Convergence rates on periodic homogenization of p-Laplace type equations

Li Wang
School of Mathematics and Statistics, Lanzhou University, Lanzhou, 710000, China.
Qiang Xu *†
Max Planck Institute for Mathematics in the Sciences, Inselstrasse 22, 04103 Leipzig, Germany
Peihao Zhao
School of Mathematics and Statistics, Lanzhou University, Lanzhou, 710000, China.

December 13, 2018

Abstract

In this paper, we find some error estimates for periodic homogenization of p-Laplace type equations under the same structure assumption on homogenized equations. The main idea is that by adjusting the size of the difference quotient of the correctors to make the convergence rate visible. In order to reach our goal, the corresponding flux corrector with some properties are developed. Meanwhile, the shift-arguments is in fact applied down to ε scale, which leads to a new weighted type inequality for smoothing operator with the weight satisfying Harnack’s inequality in small scales. As a result, it is possible to develop some large-scale estimates. We finally mention that our approach brought in a systematic error (this phenomenon will disappear in linear and non-degenerated cases), which was fortunately a quantity o(ε) here.

Key words. Homogenization; p-Laplace; convergence rates; large-scale estimates

Contents

1 Instruction and main results .......................... 2
  1.1 Motivation and relation to previous works .......... 3
  1.2 Main results and some comments .......................... 6
  1.3 Outline of the proof of Theorem 1.1 .................. 9
  1.4 Notation ........................................ 10
  1.5 Structure of the paper ................................ 10

*Corresponding author.
†Email: qiangxu@mis.mpg.de.
1 Instruction and main results

The purpose of the present paper is mainly to study the error estimates for a class of quasilinear elliptic equations, arising in the periodic homogenization theory. More precisely, let $\Omega \subset \mathbb{R}^d$ be a bounded domain, and consider the following elliptic equations in divergence form depending on a parameter $\varepsilon > 0$,

$$
\begin{aligned}
(D_\varepsilon) \left\{ \begin{array}{l}
\mathcal{L}_\varepsilon u_\varepsilon = -\text{div} A(x/\varepsilon, \nabla u_\varepsilon) = F \\
u_\varepsilon = g
\end{array} \right. \quad \text{in } \Omega, \\
u_\varepsilon = g \quad \text{on } \partial \Omega.
\end{aligned}
$$

Given three constants $\mu_0, \mu_1, \mu_2 > 0$, let us fix a function $A : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ which satisfies the following conditions.

- For every $y, \xi \in \mathbb{R}^d$, we have the periodicity $A(y + z, \xi) = A(y, \xi)$ for $z \in \mathbb{Z}^d$, and $A(y, \cdot)$ is homogeneous with respect to the second variable, i.e.,

$$
A(y, t\xi) = t^{p-1}A(y, \xi) \quad \text{for any } t \geq 0. \quad (1.1)
$$

- For any $y \in \mathbb{R}^d$, and $\xi, \xi' \in \mathbb{R}^d$, we impose the following coerciveness and growth conditions:

$$
\mu_0(|\xi| + |\xi'|)^{p-2}|\xi - \xi'|^2 \leq \langle A(y, \xi) - A(y, \xi'), \xi - \xi' \rangle \leq \mu_1(|\xi| + |\xi'|)^{p-2}|\xi - \xi'|^2. \quad (1.2)
$$

- The smoothness condition is also imposed, and there holds

$$
|A(y, \xi) - A(y', \xi)| \leq \mu_2 |y - y'| |\xi|^{p-1} \quad (1.3)
$$

for all $y, y' \in \mathbb{R}^d$ and $\xi \in \mathbb{R}^d$. 

As a class of the examples, one may consider the weighted p-Laplace equations such as \( A(y, \xi) = a(y)|\xi|^{p-2}\xi \) (see [24, pp.225]) or the model \( A_i(y, \xi) = a_{ij}(y)|\xi|^{p-2}\xi_j \) appeared in [4, pp.117]. We also mention that the Lipschitz continuity in (1.3) can be weaken into a Hölder one, while the present condition may highlight other key factors in the theory.

In terms of qualitative homogenization, our assumptions are merely special cases considered in [7, 8, 12], so the process of homogenization of (D\(_\varepsilon\)) may be understood in their way, i.e., in the sense of G-convergence. Let \( F \in W^{1,p'}(\Omega) \) with \( 1/p + 1/p' = 1 \) and assume \( u_\varepsilon \) is a weak solution to problem (D\(_\varepsilon\)). It is known that \( u_\varepsilon \rightharpoonup u_0 \) weakly in \( W^{1,p}(\Omega) \), and \( A(x, \varepsilon, \nabla u_\varepsilon) \rightharpoonup \hat{A}(\nabla u_0) \) weakly in \( L^p(\Omega; \mathbb{R}^d) \), as \( \varepsilon \to 0 \), where \( u_0 \) is the solution to the effective (homogenized) equation

\[
(D_0) \quad \begin{cases} \mathcal{L}_0 u_0 \equiv -\text{div}\hat{A}(\nabla u_0) = F & \text{in } \Omega, \\ u_0 = g & \text{on } \partial\Omega. \end{cases}
\]

The function \( \hat{A} : \mathbb{R}^d \to \mathbb{R}^d \) is defined for every \( \xi \in \mathbb{R}^d \) by

\[
\hat{A}(\xi) = \int_{Y=(-\frac{1}{2}, \frac{1}{2})^d} A(y, \xi + \nabla y N(y, \xi)) dy,
\]

where \( N(y, \xi) \) is the so-called corrector, satisfying the cell problem

\[
\begin{cases} \text{div}A(y, \xi + \nabla y N(y, \xi)) = 0 & \text{in } Y, \\ N(\cdot, \xi) \in W^{1,p}_{\text{per}}(Y), \quad \int_Y N(\cdot, \xi) = 0, \end{cases}
\]

where the definition \( W^{1,p}_{\text{per}}(Y) \) may be found in [8, 19], and \( \int_Y \) is an average integral, defined in Subsection 1.4. The existence theory may be found in [30, Theorem 26.A].

Concerning the quantitative estimates for homogenization problems, we hope the effective operator admits the same structure as the assumption (1.2),

\[
\bar{\mu}_0|\xi - \xi'|^2(|\xi| + |\xi'|)^{p-2} \leq \langle \hat{A}(\xi) - \hat{A}(\xi'), \xi - \xi' \rangle \leq \tilde{\mu}_1(|\xi| + |\xi'|)^{p-2}|\xi - \xi'|^2
\]

for any \( \xi, \xi' \in \mathbb{R}^d \), where \( \bar{\mu}_0, \tilde{\mu}_1 \) may depend on \( \mu_0, \mu_1, p \) and \( d \), which means that the homogenized operator \( \mathcal{L}_0 \) belongs to an analogy type of the operators as \( \mathcal{L}_\varepsilon \) does, and it is clear to be true in the special case \( p = 2 \) (see for example [1, 4, 5, 29]). Under the assumptions (1.1),(1.2), one just verified the growth condition of (1.7) for the case \( 2 \leq p < \infty \), and the coerciveness of (1.7) for the case \( 1 < p \leq 2 \) (see Lemma 2.4). However, there are some examples (see Section 5) to illustrate the coerciveness of (1.7) for \( 2 < p < \infty \), and the growth condition for the case \( 1 < p < 2 \).

On the other hand, it is more or little known that the quantitative estimates in homogenization theory relies on a higher regularity of the weak solution to the effective equation (a similar statement may be found in [19, pp.477]). Therefore, the purpose of this hypothesis is to ensure that the weak solution to (D\(_0\)) owns a reasonable regularity (see Lemma 3.8), otherwise we have to directly assume the effective solution \( u_0 \) admits these regularities.

### 1.1 Motivation and relation to previous works

Before recounting the main results, it is better to explain our initial ideas and source of them. We begin the introduction from the two-scale expansion argument (see [4, 19]), which is

\[
u_\varepsilon = u_0(x) + \varepsilon u_1(x, y) + \varepsilon^2 u_2(x, y) + \cdots, \quad y = x/\varepsilon.
\]
By a formal computation, we derive that
\[ \text{div}_y A(y, \nabla u_0 + \nabla_y u_1) = 0 \] (E₁) \[ \text{div}_y A(y, \nabla u_1 + \nabla_y u_2) = F - \text{div}_x A(y, \nabla u_0 + \nabla_y u_1), \]
while the solvability of the second equation implies
\[ \text{div}_x \int_y A(y, \nabla u_0 + \nabla_y u_1)dy = F. \] (E₂)

Obviously, the equations (E₁) and (E₂) inspired people to build the equation (1.6) and the formula (1.5), respectively. Thus, the image of \( \nabla u_0 \) determines \( u_1 \) and the cardinal number of the set \( \{ \nabla u_0(x) : x \in \Omega \} \subset \mathbb{R}^d \) reveals how many equations like (1.6) need to be solved, if we want the formula (1.5) to be fully understood. In this sense, one may quickly realize the benefit of the linear equation, which tells us \( u_1 \) is linearly dependent of \( \nabla u_0 \), and therefore we just solve a finite number of equations that the correctors satisfy.

Hence, from the view of reducing the number of equations like (1.6), we find that the condition (5.1) plays a similar role. And more importantly, it also works for nonlinear cases, which eventually help us to construct some examples to verify the assumption (1.7) (see Section 5). Also, this view might be used to explain the motivation of introducing numerical correctors. We refer the reader to [10, 14, 8] and their references therein for this direction.

On the other hand, the non-degenerated case \( (p = 2) \) is very special. Although it does little to decrease the number of the equations that the corrector satisfies, it has been proved that \( u_1 \) is in fact Lipschitz continuous with respect to \( \nabla u_0 \) (see for example [29, Lemma 2.1]). Thus, we may consider the quantities \( w_\varepsilon = u_\varepsilon - v \) with \( v = u_0 + \varepsilon N(y, \varphi) \), and
\[ \nabla w_\varepsilon = \nabla u_\varepsilon - \nabla v, \]
where \( \varphi \in H^1_0(\Omega; \mathbb{R}^d) \) may be fixed later. As in the linear case, one may derive a convergence rate from estimating the quantity \( \| \nabla w_\varepsilon \|_{L^2(\Omega)} \) by energy methods (see [29, Theorem 1.1]), in which we believe that the formula below reveals an important information of convergence rates for elliptic equations with the divergence structure,
\[ \| \nabla w_\varepsilon \|_{L^2(\Omega)}^2 \leq \left| \int_{\Omega} \left[ \tilde{A}(\nabla u_0) - \tilde{A}(\varphi) \right. \right. \\
\left. \left. + \tilde{A}(\varphi) - A(x/\varepsilon, \varphi + \nabla_y N(x/\varepsilon, \varphi)) \right. \right. \\
\left. \left. + A(x/\varepsilon, \varphi + \nabla_y N(x/\varepsilon, \varphi)) - A(x/\varepsilon, \nabla v) \right] \cdot \nabla w_\varepsilon dx \right| \]
(1.8)
(where \( y = x/\varepsilon \), and see [29, Lemma 3.1]). So, in order to better describe our ideas, we would like to introduce the relevant calculations on the right-hand side of (1.8). In general, the computations of \( T_3 \) is not complicated as it appears, which can be reduced to those of the other two terms. As we have claimed above, the term \( T_1 \) is in fact related to the following estimate
\[ \| S_\varepsilon(\varphi) - \varphi \|_{L^2(\Sigma_2)} \leq C\varepsilon \| \nabla \varphi \|_{L^2(\Sigma_\varepsilon)} \]
(1.9)
(a key inequality in the shift argument), where $C$ depends only on $d$ (see Lemma 2.14), and the set $\Sigma_r$ is referred to as the “co-layer” part of $\Omega$ (see Subsection 1.4). If the smoothing operator $S_\epsilon$ (see Definition 1) is replaced by Steklov averaging operator, the reader may clearly find the relationship with a shift argument developed by V. Zhikov and S. Pastukhova (see [31, 32]), while applying smoothing operator $S_\epsilon$ to error estimates was first suggested by Z. Shen in [26], and this improvement proved to be more flexibility in analysis. Moreover, let $\varphi = \psi_r \nabla u_0$ in (1.9), where $\psi_r \in C^1_0(\Omega)$ is a cut-off function defined in (1.27). Then, the computations on $T_1$ is consequently translated into the layer type and co-layer type estimates, which means we need to estimate

$$\|\nabla u_0\|_{L^2(\Omega \setminus \Sigma_r)} \quad \text{and} \quad \epsilon \|\nabla^2 u_0\|_{L^2(\Sigma_r)}$$

for the homogenized solution to (D_0), respectively. If only a small rate of convergence is observed, then Meyer’s estimate coupled with some interior estimates for higher derivatives is sufficient to serve this purpose, which in fact will benefit a large scale estimate later on. These results have been systematically stated in [25, 26] for linear systems, as well as in [29] for a non-degenerated equation ($p = 2$). Then we proceed to show two points in the calculations of $T_2$ in (1.8), which are all based upon the cell problem (1.6). The first one derived from (1.5) is $\int_Y T_2 dy = 0$, and the second one is $\text{div}_y(T_2) = 0$. Using both of them, it is not hard to construct the so-called flux corrector, and a later calculation will benefit from its antisymmetric property. Interestingly, this construction does not depend on whether the original model we studied is linear or not.

So far, we have mainly introduced the methods used in non-degenerated equations ($p = 2$), as well as, our views on correctors and error estimates. Thus, in terms of the nonlinear problem (D_\epsilon), a natural question is how to extend the above theory established for the special case $p = 2$ to more general cases $1 < p < \infty$.

The first challenge we confront with is that the first order corrector $u_1$ is only Hölder continuous with respect to $\nabla u_0$ (i.e., the corrector $N(y, \cdot)$ is uniformly Hölder continuous with respect to the second variable for a.e. $y \in \mathbb{R}^d$, and see Lemma 2.1). The idea is that we use the difference quotient as the substitute for the derivatives of the corrector, and the first-order approximating corrector is constructed in the form of

$$V(x) = \nabla u_0(x) + D^h(\varepsilon N(\cdot/\varepsilon, \varphi))(x),$$

(1.10)

where $D^h = (D_{1}^h, \ldots, D_{d}^h)$ is a difference quotient of size $h$ (see Subsection 1.4). Then we consider the quantity

$$\nabla u_\epsilon - V,$$

which will play a similar role as the quantity $\nabla w_\epsilon$ did in (1.8). Apparently, the main differences are caused by the replacement of the derivative by the difference quotient. We are also aware of the following fact: for any $x \in \Sigma_\epsilon$, there holds

$$D_i^h(f)(x) = \nabla_i f(x) + o(1),$$

(1.11)

as $h$ goes to zero, provided $f \in C^1(\Omega)$. As usual, the calculations like “$o(1) \times \varepsilon = o(\varepsilon)$” and “$o(1) \times C = o(1)$ for a constant $C$” are agreed to be true, and $o(\varepsilon)$ is referred as to the higher order infinitesimal of $\varepsilon$, as $\varepsilon$ goes to zero. Therefore, our main results may be shown by ignoring the quantity $o(\varepsilon)$, which may be regarded as a kind of systematic error. Thus, the main problem is to find the accuracy of $h$ in the formula (1.10).
1.2 Main results and some comments

For the ease of the statement, we impose the following index throughout the paper,

$$
\alpha = \begin{cases} 
1/(3-p), & 1 < p \leq 2; \\
2/p, & 2 < p < \infty,
\end{cases} \quad
\beta = 1/p - 1/q, \quad
\gamma = \begin{cases} 
(1 - \theta/(3-p)), & 1 < p \leq 2; \\
2/p, & 2 \leq p < \infty,
\end{cases}
$$

(1.12)

where \( p < q \leq p(1+\delta) \) and \( \delta \in (0,1) \) is usually very small, which is actually determined by Meyer’s estimates (see Remark 4.2). Here \( \beta \in (0, \frac{\delta}{1+\delta}] \), and \( \alpha, \gamma \in (0,1) \) are used to describe the Hölder’s continuity of \( N(y, \cdot) \) and the corresponding flux corrector \( E(y, \cdot) \) with respect to the second variable, respectively. We now state the main results of the paper.

