



# Homogenization of Parabolic Equations with Non-self-similar Scales

JUN GENG & ZHONGWEI SHEN 

Communicated by C. LE BRIS

## Abstract

This paper is concerned with quantitative homogenization of second-order parabolic systems with periodic coefficients varying rapidly in space and time, in non-self-similar scales. The homogenization problem involves two oscillating scales. We obtain large-scale interior and boundary Lipschitz estimates as well as interior  $C^{1,\alpha}$  and  $C^{2,\alpha}$  estimates by utilizing higher-order correctors. We also investigate the problem of convergence rates for initial-boundary value problems.

**Mathematics Subject Classification** 35B27 35K40

## 1. Introduction

In this paper we shall be interested in the quantitative homogenization of a parabolic operator with periodic coefficients varying rapidly in space and time, in different scales. More precisely, we consider the parabolic operator

$$\partial_t + \mathcal{L}_\varepsilon \tag{1.1}$$

in  $\mathbb{R}^{d+1}$ , where  $\varepsilon > 0$  and

$$\mathcal{L}_\varepsilon = -\operatorname{div}(A(x/\varepsilon, t/\varepsilon^k)\nabla), \tag{1.2}$$

with  $0 < k < \infty$ . We will assume that the coefficient tensor  $A = A(y, s) = (a_{ij}^{\alpha\beta}(y, s))$ , with  $1 \leq i, j \leq d$  and  $1 \leq \alpha, \beta \leq m$ , is real, bounded measurable and satisfies the ellipticity condition

$$\|A\|_\infty \leq \mu^{-1} \quad \text{and} \quad \mu|\xi|^2 \leq a_{ij}^{\alpha\beta}(y, s)\xi_i^\alpha\xi_j^\beta \tag{1.3}$$

for any  $\xi = (\xi_i^\alpha) \in \mathbb{R}^{m \times d}$  and almost everywhere  $(y, s) \in \mathbb{R}^{d+1}$ , where  $\mu > 0$  (the summation convention is used throughout). We also assume that  $A$  is 1-periodic in  $(y, s)$ ; that is

$$A(y + z, s + t) = A(y, s) \quad \text{for } (z, t) \in \mathbb{Z}^{d+1} \text{ and almost everywhere } (y, s) \in \mathbb{R}^{d+1}. \quad (1.4)$$

The qualitative homogenization theory for the operator (1.1) has been known since the 1970s (see for example [10]). As  $\varepsilon \rightarrow 0$ , the weak solution  $u_\varepsilon$  of the initial-Dirichlet problem for the parabolic system  $(\partial_t + \mathcal{L}_\varepsilon)u_\varepsilon = F$  in  $\Omega_T = \Omega \times (0, T)$  converges weakly in  $L^2(0, T; H^1(\Omega))$  and strongly in  $L^2(\Omega_T)$ . Moreover, the limit  $u_0$  is a solution of the initial-Dirichlet problem for  $(\partial_t + \mathcal{L}_0)u_0 = F$  in  $\Omega_T$ , where  $\mathcal{L}_0$  is a second-order elliptic operator with constant coefficients. Furthermore, the (homogenized) coefficients of  $\mathcal{L}_0$  as well as the first-order correctors depend on  $k$ , but only for three separated cases:  $0 < k < 2$ ;  $k = 2$ ; and  $2 < k < \infty$ .

In recent years there has been a great amount of interest in the quantitative homogenization theory for partial differential equations, where one is concerned with problems related to the large-scale regularity and convergence rates for solutions  $u_\varepsilon$ . Major progress has been made for elliptic equations and systems in the periodic and non-periodic settings (see [3–9, 11, 16, 17, 19, 27, 28] and references therein). Some of these works have been extended to parabolic equations and systems in the self-similar case  $k = 2$ . In particular, we established the large-scale Lipschitz and  $W^{1,p}$  estimates in [13] and studied the problem of convergence rates in  $L^2(\Omega_T)$  as well as error estimates for two-scale expansions in  $L^2(0, T; H^1(\Omega))$  in [14]. Also see related works in [22–24, 30]. Most recently, in [15], we have studied the asymptotic behavior of the fundamental solution and its derivatives and established sharp estimates for the remainders. We refer the reader to [2] for quantitative stochastic homogenization of parabolic equations.

If  $k \neq 2$ , the  $\varepsilon$  scaling in the coefficient tensor  $A(x/\varepsilon, t/\varepsilon^k)$  is not consistent with the intrinsic scaling  $(x, t) \rightarrow (\delta x, \delta^2 t)$  of the second-order parabolic equations. To the authors' best knowledge, very few quantitative results are known in this case. Direct extensions of the existing techniques developed for elliptic equations fail due to the fact that the homogenization problem involves two oscillating scales mentioned above. For more recent work as well as motivations on homogenization problems with more than one oscillating scale, which are referred to as reiterated homogenization, see [1, 12, 18, 26, 29] and references therein.

In this paper we develop a new approach to study homogenization of parabolic equations and systems with non-self-similar scales. This allows us to establish large-scale interior and boundary Lipschitz estimates for the parabolic operator (1.1) with any  $0 < k < \infty$ , under conditions (1.3) and (1.4).

Let  $Q_r(x_0, t_0) = B(x_0, r) \times (t_0 - r^2, t_0)$  denote a parabolic cylinder. The following is one of the main results of the paper:

**Theorem 1.1.** *Assume  $A = A(y, s)$  satisfies (1.3) and (1.4). Let  $u_\varepsilon$  be a weak solution to*

$$(\partial_t + \mathcal{L}_\varepsilon)u_\varepsilon = F \quad \text{in } Q_R = Q_R(x_0, t_0), \quad (1.5)$$

where  $R > \varepsilon + \varepsilon^{k/2}$  and  $F \in L^p(Q_R)$  for some  $p > d + 2$ . Then for any  $\varepsilon + \varepsilon^{k/2} \leq r < R$ ,

$$\left( \int_{Q_r} |\nabla u_\varepsilon|^2 \right)^{1/2} \leq C \left\{ \left( \int_{Q_R} |\nabla u_\varepsilon|^2 \right)^{1/2} + R \left( \int_{Q_R} |F|^p \right)^{1/p} \right\}, \quad (1.6)$$

where  $C$  depends only on  $d, m, p$ , and  $\mu$ .

The inequality (1.6) may be regarded as a large-scale interior Lipschitz estimate. We also obtain large-scale  $C^{1,\alpha}$  and  $C^{2,\alpha}$  excess-decay estimates, which are new even for  $k = 2$ , for solutions of  $\partial_t + \mathcal{L}_\varepsilon$  (see Sections 4 and 5). Regarding the condition  $R > r \geq \varepsilon + \varepsilon^{k/2}$ , we mention that there exists  $u_\varepsilon$  such that  $(\partial_t + \mathcal{L}_\varepsilon)u_\varepsilon = 0$  in  $\mathbb{R}^{d+1}$  and  $\nabla u_\varepsilon$  is  $\varepsilon$ -periodic in  $x$  and  $\varepsilon^k$ -periodic in  $t$  (the solution  $u_\varepsilon$  is given by  $x_j + \varepsilon \chi_j^\lambda(x/\varepsilon, t/\varepsilon^2)$  with  $\lambda = \varepsilon^{k-2}$ ; see Section 2). Note that if the periodic cell  $(0, \varepsilon)^d \times (-\varepsilon^k, 0)$  for  $\nabla u_\varepsilon$  is contained in the parabolic cylinder  $Q_r(0, 0)$ , then  $r^2 \geq \varepsilon^k$  and  $2r \geq \sqrt{d}\varepsilon$ . This implies that  $r \geq (\varepsilon + \varepsilon^{k/2})/4$ . As a result, the condition  $R > r \geq \varepsilon + \varepsilon^{k/2}$  for (1.6) is more or less necessary without additional smoothness assumptions on  $A$ .

The next theorem gives the large-scale boundary Lipschitz estimate, which is new even in the case  $k = 2$ . Let  $\Omega$  be a bounded  $C^{1,\alpha}$  domain in  $\mathbb{R}^d$  for some  $\alpha \in (0, 1)$ . Define  $D_r(x_0, t_0) = (B(x_0, r) \cap \Omega) \times (t_0 - r^2, t_0)$  and  $\Delta_r(x_0, t_0) = (B(x_0, r) \cap \partial\Omega) \times (t_0 - r^2, t_0)$ , where  $x_0 \in \partial\Omega$  and  $t_0 \in \mathbb{R}$ .

**Theorem 1.2.** *Assume  $A = A(y, s)$  satisfies (1.3) and (1.4). Suppose that  $(\partial_t + \mathcal{L}_\varepsilon)u_\varepsilon = F$  in  $D_R = D_R(x_0, t_0)$  and  $u_\varepsilon = f$  on  $\Delta_R = \Delta_R(x_0, t_0)$ , where  $\varepsilon + \varepsilon^{k/2} < R \leq 1$ ,  $F \in L^p(D_R)$  for some  $p > d + 2$ , and  $f \in C^{1+\alpha}(\Delta_R)$  for some  $\alpha \in (0, 1)$ . Then for any  $\varepsilon + \varepsilon^{k/2} \leq r < R$ ,*

$$\begin{aligned} \left( \int_{D_r} |\nabla u_\varepsilon|^2 \right)^{1/2} &\leq C \left\{ \left( \int_{D_R} |\nabla u_\varepsilon|^2 \right)^{1/2} \right. \\ &\quad \left. + R^{-1} \|f\|_{C^{1+\alpha}(\Delta_R)} + R \left( \int_{D_R} |F|^p \right)^{1/p} \right\}, \end{aligned} \quad (1.7)$$

where  $C$  depends only on  $d, m, \alpha, p, \mu$ , and  $\Omega$ .

Under the additional Hölder continuity condition on  $A$ , the large-scale estimates in Theorems 1.1 and 1.2 imply the uniform interior and boundary Lipschitz estimates for  $|\nabla u_\varepsilon(x_0, t_0)|$ . In the case  $k = 2$ , this follows readily from a simple blow-up argument by considering  $u_\varepsilon(\varepsilon x, \varepsilon^2 t)$  and using the classical Lipschitz estimates for parabolic operators with Hölder continuous coefficients. If  $k \neq 2$ , we may consider the function  $u_\varepsilon(\delta x, \delta^2 t)$  with either  $\delta = \varepsilon$  or  $\varepsilon^{k/2}$ . It leads to the problem of uniform Lipschitz estimates for parabolic operators of forms,  $\partial_t - \operatorname{div}(A(x/\varepsilon, t)\nabla)$  and  $\partial_t - \operatorname{div}(A(x, t/\varepsilon)\nabla)$ , with locally periodic coefficients. The details will appear elsewhere.

In this paper we also investigate the rate of convergence in  $L^2(\Omega_T)$  for the initial-Dirichlet problem

$$(\partial_t + \mathcal{L}_\varepsilon)u_\varepsilon = F \quad \text{in } \Omega_T \quad \text{and} \quad u_\varepsilon = f \quad \text{on } \partial_p \Omega_T, \quad (1.8)$$

where  $\partial_p \Omega_T$  denotes the parabolic boundary of  $\Omega_T$ .

**Theorem 1.3.** *Let  $\Omega$  be a bounded  $C^{1,1}$  domain in  $\mathbb{R}^d$  and  $0 < T < \infty$ . Assume  $A = A(y, s)$  satisfies (1.3) and (1.4). Also assume that  $\|\partial_s A\|_\infty < \infty$  for  $0 < k < 2$  and  $\|\nabla^2 A\|_\infty < \infty$  for  $k > 2$ . Let  $u_\varepsilon$  be a weak solution of (1.8) and  $u_0$  its homogenized solution (with the same data  $F$  and  $f$ ). Suppose that  $u_0 \in L^2(0, T; H^2(\Omega))$  and  $\partial_t u_0 \in L^2(\Omega_T)$ . Then*

$$\|u_\varepsilon - u_0\|_{L^2(\Omega_T)} \leq C \left\{ \|u_0\|_{L^2(0,T;H^2(\Omega))} + \|\partial_t u_0\|_{L^2(\Omega_T)} \right\} \cdot \begin{cases} \varepsilon^{k/2} & \text{if } 0 < k \leq 4/3, \\ \varepsilon^{2-k} & \text{if } 4/3 < k < 2, \\ \varepsilon^{k-2} & \text{if } 2 < k < 3, \\ \varepsilon & \text{if } k = 2 \text{ or } 3 \leq k < \infty, \end{cases} \quad (1.9)$$

for any  $0 < \varepsilon < 1$ , where  $C$  depends only on  $d, m, k, A, \Omega$ , and  $T$ .

The convergence rates  $\varepsilon^\gamma$  for different  $k$ 's in (1.9) are obtained as a result of the formula  $\gamma = \min(k/2, 2 - k)$  for  $0 < k < 2$ ;  $\gamma = 1$  for  $k = 2$ ; and  $\gamma = \min(1, k - 2)$  for  $k > 2$ . In Theorem 1.3 we do not specify conditions on  $F$  and  $f$ , but rather require that  $u_0 \in L^2(0, T; H^2(\Omega))$  and  $\partial_t u_0 \in L^2(\Omega_T)$ . Notice that if  $\Omega$  is  $C^{1,1}$ ,  $F \in L^2(\Omega_T)$  and  $f = 0$ , then  $\|\nabla^2 u_0\|_{L^2(\Omega_T)} + \|\partial_t u_0\|_{L^2(\Omega_T)} \leq C\|F\|_{L^2(\Omega_T)}$ . It follows that

$$\|u_\varepsilon - u_0\|_{L^2(\Omega_T)} \leq C\varepsilon^\gamma \|F\|_{L^2(\Omega_T)}, \quad (1.10)$$

where  $\gamma \in (0, 1]$  is given above.

*Remark 1.4.* Whether the convergence rate in (1.10) is sharp for  $0 < k < 2$  and  $2 < k < 3$  remains open. We point out that even though the homogenized equation does not depend on  $k$  for  $0 < k < 2$  and for  $2 < k < \infty$ , the sharp convergence rate for (1.10) may depend on  $k$ . This is already clear in the trivial case  $A^\varepsilon = A(t/\varepsilon^k)$  for  $k$  close to 0. We also note that the convergence rate given by Theorem 1.3 is not continuous in  $k$  at  $k = 2$ , and that the non-self-similar case  $k \neq 2$  requires additional smoothness conditions. These seem to be consistent with the known results in reiterated homogenization. In particular, in the elliptic case with coefficient tensor  $A^\varepsilon = A(x/\varepsilon, x/\varepsilon^k)$  for  $0 < k < 1$ , the convergence rate obtained in [25] for  $\Omega = \mathbb{R}^d$  is  $\varepsilon^\gamma$  with  $\gamma = \min(k, 1 - k)$  under the assumption that  $A(y, z)$  is Lipschitz continuous in  $z$ .

We now describe our general approach to Theorems 1.1–1.3. The key insight is to introduce a new scale  $\lambda \in (0, \infty)$  and consider the operator

$$\mathcal{L}_{\varepsilon,\lambda} = -\operatorname{div}\left(A_\lambda(x/\varepsilon, t/\varepsilon^2)\nabla\right), \quad (1.11)$$

where  $A_\lambda(y, s) = A(y, s/\lambda)$ . Observe that the coefficient tensor  $A_\lambda$  is 1-periodic in  $y$  and  $\lambda$ -periodic in  $s$ . Moreover, for each  $\lambda$  fixed, the scaling of the parameter  $\varepsilon$  in  $A_\lambda(x/\varepsilon, t/\varepsilon^2)$  is consistent with the intrinsic scaling of the second-order parabolic operator  $\partial_t + \mathcal{L}_{\varepsilon,\lambda}$ . As a result, we may extend some of recently developed techniques

for elliptic equations to the parabolic equation  $(\partial_t + \mathcal{L}_{\varepsilon,\lambda})u_{\varepsilon,\lambda} = F$ , as in the case  $k = 2$ . We point out that for the results to be useful, it is crucial that the bounding constants  $C$  in the estimates of solutions  $u_{\varepsilon,\lambda}$  do not depend on  $\lambda$  (and  $\varepsilon$ ). This allows us to use the observation  $\mathcal{L}_\varepsilon = \mathcal{L}_{\varepsilon,\lambda}$  for  $\lambda = \varepsilon^{k-2}$  and prove Theorems 1.1 and 1.2. The approach also leads to large-scale  $C^{1,\alpha}$  and  $C^{2,\alpha}$  excess-decay estimates as well as a Liouville property, expressed in terms of correctors for  $\partial_t + \mathcal{L}_{\varepsilon,\lambda}$ .

The approach described above works equally well for the problem of convergence rates. In addition to the observation  $\mathcal{L}_{\varepsilon,\lambda} = \mathcal{L}_\varepsilon$  for  $\lambda = \varepsilon^{k-2}$ , we also use the fact that as  $\lambda \rightarrow \infty$ , the homogenized coefficient matrix  $\widehat{A}_\lambda$  for  $\partial_t + \mathcal{L}_{\varepsilon,\lambda}$  converges to  $\widehat{A}_\infty$ , the homogenized coefficient matrix for  $\partial_t + \mathcal{L}_\varepsilon$  in the case  $0 < k < 2$ . If  $\lambda \rightarrow 0$ , then  $\widehat{A}_\lambda \rightarrow \widehat{A}_0$ , the homogenized coefficient matrix for  $\partial_t + \mathcal{L}_\varepsilon$  in the case  $2 < k < \infty$ .

The paper is organized as follows: in Section 2 we introduce the first-order correctors  $\chi^\lambda$  and homogenized coefficients for  $\partial_t + \mathcal{L}_{\varepsilon,\lambda}$ , with  $\lambda > 0$  fixed, as well as correctors and homogenized coefficients for  $\mathcal{L}_\varepsilon$  in (1.1) with  $0 < k < \infty$ . We also establish precise estimates of  $|\widehat{A}_\lambda - \widehat{A}_\infty|$  for  $\lambda > 1$ , and of  $|\widehat{A}_\lambda - \widehat{A}_0|$  for  $0 < \lambda < 1$ , under additional regularity assumptions on  $A$ . These estimates are used in the proof of Theorem 1.3. In Section 3 we prove an approximation result for solutions of  $(\partial_t + \mathcal{L}_{\varepsilon,\lambda})u_{\varepsilon,\lambda} = F$  in a parabolic cylinder. This is done by using  $\varepsilon$ -smoothing and dual correctors. The proof follows the approach used in [14] by the present authors for the case  $\lambda = 1$ . The proof of Theorem 1.1 is given in Section 4, where we also establish a large-scale  $C^{1,\alpha}$  estimate. In Section 5 we introduce second-order correctors for the operator  $\partial_t + \mathcal{L}_{\varepsilon,\lambda}$  and prove a large-scale  $C^{2,\alpha}$  estimate. The large-scale boundary Lipschitz estimate in Theorem 1.2 is proved in Section 6. We remark that the approaches used in Sections 4–6 are motivated by recently developed techniques for studying the large-scale regularity in the homogenization theory for elliptic equations and systems [3,4,6–8,11,16,17]. Finally, we give the proof of Theorem 1.3 in Section 7, where we also obtain error estimates for a two-scale expansion in  $L^2(0, T; H^1(\Omega))$ .

The summation convention is used throughout. We will use  $f_E u$  to denote the  $L^1$  average of  $u$  over the set  $E$ ; that is  $f_E u = \frac{1}{|E|} \int_E u$ . For notational simplicity we will assume that  $m = 1$  in the rest of the paper. However, no particular fact pertaining to the scalar case is ever used. All results and proofs extend readily to the case  $m > 1$ —the case of parabolic systems.

## 2. Correctors and Homogenized Coefficients

Let  $A = A(y, s)$  be a matrix satisfying conditions (1.3) and (1.4). For  $\lambda > 0$ , define

$$A_\lambda = A_\lambda(y, s) = A(y, s/\lambda) \quad \text{for } (y, s) \in \mathbb{R}^{d+1}. \quad (2.1)$$

The matrix  $A_\lambda$  is  $(1, \lambda)$ -periodic in  $(y, s)$ ; that is

$$A_\lambda(y + z, s + \lambda t) = A_\lambda(y, s) \quad \text{for } (z, t) \in \mathbb{Z}^{d+1}.$$

Let  $\chi^\lambda = \chi^\lambda(y, s) = (\chi_1^\lambda(y, s), \dots, \chi_d^\lambda(y, s))$ , where  $\chi_j^\lambda = \chi_j^\lambda(y, s)$  is the weak solution of the parabolic cell problem

$$\begin{cases} \partial_s \chi_j^\lambda - \operatorname{div}(A_\lambda \nabla \chi_j^\lambda) = \operatorname{div}(A_\lambda \nabla y_j) & \text{in } \mathbb{R}^{d+1}, \\ \chi_j^\lambda \text{ is } (1, \lambda)\text{-periodic in } (y, s), \\ \int_0^\lambda \int_{\mathbb{T}^d} \chi_j^\lambda(y, s) \, dy ds = 0, \end{cases} \quad (2.2)$$

where  $\mathbb{T}^d = [0, 1)^d \cong \mathbb{R}^d / \mathbb{Z}^d$ . By the energy estimates,

$$\int_0^\lambda \int_{\mathbb{T}^d} |\nabla \chi_j^\lambda|^2 \, dy ds \leq C, \quad (2.3)$$

where  $C$  depends only on  $d$  and  $\mu$ . Since

$$\partial_s \int_{\mathbb{T}^d} \chi_j^\lambda(y, s) \, dy = 0,$$

we obtain, by the integral condition in (2.2),

$$\int_{\mathbb{T}^d} \chi_j^\lambda(y, s) \, dy = 0. \quad (2.4)$$

This, together with (2.3) and Poincaré's inequality, gives

$$\int_0^\lambda \int_{\mathbb{T}^d} |\chi_j^\lambda|^2 \, dy ds \leq C, \quad (2.5)$$

where  $C$  depends only on  $d$  and  $\mu$ . Since  $\chi^\lambda$  and  $\nabla \chi^\lambda$  are  $(1, \lambda)$ -periodic in  $(y, s)$ , it follows from (2.3) and (2.5) that if  $r \geq 1 + \sqrt{\lambda}$ ,

$$\left( \int_{Q_r} (|\nabla \chi^\lambda|^2 + |\chi^\lambda|^2) \right)^{1/2} \leq C \quad (2.6)$$

for any  $Q_r = Q_r(x, t)$ , where  $C$  depends only on  $d$  and  $\mu$ .

