COMMUN. MATH. SCI.
Vol. 20, No. 8, pp. 2107–2132

HOMOGENIZATION OF PARABOLIC SYSTEMS WITH SINGULAR PERTURBATIONS*

QING MENG† AND WEISHENG NIU‡

Abstract. We investigate convergence rates in periodic homogenization of second-order parabolic systems with fourth-order singular perturbations. Different rates depending on κ and ε, which represent respectively the strength of the singular perturbation and the scale of the heterogeneities, are obtained for the problem with Dirichlet and Navier boundary conditions.

Keywords. Homogenization; Convergence rate; Parabolic systems; Singular perturbations.

AMS subject classifications. 35B27.

1. Introduction

Let Ω be a bounded domain in $\mathbb{R}^d, d\geq 2$. We consider the quantitative homogenization of second-order parabolic systems with fourth-order perturbations,

$$\partial_t u_\varepsilon + \mathcal{L}_\varepsilon u_\varepsilon = F \qquad \text{in } \Omega\times(0,T), \tag{1.1}$$

where

$$\mathcal{L}_\varepsilon = \kappa^2\Delta^2 - \operatorname{div}(A(x/\varepsilon, t/\varepsilon^2)\nabla), \qquad 0<\kappa,\varepsilon<1.$$

We assume that $A(y,s)=(A_{ij}^{\alpha\beta}(y,s)), 1\leq i,j\leq d, 1\leq \alpha,\beta\leq n$, is real, bounded measurable and satisfies the ellipticity condition,

$$\mu|\xi|^2 \leq A_{ij}^{\alpha\beta}(y,s)\xi_i^\alpha\xi_j^\beta \leq \frac{1}{\mu}|\xi|^2 \tag{1.2}$$

for any $\xi=(\xi_i^\alpha)\in\mathbb{R}^{n\times d}$ and a.e. $(y,s)\in\mathbb{R}^{d+1}$, where $\mu>0$. Furthermore, we assume that A is 1-periodic in (y,s), i.e.,

$$A(y+z,s+\tau)=A(y,s) \;\text{ for any } (z,\tau)\in\mathbb{Z}^{d+1} \text{ and } a.e.\ (y,s)\in\mathbb{R}^{d+1}. \tag{1.3}$$

Investigations on homogenization of partial differential equations with singular perturbations goes back to 1970s. In [2], Bensoussan, Lions, and Papanicolaou established the qualitative homogenization theory of (1.1) and the associated elliptic problems

$$\kappa^2\Delta^2 v_\varepsilon - \operatorname{div}(A(x/\varepsilon)\nabla v_\varepsilon) = F(x) \quad \text{in } \Omega, \quad 0<\varepsilon<1, \tag{1.4}$$

with $\kappa=\varepsilon$. Later on, in [5] Francfort and Müller conducted systematic studies on qualitative homogenization of (1.4) and the related nonlinear functionals for the case $\kappa=\varepsilon^\lambda, 0<\lambda<\infty$. The results in [2,5] show that singular perturbations play an essential role in determining the coefficients of the effective problems. With the aim to quantify the combined effect of singular perturbations and homogenization, we investigated the quantitative homogenization theory of elliptic systems in the form of (1.4) in [14,17].

*Received: January 31, 2021; Accepted (in revised form): February 23, 2022. Communicated by Markos A. Katsoulakis.

†School of Mathematical Science, Anhui University, Hefei, 230601, China (mengqing2020126@126.com).

‡School of Mathematical Science, Anhui University, Hefei, 230601, China (niuwsh@ahu.edu.cn).

Particularly, the convergence rates, which depend on the scale κ that represents the strength of the singular perturbation and on the length scale ε of the heterogeneities, were established in [14]. More recently, the convergence rate for (1.4) with $\kappa=\varepsilon$ in $\mathbb{R}^d$ was studied in [18]. The aim of this paper is to extend our previous results in [14] to the parabolic settings.

Let $\Omega_T=\Omega\times(0,T)$, and $\Gamma_T=\partial\Omega\times(0,T)$. For $F\in L^2(\Omega_T)$, $h\in L^2(\Omega)$, let u_ε be the weak solution of the problem

$$\partial_t u_\varepsilon+\mathcal{L}_\varepsilon u_\varepsilon=F \quad \text{in } \Omega_T, \qquad u_\varepsilon=h \quad \text{on } \Omega\times\{t=0\} \tag{1.5}$$

with Dirichlet boundary conditions

$$u_\varepsilon=0, \quad \frac{\partial u_\varepsilon}{\partial\nu}=\nabla u_\varepsilon\cdot\nu=0 \quad \text{on } \Gamma_T \tag{1.6}$$

(where ν is the outward unit normal to $\partial\Omega$), which means that $u_\varepsilon\in L^2(0,T;H_0^2(\Omega))\cap L^\infty(0,T;L^2(\Omega))$,

$$\begin{aligned}&-\int_{\Omega_T} u_\varepsilon\partial_t\phi\,dxdt+\kappa^2\int_{\Omega_T}\Delta u_\varepsilon\Delta\phi\,dxdt+\int_{\Omega_T}A(x/\varepsilon,t/\varepsilon^2)\nabla u_\varepsilon\nabla\phi\,dxdt\\ &=\int_0^T\langle F,\phi\rangle_{H^{-2}(\Omega)\times H_0^2(\Omega)}dt+\int_\Omega h\phi(0)dx\end{aligned}$$

for any $\phi\in C_c^\infty(\Omega\times[0,T))$. Suppose that $\kappa=\kappa(\varepsilon)$ satisfies the assumption

$$\kappa\to 0 \text{ as } \varepsilon\to 0, \quad \text{and} \quad \lim_{\varepsilon\to 0}\frac{\kappa}{\varepsilon}=\rho. \tag{1.7}$$

Under conditions (1.2), (1.3) and (1.7), we prove that as $\varepsilon\to 0$ the weak solution u_ε of (1.5) and (1.6) converges weakly in $L^2(0,T;H^1(\Omega))$ and strongly in $L^2(0,T;L^2(\Omega))$ to the solution u_0 of the following problem

$$\begin{aligned}&\partial_t u_0-\operatorname{div}(\widehat{A}\nabla u_0)=F \quad \text{in } \Omega_T,\\ &u_0=h \quad \text{on } \Omega\times\{t=0\}, \text{ and } u_0=0 \quad \text{on } \Gamma_T,\end{aligned} \tag{1.8}$$

where the effective coefficient $\widehat{A}$ depends on ρ in three cases: $\rho=\infty, 0<\rho<\infty$ and $\rho=0$ (see Section 3 for the details).

Our first result can be stated as follows.

THEOREM 1.1. *Let Ω be a bounded $C^{1,1}$ domain in $\mathbb{R}^d, d\geq 2$ and A satisfy (1.2)-(1.3). Suppose (1.7) holds, and if $\rho=0$ we also assume that $A(y,s)$ is Lipschitz continuous in y, i.e.,*

$$|A(y_1,s)-A(y_2,s)|\leq L|y_1-y_2|, \text{ for any } y_1,y_2\in\mathbb{R}^d \text{ and } s>0. \tag{1.9}$$

Let u_ε be the solution to (1.5) and (1.6), and u_0 the solution to (1.8). Then

$$\begin{aligned}\|u_\varepsilon-u_0\|_{L^2(\Omega_T)}&\leq\left\{\|F\|_{L^2(\Omega_T)}+\|h\|_{H_0^1(\Omega)}\right\}\\ &\times\begin{cases}C_1\Big(\kappa+\varepsilon+(\varepsilon/\kappa)^2\Big) & \text{if } \rho=\infty,\\ C_2\Big(\kappa+\varepsilon+\rho^{-2}|\rho^2-(\kappa/\varepsilon)^2|\Big) & \text{if } 0<\rho<\infty,\\ C_3\Big(\kappa+\varepsilon+(\kappa/\varepsilon)^2\Big) & \text{if } \rho=0,\end{cases}\end{aligned} \tag{1.10}$$

where C_1,C_2 depend only on d,n,μ,Ω and T, while C_3 depends only on d,n,μ,Ω,T and L.

Note that the error estimate in (1.10) involves three terms. The first term κ is due to the singular perturbation, and the second term ε by homogenization, while the third term is generated by the error between $\widehat{A}$ and $\widehat{A^\lambda}$ (see (3.5) for the definition of $\widehat{A^\lambda}$). There are one-dimensional examples in the elliptic case [6], which show that the perturbation error $O(\kappa)$ is optimal. It is well known that homogenization error $O(\varepsilon)$ is also optimal. Moreover, our estimate on $|\widehat{A}-\widehat{A^\lambda}|$ should also be sharp as $\lambda\to 0$ or $\lambda\to\infty$. Therefore, the convergence rate in (1.10) should be optimal.

Our next theorem provides the error estimate for the problem with the Navier boundary condition

$$u_\varepsilon=\Delta u_\varepsilon=0 \quad \text{on } \Gamma_T. \tag{1.11}$$

THEOREM 1.2. *Let Ω be a bounded $C^{1,1}$ domain in $\mathbb{R}^d, d\geq 2$ and A satisfy (1.2)-(1.3). Suppose (1.7) holds, and if $\rho=0$ we also assume that $A(y,s)$ satisfies (1.9). Let u_ε be the solution to (1.5) and (1.11), and u_0 the solution to (1.8). Then*

$$\begin{aligned}\|u_\varepsilon-u_0\|_{L^2(\Omega_T)}\leq &\left\{\|F\|_{L^2(\Omega_T)}+\|h\|_{H^2(\Omega)}\right\}\\ &\times\begin{cases} C_1\Big(\kappa^2+\varepsilon+(\varepsilon/\kappa)^2\Big) & \text{if } \rho=\infty,\\ C_2\Big(\kappa^2+\varepsilon+\rho^{-2}|\rho^2-(\kappa/\varepsilon)^2|\Big) & \text{if } 0<\rho<\infty,\\ C_3\Big(\kappa^2+\varepsilon+(\kappa/\varepsilon)^2\Big) & \text{if } \rho=0,\end{cases}\end{aligned} \tag{1.12}$$

where C_1,C_2 depend only on d,n,μ,Ω and T, while C_3 depends only on d,n,μ,Ω,T, and L in (1.9).

The difference of the convergence rates in (1.10) and (1.12) is due to the variance of optimal convergence rates for the singular perturbation problem

$$\partial_t u_\kappa+\kappa^2\Delta^2u_\kappa-\operatorname{div}(A(x,t)\nabla u_\kappa)=F \quad \text{in } \Omega_T, \quad u_\kappa=h \ \text{ on } \Omega\times\{t=0\} \tag{1.13}$$

supplemented with different (Dirichlet and Navier) boundary conditions. We note that convergence rates in singular perturbations (without homogenization) of elliptic and parabolic equations have been studied deeply [3, 6, 11–13, 19]. For problem (1.13) with Dirichlet boundary conditions, the $O(\kappa^{1/2})$ convergence rate was obtained in [12], while the interior $O(\kappa)$ convergence rate was obtained in [6]. Later on, Schuss established the $O(\kappa)$ rate in [19] for the case $d=2$ and $A=I$. It is also worth remarking that the example in [6] shows that the $O(\kappa)$ rate is optimal for the initial-Dirichlet problems. Yet, in Theorem 2.2 we shall prove that the optimal convergence rate should be $O(\kappa^2)$ for problem (1.13) with the Navier boundary condition (1.11). Moreover, our proof of Theorem 2.2 is quite different from the one in [19].

We recall that the convergence rate in homogenization of parabolic equations (systems) has been studied intensively. For the case $\kappa=0$, the $O(\varepsilon)$-order convergence rate in homogenization of Equation (1.5) has recently been derived in [8]. The result was then extended to higher order parabolic systems in [15], and to second order parabolic systems in non-smooth cylinders in [16, 24]. See also [1, 7, 9, 10, 23, 25] for more related results. Compared with the previous works for parabolic systems, the main difficulty we encounter here is that the original equation and its limit have different orders, and therefore have different number of boundary conditions. Moreover, the scaling of the singular perturbation is different from the other parts of the operators.

To prove Theorem 1.1, we first introduce a family of λ-dependent operators

$$\partial_t + \mathcal{L}_\varepsilon^\lambda = \partial_t + \lambda^2\varepsilon^2\Delta^2 - \operatorname{div}(A(x/\varepsilon, t/\varepsilon^2)\nabla), \tag{1.14}$$

for which the ε-scaling is the same in each part of the operator. Let $\partial_t - \operatorname{div}(\widehat{A^\lambda}\nabla)$ be the homogenized operator of $\partial_t + \mathcal{L}_\varepsilon^\lambda$, where $\widehat{A^\lambda}$ is given by (3.5). Let $u_{\varepsilon,\lambda}$ be the solution to

$$(\partial_t + \mathcal{L}_\varepsilon^\lambda)u_{\varepsilon,\lambda} = F \quad \text{in } \Omega_T, \quad u_{\varepsilon,\lambda}(x,0) = h \quad \text{in } \Omega\times\{t=0\}$$

with homogeneous Dirichlet boundary conditions, and $u_{0,\lambda}$ the solution to the homogenized problem. We first investigate the convergence rate in the singular perturbation of the parabolic systems with periodic boundary conditions (see Section 2), by which we establish the error estimate between $\widehat{A}$ and $\widehat{A^\lambda}$ (see Section 3 for the details), and moreover, the bound of $\|u_{0,\lambda} - u_0\|_{L^2(\Omega_T)}$. Then we introduce proper auxiliary function (see (4.6) in Section 4 for the meaning of each term)

$$w_{\varepsilon,\lambda}(x,t) = u_{\varepsilon,\lambda} - u_{0,\lambda}\eta_{1,\delta} - \varepsilon(\chi^\lambda)^\varepsilon K_\varepsilon(\nabla u_{0,\lambda}) + \varepsilon^2(\mathfrak{B}^\lambda_{k(d+1)j})^\varepsilon \frac{\partial}{\partial x_k} K_\varepsilon(\frac{\partial u_{0,\lambda}}{\partial x_j}),$$

which helps us overcome the difficulty caused by the difference of the orders of the original equations and its limit (see Section 4 for more explanations). Finally, by performing the two-scale expansion and adapting the duality argument originated in [22] (see also [8, 20, 21]), we derive the estimate of $\|u_{\varepsilon,\lambda} - u_{0,\lambda}\|_{L^2(\Omega_T)}$, which, together with the bound of $\|u_{0,\lambda} - u_0\|_{L^2(\Omega_T)}$ and the observation $u_{\varepsilon,\lambda} = u_\varepsilon$ if $\lambda = \kappa/\varepsilon$, gives the estimate 1.1.

The proof of Theorem 1.2 is slightly different. The key point is to deal with convergence rates of singular perturbation and homogenization separately. To this aim, besides (1.14) we also introduce an intermediate problem

$$\begin{aligned} &(\partial_t + \mathcal{L}_0^\lambda)v_{\varepsilon,\lambda} + \lambda^2\varepsilon^2\Delta^2 v_{\varepsilon,\lambda} = F \quad \text{in } \Omega_T,\\ &v_{\varepsilon,\lambda} = h \ \text{ on } \Omega\times\{t=0\} \quad \text{and} \quad v_{\varepsilon,\lambda} = \Delta v_{\varepsilon,\lambda} = 0 \ \text{ on } \Gamma_T, \end{aligned} \tag{1.15}$$

where $\mathcal{L}_0^\lambda$ is given by (3.4) and (3.5). Let $v_{0,\lambda}$ be the unique solution to the limit problem (as $\varepsilon \to 0$) of (1.15), i.e.,

$$(\partial_t + \mathcal{L}_0^\lambda)v_{0,\lambda} = F \quad \text{in } \Omega_T$$

with $v_{0,\lambda} = h$ on $\Omega\times\{t=0\}$ and $v_{0,\lambda} = 0$ on Γ_T. We first establish the error estimate in singular perturbation, i.e., the error estimate between $v_{\varepsilon,\lambda}$ and $v_{0,\lambda}$ (see Theorem 2.2). Then we consider the convergence rate in pure homogenization, i.e., the error estimate between $u_{\varepsilon,\lambda}$ and $v_{\varepsilon,\lambda}$. This step is quite similar to the proof of Theorem 1.1, except the auxiliary functions. Note that the error estimate between u_0 and $v_{0,\lambda}$ follows from the estimate on $|\widehat{A} - \widehat{A^\lambda}|$ and standard energy estimates. We can finally complete the proof of Theorem 1.2 by using the triangle inequality and setting $\lambda = \kappa/\varepsilon$ as in Theorem 1.1.

2. Singular perturbations

2.1. Periodic boundary conditions. Let $C_p^\infty(\mathbb{R}^d)$ be the space of C^∞, 1-periodic vector valued functions in $\mathbb{R}^d$. For $k\geq 0$ and $\mathbb{T}^d=[0,1]^d$, let $H_p^k(\mathbb{T}^d)$ denote the closure of $C_p^\infty(\mathbb{R}^d)$ in $H^k(\mathbb{T}^d)$, and $\dot{H}_p^k(\mathbb{T}^d)$ the subspace of $H_p^k(\mathbb{T}^d)$ with zero spatial mean, i.e., $\int_{\mathbb{T}^d} v\,dx=0$. In particular, $\dot{H}_p^0(\mathbb{T}^d)=\dot{L}_p^2(\mathbb{T}^d)$.

