Discontinuous Galerkin and Multiscale Variational Schemes for a Coupled Damped Nonlinear System of Schrödinger Equations

M. Asadzadeh,[1] **D. Rostamy,**[2] **F. Zabihi**[2]
[1] *Department of Mathematics, Chalmers University of Technology and Göteborg University, SE–412 96 Göteborg, Sweden*

[2] *Department of Mathematics, Imam Khomeini International University, Qazvin, Iran*

Received 5 June 2012; accepted 9 January 2013
Published online in Wiley Online Library (wileyonlinelibrary.com).
DOI 10.1002/num.21782

In this article, we study a streamline diffusion-based discontinuous Galerkin approximation for the numerical solution of a coupled nonlinear system of Schrödinger equations 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 coercive, 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. Numer Methods Partial Differential Eq 000: 000–000, 2013

Keywords: coupled nonlinear Schrödinger equations; multiscale variational scheme; discontinuous Galerkin method; streamline diffusion method; stability; convergence

I. 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 [1–4]. 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

Correspondence to: M. Asadzadeh, Department of Mathematics, Chalmers University of Technology and Göteborg University, SE–412 96 Göteborg, Sweden (e-mail: mohammad@chalmers.se)
Contract grant sponsor: Swedish Foundation of Strategic Research (SSF) in Gothenburg Mathematical Modeling Center (GMMC) and The Swedish Research Council (VR)

equations are investigated by a number of authors and a physical problem was introduced by Fushchych in Ref. [1].

As for the numerical methods for this type of problems, the discontinuous Galerkin method has been considered by several authors in various settings, see for example, Refs. [5–9]. In our study, as a result of a special streamline-diffusion-based (SD) 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. The oscillations merge from the lack of stability of the standard Galerkin method for convection dominated problems. To enhance the stability, we modify the standard Galerkin by adding a multiple of convection term (expressed in terms of the test function) to the test function. This would result adding an artificial diffusion in the characteristic (streamlines) direction motivating for the name of the method: the SD method which improves stability in the streamline direction. However, this has to be done cleverly, because additional stability is often obtained at the cost of decreased accuracy. The objective is to improve stability without sacrificing the accuracy. Therefore, the choice of the multiple in the added convection term to the test function is crucial. For the present type of equations, a safe choice for the coefficient of the added term is of the order of the mesh size. Then, for example, 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 Ref. [10] and Hughes and Mallet in Ref. [11] for the fluid problems is further studied and mathematically developed, for example, for the hyperbolic partial differential equations and convection-diffusion problems in Refs. [6–9, 12].

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. Of course, one of the major consideration of a multiscale approach is for decoupling of localized problems that results in a faster solution technique (in a parallel setting). However, focusing on convergence analysis, parallel computation is not essential in this article. Instead, in this work, a hybrid method, combining multiscale and DG methods, is given for a CNSE. 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

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

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 Eq. (1.1) is usually associated with the following initial conditions:

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

and also the Neumann-type boundary conditions:

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

where $\mathbf{n}(\mathbf{x})$ is the outward unit normal to $\partial\Omega$ at the point $\mathbf{x} \in \partial\Omega$. Finally, $\mathbf{i}$ is the complex unity: $\mathbf{i}^2 = -1$, $\mathbf{i} = \sqrt{-1}$. 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 $(u_i, i = 1, \dots, 4)$ are real functions. Thus, the system (1.1), associated with modified boundary condition, can be converted to

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

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

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

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

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

$$\widetilde{\beta} = \begin{pmatrix} \beta\cdot \\ \beta\cdot \\ \beta\cdot \\ \beta\cdot \end{pmatrix}, \quad \text{and} \quad \mathcal{F}_\varepsilon^\beta(\mathbf{u})\mathbf{u} = \begin{pmatrix} \beta \cdot \nabla u_2 + \varepsilon z_1 u_2 \\ -\beta \cdot \nabla u_1 - \varepsilon z_1 u_1 \\ \beta \cdot \nabla u_4 + \varepsilon z_2 u_4 \\ -\beta \cdot \nabla u_2 + \varepsilon z_2 u_3 \end{pmatrix}.$$

Recall that, $\mathbf{n}$ is the outward unit normal to $\partial\Omega$, defined pointwise and denoted by $\mathbf{n}(\mathbf{x})$ for $\mathbf{x} \in \partial\Omega$, and

$$\partial\Omega_\beta^+ =: \partial\Omega \setminus \partial\Omega_\beta^- = \{\mathbf{x} \in \partial\Omega : \mathbf{n}(\mathbf{x}) \cdot \beta \geq 0\}.$$

Finally, if in the initial conditions (1.2), we let $g_j(x) = g_{j1}(x) + ig_{j2}(x)$ for $j = 1, 2$ then, $\mathbf{u}_0 = (g_{11}, g_{12}, g_{21}, g_{22})^T$.

Because 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 Eq. (1.5), which is modified yet with an additional viscosity term, viz

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

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 CNSE, see, for example, Refs. [3, 4, 13–16]. In the most recent studies, however, the focus has been moved toward some aspects of finite element approach, for example local DG methods are used for solving the general nonlinear Schrödinger equation as well as the CNSE are studied in Ref. [16] 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 [12, 16, 17]. First of all, the SD-based stability estimates and convergence analysis for the DG approximation of the problem (1.6) 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 article is as follows: In Section 2, we introduce DG variational muliscale scheme (DGVMS) based on SD 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 (1.6) as $\varepsilon \to 0$. Computational results are given in Section 5, and finally our concluding remarks are included in Section 6.

II. HYBRID DG/SD AND VARIATIONAL MULTISCALE METHOD FOR CDNSEC

A. Hybrid SD Based DG

In this section, we consider DG method to solve the Eq. (1.6) using the finite element approximation of the space-time domain Ω_T. To this end, let $0 = t_0 < t_1 < \cdots < t_N = T$ be a subdivision of the time interval $[0, T]$ into the subintervals $I_n = (t_{n-1}, t_n]$, with the time steps $k_n = t_n - t_{n-1}$, $n = 1, \ldots, N$ and introduce the corresponding space-time slabs,

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

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 $k \ge 0$ on τ and k on I_n (here, we recall that $P_k(K)$ denotes a polynomial space of order k on K). Note that $\{\mathcal{C}_h\}$ should be viewed as union of the following subdivision on S_n's, $n = 1, \ldots, N$:

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

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

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

where $(\mathbf{n}, \mathbf{n}_t) = (\mathbf{n}_\mathbf{x}, \mathbf{n}_t)$ denotes the outward unit normal to $\partial K \subset \Omega_T$. Below $P_k(K)$ and $P_k(K_n)$ will denote the space of polynomials of degree $\leq k$ on K and K_n, respectively. Now we introduce, for $k \geq 0$, the following spaces;