Let

$$\widehat{A}_\lambda = \int_0^\lambda \int_{\mathbb{T}^d} (A_\lambda + A_\lambda \nabla \chi^\lambda) \, dy ds. \quad (2.7)$$

**Lemma 2.1.** *There exists  $C > 0$ , depending only on  $d$  and  $\mu$ , such that  $|\widehat{A}_\lambda| \leq C$ . Moreover,*

$$\mu |\xi|^2 \leq \xi \cdot \widehat{A}_\lambda \xi \quad (2.8)$$

for any  $\xi \in \mathbb{R}^d$ .

*Proof.* The inequality  $|\widehat{A_\lambda}| \leq C$  follows readily from (2.3). To see (2.8), we note that

$$\begin{aligned} \xi \cdot \widehat{A_\lambda} \xi &= \int_0^\lambda \int_{\mathbb{T}^d} A_\lambda \nabla(\xi \cdot y + \xi \cdot \chi^\lambda) \cdot \nabla(\xi \cdot y + \xi \cdot \chi^\lambda) \, dy ds \\ &\geq \mu \int_0^\lambda \int_{\mathbb{T}^d} |\nabla(\xi \cdot y + \xi \cdot \chi^\lambda)|^2 \, dy ds \\ &= \mu \int_0^\lambda \int_{\mathbb{T}^d} (|\xi|^2 + |\xi \nabla \chi^\lambda|^2) \, dy ds \\ &\geq \mu |\xi|^2 \end{aligned}$$

for any  $\xi \in \mathbb{R}^d$ , where we have used the fact  $\int_0^\lambda \int_{\mathbb{T}^d} \nabla \chi^\lambda \, dy ds = 0$ .  $\square$

It is well known that for a fixed  $\lambda > 0$ , the homogenized operator for the parabolic operator

$$\partial_t + \mathcal{L}_{\varepsilon, \lambda} = \partial_t - \operatorname{div}(A_\lambda(x/\varepsilon, t/\varepsilon^2) \nabla) \quad (2.9)$$

is given by  $\partial_t - \operatorname{div}(\widehat{A_\lambda} \nabla)$  [10]. In particular, if  $k = 2$ , the homogenized operator for the operator in (1.1) is given by  $\partial_t - \operatorname{div}(\widehat{A_\lambda} \nabla)$  with  $\lambda = 1$ .

To introduce the homogenized operator for  $\partial_t + \mathcal{L}_\varepsilon$  in (1.1) for  $k \neq 2$ , we first consider the case  $0 < k < 2$ . Let  $\chi^\infty = \chi^\infty(y, s) = (\chi_1^\infty(y, s), \dots, \chi_d^\infty(y, s))$ , where  $\chi_j^\infty = \chi_j^\infty(y, s)$  denotes the weak solution of the (elliptic) cell problem

$$\begin{cases} -\operatorname{div}(A \nabla \chi_j^\infty) = \operatorname{div}(A \nabla y_j) & \text{in } \mathbb{R}^{d+1}, \\ \chi_j^\infty \text{ is 1-periodic in } (y, s), \\ \int_{\mathbb{T}^d} \chi_j^\infty(y, s) \, dy = 0. \end{cases} \quad (2.10)$$

By the energy estimates and Poincaré's inequality,

$$\int_{\mathbb{T}^d} (|\nabla \chi_j^\infty(y, s)|^2 + |\chi_j^\infty(y, s)|^2) \, dy \leq C, \quad (2.11)$$

for almost everywhere  $s \in \mathbb{R}$ , where  $C$  depends only on  $d$  and  $\mu$ . Let

$$\widehat{A_\infty} = \int_0^1 \int_{\mathbb{T}^d} (A + A \nabla \chi^\infty) \, dy ds. \quad (2.12)$$

It follows from (2.11) that  $|\widehat{A_\infty}| \leq C$ , where  $C$  depends only on  $d$  and  $\mu$ . By the same argument as in the proof of Lemma 2.1, one may also show that

$$\mu |\xi|^2 \leq \xi \cdot \widehat{A_\infty} \xi \quad (2.13)$$

for any  $\xi \in \mathbb{R}^d$ . For  $0 < k < 2$ , the homogenized operator for the parabolic operator in (1.1) is given by  $\partial_t - \operatorname{div}(\widehat{A_\infty} \nabla)$  (see [10]).

Next, we consider the case  $2 < k < \infty$ . Define

$$\bar{A} = \bar{A}(y) = \int_0^1 A(y, s) \, ds. \tag{2.14}$$

Let  $\chi^0 = \chi^0(y) = (\chi_1^0(y), \dots, \chi_d^0(y))$ , where  $\chi_j^0 = \chi_j^0(y)$  is the weak solution of the (elliptic) cell problem

$$\begin{cases} -\operatorname{div}(\bar{A}\nabla\chi_j^0) = \operatorname{div}(\bar{A}\nabla y_j) & \text{in } \mathbb{R}^d, \\ \chi_j^0 \text{ is 1-periodic in } y, \\ \int_{\mathbb{T}^d} \chi_j^0 \, dy = 0. \end{cases} \tag{2.15}$$

As in the case  $0 < k < 2$ , by the energy estimates and Poincaré’s inequality,

$$\int_{\mathbb{T}^d} (|\nabla\chi_j^0(y)|^2 + |\chi_j^0(y)|^2) \, dy \leq C, \tag{2.16}$$

where  $C$  depends only on  $d$  and  $\mu$ . Let

$$\widehat{A}_0 = \int_0^1 \int_{\mathbb{T}^d} (A + A\nabla\chi^0) \, dy \, ds = \int_{\mathbb{T}^d} (\bar{A} + \bar{A}\nabla\chi^0) \, dy. \tag{2.17}$$

It follows from (2.16) that  $|\widehat{A}_0| \leq C$ , where  $C$  depends only on  $d$  and  $\mu$ . By the same argument as in the proof of Lemma 2.1, we obtain

$$\mu|\xi|^2 \leq \xi \cdot \widehat{A}_0\xi \tag{2.18}$$

for any  $\xi \in \mathbb{R}^d$ . For  $2 < k < \infty$ , the homogenized operator for  $\partial_t + \mathcal{L}_\varepsilon$  in (1.1) is given by  $\partial_t - \operatorname{div}(\widehat{A}_0\nabla)$  (see [10]).

In the remaining of this section we study the asymptotic behavior of the matrix  $\widehat{A}_\lambda$ , as  $\lambda \rightarrow \infty$  and as  $\lambda \rightarrow 0$ . We begin with a lemma on the higher integrability of  $\nabla\chi^\lambda$ .

**Lemma 2.2.** *Let  $\chi^\lambda$  be defined by (2.2). Then there exists  $q > 2$ , depending on  $d$  and  $\mu$ , such that*

$$\left( \int_0^1 \int_{\mathbb{T}^d} |\nabla\chi^\lambda|^q \, dy \, ds \right)^{1/q} \leq C, \tag{2.19}$$

where  $C$  depends only on  $d$  and  $\mu$ .

*Proof.* Let  $u(y, s) = y_j + \chi_j^\lambda$ . Then  $\partial_s u - \operatorname{div}(A_\lambda \nabla u) = 0$  in  $\mathbb{R}^{d+1}$ . By Meyers-type estimates for parabolic systems (see for example [2, Appendix]), there exist  $q > 2$  and  $C > 0$ , depending only on  $d$  and  $\mu$ , such that

$$\left( \int_{Q_r} |\nabla u|^q \, dy \, ds \right)^{1/q} \leq C \left( \int_{Q_{2r}} |\nabla u|^2 \, dy \, ds \right)^{1/2} \tag{2.20}$$



for any  $Q_r = Q_r(x, t) = B(x, r) \times (t - r^2, t)$ . It follows that

$$\left( \int_{Q_r} |\nabla \chi_j^\lambda|^q dy ds \right)^{1/q} \leq C + C \left( \int_{Q_{2r}} |\nabla \chi_j^\lambda|^2 dy ds \right)^{1/2}. \quad (2.21)$$

Choose  $r > 1 + \sqrt{\lambda}$  so large that  $\mathbb{T}^d \times (0, \lambda) \subset Q_r$ . Since  $\nabla \chi_j^\lambda$  is 1-periodic in  $y$  and  $\lambda$ -periodic in  $s$ , we obtain

$$\begin{aligned} \left( \int_0^\lambda \int_{\mathbb{T}^d} |\nabla \chi_j^\lambda|^q dy ds \right)^{1/q} &\leq C \left( \int_{Q_r} |\nabla \chi_j^\lambda|^q dy ds \right)^{1/q} \\ &\leq C + C \left( \int_{Q_{2r}} |\nabla \chi_j^\lambda|^2 dy ds \right)^{1/2} \\ &\leq C, \end{aligned}$$

where we have used (2.6) for the last step.  $\square$

**Theorem 2.3.** Assume  $A = A(y, s)$  satisfies conditions (1.3) and (1.4). Then

$$\widehat{A}_\lambda \rightarrow \widehat{A}_\infty \quad \text{as } \lambda \rightarrow \infty. \quad (2.22)$$

Moreover, if  $\|\partial_s A\|_\infty < \infty$ , then

$$|\widehat{A}_\lambda - \widehat{A}_\infty| \leq C \lambda^{-1} \|\partial_s A\|_\infty \quad (2.23)$$

for any  $\lambda > 1$ , where  $C$  depends only on  $d$  and  $\mu$ .

*Proof.* We first prove (2.23). Observe that

$$\widehat{A}_\lambda - \widehat{A}_\infty = \int_0^1 \int_{\mathbb{T}^d} A(y, s) \nabla \{ \chi^\lambda(y, \lambda s) - \chi^\infty(y, s) \} dy ds.$$

It follows by the Cauchy inequality that

$$|\widehat{A}_\lambda - \widehat{A}_\infty| \leq C \left( \int_0^1 \int_{\mathbb{T}^d} |\nabla \{ \chi^\lambda(y, \lambda s) - \chi^\infty(y, s) \}|^2 dy ds \right)^{1/2}. \quad (2.24)$$

By the definitions of  $\chi^\lambda$  and  $\chi^\infty$ ,

$$\frac{1}{\lambda} \frac{\partial}{\partial s} \{ \chi_j^\lambda(y, \lambda s) \} - \operatorname{div} \{ A(y, s) \nabla (\chi_j^\lambda(y, \lambda s) - \chi_j^\infty(y, s)) \} = 0 \quad \text{in } \mathbb{T}^{d+1}.$$

This leads to

$$\begin{aligned} &\int_0^1 \int_{\mathbb{T}^d} A(y, s) \nabla \{ \chi_j^\lambda(y, \lambda s) - \chi_j^\infty(y, s) \} \cdot \nabla \{ \chi_j^\lambda(y, \lambda s) - \chi_j^\infty(y, s) \} dy ds \\ &= -\frac{1}{\lambda} \int_0^1 \int_{\mathbb{T}^d} \frac{\partial}{\partial s} \{ \chi_j^\lambda(y, \lambda s) \} \cdot \{ \chi_j^\lambda(y, \lambda s) - \chi_j^\infty(y, s) \} dy ds \\ &= -\frac{1}{\lambda} \int_0^1 \int_{\mathbb{T}^d} \frac{\partial}{\partial s} \{ \chi_j^\infty(y, s) \} \cdot \{ \chi_j^\lambda(y, \lambda s) - \chi_j^\infty(y, s) \} dy ds, \end{aligned}$$

where we have used the fact

$$\int_0^1 \int_{\mathbb{T}^d} \frac{\partial}{\partial s} \{ \chi_j^\lambda(y, \lambda s) - \chi_j^\infty(y, s) \} \cdot \{ \chi_j^\lambda(y, \lambda s) - \chi_j^\infty(y, s) \} \, dy ds = 0$$

for the last step. Hence, by (1.3) and the Cauchy inequality,

$$\begin{aligned} & \mu \int_0^1 \int_{\mathbb{T}^d} |\nabla \{ \chi_j^\lambda(y, \lambda s) - \chi_j^\infty(y, s) \}|^2 \, dy ds \\ & \leq \frac{1}{\lambda} \left( \int_0^1 \int_{\mathbb{T}^d} |\chi_j^\lambda(y, \lambda s) - \chi_j^\infty(y, s)|^2 \, dy ds \right)^{1/2} \left( \int_0^1 \int_{\mathbb{T}^d} |\partial_s \chi_j^\infty(y, s)|^2 \, dy ds \right)^{1/2}. \end{aligned}$$

Since

$$\int_{\mathbb{T}^d} \chi_j^\lambda(y, \lambda s) \, dy = \int_{\mathbb{T}^d} \chi_j^\infty(y, s) \, dy = 0,$$

by Poincaré’s inequality, we obtain

$$\begin{aligned} & \left( \int_0^1 \int_{\mathbb{T}^d} |\nabla \{ \chi_j^\lambda(y, \lambda s) - \chi_j^\infty(y, s) \}|^2 \, dy ds \right)^{1/2} \\ & \leq \frac{C}{\lambda} \left( \int_0^1 \int_{\mathbb{T}^d} |\partial_s \chi_j^\infty(y, s)|^2 \, dy ds \right)^{1/2}. \end{aligned}$$

In view of (2.24) we have proved that

$$|\widehat{A}_\lambda - \widehat{A}_\infty| \leq \frac{C}{\lambda} \left( \int_0^1 \int_{\mathbb{T}^d} |\partial_s \chi^\infty(y, s)|^2 \, dy ds \right)^{1/2}, \tag{2.25}$$

where  $C$  depends only on  $d$  and  $\mu$ .

To bound the right-hand side of (2.25), we differentiate in  $s$  the elliptic equation for  $\chi_j^\infty$  to obtain

$$-\operatorname{div}(A \nabla \partial_s \chi_j^\infty) = \operatorname{div}(\partial_s A \nabla y_j) + \operatorname{div}(\partial_s A \nabla \chi_j^\infty).$$

It follows that

$$\int_{\mathbb{T}^d} |\nabla \partial_s \chi_j^\infty(y, s)|^2 \, dy \leq C \int_{\mathbb{T}^d} |\partial_s A(y, s)|^2 \, dy + C \int_{\mathbb{T}^d} |\partial_s A(y, s)|^2 |\nabla \chi_j^\infty(y, s)|^2 \, dy.$$

By Meyer’s estimates, there exists some  $q > 2$ , depending only on  $d$  and  $\mu$ , such that

$$\int_{\mathbb{T}^d} |\nabla \chi_j^\infty(y, s)|^q \, dy \leq C,$$

where  $C$  depends only on  $d$  and  $\mu$ . Thus, by Hölder’s inequality,

$$\left( \int_0^1 \int_{\mathbb{T}^d} |\nabla \partial_s \chi_j^\infty|^2 \, dy ds \right)^{1/2} \leq C \left( \int_0^1 \int_{\mathbb{T}^d} |\partial_s A|^{p_0} \, dy ds \right)^{1/p_0},$$

for  $p_0 = \frac{2q}{q-2}$ . In view of (2.25) this gives

$$|\widehat{A}_\lambda - \widehat{A}_\infty| \leq \frac{C}{\lambda} \left( \int_0^1 \int_{\mathbb{T}^d} |\partial_s A|^{p_0} dy ds \right)^{1/p_0}, \tag{2.26}$$

by using Poincaré’s inequality. As a consequence, we obtain (2.23).

Finally, to prove (2.22), we let  $D$  be a matrix satisfying conditions (1.3) and (1.4). Also assume that  $D$  is smooth in  $(y, s)$ . Let  $\widehat{D}_\lambda$  and  $\widehat{D}_\infty$  be defined in the same manner as  $\widehat{A}_\lambda$  and  $\widehat{A}_\infty$ , respectively. By using the energy estimates as well as (2.19), it is not hard to show that

$$|\widehat{A}_\lambda - \widehat{D}_\lambda| \leq C \left( \int_0^1 \int_{\mathbb{T}^d} |A - D|^{p_0} dy ds \right)^{1/p_0},$$

where  $C$  depends only on  $d$  and  $\mu$ . A similar argument also gives

$$|\widehat{A}_\infty - \widehat{D}_\infty| \leq C \left( \int_0^1 \int_{\mathbb{T}^d} |A - D|^{p_0} dy ds \right)^{1/p_0}.$$

Thus, by applying the estimate (2.26) to the matrix  $D$ , we obtain

$$\begin{aligned} |\widehat{A}_\lambda - \widehat{A}_\infty| &\leq |\widehat{A}_\lambda - \widehat{D}_\lambda| + |\widehat{D}_\lambda - \widehat{D}_\infty| + |\widehat{D}_\infty - \widehat{A}_\infty| \\ &\leq C \left( \int_0^1 \int_{\mathbb{T}^d} |A - D|^{p_0} dy ds \right)^{1/p_0} + \frac{C}{\lambda} \left( \int_0^1 \int_{\mathbb{T}^d} |\partial_s D|^{p_0} dy ds \right)^{1/p_0}. \end{aligned}$$

It follows that

$$\limsup_{\lambda \rightarrow \infty} |\widehat{A}_\lambda - \widehat{A}_\infty| \leq C \left( \int_0^1 \int_{\mathbb{T}^d} |A - D|^{p_0} dy ds \right)^{1/p_0}.$$

Since  $p_0 = \frac{2q}{q-2} < \infty$ , by using convolution, we may approximate  $A$  in  $L^{p_0}(\mathbb{T}^{d+1})$  by a sequence of smooth matrices satisfying (1.3) and (1.4). As a result, we conclude that  $\widehat{A}_\lambda \rightarrow \widehat{A}_\infty$  as  $\lambda \rightarrow \infty$ .  $\square$

*Remark 2.4.* It follows from the proof of Theorem 2.3 that

$$\begin{aligned} &\left( \int_0^\lambda \int_{\mathbb{T}^d} |\nabla \chi^\lambda(y, s) - \nabla \chi^\infty(y, s/\lambda)|^2 dy ds \right)^{1/2} \\ &\quad + \left( \int_0^\lambda \int_{\mathbb{T}^d} |\chi^\lambda(y, s) - \chi^\infty(y, s/\lambda)|^2 dy ds \right)^{1/2} \\ &\leq C \lambda^{-1} \|\partial_s A\|_\infty. \end{aligned}$$

By the periodicity this implies that if  $r \geq (1 + \sqrt{\lambda})\varepsilon$ , then

$$\begin{aligned} &\left( \int_{Q_r} |\nabla \chi^\lambda(y/\varepsilon, s/\varepsilon^2) - \nabla \chi^\infty(y/\varepsilon, s/(\lambda\varepsilon^2))|^2 dy ds \right)^{1/2} \\ &\quad + \left( \int_{Q_r} |\chi^\lambda(y/\varepsilon, s/\varepsilon^2) - \chi^\infty(y/\varepsilon, s/(\lambda\varepsilon^2))|^2 dy ds \right)^{1/2} \tag{2.27} \\ &\leq C \lambda^{-1} \|\partial_s A\|_\infty. \end{aligned}$$

The next theorem is concerned with the limit of  $\widehat{A}_\lambda$  as  $\lambda \rightarrow 0$ .

**Theorem 2.5.** *Assume  $A = A(y, s)$  satisfies conditions (1.3) and (1.4). Then*

$$\widehat{A}_\lambda \rightarrow \widehat{A}_0 \quad \text{as } \lambda \rightarrow 0. \tag{2.28}$$

Moreover, if  $\|\nabla^2 A\|_\infty < \infty$ , then

$$|\widehat{A}_\lambda - \widehat{A}_0| \leq C\lambda \{ \|\nabla^2 A\|_\infty + \|\nabla A\|_\infty^2 \}, \tag{2.29}$$

where  $C$  depends only on  $d$  and  $\mu$ .

