COMMUN. MATH. SCI.
Vol. 19, No. 3, pp. 585–606

ON THE CONVERGENCE OF FROZEN GAUSSIAN APPROXIMATION FOR LINEAR NON-STRICTLY HYPERBOLIC SYSTEMS*

LIHUI CHAI†, JAMES C. HATELEY‡, EMMANUEL LORIN§, AND XU YANG¶

Abstract. Frozen Gaussian approximation (FGA) has been applied and numerically verified as an efficient tool to compute high-frequency wave propagation modeled by non-strictly hyperbolic systems, such as the elastic wave equations [J.C. Hateley, L. Chai, P. Tong and X. Yang, Geophys. J. Int., 216:1394–1412, 2019] and the Dirac system [L. Chai, E. Lorin and X. Yang, SIAM J. Numer. Anal., 57:2383–2412, 2019]. However, the theory of convergence is still incomplete for non-strictly hyperbolic systems, where the latter can be interpreted as a diabatic (or more) coupling. In this paper, we establish the convergence theory for FGA for linear non-strictly hyperbolic systems, with an emphasis on the elastic wave equations and the Dirac system. Unlike the convergence theory of FGA for strictly linear hyperbolic systems, the key estimate lies in the boundedness of intraband transitions in diabatic coupling.

Keywords. Frozen Gaussian approximation; Convergence; Non-strictly hyperbolic; Elastic wave equations; Dirac equation.

AMS subject classifications. 65M12; 81Q05.

1. Introduction

The goal of this paper is to provide the convergence theory of frozen Gaussian approximation (FGA) for linear non-strictly hyperbolic systems. In [13], the authors study in detail the convergence and the accuracy of FGA's applied to linear strictly hyperbolic systems in high frequency regime. However, several fundamental hyperbolic systems are not strictly hyperbolic, such as the elastic wave equations or the Dirac system modeling in particular quantum relativistic particles, see [18]. The paper aims to precisely study the boundedness of intraband transitions in the diabatic coupling, which is specific to non-strictly hyperbolic systems.

For the sake of clarity, we then shall first consider the two fundamental examples mentioned above: elastic wave equations and the Dirac system; then we will extend the arguments for general non-strictly hyperbolic systems by mainly focusing on the technical consequences due to the multiplicity of some eigenvalues of Jacobian matrices, and will refer to the appropriate references in the strictly hyperbolic case.

We start by introducing the elastic wave equations in three dimensions which models elastic wave propagation, as in [6]. We define the elastic wave system:

$$\text{Elastic wave system (EWS):}\quad \begin{cases} (\rho(\boldsymbol{x})\partial_t^2-\mathcal{L})\boldsymbol{u}(t,\boldsymbol{x})=0,\\ \boldsymbol{u}(0,\boldsymbol{x})=\boldsymbol{u}_0^{\varepsilon}(\boldsymbol{x}),\\ \partial_t\boldsymbol{u}(0,\boldsymbol{x})=\boldsymbol{u}_1^{\varepsilon}(\boldsymbol{x}),\end{cases} \tag{1.1}$$

*Received: April 17, 2020; Accepted (in revised form): September 21, 2020. Communicated by Jianfeng Lu.

†School of Mathematics, Sun Yat-sen University, Guangzhou, 510275, China (chailihui@mail.sysu.edu.cn).

‡Department of Mathematics, Vanderbilt University, Nashville, TN 37240, USA (hateleyjc@gmail.com).

§School of Mathematics and Statistics, Carleton University, Ottawa, Canada; Centre de Recherches Mathématiques, Université de Montréal, Montréal, Canada (elorin@math.carleton.ca).

¶Department of Mathematics, University of California, Santa Barbara, CA 93106, USA (xuyang@math.ucsb.edu).

where the operator $\mathcal{L}$ is given by

$$\mathcal{L}\boldsymbol{u}(t,\boldsymbol{x}) = (\lambda(\boldsymbol{x})+2\mu(\boldsymbol{x}))\nabla\big(\nabla\cdot\boldsymbol{u}(t,\boldsymbol{x})\big) - \mu(\boldsymbol{x})\nabla\times\big(\nabla\times\boldsymbol{u}(t,\boldsymbol{x})\big), \tag{1.2}$$

with the differential operators taken in the spatial variables, $\lambda,\mu:\mathbb{R}^3\to\mathbb{R}$ being the first and second Lamé parameters and $\rho:\mathbb{R}^3\to\mathbb{R}$ is a material density. We remark that the P-, S- wave speeds (*e.g.*, [6]) are given by

$$c_{\mathrm{p}}^2(\boldsymbol{x}) = \frac{\lambda(\boldsymbol{x})+2\mu(\boldsymbol{x})}{\rho(\boldsymbol{x})} \qquad \text{and} \qquad c_{\mathrm{s}}^2(\boldsymbol{x}) = \frac{\mu(\boldsymbol{x})}{\rho(\boldsymbol{x})}, \tag{1.3}$$

respectively. In addition, if one considers the EWS and defines the following quantities

$$\Theta(t,\boldsymbol{x}) = \nabla\cdot\boldsymbol{u}(t,\boldsymbol{x}), \qquad \Psi(t,\boldsymbol{x}) = \nabla\times\boldsymbol{u}(t,\boldsymbol{x}), \qquad \boldsymbol{v}(t,\boldsymbol{x}) = \partial_t\boldsymbol{u}(t,\boldsymbol{x}), \tag{1.4}$$

with $\boldsymbol{v}=(v_1,v_2,v_3)^T$ and $\Psi=(\Psi_1,\Psi_2,\Psi_3)^T$ with $X=(v_1,v_2,v_3,\Theta,\Psi_1,\Psi_2,\Psi_3)^T$, then the elastic wave equations can be written as a matrix system,

$$\partial_t X = M_x\partial_x X + M_y\partial_y X + M_z\partial_z X, \tag{1.5}$$

where, using sparse notation; *e.g.*, $M_{ij}=v$ is denoted (i,j,v), M_x,M_y,M_z are as follows:

$$\begin{aligned} M_x &: (1,4,c_{\mathrm{p}}^2),(2,7,c_{\mathrm{s}}^2),(3,6,-c_{\mathrm{s}}^2),(4,1,1),(7,2,1),(6,3,-1),\\ M_y &: (1,7,-c_{\mathrm{s}}^2),(2,4,c_{\mathrm{p}}^2),(3,5,c_{\mathrm{s}}^2),(4,2,1),(5,3,1),(7,1,-1),\\ M_z &: (1,6,c_{\mathrm{s}}^2),(2,5,-c_{\mathrm{s}}^2),(3,4,c_{\mathrm{p}}^2),(4,3,1),(6,1,1),(5,2,-1). \end{aligned} \tag{1.6}$$

It can be seen that Equation (1.5) is a non-strictly hyperbolic system; indeed, the eigenvalues of $M_x+M_y+M_z$ are $\pm c_{\mathrm{p}},0,\pm c_{\mathrm{s}}$, where $\pm c_{\mathrm{s}}$ have a multiplicity of 2.

Another fundamental non-strictly hyperbolic system that we shall study is the Dirac system, usually referred in the physics literature as the *Dirac equation*:

$$\text{Dirac System (DS):} \quad \begin{cases} \mathrm{i}\varepsilon\partial_t\boldsymbol{\psi}^{\boldsymbol{\varepsilon}}(t,\boldsymbol{x}) = \big(-\mathrm{i}c\varepsilon\hat{\boldsymbol{\sigma}}\cdot\nabla - \hat{\boldsymbol{\sigma}}\cdot\boldsymbol{A}(\boldsymbol{x}) + m\beta c^2 + V(\boldsymbol{x})\big)\boldsymbol{\psi}^{\boldsymbol{\varepsilon}}(t,\boldsymbol{x}),\\ \boldsymbol{\psi}^{\boldsymbol{\varepsilon}}(\boldsymbol{x},0) = \boldsymbol{\varphi}_I^{\varepsilon}(\boldsymbol{x}), \end{cases} \tag{1.7}$$

where $\boldsymbol{\psi}^{\boldsymbol{\varepsilon}}=(\psi_1^\varepsilon,\psi_2^\varepsilon,\psi_3^\varepsilon,\psi_4^\varepsilon)^T$, which takes its values in $\mathbb{C}^4$, is a 4-spinor, with the initial condition $\boldsymbol{\varphi}_I\in L^2(\mathbb{R}^d;\mathbb{C}^4)$. The Dirac matrices $\hat{\boldsymbol{\sigma}}=(\alpha_x,\alpha_y,\alpha_z)$, β are defined as follows. For $\gamma=x,y,z$,

$$\alpha_\gamma = \begin{bmatrix} 0 & \sigma_\gamma\\ \sigma_\gamma & 0\end{bmatrix}, \ \beta = \begin{bmatrix}\mathbb{I}_2 & 0\\ 0 & -\mathbb{I}_2\end{bmatrix}. \tag{1.8}$$

The σ_γ's are the 2×2 Pauli matrices defined as

$$\sigma_x = \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}, \ \sigma_y = \begin{bmatrix}0 & -\mathrm{i}\\ \mathrm{i} & 0\end{bmatrix} \text{ and } \sigma_z = \begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}, \tag{1.9}$$

and $\mathbb{I}_2$ is the 2×2 unit matrix. The momemtum operator is denoted $\boldsymbol{p}=-\mathrm{i}\nabla$. The speed of light c and fermion mass m are kept explicit. This equation models a relativistic electron of mass m subject to an interaction potential V and an electromagnetic field $\boldsymbol{A}$. We set

$$B = -\hat{\boldsymbol{\sigma}}\cdot\boldsymbol{A} + m\beta c^2 + V, \tag{1.10}$$

and then the Dirac operator can be written as

$$\mathcal{D} = -\mathrm{i}c\varepsilon\hat{\boldsymbol{\sigma}}\cdot\nabla + B, \tag{1.11}$$

and its corresponding semi-classical symbol which reads

$$D(\boldsymbol{q},\boldsymbol{p}) = \hat{\boldsymbol{\sigma}}\cdot(\boldsymbol{p}c - \boldsymbol{A}(\boldsymbol{q})) + m\hat{\boldsymbol{\sigma}}_0 c^2 + V(\boldsymbol{q}) = \hat{\boldsymbol{\sigma}}\cdot\boldsymbol{p}c + B(\boldsymbol{q}), \tag{1.12}$$

is a Hermitian matrix which has two double eigenvalues

$$h_\pm(\boldsymbol{q},\boldsymbol{p}) = \pm\sqrt{|\boldsymbol{p}c - \boldsymbol{A}(\boldsymbol{q})|^2 + c^4} + V(\boldsymbol{q}), \tag{1.13}$$

with the corresponding normalized eigenvectors denoted as $\boldsymbol{\Upsilon}_{\pm 1}$ and $\boldsymbol{\Upsilon}_{\pm 2}$. Notice that the eigenvalues are usually called as *energy bands* in quantum mechanics, and as the light speed c is not zero, there is always a positive band gap, that is, $h_+ - h_- \geq 2c^2$. For more details of the computation for the eigenvalues please refer to [2].

The main result of this paper is the proof that FGA for both EWS [6] and DS [2] are first-order convergent, as for strictly hyperbolic systems [14]. The proof will follow the same machinery from [14], with however more careful estimates for the boundedness of intraband transitions in diabatic coupling. The main theorem for which we will provide a detailed proof, reads as follows,

MAIN THEOREM 1.1. *Let $\{\boldsymbol{u}_0^\varepsilon\}$ be a family of asymptotically high frequency initial conditions. Let $\boldsymbol{u}:[0,T]\times\mathbb{R}^d$ satisfies respective hyperbolic system (EWS) (1.1) or (DS) (1.7).*

- *(For elastic wave system). Let $\boldsymbol{u}_0^\varepsilon \in H_0^1(\mathbb{R}^d)$ be uniformly bounded, i.e. $\|\boldsymbol{u}_0^\varepsilon\|_{H_0^1} < M$, and let $\boldsymbol{u}_{\mathrm{F}}$ be the FGA to (1.1), then*

$$\sup_{t\in[0,T]} \|\boldsymbol{u}(t,\cdot) - \boldsymbol{u}_{\mathrm{F}}(t,\cdot)\|_{\mathrm{E}} \leq \varepsilon C_T,$$

  where the norm $\|\cdot\|_{\mathrm{E}}$ is a scaled semi-norm defined in (3.28).

- *(For Dirac system). Let $\boldsymbol{u}_0^\varepsilon \in L^2(\mathbb{R}^d)$ be uniformly bounded, and $\boldsymbol{u}_{\mathrm{F}}$ is the FGA to (1.7), then*

$$\sup_{t\in[0,T]} \|\boldsymbol{u}(t,\cdot) - \boldsymbol{u}_{\mathrm{F}}(t,\cdot)\|_{L^2} \leq \varepsilon C_T.$$

  In each case, C_T is a constant depending on the final time T.

Related works. The FGA, introduced originally in quantum mechanics as the Herman-Kluk (HK) propagator [7–9], was used to approximate the solution of the Schrödinger equation in the semi-classical regime. The mathematical analysis was then proposed in [16,17] to show the accuracy and efficiency of the HK ansatz, in particular, when the initial data are localized in phase space. HK formalism was later developed for several types of partial differential equations, such as the wave equations [12], linear hyperbolic systems of conservation laws [13], elastic wave equations, and seismic tomography [3, 4, 6]. The FGA for the elastic wave equations has been used to train neural networks for seismic interface and pocket detection [5]. Some applications and analysis on the Schrödinger equations were also proposed in [10, 11, 19].

Organization of the paper. In Section 2, we introduce the necessary notations and preliminaries needed for phase plane analysis. We present the full convergence analysis for the elastic wave system in Section 3, and for the Dirac system in Section 4. In Section 5, we provide the key arguments which allow for generalizing the convergence statements to any linear non-strictly hyperbolic system. In Section 6, we propose some concluding remarks.

2. Notations and preliminaries

In this section, we introduce some notations and preliminary results that are needed for the convergence analysis of both the elastic wave equations and the Dirac system. We denote $\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^d$ as spatial variables, $(\boldsymbol{q},\boldsymbol{p})\in\mathbb{R}^{2d}$ for the position and momentum variables, respectively, in the phase space.

We use hereafter the notation $\mathcal{O}(\varepsilon^\infty)$: $A^\varepsilon=\mathcal{O}(\varepsilon^\infty)$ meaning for any $k\in\mathbb{N}$

$$\lim_{\varepsilon\to 0}\varepsilon^{-k}|A^\varepsilon|=0.$$

Notation C will be used as a general positive constant, that can vary from line to line. The explicit value of this finite constant is however irrelevant in the analysis. We will use subscripts to denote constant dependence, e.g. C_T, is a constant that depends on the parameter T. We will respectively denote by $\mathcal{S}$, C^∞ and C_c^∞, the Schwartz class, smooth and compacted supported smooth function spaces. For generality, we will often use $\mathbb{R}^d$ as a d-dimensional Euclidean space; however, for the actual equations and computations we set $d=3$, as we deal with these differential operators on $\mathbb{R}^3$.