Consider the operator

$$\partial_t+\mathcal{L}^\lambda=\partial_t+\lambda^2\Delta^2-\operatorname{div}(A(x,t)\nabla), \tag{2.1}$$

where A is 1-periodic in x and satisfies the ellipticity condition (1.2). For $F\in L^2(0,T;\dot{L}_p^2(\mathbb{T}^d)), h\in\dot{L}_p^2(\mathbb{T}^d)$, let $u_\lambda\in L^2(0,T;\dot{H}_p^2(\mathbb{T}^d))$ be the unique solution to

$$\partial_t u_\lambda+\mathcal{L}^\lambda u_\lambda=F \quad \text{in } \mathbb{T}^d\times(0,T) \quad \text{and } u_\lambda(x,0)=h, \tag{2.2}$$

and $u_0\in L^2(0,T;\dot{H}_p^1(\mathbb{T}^d))$ the unique solution to

$$\partial_t u_0-\operatorname{div}(A(x,t)\nabla u_0)=F \quad \text{in } \mathbb{T}^d\times(0,T) \quad \text{and } u_0(x,0)=h. \tag{2.3}$$

LEMMA 2.1. *Assume that A satisfies (1.2) and is 1-periodic in x. Let u_λ,u_0 be, respectively, the weak solutions to (2.2) and (2.3), and $u_0\in L^2(0,T;\dot{H}^2(\mathbb{T}^d))$. Then*

$$\|\nabla u_\lambda-\nabla u_0\|_{L^2(\mathbb{T}^d\times(0,T))}\leq C\lambda\|u_0\|_{L^2(0,T;H^2(\mathbb{T}^d))}, \tag{2.4}$$

where C depends only on d,n,μ and T.

Proof. Let $w=u_\lambda-u_0$. Then

$$(\partial_t+\mathcal{L}^\lambda)w=-\lambda^2\Delta^2 u_0.$$

Thus for any $\psi\in L^2(0,T;\dot{H}_p^2(\mathbb{T}^d))$,

$$|\langle(\partial_t+\mathcal{L}^\lambda)w,\psi\rangle|\leq\lambda^2\|\Delta u_0\|_{L^2(0,T;L^2(\mathbb{T}^d))}\|\Delta\psi\|_{L^2(0,T;L^2(\mathbb{T}^d))}. \tag{2.5}$$

By taking $\psi=w$ in (2.5) and using the Cauchy inequality, we obtain

$$\lambda\|\Delta w\|_{L^2(\mathbb{T}^d\times(0,T))}+\|\nabla w\|_{L^2(\mathbb{T}^d\times(0,T))}\leq C\lambda\|u_0\|_{L^2(0,T;H^2(\mathbb{T}^d))}, \tag{2.6}$$

which gives (2.4). □

THEOREM 2.1. *Suppose A is 1-periodic in x and satisfies the assumptions (1.2) and (1.9). Let u_λ and u_0 be, respectively, the weak solutions to (2.2) and (2.3). Then*

$$\|u_\lambda-u_0\|_{L^2(0,T;L^2(\mathbb{T}^d))}\leq C\lambda^2\big\{\|F\|_{L^2(0,T;L^2(\mathbb{T}^d))}+\|h\|_{L^2(\mathbb{T}^d)}\big\}, \tag{2.7}$$

where C depends on d,n,μ,T, and L in (1.9).

Proof. For $H\in L^2(0,T;L^2(\mathbb{T}^d))$, let $v_\lambda\in L^2(0,T;\dot{H}_p^2(\mathbb{T}^d))$ be the weak solution to the problem

$$-\partial_t v_\lambda+\mathcal{L}^{\lambda*}v_\lambda=H \quad \text{in } \mathbb{T}^d\times(0,T) \quad \text{and} \quad v_\lambda=0 \quad \text{on } \mathbb{T}^d\times\{t=T\},$$

where $\mathcal{L}^{\lambda*}=\lambda^2\Delta^2-\operatorname{div}(A^*\nabla)$ with $A^*=((A_{ij}^{\alpha\beta})^*)=(A_{ji}^{\beta\alpha}),1\leq i,j\leq d,1\leq\alpha,\beta\leq n$. Let $v_0\in L^2(0,T;\dot{H}_p^1(\mathbb{T}^d))$ be the unique solution to the limit problem

$$-\partial_t v_0-\operatorname{div}(A^*\nabla v_0)=H \quad \text{in } \mathbb{T}^d\times(0,T) \quad \text{and} \quad v_0=0 \quad \text{on } \mathbb{T}^d\times\{t=T\}.$$

Then $v_\lambda(T-t)$ and $v_0(T-t)$ are the solutions to (2.2) and (2.3) with $h=0, F=H(x,T-t)$ and $A(x,t)$ replaced by $A^*(x,(T-t))$. Define $\widetilde{w}(x,t)=v_\lambda(x,T-t)-v_0(x,T-t)$. We derive from (2.6) that

$$\lambda\|\Delta\widetilde{w}\|_{L^2(\mathbb{T}^d\times(0,T))}+\|\nabla\widetilde{w}\|_{L^2(\mathbb{T}^d\times(0,T))}\leq C\lambda\|v_0\|_{L^2(0,T;H^2(\mathbb{T}^d))}. \tag{2.8}$$

Note that

$$\begin{aligned}\Big|\int_0^T\int_{\mathbb{T}^d} w\cdot H\,dxdt\Big|&=\Big|\int_0^T\langle(\partial_t+\mathcal{L}^\lambda)w,v_\lambda(t)\rangle\,dt\Big|\\&\leq\Big|\int_0^T\langle(\partial_t+\mathcal{L}^\lambda)w,\widetilde{w}(T-t)\rangle dt\Big|+\Big|\int_0^T\langle(\partial_t+\mathcal{L}^\lambda)w,v_0(t)\rangle dt\Big|.\end{aligned} \tag{2.9}$$

By (2.5), (2.6), and (2.8),

$$\begin{aligned}\Big|\int_0^T\langle(\partial_t+\mathcal{L}^\lambda)w,\widetilde{w}(T-t)\rangle dt\Big|&\leq\lambda^2\|\Delta u_0\|_{L^2(0,T;L^2(\mathbb{T}^d))}\|\Delta\widetilde{w}\|_{L^2(0,T;L^2(\mathbb{T}^d))}\\&\leq C\lambda^2\|u_0\|_{L^2(0,T;H^2(\mathbb{T}^d))}\|v_0\|_{L^2(0,T;H^2(\mathbb{T}^d))},\end{aligned} \tag{2.10}$$

and

$$\Big|\int_0^T\langle(\partial_t+\mathcal{L}^\lambda)w,v_0\rangle\Big|\leq\lambda^2\|\Delta u_0\|_{L^2(0,T;L^2(\mathbb{T}^d))}\|\Delta v_0\|_{L^2(0,T;L^2(\mathbb{T}^d))}. \tag{2.11}$$

Since A satisfies (1.9), we have the following H^2 estimates

$$\begin{aligned}&\|u_0\|_{L^2(0,T;H^2(\mathbb{T}^d))}\leq C\big\{\|F\|_{L^2(0,T;L^2(\mathbb{T}^d))}+\|h\|_{L^2(\mathbb{T}^d)}\big\},\\&\|v_0\|_{L^2(0,T;H^2(\mathbb{T}^d))}\leq C\|H\|_{L^2(0,T;L^2(\mathbb{T}^d))},\end{aligned}$$

where C depends on d,n,μ, and L in (1.9). From (2.9)–(2.11), it follows that

$$\Big|\int_0^T\int_{\mathbb{T}^d} w\cdot H\,dxdt\Big|\leq C\lambda^2\big\{\|F\|_{L^2(0,T;L^2(\mathbb{T}^d))}+\|h\|_{L^2(\mathbb{T}^d)}\big\}\|H\|_{L^2(0,T;L^2(\mathbb{T}^d))},$$

which, by duality, gives (2.7). □

REMARK 2.1. Let $F(x,t)$ be 1-periodic in (x,t), and A satisfy (1.2), (1.3) and (1.9). Let u_λ be the solution to

$$\begin{cases}\partial_t u_\lambda+\mathcal{L}^\lambda(u_\lambda)=F & \text{in } Q_1=\mathbb{T}^d\times(0,1),\\ u_\lambda(x,t) \text{ is 1-periodic in } (x,t), & \int_0^1\int_{\mathbb{T}^d}u_\lambda(x,t)dxdt=0,\end{cases}$$

and let u_0 be the unique solution to the limit (as $\lambda\to 0$) problem, i.e,

$$\begin{cases}\partial_t u_0-\operatorname{div}(A(x,t)\nabla u_0)=F & \text{in } Q_1,\\ u_0(x,t) \text{ is 1-periodic in } (x,t), & \int_0^1\int_{\mathbb{T}^d}u_0(x,t)dxdt=0.\end{cases}$$

With slight modifications on the proofs of Lemma 2.1 and Theorem 2.1, we can prove that

$$\|u_\lambda-u_0\|_{L^2(\mathbb{T}^{d+1})}\leq C\lambda^2\|u_0\|_{L^2(0,1;H^2(\mathbb{T}^d))}. \tag{2.12}$$

2.2. Navier boundary conditions. For $h\in L^2(\Omega),F\in L^2(\Omega_T)$, let u_λ be a weak solution to

$$\begin{aligned}&\partial_t u_\lambda+\mathcal{L}^\lambda u_\lambda=F \quad \text{in } \Omega_T,\\&u_\lambda=h \ \text{ on } \Omega\times\{t=0\} \quad \text{and} \quad u_\lambda=\Delta u_\lambda=0 \ \text{ on } \Gamma_T,\end{aligned}\tag{2.13}$$

where $\mathcal{L}_\lambda$ is defined in (2.1), and u_0 the solution to

$$\begin{aligned}&\partial_t u_0-\operatorname{div}(A(x,t)\nabla u_0)=F \quad \text{in } \Omega_T,\\&u_0=h \ \text{ on } \Omega\times\{t=0\} \quad \text{and} \quad u_0=0 \ \text{ on } \Gamma_T.\end{aligned}\tag{2.14}$$

We investigate the convergence rate of u_λ to u_0 as λ tends to zero.

LEMMA 2.2. *Assume that A satisfies conditions (1.2). Let u_λ,u_0 be, respectively, the weak solutions to (2.13) and (2.14), and $u_0\in L^2(0,T;H^2(\Omega))$. Then*

$$\|\nabla u_\lambda-\nabla u_0\|_{L^2(\Omega_T)}\le C\lambda\|u_0\|_{L^2(0,T;H^2(\Omega))},\tag{2.15}$$

where C depends only on d,n,μ and T.

Proof. Note that

$$\partial_t(u_\lambda-u_0)-\operatorname{div}(A\nabla(u_\lambda-u_0))+\lambda^2\Delta^2 u_\lambda=0.$$

Since $u_\lambda=\Delta u_\lambda=0$ on Γ_T. For any $\psi\in L^2(0,T;H_0^1(\Omega)\cap H^2(\Omega))$, we have

$$\int_0^T\langle\partial_t(u_\lambda-u_0)-\operatorname{div}(A\nabla(u_\lambda-u_0)),\psi\rangle dt+\lambda^2\int_{\Omega_T}\Delta u_\lambda\Delta\psi\,dxdt=0.$$

By setting $\psi=u_\lambda-u_0$, we get

$$\|\nabla(u_\lambda-u_0)\|^2_{L^2(\Omega_T)}+\lambda^2\|\Delta u_\lambda\|^2_{L^2(\Omega_T)}\le C\lambda^2\|\Delta u_0\|^2_{L^2(\Omega_T)},\tag{2.16}$$

from which (2.15) follows. □

REMARK 2.2. The proof above doesn't seem to work directly for the problem with Dirichlet boundary conditions. Note that in this case $u_\lambda=\nabla u_\lambda\cdot\nu=0$ on the boundary, but ∇u_0 does not necessarily equal to 0 on the boundary. To derive the convergence rate, one may first consider the error estimate between u_λ and $u_0\zeta_\varepsilon$ for proper smooth cut-off functions ζ_ε, see e.g., (4.6). Then by some boundary layer estimate and the triangle inequality to derive the error estimate between u_λ and u_0. It is also worth remarking that due to the deviations of ∇u_λ and ∇u_0 on the boundary, one can not expect the first order error $\|u_\lambda-u_0\|_{L^2(0,T;H_0^1(\Omega))}$ with the sharp order $O(\lambda)$ [6].

THEOREM 2.2. Assume that A satisfies conditions (1.2) and (1.9). Let u_λ,u_0 be, respectively, the weak solutions to (2.13) and (2.14), and $u_0\in L^2(0,T;H^2(\Omega))$. Then

$$\|u_\lambda-u_0\|_{L^2(\Omega_T)}\le C\lambda^2\|u_0\|_{L^2(0,T;H^2(\Omega))},\tag{2.17}$$

where C depends only on d,n,μ,Ω,T, and L in (1.9).

Proof. For $H\in C_0^\infty(\Omega_T)$, let $v_\lambda\in L^2(0,T;H^2(\Omega))$ be the weak solution to the problem

$$-\partial_t v_\lambda+\mathcal{L}^{\lambda *}v_\lambda=H \ \text{ in } \Omega_T,$$

$$v_\lambda = 0 \quad \text{on } \Omega\times\{t=T\}, \quad v=\Delta v = 0 \quad \text{on } \Gamma_T,$$

where $\mathcal{L}^{\lambda *} = \lambda^2\Delta^2 - \operatorname{div}(A^*\nabla)$ with $A^* = ((A_{ij}^{\alpha\beta})^*) = (A_{ji}^{\beta\alpha}), 1\le i,j\le d, 1\le \alpha,\beta\le n$. Let $v_0 \in L^2(0,T;H_0^1(\Omega))$ be the unique solution to the limit problem

$$-\partial_t v_0 - \operatorname{div}(A^*\nabla v_0) = H \quad \text{in } \Omega_T \quad \text{and} \quad v_0 = 0 \quad \text{on } \Omega\times\{t=T\}.$$

Then $v_\lambda(T-t)$ and $v_0(T-t)$ are the solutions to (2.13) and (2.14) with $h=0$ and $A(x,t)$ replaced by $A^*(x,(T-t))$. In view of (2.16), we have

$$\|\nabla(v_\lambda - v_0)\|_{L^2(\Omega_T)}^2 + \lambda^2\|\Delta v_\lambda\|_{L^2(\Omega_T)}^2 \le C\lambda^2 \|\Delta v_0\|_{L^2(\Omega_T)}^2. \tag{2.18}$$

Let $w = u_\lambda - u_0$ and $\widetilde{w}(x,t) = v_\lambda(x,T-t) - v_0(x,T-t)$. Note that

$$\begin{aligned}\Big|\int_0^T \langle w, H\rangle dt\Big| &\le \Big|\int_0^T \langle \partial_t w - \operatorname{div}(A\nabla w), v_0\rangle dt\Big| \\ &\le \Big|\int_0^T \langle \partial_t w - \operatorname{div}(A\nabla w), \widetilde{w}(T-t)\rangle dt\Big| + \Big|\int_0^T \langle \partial_t w - \operatorname{div}(A\nabla w), v_\lambda(t)\rangle dt\Big| \\ &\le C\lambda^2 \|\Delta u_\lambda\|_{L^2(\Omega_T)}\big\{\|\nabla \widetilde{w}\|_{L^2(\Omega_T)} + \|\Delta v_\lambda\|_{L^2(\Omega_T)}\big\}. \end{aligned}\tag{2.19}$$

This, together with (2.16), (2.18) and the H^2 estimates for v_0 (see [4] for the H^2 estimates of parabolic systems in nondivergence form in $\mathbb{R}_+^d \times \mathbb{R}^+$. One can derive the estimate for v_0 by using standard extension and covering argument)

$$\|\nabla^2 v_0\|_{L^2(\Omega_T)} \le C\|H\|_{L^2(\Omega_T)},$$

gives (2.17) by duality. ∎

3. Qualitative homogenization

3.1. Correctors and the effective problem. The aim of this part is to investigate qualitative homogenization of (1.1), for which the effective problem has been recognized in [2] for $\kappa = \varepsilon$. Denote $\kappa\varepsilon^{-1}$ as $\lambda = \lambda(\varepsilon)$. Then $0<\lambda<\infty$ and the first equation in (1.1) can be rewritten as

$$\partial_t u_{\varepsilon,\lambda} + \mathcal{L}_\varepsilon^\lambda u_{\varepsilon,\lambda} = F, \tag{3.1}$$

with

$$\mathcal{L}_\varepsilon^\lambda = \lambda^2\varepsilon^2\Delta^2 - \operatorname{div}(A(x/\varepsilon, t/\varepsilon^2)\nabla).$$

For fixed $0\le\lambda<\infty$ and $1\le\beta\le n$, $1\le j\le d$, let $\chi_j^{\lambda,\beta} = (\chi_j^{\lambda,1\beta},\dots,\chi_j^{\lambda,n\beta})$ be the unique weak solution to the following cell problem

$$\begin{cases} \partial_s \chi_j^{\lambda,\beta} + \lambda^2\Delta^2\chi_j^{\lambda,\beta} - \operatorname{div}\big[A(y,s)\nabla(P_j^\beta + \chi_j^{\lambda,\beta})\big] = 0 \quad \text{in } \mathbb{R}^{d+1}, \\ \chi_j^{\lambda,\beta}(y,s) \text{ is 1-periodic in } (y,s), \\ \int_{\mathbb{T}^{d+1}} \chi_j^{\lambda,\beta}(y,s)\,dyds = 0, \end{cases} \tag{3.2}$$

where $P_j^\beta = y_j(0...1...0)$ with 1 in the βth position. Note that $\int_{\mathbb{T}^d}\chi^\lambda(y,s)dy = 0$ for $s\in\mathbb{R}$. By standard energy estimates and Poincaré's inequality, we have for $0<\lambda<\infty$,

$$\begin{aligned} &\|\chi^\lambda\|_{L^2(\mathbb{T}^{d+1})} + \|\nabla\chi^\lambda\|_{L^2(\mathbb{T}^{d+1})} \le C(1+\lambda)^{-2}, \\ &\|\nabla^2\chi^\lambda\|_{L^2(\mathbb{T}^{d+1})} \le C\lambda^{-1}(1+\lambda)^{-1}, \quad \|\nabla^3\chi^\lambda\|_{L^2(\mathbb{T}^{d+1})} \le C\lambda^{-2} \end{aligned}\tag{3.3}$$

for some constant C depending only on d,n,μ. On the other hand, for $\lambda=0$ one has $\|\nabla\chi^0\|_{L^2(\mathbb{T}^{d+1})}\leq C$.

