CHALMERS | UNIVERSITY OF GOTHENBURG

PREPRINT 2012:1

On Discontinuous Galerkin Multiscale Variational Scheme for a Coupled Damped Nonlinear Schrödinger System of Equations

M. ASADZADEH
D. ROSTAMY
F. ZABIHI

Department of Mathematical Sciences
Division of Mathematics
CHALMERS UNIVERSITY OF TECHNOLOGY
UNIVERSITY OF GOTHENBURG
Gothenburg Sweden 2012

Preprint 2012:1

On Discontinuous Galerkin Multiscale Variational Scheme for a Coupled Damped Nonlinear Schrödinger System of Equations

M. Asadzadeh, D. Rostamy, and F. Zabihi

Department of Mathematical Sciences
Division of Mathematics
Chalmers University of Technology and University of Gothenburg
SE-412 96 Gothenburg, Sweden
Gothenburg, January 2012

Preprint 2012:1
ISSN 1652-9715

Matematiska vetenskaper
Göteborg 2012

ON DISCONTINUOUS GALERKIN MULTISCALE VARIATIONAL SCHEME FOR A COUPLED DAMPED NONLINEAR SCHRÖDINGER SYSTEM OF EQUATIONS

M. ASADZADEH[1], D. ROSTAMY, AND F. ZABIHI

Abstract. In this paper we study a streamline diffusion based discontinuous Galerkin approximation for the numerical solution of a coupled nonlinear Schrödinger equation system and extend the resulting method to a multiscale variational scheme. We prove stability estimates and derive optimal convergence rates due to the maximal available regularity of the exact solution. In the weak formulation, to make the underlying bilinear form, it was necessary to supply the equation system with an artificial viscosity term with a small coefficient of order proportional to a power of mesh size. We justify the theory by implementing an example of an application of the time dependent Schrödinger equation in the coupled ultrafast laser.

1. Introduction

The coupled nonlinear Schrödinger equation (CNSE) describes several interesting physical phenomena. In fiber communication system, such equations have been shown to govern pulse propagation along orthogonal polarization axes in nonlinear optical fibers and in wavelength-division-multiplexed systems [13, 24, 30, 31]. These equations also model beam propagation inside crystals or photo refractive as well as water wave interactions. Solitary waves in these equations are often called vector solutions in the literature as they generally contain two components. In all the above physical situations, collision of vector solutions is an important issue. This system has been studied intensively in recent years. It has been shown that, in addition to passing-through collision, vector solutions can also bounce off each other or trap each other. The stationary forms of these equations are investigated by a number of authors and a physical problem was introduce by Fushchych in [13].

As for the numerical methods for this type of problems, the discontinuous Galerkin method has been considered by several authors in various settings, see e.g. [2, 3, 4, 12, 22]. In our study, as a result of a special streamline-diffusion based DG scheme, artificial diffusion is added only in the characteristic direction so that internal layers are not smeared out while the added diffusion removes oscillations near boundary layers. In the streamline diffusion (SD) method, roughly speaking, the weak formulations is modified through adding the convection term in the equation expressed in terms of the test function with a small coefficient of the order of mesh size to the test functions. Then, e.g. in stability estimates, choosing test function as the trial function this would automatically correspond to add of a small diffusion term to the strong form of the equation in the streamline direction. SD method which was first introduced by Hughes and Brooks in [16] and Hughes and Mallet in [17] for the fluid problems is further studied and mathematically developed, e.g. for the hyperbolic partial differential equations and convection-diffusion problems in [5, 3, 4, 12, 22].

Key words and phrases. coupled nonlinear Schrödinger equations, multiscale variational scheme, discontinuous Galerkin method, streamline diffusion method(SD), stability, convergence.

[1] The research of this author is partially supported by the *Swedish Foundation of Strategic Research (SSF) in Gothenburg Mathematical Modeling Center (GMMC)* and *The Swedish Research Council (VR).*

We also study a multiscale method, an approach that uses both local and global information computed on different scales. We observe that the coupled multiphysics models in science and engineering are multiscale in nature such that they involve several types of physics that interact in space and time. For example, things are made up of atoms and electrons at the atomic scale, and at the same time are characterized by their own geometric dimensions which are usually several orders of magnitude larger. In this work, a hybrid method, combining multiscale and discontinuous Galerkin methods, is given for a coupled nonlinear Schrödinger equation. Stability estimates and convergence rates are derived for the DG part.

We consider the *coupled damped nonlinear system of Schrödinger equations with additional convection term* (CDNSEC) of the form

$$\begin{cases} \mathbf{i}\frac{\partial \psi_1}{\partial t} + \beta.\nabla\psi_1 + \frac{1}{2}\triangle\psi_1 + \varepsilon\left(|\psi_1|^2 + \alpha|\psi_2|^2\right)\psi_1 = 0, \\ \mathbf{i}\frac{\partial \psi_2}{\partial t} + \beta.\nabla\psi_2 + \frac{1}{2}\triangle\psi_2 + \varepsilon\left(\alpha|\psi_1|^2 + |\psi_2|^2\right)\psi_2 = 0, \end{cases} \quad (\mathbf{x},t) \in \Omega_T = \Omega\times[0,T], \tag{1.1}$$

where ψ_1 and ψ_2 are the wave amplitudes in two polarizations, $\Omega \subset \mathbb{R}^d, d = 1,2,3$ is a bounded domain with boundary $\partial\Omega$, α is the cross-phase modulation coefficient, $\varepsilon \geq 0$ is a small damping factor which controls the nonlinearity and β is a *linear* function of $\mathbf{x}$ representing the convection velocity. The equation (1.1) is associated with the following initial conditions:

$$\psi_1(\mathbf{x},0) = g_1(\mathbf{x}), \qquad \psi_2(\mathbf{x},0) = g_2(\mathbf{x}), \tag{1.2}$$

and also the Neumann type boundary conditions:

$$\nabla\psi_1(\mathbf{x},t) = \nabla\psi_2(\mathbf{x},t) = 0, \quad \text{for } \mathbf{x} \in \partial\Omega_\beta^- = \{\mathbf{x}\in\partial\Omega : n_{\mathbf{x}}.\beta < 0\}. \tag{1.3}$$

Finally $\mathbf{i}$ is the complex unity: $\mathbf{i}^2 = -1, \quad \mathbf{i} = \sqrt{-1}$. We assume that the solution of the system (1.1) is negligibly small outside the d-dimensional domain $[x_L, x_R]^d$, (otherwise one may replace ψ_1 and ψ_2 by appropriate multiplicative cut-offs).

We decompose the complex functions ψ_1 and ψ_2 in the CDNSEC into their real and imaginary parts,

$$\psi_1(\mathbf{x},t) = u_1 + \mathbf{i}u_2, \quad \psi_2(\mathbf{x},t) = u_3 + \mathbf{i}u_4,$$

where $(u_i, i = 1,...,4)$ are real functions. Thus, the system (1.1) can be converted to

$$\begin{cases} \frac{\partial u_1}{\partial t} + \beta.\nabla u_2 + \frac{1}{2}\triangle u_2 + \varepsilon z_1 u_2 = 0, \\ \frac{\partial u_2}{\partial t} - \beta.\nabla u_1 - \frac{1}{2}\triangle u_1 - \varepsilon z_1 u_1 = 0, \\ \frac{\partial u_3}{\partial t} + \beta.\nabla u_4 + \frac{1}{2}\triangle u_4 + \varepsilon z_2 u_4 = 0, \\ \frac{\partial u_4}{\partial t} - \beta.\nabla u_3 - \frac{1}{2}\triangle u_3 - \varepsilon z_2 u_3 = 0, \end{cases} \tag{1.4}$$

where

$$z_1 = u_1^2 + u_2^2 + \alpha(u_3^2 + u_4^2), \quad \text{and} \quad z_2 = \alpha(u_1^2 + u_2^2) + u_3^2 + u_4^2.$$

The system (1.4) can be written in a matrix form as

$$\begin{cases} \mathbf{u}_t - \frac{1}{2}A\Delta\mathbf{u} + \mathcal{F}_\varepsilon^\beta(\mathbf{u})\mathbf{u} = 0, & \text{in } \ (\mathbf{x},t)\in\Omega_T, \\ \mathbf{u}(\mathbf{x},0) = \mathbf{u_0}, & \text{on } \ \Omega, \\ \mathbf{u} = 0, & \text{on } \ \partial\widetilde{\Omega} \subset \partial\Omega_\beta^+, \\ -\nabla\mathbf{u}\cdot\mathbf{n} = 0, & (\mathbf{x},t)\in\partial\Omega_\beta^-, \end{cases} \tag{1.5}$$

where $\mathcal{F}_\varepsilon^\beta(\mathbf{u})\mathbf{u} = \varepsilon F(\mathbf{u})\mathbf{u} - \widetilde{\beta}^T A\nabla\mathbf{u}$, $\mathbf{u}_t = \frac{\partial\mathbf{u}}{\partial t}$, $\Delta\mathbf{u} = \sum_{i=1}^d \frac{\partial^2\mathbf{u}}{\partial {x_i}^2}$, and $\nabla\mathbf{u} = (\frac{\partial\mathbf{u}}{\partial x_1}, ..., \frac{\partial\mathbf{u}}{\partial x_n})$,

$$\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix}, \ A = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \ F(\mathbf{u}) = \begin{pmatrix} 0 & z_1 & 0 & 0 \\ -z_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & z_2 \\ 0 & 0 & -z_2 & 0 \end{pmatrix}, \widetilde{\beta} = \begin{pmatrix} \beta. \\ \beta. \\ \beta. \\ \beta. \end{pmatrix}.$$

Further, $\mathbf{n}$ is the outward unit normal to $\partial\Omega$, and $\partial\widetilde{\Omega}$ has a positive Lebesgue measure. Finally, if in the initial conditions (1.2), we let $g_j(x) = g_{j1}(x) + ig_{j2}(x)$ for $j = 1, 2$ then, $\mathbf{u}_0 = (g_{11}, g_{12}, g_{21}, g_{22})^T$.

Since A results to a vanishing viscosity (simply, due to the symmetry in the definition of A, we have $(A\mathbf{u}, \mathbf{u}) = (A\nabla\mathbf{u}, \nabla\mathbf{u}) = 0$). Therefore, we shall study the system of Schrödinger equation (1.5), with an additional viscosity term, viz

$$\begin{cases} \mathbf{u}_t - \frac{1}{2}(A + \tilde{\varepsilon}I)\Delta\mathbf{u} + \mathcal{F}_\varepsilon^\beta(\mathbf{u})\mathbf{u} = 0, & \text{in} \quad (\mathbf{x}, t) \in \Omega_T, \\ \mathbf{u}(\mathbf{x}, 0) = \mathbf{u_0}, & \text{on} \quad \Omega, \\ \mathbf{u} = 0, & \text{on} \quad \partial\widetilde{\Omega} \subset \partial\Omega_\beta^+, \\ -\nabla\mathbf{u}\cdot\mathbf{n} = 0, & (\mathbf{x}, t) \in \partial\Omega_\beta^-, \end{cases} \tag{1.6}$$

where $\tilde{\varepsilon} > 0$ is small, and I is the 4×4 identity matrix.

Finite differences have been the most dominating method in the numerical study of the coupled nonlinear Schrödinger equation, see, e.g. [18, 19, 20, 30, 31, 32]. In the most recent studies, however, the focus has been moved towards some aspects of finite element approach, e.g. local discontinuous Galerkin methods are used for solving the general nonlinear Schrödinger equation as well as the coupled nonlinear Schrödinger equation are studied in [32] where L_2 stability was obtained in both cases. However, convective terms of the type included in (1.1) are not considered elsewhere. Below, we compare the contribution of this work with related results from the existing literature. First of all, the streamline diffusion based stability estimates and convergence analysis for the discontinuous Galerkin approximation of the problem (1.5) are not studied elsewhere. Further, to construct a variational multiscale scheme (VMS) to the nonlinear problem (1.1), we follow the VMS procedure for the linear case (corresponding to $\varepsilon = 0$). Hence, we need to modify the linear scheme considering the additional contributions from the nonlinear terms.

An outline of this paper is as follows: In Section 2, we introduce discontinuous Galerkin variational muliscale scheme (DGVMS) based on streamline diffusion for CDNSEC. Section 3 is devoted to the study of stability estimates and proof of convergence rates. In Section 4 we construct both DGVMS, and SDVMS for the problem (1.5) as $\varepsilon \to 0$. Computational results are given in Section 5, and finally our concluding remarks are included in Section 6.

2. Hybrid DG/SD and variational multiscale method for CDNSEC

2.1. **Hybrid streamline-diffusion based discontinuous Galerkin.** In this section we consider discontinuous Galerkin method to solve the equation (1.5) using the finite element approximation of the space-time domain Ω_T. To this end, let $0 = t_0 < t_1 < ... < t_N = T$ be a subdivision of the time interval $[0, T]$ into the intervals $I_n = (t_{n-1}, t_n]$, with time steps $k_n = t_n - t_{n-1}$, $n = 1, ..., N$ and introduce the corresponding space-time slabs,

$$S_n = \{\ (\mathbf{x}, t) : \quad \mathbf{x} \in \Omega, \quad t_{n-1} < t \leq t_n\}, \qquad n = 1, \ldots, N. \tag{2.1}$$

We define a finite element structure on Ω_T, based on a partition $\mathcal{T}_h := \{\tau\}$ of the spatial domain Ω into triangular (or tetrahedral) elements τ satisfying the usual minimal angle condition. Then, we let $\{\mathcal{C}_h\}$ be the corresponding subdivision of Ω_T into elements $K_n := \tau \times I_n$, with the mesh parameter $h = diam(K_n)$ and $P_k(K_n) = P_k(\tau) \times P_k(I_n)$ being the set of polynomials in $\mathbf{x}$ and t of degree at most $k \geq 0$ on τ_n and k on I_n. Note that $\{\mathcal{C}_h\}$ should be viewed as union of the following subdivision on S_n's, $n = 1, ..., N$:

$$\{\mathcal{C}_{h,n}\} := \{K_n : K_n := \tau \times I_n, \tau \in \mathcal{T}_h\}.$$

To introduce a finite element method using discontinuous trial functions, we define the following notation: if $\hat{\beta}$ is a given smooth d-dimensional vector field on Ω, we define for

$K \in \mathcal{C}_h$,

$$\partial K_{\substack{-\\+}}(\hat{\beta}) = \{(\mathbf{x},t) \in \partial K : \mathbf{n}_t(\mathbf{x},t) + \mathbf{n}(x,t).\hat{\beta}(\mathbf{x},t) \lesseqgtr 0\},$$

where $(\mathbf{n}, \mathbf{n}_t) = (\mathbf{n}_\mathbf{x}, \mathbf{n}_t)$ denotes the outward unit normal to $\partial K \subset \Omega_T$. We also introduce for $k \geq 0$, the following spaces;

$$W_h = \Pi_{n=1}^N W_h^n, \qquad \mathbf{W}_h = \Pi_{n=1}^N \mathbf{W}_h^n$$

where for $n = 1, 2, ..., N$;

$$W_h^n = \{g \in [L_2(S_n)] : g|_K \in P_k(K_n), \quad \forall\ K_n \in \mathcal{C}_{h,n}\}$$

is a finite element space on S_n,

$$\mathbf{W}_h^n = \{\mathbf{w} \in [L_2(S_n)]^4 : \mathbf{w}|_{K_n} \in [P_k(K)]^4, \quad \forall\ K_n \in \mathcal{C}_{h,n}\},$$

also,

$$W_h = \{g \in [L_2(\Omega_T)] : g|_K \in P_k(K), \quad \forall\ K \in \mathcal{C}_h\},$$

and

$$\mathbf{W}_h = \{\mathbf{w} \in [L_2(\Omega_T)]^4 : \mathbf{w}|_K \in [P_k(K)]^4, \quad \forall\ K \in \mathcal{C}_h\}.$$

To derive a variational formulation, for the diffusive part of (1.5), based on discontinuous trial functions, we choose $\delta = c_0 h^\gamma$, where $1 \leq \gamma \leq 2$, and c_0 is a positive constant. We shall also use the following notation:

$$(\mathbf{u}, \mathbf{g})_{\Omega_T} = \textstyle\sum_{K \in \mathcal{C}_h} (\mathbf{u}, \mathbf{g})_K, \qquad (\mathbf{u}, \mathbf{g})_K = \int_{I_n} \int_\tau \mathbf{u}^T .\ \mathbf{g}\ d\mathbf{x}dt,$$
$$\mathbf{u}_\pm(\mathbf{x}, t) = lim_{s \to 0^\pm} \mathbf{u}(\mathbf{x}, t+s), \qquad \langle \mathbf{u}, \mathbf{g} \rangle_{n,\Omega} = \int_\Omega \mathbf{u}^T(\mathbf{x}, t_n) \mathbf{g}(\mathbf{x}, t_n) d\mathbf{x}, \quad n \in \mathbb{Z}^+.$$

Further, different L_2-based norms are denoted by $\|.\|_\Omega := \|.\|_{L_2(\Omega_T)}$, $\|.\|_{k,\Omega_T} := \|.\|_{L_k(\Omega_T)}$, $\|.\|_{\infty,\Omega_T} := \|.\|_{L_\infty(\Omega_T)}$ and $\|.\|_s := \|.\|_{s,\Omega_T} = \|.\|_{H^s(\Omega_T)}$. Recall that since $\hat{\beta} = \mathcal{F}_\varepsilon^\beta(\mathbf{u_h})$ is an arbitrary vector and $\hat{\beta} \cdot \mathbf{n}$ is continuous across the interelement boundaries in $\mathcal{C}_h$ thus, $\partial K_{\substack{+\\-}}(\hat{\beta})$ is well defined.

