligned}
|\eta|^2 &= |A^{-1}\delta\xi - u - v|^2\\
&\ge |A^{-1}\delta\xi|^2 + |u+v|^2 - 2|A^{-1}\delta\xi|\cdot|u+v|\\
&\ge \frac{|A^{-1}\delta\xi|^2}{2} - |u+v|^2\\
&\ge \frac{|A^{-1}\delta\xi|^2}{2} - (|u|+|v|)^2\\
&\ge \frac{|\delta\xi|^2}{2c_1(\lambda_1+\lambda_2)} - \frac{4c_1^2c_2^2\lambda_1^2\lambda_2^2|\delta x|^2}{(c_0(\lambda_1+\lambda_2))^3}\\
&\ge \frac{|\delta\xi|^2}{2c_1(\lambda_1+\lambda_2)} - \frac{4c_1^2c_2^2\lambda_1\lambda_2|\delta x|^2}{c_0(\lambda_1+\lambda_2)}.
\end{aligned}$$

Thus

$$J \le \sum_{|\mu|\le|\alpha|+|\beta|} |d_\mu|\cdot e^{-\frac{|\delta\xi|^2}{8c_1(\lambda_1+\lambda_2)\tau(1+\left(\frac{c_2}{c_0}\right)^2)} + \frac{c_1^2c_2^2\lambda_1\lambda_2}{c_0(\lambda_1+\lambda_2)\tau(1+\left(\frac{c_2}{c_0}\right)^2)}|\delta x|^2}.$$

As in the proof of Lemma 4.7, this combined with (65) and Lemma 4.3 gives

$$|I| \lesssim_{c_0} \frac{1}{(\lambda_1+\lambda_2)^{\frac{d}{2}}} e^{-\frac{c_0\lambda_1\lambda_2}{2(\lambda_1+\lambda_2)}|\delta x|^2 + \frac{c_1^2c_2^2\lambda_1\lambda_2}{c_0(\lambda_1+\lambda_2)\tau(1+\left(\frac{c_2}{c_0}\right)^2)}|\delta x|^2 - \frac{|\delta\xi|^2}{8c_1(\lambda_1+\lambda_2)\tau(1+\left(\frac{c_2}{c_0}\right)^2)}} \cdot \frac{1}{\lambda_1^{\frac{|\alpha|}{2}}\cdot\lambda_2^{\frac{|\beta|}{2}}}.$$

By choosing $\tau = \frac{4c_1c_2}{c_0^2+c_2^2}$ and $c_1^* = \frac{32c_1^3c_2^2}{c_0^2}$, it follows that

$$|I| \lesssim_{c_0} \frac{1}{(\lambda_1+\lambda_2)^{\frac{d}{2}}} e^{-\frac{c_0\lambda_1\lambda_2}{4(\lambda_1+\lambda_2)}|\delta x|^2 - \frac{|\delta\xi|^2}{c_1^*(\lambda_1+\lambda_2)}} \cdot \frac{1}{\lambda_1^{\frac{|\alpha|}{2}} \cdot \lambda_2^{\frac{|\beta|}{2}}},$$

which completes the proof. □

4.4. *Proof of Lemma 3.8*

Proof. Step 1. Let $a_{\gamma,\gamma'} = \lambda_\gamma^{\frac{m}{2}}\lambda_{\gamma'}^{\frac{m}{2}} < \Phi_\gamma(t,x)\cdot(x-x_\gamma(t))^{\alpha(\gamma)}, \Phi_{\gamma'}(t,x)\cdot(x-x_{\gamma'}(t))^{\alpha(\gamma')} >$. Then

$$|a_{\gamma,\gamma'}| = \lambda_\gamma^{\frac{m}{2}}\lambda_{\gamma'}^{\frac{m}{2}}|A_\gamma(t)|\cdot|A_{\gamma'}(t)|\cdot\left|\int_{\mathbb{R}^d}(x-x_\gamma(t))^{\alpha(\gamma)}\cdot(x-x_{\gamma'}(t))^{\alpha(\gamma')}\right.$$
$$\left.\cdot\, e^{\sqrt{-1}\cdot(\xi_\gamma(t)-\xi_{\gamma'}(t))\cdot x - \lambda_\gamma(x-x_\gamma(t))^T M_\gamma(t)(x-x_\gamma(t)) - \lambda_{\gamma'}(x-x_{\gamma'}(t))^T M_{\gamma'}(t)(x-x_{\gamma'}(t))}\,dx\right|.$$