**Theorem 1.1** (convergence rates). Let \( \Omega \subset \mathbb{R}^d \) be a bounded Lipschitz domain with \( r_0 = \text{diam}(\Omega) \), and \( 0 < \varepsilon \ll r_0 \). Assume that \( A \) satisfies the conditions (1.1), (1.2), (1.3). Then there exist \( \Theta \in (0,1) \) and a constant \( C \), depending on \( \mu_1, \mu_2, \mu_3, p \) and \( d \), such that

$$
\| \nabla N(\cdot, \xi) \|_{L^\infty(Y)} + [\nabla N(\cdot, \xi)]_{C^{\alpha, \theta}(Y)} \leq C \left( \int_Y |N(\cdot, \xi)|^p \right)^{1/p} \leq C|\xi|
$$

(1.13)

holds for any \( \xi \in \mathbb{R}^d \). Moreover, suppose that \( \hat{A} \) satisfies the condition (1.7), and \( u_\varepsilon, u_0 \in W^{1,p}(\Omega) \) satisfy \( L_\varepsilon u_\varepsilon = L_0 u_0 = 0 \) in \( \Omega \) and \( u_\varepsilon = u_0 \) on \( \partial \Omega \), with assumption that

$$
M := \| \nabla u_0 \|_{C^{\alpha, \theta}(\Omega)} \quad \text{with some } \Theta \in (0,1), \quad \Omega_{M,\varepsilon} = \{ x \in \Omega : |\nabla u_0| \leq M \varepsilon^\Theta \}.
$$

(1.14)

Let \( \varphi_\varepsilon = S_\varepsilon(\psi_\varepsilon \nabla u_0) \), and the first-order approximating corrector is in the form of (1.10), i.e.,

$$
V(x) = \nabla u_0(x) + D^h(\varepsilon N(\cdot, \varepsilon, \varphi_\varepsilon))(x),
$$

where we prefer \( h = \varepsilon^\tau \) and \( 1 < \tau < 1/(1 - \alpha) \) in advance, and \( \psi_\varepsilon \in C^\alpha_0(\Omega) \) is cut-off function satisfying (1.27). Then, by ignoring a systematic error \( o(\varepsilon) \), we have the following conclusions.

- For the case \( 1 < p < 2 \), one may derive that

$$
\| \nabla u_\varepsilon - V \|_{L^p(\Omega)} \lesssim |\Omega_{M,\varepsilon}|^\beta \| \nabla u_0 \|_{L^q(\Omega_{M,\varepsilon})} + \varepsilon^{\theta(1-\tau)} \| \nabla u_0 \|_{L^q(\Omega)}.
$$

(1.15)

where the range of \( \tau \) is determined by

$$
1 < \tau < \frac{1 - \gamma + \theta(p-1) + \gamma \beta}{1 - \gamma + \theta(p-1)}.
$$

(1.16)

If we further consider the connection to the non-degenerated case \( (p = 2) \), there holds

$$
\| \nabla u_\varepsilon - V \|_{L^p(\Omega)} \lesssim (1 - I_{\{p=2\}}) |\Omega_{M,\varepsilon}|^\beta \| \nabla u_0 \|_{L^q(\Omega_{M,\varepsilon})} + \varepsilon^{\alpha \beta + \gamma(1-\gamma)/(p-1)} \| \nabla u_0 \|_{L^q(\Omega)}
$$

(1.17)

for any \( 1 < p \leq 2 \), where \( I_{\{p=2\}} \) denotes the character function of the set \( \{ p = 2 \} \), and it requires the range of \( \tau \) to be

$$
\frac{1 - \gamma + \theta(p-1) + \gamma \beta}{1 - \gamma + \theta(p-1)} < \tau < \frac{1 - \gamma + \gamma \beta}{1 - \gamma}.
$$

(1.18)
If $2 < p < \infty$, then we have
\[
\|\nabla u_\varepsilon - V\|_{L^p(\Omega)} \lesssim |\Omega_{M,\varepsilon}|^\beta \|\nabla u_0\|_{L^q(\Omega_{M,\varepsilon})} + \begin{cases} 
\varepsilon^{(\frac{p-2}{p-1})[\beta+(1-\alpha)(1-\tau)]} \|\nabla u_0\|_{L^q(\Omega)} & \text{if } 2 < p < 3; \\
\varepsilon^{(\frac{p-3}{p-1})[\beta+(1-\alpha)(1-\tau)]} \|\nabla u_0\|_{L^q(\Omega)} & \text{if } 3 \leq p < \infty,
\end{cases}
\]
when $\tau$ is chosen such that
\[
\frac{1 - \alpha + \theta + \beta}{1 - \alpha + \theta} < \tau < \frac{1 - \alpha + \beta}{1 - \alpha}.
\]

Here $\lesssim$ means $\leq$ up to a multiplicative constant which depends on $\mu_0, \mu_1, \mu_2, \tilde{\mu}_0, \tilde{\mu}_1, p, q, d$ and $r_0$.

We really hope that the theorem can be written succinctly, and as a compensation, we make an effort to explain as much as possible, and some comments are as follows.

Comment 1. Let us begin from the explanation of the estimate (1.17). For the special case $p = 2$, it will become
\[
\|\nabla u_\varepsilon - V\|_{L^2(\Omega)} \lesssim \varepsilon^\beta \|\nabla u_0\|_{L^q(\Omega)}
\]
with any $\tau > (1 + \beta/\theta)$, where $q = 2(1 + \delta)$. As mentioned before the quantity $\nabla_x (\varepsilon N(x/\varepsilon, \varphi))$ exists in such the case, and so the above estimate implies
\[
\|\nabla w_\varepsilon\|_{L^2(\Omega)} \lesssim \varepsilon^\beta \|\nabla u_0\|_{L^2(\Omega)},
\]
which is exactly what we have proved in [29, Theorem 1.1]. So, this theorem may be roughly regarded as an extension of the previous one, and our purpose is also to study a uniform estimate for the solution of $(D_\varepsilon)$ through the convergence rates (see Theorem 1.2). For this consideration the presence of the $q$-energy norm (i.e., $\|\nabla u_0\|_{L^q(\Omega)}$ which is a little stronger than $p$-energy norm) in the right-hand side of error estimates is quite convenient to us even when a higher rate of convergence could be derived by a stronger norms.

Comment 2. The first term in the right-hand side of the estimates (1.17), (1.15) and (1.19) has two possible directions:
\[
|\Omega_{M,\varepsilon}|^\beta \|\nabla u_0\|_{L^p(\Omega_{M,\varepsilon})} \leq \begin{cases} 
|\Omega_{M,\varepsilon}|^\beta \|\nabla u_0\|_{L^p(\Omega)}; \\
|\Omega_{M,\varepsilon}|^{\beta + \frac{1}{p}} M^\theta \varepsilon^\beta.
\end{cases}
\]
Although the first line of (1.21) does not clearly show any convergence rate, it is proved to be useful and enlightening for investigating a large scale estimate (see Lemma 4.3 and Theorem 1.2). On the contrary, the second line of (1.21) gives the rate of convergence on account of the assumption (1.14). However, it actually indicates that the process of homogenization are not truly observable when fluctuations are limited in a small scale.

Comment 3. Two reasons leads to the assumption (1.14). On the one hand, the $C^{1,\theta}$ smoothness imposed to the effective solution is reasonable under the precondition (1.7) (although it usually admits such the regularity in a local sense), which ensures that the calculation like (1.11) conforms to the rule and ultimately produces a quantity $o(\varepsilon)$. On the other hand, the set $\Omega_{M,\varepsilon} = \{x \in \Omega : |\nabla u_0(x)| > M\varepsilon^\theta\}$ is in fact more important to us, since one may acquire a Harnack’s inequality
\[
\sup_{Y_k^\varepsilon} |\nabla u_0| \leq 2 \inf_{Y_k^\varepsilon} |\nabla u_0|.
\]
for any $Y^k_\varepsilon \subset \Omega^k_{M,\varepsilon}$ with $|Y^k_\varepsilon| = \varepsilon^d$ (where $\Omega^k_{M,\varepsilon}$ is an open set, and see Subsection 1.4 for $Y^k$), and this together with the estimate (1.9) implies the following weighted inequality

$$
\left( \int_{\Omega^k_{M,\varepsilon}} |\varphi - S_\varepsilon(\varphi)|^p |\nabla u_0|^r dx \right)^{1/p} \lesssim \varepsilon \left( \int_{\Omega} |\nabla \varphi|^p |\nabla u_0|^r dx \right)^{1/p}
$$

(1.22)

for any $1 < p, r < \infty$, where $|\nabla u_0|$ is referred to as a weight (see Remark 2.15). If familiar with the regularity theory of p-Laplace type equations, one may immediately realize that the estimate (1.22) is in help of giving the quantity of the left-hand side below,

$$
\int_{\Sigma_{2r}} |\nabla u_0|^{p-2} |\nabla^2 u_0|^2 dx \lesssim \frac{1}{r^2} \int_{\Sigma_r} |\nabla u_0|^p dx \quad \text{for } p > 2,
$$

$$
\int_{\Sigma_{2r}} |\nabla^2 u_0|^p dx \lesssim \frac{1}{r^p} \int_{\Sigma_r} |\nabla u_0|^p dx \quad \text{for } 1 < p \leq 2,
$$

(see Lemma 3.8). Thanks to them, the stated error estimates (1.17), (1.15) and (1.19) do not involve the higher derivatives any more. The above computations serve as a counterpart of those for $T_1$ in (1.8), which will be addressed in Lemmas 3.3, 3.9. Besides, the set $\Omega_{M,\varepsilon}$ includes the degenerate information of the homogenized equation $(D_0)$, and it seems that this kind of information can not be known in advance through the homogenization of the equation $(D_\varepsilon)$, which eventually leads to the separated two parts of the right-hand side in these error estimates.

Comment 4. In both periodic and aperiodic settings, error estimates have always been a hot topic in homogenization theory (see for example [1, 2, 10, 14, 16, 17, 21, 24, 25, 27, 28, 29, 31, 32] and their references therein for more details), while there is few contributions in this field concerning p-Laplace type equations, to the authors’ best acknowledge. In fact, there seems to be an interesting problem: except of the special cases in Section 5, what kind of conditions can guarantee the homogenized operator $L_0$ admits the structure (1.7).

As an application of Theorem 1.1, we derive the following result.

Theorem 1.2 (large-scale Hölder estimates). Let $B = B(0, 1) \subset \Omega$ and $1 < p < \infty$. Assume the same conditions as in Theorem 1.1. Suppose that $u_\varepsilon \in W^{1,p}(B)$ satisfies $L_0 u_\varepsilon = 0$ in $B$. Then for any $\kappa \in (0, d)$, there holds an uniform estimate

$$
\int_{B(x,r)} |\nabla u_\varepsilon|^p dz \lesssim \left( \frac{r}{R} \right)^{d-\kappa} \int_{B(x,R)} |\nabla u_\varepsilon|^p dz
$$

(1.23)

for any $\varepsilon \leq r \leq R \leq (1/2)$, where the up-to constant depends on $\mu_1, \mu_2, \mu_3, p, d$ and $\kappa$.

Comment 5. Recently, the large-scale estimates received important developments in quantitative homogenization theory, and without attempting to be exhaustive we refer the reader to [1, 2, 15, 20, 26] for more details in periodic and non-periodic settings, while the first systematic study in this area went back to M. Avellaneda, F. Lin [3]. Concerning the equation $(D_\varepsilon)$, we believe that the estimate (1.23) is still true for $\kappa = 0$ (i.e., large-scale Lipschitz estimates), and may further develop a large-scale $L^q$ estimate with $q > p$ in an average sense. However, we plan to investigate this topic in a separated work. Since the method developed for large-scale estimates are based upon an iteration argument, it is a kind of nonlinear approaches, which in fact is independent of the structure of equations. Roughly speaking, the proof of Theorem 1.2 follows the same strategy introduced in the previous work [29] (see [25, 26] for linear systems). Here, we just mention that the crucial step between Theorems 1.1 and 1.2 is based upon the following corollary of Theorem 1.1, which will be concretely known as an approximating lemma (see Lemma 4.3).
Corollary 1.3. Assume the same conditions as in Theorem 1.1. If the left-hand side of the estimates (1.15), (1.17) and (1.19) is replaced by $\|u_\varepsilon - u_0\|_{L^p(\Omega)}$, respectively. Then, we have the same conclusions.

Comment 6. Here we just mention that Corollary 1.3 cannot be derived from Poincaré’s inequality. Instead, it follows from the fact $\|u_\varepsilon - u_0\|_{L^p(\Omega)} = \|\nabla_i(u_\varepsilon - u_0)\|_{W^{-1,p}(\Omega)}$ with $i = 1, \ldots, d$, and the related details have been shown in the proof of Lemma 4.3.

1.3 Outline of the proof of Theorem 1.1

In this subsection, we show a strategy of the proof, which consists of four ingredients in order.

Ingredient 1. We manage to obtain the following formula, which is the key ingredient, analogy to the estimate (1.8) for the non-degenerated case.

$$
\|\nabla u_\varepsilon - V\|_{L^p(\Omega)}^p \leq \left| \int_\Omega \left[ \begin{array}{l}
\hat{A}(\nabla u_0) - \hat{A}(\varphi) \\
+ \hat{A}(\varphi) - A(x/\varepsilon, \varphi + \nabla_y N(x/\varepsilon, \varphi)) \\
+ A(x/\varepsilon, \varphi + \nabla_y N(x/\varepsilon, \varphi)) - A(x/\varepsilon, V) \end{array} \right] \cdot (\nabla u_\varepsilon - V) \, dx \right| + o(\varepsilon)
$$

(1.24)

(where $y = x/\varepsilon$).

Ingredient 2. Estimating the term $P_1$ forced us to find the weighted type estimate (1.22), which means that the shift argument still works for nonlinear equations, at least in the small scales.

Ingredient 3. The core idea comes from dealing with the term $P_3$ of (1.24), and we realized the importance of the following equality

$$
D_i^h(\varepsilon N(y, \varphi)) = \int_0^1 \partial_i N(y + t(h/\varepsilon)e_i, \varphi) \, dt + \varepsilon \alpha^{\alpha - 1} \Phi(y, \varphi) \left| D_i^h \varphi \right|^\alpha, \quad (1.25)
$$

and it inspires us to discover that there is an opportunity to observe the rate of convergence by adjusting the size of $h$. A natural thinking is to set $h = \varepsilon^\tau$, and reduce the problem to find out a suitable range of $\tau$, independent of $\varepsilon$. Note that we have $R_1 = 0$ by periodicity of the corrector when $\tau = 1$. Since we have to estimate the quantity

$$
\|D_i^h(\varepsilon N(y, \varphi)) - \nabla_y N(y, \varphi)\|_{L^p(\Omega)}
$$

produced by the term $P_3$ in (1.24), the higher regularity of the corrector turns to be crucial in these computations, and it is one of roles that the imposed assumption (1.3) plays. Meanwhile, the related estimate of $\Phi$ in $R_2$ leads to most of complicated parts in the paper.