The discontinuous Galerkin method for (1.5) can now be formulated as follows: Find $\mathbf{u} \in \mathbf{W}_h$ such that

$$\begin{aligned}
&\big(\mathbf{u}_{h,t} + \mathcal{F}_\varepsilon^\beta(\mathbf{u}_h)\mathbf{u}_h,\ \mathbf{g} + \delta(\mathbf{g}_t + \mathcal{F}_\varepsilon^\beta(\mathbf{u}_h)\mathbf{g})\big)_{\Omega_T} + \frac{1}{2}(A\nabla\mathbf{u}_h, \nabla\mathbf{g})_{\Omega_T} + \frac{1}{2}\tilde{\varepsilon}(I\nabla\mathbf{u}_h, \nabla\mathbf{g})_{\Omega_T} \\
&\quad - \frac{1}{2}\int_{\partial\Omega_- \times I} (A + \tilde{\varepsilon}I)(\nabla\mathbf{u}_h)\mathbf{g}\, d\sigma ds - \frac{1}{2}\delta\Big((A + \tilde{\varepsilon}I)\Delta\mathbf{u}_h, \mathbf{g}_t + \mathcal{F}_\varepsilon^\beta(\mathbf{u}_h)\mathbf{g}\Big)_{\Omega_T} \\
&\quad + \sum_{K \in \mathcal{C}_h} \int_{\partial K_-(\hat{\beta})'} [\mathbf{u}_h]\mathbf{g}_+ |\mathbf{n}_t + \mathbf{n}\cdot\hat{\beta}|\, d\sigma + \langle \mathbf{u}_{h,+}, \mathbf{g}_+ \rangle_{0,\Omega} = \langle \mathbf{u}_0, \mathbf{g}_+ \rangle_{0,\Omega},
\end{aligned} \tag{2.2}$$

where $\partial K_-(\hat{\beta})' = \partial K_-(\hat{\beta}) \backslash \Omega \times \{0\}$. Here, $\delta = 0$ is the usual discontinuous Galerkin method, and $\delta \sim C_\delta h^\mu$ has properties similar to the corresponding SD-method. Now, for $\mathbf{v} \in \mathbf{W}_h$, we introduce the bilinear form

$$\begin{aligned}
B(\mathcal{F}_\varepsilon^\beta(\mathbf{u}_h); \mathbf{v}, \mathbf{g}) = &\big(\mathbf{v}_t + \mathcal{F}_\varepsilon^\beta(\mathbf{u}_h)\mathbf{v}, \mathbf{g} + \delta(\mathbf{g}_t + \mathcal{F}_\varepsilon^\beta(\mathbf{u}_h)\mathbf{g})\big)_{\Omega_T} + \frac{1}{2}(A\nabla\mathbf{v}, \nabla\mathbf{g})_{\Omega_T} \\
&+ \frac{1}{2}\tilde{\varepsilon}(\nabla\mathbf{v}, \nabla\mathbf{g})_{\Omega_T} - \frac{1}{2}\int_{\partial\Omega_- \times I} ((A + \tilde{\varepsilon}I)\nabla\mathbf{v})\mathbf{g} d\sigma ds \\
&- \frac{1}{2}\delta\Big((A + \tilde{\varepsilon}I)\Delta\mathbf{v}, \mathbf{g}_t + \mathcal{F}_\varepsilon^\beta(\mathbf{u_h})\mathbf{g}\Big)_{\Omega_T} \\
&+ \sum_{K \in \mathcal{C}_h} \int_{\partial K_-(\hat{\beta})} [\mathbf{v}]\mathbf{g}_+ |\mathbf{n}_t + \mathbf{n}\cdot\hat{\beta}| d\sigma + \langle \mathbf{v}_+, \mathbf{g}_+ \rangle_{0,\Omega},
\end{aligned} \tag{2.3}$$

and the linear form L,

$$L(\mathbf{g}) = \langle \mathbf{u}_0, \mathbf{g}_+ \rangle_{0,\Omega},$$

Then, the finite element method for (1.5) can now be formulated in the compact form as follows: Find $\mathbf{u}_h \in \mathbf{W}_h$ such that

$$B(\mathcal{F}_\varepsilon^\beta(\mathbf{u}_h); \mathbf{u}_h, \mathbf{g}) = L(\mathbf{g}), \qquad \forall\ \mathbf{g} \in \mathbf{W}_h. \tag{2.4}$$

We observe that the scheme is based on the space time discretizations and recall that we are dealing with a nonlinear system due to the presence of $F(\mathbf{u})\mathbf{u}$ in $\mathcal{F}_\varepsilon^\beta$ (see the definitions of z_i and $F(\mathbf{u})$).

2.2. **Variational multiscale method.** Now we use a symmetric split proposed in [25], and split the solution $\mathbf{u}_h \in \mathbf{W}_h$ into the coarse part $\mathbf{u}_c \in \mathbf{W}_{\widehat{h},c}$ and the fine part $\mathbf{u}_f \in \mathbf{W}_{h,f}$ such that $\widehat{h} > h$ and

$$\mathbf{u}_h = \mathbf{u}_c + \mathbf{u}_f + \mathcal{T}\mathbf{u}_c. \tag{2.5}$$

Here $\mathbf{W}_{\widehat{h},c}$ and $\mathbf{W}_{h,f}$ are the function spaces for the coarse mesh and fine meshes on Ω_T i.e. for the meshes $K_n^c = \tau_c \times I_n$ and $K_n^f = \tau_f \times I_n$, respectively. Further, $\widehat{h} = dim\{K_n^c\}$ and we consider τ_c and τ_f are the coarse and fine mesh elements on Ω. We also introduce the following operators

$$\mathcal{T} : \mathbf{W}_{\widehat{h},c} \to \mathbf{W}_{h,f},$$
$$\mathcal{I}_c : \mathbf{W}_h \to \mathbf{W}_{\widehat{h},c},$$
$$I - \mathcal{I}_c : \mathbf{W}_h \to \mathbf{W}_{\widehat{h},c},$$

where, $\mathbf{W}_{h,f} = (I - \mathcal{I}_c)\mathbf{W}_h = \{\mathbf{u}_h \in \mathbf{W}_h : \mathcal{I}_c\mathbf{u}_h = 0\}$ and $\mathcal{T}$ is defined by the following formula, for an arbitrary $\mathbf{v}_c$.

$$B(\mathcal{T}\mathbf{v}_c, \mathbf{v}_f) = -B(\mathbf{v}_c, \mathbf{v}_f), \quad \forall \mathbf{v}_c \in \mathbf{W}_{h,c}, \quad \forall \mathbf{v}_f \in \mathbf{W}_{h,f}. \tag{2.6}$$

Here $B(\mathbf{u}, \mathbf{v}) := B(\mathcal{F}_\varepsilon^\beta(\mathbf{u}); \mathbf{u}, \mathbf{v})$. We assume the L^2-orthogonal spaces $\mathbf{W}_{\widehat{h},c}$, $\mathbf{W}_{h,f}$ such that $\mathbf{W}_h = \mathbf{W}_{\widehat{h},c} \bigoplus \mathbf{W}_{h,f}$. Therefore, we can rewrite (2.4) as the following problem: find $\mathbf{u}_f \in \mathbf{W}_{h,f}$ and $\mathbf{u}_c \in \mathbf{W}_{\widehat{h},c}$ such that

$$B(\mathbf{u}_c + \mathbf{u}_f + \mathcal{T}\mathbf{u}_c, \mathbf{v}_c + \mathbf{v}_f + \mathcal{T}\mathbf{v}_c) = L(\mathbf{v}_c + \mathbf{v}_f + \mathcal{T}\mathbf{v}_c), \quad \forall\ \mathbf{v}_c \in \mathbf{W}_{\widehat{h},c}, \quad \forall\ \mathbf{v}_f \in \mathbf{W}_{h,f}. \tag{2.7}$$

Fine-scale equations are derived by putting $\mathbf{v}_c = 0$ in (2.5):
Find $\mathcal{T}\mathbf{u}_c$ and $\mathbf{u}_f \in \mathbf{W}_{h,f}$ such that

$$B(\mathcal{T}(\mathbf{u}_c) + \mathbf{u}_f, \mathbf{v}_f) = (\mathcal{R}_\varepsilon^\beta(\mathbf{u}_c), \mathbf{v}_f), \quad \forall\ \mathbf{v}_f \in \mathbf{W}_{h,f} \tag{2.8}$$

$$B(\mathbf{u}_f, \mathbf{v}_f) = L(\mathbf{v}_f), \quad \forall\ \mathbf{v}_f \in \mathbf{W}_{h,f}, \tag{2.9}$$

$$B(\mathcal{T}(\mathbf{u}_c), \mathbf{v}_f) = -B(\mathbf{u}_c, \mathbf{v}_f) \quad \forall\ \mathbf{v}_f \in \mathbf{W}_{h,f}. \tag{2.10}$$

where

$$(\mathcal{R}_\varepsilon^\beta(\mathbf{u}_c), \mathbf{v}_f) := L(\mathbf{v}_f) - B(\mathbf{u}_c, \mathbf{v}_f).$$

The above equations are particularly easy, since the coarse and fine scales are completely separated. Also, the coarse scale solution is obtained by letting $\mathbf{v}_f = 0$ in (2.5) that is: Find $\mathbf{v}_c \in \mathbf{W}_{\widehat{h},c}$ such that

$$B(\mathbf{u}_c + \mathcal{T}\mathbf{u}_c, \mathbf{v}_c + \mathcal{T}\mathbf{v}_c) = L(\mathbf{v}_c + \mathcal{T}\mathbf{v}_c) - B(\mathbf{u}_f, \mathbf{v}_c + \mathcal{T}\mathbf{v}_c), \quad \forall\ \mathbf{v}_c \in \mathbf{W}_{\widehat{h},c}. \tag{2.11}$$

In (2.9) $\mathcal{T}\mathbf{v}_c$ and $\mathbf{u}_f$ are unknowns so they can be obtained by solving (2.7) and the definition of $\mathcal{T}$. We recall that $B(\mathbf{u}_f, \mathbf{v}_c + \mathcal{T}\mathbf{v}_c) = 0$ if $\mathcal{T}\mathbf{v}_c$ is exact.
Let now $\mathcal{N}$ be the number of coarse nodes and $\{\phi_i\}_{i\in\mathcal{N}}$ be the coarse basis functions, such

that $\mathbf{u}_c = \sum_i \mathbf{u}_c^i \phi_i$ for some coefficients $\mathbf{u}_c^i$. Then, if we introduce a partition of unity ρ_i such that $\sum_{i\in\mathcal{N}} \rho_i = 1$, we can write

$$\begin{cases} B(\mathcal{T}(\phi_i), \mathbf{v}_f) = -B(\phi_i, \mathbf{v}_f), \\ B(\mathbf{u}_f, \mathbf{v}_f) = L(\rho_i \mathbf{v}_f). \end{cases} \tag{2.12}$$

Therefore, by solving these equations for all coarse basis functions ϕ_i, we can write

$$\mathbf{u}_f = \sum_i \mathbf{u}_c^i \mathcal{T}\phi_i + \sum_i \mathbf{u}_{f,l,i}, \tag{2.13}$$

where $\mathbf{u}_{f,l} = \sum_i \mathbf{u}_{f,l,i} \in \mathbf{W}_{h,f}$. Therefore we can put this expression for $\mathbf{u}_f$ back into the coarse scale equation. Also, we can use $\{\phi_i\}_{i\in\mathcal{N}}$ as a basis functions for $\mathbf{W}_{\widehat{h},c}$ so, we have

$$B((\phi_i + \mathcal{T}\phi_i), \phi_j) = L(\phi_j) - B(\mathbf{u}_{f,l}, \phi_j), \quad \forall\ i, j \in \mathcal{N}. \tag{2.14}$$

Therefore we observe that we can obtain a best approximation for $\mathbf{u}_f$ by using (2.10) on a partition. Since, this procedure is related to supp$\{\phi_i\}$ for decaying of $\mathcal{T}\phi_i$, $\mathbf{u}_{f,l,i}$ and instead of solving the fine scale problem (2.6) on the whole domain we solve many smaller decoupled problems.

In the next sections, we should prove some optimal convergence as practical error estimate, also we construct fine scale problems both as $\varepsilon \to 0$ and also $\varepsilon > 0$ as in (2.3).

3. Stability for the SD based discontinuous Galerkin

Below, we derive stability estimates for the method (2.4). First we assume slab-wise jump discontinuities in t for the SD method. Then, the DG approach has discontinuities both in t and x. To this end we need to show that the bilinear form B introduced in (2.2) is coercive. Here we shall state the coercivity criterion both in the normal sense as well as in the multi-scale case of coarse and fine meshes. The proof is, however, given in the usual case and is easily expandable for the multi-case setting. Our stability and error estimates will be given in the following triple norm,

$$\begin{aligned} |||\mathbf{g}|||^2 =& \frac{1}{2}\Big\{2\delta||\mathbf{g}_t + \mathcal{F}_\varepsilon^\beta(\mathbf{u}_h)\mathbf{g}||_{\Omega_T}^2 + |\ \mathbf{g}_-\ |_N^2 + |\ \mathbf{g}_+\ |_0^2 \\ &+ \sum_{K\in\mathcal{C}_h} \int_{\partial K_-(\hat{\beta})'} [\mathbf{g}]^2 |\mathbf{n}_t + \mathbf{n}\cdot\hat{\beta}| d\sigma + Ch^{-1}||\nabla \mathbf{g}||_{\Omega_T}^2 \Big\}. \end{aligned} \tag{3.1}$$

For the standard continuous Galerkin method, both δ and the sum in the above norm are zero. For the streamline diffusion method, with discontinuities only in t, the sum in replaced by the sum of jumps over the discrete time levels: $\sum_{m=1}^{M-1} ||[\mathbf{g}]||_m^2$. The presence of the δ-term in the $|||\cdot|||$, is the main contribution of the SD-method to both better stability and enhanced convergence. The discontinuous Galerkin approach includes additional control of the jumps over interelement boundaries both in x and t.

Proposition 3.1. *Let B be defined as in* (2.2)*, then there exists a constant $\alpha > 0$ independent of h such that:*

$$B(\mathcal{F}_\varepsilon^\beta(\mathbf{u}_h); \mathbf{g}, \mathbf{g}) \geq \alpha |||\mathbf{g}|||^2, \quad \forall \mathbf{g} \in \mathbf{W}_h. \tag{3.2}$$

Proof. We use (2.3) and with setting $\mathbf{v}=\mathbf{g}$ it follows that

$$\begin{aligned} B(\mathcal{F}_\varepsilon^\beta(\mathbf{u}_h);\mathbf{g},\mathbf{g}) =&|\mathbf{g}_+|_0^2+\sum_{K\in\mathcal{C}_h}\int_{\partial K_-(\hat{\beta})'}[\mathbf{g}]\mathbf{g}_+|\mathbf{n}_t+\mathbf{n}.\hat{\beta}|d\sigma+(\mathbf{g}_t,\mathbf{g})_{\Omega_T}\\ &+(\mathcal{F}_\varepsilon^\beta(\mathbf{u}_h)\mathbf{g},\mathbf{g})_{\Omega_T}+\delta||\mathbf{g}_t+\mathcal{F}_\varepsilon^\beta(\mathbf{u}_h)\mathbf{g}||_{\Omega_T}^2+\frac{1}{2}(A\nabla\mathbf{g},\nabla\mathbf{g})_{\Omega_T}\\ &+\frac{1}{2}\tilde{\varepsilon}(\nabla\mathbf{g},\nabla\mathbf{g})_{\Omega_T}-\frac{1}{2}\int_{I\times\partial\Omega_-}((A+\tilde{\varepsilon}I)\nabla g)\,g\,d\sigma\,ds\\ &-\frac{1}{2}\delta((A+\tilde{\varepsilon}I)\Delta\mathbf{g},\mathbf{g}_t+\mathcal{F}_\varepsilon^\beta(\mathbf{u}_h)\mathbf{g})_{\Omega_T}:=\sum_{i=1}^{9}T_i. \end{aligned} \tag{3.3}$$

Bellow we shall estimate $T_i:\ i=1,\dots,9$, separately

$$\begin{aligned} \sum_{i=1}^{3}T_i&=(\mathbf{g}_t,\mathbf{g})_{\Omega_T}+\sum_{K\in\mathcal{C}_h}\int_{\partial K_-(\hat{\beta})'}[\mathbf{g}]\mathbf{g}_+|\mathbf{n}_t+\mathbf{n}.\hat{\beta}|d\sigma+|\mathbf{g}_+|_0^2\\ &=\frac{1}{2}\Big[|\mathbf{g}_+|_0^2+|\mathbf{g}_-|_N^2]+\sum_{K\in\mathcal{C}_h}\int_{\partial K_-(\hat{\beta})'}[\mathbf{g}]^2|\mathbf{n}_t+\mathbf{n}\cdot\hat{\beta}|d\sigma\Big]. \end{aligned} \tag{3.4}$$

By the definition of $(\varepsilon F(\mathbf{u}_h)\mathbf{g},\mathbf{g})$, we have

$$T_4=(\mathcal{F}_\varepsilon^\beta(\mathbf{u}_h)\mathbf{g},\mathbf{g})_{\Omega_T}=(\varepsilon F(\mathbf{u}_h)\mathbf{g},\mathbf{g})_{\Omega_T}-(\widetilde{\beta}^T\nabla A\mathbf{g},\mathbf{g})_{\Omega_T}:=T_{41}+T_{42}. \tag{3.5}$$

We can easily check that.

$$T_{41}=0, \tag{3.6}$$

Further, since $\widetilde{\beta}$ is linear in $\mathbf{x}$ we have $\nabla\widetilde{\beta}\equiv 0$, and therefore

$$\begin{aligned} T_{42}&=-(A\mathbf{g},\widetilde{\beta}\mathbf{g})_{\Omega_T}=(\nabla A\mathbf{g},\nabla(\widetilde{\beta}\mathbf{g}))_{\Omega_T}-\langle A\mathbf{g},\widetilde{\beta}\mathbf{g}\rangle_{\partial\Omega_T}\\ &=(A\mathbf{g},(\nabla\widetilde{\beta})\mathbf{g})_{\Omega_T}+(A\mathbf{g},\widetilde{\beta}\nabla\mathbf{g})_{\Omega_T}, \end{aligned} \tag{3.7}$$

Now, by the definition of A we have $-\langle A\mathbf{g},\widetilde{\beta}\mathbf{g}\rangle_{\partial\Omega_T}=0$, and hence

$$T_{42}=0. \tag{3.8}$$

Also, using the definition of A yields

$$T_6=0,\qquad\text{and}\quad T_7=\frac{1}{2}\tilde{\varepsilon}\|\nabla\mathbf{g}\|^2. \tag{3.9}$$