*Proof.* We first prove (2.29). Observe that

$$\begin{aligned} \widehat{A}_\lambda - \widehat{A}_0 &= \int_0^1 \int_{\mathbb{T}^d} A(y, s) \nabla (\chi^\lambda(y, \lambda s) - \chi^0(y)) \, dy ds \\ &= \int_0^1 \int_{\mathbb{T}^d} (A(y, s) - \bar{A}(y)) \nabla \chi^\lambda(y, \lambda s) \, dy ds \\ &\quad + \int_{\mathbb{T}^d} \bar{A}(y) \nabla \left( \int_0^1 \chi^\lambda(y, \lambda s) \, ds - \chi^0(y) \right) \, dy \\ &= I_1 + I_2. \end{aligned} \tag{2.30}$$

Write  $A(y, s) - \bar{A}(y) = \partial_s \tilde{A}(y, s)$ , where

$$\tilde{A}(y, s) = \int_0^s (A(y, \tau) - \bar{A}(y)) \, d\tau.$$

Since  $\tilde{A}(y, s)$  is 1-periodic in  $(y, s)$ , we may use an integration by parts and the Cauchy inequality to obtain

$$|I_1| \leq C\lambda \left( \int_0^\lambda \int_{\mathbb{T}^d} |\partial_s \nabla \chi^\lambda|^2 \, dy ds \right)^{1/2}. \tag{2.31}$$

To bound the term  $I_2$  in (2.30), we observe that

$$-\operatorname{div} \left( \int_0^1 A(y, s) \nabla \chi_j^\lambda(y, \lambda s) \, ds \right) = \operatorname{div} (\bar{A}(y) \nabla y_j) = -\operatorname{div} (\bar{A}(y) \nabla \chi_j^0(y)).$$

It follows that

$$\begin{aligned} &-\operatorname{div} \left\{ \bar{A}(y) \nabla \left( \int_0^1 \chi_j^\lambda(y, \lambda s) \, ds - \chi_j^0(y) \right) \right\} \\ &= \operatorname{div} \left\{ \int_0^1 (A(y, s) - \bar{A}(y)) \nabla \chi_j^\lambda(y, \lambda s) \, ds \right\}. \end{aligned}$$

By the energy estimates we obtain

$$\begin{aligned} & \|\nabla \left( \int_0^1 \chi_j^\lambda(y, \lambda s) \, ds - \chi_j^0(y) \right)\|_{L^2(\mathbb{T}^d)} \\ & \leq C \left\| \int_0^1 (A(y, s) - \bar{A}(y)) \nabla \chi_j^\lambda(y, \lambda s) \, ds \right\|_{L^2(\mathbb{T}^d)} \\ & \leq C \lambda \left( \int_0^1 \int_{\mathbb{T}^d} |\partial_s \nabla \chi_j^\lambda|^2 \, dy ds \right)^{1/2}, \end{aligned}$$

where, for the last step, we have used the integration by parts as in the estimate of  $I_1$ . As a result, in view of (2.30) and (2.31), we have proved that

$$|\widehat{A}_\lambda - \widehat{A}_0| \leq C \lambda \left( \int_0^1 \int_{\mathbb{T}^d} |\partial_s \nabla \chi_j^\lambda|^2 \, dy ds \right)^{1/2}. \quad (2.32)$$

To bound the right-hand side of (2.32), we differentiate in  $y$  the parabolic equation for  $\chi_j^\lambda$  to obtain

$$\partial_s \nabla \chi_j^\lambda - \operatorname{div}(A_\lambda \nabla(\nabla \chi_j^\lambda)) = \operatorname{div}(\nabla A_\lambda \cdot \nabla \chi_j^\lambda) + \operatorname{div}(\nabla A_\lambda \cdot \nabla y_j). \quad (2.33)$$

By the energy estimates,

$$\int_0^1 \int_{\mathbb{T}^d} |\nabla^2 \chi_j^\lambda|^2 \, dy ds \leq C \|\nabla A\|_\infty^2. \quad (2.34)$$

By differentiating (2.33) in  $y$  we have

$$\begin{aligned} & \partial_s \nabla^2 \chi_j^\lambda - \operatorname{div}(A_\lambda \nabla(\nabla^2 \chi_j^\lambda)) \\ & = \operatorname{div}(\nabla A_\lambda \cdot \nabla^2 \chi_j^\lambda) + \operatorname{div}(\nabla^2 A_\lambda \cdot \nabla \chi_j^\lambda) + \operatorname{div}(\nabla A_\lambda \cdot \nabla^2 \chi_j^\lambda) \\ & \quad + \operatorname{div}(\nabla^2 A_\lambda \cdot \nabla y_j). \end{aligned}$$

Again, by the energy estimates,

$$\begin{aligned} \int_0^1 \int_{\mathbb{T}^d} |\nabla^3 \chi_j^\lambda|^2 \, dy ds & \leq C \|\nabla A\|_\infty^2 \int_0^1 \int_{\mathbb{T}^d} |\nabla^2 \chi_j^\lambda|^2 \, dy ds \\ & \quad + C \int_0^1 \int_{\mathbb{T}^d} |\nabla^2 A_\lambda|^2 |\nabla \chi_j^\lambda|^2 \, dy ds \\ & \quad + C \int_0^1 \int_{\mathbb{T}^d} |\nabla^2 A_\lambda|^2 \, dy ds \\ & \leq C \left\{ \|\nabla A\|_\infty^4 + \|\nabla^2 A\|_\infty^2 \right\}. \end{aligned}$$

It follows by the equation (2.33) that

$$\int_0^1 \int_{\mathbb{T}^d} |\partial_s \nabla \chi_j^\lambda|^2 \, dy ds \leq C \left\{ \|\nabla A\|_\infty^4 + \|\nabla^2 A\|_\infty^2 \right\},$$

which, together with (2.32), gives (2.29).

Finally, to see (2.28), we let  $D$  be a smooth matrix satisfying (1.3) and (1.4). As in the proof of Theorem 2.3, we have

$$\begin{aligned} |\widehat{A}_\lambda - \widehat{A}_0| &\leq |\widehat{A}_\lambda - \widehat{D}_\lambda| + |\widehat{D}_\lambda - \widehat{D}_0| + |\widehat{D}_0 - \widehat{A}_0| \\ &\leq C \left( \int_0^1 \int_{\mathbb{T}^d} |A - D|^{p_0} dy ds \right)^{1/p_0} + C\lambda \left\{ \|\nabla^2 D\|_\infty + \|\nabla D\|_\infty^2 \right\}. \end{aligned}$$

By letting  $\lambda \rightarrow 0$  and by approximating  $A$  in the  $L^{p_0}(\mathbb{T}^{d+1})$  norm by a sequence of smooth matrices satisfying (1.3) and (1.4), we conclude that  $\widehat{A}_\lambda \rightarrow \widehat{A}_0$  as  $\lambda \rightarrow 0$ .  $\square$

*Remark 2.6.* It follows from the proof of Theorem 2.5 that if  $r \geq \varepsilon$ ,

$$\begin{aligned} &\left( \int_{Q_r} |\nabla \chi^\lambda(y/\varepsilon, s/\varepsilon^2) - \nabla \chi^0(y/\varepsilon)|^2 dy ds \right)^{1/2} \\ &\quad + \left( \int_{Q_r} |\chi^\lambda(y/\varepsilon, s/\varepsilon^2) - \chi^0(y/\varepsilon)|^2 dy ds \right)^{1/2} \tag{2.35} \\ &\leq C\lambda \left\{ \|\nabla^2 A\|_\infty + \|\nabla A\|_\infty^2 \right\} \end{aligned}$$

for  $0 < \lambda < 1$ , where  $C$  depends only on  $d$  and  $\mu$ .

### 3. Approximation

Let  $A_\lambda$  be the matrix given by (2.1) and  $\mathcal{L}_{\varepsilon,\lambda} = -\operatorname{div}(A_\lambda(x/\varepsilon, t/\varepsilon^2)\nabla)$ . Let  $\mathcal{L}_{0,\lambda} = -\operatorname{div}(\widehat{A}_\lambda \nabla)$ , where the constant matrix  $\widehat{A}_\lambda$  is given by (2.7). The goal of this section is to prove the following theorem:

**Theorem 3.1.** *Suppose  $A$  satisfies conditions (1.3) and (1.4). Let  $u_{\varepsilon,\lambda}$  be a weak solution of*

$$(\partial_t + \mathcal{L}_{\varepsilon,\lambda})u_{\varepsilon,\lambda} = F \quad \text{in } Q_{2r}, \tag{3.1}$$

where  $r > (1 + \sqrt{\lambda})\varepsilon$  and  $F \in L^p(Q_{2r})$  for some  $p > d + 2$ . Then there exists a weak solution of

$$(\partial_t + \mathcal{L}_{0,\lambda})u_{0,\lambda} = F \quad \text{in } Q_r, \tag{3.2}$$

such that

$$\left( \int_{Q_r} |\nabla u_{0,\lambda}|^2 \right)^{1/2} \leq C \left( \int_{Q_{2r}} |\nabla u_{\varepsilon,\lambda}|^2 \right)^{1/2}, \tag{3.3}$$

and

$$\begin{aligned} &\left( \int_{Q_{r/2}} |\nabla u_{\varepsilon,\lambda} - \nabla u_{0,\lambda} - (\nabla \chi^\lambda)^\varepsilon \nabla u_{0,\lambda}|^2 \right)^{1/2} \\ &\leq C \left( \frac{(1 + \sqrt{\lambda})\varepsilon}{r} \right)^\sigma \left\{ \left( \int_{Q_{2r}} |\nabla u_{\varepsilon,\lambda}|^2 \right)^{1/2} + r \left( \int_{Q_{2r}} |F|^p \right)^{1/p} \right\}, \end{aligned} \tag{3.4}$$

where  $\sigma \in (0, 1)$  and  $C > 0$  depend only on  $d$ ,  $\mu$  and  $p$ .

We begin by introducing the dual correctors  $\phi^\lambda$  for the operator  $\partial_t + \mathcal{L}_{\varepsilon,\lambda}$ . Let

$$B_\lambda = A_\lambda + A_\lambda \nabla \chi^\lambda - \widehat{A}_\lambda, \quad (3.5)$$

where the corrector  $\chi^\lambda$  is given by (2.2). Note that  $B_\lambda$  is  $(1, \lambda)$ -periodic in  $(y, s)$ .

**Lemma 3.2.** *Let  $B_\lambda = (b_{ij}^\lambda)$  be given by (3.5). Then there exist  $(1, \lambda)$ -periodic functions  $\phi_{kij}^\lambda$  and  $\phi_{k(d+1)j}^\lambda$ , with  $1 \leq i, j, k \leq d$ , in  $H_{loc}^1(\mathbb{R}^{d+1})$  such that*

$$\begin{cases} b_{ij}^\lambda = \frac{\partial}{\partial y_k} \phi_{kij}^\lambda - \partial_s \phi_{i(d+1)j}^\lambda, \\ -\chi_j^\lambda = \frac{\partial}{\partial y_k} \phi_{k(d+1)j}^\lambda. \end{cases} \quad (3.6)$$

Moreover,  $\phi_{kij}^\lambda = -\phi_{ikj}^\lambda$  and

$$\int_0^\lambda \int_{\mathbb{T}^d} (|\phi_{kij}^\lambda|^2 + |\nabla \phi_{k(d+1)j}^\lambda|^2) \leq C, \quad (3.7)$$

$$\int_0^\lambda \int_{\mathbb{T}^d} |\phi_{k(d+1)j}^\lambda|^2 \leq C(1 + \lambda)^2, \quad (3.8)$$

where  $C$  depends only on  $d$  and  $\mu$ .

*Proof.* The lemma was proved in [14] for the case  $\lambda = 1$ . The case  $\lambda \neq 1$  is similar. However, one needs to be careful with the dependence of the constants  $C$  on the parameter  $\lambda$ .

Let  $\Delta_{d+1}$  denote the Laplacian operator in  $(y, s) \in \mathbb{R}^d \times \mathbb{R}$ . By the definition of  $\widehat{A}_\lambda$ ,

$$\int_0^\lambda \int_{\mathbb{T}^d} B_\lambda(y, s) \, dy ds = 0. \quad (3.9)$$

It follows that there exist  $(1, \lambda)$ -periodic functions  $f_{ij}^\lambda \in H_{loc}^2(\mathbb{R}^{d+1})$  such that  $\Delta_{d+1} f_{ij}^\lambda = b_{ij}^\lambda$  in  $\mathbb{R}^{d+1}$  for  $1 \leq i, j \leq d$ . Similarly, there exist  $(1, \lambda)$ -periodic functions  $f_{(d+1)j}^\lambda \in H_{loc}^2(\mathbb{R}^{d+1})$  such that  $\Delta_{d+1} f_{(d+1)j}^\lambda = -\chi_j^\lambda$  in  $\mathbb{R}^{d+1}$  for  $1 \leq j \leq d$ . By the definition of  $\chi_j^\lambda$ , we have

$$\frac{\partial}{\partial y_i} b_{ij}^\lambda = \partial_s \chi_j^\lambda \quad \text{in } \mathbb{R}^{d+1}, \quad (3.10)$$

which leads to

$$\Delta_{d+1} \left( \frac{\partial f_{ij}^\lambda}{\partial y_i} + \partial_s f_{(d+1)j}^\lambda \right) = 0 \quad \text{in } \mathbb{R}^{d+1}.$$

By the periodicity and Liouville Theorem we may conclude that

$$\frac{\partial f_{ij}^\lambda}{\partial y_i} + \partial_s f_{(d+1)j}^\lambda \quad \text{is constant in } \mathbb{R}^{d+1} \text{ for } 1 \leq j \leq d. \quad (3.11)$$

This allows us to write

$$b_{ij}^\lambda = \frac{\partial}{\partial y_k} \left\{ \frac{\partial f_{ij}^\lambda}{\partial y_k} - \frac{\partial f_{kj}^\lambda}{\partial y_i} \right\} + \partial_s \left\{ \partial_s f_{ij}^\lambda - \frac{\partial f_{(d+1)j}^\lambda}{\partial y_i} \right\},$$

and

$$-\chi_j^\lambda = \frac{\partial}{\partial y_k} \left\{ \frac{\partial f_{(d+1)j}^\lambda}{\partial y_k} - \partial_s f_{kj}^\lambda \right\}.$$

We now define  $\phi_{kij}^\lambda$  and  $\phi_{k(d+1)j}^\lambda$  by

$$\begin{cases} \phi_{kij}^\lambda = \frac{\partial f_{ij}^\lambda}{\partial y_k} - \frac{\partial f_{kj}^\lambda}{\partial y_i}, \\ \phi_{k(d+1)j}^\lambda = \frac{\partial f_{(d+1)j}^\lambda}{\partial y_k} - \partial_s f_{kj}^\lambda \end{cases} \quad (3.12)$$

for  $1 \leq i, j, k \leq d$ . This gives (3.6). It is easy to see that  $\phi_{kij}^\lambda = -\phi_{ikj}^\lambda$ .

Finally, to prove estimates (3.7) and (3.8), we use the Fourier series to write

$$b_{ij}^\lambda(y, s) = \sum_{\substack{n \in \mathbb{Z}^d, m \in \mathbb{Z} \\ (n, m) \neq (0, 0)}} a_{n, m} e^{-2\pi i n \cdot y - 2\pi i m s \lambda^{-1}}.$$

Then

$$f_{ij}^\lambda(y, s) = -\frac{1}{4\pi^2} \sum_{\substack{n \in \mathbb{Z}^d, m \in \mathbb{Z} \\ (n, m) \neq (0, 0)}} \frac{a_{n, m}}{|n|^2 + |m|^2 \lambda^{-2}} e^{-2\pi i n \cdot y - 2\pi i m s \lambda^{-1}}.$$

It follows by Parseval's Theorem that

$$\begin{aligned} & \int_0^\lambda \int_{\mathbb{T}^d} \left( |\nabla f_{ij}^\lambda|^2 + |\nabla^2 f_{ij}^\lambda|^2 + |\partial_s^2 f_{ij}^\lambda|^2 + |\nabla \partial_s f_{ij}^\lambda|^2 \right) \\ & \leq C \sum_{n, m} |a_{n, m}|^2 = C \int_0^\lambda \int_{\mathbb{T}^d} |b_{ij}^\lambda|^2 \leq C, \end{aligned} \quad (3.13)$$

where  $C$  depends only on  $d$  and  $\mu$ . Also note that

$$\int_0^\lambda \int_{\mathbb{T}^d} |\partial_s f_{ij}^\lambda|^2 \leq C \lambda^2, \quad (3.14)$$

where  $C$  depends only on  $d$  and  $\mu$ . Similarly, using the estimate (2.5), we obtain

$$\int_0^\lambda \int_{\mathbb{T}^d} \left( |\nabla f_{(d+1)j}^\lambda|^2 + |\nabla^2 f_{(d+1)j}^\lambda|^2 + |\partial_s^2 f_{(d+1)j}^\lambda|^2 + |\nabla \partial_s f_{(d+1)j}^\lambda|^2 \right) \leq C. \quad (3.15)$$

The desired estimates (3.7) and (3.8) follow readily from (3.12)–(3.15).  $\square$



We fix  $\varphi = \varphi(y, s) = \theta_1(y)\theta_2(s)$ , where  $\theta_1 \in C_0^\infty(B(0, 1))$ ,  $\theta_2 \in C_0^\infty(-1, 0)$ ,  $\theta_1, \theta_2 \geq 0$ , and  $\int_{\mathbb{R}^d} \theta_1(y) dy = \int_{\mathbb{R}} \theta_2(s) ds = 1$ . Define

$$S_\delta(f)(x, t) = \int_{\mathbb{R}^{d+1}} f(x - y, t - s) \varphi_\delta(y, s) dy ds, \quad (3.16)$$

where  $\delta > 0$  and  $\varphi_\delta(y, s) = \delta^{-d-2} \varphi(y/\delta, s/\delta^2)$ .

**Lemma 3.3.** *Let  $g \in L^2_{loc}(\mathbb{R}^{d+1})$  and  $f \in L^2(\mathbb{R}^{d+1})$ . Then*

$$\|g S_\delta(f)\|_{L^2(\mathbb{R}^{d+1})} \leq C \sup_{(y,s) \in \mathbb{R}^{d+1}} \left( \int_{Q_\delta(y,s)} |g|^2 \right)^{1/2} \|f\|_{L^2(\mathbb{R}^{d+1})}, \quad (3.17)$$

$$\|g \nabla S_\delta(f)\|_{L^2(\mathbb{R}^{d+1})} \leq C \delta^{-1} \sup_{(y,s) \in \mathbb{R}^{d+1}} \left( \int_{Q_\delta(y,s)} |g|^2 \right)^{1/2} \|f\|_{L^2(\mathbb{R}^{d+1})}, \quad (3.18)$$

where  $C$  depends only on  $d$ .

*Proof.* By Hölder's inequality,

$$|S_\delta(f)(x, t)|^2 \leq \int_{\mathbb{R}^{d+1}} |f(y, s)|^2 \varphi_\delta(x - y, t - s) dy ds.$$

It follows by Fubini's Theorem that

$$\begin{aligned} & \int_{\mathbb{R}^{d+1}} |g|^2 |S_\delta(f)|^2 dx dt \\ & \leq \int_{\mathbb{R}^{d+1}} |f(y, s)|^2 \left( \int_{\mathbb{R}^{d+1}} |g(x, t)|^2 \varphi_\delta(x - y, t - s) dx dt \right) dy ds \\ & \leq C \sup_{(y,s) \in \mathbb{R}^{d+1}} \left( \int_{Q_\delta(y,s)} |g|^2 \right) \|f\|_{L^2(\mathbb{R}^{d+1})}^2, \end{aligned}$$

where  $C$  depends only on  $d$ . This gives (3.17). The estimate (3.18) follows in a similar manner.  $\square$

**Lemma 3.4.** *Let  $S_\delta$  be defined by (3.16). Then*

$$\begin{aligned} \|g \nabla f - S_\delta(g \nabla f)\|_{L^2(\mathbb{R}^{d+1})} & \leq C \delta \left\{ \|\nabla(g \nabla f)\|_{L^2(\mathbb{R}^{d+1})} + \|g \partial_t f\|_{L^2(\mathbb{R}^{d+1})} \right. \\ & \quad \left. + \delta \|(\partial_t g)(\nabla f)\|_{L^2(\mathbb{R}^{d+1})} \right. \\ & \quad \left. + \delta \|(\nabla g) \partial_t f\|_{L^2(\mathbb{R}^{d+1})} \right\}, \quad (3.19) \end{aligned}$$

where  $C$  depends only on  $d$ .

*Proof.* Write  $S_\delta = S_\delta^1 S_\delta^2$ , where

$$\begin{cases} S_\delta^1(f)(x, t) = \int_{\mathbb{R}^d} f(x-y, t) \delta^{-d} \theta_1(y/\delta) dy, \\ S_\delta^2(f)(x, t) = \int_{\mathbb{R}} f(x, t-s) \delta^{-2} \theta_2(s/\delta^2) ds. \end{cases} \quad (3.20)$$

By using the Plancherel Theorem, it is easy to see that

$$\begin{cases} \|f - S_\delta^1(f)\|_{L^2(\mathbb{R}^{d+1})} \leq C\delta \|\nabla f\|_{L^2(\mathbb{R}^{d+1})}, \\ \|f - S_\delta^2(f)\|_{L^2(\mathbb{R}^{d+1})} \leq C\delta^2 \|\partial_t f\|_{L^2(\mathbb{R}^{d+1})}, \end{cases}$$

where  $C$  depends only on  $d$ . It follows that

$$\begin{aligned} \|g\nabla f - S_\delta(g\nabla f)\|_{L^2(\mathbb{R}^{d+1})} &\leq \|g\nabla f - S_\delta^1(g\nabla f)\|_{L^2(\mathbb{R}^{d+1})} + \|S_\delta^1(g\nabla f) \\ &\quad - S_\delta(g\nabla f)\|_{L^2(\mathbb{R}^{d+1})} \\ &\leq C\delta \|\nabla(g\nabla f)\|_{L^2(\mathbb{R}^{d+1})} + C\delta^2 \|\partial_t S_\delta^1(g\nabla f)\|_{L^2(\mathbb{R}^{d+1})}. \end{aligned}$$

To bound the last term in the inequalities above, we note that

$$\partial_t(g\nabla f) = (\partial_t g)\nabla f + \nabla(g\partial_t f) - (\nabla g)\partial_t f.$$

Using the estimates

$$\|S_\delta^1(h)\|_{L^2(\mathbb{R}^{d+1})} \leq \|h\|_{L^2(\mathbb{R}^{d+1})} \quad \text{and} \quad \|\nabla S_\delta^1(h)\|_{L^2(\mathbb{R}^{d+1})} \leq C\delta^{-1} \|h\|_{L^2(\mathbb{R}^{d+1})},$$

we obtain

$$\begin{aligned} \|\partial_t S_\delta^1(g\nabla f)\|_{L^2(\mathbb{R}^{d+1})} &\leq \|(\partial_t g)\nabla f\|_{L^2(\mathbb{R}^{d+1})} + C\delta^{-1} \|g\partial_t f\|_{L^2(\mathbb{R}^{d+1})} \\ &\quad + \|(\nabla g)\partial_t f\|_{L^2(\mathbb{R}^{d+1})}. \end{aligned}$$

This completes the proof.  $\square$

Let

$$w_\varepsilon = u_{\varepsilon, \lambda} - u_{0, \lambda} - \varepsilon(\chi_j^\lambda)^\varepsilon K_\varepsilon \left( \frac{\partial u_{0, \lambda}}{\partial x_j} \right) + \varepsilon^2 \left( \phi_{i(d+1)j}^\lambda \right)^\varepsilon \frac{\partial}{\partial x_i} K_\varepsilon \left( \frac{\partial u_{0, \lambda}}{\partial x_j} \right), \quad (3.21)$$

where

$$(\chi_j^\lambda)^\varepsilon = \chi_j^\lambda(x/\varepsilon, t/\varepsilon^2), \quad (\phi_{i(d+1)j}^\lambda)^\varepsilon = \phi_{i(d+1)j}^\lambda(x/\varepsilon, t/\varepsilon^2),$$

and  $K_\varepsilon$  is a linear operator to be specified later.