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

where for $n = 1, 2, \ldots, N$;

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

is a finite element space on S_n,

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

Also,

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

and

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

To derive a variational formulation, for the diffusive part of (1.6), 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:

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

Further, different L_2-based norms are denoted by $\|\cdot\|_{\Omega_T} := \|\cdot\|_{L_2(\Omega_T)}$, $\|\cdot\|_{k,\Omega_T} := \|\cdot\|_{L_k(\Omega_T)}$, $\|\cdot\|_{\infty,\Omega_T} := \|\cdot\|_{L_\infty(\Omega_T)}$, and $\|\cdot\|_s := \|\cdot\|_{s,\Omega_T} = \|\cdot\|_{H^s(\Omega_T)}$. Recall that as $\hat{\beta} = \mathcal{F}_\varepsilon^\beta(\mathbf{u_h})$ 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 DG method for (1.6) can now be formulated as follows: find $\mathbf{u} \in \mathbf{W}_h$ such that

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

where $\partial K_-(\hat{\beta})' = \partial K_-(\hat{\beta}) \setminus \Omega \times \{0\}$. Here, $\delta = 0$ is the usual DG 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

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

and the linear form L,

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

Then, the finite element method for (1.6) can now be formulated in the compact form as follows: Find $\mathbf{u}_h \in \mathbf{W}_h$ such that

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

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 $F(\mathbf{u})$ defined in (1.4) and (1.5)).

B. Variational Multiscale Method

Now we use a symmetric split proposed in Ref. [18], 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

$$\mathbf{u}_h = \mathbf{u}_c + \mathbf{u}_f + \mathcal{T}\mathbf{u}_c. \qquad (2.5)$$

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 that is, 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 as 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 $\mathbf{v}_c$.

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

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

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

Fine-scale equations are derived by putting $\mathbf{v}_c = 0$ in (2.7):
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, because the coarse and fine scales are completely separated. Also, the coarse-scale solution is obtained by letting $\mathbf{v}_f = 0$ in (2.7) 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 (2.11), $\mathcal{T}\mathbf{v}_c$ 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. Because, 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 (2.8)-(2.10) 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 for $\varepsilon > 0$ as in (2.4).

III. STABILITY FOR THE SD BASED DG

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, follow the conventional procedure where we need to show that the bilinear form B introduced in (2.3) is coercive. In some cases, it is possible (and is more convenient) to circumvent the coercivity. However, this often ends up with a rather involved stability approach, see, for example, the penalty approach in Ref. [19]. Here, we shall state the coercivity criterion both in the normal sense as well as in the multiscale case of coarse and fine meshes. The proof is, however, given in the usual case and is easily expandable for the multiscale case setting. Our stability and error estimates will be given in the following triple norm,

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

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

Proposition 3.1. *Let B be defined as in* (2.3), *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), where setting $\mathbf{v} = \mathbf{g}$ it follows that

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

Below we shall estimate each $T_i : i = 1, \ldots, 9$, separately. Note that

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

$$= \frac{1}{2}\left[|\mathbf{g}_+|_0^2 + |\mathbf{g}_-|_N^2 + \sum_{K\in\mathcal{C}_h} \int_{\partial K_-(\hat{\beta})'} [\mathbf{g}]^2|\mathbf{n}_t + \mathbf{n}\cdot\hat{\beta}|d\sigma\right]. \tag{3.4}$$

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, as $\widetilde{\beta}$ is linear in $\mathbf{x}$ we have $\nabla\widetilde{\beta} \equiv C$ (a constant), and therefore

$$T_{42} = -(\nabla A\mathbf{g}, \widetilde{\beta}\mathbf{g})_{\Omega_T} = (A\mathbf{g}, \nabla(\widetilde{\beta}\mathbf{g}))_{\Omega_T} - \langle A\mathbf{g}, \widetilde{\beta}\mathbf{g}\rangle_{\partial\Omega_T}$$

$$= (A\mathbf{g}, (\nabla\widetilde{\beta})\mathbf{g})_{\Omega_T} + (A\mathbf{g}, \widetilde{\beta}\nabla\mathbf{g})_{\Omega_T}, \tag{3.7}$$

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

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

Likewise, using the definition of A yields

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

To estimate T_8, we have using first the trace estimate, and then the inverse estimate see Ref. [20],

$$-T_8 = \frac{1}{2}\int_{\partial\Omega_-\times I} ((A + \tilde{\varepsilon}I)\nabla\mathbf{g})\mathbf{g}d\sigma ds \le \frac{\gamma_1}{4}\|(A + \tilde{\varepsilon}I)\nabla\mathbf{g}\|^2_{L_2(\partial\Omega_-\times I)} + \frac{1}{4\gamma_1}\|\mathbf{g}\|^2_{L_2(\partial\Omega_-\times I)}$$

$$\le \sqrt[4]{8}\frac{\gamma_1}{4}(\|(A + \tilde{\varepsilon}I)\nabla\mathbf{g}\|_{\Omega_T}\|(A + \tilde{\varepsilon}I)\Delta\mathbf{g}\|_{\Omega_T}) + \sqrt[4]{8}\frac{1}{4\gamma_1}\|\mathbf{g}\|_{\Omega_T}\|\nabla\mathbf{g}\|_{\Omega_T}$$

$$\le \sqrt[4]{8}\frac{\gamma_1}{4}h^{-1}\|(A + \tilde{\varepsilon}I)\nabla\mathbf{g}\|^2_{\Omega_T} + \sqrt[4]{8}\frac{1}{8\gamma_1}\|\mathbf{g}\|^2_{\Omega_T} + \sqrt[4]{8}\frac{1}{8\gamma_1}\|\nabla\mathbf{g}\|^2_{\Omega_T}.$$

Similarly,

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

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

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

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

Now, assuming, for example, $C = \frac{\tilde{\varepsilon}}{8} + (2\sqrt{2}+1)(1+\tilde{\varepsilon})^2 + 2(1+|\Omega_T|)h$, 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 multiscale cases:

Corollary 3.1. *Let B be defined as in* (2.3)*, then there exists a constant $\alpha > 0$ independent of h such that:*

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

and

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

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 no δ-term but jumps. Hence, it does not rely on the additional viscosity term $-\tilde{\varepsilon}\Delta\mathbf{u}$ of (1.6). Thus, the proposition below is valid for the larger, but less smooth, system described by (1.5), and obviously, in a modified form, is valid for (1.6) as well.