To estimate T_8 we have using first the trace estimate, and then the inverse estimate see [7],

$$\begin{aligned} -T_8&=\frac{1}{2}\int_{\partial\Omega_-\times I}((A+\tilde{\varepsilon}I)\nabla\mathbf{g})\mathbf{g}d\sigma ds\le\frac{\gamma_1}{4}\|(A+\tilde{\varepsilon}I)\nabla\mathbf{g}\|_{L_2(\partial\Omega_-\times I)}^2+\frac{1}{4\gamma_1}\|\mathbf{g}\|_{L_2(\partial\Omega_-\times I)}^2\\ &\le\sqrt[4]{8}\frac{\gamma_1}{4}\Big(\|(A+\tilde{\varepsilon}I)\nabla\mathbf{g}\|_{\Omega_T}\|(A+\tilde{\varepsilon}I)\Delta\mathbf{g}\|_{\Omega_T}\Big)+\sqrt[4]{8}\frac{1}{4\gamma_1}\|\mathbf{g}\|_{\Omega_T}\|\nabla\mathbf{g}\|_{\Omega_T}\\ &\le\sqrt[4]{8}\frac{\gamma_1}{4}h^{-1}\|(A+\tilde{\varepsilon}I)\nabla\mathbf{g}\|_{\Omega_T}^2+\sqrt[4]{8}\frac{1}{8\gamma_1}\|\mathbf{g}\|_{\Omega_T}^2+\sqrt[4]{8}\frac{1}{8\gamma_1}\|\nabla\mathbf{g}\|_{\Omega_T}^2. \end{aligned}$$

Similarly,

$$\begin{aligned} -T_9&=\frac{1}{2}\delta((A+\tilde{\varepsilon}I)\Delta\mathbf{g},\mathbf{g}_t+\mathcal{F}_\varepsilon^\beta(\mathbf{u}_h)\mathbf{g})_{\Omega_T}\le\frac{\gamma_2}{4}\delta\|(A+\tilde{\varepsilon}I)\Delta\mathbf{g}\|_{\Omega_T}^2+\frac{1}{4\gamma_2}\delta\|\mathbf{g}_t+\mathcal{F}_\varepsilon^\beta(\mathbf{u}_h)\mathbf{g}\|_{\Omega_T}^2\\ &\le\frac{\gamma_2}{4}\delta h^{-2}\|(A+\tilde{\varepsilon}I)\nabla\mathbf{g}\|_{\Omega_T}^2+\frac{1}{4\gamma_2}\delta\|\mathbf{g}_t+\mathcal{F}_\varepsilon^\beta(\mathbf{u}_h)\mathbf{g}\|_{\Omega_T}^2. \end{aligned}$$

Now we let $\delta \sim h$, and choose $\gamma_1 = \sqrt[4]{8}$, $\gamma_2 = 1$. Then, by Poincareé inequality (the solution vanishes in a part of boundary. i.e. $\widetilde{\partial\Omega}$ with positive measure) $\|\mathbf{g}\|_{\Omega_T} \leq |\Omega_T|\|\nabla\mathbf{g}\|_{\Omega_T}$ and the fact that the matrix norm of A is identity: $\|A\|_\infty = 1$, we end up with

$$
\begin{aligned}
-(T_8 + T_9) &\leq \frac{2\sqrt{2}}{4}h^{-1}(1+\tilde{\varepsilon})^2\|\nabla\mathbf{g}\|^2_{\Omega_T} + \frac{1}{8}\|\mathbf{g}\|^2_{\Omega_T} + \frac{1}{8}\|\nabla\mathbf{g}\|^2_{\Omega_T} \\
&+ \frac{1}{4}h^{-1}(1+\tilde{\varepsilon})^2\|\nabla\mathbf{g}\|^2_{\Omega_T} + \frac{1}{4}\delta\|\mathbf{g}_t + \mathcal{F}^\beta_\varepsilon(\mathbf{u}_h)\mathbf{g}\|^2_{\Omega_T} \\
&= \Big((\frac{2\sqrt{2}}{4} + \frac{1}{4})h^{-1}(1+\tilde{\varepsilon})^2 + \frac{1}{8}\Big)\|\nabla\mathbf{g}\|^2_{\Omega_T} + \frac{1}{8}\|\mathbf{g}\|^2_{\Omega_T} + \frac{1}{4}\delta\|\mathbf{g}_t + \mathcal{F}^\beta_\varepsilon(\mathbf{u}_h)\mathbf{g}\|^2_{\Omega_T} \\
&\leq \Big(\frac{(2\sqrt{2}+1)(1+\tilde{\varepsilon})^2 + 2(1+|\Omega_T|)h}{4h}\Big)\|\nabla\mathbf{g}\|^2_{\Omega_T} + \frac{1}{4}\delta\|\mathbf{g}_t + \mathcal{F}^\beta_\varepsilon(\mathbf{u}_h)\mathbf{g}\|^2_{\Omega_T}.
\end{aligned}
$$

Now, assuming, e.g. $C = \frac{\tilde{\varepsilon}}{8} + (2\sqrt{2}+1)(1+\tilde{\varepsilon})^2 + 2(1+|\Omega_T|)h$, and using (3.2)-(3.9) together with the above estimate for $T_8 + T_9$, yields the desired result. We omit the details. □

Below we state the corresponding coercivity estimates for the multi-scale cases:

Corollary 3.1. *Let B be defined as in* (2.2)*, then there exists a constant $\alpha > 0$ independent of h such that:*

$$B(\mathcal{F}^\beta_\varepsilon(\mathbf{u}_{h,f}); \mathbf{g}, \mathbf{g}) \geq \alpha|||\mathbf{g}|||^2, \quad \forall \mathbf{g} \in \mathbf{W}_{h,f}, \tag{3.10}$$

and

$$B(\mathcal{F}^\beta_\varepsilon(\mathbf{u}_{\widehat{h},c}); \mathbf{g}, \mathbf{g}) \geq \alpha|||\mathbf{g}|||^2, \quad \forall \mathbf{g} \in \mathbf{g} \in \mathbf{W}_{\widehat{h},c}. \tag{3.11}$$

Proof. The proof is similar as that of the above proposition and therefore is omitted. □

Remark 3.2. *The additional viscosity term $-\tilde{\varepsilon}\Delta\mathbf{u}$ in* (1.6) *is also the key factor to include the L_2-norm in the error estimates involving the triple norm $|||\mathbf{g}|||$. Below we state and prove a SD approach to the L_2-norm control, which includes the δ-term and jumps.*

Proposition 3.2. *For any constant $C_1 > 0$, we have for $\mathbf{g} \in \mathbf{W}_h$,*

$$
\begin{aligned}
||\mathbf{g}||^2_{\Omega_T} \leq \Big[\ \frac{1}{C_1}||\mathbf{g}_t + \mathcal{F}^\beta_\varepsilon(\mathbf{u}_h)\mathbf{g}||^2_{\Omega_T} + \sum_{n=1}^{N} |\ \mathbf{g}_-\ |^2_n + \sum_{K\in\mathcal{C}_h} \int_{\partial K_-(\beta_1)''} [\mathbf{g}]^2|\mathbf{n}\cdot\hat{\beta}|d\sigma \\
+ \int_{\partial\Omega_+\times I} \mathbf{g}^2|\mathbf{n}\cdot\beta_1|d\sigma ds\ \Big] h e^{C_1 h},
\end{aligned}
\tag{3.12}
$$

where

$$\beta_1 = (1, \mathcal{F}^\beta_\varepsilon(\mathbf{u}_h)) := (1, \hat{\beta}), \qquad (\beta_1)' = \Big(1, (\hat{\beta})'\Big),$$

and

$$\partial K_-(\beta_1)'' = \{(x,t) \in \partial K_-(\beta_1)' : \mathbf{n}_t(x,t) = 0\}.$$

Proof. We have for $t_n < t < t_{n+1},\ K = \tau \times I_n$,

$$
\begin{aligned}
||\mathbf{g}(t)||^2_\tau &= |\mathbf{g}_-|^2_{n+1,\tau} - \int_t^{t_{n+1}} \frac{d}{ds}||\mathbf{g}(s)||^2_\tau ds = |\mathbf{g}_-|^2_{n+1,\tau} - 2\int_t^{t_{n+1}} (\mathbf{g}_s, \mathbf{g})_\tau ds \\
&= |\mathbf{g}_-|^2_{n+1,\tau} - 2\int_t^{t_{n+1}} [(\mathbf{g}_s + \mathcal{F}^\beta_\varepsilon(\mathbf{u}_h)\mathbf{g}, \mathbf{g})_\tau - (\mathcal{F}^\beta_\varepsilon(\mathbf{u}_h)\mathbf{g}, \mathbf{g})_\tau]ds,
\end{aligned}
$$

where $|\mathbf{g}_-|_{n+1,\tau}$ is the obvious restriction of $|\mathbf{g}_-|_{n+1}$, to the spatial element τ. Now, using (1.5), we have that

$$(\mathcal{F}^\beta_\varepsilon(\mathbf{u}_h)\mathbf{g}, \mathbf{g})_\tau = (\varepsilon F(\mathbf{u}_h)\mathbf{g}, \mathbf{g})_\tau - (\widetilde{\beta}^T A\nabla\mathbf{g}, \mathbf{g})_\tau = 0 - (\widetilde{\beta}^T A\nabla\mathbf{g}, \mathbf{g})_\tau.$$

Using Green's formula, yields

$$-(\widetilde{\beta}^T A\nabla \mathbf{g},\mathbf{g})_\tau = -\Big(\nabla(\widetilde{\beta}^T A\mathbf{g}),\mathbf{g}\Big) + \Big(\nabla(\widetilde{\beta}^T)A\mathbf{g},\mathbf{g}\Big) = -\Big(\nabla(\widetilde{\beta}^T A\mathbf{g}),\mathbf{g}\Big) + 0$$
$$= (\widetilde{\beta}^T A\nabla \mathbf{g},\mathbf{g})_\tau - \langle\widetilde{\beta}^T A\mathbf{g},\mathbf{g}\rangle_{\partial\tau}.$$

Hence

$$(\mathcal{F}_\varepsilon^\beta(\mathbf{u}_h)\mathbf{g},\mathbf{g})_\tau = -\frac{1}{2}\langle\widetilde{\beta}^T A\mathbf{g},\mathbf{g}\rangle_{\partial\tau},$$

and since $\langle\varepsilon F(\mathbf{u}_h)\mathbf{g},\mathbf{g}\rangle_{\partial\tau}\equiv 0$, we obtain

$$(\mathcal{F}_\varepsilon^\beta(\mathbf{u}_h)\mathbf{g},\mathbf{g})_\tau = \frac{1}{2}\langle\mathcal{F}_\varepsilon^\beta(\mathbf{u}_h)\mathbf{g},\mathbf{g}\rangle_{\partial\tau}.$$

Thus

$$||\mathbf{g}(t)||_\tau^2 = |\mathbf{g}_-|_{n+1,\tau}^2 - 2\int_t^{t_{n+1}}\Big[\,(\mathbf{g}_s+\hat{\beta}\mathbf{g},\mathbf{g})_\tau - \frac{1}{2}\int_{\partial\tau_+}\mathbf{g}^2(\mathbf{n}\cdot\hat{\beta})d\sigma + \frac{1}{2}\int_{\partial\tau_-}\mathbf{g}^2(\mathbf{n}\cdot\hat{\beta})\,d\sigma\,\Big]ds.$$

Summing over τ, and using the fact that $\hat{\beta}=\mathcal{F}_\varepsilon^\beta(\mathbf{u}_h)$ we have

$$\begin{aligned}||\mathbf{g}(t)||_{\Omega_T}^2 &\le |\mathbf{g}_-|_{n+1}^2 + 2\int_t^{t_{n+1}}|(\mathbf{g}_s+\mathcal{F}_\varepsilon^\beta(\mathbf{u}_h)\mathbf{g},\mathbf{g})_{\Omega_T}|\,ds\\
&\quad+\sum_K\int_{\partial K_-(\beta_1)''\bigcap\{s:t<s<t_{n+1}\}}[\mathbf{g}^2]|\mathbf{n}\cdot\hat{\beta}|d\sigma + \int_{\partial\Omega_+\times\{s:t<s<t_{n+1}\}}\mathbf{g}^2|\mathbf{n}\cdot\beta_1|d\sigma ds\\
&\le |\mathbf{g}_-|_{n+1}^2 + \frac{1}{C_1}||\mathbf{g}_t+\mathcal{F}_\varepsilon^\beta(\mathbf{u}_h)\mathbf{g}||_n^2 + C_1\int_t^{t_{n+1}}||\mathbf{g}(s)||_\Omega^2 dt\\
&\quad+C\sum_K\int_{\partial K_-(\beta_1)''\bigcap I_n}[\mathbf{g}^2]|\mathbf{n}\cdot\hat{\beta}|d\sigma + \int_{\partial\Omega_+\times I_n}\mathbf{g}^2|\mathbf{n}\cdot\beta_1|d\sigma ds,\end{aligned}$$

where we have also used the relationship between $\widetilde{\beta}^T$, $\hat{\beta}$, β_1, and the following inequality

$$[\mathbf{g}^2] = \mathbf{g}_+^2 - \mathbf{g}_-^2 = (\mathbf{g}_+-\mathbf{g}_-)(\mathbf{g}_++\mathbf{g}_-) = [\mathbf{g}](\mathbf{g}_++\mathbf{g}_-) \le C[\mathbf{g}]^2 + \frac{1}{C}(\mathbf{g}_++\mathbf{g}_-)^2,$$

we take C sufficiently large and hide $\frac{1}{C}(\mathbf{g}_++\mathbf{g}_-)^2$ term in the norm on the left hand side, $||\mathbf{g}||_{\Omega_T}^2$. Then, using Grönwall inequality, we have

$$\begin{aligned}||\mathbf{g}||_{\Omega_T}^2 \le [\ &|\mathbf{g}_-|_{n+1}^2 + \frac{1}{C_1}||\mathbf{g_t}+\mathcal{F}_\varepsilon^\beta(\mathbf{u}_h)\mathbf{g}||_n^2\\
&+C\sum_K\int_{\partial K_-(\beta_1)''\bigcap I_n}[\mathbf{g}^2]|\mathbf{n}\cdot\hat{\beta}|d\sigma + \int_{\partial\Omega_+\times I_n}\mathbf{g}^2|\mathbf{n}\cdot\beta_1|d\sigma ds\ ]e^{Ch}.\end{aligned}$$

Integrating over I_n and summing over n (for $n=0,1,...,N-1$, and using a shifting as $n-1\to n$), we obtain the desired result. □

Remark 3.3. *The same estimates, with similar proofs, hold for multiscale cases, assuming* $\mathbf{g}\in\mathbf{W}_{h,i}\quad i=c,f$.

4. Convergence for the SD-based discontinuous Galerkin

Here, we use the standard finite element procedure and introduce the linear nodal interpolation $I_h\mathbf{u}\in\mathbf{W}_h$ of the exact solution $\mathbf{u}$. We set $\eta=\mathbf{u}-I_h\mathbf{u}$ and $\xi=\mathbf{u}_h-I_h\mathbf{u}$. Thus, we have

$$e := \mathbf{u}-\mathbf{u}_h = (\mathbf{u}-I_h\mathbf{u}) - (\mathbf{u}_h-I_h\mathbf{u}) = \eta-\xi,$$

We also have the Galerkin orthogonality relation

$$B(\mathcal{F}_\varepsilon^\beta(\mathbf{u}_h);e,\mathbf{g}) = 0, \qquad \forall\ \mathbf{g}\in\mathbf{W}_h. \tag{4.1}$$

Below we shall use Propositions 3.1 and 3.2 and derive optimal convergence rates. We follow an error estimate procedure that is more practical in implementations/computations.

Remark 4.1. *Analogously, the formulation for multiscale case reads as follows: Introduce the interpolants:* $I_{h,f}\mathbf{u} \in \mathbf{W}_{h,f}$ *and* $I_{\widehat{h},c}\mathbf{u} \in \mathbf{W}_{\widehat{h},c}$*, and set* $\eta_f = \mathbf{u} - I_{h,f}\mathbf{u}$*,* $\xi_f = \mathbf{u}_{h,f} - I_{h,f}\mathbf{u}$*, and* $\eta_c = \mathbf{u} - I_{h,c}\mathbf{u}$ *and* $\xi_c = \mathbf{u}_{\widehat{h},c} - I_{\widehat{h},c}\mathbf{u}$*. Then,*

$$e_f := \mathbf{u} - \mathbf{u}_{h,f} = (\mathbf{u} - I_{h,f}\mathbf{u}) - (\mathbf{u}_{h,f} - I_{h,f}\mathbf{u}) = \eta_f - \xi_f,$$

$$e_c := \mathbf{u} - \mathbf{u}_{\widehat{h},c} = (\mathbf{u} - I_{\widehat{h},c}\mathbf{u}) - (\mathbf{u}_{\widehat{h},c} - I_{\widehat{h},c}\mathbf{u}) = \eta_c - \xi_c,$$

and we have the corresponding Galerkin orthogonality relations are given by

$$B(\mathcal{F}_\varepsilon^\beta(\mathbf{u}_h); e_f, \mathbf{g}) = 0, \ \forall\ \mathbf{g} \in \mathbf{W}_{h,f}, \tag{4.2}$$

and

$$B(\mathcal{F}_\varepsilon^\beta(\mathbf{u}_{\widehat{h}}); e_c, \mathbf{g}) = 0, \ \forall\ \mathbf{g} \in \mathbf{W}_{\widehat{h},c}. \tag{4.3}$$

Theorem 4.2. *If the exact solution* $\mathbf{u} \in L^\infty(H^{k+1}(\Omega))$*,* $k \geq 1$*; for* (1.5) *satisfies*

$$||\mathbf{u}||_\infty + ||\mathcal{F}_\varepsilon^\beta(\mathbf{u})||_\infty + ||div\ (\mathcal{F}_\varepsilon^\beta(\mathbf{u}) - \varepsilon F(\mathbf{u}))||_\infty + ||\nabla\eta||_\infty \leq C,$$

then, we have the following error estimate

$$|||\mathbf{u} - \mathbf{u}_h||| \leq Ch^{k+1/2}, \qquad \mathbf{u}_h \in \mathbf{W}_h. \tag{4.4}$$

Proof. Using the definition of η, we can write

$$\begin{aligned} \alpha|||\xi|||^2 &\leq B(\mathcal{F}_\varepsilon^\beta(\mathbf{u}_h); \xi, \xi) = B(\mathcal{F}_\varepsilon^\beta(\mathbf{u}_h); \mathbf{u}_h, \xi) - B(\mathcal{F}_\varepsilon^\beta(\mathbf{u}_h); I_h\mathbf{u}, \xi) \\ &= L(\xi) - B(\mathcal{F}_\varepsilon^\beta(\mathbf{u}_h); I_h\mathbf{u}, \xi) = B(\mathcal{F}_\varepsilon^\beta(\mathbf{u}); \mathbf{u}, \xi) - B(\mathcal{F}_\varepsilon^\beta(\mathbf{u}_h); I_h\mathbf{u}, \xi) \\ &= B(\mathcal{F}_\varepsilon^\beta(\mathbf{u}_h); \eta, \xi) + B(\mathcal{F}_\varepsilon^\beta(\mathbf{u}); \mathbf{u}, \xi) - B(\mathcal{F}_\varepsilon^\beta(\mathbf{u}_h); \mathbf{u}, \xi) := T_1 + T_2 - T_3. \end{aligned} \tag{4.5}$$