**Lemma 3.5.** *Suppose that*

$$(\partial_t + \mathcal{L}_{\varepsilon,\lambda})u_{\varepsilon,\lambda} = (\partial_t + \mathcal{L}_{0,\lambda})u_{0,\lambda} \quad \text{in } \Omega \times (T_0, T_1).$$

Let  $w_\varepsilon$  be defined by (3.21). Then

$$\begin{aligned} (\partial_t + \mathcal{L}_{\varepsilon,\lambda})w_\varepsilon = & -\operatorname{div}\left(\left(\widehat{A}_\lambda - A_\lambda(x/\varepsilon, t/\varepsilon^2)\right)(\nabla u_{0,\lambda} - K_\varepsilon(\nabla u_{0,\lambda}))\right) \\ & + \varepsilon \operatorname{div}\left(A_\lambda(x/\varepsilon, t/\varepsilon^2)\chi^\lambda(x/\varepsilon, t/\varepsilon^2)\nabla K_\varepsilon(\nabla u_{0,\lambda})\right) \\ & + \varepsilon \frac{\partial}{\partial x_k} \left\{ \phi_{kij}^\lambda(x/\varepsilon, t/\varepsilon^2) \frac{\partial}{\partial x_i} K_\varepsilon \left( \frac{\partial u_{0,\lambda}}{\partial x_j} \right) \right\} \\ & + \varepsilon^2 \frac{\partial}{\partial x_k} \left\{ \phi_{k(d+1)j}^\lambda(x/\varepsilon, t/\varepsilon^2) \partial_t K_\varepsilon \left( \frac{\partial u_{0,\lambda}}{\partial x_j} \right) \right\} \\ & - \varepsilon \frac{\partial}{\partial x_i} \left\{ a_{ij}^\lambda(x/\varepsilon, t/\varepsilon^2) \left( \frac{\partial}{\partial x_j} \phi_{\ell(d+1)k}^\lambda(x/\varepsilon, t/\varepsilon^2) \frac{\partial}{\partial x_\ell} K_\varepsilon \left( \frac{\partial u_{0,\lambda}}{\partial x_k} \right) \right) \right\} \\ & - \varepsilon^2 \frac{\partial}{\partial x_i} \left\{ a_{ij}^\lambda(x/\varepsilon, t/\varepsilon^2) \phi_{\ell(d+1)k}^\lambda(x/\varepsilon, t/\varepsilon^2) \frac{\partial^2}{\partial x_j \partial x_\ell} K_\varepsilon \left( \frac{\partial u_{0,\lambda}}{\partial x_k} \right) \right\}, \end{aligned} \quad (3.22)$$

where  $A_\lambda = (a_{ij}^\lambda)$ .

*Proof.* This is proved by a direct computation. See [14, Theorem 2.2] for the case  $\lambda = 1$ .  $\square$

**Lemma 3.6.** *Let  $Q_r = B(0, r) \times (-r^2, 0)$ . Suppose  $u_{\varepsilon,\lambda}$  is a weak solution of  $(\partial_t + \mathcal{L}_{\varepsilon,\lambda})u_{\varepsilon,\lambda} = F$  in  $Q_2$  for some  $F \in L^2(Q_2)$ . Then there exists a weak solution of  $(\partial_t + \mathcal{L}_{0,\lambda})u_{0,\lambda} = F$  in  $Q_1$  such that*

$$\left( \int_{Q_1} |\nabla u_{0,\lambda}|^2 \right)^{1/2} \leq C \left( \int_{Q_2} |\nabla u_{\varepsilon,\lambda}|^2 \right)^{1/2}, \quad (3.23)$$

and for  $\delta = (1 + \sqrt{\lambda})\varepsilon$ ,

$$\begin{aligned} & \left( \int_{Q_1} |\nabla(u_{\varepsilon,\lambda} - u_{0,\lambda} - \varepsilon\chi^\lambda(x/\varepsilon, t/\varepsilon^2)K_\varepsilon(\nabla u_{0,\lambda}))|^2 dxdt \right)^{1/2} \\ & \leq C\delta^\sigma \left\{ \left( \int_{Q_2} |\nabla u_{\varepsilon,\lambda}|^2 \right)^{1/2} + \left( \int_{Q_2} |F|^2 \right)^{1/2} \right\}, \end{aligned} \quad (3.24)$$

where  $\sigma \in (0, 1)$  and  $C > 0$  depend only on  $d$  and  $\mu$ . The operator  $K_\varepsilon$  is defined by (3.27).

*Proof.* We start out by defining  $u_{0,\lambda}$  to be the weak solution of the initial-Dirichlet problem

$$\begin{cases} (\partial_t + \mathcal{L}_{0,\lambda})u_{0,\lambda} = F & \text{in } Q_1, \\ u_{0,\lambda} = u_{\varepsilon,\lambda} & \text{on } \partial_p Q_1, \end{cases} \quad (3.25)$$

where  $\partial_p Q_1$  denotes the parabolic boundary of the cylinder  $Q_1$ . Note that

$$(\partial_t + \mathcal{L}_{0,\lambda})(u_{0,\lambda} - u_{\varepsilon,\lambda}) = (\mathcal{L}_{\varepsilon,\lambda} - \mathcal{L}_{0,\lambda})u_{\varepsilon,\lambda}$$

in  $Q_1$  and  $u_{\varepsilon,\lambda} - u_{0,\lambda} = 0$  on  $\partial_p Q_1$ . It follows from the standard regularity estimates for parabolic operators with constant coefficients that

$$\int_{Q_1} |\nabla(u_{\varepsilon,\lambda} - u_{0,\lambda})|^q \leq C \int_{Q_1} |\nabla u_{\varepsilon,\lambda}|^q$$

for any  $2 \leq q < \infty$ , where  $C$  depends only on  $d, \mu$  and  $q$ . This gives

$$\int_{Q_1} |\nabla u_{0,\lambda}|^q \leq C \int_{Q_1} |\nabla u_{\varepsilon,\lambda}|^q$$

for any  $2 < q < \infty$ . By the Meyers-type estimates for parabolic systems [2, Appendix], there exist some  $q > 2$  and  $C > 0$ , depending on  $d$  and  $\mu$ , such that

$$\left( \int_{Q_1} |\nabla u_{\varepsilon,\lambda}|^q \right)^{1/q} \leq C \left\{ \left( \int_{Q_2} |\nabla u_{\varepsilon,\lambda}|^2 \right)^{1/2} + \left( \int_{Q_2} |F|^2 \right)^{1/2} \right\}.$$

As a result, we obtain

$$\left( \int_{Q_1} |\nabla u_{0,\lambda}|^q \right)^{1/q} \leq C \left\{ \left( \int_{Q_2} |\nabla u_{\varepsilon,\lambda}|^2 \right)^{1/2} + \left( \int_{Q_2} |F|^2 \right)^{1/2} \right\} \quad (3.26)$$

for some  $q > 2$  and  $C > 0$ , depending only on  $d$  and  $\mu$ .

To prove (3.24), we let  $\delta = (1 + \sqrt{\lambda})\varepsilon$ . We may assume  $\delta \leq 1/8$ ; for otherwise the estimate is trivial. Choose  $\eta_\delta \in C_0^\infty(\mathbb{R}^{d+1})$  such that  $0 \leq \eta_\delta \leq 1$ ,  $|\nabla \eta_\delta| \leq C/\delta$ ,  $|\partial_t \eta_\delta| + |\nabla^2 \eta_\delta| \leq C/\delta^2$ ,

$$\eta_\delta = 1 \quad \text{in } Q_{1-3\delta} \quad \text{and} \quad \eta_\delta = 0 \quad \text{in } Q_1 \setminus Q_{1-2\delta}.$$

Let  $w_\varepsilon$  be defined by (3.21), where the operator  $K_\varepsilon$  is given by

$$K_\varepsilon(f) = S_\delta(\eta_\delta f), \quad (3.27)$$

with  $S_\delta$  defined in (3.16). Note that  $w_\varepsilon = 0$  in  $\partial_p Q_1$ . It follows from Lemma 3.5 and energy estimates that

$$\begin{aligned} \int_{Q_1} |\nabla w_\varepsilon|^2 &\leq C \int_{Q_1} |\nabla u_{0,\lambda} - K_\varepsilon(\nabla u_{0,\lambda})|^2 + C\varepsilon^2 \int_{Q_1} |(\chi^\lambda)^\varepsilon \nabla K_\varepsilon(\nabla u_{0,\lambda})|^2 \\ &\quad + C\varepsilon^2 \int_{Q_1} \sum_{k,i,j} |(\phi_{kij}^\lambda)^\varepsilon|^2 |\nabla K_\varepsilon(\nabla u_{0,\lambda})|^2 \\ &\quad + C\varepsilon^4 \int_{Q_1} \sum_{k,j} |(\phi_{k(d+1)j}^\lambda)^\varepsilon|^2 |\partial_t K_\varepsilon(\nabla u_{0,\lambda})|^2 \\ &\quad + C\varepsilon^2 \int_{Q_1} \sum_{\ell,k} |(\nabla \phi_{\ell(d+1)k}^\lambda)^\varepsilon|^2 |\nabla K_\varepsilon(\nabla u_{0,\lambda})|^2 \\ &\quad + C\varepsilon^4 \int_{Q_1} \sum_{\ell,k} |(\phi_{\ell(d+1)k}^\lambda)^\varepsilon|^2 |\nabla^2 K_\varepsilon(\nabla u_{0,\lambda})|^2 \\ &= I_1 + I_2 + I_3 + I_4 + I_5 + I_6. \end{aligned} \quad (3.28)$$

To bound  $I_1$ , we use Lemma 3.4. This gives

$$\begin{aligned} I_1 &\leq 2 \int_{Q_1} |\nabla u_{0,\lambda} - \eta_\delta(\nabla u_{0,\lambda})|^2 + 2 \int_{Q_1} |\eta_\delta(\nabla u_{0,\lambda}) - S_\delta(\eta_\delta(\nabla u_{0,\lambda}))|^2 \\ &\leq C \int_{Q_1 \setminus Q_{1-3\delta}} |\nabla u_{0,\lambda}|^2 + C\delta^2 \int_{Q_{1-2\delta}} (|\nabla^2 u_{0,\lambda}|^2 + |\partial_t u_{0,\lambda}|^2). \end{aligned}$$

By the standard regularity estimates for parabolic systems with constant coefficients [20, 21]

$$\int_{Q_{1-2\delta}} (|\nabla^2 u_{0,\lambda}|^2 + |\partial_t u_{0,\lambda}|^2) \leq C \left\{ \int_{Q_{1-\delta}} \frac{|\nabla u_{0,\lambda}(y, s)|^2 dy ds}{|\text{dist}_p((y, s), \partial_p Q_1)|^2} + \int_{Q_1} |F|^2 \right\},$$

where  $\text{dist}_p((y, s), \partial_p Q_1) = \inf \{|x - y| + |s - t|^{1/2} : (x, t) \in \partial_p Q_1\}$  denotes the parabolic distance from  $(y, s)$  to  $\partial_p Q_1$ . It follows that

$$\begin{aligned} I_1 &\leq C \int_{Q_1 \setminus Q_{1-3\delta}} |\nabla u_{0,\lambda}|^2 + C\delta^2 \left\{ \int_{Q_{1-\delta}} \frac{|\nabla u_{0,\lambda}(y, s)|^2 dy ds}{|\text{dist}_p((y, s), \partial_p Q_1)|^2} + \int_{Q_1} |F|^2 \right\} \\ &\leq C\delta^{1-\frac{2}{q}} \left( \int_{Q_1} |\nabla u_{0,\lambda}|^q \right)^{2/q} + C\delta^2 \int_{Q_1} |F|^2, \end{aligned} \tag{3.29}$$

where  $q > 2$  and we have used Hölder’s inequality for the last step.

To bound  $I_2, I_3$  and  $I_5$ , we use Lemma 3.3 and estimates (2.5) and (3.7) as well as the observation  $\nabla S_\delta(f) = S_\delta(\nabla f)$ . Note that  $(\chi^\lambda)^\varepsilon, (\phi_{kij}^\lambda)^\varepsilon$  and  $(\nabla \phi_{\ell(d+1)k}^\lambda)^\varepsilon$  are  $\varepsilon$ -periodic in  $x$  and  $\varepsilon^2\lambda$ -periodic in  $t$ . Since  $\delta = (1 + \sqrt{\lambda})\varepsilon \geq \varepsilon$  and  $\delta^2 \geq \varepsilon^2\lambda$ , we obtain

$$\begin{aligned} &\int_{Q_\delta(x,t)} \left( |(\chi^\lambda)^\varepsilon|^2 + |(\phi_{kij}^\lambda)^\varepsilon|^2 + |(\nabla \phi_{\ell(d+1)k}^\lambda)^\varepsilon|^2 \right) \\ &\leq C \int_0^\lambda \int_{\mathbb{T}^d} \left( |\chi^\lambda|^2 + |\phi_{kij}^\lambda|^2 + |\nabla \phi_{\ell(d+1)k}^\lambda|^2 \right) \\ &\leq C \end{aligned}$$

for any  $(x, t) \in \mathbb{R}^{d+1}$ . It follows that

$$\begin{aligned} I_2 + I_3 + I_5 &\leq C\varepsilon^2 \int_{Q_1} |\nabla(\eta_\delta(\nabla u_{0,\lambda}))|^2 \\ &\leq C\delta^{1-\frac{2}{q}} \left( \int_{Q_1} |\nabla u_{0,\lambda}|^q \right)^{2/q} + C\delta^2 \int_{Q_1} |F|^2. \end{aligned} \tag{3.30}$$

To bound  $I_6$ , we use the inequality (3.18) as well as the estimate (3.8). This leads to

$$\begin{aligned} I_6 &\leq C\varepsilon^4(1 + \lambda)^2\delta^{-2} \int_{Q_1} |\nabla(\eta_\delta \nabla u_{0,\lambda})|^2 \\ &\leq C\delta^{1-\frac{2}{q}} \left( \int_{Q_1} |\nabla u_{0,\lambda}|^q \right)^{2/q} + C\delta^2 \int_{Q_1} |F|^2. \end{aligned} \tag{3.31}$$

Finally, to handle  $I_4$ , we use the observation

$$\begin{aligned} \partial_t K_\varepsilon(\nabla u_{0,\lambda}) &= \partial_t S_\delta(\eta_\delta \nabla u_{0,\lambda}) \\ &= S_\delta((\partial_t \eta_\delta) \nabla u_{0,\lambda}) + S_\delta(\nabla(\eta_\delta \partial_t u_{0,\lambda})) - S_\delta((\nabla \eta_\delta) \partial_t u_{0,\lambda}). \end{aligned} \quad (3.32)$$

As in the case of  $I_6$ , we obtain

$$\begin{aligned} I_4 &\leq C\varepsilon^4(1+\lambda)^2 \int_{Q_1} \left\{ |(\partial_t \eta_\delta) \nabla u_{0,\lambda}|^2 + \delta^{-2} |\eta_\delta \partial_t u_{0,\lambda}|^2 + |(\nabla \eta_\delta) \partial_t u_{0,\lambda}|^2 \right\} \\ &\leq C\delta^{1-\frac{2}{q}} \left( \int_{Q_1} |\nabla u_{0,\lambda}|^q \right)^{2/q} + C\delta^2 \int_{Q_1} |F|^2. \end{aligned} \quad (3.33)$$

Let  $\sigma = \frac{1}{2} - \frac{1}{q} > 0$ . In view of (3.29)–(3.32), we have proved that

$$\begin{aligned} \int_{Q_1} |\nabla w_\varepsilon|^2 &\leq C\delta^{2\sigma} \left( \int_{Q_1} |\nabla u_{0,\lambda}|^q \right)^{2/q} + C\delta^2 \int_{Q_1} |F|^2 \\ &\leq C\delta^{2\sigma} \left\{ \int_{Q_2} |\nabla u_{\varepsilon,\lambda}|^2 + \int_{Q_2} |F|^2 \right\}, \end{aligned} \quad (3.34)$$

where we have used (3.26) for the last step. To finish the proof, we let  $H_\varepsilon$  be the last term in (3.21). It is easy to see that

$$\int_{Q_1} |\nabla H_\varepsilon|^2 \leq I_5 + I_6.$$

This, together with (3.34), gives the estimate (3.24).  $\square$

We are now ready to give the proof of Theorem 3.1.

*Proof of Theorem 3.1.* By translation and dilation we may assume that  $r = 1$  and  $Q_2 = B(0, 2) \times (-4, 0)$ . We may also assume that  $\delta = (1 + \sqrt{\lambda})\varepsilon \leq 1/8$ . This reduces the problem to the case considered in Lemma 3.6. Observe that  $K_\varepsilon(\nabla u_{0,\lambda}) = S_\delta(\nabla u_{0,\lambda})$  on  $Q_{1/2}$ . Thus, in view of Lemma 3.6, it suffices to show that

$$\left( \int_{Q_{1/2}} |\nabla \{ \varepsilon(\chi^\lambda)^\varepsilon S_\delta(\nabla u_{0,\lambda}) \} - (\nabla \chi^\lambda)^\varepsilon \nabla u_{0,\lambda}|^2 \right)^{1/2} \quad (3.35)$$

is bounded by the right-hand side of (3.24). Furthermore, since  $(\partial_t + \mathcal{L}_{0,\lambda})u_{0,\lambda} = F$  in  $Q_1$ , we have

$$\|\nabla^2 u_{0,\lambda}\|_{L^2(Q_{3/4})} \leq C \left\{ \left( \int_{Q_1} |\nabla u_{0,\lambda}|^2 \right)^{1/2} + \left( \int_{Q_1} |F|^2 \right)^{1/2} \right\}.$$

Also, recall that

$$\|(\chi^\lambda)^\varepsilon\|_{L^2(Q_1)} + \|(\nabla \chi^\lambda)^\varepsilon\|_{L^2(Q_1)} \leq C. \quad (3.36)$$

As a result, it is enough to show that

$$\left( \int_{Q_{1/2}} |(\nabla \chi^\lambda)^\varepsilon (S_\delta(\nabla u_{0,\lambda}) - \nabla u_{0,\lambda})|^2 \right)^{1/2} \tag{3.37}$$

is bounded by the right-hand side of (3.24). This, however, follows from (3.36) and the estimate

$$\|S_\delta(\nabla u_{0,\lambda}) - \nabla u_{0,\lambda}\|_{L^\infty(Q_{1/2})} \leq C\delta^\sigma \left\{ \left( \int_{Q_1} |\nabla u_{0,\lambda}|^2 \right)^{1/2} + \left( \int_{Q_1} |F|^p \right)^{1/p} \right\}, \tag{3.38}$$

where  $p > d + 2$  and  $\sigma = 1 - \frac{d+2}{p}$ .

Finally, we point out that (3.38) follows readily from the  $C^{1+\sigma}$  estimates for  $\partial_t + \mathcal{L}_{0,\lambda}$ ,

$$\begin{aligned} & |\nabla u_{0,\lambda}(x, t) - \nabla u_{0,\lambda}(y, s)| \\ & \leq C(|x - y| + |t - s|^{1/2})^\sigma \left\{ \left( \int_{Q_1} |\nabla u_{0,\lambda}|^2 \right)^{1/2} + \left( \int_{Q_1} |F|^p \right)^{1/p} \right\} \end{aligned} \tag{3.39}$$

for any  $(x, t), (y, s) \in Q_{1/2}$ . This completes the proof.

#### 4. Large-Scale Lipschitz and $C^{1,\alpha}$ Estimates

In this section we establish the large-scale Lipschitz and  $C^{1,\alpha}$  estimates for  $\partial_t + \mathcal{L}_{\varepsilon,\lambda}$ . As a consequence, we obtain the same estimates for the parabolic operator  $\partial_t + \mathcal{L}_\varepsilon$  in (1.1). Let

$$\begin{aligned} P_{1,\varepsilon}^\lambda &= \left\{ P = P(x, t) : P(x, t) = \beta_0 + \beta_j(x_j + \varepsilon \chi_j^\lambda(x/\varepsilon, t/\varepsilon^2)) \right. \\ & \left. \text{for some } \beta = (\beta_0, \beta_1, \dots, \beta_d) \in \mathbb{R}^{d+1} \right\}, \end{aligned} \tag{4.1}$$

where the index  $j$  is summed from 1 to  $d$ . Note that  $(\partial_t + \mathcal{L}_{\varepsilon,\lambda})P = 0$  in  $\mathbb{R}^{d+1}$  for any  $P \in P_{1,\varepsilon}^\lambda$ .