Using Lemma 3.1 and Lemma 3.7, we further get

$$|a_{\gamma,\gamma'}| \lesssim \lambda_\gamma^{\frac{m}{2}}\lambda_{\gamma'}^{\frac{m}{2}} \frac{(\lambda_\gamma\lambda_{\gamma'})^{\frac{d}{4}}}{(\lambda_\gamma+\lambda_{\gamma'})^{\frac{d}{2}}} e^{-\frac{c_0\lambda_\gamma\lambda_{\gamma'}}{2(\lambda_\gamma+\lambda_{\gamma'})}|x_\gamma(t)-x_{\gamma'}(t)|^2 - \frac{|\xi_\gamma(t)-\xi_{\gamma'}(t)|^2}{4c_1(\lambda_\gamma+\lambda_{\gamma'})}} \cdot \frac{1}{\lambda_\gamma^{\frac{|\alpha(\gamma)|}{2}}\cdot\lambda_{\gamma'}^{\frac{|\alpha(\gamma')|}{2}}}$$
$$\le \frac{(\lambda_\gamma\lambda_{\gamma'})^{\frac{d}{4}}}{(\lambda_\gamma+\lambda_{\gamma'})^{\frac{d}{2}}} e^{-\frac{c_0\lambda_\gamma\lambda_{\gamma'}}{2(\lambda_\gamma+\lambda_{\gamma'})}|x_\gamma(t)-x_{\gamma'}(t)|^2 - \frac{|\xi_\gamma(t)-\xi_{\gamma'}(t)|^2}{4c_1(\lambda_\gamma+\lambda_{\gamma'})}}, \tag{66}$$

where c_0 and c_1 are some constants.

Step 2. Denote

$$d_{\gamma,\gamma'} = \frac{(\lambda_\gamma\lambda_{\gamma'})^{\frac{d}{4}}}{(\lambda_\gamma+\lambda_{\gamma'})^{\frac{d}{2}}} e^{-\frac{c_0\lambda_\gamma\lambda_{\gamma'}}{2(\lambda_\gamma+\lambda_{\gamma'})}|x_\gamma(t)-x_{\gamma'}(t)|^2 - \frac{|\xi_\gamma(t)-\xi_{\gamma'}(t)|^2}{4c_1(\lambda_\gamma+\lambda_{\gamma'})}}. \tag{67}$$

By Lemma 3.2, we have

$$|a_{\gamma,\gamma'}| \le d_{\gamma,\gamma'} \le \frac{(\lambda_\gamma\lambda_{\gamma'})^{\frac{d}{4}}}{(\lambda_\gamma+\lambda_{\gamma'})^{\frac{d}{2}}} e^{-\frac{c_0C_3'\lambda_\gamma\lambda_{\gamma'}}{2(\lambda_\gamma+\lambda_{\gamma'})}|x_\gamma(0)-x_{\gamma'}(0)|^2 - \frac{C_3'|\xi_\gamma(0)-\xi_{\gamma'}(0)|^2}{4c_1(\lambda_\gamma+\lambda_{\gamma'})}}$$

for some $C_3' > 0$.