2.1. Wave packet decomposition.

For any $(\boldsymbol{q},\boldsymbol{p})\in\mathbb{R}^{2d}$, we define $\phi^\varepsilon_{\boldsymbol{q},\boldsymbol{p}}$ as

$$\phi^\varepsilon_{\boldsymbol{q},\boldsymbol{p}}(\boldsymbol{x})=(-2\pi\varepsilon)^{-d/2}\exp\Big(\mathrm{i}\boldsymbol{p}\cdot(\boldsymbol{x}-\boldsymbol{q})/\varepsilon-|\boldsymbol{x}-\boldsymbol{q}|^2/(2\varepsilon)\Big). \tag{2.1}$$

We recall that the Fourier-Bros-Iagolnitzer (FBI) transform on $\mathcal{S}(\mathbb{R}^d)$ [15], is defined as

$$\begin{aligned}(\mathcal{F}^\varepsilon f)(\boldsymbol{q},\boldsymbol{p})&=(\pi\varepsilon)^{-d/4}\langle\psi^\varepsilon_{\boldsymbol{q},\boldsymbol{p}},f\rangle\\ &=2^{-d/2}(\pi\varepsilon)^{-3d/4}\int_{\mathbb{R}^d}\exp\Big(\mathrm{i}\boldsymbol{p}\cdot(\boldsymbol{x}-\boldsymbol{q})/\varepsilon-|\boldsymbol{x}-\boldsymbol{q}|^2/(2\varepsilon)\Big)f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.\end{aligned}$$

The inverse transform $\big(\mathcal{F}^\varepsilon\big)^*$ defined on $\mathcal{S}(\mathbb{R}^{2d})$ is given by

$$\Big(\big(\mathcal{F}^\varepsilon\big)^*g\Big)(\boldsymbol{x})=2^{-d/2}(\pi\varepsilon)^{-3d/4}\int_{\mathbb{R}^{2d}}\exp\Big(\mathrm{i}\boldsymbol{p}\cdot(\boldsymbol{x}-\boldsymbol{q})/\varepsilon-|\boldsymbol{x}-\boldsymbol{q}|^2/(2\varepsilon)\Big)g(\boldsymbol{q},\boldsymbol{p})\,\mathrm{d}\boldsymbol{p}\mathrm{d}\boldsymbol{q}. \tag{2.2}$$

The following is a standard result from microlocal analysis, see for instance [1].

PROPOSITION 2.1. *For Schwartz class functions, the FBI transform is an isometry on $\mathbb{R}^d$, i.e., for any $f\in\mathcal{S}(\mathbb{R}^d)$,*

$$\|\mathcal{F}^\varepsilon f\|_{\mathrm{L}^{2d}}=\|f\|_{\mathrm{L}^d}. \tag{2.3}$$

Furthermore; $\big(\mathcal{F}^\varepsilon\big)^*\mathcal{F}^\varepsilon=\mathrm{Id}_{\mathrm{L}^2(\mathbb{R}^d)}$. By standard density arguments, this implies that the domain of $\mathcal{F}^\varepsilon$ and $\big(\mathcal{F}^\varepsilon\big)^*$ can be extended to $\mathrm{L}^2(\mathbb{R}^d)$ and $\mathrm{L}^2(\mathbb{R}^{2d})$ respectively.

We define the set closed set $K_\delta\subset\mathbb{R}^{2d}$ as follows. For $\delta>0$,

$$K_\delta=\Big\{(\boldsymbol{q},\boldsymbol{p})\in\mathbb{R}^{2d}:|\boldsymbol{q}|\le 1/\delta,\ \ \delta\le|\boldsymbol{p}|\le 1/\delta\Big\}. \tag{2.4}$$

DEFINITION 2.1. *Let $\{\boldsymbol{u}^\varepsilon\}\subset\mathrm{L}^2(\mathbb{R}^d)$ be a family of functions which is uniformly bounded. Given $\delta>0$, $\{\boldsymbol{u}^\varepsilon\}$ is asymptotically high frequency with cut off δ*, if

$$\int_{\mathbb{R}^{2d}\setminus K_\delta}|(\mathcal{F}^\varepsilon\boldsymbol{u}^\varepsilon)(\boldsymbol{q},\boldsymbol{p})|^2\,\mathrm{d}\boldsymbol{p}\mathrm{d}\boldsymbol{q}=\mathcal{O}(\varepsilon^\infty), \tag{2.5}$$

as $\varepsilon \to 0$.

Definition 2.2. *For $M_n \in \mathrm{L}^\infty(\mathbb{R}^{2d};\mathbb{C}^{N\times N})$ and a Schwartz function $\boldsymbol{u}\in\mathcal{S}(\mathbb{R}^d;\mathbb{C}^N)$, for each $n=1,\dots,N$ we define the Fourier integral operator (FIO) $\mathcal{I}_n^\varepsilon(t,M)\boldsymbol{u}$ as*

$$(\mathcal{I}_n^\varepsilon(t,M)\boldsymbol{u})(\boldsymbol{x}) = (2\pi\varepsilon)^{-3d/2}\int_{\mathbb{R}^{3d}} G_n^\varepsilon(t,\boldsymbol{x},\boldsymbol{y},\boldsymbol{p},\boldsymbol{q})M(\boldsymbol{q},\boldsymbol{p})\boldsymbol{u}(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}\boldsymbol{p}\,\mathrm{d}\boldsymbol{q}, \tag{2.6}$$

with

$$G_n^\varepsilon(t,\boldsymbol{x},\boldsymbol{y},\boldsymbol{p},\boldsymbol{q}) = e^{\mathrm{i}\phi_n(t,\boldsymbol{x},\boldsymbol{y},\boldsymbol{p},\boldsymbol{q})/\varepsilon}, \tag{2.7}$$

where the phase function is defined as

$$\begin{aligned}\phi_n(t,\boldsymbol{x},\boldsymbol{y},\boldsymbol{p},\boldsymbol{q}) = \frac{\mathrm{i}}{2}|\boldsymbol{y}-\boldsymbol{q}|^2 - \boldsymbol{p}\cdot(\boldsymbol{y}-\boldsymbol{q}) + \frac{\mathrm{i}}{2}|\boldsymbol{x}-\boldsymbol{Q}_n(t,\boldsymbol{q},\boldsymbol{p})|^2 \\ + \boldsymbol{P}_n(t,\boldsymbol{q},\boldsymbol{p})\cdot(\boldsymbol{x}-\boldsymbol{Q}_n(t,\boldsymbol{q},\boldsymbol{p})).\end{aligned} \tag{2.8}$$

Proposition 2.2. *If $M\in\mathrm{L}^\infty(\mathbb{R}^{2d};\mathbb{C}^{N\times N})$, for any t and each $n=1,\dots,N$, $\mathcal{I}_n^\varepsilon(t,M)$ can be extended to a bounded linear operator on $\mathrm{L}^2(\mathbb{R}^d;\mathbb{C}^N)$ with the bound*

$$\|\mathcal{I}_n^\varepsilon(t,M)\|_{\mathcal{L}(\mathrm{L}^2(\mathbb{R}^d;\mathbb{C}^N))} \le 2^{-d/2}\|M\|_{\mathrm{L}^\infty(\mathbb{R}^{2d};\mathbb{C}^{N\times N})}. \tag{2.9}$$

This is Proposition 3.7 in [13], with a more general version proved in [17, Theorem 2].

2.2. Preliminaries. The strategy for proving the convergence consists in estimating the error generated by the FGA to the correct order. There are several technical details which are required, and listed in this subsection. For the non-strict hyperbolic case, the FGA and filtered FGA possess asymptotic correction terms which allow for dealing with the multiplicity of eigenvalues. Operators derived from these asymptotic correction terms need to be bounded in the appropriate sense so that estimations can be made to the correct order.

In the following, we consider a non-strictly hyperbolic system, and assume, for any $(\boldsymbol{q},\boldsymbol{p})\in\mathbb{R}^{2d}$, that the system has N eigenvalues, denoted by the Hamiltonian $H_n(\boldsymbol{q},\boldsymbol{p})_{n=1}^N$ with the associated flow

$$\begin{cases} \dfrac{\mathrm{d}\boldsymbol{Q}_n(t,q,p)}{\mathrm{d}t} = \partial_{\boldsymbol{P}_n} H_n(\boldsymbol{Q}_n(t,\boldsymbol{q},\boldsymbol{p}),\boldsymbol{P}_n(t,\boldsymbol{q},\boldsymbol{p})), \\ \dfrac{\mathrm{d}\boldsymbol{P}_n(t,q,p)}{\mathrm{d}t} = -\partial_{\boldsymbol{Q}_n} H_n(\boldsymbol{Q}_n(t,\boldsymbol{q},\boldsymbol{p}),\boldsymbol{P}_n(t,\boldsymbol{q},\boldsymbol{p})), \end{cases} \tag{2.10}$$

with initial conditions $\boldsymbol{Q}_n(0,\boldsymbol{q},\boldsymbol{p})=\boldsymbol{q}$ and $\boldsymbol{P}_n(0,\boldsymbol{q},\boldsymbol{p})=\boldsymbol{p}$.

Definition 2.3. *A map $\kappa_n:(\boldsymbol{q},\boldsymbol{p})\to\big(\boldsymbol{Q}_n(\boldsymbol{q},\boldsymbol{p}),\boldsymbol{P}_n(\boldsymbol{q},\boldsymbol{p})\big)$ is called a canonical transformation if the associated Jacobian matrix is symplectic, i.e., for any $(\boldsymbol{q},\boldsymbol{p})$*

$$J_n(\boldsymbol{q},\boldsymbol{p}) = \begin{pmatrix} (\partial_q\boldsymbol{Q}_n)^T(\boldsymbol{q},\boldsymbol{p}) & (\partial_p\boldsymbol{Q}_n)^T(\boldsymbol{q},\boldsymbol{p}) \\ (\partial_q\boldsymbol{P}_n)^T(\boldsymbol{q},\boldsymbol{p}) & (\partial_p\boldsymbol{P}_n)^T(\boldsymbol{q},\boldsymbol{p}) \end{pmatrix},$$

is such that

$$J_n^T\begin{pmatrix} 0 & \mathrm{Id}_3 \\ -\mathrm{Id}_3 & 0\end{pmatrix}J_n = \begin{pmatrix} 0 & \mathrm{Id}_3 \\ -\mathrm{Id}_3 & 0\end{pmatrix}, \tag{2.11}$$

where Id_3 is a 3×3 identity matrix.

The following Propositions 2.3 and 2.4 are Proposition 3.1 and Proposition 3.4 in [13] respectively, the proofs are omitted. We will also use Assumption A from [13], and it is recalled below:

Assumption A. For each $n=1,\dots,N$, there exists a constant $C>0$, so that the Hamiltonian H_n satisfies for any $(\boldsymbol{q},\boldsymbol{p})\in\mathbb{R}^{2d}$ with $|\boldsymbol{p}|>0$,

$$|\boldsymbol{p}\cdot\partial_{\boldsymbol{q}}H_n(\boldsymbol{q},\boldsymbol{p})|\le C|\boldsymbol{p}|^2 \qquad \text{and} \qquad |\boldsymbol{q}\cdot\partial_{\boldsymbol{p}}H_n(\boldsymbol{q},\boldsymbol{p})|\le C|\boldsymbol{q}|^2. \tag{2.12}$$

PROPOSITION 2.3. *Given a canonical transformation κ_n, for $T>0$ and $\delta>0$, there is a constant $\delta_T>0$, such that*

$$(\boldsymbol{Q}_n(t,\boldsymbol{q},\boldsymbol{p}),\boldsymbol{P}_n(t,\boldsymbol{q},\boldsymbol{p}))\in K_{\delta_T}, \tag{2.13}$$

for any $(\boldsymbol{q},\boldsymbol{p})\in K_\delta$ and $t\in[0,T]$.

PROPOSITION 2.4. *The map κ is a canonical transform for any $T,\delta>0$; furthermore it has a bounded sup-norm.*

For a canonical transform κ_n define the quantity Z_n as

$$Z_n(t,\boldsymbol{q},\boldsymbol{p})=\partial_{\boldsymbol{z}}(\boldsymbol{Q}(t,\boldsymbol{q},\boldsymbol{p})+\mathrm{i}\boldsymbol{P}(t,\boldsymbol{q},\boldsymbol{p})), \tag{2.14}$$

with $\partial_{\boldsymbol{z}}=(\partial_{\boldsymbol{q}}-\mathrm{i}\partial_{\boldsymbol{p}})$. Note that

$$Z_n=(\mathrm{iId}_3\;\;\mathrm{Id}_3)\begin{pmatrix}\partial_{\boldsymbol{q}}\boldsymbol{Q}_n & \partial_{\boldsymbol{q}}\boldsymbol{P}_n\\ \partial_{\boldsymbol{p}}\boldsymbol{Q}_n & \partial_{\boldsymbol{p}}\boldsymbol{P}_n\end{pmatrix}\begin{pmatrix}-\mathrm{iId}_3\\ \mathrm{Id}_3\end{pmatrix}. \tag{2.15}$$

The following compact notation will be useful hereafter.

DEFINITION 2.4. *For $\mathbf{a}\in C^\infty(\Omega,\mathbb{C})$, define for $k\in\mathbb{N}$*

$$\Lambda_{k,\Omega}(\mathbf{a}):=\max_{|\alpha_{\boldsymbol{p}}|+|\alpha_{\boldsymbol{q}}|=k}\;\sup_{(\boldsymbol{q},\boldsymbol{p})\in\Omega}|\partial_{\boldsymbol{q}}^{\alpha_{\boldsymbol{q}}}\partial_{\boldsymbol{p}}^{\alpha_{\boldsymbol{p}}}\mathbf{a}(\boldsymbol{q},\boldsymbol{p})|, \tag{2.16}$$

with $\alpha_{\boldsymbol{q}}$, and $\alpha_{\boldsymbol{q}}$ being multi-indices corresponding to $\boldsymbol{q}$ and $\boldsymbol{p}$ respectively. By convention, we denote $\Lambda_k=\Lambda_{k,\mathbb{R}^{2d}}$.

We will also need the following technical lemmas; see Lemma 5.1 and Lemma 5.2 in [13] for details of the proof.

LEMMA 2.1. *Z_n is invertible for $(\boldsymbol{q},\boldsymbol{p})\in\mathbb{R}^{2d}$ with $|\boldsymbol{p}|>0$. Furthermore, for any $k\ge 0$ and $\delta>0$, there exist constants $C_{k,\delta}>0$ such that*

$$\Lambda_{k,K_\delta}\left((Z_n(t,\boldsymbol{q},\boldsymbol{p}))^{-1}\right)\le C_{k,\delta}. \tag{2.17}$$

LEMMA 2.2. *We introduce the notation $f\sim g$ to mean*

$$\int_{\mathbb{R}^{3d}}f(\boldsymbol{y})G^\varepsilon(t,\boldsymbol{x},\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}\boldsymbol{p}\,\mathrm{d}\boldsymbol{q}=\int_{\mathbb{R}^{3d}}g(\boldsymbol{y})G^\varepsilon(t,\boldsymbol{x},\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}\boldsymbol{p}\,\mathrm{d}\boldsymbol{q}. \tag{2.18}$$

For any vector $\boldsymbol{a}(\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})=(a_j)$, matrix $M(\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})=(M_{jk})$, and tensor $T(\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})=(T_{ijk})$ in Schwartz class, one has the following integration by parts formula in the component-

wise form, with $\partial_{\boldsymbol{z}} = (\partial_{z_1}, \partial_{z_2}, \partial_{z_3})$,

$$\begin{aligned}
a_j(\boldsymbol{x}-\boldsymbol{Q})_j &\sim -\varepsilon\partial_{z_m}\left(a_j Z_{jm}^{-1}\right),\\
(\boldsymbol{x}-\boldsymbol{Q})_j(\boldsymbol{x}-\boldsymbol{Q})_k M_{jk} &\sim \varepsilon\partial_{z_n} Q_j M_{jk} Z_{kn}^{-1} + \varepsilon^2\partial_{z_m}\left(\partial_{z_n}\left(M_{jk}Z_{kn}^{-1}\right)Z_{jm}^{-1}\right),\\
(\boldsymbol{x}-\boldsymbol{Q})_i(\boldsymbol{x}-\boldsymbol{Q})_j(\boldsymbol{x}-\boldsymbol{Q})_k T_{ijk} &\sim -\varepsilon^2\partial_{z_n}\left(\partial_{z_l} Q_j T_{ijk} Z_{il}^{-1} Z_{kn}^{-1}\right)\\
&\quad -\varepsilon^2\partial_{z_m}\left(\partial_{z_l} Q_k T_{ijk} Z_{il}^{-1} Z_{jm}^{-1}\right)\\
&\quad -\varepsilon^2\partial_{z_n} Q_j \partial_{z_l}\left(T_{ijk} Z_{il}^{-1}\right) Z_{kn}^{-1}\\
&\quad -\varepsilon^3\partial_{z_m}\left(\partial_{z_n}\left(\partial_{z_l}\left(T_{ijk}Z_{il}^{-1}\right)Z_{kn}^{-1}\right)Z_{jm}^{-1}\right).
\end{aligned}$$

3. Convergence analysis for the elastic wave equations

In this section, we first introduce the Hamiltonian flow associated to the FGA formulation of the elastic wave equations (EWS) described in (1.1), and related boundedness estimates on the quantities used in the FGA formulation. Then we compute the asymptotic corrections, and prove that these correction terms are bounded in proper norms, which eventually implies the convergence results.