**Theorem 4.1.** ( $C^{1,\alpha}$  estimate) *Suppose  $A$  satisfies conditions (1.3) and (1.4). Let  $u_{\varepsilon,\lambda}$  be a weak solution of  $(\partial_t + \mathcal{L}_{\varepsilon,\lambda})u_{\varepsilon,\lambda} = F$  in  $Q_R$ , where  $R > (1 + \sqrt{\lambda})\varepsilon$  and  $F \in L^p(Q_R)$  for some  $p > d + 2$ . Then, for any  $(1 + \sqrt{\lambda})\varepsilon \leq r < R$  and  $0 < \alpha < 1 - \frac{d+2}{p}$ ,*

$$\begin{aligned} \inf_{P \in P_{1,\varepsilon}^\lambda} \left( \int_{Q_r} |\nabla(u_{\varepsilon,\lambda} - P)|^2 \right)^{1/2} & \leq C \left( \frac{r}{R} \right)^\alpha \left\{ \left( \int_{Q_R} |\nabla u_{\varepsilon,\lambda}|^2 \right)^{1/2} \right. \\ & \left. + R \left( \int_{Q_R} |F|^p \right)^{1/p} \right\}, \end{aligned} \tag{4.2}$$

where  $C > 0$  depends only on  $d, \mu, p$  and  $\alpha$ .

*Proof.* The proof relies on the approximation results in Theorem 3.1 and uses classical regularity estimates for parabolic systems with constant coefficients. By translation and dilation we may assume that  $R = 2$  and  $Q_2 = B(0, 2) \times (-4, 0)$ . Let

$$(1 + \sqrt{\lambda})\varepsilon < \theta r < r < 1,$$

where  $\theta \in (0, 1/4)$  is to be chosen later. Let  $u_{0,\lambda}$  be the weak solution of  $(\partial_t + \mathcal{L}_{0,\lambda})u_{0,\lambda} = F$  in  $Q_r$ , given by Theorem 3.1. By the classical  $C^{1+\alpha}$  estimates for parabolic systems with constant coefficients [20,21],

$$\begin{aligned} |\nabla u_{0,\lambda}(x, t) - \nabla u_{0,\lambda}(0, 0)| &\leq C \left( \frac{|x| + |t|^{1/2}}{r} \right)^{\alpha_p} \left\{ \left( \int_{Q_r} |\nabla u_{0,\lambda}|^2 \right)^{1/2} \right. \\ &\quad \left. + r \left( \int_{Q_r} |F|^p \right)^{1/p} \right\} \end{aligned}$$

for any  $(x, t) \in Q_{r/2}$ , where  $\alpha_p = 1 - \frac{d+2}{p}$ . Let  $P(x, t) = \beta_j(x_j + \varepsilon \chi_j^\lambda(x/\varepsilon, t/\varepsilon^2))$  with  $\beta_j = \frac{\partial u_{0,\lambda}}{\partial x_j}(0, 0)$ . Then

$$\begin{aligned} &\left( \int_{Q_{\theta r}} |\nabla u_{0,\lambda}(x, t) + \nabla \chi^\lambda(x/\varepsilon, t/\varepsilon^2) \nabla u_{0,\lambda}(x, t) - \nabla P(x, t)|^2 dx dt \right)^{1/2} \\ &\leq C \theta^{\alpha_p} \left\{ \left( \int_{Q_r} |\nabla u_{0,\lambda}|^2 \right)^{1/2} + r \left( \int_{Q_r} |F|^p \right)^{1/p} \right\} \end{aligned}$$

for any  $(x, t) \in Q_{\theta r}$ . It follows that

$$\begin{aligned} &\left( \int_{Q_{\theta r}} |\nabla(u_{\varepsilon,\lambda} - P)|^2 \right)^{1/2} + \theta r \left( \int_{Q_{\theta r}} |F|^p \right)^{1/p} \\ &\leq C \left( \int_{Q_{\theta r}} |\nabla u_{\varepsilon,\lambda} - \nabla u_{0,\lambda} - (\nabla \chi^\lambda)^\varepsilon \nabla u_{0,\lambda}|^2 \right)^{1/2} \\ &\quad + C \theta^{\alpha_p} \left\{ \left( \int_{Q_r} |\nabla u_{0,\lambda}|^2 \right)^{1/2} + r \left( \int_{Q_r} |F|^p \right)^{1/p} \right\} + \theta r \left( \int_{Q_{\theta r}} |F|^p \right)^{1/p} \\ &\leq C_0 \left\{ \theta^{-\frac{d+2}{2}} \left( \frac{(1 + \sqrt{\lambda})\varepsilon}{r} \right)^\sigma + \theta^{\alpha_p} \right\} \left\{ \left( \int_{Q_{2r}} |\nabla u_{\varepsilon,\lambda}|^2 \right)^{1/2} \right. \\ &\quad \left. + 2r \left( \int_{Q_{2r}} |F|^p \right)^{1/p} \right\}, \end{aligned}$$

where  $C_0$  depends only  $d, \mu$  and  $p$ . Fix  $0 < \alpha < \alpha_p$ . We choose  $\theta \in (0, 1/4)$  so small that  $C_0 \theta^{\alpha_p} \leq (1/2)\theta^\alpha$ . With  $\theta$  chosen, we assume that  $r \geq C_\theta(1 + \sqrt{\lambda})\varepsilon$ , where  $C_\theta > 1$  is so large that

$$C_0 \theta^{-\frac{d+2}{2}} C_\theta^{-\sigma} < (1/2)\theta^\alpha.$$



This leads to

$$\begin{aligned} & \left( \int_{Q_{\theta r}} |\nabla(u_{\varepsilon,\lambda} - P)|^2 \right)^{1/2} + \theta r \left( \int_{Q_{\theta r}} |F|^p \right)^{1/p} \\ & \leq \theta^\alpha \left\{ \left( \int_{Q_{2r}} |\nabla u_{\varepsilon,\lambda}|^2 \right)^{1/2} + 2r \left( \int_{Q_{2r}} |F|^p \right)^{1/p} \right\}. \end{aligned}$$

Since  $(\partial_t + \mathcal{L}_{\varepsilon,\lambda})P = 0$  in  $\mathbb{R}^{d+1}$  for any  $P \in P_{1,\varepsilon}^\lambda$ , we obtain

$$\begin{aligned} & \inf_{P \in P_{1,\varepsilon}^\lambda} \left( \int_{Q_{\theta r}} |\nabla(u_{\varepsilon,\lambda} - P)|^2 \right)^{1/2} + \theta r \left( \int_{Q_{\theta r}} |F|^p \right)^{1/p} \\ & \leq \theta^\alpha \left\{ \inf_{P \in P_{1,\varepsilon}^\lambda} \left( \int_{Q_{2r}} |\nabla(u_{\varepsilon,\lambda} - P)|^2 \right)^{1/2} + 2r \left( \int_{Q_{2r}} |F|^p \right)^{1/p} \right\} \end{aligned} \quad (4.3)$$

for any  $C_\theta(1 + \sqrt{\lambda})\varepsilon \leq r < 1$ . By an iteration argument it follows that

$$\begin{aligned} & \inf_{P \in P_{1,\varepsilon}^\lambda} \left( \int_{Q_r} |\nabla(u_{\varepsilon,\lambda} - P)|^2 \right)^{1/2} + r \left( \int_{Q_r} |F|^p \right)^{1/p} \\ & \leq Cr^\alpha \left\{ \inf_{P \in P_{1,\varepsilon}^\lambda} \left( \int_{Q_2} |\nabla(u_{\varepsilon,\lambda} - P)|^2 \right)^{1/2} + \left( \int_{Q_2} |F|^p \right)^{1/p} \right\} \end{aligned} \quad (4.4)$$

for any  $(1 + \sqrt{\lambda})\varepsilon \leq r < 1$ . This gives the large-scale  $C^{1,\alpha}$  estimate (4.2).  $\square$

**Theorem 4.2.** (Lipschitz estimate) *Suppose A satisfies conditions (1.3) and (1.4). Let  $u_{\varepsilon,\lambda}$  be a weak solution of  $(\partial_t + \mathcal{L}_{\varepsilon,\lambda})u_{\varepsilon,\lambda} = F$  in  $Q_R$ , where  $R > (1 + \sqrt{\lambda})\varepsilon$  and  $F \in L^p(Q_R)$  for some  $p > d + 2$ . Then, for any  $(1 + \sqrt{\lambda})\varepsilon \leq r < R$ ,*

$$\left( \int_{Q_r} |\nabla u_{\varepsilon,\lambda}|^2 \right)^{1/2} \leq C \left\{ \left( \int_{Q_R} |\nabla u_{\varepsilon,\lambda}|^2 \right)^{1/2} + R \left( \int_{Q_R} |F|^p \right)^{1/p} \right\}, \quad (4.5)$$

where  $C > 0$  depends only on  $d$ ,  $\mu$  and  $p$ .

*Proof.* By translation and dilation we may assume that  $R = 2$  and  $Q_2 = B(0, 2) \times (-4, 0)$ . Define

$$h(r) = \left( \int_{Q_r} |\nabla H_r|^2 \right)^{1/2},$$

where  $H_r = E_r \cdot (x + \varepsilon\chi^\lambda(x/\varepsilon, t/\varepsilon^2))$ , with  $E_r \in \mathbb{R}^d$ , is a function in  $P_{1,\varepsilon}^\lambda$  such that

$$\left( \int_{Q_r} |\nabla(u_{\varepsilon,\lambda} - H_r)|^2 \right)^{1/2} = \inf_{P \in P_{1,\varepsilon}^\lambda} \left( \int_{Q_r} |\nabla(u_{\varepsilon,\lambda} - P)|^2 \right)^{1/2}.$$

Let  $C(1 + \sqrt{\lambda})\varepsilon < r < 1/2$ . Note that

$$\begin{aligned} |E_{2r} - E_r| &\leq \frac{C}{r} \inf_{\beta_0 \in \mathbb{R}} \left( \int_{Q_{r/2}} |(E_{2r} - E_r) \cdot x - \beta_0|^2 \right)^{1/2} \\ &\leq \frac{C}{r} \inf_{\beta_0 \in \mathbb{R}} \left( \int_{Q_{r/2}} |H_{2r} - H_r - \beta_0|^2 \right)^{1/2} + C|E_{2r} - E_r|r^{-1}\varepsilon, \end{aligned}$$

where  $C$  depends only on  $d$  and  $\mu$ . It follows that if  $r \geq C_1\varepsilon$  and  $C_1 > 1$  is sufficiently large, then

$$\begin{aligned} |E_{2r} - E_r| &\leq \frac{C}{r} \inf_{\beta_0 \in \mathbb{R}} \left( \int_{Q_{r/2}} |H_{2r} - H_r - \beta_0|^2 \right)^{1/2} \\ &\leq C \left( \int_{Q_r} |\nabla(H_{2r} - H_r)|^2 \right)^{1/2}. \end{aligned} \tag{4.6}$$

We remark that the last inequality follows from the fact that  $u = H_{2r} - H_r - \beta_0$  is a solution of the second-order parabolic system in divergence form  $(\partial_t + \mathcal{L}_{\varepsilon, \lambda})u = 0$  in  $\mathbb{R}^{d+1}$ . Such a solution satisfies the Poincaré-type inequality,

$$\inf_{\beta_0 \in \mathbb{R}} \int_{Q_{r/2}} |u - \beta_0|^2 \leq Cr^2 \int_{Q_r} |\nabla u|^2$$

(see for example [13, Lemma 2.2]). Hence,

$$\begin{aligned} |E_{2r} - E_r| &\leq C \left( \int_{Q_{2r}} |\nabla(u_{\varepsilon, \lambda} - H_{2r})|^2 \right)^{1/2} + C \left( \int_{Q_r} |\nabla(u_{\varepsilon, \lambda} - H_r)|^2 \right)^{1/2} \\ &\leq Cr^\alpha \left\{ \inf_{P \in P_{1, \varepsilon}^\lambda} \left( \int_{Q_2} |\nabla(u_{\varepsilon, \lambda} - P)|^2 \right)^{1/2} + \left( \int_{Q_2} |F|^p \right)^{1/p} \right\}, \end{aligned}$$

where we have used (4.4) for the last step. By a simple summation this yields

$$h(r) \leq C|E_r| \leq C \left\{ \left( \int_{Q_2} |\nabla u_{\varepsilon, \lambda}|^2 \right)^{1/2} + \left( \int_{Q_2} |F|^p \right)^{1/p} \right\},$$

which, together with (4.2), gives the large-scale Lipschitz estimate (4.5).  $\square$

*Proof of Theorem 1.1.* Recall that if  $\lambda = \varepsilon^{k-2}$ , then  $\mathcal{L}_{\varepsilon, \lambda} = \mathcal{L}_\varepsilon$ . Also note that in this case,  $(1 + \sqrt{\lambda})\varepsilon = \varepsilon + \varepsilon^{k/2}$ . As a result, Theorem 1.1 follows directly from Theorem 4.2.

*Remark 4.3.* ( $C^{1, \alpha}$  estimate) Let  $u_\varepsilon$  be a weak solution of  $(\partial_t + \mathcal{L}_\varepsilon)u_\varepsilon = F$  in  $Q_R$ , where  $R > \varepsilon + \varepsilon^{k/2}$  and  $F \in L^p(Q_R)$  for some  $p > d + 2$ . It follows from Theorem 4.1 that for  $\varepsilon + \varepsilon^{k/2} \leq r < R$  and  $0 < \alpha < 1 - \frac{d+2}{p}$ ,

$$\begin{aligned} & \inf_{E \in \mathbb{R}^d} \left( \int_{Q_r} |\nabla u_\varepsilon - E - E \nabla \chi^\lambda(x/\varepsilon, t/\varepsilon^2)|^2 \right)^{1/2} \\ & \leq C \left( \frac{r}{R} \right)^\alpha \left\{ \inf_{E \in \mathbb{R}^d} \left( \int_{Q_R} |\nabla u_\varepsilon - E - E \nabla \chi^\lambda(x/\varepsilon, t/\varepsilon^2)|^2 \right)^{1/2} \right. \\ & \quad \left. + R \left( \int_{Q_R} |F|^p \right)^{1/p} \right\}, \end{aligned} \tag{4.7}$$

where  $\lambda = \varepsilon^{k-2}$  and  $C$  depends only on  $d, \mu, p$  and  $\alpha$ . Note that  $\nabla \chi^\lambda(x/\varepsilon, t/\varepsilon^2)$  is  $\varepsilon$ -periodic in  $x$  and  $\varepsilon^k$ -periodic in  $t$ . One may regard (4.7) as a  $C^{1,\alpha}$  excess-decay estimate for the operator  $\partial_t + \mathcal{L}_\varepsilon$  in (1.1).

Let  $E_r \in \mathbb{R}^d$  be the constant for which the left-hand side of (4.7) obtains its minimum. It follows from the proof of Theorem 4.2 that

$$|E_r| \leq C \left\{ \left( \int_{Q_R} |\nabla u_\varepsilon|^2 \right)^{1/2} + R \left( \int_{Q_R} |F|^p \right)^{1/p} \right\}. \tag{4.8}$$

Let  $\chi^\infty$  be defined by (2.10). In view of (2.27) we have

$$\left( \int_{Q_r} |\nabla \chi^\lambda(x/\varepsilon, t/\varepsilon^2) - \nabla \chi^\infty(x/\varepsilon, t/\varepsilon^k)|^2 dx dt \right)^{1/2} \leq C \varepsilon^{2-k} \|\partial_s A\|_\infty. \tag{4.9}$$

This, together with (4.7) and (4.8), yields

$$\begin{aligned} & \inf_{E \in \mathbb{R}^d} \left( \int_{Q_r} |\nabla u_\varepsilon - E - E \nabla \chi^\infty(x/\varepsilon, t/\varepsilon^k)|^2 dx dt \right)^{1/2} \\ & \leq C \left\{ \left( \frac{r}{R} \right)^\alpha + \varepsilon^{2-k} \|\partial_s A\|_\infty \right\} \left\{ \left( \int_{Q_R} |\nabla u_\varepsilon|^2 \right)^{1/2} + R \left( \int_{Q_R} |F|^p \right)^{1/p} \right\}, \end{aligned} \tag{4.10}$$

for  $0 < k < 2$ . Similarly, for  $2 < k < \infty$ , we obtain

$$\begin{aligned} & \inf_{E \in \mathbb{R}^d} \left( \int_{Q_r} |\nabla u_\varepsilon - E - E \nabla \chi^0(x/\varepsilon, t/\varepsilon^k)|^2 dx dt \right)^{1/2} \\ & \leq C \left\{ \left( \frac{r}{R} \right)^\alpha + \varepsilon^{k-2} \|\nabla^2 A\|_\infty + \varepsilon^{k-2} \|\nabla A\|_\infty^2 \right\} \left\{ \left( \int_{Q_R} |\nabla u_\varepsilon|^2 \right)^{1/2} \right. \\ & \quad \left. + R \left( \int_{Q_R} |F|^p \right)^{1/p} \right\}. \end{aligned}$$

### 5. Higher-Order Correctors and $C^{2,\alpha}$ Estimates

In this section we introduce the second-order correctors and establish the large-scale  $C^{2,\alpha}$  estimates for  $\mathcal{L}_{\varepsilon,\lambda}$ .

Let  $A_\lambda = (a_{ij}^\lambda)$  and  $B_\lambda = (b_{k\ell}^\lambda)$  be the  $(1, \lambda)$ -periodic matrices given by (2.1) and (3.5), respectively. For  $1 \leq k, \ell \leq d$ , the second-order corrector  $\chi_{k\ell}^\lambda = \chi_{k\ell}^\lambda(y, s)$  is defined to be the weak solution of the cell problem

$$\begin{cases} \partial_s \chi_{k\ell}^\lambda - \operatorname{div}(A_\lambda \nabla \chi_{k\ell}^\lambda) = b_{k\ell}^\lambda + b_{\ell k}^\lambda + \frac{\partial}{\partial y_i} (a_{i\ell}^\lambda \chi_k^\lambda) + \frac{\partial}{\partial y_i} (a_{ik}^\lambda \chi_\ell^\lambda) & \text{in } \mathbb{R}^{d+1}, \\ \chi_{k\ell}^\lambda \text{ is } (1, \lambda)\text{-periodic in } (y, s), \\ \int_0^\lambda \int_{\mathbb{T}^d} \chi_{k\ell}^\lambda \, dy ds = 0, \end{cases} \quad (5.1)$$

where  $(\chi_j^\lambda)$  are the first-order correctors defined by (2.2). Since

$$\int_0^\lambda \int_{\mathbb{T}^d} b_{k\ell}^\lambda \, dy ds = 0,$$

the solution to (5.1) exists and is unique. Also, observe that  $\chi_{k\ell}^\lambda = \chi_{\ell k}^\lambda$ . Moreover, by the energy estimates,

$$\int_0^\lambda \int_{\mathbb{T}^d} |\nabla \chi_{k\ell}^\lambda|^2 \leq C, \quad (5.2)$$

where  $C$  depends only on  $d$  and  $\mu$ .

**Lemma 5.1.** *Let*

$$u(y, s) = y_k y_\ell + y_k \chi_\ell^\lambda(y, s) + y_\ell \chi_k^\lambda(y, s) + \chi_{k\ell}^\lambda(y, s).$$

*Then*

$$(\partial_s - \operatorname{div}(A_\lambda \nabla))u = (\partial_s - \operatorname{div}(\widehat{A}_\lambda \nabla))(y_k y_\ell) = -\widehat{a}_{\ell k}^\lambda - \widehat{a}_{k\ell}^\lambda$$

*in*  $\mathbb{R}^{d+1}$ , *where*  $\widehat{A}_\lambda = (\widehat{a}_{k\ell}^\lambda)$ .

*Proof.* This follows from a direct computation, using the definitions of  $\chi_j^\lambda$  and  $\chi_{k\ell}^\lambda$ .  $\square$

Let  $P_0(x, t) = \beta_0 + c_0 t + c_k x_k + c_{k\ell} x_k x_\ell$  and

$$\begin{aligned} P_\varepsilon(x, t) &= \beta_0 + c_0 t + c_k \{x_k + \varepsilon \chi_k^\lambda(x/\varepsilon, t/\varepsilon^2)\} \\ &\quad + c_{k\ell} \left\{ x_k x_\ell + \varepsilon x_k \chi_\ell^\lambda(x/\varepsilon, t/\varepsilon^2) + \varepsilon x_\ell \chi_k^\lambda(x/\varepsilon, t/\varepsilon^2) \right. \\ &\quad \left. + \varepsilon^2 \chi_{k\ell}^\lambda(x/\varepsilon, t/\varepsilon^2) \right\}, \end{aligned} \quad (5.3)$$

where  $\beta_0, c_0, c_k, c_{k\ell} = c_{\ell k} \in \mathbb{R}$ . It follows from Lemma 5.1 by rescaling that

$$(\partial_t + \mathcal{L}_{\varepsilon,\lambda})P_\varepsilon = (\partial_t + \mathcal{L}_{0,\lambda})P_0 = c_0 - 2c_{k\ell} \widehat{a_{k\ell}^\lambda} \quad \text{in } \mathbb{R}^{d+1}.$$

We shall use  $P_{2,\varepsilon}^\lambda$  to denote the set of all functions  $P_\varepsilon(x, t)$  in the form of (5.3) such that  $(\partial_t + \mathcal{L}_{\varepsilon,\lambda})P_\varepsilon = 0$ . Let  $C_p^\sigma(Q_R)$  denote the space of Hölder continuous functions  $u = u(x, t)$  such that

$$\begin{aligned} \|u\|_{C^\sigma(Q_R)} &:= R^\sigma \sup \left\{ \frac{|u(x, t) - u(y, s)|}{(|x - y| + |t - s|^{1/2})^\sigma} : (x, t), (y, s) \right. \\ &\quad \left. \in Q_R \text{ and } (x, t) \neq (y, s) \right\} < \infty, \end{aligned}$$

where  $\sigma \in (0, 1)$ .