Thanks to [2], for each fixed $\lambda\geq 0$ the homogenized operator of $\partial_t+\mathcal{L}_\varepsilon^\lambda$ is given by

$$\partial_t+\mathcal{L}_0^\lambda=\partial_t-\operatorname{div}(\widehat{A^\lambda}\nabla), \tag{3.4}$$

where the effective coefficient matrix $\widehat{A^\lambda}=(\widehat{A^\lambda}_{ij}^{\alpha\beta})$ with $1\leq\alpha,\beta\leq n$, $1\leq i,j\leq d$, is defined as

$$\widehat{A^\lambda}_{ij}^{\alpha\beta}=\fint_{\mathbb{T}^{d+1}}[A_{ij}^{\alpha\beta}(y,s)+A_{ik}^{\alpha\gamma}(y,s)\partial_{y_k}\chi_j^{\lambda,\gamma\beta}(y,s)]dyds. \tag{3.5}$$

Here and henceforth we use $\fint_E u$ to denote the L^1 average of u over the set E, i.e., $\fint_E u=\frac{1}{|E|}\int_E u$.

Lemma 3.1. *The matrix $\widehat{A^\lambda}$ is bounded and satisfies the condition (1.2) with $\frac{1}{\mu}$ replaced by some constant μ_0 depending only on $d,n,$ and μ.*

Proof. By the estimates of χ^λ in (3.3), we know that $|\widehat{A^\lambda}|\leq\mu_0$ with μ_0 depending only on $d,n,$ and μ. On the other hand, observe that for any matrix $\xi=(\xi_j^\beta)\in\mathbb{R}^{n\times d}$,

$$\begin{aligned}\langle\widehat{A^\lambda}\xi,\xi\rangle=&\fint_{\mathbb{T}^{d+1}}A_{\ell k}^{\zeta\gamma}\partial_{y_k}(\xi_j^\beta P_j^{\gamma\beta}+\xi_j^\beta\chi_j^{\lambda,\gamma\beta})\partial_{y_\ell}(\xi_i^\alpha P_i^{\zeta\alpha}+\xi_i^\alpha\chi_i^{\lambda,\zeta\alpha})dyds\\&+\lambda^2\fint_{\mathbb{T}^{d+1}}\xi_j^\beta\Delta\chi_j^{\lambda,\beta}\xi_i^\alpha\Delta\chi_i^{\lambda,\alpha}dyds+\int_0^1\langle\partial_s\chi_j^{\lambda,\beta}\xi_j^\beta,\chi_i^{\lambda,\alpha}\xi_i^\alpha\rangle ds.\end{aligned} \tag{3.6}$$

Note that

$$\int_0^1\langle\partial_s\chi_j^{\lambda,\beta}\xi_j^\beta,\chi_i^{\lambda,\alpha}\xi_i^\alpha\rangle ds=0.$$

We therefore obtain from (3.6) that

$$\begin{aligned}\langle\widehat{A^\lambda}\xi,\xi\rangle\geq&\fint_{\mathbb{T}^{d+1}}A_{\ell k}^{\zeta\gamma}\partial_{y_k}(\xi_j^\beta P_j^{\gamma\beta}+\xi_j^\beta\chi_j^{\lambda,\gamma\beta})\partial_{y_\ell}(\xi_i^\alpha P_i^{\zeta\alpha}+\xi_i^\alpha\chi_i^{\lambda,\zeta\alpha})dyds.\\\geq&\mu\fint_{\mathbb{T}^{d+1}}|\nabla(\xi_j^\beta P_j^\beta+\xi_j^\beta\chi_j^{\lambda,\beta})|^2dyds\\=&\mu|\xi|^2+\mu\fint_{\mathbb{T}^{d+1}}|\nabla(\xi_i^\beta\chi_i^{\lambda\beta})|^2dyds\\\geq&\mu|\xi|^2,\end{aligned} \tag{3.7}$$

where we have used the fact $\int_{\mathbb{T}^{d+1}}\nabla\chi^\lambda dyds=0$ for the third step. ∎

Define

$$\widehat{A}=\begin{cases}\overline{A}=\fint_{\mathbb{T}^{d+1}}A(y,s)dyds & \text{if } \rho=\infty,\\ \widehat{A^\rho} & \text{if } 0\leq\rho<\infty,\end{cases} \tag{3.8}$$

where $\widehat{A^\rho}$ is defined as in (3.5).

LEMMA 3.2. *Suppose A satisfy (1.2) and (1.3). Then $\widehat{A^\lambda}\to\widehat{A}$ as $\lambda\to\rho$. More precisely, we have*

$$|\widehat{A^\lambda}-\widehat{A}|\leq\begin{cases}C\lambda^{-2} & \text{for } 1\leq\lambda<\infty,\rho=\infty,\\ C|1-(\lambda/\rho)^2| & \text{for } 0<\rho,\lambda<\infty,\\ C'\lambda^2 & \text{for } 0<\lambda\leq 1,\rho=0, \text{if in addition } \|\nabla_y A\|_\infty<L,\end{cases}\tag{3.9}$$

where C depends only on μ,n and d, while C' depends on μ,n,d and L.

Proof. The estimate for $\rho=\infty$ in (3.9) is a direct consequence of (3.3) and Hölder's inequality. We pass to the case $0<\rho<\infty$. Note that

$$\partial_s(\chi^\rho-\chi^\lambda)-\operatorname{div}\big(A(y,s)\nabla(\chi^\rho-\chi^\lambda)\big)+\lambda^2\Delta^2(\chi^\rho-\chi^\lambda)=(\lambda^2-\rho^2)\Delta^2\chi^\rho.$$

By standard energy estimates,

$$\begin{aligned}&\|\nabla(\chi^\rho-\chi^\lambda)\|^2_{L^2(\mathbb{T}^{d+1})}+\lambda^2\|\Delta(\chi^\rho-\chi^\lambda)\|^2_{L^2(\mathbb{T}^{d+1})}\\ \leq& C|\lambda^2-\rho^2|\|\nabla^3\chi^\rho\|_{L^2(\mathbb{T}^{d+1})}\|\nabla(\chi^\rho-\chi^\lambda)\|_{L^2(\mathbb{T}^{d+1})}.\end{aligned}$$

This, combined with (3.3), gives

$$\|\nabla\chi^\rho-\nabla\chi^\lambda\|_{L^2(\mathbb{T}^{d+1})}\leq C\rho^{-2}|\lambda^2-\rho^2|,$$

which, together with the definitions of $\widehat{A^\lambda}$ and $\widehat{A}$, implies the desired estimate. Finally, the estimate in (3.9) for $\rho=0$ follows from (2.12) and the observation

$$|\widehat{A^\lambda}-\widehat{A^0}|=\Big|\fint_{\mathbb{T}^{d+1}}A(y,s)\nabla(\chi^\lambda-\chi^0)dyds\Big|\leq\|\nabla_y A\|_\infty\|\chi^\lambda-\chi^0\|_{L^2(\mathbb{T}^{d+1})}.\tag{3.10}$$

The convergence of $\widehat{A^\lambda}$ to $\widehat{A}$ (as $\lambda\to\rho$) for the case $0<\rho\leq\infty$ follows directly from the first two estimates in (3.9). By approximating A with a sequence of smooth matrices, we can obtain the convergence of $\widehat{A^\lambda}$ for $\rho=0$ from the third estimate in (3.9). Let us omit the details for concision. □

THEOREM 3.1. *Suppose A satisfies (1.2)-(1.3) and κ satisfies (1.7). Let Ω be a bounded Lipschitz domain in $\mathbb{R}^d$. For $F\in L^2(0,T;H^{-1}(\Omega)),h\in L^2(\Omega)$, let u_ε be the solution to (1.5) subjected to the boundary condition (1.6) or (1.11), and u_0 the unique solution to (1.8), with $\widehat{A}$ being defined as (3.8). Then u_ε converges to u_0 weakly in $L^2(0,T;H^1(\Omega))$ and $A\nabla u_\varepsilon$ converges to $\widehat{A}\nabla u_0$ weakly in $L^2(\Omega_T)$ as $\varepsilon\to 0$.*

Proof. By standard energy estimate,

$$\kappa\|\Delta u_\varepsilon\|_{L^2(\Omega_T)}+\|u_\varepsilon\|_{L^2(0,T;H^1(\Omega))}\leq C\big\{\|F\|_{L^2(0,T;H^{-1}(\Omega))}+\|h\|_{L^2(\Omega)}\big\}.\tag{3.11}$$

Thus $\kappa\Delta u_\varepsilon$ and $A^\varepsilon\nabla u_\varepsilon$ are uniformly bounded in $L^2(\Omega_T)$, where $A^\varepsilon(x,t)=A(x/\varepsilon,t/\varepsilon^2)$, while $u_\varepsilon,\partial_t u_\varepsilon$ are uniformly (in ε) bounded in $L^2(0,T;H^1(\Omega))$ and $L^2(0,T;H^{-2}(\Omega))$, respectively. And there exists a function u_0 such that, up to subsequences,

$$\begin{aligned}&u_\varepsilon\longrightarrow u_0 \text{ weakly in } L^2(0,T;H^1(\Omega)),\\ &\partial_t u_\varepsilon\longrightarrow\partial_t u_0 \text{ weakly in } L^2(0,T;H^{-2}(\Omega)),\\ &A^\varepsilon\nabla u_\varepsilon\longrightarrow\mathcal{M}(x,t) \text{ weakly in } L^2(\Omega_T),\\ &u_\varepsilon\longrightarrow u_0 \text{ strongly in } L^2(\Omega_T).\end{aligned}\tag{3.12}$$

Moreover, it is not difficult to see that $u_0(0)=h$, and

$$\partial_t u_0 - \operatorname{div}\mathcal{M} = F \quad \text{ in } L^2(0,T;H^{-1}(\Omega)).$$

Next, we show that $\mathcal{M}=\widehat{A}\nabla u_0$, which implies that u_0 is the weak solution to (1.8), and the whole sequence u_ε converges to u_0 weakly in $L^2(0,T;H^1(\Omega))$.

Let $\chi_j^{*\lambda,\beta}(y,s)$ be the corrector of the dual operator of $\partial_t+\mathcal{L}_\varepsilon^\lambda$ with $\lambda=\kappa/\varepsilon$. Let

$$\theta_\varepsilon = P_j^\beta + \varepsilon\chi_j^{*\lambda,\beta}(x/\varepsilon,t/\varepsilon^2), \quad 1\le j\le d, \quad 1\le\beta\le n.$$

We have for any $\phi\in C_c^\infty(\Omega_T)$,

$$-\int_0^T\langle\partial_t\theta_\varepsilon,u_\varepsilon\phi\rangle dt+\lambda^2\varepsilon^2\int_{\Omega_T}\Delta\theta_\varepsilon\Delta(u_\varepsilon\phi)dxdt+\int_{\Omega_T}(A^\varepsilon)^*\nabla\theta_\varepsilon\nabla(u_\varepsilon\phi)dxdt=0. \tag{3.13}$$

On the other hand, by (1.1),

$$\int_0^T\langle\partial_t u_\varepsilon,\phi\theta_\varepsilon\rangle dt+\lambda^2\varepsilon^2\int_{\Omega_T}\Delta u_\varepsilon\Delta(\phi\theta_\varepsilon)dxdt+\int_{\Omega_T}A^\varepsilon\nabla u_\varepsilon\nabla(\phi\theta_\varepsilon)dxdt=\int_0^T\langle F,\phi\theta_\varepsilon\rangle dt,$$

from which, we subtract (3.13) to obtain that

$$\begin{aligned}
&2\lambda^2\varepsilon^2\int_{\Omega_T}\Delta u_\varepsilon\nabla\theta_\varepsilon\nabla\phi dxdt-2\lambda^2\varepsilon^2\int_{\Omega_T}\Delta\theta_\varepsilon\nabla u_\varepsilon\nabla\phi dxdt\\
&\quad+\lambda^2\varepsilon^2\int_{\Omega_T}\Delta u_\varepsilon\theta_\varepsilon\Delta\phi dxdt-\lambda^2\varepsilon^2\int_{\Omega_T}\Delta\theta_\varepsilon u_\varepsilon\Delta\phi dxdt\\
&\quad+\int_{\Omega_T}A^\varepsilon\nabla u_\varepsilon\theta_\varepsilon\nabla\phi dxdt-\int_{\Omega_T}(A^\varepsilon)^*\nabla\theta_\varepsilon u_\varepsilon\nabla\phi dxdt\\
&=\int_0^T\langle u_\varepsilon,\theta_\varepsilon\partial_t\phi\rangle dt+\int_0^T\langle F,\phi\theta_\varepsilon\rangle dt.
\end{aligned}\tag{3.14}$$

Denote the right-hand side of (3.14) as $(3.14)_r$, and the terms in the left-hand side of (3.14) by $(3.14)_1,\ldots,(3.14)_6$ sequentially. By (3.3) and (3.11), we know that

$$(3.14)_1+\ldots+(3.14)_4\longrightarrow 0 \quad \text{ as } \varepsilon\to 0. \tag{3.15}$$

Note that by (3.3)

$$\theta_{\varepsilon,j}^\beta\longrightarrow P_j^\beta \text{ strongly in } L^2(\Omega_T) \text{ and weakly in } L^2(0,T;H^1(\Omega)),$$

which implies that

$$\begin{aligned}
&(3.14)_5\longrightarrow\int_{\Omega_T}\mathcal{M}_i^\alpha P_j^{\alpha\beta}\partial_{x_i}\phi dxdt,\\
&(3.14)_r\longrightarrow\int_0^T\langle u_0,P_j^\beta\partial_t\phi\rangle dt+\int_0^T\langle F,P_j^\beta\phi\rangle dt=\int_{\Omega_T}\mathcal{M}\nabla(P_j^\beta\phi)dxdt.
\end{aligned}\tag{3.16}$$

Furthermore, since

$$\widehat{A^\lambda}_{ij}^{\alpha\beta}=\fint_{\mathbb{T}^{d+1}}A^*\nabla(P_i^\alpha+\chi_i^{*\lambda,\alpha})\nabla P_j^\beta dyds,$$

the weak convergence result for periodic functions implies that for each fixed $\lambda>0$,

$$(A^\varepsilon)^*\nabla\theta_\varepsilon \longrightarrow \widehat{A^\lambda}^{\beta\alpha}_{ji} \text{ weakly in } L^2(\Omega_T).$$

Noticing that $u_\varepsilon \to u_0$ strongly in $L^2(\Omega_T)$ and $\widehat{A^\lambda}\to\widehat{A}$ as ε tends to zero (see Lemma 3.2), we get

$$(3.14)_6 \longrightarrow -\int_{\Omega_T} \widehat{A}^{\beta\alpha}_{ji} u_0^\alpha \partial_{x_i}\phi\, dxdt = \int_{\Omega_T} \widehat{A}^{\beta\alpha}_{ji}\partial_{x_i} u_0^\alpha \phi\, dxdt. \tag{3.17}$$

By taking (3.15)-(3.17) into (3.14), we get

$$\int_{\Omega_T} \mathcal{M}^\beta_j \phi\, dxdt = \int_{\Omega_T} \widehat{A}^{\beta\alpha}_{ji}\partial_{x_i} u_0^\alpha \phi\, dxdt$$

for any $\phi\in C_c^\infty(\Omega_T)$. It follows that $\mathcal{M}=\widehat{A}\nabla u_0$. The proof is complete. □

3.2. Flux correctors and an ε smoothing operator. Let χ^λ be the correctors given by (3.2). For $1\le\overline{\tau}\le d+1$, $1\le i\le d$, we define

$$B^{\lambda,\alpha\beta}_{\overline{\tau}j} = \begin{cases} -\lambda^2\frac{\partial}{\partial y_i}\Delta\chi^{\lambda,\alpha\beta}_j + A^{\alpha\beta}_{ij} + A^{\alpha\gamma}_{ik}\frac{\partial\chi^{\lambda,\gamma\beta}_j}{\partial y_k} - \widehat{A^\lambda}^{\alpha\beta}_{ij} & \text{if } \overline{\tau}=i,\\ -\chi^{\lambda,\alpha\beta}_j & \text{if } \overline{\tau}=d+1. \end{cases} \tag{3.18}$$

The following lemma provides the existence of flux correctors for the operator $\partial_t+\mathcal{L}^\lambda_\varepsilon$.