Now we estimate the terms T_1 and $T_2 - T_3$, separately. For the term T_1 we use the inverse inequality and the above assumptions to obtain

$$\begin{aligned} T_1 =& B(\mathcal{F}_\varepsilon^\beta(\mathbf{u}_h); \eta, \xi) = (\eta_t + \mathcal{F}_\varepsilon^\beta(\mathbf{u}_h)\eta, \xi + \delta(\xi_t + \mathcal{F}_\varepsilon^\beta(\mathbf{u_h})\xi))_{\Omega_T} \\ &+ \sum_{K\in\mathcal{C}_h} \int_{\partial K_-(\hat\beta)'} [\eta]\xi_+|\mathbf{n}_t + \mathbf{n}\cdot\hat\beta| d\sigma + \frac{1}{2}(A\nabla\eta, \nabla\xi)_{\Omega_T} + \frac{1}{2}\tilde\varepsilon(\nabla\eta, \nabla\xi)_{\Omega_T} \\ &- \frac{1}{2}\int_{\partial\Omega_-\times I} ((A + \tilde\varepsilon I)\nabla\eta)\xi d\sigma ds - \frac{1}{2}\delta((A + \tilde\varepsilon I)\Delta\eta, \xi_t + \mathcal{F}_\varepsilon^\beta(\mathbf{u}_h)\xi)_{\Omega_T} + < \eta_+, \xi_+ >_0 \\ :=& \sum_{i=1}^{7} S_i. \end{aligned}$$

Thus, we need to estimate S_i, $1 \leq i \leq 7$,

$$\begin{aligned} S_1 =& (\eta_t + \mathcal{F}_\varepsilon^\beta(\mathbf{u}_h)\eta, \xi)_{\Omega_T} + \delta(\eta_t + \mathcal{F}_\varepsilon^\beta(\mathbf{u}_h)\eta, \xi_t + \mathcal{F}_\varepsilon^\beta(\mathbf{u_h})\xi)_{\Omega_T} \\ =& \sum_{K\in\mathcal{C}_h} \int_{\partial K} \eta\xi(\mathbf{n}_t + \mathbf{n}\cdot\hat\beta) d\sigma - \sum_{K\in\mathcal{C}_h} (\eta, \xi_t + \hat\beta\xi)_K - (\eta, \xi \mathrm{div}(\beta_1))_{\Omega_T} \\ &+ \delta(\eta_t + \mathcal{F}_\varepsilon^\beta(\mathbf{u}_h)\eta, \xi_t + \mathcal{F}_\varepsilon^\beta(\mathbf{u_h})\xi)_{\Omega_T}. \end{aligned} \tag{4.6}$$

We may write the first term in S_1 as

$$\begin{aligned}\sum_{K\in\mathcal{C}_h}\int_{\partial K}\eta\xi(n_t+\mathbf{n}\cdot\hat{\beta})d\sigma &= \sum_{K\in\mathcal{C}_h}\int_{\partial K_-}\eta_+\xi_+(\mathbf{n}_t+\mathbf{n}\cdot\hat{\beta})d\sigma\\ &+\sum_{K\in\mathcal{C}_h}\int_{\partial K_+}\eta_-\xi_-(\mathbf{n}_t+\mathbf{n}\cdot\hat{\beta})d\sigma = -\sum_{K\in\mathcal{C}_h}\int_{\partial K_-(\beta_1)'}\eta_+\xi_+|\mathbf{n}_t+\mathbf{n}\cdot\hat{\beta}|d\sigma\\ &+\sum_{K\in\mathcal{C}_h}\int_{\partial K_-(\beta_1)'}\eta_-\xi_-|\mathbf{n}_t+\mathbf{n}\cdot\hat{\beta}|d\sigma+\int_{\partial\Omega_+\times I}\eta_-\xi_-|\mathbf{n}\cdot\beta_1|d\sigma ds.\end{aligned}$$

Therefore, combining $\partial K_-(\beta_1)'$ and $\partial\Omega_+$ terms of S_1 and S_2, we have

$$\begin{aligned}\sum_{K\in\mathcal{C}_h}\Big[\int_{\partial K_-(\beta_1)'}&\eta_-\xi_-|\mathbf{n}_t+\mathbf{n}\cdot\hat{\beta}|d\sigma-\int_{\partial K_-(\beta_1)'}\eta_+\xi_+|\mathbf{n}_t+\mathbf{n}\cdot\hat{\beta}|d\sigma\\ &+\int_{\partial K_-(\beta_1)'}[\eta]\xi_+|\mathbf{n}_t+\mathbf{n}\cdot\hat{\beta}|d\sigma\Big]\\ &=\sum_{K\in\mathcal{C}_h}\Big[\int_{\partial K_-(\beta_1)'}\eta_-\xi_-|\mathbf{n}_t+\mathbf{n}\cdot\hat{\beta}|d\sigma-\int_{\partial K_-(\beta_1)'}\eta_-\xi_+|\mathbf{n}_t+\mathbf{n}\cdot\hat{\beta}|d\sigma\Big]\\ &-\sum_{K\in\mathcal{C}_h}\int_{\partial K_-(\beta_1)'}\eta_-[\xi]|\mathbf{n}_t+\mathbf{n}\cdot\hat{\beta}|d\sigma\\ &=\sum_{K\in\mathcal{C}_h}\int_{\partial K_-(\beta_1)''}\eta_-[\xi]|\beta_1\cdot\mathbf{n}|d\sigma-\sum_{n=1}^{N-1}<\eta_-,[\xi]>_n .\end{aligned}\tag{4.7}$$

Note that, in the last equality above, by transferring the boundary integrals from $\partial K_-(\beta_1)'$ to $\partial K_-(\beta_1)''$ we should keep the track of the jump terms in the time direction separately.

To bound the first term on the right-hand side above:

$$T:=\sum_{K\in\mathcal{C}_h}\int_{\partial K_-(\beta_1)''}\eta_-[\xi]|\beta_1\cdot\mathbf{n}|d\sigma,\tag{4.8}$$

we use Cauchy-Schwarz inequality, and with $\delta>0$, we may write

$$|T|\le\frac{C}{\delta}\sum_{K\in\mathcal{C}_h}\int_{\partial K_-(\beta_1)''}|\eta_-|^2|\mathbf{n}\cdot\beta_1|d\sigma+C\delta\sum_{K\in\mathcal{C}_h}\int_{\partial K_-(\beta_1)''}[\xi]^2|\mathbf{n}\cdot\beta_1|d\sigma.$$

Here the last sum can be hidden in $|||\xi|||^2$. We estimate the first one as follows

$$\begin{aligned}\sum_{K\in\mathcal{C}_h}\int_{\partial K_-(\beta_1)''}|\eta_-|^2|\mathbf{n}\cdot\beta_1|d\sigma &\le ||\eta||_\infty^2\sum_{K\in\mathcal{C}_h}\Big\{\int_{\partial K_-(\beta_1)''}|\mathbf{n}\cdot\hat{\beta}|^2d\sigma+\int_{\partial K_-(\beta_1)''}d\sigma\Big\}\\ &\le ||\eta||_\infty^2\sum_{K\in\mathcal{C}_h}[Ch^{-1}||\hat{\beta}||_K^2+Ch^{2d}],\end{aligned}\tag{4.9}$$

where we have used the fact that

$$\int_{\partial K}\mathbf{g}^2d\sigma\le Ch^{-1}\int_K\mathbf{g}^2dx,\qquad \mathbf{g}\in[P_k(K)]^4.\tag{4.10}$$

By using the definition of $\mathcal{F}_\varepsilon^\beta$, and the assumption on Fréchet differentiability of F we have

$$\begin{aligned}||\mathcal{F}_\varepsilon^\beta(\mathbf{u})-\mathcal{F}_\varepsilon^\beta(\mathbf{u}_h)||_{\Omega_T}&=\varepsilon||F(\mathbf{u})-F(\mathbf{u}_h)||_{\Omega_T}\le C\varepsilon(1+\alpha)||\mathbf{u}-\mathbf{u}_h||_{\Omega_T}||F'(\mathbf{u})||_\infty\\ &\le C\varepsilon(1+\alpha)||\mathbf{u}-\mathbf{u}_h||_{\Omega_T}||\mathbf{u}||_\infty\le C_{\alpha,\varepsilon}(||\xi||_{\Omega_T}+||\eta||_{\Omega_T}),\end{aligned}\tag{4.11}$$

where we have used the assumption $||\mathbf{u}||_\infty < C$. Further

$$||\hat{\beta}||_{\Omega_T} = ||\mathcal{F}_\varepsilon^\beta(\mathbf{u}_h)||_{\Omega_T} \leq C(||\xi||_{\Omega_T} + ||\eta||_{\Omega_T}) + ||\mathcal{F}_\varepsilon^\beta(\mathbf{u})||_{\Omega_T}, \tag{4.12}$$

yields

$$||\eta_t + \hat{\beta}\eta||_{\Omega_T} \leq ||\eta_t + \mathcal{F}_\varepsilon^\beta(\mathbf{u})\eta||_{\Omega_T} + ||(\mathcal{F}_\varepsilon^\beta(\mathbf{u}_h) - \mathcal{F}_\varepsilon^\beta(\mathbf{u}))\eta||_{\Omega_T}. \tag{4.13}$$

Moreover, the interpolation error η satisfies

$$||\eta||_\infty = ||\mathbf{u} - I_h\mathbf{u}||_\infty \leq Ch^{k+1}||\mathbf{u}||_{k+1,\infty}. \tag{4.14}$$

Thus (4.2)-(4.14) imply that

$$|T| \leq c|||\xi|||^2 + Ch^{2k+2}||\mathbf{u}||_{k+1,\infty}^2 \times [h^{-1}(||\xi||_{\Omega_T}^2 + ||\eta||_{\Omega_T}^2 + ||\mathcal{F}_\varepsilon^\beta(\mathbf{u})||_{\Omega_T}^2) + h^{2d}], \tag{4.15}$$

where we assume that $c \ll 1$ and by the above assumptions:

$$|T| \leq c|||\xi|||^2 + Ch^{2k+1}||\mathbf{u}||_{k+1,\infty} \leq c|||\xi|||^2 + C_1h^{2k+1}.$$

Now, we return to the remaining boundary terms involving both positive boundaries and the jumps in the time direction:

$$T' := -\sum_{n=1}^{N-1} < \eta_-, [\xi] >_n + < \eta_-, \xi_- >_N + \int_{\partial\Omega_+ \times I} \xi\eta(\mathbf{n}\cdot\beta_1)d\sigma ds.$$

Once again, using the Cauchy-Schwarz inequality, we have

$$\begin{aligned} |T'| &\leq \frac{1}{C}\Big[\sum_{n=1}^{N-1} |[\xi]|_n^2 + |\xi|_N^2 \int_{\partial\Omega_+ \times I} |\xi|^2(\mathbf{n}\cdot\beta_1)d\sigma ds\Big] \\ &\quad + C\Big(\sum_{n=1}^{N} |\eta_-|_n^2 + \int_{\partial\Omega_+ \times I} |\eta|^2(\mathbf{n}\cdot\beta_1)d\sigma ds\Big) \\ &\leq c|||\xi|||^2 + C\Big(\sum_{n=1}^{N} |\eta_-|_n^2 + \int_{\partial\Omega_+ \times I} |\eta|^2(\mathbf{n}\cdot\beta_1)d\sigma ds\Big). \end{aligned} \tag{4.16}$$

For second term on the right hand side of (4.6) (the mixed terms), we have

$$\sum_{K\in\mathcal{C}_h} (\eta, \xi_t + \hat{\beta}\xi)_K \leq \sum_{K\in\mathcal{C}_h} [\frac{C}{h}||\eta||_K^2 + \frac{h}{C}||\xi_t + \hat{\beta}\xi||_K^2], \tag{4.17}$$

where, we may hide the term $\frac{h}{C}||\xi_t+\hat{\beta}\xi||_K^2$ in the triple norm $|||\xi|||^2$. Similarly, using previous estimates and the above assumptions, we estimate the third term on the right hand side of (4.6) as

$$(\eta, \xi \ \mathrm{div}\beta_1)_{\Omega_T} \leq ||\eta||_{\Omega_T}||\xi||_{\Omega_T}|| \ \mathrm{div} \ \beta_1||_\infty \leq Ch^{k+1}||\mathbf{u}||_{k+1,\infty}||\xi||_{\Omega_T}. \tag{4.18}$$

For the remaining term in S_1, we use once again, the Cauchy-Schwarz inequality to get

$$\delta(\eta_t+\hat{\beta}\eta, \xi_t+\hat{\beta}\xi)_{\Omega_T} \leq Ch||\eta_t+\hat{\beta}\eta||_{\Omega_T}^2 + Ch||\xi_t+\hat{\beta}\xi||_{\Omega_T}^2 \leq Ch||\eta_t+\hat{\beta}\eta||_{\Omega_T}^2 + C|||\xi|||^2. \tag{4.19}$$

Combining the estimates for all S_1 and S_2 terms, we have

$$|S_1 + S_2| \leq c|||\xi|||^2 + C_1h^{2k+1} + C\Big(\int_{\partial\Omega_+ \times I} \eta^2|\mathbf{n}\cdot\beta_1|d\sigma ds + \sum_{n=1}^{N} |\eta|_n^2\Big). \tag{4.20}$$

The terms S_3, S_4, S_5 and S_6 are estimated by using $||A||_\infty = 1$ and inverse inequality

$$\begin{aligned} |S_3| &= \frac{1}{2}|(A\nabla\eta, \nabla\xi)_{\Omega_T}| \leq \frac{1}{2}||A\nabla\eta||_{\Omega_T}||\nabla\xi||_{\Omega_T} \\ &\leq \frac{1}{4}h^{-1}||A||_\infty^2||\eta||_{\Omega_T}^2 + \frac{1}{4}||\nabla\xi||_{\Omega_T}^2 \leq Ch^{-1}||\eta||_{\Omega_T}^2 + C||\nabla\xi||_{\Omega_T}^2. \end{aligned} \tag{4.21}$$

Similarly

$$|S_4| = \frac{1}{2}\tilde{\varepsilon}|(\nabla\eta, \nabla\xi)_{\Omega_T}| \le C\tilde{\varepsilon}h^{-1}||\eta||^2_{\Omega_T} + C\tilde{\varepsilon}||\nabla\xi||^2_{\Omega_T}. \tag{4.22}$$

Thus for sufficiently small $\tilde{\varepsilon}$

$$|S_3 + S_4| \le Ch^{-1}||\eta||^2_{\Omega_T} + C||\nabla\xi||^2_{\Omega_T}. \tag{4.23}$$

Further, since

$$\begin{aligned}\frac{1}{2}\Big|\int_{\partial\Omega_-\times I}(A\nabla\eta)\xi d\sigma ds\Big| &\le \frac{1}{4}\|A\nabla\eta\|^2_{L_2(\partial\Omega_-\times I)} + \frac{1}{4}\|\xi\|^2_{L_2(\partial\Omega_-\times I)}\\ &\le \frac{1}{4}\|A\nabla\eta\|_{\Omega_T}\|A\triangle\eta\|_{\Omega_T} + \frac{1}{4}\|\xi\|_{\Omega_T}\|\nabla\xi\|_{\Omega_T}\\ &\le C\Big(h^{-1}\|\eta\|^2_{\Omega_T} + \|\xi\|^2_{\Omega_T} + \|\nabla\xi\|^2_{\Omega_T}\Big),\end{aligned} \tag{4.24}$$

hence

$$\begin{aligned}|S_5| &= \frac{1}{2}\Big|\int_{\partial\Omega_-\times I}((A+\tilde{\varepsilon}I)\nabla\eta)\xi d\sigma ds\Big|\\ &\le C\Big(h^{-1}\|\eta\|^2_{\Omega_T} + \|\xi\|^2_{\Omega_T} + \|\nabla\xi\|^2_{\Omega_T}\Big) + C\tilde{\varepsilon}\Big(h^{-1}\|\eta\|^2_{\Omega_T} + \|\xi\|^2_{\Omega_T} + \tilde{\varepsilon}\|\nabla\xi\|^2_{\Omega_T}\Big)\\ &\le C\Big(h^{-1}\|\eta\|^2_{\Omega_T} + \|\xi\|^2_{\Omega_T} + \|\nabla\xi\|^2_{\Omega_T}\Big).\end{aligned} \tag{4.25}$$

Similar argument yields

$$\begin{aligned}|S_6| &= \frac{1}{2}\delta|((A+\tilde{\varepsilon}I)\Delta\eta, \xi_t + \mathcal{F}^\beta_\varepsilon(\mathbf{u}_h)\xi)_{\Omega_T}| \le \frac{1}{2}\delta||(A+\tilde{\varepsilon}I)\Delta\eta||_{\Omega_T}||\xi_t + \mathcal{F}^\beta_\varepsilon(\mathbf{u}_h)\xi||_{\Omega_T}\\ &\le \frac{1}{2}h^{-1}\delta(||A||_\infty + \tilde{\varepsilon})||\eta||_{\Omega_T}||\xi_t + \mathcal{F}^\beta_\varepsilon(\mathbf{u}_h)\xi||_{\Omega_T}\\ &\le Ch^{-1}||\eta||^2_{\Omega_T} + \frac{\delta}{2}||\xi_t + \mathcal{F}^\beta_\varepsilon(\mathbf{u}_h)\xi||^2_{\Omega_T}.\end{aligned} \tag{4.26}$$

