Convergence rates in homogenization of higher order parabolic systems

Weisheng Niu * Yao Xu

Abstract

This paper is concerned with the optimal convergence rate in homogenization of higher order parabolic systems with bounded measurable, rapidly oscillating periodic coefficients. The sharp $O(\varepsilon)$ convergence rate in the space $L^2(0, T; H^{m-1}(\Omega))$ is obtained for both the initial-Dirichlet problem and the initial-Neumann problem. The duality argument inspired by [25] is used here.

1 Introduction

We consider the sharp convergence rate in periodic homogenization of initial value problems

\[
\begin{aligned}
    \partial_t u_\varepsilon + L_\varepsilon u_\varepsilon &= f & \text{in } \Omega \times (0, T), \\
    u_\varepsilon &= h & \text{on } \Omega \times \{t = 0\},
\end{aligned}
\]  

with homogeneous Dirichlet and Neumann boundary data. The linear operator $L_\varepsilon$ is defined as

\[
L_\varepsilon = (-1)^m \sum_{|\alpha|=|\beta|=m} D^\alpha (A^{\alpha\beta}(x/\varepsilon, t/\varepsilon^{2m}) D^{\beta}),
\]

where $\alpha, \beta$ are $d$-dimensional multi-indices with components $\alpha_k, \beta_k, k = 1, 2, ..., d$, and

\[
|\alpha| = \sum_{k=1}^d \alpha_k, \quad D^\alpha = D_{x_1}^{\alpha_1}D_{x_2}^{\alpha_2} \cdots D_{x_d}^{\alpha_d}.
\]

The coefficients matrix $A(y, s) = (A_{ij}^{\alpha\beta}(y, s)), 1 \leq i, j \leq n$, is real, bounded measurable with

\[
\sum_{|\alpha|=|\beta|=m} A^{\alpha\beta}(y, s)\xi_\alpha \xi_\beta \geq \mu |\xi|^2, \quad |A(y, s)| \leq \frac{1}{\mu} \quad \text{for a.e. } (y, s) \in \mathbb{R}^{d+1},
\]

where $\mu > 0$, $\xi = (\xi_\alpha)_{|\alpha|=m}$, $\xi_\alpha = (\xi^1_\alpha, ..., \xi^n_\alpha) \in \mathbb{R}^n$. We also assume that $A$ is 1-periodic in $(y, s)$, that is,

\[
A(y + z, s + t) = A(y, s) \quad \text{for any } (z, t) \in \mathbb{Z}^{d+1} \text{ and a.e. } (y, s) \in \mathbb{R}^{d+1}.
\]

*Supported partly by the NSF of China (11701002,11301003) and NSF of Anhui Province (1708085MA02).
Let $H^n_0(\Omega)$ be the conventional $\mathbb{R}^n$-valued Sobolev spaces with dual $H^{-m}(\Omega)$. For $0 < T < \infty$, it is known that under the uniform parabolic condition \((1.3)\), for any $f \in L^2(0; T; H^{-m}(\Omega))$ and any $h \in L^2(\Omega)$, initial value problem \((1.1)\) with homogeneous Dirichlet data is given by

\[
\begin{aligned}
-\int_{\Omega_T} u_\varepsilon \partial_t \phi dx dt + \sum_{|\alpha|=|\beta|=m} \int_{\Omega_T} A^0_{ij}(x/\varepsilon, t/\varepsilon^{2m}) D^\beta u_{\varepsilon j} D^\alpha \phi_i dx dt \\
= \int_0^T \langle f(t), \phi(t) \rangle_{H^{-m}(\Omega) \times H^m_0(\Omega)} dt + \int_{\Omega} h \phi(0) dx
\end{aligned}
\]

for any $\phi \in C_c^\infty(\Omega \times [0, T))$.

As we shall prove in Section 2, under periodicity condition \((1.4)\) the homogenized problem of \((1.1)\) with homogeneous Dirichlet boundary data is given by

\[
\begin{aligned}
\partial_t u_0 + \mathcal{L}_0 u_0 &= f & \text{in } \Omega \times (0, T), \\
\text{Tr}(D^\gamma u_0) &= 0 & \text{on } \partial \Omega \times (0, T), 0 \leq |\gamma| \leq m - 1, \\
u_0 &= h & \text{on } \Omega \times \{t = 0\},
\end{aligned}
\]

(1.5)

where

\[
\mathcal{L}_0 = (-1)^m \sum_{|\alpha|=|\beta|=m} D^\alpha (\hat{A}^\alpha \cdot D^\beta),
\]

\[
\hat{A}^\alpha_{ij} = \sum_{|\gamma|=m} \int_Y \left[ A_{ij}^\alpha(y, s) + A_{i\ell}^\alpha(y, s) D^\gamma \chi_{\ell j}^\beta(y, s) \right] dy ds,
\]

(1.6)

with $Y = [-1/2, 1/2]^{d+1}$ and $\chi = (\chi_{ij}^\gamma)$ being the matrix of correctors for the operator $\partial_t + \mathcal{L}_\varepsilon$.

Moreover, as $\varepsilon$ tends to zero $u_\varepsilon$ converges strongly to $u_0$ in $L^2(0, T; H^{m-1}_0(\Omega)).$

Our first objective is to derive the optimal convergence rate of $u_\varepsilon$ to $u_0$.

**Theorem 1.1.** Let $\Omega$ be a bounded $C^{m1}$ domain in $\mathbb{R}^d$ and $0 < T < \infty$. Assume that $A$ satisfies \((1.3)\) and \((1.4)\), $f \in L^2(0; T; H^{-m+1}(\Omega))$, $h \in L^2(\Omega)$. Let $u_\varepsilon, u_0 \in L^2(0, T; H^m_0(\Omega))$ be the weak solutions to initial-Dirichlet problems \((1.1)\) and \((1.5)\), respectively. If in addition $u_0 \in L^2(0, T; H^{m+1}(\Omega))$, then

\[
\|u_\varepsilon - u_0\|_{L^2(0, T; H^{m-1}_0(\Omega))} \leq C\varepsilon \left\{ \|u_0\|_{L^2(0, T; H^{m+1}(\Omega))} + \|f\|_{L^2(0, T; H^{-m+1}(\Omega))} + \|h\|_{L^2(\Omega)} \right\},
\]

(1.7)

where $C$ depends only on $d, n, m, \mu, T$ and $\Omega$.

The second objective of this paper is to obtain the sharp convergence rate in the homogenization of \((1.1)\) with homogeneous Neumann boundary data. Let $H^m(\Omega)$ be the conventional $\mathbb{R}^n$-valued Sobolev spaces with dual $\hat{H}^{-m}(\Omega)$ and let $0 < T < \infty$. Write out the complete variational form of \((1.1)\) with boundary term formally. Assume that it possesses the following form

\[
-\int_{\Omega_T} u_\varepsilon \partial_t \phi dx dt + \sum_{|\alpha|=|\beta|=m} \int_{\Omega_T} A^0_{ij}(x/\varepsilon, t/\varepsilon^{2m}) D^\beta u_{\varepsilon j}(x, t) D^\alpha \phi_i(x, t) dx dt
\]

2
for any $\phi \in C^\infty(\Omega \times [0, T])$ with $\phi(T) = 0$, where $\partial_\nu$ is the derivative along the unit outward normal $\nu$ of $\Omega$. By homogeneous Neumann boundary data, we mean that, formally,

$$
N_{m-1-j}u_\varepsilon = 0 \quad \text{on } \partial \Omega \times (0, T), \quad j = 0, 1, \ldots, m - 1. \tag{1.9}
$$

For $f \in L^2(0, T; \tilde{H}^{-m}(\Omega)), h \in L^2(\Omega)$, problem (1.1) with homogeneous Neumann boundary data admits a unique weak solution $u_\varepsilon \in L^2(0, T; H^m(\Omega)) \cap L^\infty(0, T; L^2(\Omega))$ [5].

Similar to the initial-Dirichlet problem, under the periodicity condition [14] $u_\varepsilon$ converges strongly in $L^2(0, T; H^{m-1}(\Omega))$ to the solution $u_0$ of the following initial-Neumann problem

$$
\begin{aligned}
&\frac{\partial}{\partial t}u_0 + \mathcal{L}_0 u_0 = f \quad \text{in } \Omega \times (0, T), \\
&N_{m-1-j}(u_0) = 0 \quad \text{on } \partial \Omega \times (0, T), \\
&u_0 = h \quad \text{on } \Omega \times \{t = 0\},
\end{aligned} \tag{1.10}
$$

where $\mathcal{L}_0$ is defined as in (1.6) (see Theorem 2.2). Parallel to Theorem 1.1, we have the following sharp convergence rate in homogenization of the initial-Neumann problem.

**Theorem 1.2.** Let $\Omega$ be a bounded $C^{m,1}$ domain in $\mathbb{R}^d$ and $0 < T < \infty$. Assume that $A$ satisfies (L.3) and (L.4), and $f \in L^2(0, T; \tilde{H}^{-m+1}(\Omega)), h \in L^2(\Omega)$. Let $u_\varepsilon, u_0 \in L^2(0, T; H^m(\Omega))$ be, respectively, the weak solutions to the initial-Neumann problems (1.1) and (1.10). If in addition $u_0 \in L^2(0, T; H^{m+1}(\Omega))$, then

$$
\|u_\varepsilon - u_0\|_{L^2(0, T; H^{m-1}(\Omega))} \leq C \varepsilon \left\|\|u_0\|_{L^2(0, T; H^{m+1}(\Omega))} + \|f\|_{L^2(0, T; \tilde{H}^{-m+1}(\Omega))} + \|h\|_{L^2(\Omega)}\right\}, \tag{1.11}
$$

where $C$ depends only on $d, n, m, \mu, T$ and $\Omega$.

The proof of Theorem 1.1 is mainly based on the duality argument initiated in [25]. To adapt the ideas, we first provide the existence results for the matrix of correctors $\chi(y, s)$ and flux correctors $\mathcal{B}(y, s)$ (also referred as dual correctors) for operators $\partial_\nu + \mathcal{L}_\varepsilon (\varepsilon > 0)$ in Section 2. Recall that flux correctors play an essential role in the investigation on sharp convergence rate in the homogenization of second order elliptic or parabolic systems [10, 11, 12, 21, 7]. In [7], the flux correctors are obtained by considering a harmonic system with periodic boundary conditions in $\mathbb{R}^{d+1}$ (see Lemma 2.1 therein), which is a modification of the approach for second order elliptic systems. The process however seems not applicable to higher order parabolic systems. Indeed, following the process we will obtain a degenerate elliptic system in $\mathbb{R}^{d+1}$, which is hard to cope with. Instead, we will modify the construction of flux correctors for elliptic systems in another manner to construct the flux correctors for high order parabolic systems, see Lemmas 2.1 and 2.2. This approach also provides us further regularity results on the flux correctors (see (2.5)2). As we can see from the definition of $w_\varepsilon$ in (1.12), the higher regularity ($H^{2m-1}$) on $\mathcal{B}^{\gamma(d+1)\beta}$ (or $\mathcal{B}^{\gamma(d+1)\beta}$) are essential, which however is trivial for second order parabolic systems ($m = 1$) [7].

Since we consider the systems with coarse coefficients, the correctors $\chi(y, s)$ and the flux correctors $\mathcal{B}(y, s)$ may be unbounded. Therefore, similar to [7], in Section 3 we introduce the smoothing
operator $S_\varepsilon$ with respect to the space and time variables $x, t$ and establish proper estimates for the smoothing operator. However, to deal with the higher order operators, more general estimates are proved by using an approach quite different from [7].

With preparations in Sections 2 and 3, in Section 4 we introduce the function

$$w_\varepsilon(x, t) = u_\varepsilon(x, t) - u_0(x, t) - \varepsilon^m \sum_{|\gamma| = m} \chi^\gamma(x/\varepsilon, t/\varepsilon^{2m})S_\varepsilon^2(D^\gamma u_0)\rho_\varepsilon(x)\varrho_\varepsilon(t)$$

$$+ (-1)^{m+1} \sum_{|\beta|=|\gamma|=m} \varepsilon^{m+|\eta|} D^{\eta} \mathcal{B}\gamma(d+1) (x/\varepsilon, t/\varepsilon^{2m})D^{\eta} \left[S_\varepsilon^2(D^\gamma u_0)\rho_\varepsilon(x)\varrho_\varepsilon(t)\right]$$

$$+ (-1)^m \varepsilon^{2m} \sum_{|\alpha|=|\beta|=m} \mathcal{B}^{\alpha\beta}(x/\varepsilon, t/\varepsilon^{2m})D^{\alpha} \left[S_\varepsilon^2(D^\gamma u_0)\rho_\varepsilon(x)\varrho_\varepsilon(t)\right],$$

(1.12)

where $\rho_\varepsilon, \varrho_\varepsilon$ are proper cut-off functions, $S_\varepsilon^2 = S_\varepsilon \circ S_\varepsilon$, see (4.4) for the details. Then through some delicate analysis and proper use of preparations aforementioned, we prove the following $O\sqrt{\varepsilon}$ estimate in $L^2(0, T; H^{m-1}(\Omega))$,

$$\|\nabla^m w_\varepsilon\|_{L^2(\Omega)} \leq C\varepsilon^{1/2} \left\{ ||u_0||_{L^2(0, T; H^{m+1}(\Omega))} + ||f||_{L^2(0, T; H^{-m+1}(\Omega))} + ||h||_{L^2(\Omega)} \right\}.$$  

(1.13)

The above estimate should be comparable to (3.17) or (3.20) in [7]. Yet, we point out that the auxiliary function $w_\varepsilon$ is much more complicated than that for second order parabolic systems constructed in [7]. And compared to the proof of (3.17) in [7], the proof of (1.13) needs more delicate analysis. Whence (1.13) is obtained, the desired estimate (1.7) follows directly by the duality argument motivated by [20, 27]. The proof of Theorem 1.2 is completely parallel, and is sketched in Section 5.

As the end of the introduction, let us provide a brief review on the background of convergence rates in quantitative homogenization, which is currently a quite active area of research. Sharp convergence rates for second order elliptic equations (systems) has been studied extensively in various circumstances in the past years. To name but a few, in [8, 23, 24] the optimal $O(\varepsilon)$ convergence rate was obtained for second order elliptic equations with Dirichlet or Neumann boundary conditions in $C^{1,1}$ domains. In [11, 21], the optimal $O(\varepsilon)$ and suboptimal convergence rates (like $O(\varepsilon \ln^{1/2})$) were derived for second order elliptic systems with Dirichlet or Neumann boundary conditions in Lipschitz domains. See also [2, 13, 9, 22, 23] and references therein for more related results.