3.1. Hamiltonian flow for EWS. According to the results in [6], the Hamiltonian associated with Θ, Ψ in the FGA formulation for both P- and S-waves are

$$H_{\mathrm{p}\pm,\mathrm{s}\pm}(t,\boldsymbol{Q},\boldsymbol{P}) = \pm c_{\mathrm{p,s}}(\boldsymbol{Q}_{\mathrm{p}\pm,\mathrm{s}\pm}(t,\boldsymbol{q},\boldsymbol{p}))\boldsymbol{P}_{\mathrm{p}\pm,\mathrm{s}\pm}(t,\boldsymbol{q},\boldsymbol{p}), \tag{3.1}$$

where the wave speeds $c_{\mathrm{p,s}}$ are given in (1.3). The corresponding flows are given by

$$\begin{cases}
\dfrac{\mathrm{d}\boldsymbol{Q}_{\mathrm{p}\pm,\mathrm{s}\pm}}{\mathrm{d}t}(0,\boldsymbol{q},\boldsymbol{p}) & = \pm c_{\mathrm{p,s}}(\boldsymbol{Q}_{\mathrm{p}\pm,\mathrm{s}\pm}(t,\boldsymbol{q},\boldsymbol{p}))\dfrac{\boldsymbol{P}_{\mathrm{p}\pm,\mathrm{s}\pm}(t,\boldsymbol{q},\boldsymbol{p})}{|\boldsymbol{P}_{\mathrm{p}\pm,\mathrm{s}\pm}(t,\boldsymbol{q},\boldsymbol{p})|},\\
\dfrac{\mathrm{d}\boldsymbol{P}_{\mathrm{p}\pm,\mathrm{s}\pm}}{\mathrm{d}t}(0,\boldsymbol{q},\boldsymbol{p}) & = \mp\partial_{\boldsymbol{Q}} c_{\mathrm{p,s}}(\boldsymbol{Q}_{\mathrm{p}\pm,\mathrm{s}\pm}(t,\boldsymbol{q},\boldsymbol{p}))|\boldsymbol{P}_{\mathrm{p}\pm,\mathrm{s}\pm}(t,\boldsymbol{q},\boldsymbol{p})|,
\end{cases} \tag{3.2}$$

with initial conditions

$$\boldsymbol{Q}_{\mathrm{p}\pm,\mathrm{s}\pm}(0,\boldsymbol{q},\boldsymbol{p}) = \boldsymbol{q} \text{ and } \boldsymbol{P}_{\mathrm{p}\pm,\mathrm{s}\pm}(0,\boldsymbol{q},\boldsymbol{p}) = \boldsymbol{p}. \tag{3.3}$$

We remark that

$$|\boldsymbol{p}\cdot\partial_{\boldsymbol{q}} H(t,\boldsymbol{q},\boldsymbol{p})| \lesssim |\boldsymbol{p}|^2, \qquad \text{and} \qquad |\boldsymbol{q}\cdot\partial_{\boldsymbol{p}} H(t,\boldsymbol{q},\boldsymbol{p})| \lesssim |\boldsymbol{q}|^2, \tag{3.4}$$

so that the global Lipschitz assumption A is satisfied.

For all practical purposes, the set K_δ is bounding the position and the magnitude for the direction of propagation of the wave packets. For the elastic wave system Equation (1.1), for the set K_δ, upper bounds on $|\boldsymbol{q}|$ and $|\boldsymbol{p}|$ are reasonable as any computational domain will be a finite domain. Furthermore $\boldsymbol{p}$ is bounded away from zero; as if $\boldsymbol{p}=0$, the wave packet does not propagate and the Hamiltonian system is degenerate, i.e., $H=0$.

Thus far, notations and relatively standard estimates for the analysis of FGA have been introduced. In the following, we present new estimates valid for the non-strictly hyperbolic elastic wave equations.

3.2. Next order corrections of FGA for EWS. According to [6], the first-order FGA is

$$\boldsymbol{u}_{\mathrm{F},0}(t,\boldsymbol{x}) = (2\pi\varepsilon)^{-3d/2}\int_{\mathbb{R}^{3d}}\sum_{b=\pm}\Big[\mathbf{a}_{\mathrm{p},b,0}(t,\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})G^{\varepsilon}_{\mathrm{p},b}(t,\boldsymbol{x},\boldsymbol{y},\boldsymbol{q},\boldsymbol{p}) + (\mathbf{a}_{\mathrm{sh},b,0}(t,\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})+\mathbf{a}_{\mathrm{sv},b,0}(t,\boldsymbol{y},\boldsymbol{q},\boldsymbol{p}))G^{\varepsilon}_{\mathrm{s},b}(t,\boldsymbol{x},\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})\Big]\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}\boldsymbol{p}\,\mathrm{d}\boldsymbol{q}. \tag{3.5}$$

As the computations for the branches and P-,S- wavefields are similar, in the following we will either omit the subscript, or will simply subscript using the index $n=\{1,\cdots,6\}$ instead of (p±,sh±,sh±) or (p±,s±). With this notation, we can define Equation (3.5) more compactly, as

$$\boldsymbol{u}_{\mathrm{F},0}(t,\boldsymbol{x}) = (2\pi\varepsilon)^{-3d/2}\int_{\mathbb{R}^{3d}}\sum_{n=1}^{6}\mathbf{a}_{n,0}(t,\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})G^{\varepsilon}_{n}(t,\boldsymbol{x},\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}\boldsymbol{p}\,\mathrm{d}\boldsymbol{q}. \tag{3.6}$$

For $k>1$, define the k-th ordered FGA with a correction term as

$$\boldsymbol{u}_{\mathrm{F},k}(t,\boldsymbol{x}) = \boldsymbol{u}_{\mathrm{F},0}(t,\boldsymbol{x}) + (2\pi\varepsilon)^{-3d/2}\int_{\mathbb{R}^{3d}}\sum_{j=1,n=1}^{k,6}\varepsilon^{j}\left(\mathbf{a}_{n,j}(t,\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})+\mathbf{a}^{\perp}_{n,j}(t,\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})\right) \times G^{\varepsilon}_{n}(t,\boldsymbol{x},\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}\boldsymbol{p}\,\mathrm{d}\boldsymbol{q}, \tag{3.7}$$

where the terms $\mathbf{a}^{\perp}_{n,1}$ will be defined later. We next define a standard smooth cutoff function $\chi_\delta:\mathbb{R}^{2d}\to[0,1]$ for the set K_δ as

$$\chi_\delta(\boldsymbol{q},\boldsymbol{p}) = \begin{cases} 1, & (\boldsymbol{q},\boldsymbol{p})\in K_\delta, \\ 0, & (\boldsymbol{q},\boldsymbol{p})\in\mathbb{R}\backslash K_{\delta/2}, \end{cases} \tag{3.8}$$

and such that for any $k\in\mathbb{N}$, there exists a constant $C_{K,\delta}$ such that

$$\Lambda_k\left(\chi_\delta(\boldsymbol{q},\boldsymbol{p})\right) < C_{K,\delta}. \tag{3.9}$$

We define the filtered version of the FGA, as follows for $k\in\mathbb{N}$,

$$\tilde{u}_{\mathrm{F},k}(t,\boldsymbol{x}) = (2\pi\varepsilon)^{-3d/2}\int_{\mathbb{R}^{3d}}\chi_\delta(\boldsymbol{q},\boldsymbol{p})\sum_{j=0,n=1}^{k,6}\varepsilon^{j}\left(\mathbf{a}_{n,j}(t,\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})+\mathbf{a}^{\perp}_{n,j}(t,\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})\right) \times G^{\varepsilon}_{n}(t,\boldsymbol{x},\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}\boldsymbol{p}\,\mathrm{d}\boldsymbol{q}, \tag{3.10}$$

with $\mathbf{a}^{\perp}_{n,0}=0$. In the following we detail the construction of $\mathbf{a}^{\perp}_{n,1}$. Define the unit vectors $\hat{\boldsymbol{N}}_{\mathrm{p}\pm},\hat{\boldsymbol{N}}_{\mathrm{sv}\pm},\hat{\boldsymbol{N}}_{\mathrm{sh}\pm}$ that point in the direction of $\boldsymbol{P}$, $\boldsymbol{SV}$, or $\boldsymbol{SH}$ respectively. Then $\mathbf{a}_{n,0}(t,\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})$ is defined as follows,

$$\mathbf{a}_{n,0}(t,\boldsymbol{y},\boldsymbol{q},\boldsymbol{p}) = a_{n,0}(t,\boldsymbol{q},\boldsymbol{p})\alpha^{\varepsilon}_{n}(\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})\hat{\boldsymbol{N}}_n(t,\boldsymbol{q},\boldsymbol{p}), \tag{3.11}$$

where $\hat{\boldsymbol{N}}_n(0,\boldsymbol{q},\boldsymbol{p})=\hat{\boldsymbol{n}}_n$ and α^{ε} incorporates the initial conditions,

$$\alpha^{\varepsilon}_{n}(\boldsymbol{y},\boldsymbol{q},\boldsymbol{p}) = \frac{1}{2c_n|\boldsymbol{p}|^3}\left(\boldsymbol{u}^{\varepsilon}_0(\boldsymbol{y})c_n|\boldsymbol{p}|\pm\mathrm{i}\varepsilon\boldsymbol{u}^{\varepsilon}_1(\boldsymbol{y})\right)\cdot\hat{\boldsymbol{n}}_n. \tag{3.12}$$

The scalar functions $a_{n,0}$, with n representing for (p±,sh±,sh±) satisfy the following evolution equations [6],

$$\frac{\mathrm{d}a_{\mathrm{p}}}{\mathrm{d}t} = a_{\mathrm{p}}\left(\pm\frac{\partial_{\boldsymbol{Q}_{\mathrm{p}}}c_{\mathrm{p}}\cdot\boldsymbol{P}_{\mathrm{p}}}{|\boldsymbol{P}_{\mathrm{p}}|}+\frac{1}{2}\mathrm{Tr}\left(Z_{\mathrm{p}}^{-1}\frac{\mathrm{d}Z_{\mathrm{p}}}{\mathrm{d}t}\right)\right), \tag{3.13}$$

$$\frac{\mathrm{d}a_{\mathrm{sv}}}{\mathrm{d}t} = a_{\mathrm{sv}}\left(\pm\frac{\partial_{\boldsymbol{Q}_{\mathrm{s}}} c_{\mathrm{s}}\cdot \boldsymbol{P}_{\mathrm{s}}}{|\boldsymbol{P}_{\mathrm{s}}|} + \frac{1}{2}\mathrm{Tr}\left(Z_{\mathrm{s}}^{-1}\frac{\mathrm{d}Z_{\mathrm{s}}}{\mathrm{d}t}\right)\right) - a_{\mathrm{sh}}\frac{\mathrm{d}\hat{\boldsymbol{N}}_{\mathrm{sh}}}{\mathrm{d}t}\cdot\hat{\boldsymbol{N}}_{\mathrm{sv}}, \tag{3.14}$$

$$\frac{\mathrm{d}a_{\mathrm{sh}}}{\mathrm{d}t} = a_{\mathrm{sh}}\left(\pm\frac{\partial_{\boldsymbol{Q}_{\mathrm{s}}} c_{\mathrm{s}}\cdot \boldsymbol{P}_{\mathrm{s}}}{|\boldsymbol{P}_{\mathrm{s}}|} + \frac{1}{2}\mathrm{Tr}\left(Z_{\mathrm{s}}^{-1}\frac{\mathrm{d}Z_{\mathrm{s}}}{\mathrm{d}t}\right)\right) + a_{\mathrm{sv}}\frac{\mathrm{d}\hat{\boldsymbol{N}}_{\mathrm{sh}}}{\mathrm{d}t}\cdot\hat{\boldsymbol{N}}_{\mathrm{sv}}. \tag{3.15}$$

Equation (3.12) is derived from $\boldsymbol{u}_0(\boldsymbol{x})$, $\boldsymbol{u}_1(\boldsymbol{x})$ written in terms of FBI and inverse FBI transform, *i.e.*,

$$\boldsymbol{u}_0(\boldsymbol{x}) = (2\pi\varepsilon)^{-3d/2}\int_{\mathbb{R}^{3d}} \boldsymbol{u}_0(\boldsymbol{y}) G^{\varepsilon}(0,\boldsymbol{x},\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}\boldsymbol{p}\,\mathrm{d}\boldsymbol{q}, \tag{3.16}$$

and decomposing the integrand in terms of the basis $\{\hat{n}_{\mathrm{p}},\hat{n}_{\mathrm{sv}},\hat{n}_{\mathrm{sh}}\}$.

REMARK 3.1. It is easy to check that $\boldsymbol{u}_{\mathrm{F}}(0,\boldsymbol{x}) = \boldsymbol{u}_0(\boldsymbol{x})$ as $\mathcal{F}^*(\mathcal{F}(\boldsymbol{u}_0)) = \boldsymbol{u}_0$.