Lemma 3.3. *Let $1\le\alpha,\beta\le n$, $1\le i,k,j\le d$ and $1\le\overline{\varsigma},\overline{\tau}\le d+1$. There exist 1-periodic functions $\mathfrak{B}^{\lambda,\alpha\beta}_{\overline{\varsigma}\overline{\tau}j}(y,s)$ in $\mathbb{R}^{d+1}$ such that ,*

$$\begin{aligned} \mathfrak{B}^{\lambda,\alpha\beta}_{\overline{\varsigma}\overline{\tau}j} &= -\mathfrak{B}^{\lambda,\alpha\beta}_{\overline{\tau}\overline{\varsigma}j},\\ B^{\lambda,\alpha\beta}_{ij}(y,s) &= \partial_{y_k}\mathfrak{B}^{\lambda,\alpha\beta}_{kij}(y,s)+\partial_s\mathfrak{B}^{\lambda,\alpha\beta}_{(d+1)ij}(y,s). \end{aligned} \tag{3.19}$$

Furthermore, there exists a constant C, depending only on d, n,μ, such that

$$\begin{aligned} &\|\mathfrak{B}^{\lambda,\alpha\beta}_{\overline{\varsigma}\overline{\tau}j}\|_{L^2(0,1;H^1(\mathbb{T}^d))}\le C, && \text{if } 1\le\overline{\varsigma},\overline{\tau}\le d,\\ &\|\nabla\mathfrak{B}^{\lambda,\alpha\beta}_{\overline{\varsigma}\overline{\tau}j}\|_{L^2(\mathbb{T}^{d+1})}+\|\nabla^2\mathfrak{B}^{\lambda,\alpha\beta}_{\overline{\varsigma}\overline{\tau}j}\|_{L^2(\mathbb{T}^{d+1})}\le C(1+\lambda)^{-2}, && \text{if } \overline{\varsigma} \text{ or } \overline{\tau}=d+1,\\ &\|\nabla^3\mathfrak{B}^{\lambda,\alpha\beta}_{\overline{\varsigma}\overline{\tau}j}\|_{L^2(\mathbb{T}^{d+1})}\le C\lambda^{-1}(1+\lambda)^{-1}, && \text{if } \overline{\varsigma} \text{ or } \overline{\tau}=d+1. \end{aligned} \tag{3.20}$$

Proof. The construction of $\mathfrak{B}^{\lambda,\alpha\beta}_{\overline{\varsigma}\overline{\tau}j}$ is completely the same as in [8]. The estimates in (3.20) are direct consequences of the estimates for χ^λ in (3.3). Let us omit the details. □

Let $\varphi_1(s)\in C_c^\infty(-1/2,1/2), \varphi_2(y)\in C_c^\infty(B(0,\frac12))$ be fixed nonnegative functions such that

$$\int_{\mathbb{R}}\varphi_1(s)=1 \quad\text{and}\quad \int_{\mathbb{R}^d}\varphi_2(y)=1.$$

Set $\varphi_{1,\varepsilon}(s)=\frac{1}{\varepsilon^2}\varphi_1(s/\varepsilon^2), \varphi_{2,\varepsilon}(y)=\frac{1}{\varepsilon^d}\varphi_2(y/\varepsilon)$, and define

$$S_\varepsilon(f)(x,t)=\int_{\mathbb{R}^{d+1}}\varphi_{1,\varepsilon}(s)\varphi_{2,\varepsilon}(y)f(x-y,t-s)\,dyds. \tag{3.21}$$

LEMMA 3.4. *Let S_ε be defined as above. Then*

$$\|S_\varepsilon(\nabla f)-\nabla f\|_{L^2(\mathbb{R}^{d+1})}\leq C\varepsilon\big\{\|\nabla^2 f\|_{L^2(\mathbb{R}^{d+1})}+\|\partial_t f\|_{L^2(\mathbb{R}^{d+1})}\big\}, \tag{3.22}$$

where C depends only on d. Moreover, let $g(y,s)$ be 1-periodic in (y,s). Then

$$\|g(x/\varepsilon,t/\varepsilon^2)S_\varepsilon(\nabla^k f)(x,t)\|_{L^2(\mathbb{R}^{d+1})}\leq C\varepsilon^{-k}\|g\|_{L^2(\mathbb{T}^{d+1})}\|f\|_{L^2(\mathbb{R}^{d+1})},\ \text{for } k=0,1,2..., \tag{3.23}$$

where C depends only on d.

Proof. See Lemmas 3.2 and 3.3 in [8], or Lemma 3.2 in [15] for the proof. □

For $\Omega\subseteq\mathbb{R}^d$ and $0<\delta<c_0\mathrm{diam}(\Omega)$, let

$$\Omega_{T,\delta}=\big(\Omega_\delta\times(0,T)\big)\cup\big(\Omega\times(0,\delta^2)\big), \tag{3.24}$$

where $\Omega_\delta=\{x\in\Omega:\mathrm{dist}(x,\partial\Omega)<\delta\}$.

The following lemma is a direct consequence of Lemma 2.8 in [14].

LEMMA 3.5. *Let Ω be a bounded Lipschitz domain in $\mathbb{R}^d$. Then,*

$$\begin{aligned}&\|u\|_{L^2(\Omega_\delta\times(0,T))}\leq C\delta\|\nabla u\|_{L^2(\Omega_{2\delta}\times(0,T))} && \text{for } u\in L^2(0,T;H_0^1(\Omega)),\\ &\|u\|_{L^2(\Omega_\delta\times(0,T))}\leq C\delta^{1/2}\|u\|^{1/2}_{L^2(\Omega_T)}\|u\|^{1/2}_{L^2(0,T;H^1(\Omega))} && \text{for } u\in L^2(0,T;H^1(\Omega)),\end{aligned} \tag{3.25}$$

and for $u\in L^2(0,T;H^2(\Omega)\cap H_0^1(\Omega))$,

$$\|u\|_{L^2(\Omega_\delta\times(0,T))}\leq C\delta^{3/2}\|u\|^{1/2}_{L^2(0,T;H^1(\Omega))}\|u\|^{1/2}_{L^2(0,T;H^2(\Omega))}, \tag{3.26}$$

where C depends on d and Ω.

4. Convergence rate for the initial-Dirichlet problem

For $F\in L^2(\Omega_T),h\in L^2(\Omega)$, let $u_{\varepsilon,\lambda},u_{0,\lambda}$ be respectively the weak solutions to

$$\begin{aligned}&(\partial_t+\mathcal{L}_\varepsilon^\lambda)u_{\varepsilon,\lambda}=F \quad \text{in } \Omega_T,\\ &u_{\varepsilon,\lambda}=h \quad \text{on } \Omega\times\{t=0\}, \quad u_{\varepsilon,\lambda}=\frac{\partial u_{\varepsilon,\lambda}}{\partial\nu}=0 \quad \text{on } \Gamma_T,\end{aligned} \tag{4.1}$$

and

$$\begin{aligned}&(\partial_t+\mathcal{L}_0^\lambda)u_{0,\lambda}=F \quad \text{in } \Omega_T,\\ &u_{0,\lambda}=h \quad \text{on } \Omega\times\{t=0\}, \quad u_{0,\lambda}=0 \quad \text{on } \Gamma_T,\end{aligned} \tag{4.2}$$

where $\mathcal{L}_\varepsilon^\lambda,\mathcal{L}_0^\lambda(0<\lambda<\infty)$ are defined as in (3.1) and (3.4).

We now investigate the error estimate between $u_{\varepsilon,\lambda}$ and $u_{0,\lambda}$ by the classical two-scale expansion method [8, 12]. The key step is to find the proper first-order approximation of $u_{\varepsilon,\lambda}$, say $\phi_{\varepsilon,\lambda}$, and derive the $\sqrt{\varepsilon}$-order error estimate between $u_{\varepsilon,\lambda}$ and $\phi_{\varepsilon,\lambda}$ in $L^2(0,T;H^1(\Omega))$. By formal expansions, one might expect the function

$$\phi_{\varepsilon,\lambda}=u_{0,\lambda}+\varepsilon(\chi^\lambda)^\varepsilon\nabla u_{0,\lambda}-\varepsilon^2(\mathfrak{B}^\lambda_{k(d+1)j})^\varepsilon\frac{\partial^2 u_{0,\lambda}}{\partial x_j\partial x_k}$$

to be the right approximation of $u_{\varepsilon,\lambda}$ in $L^2(0,T;H^1(\Omega))$. Unfortunately, it can not be the candidate as it does not necessarily belong to $L^2(0,T;H^1(\Omega))$. Actually, since the

Equation (4.1) is fourth order, to proceed the two-scale expansion strictly and effectively, the right approximation should belong to $L^2(0,T;H^2(\Omega))$. Moreover, $\phi_{\varepsilon,\lambda}$ given above does not satisfy the Dirichlet boundary condition, and therefore cannot approximate well the function $u_{\varepsilon,\lambda}$ near the boundary. To overcome these difficulties, we need to introduce the smoothing operators and cut-off functions to make proper modifications on $\phi_{\varepsilon,\lambda}$.

For $0<\delta<1$ to be determined, define the operator $K_\varepsilon = K_{\varepsilon,\delta}$ as following

$$K_\varepsilon(f)(x,t) = S_\varepsilon(f)(x,t)\eta_\delta(x,t), \tag{4.3}$$

where S_ε is the smoothing operator defined in (3.21), and $\eta_\delta = \eta_{1,\delta}(x)\eta_{2,\delta}(t)$ is the cut-off function with $\eta_{1,\delta}\in C_c^\infty(\mathbb{R}^d), \eta_{2,\delta}\in C_c^\infty(\mathbb{R})$, and

$$\begin{aligned}
&0\le \eta_{1,\delta}\le 1, \quad \eta_{1,\delta}=0 \text{ in } \Omega_{3\delta}, \quad \eta_{1,\delta}=1 \text{ in } \Omega\setminus\Omega_{4\delta},\\
&0\le \eta_{2,\delta}\le 1, \quad \eta_{2,\delta}=0 \text{ in } (0,3\delta^2), \quad \eta_{2,\delta}=1 \text{ in } (4\delta^2,T),\\
&|\partial_t\eta_{2,\delta}|\le C\delta^{-2} \quad \text{and} \quad |\nabla^k\eta_{1,\delta}|\le C\delta^{-k}, k=1,2.
\end{aligned} \tag{4.4}$$

By (4.4), it is not difficult to find that

$$\begin{aligned}
&\eta_\delta = 1 \text{ in } \Omega_T\setminus\Omega_{T,4\delta}, \quad \eta_\delta=0 \text{ in } \Omega_{T,3\delta},\\
&|\nabla\eta_\delta|\le C\delta^{-1}, \quad |\partial_t\eta_\delta|+|\nabla^2\eta_\delta|\le C\delta^{-2},
\end{aligned} \tag{4.5}$$

where $\Omega_{T,\delta}$ is given by (3.24).

We now introduce the right approximation of $u_{\varepsilon,\lambda}$ in $L^2(0,T;H^1(\Omega))$,

$$\widetilde{\phi}_{\varepsilon,\lambda} = u_{0,\lambda}\eta_{1,\delta} + \varepsilon(\chi^\lambda)^\varepsilon K_\varepsilon(\nabla u_{0,\lambda}) - \varepsilon^2(\mathfrak{B}^\lambda_{k(d+1)j})^\varepsilon \frac{\partial}{\partial x_k} K_\varepsilon(\frac{\partial u_{0,\lambda}}{\partial x_j}).$$

Note that the smoothing operator S_ε contained in K_ε ensures that $\widetilde{\phi}_{\varepsilon,\lambda}\in L^2(0,T;H^2(\Omega))$, which allows us to perform the two-scale expansion strictly in the weak sense in Lemma 4.1. See also [26] for similar techniques involving the Steklov smoothing. On the other hand, thanks to the cut-off functions η_δ and in particular $\eta_{1,\delta}$ in (4.6), $w_{\varepsilon,\lambda} = u_{\varepsilon,\lambda} - \widetilde{\phi}_{\varepsilon,\lambda}$ belongs to the space $L^2(0,T;H_0^2(\Omega))$. We can therefore take $w_{\varepsilon,\lambda}$ as a test function in (4.11) to derive the desirable error estimate (4.13).

LEMMA 4.1. *Let Ω be a bounded Lipschitz domain in $\mathbb{R}^d$ and $0<T<\infty$. Suppose A satisfies conditions (1.2) and (1.3). Define*

$$w_{\varepsilon,\lambda}(x,t) = u_{\varepsilon,\lambda} - u_{0,\lambda}\eta_{1,\delta} - \varepsilon(\chi^\lambda)^\varepsilon K_\varepsilon(\nabla u_{0,\lambda}) + \varepsilon^2(\mathfrak{B}^\lambda_{k(d+1)j})^\varepsilon\frac{\partial}{\partial x_k}K_\varepsilon(\frac{\partial u_{0,\lambda}}{\partial x_j}). \tag{4.6}$$

Then we have

$$\begin{aligned}
&(\partial_t+\mathcal{L}_\varepsilon^\lambda)w_{\varepsilon,\lambda}\\
=&(\partial_t+\mathcal{L}_\varepsilon^\lambda)\{u_{0,\lambda}(1-\eta_{1,\delta})\} - \lambda^2\varepsilon^2\Delta^2 u_{0,\lambda}\\
&-div\{(\widehat{A^\lambda}-A^\varepsilon)(\nabla u_{0,\lambda}-K_\varepsilon(\nabla u_{0,\lambda}))\} - \varepsilon\, div\{(\mathfrak{B}^\lambda)^\varepsilon\nabla K_\varepsilon(\nabla u_{0,\lambda})\}\\
&-\varepsilon^2\frac{\partial}{\partial x_i}\Big\{(\mathfrak{B}^\lambda_{(d+1)ij})^\varepsilon\partial_t K_\varepsilon\Big(\frac{\partial u_{0,\lambda}}{\partial x_j}\Big)\Big\} + \varepsilon\, div\{A^\varepsilon(\chi^\lambda)^\varepsilon\nabla K_\varepsilon(\nabla u_{0,\lambda})\}\\
&-\lambda^2\varepsilon\, div\{(\Delta\chi^\lambda)^\varepsilon\nabla K_\varepsilon(\nabla u_{0,\lambda})\} - 2\lambda^2\varepsilon^2\Delta\{(\nabla\chi^\lambda)^\varepsilon\nabla K_\varepsilon(\nabla u_{0,\lambda})\}
\end{aligned}$$

$$-\lambda^2\varepsilon^3\Delta\{(\chi^\lambda)^\varepsilon\Delta K_\varepsilon(\nabla u_{0,\lambda})\}-\varepsilon\, div\Big\{A^\varepsilon(\nabla\mathfrak{B}^\lambda_{k(d+1)j})^\varepsilon\frac{\partial}{\partial x_k}K_\varepsilon\Big(\frac{\partial u_{0,\lambda}}{\partial x_j}\Big)\Big\}$$
$$-\varepsilon^2\, div\Big\{A^\varepsilon(\mathfrak{B}^\lambda_{k(d+1)j})^\varepsilon\nabla\frac{\partial}{\partial x_k}K_\varepsilon\Big(\frac{\partial u_{0,\lambda}}{\partial x_j}\Big)\Big\}$$
$$+\lambda^2\varepsilon^2\Delta\Big\{(\Delta\mathfrak{B}^\lambda_{k(d+1)j})^\varepsilon\frac{\partial}{\partial x_k}K_\varepsilon\Big(\frac{\partial u_{0,\lambda}}{\partial x_j}\Big)+2\varepsilon(\nabla\mathfrak{B}^\lambda_{k(d+1)j})^\varepsilon\nabla\frac{\partial}{\partial x_k}K_\varepsilon\Big(\frac{\partial u_{0,\lambda}}{\partial x_j}\Big)$$
$$+\varepsilon^2(\mathfrak{B}^\lambda_{k(d+1)j})^\varepsilon\Delta\frac{\partial}{\partial x_k}K_\varepsilon\Big(\frac{\partial u_{0,\lambda}}{\partial x_j}\Big)\Big\} \tag{4.7}$$

in the weak sense.

Proof. Since $w_{\varepsilon,\lambda}(x,t)\in L^2(0,T;H^2_0(\Omega))$, by direct calculations we have

$$\begin{aligned}(\partial_t+\mathcal{L}^\lambda_\varepsilon)w_{\varepsilon,\lambda}=&(\partial_t+\mathcal{L}^\lambda_\varepsilon)[u_{0,\lambda}(1-\eta_{1,\delta})]+(\mathcal{L}^\lambda_0-\mathcal{L}^\lambda_\varepsilon)u_{0,\lambda}\\&-(\partial_t+\mathcal{L}^\lambda_\varepsilon)\{\varepsilon(\chi^\lambda)^\varepsilon K_\varepsilon(\nabla u_{0,\lambda})\}\\&+(\partial_t+\mathcal{L}^\lambda_\varepsilon)\Big\{\varepsilon^2(\mathfrak{B}^\lambda_{k(d+1)j})^\varepsilon\frac{\partial}{\partial x_k}K_\varepsilon\Big(\frac{\partial u_{0,\lambda}}{\partial x_j}\Big)\Big\}\\=&(\partial_t+\mathcal{L}^\lambda_\varepsilon)\{u_{0,\lambda}(1-\eta_{1,\delta})\}-\lambda^2\varepsilon^2\Delta^2u_{0,\lambda}\\&-\mathrm{div}\{(\widehat{A^\lambda}-A^\varepsilon)(\nabla u_{0,\lambda}-K_\varepsilon(\nabla u_{0,\lambda}))\}+\mathrm{div}\{(B^\lambda)^\varepsilon K_\varepsilon(\nabla u_{0,\lambda})\}\\&+\mathrm{div}\{\big(\lambda^2(\nabla\Delta\chi^\lambda)^\varepsilon-A^\varepsilon(\nabla\chi^\lambda)^\varepsilon\big)K_\varepsilon(\nabla u_{0,\lambda})\}\\&-(\partial_t+\mathcal{L}^\lambda_\varepsilon)\{\varepsilon(\chi^\lambda)^\varepsilon K_\varepsilon(\nabla u_{0,\lambda})\}\\&+(\partial_t+\mathcal{L}^\lambda_\varepsilon)\Big\{\varepsilon^2(\mathfrak{B}^\lambda_{k(d+1)j})^\varepsilon\frac{\partial}{\partial x_k}K_\varepsilon\Big(\frac{\partial u_{0,\lambda}}{\partial x_j}\Big)\Big\}\end{aligned}$$

in the weak sense, where $B^\lambda=B^\lambda_{ij},1\le i,j\le d$, is defined as (3.18). Since

$$\begin{aligned}&(\partial_t+\mathcal{L}^\lambda_\varepsilon)\{\varepsilon(\chi^\lambda)^\varepsilon K_\varepsilon(\nabla u_{0,\lambda})\}\\=&\varepsilon\partial_t\{(\chi^\lambda)^\varepsilon K_\varepsilon(\nabla u_{0,\lambda})\}-\mathrm{div}\{A^\varepsilon(\nabla\chi^\lambda)^\varepsilon K_\varepsilon(\nabla u_{0,\lambda})\}\\&-\varepsilon\mathrm{div}\{A^\varepsilon(\chi^\lambda)^\varepsilon\nabla K_\varepsilon(\nabla u_{0,\lambda})\}+\lambda^2\mathrm{div}\{(\nabla\Delta\chi^\lambda)^\varepsilon K_\varepsilon(\nabla u_{0,\lambda})\}\\&+\lambda^2\varepsilon\mathrm{div}\{(\Delta\chi^\lambda)^\varepsilon\nabla K_\varepsilon(\nabla u_{0,\lambda})\}+2\lambda^2\varepsilon^2\Delta\{(\nabla\chi^\lambda)^\varepsilon\nabla K_\varepsilon(\nabla u_{0,\lambda})\}\\&+\lambda^2\varepsilon^3\Delta\{(\chi^\lambda)^\varepsilon\Delta K_\varepsilon(\nabla u_{0,\lambda})\}.\end{aligned}$$