For second order parabolic equations with time-independent coefficients, the sharp convergence rate has also been studied widely, see [11, 24] for the Cauchy problems on the whole space, and [14, 16] for the initial boundary value problems in $C^{1,1}$ cylinders. Quantitative estimates for parabolic equations with time dependent coefficients are a bit more intricate and little progress was made until very recently. In [6, 7, 4, 29, 1] the optimal $O(\varepsilon)$ convergence rate in $L^2(0, T; L^2(\Omega))$ was obtained in homogenization of second order parabolic systems in $C^{1,1}$ cylinders, while in [29] the suboptimal $O(\varepsilon \ln(1/\varepsilon))$ convergence rate in $L^2(0, T; L^2(\Omega))$ was obtained for parabolic systems of elasticity in Lipschitz cylinders. More recently, in [11] the convergence rate and uniform regularity estimates in homogenization of second order stochastic parabolic equations were deeply studied. See also [6, 4] for more results on the uniform regularity estimates in the periodic setting.
Homogenization of higher order elliptic equations arises in the study of linear elasticity [3,10,20], for which the qualitative results have been obtained for many years [3,10]. Few quantitative results were known in the homogenization of higher order elliptic or parabolic equations until very recently. In [19,20,14], the optimal $O(\varepsilon)$ convergence rate was established in homogenization of higher-order elliptic equations in the whole space. In [27,28], some $O(\varepsilon)$ two-parameter resolvent estimates were obtained for more general higher order elliptic systems with homogeneous Dirichlet or Neumann boundary data in bounded $C^{2m}$ domains. Shortly, the sharp convergence rate and uniform regularity estimates in the homogenization of higher order elliptic systems with symmetric or nonsymmetric coefficients were further studied in [17,18], see also [30] for the results in the almost-periodic setting.

As far as we know, quantitative estimates in homogenization of higher order parabolic equations have not been studied, especially for those with time dependent coefficients. The present paper seems to be the first attempt in this direction. Our results in Theorems 1.1 and 1.2 extend the convergence results for higher order elliptic equations in [19,20,27,28] to parabolic systems on the one hand, and on the other hand they extend the results for second order parabolic systems in [7] to higher order parabolic systems.

## 2 Qualitative homogenization

### 2.1 Correctors and flux correctors

Set $Y = [-\frac{1}{2},\frac{1}{2}]^{d+1}$. For $1 \leq i,j \leq n$ and $d$-dimensional multi-index $\gamma$ of degree $m$, i.e. $|\gamma| = m$, we introduce the matrix of correctors $\chi = (\chi^\gamma_j(y,s)) = (\chi^\gamma_{ij}(y,s))$ for the family of operators $\partial_s + \mathcal{L}(\varepsilon > 0)$, given by the following cell problem in $Y$,

$$
\frac{\partial(\chi^\gamma_{ij})}{\partial s} + (-1)^m \sum_{|\alpha|=|\beta|=m} D^\alpha \left\{ A^\alpha_{ik}(y,s) D^\beta \chi^\gamma_{kj} \right\} = (-1)^{m+1} \sum_{|\alpha|=m} D^\alpha A^\alpha_{ij}(y,s) \quad \text{in } Y, \tag{2.1}
$$

where $\chi^\gamma_j(y,s)$ is 1-periodic in $(y,s)$, and $\int_Y \chi^\gamma_j(y,s) dy ds = 0$, $1 \leq j \leq n$,

\[\chi^\gamma_j(y,s)\] is obtained for each fixed $\gamma$ and $j$. Under conditions (1.3) and (1.4), the existence of $\chi^\gamma_j$ follows from standard existence results of general parabolic systems.

For $1 \leq i,j \leq n$ and $d$-dimensional multi-indices $\alpha, \beta$ of degree $m$, we set

$$
A^\alpha_{ij} = \int_Y \left\{ A^\alpha_{ij}(y,s) + \sum_{|\gamma|=m} A^\alpha_{ik}(y,s) D^\gamma \chi^\beta_{kj}(y,s) \right\} dy ds
$$

$$
= \sum_{|\gamma|=m} \int_Y A^\gamma(y,s) D^\gamma \left( \frac{1}{\beta!} y^\beta e_j + \chi^\beta_j(y,s) \right) \cdot D^\gamma \left( \frac{1}{\gamma!} y^\gamma e_i \right) dy ds, \tag{2.2}
$$

where $y^\alpha = y_1^{a_1} y_2^{a_2} \cdots y_d^{a_d}$.

In this section, the symbols $i,j$ may equal $(d+1)$ or represent $d$-dimensional multi-indices of degree $m$. We define $B^\alpha_{ij}$ for $|\beta| = m$ by

$$
B^\alpha_{ij} = \begin{cases} 
A^\alpha_{ij} + \sum_{|\gamma|=m} A^\alpha_{ik}(y,s) D^\gamma \chi^\beta_{kj} - \tilde{A}^\alpha_{ij}, & \text{if } i = \alpha, \\
(-1)^m \chi^\alpha_{ij}, & \text{if } i = d+1, \tag{2.3}
\end{cases}
$$

[...remaining text...]
where $\alpha$ is a $d$-dimensional multi-index of degree $m$. Also, we define

$$\hat{u}(s) = \int_{Y_d} u(y, s) dy,$$

for 1-periodic function $u(y, s)$ in $\mathbb{R}^{d+1}$, where $Y_d = [-1/2, 1/2]^d$.

The following lemma gives the existence of the matrix of flux correctors for the family of operators $\partial_s + L_\varepsilon(\varepsilon > 0)$.

**Lemma 2.1.** For any $1 \leq i, j \leq n$ and $d$-dimensional multi-index $\beta$ with $|\beta| = m$, there exist 1-periodic functions $\mathcal{B}^{ij}_\beta(y, s)$ in $\mathbb{R}^{d+1}$ such that

$$\mathcal{B}^{ij}_\beta = -\mathcal{B}^{ji}_\beta, \quad \sum_{|\gamma| = m} D^\gamma \mathcal{B}^{ij\gamma}(y, s) + \partial_s \mathcal{B}^{(d+1)\beta}(y, s) = B^{ij\beta}(y, s) - \hat{B}^{ij\beta}(s).$$

Furthermore, there exists a constant $C$ depending only on $d, n, m, \mu$ such that

$$\|\mathcal{B}^{ij\beta}\|_{L^2(-1/2, 1/2, H^m(Y_d))} \leq C \quad \text{if } i, j \text{ are } d\text{-dimensional multi-indices of degree } m,$$

$$\|\mathcal{B}^{ij\beta}\|_{L^2(-1/2, 1/2, H^{2m}(Y_d))} \leq C \quad \text{if } i \text{ or } j = d + 1. \quad (2.5)$$

**Proof.** For simplicity of presentations, let us suppress the subscripts $i, j$. Since for $|\beta| = m$, $B^{i\beta}(y, s)$ are 1-periodic in $\mathbb{R}^{d+1}$, and for any $s \in \mathbb{R}$,

$$\int_{Y_d} [B^{i\beta}(y, s) - \hat{B}^{i\beta}(s)] dy = 0,$$

there exist $f^{i\beta}(\cdot, s) \in H^{2m}(Y_d)$ such that

$$\begin{cases} \Delta_d f^{i\beta}(y, s) = B^{i\beta}(y, s) - \hat{B}^{i\beta}(s) \quad \text{in } \mathbb{R}^d, \\ f^{i\beta}(\cdot, s) \text{ is } 1\text{-periodic} \quad \text{in } \mathbb{R}^d, \end{cases}$$

where $\Delta_d$ denotes the Laplacian in $\mathbb{R}^d$. We define for $|\beta| = m$,

$$\mathcal{B}^{ij\beta} = \begin{cases} D^\gamma f^{i\alpha - D^\alpha f^{i\beta}}, & \text{if } i = \gamma, j = \alpha, \\ D^\alpha f^{(d+1)\beta}_d, & \text{if } i = \alpha, j = d + 1, \\ -D^\alpha f^{(d+1)\beta}_d, & \text{if } i = d + 1, j = \alpha, \\ 0, & \text{if } i = j = d + 1, \end{cases}$$

where $\alpha, \gamma$ are $d$-dimensional multi-indices of degree $m$. Obviously, $\mathcal{B}^{ij\beta} = -\mathcal{B}^{ji\beta}$. Since

$$\sum_{|\alpha| = m} D^\alpha B^{i\alpha\beta}(y, s) + \partial_s B^{(d+1)\beta}(y, s) = 0, \quad (2.6)$$

we have

$$\Delta_d^{m} \left[ \sum_{|\alpha| = m} D^\alpha f^{i\alpha\beta}(y, s) + \partial_s f^{(d+1)\beta}(y, s) \right] = 0.$$
By the Liouville property for $\Delta_d^\alpha$ and the periodicity of $f$, we know that $\sum_{|\alpha|=m} D^\alpha f^\alpha(\cdot, s) + \partial_s f^{(d+1)\beta}(\cdot, s)$ is a constant. Consequently, it’s not hard to verify that

$$\sum_{|\gamma|=m} D^\gamma B^\gamma(y, s) + \partial_s B^{(d+1)\beta}(y, s) = B^\beta(y, s) - \hat{B}^\beta(s).$$

Moreover, note that for any $|\alpha| = |\beta| = |\gamma| = m$,

$$\|B^\gamma y^\beta(s)\|_{H^m(Y_d)} \leq C(\|f^\alpha(s)\|_{H^2m(Y_d)} + \|f^\gamma(s)\|_{H^{2m}(Y_d)}) \leq C(\|B^\alpha(s)\|_{L^2(Y_d)} + \|B^\gamma(s)\|_{L^2(Y_d)}),$$

which implies the first estimate in (2.5). Similarly, the second part of (2.5) follows from setting $\gamma = 0$.

Moreover, note that for any $|\alpha| = |\beta| = m$,

$$\|B^\alpha\|_{H^1([-1/2, 1/2])} \leq C\|\hat{B}^\alpha\|_{L^2([-1/2, 1/2])}$$

for some positive constant $C$.

Proof. Integrating equation (2.6) over $Y_d$, we get $\partial_s \hat{B}^{(d+1)\beta}(s) = 0$, which together with the fact $\int_{-1/2}^{1/2} \hat{B}^{(d+1)\beta}(s) ds = 0$ implies that $\hat{B}^{(d+1)\beta}(s) \equiv 0$. Moreover, note that $\int_{-1/2}^{1/2} \hat{B}^\alpha(s) ds = 0$. By setting $B^\alpha(s) = \int_0^s \hat{B}^\alpha(s) ds$, we get the desired function. The proof is complete. \qed

Lemma 2.2. Let $B^\beta$ be defined as in (2.3). Then $\hat{B}^{(d+1)\beta}(s) \equiv 0$, and moreover for $|\alpha| = |\beta| = m$, there exist 1-periodic functions $\mathcal{B}^{\alpha\beta}(s)$ in $\mathbb{R}$ such that $\partial_s \mathcal{B}^{\alpha\beta}(s) = \hat{B}^{\alpha\beta}(s)$ and

$$\|\mathcal{B}^{\alpha\beta}\|_{H^1([-1/2, 1/2])} \leq C\|B^\alpha\|_{L^2([-1/2, 1/2])}$$

for some positive constant $C$.

Proof. Integrating equation (2.6) over $Y_d$, we get $\partial_s \hat{B}^{(d+1)\beta}(s) = 0$, which together with the fact $\int_{-1/2}^{1/2} \hat{B}^{(d+1)\beta}(s) ds = 0$ implies that $\hat{B}^{(d+1)\beta}(s) \equiv 0$. Moreover, note that $\int_{-1/2}^{1/2} \hat{B}^\alpha(s) ds = 0$. By setting $\mathcal{B}^{\alpha\beta}(s) = \int_0^s \hat{B}^{\alpha\beta}(s) ds$, we get the desired function. The proof is complete. \qed

Let $\mathcal{L}_\epsilon^* = \mathcal{L}_\epsilon$, i.e.,

$$\mathcal{L}_\epsilon^* = (-1)^m \sum_{|\alpha|=|\beta|=m} D^\alpha \left( A^{\alpha\beta}(x/\epsilon, t/\epsilon^{2m}) D^\beta \right),$$

where $A^* = A^{*\alpha\beta} = (A_{ij}^*)$. Parallel to (2.1), we can introduce the matrix of correctors $\chi^* = (\chi^*_{ij})$ for the operator $-\partial_s + \mathcal{L}_\epsilon^*$, where $(\chi^*_{ij}) = (\chi^*_{11}(y), \chi^*_{22}(y), ..., \chi^*_{nn}(y))$ is the solution to the following cell problem,

$$-\frac{\partial(\chi^*_{ij})}{\partial s} + (-1)^m \sum_{|\gamma|=|\beta|=m} D^\gamma \left\{ A^{*\gamma\beta}_{ij} (y, s) \right\} = (-1)^{m+1} \sum_{|\eta|=m} D^\eta A^{*\eta\alpha}_{i\ell} (y, s) \text{ in } Y,$$

$$\chi^*_{ij}(y, s) \text{ is 1-periodic in } (y, s), \text{ and } \int_Y \chi^*_{ij}(y, s) dy ds = 0, \quad 1 \leq i \leq n.$$
\[ \int_{-1/2}^{1/2} \left( \frac{\partial (\chi_j^\alpha)}{\partial s}, \chi_j^\gamma \right) ds = \sum_{|\eta|=|\beta|=m} \int_Y \left\{ A^{\eta\beta} D^\eta \chi_j^\alpha D^\gamma \chi_j^\gamma + A^{\eta\beta} D^{\eta} \left( \frac{y^\alpha}{\alpha!} e_i \right) D^\gamma \chi_j^\gamma \right\} dy ds \]

which by summation implies that

\[ \sum_{|\eta|=|\beta|=m} \int_Y A^{\eta\beta}(y, s) D^\eta \chi_j^\gamma (y, s) D^{\gamma} \left( \frac{y^\alpha}{\alpha!} e_i \right) dy ds \]

\[ = \sum_{|\eta|=|\beta|=m} \int_Y A^{\eta\beta}(y, s) D^{\eta} \left( \frac{y^\gamma}{\gamma!} e_j \right) D^\gamma \chi_j^\alpha (y, s) dy ds \]

\[ = \sum_{|\eta|=|\beta|=m} \int_Y A^{\eta\beta}(y, s) D^\eta \chi_j^\alpha (y, s) D^{\gamma} \left( \frac{y^\gamma}{\gamma!} e_j \right) dy ds. \]

In view of (2.2), this provides another expression of \( \tilde{A} \) in terms of \( A^* \) and \( \chi^* \),

\[ A_{ij}^{\alpha \gamma} = \sum_{|\eta|=|\beta|=m} \int_Y A^{\eta\beta}(y, s) D^\eta \left( \chi_j^\gamma (y, s) + \frac{y^\gamma}{\gamma!} e_j \right) D^{\gamma} \left( \frac{y^\alpha}{\alpha!} e_i \right) dy ds \]

\[ = \sum_{|\eta|=|\beta|=m} \int_Y A^{\eta\beta}(y, s) D^\eta \chi_j^\alpha (y, s) D^{\gamma} \left( \frac{y^\gamma}{\gamma!} e_j \right) dy ds. \quad (2.9) \]