As for the term S_7 we have

$$|S_7| \le |\eta_+|^2_0 + |\xi_+|^2_0. \tag{4.27}$$

We combine the estimates for all S_i, $i = 1, \dots, 7$ terms and hide all ξ-terms, including $||\xi_t + \mathcal{F}^\beta_\varepsilon(\mathbf{u}_h)\xi||^2_{\Omega_T}$ in the triple norm $|||\xi|||$. To estimate the term $T_2 - T_3$, we write

$$\begin{aligned}T_2 - T_3 &= \Big(\mathbf{u}_t + \mathcal{F}^\beta_\varepsilon(\mathbf{u})\mathbf{u}, \xi + \delta(\xi_t + \mathcal{F}^\beta_\varepsilon(\mathbf{u}_h)\xi)\Big)_{\Omega_T} - \Big(\mathbf{u}_t + \mathcal{F}^\beta_\varepsilon(\mathbf{u}_h)\mathbf{u}, \xi + \delta(\xi_t + \mathcal{F}^\beta_\varepsilon(\mathbf{u}_h)\xi)\Big)_{\Omega_T}\\ &= \Big(\mathcal{F}^\beta_\varepsilon(\mathbf{u}) - \mathcal{F}^\beta_\varepsilon(\mathbf{u}_h))\mathbf{u}, \xi + \delta(\xi_t + \mathcal{F}^\beta_\varepsilon(\mathbf{u}_h)\xi)\Big)_{\Omega_T} \le ||\mathcal{F}^\beta_\varepsilon(\mathbf{u}) - \mathcal{F}^\beta_\varepsilon(\mathbf{u}_h)||_{\Omega_T}||\mathbf{u}||_\infty||\xi||_{\Omega_T}\\ &+ Ch||\mathcal{F}^\beta_\varepsilon(\mathbf{u}) - \mathcal{F}^\beta_\varepsilon(\mathbf{u}_h)||^2_{\Omega_T}\ ||\mathbf{u}||^2_\infty + \frac{Ch}{8}||\xi_t + \mathcal{F}^\beta_\varepsilon(\mathbf{u}_h)\xi||^2_{\Omega_T}\\ &\le C(||\xi||_{\Omega_T} + ||\eta||_{\Omega_T})||\mathbf{u}||_\infty||\xi||_{\Omega_T} + Ch(||\xi||_{\Omega_T} + ||\eta||_{\Omega_T})^2||\mathbf{u}||^2_\infty + c|||\xi|||^2.\end{aligned}$$

Finally, combining the estimates for the T_1 and $T_2 - T_3$ terms we obtain

$$\begin{aligned}\alpha|||\xi|||^2 \le Ch^{2k+1} + C\Big[\int_{\partial\Omega_+\times I}\eta^2|\beta_1.\mathbf{n}|d\sigma ds + h^{-1}||\eta||^2_{\Omega_T} + \sum_{n=0}^{N}|\eta|^2_n\Big]\\ +C(||\xi||_{\Omega_T} + ||\eta||_{\Omega_T})||\xi||_{\Omega_T} + Ch(||\xi||_{\Omega_T} + ||\eta||_{\Omega_T})^2 + c|||\xi|||^2.\end{aligned}$$

We have the following estimation for third term as

$$(||\xi||_{\Omega_T} + ||\eta||_{\Omega_T})||\xi||_{\Omega_T} = ||\xi||^2_{\Omega_T} + ||\eta||_{\Omega_T}||\xi||_{\Omega_T} \le ||\xi||^2_{\Omega_T} + ch||\xi||^2_{\Omega_T} + ch^{-1}||\eta||^2_{\Omega_T},$$

and the estimate of $||\xi||^2_{\Omega_T}$ by Proposition 3.2. All terms involving ξ, except the term $h\sum_{n=1}^{N}|\xi_-|^2_n$, can be hidden in the triple norm $|||\xi|||^2_{\Omega_T}$. So can conclude that

$$|||\xi|||^2 \le Ch^{2k+1} + C\Big[\int_{\partial\Omega_+\times I}\eta^2|\beta_1\cdot\mathbf{n}|d\sigma ds + h^{-1}||\eta||^2_{\Omega_T} + \sum_{n=0}^{N}|\eta|^2_n + h\sum_{n=1}^{N}|\xi_-|^2_n\Big].$$

Finally, by standard interpolation theory we have (see [8], p.123)

$$\Big[\int_{\partial\Omega_+\times I}\eta^2|\beta_1.\mathbf{n}|d\sigma ds + h^{-1}||\eta||^2_{\Omega_T} + \sum_{n=0}^{N}|\eta|^2_n\Big]^{\frac{1}{2}} \le Ch^{2k+1}||\mathbf{u}||_{k+1,\Omega}.$$

Thus, using the assumptions, together with $||u||_{k+1,\infty} \le \infty$ we have

$$|||\xi|||^2 \le Ch^{2k+1} + C_1 h\sum_{n=1}^{N}|\xi_-|^2_n. \tag{4.28}$$

Now, we use the following Grönwall's estimate. If

$$y(\cdot,t_n) \le C + C_1 h\sum_{n=1}^{N}|y(\cdot,t_n)|^2_n,$$

then

$$y(t_n) \le Ce^{C_1 t} \le Ce^{C_1 T}.$$

Obviously (4.28) implies that

$$|\xi_-|^2_n \le Ch^{2k+1} + C_1 h\sum_{n=1}^{N}|\xi_-|^2_n,$$

therefore, according to the above inequality and Grönwall estimate we have

$$|\xi_-|^2_n \le Ch^{2k+1}e^{C_1 T}. \tag{4.29}$$

By (4.28) and (4.29)

$$|||\xi|||^2 \le Ch^{2k+1} + C_1 h\sum_{n=1}^{N}(Ch^{2k+1}e^{C_1 T}) \le C(T)h^{2k+1}, \tag{4.30}$$

where $C(T) = Ce^{C_1 T}$. Also similar to (4.30), we may conclude that $|||\eta|||^2 \le C(T)h^{2k+1}$. Now using

$$|||e|||^2 \le |||\xi|||^2 + |||\eta|||^2,$$

completes the proof and we have

$$|||e|||^2 \le Ch^{2k+1}.$$

□

Analogously, the multiscale variant of the convergence theorem reads as follows:

Corollary 4.3. *If* $\mathbf{u}_c \in \mathbf{W}_{\widehat{h},c}$, $\mathbf{u}_f \in \mathbf{W}_{h,f}$ *satisfy in (2.5),* $\mathbf{u} \in L^\infty(H^{k+1}(\Omega))$ *for* $k \ge 1$ *is exact solution* (1.5) *and*

$$||\mathbf{u}||_\infty + ||\mathcal{F}^\beta_\varepsilon(\mathbf{u})||_\infty + ||div\ (\mathcal{F}^\beta_\varepsilon(\mathbf{u}) - \varepsilon F(\mathbf{u}))||_\infty + ||\nabla\eta||_\infty \le C,$$

then we have the following error estimate

$$|||\mathbf{u} - \mathbf{u}_{h,f}||| \le C_f h^{k+1/2}, \quad \mathbf{u}_{h,f} \in \mathbf{W}_{h,f}, \tag{4.31}$$

$$|||\mathbf{u} - \mathbf{u}_{\widehat{h},c}||| \le C_c \widehat{h}^{k+1/2}, \quad \mathbf{u}_{\widehat{h},c} \in \mathbf{W}_{\widehat{h},c}. \tag{4.32}$$

5. Construction of MV on DG/SD

In this section to be concise we start with the slabwise construction of the variational multiscale scheme, assuming that $\mathbf{u}_h^n = \mathbf{u}_h(t_n)$. To this end, we separate (2.2) into the following, slabwise formulated, time dependent variational problem:

$$\begin{cases} (\mathbf{u}_{h,t}^n, \mathbf{g})_{\Omega_T} + \delta(\mathbf{u}_{h,t}^n, \mathbf{g}_t)_{\Omega_T} + \delta(\mathbf{u}_{h,t}^n, \mathcal{F}_\varepsilon^\beta(\mathbf{u}_h^n)\mathbf{g})_{\Omega_T} + \overline{B}_{\delta,\varepsilon,\beta}(\mathbf{u}_h^n, \mathbf{g}) = L(\mathbf{g}), \\ (\mathbf{u}_h(t_{n-1}), \mathbf{g})_{\Omega_T} = (\mathbf{u}_h^{n-1}, \mathbf{g})_{\Omega_T}, \qquad \forall\, \mathbf{g} \in \mathbf{W}_h, \end{cases} \tag{5.1}$$

where $\mathbf{u}_h^{n-1}$ at time t_{n-1} is given, and the new bilinear form $\overline{B}_{\delta,\varepsilon,\beta}$ is defined by

$$\begin{aligned} \overline{B}_{\delta,\varepsilon,\beta}(\mathbf{u}_h, g) = & \big(\mathcal{F}_\varepsilon^\beta(\mathbf{u}_h)\mathbf{u_h}, \mathbf{g} + \delta(\mathbf{g}_t + \mathcal{F}_\varepsilon^\beta(\mathbf{u_h})\mathbf{g}\big)_{\Omega_T} + \frac{1}{2}\,(A\nabla \mathbf{u}_h, \nabla \mathbf{g})_{\Omega_T} + \frac{1}{2}\tilde{\varepsilon}(\nabla \mathbf{u}_h, \nabla \mathbf{g})_{\Omega_T} \\ & - \frac{1}{2}\delta((A+\tilde{\varepsilon}I)\Delta \mathbf{u}_h, \mathbf{g}_t + \mathcal{F}_\varepsilon^\beta(\mathbf{u_h})\mathbf{g})_{\Omega_T} + \sum_{K\in\mathcal{C}_h} \int_{\partial K_-(\hat{\beta})} [\mathbf{u}_h]\mathbf{g}_+ |\mathbf{n}_t + \mathbf{n}.\hat{\beta}| d\sigma \\ & - \frac{1}{2}\int_{\partial\Omega_-\times I} ((A+\tilde{\varepsilon}I)\nabla \mathbf{u}_h)\mathbf{g} d\sigma ds + \langle \mathbf{u}_+, \mathbf{g}_+\rangle_{0,\Omega}. \end{aligned}$$

Below we shall suppress the superscript n. Then, without loss of generality, the variational multiscale method for this problem is formulated as follows [25]:

Find $\mathbf{u}_h = \mathbf{u}_f + \mathbf{u}_c$, where $\mathbf{u}_c \in \mathbf{W}_{\widehat{h},c}$ and $\mathbf{u}_f \in \mathbf{W}_{h,f}$ such that $\forall\, \mathbf{g}_c \in \mathbf{W}_{\widehat{h},c}$, $\forall\, \mathbf{g}_f \in \mathbf{W}_{h,f}$

$$\begin{cases} \big(\mathbf{u}_{f,t} + \mathbf{u}_{c,t}, \mathbf{g}_f + \mathbf{g}_c\big)_{\Omega_T} + \delta\big(\mathbf{u}_{f,t} + \mathbf{u}_{c,t}, \mathbf{g}_{f,t} + \mathbf{g}_{c,t}\big)_{\Omega_T} \\ +\delta\big(\mathbf{u}_{f,t} + \mathbf{u}_{c,t}, \mathcal{F}_\varepsilon^\beta(\mathbf{u}_f + \mathbf{u}_c)(\mathbf{g}_f + \mathbf{g}_c)\big)_{\Omega_T} + \overline{B}_{\delta,\varepsilon,\beta}(\mathbf{u}_f + \mathbf{u}_c, \mathbf{g}_f + \mathbf{g}_c) = L(\mathbf{g}_f + \mathbf{g}_c) \\ \big(\mathbf{u}_c(t_{n-1}) + \mathbf{u}_f(t_{n-1}), \mathbf{g}_c + \mathbf{g}_f\big)_{\Omega_T} = (\mathbf{u}_h^{n-1}, \mathbf{g}_c + \mathbf{g}_f)_{\Omega_T}. \end{cases} \tag{5.2}$$

We may split this equation into two parts and use an L_2-orthogonal split of the coarse and fine scales which cancels cross terms as $(\mathbf{u}_{f,t}, \mathbf{g}_c)$ and $(\mathbf{u}_{c,t}, \mathbf{g}_f)$ as $\varepsilon \to 0$, i.e. we have

$$\begin{cases} \big(\mathbf{u}_{c,t}, \mathbf{g}_c\big)_{\Omega_T} + \delta\big(\mathbf{u}_{c,t}, \mathbf{g}_{c,t}\big)_{\Omega_T} \\ \quad +\delta\big(\mathbf{u}_{c,t}, \mathcal{F}_{\varepsilon\to 0}^\beta(\mathbf{u}_c)\mathbf{g}_c\big)_{\Omega_T} + \overline{B}_{\delta,\varepsilon\to 0,\beta}(\mathbf{u}_c, \mathbf{g}_c) = L(\mathbf{g}_c), \quad \forall\, \mathbf{g}_c \in \mathbf{W}_{\widehat{h},c}, \\ \big(\mathbf{u}_c(t_{n-1}), \mathbf{g}_c\big)_{\Omega_T} = (\mathbf{u}_h^{n-1}, \mathbf{g}_c)_{\Omega_T}, \qquad \forall\, \mathbf{g}_c \in \mathbf{W}_{\widehat{h},c}, \end{cases} \tag{5.3}$$

and

$$\begin{cases} \big(\mathbf{u}_{f,t}, \mathbf{g}_f\big)_{\Omega_T} + \delta\big(\mathbf{u}_{f,t}, \mathbf{g}_{f,t}\big)_{\Omega_T} + \delta\big(\mathbf{u}_{f,t}, \mathcal{F}_{\varepsilon\to 0}^\beta(\mathbf{u}_f)\mathbf{g}_f\big)_{\Omega_T} \\ \quad +\overline{B}_{\delta,\varepsilon\to 0,\beta}(\mathbf{u}_f, \mathbf{g}_f) = L(\mathbf{g}_f) - \overline{B}_{\delta,\varepsilon\to 0,\beta}(\mathbf{u}_c, \mathbf{g}_f), \qquad \forall\, \mathbf{g}_f \in \mathbf{W}_{h,f}, \\ \big(\mathbf{u}_f(t_{n-1}), \mathbf{g}_f\big)_{\Omega_T} = (\mathbf{u}_h^{n-1}, \mathbf{g}_f)_{\Omega_T}, \qquad \forall\, \mathbf{g}_f \in \mathbf{W}_{h,f}. \end{cases} \tag{5.4}$$

We use the partition of unities $\{\phi_i\}_{i\in\mathcal{N}}$ and $\{\chi_i\}_{i\in\mathcal{N}}$, where $\chi_i = \frac{1}{d+1}$ supp (ϕ_i), and split (5.4) into the following three equations

$$\begin{cases} \big(\mathbf{u}_{f,l,i,t}, \mathbf{g}_f\big)_{\Omega_T} + \delta\big(\mathbf{u}_{f,l,i,t}, \mathbf{g}_{f,t}\big)_{\Omega_T} + \delta\big(\mathbf{u}_{f,t}, \mathcal{F}_{\varepsilon\to 0}^\beta(\mathbf{u}_{f,l,i})\mathbf{g}_f\big)_{\Omega_T} \\ \quad +\overline{B}_{\delta,\varepsilon\to 0,\beta}(\mathbf{u}_{f,l,i}, \mathbf{g}_f) = L(\chi_i \mathbf{g}_f), \qquad \forall\, \mathbf{g}_f \in \mathbf{W}_{h,f}, \\ \big(\mathbf{u}_{f,l,i}(t_{n-1}), \mathbf{g}_f\big)_{\Omega_T} = 0, \qquad \forall\, \mathbf{g}_f \in \mathbf{W}_{h,f}, \end{cases} \tag{5.5}$$

$$\begin{cases} \big(\mathbf{u}_{f,0,i,t}, \mathbf{g}_f\big)_{\Omega_T} + \delta\big(\mathbf{u}_{f,0,i,t}, \mathbf{g}_{f,t}\big)_{\Omega_T} + \delta\big(\mathbf{u}_{f,0,i,t}, \mathcal{F}_{\varepsilon\to 0}^\beta(\mathbf{u}_{f,l,i})\mathbf{g}_f\big)_{\Omega_T} \\ \quad +\overline{B}_{(\delta,\varepsilon\to 0,\beta)}(\mathbf{u}_{f,l,i}, \mathbf{g}_f) = 0, \qquad \forall\, \mathbf{g}_f \in \mathbf{W}_{h,f}, \\ \big(\mathbf{u}_{f,0,i}(t_{n-1}), \mathbf{g}_f\big)_{\Omega_T} = (\chi_i \mathbf{u}^{n-1}, \mathbf{g}_f)_{\Omega_T}, \qquad \forall\, \mathbf{g}_f \in \mathbf{W}_{h,f}, \end{cases} \tag{5.6}$$

and

$$\begin{cases} \big(\mathcal{T}_t\phi_i, \mathbf{g}_f\big)_{\Omega_T} + \delta\big(\mathcal{T}_t\phi_i, \mathbf{g}_{f,t}\big)_{\Omega_T} + \delta\big(\mathcal{T}_t\phi_i, \mathcal{F}^{\beta}_{\varepsilon\to 0}(\phi_i)\mathbf{g}_f\big)_{\Omega_T} \\ \quad +\overline{B}_{\delta,\varepsilon\to 0,\beta}(\mathcal{T}\phi_i, \mathbf{g}_f) = -\overline{B}_{\delta,\varepsilon\to 0,\beta}(\phi_i, \mathbf{g}_f), & \forall\ \mathbf{g}_f \in \mathbf{W}_{h,f}, \\ \big(\mathcal{T}\phi(t_{n-1}), \mathbf{g}_f\big)_{\Omega_T} = 0, & \forall\ \mathbf{g}_f \in \mathbf{W}_{h,f}, \end{cases} \tag{5.7}$$

where $\mathcal{T}_t = \mathcal{T}\frac{\partial}{\partial t}$ (we recall that $\mathcal{T}$ is defined in (2.5)). Now, if $\mathbf{u}_c = \sum_{i\in\mathcal{N}} \alpha_i\phi_i$ then $\mathbf{u}_f = \sum_{i\in\mathcal{N}}(\mathbf{u}_{f,l,i} + \mathbf{u}_{f,0,i}\alpha_i\mathcal{T}\phi_i)$. Hence, for all $\mathbf{g}_f \in \mathbf{W}_{\widehat{h},f}$, we may rewrite (5.3) as the following for the coarse scale equation:

$$\begin{cases} \big(\mathbf{u}_{c,t}, \mathbf{g}_c\big)_{\Omega_T} + \delta\big(\mathbf{u}_{c,t}, \mathbf{g}_{c,t}\big)_{\Omega_T} + \delta\big(\mathbf{u}_{c,t}, \mathcal{F}^{\beta}_{\varepsilon\to 0}(\mathbf{u}_c)\mathbf{g}_c\big)_{\Omega_T} \\ \quad +\overline{B}_{\delta,\varepsilon\to 0,\beta}(\mathbf{u}_c + \mathcal{T}\mathbf{u}_c, \mathbf{g}_c) = L(\mathbf{g}_c) - \delta\big(\mathbf{u}_{f,l,t} + \mathbf{u}_{f,0,t}, \mathbf{g}_{f,t}\big)_{\Omega_T} \\ \quad +\delta\big(\mathbf{u}_{f,l,t} + \mathbf{u}_{f,0,t}, \mathcal{F}^{\beta}_{\varepsilon\to 0}(\mathbf{u}_f + \mathbf{u}_{f,0,t})\mathbf{g}_c\big)_{\Omega_T} + \overline{B}_{\delta,\varepsilon\to 0,\beta}(\mathbf{u}_{f,l} + \mathbf{u}_{f,0,t}, \mathbf{g}_f), \\ \big(\mathbf{u}_c(t_{n-1}), \mathbf{g}_c\big)_{\Omega_T} = (\mathbf{u}_h^{n-1}, \mathbf{g}_c)_{\Omega_T}, \end{cases} \tag{5.8}$$

where $\mathcal{T}\mathbf{u}_c = \sum_{i\in\mathcal{N}} \alpha_i\mathcal{T}\phi_i$, $\mathbf{u}_{f,l} = \sum_{i\in\mathcal{N}} \mathbf{u}_{f,l,i}$ and $\mathbf{u}_{f,0} = \sum_{i\in\mathcal{N}} \mathbf{u}_{f,0,i}$. We shall use backward Euler in time to solve the above equation.