Step 3. We claim that

$$e^{-\frac{c_0C_3'\lambda_\gamma\lambda_{\gamma'}}{2(\lambda_\gamma+\lambda_{\gamma'})}|x_\gamma(0)-x_{\gamma'}(0)|^2} \lesssim \min_{x\in D_\gamma,\xi\in B_\gamma,x'\in D_{\gamma'},\xi'\in B_{\gamma'}} e^{-\frac{c_0'|\xi||\xi'|}{|\xi|+|\xi'|}|x-x'|^2}, \tag{68}$$

$$e^{-\frac{C_3'|\xi_\gamma(0)-\xi_{\gamma'}(0)|^2}{4c_1(\lambda_\gamma+\lambda_{\gamma'})}} \lesssim \min_{x\in D_\gamma,\xi\in B_\gamma,x'\in D_{\gamma'},\xi'\in B_{\gamma'}} e^{-\frac{|\xi-\xi'|^2}{c_1'(|\xi|+|\xi'|)}}, \tag{69}$$

for some properly chosen c_0' and c_1'.

To see (68), we recall that $|\xi| \sim \lambda_\gamma$, $|\xi'| \sim \lambda_{\gamma'}$. Thus

$$\frac{|\xi||\xi'|}{(|\xi| + |\xi'|)} \sim \frac{\lambda_\gamma \lambda_{\gamma'}}{(\lambda_\gamma + \lambda_{\gamma'})}.$$

Moreover,

$$\begin{aligned} |x - x'|^2 &= |x - x_\gamma + x_\gamma - x_{\gamma'} + x_{\gamma'} - x'|^2 \\ &\le 3|x_\gamma - x_{\gamma'}|^2 + 3|x_\gamma - x|^2 + 3|x_{\gamma'} - x'|^2 \\ &\le 3|x_\gamma - x_{\gamma'}|^2 + \frac{3}{L_\gamma^2} + \frac{3}{L_{\gamma'}^2}. \end{aligned}$$

Since $\lambda_\gamma = O(4^\ell)$, $\lambda_{\gamma'} = O(4^{\ell'})$, $L_\gamma = O(2^\ell)$, and $L_{\gamma'} = O(2^{\ell'})$, we have $\frac{\lambda_\gamma \lambda_{\gamma'}}{\lambda_\gamma + \lambda_{\gamma'}}(\frac{1}{L_\gamma} + \frac{1}{L_{\gamma'}}) = O(1)$. The inequality (68) follows. Similarly, we can prove (69).

Step 4. For the given b_γ, we define function $\tilde{b}(x, \xi)$ by letting $\tilde{b}(x, \xi) = b_\gamma$ for all $x \in D_\gamma$, $\xi \in B_\gamma$. We can check that $\|\tilde{b}\|^2_{L^2(\mathbb{R}^{2d})} = \frac{1}{2^d} \sum_\gamma |b_\gamma|^2$.

Define

$$f(x, \xi, x', \xi') = \tilde{b}(x, \xi)\tilde{b}(x', \xi') \frac{|\xi|^{\frac{d}{4}} |\xi'|^{\frac{d}{4}}}{(|\xi| + |\xi'|)^{\frac{d}{2}}} e^{-\frac{c_0' |\xi||\xi'|}{|\xi| + |\xi'|}|x - x'|^2 - \frac{|\xi - \xi'|^2}{c_1'(|\xi| + |\xi'|)}}.$$

The inequalities (68) and (69) yield

$$|b_\gamma b_{\gamma'} d_{\gamma,\gamma'}| \lesssim \min_{x \in D_\gamma, \xi \in B_\gamma, x' \in D_{\gamma'}, \xi' \in B_{\gamma'}} |f(x, \xi, x', \xi')|.$$

It follows that

$$J = \sum_{\gamma,\gamma'} |b_\gamma b_{\gamma'} d_{\gamma,\gamma'}| \lesssim \int_{\mathbb{R}^{4d}} |f(x, \xi, x', \xi')| \, dx d\xi dx' d\xi'. \tag{70}$$