For the detailed derivation of Equations (3.13)–(3.15), see [6]. These lengthy calculations are necessary to arrive at the following operators for estimating the intraband transitions. Omitting the n index, we define the operators $\mathcal{L}_0,\mathcal{L}_1,\mathcal{L}_2$ acting on $\boldsymbol{A}$ as:

$$\mathcal{L}_0(\boldsymbol{A}) := -\rho\boldsymbol{A}(\boldsymbol{P}\cdot\boldsymbol{Q}_t)^2 + (\lambda+\mu)(\boldsymbol{A}\cdot\boldsymbol{P})\boldsymbol{P} + \mu(\boldsymbol{P}\cdot\boldsymbol{P})\boldsymbol{A}, \tag{3.17}$$

$$\begin{aligned}\mathcal{L}_1(\boldsymbol{A}) :=\, & \lambda\boldsymbol{A} + 2\mathrm{i}\boldsymbol{A}_t\boldsymbol{P}\cdot\boldsymbol{Q}_t - \mathrm{i}\rho\boldsymbol{A}(\boldsymbol{P}_t - \mathrm{i}\boldsymbol{Q}_t)\cdot\boldsymbol{Q}_t\\
& - \mathrm{i}(\lambda+\mu)\partial_{\boldsymbol{z}} : ((Z^{-1})^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}) - \mathrm{i}(\lambda+\mu)\partial_{\boldsymbol{z}} : (\boldsymbol{A}\cdot\boldsymbol{P}(Z^{-1})^{\mathrm{T}})\\
& + \partial_{\boldsymbol{z}} : ((Z^{-1})^{\mathrm{T}}\partial_{\boldsymbol{Q}}\rho\boldsymbol{A}(\boldsymbol{P}\cdot\boldsymbol{Q}_t)^2) - 2\mu\mathrm{i}\partial_{\boldsymbol{z}} : ((Z^{-1})^{\mathrm{T}}\boldsymbol{P}\boldsymbol{A})\\
& - \rho\boldsymbol{A}\mathrm{Tr}\left((\boldsymbol{P}_t - \mathrm{i}\boldsymbol{Q}_t)\otimes(\boldsymbol{P}_t - \mathrm{i}\boldsymbol{Q}_t)Z^{-1}\partial_{\boldsymbol{z}}\boldsymbol{Q}\right)\\
& - (\lambda+\mu)Z^{-1} : (\partial_{\boldsymbol{z}}\boldsymbol{Q}\boldsymbol{A}) - \mu\boldsymbol{A}\mathrm{Tr}(Z^{-1}\partial_{\boldsymbol{z}})\\
& - \frac{1}{2}\partial^2_{\boldsymbol{Q}\boldsymbol{Q}}\rho\boldsymbol{A}(\boldsymbol{P}\cdot\boldsymbol{Q}_t)^2\mathrm{Tr}(Z^{-1}\partial_{\boldsymbol{z}}),\end{aligned} \tag{3.18}$$

$$\begin{aligned}\mathcal{L}_2(\boldsymbol{A}) :=\, & \rho\boldsymbol{A}_{tt} - \rho\boldsymbol{A}_t\boldsymbol{P}\cdot\boldsymbol{Q}_t - \rho\partial_{\boldsymbol{z}} : (\rho(Z^{-1})^{\mathrm{T}}M_1) + \partial_{\boldsymbol{z}} : ((Z^{-1})^{\mathrm{T}}\partial_{\boldsymbol{Q}}\rho(\boldsymbol{P}_t - \mathrm{i}\boldsymbol{Q}_t)\cdot\boldsymbol{Q}_t)\boldsymbol{A})\\
& + \partial_{\boldsymbol{z}} : (\partial_{\boldsymbol{z}} : (\rho\boldsymbol{A}(\boldsymbol{P}_t - \mathrm{i}\boldsymbol{Q}_t)\otimes(\boldsymbol{P}_t - \mathrm{i}\boldsymbol{Q}_t)Z^{-1})Z^{-1}))\\
& - (\lambda+\mu)\partial_{\boldsymbol{z}} : (\partial_{\boldsymbol{z}} : (\partial_{\boldsymbol{z}}(\boldsymbol{Q}\boldsymbol{A})Z^{-1})Z^{-1})\\
& - \mu\partial_{\boldsymbol{z}} : (\partial_{\boldsymbol{z}} : (Z^{-1})\boldsymbol{A}Z^{-1}) + \partial_{\boldsymbol{Q}}\rho M_1\mathrm{Tr}(Z^{-1}\partial_{\boldsymbol{z}})\\
& - \mathrm{i}\frac{1}{2}\partial^2_{\boldsymbol{Q}\boldsymbol{Q}}\rho\boldsymbol{A}(\boldsymbol{P}_t - \mathrm{i}\boldsymbol{Q}_t)\cdot\boldsymbol{Q}_t\mathrm{Tr}(Z^{-1}\partial_{\boldsymbol{z}})\\
& - \partial_{\boldsymbol{z}} : (\partial_{\boldsymbol{z}} : (Z^{-1}\partial^2_{\boldsymbol{Q}\boldsymbol{Q}}\rho)\boldsymbol{A}(\boldsymbol{P}\cdot\boldsymbol{Q}_t)^2Z^{-1})\\
& + \partial_{\boldsymbol{z}} : \left(\mathrm{Tr}(\partial^2_{\boldsymbol{Q}\boldsymbol{Q}}\rho(Z^{-1})^{\mathrm{T}}\partial_{\boldsymbol{z}}\boldsymbol{Q})(\boldsymbol{P}\cdot\boldsymbol{Q}_t)\left[(\boldsymbol{P}_t - \mathrm{i}\boldsymbol{Q}_t)\right]\boldsymbol{A}\right),\end{aligned} \tag{3.19}$$

with the notation : being a contraction and M_1 defined as

$$M_1 = 2\mathrm{i}\boldsymbol{A}_t\otimes(\boldsymbol{P}_t - \mathrm{i}\boldsymbol{Q}_t) + \mathrm{i}\boldsymbol{A}\otimes(\boldsymbol{P}_{tt} - \mathrm{i}\boldsymbol{Q}_{tt}) + 2(\boldsymbol{P}\cdot\boldsymbol{Q}_t)\boldsymbol{A}\otimes(\boldsymbol{P}_t - \mathrm{i}\boldsymbol{Q}_t).$$

Now $(\partial_t^2 - \mathcal{L})\boldsymbol{u}_{\mathrm{F},1}$ can be written as, with $\mathcal{L}$ given by (1.2),

$$\begin{aligned}(\partial_t^2 - \mathcal{L})\boldsymbol{u}_{\mathrm{F},1} = (2\pi\varepsilon)^{-3/2}\sum_n\int_{\mathbb{R}^{3d}} & \Big(\varepsilon^{-2}\mathcal{L}_{n,0}\big(\mathbf{a}_{n,0} + \varepsilon\bar{\mathbf{a}}_{n,1}\big)\\
& \varepsilon^{-1}\mathcal{L}_{n,1}\big(\mathbf{a}_{n,0} + \varepsilon\bar{\mathbf{a}}_{n,1}\big) + \mathcal{L}_{n,2}\big(\mathbf{a}_{n,0} + \varepsilon\bar{\mathbf{a}}_{n,1}\big)\Big)G_n^{\varepsilon}\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}\boldsymbol{p}\,\mathrm{d}\boldsymbol{q}.\end{aligned} \tag{3.20}$$

where $\bar{\mathbf{a}}_{n,1} = \mathbf{a}_{n,1} + \mathbf{a}_{n,1}^{\perp}$. Substituting the dynamics for $\mathcal{L}_{n,0}$ reveals that $\mathcal{L}_{n,0}(\mathbf{a}_{n,0}) = 0$. Looking at the $\mathcal{O}(1/\varepsilon)$ term and equating to zero gives

$$\mathcal{L}_{n,1}(\mathbf{a}_{n,0}) = -\mathcal{L}_{n,0}(\bar{\mathbf{a}}_{n,1}). \tag{3.21}$$

Now $\mathcal{L}_{n,0}$ can be written as

$$\mathcal{L}_{n,0} = \left(\mu|\boldsymbol{P}_n|^2 - \rho(\boldsymbol{P}_n \cdot \partial_t \boldsymbol{Q}_n)^2\right)\mathrm{Id}_3 + (\lambda+\mu)\boldsymbol{P}_n \otimes \boldsymbol{P}_n, \tag{3.22}$$

which is a symmetric matrix with eigenvalues

$$\beta_{n,1} = (\lambda+2\mu)|\boldsymbol{P}_n|^2 - \rho|\boldsymbol{P}\cdot\partial_t \boldsymbol{Q}_n|^2, \qquad \beta_{n,2} = \beta_{n,3} = \mu|\boldsymbol{P}|^2 - \rho|\boldsymbol{P}_n \cdot \partial_t \boldsymbol{Q}_n|^2,$$

and the corresponding eigenvectors

$$\boldsymbol{P}_n = (p_{n,1}, p_{n,1}, p_{n,1}), \qquad \boldsymbol{d}_{n,1} = (-p_{n,2}, p_{n,1}, 0), \qquad \boldsymbol{d}_{n,2} = (-p_{n,3}, 0, p_{n,1}).$$

For the P-wave, $n = \mathrm{p}$, taking inner product of (3.21) with the eigenvector $\boldsymbol{P}_{\mathrm{p}}$ brings $\langle \boldsymbol{P}_{\mathrm{p}}, \mathcal{L}_{\mathrm{p},0}(\mathbf{a}_{\mathrm{p},1})\rangle = -\langle \boldsymbol{P}_{\mathrm{p}}, \mathcal{L}_{\mathrm{p},1}(\mathbf{a}_{\mathrm{p},0})\rangle$, which yields

$$\langle \mathcal{L}_{\mathrm{p},0}^{*}(\boldsymbol{P}_{\mathrm{p}}), \bar{\mathbf{a}}_{\mathrm{p},1}\rangle = \langle \mathcal{L}_{\mathrm{p},0}(\boldsymbol{P}), \bar{\mathbf{a}}_{\mathrm{p},1}\rangle = ((\lambda+2\mu)|\boldsymbol{P}_{\mathrm{p}}|^2 - \rho|\boldsymbol{P}_{\mathrm{p}}\cdot\partial_t \boldsymbol{Q}_{\mathrm{p}}|^2)\langle \boldsymbol{P}_{\mathrm{p}}, \bar{\mathbf{a}}_{\mathrm{p},1}\rangle = 0.$$

After plugging in Equation (3.2) one can recover the Equation (3.13) by $\langle \boldsymbol{P}_{\mathrm{p}}, \mathcal{L}_{\mathrm{p},1}(\mathbf{a}_{\mathrm{p},0})\rangle = 0$.

Considering $\boldsymbol{d}_{1,2}$ with a similar strategy shows,

$$\langle \mathcal{L}_{\mathrm{p},0}^{*}(\boldsymbol{d}_{1,2}), \bar{\mathbf{a}}_{\mathrm{p},1}\rangle = \langle \mathcal{L}_{\mathrm{p},0}(\boldsymbol{d}_{1,2}), \bar{\mathbf{a}}_{\mathrm{p},1}\rangle = (\mu|\boldsymbol{P}|^2 - \rho|\boldsymbol{P}\cdot\partial_t \boldsymbol{Q}|^2)\langle \boldsymbol{d}_{1,2}, \bar{\mathbf{a}}_{\mathrm{p},1}\rangle.$$

Plugging in the Hamiltonian flow (3.2) gives

$$\langle \boldsymbol{d}_{1,2}, \bar{\mathbf{a}}_{\mathrm{p},1}\rangle = \frac{1}{\rho(c_{\mathrm{s}}^2 - c_{\mathrm{p}}^2)|\boldsymbol{P}|^2}\langle \boldsymbol{d}_{1,2}, \mathcal{L}_{\mathrm{p},1}(\mathbf{a}_{\mathrm{p},0})\rangle.$$

Define the pseudo-inverse, for $\boldsymbol{v} \in \mathcal{S}(\mathbb{R}^3)$,

$$\mathcal{L}_{\mathrm{p},0}^{-1}(\boldsymbol{v}) = \frac{1}{\rho(c_{\mathrm{s}}^2 - c_{\mathrm{p}}^2)|\boldsymbol{P}|^2}\left(\langle \hat{\boldsymbol{d}}_1, \boldsymbol{v}\rangle \hat{\boldsymbol{d}}_1 + \langle \hat{\boldsymbol{d}}_2, \boldsymbol{v}\rangle \hat{\boldsymbol{d}}_2\right), \tag{3.23}$$

and define

$$\mathbf{a}_{\mathrm{p},1}^{\perp} = \mathcal{L}_{\mathrm{p},0}^{-1}\left((\mathrm{Id} - \Pi_{\mathrm{p}})\mathcal{L}_{\mathrm{p},1}(\mathbf{a}_{\mathrm{p},0})\right), \tag{3.24}$$

where Π_{p} is projection onto $\boldsymbol{P}_{\mathrm{p}}$.

For the S-wave, with $n = \mathrm{sv}, \mathrm{sh}$, from (3.21) one has

$$\mathcal{L}_{\mathrm{s},1}(\mathbf{a}_{\mathrm{s},0}) = -\mathcal{L}_{\mathrm{s},0}(\bar{\mathbf{a}}_{\mathrm{s},1}). \tag{3.25}$$

Let $\boldsymbol{d}_{s,1} = \hat{\boldsymbol{N}}_{\mathrm{sh}}$, taking inner product with (3.25) gives $\langle \mathcal{L}_{\mathrm{s},0}(\hat{\boldsymbol{N}}_{\mathrm{sv}}), \bar{\mathbf{a}}_{\mathrm{s},1}\rangle = (\mu|\boldsymbol{P}_{\mathrm{s}}|^2 - \rho|\boldsymbol{P}_{\mathrm{s}} \cdot \partial_t \boldsymbol{Q}_{\mathrm{s}}|^2)\langle \hat{\boldsymbol{N}}_{\mathrm{sv}}, \bar{\mathbf{a}}_{\mathrm{s},1}\rangle$, which is zero when the dynamics are substituted. From this one can get $\langle \hat{\boldsymbol{N}}_{\mathrm{sv}}, \mathcal{L}_{\mathrm{s},1}\mathbf{a}_{\mathrm{sv},0}\rangle = -\langle \hat{\boldsymbol{N}}_{\mathrm{sv}}, \mathcal{L}_{\mathrm{s},1}\mathbf{a}_{\mathrm{sh},0}\rangle$, which gives us Equation (3.14). Note that Equation (3.15) can be recovered in a similar manner.

Taking inner product with $\boldsymbol{P}_{\mathrm{s}}$ of (3.25) leads to

$$\langle \boldsymbol{P}, \bar{\mathbf{a}}_{\mathrm{s},1}\rangle = -\frac{1}{(\lambda+\mu)|\boldsymbol{P}|^2}\langle \boldsymbol{P}, \mathcal{L}_{\mathrm{s},1}(\mathbf{a}_{\mathrm{sv},0} + \mathbf{a}_{\mathrm{sh},0})\rangle.$$

Define the pseudo-inverse, for $\boldsymbol{v}\in\mathcal{S}(\mathbb{R}^3)$,

$$\mathcal{L}_{\mathrm{s},0}^{-1}(\boldsymbol{v})=-\frac{1}{(\lambda+\mu)|\boldsymbol{P}_{\mathrm{s}}|^2}\langle\hat{\boldsymbol{P}}_{\mathrm{s}},\boldsymbol{v}\rangle\hat{\boldsymbol{P}}_{\mathrm{s}}, \tag{3.26}$$

and then define

$$\mathbf{a}_{\mathrm{s},1}^{\perp}=\mathcal{L}_{\mathrm{s},0}^{-1}\big((\mathrm{Id}-\Pi_{\mathrm{s}})\mathcal{L}_{\mathrm{s},1}(\mathbf{a}_{\mathrm{s},0})\big), \tag{3.27}$$

with Π_{s} a projection onto the spanned space by $\boldsymbol{d}_{\mathrm{s},1}$ and $\boldsymbol{d}_{\mathrm{s},2}$.

REMARK 3.2. The existence of multiple eigenvalues makes the FGA different from the hyperbolic case. For the elastic wave equations, the S-wave mode has two orthogonal directions $\boldsymbol{SV}$ and $\boldsymbol{SH}$, and one can see that: (i) Equation (3.25) results in a coupled equation system (3.14) and (3.15) which presents the transitions between the $\boldsymbol{SV}$ and $\boldsymbol{SH}$ modes; (ii) the P-perpendicular term, *i.e.* , $\mathbf{a}_{\mathrm{p},1}^{\perp}$ should then belong to a double dimension subspace spanned by the two S-wave modes, and the convergence analysis will partially rely on the boundedness of this term.

In the next section, we will bound these pseudo-inverses and then further reach the convergence results.

3.3. Error estimates and main result for EWS. Thanks to the above computation in particular the explicit expression of $\mathbf{a}_{\mathrm{p}}^{\perp}$, $\mathbf{a}_{\mathrm{s}}^{\perp}$, $\mathcal{L}_{\mathrm{p},0}^{-1}$, and $\mathcal{L}_{\mathrm{s},0}^{-1}$, we derive some error estimates and eventually conclude on the convergence of FGA for the EWS.