Remark 5.1. *We recall that if we put $\mathcal{F}^{\beta,L}_{\varepsilon}(\mathbf{u}_h)$ as an approximation of linear form defined with $\mathcal{F}^{\beta}_{\varepsilon}(\mathbf{u}_h)$, then we can continue considering the above system without removing the ε term, corresponds to a non-vanishig, but per small value for ε ($\varepsilon \nrightarrow 0$). We denote this phenomenon replacing $\overline{B}_{\delta,\varepsilon\to 0,\beta}$ by $\overline{B}_{\delta,\varepsilon\rightsquigarrow 0,\beta}$.*

We emphasize that, in principle, the analysis of the variational multiscale method is similar to that of the original SD based DG approach studied along the previous sections. For a more detailed study we refer the reader to the monograph (PhD thesis) by R. Söderlind [29].

In the next section, we shall describe some aspects of implementing the above systems.

6. Numerical considerations

We let $\{\phi_i\}_{i\in\mathcal{N}}$ be the standard piecewise linear discontinuous basis functions such that $\text{span}\{\phi_i\}_{i\in\mathcal{N}} = \mathbf{W}_{\widehat{h},c}$. This means that the support of ϕ_i is exactly one coarse element in the mesh K_n^c. Around the support of each basis function ϕ_i, we construct a patch w_i of coarse elements for solving localized fine scale problems. In the definition of the patches we will also be referring to the standard piecewise linear continuous functions on the mesh K_n^c, which we denote by $\{\theta_i\}$. The patches will include the support of the associated basis function. In the following , we recall the standard definition of patch:

Definition 6.1. *We define w_i^m as a symmetric m layer patch for $m = 2, 3, \ldots$, around the support of basis function ϕ_i,*

$$w_i^m = \bigcup_{\{j:supp(\theta_j)\bigcap w_i^{m-1}\neq\emptyset\}} supp\ (\theta_j), \quad m = 2, 3, \ldots \tag{6.1}$$

where θ_j is a coarse scale piecewise linear continues basis function. On the other hand, we can introduce, w_i^1 as a symmetric 1-layer patch if $w_i^1 =$ supp ϕ_i, where ϕ_i is a coarse basis function, with support on one coarse element. Moreover, we can show the directed m layer version of patch by $\overrightarrow{w}_i^m$ such that $\overrightarrow{w}_i^1 = w_i^1$, $\overrightarrow{w}_i^m \subset w_i^m$ possibly equal, and any $K_n \in \overrightarrow{w}_i^m$ can be reached from $\overrightarrow{w}_i^{m-1}$ by passing from element K_n^- to element K_n^+ (starting from any element in $\overrightarrow{w}_i^{m-1}$).

In the next subsection, we use the discrete function spaces on patches to formulate an approximate method.

6.1. **Matrix equation by Spatial discretization.** In the following, we will observe three systems of fine scale equations and a coarse scale equation and then coarse scale equation in detail. Let $\mathbf{U}_{f,l,i}$, $\mathbf{U}_{f,0,i}$, $\mathcal{T}\phi_i$ belong to $\mathbf{W}_{h,f}(w_i)$ ($\mathbf{W}_{h,c}(w_i)$) i.e. the piecewise linear polynomials fine (coarse) scale spaces on the patch w_i in time such that they solve the following systems:

$$\begin{cases} \big(\mathbf{U}_{f,l,i,t},\mathbf{g}_f\big)_{\Omega_T}+\delta\big(\mathbf{U}_{f,l,i,t},\mathbf{g}_{f,t}\big)_{\Omega_T}+\delta\big(\mathbf{U}_{f,t},\mathcal{F}^{\beta}_{\varepsilon\to 0}(\mathbf{U}_{f,l,i})\mathbf{g}_f\big)_{\Omega_T} \\ \quad +\overline{B}_{\delta,\varepsilon\to 0,\beta}(\mathbf{U}_{f,l,i},\mathbf{g}_f)=L(\chi_i\mathbf{g}_f), \qquad \forall\ \mathbf{g}_f\in\mathbf{W}_{h,f}(w_i), \\ \big(\mathbf{U}_{f,l,i}(t_{n-1}),\mathbf{g}_f\big)_{\Omega_T}=0, \end{cases} \tag{6.2}$$

$$\begin{cases} \big(\mathbf{U}_{f,0,i,t},\mathbf{g}_f\big)_{\Omega_T}+\delta\big(\mathbf{U}_{f,0,i,t},\mathbf{g}_{f,t}\big)_{\Omega_T}+\delta\big(\mathbf{U}_{f,0,i,t},\mathcal{F}^{\beta}_{\varepsilon\to 0}(\mathbf{U}_{f,l,i})\mathbf{g}_f\big)_{\Omega_T} \\ \quad +\overline{B}_{(\delta,\varepsilon\to 0,\beta)}(\mathbf{U}_{f,l,i},\mathbf{g}_f)=0, \qquad \forall\ \mathbf{g}_f\in\mathbf{W}_{h,f}(w_i), \\ \big(\mathbf{U}_{f,0,i}(t_{n-1}),\mathbf{g}_f\big)_{\Omega_T}=(\chi_i\mathbf{U}^{n-1},\mathbf{g}_f)_{\Omega_T}, \end{cases} \tag{6.3}$$

and

$$\begin{cases} \big(\mathcal{T}_t\phi_i,\mathbf{g}_f\big)_{\Omega_T}+\delta\big(\mathcal{T}_t\phi_i,\mathbf{g}_{f,t}\big)_{\Omega_T}+\delta\big(\mathcal{T}_t\phi_i,\mathcal{F}^{\beta}_{\varepsilon\to 0}(\phi_i)\mathbf{g}_f\big)_{\Omega_T} \\ \quad +\overline{B}_{\delta,\varepsilon\to 0,\beta}(\mathcal{T}\phi_i,\mathbf{g}_f)=-\overline{B}_{\delta,\varepsilon\to 0,\beta}(\phi_i,\mathbf{g}_f), \qquad \forall\ \mathbf{g}_f\in\mathbf{W}_{h,f}(w_i), \\ \big(\mathcal{T}\phi(t_{n-1}),\mathbf{g}_f\big)_{\Omega_T}=0. \end{cases} \tag{6.4}$$

for all $i\in\mathcal{N}$. In the following, we observe the results by using the backward Euler method (or the 4th order Runge-Kutta method). Thus if we have $\mathbf{U}_c=\sum_{i\in\mathcal{N}}\alpha_i\phi_i$, then $\mathbf{U}_f=\sum_{i\in\mathcal{N}}(\mathbf{U}_{f,l,i})+\mathbf{U}_{f,0,i}+\alpha_i\mathcal{T}\phi_i$, and we can obtain the following coarse scale equation (4.9):

$$\begin{cases} \big(\mathbf{U}_{c,t},\mathbf{g}_c\big)_{\Omega_T}+\delta\big(\mathbf{U}_{c,t},\mathbf{g}_{c,t}\big)_{\Omega_T}+\delta\big(\mathbf{U}_{c,t},\mathcal{F}^{\beta}_{\varepsilon\to 0}(\mathbf{U}_c)\mathbf{g}_c\big)_{\Omega_T}+\overline{B}_{\delta,\varepsilon\to 0,\beta}(\mathbf{U}_c+\mathcal{T}\mathbf{U}_c,\mathbf{g}_c) \\ \quad =L(\mathbf{g}_c)-\delta\big(\mathbf{U}_{f,l,t}+\mathbf{U}_{f,0,t},\mathbf{g}_{f,t}\big)_{\Omega_T}+\delta\big(\mathbf{U}_{f,l,t}+\mathbf{U}_{f,0,t},\mathcal{F}^{\beta}_{\varepsilon\to 0}(\mathbf{U}_f+\mathbf{u}_{f,0,t})\mathbf{g}_c\big)_{\Omega_T} \\ \quad +\overline{B}_{\delta,\varepsilon\to 0,\beta}(\mathbf{U}_{f,l}+\mathbf{U}_{f,0,t},\mathbf{g}_f), \qquad \forall\ \mathbf{g}_c\in\mathbf{W}_{\widehat{h},c}(w_i), \\ \big(\mathbf{U}_c(t_{n-1}),\mathbf{g}_c\big)_{\Omega_T}=(\mathbf{U}_h^{n-1},\mathbf{g}_c)_{\Omega_T}, \end{cases} \tag{6.5}$$

where $\mathbf{U}_{f,l}=\sum_{i\in\mathcal{N}}\alpha_i\phi_i$, then $\mathbf{U}_{f,0}=\sum_{i\in\mathcal{N}}\mathbf{U}_{f,0,i}$.

Remark 6.2. *We denote by $\mathcal{E}^i_I$ the set of all interior edges and by $\mathcal{E}^i_\Gamma$ the set of all boundary edges in the mesh $\mathbf{W}_{\widehat{h},c}(w_i)$. Further, we assume that Γ^i_- is the inflow part of the boundary ∂w_i.*

After using the backward Euler method for equations of (6.2), we get the following matrix equation:

$$\begin{cases} \Big(A^i_1+\delta(A^i_2+A^i_3)+\Delta t_i(B^i_1+B^i_2+B^i_3+B^i_4)\Big)\overrightarrow{\mathbf{U}}_{f,l,i}(t_n)=\Delta t_i b^i, \\ H\overrightarrow{\mathbf{U}}_{f,l,i}=0, \qquad i=1,...,N, \end{cases} \tag{6.6}$$

where we have used the notation,

$$
\begin{aligned}
&\triangle t_i = t_i - t_{i-1},\\
&A_1^i = (a_{1,j,k}^i = (\phi_k, \phi_j)_{w_i}),\\
&A_2^i = (a_{2,j,k}^i = (\phi_k, 1)_{w_i}),\\
&A_3^i = (a_{3,j,k}^i = (\phi_k, \mathcal{F}_{\varepsilon\to 0}^{\beta}(\phi_j))_{w_i}),\\
&B_1^i = \Big(b_{1,j,k}^i = (\phi_k \mathcal{F}_{\varepsilon\to 0}^{\beta}(\phi_j), \phi_j + \delta(1+\phi_j \mathcal{F}_{\varepsilon\to 0}^{\beta}(\phi_k)))_{\mathcal{E}_\Gamma^i \setminus \Gamma_-^i}\Big),\\
&B_2^i = \Big(b_{2,j,k}^i = \Big(\phi_k \mathcal{F}_{\varepsilon\to 0}^{\beta}(\mathcal{T}\rho_k)\rho_k, \phi_j + \delta(1+\phi_j \mathcal{F}_{\varepsilon\to 0}^{\beta}(\mathcal{T}\phi_k))\Big)_{\mathcal{E}_\Gamma^i \setminus \Gamma_-^i}\\
&\qquad\quad + \int_{\partial w_i^- \times I} [\mathcal{T}\rho_k]\rho_{k,+}|\mathbf{n}_t + \mathbf{n}.\beta|ds\Big),\\
&B_3^i = \Big(b_{3,j,k}^i = \int_{\partial w_i^- \times I}[\phi_k]\phi_{j,+}|\mathbf{n}_t + \mathbf{n}.\hat{\beta}|ds + \langle \phi_{k,+}, \phi_j\rangle_{0,w_i}\Big),\\
&B_4^i = \Big(b_{4,j,k}^i = (\tfrac{-1}{2}\int_{\partial w_i^- \times I} A\nabla\phi_k\phi_j d\sigma ds)\Big),\\
&\text{and}\\
&b_i = \Big(b_{4,j,k}^i = \langle \phi_0, \chi_i\phi_j\rangle_{0,w_i}\Big).
\end{aligned}
$$

Also, $\overrightarrow{\mathbf{U}}_{f,l,i}$ is the vector of nodal values of $\mathbf{U}_{f,l,i}$ and H is the matrix of the boundary condition. Again, we discretize in time and use equations (6.2)-(6.4) to get the fine scale contribution. Therefore, we can obtain the backward Euler method such that given $\mathbf{U}_c(t_{n-1})$, we can compute $\mathbf{U}_c(t_n)$ by solving the following matrix equation

$$
\begin{aligned}
&\Big(A_1 + \delta(A_2 + A_3) + \Delta t_i(B_1 + B_2 + B_3 + B_4)\Big)\overrightarrow{\mathbf{U}}_c(t_n)\\
&\qquad\qquad = (A_1 + \delta(A_2 + A_3))\overrightarrow{\mathbf{U}}_c(t_{n-1}) + \triangle t_i b,
\end{aligned}
\tag{6.7}
$$

where $\overrightarrow{\mathbf{U}_c}$ is the vector of nodal values of $\mathbf{U}_c$ and,

$$
\begin{aligned}
&A_1 = \Big(a_{1,j,k} = (\phi_k, \phi_j)_{\Omega_T}\Big),\\
&A_2 = \Big(a_{2,j,k} = (\phi_k, 1)_{\Omega_T}\Big),\\
&A_3 = \Big(a_{3,j,k} = (\phi_k, \mathcal{F}_{\varepsilon\to 0}^{\beta}(\phi_j))_{\Omega_T}\Big),\\
&B_1 = \Big(b_{1,j,k} = (\phi_k \mathcal{F}_{\varepsilon\to 0}^{\beta}(\phi_j), \phi_j + \delta(1+\phi_j \mathcal{F}_{\varepsilon\to 0}^{\beta}(\phi_k)))_{\mathcal{E}_\Gamma \setminus \Gamma_-}\Big),\\
&B_2 = \Big(b_{2,j,k} = \Big(\phi_k \mathcal{F}_{\varepsilon\to 0}^{\beta}(\mathcal{T}\rho_k)\rho_k, \phi_j + \delta(1+\phi_j \mathcal{F}_{\varepsilon\to 0}^{\beta}(\mathcal{T}\phi_k))\Big)_{\mathcal{E}_\Gamma \setminus \Gamma_-}\\
&\qquad\quad + \int_{\partial\Omega^- \times I} [\mathcal{T}\rho_k]\rho_{k,+}|\mathbf{n}_t + \mathbf{n}.\beta|ds\Big),\\
&B_3 = \Big(b_{3,j,k} = \int_{\partial\Omega^- \times I}[\phi_k]\phi_{j,+}|\mathbf{n}_t + \mathbf{n}.\hat{\beta}|ds + \langle \phi_{k,+}, \phi_j\rangle_{0,\Omega}\Big),\\
&B_4 = \Big(b_{4,j,k} = \tfrac{-1}{2}\int_{\partial\Omega^- \times I} A\nabla\phi_k\phi_j d\sigma ds\Big),\\
&b = \Big(b_{4,j,k} = \langle \phi_0, \chi_i\phi_j\rangle_{0,\Omega}\Big).
\end{aligned}
$$

Remark 6.3. *We note that if we have $w_i = \Omega$ and the same resolution is used in all patches the reference solution on the fine mesh in recovered. In the numerical results, we can use this concept when studying how the truncated domains w_i affects the approximate solution $U_c + U_f$.*

Below we describe a numerical algorithm for the entire method.

Alagorithm 6.4. *DGVMS for solving (1.1)*

Step 0- *Input* c_0 ,α, β, ε, $1 \leq \gamma \leq 2$, Ω, $\mathbf{u}_0$, N, $\widehat{h}$, m *and* j.
Step 1- *If* $\varepsilon \neq 0$ *then we replace* $\overline{B}_{\delta,\varepsilon\to 0,\beta}$ *by* $\overline{B}_{\delta,\varepsilon\rightsquigarrow 0,\beta}$
Step 2- *Discretization of* $[0,T]$, $0 = t_0 < t_1 < ... < t_N = T$ *with* $k_n = t_{n+1} - t_n$
Step 3- *Compute* $h := \frac{\widehat{h}}{2^j}$, $\delta := c_0 h^\gamma$, $\{\phi_i\}_{i\in\mathcal{N}}$ *and* $\{\rho_i\}_{i\in\mathcal{N}}$.
Step 4-*Assemble the local fine scale matrices* $A^i_1, A^i_2, A^i_3, B^i_1, B^i_2, B^i_3, B^i_4$ *and vector* b^i *on each patch.*
Step 5-*Produce of the global matrices* $A_1, A_2, A_3, B_1, B_2, B_3, B_4$ *and vector* b *on each patch.*
Step 6- *Compute the time independent fine scale solution* $\overrightarrow{\mathbf{U}}_{f,l,i} \quad i \in \mathcal{N}$.
Step 7-*Obtain* $\overrightarrow{\mathbf{U}}_c$ *by solving* (6.7).
Step 8- *Construct* $\overrightarrow{\mathbf{U}} = \overrightarrow{\mathbf{U}}_c + \overrightarrow{\mathbf{U}}_f$
Step 9-*Change* j, N *and go to step 2, otherwise stop.*

6.2. **Experimental results and some realistic applications.** The Schrödinger equation plays the role of Newton's laws and conservation of energy in classical mechanics - i.e., it predicts the future behavior of a dynamical system [6, 15]. It is a wave equation in terms of the wave function which predicts analytically and precisely the probability of events or outcome. The detailed outcome is not strictly determined, but given a large number of events, the Schrödinger equation will predict the distribution of results. In this section, we present some numerical results for the proposed Algorithm 6.4. with $\delta = h$, $\varepsilon = 0.01$ and $\beta = (1,1)^T$. The accuracy of this method with the reference solutions $\psi_{R,i}(x,t) \quad i = 1,2$ is tested in the L_2-norm. We recall that $\psi_{R,i}(x,t)$ is a reference solution computed on a fully mesh.