Step 5. To estimate the integral $\int_{\mathbb{R}^{4d}} f(x, \xi, x', \xi') \, dx d\xi dx' d\xi'$, we define

$$I(\xi, \xi') = \int \tilde{b}(x, \xi)\tilde{b}(x', \xi') e^{-\lambda |x - x'|^2} \, dx dx',$$

where $\lambda = \frac{c_0' |\xi||\xi'|}{|\xi| + |\xi'|}$. By the change of variable, $x = u + v$ and $x' = v - u$, we have

$$\begin{aligned} |I(\xi, \xi')| &= \left| \int \tilde{b}(u + v, \xi)\tilde{b}(v - u, \xi') e^{-2^{2d}\lambda|u|^2} du dv \right| \\ &= \left| 2^d \int e^{-2^{2d}\lambda|u|^2} \, du \int \tilde{b}(u + v, \xi)\tilde{b}(v - u, \xi') dv \right| \\ &\le 2^d \int e^{-2^{2d}\lambda|u|^2} \, du \left(\int |\tilde{b}|^2(u + v, \xi) dv + \int |\tilde{b}|^2(v - u, \xi') dv \right) \\ &\le \frac{1}{\lambda^{\frac{d}{2}}} \left(\int |\tilde{b}|^2(x, \xi) \, dx + \int |\tilde{b}|^2(x, \xi') dx \right) \\ &= \left(\frac{c_0'(|\xi| + |\xi'|)}{|\xi||\xi'|} \right)^{\frac{d}{2}} \left(\int |\tilde{b}|^2(x, \xi) dx + \int |\tilde{b}|^2(x, \xi') dx \right). \end{aligned}$$

Thus

$$
\begin{aligned}
&\left|\int_{\mathbb{R}^{4d}} f(x,\xi,x',\xi')\,dxd\xi dx'd\xi'\right| \\
&= \left|\int I(\xi,\xi')\frac{|\xi|^{\frac{d}{4}}|\xi'|^{\frac{d}{4}}}{(|\xi|+|\xi'|)^{\frac{d}{2}}} e^{-\frac{|\xi-\xi'|^2}{c_1'(|\xi|+|\xi'|)}}\,d\xi d\xi'\right| \\
&\lesssim \int \frac{1}{|\xi|^{\frac{d}{4}}|\xi'|^{\frac{d}{4}}} e^{-\frac{|\xi-\xi'|^2}{c_1'(|\xi|+|\xi'|)}}\,d\xi d\xi' \left(\int |\tilde{b}|^2(x,\xi)dx + \int |\tilde{b}|^2(x,\xi')dx\right) \\
&= 2\int \frac{|\tilde{b}|^2(x,\xi)}{|\xi|^{\frac{d}{4}}} dxd\xi \int e^{-\frac{|\xi-\xi'|^2}{c_1'(|\xi|+|\xi'|)}} |\xi'|^{-\frac{d}{4}}\,d\xi' \\
&\lesssim \int |\tilde{b}|^2(x,\xi)dxd\xi \qquad \text{by Lemma 4.5} \\
&= \frac{1}{2^d}\sum_\gamma |b_\gamma|^2. \qquad (71)
\end{aligned}
$$

Step 6. Finally, combining (66), (67), (70), and (71), we have

$$
\left\| \sum_\gamma b_\gamma \lambda_\gamma^{\frac{m}{2}} \Phi_\gamma(t,\cdot)\cdot(\cdot - x_\gamma(t))^{\alpha(\gamma)} \right\|_{L^2(\mathbb{R}^d)}^2 \le \sum_{\gamma,\gamma'} b_\gamma b_{\gamma'} |a_{\gamma,\gamma'}| \lesssim \sum_{\gamma,\gamma'} b_\gamma b_{\gamma'} d_{\gamma,\gamma'} \lesssim \sum_\gamma |b_\gamma|^2.
$$

The lemma is proved. □

Acknowledgment

The research of G. Bao and H. Zhang was supported in part by the NSF grant DMS-0908325. This research of G. Bao was also supported in part by the NSF grants CCF-0830161, EAR-0724527, the ONR grant N00014-09-1-0384 and a special research grant from Zhejiang University. J. Qian would like to thank Professors R. Burridge, S. Osher, J. Ralston and W. W. Symes for their interest and encouragement in this work. J. Qian is partially supported by the NSF Grant DMS-0810104, NSF Grant CCF-0830161, NSF Grant DMS-1115363, and NSF Grant DMS-1222368. L. Ying is partially supported by a Sloan Research Fellowship and NSF Career Award DMS-0846501.