**Theorem 5.2.** ( $C^{2,\alpha}$  estimate) *Suppose  $A$  satisfies conditions (1.3) and (1.4). Let  $u_{\varepsilon,\lambda}$  be a weak solution of  $(\partial_t + \mathcal{L}_{\varepsilon,\lambda})u_{\varepsilon,\lambda} = F$  in  $Q_R$ , where  $R > (1 + \sqrt{\lambda})\varepsilon$  and  $F \in C^\sigma(Q_R)$  for some  $\sigma \in (0, 1)$ . Then, for any  $(1 + \sqrt{\lambda})\varepsilon \leq r < R$  and  $0 < \alpha < \sigma$ ,*

$$\begin{aligned} &\inf_{P \in P_{2,\varepsilon}^\lambda} \left( \int_{Q_r} |\nabla(u_{\varepsilon,\lambda} - P)|^2 \right)^{1/2} \\ &\leq C \left( \frac{r}{R} \right)^{1+\alpha} \left\{ \inf_{P \in P_{2,\varepsilon}^\lambda} \left( \int_{Q_r} |\nabla(u_{\varepsilon,\lambda} - P)|^2 \right)^{1/2} + R \|F\|_{C^\sigma(Q_R)} \right\}, \end{aligned} \tag{5.4}$$

where  $C$  depends only on  $d, \sigma, \mu$ , and  $\alpha$ .

*Proof.* By translation and dilation we may assume that  $R = 2$  and  $Q_2 = B(0, 2) \times (-4, 0)$ . By subtracting  $c_0 t$  from  $u_{\varepsilon,\lambda}$ , we may also assume that  $F(0, 0) = 0$ , which implies  $\|F\|_{L^\infty(Q_r)} \leq C \|F\|_{C^\sigma(Q_r)}$ . Let  $(1 + \sqrt{\lambda})\varepsilon < \theta r < r < 1$ , where  $\theta \in (0, 1/4)$  is to be chosen later. Let  $u_{0,\lambda}$  be the weak solution of  $(\partial_t + \mathcal{L}_{0,\lambda})u_{0,\lambda} = F$  in  $Q_r$ , given by Theorem 3.1. By the classical  $C^{2+\alpha}$  estimates for parabolic systems with constant coefficients [20, 21],

$$\begin{aligned} &\left| \frac{\partial u_{0,\lambda}}{\partial x_i}(x, t) - \frac{\partial u_{0,\lambda}}{\partial x_i}(0, 0) - \frac{\partial^2 u_{0,\lambda}}{\partial x_j \partial x_i}(0, 0)x_j \right| \\ &\leq \left| \frac{\partial u_{0,\lambda}}{\partial x_i}(x, t) - \frac{\partial u_{0,\lambda}}{\partial x_i}(x, 0) \right| + \left| \frac{\partial u_{0,\lambda}}{\partial x_i}(x, 0) - \frac{\partial u_{0,\lambda}}{\partial x_i}(0, 0) - \frac{\partial^2 u_{0,\lambda}}{\partial x_j \partial x_i}(0, 0)x_j \right| \\ &\leq C\theta^{1+\sigma} \left\{ \left( \int_{Q_r} |\nabla u_{0,\lambda}|^2 \right)^{1/2} + r \|F\|_{C^\sigma(Q_r)} \right\} \\ &\leq C\theta^{1+\sigma} \left\{ \left( \int_{Q_{2r}} |\nabla u_{\varepsilon,\lambda}|^2 \right)^{1/2} + r \|F\|_{C^\sigma(Q_r)} \right\} \end{aligned} \tag{5.5}$$

for any  $(x, t) \in Q_{\theta r}$ , where we have used (3.3) for the last inequality. Let

$$P_0(x, t) = c_0 t + c_i x_i + c_{ij} x_i x_j,$$

where

$$c_0 = \partial_t u_{0,\lambda}(0, 0), \quad c_i = \frac{\partial u_{0,\lambda}}{\partial x_i}(0, 0), \quad \text{and} \quad c_{ij} = \frac{1}{2} \frac{\partial^2 u_{0,\lambda}}{\partial x_i \partial x_j}(0, 0). \quad (5.6)$$

Note that

$$(\partial_t + \mathcal{L}_{0,\lambda})P_0 = c_0 - 2c_{ij}\widehat{a_{ij}^\lambda} = (\partial_t + \mathcal{L}_{0,\lambda})u_0(0, 0) = F(0, 0) = 0, \quad (5.7)$$

and by (5.5),

$$\|\nabla(u_{0,\lambda} - P_0)\|_{L^\infty(Q_{\theta r})} \leq C\theta^{1+\sigma} \left\{ \left( \int_{Q_r} |\nabla u_{0,\lambda}|^2 \right)^{1/2} + r\|F\|_{C^\sigma(Q_r)} \right\}. \quad (5.8)$$

This, together with the inequality (3.4), gives

$$\begin{aligned} & \left( \int_{Q_{\theta r}} |\nabla u_{\varepsilon,\lambda} - \nabla P_0 - (\nabla \chi^\lambda)^\varepsilon (\nabla P_0)|^2 \right)^{1/2} \\ & \leq C \left\{ \theta^{1+\sigma} + \left( \frac{(1 + \sqrt{\lambda})\varepsilon}{r} \right)^\sigma \right\} \left\{ \left( \int_{Q_{2r}} |\nabla u_{\varepsilon,\lambda}|^2 \right)^{1/2} + r\|F\|_{C^\sigma(Q_{2r})} \right\}. \end{aligned} \quad (5.9)$$

Let  $P_\varepsilon = P_\varepsilon(x, t)$  be given by (5.3) with the same coefficients as those of  $P_0$  in (5.6). Then  $(\partial_t + \mathcal{L}_{\varepsilon,\lambda})P_\varepsilon = (\partial_t + \mathcal{L}_{0,\lambda})P_0 = 0$ , and

$$|\nabla P_\varepsilon - \nabla P_0 - (\nabla \chi^\lambda)^\varepsilon (\nabla P_0)| \leq \varepsilon |c_{k\ell} \nabla \chi_{k\ell}^\lambda(x/\varepsilon, t/\varepsilon^2)|. \quad (5.10)$$

In view of (5.9), we obtain

$$\begin{aligned} & \left( \int_{Q_{\theta r}} |\nabla(u_{\varepsilon,\lambda} - P_\varepsilon)|^2 \right)^{1/2} \\ & \leq C \left\{ \theta^{1+\sigma} + \left( \frac{(1 + \sqrt{\lambda})\varepsilon}{r} \right)^\sigma \right\} \left\{ \left( \int_{Q_{2r}} |\nabla u_{\varepsilon,\lambda}|^2 \right)^{1/2} + r\|F\|_{C^\sigma(Q_{2r})} \right\}, \end{aligned} \quad (5.11)$$

where we have used (5.2) and the assumption that  $\theta r \geq (1 + \sqrt{\lambda})\varepsilon$ .

To proceed, we let

$$\Psi(r) = \inf_{P \in P_{2,\varepsilon}^\lambda} \left( \int_{Q_r} |\nabla(u_{\varepsilon,\lambda} - P)|^2 \right)^{1/2} + r\|F\|_{C^\sigma(Q_r)}.$$

It follows from (5.11) that

$$\Psi(\theta r) \leq C_0 \left\{ \theta^{1+\sigma} + \left( \frac{(1 + \sqrt{\lambda})\varepsilon}{r} \right)^\sigma \right\} \Psi(2r)$$

for  $(1 + \sqrt{\lambda})\varepsilon < \theta r < r < 1$ , where  $C_0$  depends only on  $d, \mu$  and  $\sigma$ . Fix  $\alpha \in (0, \sigma)$ . Choose  $\theta \in (0, 1/4)$  so small that  $C_0\theta^{1+\sigma} \leq (1/2)(\theta/2)^{1+\alpha}$ . With  $\theta$

chosen, we may choose  $C_1 > 1$  so large that  $C_0 C_1^{-\sigma} \leq (1/2)(\theta/2)^{1+\alpha}$ . As a result, for  $C_1(1 + \sqrt{\lambda})\varepsilon < \theta r < r < 1$ , we have

$$\Psi(\theta r) \leq (\theta/2)^{1+\alpha} \Psi(2r).$$

By a simple iteration argument this gives  $\Psi(r) \leq C r^{1+\alpha} \Psi(2)$  for any  $(1 + \sqrt{\lambda})\varepsilon \leq r < 2$ .  $\square$

*Remark 5.3.* (Liouville property) By letting  $\lambda = \varepsilon^{k-2}$  in Theorem 5.2 we obtain a  $C^{2,\alpha}$  excess-decay estimate for  $\partial_t + \mathcal{L}_\varepsilon$  in (1.1) for any  $0 < k < \infty$ . The estimate may be used to establish a Liouville property for the operator. Indeed, let  $u_\varepsilon$  be a solution of  $(\partial_t + \mathcal{L}_\varepsilon)u_\varepsilon = 0$  in  $\mathbb{R}^d \times (-\infty, t_0)$  for some  $t_0 \in \mathbb{R}$ . Suppose there exist  $C_u > 0$  and  $\alpha \in (0, 1)$  such that

$$\left( \int_{Q_R(0,t_0)} |u_\varepsilon|^2 \right)^{1/2} \leq C_u R^{2+\alpha} \tag{5.12}$$

for any  $R > 1$ . By Caccioppoli’s inequality it follows that

$$\left( \int_{Q_R(0,t_0)} |\nabla u_\varepsilon|^2 \right)^{1/2} \leq C R^{1+\alpha}$$

for any  $R > 1$ . This, together with (5.4), implies that  $u_\varepsilon = P$  in  $\mathbb{R}^d \times (-\infty, t_0)$  for some  $P \in P_{2,\varepsilon}^\lambda$ .

### 6. Boundary Lipschitz Estimates

In this section we establish large-scale boundary Lipschitz estimates for the operator  $\partial_t + \mathcal{L}_{\varepsilon,\lambda}$ , where  $\mathcal{L}_{\varepsilon,\lambda} = -\operatorname{div}(A_\lambda(x/\varepsilon, t/\varepsilon^2)\nabla)$ . As a consequence, we obtain the large-scale boundary Lipschitz estimate for  $\partial_t + \mathcal{L}_\varepsilon$  in Theorem 1.2.

Throughout this section we will assume that  $\Omega$  is a bounded  $C^{1,\alpha}$  domain for some  $\alpha \in (0, 1)$ . Let

$$\begin{aligned} D_r(x_0, t_0) &= (B(x_0, r) \cap \Omega) \times (t_0 - r^2, t_0), \\ \Delta_r(x_0, t_0) &= (B(x_0, r) \cap \partial\Omega) \times (t_0 - r^2, t_0), \end{aligned} \tag{6.1}$$

where  $x_0 \in \partial\Omega$  and  $t_0 \in \mathbb{R}$ . For  $\alpha \in (0, 1)$  and  $\Delta_r = \Delta_r(x_0, t_0)$ , we use  $C^{1+\alpha}(\Delta_r)$  to denote the parabolic  $C^{1+\alpha}$  space of functions on  $\Delta_r$  with the scale-invariant norm,

$$\|f\|_{C^{1+\alpha}(\Delta_r)} := \|f\|_{L^\infty(\Delta_r)} + r \|\nabla_{\tan} f\|_{L^\infty(\Delta_r)} + r \|\nabla_{\tan} f\|_{C^\alpha(\Delta_r)} + \|f\|_{C_t^{\frac{1+\alpha}{2}}(\Delta_r)},$$

where  $\|g\|_{C^\alpha(\Delta_r)}$  is the smallest constant  $C_0$  such that

$$|g(x, t) - g(y, s)| \leq C_0 r^{-\alpha} (|x - y| + |t - s|^{1/2})^\alpha$$

for any  $(x, t), (y, s) \in \Delta_r$ , and

$$\begin{aligned} \|f\|_{C_t^{\frac{1+\alpha}{2}}(\Delta_r)} &= \inf \{C : |f(y, \tau) - f(y, s)| \\ &\leq C r^{-1-\alpha} |\tau - s|^{\frac{1+\alpha}{2}} \text{ for any } (y, \tau), (y, s) \in \Delta_r\}. \end{aligned}$$

**Theorem 6.1.** *Assume  $A = A(y, s)$  satisfies (1.3) and (1.4). Suppose that  $(\partial_t + \mathcal{L}_{\varepsilon,\lambda})u_{\varepsilon,\lambda} = F$  in  $D_R = D_R(x_0, t_0)$  and  $u_{\varepsilon,\lambda} = f$  on  $\Delta_R = \Delta_R(x_0, t_0)$ , where  $x_0 \in \partial\Omega$ ,  $(1 + \sqrt{\lambda})\varepsilon < R \leq 1$ , and  $F \in L^p(D_R)$  for some  $p > d + 2$ . Then, for any  $(1 + \sqrt{\lambda})\varepsilon \leq r < R$ ,*

$$\begin{aligned} & \left( \int_{D_r} |\nabla u_{\varepsilon,\lambda}|^2 \right)^{1/2} \\ & \leq C \left\{ \left( \int_{D_R} |\nabla u_{\varepsilon,\lambda}|^2 \right)^{1/2} + R^{-1} \|f\|_{C^{1+\alpha}(\Delta_R)} + R \left( \int_{D_R} |F|^p \right)^{1/p} \right\}, \end{aligned} \tag{6.2}$$

where  $C$  depends only on  $d, \mu, p, \alpha$ , and  $\Omega$ .

To prove Theorem 6.1, we localize the boundary of  $\Omega$ . Let  $\psi : \mathbb{R}^{d-1} \rightarrow \mathbb{R}$  be a  $C^{1,\alpha}$  function such that  $\psi(0) = 0$  and  $\|\psi\|_{C^{1,\alpha}(\mathbb{R}^{d-1})} \leq M$ . Define

$$\begin{aligned} T_r &= \{(x', x_d) : |x'| < r \text{ and } \psi(x') < x_d < 100\sqrt{d}(M + 1)\} \times (-r^2, 0), \\ I_r &= \{(x', \psi(x')) : |x'| < r\} \times (-r^2, 0), \end{aligned} \tag{6.3}$$

where  $0 < r < \infty$

We begin with an approximation lemma.

**Lemma 6.2.** *Assume  $A$  satisfies (1.3) and (1.4). Suppose that  $(\partial_t + \mathcal{L}_{\varepsilon,\lambda})u_{\varepsilon,\lambda} = F$  in  $T_{2r}$  and  $u_{\varepsilon,\lambda} = f$  on  $I_{2r}$  for some  $0 < r \leq 1$ . Then there exists a function  $u_{0,\lambda}$  such that  $(\partial_t + \mathcal{L}_{0,\lambda})u_{0,\lambda} = F$  in  $T_r$ ,  $u_{0,\lambda} = f$  on  $I_r$ , and*

$$\begin{aligned} & \left( \int_{T_r} |u_{\varepsilon,\lambda} - u_{0,\lambda}|^2 \right)^{1/2} \\ & \leq C \left( \frac{(1 + \sqrt{\lambda})\varepsilon}{r} \right)^\sigma \left\{ \left( \int_{T_{2r}} |u_{\varepsilon,\lambda}|^2 \right)^{1/2} + \|f\|_{C^{1+\alpha}(I_{2r})} + r^2 \left( \int_{T_{2r}} |F|^2 \right)^{1/2} \right\}, \end{aligned} \tag{6.4}$$

where  $\sigma \in (0, 1)$  and  $C > 0$  depend only on  $d, \mu, p$ , and  $M$ .

*Proof.* The proof is similar to that of Theorem 3.1. By dilation we may assume  $r = 1$ . Let  $u_{0,\lambda}$  be the weak solution to the initial-Dirichlet problem,

$$(\partial_t + \mathcal{L}_{0,\lambda})u_{0,\lambda} = F \quad \text{in } T_1 \quad \text{and} \quad u_{0,\lambda} = u_{\varepsilon,\lambda} \quad \text{on } \partial_p T_1.$$

It follows by the Meyers-type estimates and Caccioppoli's inequality for parabolic systems that

$$\begin{aligned} & \left( \int_{T_1} |\nabla u_{0,\lambda}|^q \right)^{1/q} \leq C \left( \int_{T_1} |u_{\varepsilon,\lambda}|^q \right)^{1/q} \\ & \leq C \left\{ \left( \int_{T_2} |u_{\varepsilon,\lambda}|^2 \right)^{1/2} + \left( \int_{T_2} |F|^2 \right)^{1/2} + \|f\|_{C^{1+\alpha}(I_2)} \right\}, \end{aligned} \tag{6.5}$$



where  $q > 2$  and  $C > 0$  depend only on  $d, \mu, \alpha$  and  $M$ . To see (6.4), we define  $w_\varepsilon$  as in (3.21). Using the same argument as in the proof of Theorem 3.1, we may show that

$$\left( \int_{T_1} |\nabla w_\varepsilon|^2 \right)^{1/2} \leq C \delta^\sigma \left( \int_{T_1} |\nabla u_{0,\lambda}|^q \right)^{1/q}, \quad (6.6)$$

where  $\delta = (1 + \sqrt{\lambda})\varepsilon$  and  $\sigma = \frac{1}{2} - \frac{1}{q} > 0$ . Since  $w_\varepsilon = 0$  on  $\partial_p T_1$ , it follows from Poincaré's inequality and (6.5) that

$$\begin{aligned} & \left( \int_{T_1} |w_\varepsilon|^2 \right)^{1/2} \\ & \leq C \delta^\sigma \left\{ \left( \int_{T_2} |u_{\varepsilon,\lambda}|^2 \right)^{1/2} + \left( \int_{T_2} |F|^2 \right)^{1/2} + \|f\|_{C^{1+\alpha}(I_2)} \right\}. \end{aligned} \quad (6.7)$$

This yields (6.4), as  $\|w_\varepsilon - (u_{\varepsilon,\lambda} - u_{0,\lambda})\|_{L^2(T_1)}$  is also bounded by the right-hand side of (6.7).  $\square$

For a function  $u$  in  $T_r$ , define

$$\Psi(r; u) = \frac{1}{r} \inf_{\substack{E \in \mathbb{R}^d \\ \beta_0 \in \mathbb{R}}} \left\{ \left( \int_{T_r} |u - E \cdot x - \beta_0|^2 \right)^{1/2} + \|u - E \cdot x - \beta_0\|_{C^{1+\alpha}(I_r)} \right\}. \quad (6.8)$$

**Lemma 6.3.** *Suppose that  $(\partial_t + \mathcal{L}_{0,\lambda})u = F$  in  $T_r$ , where  $0 < r \leq 1$  and  $F \in L^p(T_r)$  for some  $p > d + 2$ . Then there exists  $\theta \in (0, 1/4)$ , depending only on  $d, \mu, \alpha, p$ , and  $M$ , such that*

$$\Psi(\theta r; u) + \theta r \left( \int_{T_{\theta r}} |F|^p \right)^{1/p} \leq \frac{1}{2} \left\{ \Psi(r; u) + r \left( \int_{T_r} |F|^p \right)^{1/p} \right\}. \quad (6.9)$$

*Proof.* Choose  $\sigma \in (0, 1)$  such that  $\sigma < \min(\alpha, 1 - \frac{d+2}{p})$ . The proof uses the boundary  $C^{1+\sigma}$  estimate for second-order parabolic systems with constant coefficients in  $C^{1,\alpha}$  cylinders. Let  $E_0 = \nabla u(0, 0)$  and  $\beta_0 = u(0, 0)$ . Then, for any  $(x, t) \in T_{r/2}$ ,

$$\begin{aligned} & |u(x, t) - E_0 \cdot x - \beta_0| \\ & \leq C(|x| + |t|^{1/2})^{1+\sigma} \left\{ \left( \int_{T_r} |u|^2 \right)^{1/2} + \|u\|_{C^{1+\alpha}(\Delta_r)} + r^2 \left( \int_{T_r} |F|^p \right)^{1/p} \right\}, \end{aligned}$$

where  $C$  depends only on  $d, \mu, \alpha, p$ , and  $M$ . It follows that the left-hand side of (6.9) is bounded by

$$\frac{C_0 \theta^\sigma}{r} \left\{ \left( \int_{T_r} |u|^2 \right)^{1/2} + \|u\|_{C^{1+\alpha}(\Delta_r)} + r^2 \left( \int_{T_r} |F|^p \right)^{1/p} \right\}.$$

Since  $(\partial_t + \mathcal{L}_{0,\lambda})(E \cdot x + \beta_0) = 0$  for any  $E \in \mathbb{R}^d$  and  $\beta_0 \in \mathbb{R}$ , we may replace  $u$  by  $u - E \cdot x - \beta_0$ . As a result, we see that the left-hand side of (6.9) is bounded by

$$C_0\theta^\sigma \left\{ \Psi(r; u) + r \left( \int_{T_r} |F|^p \right)^{1/p} \right\}.$$

To finish the proof, we choose  $\theta \in (0, 1/4)$  so small that  $C_0\theta^\sigma \leq (1/2)$ .  $\square$

**Lemma 6.4.** *Suppose that  $(\partial_t + \mathcal{L}_{\varepsilon,\lambda})u_{\varepsilon,\lambda} = F$  in  $T_2$  and  $u_{\varepsilon,\lambda} = f$  on  $I_2$ , where  $(1 + \sqrt{\lambda})\varepsilon < 1$  and  $F \in L^p(T_2)$  for some  $p > d + 2$ . Let  $\theta \in (0, 1/4)$  be given by Lemma 6.3. Then for any  $(1 + \sqrt{\lambda})\varepsilon \leq r \leq 1$ ,*

$$\begin{aligned} & \Psi(\theta r; u_{\varepsilon,\lambda}) + \theta r \left( \int_{T_{\theta r}} |F|^p \right)^{1/p} \\ & \leq \frac{1}{2} \left\{ \Psi(r; u_{\varepsilon,\lambda}) + r \left( \int_{T_r} |F|^p \right)^{1/p} \right\} \\ & \quad + C \left( \frac{(1 + \sqrt{\lambda})\varepsilon}{r} \right)^\sigma \left\{ \frac{1}{r} \left( \int_{T_{2r}} |u_{\varepsilon,\lambda}|^2 \right)^{1/2} \right. \\ & \quad \left. + r \left( \int_{T_{2r}} |F|^p \right)^{1/p} + r^{-1} \|f\|_{C^{1+\alpha}(I_{2r})} \right\}, \end{aligned} \tag{6.10}$$

where  $C$  depends only on  $d, \mu, p, \alpha$  and  $M$ .