### 2.2 Effective operators and homogenized systems

In this part, we prove that the effective operator for \( \partial_t + L \), is \( \partial_t + L_0 \), where \( L_0 \) is defined as in (1.6). Let \( \tilde{A} = (A_{ij}^{\alpha \beta}) \) be defined as in (2.2). In view of (2.1), we have

\[ \tilde{A}_{ij}^{\alpha \beta} = \int_Y \sum_{|\zeta|=|\eta|=m} A^{\eta\zeta} D^\eta \left( \frac{y^\beta}{\beta!} e_j + \chi_j^\beta \right) \cdot D^{\zeta} \left( \frac{y^\alpha}{\alpha!} e_i + \chi_i^\alpha \right) dy ds + \int_{-1/2}^{1/2} \langle \partial_s \chi_j^\beta, \chi_i^\alpha \rangle ds. \]

Therefore,

\[ \tilde{A}_{ij}^{\alpha \beta} \xi_{\alpha}^\beta \xi_{\beta}^j = \sum_{|\zeta|=|\eta|=m} \int_{-1/2}^{1/2} A^{\eta\zeta} D^\eta \left( \frac{y^\beta}{\beta!} e_j \xi_{\beta}^j + \chi_j^\beta \xi_{\beta}^j \right) \cdot D^{\zeta} \left( \frac{y^\alpha}{\alpha!} e_i \xi_{\alpha}^j + \chi_i^\alpha \xi_{\alpha}^i \right) dy ds \]

\[ + \int_{-1/2}^{1/2} \langle \partial_s \chi_j^\beta \xi_{\beta}^j, \chi_i^\alpha \xi_{\alpha}^i \rangle ds, \]

for any \( \xi = (\xi_{\alpha})_{|\alpha|=m} \) with \( \xi_{\alpha} = (\xi_{\alpha}^j) \in \mathbb{R}^n \). Note that

\[ \sum_{|\alpha|=|\beta|=m} \int_{-1/2}^{1/2} \langle \partial_s \chi_j^\beta \xi_{\beta}^j, \chi_i^\alpha \xi_{\alpha}^i \rangle ds = \frac{1}{2} \sum_{|\alpha|=|\beta|=m} \int_{-1/2}^{1/2} \partial_s \langle \chi_j^\beta \xi_{\beta}^j, \chi_i^\alpha \xi_{\alpha}^i \rangle ds = 0. \]
Using (1.3) and integration by parts we obtain that,

$$
\sum_{|\alpha|=|\beta|=m} A_{ij}^{\alpha \beta} \xi_i^{\alpha} \xi_j^{\beta} \geq \mu \sum_{|\alpha|=m} \int_Y D^{\alpha} \left\{ \sum_{|\beta|=m} \frac{\chi^\beta_j(y,s)}{\beta!} \xi_j^{\beta} \right\} \cdot \sum_{|\alpha|=m} D^{\alpha} \left\{ \sum_{|\beta|=m} \chi^\alpha_i(y,s) \xi_i^{\alpha} \right\}
$$

$$
= \mu \sum_{|\alpha|=m} |\xi^{\alpha}|^2 + \mu \sum_{|\alpha|=m} \int_Y D^{\alpha} \left\{ \sum_{|\beta|=m} \chi^\beta_j(y,s) \xi_j^{\beta} \right\} D^{\alpha} \left\{ \sum_{|\alpha|=m} \chi^\alpha_i(y,s) \xi_i^{\alpha} \right\}
$$

$$
\geq \mu \sum_{|\alpha|=m} |\xi^{\alpha}|^2,
$$

(2.10)

which, combined with (2.2), implies that $\bar{A}$ satisfies the ellipticity condition (1.3) with $1/\mu$ replaced by some constant $\mu_0$, depending only on $d, n, m$ and $\mu$. Therefore, for any $f \in L^2(0, T; H^{-m}(\Omega)), h \in L^2(\Omega)$, problem (1.5) admits a unique solution $u_0 \in L^2(0, T; H^0_m(\Omega)) \cap L^\infty(0, T; L^2(\Omega))$.

**Theorem 2.1.** Let $\Omega$ be a bounded Lipschitz domain in $\mathbb{R}^d$. Assume that $A$ satisfies conditions (1.3) – (1.4), and $f \in L^2(0, T; H^{-m}(\Omega)), h \in L^2(\Omega)$. Let $u_\varepsilon, u_0 \in L^2(0, T; H^0_m(\Omega)) \cap L^\infty(0, T; L^2(\Omega))$ be the unique weak solutions to initial-Dirichlet problems (1.1) and (1.5), respectively. Then as $\varepsilon \to 0$,

$$
u_\varepsilon \to u_0 \text{ weakly in } L^2(0, T; H^0_m(\Omega)) \text{ and strongly in } L^2(0, T; H^{m-1}(\Omega)).
$$

(2.11)

**Proof.** The proof is adapted from Theorem 2.1 in [3, p.140], where similar results was proved for second order parabolic equations. Note that $u_\varepsilon, \partial_t u_\varepsilon$ are uniformly (in $\varepsilon$) bounded in $L^2(0, T; H^0_m(\Omega))$ and $L^2(0, T; H^{-m}(\Omega))$, respectively, $\sum_{|\beta|=m} A_{ij}^{\alpha \beta} D^\beta u_\varepsilon$ is uniformly bounded in $L^2(0, T; L^2(\Omega))$ for all $|\alpha|=m$, where $A_{ij}^{\alpha \beta}(x,t) = A^{\alpha \beta}(x/\varepsilon, t/\varepsilon^2m)$. Up to subsequences, we may assume that there exists a function $u_0$ such that

$$
u_\varepsilon \to u_0 \text{ weakly in } L^2(0, T; H^0_m(\Omega)),
$$

$$
\partial_t u_\varepsilon \to \partial_t u_0 \text{ weakly in } L^2(0, T; H^{-m}(\Omega)),
$$

$$
\sum_{|\beta|=m} A_{ij}^{\alpha \beta} D^\beta u_\varepsilon \to \Psi^\alpha(x,t) \text{ weakly in } L^2(0, T; L^2(\Omega)),
$$

(2.12)

$$
u_\varepsilon \to u_0 \text{ strongly in } L^2(0, T; H^{m-1}(\Omega)),
$$

where the last convergence result in (2.12) follows from the well-known Aubin-Simon type compactness result. Moreover,

$$
\partial_t u_0 + (-1)^m \sum_{|\alpha|=m} D^\alpha \Psi^\alpha = f \quad \text{in } L^2(0, T; H^{-m}(\Omega)).
$$

Taking $\phi \in C^1([0, T]; H^m(\Omega))$ with $\phi(T) = 0$ as a test function, we obtain that

$$
- \int_\Omega u_0(0)\phi(0) \, dx - \int_{\Omega_T} u_0 \partial_t \phi \, dx \, dt + \sum_{|\alpha|=m} \int_{\Omega_T} \Psi^\alpha D^\alpha \phi \, dx \, dt = \int_0^T \langle f(t), \phi(t) \rangle \, dt,
$$

(2.13)
where $\Omega_T = \Omega \times (0, T)$. Hereafter, let us denote the product of $H^{-m}(\Omega)$ and $H^m_0(\Omega)$ as $\langle , \rangle$ for short. On the other hand, taking such a $\phi$ as a test function in (1.1) and passing to the limits, we obtain that

$$- \int_\Omega h\phi(0) \, dx - \int_{\Omega_T} u_0 \partial_t \phi \, dx dt + \sum_{|\alpha| = |\beta| = m} \int_{\Omega_T} \Psi^\alpha D^\alpha \phi \, dx dt = \int_0^T \langle f(t), \phi(t) \rangle \, dt,$$

which, combined with (2.13), implies that $u_0(0) = h$. Therefore, to verify that $u_0$ is a weak solution of (1.5), it remains to prove

$$\Psi^\alpha = \sum_{|\beta| = m} A^{\alpha\beta} D^\beta u_0. \quad (2.14)$$

For positive integer $k$, let

$$\Psi_k = \{(P^1_k, P^2_k, \ldots, P^n_k) \mid P^i_k \text{ are homogeneous polynomials of } y \text{ of degree } k\}.$$  

For $P_m \in \Psi_m$, let $\omega$ be the weak solution to the cell problem

$$- \partial_t \omega + (-1)^m \sum_{|\alpha| = |\beta| = m} D^\alpha (A^{\alpha\beta} D^\beta \omega) = (-1)^{m+1} \sum_{|\alpha| = m} D^\alpha (A^{\alpha\beta} D^\beta P_m) \quad \text{in } Y,$$

$$\omega(y, s) \text{ is } 1\text{-periodic in } (y, s) \quad \text{and } \int_Y \omega(y, s) \, dy \, ds = 0. \quad (2.15)$$

Setting

$$\theta_\varepsilon = \theta_\varepsilon(x, t) = \varepsilon^m \left\{ \omega(x/\varepsilon, t/\varepsilon^{2m}) + P_m(x/\varepsilon) \right\},$$

it is not difficult to find that, as $\varepsilon$ tend to zero,

$$\theta_\varepsilon \longrightarrow P_m \quad \text{strongly in } L^2(0, T; H^{m-1}(\Omega)) \quad \text{and weakly in } L^2(0, T; H^m(\Omega)), \quad (2.16)$$

$$- \partial_t \theta_\varepsilon + (-1)^m \sum_{|\alpha| = |\beta| = m} D^\alpha (A^{\alpha\beta}_\varepsilon D^\beta \theta_\varepsilon) = 0 \quad \text{in } \mathbb{R}^{d+1}, \quad (2.17)$$

where $A^{\alpha\beta}_\varepsilon(x, t) = A^{\alpha\beta}(x/\varepsilon, t/\varepsilon^{2m})$. For any $\phi \in C_c^\infty(\Omega \times (0, T))$, we deduce from (1.1) and (2.17) that

$$\int_0^T \langle \partial_t u_\varepsilon, \phi \theta_\varepsilon \rangle \, dt + \sum_{|\alpha| = |\beta| = m} \int_{\Omega_T} A^{\alpha\beta}_\varepsilon D^\beta u_\varepsilon \phi \theta_\varepsilon \, dx dt = \int_0^T \langle f, \phi \theta_\varepsilon \rangle \, dt,$$

$$- \int_0^T \langle \partial_t \theta_\varepsilon, \phi u_\varepsilon \rangle \, dt + \sum_{|\alpha| = |\beta| = m} \int_{\Omega_T} A^{\alpha\beta}_\varepsilon D^\beta \theta_\varepsilon \phi u_\varepsilon \, dx dt = 0. \quad (2.18)$$

Subtracting the second equality from the first one in (2.18), it yields

$$\sum_{|\alpha| = |\beta| = m} C(\zeta) \left\{ \int_{\Omega_T} A^{\alpha\beta}_\varepsilon D^\beta u_\varepsilon D^n \phi D^\zeta \theta_\varepsilon \, dx dt - \int_{\Omega_T} A^{\alpha\beta}_\varepsilon D^\beta \theta_\varepsilon D^n \phi D^\zeta u_\varepsilon \, dx dt \right\}$$
\[
\int_0^T \langle u_\varepsilon, \theta_\varepsilon \partial_t \phi \rangle dt + \int_0^T \langle f, \phi \theta_\varepsilon \rangle dt,
C(\zeta) = \frac{\alpha!}{\zeta!(\alpha - \zeta)!}
\]  \tag{2.19}

Thanks to the convergence results for \(u_\varepsilon\) and \(\theta_\varepsilon\) (see (2.12) and (2.16)), up to subsequences,
\[
\int_0^T \langle u_\varepsilon, \theta_\varepsilon \partial_t \phi \rangle dt + \int_0^T \langle f, \phi \theta_\varepsilon \rangle dt \rightarrow \int_0^T \langle u_0, \partial_t \phi P_m \rangle dt + \int_0^T \langle f, \phi P_m \rangle dt,
\]
\[
\sum_{|\beta|=m} \int_{\Omega_T} A^{\alpha \beta}_\varepsilon D^\beta u_\varepsilon D^\alpha \phi D^\zeta \partial_\varepsilon \phi dxdt \rightarrow \int_{\Omega_T} \Psi^\alpha D^\alpha P_m \phi dxdt,
\]  \tag{2.20}

for \(\zeta + \eta = \alpha, |\zeta| = m - 1\). On the other hand, taking \(\phi P_m\) as a test function in (1.1) and passing to the limit in \(\varepsilon\), we get
\[
- \int_0^T \langle u_0, \partial_t \phi P_m \rangle dt + \sum_{|\alpha|=m} \int_{\Omega_T} \Psi^\alpha D^\alpha (\phi P_m) dxdt = \int_0^T \langle f, \phi P_m \rangle dt. \tag{2.21}
\]

By combining (2.19), (2.20) and (2.21), we obtain that
\[
\sum_{|\alpha|=|\beta|=m, \eta + \zeta = \alpha, |\zeta| \leq m-1} C(\zeta) \int_{\Omega_T} A^{\alpha \beta}_\varepsilon D^\beta \theta_\varepsilon D^\alpha \phi D^\zeta \partial_\varepsilon \phi dxdt \rightarrow - \sum_{|\alpha|=m} \int_{\Omega_T} \Psi^\alpha D^\alpha P_m \phi dxdt.
\]  \tag{2.22}

Note that \(A^{\alpha \beta}(y, s) D^\beta \theta(y)\) is periodic and uniformly bounded in \(L^2(0, T; L^2(\Omega))\). By the well-known weak convergence result of rapidly oscillating periodic functions [10, p.5],
\[
\sum_{|\beta|=m} A^{\alpha \beta}_\varepsilon D^\beta \theta_\varepsilon \text{ converges weakly to } \sum_{|\beta|=m} \int_Y A^{\alpha \beta}(y, s) (D^\beta \omega(y, s) + D^\beta P_m) = \mathcal{M}^\alpha
\]  \tag{2.23}

in \(L^2(0, T; L^2(\Omega))\), which implies that the l.h.s. of (2.22) converges to
\[
\sum_{|\alpha|=m, |\zeta| \leq m-1} C(\zeta) \int_{\Omega_T} \mathcal{M}^\alpha D^\alpha \phi D^\zeta u_0 dxdt
\]
\[
= \sum_{|\alpha|=m} \int_{\Omega_T} \mathcal{M}^\alpha D^\alpha u_0 \phi dxdt - \sum_{|\alpha|=m} \int_{\Omega_T} \mathcal{M}^\alpha D^\alpha u_0 \phi dxdt
\]
\[
= - \sum_{|\alpha|=m} \int_{\Omega_T} \mathcal{M}^\alpha D^\alpha u_0 \phi dxdt.
\]  \tag{2.24}

In view of (2.22), we get
\[
\sum_{|\alpha|=m} \int_{\Omega_T} \Psi^\alpha D^\alpha P_m \phi dxdt = \sum_{|\alpha|=m} \int_{\Omega_T} \mathcal{M}^\alpha D^\alpha u_0 \phi dxdt, \quad \forall \phi \in C_c(\Omega \times (0, T)). \tag{2.25}
\]

Hence
\[
\sum_{|\alpha|=m} \Psi^\alpha D^\alpha P_m = \sum_{|\alpha|=m} \mathcal{M}^\alpha D^\alpha u_0.
\]
For any $d$-dimensional multi-index $\gamma$ with $|\gamma| = m$, set $P_m = \frac{1}{\gamma!}y^\gamma e_i$. Then by (2.15) and (2.23), we have $\omega = \chi^s_{i\gamma}$ and

$$
\Psi^s_i = \sum_{|\alpha|=m} \left\{ \sum_{|\beta|=m} \int_\gamma A^{s\gamma}(y, s)D^\alpha(\chi^s_{ij}) + \frac{y^\gamma}{\gamma!}e_i D^\alpha(\frac{y^\alpha}{\alpha!}) dy ds \right\} D^\alpha u_{0j} = \sum_{|\alpha|=m} \tilde{A}^{s\alpha}_{ij} D^\alpha u_{0j},
$$

which is exactly (2.14). The proof is thus complete. \hfill \Box

Similar to Theorem 2.1, we can prove that the homogenized operator for $-\partial_t + L^*_\varepsilon$ is given by

$$
-\partial_t + L^*_0 = -\partial_t + (-1)^m \sum_{|\alpha|=|\beta|=m} D^\alpha(\tilde{A}^{s\alpha\beta}_{ij} D^\beta),
$$

(2.26)

where $\tilde{A}^s = (\tilde{A}^{s\alpha\beta}_{ij}) = (\tilde{A}^{\beta\alpha}_{ji})$. Furthermore, the same argument also gives the homogenized system for the initial value problem (1.1) with homogeneous Neumann boundary data.