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

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

where

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

and

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

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

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

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

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

Using Green's formula yields

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

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

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

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

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

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

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

Integrating over I_n and summing over n (for $n = 0, 1, \ldots, N-1$, 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 and assuming $\mathbf{g} \in \mathbf{W}_{h,i}$ with $i = c, f$.

IV. CONVERGENCE FOR THE SD-BASED DG

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

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

Below we shall use Propositions 3.1 and 3.2 and derive optimal convergence rates for the problem (1.6). Similar convergence analysis for the problem (1.5) is a rather cumbersome procedure. This, as we mentioned in the previous section, is due to the lack of a coercivity estimate like the one in the Proposition 3.1. An estimate as in Proposition 3.2 works only for the L_2- stability. 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 interpolant: $I_{h,f}\mathbf{u} \in \mathbf{W}_{h,f}$ 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,

$$\begin{aligned}e_f &:= \mathbf{u} - \mathbf{u}_{h,f} = (\mathbf{u} - I_{h,f}\mathbf{u}) - (\mathbf{u}_{h,f} - I_{h,f}\mathbf{u}) = \eta_f - \xi_f, \\ e_c &:= \mathbf{u} - \mathbf{u}_{\widehat{h},c} = (\mathbf{u} - I_{\widehat{h},c}\mathbf{u}) - (\mathbf{u}_{\widehat{h},c} - I_{\widehat{h},c}\mathbf{u}) = \eta_c - \xi_c,\end{aligned}$$

and we have the corresponding Galerkin orthogonality relations given by

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

and

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

Theorem 4.2. *If the exact solution* $\mathbf{u} \in L^\infty(H^{k+1}(\Omega))$, $k \geq 1$*; for* (1.6) *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

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

Proof. Using the definition of η we can, formally, write

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

Here, we invoke for $\mathcal{F}_\varepsilon^\beta(\mathbf{u}_h)$ from (1.6) and estimate the terms T_1 and $T_2 - T_3$, separately. Then, the additional viscosity term is explicitly present in the terms S_3 and S_4 of T_1 below. To estimate the term T_1, we use the inverse inequality and the above assumptions to obtain

$$\begin{aligned}
T_1 &= B(\mathcal{F}_\varepsilon^\beta(\mathbf{u}_h); \eta, \xi) = (\eta_t + \mathcal{F}_\varepsilon^\beta(\mathbf{u}_h)\eta, \xi + \delta(\xi_t + \mathcal{F}_\varepsilon^\beta(\mathbf{u_h})\xi))_{\Omega_T} \\
&\quad + \sum_{K\in\mathcal{C}_h} \int_{\partial K_-(\hat{\beta})'} [\eta]\xi_+|\mathbf{n}_t + \mathbf{n}\cdot\hat{\beta}|d\sigma + \frac{1}{2}(A\nabla\eta, \nabla\xi)_{\Omega_T} + \frac{1}{2}\tilde{\varepsilon}(\nabla\eta, \nabla\xi)_{\Omega_T} \\
&\quad - \frac{1}{2}\int_{\partial\Omega_-\times I} ((A + \tilde{\varepsilon}I)\nabla\eta)\xi d\sigma ds - \frac{1}{2}\delta((A + \tilde{\varepsilon}I)\Delta\eta, \xi_t + \mathcal{F}_\varepsilon^\beta(\mathbf{u}_h)\xi)_{\Omega_T} + < \eta_+, \xi_+ > 0 \\
&:= \sum_{i=1}^{7} S_i.
\end{aligned}$$

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

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

We may write the first term in S_1 as

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

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

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

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

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

where we have used the fact that

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

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

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

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

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

where we assume that c is sufficiently small, then 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

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

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

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

where, we may hide the term $\frac{h}{C}||\xi_t + \hat{\beta}\xi||^2_K$ 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

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

Combining the estimates for all S_1 and S_2 terms, we end up with the estimate

$$|S_1 + S_2| \leq c|||\xi|||^2 + C_1 h^{2k+1} + C\left(\int_{\partial\Omega_+ \times I} \eta^2 |\mathbf{n} \cdot \beta_1| d\sigma ds + \sum_{n=1}^{N} |\eta|_n^2\right). \tag{4.20}$$

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

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

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

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

Further, as

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

hence

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

Similar argument yields

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

As for the term S_7 we have

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

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

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

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

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

where we estimate the third term on the right-hand side above as

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

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

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

Finally, by standard interpolation theory we have also (see Ref. [21], p. 123)

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

Thus, using the assumptions, together with $||u||_{k+1,\infty} \leq \infty$ we end up with

$$|||\xi|||^2 \leq Ch^{2k+1} + C_1h\sum_{n=1}^N|\xi_-|_n^2. \tag{4.28}$$

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

$$y(\cdot, t_n) \leq C + C_1h\sum_{n=1}^N|y(\cdot, t_n)|_n^2,$$

then

$$y(t_n) \leq Ce^{C_1t} \leq Ce^{C_1T}.$$

Obviously (4.28) implies that

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

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

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

Thus, by (4.28) and (4.29),

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

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

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

completes the proof and we have

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

■

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* $k \geq 1$ *is exact solution* (1.6) *and*

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

then we have the following error estimate

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

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

V. CONSTRUCTION OF MV ON DG/SD

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

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

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

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

Below we shall suppress the superscript n. Then, without loss of generality, the variational multiscale method for this problem is formulated as follows [18]:

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

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

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 $\varepsilon \to 0$, that is, we have

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

and

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

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

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

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

and

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

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), for the coarse-scale equation, as follows:

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

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 by $\mathcal{F}^{\beta}_{\varepsilon}(\mathbf{u}_h)$, then we may continue considering the above system without removing the ε terms. This corresponds to a nonvanishing, but small value for ε ($\varepsilon \nrightarrow 0$). We denote this phenomenon by replacing $\overline{B}_{\delta,\varepsilon\to 0,\beta}$ with a new notation: $\overline{B}_{\delta,\varepsilon\rightsquigarrow 0,\beta}$.

We emphasize that, in principle, the analysis of the variational multiscale method is similar to that for the original SD-based DG approach studied along the previous sections. Some detailed analysis can be found in Ref. [22] dealing with convergence of a nonconforming multiscale approach. A swift follow up is given by Bochevi et al. in Ref. [23], where a multiscale DG method is introduced. For further studies and applications, we refer the reader to the PhD thesis by Söderlind [24] and the references therein. As we mentioned earlier in our approach, the SD modification, being involved in the analysis, improves the stability of the considered convection-diffusion problem without causing a significant reduction of the convergence rate of the approximation scheme. In most studies, for such an involved model, a stability parameter is computationally determined to fit some canonical examples.