*Proof.* Fix  $(1 + \sqrt{\lambda})\varepsilon \leq r \leq 1$ . Let  $u_{0,\lambda}$  be the solution of  $(\partial_t + \mathcal{L}_{0,\lambda})u_{0,\lambda} = F$  in  $T_r$  with  $u_{0,\lambda} = f$  on  $I_r$ , given by Lemma 6.2. Observe that

$$\begin{aligned} & \Psi(\theta r; u_{\varepsilon,\lambda}) + \theta r \left( \int_{T_{\theta r}} |F|^p \right)^{1/p} \\ & \leq \Psi(\theta r; u_{0,\lambda}) + \theta r \left( \int_{T_{\theta r}} |F|^p \right)^{1/p} + \frac{1}{\theta r} \left( \int_{T_{\theta r}} |u_{\varepsilon,\lambda} - u_{0,\lambda}|^2 \right)^{1/2} \\ & \leq \frac{1}{2} \left\{ \Psi(r; u_{0,\lambda}) + r \left( \int_{T_r} |F|^p \right)^{1/p} \right\} + \frac{C_\theta}{r} \left( \int_{T_r} |u_{\varepsilon,\lambda} - u_{0,\lambda}|^2 \right)^{1/2} \\ & \leq \frac{1}{2} \left\{ \Psi(r; u_{\varepsilon,\lambda}) + r \left( \int_{T_r} |F|^p \right)^{1/p} \right\} + \frac{C_\theta}{r} \left( \int_{T_r} |u_{\varepsilon,\lambda} - u_{0,\lambda}|^2 \right)^{1/2}, \end{aligned}$$

where we have used Lemma 6.3 for the second inequality. This, together with Lemma 6.2, gives (6.10).  $\square$

The proof of the next lemma may be found in [27, pp.157-158].

**Lemma 6.5.** *Let  $H(r)$  and  $h(r)$  be two nonnegative and continuous functions on the interval  $[0, 1]$ . Let  $0 < \delta < (1/4)$ . Suppose that there exists a constant  $C_0$  such that*

$$\max_{r \leq t \leq 2r} H(t) \leq C_0 H(2r) \quad \text{and} \quad \max_{r \leq t, s \leq 2r} |h(t) - h(s)| \leq C_0 H(2r) \tag{6.11}$$

for any  $r \in [\delta, 1/2]$ . Suppose further that

$$H(\theta r) \leq \frac{1}{2}H(r) + C_0\eta(\delta/r)\{H(2r) + h(2r)\} \quad (6.12)$$

for any  $r \in [\delta, 1/2]$ , where  $\theta \in (0, 1/4)$  and  $\eta(t)$  is a nonnegative and nondecreasing function on  $[0, 1]$  such that  $\eta(0) = 0$  and

$$\int_0^1 \frac{\eta(t)}{t} dt < \infty. \quad (6.13)$$

Then

$$\max_{\delta \leq r \leq 1} \{H(r) + h(r)\} \leq C\{H(1) + h(1)\}, \quad (6.14)$$

where  $C$  depends only on  $C_0$ ,  $\theta$ , and the function  $\eta(t)$ .

We are now ready to give the proof of Theorem 6.1

*Proof of Theorem 6.1.* By translation and dilation we may assume that  $(x_0, t_0) = (0, 0)$  and  $R = 1$ . Moreover, it suffices to show that for  $(1 + \sqrt{\lambda})\varepsilon \leq r < 2$ ,

$$\begin{aligned} & \left( \int_{T_r} |\nabla u_{\varepsilon, \lambda}|^2 \right)^{1/2} \\ & \leq C \left\{ \left( \int_{T_2} |\nabla u_{\varepsilon, \lambda}|^2 \right)^{1/2} + \|f\|_{C^{1+\alpha}(I_2)} + \left( \int_{T_2} |F|^p \right)^{1/p} \right\}, \end{aligned} \quad (6.15)$$

where  $(\partial_t + \mathcal{L}_{\varepsilon, \lambda})u_{\varepsilon, \lambda} = F$  in  $T_2$  and  $u_{\varepsilon, \lambda} = f$  on  $I_2$ . To this end, we apply Lemma 6.5 with

$$H(r) = \Psi(r; u_{\varepsilon, \lambda}) + r \left( \int_{T_r} |F|^p \right)^{1/p}$$

and  $h(t) = |E_r|$ , where  $E_r$  is a vector in  $\mathbb{R}^d$  such that

$$\Psi(r; u_{\varepsilon, \lambda}) = \frac{1}{r} \inf_{\beta_0 \in \mathbb{R}} \left\{ \left( \int_{T_r} |u_{\varepsilon, \lambda} - E_r \cdot x - \beta_0|^2 \right)^{1/2} + \|f - E_r \cdot x - \beta_0\|_{C^{1+\alpha}(I_r)} \right\}.$$

Note that, by (6.10),

$$H(\theta r) \leq \frac{1}{2}H(r) + C_0 \left( \frac{\delta}{r} \right)^\sigma \{H(2r) + h(2r)\}$$

for  $r \in [\delta, 1]$ , where  $\delta = (1 + \sqrt{\lambda})\varepsilon$ . This gives (6.12) with  $\eta(t) = t^\sigma$ , which satisfies (6.13).

It is easy to see that  $H(r)$  satisfies the first inequality in (6.11). To verify the second, we note that, for  $r \leq t$ ,  $s \leq 2r$ ,

$$\begin{aligned} |h(t) - h(s)| &\leq |E_t - E_s| \\ &\leq \frac{C}{r} \inf_{\beta_0 \in \mathbb{R}} \left( \int_{T_r} |(E_t - E_s) \cdot x - \beta_0|^2 \right)^{1/2} \\ &\leq \frac{C}{r} \inf_{\beta_0 \in \mathbb{R}} \left( \int_{T_r} |u_{\varepsilon, \lambda} - E_t \cdot x - \beta_0|^2 \right)^{1/2} \\ &\quad + \frac{C}{r} \inf_{\beta_0 \in \mathbb{R}} \left( \int_{T_r} |u_{\varepsilon, \lambda} - E_s \cdot x - \beta_0|^2 \right)^{1/2} \\ &\leq C\{H(t) + H(s)\} \\ &\leq CH(2r), \end{aligned}$$

where  $C$  depends only on  $d$ ,  $\alpha$  and  $M$ . Thus, by Lemma 6.5, we obtain

$$\begin{aligned} \frac{1}{r} \inf_{\beta \in \mathbb{R}} \left( \int_{T_r} |u_{\varepsilon, \lambda} - \beta|^2 \right)^{1/2} &\leq H(r) + h(r) \\ &\leq C\{H(1) + h(1)\} \\ &\leq C \left\{ \left( \int_{T_1} |u_{\varepsilon, \lambda}|^2 \right)^{1/2} + \|f\|_{C^{1+\alpha}(I_1)} \right. \\ &\quad \left. + \left( \int_{T_1} |F|^p \right)^{1/p} \right\}. \end{aligned}$$

By Caccioppoli's inequality for parabolic systems (see for example [2, Appendix]),

$$\left( \int_{T_{r/2}} |\nabla u_{\varepsilon, \lambda}|^2 \right)^{1/2} \leq C \left\{ \left( \int_{T_1} |u_{\varepsilon, \lambda}|^2 \right)^{1/2} + \|f\|_{C^{1+\alpha}(I_1)} + \left( \int_{T_1} |F|^p \right)^{1/p} \right\}.$$

Since  $(\partial_t + \mathcal{L}_{\varepsilon, \lambda})(\beta_0) = 0$  for any  $\beta_0 \in \mathbb{R}$ , we may replace  $u_{\varepsilon, \lambda}$  in the right-hand side of the inequality above by  $u_{\varepsilon, \lambda} - \beta_0$ . This, together with Poincaré-type inequality for parabolic systems, yields (6.15).  $\square$

*Proof of Theorem 1.2.* Since  $\mathcal{L}_\varepsilon = \mathcal{L}_{\varepsilon, \lambda}$  for  $\lambda = \varepsilon^{k-2}$ , Theorem 1.2 follows readily from Theorem 6.1.

## 7. Convergence Rates

In this section we investigate the problem of convergence rates for the initial-Dirichlet problem,

$$\begin{cases} (\partial_t + \mathcal{L}_{\varepsilon, \lambda})u_{\varepsilon, \lambda} = F & \text{in } \Omega_T, \\ u_{\varepsilon, \lambda} = f & \text{on } \partial_p \Omega_T, \end{cases} \quad (7.1)$$

where  $\Omega$  is a bounded domain in  $\mathbb{R}^d$  and  $\Omega_T = \Omega \times (0, T)$ . As a consequence, we obtain rates of convergence for the operator  $\partial_t + \mathcal{L}_\varepsilon$  in (1.1).

Let  $u_{0,\lambda}$  be the solution of the homogenized problem for (7.1),

$$\begin{cases} (\partial_t + \mathcal{L}_{0,\lambda})u_{0,\lambda} = F & \text{in } \Omega_T, \\ u_{0,\lambda} = f & \text{on } \partial_\rho \Omega_T. \end{cases} \tag{7.2}$$

Let  $w_\varepsilon$  be the two-scale expansion given by (3.21). As before, the operator  $K_\varepsilon$  is defined by  $K_\varepsilon(f) = S_\delta(\eta_\delta f)$  with  $\delta = (1 + \sqrt{\lambda})\varepsilon$ . The cut-off function  $\eta_\delta = \eta_\delta^1(x)\eta_\delta^2(t)$  is chosen so that  $0 \leq \eta_\delta \leq 1$ ,  $|\nabla \eta_\delta| \leq C/\delta$ ,  $|\partial_t \eta_\delta| + |\nabla^2 \eta_\delta| \leq C/\delta^2$ , and

$$\eta_\delta = 1 \quad \text{in } \Omega_T \setminus \Omega_{T,3\delta} \quad \text{and} \quad \eta_\delta = 0 \quad \text{in } \Omega_{T,2\delta},$$

where  $\Omega_{T,\rho}$  denotes the (parabolic) boundary layer

$$\Omega_{T,\rho} = \left( \{x \in \Omega : \text{dist}(x, \partial\Omega) < \rho\} \times (0, T) \right) \cup \left( \Omega \times (0, \rho^2) \right) \tag{7.3}$$

for  $0 < \rho \leq c$ .

**Lemma 7.1.** *Let  $\Omega$  be a bounded Lipschitz domain in  $\mathbb{R}^d$ . Let  $\Omega_{T,\rho}$  be defined by (7.3). Then*

$$\|\nabla g\|_{L^2(\Omega_{T,\rho})} \leq C\sqrt{\rho} \left\{ \|\nabla g\|_{L^2(\Omega_T)} + \|\nabla^2 g\|_{L^2(\Omega_T)} + \|\partial_t g\|_{L^2(\Omega_T)} \right\}, \tag{7.4}$$

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

*Proof.* Let  $\Omega_\rho = \{x \in \Omega : \text{dist}(x, \partial\Omega) < \rho\}$ . Then

$$\|\nabla g(\cdot, t)\|_{L^2(\Omega_\rho)} \leq C\sqrt{\rho} \|\nabla g(\cdot, t)\|_{H^1(\Omega)}.$$

It follows that

$$\|\nabla g\|_{L^2(\Omega_\rho \times (0, T))} \leq C\sqrt{\rho} \left\{ \|\nabla g\|_{L^2(\Omega_T)} + \|\nabla^2 g\|_{L^2(\Omega_T)} \right\}.$$

To estimate  $\|\nabla g\|_{L^2((\Omega \setminus \Omega_\rho) \times (0, \rho^2))}$ , we choose a cut-off function  $\theta \in C_0^\infty(\Omega)$  such that  $0 \leq \theta \leq 1$ ,  $\theta = 1$  on  $\Omega \setminus \Omega_\rho$ , and  $|\nabla \theta| \leq C/\rho$ . By Fubini's Theorem we may also choose  $t_0 \in (T/2, T)$  such that

$$\int_\Omega |\nabla g(x, t_0)|^2 dx \leq \frac{2}{T} \int_{\Omega_T} |\nabla g|^2 dx dt.$$

Note that for any  $t \in (0, \rho^2)$ ,

$$\begin{aligned} \int_\Omega |\nabla g(x, t)|^2 \theta(x) dx &\leq \int_\Omega |\nabla g(x, t_0)|^2 \theta(x) dx + \left| \int_t^{t_0} \int_\Omega \partial_s (|\nabla g(x, s)|^2 \theta(x)) dx ds \right| \\ &\leq \frac{2}{T} \int_{\Omega_T} |\nabla g|^2 + \int_{\Omega_T} |\nabla^2 g| |\partial_t g| + 2 \int_{\Omega_T} |\nabla g| |\partial_t g| |\nabla \theta|, \end{aligned}$$

where we have used an integration by parts in  $x$  for the last step. By integrating the inequality above in the variable  $t$  over the interval  $(0, \rho^2)$ , we obtain

$$\int_0^{\rho^2} \int_\Omega |\nabla g|^2 \theta dx dt \leq C\rho \int_{\Omega_T} \left\{ |\nabla g|^2 + |\nabla^2 g|^2 + |\partial_t g|^2 \right\},$$

where we also used the Cauchy inequality. This completes the proof.  $\square$

**Lemma 7.2.** *Let  $\Omega$  be a bounded Lipschitz domain in  $\mathbb{R}^d$  and  $0 < T < \infty$ . Let  $u_{\varepsilon,\lambda}$  be a weak solution of (7.1) and  $u_{0,\lambda}$  the homogenized problem (7.2). Let  $w_\varepsilon$  be defined by (3.21). Then, for any  $\psi \in L^2(0, T; H_0^1(\Omega))$ ,*

$$\begin{aligned} & \left| \int_0^T (\partial_t w_\varepsilon, \psi)_{H^{-1}(\Omega) \times H_0^1(\Omega)} + \int_{\Omega_T} A_\lambda(x/\varepsilon, t/\varepsilon^2) \nabla w_\varepsilon \cdot \nabla \psi \right| \\ & \leq C \left\{ \|u_{0,\lambda}\|_{L^2(0,T;H^2(\Omega))} + \|\partial_t u_{0,\lambda}\|_{L^2(\Omega_T)} \right\} \left\{ \delta \|\nabla \psi\|_{L^2(\Omega_T)} + \delta^{1/2} \|\nabla \psi\|_{L^2(\Omega_{T,3\delta})} \right\}, \end{aligned} \quad (7.5)$$

where  $\delta = (1 + \sqrt{\lambda})\varepsilon$  and  $C$  depends only on  $d, \mu, \Omega$  and  $T$ .

*Proof.* In view of Lemma 7.1, the case  $\lambda = 1$  follows from [14, Lemma 3.5]. The case  $\lambda \neq 1$  is proved in a similar manner. Indeed, by (3.22), the left-hand side of (7.5) is bounded by

$$\begin{aligned} & C \int_{\Omega_T} |\nabla u_{0,\lambda} - K_\varepsilon(\nabla u_{0,\lambda})| |\nabla \psi| + C\varepsilon \int_{\Omega_T} |(\chi^\lambda)^\varepsilon| |\nabla K_\varepsilon(\nabla u_{0,\lambda})| |\nabla \psi| \\ & + C\varepsilon \int_{\Omega_T} \sum_{k,i,j} |(\phi_{kij}^\lambda)^\varepsilon| |\nabla K_\varepsilon(\nabla u_{0,\lambda})| |\nabla \psi| \\ & + C\varepsilon^2 \int_{\Omega_T} \sum_{k,j} |(\phi_{k(d+1)j}^\lambda)^\varepsilon| |\partial_t K_\varepsilon(\nabla u_{0,\lambda})| |\nabla \psi| \\ & + C\varepsilon \int_{\Omega_T} \sum_{k,j} |(\nabla \phi_{k(d+1)j}^\lambda)^\varepsilon| |\nabla K_\varepsilon(\nabla u_{0,\lambda})| |\nabla \psi| \\ & + C\varepsilon^2 \int_{\Omega_T} \sum_{k,j} |(\phi_{k(d+1)j}^\lambda)^\varepsilon| |\nabla^2 K_\varepsilon(\nabla u_{0,\lambda})| |\nabla \psi| \\ & = I_1 + I_2 + I_3 + I_4 + I_5 + I_6. \end{aligned}$$

The estimates of  $I_j$  for  $j = 1, \dots, 6$  are exactly the same as in the proof of Lemma 3.5 in [14]. Also see the proof of Lemma 3.6 in Section 3. We point out that in the cases of  $I_4$  and  $I_6$ , the estimate

$$\sup_{(x,t) \in \mathbb{R}^{d+1}} \left( \int_{Q_\delta(x,t)} |(\phi_{k(d+1)j}^\lambda)^\varepsilon|^2 \right)^{1/2} \leq C(1 + \lambda)$$

is used. We omit the details.  $\square$

The next theorem gives an error estimate for the two-scale expansion

$$\tilde{w}_\varepsilon(x, t) = u_{\varepsilon,\lambda} - u_{0,\lambda} - \varepsilon \chi^\lambda(x/\varepsilon, t/\varepsilon^2) K_\varepsilon(\nabla u_{0,\lambda}) \quad (7.6)$$

in  $L^2(0, T; H^1(\Omega))$ .

**Theorem 7.3.** *Let  $\tilde{w}_\varepsilon$  be defined by (7.6). Under the same conditions as in Lemma 7.2, we have*

$$\|\nabla \tilde{w}_\varepsilon\|_{L^2(\Omega_T)} \leq C\sqrt{\delta} \left\{ \|u_{0,\lambda}\|_{L^2(0,T;H^2(\Omega))} + \|\partial_t u_{0,\lambda}\|_{L^2(\Omega_T)} \right\}, \quad (7.7)$$

where  $\delta = (1 + \sqrt{\lambda})\varepsilon \leq 1$  and  $C$  depends only on  $d, \mu, \Omega$  and  $T$ .

*Proof.* Let  $\psi = w_\varepsilon$  in (7.5), where  $w_\varepsilon$  is given by (3.21). Since  $w_\varepsilon = 0$  on  $\partial_p \Omega_T$ , we see that  $\int_0^T \langle \partial_t w_\varepsilon, w_\varepsilon \rangle \geq 0$ . It follows that  $\|\nabla w_\varepsilon\|_{L^2(\Omega_T)}$  is bounded by the right-hand side of (7.7). It is not hard to show that  $\|\nabla(w_\varepsilon - \tilde{w}_\varepsilon)\|_{L^2(\Omega_T)}$  is also bounded by the right-hand side of (7.7). This gives the inequality (7.7).  $\square$

We now move on to the convergence rate of  $u_{\varepsilon,\lambda} - u_{0,\lambda}$  in  $L^2(\Omega_T)$ .

**Theorem 7.4.** *Suppose  $A$  satisfies (1.3) and (1.4). Let  $\Omega$  be a bounded  $C^{1,1}$  domain in  $\mathbb{R}^d$ . Let  $u_{\varepsilon,\lambda}$  be a weak solution of (7.1) and  $u_{0,\lambda}$  the solution of the homogenized problem (7.2). Then*

$$\|u_{\varepsilon,\lambda} - u_{0,\lambda}\|_{L^2(\Omega_T)} \leq C\delta \left\{ \|u_{0,\lambda}\|_{L^2(0,T;H^2(\Omega))} + \|\partial_t u_{0,\lambda}\|_{L^2(\Omega_T)} \right\}, \quad (7.8)$$

where  $\delta = (1 + \sqrt{\lambda})\varepsilon$  and  $C$  depends only on  $d, \mu, \Omega$  and  $T$ .

*Proof.* In view of Lemma 7.1, this theorem was proved in [14, Theorem 1.1] for the case  $\lambda = 1$ . With Lemma 7.2 at our disposal, the case  $\lambda \neq 1$  follows by a similar duality argument. We omit the details.  $\square$

Finally, we study the problem of convergence rates for the parabolic operator  $\partial_t + \mathcal{L}_\varepsilon$ , where  $\mathcal{L}_\varepsilon = -\operatorname{div}(A(x/\varepsilon, t/\varepsilon^k)\nabla)$  and  $0 < k < \infty$ . Note that the case  $k = 2$  is already treated in Theorems 7.3 and 7.4 with  $\lambda = 1$ .