DEFINITION 3.1. *Define the scaled semi-norm*

$$\|\boldsymbol{u}(t,\cdot)\|_{\mathrm{E}}=\varepsilon(\|\partial_t\boldsymbol{u}(t,\cdot)\|_{\mathrm{L}^2}+\|\nabla\cdot\boldsymbol{u}(t,\cdot)\|_{\mathrm{L}^2}+\|\nabla\times\boldsymbol{u}(t,\cdot)\|_{\mathrm{L}^2}). \tag{3.28}$$

PROPOSITION 3.1. *Let* $\mathbf{a}_{\mathrm{s}}=a_{\mathrm{sv}}\alpha_{\mathrm{sv}}\hat{\boldsymbol{N}}_{\mathrm{sv}}+a_{\mathrm{sh}}\alpha_{\mathrm{sh}}\hat{\boldsymbol{N}}_{\mathrm{sh}}$ *and* $\mathbf{a}_{\mathrm{p}}=a_{\mathrm{p}}\alpha_{\mathrm{p}}\hat{\boldsymbol{N}}_{\mathrm{p}}$. *The terms* $\mathbf{a}_{\mathrm{p}}$, $\mathbf{a}_{\mathrm{s}}$ *are bounded in the* L^2 *sense; furthermore,*

$$\|\boldsymbol{u}_{\mathrm{F},1}-\boldsymbol{u}_{\mathrm{F},0}\|_{\mathrm{E}}\le\varepsilon C_{T,\delta}.$$

Proof. First, we remark that

$$\|\mathbf{a}_n(t,\cdot)\|_{\mathrm{L}^2}\le\|\alpha_n\|_{\mathrm{L}^2}\|a_n(t,\cdot)\|_{\mathrm{L}^\infty}\ \text{and}\ \|\mathbf{a}_n(t,\cdot)\|_{\mathrm{L}^\infty}\lesssim\|a_n(t,\cdot)\|_{\mathrm{L}^\infty}.$$

From the definitions we have an immediate bound

$$\|\boldsymbol{u}_{\mathrm{F},1}(t,\cdot)-\boldsymbol{u}_{\mathrm{F},0}(t,\cdot)\|_{\mathrm{E}}\le(2\pi\varepsilon)^{-3d/2}\sum_n\varepsilon\|\int_{\mathbb{R}^{3d}}(\mathbf{a}_{n,1}^{\perp}+\mathbf{a}_{n,1})G_n\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}\boldsymbol{p}\,\mathrm{d}\boldsymbol{q}\|_{\mathrm{E}}. \tag{3.29}$$

Applying the derivatives with Proposition 2.2, we have the estimate

$$\|\boldsymbol{u}_{\mathrm{F},1}(t,\cdot)-\boldsymbol{u}_{\mathrm{F},0}(t,\cdot)\|_{\mathrm{E}}\le\varepsilon C\sum_n\|\mathbf{a}_{n,1}^{\perp}(t,\cdot)+\mathbf{a}_{n,1}(t,\cdot)\|_{\mathrm{L}^\infty}.$$

The estimate of (3.29) then follows directly from Proposition 2.2. We need to bound the prefactor terms, we note that on the compact set K_δ the bound for the prefactor terms comes from Lemma 5.4 in [13]. We go through several of the bounds here, starting with the P-wave and dropping the subscripts as the calculations are the same, and setting $\boldsymbol{P}=\boldsymbol{P}_p$, $\boldsymbol{Q}=\boldsymbol{Q}_p$,

$$\partial_t a_{\mathrm{p},0}=a_{\mathrm{p},0}\left(\frac{\partial_{\boldsymbol{Q}}c_{\mathrm{p}}\cdot\boldsymbol{P}}{|\boldsymbol{P}|}+\frac{1}{2}\mathrm{Tr}\left(Z^{-1}\partial_t Z\right)\right), \tag{3.30}$$

$$\partial_t a_{\mathrm{p},1} = a_{\mathrm{p},1}\left(\frac{\partial_{\boldsymbol{Q}} c_{\mathrm{p}}\cdot\boldsymbol{P}}{|\boldsymbol{P}|}+\frac{1}{2}\mathrm{Tr}\left(Z^{-1}\partial_t Z\right)\right)+F_{\mathrm{p}}(a_{\mathrm{p},0},\partial_{\boldsymbol{z}} a_{\mathrm{p},0},\boldsymbol{Q},\boldsymbol{P},c_{\mathrm{p}}). \tag{3.31}$$

With F_{p} being a continuously differentiable function in its arguments for $\boldsymbol{P},\boldsymbol{Q}\in K_{\delta_T}$. Equation (3.30) immediately implies:

$$\partial_t|a_{\mathrm{p},0}| \le |a_{\mathrm{p},0}|\left|\frac{\partial_{\boldsymbol{Q}} c_{\mathrm{p}}\cdot\boldsymbol{P}}{|\boldsymbol{P}|}+\frac{1}{2}\mathrm{Tr}\left(Z^{-1}\partial_t Z\right)\right|. \tag{3.32}$$

An application of Grönwall's inequality gives

$$\sup_{t\in[0,T]} \Lambda_{0,K_{\delta/2}}\left(a_{\mathrm{p},0}(t,\boldsymbol{q},\boldsymbol{p})\right)\le C_{\delta,T}. \tag{3.33}$$

To bound Equation (3.31), $\partial_{\boldsymbol{z}} a_{\mathrm{p},0}$ needs to be bounded, but with partial $\boldsymbol{z}$ of (3.30) using a similar inequality as Equation (3.32) and taking Grönwall's inequality we have

$$\sup_{t\in[0,T]} \Lambda_{1,K_{\delta/2}}\left(a_{\mathrm{p},0}(t,\boldsymbol{q},\boldsymbol{p})\right)\le C_{\delta,T}. \tag{3.34}$$

The function $F_{\mathrm{p}}(a_{\mathrm{p},0},\partial_{\boldsymbol{z}} a_{\mathrm{p},0},\boldsymbol{Q},\boldsymbol{P},c_{\mathrm{p}})$ is differentiable with differentiable arguments on the compact set $K_{\delta/2}$. Combining (3.33), (3.34) and Grönwall's inequality to Equation (3.31) we see that $\|a_{\mathrm{p}}\boldsymbol{P}\|_{\mathrm{L}^2}$ is bounded on $[0,T]\times K_{\delta/2}$.

For the S-wave terms, again using the short notation $\boldsymbol{P}=\boldsymbol{P}_s$, $\boldsymbol{Q}=\boldsymbol{Q}_s$ and dropping the subscript, we can write the system ((3.14), (3.15)) as

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix}\mathbf{a}_{\pm}^{sv}\\ \mathbf{a}_{\pm}^{sh}\end{pmatrix}=\frac{1}{2}\begin{pmatrix}h_{\pm} & m_{\pm}\\ -m_{\pm} & h_{\pm}\end{pmatrix}\begin{pmatrix}\mathbf{a}_{\pm}^{sv}\\ \mathbf{a}_{\pm}^{sh}\end{pmatrix}, \tag{3.35}$$

where $m_{\pm}=\partial_t\hat{\boldsymbol{N}}_{\mathrm{sh}}\cdot\hat{\boldsymbol{N}}_{\mathrm{sv}}$ and $h_{\pm}=2\partial_{\boldsymbol{Q}_{\mathrm{s},\pm}} c_s\cdot\hat{\boldsymbol{N}}_{\mathrm{sh}}+\mathbf{a}^s\mathrm{Tr}\left(Z_s^{-1}\partial_t Z^s\right)$. Denote M as the matrix in Equation (3.35), and $\mathbf{a}=(\mathbf{a}^{sv},\mathbf{a}^{sh})^T$. Then the system can be recast as $\frac{\mathrm{d}\mathbf{a}}{\mathrm{d}t}=M(t)\mathbf{a}$. Solving for the eigenvalues:

$$\lambda_{\mathrm{sh,sv}}(t)=-\partial_{\boldsymbol{Q}} c_{\mathrm{s}}\cdot\hat{\boldsymbol{N}}_{\mathrm{sh,sv}}-\frac{1}{2}\mathrm{Tr}\left(Z_{\mathrm{s}}^{-1}\frac{\mathrm{d}Z_{\mathrm{s}}}{\mathrm{d}t}\right)\mp\mathrm{i}\frac{\mathrm{d}\hat{\boldsymbol{N}}_{\mathrm{sv,sv}}}{\mathrm{d}t}\cdot\hat{\boldsymbol{N}}_{\mathrm{sh,sv}}. \tag{3.36}$$

To see that the latter are bounded, simply note that smooth $\{\hat{N}_{\mathrm{p}},\hat{N}_{\mathrm{sh}},\hat{N}_{\mathrm{sh}}\}$ form an orthonormal frame, and hence the last term in (3.36) is bounded for all $t\ge 0$.

Note that

$$\mathrm{Tr}\left(Z_{\mathrm{s}}^{-1}\frac{\mathrm{d}Z_{\mathrm{s}}}{\mathrm{d}t}\right)=\frac{1}{\det(Z_{\mathrm{s}})}\frac{\mathrm{d}\det(Z_{\mathrm{s}})}{\mathrm{d}t}, \tag{3.37}$$

then by (2.1) we have a bound for $\det(Z_{\mathrm{s}})$ so Equation (3.37) is bounded for all $t\ge 0$. Notice that

$$\frac{\partial_{\boldsymbol{q}} H_{\mathrm{s}}\cdot\partial_{\boldsymbol{p}} H_{\mathrm{s}}}{H_{\mathrm{s}}}=-\frac{\partial_{\boldsymbol{Q}_{\mathrm{s}}} c_{\mathrm{s}}\cdot\boldsymbol{P}_{\mathrm{s}}}{|\boldsymbol{P}_{\mathrm{s}}|}, \tag{3.38}$$

then with (3.4), Equation (3.38) is bounded for all $t\ge 0$. Now the eigenvalues in Equation (3.36) are bounded for all $t\ge 0$. So we have

$$\sup_{t\in[0,T]} \Lambda_{0,K_{\delta/2}}\left(\mathbf{a}_{\mathrm{s},0}(t,\boldsymbol{q},\boldsymbol{p})\right)\le C_{\delta,T}.$$

For $\partial_t \mathbf{a}_{s,1}$, we can write the system

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} a_{sv,1} \\ a_{sh,1}\end{pmatrix} = \frac{1}{2}\begin{pmatrix} h & m \\ -m & h\end{pmatrix}\begin{pmatrix} a_{\pm}^{sv} \\ a_{\pm}^{sh}\end{pmatrix} + \mathbf{F}_s(\mathbf{a}_{s,0}, \partial_{\boldsymbol{z}}\mathbf{a}_{s,0}, \boldsymbol{Q}, \boldsymbol{P}, c_s), \tag{3.39}$$

with $\mathbf{F}_s$ being a continuously differentiable function in its arguments for $\boldsymbol{P},\boldsymbol{Q} \in K_{\delta_T}$. The bounds follow in a similar fashion from previous work, and we arrive at

$$\sup_{t\in[0,T]} \Lambda_{1,K_{\delta/2}}(\mathbf{a}_{s,1}(t,\boldsymbol{q},\boldsymbol{p})) \le C_{\delta,T},$$

which gives the needed result. ∎

PROPOSITION 3.2. *For any $T>0$ and $t\in[0,T]$*

$$\|\boldsymbol{u}_{F,1}(t,\cdot) - \tilde{\boldsymbol{u}}_{F,1}(t,\cdot)\|_E = \mathcal{O}(\varepsilon^\infty).$$

Proof. We present the main argument, for full details we refer to Lemma 5.6 in [13] as the statements are similar. First observe that $\|\partial_x \boldsymbol{u}_{F,1}\|_{L^2} = \mathcal{O}(1/\varepsilon)$ (see the proof in Proposition 3.7), with the scaling in $\|\cdot\|_E$ from (3.28) we have $\mathcal{O}(\|\boldsymbol{u}_{F,1}(t,\cdot)\|_E) = \mathcal{O}(\|\boldsymbol{u}_{F,1}(t,\cdot)\|_{L^2})$. It follows that

$$\mathcal{O}(\|\boldsymbol{u}_{F,1}(t,\cdot) - \tilde{\boldsymbol{u}}_{F,1}(t,\cdot)\|_E) = \mathcal{O}(\|\boldsymbol{u}_{F,1}(t,\cdot) - \tilde{\boldsymbol{u}}_{F,1}(t,\cdot)\|_{L^2}).$$

Directly bounding (3.12) gives

$$\|\alpha_n^\varepsilon\|_{L^2(\mathbb{R}^{3d}\setminus K_\delta)} \le C_\delta(\|\boldsymbol{u}_0^\varepsilon\|_{L^2} + \epsilon\|\boldsymbol{u}_1^\varepsilon\|_{L^2}). \tag{3.40}$$

Now from the definition of $\boldsymbol{u}_{F,1}$ in (3.7) and $\tilde{\boldsymbol{u}}_{F,1}$ in (3.10), we get

$$\begin{aligned}\|\boldsymbol{u}_{F,1} - \tilde{\boldsymbol{u}}_{F,1}\|_{L^2} &\le 2^{-d/2}\sum_n \|(1-\chi_\delta)(a_{n,0}+\varepsilon a_{n,1})\hat{\boldsymbol{N}}_n \mathcal{F}^\varepsilon \alpha_n^\varepsilon\|_{L^2} \\ &\le C_{\delta,T}\left(\|\mathcal{F}^\varepsilon \boldsymbol{u}_0^\varepsilon\|_{L^2(\mathbb{R}^{2d}\setminus K_\delta)} + \varepsilon\|\mathcal{F}^\varepsilon \boldsymbol{u}_1^\varepsilon\|_{L^2(\mathbb{R}^{2d}\setminus K_\delta)}\right) \\ &\quad \times \sum_n \|(1-\chi_\delta)(a_{n,0}+\varepsilon a_{n,1})\hat{\boldsymbol{N}}_n\|_{L^\infty} \\ &\le C_{\delta,T}\left(\|\mathcal{F}^\varepsilon \boldsymbol{u}_0^\varepsilon\|_{L^2(\mathbb{R}^{2d}\setminus K_\delta)} + \varepsilon\|\mathcal{F}^\varepsilon \boldsymbol{u}_1^\varepsilon\|_{L^2(\mathbb{R}^{2d}\setminus K_\delta)}\right) = \mathcal{O}(\varepsilon^\infty).\end{aligned}$$

The first equality inequality is obtained by similar arguments for Proposition 2.2 while the second obtained by similar arguments found in Proposition 3.1. ∎

PROPOSITION 3.3. *The operators $\mathcal{L}_0, \mathcal{L}_1, \mathcal{L}_2$ are bounded. That is, for a given T and any $t\in[0,T]$, $\mathbf{a}\in C^\infty([0,T])\times S(\mathbb{R}^{2d})$ and for $k=0,1$, $j=0,1,2$,*

$$\sup_{t\in[0,T]} \Lambda_{k,K_\delta}(\mathcal{L}_j(\mathbf{a})) < C_{T,K_\delta} \text{ and } \|\mathcal{L}_j(\mathbf{a}(t,\cdot))\|_{L^\infty} < C_{T,\delta}.$$

Proof. Notice $\mathcal{L}_{n,j}$ depend on $\boldsymbol{P}_n$, $\boldsymbol{Q}_n$ Z_n^{-1} and its derivatives, which are all bounded on $[0,T]\times K_\delta$. This gives the result. ∎

PROPOSITION 3.4. *For a given T and any $t\in[0,T]$, for $k=0,1$ and $\mathbf{a}\in C^\infty([0,T])\times S(\mathbb{R}^{2d})$, we have*

$$\Lambda_{k,K_\delta}(\mathbf{a}_{n,1}^\perp(\mathbf{a}(t,\cdot))) < C_{T,K_\delta} \text{ and } \|\mathbf{a}_{n,1}^\perp(\mathbf{a}(t,\cdot))\|_{L^\infty} < C_{T,\delta}. \tag{3.41}$$

Proof. For both $\mathbf{a}_{\mathrm{p}}^{\perp}$ and $\mathbf{a}_{\mathrm{s}}^{\perp}$, the pseudo-operators $\mathcal{L}_{\mathrm{p},0}^{-1}$, and $\mathcal{L}_{\mathrm{s},0}^{-1}$ are bounded on K_δ as $|P|>0$ and $|\mathcal{L}_{n,0}^{-1}|_{\mathrm{L}^\infty}\le C\delta^{-1}$. Then by Proposition 3.3, we have the following result. ∎