We obtain that

$$\begin{aligned}(\partial_t+\mathcal{L}^\lambda_\varepsilon)w_{\varepsilon,\lambda}=&(\partial_t+\mathcal{L}^\lambda_\varepsilon)\{u_{0,\lambda}(1-\eta_{1,\delta})\}-\lambda^2\varepsilon^2\Delta^2u_{0,\lambda}\\&-\mathrm{div}\{(\widehat{A^\lambda}-A^\varepsilon)(\nabla u_{0,\lambda}-K_\varepsilon(\nabla u_{0,\lambda}))\}\\&+\mathrm{div}\{(B^\lambda)^\varepsilon K_\varepsilon(\nabla u_{0,\lambda})\}-\varepsilon\partial_t\{(\chi^\lambda)^\varepsilon K_\varepsilon(\nabla u_{0,\lambda})\}\\&+\varepsilon\mathrm{div}\{A^\varepsilon(\chi^\lambda)^\varepsilon\nabla K_\varepsilon(\nabla u_{0,\lambda})\}-\lambda^2\varepsilon\mathrm{div}\{(\Delta\chi^\lambda)^\varepsilon\nabla K_\varepsilon(\nabla u_{0,\lambda})\}\\&-2\lambda^2\varepsilon^2\Delta\{(\nabla\chi^\lambda)^\varepsilon\nabla K_\varepsilon(\nabla u_{0,\lambda})-\lambda^2\varepsilon^3\Delta\{(\chi^\lambda)^\varepsilon\Delta K_\varepsilon(\nabla u_{0,\lambda})\}\\&+(\partial_t+\mathcal{L}^\lambda_\varepsilon)\Big\{\varepsilon^2(\mathfrak{B}^\lambda_{k(d+1)j})^\varepsilon\frac{\partial}{\partial x_k}K_\varepsilon\Big(\frac{\partial u_{0,\lambda}}{\partial x_j}\Big)\Big\}.\end{aligned}\tag{4.8}$$

In view of Lemma 3.3,

$$\begin{aligned}
&\frac{\partial}{\partial x_i}\Big\{(B_{ij}^\lambda)^\varepsilon K_\varepsilon\Big(\frac{\partial u_{0,\lambda}}{\partial x_j}\Big)\Big\}-\varepsilon\partial_t\Big\{(\chi_j^\lambda)^\varepsilon K_\varepsilon\Big(\frac{\partial u_{0,\lambda}}{\partial x_j}\Big)\Big\}\\
=&\frac{\partial}{\partial x_i}\Big\{\Big(\varepsilon\frac{\partial}{\partial x_k}(\mathfrak{B}_{kij}^\lambda)^\varepsilon+\varepsilon^2\partial_t(\mathfrak{B}_{(d+1)ij}^\lambda)^\varepsilon\Big)K_\varepsilon\Big(\frac{\partial u_{0,\lambda}}{\partial x_j}\Big)\Big\}\\
&+\varepsilon^2\partial_t\Big\{\frac{\partial}{\partial x_k}(\mathfrak{B}_{k(d+1)j}^\lambda)^\varepsilon K_\varepsilon\Big(\frac{\partial u_{0,\lambda}}{\partial x_j}\Big)\Big\}\\
=&\varepsilon\frac{\partial}{\partial x_i}\frac{\partial}{\partial x_k}\Big\{(\mathfrak{B}_{kij}^\lambda)^\varepsilon K_\varepsilon\Big(\frac{\partial u_{0,\lambda}}{\partial x_j}\Big)\Big\}-\varepsilon\frac{\partial}{\partial x_i}\Big\{(\mathfrak{B}_{kij}^\lambda)^\varepsilon\frac{\partial}{\partial x_k}K_\varepsilon\Big(\frac{\partial u_{0,\lambda}}{\partial x_j}\Big)\Big\}\\
&+\varepsilon^2\frac{\partial}{\partial x_i}\partial_t\Big\{(\mathfrak{B}_{(d+1)ij}^\lambda)^\varepsilon K_\varepsilon\Big(\frac{\partial u_{0,\lambda}}{\partial x_j}\Big)\Big\}-\varepsilon^2\frac{\partial}{\partial x_i}\Big\{(\mathfrak{B}_{(d+1)ij}^\lambda)^\varepsilon\partial_t K_\varepsilon\Big(\frac{\partial u_{0,\lambda}}{\partial x_j}\Big)\Big\}\\
&+\varepsilon^2\frac{\partial}{\partial x_k}\partial_t\Big\{(\mathfrak{B}_{k(d+1)j}^\lambda)^\varepsilon K_\varepsilon\Big(\frac{\partial u_{0,\lambda}}{\partial x_j}\Big)\Big\}-\varepsilon^2\partial_t\Big\{(\mathfrak{B}_{k(d+1)j}^\lambda)^\varepsilon\frac{\partial}{\partial x_k}K_\varepsilon\Big(\frac{\partial u_{0,\lambda}}{\partial x_j}\Big)\Big\}.
\end{aligned}$$

By the skew-symmetry of $\mathfrak{B}^\lambda$, we derive that

$$\begin{aligned}
&\operatorname{div}\big\{(B^\lambda)^\varepsilon K_\varepsilon(\nabla u_{0,\lambda})\big\}-\varepsilon\partial_t\big\{(\chi^\lambda)^\varepsilon K_\varepsilon(\nabla u_{0,\lambda})\big\}\\
=&-\varepsilon\operatorname{div}\big\{(\mathfrak{B}^\lambda)^\varepsilon\nabla K_\varepsilon(\nabla u_{0,\lambda})\big\}-\varepsilon^2\frac{\partial}{\partial x_i}\Big\{(\mathfrak{B}_{(d+1)ij}^\lambda)^\varepsilon\partial_t K_\varepsilon\Big(\frac{\partial u_{0,\lambda}}{\partial x_j}\Big)\Big\}\\
&-\varepsilon^2\partial_t\Big\{(\mathfrak{B}_{k(d+1)j}^\lambda)^\varepsilon\frac{\partial}{\partial x_k}K_\varepsilon\Big(\frac{\partial u_{0,\lambda}}{\partial x_j}\Big)\Big\}.
\end{aligned}\tag{4.9}$$

Finally, note that

$$\begin{aligned}
&(\partial_t+\mathcal{L}_\varepsilon^\lambda)\Big\{\varepsilon^2(\mathfrak{B}_{k(d+1)j}^\lambda)^\varepsilon\frac{\partial}{\partial x_k}K_\varepsilon\Big(\frac{\partial u_{0,\lambda}}{\partial x_j}\Big)\Big\}\\
=&\varepsilon^2\partial_t\Big\{(\mathfrak{B}_{k(d+1)j}^\lambda)^\varepsilon\frac{\partial}{\partial x_k}K_\varepsilon\Big(\frac{\partial u_{0,\lambda}}{\partial x_j}\Big)\Big\}-\varepsilon\operatorname{div}\Big\{A^\varepsilon(\nabla\mathfrak{B}_{k(d+1)j}^\lambda)^\varepsilon\frac{\partial}{\partial x_k}K_\varepsilon\Big(\frac{\partial u_{0,\lambda}}{\partial x_j}\Big)\Big\}\\
&-\varepsilon^2\operatorname{div}\Big\{A^\varepsilon(\mathfrak{B}_{k(d+1)j}^\lambda)^\varepsilon\nabla\frac{\partial}{\partial x_k}K_\varepsilon\Big(\frac{\partial u_{0,\lambda}}{\partial x_j}\Big)\Big\}+\lambda^2\varepsilon^2\Delta\Big\{(\Delta\mathfrak{B}_{k(d+1)j}^\lambda)^\varepsilon\frac{\partial}{\partial x_k}K_\varepsilon\Big(\frac{\partial u_{0,\lambda}}{\partial x_j}\Big)\\
&+2\varepsilon(\nabla\mathfrak{B}_{k(d+1)j}^\lambda)^\varepsilon\nabla\frac{\partial}{\partial x_k}K_\varepsilon\Big(\frac{\partial u_{0,\lambda}}{\partial x_j}\Big)+\varepsilon^2(\mathfrak{B}_{k(d+1)j}^\lambda)^\varepsilon\Delta\frac{\partial}{\partial x_k}K_\varepsilon\Big(\frac{\partial u_{0,\lambda}}{\partial x_j}\Big)\Big\}.
\end{aligned}\tag{4.10}$$

By taking (4.9) and (4.10) into (4.8), we get (4.7) immediately. ∎

LEMMA 4.2. *Let Ω be a bounded Lipschitz domain in $\mathbb{R}^d$ and $0<T<\infty$. Suppose that A satisfies conditions (1.2)-(1.3). Let $u_{\varepsilon,\lambda},u_{0,\lambda}$ be, respectively, the solutions to (4.1) and (4.2). Let $w_{\varepsilon,\lambda}$ be defined as (4.6) with K_ε given by (4.3) and $\delta=(1+\lambda)\varepsilon<1$. Then for any $\psi\in L^2(0,T;H_0^2(\Omega))$,*

$$\begin{aligned}
&\Big|\int_0^T\langle(\partial_t+\mathcal{L}_\varepsilon^\lambda)w_{\varepsilon,\lambda},\psi\rangle_{H^{-2}(\Omega)\times H_0^2(\Omega)}dt\Big|\\
\le& C\varepsilon\big\{\|\nabla^2u_{0,\lambda}\|_{L^2(\Omega_T\setminus\Omega_{T,2\delta})}+\|\partial_tu_{0,\lambda}\|_{L^2(\Omega_T\setminus\Omega_{T,2\delta})}\big\}\|\nabla\psi\|_{L^2(\Omega_T)}\\
&+C\|\partial_tu_{0,\lambda}\|_{L^2(\Omega_T)}\|\psi\|_{L^2(\Omega_{4\delta}\times(0,T))}\\
&+C\lambda^2\varepsilon^2\delta^{-1/2}\|\Delta u_{0,\lambda}\|_{L^2(\Omega_T)}\|\Delta\psi\|_{L^2(\Omega_{5\delta}\times(0,T))}\\
&+C\lambda^2\varepsilon^2\delta^{-1}\|\nabla u_{0,\lambda}\|_{L^2(\Omega_{5\delta}\times(0,T))}\|\Delta\psi\|_{L^2(\Omega_{5\delta}\times(0,T))}
\end{aligned}$$

$$+C\lambda^2\varepsilon^2\|\nabla^2 u_{0,\lambda}\|_{L^2(\Omega_T\setminus\Omega_{T,2\delta})}\|\Delta\psi\|_{L^2(\Omega_T)}$$
$$+C\|\nabla u_{0,\lambda}\|_{L^2(\Omega_{T,5\delta})}\|\nabla\psi\|_{L^2(\Omega_{T,5\delta})}, \tag{4.11}$$

where $C>0$ depends only on d,n,μ,T and Ω.

Proof. By (4.7), we can take $\psi\in L^2(0,T;H_0^2(\Omega))$ as a test function to deduce that

$$\begin{aligned}
&\Big|\int_0^T\langle(\partial_t+\mathcal{L}_\varepsilon^\lambda)w_{\varepsilon,\lambda},\psi\rangle_{H^{-2}(\Omega)\times H_0^2(\Omega)}dt\Big|\\
&\le C\int_{\Omega_T}|\partial_t u_{0,\lambda}(1-\eta_{1,\delta})\psi|+C\int_{\Omega_T}|\nabla\big(u_{0,\lambda}(1-\eta_{1,\delta})\big)||\nabla\psi|\\
&\quad+C\lambda^2\varepsilon^2\int_{\Omega_T}|\Delta\big(u_{0,\lambda}(1-\eta_{1,\delta})\big)||\Delta\psi|+C\lambda^2\varepsilon^2\int_{\Omega_T}|\Delta u_{0,\lambda}||\Delta\psi|\\
&\quad+C\int_{\Omega_T}|\nabla u_{0,\lambda}-K_\varepsilon(\nabla u_{0,\lambda})||\nabla\psi|+C\varepsilon\int_{\Omega_T}|(\chi^\lambda)^\varepsilon\nabla K_\varepsilon(\nabla u_{0,\lambda})||\nabla\psi|\\
&\quad+C\varepsilon\int_{\Omega_T}|(\mathfrak{B}^\lambda)^\varepsilon\nabla K_\varepsilon(\nabla u_{0,\lambda})||\nabla\psi|+C\varepsilon^2\int_{\Omega_T}\Big|(\mathfrak{B}^\lambda_{(d+1)ij})^\varepsilon\partial_t K_\varepsilon\Big(\frac{\partial u_{0,\lambda}}{\partial x_j}\Big)\Big|\Big|\frac{\partial\psi}{\partial x_i}\Big|\\
&\quad+C\lambda^2\varepsilon\int_{\Omega_T}|(\Delta\chi^\lambda)^\varepsilon\nabla K_\varepsilon(\nabla u_{0,\lambda})||\nabla\psi|+C\lambda^2\varepsilon^2\int_{\Omega_T}|(\nabla\chi^\lambda)^\varepsilon\nabla K_\varepsilon(\nabla u_{0,\lambda})||\Delta\psi|\\
&\quad+C\varepsilon\int_{\Omega_T}\Big|(\nabla\mathfrak{B}^\lambda_{k(d+1)j})^\varepsilon\frac{\partial}{\partial x_k}K_\varepsilon\Big(\frac{\partial u_{0,\lambda}}{\partial x_j}\Big)\Big||\nabla\psi|+C\lambda^2\varepsilon^3\int_{\Omega_T}|(\chi^\lambda)^\varepsilon\Delta K_\varepsilon(\nabla u_{0,\lambda})||\Delta\psi|\\
&\quad+C\varepsilon^2\int_{\Omega_T}\Big|(\mathfrak{B}^\lambda_{k(d+1)j})^\varepsilon\frac{\partial}{\partial x_k}\nabla K_\varepsilon\Big(\frac{\partial u_{0,\lambda}}{\partial x_j}\Big)\Big||\nabla\psi|\\
&\quad+C\lambda^2\varepsilon^2\int_{\Omega_T}\Big|(\Delta\mathfrak{B}^\lambda_{k(d+1)j})^\varepsilon\frac{\partial}{\partial x_k}K_\varepsilon\Big(\frac{\partial u_{0,\lambda}}{\partial x_j}\Big)\Big||\Delta\psi|\\
&\quad+C\lambda^2\varepsilon^3\int_{\Omega_T}\Big|(\nabla\mathfrak{B}^\lambda_{k(d+1)j})^\varepsilon\nabla\frac{\partial}{\partial x_k}K_\varepsilon\Big(\frac{\partial u_{0,\lambda}}{\partial x_j}\Big)\Big||\Delta\psi|\\
&\quad+C\lambda^2\varepsilon^4\int_{\Omega_T}\Big|(\mathfrak{B}^\lambda_{k(d+1)j})^\varepsilon\Delta\frac{\partial}{\partial x_k}K_\varepsilon\Big(\frac{\partial u_{0,\lambda}}{\partial x_j}\Big)\Big||\Delta\psi|\\
&=I_1+\ldots+I_{16},
\end{aligned}\tag{4.12}$$

where C depends only on d,n and μ.