**Theorem 2.2.** Let $\Omega$ be a bounded Lipschitz domain in $\mathbb{R}^d$. Assume that $A$ satisfies conditions (1.3)–(1.4), and $f \in L^2(0, T; \tilde{H}^{-m}(\Omega))$, $h \in L^2(\Omega)$. Let $u_\varepsilon, u_0 \in L^2(0, T; H^m(\Omega)) \cap L^\infty(0, T; L^2(\Omega))$ be the weak solutions to problem (1.1) and problem (1.10), respectively. Then as $\varepsilon \rightarrow 0$,

$$
u_\varepsilon \rightarrow u_0 \text{ weakly in } L^2(0, T; H^m(\Omega)) \text{ and strongly in } L^2(0, T; H^{m-1}(\Omega)).
$$

(2.27)

### 3 Smoothing operators and auxiliary estimates

We fix nonnegative functions $\varphi_1(s) \in C^\infty_c(-1/2, 1/2)$, $\varphi_2(y) \in C^\infty_c(B(0, 1/2))$, such that $\int_\mathbb{R} \varphi_1(s) = 1$ and $\int_{\mathbb{R}^d} \varphi_2(y) = 1$. Set $\varphi_{1,\varepsilon}(s) = \frac{1}{\varepsilon^m} \varphi_1(s/\varepsilon^m)$, $\varphi_{2,\varepsilon}(y) = \frac{1}{\varepsilon^m} \varphi_2(y/\varepsilon)$, and define

$$
\tilde{S}_\varepsilon(f)(x, t) = \int_{\mathbb{R}^d} \varphi_{2,\varepsilon}(y)f(x-y, t) dy = \int_{\mathbb{R}^d} \varphi_2(y)f(x-\varepsilon y, t) dy,
$$

$$
S_\varepsilon(f)(x, t) = \int_{\mathbb{R}^{d+1}} \varphi_{1,\varepsilon}(s)\varphi_{2,\varepsilon}(y)f(x-y, t-s) = \int_{\mathbb{R}^{d+1}} \varphi_1(s)\varphi_2(y)f(x-\varepsilon y, t-\varepsilon^2 s).
$$

(3.1)

By the definition, it is obvious that

$$
S_\varepsilon(f)(x, t) = \int_\mathbb{R} \varphi_{1,\varepsilon}(s)\tilde{S}_\varepsilon(f)(x, t-s) ds
$$

(3.2)

**Lemma 3.1.** Let $f \in L^p(\mathbb{R}^{d+1})$ for some $1 \leq p < \infty$, $g \in L^p_{\text{loc}}(\mathbb{R}^{d+1})$ and $h \in L^\infty(\mathbb{R}^{d+1})$ with support $\mathcal{O}$. Then

$$
\|gS_\varepsilon f\|_{L^p(\mathbb{R}^{d+1})} \leq C \sup_{z \in \mathbb{R}^{d+1}} \left( \int_{B(z, 1)} |g|^p \right)^{1/p} \|f\|_{L^p(\mathcal{O}(\varepsilon))} \|h\|_{L^\infty},
$$

where $g_\varepsilon(x, t) = g(x/\varepsilon, t/\varepsilon^2)$, $\mathcal{O}(\varepsilon) = \{ z \in \mathbb{R}^{d+1} : \text{dist}(z, \mathcal{O}) < \varepsilon \}$. If in addition $g(y, s)$ is 1-periodic in $(y, s)$, then

$$
\|g_\varepsilon S_\varepsilon f\|_{L^p(\mathbb{R}^{d+1})} \leq C \|g\|_{L^p(\mathbb{R})} \|f\|_{L^p(\mathcal{O}(\varepsilon))} \|h\|_{L^\infty}.
$$
Proof. See [7, Lemma 3.3] and also [21, Lemma 2.1].

For $k > 0$, let $\Omega_{k\varepsilon} = \{x \in \Omega : \text{dist}(x, \partial \Omega) \leq k\varepsilon\}$ and

$$\Omega_{T,k\varepsilon} = \Omega_{k\varepsilon} \times (0, T) \cup \Omega \times (0, k\varepsilon^{2m}] \cup \Omega \times [T-k\varepsilon^{2m}, T).$$

Lemma 3.2. Let $S_\varepsilon(f)$ be defined as in (3.1). Then for any $1 < q < \infty$, and any integer $\ell > 0$,

$$\|\partial_\varepsilon S_\varepsilon(f)\|_{L^q(\Omega_{T\setminus \Omega_{T,2\varepsilon}})} \leq C\varepsilon^{-2\ell} \|f\|_{L^q(\Omega_{T\setminus \Omega_{T,2\varepsilon}})}, \quad (3.3)$$

$$\|\nabla^\ell S_\varepsilon(f)\|_{L^q(\Omega_{T\setminus \Omega_{T,2\varepsilon}})} \leq C\varepsilon^{-\ell} \|f\|_{L^q(\Omega_{T\setminus \Omega_{T,2\varepsilon}})}, \quad (3.4)$$

$$\|S_\varepsilon(\nabla^\ell f) - \nabla^\ell f\|_{L^q(\Omega_{T\setminus \Omega_{T,4\varepsilon}})} \leq C\varepsilon^\ell \|\nabla^{\ell+1}f\|_{L^q(\Omega_{T\setminus \Omega_{T,2\varepsilon}})} + C\varepsilon^{m-\ell+1} \|\partial_\varepsilon f\|_{L^q(\varepsilon^{2m}, T; \varepsilon^{2m}; W^{m+1,q}(\Omega_{T,2\varepsilon}))}, \quad (3.5)$$

where $C$ depends only on $d, q, \Omega, T$.

Proof. Note that

$$\|\partial_\varepsilon S_\varepsilon(f)\|^q_{L^q(\Omega_{T\setminus \Omega_{T,2\varepsilon}})} = \int_{\Omega_{T\setminus \Omega_{T,2\varepsilon}}} \left| \int_{\mathbb{R}^{d+1}} \partial_\varepsilon \varphi(x - y, t - s) f(y, s) dy ds \right|^q dx dt, \quad (3.6)$$

where $\varphi(x - y, t - s) = \varphi_{1,\varepsilon}(t-s)\varphi_{2,\varepsilon}(x-y)$. Using Hölder’s inequality, we deduce that

$$\left| \int_{\mathbb{R}^{d+1}} \partial_\varepsilon \varphi(x - y, t - s) f(y, s) dy ds \right|^q \leq \left\{ \int_{\mathbb{R}^{d+1}} |\partial_\varepsilon \varphi(x - y, t - s)| |f(y, s)| dy ds \right\} \left\{ \int_{\mathbb{R}^{d+1}} |\partial_\varepsilon \varphi(x - y, t - s)| dy ds \right\}^{q-1} \leq C\varepsilon^{2q-1} \left\{ \int_{\mathbb{R}^{d+1}} |\partial_\varepsilon \varphi(x - y, t - s)| |f(y, s)| dy ds \right\}. \quad (3.7)$$

Taking (3.7) into (3.6) and using Fubini’s theorem, it yields

$$\|\partial_\varepsilon S_\varepsilon(f)\|^q_{L^q(\Omega_{T\setminus \Omega_{T,2\varepsilon}})} \leq C\varepsilon^{-2\ell q} \int_{\Omega_{T\setminus \Omega_{T,2\varepsilon}}} |f(y, s)|^q dy ds,$$

which is exactly (3.3). The proof of (3.4) is similar and we therefore pass to (3.5).

Recall that (3.5) was essentially proved in [7, Lemma 3.2] for the case $m = \ell = 1, q = 2$ by the Plancherel Theorem, which is not applicable for general $q$. By (3.2),

$$\|S_\varepsilon(\nabla^\ell f) - \nabla^\ell f\|^q_{L^q(\Omega_{T\setminus \Omega_{T,4\varepsilon}})} \leq \int_{\Omega_{T\setminus \Omega_{T,4\varepsilon}}} \left| \int_{\mathbb{R}} \varphi_1(s) \tilde{S}_\varepsilon(\nabla^\ell f)(x, t - \varepsilon^{2m}s) ds - \tilde{S}_\varepsilon(\nabla^\ell f)(x, t) \right|^q dx dt + \int_{\Omega_{T\setminus \Omega_{T,4\varepsilon}}} |\tilde{S}_\varepsilon(\nabla^\ell f)(x, t) - \nabla^\ell f(x, t)|^q dx dt \equiv (I) + (II). \quad (3.8)$$

Similar to (3.7), we may use Hölder’s inequality and Fubini’s theorem to deduce that

$$(II) \leq C \int_{\Omega_{T\setminus \Omega_{T,4\varepsilon}}} \int_{\mathbb{R}^d} \varphi_2(y) \left| \nabla^\ell f(x - \varepsilon y, t) - \nabla^\ell f(x, t) \right|^q dy dx dt \leq C\varepsilon^q \|\nabla^{\ell+1}f\|^q_{L^q(\Omega_{T\setminus \Omega_{T,2\varepsilon}})}.$$
And also,

\[
(I) \leq \int_{\Omega_T \setminus \Omega_{2\varepsilon}} \int_\mathbb{R} \varphi_1(s) \left| \tilde{S}_\varepsilon(\nabla^\ell f)(x, t - \varepsilon^{2m}s) - \tilde{S}_\varepsilon(\nabla^\ell f)(x, t) \right|^q ds \, dx \, dt
\]

\[
\leq C \varepsilon^{2mq} \left\| \partial_t \tilde{S}_\varepsilon(\nabla^\ell f) \right\|_{L^q(\Omega_T \setminus \Omega_{2\varepsilon})}^q \int_\mathbb{R} \varphi_1(s) ds
\]

\[
\leq C \varepsilon^{(m+1-\ell)q} \left\| \partial_t f \right\|_{L^q(\varepsilon^{2m}, T, \varepsilon^{-2m}; W^{m+1, q}(\Omega, \Omega_\varepsilon))}^q
\]

which, together with (3.8) and (3.9), gives (3.10). For the last step of (3.11), we have used the following observation

\[
\left\| \partial_t \tilde{S}_\varepsilon(\nabla^\ell f) \right\|_{L^q(\Omega_{T \setminus \Omega_{2\varepsilon}})} \leq C \varepsilon^{-m-\ell+1} \left\| \partial_t f \right\|_{L^q(\varepsilon^{2m}, T, \varepsilon^{-2m}; W^{m+1, q}(\Omega, \Omega_\varepsilon))}^q.
\]

To see this, we note that for any \( g \in C^\infty_c(\Omega_T \setminus \Omega_{2\varepsilon}) \) and \( \ell \)-dimensional multi-index \( \eta \) with \( |\eta| = \ell \),

\[
\int_{\Omega_T \setminus \Omega_{2\varepsilon}} \partial_t \tilde{S}_\varepsilon(\nabla^\ell f)(x, t)g(x, t) \, dx \, dt
\]

\[
= \int_{2\varepsilon^{2m}}^{T-2\varepsilon^{2m}} \int_{\Omega \setminus \Omega_{2\varepsilon}} \partial_t \left\{ \int_\mathbb{R}^d \varphi_{2, \varepsilon}(x - y) D_\eta^\ell f(y, t) \, dy \right\} g(x, t) \, dx \, dt
\]

\[
= (-1)^{\ell+1} \int_{2\varepsilon^{2m}}^{T-2\varepsilon^{2m}} \left\{ \int_{\Omega \setminus \Omega_{2\varepsilon}} \int_\mathbb{R}^d D_\eta^\ell \varphi_{2, \varepsilon}(x - y) f(y, t) \partial_t g(x, t) \, dy \, dx \right\} dt
\]

\[
= (-1)^{\ell+1} \int_{2\varepsilon^{2m}}^{T-2\varepsilon^{2m}} \left\{ \int_{\Omega \setminus \Omega_{\varepsilon}} f(y, t) \partial_t \left[ \int_\mathbb{R}^d D_\eta^\ell \varphi_{2, \varepsilon}(x - y) g(x, t) \, dx \right] \right\} dt
\]

\[
= (-1)^\ell \int_{2\varepsilon^{2m}}^{T-2\varepsilon^{2m}} (\partial_t f(y, t), D_\eta^\ell \tilde{S}_\varepsilon(g))_{W^{-m-1, q}((\Omega, \Omega_\varepsilon) \times W_0^{m-1, q'}((\Omega, \Omega_\varepsilon)) dt, \quad q' = q/(q - 1),
\]

where the integration by parts and Fubini's theorem have been used for the second and the third equalities respectively. Therefore, using estimates similar to (3.4) we may deduce that

\[
\left| \int_{\Omega_T \setminus \Omega_{2\varepsilon}} \partial_t \tilde{S}_\varepsilon(\nabla^\ell f)(x, t)g(x, t) \right| \leq C \left\| \partial_t f \right\|_{L^q(\varepsilon^{2m}, T, \varepsilon^{-2m}; W^{m+1, q}(\Omega, \Omega_\varepsilon))} \left\| \nabla^{m+1-\ell} \tilde{S}_\varepsilon(g) \right\|_{L^{q'}(\Omega_T \setminus \Omega_{2\varepsilon})}
\]

\[
\leq C \varepsilon^{-m-\ell+1} \left\| \partial_t f \right\|_{L^q(\varepsilon^{2m}, T, \varepsilon^{-2m}; W^{m+1, q}(\Omega, \Omega_\varepsilon))} \left\| g \right\|_{L^{q'}(\Omega_T \setminus \Omega_{2\varepsilon})},
\]

which implies (3.11) directly. The proof is complete. \( \square \)

### 4 Convergence rates for the initial-Dirichlet problem

#### 4.1 \( O(\sqrt{\varepsilon}) \) error estimate in \( L^2(0, T; H^m(\Omega)) \)

Let \( \rho_\varepsilon \) be a function in \( C^\infty_c(\Omega) \) such that

\[supp(\rho_\varepsilon) \subset \Omega \setminus \Omega_{8\varepsilon}, \quad 0 \leq \rho_\varepsilon \leq 1, \quad \rho_\varepsilon = 1 \text{ on } \Omega \setminus \Omega_{8\varepsilon} \quad \text{and} \quad |\nabla^k \rho_\varepsilon| \leq C_\varepsilon^{-k} \text{ for } 1 \leq k \leq m.\]
For $0 < T < \infty$, let $\varrho_\varepsilon$ be a function in $C^\infty_c(0, T)$ with $\text{supp}(\varrho_\varepsilon) \subseteq (6\varepsilon^{2m}, T - 6\varepsilon^{2m})$ and

$$0 \leq \varrho_\varepsilon \leq 1, \quad \varrho_\varepsilon = 1 \text{ in } (8\varepsilon^{2m}, T - 8\varepsilon^{2m}), \quad |\varrho'_\varepsilon| \leq C\varepsilon^{-2m}.$$ 

Define

$$\varpi_\varepsilon(x, t) = u_\varepsilon(x, t) - u_0(x, t) - \varepsilon^m \sum_{|\gamma| = m} \chi^\gamma_\varepsilon(x, t)K_\varepsilon(D^\gamma u_0)(x, t), \quad (4.1)$$

where $K_\varepsilon(u) = S^2_\varepsilon(u)\varrho_\varepsilon, S^2_\varepsilon = S_\varepsilon \circ S_\varepsilon$ and $\chi^\gamma_\varepsilon(x, t) = \chi^\gamma(x/\varepsilon, t/\varepsilon^{2m})$.