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

VI. 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^c_n. 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^c_n, which we denote by $\{\theta_i\}$. The patches will include the support of the associated basis functions. Below we recall the standard definition of a patch:

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

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

where θ_j is a coarse-scale piecewise linear continuous basis function. Further in (6.1) *we need w_i^1 which is a symmetric 1-layer patch defined by $w_i^1 = supp\ \phi_i$, 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 the oriented 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.

A. Matrix Equation by Spatial Discretization

In this part, we introduce three systems of fine scale equations and one coarse-scale equation and formulate the coarse-scale equation in full details. Let $\mathbf{U}_{f,l,i}$, $\mathbf{U}_{f,0,i}$, $\mathcal{T}\phi_i$ belong to $\mathbf{W}_{h,f}(w_i)$ $(\mathbf{W}_{h,c}(w_i))$ that is, the piecewise linear polynomials fine (coarse) scale spaces on the patch w_i in time such that they solve the following systems:

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

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

and

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

for all $i \in \mathcal{N}$. Below, we present the scheme based on 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 obtain the following coarse-scale equation:

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

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

Remark 6.2. We shall denote by $\mathcal{E}_I^i$ the set of all interior edges in the mesh $\mathbf{W}_{\widehat{h},c}(w_i)$ and by $\mathcal{E}_\Gamma^i$ we shall mean 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.

Applying the backward Euler method for equations of (6.2), yields the following matrix equation:

$$\begin{cases} (A_1^i + \delta(A_2^i + A_3^i) + \Delta t_i(B_1^i + B_2^i + B_3^i + B_4^i))\overrightarrow{\mathbf{U}}_{f,l,i}(t_n) = \Delta t_i b^i, \\ H\overrightarrow{\mathbf{U}}_{f,l,i} = 0, \quad i = 1, \dots, N, \end{cases} \tag{6.6}$$

where we have used the notation,

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

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 Eqs. (6.2)–(6.4) to get the fine scale contribution. Consequently, we obtain the backward Euler method where for a given $\mathbf{U}_c(t_{n-1})$, we can compute $\mathbf{U}_c(t_n)$ by solving the following matrix equation

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

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

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

Remark 6.3. Note that if we have $w_i = \Omega$ and the same resolution is used in all patches the reference solution on the fine mesh is 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.

Algorithm 6.4. *DGVMS for solving* (1.1)

Step 0- *Input* c_0 ,α, β, ε, $\tilde{\varepsilon}$, $1 \leq \gamma \leq 2$, Ω, $\mathbf{u}_0$, N, $\widehat{h}$, m, j *and an error tolerance TOL.*
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- *Discretize* $[0, T]$, $0 = t_0 < t_1 < \cdots < t_N = T$ *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* $A_1^i, A_2^i, A_3^i, B_1^i, B_2^i, B_3^i, B_4^i$ *and vector* b^i *on each patch.*
Step 5- *Compute 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* $\overrightarrow{\mathbf{U}}_{f,l,i}$ $i \in \mathcal{N}$.
Step 7- *Compute* $\overrightarrow{\mathbf{U}}_c$ *by solving* (6.7).
Step 8- *Construct* $\overrightarrow{\mathbf{U}} = \overrightarrow{\mathbf{U}}_c + \overrightarrow{\mathbf{U}}_f$ *and check: If the TOL is reached stop. Otherwise,*
Step 9- *Relable* j, N *and go to step 2.*

B. Experimental Results and Some Realistic Applications

The Schrödinger equation plays the role of Newton's laws and conservation of energy in classical mechanics—that is, it predicts the future behavior of a dynamical system [25, 26]. 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 $\delta = h$, $\varepsilon = \tilde{\varepsilon} = 0.01$ and $\beta = (1, 1)^T$. The accuracy of this method with the reference solutions $\psi_{R,i}(x, t)$ $i = 1, 2$ is tested both pointwise, and in the triple norm. (The first five rows in the tables are the results of pointwise estimates, and the last line in the tables concern results for the discrete triple norm. The ℓ_2 estimates are included in the discrete triple norm). We recall that $\psi_{R,i}(x, t)$ is a reference solution computed on a fully resolved mesh and we use linear elements for the approximation.

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

*(***Case 1***)*

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

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

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

where

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

TABLE I. 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
$\|\|\|Im(e_1)\|\|\|$	0.218e-5				
order	3.4261				

TABLE II. 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
$\|\|\|Re(e_1)\|\|\|$	0.301e-5				
order	3.9624				

TABLE III. 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
$\|\|\|Im(e_2)\|\|\|$	0.341e-6				
order	3.9653				

$x_{1,L} = x_{2,L} = -1$, $x_{1,R} = x_{2,R} = 1$, $\widehat{h_c} = \alpha = 0.1$ and $N = 100$. The accuracy is measured in the L_2 and triple norms defined by (3.1)*:*

$$|||Re(e_i(h))||| = |||Re(\psi_{R,i}) - Re(\psi_{h,i})|||$$

and

$$|||Im(e_i(h))||| = |||Im(\psi_{R,i}) - Im(\psi_{h,i})|||$$

TABLE IV. 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
$\|\|\|Re(e_2)\|\|\|$	0.242e-6				
order	3.6501				

TABLE V. 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
$\|\|\|Im(e_1)\|\|\|$	0.276e-5				
order	3.2482				

TABLE VI. 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
$\|\|\|Re(e_1)\|\|\|$	0.576e-5				
order	3.4698				

TABLE VII. 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
$\|\|\|Im(e_2)\|\|\|$	0.984e-6				
order	3.3491				

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

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

and

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

TABLE VIII. 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
$\|\|\|Re(e_2)\|\|\|$	0.354e-6				
order	3.1280				