It is easy to see that

$$I_1\le C\|\partial_t u_{0,\lambda}\|_{L^2(\Omega_T)}\|\psi\|_{L^2(\Omega_{4\delta}\times(0,T))},$$
$$I_4\le C\lambda^2\varepsilon^2\|\nabla^2 u_{0,\lambda}\|_{L^2(\Omega_T)}\|\Delta\psi\|_{L^2(\Omega_T)}.$$

By Lemma 3.5, we deduce that

$$\begin{aligned}
I_2&\le C\big\{\|\nabla u_{0,\lambda}\|_{L^2(\Omega_{4\delta}\times(0,T))}+\delta^{-1}\|u_{0,\lambda}\|_{L^2(\Omega_{4\delta}\times(0,T))}\big\}\|\nabla\psi\|_{L^2(\Omega_{4\delta}\times(0,T))}\\
&\le C\|\nabla u_{0,\lambda}\|_{L^2(\Omega_{5\delta}\times(0,T))}\|\nabla\psi\|_{L^2(\Omega_{4\delta}\times(0,T))}.
\end{aligned}$$

Likewise

$$\begin{aligned}
I_3\le{}& C\lambda^2\varepsilon^2\|\Delta u_{0,\lambda}\|_{L^2(\Omega_T)}\|\Delta\psi\|_{L^2(\Omega_{4\delta}\times(0,T))}\\
&+C\lambda^2\varepsilon^2\delta^{-1/2}\|\Delta u_{0,\lambda}\|_{L^2(\Omega_T)}\|\Delta\psi\|_{L^2(\Omega_{4\delta}\times(0,T))}.
\end{aligned}$$

Note that

$$\nabla u_{0,\lambda}-K_\varepsilon(\nabla u_{0,\lambda})=\nabla u_{0,\lambda}(1-\eta_\delta)+\eta_\delta\big(\nabla u_{0,\lambda}-S_\varepsilon(\nabla u_{0,\lambda})\big).$$

By Lemma 3.4, we can bound I_5 as following

$$\begin{aligned}I_5\le C\varepsilon\big\{&\|\nabla^2 u_{0,\lambda}\|_{L^2(\Omega_T\setminus\Omega_{T,2\delta})}+\|\partial_t u_{0,\lambda}\|_{L^2(\Omega_T\setminus\Omega_{T,2\delta})}\big\}\|\nabla\psi\|_{L^2(\Omega_T)}\\&+C\|\nabla u_{0,\lambda}\|_{L^2(\Omega_{T,5\delta})}\|\nabla\psi\|_{L^2(\Omega_{T,5\delta})}.\end{aligned}$$

Observe that

$$\begin{aligned}\nabla K_\varepsilon(\nabla u_{0,\lambda})&=S_\varepsilon(\nabla u_{0,\lambda})\nabla\eta_\delta+S_\varepsilon(\nabla^2 u_{0,\lambda})\eta_\delta,\\ \partial_t K_\varepsilon(\nabla u_{0,\lambda})&=S_\varepsilon(\nabla u_{0,\lambda})\partial_t\eta_\delta+S_\varepsilon(\nabla\partial_t u_{0,\lambda})\eta_\delta.\end{aligned}$$

We can use Lemma 3.4 and the estimates on χ^λ and $\mathfrak{B}^\lambda$ to deduce that

$$\begin{aligned}I_6+\ldots+I_9+I_{11}\le{}& C\varepsilon\big\{\|\nabla^2 u_{0,\lambda}\|_{L^2(\Omega_T\setminus\Omega_{T,2\delta})}+\|\partial_t u_{0,\lambda}\|_{L^2(\Omega_T\setminus\Omega_{T,2\delta})}\big\}\|\nabla\psi\|_{L^2(\Omega_T)}\\&+C\|\nabla u_{0,\lambda}\|_{L^2(\Omega_{T,5\delta})}\|\nabla\psi\|_{L^2(\Omega_{T,5\delta})},\\ I_{10}+I_{14}\le{}& C\lambda^2\varepsilon^2\|\nabla^2 u_{0,\lambda}\|_{L^2(\Omega_T\setminus\Omega_{T,2\delta})}\|\Delta\psi\|_{L^2(\Omega_T)}\\&+C\lambda^2\varepsilon^2\delta^{-1}\|\nabla u_{0,\lambda}\|_{L^2(\Omega_{5\delta}\times(0,T))}\|\Delta\psi\|_{L^2(\Omega_{5\delta}\times(0,T))}.\end{aligned}$$

Also since

$$\nabla^2 K_\varepsilon(\nabla u_{0,\lambda})=S_\varepsilon(\nabla^3 u_{0,\lambda})\eta_\delta+2S_\varepsilon(\nabla^2 u_{0,\lambda})\nabla\eta_\delta+S_\varepsilon(\nabla u_{0,\lambda})\nabla^2\eta_\delta$$

and $\varepsilon\delta^{-1}<1$, we can perform similar analysis to derive that

$$\begin{aligned}I_{12}+I_{15}+I_{16}\le{}& C\lambda^2\varepsilon^2\|\nabla^2 u_{0,\lambda}\|_{L^2(\Omega_T\setminus\Omega_{T,2\delta})}\|\Delta\psi\|_{L^2(\Omega_T)}\\&+C\lambda^2\varepsilon^2\delta^{-1}\|\nabla u_{0,\lambda}\|_{L^2(\Omega_{5\delta}\times(0,T))}\|\Delta\psi\|_{L^2(\Omega_{5\delta}\times(0,T))},\end{aligned}$$

and

$$\begin{aligned}I_{13}\le{}& C\varepsilon\|\nabla^2 u_{0,\lambda}\|_{L^2(\Omega_T\setminus\Omega_{T,2\delta})}\|\nabla\psi\|_{L^2(\Omega_T)}\\&+C\|\nabla u_{0,\lambda}\|_{L^2(\Omega_{5\delta}\times(0,T))}\|\nabla\psi\|_{L^2(\Omega_{5\delta}\times(0,T))}.\end{aligned}$$

The desired estimate (4.11) follows directly from (4.12) and the estimates on I_1–I_{16}. ☐

LEMMA 4.3. *In addition to the assumptions of Lemma 4.2, we assume that* $h\in H^1(\Omega)$. *Then*

$$\lambda\varepsilon\|\Delta w_{\varepsilon,\lambda}\|_{L^2(\Omega_T)}+\|\nabla w_{\varepsilon,\lambda}\|_{L^2(\Omega_T)}\le C(1+\lambda)^{1/2}\varepsilon^{1/2}\big\{\|F\|_{L^2(\Omega_T)}+\|h\|_{H^1(\Omega)}\big\},\quad(4.13)$$

where C *depends only on* d,n,μ,T *and* Ω.

Proof. Note that $\delta=(\lambda+1)\varepsilon$ and $w_{\varepsilon,\lambda}\in L^2(0,T;H^2_0(\Omega))$ with $w_{\varepsilon,\lambda}(x,0)=h(1-\eta_{1,\delta})$. By taking $\psi=w_{\varepsilon,\lambda}$ in (4.11) and using Cauchy's inequality as well as Lemma 3.5, we obtain that

$$\begin{aligned}&\lambda^2\varepsilon^2\|\Delta w_{\varepsilon,\lambda}\|^2_{L^2(\Omega_T)}+\|\nabla w_{\varepsilon\lambda}\|^2_{L^2(\Omega_T)}\\&\le C(\lambda+1)\varepsilon\big\{\|\nabla^2 u_{0,\lambda}\|^2_{L^2(\Omega_T)}+\|\partial_t u_{0,\lambda}\|^2_{L^2(\Omega_T)}\big\}\\&\quad+C\|\nabla u_{0,\lambda}\|^2_{L^2(\Omega_{T,5\delta})}+C\|h(1-\eta_{1,\delta})\|^2_{L^2(\Omega)},\end{aligned}\quad(4.14)$$

Note that

$$\|\nabla u_{0,\lambda}\|^2_{L^2(\Omega_{T,5\delta})} \le \|\nabla u_{0,\lambda}\|^2_{L^2(\Omega_{5\delta}\times(0,T))} + \int_0^{25\delta^2} \|\nabla u_{0,\lambda}(s)\|^2_{L^2(\Omega)} ds. \tag{4.15}$$

By Lemma 3.5,

$$\|h(1-\eta_{1,\delta})\|_{L^2(\Omega)} \le C\delta^{1/2}\|h\|_{H^1(\Omega)}, \tag{4.16}$$

$$\|\nabla u_{0,\lambda}\|_{L^2(\Omega_{5\delta}\times(0,T))} \le C\delta^{1/2}\|\nabla^2 u_{0,\lambda}\|^{1/2}_{L^2(\Omega_T)}\|\nabla u_{0,\lambda}\|^{1/2}_{L^2(\Omega_T)}. \tag{4.17}$$

On the other hand, for $F\in L^2(\Omega_T)$ and $h\in L^2(\Omega)$, standard energy estimates imply that

$$\begin{aligned}
&\sup_{25\delta^2\le t\le T}\int_{t-25\delta^2}^{t} \|\nabla u_{0,\lambda}\|^2_{L^2(\Omega)} ds\\
&\le C\delta\big\{\|\partial_t u_{0,\lambda}\|_{L^2(\Omega_T)} + \|F\|_{L^2(\Omega_T)}\big\}\sup_{0<t<T}\|u_{0,\lambda}(\cdot,t)\|_{L^2(\Omega)}\\
&\le C\delta\big\{\|\partial_t u_{0,\lambda}\|^2_{L^2(\Omega_T)} + \|F\|^2_{L^2(\Omega_T)} + \|h\|^2_{L^2(\Omega)}\big\}\\
&\le C\delta\big\{\|\nabla^2 u_{0,\lambda}\|^2_{L^2(\Omega_T)} + \|F\|^2_{L^2(\Omega_T)} + \|h\|^2_{L^2(\Omega)}\big\},
\end{aligned} \tag{4.18}$$

where for the last step we have used the estimate

$$\|\partial_t u_{0,\lambda}\|^2_{L^2(\Omega_T)} \le C\big\{\|\nabla^2 u_{0,\lambda}\|^2_{L^2(\Omega_T)} + \|F\|^2_{L^2(\Omega_T)}\big\}. \tag{4.19}$$

Taking (4.17) and (4.18) into (4.15), we derive that

$$\|\nabla u_{0,\lambda}\|^2_{L^2(\Omega_{T,5\delta})} \le C\delta\big\{\|u_{0,\lambda}\|^2_{L^2(0,T;H^2(\Omega))} + \|F\|^2_{L^2(\Omega_T)} + \|h\|^2_{L^2(\Omega)}\big\}, \tag{4.20}$$

which, together with (4.14), (4.16) and the H^2 estimate for $u_{0,\lambda}$

$$\|u_{0,\lambda}\|^2_{L^2(0,T;H^2(\Omega))} \le C\big\{\|F\|^2_{L^2(\Omega_T)} + \|h\|^2_{H^1(\Omega)}\big\}, \tag{4.21}$$

gives (4.13). ∎

To prove Theorem 1.1, we shall use the following sharp convergence rate for the operator $\partial_t + \mathcal{L}^\lambda_\varepsilon$ with fixed λ.

THEOREM 4.1. *Suppose Ω is a bounded $C^{1,1}$ domain in $\mathbb{R}^d$ and $0<T<\infty$. Assume A satisfies conditions (1.2)-(1.3). Let $u_{\varepsilon,\lambda}, u_{0,\lambda}$ be, respectively, the solutions to (4.1) and (4.2) with $F\in L^2(\Omega_T)$ and $h\in H^1_0(\Omega)$. Then for any fixed $0<\lambda<\infty$,*

$$\|u_{\varepsilon,\lambda} - u_{0,\lambda}\|_{L^2(\Omega_T)} \le C(1+\lambda)\varepsilon\big\{\|F\|_{L^2(\Omega_T)} + \|h\|_{H^1_0(\Omega)}\big\}, \tag{4.22}$$

where C depends only on d, n, μ, Ω and T.

We now give the proof of Theorem 1.1 by using Theorem 4.1, the proof of which is left as the end of this section.

Proof. **(Proof of Theorem 1.1.)** Let u_ε be the solution of (1.5) and (1.6), and u_0 the solution of (1.8). Note that u_ε is the solution to (4.1) with $\lambda=\kappa/\varepsilon$. Therefore, by (4.22),

$$\begin{aligned}
\|u_\varepsilon - u_0\|_{L^2(\Omega_T)} &\le \|u_{\varepsilon,\lambda} - u_{0,\lambda}\|_{L^2(\Omega_T)} + \|u_{0,\lambda} - u_0\|_{L^2(\Omega_T)}\\
&\le C(1+\lambda)\varepsilon\big\{\|F\|_{L^2(\Omega_T)} + \|h\|_{H^1_0(\Omega)}\big\} + \|u_{0,\lambda} - u_0\|_{L^2(\Omega_T)}.
\end{aligned} \tag{4.23}$$

Since $u_{0,\lambda}-u_0=0$ on $\partial\Omega_T$ and

$$\partial_t(u_{0,\lambda}-u_0)-\operatorname{div}\big(\widehat{A^\lambda}\nabla(u_{0,\lambda}-u_0)\big)=\operatorname{div}\big((\widehat{A^\lambda}-\widehat{A})\nabla u_0\big)\quad \text{in } \Omega_T.$$

We have

$$\begin{aligned}\|u_{0,\lambda}-u_0\|_{L^2(0,T;H_0^1(\Omega))}&\le C|\widehat{A^\lambda}-\widehat{A}|\|u_0\|_{L^2(0,T;H^1(\Omega))}\\&\le C|\widehat{A^\lambda}-\widehat{A}|\big\{\|F\|_{L^2(\Omega_T)}+\|h\|_{L^2(\Omega)}\big\}.\end{aligned}\tag{4.24}$$

By taking (4.24) into (4.23), and using (3.9), we derive (1.10) and complete the proof. ∎

Finally, let us prove Theorem 4.1 following the idea of [8, 22].

Proof. **(Proof of Theorem 4.1.)** Let $\mathcal{L}_\varepsilon^{\lambda*}=\lambda^2\varepsilon^2\Delta^2-\operatorname{div}(A^*(x/\varepsilon,t/\varepsilon^2)\nabla)$ be the adjoint operator of $\mathcal{L}_\varepsilon^\lambda$, where $A^*=((A_{ij}^{\alpha\beta})^*)=(A_{ji}^{\beta\alpha}),1\le i,j\le d,1\le\alpha,\beta\le n$. For $H\in L^2(\Omega_T)$, let $v_{\varepsilon,\lambda}\in L^2(0,T;H_0^2(\Omega))$, be the weak solution to the problem

$$-\partial_t v_{\varepsilon,\lambda}+\mathcal{L}_\varepsilon^{\lambda*}v_{\varepsilon,\lambda}=H\ \text{ in } \Omega_T\quad\text{and}\quad v_{\varepsilon,\lambda}=0\ \text{ on } \Omega\times\{t=T\},$$

and $v_{0,\lambda}$ the homogenized solution. Then $v_{\varepsilon,\lambda}(T-t)$ and $v_{0,\lambda}(T-t)$ are the solutions to (1.5) and (1.8) with $h=0,F=H(x,T-t)$ and $A(x/\varepsilon,t/\varepsilon^2)$ replaced by $A^*(x/\varepsilon,(T-t)/\varepsilon^2)$. Let $\chi_T^{\lambda*}$ and $\mathfrak{B}_T^{\lambda*}$ be respectively the matrix of correctors and flux correctors for the family of parabolic operators $\partial_t+\lambda^2\varepsilon^2\Delta^2-\operatorname{div}\{A^*(x/\varepsilon,(T-t)/\varepsilon^2)\nabla\}$.

Similar to (4.6), we define

$$\begin{aligned}\widetilde{w}_{\varepsilon,\lambda}(t)=&v_{\varepsilon,\lambda}(T-t)-v_{0,\lambda}(T-t)\widetilde{\eta}_1+\varepsilon(\chi_T^{\lambda*})^\varepsilon\widetilde{K}_\varepsilon(\nabla v_{0,\lambda}(T-t))\\&+\varepsilon^2(\mathfrak{B}_{T,k(d+1)j}^{\lambda*})^\varepsilon\frac{\partial}{\partial x_k}\widetilde{K}_\varepsilon(\frac{\partial v_{0,\lambda}}{\partial x_j}(T-t)),\end{aligned}$$

where $\widetilde{K}_\varepsilon(f)=S_\varepsilon(f)(x,t)\widetilde{\eta}_\delta$, and $\widetilde{\eta}_\delta=\widetilde{\eta}_1(x)\widetilde{\eta}_2(t)$ is the smooth cut-off function, such that $0\le\widetilde{\eta}_1,\widetilde{\eta}_2\le 1$ and

$$\begin{aligned}&\widetilde{\eta}_\delta=1\ \text{ in } \Omega_T\backslash\Omega_{T,10\delta},\quad \widetilde{\eta}_\delta=0\ \text{ in } \Omega_{T,8\delta},\\&|\nabla\widetilde{\eta}_\delta|\le C\delta^{-1}\quad\text{and}\quad|\partial_t\widetilde{\eta}_\delta|+|\nabla^2\widetilde{\eta}_\delta|\le C\delta^{-2},\end{aligned}\tag{4.25}$$

where $\delta=(1+\lambda)\varepsilon$. Thanks to (4.13), we have

$$\lambda\varepsilon\|\Delta\widetilde{w}_{\varepsilon,\lambda}\|_{L^2(\Omega_T)}+\|\nabla\widetilde{w}_{\varepsilon,\lambda}\|_{L^2(\Omega_T)}\le C(1+\lambda)^{1/2}\varepsilon^{1/2}\|H\|_{L^2(\Omega_T)},\tag{4.26}$$

where C depends only on d,μ,n,T and Ω.

Note that

$$\begin{aligned}\Big|\int_{\Omega_T}w_{\varepsilon,\lambda}\cdot H\,dz\Big|=&\Big|\int_0^T\langle(\partial_t+\mathcal{L}_\varepsilon^\lambda)w_{\varepsilon,\lambda},\ v_{\varepsilon,\lambda}\rangle dt\Big|+\int_\Omega|h(1-\eta_{1,\delta})v_{\varepsilon,\lambda}(0)|\\\le&\Big|\int_0^T\langle(\partial_t+\mathcal{L}_\varepsilon^\lambda)w_{\varepsilon,\lambda},\widetilde{w}_{\varepsilon,\lambda}(T-t)\rangle dt\Big|\\&+\Big|\int_0^T\langle(\partial_t+\mathcal{L}_\varepsilon^\lambda)w_{\varepsilon,\lambda},\ v_{0,\lambda}(t)\widetilde{\eta}_1\rangle dt\Big|\\&+\Big|\int_0^T\langle(\partial_t+\mathcal{L}_\varepsilon^\lambda)w_{\varepsilon,\lambda},\ v_{\varepsilon,\lambda}(t)-v_{0,\lambda}(t)\widetilde{\eta}_1-\widetilde{w}_{\varepsilon,\lambda}(T-t)\rangle dt\Big|\end{aligned}$$

$$+\int_\Omega |h(1-\eta_{1,\delta})v_{\varepsilon,\lambda}(0)|. \tag{4.27}$$

Denote the terms in the right-hand side of (4.27) as J_1, J_2, J_3 and J_4 sequentially. Thanks to (4.11) and Lemma 3.5, we deduce that

$$\begin{aligned}
J_1 \le\, & C\varepsilon\big\{\|\nabla^2 u_{0,\lambda}\|_{L^2(\Omega_T\setminus\Omega_{T,2\delta})} + \|\partial_t u_{0,\lambda}\|_{L^2(\Omega_T\setminus\Omega_{T,2\delta})}\big\}\|\nabla\widetilde{w}_{\varepsilon,\lambda}(T-t)\|_{L^2(\Omega_T)}\\
& + C\delta^{3/2}\|\partial_t u_{0,\lambda}\|_{L^2(\Omega_T)}\|\nabla^2\widetilde{w}_{\varepsilon,\lambda}(T-t)\|_{L^2(\Omega_T)}\\
& + C\lambda^2\varepsilon^2\delta^{-1/2}\big\{\|\nabla^2 u_{0,\lambda}\|_{L^2(\Omega_T)} + \|\nabla u_{0,\lambda}\|_{L^2(\Omega_T)}\big\}\|\Delta\widetilde{w}_{\varepsilon,\lambda}(T-t)\|_{L^2(\Omega_{5\delta}\times(0,T))}\\
& + C\lambda^2\varepsilon^2\|\nabla^2 u_{0,\lambda}\|_{L^2(\Omega_T\setminus\Omega_{T,2\delta})}\|\Delta\widetilde{w}_{\varepsilon,\lambda}(T-t)\|_{L^2(\Omega_T)}\\
& + C\|\nabla u_{0,\lambda}\|_{L^2(\Omega_{T,5\delta})}\|\nabla\widetilde{w}_{\varepsilon,\lambda}(T-t)\|_{L^2(\Omega_{T,5\delta})}\\
\le\, & C(1+\lambda)\varepsilon\big\{\|F\|_{L^2(\Omega_T)} + \|h\|_{H^1(\Omega)}\big\}\|H\|_{L^2(\Omega_T)},
\end{aligned}$$

where we have used (4.19)-(4.21), and (4.26) for the last step.