Ingredient 4. Concerning $P_2$ in (1.24), we also employ the antisymmetry property of the flux corrector (see Lemma 3.4). The difference quotient presented here brings us more technical difficulties compared to the computation of $T_2$ in (1.8) (see Lemma 3.4). We finally mention that those arguments developed for the corrector in (1.25) and (1.26) also worked for the flux corrector, although a little more efforts and careforness have to be paid.
1.4 Notation

Throughout the paper, we make use of the following notation:

- \( d \geq 1 \) is the dimension, and \( 1 < p < \infty \) is the characterization of the equation type, and \( \alpha, \beta, \gamma \) are defined in (1.12), and \( \delta \in (0,1) \) is known as a Meyer’s index;

- \( \Omega \subset \mathbb{R}^d \) is a bounded Lipschitz domain with \( r_0 \) being the diameter of \( \Omega \), and \( \Sigma_r = \{ x \in \Omega : \text{dist}(x, \partial \Omega) > r \} \) represents the so-called “co-layer” part of \( \Omega \), and its layer part is denoted by \( \Omega \setminus \Sigma_r \), and \( \psi_r \in C_0^1(\Omega) \) is a cut-off function associated with \( \Sigma_r \), satisfying
  \[
  \psi_r = 1 \quad \text{in} \quad \Sigma_r, \quad \psi_r = 0 \quad \text{outside} \quad \Sigma_r, \quad \| \nabla \psi_r \| \leq C/r; \quad (1.27)
  \]

- \( D^h = (D^h_1, \ldots, D^h_d) \), where the \( i \)-th difference quotient of size \( h \) is defined by \( D^h_i(f)(x) = \dfrac{f(x+he_i)-f(x)}{h} \) for any \( x \in \Sigma_r \) with \( 0 < h < \varepsilon \);

- \( \nabla = (\nabla_1, \ldots, \nabla_d) \) with \( \nabla_i = \nabla x_i = \partial / \partial x_i \), and \( \nabla_y \) (or div \( y \)) means we take a gradient (or divergence) with respective to the variable \( y \), when we need to emphasize that it is different from the default case \( \nabla \). The notation \( \text{div}(F) = \nabla \cdot F = \sum_{i=1}^d \nabla_i F_i \), and \( \nabla^2 f = \nabla^2_{ij} f = \dfrac{\partial^2 f}{\partial x_i \partial x_j} \) denotes the divergence of \( F \), and Hessian matrix of \( f \), respectively. In general, we denote \( \nabla_z N(x/\varepsilon + \omega, \varphi) \) by \( \partial_1 N(x/\varepsilon + \omega, \varphi) \) without confusion, where \( z = x/\varepsilon + \omega \);

- \( o(1) \) represents an infinitesimal quantity, and \( o(\varepsilon) \) is a higher order infinitesimal of \( \varepsilon \) as \( \varepsilon \to 0 \);

- \( \langle A, B \rangle = \sum_{i=1}^d A_i B_i \), and \( |\xi| = \langle \xi, \xi \rangle^{\frac{1}{2}} \), and \( |\Omega| \) represents the volume of \( \Omega \), and \( \int_{\Omega} = \frac{1}{|\Omega|} \int_{\Omega} \); and \( Y^k = \varepsilon(k + Y) \) for \( k \in \mathbb{Z}^d \), where \( Y = (-\frac{1}{2}, \frac{1}{2}], 2 \]

- \( \lesssim \) and \( \gtrsim \) stand for \( \leq \) and \( \geq \) up to a multiplicative constant which may depend on any given constants such as \( \mu_0, \mu_1, \mu_2, d, p, q, \theta, \vartheta, \kappa, \delta, \zeta_0 \), but independent of \( \varepsilon \), and rescaling parameters such as \( r, R \in (0,1) \), and we use \( \ll \) instead of \( \lesssim \) to indicate that multiplicative constant is quite close to zero (although it is still a positive number).

1.5 Structure of the paper

In Section 2, we mainly addressed the correctors and flux correctors. Concerned with the properties of the corrector \( N(\cdot, \xi) \), we verified that \( N(\cdot, \xi) \in C^{1,\theta} \times C^\alpha \), and in particular, gave a proof of (1.13). As an application, the condition (1.7) was partially proved. Some new tricks were dedicated to estimating the quantity \( \| D^h(\varepsilon N(\cdot, \xi, \varphi)) \|_{L^p(\Omega)} \). In terms of the flux corrector \( E(\cdot, \xi) \), we similarly proved that \( E(\cdot, \xi) \in C^{1,\rho} \times C^\gamma \), where \( \rho = \min\{ \theta, (p - 1)\theta \} \), while we merely demonstrated \( D^h(\varepsilon E(\cdot, \xi, \varphi)) \in L^p(\Omega) \) with \( 1 < r \leq p/(p-1) \). In the end of this section, we defined the smoothing operator and stated its properties. In Section 3, the purpose was to show a proof of Theorem 1.1. The structure of this section was in fact consistent with that shown in Subsection 1.3. In Section 4, a main work was to show the so-called approximating lemma (see Lemma 4.3), and then we gave a proof of Theorem 1.2, as well as, Caccioppolli’s inequality (see Lemma 4.1). Finally, in Section 5 we constructed some special examples to satisfy the assumption (1.7).

The interest of the present paper is finding a new way (independent of numerical correctors) to obtain the rate of convergence for nonlinear homogenization problems, which could be further applied to studying large scale estimates. As a theory, it is more or less self-enclosed as we hoped. However, it can not directly help us understand the mechanism of homogenization for nonlinear problems in a little larger classes of operators, since this method does not tell us why the homogenized operator runs out of the class (1.2) in most cases, and we will make an effort for this direction in future.
2 Preliminaries

2.1 Corrector and its properties

Lemma 2.1. Assume that $A$ satisfies the conditions (1.1), (1.2). Let $N(\cdot, \xi) \in W^{1,p}_{\text{per}}(Y)$ be the weak solution to the equation (1.6), and then for any $\xi \in \mathbb{R}^d$, we have the following estimate

$$\int_Y |N(\cdot, \xi)|^p + \int_Y |\nabla N(\cdot, \xi)|^p \leq C|\xi|^p. \quad (2.1)$$

Moreover, by setting

$$P(y, \xi) = \xi + \nabla_y N(y, \xi) \quad (2.2)$$

one may derive

$$\left(\int_Y |P(\cdot, \xi) - P(\cdot, \xi')|^p\right)^{\frac{1}{p}} \leq C \begin{cases} |\xi - \xi'|^\frac{2}{p} \left(|\xi| + |\xi'|\right)^{1-\frac{2}{p}} & \text{if } 2 \leq p < \infty; \\ |\xi - \xi'|^\frac{2}{p} \left(|\xi| + |\xi'|\right)^{\frac{2-p}{p}} & \text{if } 1 < p \leq 2, \end{cases} \quad (2.3)$$

where $C$ depends only on $\mu_0, \mu_1, p$ and $d$.

Proof. These results have already been in [8, Lemma 3.4].

Lemma 2.2. Suppose that $A$ satisfies the conditions (1.1) and (1.2). Then we conclude that $N(\cdot, \xi)$ is Hölder continuous with an upper bound estimate

$$\|N(\cdot, \xi)\|_{L^\infty(Y)} \lesssim |\xi| \quad (2.4)$$

for any $\xi \in \mathbb{R}^d$. Moreover, if we additionally impose the smoothness condition (1.3), then there exist $\theta \in (0, 1)$ and a constant $C$, depending on $\mu_0, \mu_1, \mu_2, p$ and $d$, such that

$$\|P(\cdot, \xi)\|_{L^\infty(Y)} + \|P(\cdot, \xi)\|_{C^{0,\theta}(Y)} \leq C \left(\int_Y |P(\cdot, \xi)|^p\right)^{\frac{1}{p}} \leq C|\xi| \quad (2.5)$$

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

Proof. The estimate (2.4) have be shown in [5, Theorem 4] for the case $p = 2$. In terms of $p \neq 2$, by setting $u = N(y, \xi) + y \cdot \xi$, it follows from [22, Theorem 3.9] that

$$\|u\|_{L^\infty(B(0, r))} \lesssim \left(\int_{B(0, 2r)} |u|^p\right)^{\frac{1}{p}} \quad (2.6)$$

for any $0 < r < (1/2)$. Obviously, this implies the stated estimate (2.4). Then we turn to address the estimate (2.5). Under the smoothness assumption (1.3), it is known from [9, Theorems 1,2] that

$$\|\nabla u\|_{L^\infty(B(0, r))} + \|\nabla u\|_{C^{0,\theta}(B(0, r))} \leq C(r, \mu_0, \mu_1, \mu_2, p, \|u\|_{L^\infty(B(0, 2/3))}) \leq C(r, \mu_0, \mu_1, \mu_2, p, \|u\|_{L^p(B(0, 1)))}).$$

where $r \in (0, 1/2)$, and we use the estimate (2.6) in the second inequality. Thus, it is fine to assume $\|u\|_{L^p(B(0, 1))} = 1$, and there holds

$$\|\nabla u\|_{L^\infty(B(0, 1/4))} + \|\nabla u\|_{C^{0,\theta}(B(0, 1/4))} \leq C(\mu_0, \mu_1, d, p).$$

Then, let $\bar{u} = u/\|u\|_{L^p(B(0, 1))}$, and it follows from the homogeneity assumption (1.1) that $\text{div}A(y, \nabla \bar{u}) = 0$ in $B(0, 1)$. Therefore,

$$\|\nabla \bar{u}\|_{L^\infty(B(0, 1/4))} + \|\nabla \bar{u}\|_{C^{0,\theta}(B(0, 1/4))} \lesssim \|u\|_{L^p(B(0, 1))}.$$

This together with a covering argument leads to the stated result (2.5).
Remark 2.3. In view of the estimate (2.1), one may conclude that \( N(y, 0) = 0 \) for any \( y \in \mathbb{R}^d \).

Lemma 2.4. Suppose \( A \) satisfies the assumptions (1.1) and (1.2). Let \( \hat{A} \) be given in (1.5). Then the effective operator \( L_0 \) satisfies \( \hat{A}(0) = 0 \), and

\[
\begin{cases}
\langle \hat{A}(\xi) - \hat{A}(\xi'), \xi - \xi' \rangle \geq C_0 |\xi - \xi'|^p, \\
|\hat{A}(\xi) - \hat{A}(\xi')| \leq C_1 |\xi - \xi'|(|\xi| + |\xi'|)^{p-2}
\end{cases}
\]

for \( 2 \leq p < \infty \), and

\[
\begin{cases}
\langle \hat{A}(\xi) - \hat{A}(\xi'), \xi - \xi' \rangle \geq C_0 (|\xi| + |\xi'|)^{p-2} |\xi - \xi'|^2, \\
|\hat{A}(\xi) - \hat{A}(\xi')| \leq C_2 |\xi - \xi'|^{\frac{p-1}{3-p}} (|\xi| + |\xi'|)^{\frac{(p-1)(2-p)}{3-p}}
\end{cases}
\]

for \( 1 < p \leq 2 \), in which \( C_0, C_1, C_2 \) depend on \( \mu_0, \mu_1, p \) and \( d \).

**Proof.** The proof may be found in [8, Remark 1.3] or [12, Proposition 2.7]. \( \square \)

Remark 2.5. In order to obtain some higher regularity from homogenized equations, we have to impose the following conditions

\[
\langle \hat{A}(\xi) - \hat{A}(\xi'), \xi - \xi' \rangle \geq \tilde{\mu}_0 |\xi - \xi'|^2 (|\xi| + |\xi'|)^{p-2}
\]

in the case of \( 2 < p < \infty \) (since there merely holds the first line of (2.7) in general), and

\[
|\hat{A}(\xi) - \hat{A}(\xi')| \leq \tilde{\mu}_1 |\xi - \xi'| (|\xi| + |\xi'|)^{p-2}
\]

for the case \( 1 < p < 2 \), where \( \xi, \xi' \in \mathbb{R}^d \setminus \{0\} \). Obviously, the assumptions (2.9) and (2.10) implies the first line of (2.7) and the second one of (2.8), respectively.

Lemma 2.6. Suppose that \( A \) satisfies (1.1) and (1.2). Let \( \varphi \in C_0^1(\Sigma_{2h}; \mathbb{R}^d) \), and \( 0 < h < \varepsilon \ll 1 \). Then, there exists \( \Phi(\cdot, \varphi) \in L^r(Y) \) with \( 1 < r \leq \infty \), satisfying

\[
\left( \int_Y |\Phi(\cdot, \varphi)|^r \right)^{\frac{1}{r}} \lesssim \begin{cases} |\varphi|^{\frac{p-2}{p}} & \text{if } 2 \leq p < \infty; \\
|\varphi|^{\frac{2-p}{3-p}} & \text{if } 1 < p \leq 2,
\end{cases}
\]

such that

\[
D^h(\varepsilon N(y, \varphi)) = \int_0^1 \partial_t N(y + t(h/\varepsilon)\xi_i, \tilde{\varphi}) dt + \varepsilon h^{\alpha-1} \Phi(y, \varphi) |D^h \varphi|^\alpha,
\]

where \( \tilde{\varphi} = \varphi(\cdot + he_i) \), and \( \alpha \) is defined in (1.12). Moreover, let \( h = \varepsilon^\tau \) with \( 1 < \tau < 1/(1 - \alpha) \) and \( p \neq 2 \), we obtain

\[
\int_\Omega |D^h(\varepsilon N(x/\varepsilon, \varphi))|^p dx \lesssim \int_\Omega |\varphi|^p dx + \varepsilon^{p-p(1-\alpha)\tau} \begin{cases} \int_\Omega |\varphi|^{p-2} |D^h \varphi|^2 dx & \text{if } 2 < p < \infty; \\
\int_\Omega |\varphi|^{\frac{p(p-1)}{3-p}} |D^h \varphi|^{\frac{p}{3-p}} dx & \text{if } 1 < p < 2,
\end{cases}
\]