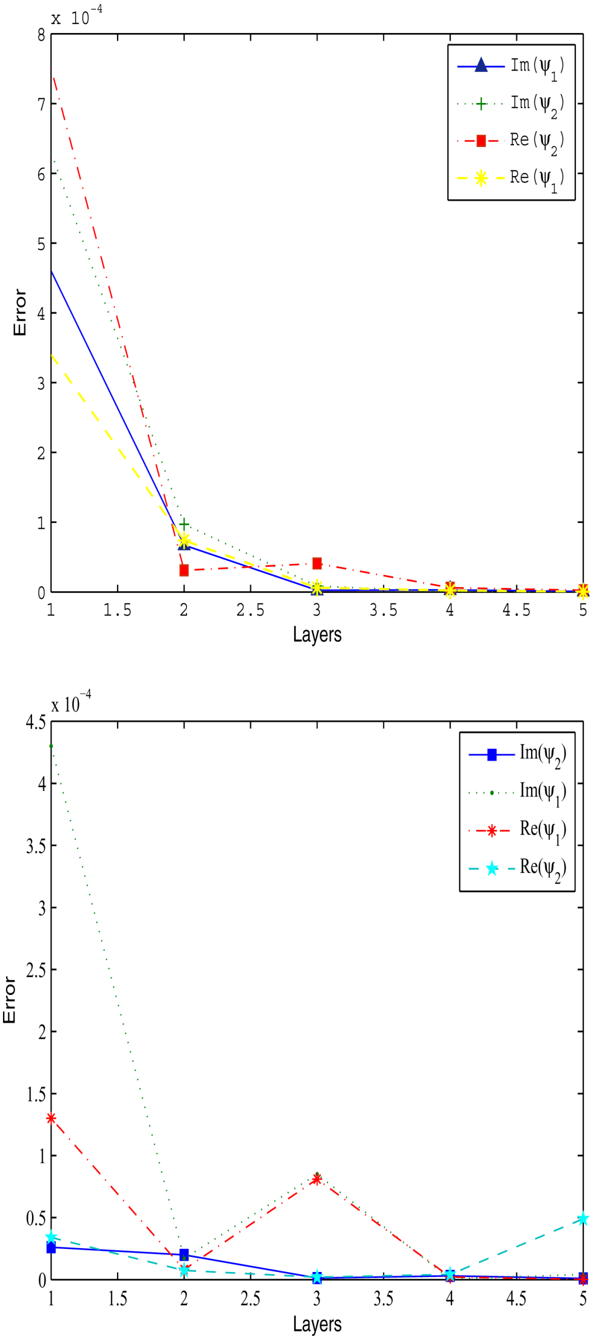

FIG. 1. Decay of the approximation solution on boundary for different layers for Case 1 at $t = 0.06$ (in the top) and Case 2 at $t = 1$ (in the bottom). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

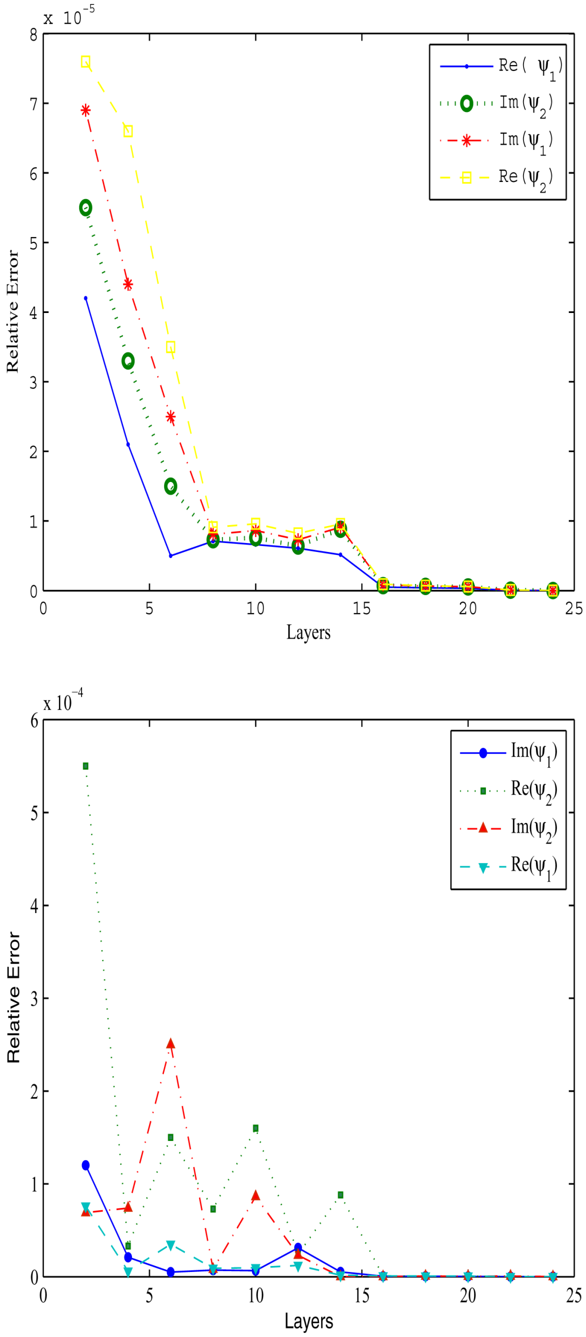

FIG. 2. Convergence $E(Re(\psi_{m,1}))$, $E(Re(\psi_{m,1}))$, $E(Im(\psi_{m,1}))$, and $E(Im(\psi_{m,1}))$ when m increases for Case 1 at $t = 0.2$ (in the top) and Case 2 at $t = 0.5$ (in the bottom). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