**Lemma 4.1.** Let $\Omega$ be a bounded Lipschitz domain in $\mathbb{R}^d$, and $A$ satisfy conditions $(1.3)$ and $(1.4)$. Let $u_\varepsilon, u_0$ be weak solutions to initial-Dirichlet problems $(1.1)$ and $(1.3)$, respectively. Then we have

$$(\partial_t + L_\varepsilon)\varpi_\varepsilon = (-1)^{m+1} \sum_{|\alpha| = |\beta| = m} D^\alpha \left\{ (A^\alpha_\varepsilon - \bar{A}^\alpha_\varepsilon)(D^\beta u_0 - K_\varepsilon(D^\beta u_0)) \right\}$$

$$+ (-1)^{m+1} \sum_{|\alpha| = |\beta| = |\gamma| = m \atop 0 \leq |\xi + \eta - \gamma| \leq m-1} C(\gamma, \zeta)\varepsilon^{|\eta|} D^\alpha \left\{ [A^\alpha_\varepsilon\gamma(D^\zeta \chi^\gamma_\varepsilon - (D^\zeta \omega^\alpha_\varepsilon)) \zeta]D^\eta K_\varepsilon(D^\beta u_0) \right\}$$

$$+ (-1)^m \varepsilon^{2m} \sum_{|\beta| = |\gamma| = m \atop 0 \leq |\xi + \eta - \gamma| \leq m-1} D^\alpha \left\{ (\omega^{(d+1)\alpha_\varepsilon} + B^\alpha_\varepsilon) \partial_t K_\varepsilon(D^\beta u_0) \right\}$$

$$+ (-1)^m \varepsilon^{m+|\eta|} \partial_t \left\{ (D^\zeta \omega^{(d+1)\beta_\varepsilon})\varepsilon D^\eta K_\varepsilon(D^\beta u_0) \right\}$$

$$+ (-1)^{m+1} \varepsilon^{2m} \sum_{|\alpha| = |\beta| = m} \partial_t \left\{ B^\alpha_\varepsilon D^\alpha K_\varepsilon(D^\beta u_0) \right\}. \quad (4.2)$$

**Proof.** For the simplicity of presentations, let us omit the subscripts $i, j$ in $(4.3)$ and hereafter. Using definitions of $\varpi_\varepsilon$ and $B^\alpha_\varepsilon$ (see $(4.1)$ and $(2.3)$ respectively), a direct computation yields that

$$(\partial_t + L_\varepsilon)\varpi_\varepsilon = (-1)^{m+1} \sum_{|\alpha| = |\beta| = m} D^\alpha \left\{ (A^\alpha_\varepsilon - \bar{A}^\alpha_\varepsilon)(D^\beta u_0 - K_\varepsilon(D^\beta u_0)) \right\}$$

$$+ (-1)^{m+1} \varepsilon^{m} \sum_{|\alpha| = |\beta| = |\gamma| = m \atop 0 \leq |\xi + \eta - \gamma| \leq m-1} C(\alpha, \zeta)\varepsilon D^\alpha \left\{ A^\alpha_\varepsilon\zeta D^\zeta \chi^\gamma_\varepsilon(D^\gamma u_0) \right\}$$

$$+ (-1)^{m+1} \sum_{|\alpha| = |\beta| = m} D^\alpha \left\{ B^\alpha_\varepsilon K_\varepsilon(D^\beta u_0) \right\} - \varepsilon^{m} \sum_{|\alpha| = m} \partial_t \left\{ \chi^\gamma_\varepsilon K_\varepsilon(D^\gamma u_0) \right\}, \quad (4.3)$$

where $B^\alpha_\varepsilon(x, t) = B^\alpha_\varepsilon(x/\varepsilon, t/\varepsilon^{2m})$ and $C(\alpha, \zeta) = \frac{\alpha^T}{\zeta(\alpha - \zeta)}$. In view of $(2.4)$, we may deduce that

$$(-1)^{m+1} \sum_{|\alpha| = |\beta| = m} D^\alpha \left\{ B^\alpha_\varepsilon K_\varepsilon(D^\beta u_0) \right\} - \varepsilon^{m} \sum_{|\alpha| = m} \partial_t \left\{ \chi^\gamma_\varepsilon K_\varepsilon(D^\gamma u_0) \right\}$$
\[ (-1)^{m+1} \sum_{|\alpha| = |\beta| = m} D^\alpha \left\{ \left( \sum_{|\gamma| = m} \varepsilon^m D^\gamma \mathfrak{B}_\varepsilon^{\alpha\beta} + \varepsilon^{2m} \partial_{\varepsilon} \mathfrak{B}_\varepsilon^{(d+1)\alpha\beta} + \widehat{B}^{\alpha\beta}_\varepsilon \right) K_\varepsilon(D^\beta u_0) \right\} \]

\[ + \varepsilon^m \sum_{|\alpha| = |\beta| = |\gamma| = m} \partial_{\varepsilon} \left\{ \left( \sum_{|\gamma| = m} \varepsilon^m D^\gamma \mathfrak{B}_\varepsilon^{(d+1)\beta} \right) K_\varepsilon(D^\beta u_0) \right\} \]

\[ = (-1)^{m+1} \varepsilon^m \sum_{|\alpha| = |\beta| = |\gamma| = m} D^\alpha D^\gamma \left\{ \mathfrak{B}_\varepsilon^{\alpha\beta} K_\varepsilon(D^\beta u_0) \right\} \]

\[ + (-1)^m \varepsilon^m \sum_{|\alpha| = |\beta| = |\gamma| = m} C(\gamma, \zeta) D^\alpha \left\{ D^\zeta \mathfrak{B}_\varepsilon^{\alpha\beta} D^\eta K_\varepsilon(D^\beta u_0) \right\} \]

\[ + (-1)^{m+1} \varepsilon^{2m} \sum_{|\alpha| = |\beta| = m} D^\alpha \partial_{\varepsilon} \left\{ \mathfrak{B}_\varepsilon^{(d+1)\alpha\beta} K_\varepsilon(D^\beta u_0) \right\} \]

\[ + (-1)^m \varepsilon^{2m} \sum_{|\alpha| = |\beta| = m} D^\alpha \left\{ \mathfrak{B}_\varepsilon^{(d+1)\alpha\beta} \partial_{\varepsilon} K_\varepsilon(D^\beta u_0) \right\} \]

\[ + (-1)^{m+1} \varepsilon^{2m} \sum_{|\beta| = |\gamma| = m} \partial_{\varepsilon} \left\{ \mathfrak{B}_\varepsilon^{\gamma(d+1)\beta} K_\varepsilon(D^\beta u_0) \right\} \]

\[ + (-1)^m \varepsilon^{2m} \sum_{|\beta| = |\gamma| = m} \partial_{\varepsilon} \left\{ D^\zeta \mathfrak{B}_\varepsilon^{\gamma(d+1)\beta} D^\eta K_\varepsilon(D^\beta u_0) \right\} \]

\[ + (-1)^m \varepsilon^m \sum_{|\alpha| = |\beta| = m} D^\alpha \left\{ \widehat{B}^{\alpha\beta}_\varepsilon K_\varepsilon(D^\beta u_0) \right\} \]

\[ = (-1)^m \varepsilon^m \sum_{|\alpha| = |\beta| = |\gamma| = m} C(\gamma, \zeta) D^\alpha \left\{ D^\zeta \mathfrak{B}_\varepsilon^{\alpha\beta} D^\eta K_\varepsilon(D^\beta u_0) \right\} \]

\[ + (-1)^m \varepsilon^m \sum_{|\alpha| = |\beta| = m} D^\alpha \left\{ \mathfrak{B}_\varepsilon^{(d+1)\alpha\beta} \partial_{\varepsilon} K_\varepsilon(D^\beta u_0) \right\} \]

\[ + (-1)^m \varepsilon^m \sum_{|\beta| = |\gamma| = m} \partial_{\varepsilon} \left\{ \mathfrak{B}_\varepsilon^{\gamma(d+1)\beta} D^\eta K_\varepsilon(D^\beta u_0) \right\} \]

\[ + (-1)^m \varepsilon^m \sum_{|\alpha| = |\beta| = m} D^\alpha \left\{ \widehat{B}^{\alpha\beta}_\varepsilon K_\varepsilon(D^\beta u_0) \right\}, \quad (4.4) \]

where we have used the facts \( \mathfrak{B}^{(d+1)(d+1)\beta} = 0 \) and \( \widehat{B}^{(d+1)\beta} = 0 \) for the second step and the skew-symmetry of \( \mathfrak{B} \) for the last step. Substituting (4.4) into (4.3), we get

\[ (\partial_{\varepsilon} + \mathcal{L}_{\varepsilon}) \varphi_\varepsilon = (-1)^{m+1} \sum_{|\alpha| = |\beta| = m} D^\alpha \left\{ \left( A^{\alpha\beta}_\varepsilon - \widehat{A}^{\alpha\beta}_\varepsilon \right) (D^\beta u_0 - K_\varepsilon(D^\beta u_0)) \right\} \]
This, together with (4.5), gives (4.2).

Moreover, thanks to Lemma 2.2, we have

\[
(4.5)
\]

\[
\begin{align*}
(-1)^{m+1} \sum_{\alpha = |\beta| = m} D^\alpha \left\{ \hat{B}_\varepsilon^{\alpha \beta} K_\varepsilon(D^\beta u_0) \right\} &= (-1)^{m+1} \sum_{\alpha = |\beta| = m} D^\alpha \left\{ \varepsilon^{2m} \partial_t \hat{B}_\varepsilon^{\alpha \beta} K_\varepsilon(D^\beta u_0) \right\} \\
&= (-1)^{m+1} \varepsilon^{2m} \sum_{\alpha = |\beta| = m} \partial_t \left\{ B_\varepsilon^{\alpha \beta} D^\alpha K_\varepsilon(D^\beta u_0) \right\} \\
&\quad + (-1)^{m+1} \varepsilon^{2m} \sum_{\alpha = |\beta| = m} D^\alpha \left\{ B_\varepsilon^{\alpha \beta} \partial_t K_\varepsilon(D^\beta u_0) \right\}.
\end{align*}
\]

This, together with (4.5), gives (4.2).

In the following, we define

\[
w_\varepsilon(x, t) = u_\varepsilon(x, t) - u_0(x, t) - \varepsilon^m \sum_{|\gamma| = m} \chi_\varepsilon^\gamma(x, t) K_\varepsilon(D^\gamma u_0)(x, t) \\
+ (-1)^{m+1} \sum_{|\beta| = |\gamma| = m} \varepsilon^{m+|\eta|} (D^\varepsilon \mathcal{Q}_\varepsilon^{(d+1) \gamma \beta}) D^\gamma K_\varepsilon(D^\beta u_0) \\
+ (-1)^{m} \varepsilon^{2m} \sum_{\alpha = |\beta| = m} B_\varepsilon^{\alpha \beta} D^\alpha K_\varepsilon(D^\beta u_0). \tag{4.6}
\]

**Lemma 4.2.** In addition of the assumptions of Lemma 4.1, if \( u_0 \in L^2(0, T; H^{m+1}(\Omega)) \) we have for any \( \phi \in L^2(0, T; H_0^{m+1}(\Omega)) \),

\[
\left| \int_0^T \langle \partial_t w_\varepsilon, \phi \rangle dt + \sum_{\alpha = |\beta| = m} \int_{\Omega_T} D^\alpha \phi A_\varepsilon^{\alpha \beta} D^\beta w_\varepsilon \right|
\]

\[
\leq C \left\{ \| u_0 \|_{L^2(0, T; H^{m+1}(\Omega))} + \| \partial_t u_0 \|_{L^2(0, T; H^{-m+1}(\Omega))} + \sup_{10\varepsilon < t < T} \left( \frac{1}{\varepsilon^m} \int_{t-10\varepsilon}^t \| \nabla^m u_0(t) \|_{L^2(\Omega)}^2 \right)^{1/2} \right\}
\times \left\{ \varepsilon^{1/2} \| \nabla^m \phi \|_{L^2(\Omega_T)} + \varepsilon \| \nabla^m \phi \|_{L^2(\Omega_{T, R})} \right\}, \tag{4.7}
\]