References

[1] Andersson, F., de Hoop, M.V., Smith, H., Uhlmann, G. (2008). A multi-scale approach to hyperbolic equations with limited smoothness. *Comm. Part. Diff. Eq.* 33:988–1017.

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

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

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

[5] Candès, E., Demanet, L., Donoho, D., Ying, L. (2006). Fast discrete curvelet transforms. *Multiscale Model. Simul.* 5:861–899.
[6] Candès, E., Donoho, D. (2004). New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities. *Comm. Pure Appl. Math.* 57:219–266.
[7] Cerveny, V., Popov, M., Psencik, I. (1982). Computation of wave fields in inhomogeneous media-Gaussian beam approach. *Geophys. J. R. Astr. Soc.* 70:109–128.
[8] Cordoba, A., Fefferman, C. (1978). Wave packets and fourier integral operators. *Comm. Part. Diff. Eqs.* 3:979–1005.
[9] de Hoop, M., Smith, H., Uhlmann, G., Van der Hilst, R. (2009). Seismic imaging with the generalized Radon transform: A curvelet transform perspective. *Inverse Problems* 25:025005.
[10] Demanet, L., Ying, L. (2007). Wave atoms and sparsity of oscillatory patterns. *Appl. Comput. Harmon. Anal.* 23:368–387.
[11] Demanet, L., Ying, L. (2009). Wave atoms and time upscaling of wave equations. *Numerische Mathematik* 113:1–71.
[12] Geba, D., Tataru, D. (2007). A phase space transform adapted to the wave equation. *Comm. Part. Diff. Eqs.* 32:1065–1101.
[13] Hill, N. (1990). Gaussian beam migration. *Geophysics* 55:1416–1428.
[14] Hill, N. (2001). Prestack Gaussian-beam depth migration. *Geophysics* 66: 1240–1250.
[15] Katchalov, A., Kurylev, Y., Lassas, M. (2001). *Inverse Boundary Spectral Problems*. Boca Raton, FL: Chapman & Hall/CRC Press.
[16] Lasiecka, I., Lions, J.L., Triggiani, R. (1986). Non-homogeneous boundary value problems for second order hyperbolic problems. *J. Math. Pures Appl.* 65:149–192.
[17] Leung, S., Qian, J. (2009). Eulerian Gaussian beam methods for Schrödinger equations in the semi-classical regime. *J. Comp. Phys.* 228:2951–2977.
[18] Leung, S. Qian, J. (2010). The backward phase flow and FBI-transform-based Eulerian Gaussian beams for the Schrödinger equation. *J. Comp. Phys.* 229:8888–8917.
[19] Leung, S., Qian, J., Burridge, R. (2007). Eulerian Gaussian beams for high frequency wave propagation. *Geophysics* 72:SM61–SM76.
[20] Maslov, V.P. (1994). *The Complex WKB Method for Nonlinear Equations I: Linear Theory*. Basel: Birkhäuser Verlag.
[21] Popov, M.M. (1982). A new method of computation of wave fields using Gaussian beams. *Wave Motion* 4:85–97.
[22] Qian, J., Ying, L. (2010). Fast Gaussian wavepacket transforms and Gaussian beams for the Schrödinger equation. *J. Comput. Phys.* 229:7848–7873.
[23] Qian, J., Ying, L. (2010). Fast multiscale Gaussian wavepacket transforms and multiscale Gaussian beams for the wave equation. *SIAM J. Multiscale Model. Sim.* 8:1803–1837.
[24] Ralston, J. (1983). Gaussian beams and the propagation of singularities. In: Littman, W., ed., *Studies in Partial Differential Equations.* MAA Studies in Mathematics, Washington, DC: Mathematics Association of America, pp. 206–248.
[25] Smith, H. (1998). A parametrix construction for wave equations with $C^{1,1}$ coefficients. *Ann. Inst. Fourier Grenoble* 48:797–835.