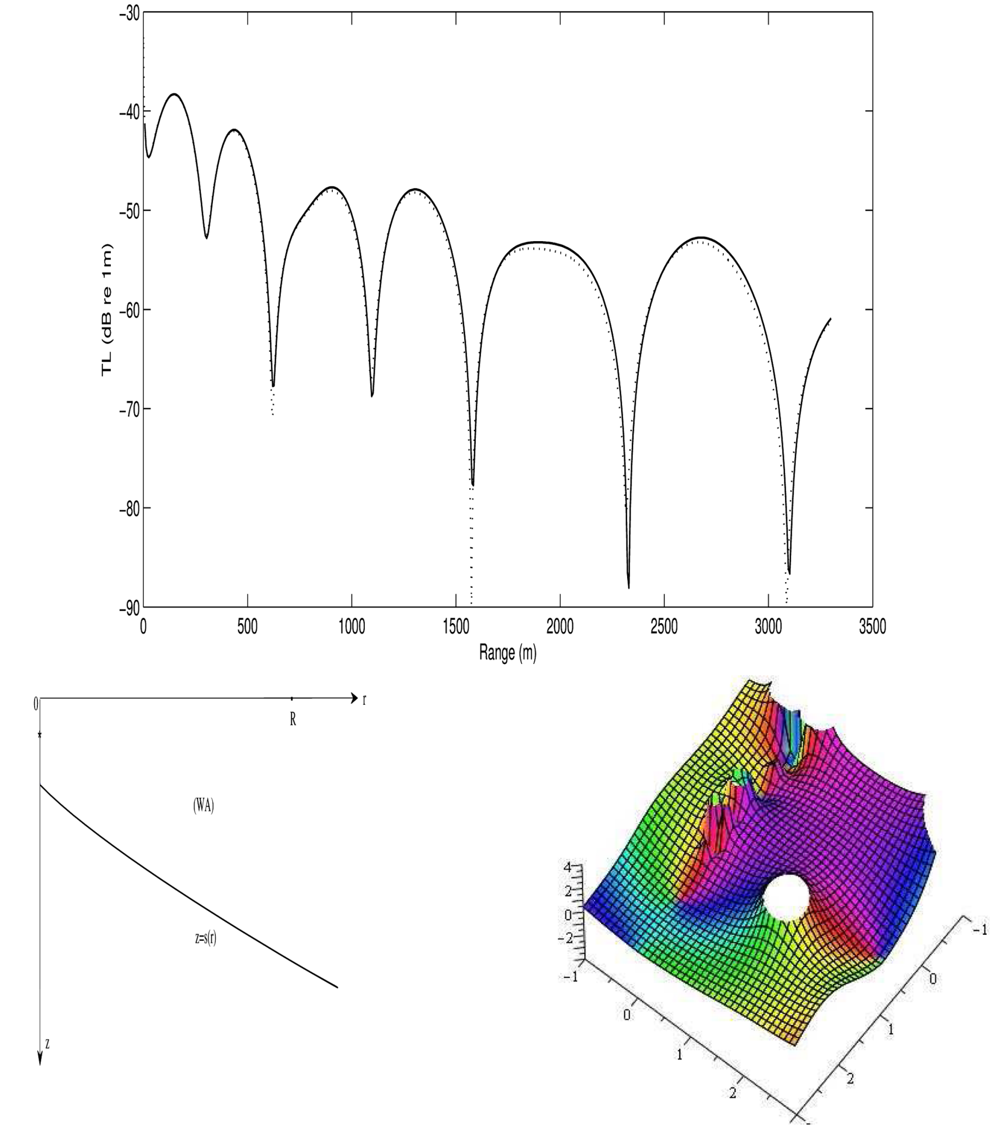

FIG. 3. Top: The multiscale solution $\psi_1 = \psi_{c,1} + \psi_{f,1}$ at layer patch $m = 10$, below (left): plot of the coarse scale solution $\psi_{c,1}$, (right): plot of the fine scale solution $\psi_{f,1}$ for Case 1 after 100 time steps. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

where $i = 1, 2$. *We carry out the above algorithm, by an AMD Opteron computer where* 15-*GB RAM memory with* 2.2-*GHz CPU has been used for the experiments. The pointwise error quantity (all entries of the tables except the bottom lines) and the discrete triple norm (the last lines on each table indicated by the triple norm for the real and imaginary part of the error and the wording "order") of error for the approximation method with reference solution for the real and imaginary parts are given in Tables I–VIII.*

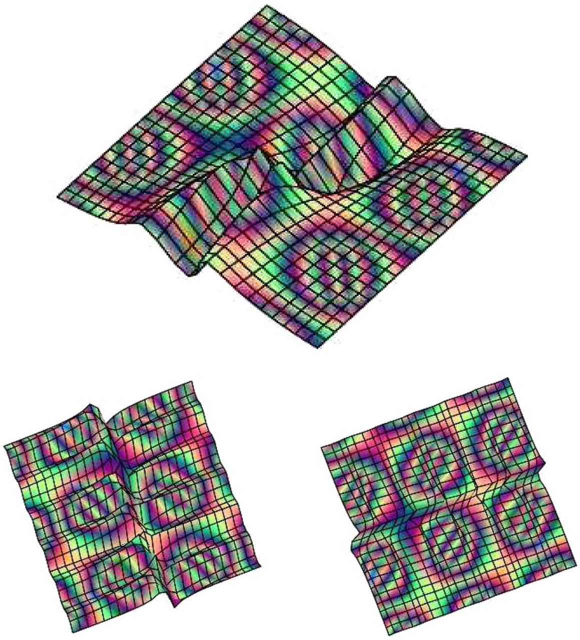

FIG. 4. Top: the multiscale solution $\psi_2 = \psi_{c,2} + \psi_{f,2}$ at layer patch $m = 10$, Below (left): plot of the coarse scale solution $\psi_{c,1}$, (right: plot the fine scale solution $\psi_{f,2}$ for Case 1 after 100 time steps. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

More specifically, the magnitude of the real and imaginary parts of the pointwise (local) errors are computed at the geometric singularities (corners of a reference domain $\Omega =: (-1, 1) \times (-1, 1)$*) at the discrete times* $t = 0.00, 0.01, 0.05, 0.10,$ *and* 0.15*. The errors are also computed at a point inside the domain:* $(0.5, 0.5) \in (-1, 1) \times (-1, 1)$*; which together with assuming an inflow in a diagonal direction (*$(-1, -1) \to (1, 1)$*, not aligned with the mesh) shows, for example, that the method efficiently damps the accumulation errors in time. Finally, global error estimates are represented by the triple norms of the real and imaginary parts of the relative error.*

It can be shown 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 the standard Galerkin finite element methods. The agreement, and the small triple norm error between exact and numerical solution in the considered example show that the method is accurate. The order of error in this method is close to 4*, whereas the the finite difference methods has an error of order* 2 *and that of the standard Galerkin finite element method is* 3 *(see Refs.* [27], [28] *and* [29] *for further details). We recall that the convergence rates of errors by triple norm is addressed in Theorem* 4.2*. In Fig. 1, we report decay of error of approximate solutions in the triple norm for layers. The convergence is measured in the relative error*

$$E(Re(\psi_{R,i})) = \frac{|||Re(\psi_{m,i}) - Re(\psi_{R,i})|||}{|||Re(\psi_{R,i})|||}, \quad and \quad E(Im(\psi_{R,i})) = \frac{|||Im(\psi_{m,i}) - Im(\psi_{R,i})|||}{|||Im(\psi_{R,i})|||}$$