On the other hand, since $\widetilde{\eta}_1 = 0$ in $\Omega_{8\delta}$, we derive from (4.11) that

$$\begin{aligned}
J_2 \le\, & C\varepsilon\big\{\|\nabla^2 u_{0,\lambda}\|_{L^2(\Omega_T\setminus\Omega_{T,2\delta})} + \|\partial_t u_{0,\lambda}\|_{L^2(\Omega_T\setminus\Omega_{T,2\delta})}\big\}\|\nabla(v_{0,\lambda}(t)\widetilde{\eta}_1)\|_{L^2(\Omega_T)}\\
& + C\lambda^2\varepsilon^2\|\nabla^2 u_{0,\lambda}\|_{L^2(\Omega_T\setminus\Omega_{T,2\delta})}\|\Delta(v_{0,\lambda}(t)\widetilde{\eta}_1)\|_{L^2(\Omega_T)}\\
& + C\|\nabla u_{0,\lambda}\|_{L^2(\Omega_{T,5\delta})}\|\nabla(v_{0,\lambda}(t)\widetilde{\eta}_1)\|_{L^2(\Omega_{T,5\delta})}\\
\doteq\, & J_{21} + J_{22} + J_{23},
\end{aligned} \tag{4.28}$$

where $\delta = (1+\lambda)\varepsilon$. By Lemma 3.5, (4.19) and (4.21), we deduce that

$$\begin{aligned}
J_{21} \le\, & C\varepsilon\big\{\|u_{0,\lambda}\|_{L^2(0,T;H^2(\Omega))} + \|\partial_t u_{0,\lambda}\|_{L^2(\Omega_T)}\big\}\\
& \times\big\{\|\nabla v_{0,\lambda}\|_{L^2(\Omega_T)} + \delta^{-1}\|v_{0,\lambda}\|_{L^2(\Omega_{4\delta}\times(0,T))}\big\}\\
\le\, & C\varepsilon\big\{\|F\|_{L^2(\Omega_T)} + \|h\|_{H^1(\Omega)}\big\}\|H\|_{L^2(\Omega_T)}.
\end{aligned} \tag{4.29}$$

Likewise, J_{22} can be bounded as following

$$\begin{aligned}
J_{22} \le\, & C(\lambda+1)\varepsilon\|u_{0,\lambda}\|_{L^2(0,T;H^2(\Omega))}\|\nabla^2 v_{0,\lambda}\|_{L^2(\Omega_T)}\\
\le\, & C(\lambda+1)\varepsilon\big\{\|F\|_{L^2(\Omega_T)} + \|h\|_{H^1(\Omega)}\big\}\|H\|_{L^2(\Omega_T)}.
\end{aligned} \tag{4.30}$$

For J_{23}, we note that by Lemma 3.5 and the definition of $\widetilde{\eta}_1$,

$$\|\nabla(v_{0,\lambda}(t)\widetilde{\eta}_1)\|_{L^2(\Omega_{T,5\delta})} \le C\|\nabla v_{0,\lambda}\|_{L^2(\Omega\times(0,(10\delta)^2))}.$$

In view of (4.20) and (4.21), we get

$$J_{23} \le C(1+\lambda)\varepsilon\big\{\|F\|_{L^2(\Omega_T)} + \|h\|_{H^1(\Omega)}\big\}\|H\|_{L^2(\Omega_T)},$$

which, together with (4.29) and (4.30), implies that

$$J_2 \le C(1+\lambda)\varepsilon\big\{\|F\|_{L^2(\Omega_T)} + \|h\|_{H^1(\Omega)}\big\}\|H\|_{L^2(\Omega_T)}.$$

To estimate J_3, we take $\psi = v_{\varepsilon,\lambda}(t) - v_{0,\lambda}(t)\widetilde{\eta}_1 - \widetilde{w}_{\varepsilon,\lambda}(T-t)$, which is zero in $\Omega_{T,8\delta}$, in (4.11). In view of the estimates on $\chi_T^{\lambda*}$ and $\mathfrak{B}_T^{\lambda*}$, we can perform similar analysis as we did for J_2 to derive that

$$J_3 \le C(1+\lambda)\varepsilon\big\{\|F\|_{L^2(\Omega_T)} + \|h\|_{H^1(\Omega)}\big\}\|H\|_{L^2(\Omega_T)}.$$

Finally, by Lemma 3.5,

$$J_4 \le C\|h\|_{L^2(\Omega_{10\delta})}\|v_{\varepsilon,\lambda}(0)\|_{L^2(\Omega)} \le C\delta\|h\|_{H_0^1(\Omega)}\|H\|_{L^2(\Omega_T)},$$

which, combined with the estimates on J_1-J_3 and (4.27), gives

$$\|w_\varepsilon\|_{L^2(\Omega_T)} \le C(1+\lambda)\varepsilon\big\{\|F\|_{L^2(\Omega_T)} + \|h\|_{H_0^1(\Omega)}\big\}. \tag{4.31}$$

Note that

$$\begin{aligned}
&\|u_{0,\lambda}(1-\eta_{1,\delta})\|_{L^2(\Omega_T)} \le C\|u_{0,\lambda}\|_{L^2(\Omega_{5\delta}\times(0,T))} \le C\delta\|\nabla u_{0,\lambda}\|_{L^2(\Omega_T)},\\
&\varepsilon\|(\chi^\lambda)^\varepsilon K_\varepsilon(\nabla u_{0,\lambda})\|_{L^2(\Omega_T)} \le C\varepsilon\|\nabla u_{0,\lambda}\|_{L^2(\Omega_T)},\\
&\varepsilon^2\Big\|(\mathfrak{B}^\lambda_{k(d+1)j})^\varepsilon \frac{\partial}{\partial x_k} K_\varepsilon\Big(\frac{\partial u_{0,\lambda}}{\partial x_j}\Big)\Big\|_{L^2(\Omega_T)} \le C\varepsilon\|\nabla u_{0,\lambda}\|_{L^2(\Omega_T)},
\end{aligned}$$

where (3.23) has been used for the last two estimates. We obtain (4.22) from (4.31) and complete the proof immediately. ∎

5. Convergence rate for the initial-Navier problem

This part is devoted to the convergence rate of problem (1.5) with the Navier boundary condition (1.11). For fixed $0<\lambda<\infty$, let $\mathcal{L}_\varepsilon^\lambda$ and $\mathcal{L}_0^\lambda$ be defined as in (3.1) and (3.4) respectively. Let $u_{\varepsilon,\lambda}$ be the weak solution of

$$\begin{aligned}
&(\partial_t + \mathcal{L}_\varepsilon^\lambda)u_{\varepsilon,\lambda} = F \quad \text{in } \Omega_T,\\
&u_{\varepsilon,\lambda} = h \ \text{ on } \Omega\times\{t=0\} \quad \text{and} \quad u_{\varepsilon,\lambda} = \Delta u_{\varepsilon,\lambda} = 0 \ \text{ on } \Gamma_T.
\end{aligned} \tag{5.1}$$

To prove Theorem 1.2, we introduce the following intermediate problem

$$\begin{aligned}
&(\partial_t + \mathcal{L}_0^\lambda)v_{\varepsilon,\lambda} + \lambda^2\varepsilon^2\Delta^2 v_{\varepsilon,\lambda} = F \quad \text{in } \Omega_T,\\
&v_{\varepsilon,\lambda} = h \ \text{ on } \Omega\times\{t=0\} \quad \text{and} \quad v_{\varepsilon,\lambda} = \Delta v_{\varepsilon,\lambda} = 0 \ \text{ on } \Gamma_T.
\end{aligned} \tag{5.2}$$

Let χ^λ and $\mathfrak{B}^\lambda$ be the correctors and flux correctors introduced in (3.2) and Lemma 3.3. Similar to (4.6), we define

$$w^1_{\varepsilon,\lambda}(x,t) = u_{\varepsilon,\lambda} - v_{\varepsilon,\lambda} - \varepsilon(\chi^\lambda)^\varepsilon K_\varepsilon(\nabla v_{\varepsilon,\lambda}) + \varepsilon^2(\mathfrak{B}^\lambda_{k(d+1)j})^\varepsilon \frac{\partial}{\partial x_k} K_\varepsilon\Big(\frac{\partial v_{\varepsilon,\lambda}}{\partial x_j}\Big), \tag{5.3}$$

where $K_\varepsilon(f)(x,t) = S_\varepsilon(f)(x,t)\eta_\varepsilon(x,t)$, with $\eta_\varepsilon(x,t) = \eta_{1,\varepsilon}(x)\eta_{2,\varepsilon}(t)$ being the smooth cut-off function, such that $0 \le \eta_{1,\varepsilon}, \eta_{2,\varepsilon} \le 1$ and

$$\begin{aligned}
&\eta_\varepsilon = 1 \ \text{ in } \Omega_T\setminus\Omega_{T,4\varepsilon}, \quad \eta_\varepsilon = 0 \ \text{ in } \Omega_{T,3\varepsilon},\\
&|\nabla\eta_\varepsilon| \le C\varepsilon^{-1}, \quad |\partial_t\eta_\varepsilon| + |\nabla^2\eta_\varepsilon| \le C\varepsilon^{-2}.
\end{aligned}$$

We remark that since $v_{\varepsilon,\lambda}$ satisfies the same boundary condition as $u_{\varepsilon,\lambda}$, we do not need to multiply it with the cut-off function $\eta_{1,\delta}$ as we did for $u_{0,\lambda}$ in (4.6).

LEMMA 5.1. *Let Ω be a bounded Lipschitz domain in $\mathbb{R}^d$ and $0<T<\infty$. Suppose A satisfies conditions (1.2)-(1.3). Let $u_{\varepsilon,\lambda}, v_{\varepsilon,\lambda}$ be, respectively, the solutions to (5.1) and (5.2). Let $w^1_{\varepsilon,\lambda}$ be defined as in (5.3). Then we have*

$$\begin{aligned}
&(\partial_t + \mathcal{L}_\varepsilon^\lambda)w^1_{\varepsilon,\lambda}\\
&= -\,div\big\{(\widehat{A^\lambda} - A^\varepsilon)(\nabla v_{\varepsilon,\lambda} - K_\varepsilon(\nabla v_{\varepsilon,\lambda}))\big\} - \varepsilon\, div\big\{(\mathfrak{B}^\lambda)^\varepsilon \nabla K_\varepsilon(\nabla v_{\varepsilon,\lambda})\big\}
\end{aligned}$$

$$
\begin{aligned}
&+\varepsilon\, div\big\{A^\varepsilon(\chi^\lambda)^\varepsilon\nabla K_\varepsilon(\nabla v_{\varepsilon,\lambda})\big\}-\varepsilon^2\frac{\partial}{\partial x_i}\Big\{(\mathfrak{B}^\lambda_{(d+1)ij})^\varepsilon\partial_t K_\varepsilon\Big(\frac{\partial v_{\varepsilon,\lambda}}{\partial x_j}\Big)\Big\}\\
&-\lambda^2\varepsilon\, div\big\{(\Delta\chi^\lambda)^\varepsilon\nabla K_\varepsilon(\nabla v_{\varepsilon,\lambda})\big\}-2\lambda^2\varepsilon^2\Delta\big\{(\nabla\chi^\lambda)^\varepsilon\nabla K_\varepsilon(\nabla v_{\varepsilon,\lambda})\big\}\\
&-\lambda^2\varepsilon^3\Delta\big\{(\chi^\lambda)^\varepsilon\Delta K_\varepsilon(\nabla v_{\varepsilon,\lambda})\big\}-\varepsilon\, div\Big\{A^\varepsilon(\nabla\mathfrak{B}^\lambda_{k(d+1)j})^\varepsilon\frac{\partial}{\partial x_k}K_\varepsilon\Big(\frac{\partial v_{\varepsilon,\lambda}}{\partial x_j}\Big)\Big\}\\
&-\varepsilon^2\, div\Big\{A^\varepsilon(\mathfrak{B}^\lambda_{k(d+1)j})^\varepsilon\nabla\frac{\partial}{\partial x_k}K_\varepsilon\Big(\frac{\partial v_{\varepsilon,\lambda}}{\partial x_j}\Big)\Big\}\\
&+\lambda^2\varepsilon^2\Delta\Big\{(\Delta\mathfrak{B}^\lambda_{k(d+1)j})^\varepsilon\frac{\partial}{\partial x_k}K_\varepsilon\Big(\frac{\partial v_{\varepsilon,\lambda}}{\partial x_j}\Big)+2\varepsilon(\nabla\mathfrak{B}^\lambda_{k(d+1)j})^\varepsilon\nabla\frac{\partial}{\partial x_k}K_\varepsilon\Big(\frac{\partial v_{\varepsilon,\lambda}}{\partial x_j}\Big)\\
&\qquad+\varepsilon^2(\mathfrak{B}^\lambda_{k(d+1)j})^\varepsilon\Delta\frac{\partial}{\partial x_k}K_\varepsilon\Big(\frac{\partial v_{\varepsilon,\lambda}}{\partial x_j}\Big)\Big\}. \qquad (5.4)
\end{aligned}
$$

Proof. Similar to (4.8), we can prove that

$$
\begin{aligned}
(\partial_t+\mathcal{L}^\lambda_\varepsilon)w^1_{\varepsilon,\lambda}=&-\mathrm{div}\big\{(\widehat{A^\lambda}-A^\varepsilon)(\nabla v_{\varepsilon,\lambda}-K_\varepsilon(\nabla v_{\varepsilon,\lambda}))\big\}\\
&+\mathrm{div}\big\{(B^\lambda)^\varepsilon K_\varepsilon(\nabla v_{\varepsilon,\lambda})\big\}-\varepsilon\partial_t\big\{(\chi^\lambda)^\varepsilon K_\varepsilon(\nabla v_{\varepsilon,\lambda})\big\}\\
&+\varepsilon\mathrm{div}\big\{A^\varepsilon(\chi^\lambda)^\varepsilon\nabla K_\varepsilon(\nabla v_{\varepsilon,\lambda})\big\}-\lambda^2\varepsilon\mathrm{div}\big\{(\Delta\chi^\lambda)^\varepsilon\nabla K_\varepsilon(\nabla v_{\varepsilon,\lambda})\big\}\\
&-2\lambda^2\varepsilon^2\Delta\big\{(\nabla\chi^\lambda)^\varepsilon\nabla K_\varepsilon(\nabla v_{\varepsilon,\lambda})-\lambda^2\varepsilon^3\Delta\big\{(\chi^\lambda)^\varepsilon\Delta K_\varepsilon(\nabla v_{\varepsilon,\lambda})\big\}\\
&+(\partial_t+\mathcal{L}^\lambda_\varepsilon)\Big\{\varepsilon^2(\mathfrak{B}^\lambda_{k(d+1)j})^\varepsilon\frac{\partial}{\partial x_k}K_\varepsilon\Big(\frac{\partial v_{\varepsilon,\lambda}}{\partial x_j}\Big)\Big\}.
\end{aligned}
$$

The remaining proof is the same as Lemma 4.1. We therefore omit the details. ∎

LEMMA 5.2. *Under the assumption of Lemma 5.1, for any $\psi\in L^2(0,T;H^1_0(\Omega)\cap H^2(\Omega))$ we have*

$$
\begin{aligned}
&\Big|\int_0^T\langle(\partial_t+\mathcal{L}^\lambda_\varepsilon)w^1_{\varepsilon,\lambda},\psi\rangle dt\Big|\\
&\le C\varepsilon\big\{\|\nabla^2 v_{\varepsilon,\lambda}\|_{L^2(\Omega_T\setminus\Omega_{T,2\varepsilon})}+\|\partial_t v_{\varepsilon,\lambda}\|_{L^2(\Omega_T\setminus\Omega_{T,2\varepsilon})}\big\}\|\nabla\psi\|_{L^2(\Omega_T)}\\
&\quad+C\lambda^2(1+\lambda)^{-2}\varepsilon\|\nabla v_{\varepsilon,\lambda}\|_{L^2(\Omega_{5\varepsilon}\times(0,T))}\|\Delta\psi\|_{L^2(\Omega_{5\varepsilon}\times(0,T))}\\
&\quad+C\lambda^2(1+\lambda)^{-2}\varepsilon^2\|\nabla^2 v_{\varepsilon,\lambda}\|_{L^2(\Omega_T\setminus\Omega_{T,2\varepsilon})}\|\Delta\psi\|_{L^2(\Omega_T)}\\
&\quad+C\|\nabla v_{\varepsilon,\lambda}\|_{L^2(\Omega_{T,5\varepsilon})}\|\nabla\psi\|_{L^2(\Omega_{T,5\varepsilon})}, \qquad (5.5)
\end{aligned}
$$

where $\Omega_{T,\delta}$ is given by (3.24), and C depends only on d,n,μ,T and Ω.