where \( C \) depends only on \( d, n, m, \mu, T \) and \( \Omega \).
Proof. According to (4.2), we have
\[
(\partial_t + \mathcal{L}_\varepsilon)w_\varepsilon = (-1)^{m+1} \sum_{|\alpha|=|\beta|=m} D^\alpha \left\{ (A_\varepsilon^{\alpha\beta} - \tilde{A}^{\alpha\beta}) (D^\beta u_0 - K_\varepsilon(D^\beta u_0)) \right\} \\
+ (-1)^{m+1} \sum_{|\alpha|=|\beta|=|\gamma|=m} C(\gamma, \zeta, \xi).Callback\{D^\alpha \left\{ (A_\varepsilon^{\alpha\gamma}(D^\xi \chi_\varepsilon) - \tilde{D}^\xi \mathfrak{A}^{\alpha\gamma}_\varepsilon(D^\beta u_0)) \right\} \\
- \sum_{|\alpha|=|\xi|=m} D^\alpha \left\{ A_\varepsilon^{\alpha\xi} D^\xi \left\{ \sum_{|\beta|=|\gamma|=m} \varepsilon^{m+|\eta|} (D^\xi \mathfrak{A}^{\gamma(\xi+1)^\beta}_\varepsilon) D^\eta K_\varepsilon(D^\beta u_0) \right\} \right\} \\
+ \sum_{|\alpha|=|\xi|=m} D^\alpha \left\{ A_\varepsilon^{\alpha\xi} D^\xi \left\{ \varepsilon^{2m} \sum_{|\beta|=|\gamma|=m} B_\varepsilon^\gamma D^\gamma K_\varepsilon(D^\beta u_0) \right\} \right\} \\
+ (-1)^m \varepsilon^{2m} \sum_{|\alpha|=|\beta|=m} D^\alpha \left\{ (\mathfrak{A}^{(d+1)\alpha\beta}_\varepsilon + B_\varepsilon^{\alpha\beta}) \partial_t K_\varepsilon(D^\beta u_0) \right\},
\]
which yields that for any \( \phi \in L^2(0, T; H^m_0(\Omega)) \)
\[
\int_0^T \langle \partial_t w_\varepsilon, \phi \rangle dt + \sum_{|\alpha|=|\beta|=m} \int_{\Omega_T} A_\varepsilon^{\alpha\beta} D^\beta w_\varepsilon D^\alpha \phi \\
= - \sum_{|\alpha|=|\beta|=m} \int_{\Omega_T} (A_\varepsilon^{\alpha\beta} - \tilde{A}^{\alpha\beta}) (D^\beta u_0 - K_\varepsilon(D^\beta u_0)) D^\alpha \phi \\
- \sum_{|\alpha|=|\beta|=|\gamma|=m} \varepsilon^{n|\eta|} C(\gamma, \zeta, \xi) \int_{\Omega_T} \left( A_\varepsilon^{\alpha\gamma}(D^\xi \chi_\varepsilon) - \tilde{D}^\xi \mathfrak{A}^{\alpha\gamma}_\varepsilon(D^\beta u_0) \right) D^\eta K_\varepsilon(D^\beta u_0) D^\alpha \phi \\
+ (-1)^{m+1} \sum_{|\alpha|=|\xi|=m} \varepsilon^{n|\eta|} C(\zeta, \xi, \alpha) \int_{\Omega_T} A_\varepsilon^{\alpha\xi} D^\xi \left( \sum_{|\beta|=|\gamma|=m} \varepsilon^{m+|\eta|} (D^\xi \mathfrak{A}^{\gamma(\xi+1)^\beta}_\varepsilon) D^\eta K_\varepsilon(D^\beta u_0) D^\alpha \phi \right) \\
+ (-1)^m \varepsilon^{m} \sum_{|\alpha|=|\xi|=m} \int_{\Omega_T} A_\varepsilon^{\alpha\xi} D^\xi \left( \sum_{|\beta|=|\gamma|=m} \varepsilon^{m} (D^\xi \mathfrak{A}^{\gamma(\xi+1)^\beta}_\varepsilon + B_\varepsilon^{\alpha\beta}) \partial_t K_\varepsilon(D^\beta u_0) D^\alpha \phi. \right)
\]
Denote the terms in the r.h.s of (4.8) as \( I_i, i = 1, \ldots, 5 \), in turn. Note that
\[
|D^\beta u_0 - K_\varepsilon(D^\beta u_0)| \\
\leq |[D^\beta u_0 - S_\varepsilon(D^\beta u_0)] \rho_\varepsilon \phi_\varepsilon| + |[S_\varepsilon(D^\beta u_0) - S^2_\varepsilon(D^\beta u_0)] \rho_\varepsilon \phi_\varepsilon| + |D^\beta u_0(1 - \rho_\varepsilon \phi_\varepsilon)|,
\]
which implies that
\[
|I_1| \leq C\|\nabla^m \phi\|_{L^2(\Omega_T)} \|\nabla^m u_0 - S_\varepsilon(\nabla^m u_0)\|_{L^2(\Omega_T \setminus \Omega_{T,\varepsilon})} + C\|\nabla^m \phi\|_{L^2(\Omega_{T,\varepsilon})} \|\nabla^m u_0\|_{L^2(\Omega_{T,\varepsilon})}.
\]
To deal with $I_2$, we observe that for $|\eta| \geq 1$,
\[
D^nK_\varepsilon(D^\beta u_0) = \sum_{\eta' + \eta'' = \eta} C(\eta, \eta') D^{n'} S^2_{\varepsilon}(D^\beta u_0) D^{n''} \rho_\varepsilon \Theta_e \\
= \sum_{\eta' + \eta'' = \eta} C(\eta, \eta') D^{n'} S^2_{\varepsilon}(D^\beta u_0) D^{n''} \rho_\varepsilon \Theta_e + S^2_{\varepsilon}(D^\beta D^n u_0) \rho_\varepsilon \Theta_e, \tag{4.9}
\]
which implies by Lemma [3.1] that
\[
|I_2| \leq C \sum_{1 \leq k \leq m} \varepsilon^k \|\nabla^m \phi\|_{L^2(\Omega_T)} \|S_\varepsilon(\nabla^{m+k} u_0)\|_{L^2(\Omega_T \setminus \Omega_{T,4}\eta)} \\
+ C \sum_{0 \leq k \leq m-1} \varepsilon^k \|\nabla^m \phi\|_{L^2(\Omega_{T,8\eta})} \|S_\varepsilon(\nabla^{m+k} u_0)\|_{L^2(\Omega_{T,8\eta} \setminus \Omega_{T,4}\eta)}, \tag{4.10}
\]
where $C$ depends only on $d, n, m, \mu, T$ and $\Omega$. In a similar manner, we also have
\[
|I_3| + |I_4| \leq C \sum_{1 \leq k \leq 2m} \varepsilon^k \|\nabla^m \phi\|_{L^2(\Omega_T)} \|S_\varepsilon(\nabla^{m+k} u_0)\|_{L^2(\Omega_T \setminus \Omega_{T,4}\eta)} \\
+ C \sum_{0 \leq k \leq 2m-1} \varepsilon^k \|\nabla^m \phi\|_{L^2(\Omega_{T,8\eta})} \|S_\varepsilon(\nabla^{m+k} u_0)\|_{L^2(\Omega_{T,8\eta} \setminus \Omega_{T,4}\eta)}, \tag{4.11}
\]
where $C$ depends only on $d, n, m, \mu, T$ and $\Omega$.

Finally, by Lemma [3.1] it is not difficult to find that
\[
|I_5| \leq C \varepsilon^{2m} \|\nabla^m \phi\|_{L^2(\Omega_T)} \|S_\varepsilon(\nabla^m \partial_t u_0)\|_{L^2(\Omega_T \setminus \Omega_{T,4}\eta)} \\
+ C \|\nabla^m \phi\|_{L^2(\Omega_{T,8\eta})} \|S_\varepsilon(\nabla^m u_0)\|_{L^2(\Omega_{T,8\eta} \setminus \Omega_{T,4}\eta)}. \tag{4.12}
\]
Taking the estimates for $I_1 - I_5$ into (4.8), we obtain that
\[
\left| \int_0^T \left< \partial_t w_\varepsilon, \phi \right>_{H^{-m}(\Omega) \times H^0(\Omega)} \, dt + \sum_{[\alpha] = [\beta] = m} \int_{\Omega_T} D^\alpha \phi A^\alpha_{\varepsilon} D^\beta w_\varepsilon \, dx \, dt \right| \\
\leq C \|\nabla^m \phi\|_{L^2(\Omega_{T,8\eta})} \|\nabla^m u_0\|_{L^2(\Omega_{T,8\eta})} + C \|\nabla^m \phi\|_{L^2(\Omega_T)} \|\nabla^m u_0 - S_\varepsilon(\nabla^m u_0)\|_{L^2(\Omega_T \setminus \Omega_{T,4}\eta)} \\
+ C \sum_{0 \leq k \leq 2m-1} \varepsilon^k \|\nabla^m \phi\|_{L^2(\Omega_{T,8\eta})} \|S_\varepsilon(\nabla^{m+k} u_0)\|_{L^2(\Omega_{T,8\eta} \setminus \Omega_{T,4}\eta)} \\
+ C \sum_{1 \leq k \leq 2m} \varepsilon^k \|\nabla^m \phi\|_{L^2(\Omega_T)} \|S_\varepsilon(\nabla^{m+k} u_0)\|_{L^2(\Omega_T \setminus \Omega_{T,4}\eta)} \\
+ C \varepsilon^{2m} \|\nabla^m \phi\|_{L^2(\Omega_T)} \|S_\varepsilon(\nabla^m \partial_t u_0)\|_{L^2(\Omega_T \setminus \Omega_{T,4}\eta)}, \tag{4.13}
\]
where $C$ depends only on $d, n, m, \mu, T$ and $\Omega$.

Denote the terms in the r.h.s. of (4.13) as $J_1, J_2, J_3, J_4, J_5$ in turn. Similar to [21] p.664, we may prove that
\[
\|\nabla^m u_0(t)\|_{L^2(\Omega_T)} \leq C \varepsilon^{1/2} \|u_0(t)\|_{H^{m+1}(\Omega)} \quad \text{for a.e. } t \in (0, T).
\]
This leads to the following estimate
\[
\|\nabla^m u_0(t)\|_{L^2(\Omega_{T,8\varepsilon})}^2 \leq \int_0^T \|\nabla^m u_0(t)\|_{L^2(\Omega_{T,8\varepsilon})}^2 dt + \left( \int_0^{8\varepsilon^2m} + \int_{T-8\varepsilon^2m}^T \|\nabla^m u_0(t)\|_{L^2(\Omega)}^2 dt \right)
\]
which implies that
\[
J_1 \leq C\varepsilon^{1/2}\|\nabla^m \phi\|_{L^2(\Omega_T,8\varepsilon)} \left\{ \left( \int_0^T \|u_0(t)\|_{H^{m+1}(\Omega)}^2 \right)^{1/2} \right. \\
+ \left. \sup_{8\varepsilon^2m < t < T} \left( \frac{1}{\varepsilon} \int_{t-8\varepsilon^2m}^t \|\nabla^m u_0(t)\|_{L^2(\Omega)}^2 dt \right)^{1/2} \right\}. \tag{4.14}
\]
By (3.5), we obtain that
\[
J_2 \leq C\varepsilon\|\nabla^m \phi\|_{L^2(\Omega_T)} \left\{ \|\partial_t u_0\|_{L^2(0,T;H^{-m+1}(\Omega))} + \|\nabla^{m+1} u_0\|_{L^2(\Omega_T)} \right\}, \tag{4.15}
\]
where the fact \(\|\partial_t u_0\|_{L^2(0,T;H^{-m+1}(\Omega))} \leq \|\partial_t u_0\|_{L^2(0,T;H^{-m+1}(\Omega))}\) is used. By (3.4) and (3.11), we have
\[
\|S\varepsilon(\nabla^{m+k} u_0)\|_{L^2(\Omega_{T,9\varepsilon}/\Omega_{T,4\varepsilon})} \leq C\varepsilon^{-k}\|\nabla^m u_0\|_{L^2(\Omega_{T,10\varepsilon})}, \\
\|S\varepsilon(\nabla^{m+k} u_0)\|_{L^2(\Omega_{T}/\Omega_{T,4\varepsilon})} \leq C\varepsilon^{-k+1}\|\nabla^{m+1} u_0\|_{L^2(\Omega_T)}, \\
\|S\varepsilon(\nabla^m \partial_t u_0)\|_{L^2(\Omega_{T}/\Omega_{T,4\varepsilon})} \leq C\varepsilon^{-2m+1}\|\partial_t u_0\|_{L^2(0,T;H^{-m+1}(\Omega))}.
\]
This implies that
\[
J_3 \leq C\varepsilon^{1/2}\|\nabla^m \phi\|_{L^2(\Omega_T,8\varepsilon)} \left\{ \left( \int_0^T \|u_0(t)\|_{H^{m+1}(\Omega)}^2 \right)^{1/2} \right. \\
+ \left. \sup_{10\varepsilon^2m < t < T} \left( \frac{1}{\varepsilon} \int_{t-10\varepsilon^2m}^t \|\nabla^m u_0(t)\|_{L^2(\Omega)}^2 dt \right)^{1/2} \right\}, \tag{4.17}
\]
\[
J_4 \leq C\varepsilon\|\nabla^m \phi\|_{L^2(\Omega_T)} \|\nabla^{m+1} u_0\|_{L^2(\Omega_T)}, \\
J_5 \leq C\varepsilon\|\nabla^m \phi\|_{L^2(\Omega_T)} \|\partial_t u_0\|_{L^2(0,T;H^{-m+1}(\Omega))}.
\]
Note that (4.17) follows directly from (4.13) and (4.15)–(4.17). The proof is complete. \(\square\)

Now we are in the position to establish the error estimate in \(L^2(0,T;H^m(\Omega))\).

**Theorem 4.1.** Let \(\Omega\) be a bounded Lipschitz domain in \(\mathbb{R}^d\), and \(A\) satisfy conditions (1.3) and (1.4). Let \(u_\varepsilon, u_0\) be weak solutions to initial-Dirichlet problems (1.1) and (1.5), respectively, and moreover \(u_0 \in L^2(0,T;H^{m+1}(\Omega))\). Let \(w_\varepsilon(x,t)\) be defined by (3.6). Then
\[
\|\nabla^m w_\varepsilon\|_{L^2(\Omega_T)} \leq C\varepsilon^{1/2} \left\{ \|u_0\|_{L^2(0,T;H^{m+1}(\Omega))} + \|f\|_{L^2(0,T;H^{-m+1}(\Omega))} + \|h\|_{L^2(\Omega)} \right\}. \tag{4.18}
\]
Proof. Note that \( w_\varepsilon \in L^2(0, T; H^m_0(\Omega)) \). Taking \( \phi = w_\varepsilon \) in (4.7), it yields
\[
\| \nabla^m w_\varepsilon \|_{L^2(\Omega_T)} \leq C\varepsilon^{1/2} \left\{ \| u_0 \|_{L^2(0,T; H^{m+1}(\Omega))} + \| \partial_t u_0 \|_{L^2(0,T; H^{-m+1}(\Omega))} + \sup_{0<\varepsilon<\varepsilon_0} \left( \frac{1}{\varepsilon} \int_{t_{-10\varepsilon^2}}^t \| \nabla^m u_0(t) \|_{L^2(\Omega)}^2 dt \right)^{1/2} \right\}. 
\]  (4.19)

Since
\[
\int_{t_1}^{t_2} \int_\Omega A^{\alpha\beta} D^\alpha u_0 D^\beta u_0 = \int_{t_1}^{t_2} (\partial_t u_0 + f, u_0) dt \quad \text{for } 0 \leq t_1, t_2 \leq T.
\]
By (1.3), we have
\[
\int_{t_{-10\varepsilon^2}}^t \| \nabla^m u_0(t) \|_{L^2(\Omega)}^2 dt \leq C \left( \int_0^T \| \partial_t u_0 + f(t) \|_{H^{-m+1}(\Omega)}^2 dt \right)^{1/2} \left( \int_{t_{-10\varepsilon^2}}^t \| u_0(t) \|_{H^{m+1}(\Omega)}^2 dt \right)^{1/2}. 
\]  (4.20)