in which the up to constant depends on \( \mu_0, \mu_1, p, d \).
Proof. According to the definition of the difference quotients, we have
\[
D^h_i(\varepsilon N(y, \varphi)) = \frac{\varepsilon}{h} \left[ N(y + (h/\varepsilon)e_i, \varphi) - N(y, \varphi) \right] + \frac{\varepsilon N(y, \varphi(x + he_i)) - \varepsilon N(y, \varphi(x))}{h}
\]
\[
= \frac{\varepsilon}{h} \int_0^1 \frac{d}{dt} \left( N(y + t(h/\varepsilon)e_i, \varphi) \right) + \varepsilon h^{-1} \Phi(y, \varphi)|D^h_i(\varphi)|^\alpha,
\]
where we note that \( \tilde{\varphi}(x) = \varphi(x + he_i) \) and we omit the bar since the scale of \( h \) is even smaller than \( \varepsilon \), and it will not essentially change the location of \( \varphi \) even measured in \( \varepsilon \) scales. By letting
\[
\Phi(y, \varphi) = \frac{N(y, \varphi(x + he_i)) - N(y, \varphi(x))}{|\varphi(x + he_i) - \varphi(x)|^\alpha},
\]
it is not hard to observe the equality (2.12). Then we proceed to verify the estimate (2.11). To do so, we denoted by \( \xi = \varphi(x) \) and \( \xi' = \varphi(x + he_i) \), and it follows from Poincaré’s inequality that
\[
\left( \int_Y |N(...)| \right)^{\frac{1}{p}} \leq \left( \int_Y |\nabla N(...)| \right)^{\frac{1}{p}} \lesssim \left( \int_Y |P(...)| \right)^{\frac{1}{p}} + |\xi - \xi'|
\]
This together with the estimate (2.3) leads to
\[
\left( \int_Y |N(...)| \right)^{\frac{1}{p}} \lesssim \left\{ \begin{array}{ll}
|\xi - \xi'|^{\frac{1}{p}} |\xi|^{\frac{1}{r}}, & \text{if } 2 \leq p < \infty; \\
|\xi - \xi'|^{\frac{1}{p}} |\xi|^{\frac{2}{1 - p}}, & \text{if } 1 < p \leq 2,
\end{array} \right.
\]
where we use the fact that \( |\xi'| \leq 2|\xi| \) whenever \( h \) is sufficient small. This proved the case \( 1 < r < \infty \) of (2.11). In the following, we proceed to show the case \( r = \infty \). Let \( u(y, \xi) = N(y, \xi) + y \cdot \xi \) and \( \tilde{u}(y, \xi) = u(y, \xi) + \tilde{M} \), in which \( \tilde{M} \) is such that \( \tilde{u} \) is positive in \( B(0, 1) \). According to the estimate (2.4), it is not hard to see the existence of \( \tilde{M} \) whenever \( |\xi| \) is bounded, and such the boundedness is in fact given by \( \|\varphi\|_{L^\infty(\Omega)} \). Note that \( \tilde{u} \) still satisfies the same type equation
\[
\text{div} A(y, \nabla u(y, \xi)) = 0 \quad \text{in} \quad B(0, 1).
\]
Thus, it follows from the local boundedness estimate and the weak Harnack inequality (see for example [22, Corollary 3.10, Theorem 3.13]) that
\[
\sup_{y \in B(0,2/3)} \tilde{u}(y, \xi) \lesssim \int_{B(0,1/5)} \tilde{u}(\cdot, \xi) \quad \text{and} \quad \inf_{y \in B(0,2/3)} \tilde{u}(y, \xi) \gtrsim \int_{B(0,1/5)} \tilde{u}(\cdot, \xi),
\]
(2.14)
respectively. Then for any \( y \in Y \) such that \( \tilde{u}(y, \xi) - \tilde{u}(y, \xi') > 0 \), there holds
\[
\tilde{u}(y, \xi) - \tilde{u}(y, \xi') \leq \sup_{y \in Y} \tilde{u}(y, \xi) - \inf_{y \in Y} \tilde{u}(y, \xi') \lesssim \int_{B(0,1/5)} |\tilde{u}(\cdot, \xi) - \tilde{u}(\cdot, \xi')|,
\]
where we use the estimates (2.14) in the second inequality. Similarly, for any \( y \in Y \) such that \( \tilde{u}(y, \xi') - \tilde{u}(y, \xi) > 0 \), one may derive
\[
\tilde{u}(y, \xi') - \tilde{u}(y, \xi) \leq \sup_{y \in Y} \tilde{u}(y, \xi') - \inf_{y \in Y} \tilde{u}(y, \xi) \lesssim \int_{B(0,1/5)} |\tilde{u}(\cdot, \xi) - \tilde{u}(\cdot, \xi')|.
\]
Then we actually obtain
\[ |u(y, \xi) - u(y, \xi')| \lesssim \int_{B(0, 4/5)} |u(\cdot, \xi) - u(\cdot, \xi')| \]
\[ \lesssim \int_Y |N(\cdot, \xi) - N(\cdot, \xi')| + |\xi - \xi'| \]
for any \( y \in Y \), where we use a covering argument in the second inequality. This further implies
\[ \|N(\cdot, \xi) - N(\cdot, \xi')\|_{L^\infty(Y)} \lesssim \begin{cases} |\xi - \xi'|^{\frac{2}{p}}, & \text{if } 2 \leq p < \infty; \\ |\xi - \xi'|^{\frac{2}{2-p}}, & \text{if } 1 < p \leq 2, \end{cases} \]
whenever \( h \) is sufficiently small, which will show \( \Phi(\cdot, \varphi) \in L^\infty(Y) \) with
\[ \|\Phi(\cdot, \varphi(x))\|_{L^\infty(Y)} \lesssim \begin{cases} |\varphi(x)|^{\frac{2}{p}}, & \text{if } 2 \leq p < \infty; \\ |\varphi(x)|^{\frac{2}{2-p}}, & \text{if } 1 < p \leq 2, \end{cases} \] (2.15)
for any \( x \in \Omega \).

Recall \( Y_i^\varepsilon = \varepsilon(i + Y) \) for \( i \in \mathbb{Z}^d \), and we collect a family of \( \{Y_i^\varepsilon\} \), whose index set is denoted by \( I_\varepsilon \), such that \( \text{supp} \varphi \subset \bigcup_{i \in I_\varepsilon} Y_i^\varepsilon \subset \Omega \) and \( Y_i^\varepsilon \cap Y_j^\varepsilon = \emptyset \) if \( i \neq j \). Thus the left-hand side of (2.13) equals
\[ \sum_{k \in I_\varepsilon} \int_{Y_i^\varepsilon} \int_0^1 \left| D^h_i(\varepsilon N(x/\varepsilon, \varphi)) \right|^p dx \leq \sum_{k \in I_\varepsilon} \int_{Y_i^\varepsilon} \left| \int_0^1 \nabla_y N(y + t(h/\varepsilon) e_i, \varphi) dt \right|^p dx \\
+ \varepsilon^{p-p(1-\alpha)\tau} \int_{\Omega} |\Phi(y, \varphi)| \left| D^h_i \varphi \right|^\alpha dx. \]
The first term in the right-hand side above is controlled by
\[ \sum_{k \in I_\varepsilon} \int_{Y_i^\varepsilon} \int_0^1 \left| \nabla_y N(y + t(h/\varepsilon) e_i, \varphi) \right|^p dt dx = \sum_{k \in I_\varepsilon} \int_{Y_i^\varepsilon} \left| \nabla_y N(y + t(h/\varepsilon) e_i, \varphi) \right|^p dt dx \\
\lesssim \sum_{k \in I_\varepsilon} |Y_i^\varepsilon| \int_0^1 \int_{Y_i^\varepsilon} \left| \nabla_y N(y + t(h/\varepsilon) e_i, \varphi) \right|^p dx dt \\
\lesssim \sum_{k \in I_\varepsilon} |Y_i^\varepsilon| \int_0^1 \int_{Y_i^\varepsilon} \left| \nabla_z N(z, \tilde{\varphi}) \right|^p dz dt \\
\lesssim \sum_{k \in I_\varepsilon} |Y_i^\varepsilon| \int_{Y_i^\varepsilon} \left| \nabla_z N(z, \tilde{\varphi}) \right|^p dz, \]
where \( \tilde{\varphi}(z) = \varphi(\varepsilon z - \text{the}_i) \) with \( z \in (k + \text{the}_i) + Y \), and \( \tilde{\varphi}(z) = \varphi(\varepsilon z - \text{the}_i - k - t(h/\varepsilon) e_i) \) with \( z \in Y \). Let \( \hat{x} = \varepsilon z - \text{the}_i - k - t(h/\varepsilon) e_i \), and one may conclude that \( \hat{x} \in Y_{2\varepsilon} \). Note that the periodicity of the corrector is employed in the last inequality. Therefore, it follows from the estimate (2.1) that
\[ \sum_{k \in I_\varepsilon} |Y_i^\varepsilon| \int_{Y_i^\varepsilon} \left| \nabla_z N(z, \tilde{\varphi}) \right|^p dz \lesssim \sum_{k \in I_\varepsilon} |Y_i^\varepsilon| \inf_{\hat{x} \in Y_{2\varepsilon}} \varphi(\hat{x})^p \lesssim \sum_{k \in I_\varepsilon} \int_{Y_i^\varepsilon} \left| \varphi \right|^p dx \lesssim \int_{\Omega} \left| \varphi \right|^p dx. \]
Here we use Chebyshev’s inequality in the second inequality. Thus, we arrive at
\[ \sum_{k \in I_\varepsilon} \int_{Y_i^\varepsilon} \int_0^1 \left| \nabla_y N(y + t(h/\varepsilon) e_i, \varphi) \right|^p dt dx \lesssim \int_{\Omega} \left| \varphi \right|^p dx. \] (2.16)
In terms of the estimate (2.16) and the periodicity of \( \Phi \) with respective to the first variable, it is not hard to see that
\[
|\Phi(x/\varepsilon, \varphi(x))| \lesssim |\varphi(x)|^\nu \quad \forall x \in \Omega,
\]
where \( \nu = (p - 2)/p \) if \( 2 < p < \infty \), and \( \nu = (2 - p)/(p - 3) \) if \( 1 < p < 2 \). Thus, we have
\[
\int_{\Omega} |\Phi(y, \varphi)| D_i^h \varphi|^{\alpha} |p| dx \lesssim \int_{\Omega} |\varphi|^{\nu p} |(D_i^h \varphi)|^{\nu p} dx.
\] (2.17)

Finally, the estimates (2.16) and (2.17) imply the desired estimate (2.13), and we have completed the proof.

**Lemma 2.7.** Let \( \varphi \in C^1_0(\Sigma_2; \mathbb{R}^d) \). Let \( \theta \in (0, 1) \) be given in Lemma 2.2, and assume the same conditions and notation as in Lemma 2.6. Then, under the additional assumption (1.3), there holds the following estimate
\[
\int_{\Omega} |D_i^h (\varepsilon N(x/\varepsilon, \varphi)) - \partial_i N(x/\varepsilon, \varphi)|^p dx \\
\lesssim \varepsilon^\theta(\tau - 1) \int_{\Omega} |\varphi|^p dx + \varepsilon^{p(1-\alpha)} \left\{ \begin{array}{ll} 
\int_{\Omega} |\varphi|^{p-2} |D_i^h \varphi|^2 dx & \text{if } 2 < p < \infty; \\
\int_{\Omega} |\varphi|^{p(2-p)/3} |D_i^h \varphi|^{\frac{p}{3-p}} dx & \text{if } 1 < p < 2,
\end{array} \right.
\] (2.18)

where the up to constant depends on \( \mu_0, \mu_1, \mu_2, p, d \).

**Proof.** According to the expression (2.12) and the estimate (2.13), it suffices to estimate the following quantity
\[
\int_{\Omega} \int_{I_0} |\partial_i N(y + t(h/\varepsilon)e_j, \varphi) - \partial_i N(y, \varphi)|^p dt dx \\
\lesssim \int_{\Omega} \int_{I_1} |\partial_i N(y + t(h/\varepsilon)e_j, \varphi) - \partial_i N(y, \varphi)|^p dt dx + \int_{\Omega} |\partial_i N(y, \varphi) - \partial_i N(y, \varphi)|^p dx.
\] (2.19)

Then, it follows from the estimate (2.5) that
\[
I_1 \leq (h/\varepsilon)^\theta \int_{\Omega} [\nabla N(\cdot, \varphi)]_{C^{0,\theta}(\mathbb{R}^d)}^p dx \int_0^1 t^\theta dt \lesssim \varepsilon^\theta(\tau - 1) \int_{\Omega} |\varphi|^p dx,
\] (2.20)

where we note that the compact support of \( \varphi \) is still included in \( \Omega \) since \( 0 < h < \varepsilon \). Moreover, we claim that \( I_1 \) will be the main term, since we have the following computations. In view of (2.3),
\[
I_2 \leq \sum_{k \in I_2} |Y^k_\varepsilon| \int_{Y^k_\varepsilon} |\nabla y N(y, \varphi) - \nabla y N(y, \varphi)|^p dx \\
\lesssim h^\alpha \sum_{k \in I_2} |Y^k_\varepsilon| |\nabla \varphi(\hat{x})|^{p\alpha} |\varphi(\hat{x})|^{p\nu} \lesssim \varepsilon^{p\alpha} \left\{ \begin{array}{ll} 
\int_{\Omega} |\nabla \varphi|^2 |\varphi|^{p-2} dx & \text{if } 2 \leq p < \infty; \\
\int_{\Omega} |\nabla \varphi|^{\frac{p}{3-p}} |\varphi|^{\frac{p(2-p)}{3-p}} dx & \text{if } 1 < p \leq 2,
\end{array} \right.
\] (2.21)

where the index \( \nu \) is defined in (2.17), and \( \hat{x} \in Y^k_\varepsilon \) is due to mean value theorem, and the last step follows from the definition of Riemann integral for \( \varphi \in C^1_0(\Omega) \). As in the previous lemma, we set \( h = \varepsilon^\tau \) with \( 1 < \tau < 1/(1-\alpha) \), and it is not hard to find that the term \( I_2 \) will be eaten by the second term in the right-hand side of (2.13) in terms of the order of \( \varepsilon \). That is the reason why we say \( I_1 \) is main term here, and the proof is complete. \( \square \)
Remark 2.8. In terms of $p=2$, on account of Lemma 2.6 and the above calculations, there holds
\[ \int_{\Omega} |D_i^b(\varepsilon N(x/\varepsilon, \varphi)) - \partial_i N(x/\varepsilon, \varphi)|^2 dx \leq \varepsilon^2 \int_{\Omega} |\nabla \varphi|^2 dx \]
for any $0 < h < \varepsilon$. If we further assume $\varphi = \psi_{4\varepsilon} \nabla u_0$, then layer and co-layer type estimates (see Lemma 3.9) leads to
\[ \left( \int_{\Omega} |D_i^b(\varepsilon N(x/\varepsilon, \varphi)) - \partial_i N(x/\varepsilon, \varphi)|^2 dx \right)^{1/2} \leq \varepsilon^\beta \left( \int_{\Omega} |\nabla u_0|^q dx \right)^{1/q} \tag{2.22} \]
where $q = 2(1 + \delta)$ and $\delta \in (0, 1)$ is the so-called Meyer’s index.

2.2 Flux corrector and its properties

Lemma 2.9 (Flux correctors). Suppose $A$ satisfies (1.1) and (1.2). Let $b(y, \xi) = A(y, \xi + \nabla N(y, \xi)) - \tilde{A}(\xi)$, where $y \in Y$ and $\xi \in \mathbb{R}^d$. Then we have two properties: (i) $\int_Y b(\cdot, \xi) = 0$; (ii) $\text{div} b(\cdot, \xi) = 0$ in $Y$. Moreover, there exists $\hat{\rho} \in (1, 2)$, depending only on $p, d$, such that the so-called flux corrector $E_{ji}(\cdot, \xi) \in W^{1, \hat{\rho}}_{\text{per}}(Y) \cap L^2_{\text{per}}(Y)$, and
\[ b_i(y, \xi) = \frac{\partial}{\partial y_j} \{ E_{ji}(y, \xi) \} \quad \text{and} \quad E_{ji} = -E_{ij}, \tag{2.23} \]
and there holds
\[ \left( \int_Y |E_{ji}(\cdot, \xi) - E_{ji}(\cdot, \xi')|^2 \right)^{1/2} \lesssim \begin{cases} |\xi - \xi'|^{2/\hat{\rho}} (|\xi| + |\xi'|) & \text{if } 2 < p < \infty; \\ |\xi - \xi'|^{2 - p} (|\xi| + |\xi'|)^{(2-p)(p-1)/q} & \text{if } 1 < p \leq 2. \end{cases} \tag{2.24} \]
Moreover, if we additionally assume the smoothness condition (1.3), then we have
\[ ||\nabla E_{ji}(\cdot, \xi)||_{C^{0, \rho}(Y)} \leq C|\xi|^p-1, \tag{2.25} \]
where $\rho = \min\{\theta, (p - 1)\theta\}$, and $C$ depends only on $\mu_0, \mu_2, \mu_3, p$ and $d$.