PROPOSITION 3.5. *Consider the elastic wave equations with a forcing term,* $\boldsymbol{x}\in\mathbb{R}^d$,

$$\begin{cases}\rho(x)\partial_t^2\boldsymbol{u}-(\lambda+2\mu)\nabla(\nabla\cdot\boldsymbol{u})+\mu\nabla\times\nabla\times\boldsymbol{u}=\mathbf{F}(t,\boldsymbol{x}),\\ u(0,\boldsymbol{x})=\boldsymbol{u}_0^\varepsilon,\\ u(t,\boldsymbol{x})=\boldsymbol{u}_1^\varepsilon.\end{cases}\tag{3.42}$$

Let $T>0$*, and let* $\boldsymbol{u}_0(t,\boldsymbol{x})\in C^\infty([0,T])\times H_0^1(\mathbb{R}^d)$*. For each* $t\in[0,T]$*, we have the following estimate:*

$$\|\rho\partial_t\boldsymbol{u}\|_{\mathrm{L}^2}+\|(\lambda+2\mu)\nabla\cdot\boldsymbol{u}\|_{\mathrm{L}^2}+\|\mu\nabla\times\boldsymbol{u}\|_{\mathrm{L}^2}\le C_T\left(\frac{1}{\varepsilon}\|\boldsymbol{u}(0,\cdot)\|_{\mathrm{E}}+\int_0^t\|\mathbf{F}(s,\cdot)\|_{\mathrm{L}^2}\,\mathrm{d}s\right).$$

In particular,

$$\sup_{t\in[0,T]}\|\boldsymbol{u}(t,\cdot)\|_{\mathrm{E}}\le C_T\left(\|\boldsymbol{u}_0^\varepsilon\|_{\mathrm{E}}+\varepsilon\int_0^T\|\mathbf{F}(s,\cdot)\|_{\mathrm{L}^2}\,\mathrm{d}s\right).\tag{3.43}$$

Proof. This is a standard estimate. Dotting Equation (3.43) with $\partial_t\boldsymbol{u}$ and integrating over space, we have

$$\frac{\partial_t}{2}\int_{\mathbb{R}^d}\rho(x)|\partial_t\boldsymbol{u}|^2+(\lambda+2\mu)|\nabla\cdot\boldsymbol{u}|^2+\mu|\nabla\times\boldsymbol{u}|^2\,\mathrm{d}\boldsymbol{x}\le\int_{\mathbb{R}^d}|\partial_t\boldsymbol{u}\cdot\mathbf{F}|\,\mathrm{d}\boldsymbol{x}.$$

The right-hand side can then be estimated by

$$\int_{\mathbb{R}^d}|\partial_t\boldsymbol{u}\cdot\mathbf{F}|\,\mathrm{d}\boldsymbol{x}\le\frac{1}{2}\int_{\mathbb{R}^d}|\partial_t\boldsymbol{u}|^2+|\mathbf{F}|^2\,\mathrm{d}\boldsymbol{x}.$$

Adding the missing terms to apply Grönwall's inequality gives the bound

$$e^t(\|\rho(x)\boldsymbol{u}_1\|_{\mathrm{L}^2}^2+\|(\lambda+\mu)\nabla\cdot\boldsymbol{u}_0\|_{\mathrm{L}^2}^2+\|\mu\nabla\times\boldsymbol{u}\|_{\mathrm{L}^2}^2)+\int_0^t e^{t-s}\int_{\mathbb{R}^d}|\mathbf{F}(s,\boldsymbol{x})|^2\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}s.$$

Taking the maximum over ρ,λ,μ and over T, we arrive at the estimate. ∎

PROPOSITION 3.6. *We have*

$$\|(\partial_t^2-\mathcal{L})\tilde{\boldsymbol{u}}_{\mathrm{F},1}\|_{\mathrm{E}}\le\varepsilon C_{T,\delta}.$$

Proof. Plugging $\tilde{\boldsymbol{u}}_{\mathrm{F},1}$ into (1.1) gives,

$$\begin{aligned}&(\partial_t^2-\mathcal{L})\tilde{\boldsymbol{u}}_{\mathrm{F},1}(t,\boldsymbol{x})\\&=(2\pi\varepsilon)^{-3d/2}\int_{\mathbb{R}^{3d}}\chi_\delta(\boldsymbol{q},\boldsymbol{p})\sum_{n,m}\varepsilon^{m-2}\mathcal{L}_{n,m}\Big(\mathbf{a}_{n,0}+\varepsilon(\mathbf{a}_{n,1}+\varepsilon\mathbf{a}_{n,1}^{\perp})\Big)G_n^\varepsilon\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}\boldsymbol{p}\,\mathrm{d}\boldsymbol{q}.\end{aligned}$$

Expanding and simplifying the above equation yields

$$\begin{aligned}(\partial_t^2-\mathcal{L})\tilde{\boldsymbol{u}}_{\mathrm{F},1}(t,\boldsymbol{x})=(2\pi\varepsilon)^{-3/2}\sum_n\int_{\mathbb{R}^{3d}}\chi_\delta(\boldsymbol{q},\boldsymbol{p})G_n^\varepsilon\Big[\varepsilon^{-2}\mathcal{L}_{n,0}(\mathbf{a}_{n,0})\\+\varepsilon^{-1}\mathcal{L}_{n,0}(\mathbf{a}_{n,1})+\varepsilon^{-1}\mathcal{L}_{n,1}(\mathbf{a}_{n,0})+\mathcal{L}_{n,1}(\mathbf{a}_{n,1})+\mathcal{L}_{n,2}(\mathbf{a}_{n,0}+\varepsilon\mathbf{a}_{n,1})\\-(\varepsilon^{-1}\mathcal{L}_{n,0}+\mathcal{L}_{n,1}+\varepsilon\mathcal{L}_{n,2})\mathcal{L}_{n,0}^{-1}((\mathrm{Id}-\Pi_n)(\mathcal{L}_{n,1}(\mathbf{a}_{n,0}\hat{\boldsymbol{N}}_n)))\Big]\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}\boldsymbol{p}\,\mathrm{d}\boldsymbol{q}.\end{aligned}$$

Direct cancellation gives

$$(\partial_t^2-\mathcal{L})\tilde{\boldsymbol{u}}_{\mathrm{F},1}(t,\boldsymbol{x})=(2\pi\varepsilon)^{-3/2}\sum_n\int_{\mathbb{R}^{3d}}G_n^\varepsilon\Big[R_0(t,\boldsymbol{p},\boldsymbol{q})+\varepsilon R_1(t,\boldsymbol{p},\boldsymbol{q})\Big]\mathrm{d}\boldsymbol{y}\,\mathrm{d}\boldsymbol{p}\,\mathrm{d}\boldsymbol{q},$$

where

$$R_0(t,\boldsymbol{p},\boldsymbol{q})=\mathcal{L}_{n,1}(\mathbf{a}_{n,1})+\mathcal{L}_{n,2}(\mathbf{a}_{n,0})-\mathcal{L}_{n,1}\mathbf{a}_{n,1}^\perp,\qquad R_1(t,\boldsymbol{p},\boldsymbol{q})=\mathcal{L}_{n,2}(\mathbf{a}_{n,1})-\mathcal{L}_{n,2}\mathbf{a}_{n,1}^\perp.$$

Then by Propositions 2.2 and 3.4,

$$\|(\partial_t^2-\mathcal{L})\tilde{\boldsymbol{u}}_{\mathrm{F},1}(t,\cdot)\|_{\mathrm{L}^2}\le C_{T,\delta}(\|R_0(t,\cdot)\|_{\mathrm{L}^\infty}+\varepsilon\|R_1(t,\cdot)\|_{\mathrm{L}^\infty}).$$

By Propositions 3.3 and 3.4, R_0 and R_1 are hence bounded. ∎

PROPOSITION 3.7. *Let $\boldsymbol{u}$ solve the Cauchy problem (1.1). If $\boldsymbol{u}_{\mathrm{F},0}$ is the first-order FGA (3.5), then we have the following estimate on the initial conditions:*

$$\|\boldsymbol{u}(0,\boldsymbol{x})-\boldsymbol{u}_{\mathrm{F}}(0,\boldsymbol{x})\|_{\mathrm{E}}\le\varepsilon C_T.$$

Proof. First computing the following

$$\partial_t\mathbf{a}_n(0,\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})=\alpha(\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})\left(\frac{\partial c(\boldsymbol{q})\cdot\boldsymbol{p}}{|\boldsymbol{p}|}-d\right)\hat{\boldsymbol{n}}_n.$$

For estimating $\boldsymbol{u}(0,\boldsymbol{x})-\boldsymbol{u}_{\mathrm{F},0}(0,\boldsymbol{x})$ in the energy norm, we can write $\partial_t\boldsymbol{u}(0,x)=\boldsymbol{u}_1^\varepsilon(\boldsymbol{x})$. This gives

$$\begin{aligned}&\int_{\mathbb{R}^{3d}}\boldsymbol{u}_0^\varepsilon(\boldsymbol{y})\frac{\mathrm{i}}{\varepsilon}\Phi_n(0,\boldsymbol{x},\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})G_n^\varepsilon(0,\boldsymbol{x},\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}\boldsymbol{p}\,\mathrm{d}\boldsymbol{q}\\&=\int_{\mathbb{R}^{3d}}\boldsymbol{u}_1^\varepsilon(\boldsymbol{y})G_n^\varepsilon(0,\boldsymbol{x},\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}\boldsymbol{p}\,\mathrm{d}\boldsymbol{q}.\end{aligned}\tag{3.44}$$

Then

$$\begin{aligned}|\boldsymbol{u}_1^\varepsilon(\boldsymbol{x})-\partial_t\boldsymbol{u}_{\mathrm{F},0}(0,\boldsymbol{x})|=(2\pi\varepsilon)^{-d/2}\Big|\sum_n\int_{\mathbb{R}^{3d}}\Big[\frac{1}{2}(\boldsymbol{u}_1^\varepsilon(y)\cdot\hat{\boldsymbol{n}})\hat{\boldsymbol{n}}\\-\partial_t\mathbf{a}_n(0,\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})-\mathbf{a}_n\frac{\mathrm{i}}{\varepsilon}\Phi_n(0,\boldsymbol{x},\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})\Big]G_n^\varepsilon(0,\boldsymbol{x},\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}\boldsymbol{p}\,\mathrm{d}\boldsymbol{q}\Big|.\end{aligned}$$

For one of the right terms, it becomes, after plugging in (3.44),

$$\begin{aligned}\int_{\mathbb{R}^{3d}}\Big[\boldsymbol{u}_0^\varepsilon(\boldsymbol{y})\frac{\mathrm{i}}{2\varepsilon}\Phi_n(0,\boldsymbol{x},\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})-\alpha(\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})\left(\frac{\partial c(\boldsymbol{q})\cdot\boldsymbol{p}}{|\boldsymbol{p}|}-d\right)\hat{\boldsymbol{n}}_n\\-2^{d/2}\alpha(\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})\hat{\boldsymbol{n}}_n\frac{\mathrm{i}}{\varepsilon}\Phi_n(0,\boldsymbol{x},\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})\Big]G_n^\varepsilon(0,\boldsymbol{x},\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}\boldsymbol{p}\,\mathrm{d}\boldsymbol{q}.\end{aligned}$$

Using (3.12) and summing over the wavefields and branches gives

$$\begin{aligned}\boldsymbol{u}_1^\varepsilon(\boldsymbol{x})-\partial_t\boldsymbol{u}_{\mathrm{F},0}(0,\boldsymbol{x})=-\sum_n\int_{\mathbb{R}^{3d}}\frac{1}{2c_n|\boldsymbol{p}|^3}\Big(\boldsymbol{u}_0^\varepsilon(\boldsymbol{y})c_n|\boldsymbol{p}|\pm\mathrm{i}\varepsilon\boldsymbol{u}_1^\varepsilon(\boldsymbol{y})\Big)\cdot\hat{\boldsymbol{n}}_n\\\times\left(\frac{\partial c(\boldsymbol{q})\cdot\boldsymbol{p}}{|\boldsymbol{p}|}-d\right)\hat{\boldsymbol{n}}_nG_n^\varepsilon(0,\boldsymbol{x},\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}\boldsymbol{p}\,\mathrm{d}\boldsymbol{q}.\end{aligned}$$

By Proposition 2.2, we then have the estimate

$$\|\boldsymbol{u}_1^\varepsilon - \partial_t \boldsymbol{u}_{\mathrm{F},0}(0,\cdot)\|_{\mathrm{L}^2} \le C_T.$$

For the divergence term,

$$\begin{aligned}\nabla\cdot \boldsymbol{u}_{\mathrm{F}}(0,\boldsymbol{x}) &= (2\pi\varepsilon)^{-d/2}\sum_n \int_{\mathbb{R}^{3d}} \nabla\cdot\big(\mathbf{a}_n(0,\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})G_n^\varepsilon(0,\boldsymbol{x},\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})\big)\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}\boldsymbol{p}\,\mathrm{d}\boldsymbol{q}\\ &= (2\pi\varepsilon)^{-d/2}\sum_n \int_{\mathbb{R}^{3d}} \frac{\mathrm{i}}{\varepsilon}\mathbf{a}_n\cdot(\boldsymbol{P}_n + (\boldsymbol{x}-\boldsymbol{Q}_n))\,G_n^\varepsilon(0,\boldsymbol{x},\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}\boldsymbol{p}\,\mathrm{d}\boldsymbol{q}.\end{aligned}$$

Applying the operators and integration by parts gives, for one term,

$$\int_{\mathbb{R}^{3d}} \left(\frac{\mathrm{i}}{\varepsilon}\mathbf{a}_n\cdot\boldsymbol{p}_n - \partial_{\boldsymbol{z}}(Z^{-1}\mathbf{a})\right) G_n^\varepsilon(0,\boldsymbol{x},\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}\boldsymbol{p}\,\mathrm{d}\boldsymbol{q}.$$

Writing the difference $\nabla\cdot\boldsymbol{u}_0^\varepsilon - \nabla\cdot\boldsymbol{u}_{\mathrm{F},0}(t,\boldsymbol{x})$ in terms of the FIO; see Equation (2.6), leads to

$$(2\pi\varepsilon)^{-d/2}\int_{\mathbb{R}^{3d}} \left(\frac{\mathrm{i}}{\varepsilon}\big(\boldsymbol{u}_0^\varepsilon(\boldsymbol{y}) - \mathbf{a}(0,\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})\big)\cdot\boldsymbol{P}_n + \partial_{\boldsymbol{z}}(Z^{-1}(\boldsymbol{u}_0^\varepsilon(\boldsymbol{y}) - \mathbf{a}(0,\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})))\right) \times G_n^\varepsilon(0,\boldsymbol{x},\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}\boldsymbol{p}\,\mathrm{d}\boldsymbol{q}.$$

With $\mathbf{a}(0,\boldsymbol{y},\boldsymbol{q},\boldsymbol{p}) = \alpha_n(\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})\hat{\boldsymbol{n}}$ and summing over n we have

$$\begin{aligned}&\nabla\cdot\boldsymbol{u}_0^\varepsilon(\boldsymbol{x}) - \nabla\cdot\boldsymbol{u}_{\mathrm{F},0}(0,\boldsymbol{x})\\ &= (2\pi\varepsilon)^{-d/2}\sum_n\int_{\mathbb{R}^{3d}} \partial_{\boldsymbol{z}}(Z^{-1}(\boldsymbol{u}_0^\varepsilon(\boldsymbol{y}) - \alpha_n(\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})\hat{\boldsymbol{n}})G_n^\varepsilon(0,\boldsymbol{x},\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}\boldsymbol{p}\,\mathrm{d}\boldsymbol{q}.\end{aligned}$$

Again, by Proposition 2.2 we arrive at the estimate

$$\|\nabla\cdot\boldsymbol{u}_0^\varepsilon - \nabla\cdot\boldsymbol{u}_{\mathrm{F},0}(t,\cdot)\|_{\mathrm{L}^2} \le C_T.$$

The curl term has a similar estimate as the divergence term. These three estimates show the result. ∎

THEOREM 3.1. *Let $\{\boldsymbol{u}_0^\varepsilon\}$ be a family of asymptotically high frequency initial conditions, and let $\boldsymbol{u}$ be solution to the Cauchy problem (1.1). If $\boldsymbol{u}_{\mathrm{F},0}$ is the first-order FGA (3.5), then for a given T and any $t\in[0,T]$, $\delta>0$ and sufficiently small ε, we have*

$$\sup_{t\in[0,T]}\|\boldsymbol{u}(t,\cdot) - \boldsymbol{u}_{\mathrm{F},0}(t,\cdot)\|_{\mathrm{E}} \le \varepsilon C_{T,\delta}.$$

Proof. By the triangle inequality,

$$\|\boldsymbol{u} - \boldsymbol{u}_{\mathrm{F},0}\|_{\mathrm{E}} \le \|\boldsymbol{u} - \tilde{\boldsymbol{u}}_{\mathrm{F},1}\|_{\mathrm{E}} + \|\tilde{\boldsymbol{u}}_{\mathrm{F},1} - \boldsymbol{u}_{\mathrm{F},1}\|_{\mathrm{E}} + \|\boldsymbol{u}_{\mathrm{F},1} - \boldsymbol{u}_{\mathrm{F},0}\|_{\mathrm{E}}. \tag{3.45}$$

For the first term, we define the quantity $e = \boldsymbol{u} - \tilde{\boldsymbol{u}}_{\mathrm{F},1}$ and by Propositions 3.5 and 3.6

$$\|e\|_{\mathrm{E}} \le C_{T,\delta}\big(\|e(0,\cdot)\|_{\mathrm{E}} + \varepsilon\int_0^t \|R_0\|_{\mathrm{L}^2}\,\mathrm{d}s\big) + \mathcal{O}(\varepsilon^2).$$

Proposition 3.7 then shows that $\|e(0,\cdot)\|_{\mathrm{E}} \le \varepsilon C_{T,\delta}$, and Propositions 3.3 and 3.4 show that $\|R_0\|_{\mathrm{L}^2} \le C_{T,\delta}$. Thus, in Equation (3.45), the first term is estimated at the correct order, the second term is a $\mathcal{O}(\varepsilon^\infty)$ thanks to Proposition 3.2, and the last term is estimated to the desired order by Proposition 3.1. ∎

REMARK 3.3. The requirement of high frequency initial conditions is necessary for the convergence estimate in Theorem 3.1. Otherwise, one may lose accuracy as seen in a similar example in [13, Example 4.4] for P-wave.