Example 6.5. Couple ultrafast laser dynamics.
The couple time-dependent Schrödinger equation arises in ultrafast laser dynamics. In this example, we will consider two different cases for the initial conditions:
*(***Case 1***)*

$$\psi_1(x,0) = \psi_2(x,0) = \sqrt{\frac{2\alpha}{1+\pi}}\sec h\Big[(x_1 - x_{1,L})(x_2 - x_{2,L})(x_1 - x_{1,R})(x_2 - x_{2,R})\Big], \quad x \in \Omega$$

*and (***Case 2***)*

$$\psi_1(x,0) = \psi_2(x,0) = \sqrt{\frac{2\alpha}{1+\pi}}\cos h\Big[(x_1 - x_{1,L})(x_2 - x_{2,L})(x_1 - x_{1,R})(x_2 - x_{2,R})\Big], \quad x \in \Omega$$

where

$$\Omega = \Big\{x = (x_1,x_2)^T | \quad x_{1,L} \leq x_1 \leq x_{1,R}, \quad x_{2,L} \leq x_2 \leq x_{2,R}\Big\},$$

$x_{1,L} = x_{2,L} = -1$, $x_{1,R} = x_{2,R} = 1$, $\widehat{h}_c = \alpha = 0.1$ *and* $N = 100$. *Also, the accuracy is measured in the* L_2*-norm error defined by:*

$$\begin{aligned}\|Re(e_i(h))\| &= \|Re(\psi_{R,i}) - Re(\psi_{h,i})\| \\ &= \sqrt{h\sum_{j=0}^{m}|Re(\psi_{R,i}(x_{1,j},x_{2,j},t)) - Re(\psi_{h,i}(x_{1,j},x_{2,j},t))|^2}\end{aligned}$$

Table 1. Error of the method for the imaginary part of $\psi_1(x,t)$ at the given times for case 1.

$(x_{1,j}, x_{2,j})$	$t = 0.00$	$t = 0.01$	$t = 0.05$	$t = 0.10$	$t = 0.15$
$(-1,-1)$	0.841e-6	0.453e-6	0.433e-8	0.731e-5	0.521e-4
$(-1,1)$	0.231e-7	0.423e-4	0.323e-5	0.363e-9	0.242e-5
$(1,-1)$	0.206e-6	0.711e-5	0.654e-8	0.383e-6	0.542e-7
$(1,1)$	0.234e-4	0.153e-8	0.213e-8	0.401e-5	0.411e-6
$(0.5,0.5)$	0.834e-4	0.557e-5	0.833e-4	0.751e-5	0.931e-4
$\|Im(e_1)\|$	0.234e-5				
order	3.7543				

and

$$\begin{aligned}\|Im(e_i(h))\| &= \|Im(\psi_{R,i}) - Im(\psi_{h,i})\| \\ &= \sqrt{h\sum_{j=0}^{m} |Im(\psi_{R,i}(x_{1,j},x_{2,j},t)) - Im(\psi_{h,i}(x_{1,j},x_{2,j},t))|^2},\end{aligned}$$

where $i = 1,2$. *The order of error is calculated using the following formula:*

$$\text{Order of the real part of the error} \approx \frac{1}{\log 2}\log\frac{\|Re(e_i(h)\|}{h},$$

and

$$\text{Order of the imaginary part of the error} \approx \frac{1}{\log 2}\log\frac{\|Im(e_i(h))\|}{h},$$

where $i = 1,2$. *We carry out the above algorithm, by an AMD Opteron computer where 15 Gigabytes RAM memory with 2.2 GHz CPU has been used for these experiments. The error of this approximation method with reference solution for the real and imaginary parts are given in Tables* 1-6. *The evolution of the errors are given in Tables 7-8.*

It can be found that our method only induces a very small numerical reflection. Comparing with existing numerical results, this scheme was performing better than the finite differences and standard Galerkin finite element methods. The agreement, and the small L_2*-norm of the error between theoretical analysis and numerical results shows that the method is efficient. The order of error in this method is close to 4, while the error order the finite difference methods is 2 and that of the standard Galerkin finite element method is 3. In Fig. 1, we report decay of error of approximate solutions in* L_2 *norm for layers. The convergence is measured in the relative error*

$$E(Re(\psi_{R,i})) = \frac{\|Re(\psi_{m,i}) - Re(\psi_{R,i})\|_{L_2}}{\|Re(\psi_{R,i})\|_{L_2}}$$

and

$$E(Im(\psi_{R,i})) = \frac{\|Im(\psi_{m,i}) - Im(\psi_{R,i})\|_{L_2}}{\|Im(\psi_{R,i})\|_{L_2}}$$

(see Fig. 2*) where* $\psi_{m,i}, \quad m = 2,3,4,... \quad i = 1,2,$ *is the solution using the computational domain for each patch when solving the fine-scale problems. Further, we give the multiscale solution by the coarse and the fine scale solutions in Figs. 3 and 4. Finally, we also investigated the above results by behavior of damping with* $\varepsilon \in [0,\ 0.01]$ *but the variations of results were very small and therefore are not presented in here.*

TABLE 2. Error of this method for the real part of $\psi_1(x,t)$ at the given times for case 1.

$(x_{1,j}, x_{2,j})$	$t = 0.00$	$t = 0.01$	$t = 0.05$	$t = 0.10$	$t = 0.15$
$(-1,-1)$	0.621e-7	0.263e-6	0.439e-6	0.432e-6	0.768e-6
$(-1,1)$	0.235e-7	0.421e-7	0.523e-5	0.462e-6	0.542e-5
$(1,-1)$	0.716e-7	0.741e-8	0.253e-4	0.223e-5	0.765e-6
$(1,1)$	0.274e-7	0.103e-8	0.973e-5	0.761e-9	0.613e-8
$(0.5,0.5)$	0.532e-5	0.673e-4	0.3413e-3	0.431e-6	0.761e-5
$\|Re(e_1)\|$	0.334e-5				
order	4.1213				

TABLE 3. Error of this method for the imaginary part of $\psi_2(x,t)$ at the given times for case 1.

$(x_{1,j}, x_{2,j})$	$t = 0.00$	$t = 0.01$	$t = 0.05$	$t = 0.10$	$t = 0.15$
$(-1,-1)$	0.381e-6	0.873e-7	0.253e-7	0.141e-8	0.521e-7
$(-1,1)$	0.431e-7	0.223e-5	0.323e-6	0.313e-5	0.242e-5
$(1,-1)$	0.206e-4	0.451e-8	0.654e-8	0.223e-6	0.542e-6
$(1,1)$	0.234e-7	0.443e-6	0.213e-8	0.651e-8	0.871e-7
$(0.5,0.5)$	0.234e-7	0.153e-8	0.213e-8	0.201e-8	0.411e-7
$\|Im(e_2)\|$	0.354e-6				
order	4.0221				

TABLE 4. Error of this method for the real part of $\psi_2(x,t)$ at the given times for case 1.

$(x_{1,j}, x_{2,j})$	$t = 0.00$	$t = 0.01$	$t = 0.05$	$t = 0.10$	$t = 0.15$
$(-1,-1)$	0.881e-7	0.6873e-7	0.433e-6	0.631e-6	0.521e-4
$(-1,1)$	0.431e-7	0.409e-7	0.323e-8	0.363e-8	0.242e-5
$(1,-1)$	0.206e-7	0.113e-8	0.654e-7	0.863e-7	0.565e-6
$(1,1)$	0.434e-7	0.190e-9	0.133e-5	0.291e-6	0.431e-8
$(0.5,0.5)$	0.454e-6	0.433e-7	0.363e-6	0.401e-7	0.411e-6
$\|Re(e_2)\|$	0.2834e-7				
order	3.8543				

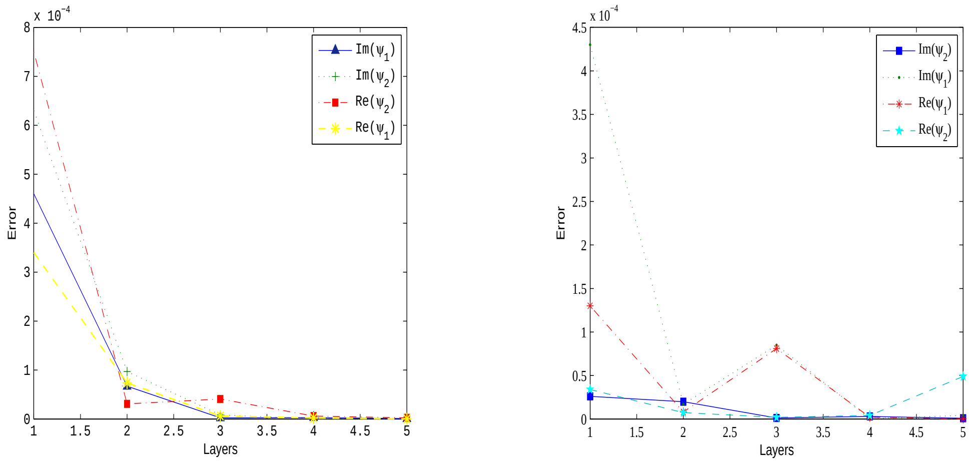


Figure 1. Decay of the approximation solution on boundary for different layers for case 1 at $t = 0.06$(in the left) and case 2 at $t = 1$ (in the right).

TABLE 5. Error of this method for the imaginary part of $\psi_1(x,t)$ at the given times for case 2.

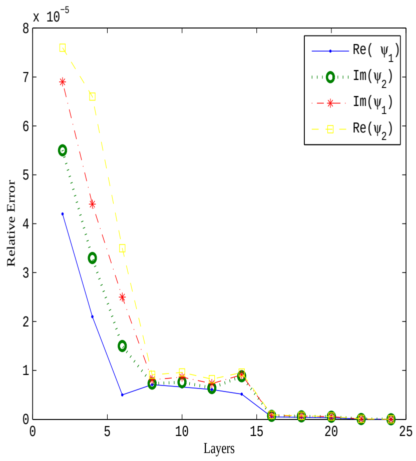
$(x_{1,j}, x_{2,j})$	$t = 0.00$	$t = 0.01$	$t = 0.05$	$t = 0.10$	$t = 0.15$
$(-1,-1)$	0.132e-5	0.876e-5	0.765e-8	0.781e-5	0.131e-5
$(-1,1)$	0.162e-6	0.523e-5	0.653e-5	0.233e-8	0.102e-5
$(1,-1)$	0.705e-7	0.871e-6	0.213e-7	0.103e-6	0.342e-6
$(1,1)$	0.134e-5	0.983e-7	0.132e-8	0.761e-6	0.715e-6
$(0.5,0.5)$	0.114e-4	0.557e-5	0.833e-4	0.751e-5	0.330e-4
$\|Im(e_1)\|$	0.134e-4				
order	3.1043				

TABLE 6. Error of this method for the real part of $\psi_1(x,t)$ at the given times for case 2.

$(x_{1,j}, x_{2,j})$	$t = 0.00$	$t = 0.01$	$t = 0.05$	$t = 0.10$	$t = 0.15$
$(-1,-1)$	0.101e-7	0.243e-6	0.466e-6	0.892e-6	0.581e-6
$(-1,1)$	0.430e-6	0.730e-6	0.743e-5	0.672e-6	0.102e-5
$(1,-1)$	0.812e-7	0.286e-7	0.509e-4	0.103e-6	0.794e-6
$(1,1)$	0.194e-6	0.534e-8	0.254e-5	0.332e-8	0.613e-8
$(0.5,0.5)$	0.301e-6	0.198e-5	0.587e-3	0.541e-6	0.702e-5
$\|Re(e_1)\|$	0.434e-4				
order	3.2013				

TABLE 7. Error of this method for the imaginary part of $\psi_2(x,t)$ at the given times for case 2.

$(x_{1,j}, x_{2,j})$	$t = 0.00$	$t = 0.01$	$t = 0.05$	$t = 0.10$	$t = 0.15$
$(-1,-1)$	0.761e-6	0.275e-6	0.232e-7	0.043e-7	0.741e-7
$(-1,1)$	0.213e-7	0.400e-6	0.876e-6	0.543e-6	0.752e-6
$(1,-1)$	0.843e-4	0.320e-7	0.109e-8	0.429e-6	0.302e-6
$(1,1)$	0.634e-7	0.203e-5	0.274e-8	0.675e-8	0.101e-8
$(0.5,0.5)$	0.234e-7	0.153e-8	0.013e-7	0.543e-7	0.651e-7
$\|Im(e_2)\|$	0.532e-6				
order	3.5430				



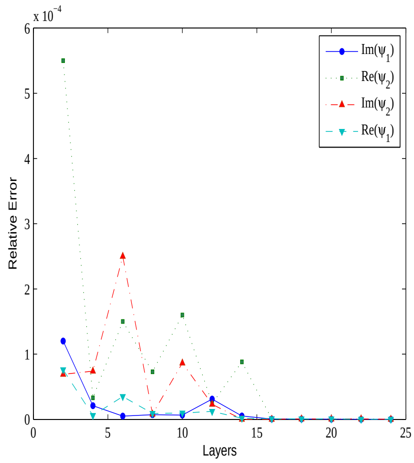


Figure 2. Convergence $E(Re(\psi_{m,1}))$, $E(Re(\psi_{m,1}))$, $E(Im(\psi_{m,1}))$ and $E(Im(\psi_{m,1}))$ when m increases for case 1 at $t = 0.2$ (in the left) and case 2 at $t = 0.5$ (in the right).

Table 8. Error of this method for the real part of $\psi_2(x,t)$ at the given times for case 2.

$(x_{1,j}, x_{2,j})$	$t = 0.00$	$t = 0.01$	$t = 0.05$	$t = 0.10$	$t = 0.15$
$(-1,-1)$	0.342e-7	0.654e-7	0.123e-6	0.067e-6	0.875e-4
$(-1,1)$	0.503e-7	0.476e-6	0.301e-8	0.342e-8	0.654e-5
$(1,-1)$	0.236e-6	0.709e-8	0.234e-7	0.543e-7	0.632e-6
$(1,1)$	0.432e-7	0.636e-7	0.532e-6	0.732e-6	0.535e-8
$(0.5,0.5)$	0.404e-6	0.637e-7	0.363e-6	0.401e-7	0.761e-6
$\|Re(e_2)\|$	0.654e-7				
order	2.9783				

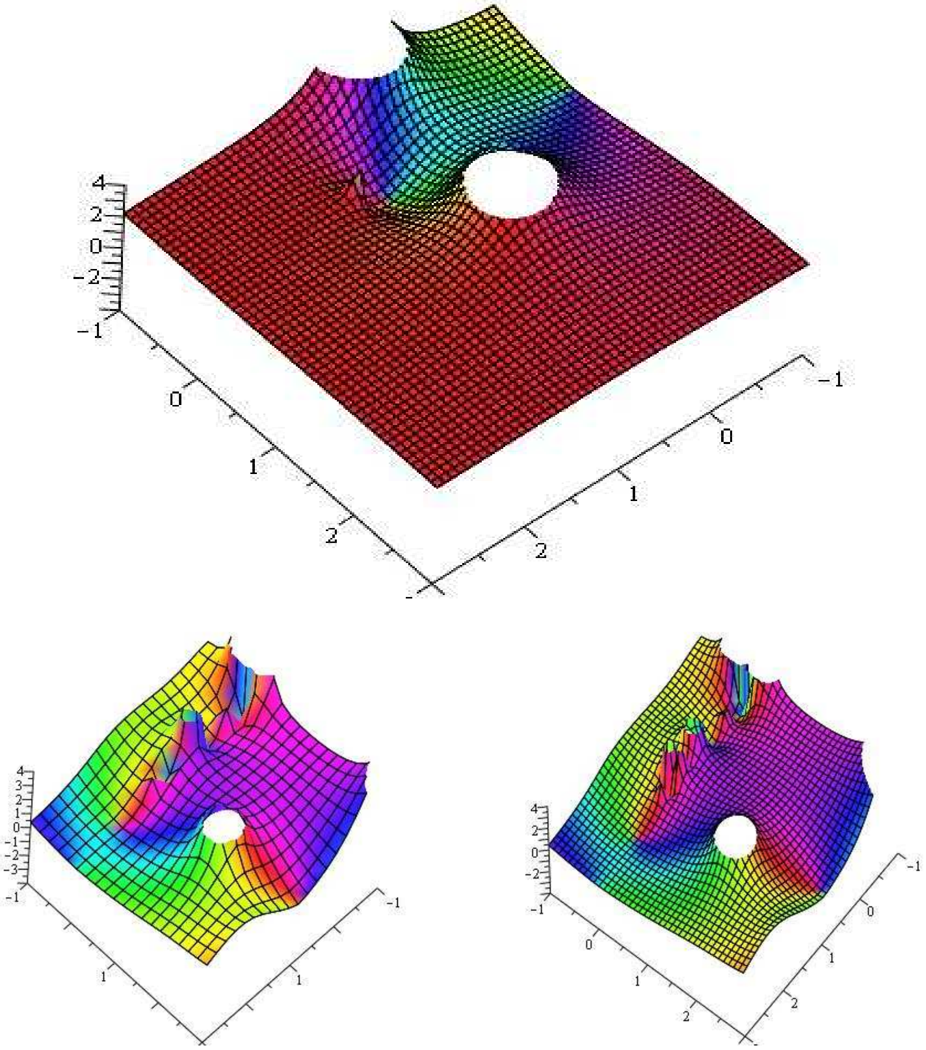


Figure 3. Top: The multiscale solution $\psi_1 = \psi_{c,1} + \psi_{f,1}$ at layer patch $m = 10$, below (heft): plot of the coarse scale solution $\psi_{c,1}$, (right): plot pf the fine scale solution $\psi_{f,1}$ for case 1 after 100 time steps.