Proof. The proof is quite similar to the linear case (see for example [25, 32]). It is clear to see that (i) and (ii) follow from the formula (1.5) and the equation (1.6), respectively. To obtain more results, we first need to verify that $b_i(\cdot, \xi) \in L^q(Y)$ with the index $q$ being as follows:
\[ \hat{q} = \begin{cases} (d=1, 2) \quad p/(p - 1), \\ (d \geq 3) \quad (d/(d + 2) + \delta p)/(p - 1) \end{cases} \]
\[ \quad \text{if } 1 < p \leq 2d/(2d - 2); \\ \quad \text{if } 2d/(2d - 2) < p < \infty. \tag{2.26} \]
Note that $\delta \in (0, 1)$ is a small number given in Remark 4.2, and we just employ Meyer’s estimate (4.3) (note that this estimate is also true for $\varepsilon = 1$.) in the case of $2d/(2d - 2) < p < \infty$ with $d \geq 3$. In fact, a routine computation leads to
\[ \left( \int_Y |b_i(\cdot, \xi)|^q \right)^{1/q} \lesssim |\xi|^{p-1}, \tag{2.27} \]
whose proof is similar to that given for (2.30). Then, it is not hard to see that $\hat{q} \geq q$ in each case, where $q$ is shown in (2.31). This together with the property (i) guarantees the existence of the weak solution of
\[ \begin{cases} \Delta f_i(\cdot, \xi) = b_i(\cdot, \xi) & \text{in } Y, \\ f_i(\cdot, \xi) \in H^1_{\text{per}}(Y), \quad \int_Y f_i(\cdot, \xi) = 0 \end{cases} \]
for each $i$ and $\xi \in \mathbb{R}^d$. Also, we have the following energy estimate

$$\|E_{ji}\|_{L^2(Y)} \lesssim \|\nabla f_i\|_{L^2(Y)} \lesssim \|b_i\|_{L^4(Y)} \lesssim |\xi|^{p-1}. \quad (2.28)$$

Let $E_{ji}(y, \xi) = \frac{\partial}{\partial y_j} \{f_i(y, \xi)\} - \frac{\partial}{\partial y_j} \{f_j(y, \xi)\}$. Obviously, $E_{ji} = -E_{ij}$, and one may derive the first expression in (2.23) from the fact (ii) in a weak sense. In fact, one may further show that $E_{ji}(\cdot, \xi) \in W_{p\hat{p}}^1(Y)$ with some $\hat{p} \in (0, 2)$ given by

$$\hat{p} = \begin{cases} \hat{q} \in (1, 2), & 1 < p \leq 2, \text{ and } d = 1, 2; \\ \hat{q}, & \text{other cases}. \end{cases}$$

Let $Y \subset B(0, 2/3)$, and it follows from the Calderón-Zygmund theorem coupled with a localization argument that,

$$\|\nabla^2 f(\cdot, \xi)\|_{L^p(Y)} \lesssim \||\nabla f(\cdot, \xi)\|_{L^p(B(0,1))} + \|b(\cdot, \xi)\|_{L^p(B(0,1))}$$

$$\lesssim \||\nabla f(\cdot, \xi)\|_{L^1(Y)} + \|b(\cdot, \xi)\|_{L^1(Y)}$$

$$\lesssim \|b(\cdot, \xi)\|_{L^1(B(Y))},$$

in which we use the periodicity of $f(\cdot, \xi)$ and a covering argument in the second inequality, and the estimate (2.28) in the last one. Note that $1 < \hat{p} \leq \hat{q}$ in each case. Collecting the estimates (2.27) and (2.29), we obtain the estimate

$$\left(\int_Y |\nabla E(\cdot, \xi)|^\hat{p} \right)^{1/\hat{p}} \lesssim |\xi|^{p-1}. $$

Then, we turn to show the estimate (2.24). For any $\xi, \xi' \in \mathbb{R}^d$, note that

$$\left(\int_Y |E_{ji}(y, \xi) - E_{ji}(y, \xi')|^2 dy\right)^{1/2} \lesssim \left(\int_Y |\nabla (f_i(y, \xi) - f_i(y, \xi'))|^2 dy\right)^{1/2}$$

$$\lesssim \int_Y \left(|b_i(y, \xi) - b_i(y, \xi')|^q dy\right)^{1/q},$$

where we employ an $H^1$ estimate in the second step, and

$$q = \begin{cases} 2d/(d+2) & \text{if } d \geq 3; \\ 1 < q < 2 & \text{if } d = 2; \\ q = 1 & \text{if } d = 1. \end{cases} \quad (2.31)$$

Then, we carry out a computation in the case of $p > 2$, which in fact involves higher regularity of $P(\cdot, \xi)$. By definition, the right-hand side of (2.30) is controlled by

$$\left(\int_Y |A(\cdot, P(\cdot, \xi)) - A(\cdot, P(\cdot, \xi'))|^q \right)^{1/q} + |\hat{A}(\xi) - \hat{A}(\xi')|$$

$$\lesssim \left(\int_Y |P(\cdot, \xi) - P(\cdot, \xi')|^q \left(|P(\cdot, \xi)| + |P(\cdot, \xi')|\right)^{(p-2)}\right)^{1/q} + |\xi - \xi'|(|\xi| + |\xi'|)^{p-2}.$$
where we use the estimates (2.3), (2.5) in the second inequality. This is the first line of (2.24), and we continue to show the case \(1 < p \leq 2\). On account of (2.8), it suffices to compute the following quantity, and by (1.2) we have

\[
\left( \int_Y |A(\cdot, P(\cdot, \xi)) - A(\cdot, P(\cdot, \xi'))|^q \right)^{1/q} \lesssim \left( \int_Y |P(\cdot, \xi) - P(\cdot, \xi')|^{q(p-1)} \right)^{1/q} \\
\lesssim \left( \int_Y |P(\cdot, \xi) - P(\cdot, \xi')|^p \right)^{\frac{p-1}{p}} \lesssim |\xi - \xi'| \frac{p-1}{p(p-1)} (|\xi| + |\xi'|)^{(2-p)(p-1)},
\]

where we use the fact that \(\frac{p}{p-1} \geq 2\) in the second inequality (which gives \(\frac{p}{q(p-1)} \geq 1\), and (2.3) in the last one. This ends the proof of the stated estimate (2.24).

We now proceed to prove the estimate (2.25). On account of Lemma 2.4, the assumptions (1.1), (1.2), one may have

\[
\|b(\cdot, \xi)\|_{L^\infty(\mathbb{R}^d)} \leq \|P(\cdot, \xi)\|_{L^{p-1}(\mathbb{R}^d)}^{p-1} + |\tilde{A}(\xi)| \lesssim |\xi|^p.
\]

where we also use the estimate (2.5) in the second inequality. For any \(y_1, y_2 \in \mathbb{R}^d\) with \(|y_1 - y_2| \ll 1\), we first have

\[
|b(y_1, \xi) - b(y_2, \xi)| \lesssim |A(y_1, P(y_1, \xi)) - A(y_2, P(y_1, \xi))| + |A(y_2, P(y_1, \xi)) - A(y_2, P(y_2, \xi))|.
\]

Then, it follows from the assumption (1.3) and the estimate (2.5) that

\[
I_1 \leq \mu_3 |y_1 - y_2| \|P(\cdot, \xi)\|_{L^{p-1}(\mathbb{R}^d)}^{p-1} \lesssim |\xi|^{p-1} |y_1 - y_2|.
\]

Similarly, by the assumption (1.2), and the estimate (2.5) we derive that

\[
I_2 \leq \mu_1 \left\{ \left\| P(\cdot, \xi) \right\|_{L^\infty(\mathbb{R}^d)}^{p-2} \| P(\cdot, \xi) \|_{C^{0,\theta}(\mathbb{R}^d)} |y_1 - y_2|^\theta \right\} \lesssim \left\{ \begin{array}{ll} |\xi|^{p-1} |y_1 - y_2|^\theta & 2 \leq p < \infty; \\ |\xi|^{p-1} |y_1 - y_2|^{(p-1)\theta} & 1 < p < 2. \end{array} \right.
\]

Hence, by setting \(\rho = \min\{\theta, (p-1)\theta\}\), combining the above estimates leads to

\[
|b(\cdot, \xi)|_{C^{0,\rho}(\mathbb{R}^d)} + \|b(\cdot, \xi)\|_{L^\infty(\mathbb{R}^d)} \lesssim |\xi|^p. \tag{2.32}
\]

This together with interior Schauder estimate (see for example [18, Theorem 5.19]) leads to

\[
[\nabla^2 f_i]_{C^{0,\rho}(Y)} \lesssim \|\nabla f_i\|_{L^2(B(0,2/3))} + \|b_i\|_{C^{0,\rho}(B(0,2/3))} \lesssim |\xi|^p,
\]

in which we use the estimates (2.28) and (2.32) in the second inequality. This implies the desired estimate (2.25), and we have completed the whole proof. \(\square\)

**Remark 2.10.** In fact, the estimate (2.24) is still true if the integrand \(|E_{ji}(y, \xi) - E_{ji}(y, \xi')|\) is replaced by \(|\nabla_y E_{ji}(y, \xi) - \nabla_y E_{ji}(y, \xi')|\). Since \(E_{ji}(y, \xi)\) satisfies

\[
\Delta E_{ji}(\cdot, \xi) = (\nabla \times b(\cdot, \xi))_{ji} =: \nabla_j b_i(\cdot, \xi) - \nabla_i b_j(\cdot, \xi) \tag{2.33}
\]

the result follows from the linearity and Caccioppoli’s inequality, i.e.,

\[
\left( \int_Y |\nabla E_{ji}(\cdot, \xi) - \nabla E_{ji}(\cdot, \xi')|^2 \right)^{1/2} \lesssim \left( \int_{2Y} |E_{ji}(\cdot, \xi) - E_{ji}(\cdot, \xi')|^2 \right)^{1/2} + \left( \int_{2Y} |b_i(\cdot, \xi) - b_i(\cdot, \xi')|^2 \right)^{1/2}.
\]
Thus, in view of (2.24), it suffices to estimate the last term above, while it admits the same computations as in the proof of (2.30). We do not repeat here. Finally, we mention that the equation (2.33) may be referred to as the definition of the flux corrector (see [15, Lemma 1]).

**Lemma 2.11.** Let \( \varphi \in C^1_0(\Omega; \mathbb{R}^d) \). Assume the same conditions as in Lemma 2.9. Let \( E_{ji}(\cdot, \xi) \in W^{1,\hat{p}}(Y) \cap L^2_{per}(Y) \) with \( \hat{p} \in (1, 2) \) be the flux corrector, then there exists \( \Psi_{ji}(\cdot, \varphi) \in L^2(Y) \) satisfying

\[
\left( \int_Y |\Psi_{ji}(\cdot, \varphi)|^2 \right)^{\frac{1}{2}} \lesssim \begin{cases} 
|\varphi| \left( \frac{(p-2)(1+p)}{p} \right) & \text{if } 2 \leq p < \infty; \\
|\varphi| \left( \frac{(2-p)(p-1)}{4-p} \right) & \text{if } 1 < p \leq 2,
\end{cases}
\]  

(2.34)

such that

- in the case of \( p = 2 \), we have

\[
\varepsilon \frac{\partial}{\partial x_j} [E_{ji}(x/\varepsilon, \varphi)] = \frac{\partial}{\partial y_j} [E_{ji}(x/\varepsilon, \varphi)] + \varepsilon \frac{\partial}{\partial \xi_k} [E_{ji}(\cdot, \varphi)] \frac{\partial \varphi_k}{\partial x_j}
\]  

(2.35)

with the estimate

\[
\left( \int_Y |\nabla E_{ji}(\cdot, \varphi)|^2 \right)^{\frac{1}{2}} \leq C(\mu_0, \mu_1, \delta);
\]  

(2.36)

- in the case of \( p \in (1, 2) \cup (2, \infty) \), there holds

\[
D^{h}_j(\varepsilon E_{ji}(y, \varphi)) = \int_0^1 \partial_j E_{ji}(y + t(h/\varepsilon)e_j, \varphi) dt + \varepsilon h^{-1} \Psi_{ji}(y, \varphi) |D^{h}_j \varphi|^\gamma,
\]  

(2.37)

where \( \varphi = \varphi(\cdot + he_j) \), and \( \gamma \) is defined in (1.12).

Moreover, let \( h = \varepsilon^\tau \) with \( 1 < \tau < 1/(1 - \gamma) \), and \( 1 < r \leq p/(p - 1) \), we obtain

\[
\int_{\Omega} \left| D^{h}_j(\varepsilon E_{ji}(x/\varepsilon, \varphi)) - \partial_j E_{ji}(x/\varepsilon, \varphi) \right|^r dx
\]

\[
\lesssim \varepsilon^{r(\rho - 1)} \int_{\Omega} |\varphi|^{(p-1)} dx + \varepsilon^{r - r(1 - \gamma)} \begin{cases} 
\int_{\Omega} \left| \varphi \left( \frac{r(p-2)(1+p)}{p} \right) |D^{h}_j \varphi|^{\frac{2r}{p}} dx & \text{if } 2 \leq p < \infty; \\
\int_{\Omega} \left| \varphi \left( \frac{(2-p)(p-1)}{4-p} \right) |D^{h}_j \varphi|^{\frac{(p-1)r}{3-p}} dx & \text{if } 1 < p < 2,
\end{cases}
\]  

(2.38)

in which \( \rho = \min\{\theta, \theta(p-1)\} \), and the up to constant depends on \( \mu_0, \mu_1, \mu_2, p \) and \( d \).

**Remark 2.12.** Although one can not expect that the estimate (2.34) is true in a \( L^\infty \)-norm, the quantities \( \Psi(\cdot/\varepsilon, \varphi(\cdot)) \) and \( \partial_j E_{ji}(\cdot/\varepsilon, \varphi(\cdot)) \) are still continuous under the smoothness assumption (1.3). In the end, we mention that the estimate (2.38) is also valid for the operator \( D^{-h}_j \).

**Proof.** Obviously, using the same argument as in the proof of Lemma 2.6, we may easily carry out the proof for the estimate (2.34) and the formula (2.37), which are all based upon the estimate (2.24). Meanwhile, (2.35) and (2.36) have been shown in the previous work [29]. Here we just give the expression of

\[
\Psi_{ji}(y, \varphi) = \frac{E_{ji}(y, \varphi(\cdot + he_i)) - E_{ji}(y, \varphi)}{|\varphi(\cdot + he_i)| - \varphi}\]

and note that \( |\varphi(\cdot + he_i)| \lesssim |\varphi| \) as long as \( h \) is sufficiently small, and this inequality is independent of the variable of \( \varphi \), and this has shown the equality (2.37).
Then we turn to show the estimate (2.38), and use the same notation as in the proof of Lemma 2.6. Let $Y^i_\varepsilon = \varepsilon(i + Y)$ for $i \in \mathbb{Z}^d$, and we collect a family of $\{Y^i_\varepsilon\}$, whose index set is denoted by $I_\varepsilon$, such that $\text{supp}\varphi \subset \bigcup_{i \in I_\varepsilon} Y^i_\varepsilon \subset \Omega$ and $Y^i_\varepsilon \cap Y^j_\varepsilon = \emptyset$ if $i \neq j$. Thus, the left-hand side of (2.38) may be controlled by

\[ \sum_{k \in I_\varepsilon} \int_{Y^k_\varepsilon} \left| D_j^b (\varepsilon E_{ji}(x/\varepsilon, \varphi)) - \partial_j E_{ji}(x/\varepsilon, \varphi) \right|^r dx \leq \varepsilon^{r-r(1-\rho)} \sum_{k \in I_\varepsilon} \int_{Y^k_\varepsilon} \left| \Psi_{ji}(y, \varphi) \right|^r |D_j^b \varphi|^{r\gamma} dx \]

\[ + \int_{\Omega} \int_0^1 \left| \partial_j E_{ji}(y + t(h/\varepsilon)e_j, \varphi) - \partial_j E_{ji}(y, \varphi) \right|^r dt dx, \]

where $\varphi^k = \inf_{x \in Y^k_\varepsilon} |\varphi|$, and $\hat{x} \in Y^k_\varepsilon$. Then, by a similar computation given in the proof of (2.20), we have

\[ I_2 \lesssim (h/\varepsilon)^r p \int_0^1 t^\rho dt \int_{\Omega} \left| \nabla E(\cdot, \varphi) \right|^r dx + \int_{\Omega} \left| \partial_j E_{ji}(y, \varphi) - \partial_j E_{ji}(y, \varphi) \right|^r dx \]

\[ \lesssim \varepsilon^{r\rho(\tau-1)} \int_{\Omega} |\varphi|^{r(p-1)} dx + \varepsilon^{r\gamma} \int_{\Omega} |\nabla \varphi|^{r\gamma} |\varphi|^{r\beta} dx, \]

where we use the estimate (2.25) to derive the first term in the second inequality. Note that the calculation for $I_3$ is totally similar to that given in (2.21), and a required estimate could be found in Remark 2.10. The notation $\bar{\beta}$ is given after the estimate (2.40). Also, an analogy computation leads to

\[ I_1 \lesssim \sum_{k \in I_\varepsilon} |Y^k_\varepsilon| \int_{Y^k_\varepsilon} \left| \Psi(y, \varphi) \right|^r dx \left| (D_j^b \varphi)(\hat{x}) \right|^{r\gamma} \]

\[ \lesssim \sum_{k \in I_\varepsilon} |Y^k_\varepsilon| |\varphi|^{|r\beta|} \left| (D_j^b \varphi)(\hat{x}) \right|^{r\gamma} \lesssim \int_{\Omega} |\varphi|^{r\beta} \left| (D_j^b \varphi) \right|^{r\gamma} dx, \]

where $\bar{\beta} = (p-2)(1+p)/p$ if $2 < p < \infty$, and $\bar{\beta} = (2-p)(p-1)/(3-p)$ if $1 < p < 2$. Here the point $\hat{x} \in Y^k_\varepsilon$ is such that $|D_j^b \varphi(\hat{x})| = \sup_{x \in Y^k_\varepsilon} |D_j^b \varphi(x)|$, and then we choose $\varphi^k = \varphi(\hat{x})$. In fact, we employ the estimate (2.34) in the second inequality, and the last one is due to the definition of Riemann integral for $\varphi \in C^0_0(\Omega)$. Finally, the estimates (2.39) and (2.40) imply the desired estimate (2.38), and the proof is complete. \hfill \square

### 2.3 Smoothing operator and its properties

**Definition 1.** Fix a nonnegative function $\zeta \in C^\infty_0(B(0,1/2))$, and $\int_{\mathbb{R}^d} \zeta(x) dx = 1$. Define the smoothing operator

\[ S_\varepsilon(f)(x) = f \ast \zeta_\varepsilon(x) = \int_{\mathbb{R}^d} f(x-y) \zeta_\varepsilon(y) dy, \]

where $\zeta_\varepsilon = \varepsilon^{-d} \zeta(x/\varepsilon)$.