For the case  $k \neq 2$ , we use the fact that  $\mathcal{L}_\varepsilon = \mathcal{L}_{\varepsilon,\lambda}$  with  $\lambda = \varepsilon^{k-2}$ . Recall that the homogenized operator for  $\partial_t + \mathcal{L}_\varepsilon$  is given by  $\partial_t - \operatorname{div}(\widehat{A}_\infty \nabla)$  for  $0 < k < 2$ , and by  $\partial_t - \operatorname{div}(\widehat{A}_0 \nabla)$  for  $2 < k < \infty$ , where  $\widehat{A}_\infty$  and  $\widehat{A}_0$  are defined in (2.12) and (2.17), respectively.

**Theorem 7.5.** *Assume  $A$  satisfies (1.3) and (1.4). Also assume that  $\|\partial_s A\|_\infty \leq M$ . Let  $0 < k < 2$ . Let  $u_\varepsilon$  be the weak solution of the initial-Dirichlet problem,*

$$\partial_t u_\varepsilon - \operatorname{div}(A(x/\varepsilon, t/\varepsilon^k)\nabla u_\varepsilon) = F \quad \text{in } \Omega_T \quad \text{and} \quad u_\varepsilon = f \quad \text{on } \partial_p \Omega_T, \quad (7.9)$$

where  $\Omega$  is a bounded  $C^{1,1}$  domain in  $\mathbb{R}^d$  and  $0 < T < \infty$ . Let  $u_0$  be the solution of the homogenized problem. Then

$$\|u_\varepsilon - u_0\|_{L^2(\Omega_T)} \leq C(\varepsilon^{k/2} + \varepsilon^{2-k}) \left\{ \|u_0\|_{L^2(0,T;H^2(\Omega))} + \|\partial_t u_0\|_{L^2(\Omega_T)} \right\} \quad (7.10)$$

for  $0 < \varepsilon \leq 1$ , where  $C$  depends only on  $d, \mu, \Omega, T$ , and  $M$ .

*Proof.* Let  $\lambda = \varepsilon^{k-2}$  and  $u_{0,\lambda}$  be the solution of the initial-Dirichlet problem,

$$\partial_t u_{0,\lambda} - \operatorname{div}(\widehat{A}_\lambda \nabla u_{0,\lambda}) = F \quad \text{in } \Omega_T \quad \text{and} \quad u_{0,\lambda} = f \quad \text{on } \partial_p \Omega_T. \quad (7.11)$$

Note that  $(1 + \sqrt{\lambda})\varepsilon = \varepsilon + \varepsilon^{k/2} \leq 2\varepsilon^{k/2}$  for  $0 < \varepsilon \leq 1$ . It follows by Theorem 7.4 that

$$\|u_\varepsilon - u_{0,\lambda}\|_{L^2(\Omega_T)} \leq C\varepsilon^{k/2} \left\{ \|u_{0,\lambda}\|_{L^2(0,T;H^2(\Omega))} + \|\partial_t u_{0,\lambda}\|_{L^2(\Omega_T)} \right\}. \quad (7.12)$$

Next, we observe that  $u_{0,\lambda} - u_0 = 0$  on  $\partial_p \Omega_T$  and

$$\partial_t(u_{0,\lambda} - u_0) - \operatorname{div}(\widehat{A}_\lambda \nabla(u_{0,\lambda} - u_0)) = \operatorname{div}((\widehat{A}_\lambda - \widehat{A}_\infty)\nabla u_0)$$

in  $\Omega_T$ . Since  $\Omega$  is  $C^{1,1}$ , it follows by the standard regularity estimates for parabolic systems with constant coefficients that

$$\begin{aligned} \|\partial_t(u_0 - u_{0,\lambda})\|_{L^2(\Omega_T)} + \|u_0 - u_{0,\lambda}\|_{L^2(0,T;H^2(\Omega))} \\ \leq C|\widehat{A}_\lambda - \widehat{A}_\infty| \|\nabla^2 u_0\|_{L^2(\Omega_T)} \\ \leq C\lambda^{-1} \|\partial_s A\|_\infty \|\nabla^2 u_0\|_{L^2(\Omega_T)}, \end{aligned}$$

where we have used (2.23) for the last step. This, together with (7.12), yields the estimate (7.10).  $\square$

The next theorem treats the case  $2 < k < \infty$ .

**Theorem 7.6.** *Assume  $A$  satisfies (1.3) and (1.4). Also assume that  $\|\nabla^2 A\|_\infty \leq M$ . Let  $2 < k < \infty$ . Let  $u_\varepsilon$  be the weak solution of the initial-Dirichlet problem (7.9), where  $\Omega$  is a bounded  $C^{1,1}$  domain in  $\mathbb{R}^d$  and  $0 < T < \infty$ . Let  $u_0$  be the solution of the homogenized problem. Then*

$$\|u_\varepsilon - u_0\|_{L^2(\Omega_T)} \leq C(\varepsilon + \varepsilon^{k-2}) \left\{ \|u_0\|_{L^2(0,T;H^2(\Omega))} + \|\partial_t u_0\|_{L^2(\Omega_T)} \right\} \tag{7.13}$$

for  $0 < \varepsilon < 1$ , where  $C$  depends only on  $d, \mu, \Omega, T$ , and  $M$ .

*Proof.* The proof is similar to that of Theorem 7.5. The only modification is that in the place of (2.24), we use the estimate (2.29) to bound  $|\widehat{A}_\lambda - \widehat{A}_0|$ . Also, note that  $\|\nabla A\|_\infty$  may be bounded by a constant depending on  $\mu$  and  $M$ . We omit the details.  $\square$

*Proof of Theorem 1.3.* Let  $0 < \varepsilon < 1$ . Note that  $\varepsilon^{2-k} \leq \varepsilon^{k/2}$  if  $0 < k \leq 4/3$ , and  $\varepsilon^{k/2} \leq \varepsilon^{2-k}$  if  $4/3 < k < 2$ . Also,  $\varepsilon \leq \varepsilon^{k-2}$  if  $2 < k < 3$ , and  $\varepsilon^{k-2} \leq \varepsilon$  if  $k \geq 3$ . Thus, by Theorems 7.5 and 7.6,

$$\begin{aligned} \|u_\varepsilon - u_0\|_{L^2(\Omega_T)} \\ \leq C \left\{ \|u_0\|_{L^2(0,T;H^2(\Omega))} + \|\partial_t u_0\|_{L^2(\Omega_T)} \right\} \cdot \begin{cases} \varepsilon^{k/2} & \text{if } 0 < k \leq 4/3, \\ \varepsilon^{2-k} & \text{if } 4/3 < k < 2, \\ \varepsilon^{k-2} & \text{if } 2 < k < 3, \\ \varepsilon & \text{if } k = 2 \text{ or } 3 \leq k < \infty. \end{cases} \end{aligned}$$

*Remark 7.7.* The results on convergence rates in Theorems 7.5 and 7.6 also hold for initial-Neumann problems. The proof is almost identical to the case of the initial-Dirichlet problem. See [14] for the case  $k = 2$ .

Using Theorem 7.3 we may obtain an error estimate in  $L^2(0, T; H^1(\Omega))$  for a two-scale expansion for  $\partial_t + \mathcal{L}_\varepsilon$  in (1.1) in terms of its own correctors. The case  $k = 2$  is contained in Theorem 7.3 with  $\lambda = 1$ . For  $k \neq 2$ , we let

$$v_\varepsilon = \begin{cases} u_\varepsilon - u_0 - \varepsilon \chi^\infty(x/\varepsilon, t/\varepsilon^k) \widetilde{K}_\varepsilon(\nabla u_0) & \text{if } 0 < k < 2, \\ u_\varepsilon - u_0 - \varepsilon \chi^0(x/\varepsilon) \widetilde{K}_\varepsilon(\nabla u_0) & \text{if } 2 < k < \infty. \end{cases} \tag{7.14}$$



In (7.14),  $\chi^\infty$  and  $\chi^0$  are the correctors defined by (2.10) and (2.15), respectively, for  $\partial_t + \mathcal{L}_\varepsilon$ . Since they satisfy the estimates (2.11) and (2.16), only smoothing in the space variable is needed for the operator  $\tilde{K}_\varepsilon$ . More precisely, we let  $\tilde{K}_\varepsilon(f) = S_\delta^1(\eta_\delta f)$ , where

$$S_\delta^1(f)(x, t) = \int_{\mathbb{R}^d} f(x - y, t) \delta^{-d} \theta_1(y/\delta) dy,$$

$\delta = \varepsilon + \varepsilon^{k/2}$ , and the cut-off function  $\eta_\delta$  is the same as in  $K_\varepsilon$ .

**Theorem 7.8.** *Suppose that  $A$  and  $\Omega$  satisfy the same conditions as in Theorem 7.5. Let  $u_\varepsilon$  be the weak solution of (7.9) and  $u_0$  the homogenized solution. Let  $v_\varepsilon$  be given by (7.14). Then*

$$\begin{aligned} & \|\nabla v_\varepsilon\|_{L^2(\Omega_T)} \\ & \leq C \left\{ \|u_0\|_{L^2(0,T;H^2(\Omega))} + \|\partial_t u_0\|_{L^2(\Omega_T)} \right\} \cdot \begin{cases} \varepsilon^{k/4} & \text{if } 0 < k \leq 8/5, \\ \varepsilon^{2-k} & \text{if } 8/5 < k < 2, \\ \varepsilon^{k-2} & \text{if } 2 < k < 5/2, \\ \varepsilon^{1/2} & \text{if } 5/2 \leq k < \infty. \end{cases} \end{aligned} \tag{7.15}$$

*Proof.* The proof uses Theorem 7.3 and the estimates of  $u_{0,\lambda} - u_0$  in the proof of Theorems 7.5 and 7.6, where  $u_{0,\lambda}$  is the solution of (7.11) with  $\lambda = \varepsilon^{2-k}$ .

Let  $\lambda = \varepsilon^{k-2}$ . Suppose  $0 < k < 2$ . In view of (7.7) it suffices to bound

$$I = \|\nabla \left\{ \varepsilon \chi^\lambda(x/\varepsilon, t/\varepsilon^2) K_\varepsilon(\nabla u_{0,\lambda}) - \varepsilon \chi^\infty(x/\varepsilon, t/\varepsilon^k) \tilde{K}_\varepsilon(\nabla u_0) \right\}\|_{L^2(\Omega_T)}.$$

Note that

$$\begin{aligned} I & \leq \left\| \left( \nabla \chi^\lambda(x/\varepsilon, t/\varepsilon^2) - \nabla \chi^\infty(x/\varepsilon, t/\varepsilon^k) \right) K_\varepsilon(\nabla u_{0,\lambda}) \right\|_{L^2(\Omega_T)} \\ & \quad + \left\| \nabla \chi^\infty(x/\varepsilon, t/\varepsilon^k) \left( K_\varepsilon(\nabla u_{0,\lambda}) - \tilde{K}_\varepsilon(\nabla u_0) \right) \right\|_{L^2(\Omega_T)} \\ & \quad + \varepsilon \left\| \chi^\lambda(x/\varepsilon, t/\varepsilon^2) \nabla K_\varepsilon(\nabla u_{0,\lambda}) \right\|_{L^2(\Omega_T)} \\ & \quad + \varepsilon \left\| \chi^\infty(x/\varepsilon, t/\varepsilon^k) \nabla \tilde{K}_\varepsilon(\nabla u_0) \right\|_{L^2(\Omega_T)} \\ & = I_1 + I_2 + I_3 + I_4. \end{aligned}$$

To bound  $I_1$ , we use the inequality (3.17). This gives

$$\begin{aligned} I_1 & \leq C \sup_{(y,s) \in \mathbb{R}^{d+1}} \left( \int_{Q_\delta} |\nabla \chi^\lambda(x/\varepsilon, t/\varepsilon^2) - \nabla \chi^\infty(x/\varepsilon, t/\varepsilon^k)|^2 dx dt \right)^{1/2} \|\nabla u_{0,\lambda}\|_{L^2(\Omega_T)} \\ & \leq C \lambda^{-1} \|\partial_s A\|_\infty \|\nabla u_{0,\lambda}\|_{L^2(\Omega_T)} \\ & \leq C \varepsilon^{2-k} \|\partial_s A\|_\infty \|\nabla u_0\|_{L^2(\Omega_T)}, \end{aligned} \tag{7.16}$$

where we have used (2.27) for the second inequality. To estimate  $I_2$ , we assume that the function  $\theta_1$  is chosen so that  $\theta_1 = \theta_{11} * \theta_{11}$ , where  $\theta_{11} \in C_0^\infty(B(0, 1))$ ,  $\theta_{11} \geq 0$

and  $\int_{\mathbb{R}^d} \theta_{11} = 1$ . This allows us to write  $S_\delta^1 = S_\delta^{11} \circ S_\delta^{11}$ , where  $S_\delta^{11}(f) = f * (\theta_{11})_\delta$ . As a result, we obtain

$$\begin{aligned} I_2 &\leq C \|S_\delta^2 [S_\delta^{11}(\eta_\delta \nabla u_0)] - S_\delta^{11}(\eta_\delta \nabla u_0)\|_{L^2(\Omega_T)} \\ &\leq C \delta^2 \|\partial_t S_\delta^{11}(\eta_\delta \nabla u_0)\|_{L^2(\Omega_T)} \\ &= C \delta^2 \|S_\delta^{11}\{(\partial_t \eta_\delta)(\nabla u_0) + \nabla(\eta_\delta \partial_t u_0) - (\nabla \eta_\delta) \partial_t u_0\}\|_{L^2(\Omega_T)} \\ &\leq C \delta^{1/2} \left\{ \|\nabla u_0\|_{L^2(\Omega_T)} + \|\nabla^2 u_0\|_{L^2(\Omega_T)} + \|\partial_t u_0\|_{L^2(\Omega_T)} \right\}. \end{aligned}$$

It is not hard to see that

$$\begin{aligned} I_3 + I_4 &\leq C \varepsilon \left\{ \|\nabla(\eta_\delta \nabla u_{0,\lambda})\|_{L^2(\Omega_T)} + \|\nabla(\eta_\delta \nabla u_0)\|_{L^2(\Omega_T)} \right\} \\ &\leq C \delta^{1/2} \left\{ \|\nabla u_0\|_{L^2(\Omega_T)} + \|\nabla^2 u_0\|_{L^2(\Omega_T)} \right\}. \end{aligned}$$

In summary, we have proved that

$$\|\nabla v_\varepsilon\|_{L^2(\Omega_T)} \leq C \left\{ \varepsilon^{k/4} + \varepsilon^{2-k} \right\} \left\{ \|u_0\|_{L^2(0,T;H^1(\Omega))} + \|\partial_t u_0\|_{L^2(\Omega_T)} \right\} \tag{7.17}$$

for  $0 < k < 2$ . A similar argument gives

$$\|\nabla v_\varepsilon\|_{L^2(\Omega_T)} \leq C \left\{ \varepsilon^{1/2} + \varepsilon^{k-2} \right\} \left\{ \|u_0\|_{L^2(0,T;H^1(\Omega))} + \|\partial_t u_0\|_{L^2(\Omega_T)} \right\} \tag{7.18}$$

for  $2 < k < \infty$ . The error estimate (7.15) follows readily from (7.17) and (7.18).  $\square$

*Acknowledgements.* Both authors thank the anonymous referees for their very helpful comments, corrections, and suggestions.

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Compliance with ethical standards**

**Conflict of interest** Authors declare that they have no conflict of interest.

### References

1. ALLAIRE, G., BRIANE, M.: Multiscale convergence and reiterated homogenisation. *Proc. R. Soc. Edinb. Sect. A* **126**(2), 297–342, 1996
2. ARMSTRONG, S.N., BORDAS, A., MOURRAT, J.-C.: Quantitative stochastic homogenization and regularity theory of parabolic equations. *Anal. PDE* **11**(8), 1945–2014, 2018
3. ARMSTRONG, S.N., KUUSI, T., MOURRAT, J.-C.: Mesoscopic higher regularity and sub-additivity in elliptic homogenization. *Commun. Math. Phys.* **347**(2), 315–361, 2016
4. ARMSTRONG, S.N., KUUSI, T., MOURRAT, J.-C.: The additive structure of elliptic homogenization. *Invent. Math.* **208**(3), 999–1154, 2017

5. ARMSTRONG, S.N., KUUSI, T., MOURRAT, J.-C.: *Quantitative Stochastic Homogenization and Large-Scale Regularity. Grundlehren der mathematischen Wissenschaften*, vol. 352. Springer, Berlin 2019
6. ARMSTRONG, S.N., MOURRAT, J.-C.: Lipschitz regularity for elliptic equations with random coefficients. *Arch. Ration. Mech. Anal.* **219**(1), 255–348, 2016
7. ARMSTRONG, S.N., SHEN, Z.: Lipschitz estimates in almost-periodic homogenization. *Commun. Pure Appl. Math.* **69**(10), 1882–1923, 2016
8. ARMSTRONG, S.N., SMART, C.: Quantitative stochastic homogenization of convex integral functionals. *Ann. Sci. Éc. Norm. Supér. (4)* **49**(2), 423–481, 2016
9. AVELLANEDA, M., LIN, F.: Compactness methods in the theory of homogenization. *Commun. Pure Appl. Math.* **40**(6), 803–847, 1987
10. BENSOUSSAN, A., LIONS, J.-L., PAPANICOLAOU, G.: *Asymptotic Analysis for Periodic Structures*. AMS Chelsea Publishing, Providence, RI (2011). Corrected reprint of the 1978 original
11. FISCHER, J., OTTO, F.: A higher-order large-scale regularity theory for random elliptic operators. *Commun. Partial Differ. Equ.* **41**(7), 1108–1148, 2016
12. FLODÉN, L., OLSSON, M.: Reiterated homogenization of some linear and nonlinear monotone parabolic operators. *Can. Appl. Math. Q.* **14**(2), 149–183, 2006
13. GENG, J., SHEN, Z.: Uniform regularity estimates in parabolic homogenization. *Indiana Univ. Math. J.* **64**(3), 697–733, 2015
14. GENG, J., SHEN, Z.: Convergence rates in parabolic homogenization with time-dependent periodic coefficients. *J. Funct. Anal.* **272**(5), 2092–2113, 2017
15. GENG, J., SHEN, Z.: Asymptotic expansions of fundamental solutions in parabolic homogenization. *Anal. PDE* (to appear)
16. GLORIA, A., NEUKAMM, S., OTTO, F.: Quantification of ergodicity in stochastic homogenization: optimal bounds via spectral gap on Glauber dynamics. *Invent. Math.* **199**(2), 455–515, 2015
17. GLORIA, A., OTTO, F.: Quantitative results on the corrector equation in stochastic homogenization. *JEMS* **19**(11), 3489–3548, 2017
18. HOLMBOM, A., SVANSTEDT, N., WELLANDER, N.: Multiscale convergence and reiterated homogenization of parabolic problems. *Appl. Math.* **50**(2), 131–151, 2005
19. KENIG, C., LIN, F., SHEN, Z.: Homogenization of elliptic systems with Neumann boundary conditions. *J. Am. Math. Soc.* **26**(4), 901–937, 2013
20. LADYŽENSKAJA, O.A., SOLONNIKOV, V.A., URALCEVA, N.N.: *Linear and Quasilinear Equations of Parabolic Type*. Translated from the Russian by S. Smith. *Translations of Mathematical Monographs*, vol. 23. American Mathematical Society, Providence, RI (1968)
21. LIEBERMAN, G.: *Second Order Parabolic Differential Equations*. World Scientific Publishing Co., Inc, River Edge 1996
22. MESHKOVA, Y.M., SUSLINA, T.A.: Homogenization of solutions of initial boundary value problems for parabolic systems. *Funct. Anal. Appl.* **49**(1), 72–76 (2015). Translation of *Funktional. Anal. i Prilozhen.* **49**(1), 88–93 (2015)
23. NIU, W., XU, Y.: Convergence rates in homogenization of higher-order parabolic systems. *Discrete Contin. Dyn. Syst.* **38**(8), 4203–4229, 2018
24. NIU, W., XU, Y.: A refined convergence result in homogenization of second order parabolic systems. *J. Differ. Equ.* **266**(12), 8294–8319, 2019
25. PASTUKHOVA, S.E., TIKHOMIROV, R.N.: Operator estimates in reiterated and locally periodic homogenization. *Dokl. Akad. Nauk* **415**(3), 304–309, 2007
26. PASTUKHOVA, S.E., TIKHOMIROV, R.N.: Estimates for locally periodic and reiterated homogenization: parabolic equations. *Dokl. Akad. Nauk* **428**(2), 166–170, 2009
27. SHEN, Z.: *Periodic Homogenization of Elliptic Systems. Operator Theory: Advances and Applications. Advances in Partial Differential Equations (Basel)*, vol. 269. Birkhäuser, Cham 2018
28. SUSLINA, T.A.: Homogenization of the Dirichlet problem for elliptic systems:  $L_2$ -operator error estimates. *Mathematika* **59**(2), 463–476, 2013

29. WOUKENG, J.-L.:  $\Sigma$ -convergence and reiterated homogenization of nonlinear parabolic operators. *Commun. Pure Appl. Anal.* **9**(6), 1753–1789, 2010
30. XU, Q., ZHOU, S.: Quantitative estimates in homogenization of parabolic systems of elasticity in Lipschitz cylinders. Preprint

JUN GENG

School of Mathematics and Statistics,  
Lanzhou University,  
Lanzhou,  
People's Republic of China.  
e-mail: gengjun@lzu.edu.cn

and

ZHONGWEI SHEN

Department of Mathematics,  
University of Kentucky,  
Lexington,  
KY,  
40506 USA.  
e-mail: zshen2@uky.edu

(Received March 22, 2019 / Accepted October 29, 2019)

Published online November 8, 2019

© Springer-Verlag GmbH Germany, part of Springer Nature (2019)