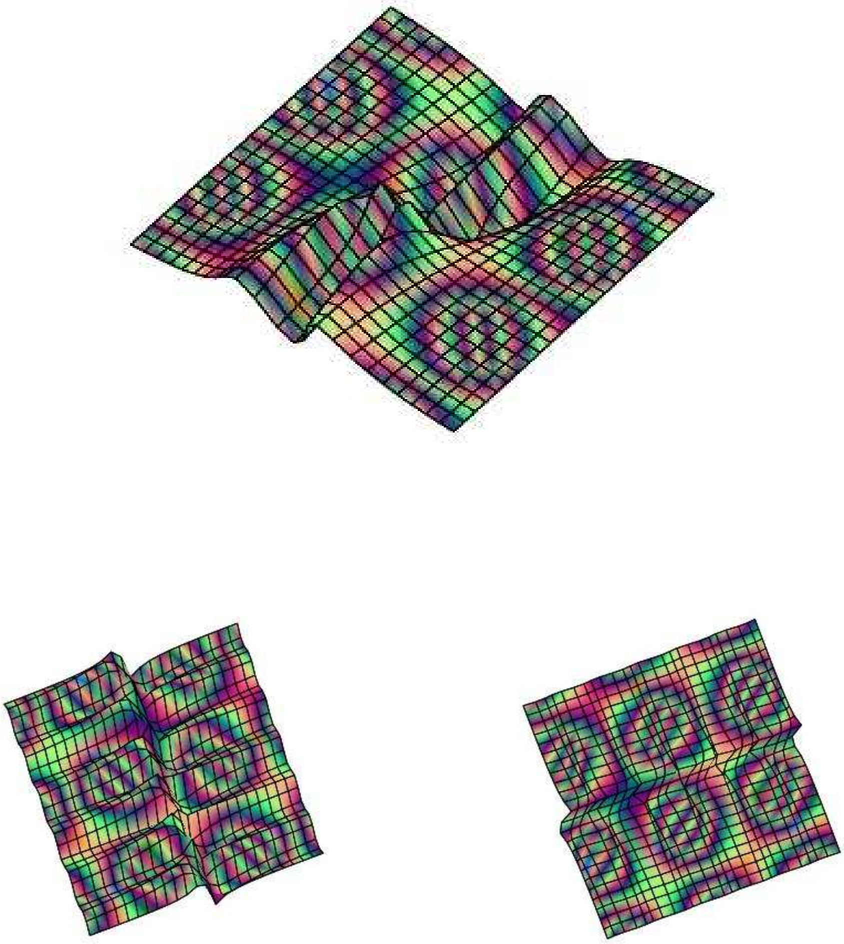


Figure 4. Top: the multiscale solution $\psi_2 = \psi_{c,2} + \psi_{f,2}$ at layer patch $m = 10$, Below (left): plot of the coarse scale solution $\psi_{c,1}$, (right: plot the fine scale solution $\psi_{f,2}$ for case 1 after 100 time steps.

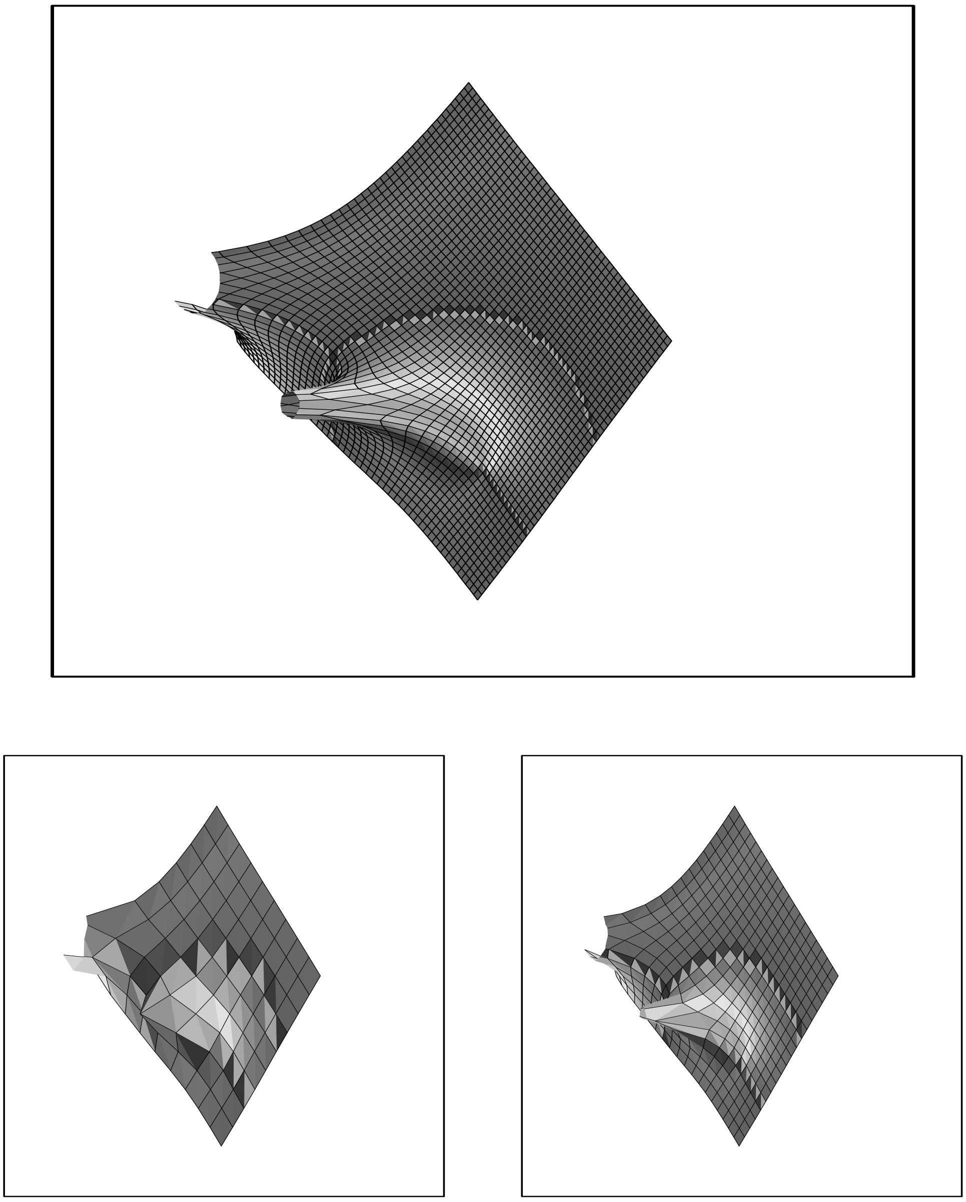


Figure 5. Top: the multiscale solution $\psi_1 = \psi_{c,1} + \psi_{f,1}$ at layer patch $m = 10$. Below (left): plot of the coarse scale solution $\psi_{c,1}$, (right): plot of the fine scale solution $\psi_{f,1}$ for case 2 after 65 time steps.

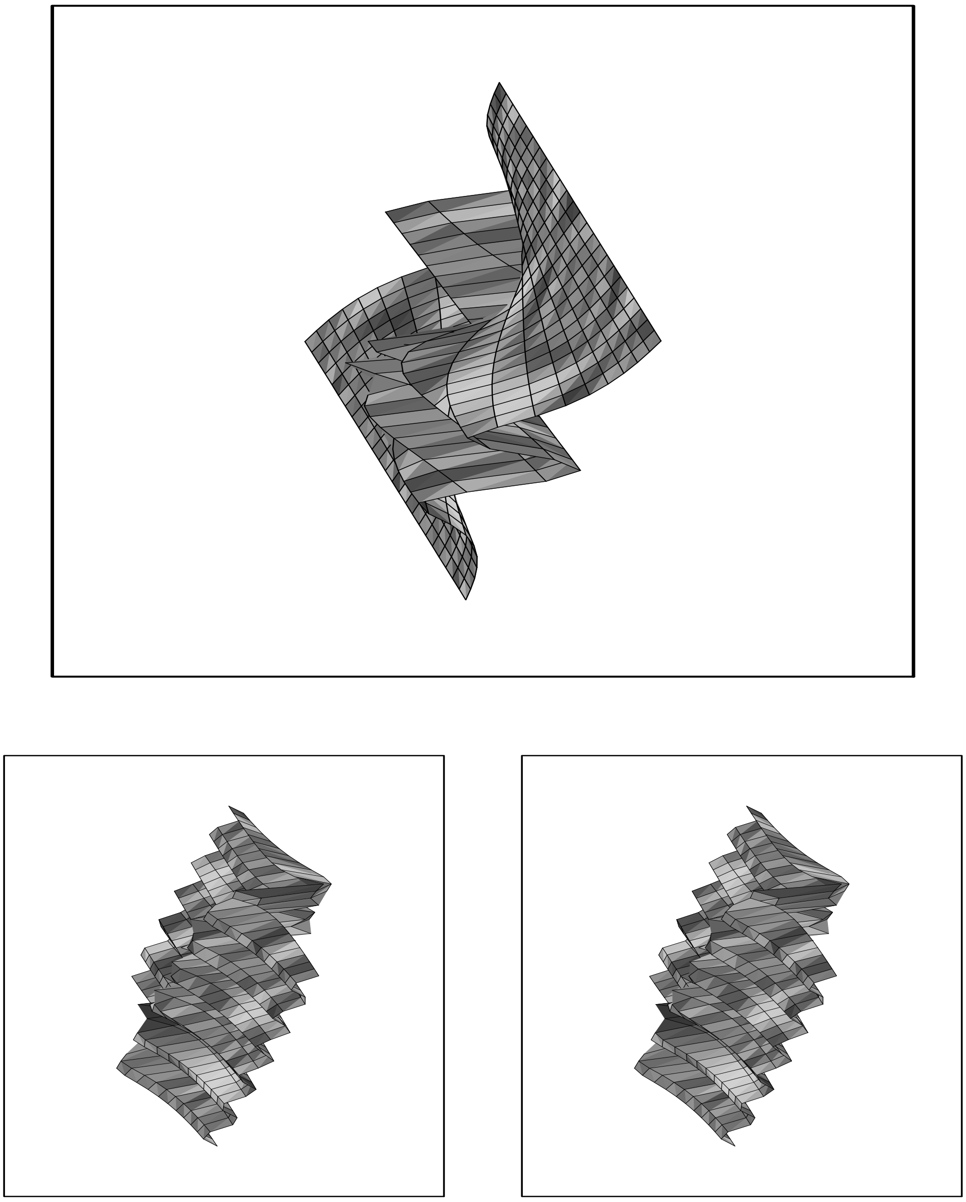


Figure 6. Top: the multiscale solution i.e. $\psi_2 = \psi_{c,2} + \psi_{f,2}$ at layer patch $m = 15$,
Below (left): plot of the coarse scale solution $\psi_{c,1}$, (right): plot of the fine scale solution $\psi_{f,2}$
for case 2 after 85 time steps.

7. Conclusion

We have constructed a streamline diffusion based, discontinuous Galerkin (DG) finite element scheme, for solving a coupled nonlinear Schrödinger system of equations. The scheme is extended to cover the multiscale variational cases which inherits the crucial stability and convergence properties of the original DG approximation. To prove a coercivity estimate, the original system is truncated through adding artificial viscosity (diffusion terms) to the equations. This viscosity terms is however, of small order of magnitude and falls in the framework of finite element approximations for convection dominated convection-diffusion problems. Except the coercivity estimates, major part of the analysis can be done without adding the extra (small) diffusion term. For the truncated system, we prove coercivity, stability, and convergence estimates. The convergence estimates are of optimal order $\mathcal{O}(h^{k+\frac{1}{2}})$ due to the maximal available regularity of the exact solution (here provide that the exact solution) $\mathbf{u}$ is in the Sobolev space $H^{k+1}(\Omega)$, where h is the global mesh size and k is the order of approximation polynomial. The original and multiscale schemes are numerically tested implementing an example of an application of the time dependent Schrödinger equation to the coupled ultrafast laser beam.

References

[1] R.A. Adams, *Sobolev Spaces*, Academic Press, New York (1975).

[2] M. Asadzadeh, D. Rostamy, and F. Zabihi, *A Posteriori Error Estimates for a Coupled Wave System with a Local Damping* , Journal of Mathematical Sciences, 175 (2011), pp. 228–248.

[3] M. Asadzadeh, *Streamline Diffusion Methods for the Vlasov-Poisson Equations*, RAIRO Math. Mod. and Numer. Anal., 24 (1990), pp. 177 - 196.

[4] M. Asadzadeh, *Streamline Diffusion Methods for the Fermi and Fokker-Planck Equations*, Transport Theory and Statistical Physics, (1997), pp. 319-340.

[5] M. Asadzadeh and P. Kowalczyk, *Convergence of Streamline Diffusion Methods for the Vlasov-Poisson-Fokker-Planck System*, Numer. Meth. Part. Diff. Eqs., 21 (2005), 472-495.

[6] R. Becerrila, F.S. Guzmana, A. Rendon-Romerob, and S. Valdez-Alvarado, *Solving the time-dependent Schrödinger equation using finite difference methods*, REVISTA MEXICANA DE FISICA E, 54 (2) (2008), pp. 120-132.

[7] S.C. Brenner and L.R. Scott, *The Mathematical Theory of Finite Element Method*, Springer-Verlag, New York, (1994).

[8] P. G. Ciarlet, The Finite Element Method for Elliptic Problems, North Holland, Amesterdam, (1987).

[9] K. Eriksson and C. Johnson, *Adaptive finite element methods for parabolic problems I: A linear model problem*, SIAM J. Numer. Anal., 28 (1991), pp. 43-77.

[10] K. Eriksson and C. Johnson, *Adaptive finite element methods for parabolic problems II: Optimal error estimates in $L_\infty L_2$ and $L_\infty L_\infty$*, SIAM J. Numer. Anal., 32 (1995), pp. 706-740.

[11] K. Eriksson and C. Johnson, *Adaptive finite element methods for parabolic problems IV: A nonlinear problem*, SIAM J. Numer. Anal., 32 (1995), pp. 1729-1749.

[12] K. Eriksson, C. Johnson and V. Thomee, *Time discretization of parabolic problems by the Discontinuous Galerkin method*, RAIRO, Anal. Numer., 19 (1985), pp. 611-643.

[13] W.I. Fushchych, Z.I. Symenoh, *Symmetry of Equations with Convection Terms*, Nonlinear Mathematical Physics, 4 (3-4) (1997), pp. 470-479.

[14] E. Godlewski and P. Raviart, *The numerical interface coupling of nonlinear hyperbolic systems of conservation laws:I. The scalar case*, Numer. Math., 97 (2004), pp. 81–130.

[15] H. Houde, J. Jicheng, W. Xiaonan, *A finite-difference method for the one-dimensional time-dependent schrödinger equation on unbounded domain*, Computers & Mathematics with Applications, 50 (8-9) (2005), pp. 1345-1362.

[16] Hughes T. J. R., and Brooks A., *A Multidimensional Upwind Scheme With No Crosswind Diffusion*, Finite Element Methods for Convection Dominated Flows, Ed. T. J. R. Hughes, AMD Vol. 34, ASME, New York (1979).

[17] Hughes T. J. R., and Mallet M., *A New Finite Element Formulation for Computational Fluid Dynamics III, The Generalized Streamline Operator for Multidimensional Advective—Diffusive Systems*, Comput. Methods Appl. Mech. Engrg. 58 pp. 305-328 (1986).

[18] M. S. Ismail, *Numerical solution of coupled nonlinear Schrödinger equation by Galerkin method*, Mathematics and Computers in Simulation, 78 (2008), pp. 532-547.

[19] M. S. Ismail and S. Z. Alamri, *Highly accurate finite difference method for coupled nonlinear Schrödinger equation*, Int. J. Comput. Math., 81 (3) (2004), pp. 333-351.

[20] M. S. Ismail, T. Taha, *A finite element solution for the coupled Schrödinger equation*, in: Proceedings of the 16th IMACS World Congress on Scientific Computation, Lausanne, (2000).

[21] P. Jamet, *Galerkin-type approximations which are discontinuous in time for parabolic equations in a variable domain*, SIAM J. Numer. Anal., 15 (1978), pp. 912–928.

[22] C. Johnson and J. Pitkaranta, *An analysis of the Discontinous Galerkin method for a scalar hyperbolic equation*, Math. comput., 46 (1986), pp. 1-26.

[23] C. Johnson, *Discontinous Galerkin finite element methods for second order hyperbolic problems*, Comput. Methods Appl. Mech. Enggrg., 107 (1993), pp. 117-129.

[24] O. Karakashian and C. Makridakis, *A space-time finite element mathod for the nonlinear Schrödinger equation: The discontinuous Galerkin method*, Math. Comput., 67 (222) (1998), pp. 479–499.

[25] M. G. Larson and A. Målqvist, *Adaptive variational multiscale methods based on a posteriori error estimation, Energy norm estimates for elliptic problems*, Computer Methods in Applied Mechanics and Engineering, 196 (21-24) (2007), pp. 2313-2324.

[26] P. Lesaint and P. A. Raviart, *On a finite element method for solving the neutron transport equation, Mathematical Aspects of Finite Elements in Partial Differential Equations (C. deBoor, ed.)*, Academic Press, New York, 1974, pp. 89–123.

[27] C. A. Raposo and W. D. Bastos, *Energe decay for the solutions of a coupled wave system*, TEMA Tend. Math. Appl. Comput., 10 (2) (2009), pp. 203-209.

[28] D. Rostamy, F. Zabihi, *A Streamline Diffusion Method for the Mass-Spring System*, J. Numer. Math. Stoch., 2 (1) (2010), pp. 76-89.

[29] R. Söderlind, *Finite Elemenet Methods for Multiscale/Multiphysics Problems*, PhD thesis, Department of Mathematics and Mathematical Statistics, Umeå University, (2011).

[30] J. Q. Sun, X. Y. Gu, Z. Q. Ma, *Numerical study of the soliton waves of the coupled nonlinear Schrödinger system*, Physica D, 196 (2004), pp. 311-328.

[31] M. Wadati, T. Izuka, M. Hisakado, *A coupled nonlinear Schrödirnger equation and optical solitons*, J. Phys. Soc. Jpn. 61 (7) (1992), pp. 2241-2245.

[32] Y. Xu, C. Shu, *Local discontinuous Galerkin method for nonlinear Schrödinger equations*, J. Comput. Phys. 205 (2005), pp. 72-97.

Department of Mathematics, Chalmers University of Technology and Göteborg University, SE–412 96, Göteborg, Sweden

Department of Mathematics, Imam Khomeini International University, Qazvin, Iran
E-mail address: `E-mail address: rostamyd@yahoo.com`

Department of Mathematics, Imam Khomeini International University, Qazvin, Iran
E-mail address: `E-mail address: fatemeh.zabihi@yahoo.com`