**Lemma 2.13.** Let $f \in L^p(\Omega)$ for some $1 \leq p < \infty$. Then for any $\varpi \in L^p_{\text{per}}(Y)$,

\[ \| \varpi \ast (\cdot/\varepsilon) S_\varepsilon(f) \|_{L^p(\Omega)} \leq C \| \varpi \|_{L^p(Y)} \| \varpi \|_{L^p(\Omega)}, \]

where $C$ depends only on $\zeta$ and $d$. 
Proof. The proof may be find in [26, Lemma 2.1] or [28, Lemma 3.2]. □

Lemma 2.14. Let \( f \in W^{1,p}(\Sigma_\varepsilon) \) for some \( 1 < p < \infty \). Then we have

\[
\| S_\varepsilon(f) - f \|_{L^p(\Sigma_{2\varepsilon})} \leq C\varepsilon \| \nabla f \|_{L^p(\Sigma_\varepsilon)},
\]

where \( C \) depends only on \( d \).

Proof. The proof may be found in [26, Lemma 2.2] and [28, Lemma 3.3], which is based upon the characterization of \( W^{1,p} \) functions, i.e., \( \| f(\cdot - \varepsilon y) - f(\cdot) \|_{L^p(\Sigma_{2\varepsilon})} \leq \varepsilon \| \nabla f \|_{L^p(\Sigma_\varepsilon)} \), where \( |y| \leq 1 \). This together with Minkowskis inequality consequently leads to the stated estimate (2.43). □

Remark 2.15. To obtain the inequality (1.22), we need to apply the estimate (2.43) in each of small cells \( Y_k^\varepsilon \) (see the proof of (3.13)). So, it is crucial in applications that the constant \( C \) is independent of the size of domain. Compared to [28, Lemma 3.3], it is not hard to see that \( |\nabla u_0|^r \) with some \( r > 1 \) in (1.22) plays a role as a weight.

3 Convergence rates

3.1 Variational formula

Lemma 3.1. Let \( \Omega \subset \mathbb{R}^d \) be a bounded Lipschitz domain. Suppose that \( u_\varepsilon, u_0 \in W^{1,p}(\Omega) \) satisfy \( \mathcal{L}_\varepsilon u_\varepsilon = \mathcal{L}_0 u_0 \) in \( \Omega \) with \( u_\varepsilon = u_0 \) on \( \partial \Omega \). For any \( \varphi \in C_0^1(\Omega; \mathbb{R}^d) \), the first-order approximating corrector is given by \( v_\varepsilon = u_0 + \varepsilon N(x/\varepsilon, \varphi) \), and let \( V = (V_i) \in L^p(\Omega; \mathbb{R}^d) \), with

\[
V_i(x) = \nabla_i u_0(x) + D_i^b(\varepsilon N(x/\varepsilon, \varphi))
\]

for some \( 0 < h < \varepsilon \). Assume that \( A \) satisfies (1.1) and (1.2), then we have

\[
\int_\Omega (A(x/\varepsilon, \nabla u_\varepsilon) - A(x/\varepsilon, V)) : (\nabla u_\varepsilon - V) dx
= \int_\Omega [\tilde{A}(\nabla u_0) - A(x/\varepsilon, V)] : (\nabla u_\varepsilon - V) dx - o(\varepsilon) \int_\Omega N(x/\varepsilon, \varphi) dx
\]

as \( h \) is sufficiently small.

Proof. In fact, it suffices to show

\[
\int_\Omega A(x/\varepsilon, \nabla u_\varepsilon) : (\nabla u_\varepsilon - V) dx = \int_\Omega \tilde{A}(\nabla u_0) : (\nabla u_\varepsilon - V) dx - o(\varepsilon) \int_\Omega N(x/\varepsilon, \varphi) dx.
\]

In view of \( \mathcal{L}_\varepsilon u_\varepsilon = \mathcal{L}_0 u_0 \) in \( \Omega \), one may have

\[
D_i^{-b}[A(x/\varepsilon, \nabla u_\varepsilon) - \tilde{A}(\nabla u_0)] = o(1),
\]

as \( h \to 0 \). This gives

\[
\int_\Omega D_i^{-b}[A_i(x/\varepsilon, \nabla u_\varepsilon) - \tilde{A}_i(\nabla u_0)] \varepsilon N(x/\varepsilon, \varphi) dx = o(\varepsilon) \int_\Omega N(x/\varepsilon, \varphi) dx,
\]

and using integration by parts for difference quotients,

\[
\int_\Omega A_i(x/\varepsilon, \nabla u_\varepsilon) D_i^b(\varepsilon N(x/\varepsilon, \varphi)) dx = \int_\Omega \tilde{A}_i(\nabla u_0) D_i^b(\varepsilon N(x/\varepsilon, \varphi)) dx - o(\varepsilon) \int_\Omega N(x/\varepsilon, \varphi) dx.
\]
Note that $N(x/\varepsilon, \varphi) = 0$ on $\partial \Omega$ due to Remark (2.3). On the other hand, we have

$$\int_{\Omega} A(x/\varepsilon, \nabla u_\varepsilon) \cdot (\nabla u_\varepsilon - \nabla u_0) \, dx = \int_{\Omega} \tilde{A}(\nabla u_0) \cdot (\nabla u_\varepsilon - \nabla u_0) \, dx,$$

where we also use the fact that $u_\varepsilon - u_0 \in W^{1,p}_0(\Omega)$. Obviously, the above two equalities imply the desired equality (3.2). We are done. \(\square\)

**Lemma 3.2.** Assume the same conditions as in Lemma 3.1. Let $K = (K_i)$ be with

$$K_i = \nabla_i u_\varepsilon - \nabla_i u_0 - D_i^h(\varepsilon N(y, \varphi_\varepsilon)), \quad y = x/\varepsilon.$$  \hfill (3.3)

Also assume $\varphi_\varepsilon = S_\varepsilon(\psi_4 \nabla u_0)$, where $\psi_4 \in C^1_0(\Omega)$ is defined in (1.27). Then we have

$$\|K\|_{L^p(\Omega)}^p \leq \int_{\Omega} [\tilde{A}(\nabla u_0) - \tilde{A}(\varphi_\varepsilon)
+ \tilde{A}(\varphi_\varepsilon) - A(x/\varepsilon, P(x/\varepsilon, \varphi_\varepsilon))
+ A(x/\varepsilon, P(x/\varepsilon, \varphi_\varepsilon)) - A(x/\varepsilon, V)] \cdot K \, dx,$$ \hfill (3.4)

by ignoring a quantity $o(\varepsilon)$.

**Proof.** In view of the assumption (1.2), the left-hand side of the equality (3.2) gives

$$\int_{\Omega} (A(x/\varepsilon, \nabla u_\varepsilon) - A(x/\varepsilon, V)) \cdot K \, dx \geq \mu_0 \|K\|_{L^p(\Omega)}^p,$$ \hfill (3.5)

while its right-hand side is equal to

$$\int_{\Omega} [\tilde{A}(\nabla u_0) - \tilde{A}(\varphi_\varepsilon) + \tilde{A}(\varphi_\varepsilon) - A(x/\varepsilon, P(x/\varepsilon, \varphi_\varepsilon))
+ A(x/\varepsilon, P(x/\varepsilon, \varphi_\varepsilon)) - A(x/\varepsilon, V)] \cdot K \, dx + o(\varepsilon) \int_{\Omega} N(x/\varepsilon, \varphi_\varepsilon) \, dx.$$ \hfill (3.6)

Note that

$$\int_{\Omega} |N(x/\varepsilon, \varphi_\varepsilon)| \, dx \lesssim \sum_{k \in I_\varepsilon} |Y_k^k| \int_{Y_k^k} |N(x/\varepsilon, \varphi_\varepsilon)| \, dx \lesssim \sum_{k \in I_\varepsilon} |Y_k^k| \inf_{x \in Y_k^k} |\varphi_\varepsilon| \lesssim \int_{\Omega} |\varphi_\varepsilon| \, dx \lesssim \int_{\Sigma_{4\varepsilon}} |\nabla u_0| \, dx$$ \hfill (3.7)

where we use the estimate (2.1) in the second inequality, and the last one follows from the estimate (2.42) by recalling $\varphi_\varepsilon = S_\varepsilon(\psi_4 \nabla u_0)$. Thus, due to the estimate (3.7) we may omit the second term of (3.6), otherwise there is nothing to prove. Combining the estimates (3.5) and (3.6) leads to our desired estimate (3.4), and we have completed the proof. \(\square\)
3.2 Estimates involving a shift argument

Lemma 3.3. Assume the same conditions as in Lemma 3.2. Let $K$ be defined in (3.3), and $u_0 \in W^{1,p}(\Omega)$ satisfies $L_0(u_0) = \mathcal{L}(u_\varepsilon) = 0$ in $\Omega$ with the assumption (1.14). Then we have

$$
\left| \int_\Omega [\tilde{A}(\nabla u_0) - \tilde{A}(\varphi_\varepsilon)] \cdot \tilde{K} \, dx \right| \lesssim \left( 1 - I_{\{p=2\}} \right) |\Omega_{M,\varepsilon}|^{\beta(p-1)} \left( \int_{\Omega \setminus \Sigma_{3\varepsilon}} |\nabla u_0|^p \, dx \right)^{\frac{p-1}{p}}
$$

$$
+ \left\{ \begin{array}{ll}
\left( \int_{\Omega \setminus \Sigma_{3\varepsilon}} |\nabla u_0|^p \, dx \right)^{1-\frac{1}{p}} + \varepsilon \left( \int_{\Sigma_{3\varepsilon}} |\nabla^2 u_0|^2 |\nabla u_0|^{p-2} \, dx \right)^{\frac{1}{2}} \left( \int_{\Omega} |\nabla u_0|^p \right)^{\frac{p-2}{2p}} & \text{if } p \in [2, \infty);
\left( \int_{\Omega \setminus \Sigma_{3\varepsilon}} |\nabla u_0|^p \, dx \right)^{1-\frac{1}{p}} + \varepsilon^{\frac{1}{p-1}} \left( \int_{\Sigma_{3\varepsilon}} |\nabla^2 u_0|^p \, dx \right)^{\frac{1}{p-1}} \left( \int_{\Omega} |\nabla u_0|^p \right)^{\frac{2-p(p-1)}{p(3-p)}} & \text{if } p \in (1, 2),
\end{array} \right.
$$

(3.8)

where the upper constant depends on $\mu_0, \mu_1, p$ and $d$.

Proof. Obviously, the proof should be divided into two parts. We first handle the case $2 \leq p < \infty$. It follows from the estimate (2.7) that

$$
\left| \int_\Omega [\tilde{A}(\nabla u_0) - \tilde{A}(\varphi_\varepsilon)] \cdot \tilde{K} \, dx \right| \lesssim \int_\Omega |\nabla u_0 - \varphi_\varepsilon| (|\nabla u_0| + |\varphi_\varepsilon|)^{p-2} |\tilde{K}| \, dx
$$

$$
\lesssim \int_{\Omega \setminus \Sigma_{3\varepsilon}} |\nabla u_0|^p |\tilde{K}| \, dx + \int_{\Sigma_{3\varepsilon}} |\nabla u_0 - \varphi_\varepsilon||\nabla u_0|^{p-2} |\tilde{K}| \, dx
$$

(3.9)

in which we recall that $\varphi_\varepsilon = S_\varepsilon(\psi_{4\varepsilon} \nabla u_0)$, and employ the facts:

$$
\text{supp} S_\varepsilon(\psi_{4\varepsilon} \nabla u_0) \subset \Sigma_{3\varepsilon}, \quad \text{and} \quad |S_\varepsilon(\psi_{4\varepsilon} \nabla u_0)(x)| \leq 2|\nabla u_0(x)| \quad \forall x \in \Sigma_{3\varepsilon},
$$

(3.10)

whenever $\varepsilon$ is sufficiently small. Also, the last step is due to the definition of the cut-off function. The easier term is

$$
I_1 \leq \left( \int_{\Omega \setminus \Sigma_{3\varepsilon}} |\nabla u_0|^p \, dx \right)^{1-\frac{1}{p}}.
$$

(3.11)

Then we show the estimate on $I_2$,

$$
I_2 \leq \left( \int_{\Sigma_{3\varepsilon} \setminus \Omega_{M,\varepsilon}} |\psi_{4\varepsilon} \nabla u_0 - \varphi_\varepsilon|^{\frac{p}{p-1}} |\nabla u_0|^{\frac{p(p-2)}{p-1}} \, dx \right)^{1-\frac{1}{p}}
$$

$$
\leq \left( \int_{\Sigma_{3\varepsilon} \setminus \Omega_{M,\varepsilon}} |\psi_{4\varepsilon} \nabla u_0 - \varphi_\varepsilon|^{\frac{p}{p-1}} |\nabla u_0|^{\frac{p(p-2)}{p-1}} \, dx \right)^{1-\frac{1}{p}} + \left( \int_{\Omega \setminus \Sigma_{3\varepsilon}} |\psi_{4\varepsilon} \nabla u_0 - \varphi_\varepsilon|^{\frac{p}{p-1}} |\nabla u_0|^{\frac{p(p-2)}{p-1}} \, dx \right)^{1-\frac{1}{p}}.
$$

Let $\Sigma_{3\varepsilon} \setminus \Omega_{M,\varepsilon} \subset \cup_{k \in I'_\varepsilon} Y_{\varepsilon}^k$. Clearly, $I'_\varepsilon \subset I_\varepsilon$, and the set $\Sigma_{3\varepsilon} \setminus \Omega_{M,\varepsilon}$ is open. Here the notation $Y_{\varepsilon}^k$ and $I_\varepsilon$ are given in Lemma 2.6. Note that for any $k \in I'_\varepsilon$, there holds a Harnack’s inequality

$$
\sup_{Y_{\varepsilon}^k} |\nabla u_0| \leq 4 \inf_{Y_{\varepsilon}^k} |\nabla u_0|
$$