By Gagliardo-Nirenberg inequality and Young’s inequality, we have
\[
\| u \|_{H^{-m+1}(\Omega)} \leq C \| u \|_{H^m(\Omega)}^{(2m-2)/m} \| u \|_{L^2(\Omega)}^{2/m} \leq C\varepsilon^2 \| u \|_{H^m(\Omega)}^2 + \varepsilon^{2m+2} \| u \|_{L^2(\Omega)}^2.
\]
We therefore obtain that
\[
\left( \int_{t_{-10\varepsilon^2}}^t \| u_0(t) \|_{H^{-m+1}(\Omega)}^2 dt \right)^{1/2} \leq C\varepsilon \left\{ \| u_0 \|_{L^2(0,T; H^m(\Omega))} + \| u_0 \|_{L^\infty(0,T; L^2(\Omega))} \right\} 
\leq C\varepsilon \left\{ \| f \|_{L^2(0,T; H^{-m+1}(\Omega))} + \| h \|_{L^2(\Omega)} \right\}. 
\]  (4.21)

On the other hand, from the equation of \( u_0 \) we note that
\[
\| \partial_t u_0 \|_{L^2(0,T; H^{-m+1}(\Omega))} \leq C \left\{ \| u_0 \|_{L^2(0,T; H^{m+1}(\Omega))} + \| f \|_{L^2(0,T; H^{-m+1}(\Omega))} \right\}. 
\]  (4.22)

We then conclude from (4.20)–(4.22) that
\[
\sup_{0<\varepsilon<\varepsilon_0} \left( \frac{1}{\varepsilon} \int_{t_{-10\varepsilon^2}}^t \| \nabla^m u_0(t) \|_{L^2(\Omega)}^2 dt \right)^{1/2} \leq C \left\{ \| u_0 \|_{L^2(0,T; H^{m+1}(\Omega))} + \| f \|_{L^2(0,T; H^{-m+1}(\Omega))} + \| h \|_{L^2(\Omega)} \right\}. 
\]  (4.23)

This, combined with (4.19) and (4.22), gives (4.18). \( \square \)

**Remark 4.1.** In addition to the assumptions of Theorem 4.1, if \( \Omega \) is \( C^{m,1} \) and \( h = 0 \), then as a consequence of (4.18) we have
\[
\| \nabla^m w_\varepsilon \|_{L^2(\Omega_T)} \leq C\varepsilon^{1/2} \| f \|_{L^2(0,T; H^{-m+1}(\Omega))}. 
\]  (4.24)

This follows from the estimate
\[
\| u_0 \|_{L^2(0,T; H^{m+1}(\Omega))} \leq C \| f \|_{L^2(0,T; H^{-m+1}(\Omega))}, 
\]  (4.25)
which may be proved by time discretization and reducing the estimate to the well-known \( H^{m+1} \) estimate for \( m \)-order elliptic systems with constant coefficients in \( C^{m,1} \) domains.
4.2 $O(\varepsilon)$ error estimate in $L^2(0,T;H^{m-1}(\Omega))$

With preparations in the last sections, we are now ready to prove Theorems 1.1. Let $L_\varepsilon^*, L_0^*$ be the adjoint operators of $L_\varepsilon$ and $L_0$, respectively. Let $\Omega$ be a bounded $C^{m,1}$ domain and $F \in L^2(0,T;H^{-m+1}(\Omega))$. Suppose that $v_\varepsilon, v_0$ are, respectively, weak solutions to

$$\begin{cases}
-\partial_t v_\varepsilon + L_\varepsilon^* v_\varepsilon = F & \text{in } \Omega_T, \\
Tr(D^\gamma v_\varepsilon) = 0 & \text{on } \partial \Omega \times (0,T), \\
v_\varepsilon = 0 & \text{on } \Omega \times \{t = T\},
\end{cases} \quad (4.26)$$

and

$$\begin{cases}
-\partial_t v_0 + L_0^* v_0 = F & \text{in } \Omega_T, \\
Tr(D^\gamma v_0) = 0 & \text{on } \partial \Omega \times (0,T), \\
v_0 = 0 & \text{on } \Omega \times \{t = T\}.
\end{cases} \quad (4.27)$$

Then it is not difficult to find that $v_\varepsilon(x, T-t), v_0(x, T-t)$ are solutions to problem (1.1) (with homogeneous Dirichlet boundary data) and problem (1.5), respectively, with $f(x, t) = F(x, T-t), h = 0$ as well as the coefficient matrix $A(x/\varepsilon, (T-t)/\varepsilon^2m)$ replaced by $A'(x/\varepsilon, (T-t)/\varepsilon^2m)$. Note that $A^*(y, T-s)$ satisfies the conditions (1.13) and (1.4) as $A(y, s)$. Similar to Section 2, we can introduce the matrix of correctors $\chi^*_{T, \varepsilon}$ and flux correctors $B^*_{T, \varepsilon}(y, s)$ and also $B^*_{T, \varepsilon}$ for the family of parabolic operators

$$-\partial_t + (-1)^m \sum_{|\alpha| = |\beta| = m} D^\alpha \{ A^{*\alpha\beta}(x/\varepsilon, (T-t)/\varepsilon^2m)D^\beta \}, \quad \varepsilon > 0.$$  

Proof of Theorem [1.1] We now use the duality argument [28, 7] to prove Theorem 1.1. For simplicity, we assume that

$$\|u_0\|_{L^2(0,T;H^{m-1}(\Omega))} + \|f\|_{L^2(0,T;H^{-m+1}(\Omega))} + \|h\|_{L^2(\Omega)} \leq 1.$$  

In view of the definition of $w_\varepsilon$, to prove (1.7) it is sufficient to prove the following estimates,

$$\varepsilon^m \left\| \sum_{|\gamma| = m} \chi^\gamma_{\varepsilon} K_{\varepsilon}(D^\gamma u_0) \right\|_{L^2(0,T;H^{m-1}_0(\Omega))} \leq C\varepsilon, \quad (4.28)$$

$$\varepsilon^{m+|\eta|} \left\| \sum_{|\beta| = |\gamma| = m} \chi^\eta_{\varepsilon} K_{\varepsilon}(D^\gamma D^\beta u_0) \right\|_{L^2(0,T;H^{m-1}_0(\Omega))} \leq C\varepsilon, \quad (4.29)$$

$$\varepsilon^{2m} \left\| \sum_{|\alpha| = |\beta| = m} B^\alpha_{\varepsilon} K_{\varepsilon}(D^\beta u_0) \right\|_{L^2(0,T;H^{m-1}_0(\Omega))} \leq C\varepsilon, \quad (4.30)$$

$$\|w_\varepsilon\|_{L^2(0,T;H^{m-1}_0(\Omega))} \leq C\varepsilon. \quad (4.31)$$

Thanks to the Poincaré inequality and Lemmas 3.1, 3.2, we deduce that

$$\left\| \sum_{|\gamma| = m} \chi^\gamma_{\varepsilon} K_{\varepsilon}(D^\gamma u_0) \right\|_{L^2(0,T;H^{m-1}_0(\Omega))}.$$  

22
Using (4.26) and (4.34), we deduce that

\[
\| \sum_{|\gamma|=m} D^\gamma \chi_{T,\varepsilon}^\gamma D^\xi \varepsilon^2 (D^\gamma u_0) D^\zeta \rho_\varepsilon \eta \|_{L^2(\Omega_T)} \leq C \varepsilon^{-m+1} \| \nabla^m u_0 \|_{L^2(\Omega_T)} \leq C \varepsilon^{-m+1},
\]

which implies (4.28). In a similar way, we can also get (4.29) and (4.30). To prove (4.31), it suffices to verify

\[
\left| \int_0^T \langle F, w_\varepsilon \rangle_{H^{-m+1}(\Omega) \times H^{m-1}_0(\Omega)} dt \right| \leq C \varepsilon \| F \|_{L^2(0,T;H^{-m+1}(\Omega))}, \quad \forall F \in L^2(0,T;H^{-m+1}(\Omega)).
\]

(4.33)

Similar to (4.3), we define

\[
\Phi_\varepsilon(x, t) = v_\varepsilon(x, T - t) - v_0(x, T - t) - \varepsilon^m \sum_{|\gamma|=m} \chi_{T,\varepsilon}^{\gamma} \tilde{K}_\varepsilon(D^\gamma v_0(x, T - t))
\]

\[
+ (-1)^{m+1} \sum_{|\beta|=|\gamma|=m} \varepsilon^m (\xi_\varepsilon^\gamma D^\xi_\varepsilon^\gamma) \tilde{K}_\varepsilon(D^\beta v_0(x, T - t))
\]

\[
+ (-1)^m \varepsilon^{2m} \sum_{|\alpha|=|\beta|=m} B_{T,\varepsilon}^{\alpha\beta} D^\alpha \tilde{K}_\varepsilon(D^\beta v_0(x, T - t)),
\]

(4.34)

where \( \tilde{K}_\varepsilon(u) = S_\varepsilon^2(u) \tilde{\rho}_\varepsilon \tilde{\eta}_\varepsilon \), \( \tilde{\rho}_\varepsilon \in C^\infty_c(\Omega) \) and \( \tilde{\eta}_\varepsilon \in C^\infty_c(0, T) \) such that

\[
\text{supp}(\tilde{\rho}_\varepsilon) \subset \Omega \setminus \Omega_{12\varepsilon}, \quad \text{supp}(\tilde{\eta}_\varepsilon) \subset (12\varepsilon^{2m}, T - 12\varepsilon^{2m}),
\]

\[
0 \leq \tilde{\rho}_\varepsilon \leq 1, \quad \tilde{\rho}_\varepsilon = 1 \text{ in } \Omega \setminus \Omega_{16\varepsilon} \quad \text{and} \quad |\nabla^k \tilde{\rho}_\varepsilon| \leq C \varepsilon^{-k}, 1 \leq k \leq m,
\]

\[
0 \leq \tilde{\eta}_\varepsilon \leq 1, \quad \tilde{\eta}_\varepsilon = 1 \text{ in } (16\varepsilon^{2m}, T - 16\varepsilon^{2m}) \quad \text{and} \quad |\tilde{\eta}'_\varepsilon| \leq C \varepsilon^{-2m}.
\]

Using (4.26) and (4.34), we deduce that

\[
\int_0^T \langle F, w_\varepsilon \rangle_{H^{-m+1}(\Omega) \times H^{m-1}_0(\Omega)} dt
\]

\[
= \int_0^T \langle \partial_t w_\varepsilon, v_\varepsilon(t) \rangle_{H^{-m+1}(\Omega) \times H^{m-1}_0(\Omega)} dt + \sum_{|\alpha|=|\beta|=m} \int_{\Omega_T} A_\varepsilon^{\beta\alpha} D^\alpha w_\varepsilon D^\beta v_\varepsilon(x, t)
\]

\[
= \int_0^T \langle \partial_t w_\varepsilon, \Phi_\varepsilon(T - t) \rangle_{H^{-m+1}(\Omega) \times H^{m-1}_0(\Omega)} dt + \sum_{|\alpha|=|\beta|=m} \int_{\Omega_T} A_\varepsilon^{\beta\alpha} D^\alpha w_\varepsilon D^\beta \Phi_\varepsilon(x, T - t)
\]

\[
+ \int_0^T \langle \partial_t w_\varepsilon, v_0(t) \rangle_{H^{-m+1}(\Omega) \times H^{m-1}_0(\Omega)} dt + \sum_{|\alpha|=|\beta|=m} \int_{\Omega_T} A_\varepsilon^{\beta\alpha} D^\alpha w_\varepsilon D^\beta v_0(x, t)
\]

\[
+ \int_0^T \langle \partial_t w_\varepsilon, v_\varepsilon(t) - v_0(t) - \Phi_\varepsilon(T - t) \rangle_{H^{-m+1}(\Omega) \times H^{m-1}_0(\Omega)} dt
\]

\[
+ \sum_{|\alpha|=|\beta|=m} \int_{\Omega_T} A_\varepsilon^{\beta\alpha} D^\alpha w_\varepsilon D^\beta \left\{ v_\varepsilon(x, t) - v_0(x, t) - \Phi_\varepsilon(x, T - t) \right\}
\]

23
\[ \pm \mathcal{I}_1 + \mathcal{I}_2 + \cdots + \mathcal{I}_6. \]  

By Lemma 4.2, we have
\[ \mathcal{I}_1 + \mathcal{I}_2 \leq C \left\{ \varepsilon \left\| \nabla^m \Phi_\varepsilon \right\|_{L^2(\Omega_T)} + \varepsilon^{1/2} \left\| \nabla^m \Phi_\varepsilon \right\|_{L^2(\Omega_{T_{4\varepsilon}})} \right\} \times \left\{ \left\| \partial_t u_0 \right\|_{L^2(0,T;H^{-m+1}(\Omega))} + \left\| u_0 \right\|_{L^2(0,T;H^{-m+1}(\Omega))} + \sup_{10^{-2m} t < T} \left( \frac{1}{\varepsilon} \int_{t-10^{-2m}}^t \left\| \nabla^m u_0(t) \right\|_{L^2(\Omega)}^2 dt \right)^{1/2} \right\} \leq C\varepsilon^{1/2} \left\| \nabla^m \Phi_\varepsilon \right\|_{L^2(\Omega_T)}, \]  

where we have used (4.22) and (4.23) for the second inequality. Since \( A(y, T - s) \) satisfies conditions (1.3) and (1.4) as \( A(y, s) \), and \( \Omega \) is \( C^{m,1} \). By (4.24), we have
\[ \left\| \nabla^m \Phi_\varepsilon \right\|_{L^2(\Omega_T)} \leq C\varepsilon^{1/2} \left\| F \right\|_{L^2(0,T;H^{-m+1}(\Omega))}, \]  

which combined with (4.36), implies that
\[ \mathcal{I}_1 + \mathcal{I}_2 \leq C\varepsilon \left\| F \right\|_{L^2(0,T;H^{-m+1}(\Omega))}. \]  

Likewise, using Lemma 4.2, (4.22) and (4.23), we deduce that
\[ \mathcal{I}_3 + \mathcal{I}_4 \leq C \left\{ \varepsilon \left\| \nabla^m v_0 \right\|_{L^2(\Omega_T)} + \varepsilon^{1/2} \left\| \nabla^m v_0 \right\|_{L^2(\Omega_{T_{4\varepsilon}})} \right\} \leq C\varepsilon \left\| F \right\|_{L^2(0,T;H^{-m+1}(\Omega))}, \]  

where, for the last step, we have used estimates (4.14), (4.23) and (4.25) for \( v_0 \). The proof of these estimates are completely the same as those for \( u_0 \), since \( A'(y, T - s) \) satisfies conditions as \( A(y, s) \).