Proof. By (5.4), we have

$$
\begin{aligned}
&\Big|\int_0^T\langle(\partial_t+\mathcal{L}_\varepsilon)w^1_{\varepsilon,\lambda},\psi\rangle dt\Big|\\
&\le C\int_{\Omega_T}|\nabla v_{\varepsilon,\lambda}-K_\varepsilon(\nabla v_{\varepsilon,\lambda})||\nabla\psi|+C\varepsilon\int_{\Omega_T}|(\mathfrak{B}^\lambda)^\varepsilon\nabla K_\varepsilon(\nabla v_{\varepsilon,\lambda})||\nabla\psi|\\
&\quad+C\varepsilon\int_{\Omega_T}|(\chi^\lambda)^\varepsilon\nabla K_\varepsilon(\nabla v_{\varepsilon,\lambda})||\nabla\psi|+C\varepsilon^2\int_{\Omega_T}\Big|(\mathfrak{B}^\lambda_{(d+1)ij})^\varepsilon\partial_t K_\varepsilon(\frac{\partial v_{\varepsilon,\lambda}}{\partial x_j})\Big||\partial_{x_i}\psi|\\
&\quad+C\lambda^2\varepsilon\int_{\Omega_T}|(\Delta\chi^\lambda)^\varepsilon\nabla K_\varepsilon(\nabla v_{\varepsilon,\lambda})||\nabla\psi|+C\lambda^2\varepsilon^2\int_{\Omega_T}|(\nabla\chi^\lambda)^\varepsilon\nabla K_\varepsilon(\nabla v_{\varepsilon,\lambda})||\Delta\psi|\\
&\quad+C\varepsilon\int_{\Omega_T}\Big|(\nabla\mathfrak{B}^\lambda_{k(d+1)j})^\varepsilon\frac{\partial}{\partial x_k}K_\varepsilon(\frac{\partial v_{\varepsilon,\lambda}}{\partial x_j})\Big||\nabla\psi|+C\lambda^2\varepsilon^3\int_{\Omega_T}|(\chi^\lambda)^\varepsilon\Delta K_\varepsilon(\nabla v_{\varepsilon,\lambda})||\Delta\psi|
\end{aligned}
$$

$$
\begin{aligned}
&+C\varepsilon^2\int_{\Omega_T}\Big|(\mathfrak{B}^{\lambda}_{k(d+1)j})^{\varepsilon}\frac{\partial}{\partial x_k}\nabla K_\varepsilon(\frac{\partial v_{\varepsilon,\lambda}}{\partial x_j})\Big||\nabla\psi|\\
&+C\lambda^2\varepsilon^2\int_{\Omega_T}\Big|(\Delta\mathfrak{B}^{\lambda}_{k(d+1)j})^{\varepsilon}\frac{\partial}{\partial x_k}K_\varepsilon(\frac{\partial v_{\varepsilon,\lambda}}{\partial x_j})\Big||\Delta\psi|\\
&+C\lambda^2\varepsilon^3\int_{\Omega_T}\Big|(\nabla\mathfrak{B}^{\lambda}_{k(d+1)j})^{\varepsilon}\nabla\frac{\partial}{\partial x_k}K_\varepsilon(\frac{\partial v_{\varepsilon,\lambda}}{\partial x_j})\Big||\Delta\psi|\\
&+C\lambda^2\varepsilon^4\int_{\Omega_T}\Big|(\mathfrak{B}^{\lambda}_{k(d+1)j})^{\varepsilon}\Delta\frac{\partial}{\partial x_k}K_\varepsilon\Big(\frac{\partial v_{\varepsilon,\lambda}}{\partial x_j}\Big)\Big||\Delta\psi|, \qquad (5.6)
\end{aligned}
$$

where C depends only on d,n and μ. Note that (3.20) implies that for $\lambda\geq 1$,

$$\|\nabla^k\mathfrak{B}^{\lambda,\alpha\beta}_{\varsigma\tau j}\|_{L^2(\mathbb{T}^{d+1})}\leq C(1+\lambda)^{-2},$$

for $k=0,1,2$. By performing the same analysis as in Lemma 4.2, we derive (5.5) immediately. ∎

Lemma 5.3. *Let Ω be a bounded $C^{1,1}$ domain in $\mathbb{R}^d$ and $0<T<\infty$. Suppose A satisfies conditions (1.2)-(1.3). Let $u_{\varepsilon,\lambda},v_{\varepsilon,\lambda}$ be, respectively, the solutions to (5.1) and (5.2) with $F\in L^2(\Omega_T),h\in H^2(\Omega)$. Then*

$$\lambda\varepsilon\|\Delta w^1_{\varepsilon,\lambda}\|_{L^2(\Omega_T)}+\|\nabla w^1_{\varepsilon,\lambda}\|_{L^2(\Omega_T)}\leq C\varepsilon^{1/2}\big\{\|h\|_{H^2(\Omega)}+\|F\|_{L^2(\Omega_T)}\big\}, \qquad (5.7)$$

where $w^1_{\varepsilon,\lambda}$ is defined as in (5.3), and C depends only on d,n,μ,T and Ω.

Proof. Since $w^1_{\varepsilon,\lambda}\in L^2(0,T;H^1_0(\Omega))$, $w^1_{\varepsilon,\lambda}(x,0)=0$, and $\Delta w^1_{\varepsilon,\lambda}=0$ on Γ_T. By taking $\psi=w^1_{\varepsilon,\lambda}$ in (5.5) and the Cauchy inequality, we obtain that

$$
\begin{aligned}
&\lambda^2\varepsilon^2\|\Delta w^1_{\varepsilon,\lambda}\|^2_{L^2(\Omega_T)}+\|\nabla w^1_{\varepsilon\lambda}\|^2_{L^2(\Omega_T)}\\
&\leq C\varepsilon\big\{\|\nabla^2 v_{\varepsilon,\lambda}\|^2_{L^2(\Omega_T)}+\|\partial_t v_{\varepsilon,\lambda}\|^2_{L^2(\Omega_T)}\big\}+C\|\nabla v_{\varepsilon,\lambda}\|^2_{L^2(\Omega_{T,5\varepsilon})}. \qquad (5.8)
\end{aligned}
$$

Let $v_{0,\lambda}$ be the solution to the limit problem of (5.2), i.e.,

$$
\begin{aligned}
&(\partial_t+\mathcal{L}^{\lambda}_0)v_{0,\lambda}=F \quad \text{in } \Omega_T,\\
&v_{0,\lambda}=h \ \text{ on } \Omega\times\{t=0\} \quad \text{and} \quad v_{0,\lambda}=0 \ \text{ on } \Gamma_T. \qquad (5.9)
\end{aligned}
$$

We have $v_{\varepsilon,\lambda}=v_{0,\lambda}$ on $\partial\Omega_T$, and

$$\partial_t(v_{\varepsilon,\lambda}-v_{0,\lambda})-\operatorname{div}(\widehat{A^\lambda}\nabla(v_{\varepsilon,\lambda}-v_{0,\lambda}))+\lambda^2\varepsilon^2\Delta^2 v_{\varepsilon,\lambda}=0 \quad \text{in } \Omega_T.$$

Taking $v_{\varepsilon,\lambda}-v_{0,\lambda}$ as a test function, we deduce that

$$\|\nabla(v_{\varepsilon,\lambda}-v_{0,\lambda})\|_{L^2(\Omega_T)}+\lambda\varepsilon\|\Delta v_{\varepsilon,\lambda}\|_{L^2(\Omega_T)}\leq C\lambda\varepsilon\|\Delta v_{0,\lambda}\|_{L^2(\Omega_T)}. \qquad (5.10)$$

Furthermore, taking $\partial_t v_{\varepsilon,\lambda}$ as the test function in (5.2), it yields

$$
\begin{aligned}
&\|\partial_t v_{\varepsilon,\lambda}\|_{L^2(\Omega_T)}+\lambda\varepsilon\sup_{0\leq t\leq T}\|\Delta v_{\varepsilon,\lambda}(t)\|_{L^2(\Omega)}\\
&\leq C\{\|\nabla^2 v_{\varepsilon,\lambda}\|_{L^2(\Omega_T)}+\|F\|_{L^2(\Omega_T)}+\lambda\varepsilon\|h\|_{H^2(\Omega)}\}. \qquad (5.11)
\end{aligned}
$$

Note that similar to (4.18), we have

$$
\begin{aligned}
\int_0^{\varepsilon^2}\|\nabla v_{\varepsilon,\lambda}\|^2_{L^2(\Omega)}ds&\leq C\varepsilon\big\{\|\partial_t v_{\varepsilon,\lambda}\|_{L^2(\Omega_T)}+\|F\|_{L^2(\Omega_T)}\big\}\sup_{0<t<T}\|v_{\varepsilon,\lambda}(\cdot,t)\|_{L^2(\Omega)}\\
&\leq C\varepsilon\big\{\|\nabla^2 v_{\varepsilon,\lambda}\|^2_{L^2(\Omega_T)}+\|F\|^2_{L^2(\Omega_T)}+\|h\|^2_{H^2(\Omega)}\big\},
\end{aligned}
$$

where we have used (5.11) for the last step. This, combined with (4.17) (for $v_{\varepsilon,\lambda}$), gives

$$\|\nabla v_{\varepsilon,\lambda}\|^2_{L^2(\Omega_{T,5\varepsilon})} \leq C\varepsilon\big\{\|\nabla^2 v_{\varepsilon,\lambda}\|^2_{L^2(\Omega_T)} + \|F\|^2_{L^2(\Omega_T)} + \|h\|^2_{H^2(\Omega)}\big\}. \tag{5.12}$$

By taking (5.10)-(5.12) into (5.8) and using the H^2 estimate for $v_{0,\lambda}$ (see e.g., (4.21)), one gets (5.7) immediately. □

With Lemmas 5.2 and 5.3 at our disposal, we can prove the optimal error estimate between $u_{\varepsilon,\lambda}$ and $v_{\varepsilon,\lambda}$ by using the duality argument as in Theorem 4.1.

THEOREM 5.1. *Suppose Ω is a bounded $C^{1,1}$ domain in $\mathbb{R}^d$ and $0<T<\infty$. Assume A satisfies conditions (1.2)-(1.3). Let $u_{\varepsilon,\lambda}, v_{\varepsilon,\lambda}$ be, respectively, the solutions to (5.1) and (5.2) with $F\in L^2(\Omega_T)$ and $h\in H^2(\Omega)$. Then for any fixed $0<\lambda<\infty$,*

$$\|u_{\varepsilon,\lambda} - v_{\varepsilon,\lambda}\|_{L^2(\Omega_T)} \leq C\varepsilon\big\{\|h\|_{H^2(\Omega)} + \|F\|_{L^2(\Omega_T)}\big\}, \tag{5.13}$$

where C depends only on d, n, μ, Ω and T.

Proof. The proof is completely parallel to Theorem 4.1. Let us omit the details. □

Proof. **(Proof of Theorem 1.2.)** Let u_ε be the solution of (1.5) and (1.11), and u_0 the solution to (1.8). Note that u_ε is the solution to (5.1) with $\lambda=\kappa/\varepsilon$, and

$$\|u_\varepsilon - u_0\|_{L^2(\Omega_T)} \leq \|u_{\varepsilon,\lambda} - v_{\varepsilon,\lambda}\|_{L^2(\Omega_T)} + \|v_{\varepsilon,\lambda} - v_{0,\lambda}\|_{L^2(\Omega_T)} + \|v_{0,\lambda} - u_0\|_{L^2(\Omega_T)}, \tag{5.14}$$

where $v_{0,\lambda}$ is the solution to (5.9). Thanks to (4.24), we know that

$$\|v_{0,\lambda} - u_0\|_{L^2(0,T;H^1_0(\Omega))} \leq C|\widehat{A^\lambda} - \widehat{A}|\|u_0\|_{L^2(0,T;H^1(\Omega))}. \tag{5.15}$$

In view of Theorem 2.2 and (4.21), we have

$$\|v_{\varepsilon,\lambda} - v_{0,\lambda}\|_{L^2(\Omega_T)} \leq C\kappa^2\|v_{0,\lambda}\|_{L^2(0,T;H^2(\Omega))} \leq C\kappa^2\big\{\|h\|_{H^1(\Omega)} + \|F\|_{L^2(\Omega_T)}\big\}. \tag{5.16}$$

By taking (5.13), (5.15) and (5.16) into (5.14), and using (3.9), we obtain (1.12) and complete the proof. □

REMARK 5.1. We have mentioned that the proof of Theorem 1.2 is slightly different from the one of Theorem 1.1. Since we have observed that the convergence rate of the pure singular perturbation problem with Navier boundary conditions admits better error estimate, we therefore introduced the intermediate problem (5.2) to separate the settings of singular perturbation and homogenization. By calculating the error estimates in each process individually, we eventually derive the desirable error estimate (1.12).

For the initial Dirichlet problem (1.5)-(1.6), one may also perform similar analysis (with some modifications as stated in Remark 2.2) to consider the error estimates in homogenization and singular perturbation separately. But there is no need to do so, as we mentioned before that the optimal convergence rate for the pure singular perturbation problem with Dirichlet boundary conditions is $O(\kappa)$ [6]. One can therefore consider the two processes in a deal as in the proof of Theorem 1.1 without any loss.

Acknowledgement. The second author would like to thank Professor Zhongwei Shen for the enlightening discussions. This work is supported by the NSF of China (11971031) and NSF of Anhui Province(2108085Y01).

REFERENCES

[1] S. Armstrong, A. Bordas, and J.C. Mourrat, *Quantitative stochastic homogenization and regularity theory of parabolic equations*, Anal. PDE, 11:1945–2014, 2018. 1

[2] A. Bensoussan, J.L. Lions, and G. Papanicolaou, *Asymptotic Analysis for Periodic Structures*, North-Holland Publishing Company Amsterdam, 5, 1978. 1, 1, 3.1, 3.1

[3] J.G. Besjes, *Singular perturbation problems for linear parabolic differential operators of arbitrary order*, J. Math. Anal. Appl., 48:594–609, 1974. 1

[4] H. Dong and D. Kim, *On L_p-estimates for elliptic and parabolic equations with A_p weights*, Trans. Amer. Math. Soc., 370:5081–5130, 2018. 2.2

[5] G.A. Francfort and S. Müller, *Combined effects of homogenization and singular perturbations in elasticity*, J. Reine Angew. Math., 454:1–35, 1994. 1

[6] A. Friedman, *Singular perturbations for partial differential equations*, Arch. Ration. Mech. Anal., 29:289–303, 1968. 1, 1, 2.2, 5.1

[7] J. Geng and W. Niu, *Homogenization of locally periodic parabolic operators with non-self-similar scales*, arXiv preprint, arXiv:2103.01418, 2021. 1

[8] J. Geng and Z. Shen, *Convergence rates in parabolic homogenization with time-dependent periodic coefficients*, J. Funct. Anal., 272:2092–2113, 2017. 1, 1, 3.2, 3.2, 4, 4

[9] J. Geng and Z. Shen, *Homogenization of parabolic equations with non-self-similar scales*, Arch. Ration. Mech. Anal., 236:145–188, 2020. 1

[10] J. Geng and Z. Shen, *Asymptotic expansions of fundamental solutions in parabolic homogenization*, Anal. PDE, 13:147–170, 2020. 1

[11] W.M. Greenlee, *Rate of convergence in singular perturbations*, Ann. Inst. Fourier (Grenoble), 18:135–191, 1968. 1

[12] J.L. Lions, *Perturbations Singulières dans les Problèmes aux Limites et en Contrôle Optimal*, Lecture Notes in Mathematics, Springer-Verlag, Berlin-New York, 323, 1973. 1, 4

[13] B. Najman, *The rate of convergence in singular perturbations of parabolic equations*, in F. Kappel and W. Schappacher (eds.), Infinite-dimensional Systems (Retzhof, 1983), Lecture Notes in Math., Springer, Berlin, 1076:147–167, 1984. 1

[14] W. Niu and Z. Shen, *Combined effects of homogenization and singular perturbations: Quantitative estimates*, Asymptot. Anal., 128(3):351–384, 2022. 1, 3.2

[15] W. Niu and Y. Xu, *Convergence rates in homogenization of higher order parabolic systems*, Discrete Contin. Dyn. Syst. Ser. A, 38:4203–4229, 2018. 1, 3.2

[16] W. Niu and Y. Xu, *A refined convergence result in homogenization of second order parabolic systems*, J. Differ. Equ., 266:8294–8319, 2019. 1

[17] W. Niu and Y. Yuan, *Convergence rate in homogenization of elliptic systems with singular perturbations*, J. Math. Phys., 60:111509, 2019. 1

[18] S.E. Pastukhova, *Homogenization estimates for singularly perturbed operators*, J. Math. Sci., 251:724–747, 2020. 1

[19] Z. Schuss, *Singular perturbations and the transition from thin plate to membrane*, Proc. Amer. Math. Soc., 58:139–147, 1976. 1

[20] Z. Shen, *Periodic Homogenization of Elliptic Systems*, Birkhäuser/Springer, Cham., 2018. 1

[21] Z. Shen and J. Zhuge, *Convergence rates in periodic homogenization of systems of elasticity*, Proc. Amer. Math. Soc., 145:1187–1202, 2017. 1

[22] T.A. Suslina, *Homogenization of the Dirichlet problem for elliptic systems: L^2-operator error estimates*, Mathematika, 59:463–476, 2013. 1, 4

[23] T.A. Suslina, *Homogenization of higher-order parabolic systems in a bounded domain*, Appl. Anal., 98:3–31, 2019. 1

[24] Q. Xu and S. Zhou, *Quantitative estimates in homogenization of parabolic systems of elasticity in Lipschitz cylinders*, arXiv preprint, arXiv:1705.01479, 2017. 1

[25] Y. Xu, *Convergence rates in homogenization of parabolic systems with locally periodic coefficients*, arXiv preprint, arXiv:2007.03853, 2020. 1

[26] V.V. Zhikov and S.E. Pastukhova, *On operator estimates for some problems in homogenization theory*, Russ. J. Math. Phys., 12:515–524, 2005. 4