(3.12)
in small scales, since $|\nabla u_0| > M\varepsilon^\theta$ in $\Sigma_{3\varepsilon} \setminus \Omega_{M,\varepsilon}$ under the assumption (1.14), Thus, where we use the estimate (2.43) in the small scales. Then, one may have the following computation:

\[ I_{21} \leq \sum_{k \in I'_c} \left( \int_{Y^\varepsilon_k} |\psi_{4\varepsilon} \nabla u_0 - \varphi_{\varepsilon}|^{p-1} dx \right) \sup_{Y^\varepsilon_k} |\nabla u_0|^{\frac{p(p-2)}{p-1}} \]

\[ \lesssim \varepsilon^{\frac{p}{p-1}} \sum_{k \in I'_c} \left( \int_{Y^\varepsilon_k} |\nabla (\psi_{4\varepsilon} \nabla u_0)|^{\frac{p}{p-1}} dx \right) \inf_{Y^\varepsilon_k} |\nabla u_0|^{\frac{p(p-2)}{p-1}} \]

\[ \lesssim \varepsilon^{\frac{p}{p-1}} \sum_{k \in I'_c} \int_{Y^\varepsilon_k} |\nabla (\psi_{4\varepsilon} \nabla u_0)|^{\frac{p}{p-1}} |\nabla u_0|^{\frac{p(p-2)}{p-1}} dx \]

\[ \lesssim \varepsilon^{\frac{p}{p-1}} \int_{\Omega} |\nabla (\psi_{4\varepsilon} \nabla u_0)|^{\frac{p}{p-1}} |\nabla u_0|^{\frac{p(p-2)}{p-1}} dx \]

where we use the estimate (2.43) in the small scales, as well as (3.12) in the second inequality. Then, the last line above will be immediately controlled by

\[ \int_{\Omega \setminus \Sigma_{8\varepsilon}} |\nabla u_0|^p dx + \varepsilon^{\frac{p}{p-1}} \int_{\Sigma_{4\varepsilon}} |\nabla^2 u_0|^{\frac{p}{p-1}} |\nabla u_0|^{\frac{p(p-2)}{p-1}} dx. \]

Therefore, we arrive at

\[ I_{21} \lesssim \left\{ \int_{\Omega \setminus \Sigma_{8\varepsilon}} |\nabla u_0|^p dx + \varepsilon^{\frac{p}{p-1}} \left( \int_{\Sigma_{4\varepsilon}} |\nabla^2 u_0|^2 |\nabla u_0|^{p-2} dx \right)^{\frac{p}{2(p-1)}} \left( \int_{\Omega} |\nabla u_0|^p dx \right)^{\frac{p-2}{2(p-1)}} \right\}, \]

where we just merely use Hölder’s inequality and it asks for the condition $p \geq 2$. We now proceed to show the estimate on $I_{22}$. Note that it is fine to assume that there are finite connected components, denoted by $\{\Omega_i\}_{i=1}^{k_0}$ such that $\Omega_{M,\varepsilon} = \bigcup_{i=1}^{k_0} \Omega_i$ since the geometry of $\Omega_{M,\varepsilon}$ is not bad in general on account of the assumption (1.14).

\[ I_{22} \leq \sum_{i=1}^{k_0} \int_{\Omega_i} |\nabla u_0|^p dx \leq \sum_{i=1}^{k_0} |\Omega_i|^{1-\frac{p}{q}} \left( \int_{\Omega_i} |\nabla u_0|^q dx \right)^{\frac{p}{q}} \lesssim |\Omega_{M,\varepsilon}|^{1-\frac{p}{q}} \left( \int_{\Omega_{M,\varepsilon}} |\nabla u_0|^q dx \right)^{\frac{p}{q}}, \]

where we just use Hölder’s inequality. Thus, Plugging the above expression back into $I_2$, we have

\[ I_2 \lesssim |\Omega_{M,\varepsilon}|^{\beta(p-1)} \left( \int_{\Omega_{M,\varepsilon}} |\nabla u_0|^q dx \right)^{\frac{p-1}{q}} + \left( \int_{\Omega \setminus \Sigma_{8\varepsilon}} |\nabla u_0|^p dx \right)^{1-\frac{p}{q}} \]

\[ + \varepsilon \left( \int_{\Sigma_{4\varepsilon}} |\nabla^2 u_0|^2 |\nabla u_0|^{p-2} dx \right)^{\frac{1}{2}} \left( \int_{\Omega} |\nabla u_0|^p dx \right)^{\frac{p-2}{2p}}, \]

where we use Hölder’s inequality in the second line, which asks for the condition $p \geq 2$. This together with the estimates (3.9) and (3.11) consequently gives the first line of the desired estimate (3.8).

Now, we turn to show the proof in the case of $1 < p < 2$. In view of the estimate (2.8),

\[ \left| \int_{\Omega} [\tilde{A}(\nabla u_0) - \tilde{A}(\varphi_{\varepsilon})] \cdot \tilde{K} dx \right| \lesssim \int_{\Omega} |\nabla u_0 - \varphi_{\varepsilon}|^{\frac{q(p-1)}{p-2}} (|\nabla u_0| + |\varphi_{\varepsilon}|)^{\frac{(p-1)(2-p)}{p}} |\tilde{K}| dx \]

\[ \lesssim \left( \int_{\Omega} |\nabla u_0 - \varphi_{\varepsilon}|^{\frac{q(p-1)}{p-2}} (|\nabla u_0| + |\varphi_{\varepsilon}|)^{\frac{(p-1)(2-p)}{p}} dx \right)^{1-\frac{1}{p}}. \]
In fact, it suffices to estimate the term $I_3$, and we have
\[ I_3 \leq \left( \int_{\Omega \setminus \Sigma_{3\epsilon}} |\nabla u_0|^p \, dx \right)^{1 - \frac{1}{p}} + \left( \int_{\Sigma_{3\epsilon}} |\nabla u_0 - \varphi_\epsilon\|_{p,\epsilon}^p |\nabla u_0|^{p(2-p)} \, dx \right)^{\frac{1}{p}} \]
\[ \lesssim \left( \int_{\Omega \setminus \Sigma_{3\epsilon}} |\nabla u_0|^p \, dx \right)^{1 - \frac{1}{p}} + \left( \int_{\Sigma_{3\epsilon}} |\psi_{\epsilon} \nabla u_0 - \varphi_\epsilon\|_{p,\epsilon}^p |\nabla u_0|^{p(2-p)} \, dx \right)^{\frac{1}{p}} \]
by noting the facts (3.10). Then using the same argument as in the proof of the estimates (3.13), (3.14) we can derive that
\[ I_3 \lesssim |\Omega_{M,\epsilon}|^{\beta(p-1)} \left( \int_{\Omega_{M,\epsilon}} |\nabla u_0|^q \, dx \right)^{\frac{p-1}{q}} + \left( \int_{\Omega \setminus \Sigma_{3\epsilon}} |\nabla u_0|^p \, dx \right)^{1 - \frac{1}{p}} \]
\[ + \varepsilon^{\frac{p-1}{3-p}} \left( \int_{\Sigma_{3\epsilon}} |\nabla^2 u_0|^{\frac{3}{3-p}} \, dx \right)^{\frac{p-1}{3-p}} \left( \int_{\Omega} |\nabla u_0|^{\frac{(p-1)(2-p)}{(p-3)(p-2)}} \right), \]
where we also use the Hölder’s inequality in the above estimate, which requires the condition $1 < p < 2$. This will give the second line of (3.8), and we have completed the whole proof. \(\square\)

3.3 Applications of the flux corrector

Lemma 3.4. Assume the same conditions as in Lemma 3.2. Let $K$ be defined in (3.3), and $E_{ji}$ be the flux corrector defined in Lemma 2.9. Then we obtain
\[ \left| \int_{\Omega} \left[ \hat{A}(\varphi_\epsilon) - A(y, P(y, \varphi_\epsilon)) \right] \cdot K \, dx \right| \]
\[ \begin{cases} \int_{\Omega} \left| D_j^{-h}(\varepsilon E_{ji}(y, \varphi_\epsilon)) - \partial_j E_{ji}(y, \varphi_\epsilon) \right| |K_i| \, dx & \text{if } p \in (1, 2) \cup (2, \infty); \quad (3.15) \\
\int_{\Omega} \left| \left\{ \frac{\partial}{\partial x_k} [E_{ji}(x/\varepsilon, \varphi_\epsilon)] \frac{\partial}{\partial x_j} [\varphi_\epsilon^k] \right\} K_i \right| \, dx & \text{if } p = 2; \end{cases} \]
by ignoring a systematic error $o(\varepsilon)$, where the notation $(\varphi_\epsilon)^k$ denotes the $k^{th}$ component of $\varphi_\epsilon$, and $\varphi_\epsilon$ is given in Lemma 3.2.

Remark 3.5. In the case $p = 2$, there is in fact no real systematic error.

Proof. The main idea is to use the antisymmetry of the flux corrector $E$ as one did in the linear equations, which play an important role in quantitative estimates. We first show a proof of the first case of (3.15), and inspired by the equation (2.37),
\[ \int_{\Omega} \left[ \hat{A}(\varphi_\epsilon) - A(y, P(y, \varphi_\epsilon)) \right] \cdot K \, dx = -\int_{\Omega} \frac{\partial}{\partial y_j} [E_{ji}(y, \varphi_\epsilon)] K_i \, dx \]
\[ = \int_{\Omega} \left( D_j^{-h}[\varepsilon E_{ji}(y, \varphi_\epsilon)] - \frac{\partial}{\partial y_j} [E_{ji}(y, \varphi_\epsilon)] \right) K_i \, dx + \varepsilon \int_{\Omega} E_{ji}(y, \varphi_\epsilon) D_j^{-h}(K_i) \, dx, \quad (3.16) \]
where we use “integration by parts” formula for difference quotients,
\[ -\int_{\Omega} D_j^{-h}[\varepsilon E_{ji}(y, \varphi_\epsilon)] K_i \, dx = \varepsilon \int_{\Omega} E_{ji}(y, \varphi_\epsilon) D_j^{-h}(K_i) \, dx. \]
Since $\nabla u_\varepsilon, \nabla u_0 \in L^p(\Omega; \mathbb{R}^d)$ with higher regularities, there may hold the following facts
\[
K_i = D_i^h[u_\varepsilon + u_0 + \varepsilon N(y, \varphi_\varepsilon)] + o(1)\varepsilon_i \quad \text{and} \quad D_j^h D_i^h f = D_i^h D_j^h f \quad (3.17)
\]
in a weak sense, whenever $0 < h \ll 1$. This together with the property (2.23) leads to
\[
\left| \varepsilon \int_{\Omega} E_{ji}(y, \varphi_\varepsilon) D_j^h(K_i)dx \right| \leq o(\varepsilon) \int_{\Omega} |E(y, \varphi_\varepsilon)|dx.
\]
Note that
\[
\int_{\Omega} |E(y, \varphi_\varepsilon)|dx \leq \sum_{k \in I_\varepsilon} |Y_k^i| \int_{Y_\varepsilon} |E(y, \varphi_\varepsilon)|dx \lesssim \int_{\Omega} |\varphi_\varepsilon|dx \lesssim \int_{\Omega} |\nabla u_0|dx,
\]
where we use the estimate (2.28) in the second inequality. Thus, ignoring the quantity
\[
o(\varepsilon) \int_{\Omega} |\nabla u_0|dx,
\]
the estimate (3.16) gives the first line of (3.15). Then we turn to show a proof for its second line.

In terms of the equation (2.35), there holds
\[
\int_{\Omega} [\tilde{A}(\varphi_\varepsilon) - A(y, P(y, \varphi_\varepsilon))] \cdot K dx = -\int_{\Omega} \frac{\partial}{\partial y_j} [E_{ji}(y, \varphi_\varepsilon)] K_i dx \\
= -\int_{\Omega} \frac{\partial}{\partial x_j} [\varepsilon E_{ji}(y, \varphi_\varepsilon)] K_i dx + \varepsilon \int_{\Omega} \frac{\partial}{\partial \xi_k} [E_{ji}(y, \varphi_\varepsilon)] \frac{\partial}{\partial x_j} [(\varphi_\varepsilon)_j] K_i dx. \quad (3.18)
\]
Note that we also have
\[
\frac{\partial}{\partial x_j} [E_{ji}(y, \varphi_\varepsilon)] = D_j^{-h} [E_{ji}(y, \varphi_\varepsilon)] + o(1)\varepsilon_i,
\]
as $h$ is sufficiently small (in such the case, $h$ could be independent of $\varepsilon$). Thus, the above formula together with the formula (3.17) and the antisymmetry property (2.23) gives
\[
\int_{\Omega} \frac{\partial}{\partial x_j} [\varepsilon E_{ji}(y, \varphi_\varepsilon)] K_i dx = \varepsilon \int_{\Omega} D_j^{-h} [E_{ji}(y, \varphi_\varepsilon)] K_i dx + o(\varepsilon)\varepsilon_i \int_{\Omega} K_i dx \\
= \varepsilon \int_{\Omega} D_j^{-h} [E_{ji}(y, \varphi_\varepsilon)] D_i^h f dx + o(\varepsilon) \int_{\Omega} D_j^{-h} [E_{ji}(y, \varphi_\varepsilon)] dx + o(\varepsilon)\varepsilon_i \int_{\Omega} K_i dx \\
= o(\varepsilon)\varepsilon_i \int_{\Omega} K_i dx.
\]
Then, plugging this back into the formula (3.18) leads to our desired estimate in (3.15). We also mention that in the last step above, we employ the following fact that
\[
\int_{\Omega} D_j^{-h} [E_{ji}(y, \varphi_\varepsilon)] dx = \int_{\Omega} D_j^{-h} [E_{ji}(y, \varphi_\varepsilon)] \psi_\varepsilon dx = -\int_{\Omega} [E_{ji}(y, \varphi_\varepsilon)] D_j^h \psi_\varepsilon dx = 0,
\]
where $\varphi_\varepsilon = S_\varepsilon(\psi_4 \varepsilon \nabla u_0)$, and both of $\psi_\varepsilon$ and $\psi_4 \varepsilon$ are cut-off functions defined in (1.27). This ends the proof. $\square$
3.4 Estimates on remainder terms

**Lemma 3.6** \((p > 2)\). Assume the same conditions as in Lemma 3.2. Let \(K\) be defined in (3.3), and \(u_0 \in W^{1,p}(\Omega)\) be with the assumption (1.14). Then by setting \(\bar{K} = K/\|K\|_{L^p(\Omega)}\), we obtain

\[
\left| \int_\Omega [A(y, P(y, \varphi_\varepsilon)) - A(y, V)] \cdot \bar{K} dx \right| \leq \text{the first line of r.h.s. of (3.8)} + \left( \int_{\Omega \setminus \Sigma_{3\varepsilon}} |\nabla u_0|^p dx \right)^{\frac{p-1}{p}}
\]

\[\quad \quad \quad \quad \quad \quad \quad + \left( \int_\Omega |D^{h}(\varepsilon N(y, \varphi_\varepsilon)) - \nabla y N(y, \varphi_\varepsilon)|^p dx \right)^{\frac{1-p}{2}} (\int_\Omega |\nabla u_0|^p dx)^{\frac{p}{2}}
\]

\[\quad \quad \quad \quad \quad \quad \quad + \left( \int_\Omega |D^{h}(\varepsilon N(y, \varphi_\varepsilon)) - \nabla y N(y, \varphi_\varepsilon)|^p dx \right)^{\frac{1-p}{2}} (\int_\Omega |\nabla u_0|^p dx)^{\frac{p}{2}}
\]

\[\quad \quad \quad \quad \quad \quad \quad + \left( \int_\Omega |D^{h}(\varepsilon N(y, \varphi_\varepsilon)) - \nabla y N(y, \varphi_\varepsilon)|^p dx \right)^{\frac{1-p}{2}},
\]

(3.19)

where the up to constant depends only on \(\mu_0, \mu_1, p\) and \(d\).