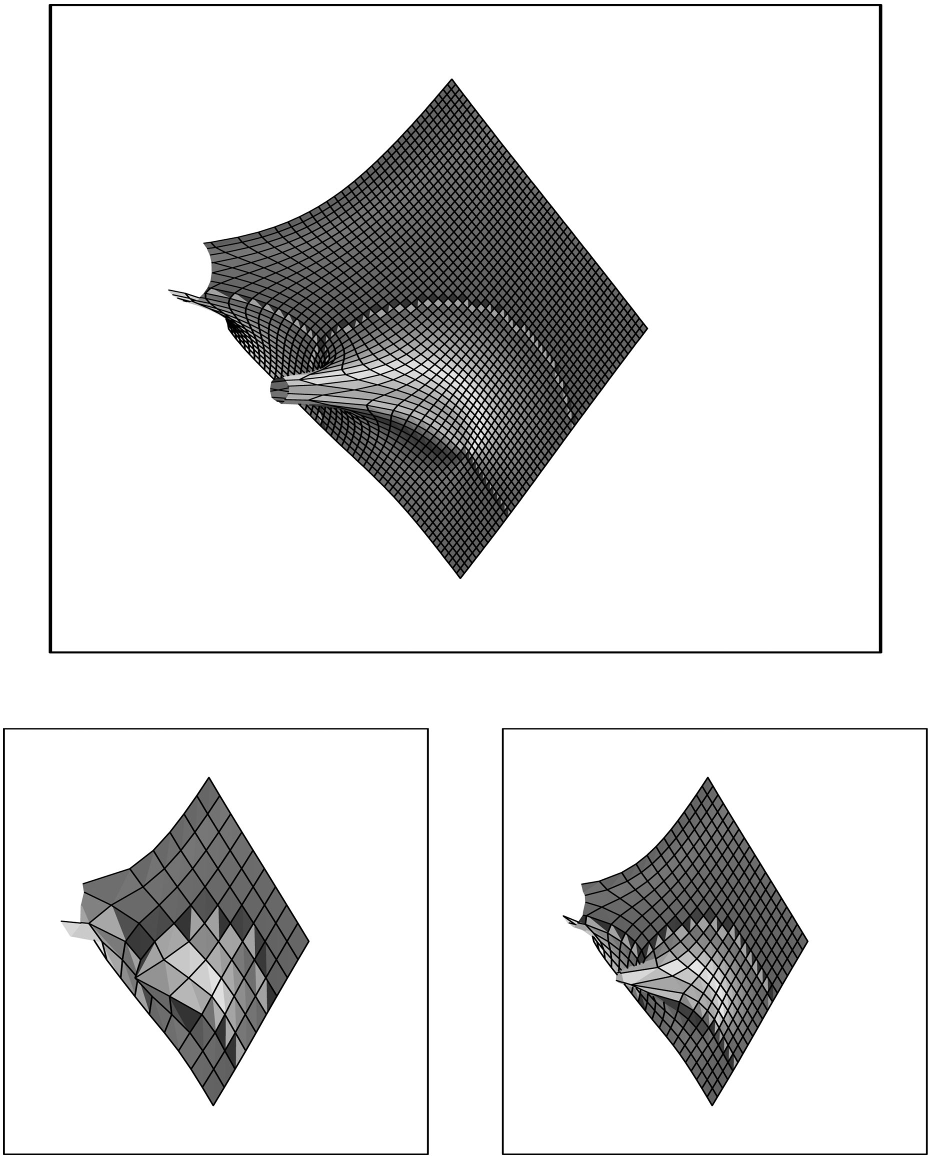

FIG. 5. Top: the multiscale solution $\psi_1 = \psi_{c,1} + \psi_{f,1}$ at layer patch $m = 10$. Below (left): plot of the coarse scale solution $\psi_{c,1}$, (right): plot of the fine scale solution $\psi_{f,1}$ for Case 2 after 65 time steps.

(see Fig. 2) where $\psi_{m,i}$, $m = 2, 3, 4, \ldots, i = 1, 2$*, 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, 4, 5 and 6. Finally, we also investigated the above convergence rates altering the behavior of the damping factor by varying* ε *in the range* $[0, 0.01]$. *But the variations in the results were very small and therefore are not reported in here.*

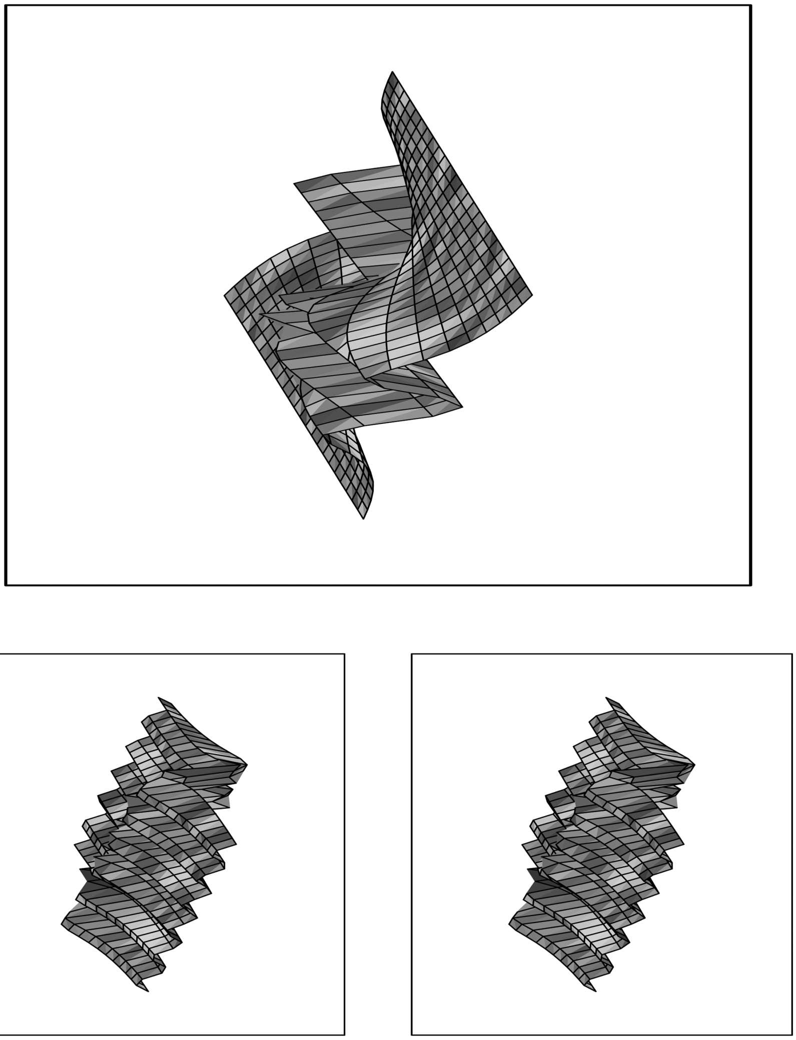

FIG. 6. Top: the multiscale solution that is, $\psi_2 = \psi_{c,2} + \psi_{f,2}$ at layer patch $m = 15$, Below (left): plot of the coarse scale solution $\psi_{c,1}$, (right): plot of the fine scale solution $\psi_{f,2}$ for Case 2 after 85 time steps.

VII. CONCLUSIONS

We have constructed a, SD-based, 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 into 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 provided that the exact solution) $\mathbf{u}$ which is assumed to be in the Sobolev space $H^{k+1}(\Omega)$, where h is the global mesh size and k is spectral order (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.