Finally, note that \( v_\varepsilon(t) - v_0(t) - \Phi_\varepsilon(T - t) \) is supported on \( \Omega_T \setminus \bar{\Omega}_{T,12\varepsilon} \). From Lemma 4.2 (4.22) and (4.23), it follows that
\[ \mathcal{I}_5 + \mathcal{I}_6 \leq C\varepsilon \left\| \nabla^m [v_\varepsilon - v_0 - \Phi_\varepsilon(T - t)] \right\|_{L^2(\Omega_T)} \times \left\{ \left\| \partial_t u_0 \right\|_{L^2(0,T;H^{-m+1}(\Omega))} + \left\| u_0 \right\|_{L^2(0,T;H^{-m+1}(\Omega))} + \sup_{10^{-2m} t < T} \left( \frac{1}{\varepsilon} \int_{t-10^{-2m}}^t \left\| \nabla^m u_0(t) \right\|_{L^2(\Omega)}^2 dt \right)^{1/2} \right\} \leq C\varepsilon \left\| \nabla^m [v_\varepsilon - v_0 - \Phi_\varepsilon(T - t)] \right\|_{L^2(\Omega_T)}. \]  

Similar to (4.32), using Poincaré’s inequality and Lemmas 3.1 3.2, we can prove that
\[ \varepsilon^m \left| \sum_{|\gamma| = m} \lambda_{T,\varepsilon}^{\gamma \gamma} K_\varepsilon(D^\gamma v_0) \right|_{L^2(0,T;H_0^m(\Omega))} \leq C\varepsilon \left\| v_0 \right\|_{L^2(0,T;H^{m+1}(\Omega))}, \]
\[ \left| \sum_{|\beta| = m} \varepsilon^{|\beta|} \left( D^\beta \mathfrak{B}_T^{\gamma(m+1),\beta} \right)_{\varepsilon} D^\beta \Phi_\varepsilon(D^\beta v_0) \right|_{L^2(0,T;H_0^m(\Omega))} \leq C\varepsilon \left\| v_0 \right\|_{L^2(0,T;H^{m+1}(\Omega))}, \]
and
\[ \varepsilon^{2m} \left| \sum_{|\alpha| = |\beta| = m} \mathcal{B}_T^{\alpha \beta} D^\alpha \Phi_\varepsilon(D^\beta v_0) \right|_{L^2(0,T;H_0^m(\Omega))} \leq C\varepsilon \left\| v_0 \right\|_{L^2(0,T;H^{m+1}(\Omega))}, \]

which, together with (4.34), (4.40) and (4.25) for \( v_0 \), gives
\[ \mathcal{I}_5 + \mathcal{I}_6 \leq C\varepsilon \left\| F \right\|_{L^2(0,T;H^{-m+1}(\Omega))}. \]  

Note that (4.33) follows directly from (4.35), (4.38), (4.39) and (4.41). The proof is thus complete. \( \square \)
5 Convergence rates for the initial-Neumann problem

In this section, we provide a concise discussion on the convergence rate in the homogenization of initial-Neumann problem with homogeneous boundary data.

Lemma 5.1. Let $\Omega$ be a bounded Lipschitz domain in $\mathbb{R}^d$, $A$ satisfy conditions (1.3) and (1.4). Let $u_\varepsilon \in L^2(0, T; H^m(\Omega))$, $u_0 \in L^2(0, T; H^{m+1}(\Omega))$ be weak solutions to the initial-Neumann problems (1.1) and (1.10), respectively. Let $w_\varepsilon$ be defined as in (4.6). Then for any $\phi \in L^2(0, T; H^m(\Omega))$,

$$
\int_0^T \langle \partial_t w_\varepsilon, \phi \rangle_{\tilde{H}^{-m}(\Omega) \times H^m(\Omega)} dt + \sum_{|\alpha|=|\beta|=m} \int_{\Omega_T} D^\alpha \phi A_\varepsilon^\beta D^\beta w_\varepsilon = (\text{the r.h.s of (4.8)}) \quad (5.1)
$$

Proof. The proof is parallel to that of Lemma 4.1 with slight adaptations to Neumann problem. \qed

Lemma 5.2. Under the assumption of Lemma 5.1, we have for any $\phi \in L^2(0, T; H^m(\Omega))$,

$$
\int_0^T \langle \partial_t w_\varepsilon, \phi \rangle_{\tilde{H}^{-m}(\Omega) \times H^m(\Omega)} dt + \sum_{|\alpha|=|\beta|=m} \int_{\Omega_T} D^\alpha \phi A_\varepsilon^\beta D^\beta w_\varepsilon dx dt
\leq C \left\{ \|u_0\|_{L^2(0, T; H^{m+1}(\Omega))} + \|\partial_t u_0\|_{L^2(0, T; \tilde{H}^{-m}(\Omega))} + \sup_{10\varepsilon^{2m}<t<T} \left( \frac{1}{\varepsilon} \int_{t-10\varepsilon^{2m}}^t \|\nabla^m u_0(t)\|_{L^2(\Omega)}^2 \right)^{1/2} \right\}
\times \left\{ \varepsilon \|\nabla^m \phi\|_{L^2(\Omega_T)} + \varepsilon^{1/2} \|\nabla^m \phi\|_{L^2(\Omega, \Sigma)} \right\}, \quad (5.2)
$$

where $C$ depends only on $d, n, m, \mu, T$ and $\Omega$.

Proof. The proof is almost the same as that of Lemma 4.2 let us omit the details. \qed

With Lemmas 5.1 and 5.2 as preparations, we obtain the following theorem parallel to Theorem 4.1. Since the proof is almost the same to the one of Theorem 4.1 we omit the details for brevity.

Theorem 5.1. Let $\Omega$ be a bounded Lipschitz domain in $\mathbb{R}^d$, $A$ satisfy conditions (1.3) and (1.4). Let $u_\varepsilon, u_0 \in L^2(0, T; H^m(\Omega))$ be weak solutions to initial problems (1.1) and (1.10) with homogeneous Neumann boundary data respectively. Moreover, assume that $u_0 \in L^2(0, T; H^{m+1}(\Omega))$ and $w_\varepsilon$ is defined as (4.6). Then

$$
\|\nabla^m w_\varepsilon\|_{L^2(\Omega_T)} \leq C \varepsilon^{1/2} \left\{ \|u_0\|_{L^2(0, T; H^{m+1}(\Omega))} + \|f\|_{L^2(0, T; H^{-m+1}(\Omega))} + \|h\|_{L^2(\Omega)} \right\}, \quad (5.3)
$$

where $C$ depends only on $d, n, m, \mu, T$ and $\Omega$. If in addition $\Omega$ is $C^{m,1}$ and $h=0$, then

$$
\|\nabla^m w_\varepsilon\|_{L^2(\Omega_T)} \leq C \varepsilon^{1/2} \|f\|_{L^2(0, T; \tilde{H}^{-m+1}(\Omega))}, \quad (5.4)
$$

where $C$ depends only on $d, n, m, \mu, T$ and $\Omega$.

Proof of Theorem 1.2. The proof is completely parallel to that of Theorem 1.1. Indeed, let $F \in L^2(0, T; H^{-m+1}(\Omega))$. We consider the following initial-Neumann problems

$$
\begin{cases}
-\partial_t v_\varepsilon + C_\varepsilon v_\varepsilon = F & \text{in } \Omega_T, \\
N_{m-1-j}(v_\varepsilon) = 0 & \text{on } \partial\Omega \times (0, T), \quad j = 0, 1, \ldots, m-1, \\
v_\varepsilon = 0 & \text{on } \Omega \times \{t = T\},
\end{cases} \quad (5.5)
$$
and
\[
\begin{cases}
-\partial_t v_0 + L_0^* v_0 = F & \text{in } \Omega_T, \\
N_{m-1-j}(v_0) = 0 & \text{on } \partial \Omega \times (0, T), \\
v_0 = 0 & \text{on } \Omega \times \{t = T\},
\end{cases}
\]  
(5.6)

It is obvious that \(v_\varepsilon(x, T - t), v_0(x, T - t)\) are solutions respectively to (1.1) and (1.10) with homogeneous Neumann boundary data, and also with \(f(x, t) = F(x, T - t), h = 0\) and \(A(x/\varepsilon, t/\varepsilon^{2m})\) replaced by \(A^\ast(x/\varepsilon, (T - t)/\varepsilon^{2m})\). Moreover, \(v_0\) still satisfy estimates (4.14), (4.23) and (4.25).

Define \(\Phi_\varepsilon\) as (4.34). Observe that
\[
\mathcal{H} \int_0^T \left\langle F, w_\varepsilon \right\rangle_{\tilde{H}^{-m+1}(\Omega) \times \tilde{H}^{m-1}(\Omega)} dt = \mathcal{H} \int_0^T \left\langle \partial_t w_\varepsilon, v_\varepsilon(t) \right\rangle_{\tilde{H}^{-m+1}(\Omega) \times \tilde{H}^{m-1}(\Omega)} dt + \sum_{|\alpha| = |\beta| = m} \int_{\Omega_T} \tilde{A}_\varepsilon^{\alpha\beta} D^\beta w_\varepsilon D^\alpha v_\varepsilon.
\]

With Lemmas 5.1, 5.2 and Theorem 5.1 at our disposal, we can perform the same analysis as we did for Theorem 1.1 to derive (1.11).

\section*{Acknowledgments}
The authors are much obliged to Professor Zhongwei Shen for the guidance and enlightening discussions. They would also like to express deep gratitude to Professor Russell Brown and the Department of Mathematics of University of Kentucky for the support and the warm hospitality during the authors’ visit.

\section*{References}
[1] S. N. Armstrong, A. Bordas, and J. C. Mourrat, \textit{Quantitative stochastic homogenization and regularity theory of parabolic equations}, Preprint arXiv:1705.07672 (2017).
[2] M. Avellaneda and F. Lin, \textit{Compactness methods in the theory of homogenization}, Comm. Pure Appl. Math. 40 (1987), no. 6, 803–847.
[3] A. Bensoussan, J. L. Lions, and G. Papanicolaou, \textit{Asymptotic analysis for periodic structures}, vol. 5, North-Holland Publishing Company Amsterdam, 1978.
[4] S. Byun and Y. Jang, \textit{Calderón-Zygmund estimate for homogenization of parabolic systems}, Discrete Contin. Dyn. Syst. 36 (2016), no. 12, 6689–6714.
[5] H. Dong and H. Zhang, \textit{Conormal problem of higher-order parabolic systems with time irregular coefficients}, Trans. Amer. Math. Soc. 368 (2016), no. 10, 7413–7460.
[6] J. Geng and Z. Shen, \textit{Uniform regularity estimates in parabolic homogenization}, Indiana Univ. Math. J. 64 (2015), no. 3, 697–733.
[7] J. Geng and Z. Shen, \textit{Convergence rates in parabolic homogenization with time-dependent periodic coefficients}, J. Funct. Anal. 272 (2017), no. 5, 2092–2113.
[8] G. Griso, *Error estimate and unfolding for periodic homogenization*, Asymptot. Anal. 40 (2004), no. 3-4, 269–286.

[9] S. Gu, *Convergence rates in homogenization of Stokes systems*, J. Differential Equations 260 (2016), no. 7, 5796–5815.

[10] V. V. Jikov, S. M. Kozlov, and O. A. Oleinik, *Homogenization of differential operators and integral functionals*, Springer Berlin Heidelberg, 1994.

[11] C. E. Kenig, F. Lin, and Z. Shen, *Convergence rates in $L^2$ for elliptic homogenization problems*, Arch. Ration. Mech. Anal. 203 (2012), no. 3, 1009–1036.

[12] ———, *Homogenization of elliptic systems with Neumann boundary conditions*, J. Amer. Math. Soc. 26 (2013), no. 4, 901–937.

[13] ———, *Periodic homogenization of Green and Neumann functions*, Comm. Pure Appl. Math. 67 (2014), no. 8, 1219–1262.

[14] A. A. Kukushkin and T. A. Suslina, *Homogenization of high-order elliptic operators with periodic coefficients*, Algebra i Analiz 28 (2016), no. 1, 89–149.

[15] Yu. M. Meshkova and T. A. Suslina, *Homogenization of solutions of initial boundary value problems for parabolic systems*, Funct. Anal. Appl. 49 (2015), no. 1, 72–76.

[16] ———, *Homogenization of initial boundary value problems for parabolic systems with periodic coefficients*, Appl. Anal. 95 (2016), no. 8, 1736–1775.

[17] W. Niu, Z. Shen, and Y. Xu, *Convergence Rates and Interior Estimates in Homogenization of Higher Order Elliptic Systems*, Preprint, arXiv:1706.02072 (2017).

[18] W. Niu and Y. Xu, *Uniform Boundary Estimates in Homogenization of Higher Order Elliptic Systems*, Preprint, arXiv:1709.04097 (2017).

[19] S. E. Pastukhova, *Estimates in homogenization of higher-order elliptic operators*, Appl. Anal. 95 (2016), no. 7, 1449–1466.

[20] ———, *Operator error estimates for homogenization of fourth order elliptic equations*, St. Petersburg Math. J. 28 (2017), no. 2, 273–289.

[21] Z. Shen, *Boundary estimates in elliptic homogenization*, Anal. PDE 10 (2017), no. 3, 653–694.

[22] ———, *Lectures on Periodic Homogenization of Elliptic Systems*, Preprint arXiv:1710.11257 (2017).

[23] Z. Shen and J. Zhuge, *Convergence rates in periodic homogenization of systems of elasticity*, Proc. Amer. Math. Soc. 145 (2017), no. 3, 1187–1202.

[24] T. A. Suslina, *Homogenization of the parabolic Cauchy problem in the Sobolev class $H^1(\mathbb{R}^d)$*, Funct. Anal. Appl. 44 (2010), no. 4, 318–322.
[25] ______, *Homogenization of the Dirichlet problem for elliptic systems: $L^2$-operator error estimates*, Mathematika 59 (2013), no. 2, 463–476.

[26] ______, *Homogenization of the Neumann problem for elliptic systems with periodic coefficients*, SIAM J. Math. Anal. 45 (2013), no. 6, 3453–3493.

[27] ______, *Homogenization of the Dirichlet problem for higher-order elliptic equations with periodic coefficients*, Algebra i Analiz 29 (2017), 139–192.

[28] ______, *Homogenization of the Neumann problem for higher-order elliptic equations with periodic coefficients*, Preprint, arXiv1705.08295 (2017).

[29] Q. Xu and S. Zhou, *Quantitative estimates in homogenization of parabolic systems of elasticity in lipschitz cylinders*, Preprint arXiv:1705.01479 (2017).

[30] Y. Xu and W. Niu, *Convergence rates in almost-periodic homogenization of higher-order elliptic systems*, arXiv preprint arXiv:1712.01744 (2017).

[31] V. V. Zhikov and S. E. Pastukhova, *Estimates of homogenization for a parabolic equation with periodic coefficients*, Russ. J. Math. Phys. 13 (2006), no. 2, 224–237.

Weisheng Niu  
School of Mathematical Science, Anhui University, Hefei, 230601, P. R. China  
E-mail:weisheng.niu@gmail.com

Yao Xu  
Department of Mathematics, Nanjing University, Nanjing, 210093, P. R. China  
E-mail:dg1421012@smail.nju.edu.cn