4. Convergence analysis for the Dirac system

Since the derivation of the boundedness estimates related to the Hamiltonian flow associated to the FGA formulation of EWS is essentially the same as the FGA formulation of the Dirac system (DS), we shall omit this derivation in this section, and only present the asymptotic corrections and the convergence results for the Dirac system.

4.1. Next order corrections of FGA for the Dirac system.

According to [2], the FGA is of the form

$$\boldsymbol{u}_{\mathrm{F},0}(t,\boldsymbol{x}) = (2\pi\varepsilon)^{-3d/2}\int_{\mathbb{R}^{3d}}\sum_{\pm}\mathbf{a}_{\pm,0}(t,\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})G^{\varepsilon}_{\pm}(t,\boldsymbol{x},\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}\boldsymbol{p}\,\mathrm{d}\boldsymbol{q}, \tag{4.1}$$

where $\mathbf{a}_{\pm,0} = \mathbf{a}_{\pm1,j} + \mathbf{a}_{\pm2,j}$, $\pm$ indicates the positive/negative eigenvalue. For $k>1$, define the k-th order FGA with a correction term as

$$\boldsymbol{u}_{\mathrm{F},k}(t,\boldsymbol{x}) = (2\pi\varepsilon)^{-3d/2}\int_{\mathbb{R}^{3d}}\sum_{j=0}^{k}\sum_{\pm}\varepsilon^j\bar{\mathbf{a}}_{\pm,j}(t,\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})G^{\varepsilon}_{\pm}(t,\boldsymbol{x},\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}\boldsymbol{p}\,\mathrm{d}\boldsymbol{q}, \tag{4.2}$$

where $\bar{\mathbf{a}}_{\pm,j} = \bar{\mathbf{a}}_{\pm1,j} + \bar{\mathbf{a}}_{\pm2,j}$, $\bar{\mathbf{a}}_{\pm m,0} = \mathbf{a}_{\pm m,0}$, and the terms $\bar{\mathbf{a}}_{\pm m,j} = \mathbf{a}_{\pm m,j} + \mathbf{a}^{\perp}_{\pm m,j}$ will be defined later for $j\geq 1$. Let $\boldsymbol{\Upsilon}_{\pm1}$ and $\boldsymbol{\Upsilon}_{\pm2}$ be the normalized eigenvectors corresponding to the eigenvalue $h_{\pm}$ of the Dirac symbol defined in (1.12). Then $\mathbf{a}_{\pm m,j}(t,\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})$ is defined as follows,

$$\mathbf{a}_{\pm m,j}(t,\boldsymbol{y},\boldsymbol{q},\boldsymbol{p}) = a_{\pm m,j}(t,\boldsymbol{q},\boldsymbol{p})\alpha^{\varepsilon}_{\pm m}(\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})\boldsymbol{\Upsilon}_{\pm m}(t,\boldsymbol{q},\boldsymbol{p}), \tag{4.3}$$

where α^{ε} incorporates the initial conditions,

$$\alpha^{\varepsilon}_{\pm m}(\boldsymbol{y},\boldsymbol{q},\boldsymbol{p}) = \boldsymbol{\varphi}^{\varepsilon}_{I}(\boldsymbol{y})\cdot\boldsymbol{\Upsilon}_{\pm m}(0,\boldsymbol{q},\boldsymbol{p}). \tag{4.4}$$

The computations for the $\pm$ branches will be similar, so we use subscript m instead of $\pm m$ when there is no misunderstanding. The scalar functions $a_{m,0}$, with $m=1,2$, satisfy the following evolution equations [2]

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix}a_{1,0}\\ a_{2,0}\end{pmatrix} + \Xi\begin{pmatrix}a_{1,0}\\ a_{2,0}\end{pmatrix} = \mathbf{0}, \tag{4.5}$$

where Ξ is 2 by 2 matrix with elements

$$\Xi_{mn} = \delta_{mn}\boldsymbol{\Upsilon}^{\dagger}_{m}\frac{\mathrm{d}\boldsymbol{\Upsilon}_n}{\mathrm{d}t} - \delta_{mn}\frac{\mathrm{i}}{2}\partial_{z_k}Q_l\,\partial_{Q_l}\partial_{Q_j}V\,Z^{-1}_{jk} + \partial_{z_k}\boldsymbol{\Upsilon}^{\dagger}_{m}\mathbf{F}^{n}_{j}Z^{-1}_{jk},$$

and

$$\mathbf{F}^{n}_{j} = \left(\partial_{P_j}h(\boldsymbol{Q},\boldsymbol{P}) - c\hat{\boldsymbol{\sigma}}_j - \mathrm{i}\partial_{Q_j}h(\boldsymbol{Q},\boldsymbol{P}) + \mathrm{i}\partial_{Q_j}B\right)\boldsymbol{\Upsilon}_n.$$

We define the operators $\mathcal{L}_{\pm,0},\mathcal{L}_{\pm,1},\mathcal{L}_{\pm,2}$ acting on $\mathbf{a}$ as

$$\mathcal{L}_{\pm,0}\mathbf{a} = \mathrm{i}\left(c\hat{\boldsymbol{\sigma}}\cdot\boldsymbol{P}_{\pm} + B(\boldsymbol{Q}_{\pm}) + \partial_t S_{\pm} - \boldsymbol{P}_{\pm}\cdot\partial_t\boldsymbol{Q}_{\pm}\right)\mathbf{a}, \tag{4.6}$$

$$\mathcal{L}_{\pm,1}\mathbf{a} = \partial_t\mathbf{a} - \partial_{z_k}\left[\left(\partial_t(Q_{\pm,j} + \mathrm{i}P_{\pm,j}) - c\hat{\boldsymbol{\sigma}}_j + \mathrm{i}\partial_{Q_j}B\right)Z^{-1}_{jk}\mathbf{a}\right] + \frac{\mathrm{i}}{2}\partial_{z_k}Q_{\pm,l}\,\partial_{Q_l}\partial_{Q_j}B\,Z^{-1}_{\pm,jk}\mathbf{a}, \tag{4.7}$$

$$\mathcal{L}_{\pm,2}\mathbf{a} = \frac{1}{3\mathrm{i}}\partial_{z_n}\left(\partial_{z_l}Q_j\partial_{Q_i}\partial_{Q_j}\partial_{Q_k}BZ^{-1}_{il}Z^{-1}_{kn}\mathbf{a}\right) + \frac{1}{6\mathrm{i}}\partial_{z_n}Q_j\partial_{z_l}\left(\partial_{Q_i}\partial_{Q_j}\partial_{Q_k}BZ^{-1}_{il}\mathbf{a}\right)Z^{-1}_{kn}$$

$$+\frac{\varepsilon}{6\mathrm{i}}\partial_{z_m}\left(\partial_{z_n}\left(\partial_{z_l}\left(\partial_{Q_i}\partial_{Q_j}\partial_{Q_k}BZ_{il}^{-1}\mathbf{a}\right)Z_{kn}^{-1}\right)Z_{jm}^{-1}\right). \tag{4.8}$$

Looking at the $O(\varepsilon^0)$ terms, and equating to zero gives

$$\mathcal{L}_{\pm,0}\mathbf{a}_{\pm,0}=\mathbf{0}, \tag{4.9}$$

thus

$$\partial_t S_\pm - \boldsymbol{P}_\pm\cdot\partial_t\boldsymbol{Q}_\pm = -h_\pm,$$

then, together with the Hamiltonian flow, one recovers the evolution equation for the action:

$$\frac{\mathrm{d}}{\mathrm{d}t}S=\boldsymbol{P}\cdot\partial_{\boldsymbol{P}}h(\boldsymbol{Q},\boldsymbol{P})-h(\boldsymbol{Q},\boldsymbol{P}). \tag{4.10}$$

Looking at the $O(\varepsilon^1)$ terms and equating to zero gives

$$\mathcal{L}_{\pm,1}\mathbf{a}_{\pm,0}+\mathcal{L}_{\pm,0}\overline{\mathbf{a}}_{\pm,1}=\mathbf{0}. \tag{4.11}$$

Let us take the + branch for example and inner product of the above equation with $\boldsymbol{\Upsilon}_{\pm m}$,

$$\boldsymbol{\Upsilon}_{\pm m}^\dagger\mathcal{L}_{\pm,1}\mathbf{a}_{\pm,0}=-\boldsymbol{\Upsilon}_{\pm m}^\dagger\mathcal{L}_{\pm,0}\overline{\mathbf{a}}_{\pm,1}=\left(\mathcal{L}_{\pm,0}\boldsymbol{\Upsilon}_{\pm m}\right)^\dagger\overline{\mathbf{a}}_{\pm,1}=0, \tag{4.12}$$

from which one recovers Equation (4.5). Next, define

$$\mathcal{L}_{\pm,0}^{-1}=\frac{\mathrm{i}}{h_\pm-h_\mp}\left(\boldsymbol{\Upsilon}_{\mp1}\boldsymbol{\Upsilon}_{\mp1}^\dagger+\boldsymbol{\Upsilon}_{\mp2}\boldsymbol{\Upsilon}_{\mp2}^\dagger\right), \tag{4.13}$$

which is a pseudo-inverse of $\mathcal{L}_{\pm,0}$. We then define

$$\mathbf{a}_{\pm,1}^\perp=-\mathcal{L}_{\pm,0}^{-1}\mathcal{L}_{\pm,1}\mathbf{a}_{\pm,0}\,. \tag{4.14}$$

Looking at the $O(\varepsilon^2)$ terms and equating to zero give

$$\mathcal{L}_{\pm,1}\left(\mathbf{a}_{\pm,1}+\mathbf{a}_{\pm,1}^\perp\right)+\mathcal{L}_{\pm,2}\mathbf{a}_{\pm,0}+\mathcal{L}_{\pm,0}\overline{\mathbf{a}}_{\pm,2}=\mathbf{0}, \tag{4.15}$$

which implies

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix}a_{1,1}\\a_{2,1}\end{pmatrix}+\Xi\begin{pmatrix}a_{1,0}\\a_{2,0}\end{pmatrix}+\begin{pmatrix}\boldsymbol{\Upsilon}_{\pm1}^\dagger\\\boldsymbol{\Upsilon}_{\pm2}^\dagger\end{pmatrix}\left(\mathcal{L}_{\pm,1}\mathbf{a}_{\pm,1}^\perp+\mathcal{L}_{\pm,2}\mathbf{a}_{\pm,0}\right)=\mathbf{0}. \tag{4.16}$$

REMARK 4.1. Each energy band has two eigenvectors, so we that obtain a coupled system (4.5) which presents the intralband diabatic transitions. The transitions only take place in the same energy band and there is no interband transition between the positive and negative energy bands since $\mathbf{a}_+$ and $\mathbf{a}_-$ are fully decoupled.

4.2. Error estimate and main result for the Dirac system.

LEMMA 4.1. *The pseudo-inverse operator defined in (4.13) is bounded, that is, for each $k\in\mathbb{N}$, there exists a constant C_k*

$$\Lambda_k\left(\mathcal{L}_{\pm m,0}^{-1}(\boldsymbol{q},\boldsymbol{p})\right)\le C_k. \tag{4.17}$$

Proof. Recall

$$\mathcal{L}_{\pm,0}^{-1}(\boldsymbol{q},\boldsymbol{p})=\frac{\mathrm{i}}{h_\pm(\boldsymbol{q},\boldsymbol{p})-h_\mp(\boldsymbol{q},\boldsymbol{p})}\Big(\boldsymbol{\varUpsilon}_{\mp1}(\boldsymbol{q},\boldsymbol{p})\boldsymbol{\varUpsilon}_{\mp1}^\dagger(\boldsymbol{q},\boldsymbol{p})+\boldsymbol{\varUpsilon}_{\mp2}(\boldsymbol{q},\boldsymbol{p})\boldsymbol{\varUpsilon}_{\mp2}^\dagger(\boldsymbol{q},\boldsymbol{p})\Big).$$

Since $h_+-h_-\ge 2c^2$, and $h_\pm$ and $\boldsymbol{\varUpsilon}_{\pm1,\pm2}$ are smooth functions of $(\boldsymbol{q},\boldsymbol{p})$, the estimate follows easily. ∎

LEMMA 4.2. *For any $T>0$, $k\in\mathbb{N}$, $j=0,1$, there exists a constant $C_{k,j,T}$*

$$\sup_{0\le t\le T}\Lambda_k(\overline{\mathbf{a}}_{\pm,j})\le C_{k,j,T}, \tag{4.18}$$

$$\sup_{0\le t\le T}\Lambda_k(\partial_t\overline{\mathbf{a}}_{\pm,j})\le C_{k,j,T}. \tag{4.19}$$

Proof. Noticing that Ξ, $\boldsymbol{\varUpsilon}$'s, and $\mathcal{L}$'s are smooth and bounded, the estimates follow from Equations (4.5), (4.16), and Grönwall's lemma. ∎

PROPOSITION 4.1. *Let $\boldsymbol{u}_{\mathrm{F},0}$ and $\boldsymbol{u}_{\mathrm{F},1}$ be the zeroth- and first-order FGA solution in Equations (4.1) and (4.2), then for any $T>0$, there exists a constant C_T, such that*

$$\sup_{0\le t\le T}\|\boldsymbol{u}_{\mathrm{F},1}-\boldsymbol{u}_{\mathrm{F},0}\|_{L^2}\le\varepsilon C_T. \tag{4.20}$$