Finally, it is important to mention that here the purpose of our study is to derive optimal convergence rates for both single- as well as multiscale cases. Of course, one of the major consideration of a multiscale approach is for decoupling of localized problems that results in a faster solution technique (in a parallel setting). However, focusing on theoretical aspects, we do not deal with parallel computation in this article. We plan to address this and some other related issues in a forthcoming article.

The authors would like to thank the referees for their valuable comments/suggestions and constructive criticism.

References

1. W. I. Fushchych and Z. I. Symenoh, Symmetry of equations with convection terms, Nonlin Math Phys 4 (1997), 470–479.
2. O. Karakashian and C. Makridakis, A space-time finite element method for the nonlinear Schrödinger equation: the discontinuous Galerkin method, Math Comput 67 (1998), 479–499.
3. J. Q. Sun, X. Y. Gu, and Z. Q. Ma, Numerical study of the soliton waves of the coupled nonlinear Schrödinger system, Physica D 196 (2004), 311–328.
4. M. Wadati, T. Izuka, and M. Hisakado, A coupled nonlinear Schrödirnger equation and optical solitons, J Phys Soc Jpn 61 (1992), 2241–2245.
5. M. Asadzadeh, D. Rostamy, and F. Zabihi, A posteriori error estimates for a coupled wave system with a local damping, J Math Sci 175, 228–248.
6. M. Asadzadeh, Streamline diffusion methods for the Vlasov-Poisson equations, RAIRO Math Mod Numer Anal 24 (1990), 177–196.
7. M. Asadzadeh, Streamline diffusion methods for the Fermi and Fokker-Planck equations, Trans Theory Stat Phys 26 (1997), 319–340.
8. K. Eriksson, C. Johnson, and V. Thomee, Time discretization of parabolic problems by the Discontinuous Galerkin method, RAIRO, Anal Numer 19 (1985), 611–643.
9. C. Johnson and J. Pitkaranta, An analysis of the discontinuous Galerkin method for a scalar hyperbolic equation, Math Comput 46 (1986), 1–26.
10. T. J. R. Hughes and A. Brooks, A multidimensional upwind scheme with no crosswind diffusion, T. J. R. Hughes, editor, Finite element methods for convection dominated flows, AMD Vol. 34, ASME, New York, 1979.
11. T. J. R. Hughes and M. Mallet, A new finite element formulation for computational fluid dynamics III, the generalized streamline operator for multidimensional advective—diffusive systems, Comput Meth Appl Mech Eng 58 (1986), 305–328.
12. 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.

13. M. S. Ismail, Numerical solution of coupled nonlinear Schrödinger equation by Galerkin method, Math Comput Simul 78 (2008), 532–547.

14. M. S. Ismail and S. Z. Alamri, Highly accurate finite difference method for coupled nonlinear Schrödinger equation, Int J Comput Math 81 (2004), 333–351.

15. M. S. Ismail and T. Taha, A finite element solution for the coupled Schrödinger equation, proceedings of the 16th IMACS World Congress on Scientific Computation, Lausanne, 2000.

16. Y. Xu and C. Shu, Local discontinuous Galerkin method for nonlinear Schrödinger equations, J Comput Phys 205 (2005), 72–97.

17. M. Asadzadeh and A. Sopasakis, Convergence of a hp streamline diffusion scheme for Vlasov-Fokker-Planck system, Math Mod Meth Appl Sci 17 (2007), 1159–1182.

18. M. G. Larson and A. Målqvist, Adaptive variational multiscale methods based on a posteriori error estimation, Energy norm estimates for elliptic problems, Comput Meth Appl Mechan Eng 196 (2007), 2313–2324.

19. M. Asadzadeh and E. Kazemi, On convergence of h and hp finite element methods for three dimensional Fermi equation, Numer Anal Mod, to appear.

20. S. C. Brenner and L. R. Scott, The mathematical theory of finite element method, Springer-Verlag, New York, 1994.

21. P. G. Ciarlet, The finite element method for elliptic problems, Amsterdam, North Holland, 1987.

22. Y. Efendiev, T. Y. Hou, and X. H. Wu, Convergence of a nonconforming multiscale finite element method, SIAM J Numer Anal 37 (2000), 888–910.

23. P. Bochev, T. J. R. Hughes, and G. Scovazzi, A multiscale discontinuous Galerkin method, Lect Note Comput Sci 3743 (2006), 84–93.

24. R. Söderlind, Finite element methods for multiscale/multiphysics problems, PhD Thesis, Department of Mathematics and Mathematical Statistics, Umeå University, Sweden, 2011.

25. R. Becerrila, F. S. Guzmana, A. Rendon-Romerob, and S. Valdez-Alvarado, Solving the time-dependent Schrödinger equation using finite difference methods, Rev Mex De Fisica E 54 (2008), 120–132.

26. H. Han, J. Jin, and X. Wu, A finite-difference method for the one-dimensional time-dependent schrödinger equation on unbounded domain, Comput Math Appl 50 (2005), 1345–1362.

27. V. Thomee, Finite difference methods for linear parabolic equations, P. G. Ciarlet and J. L. Lions, editors, Handbook of numerical analysis I, Part 1, Amsterdam, North-Holland, 1990, pp. 3–194.

28. O. C. Zienkiewicz and R. L. Taylor, The finite element method. Vol. 1. The basis., 5th Ed., Butterworth-Heinemann, Oxford, 2000.

29. B. Riviére, Discontinuous Galerkin methods for solving elliptic and parabolic equations, Theory and Implementation, SIAM, 2008.

30. R. A. Adams, Sobolev spaces, Academic Press, New York, 1975.

31. K. Eriksson and C. Johnson, Adaptive finite element methods for parabolic problems I: a linear model problem, SIAM J Numer Anal 28 (1991), 43–77.

32. 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), 706–740.

33. K. Eriksson and C. Johnson, Adaptive finite element methods for parabolic problems IV: a nonlinear problem, SIAM J Numer Anal 32 (1995), 1729–1749.

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

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

36. C. Johnson, Discontinuous Galerkin finite element methods for second order hyperbolic problems, Comput Meth Appl Mech Eng 107 (1993), 117–129.

37. 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 (1974), 89–123.

38. C. A. Raposo and W. D. Bastos, Energy decay for the solutions of a coupled wave system, TEMA Tend Math Appl Comput 10 (2009), 203–209.

39. D. Rostamy and F. Zabihi, A streamline diffusion method for the mass-spring system, J Numer Math Stoch 2 (2010), 76–89.