Proof. From the definitions we have

$$\|\boldsymbol{u}_{\mathrm{F},1}-\boldsymbol{u}_{\mathrm{F},0}\|_{L^2}\le(2\pi\varepsilon)^{-3d/2}\sum_\pm\varepsilon\left\|\int_{\mathbb{R}^{3d}}\overline{\mathbf{a}}_{\pm,1}(t,\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})G_\pm^\varepsilon(t,\boldsymbol{x},\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}\boldsymbol{p}\,\mathrm{d}\boldsymbol{q}\right\|_{L^2}, \tag{4.21}$$

then by Proposition 2.2,

$$\|\boldsymbol{u}_{\mathrm{F},1}-\boldsymbol{u}_{\mathrm{F},0}\|_{L^2}\le\varepsilon C\sum_\pm\|\overline{\mathbf{a}}_{\pm,1}(t,\boldsymbol{y},\boldsymbol{q},\boldsymbol{p})\|_{L^\infty}. \tag{4.22}$$

Therefore by Lemma 4.2, we arrive at the estimate (4.20). ∎

LEMMA 4.3. *For any $T>0$, there exists a constant C_T, such that for any $\varepsilon>0$,*

$$\|(\mathrm{i}\varepsilon\partial_t-\mathcal{D})\boldsymbol{u}_{\mathrm{F},1}\|_{L^2}\le\varepsilon^2C_T. \tag{4.23}$$

Proof. Plugging $\boldsymbol{u}_{\mathrm{F},1}$ into DS gives

$$\begin{aligned}&(\mathrm{i}\varepsilon\partial_t-\mathcal{D})\boldsymbol{u}_{\mathrm{F},1}\\ =&(2\pi\varepsilon)^{-3d/2}\int_{\mathbb{R}^{3d}}\sum_\pm\big(\mathcal{L}_{\pm,0}\mathbf{a}_{\pm,0}+\varepsilon\mathcal{L}_{\pm,0}\overline{\mathbf{a}}_{\pm,1}+\varepsilon\mathcal{L}_{\pm,1}\mathbf{a}_{\pm,0}\\ &+\varepsilon^2\mathcal{L}_{\pm,1}\overline{\mathbf{a}}_{\pm,1}+\varepsilon^2\mathcal{L}_{\pm,2}\mathbf{a}_{\pm,0}+\varepsilon^3\mathcal{L}_{\pm,2}\overline{\mathbf{a}}_{\pm,1}\big)G_\pm^\varepsilon\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}\boldsymbol{p}\,\mathrm{d}\boldsymbol{q}.\end{aligned} \tag{4.24}$$

Note that $\mathcal{L}_{\pm,0}\mathbf{a}_{\pm,0}=\mathbf{0}$ since the action Equation (4.10) and $\mathbf{a}_{\pm,0}$ is in the eigenspace of the Dirac symbol.

To show $\mathcal{L}_{\pm,0}\overline{\mathbf{a}}_{\pm,1}+\mathcal{L}_{\pm,1}\mathbf{a}_{\pm,0}=\mathbf{0}$, taking the + for example, it is suffices to show that, for each $m=\pm1,\pm2$, we have

$$\boldsymbol{\varUpsilon}_m^\dagger(\mathcal{L}_{+,0}\overline{\mathbf{a}}_{+,1}+\mathcal{L}_{+,1}\mathbf{a}_{+,0})=0. \tag{4.25}$$

When m is negative, since $\bar{\mathbf{a}}_{+,1} = \mathbf{a}_{+,1} + \mathbf{a}_{+,1}^{\perp}$ and $\mathbf{a}_{+,1} \in \ker \mathcal{L}_{+,0}$, Equation (4.25) is equivalent to

$$\boldsymbol{\Upsilon}_m^{\dagger} \left(\mathcal{L}_{+,0} \mathbf{a}_{+,1}^{\perp} + \mathcal{L}_{+,1} \mathbf{a}_{+,0} \right) = 0, \tag{4.26}$$

which is valid by the construction of $\mathbf{a}_{+,1}^{\perp}$. When m is positive, Equation (4.25) is valid by the evolutionary equation of $a_{m,0}$.

Therefore,

$$\|(\mathrm{i}\varepsilon\partial_t - \mathcal{D}) \boldsymbol{u}_{\mathrm{F},1}\|_{L^2} \le \varepsilon^2 C_{T,\delta} \left(\|\mathcal{L}_{\pm,1} \bar{\mathbf{a}}_{\pm,1} + \mathcal{L}_{\pm,2} \mathbf{a}_{\pm,0}\|_{L^\infty} + \varepsilon \|\mathcal{L}_{\pm,2} \bar{\mathbf{a}}_{\pm,1}\|_{L^\infty} \right). \tag{4.27}$$

By Lemma 4.2, $\|\mathcal{L}_{\pm,1} \bar{\mathbf{a}}_{\pm,1} + \mathcal{L}_{\pm,2} \mathbf{a}_{\pm,0}\|_{L^\infty}$ and $\|\mathcal{L}_{\pm,2} \bar{\mathbf{a}}_{\pm,1}\|_{L^\infty}$ are bounded. ∎

PROPOSITION 4.2. *Let $\boldsymbol{u}$ be the solution of the DS (1.7) and $\boldsymbol{u}_{\mathrm{F},1}$ be the corresponding FGA solution, then for any $T > 0$, there exists a constant C_T, so that for any $\varepsilon > 0$*

$$\sup_{0 \le t \le T} \|\boldsymbol{u} - \boldsymbol{u}_{\mathrm{F},1}\|_{L^2} \le \varepsilon C_T. \tag{4.28}$$

Proof. Let $e = \boldsymbol{u} - \tilde{\boldsymbol{u}}_{\mathrm{F},1}$, and then by Lemma 4.3,

$$\|e(t,\cdot)\|_{L^2} \le \|e(0,\cdot)\|_{L^2} + \varepsilon^{-1} \int_0^t \|(\mathrm{i}\varepsilon\partial_t - \mathcal{D}) \boldsymbol{u}_{\mathrm{F},1}(s,\cdot)\|_{L^2} \,\mathrm{d}s \le \varepsilon C_T. \tag{4.29}$$

∎

Noticing that $\|\boldsymbol{u} - \boldsymbol{u}_{\mathrm{F},0}\|_{L^2} \le \|\boldsymbol{u} - \boldsymbol{u}_{\mathrm{F},1}\|_{L^2} + \|\boldsymbol{u}_{\mathrm{F},1} - \boldsymbol{u}_{\mathrm{F},0}\|_{L^2}$, from Propositions 4.1 and 4.2, we can then state our main theorem on the accuracy of the FGA for the Dirac system,

THEOREM 4.1. *Let $\boldsymbol{u}$ be the solution of the DS (1.7) and $\boldsymbol{u}_{\mathrm{F},0}$ be the corresponding FGA solution, then for any $T > 0$, there exists a constant C_T, so that for any $\varepsilon > 0$*

$$\sup_{0 \le t \le T} \|\boldsymbol{u} - \boldsymbol{u}_{\mathrm{F},0}\|_{L^2} \le \varepsilon C_T. \tag{4.30}$$

Proof. The result follows from Propositions 4.1 and 4.2, and

$$\|\boldsymbol{u} - \boldsymbol{u}_{\mathrm{F},0}\|_{L^2} \le \|\boldsymbol{u} - \boldsymbol{u}_{\mathrm{F},1}\|_{L^2} + \|\boldsymbol{u}_{\mathrm{F},1} - \boldsymbol{u}_{\mathrm{F},0}\|_{L^2}. \tag{4.31}$$

∎

REMARK 4.2. We only need the initial conditions in Theorem 4.1 to be in L^2 for convergence, which is consistent with the previous accuracy estimate of Herman-Kluk propagator for the one-body Schrödinger equation [17]. The difference from Theorem 3.1 can be understood intuitively by the fact that the elastic wave system itself does not contain any high-frequency information, *i.e.*, the equations in (1.1) do not contain ε, therefore one can only introduce the high-frequency component via the initial conditions.

5. Generalization to linear non-strictly hyperbolic systems

In the previous sections, we have precisely analyzed the order of convergence of the FGA for two fundamental examples of linear non-strictly hyperbolic systems. The analysis for strictly hyperbolic systems was proposed in [13]; rather than a complete analysis of convergence of the FGA for linear non-strictly hyperbolic systems which would require reiteration of results from [13], we discuss hereafter the extension to non-strictly hyperbolic systems using the same arguments as the ones used in Sections 3 and 4. Consider a linear hyperbolic system

$$\partial_t \boldsymbol{u} + \sum_{i=1}^{d} A_i(\boldsymbol{x}) \partial_{x_i} \boldsymbol{u} = \boldsymbol{0},$$

for $\boldsymbol{u}: \mathbb{R}^d \to \mathbb{R}^N$, with A_i smooth, and $\sum_{i=1}^{d} p_i A_i(\boldsymbol{q})$ having eigenvalues not all distinct $\{\boldsymbol{H}_n\}_{n=1}^{N}$. The general strategy for proving the convergence consists in estimating

$$\sup_{t \in [0,T]} \|\boldsymbol{u}(t,\cdot) - \boldsymbol{u}_{\mathrm{F},0}(t,\cdot)\|_{\mathrm{E}} \leq \varepsilon C_{T,\delta}, \tag{5.1}$$

where $\boldsymbol{u}$ is the exact solution and $\boldsymbol{u}_{\mathrm{F},0}$ the FGA at order 0. In order to estimate (5.1), one must estimate the following 3 terms;

$$\|\boldsymbol{u} - \tilde{\boldsymbol{u}}_{\mathrm{F},1}\|_{\mathrm{L}^2}, \|\tilde{\boldsymbol{u}}_{\mathrm{F},1} - \boldsymbol{u}_{\mathrm{F},1}\|_{\mathrm{L}^2}, \|\boldsymbol{u}_{\mathrm{F},1} - \boldsymbol{u}_{\mathrm{F},0}\|_{\mathrm{L}^2}, \tag{5.2}$$

where $\tilde{\boldsymbol{u}}_{\mathrm{F},1}$ is a filtered version of FGA, and $\boldsymbol{u}_{\mathrm{F},1}$ is the first-order FGA. This is in particular the strategy which is used in [13, Theorem 4.1] (the corresponding notation in the latter reference, are $\mathcal{P}_t \boldsymbol{u}_0^{\varepsilon}$, $\mathcal{P}_{t,K,\delta}^{\varepsilon} \boldsymbol{u}$, $\widetilde{\mathcal{P}}_{t,K,\delta}^{\varepsilon} \boldsymbol{u}_0^{\varepsilon}$). However, in the non-strictly hyperbolic case, the FGA and filtered FGA possess asymptotic *correction terms* $\mathbf{a}_n^{\perp}$ which allow for dealing with the multiplicity of eigenvalues. This correction term, $\mathbf{a}_n^{\perp}$, can be explicitly evaluated thanks to bounded pseudo-inverse operators $\mathcal{L}_n^{-1}$, using similar compactness arguments on the eigenvalues and eigenvectors as in Proposition 3.4 for elastic wave system and Lemma 4.1 for the Dirac system. Then proving $\|\boldsymbol{u} - \boldsymbol{u}_{\mathrm{F},0}\|_{\mathrm{L}^2} \leq \varepsilon C_{T,\delta}$ is identical in the strict and non-strict hyperbolic cases as in [13, Theorem 4.1]. The estimates on the other terms in (5.2), are a consequence of the boundeness of the correction terms and arguments from [13], which were reiterated in Sections 3 and 4 for the elastic wave equations and the Dirac system respectively.

6. Conclusion and discussion

In this paper, we established the convergence theory of FGA for elastic wave system (EWS) and Dirac system (DS), which has been numerically verified as an efficient tool to compute high-frequency wave propagation. Unlike the convergence theory of FGA for strictly linear hyperbolic systems [13], we needed to analyze the boundedness of intraband transitions in diabatic coupling, which only appears when the system is non-strictly hyperbolic. The techniques we have developed for proving the convergence of FGA for both EWS and DS can be straightforwardly used to prove the convergence of FGA for other non-strictly hyperbolic systems. Extensions to high order approximation is possible by including correction terms in the amplitude function of the FGA formulation, following essentially the same strategy introduced in [13]. However, the calculations will be much more complex due to the non-strict hyperbolicity, and the resulting governing equations for the correction terms are not straightforward to develop parallel algorithms. Therefore, we shall leave it as our future work.

Acknowledgments. L.C. was partially supported by the NSFC grant 11901601. J.C.H. and X.Y. were partially supported by the NSF grant DMS-1818592. E.L. was partially supported through the NSERC Discovery Grant program.

REFERENCES

[1] A. Bach, *An Introduction to Semiclassical and Microlocal Analysis*, Universitext, Springer New York, 2011. 2.1

[2] L. Chai, E. Lorin, and X. Yang, *Frozen Gaussian approximation for the Dirac equation in semi-classical regime*, SIAM J. Numer. Anal., 57:2383–2412, 2019. 1, 4.1, 4.1

[3] L. Chai, P. Tong, and X. Yang, *Frozen Gaussian approximation for 3-D seismic wave propagation*, Geophys. J. Int., 208(1):59–74, 2017. 1

[4] L. Chai, P. Tong, and X. Yang, *Frozen Gaussian approximation for 3-D seismic tomography*, Inverse Probl., 34:055004, 2018. 1

[5] J.C. Hateley, J. Roberts, K. Mylonakis, and X. Yang, *Deep learning seismic substructure detection using the frozen Gaussian approximation*, J. Comput. Phys., 409:109313, 2020. 1

[6] J.C. Hateley, X. Yang, L. Chai, and P. Tong, *Frozen Gaussian approximation for 3-D elastic wave equation and seismic tomography*, Geophys. J. Int., 216(2):1394–1412, 2018. 1, 1, 1, 1, 3.1, 3.2, 3.2, 3.2

[7] M.F. Herman and E. Kluk, *A semiclassical justification for the use of non-spreading wavepackets in dynamics calculations*, Chem. Phys., 91:27–34, 1984. 1

[8] K. Kay, *Integral expressions for the semi-classical time-dependent propagator*, J. Chem. Phys., 100:4377–4392, 1994. 1

[9] K. Kay, *The Herman-Kluk approximation: Derivation and semiclassical corrections*, Chem. Phys., 322:3–12, 2006. 1

[10] E. Lorin and X. Yang, *Frozen Gaussian approximation-based two-level methods for multi-frequency Schrödinger equation*, Comput. Phys. Commun., 207:145–159, 2016. 1

[11] E. Lorin, X. Yang, and X. Antoine, *Frozen Gaussian approximation based domain decomposition methods for the linear Schrödinger equation beyond the semi-classical regime*, J. Comput. Phys., 315:221–237, 2016. 1

[12] J. Lu and X. Yang, *Frozen Gaussian approximation for high frequency wave propagation*, Commun. Math. Sci., 9:663–683, 2011. 1

[13] J. Lu and X. Yang, *Convergence of frozen Gaussian approximation for high frequency wave propagation*, Comm. Pure Appl. Math., 65:759–789, 2012. 1, 1, 2.2, 2.2, 2.2, 3.3, 3.3, 3.3, 5, 5, 6

[14] J. Lu and X. Yang, *Frozen Gaussian approximation for general linear strictly hyperbolic systems: Formulation and Eulerian methods*, Multiscale Model. Simul., 10:451–472, 2012. 1

[15] A. Martinez, *An Introduction to Semiclassical and Microlocal Analysis*, Springer-Verlag, New York, 2002. 2.1

[16] D. Robert, *On the Herman-Kluk semiclassical approximation*, Rev. Math. Phys., 22:1123–1145, 2010. 1

[17] T. Swart and V. Rousse, *A mathematical justification of the Herman-Kluk propagator*, Commun. Math. Phys., 286:725–750, 2009. 1, 2.2, 4.2

[18] B. Thaller, *The Dirac Equation*, Texts and Monographs in Physics, Springer-Verlag, Berlin, 1992. 1

[19] X. Yang and J. Zhang, *Computation of the Schrödinger equation in the semiclassical regime on unbounded domain*, SIAM J. Numer. Anal., 52:808–831, 2014. 1