**Proof.** It follows from the assumption (1.2) that

\[
\left| \int_\Omega [A(y, P(y, \varphi_\varepsilon)) - A(y, V)] \cdot \bar{K} dx \right| \leq \int_\Omega |P(y, \varphi_\varepsilon) - V|(|P(y, \varphi_\varepsilon)| + |V|)^{p-2}|\bar{K}| dx.
\]

(3.20)

Recalling the expression \(V\) in (3.1), the right-hand side above will be controlled by

\[
\int_\Omega |\varphi_\varepsilon - \nabla u_0|(|P(y, \varphi_\varepsilon)| + |\nabla u_0|)^{p-2}|\bar{K}| dx := I_1
\]

\[
+ \int_\Omega |\varphi_\varepsilon - \nabla u_0| |D^{h}(\varepsilon N(y, \varphi_\varepsilon))|^p |\bar{K}| dx := I_2
\]

\[
+ \int_\Omega |\nabla y N(y, \varphi_\varepsilon) - D^{h}(\varepsilon N(y, \varphi_\varepsilon))| |P(y, \varphi_\varepsilon)| + |V|)^{p-2}|\bar{K}| dx := I_3
\]

Before proceeding further, we state the following facts: there holds

\[
|P(x/\varepsilon, \varphi_\varepsilon(x))| \lesssim |\varphi_\varepsilon(x)| \lesssim |\nabla u_0(x)|
\]

(3.21)

for any \(x \in \Sigma_{3\varepsilon}\), according to the estimates (2.5), (3.10) and, Remark 2.3 implies

\[
P(x/\varepsilon, \varphi_\varepsilon(x)) = 0 \quad \forall x \in \Omega \setminus \Sigma_{3\varepsilon}.
\]

(3.22)

Hence, using (3.21), (3.22) one may have

\[
I_1 \lesssim \int_{\Sigma_{3\varepsilon}} |\varphi_\varepsilon - \nabla u_0||\nabla u_0|^{p-2}|\bar{K}| dx + \int_{\Omega \setminus \Sigma_{3\varepsilon}} |\nabla u_0|^p |\bar{K}| dx
\]

\[\quad \quad \quad \quad \quad \quad \quad \lesssim \text{r.h.s. of (3.8)},
\]

(3.23)

where the second inequality is due to the estimate (3.9). By the same token, we also have

\[
\int_\Omega |\varphi_\varepsilon - \nabla u_0||\nabla y N(y, \varphi_\varepsilon)|^{p-2}|\bar{K}| dx
\]

\[\lesssim \int_{\Sigma_{3\varepsilon}} |\varphi_\varepsilon - \nabla u_0||\nabla u_0|^{p-2}|\bar{K}| dx + \int_{\Omega \setminus \Sigma_{3\varepsilon}} |\nabla u_0|^p |\bar{K}| dx \lesssim \text{r.h.s. of (3.8)}.
\]

(3.24)
We now turn to show the estimate of $I_2$, and

$$I_2 \leq \int_{\Omega} |\varphi - \nabla u_0| \left| D^h(\varepsilon N(y, \varphi)) - \nabla_y N(y, \varphi) \right|^{p-2} |\tilde{K}| dx$$

$$+ \int_{\Omega} |\varphi - \nabla u_0| |\nabla_y N(y, \varphi)|^{p-2} |\tilde{K}| dx.$$ 

Note that the estimate of the second term in the right-hand side above has been shown in (3.24). Thus we merely handle its first term, and

$$\int_{\Omega} |\varphi - \nabla u_0| \left| D^h(\varepsilon N(y, \varphi)) - \nabla_y N(y, \varphi) \right|^{p-2} |\tilde{K}| dx$$

$$\lesssim \int_{\Omega \setminus \Sigma_{3c}} |\nabla u_0| \tilde{K} dx + \int_{\Sigma_{3c}} |\nabla u_0| \left| D^h(\varepsilon N(y, \varphi)) - \nabla_y N(y, \varphi) \right|^{p-2} |\tilde{K}| dx$$

by using the facts (3.21) and (3.22) in the first inequality, and the last one follows from Hölder’s inequality. This together with the estimate (3.24) gives

$$I_2 \lesssim \text{r.h.s. of (3.8)} + \left( \int_{\Omega \setminus \Sigma_{3c}} |\nabla u_0|^{\frac{p-1}{p}} dx \right)^{1-\frac{1}{p}}$$

$$\quad + \left( \int_{\Sigma_{3c}} |\nabla u_0| \left| D^h(\varepsilon N(y, \varphi)) - \nabla_y N(y, \varphi) \right|^{p} dx \right)^{1-\frac{2}{p}} \left( \int_{\Omega} |\nabla u_0|^{p} dx \right)^{\frac{1}{p}}. \quad (3.25)$$

We proceed to estimate $I_3$, and it follows from (3.21) that

$$I_3 \lesssim \int_{\Sigma_{3c}} |\nabla_y N(y, \varphi) - D^h(\varepsilon N(y, \varphi))| \left| \nabla u_0 \right|^{p-2} |\tilde{K}| dx := I_{31}$$

$$+ \int_{\Sigma_{3c}} |\nabla_y N(y, \varphi) - D^h(\varepsilon N(y, \varphi))| \left| D^h(\varepsilon N(y, \varphi)) \right|^{p-2} |\tilde{K}| dx. := I_{32}$$

Then we will compute one by one, and the easier one is

$$I_{31} \leq \left( \int_{\Sigma_{3c}} |\nabla_y N(y, \varphi) - D^h(\varepsilon N(y, \varphi))|^{p} dx \right)^{\frac{1}{p}} \left( \int_{\Omega} |\nabla u_0|^{p} dx \right)^{1-\frac{2}{p}}. \quad (3.26)$$

By using triangle’s inequality coupled with Hölder’s inequality, we obtain

$$I_{32} \lesssim \int_{\Sigma_{3c}} |\nabla_y N(y, \varphi) - D^h(\varepsilon N(y, \varphi))|^{p-1} |\tilde{K}| dx$$

$$\quad + \int_{\Sigma_{3c}} |\nabla_y N(y, \varphi) - D^h(\varepsilon N(y, \varphi))| \left| \nabla_y N(y, \varphi) \right|^{p-2} |\tilde{K}| dx$$

$$\lesssim \left( \int_{\Sigma_{3c}} |\nabla_y N(y, \varphi) - D^h(\varepsilon N(y, \varphi))|^{p} dx \right)^{1-\frac{1}{p}} + \text{r.h.s. of (3.26)},$$

where we also use the fact (3.21) in the last step. Obviously, $I_3 \lesssim I_{32}$. Plugging it with (3.23), (3.25) back into the estimate (3.20), we will have the desired estimate (3.19), and have completed the whole proof. □
Lemma 3.7 \((1 < p \leq 2)\). Assume the same conditions as in Lemma 3.6. Then we have

\[
\left| \int_{\Omega} \left[ A(y, P(y, \varphi_\varepsilon)) - A(y, V) \right] \cdot \tilde{K} dx \right| \lesssim \left( \int_{\Omega_{\Sigma_0}} \left| \nabla u_0 \right|^p dx \right)^{\frac{1}{p}} + \varepsilon^{p-1} \left( \int_{\Sigma_4} \left| \nabla^2 u_0 \right|^p dx \right)^{\frac{1}{p}}
\]

\[+ \left( \int_{\Omega_{\Sigma_0}} |D^h(\varepsilon N(y, \varphi_\varepsilon)) - \nabla N(y, \varphi_\varepsilon)|^p dx \right)^{\frac{1}{p}}.\]  

(3.27)

Proof. In terms of the assumption \((1.2)\), we have

\[
\left| \int_{\Omega} \left[ A(y, P(y, \varphi_\varepsilon)) - A(y, V) \right] \cdot \tilde{K} dx \right| \lesssim \int_{\Omega} |P(y, \varphi_\varepsilon) - V|^{p-1} |\tilde{K}| dx
\]

\[\lesssim \left( \int_{\Omega} |P(y, \varphi_\varepsilon) - V|^p dx \right)^{\frac{1}{p}},\]

and the right-hand side is further controlled by

\[
\left( \int_{\Omega} \left| \varphi_\varepsilon - \nabla u_0 \right|^p dx \right)^{\frac{1}{p}} + \left( \int_{\Omega} \left| \nabla y N(y, \varphi_\varepsilon) - D^h(\varepsilon N(y, \varphi_\varepsilon)) \right|^p dx \right)^{\frac{1}{p}}.
\]

Recall that \(\varphi_\varepsilon = S_\varepsilon (\psi_{4\varepsilon} \nabla u_0)\). By \((2.43)\), one may derive that

\[
I \lesssim \left( \int_{\Omega_{\Sigma_0}} \left| \nabla u_0 \right|^p dx \right)^{\frac{1}{p}} + \varepsilon^{p-1} \left( \int_{\Sigma_4} \left| \nabla^2 u_0 \right|^p dx \right)^{\frac{1}{p}},
\]

and this together with the above estimate shows the estimate \((3.27)\). The proof is complete. \(\Box\)

3.5 Layer and co-layer type estimates

Lemma 3.8 (higher regularity estimates for \(u_0\)). Suppose that \(\hat{A}\) satisfies the structure assumption \((1.7)\). Let \(u_0 \in W^{1,p}_0(\Omega)\) be a weak solution of \(L_0 u_0 = 0\) in \(\Omega\), then there hold

\[
\int_{B(0,r)} |\nabla^2 u_0|^2 |\nabla u_0|^{p-2} \lesssim \frac{C}{r^2} \int_{B(0,2r)} \left| \nabla u_0 \right|^p
\]

(3.28)

for the case \(2 \leq p < \infty\), and

\[
\left( \int_{B(0,r)} \left| \nabla^2 u_0 \right|^p \right)^{\frac{1}{p}} \lesssim \frac{C}{r} \left( \int_{B(0,2r)} \left| \nabla u_0 \right|^p \right)^{\frac{1}{p}}
\]

(3.29)

for the case \(1 < p \leq 2\), where \(B(0,2r) \subset \Omega\) for any \(0 < r < 1\). Also, we have

\[
\sup_{B(0,r/2)} \left| \nabla u_0 \right| \lesssim \left( \int_{B(0,r)} \left| \nabla u_0 \right|^p \right)^{\frac{1}{p}}.
\]

(3.30)

Here the up to constant depends on \(\tilde{\mu}_0, \tilde{\mu}_1, p\) and \(d\).

Proof. The proof of the estimates (3.28), (3.29) may be found in [13, pp.267-271], and the estimate (3.30) is included in [22, Theorem 3.19]. \(\Box\)
Lemma 3.9 (layer and co-layer estimates). Let \( u_0 \in W^{1,p}_{\text{loc}}(\Omega) \) be a weak solution of \( \mathcal{L}_0 u_0 = 0 \) in \( \Omega \), with the construction assumption (1.7). Then there holds the layer type estimate
\[
\left( \int_{\Omega \setminus \Sigma_{8\varepsilon}} \left| \nabla u_0 \right|^p dx \right)^{1/p} \leq \varepsilon^{1/p-1/q} \left( \int_{\Omega} \left| \nabla u_0 \right|^q dx \right)^{1/q}
\]
for some \( q > p \). Meanwhile, we have the co-layer type estimate
\[
\left( \int_{\Sigma_{3\varepsilon}} \left| \nabla^2 u_0 \right|^2 \left| \nabla u_0 \right|^{p-2} dx \right)^{1/p} \leq \varepsilon^{1/p-1/q} \left( \int_{\Omega} \left| \nabla u_0 \right|^q dx \right)^{1/q}
\]
for the case \( 2 \leq p < \infty \), and
\[
\left( \int_{\Sigma_{3\varepsilon}} \left| \nabla^2 u_0 \right|^p dx \right)^{1/p} \leq \varepsilon^{1/p-1/q-1} \left( \int_{\Omega} \left| \nabla u_0 \right|^q dx \right)^{1/q}
\]
for the case \( 1 < p \leq 2 \). Here the up to constant depends on \( \tilde{\mu}_0, \tilde{\mu}_1, p, d \) and \( r_0 \).

Proof. These estimates have already been appeared in the linear case, which are based upon Mayer’s estimate and Lemma 3.8. It follows from Hölder’s inequality that
\[
\int_{\Omega \setminus \Sigma_{8\varepsilon}} \left| \nabla u_0 \right|^{p} dx \leq \left| \Omega \setminus \Sigma_{8\varepsilon} \right|^{1-\frac{p}{q}} \left( \int_{\Omega} \left| \nabla u_0 \right|^q dx \right)^{\frac{p}{q}}
\]
and this gives the estimate (3.31). Since \( \delta(x) \approx \delta(y) \), there holds
\[
\int_{\Sigma_{3\varepsilon}} \left| \nabla^2 u_0 \right|^2 \left| \nabla u_0 \right|^{p-2} dx \leq \int_{\Sigma_{3\varepsilon}} \int_{B(x, \delta(x)/4)} \left| \nabla^2 u_0 \right|^2 \left| \nabla u_0 \right|^{p-2} dy dx \leq \int_{\Sigma_{2\varepsilon}} \left| \nabla^2 u_0 \right|^2 \left| \nabla u_0 \right|^{p-2} dx.
\]
Then, we have
\[
\int_{\Sigma_{3\varepsilon}} \left| \nabla^2 u_0 \right|^2 \left| \nabla u_0 \right|^{p-2} dx \leq \int_{\Sigma_{3\varepsilon}} \int_{B(x, \delta(x)/4)} \left| \nabla^2 u_0 \right|^2 \left| \nabla u_0 \right|^{p-2} dy dx
\]
\[
\leq \int_{\Sigma_{3\varepsilon}} \left[ \delta(x) \right]^{-2} \int_{B(x, \delta(x)/2)} \left| \nabla u_0 \right|^{p} dx \leq \int_{\Sigma_{2\varepsilon}} \left[ \delta(x) \right]^{-2} \left| \nabla u_0 \right|^{p} dx
\]
\[
\leq \varepsilon^{1-\frac{p}{q}} \left( \int_{\Omega} \left| \nabla u_0 \right|^q dx \right)^{\frac{p}{q}},
\]
where we also use the estimate (3.28) in the second inequality, and this will imply the estimate (3.32). Using the same argument as in the proof above, we may obtain the estimate (3.33), which is based upon (3.29). We have completed the proof. \( \square \)

Lemma 3.10. Assume the same conditions as in Lemma 2.7. Let \( \varphi_{\varepsilon} = S_{\varepsilon} \psi_{4\varepsilon} \nabla u_0 \), and \( \beta = 1/p - 1/q \) with some \( q > p \), then we have
\[
\left( \int_{\Omega} \left| D_1^h (\varepsilon N(y, \varphi_{\varepsilon})) - \partial_i N(y, \varphi_{\varepsilon}) \right|^p dx \right)^{1/p} \leq \begin{cases} \varepsilon^{(\tau-1)} + \varepsilon^{(1-\alpha)(1-\tau)+\beta} \left( \int_{\Omega} \left| \nabla u_0 \right|^q dx \right)^{\frac{1}{q}} & \text{if } 2 < p < \infty; \\ \varepsilon^{(\tau-1)} + \varepsilon^{(1-\alpha)(1-\tau)+\alpha\beta} \left( \int_{\Omega} \left| \nabla u_0 \right|^q dx \right)^{\frac{1}{q}} & \text{if } 1 < p < 2, \end{cases}
\]
where \( 1 < \tau < (1 + \beta - \alpha)/(1 - \alpha) \) if \( 2 < p < \infty \), and \( 1 < \tau < (1 + \alpha\beta - \alpha)/(1 - \alpha) \) if \( 1 < p < 2 \). Here the up to multiplicative constant will additionally depend on \( \tilde{\mu}_0, \tilde{\mu}_1 \) and \( r_0 \).
Proof. We first prove the first line of the estimate (3.35), and it follows from the estimate (2.18) that
\[
\left( \int_{\Omega} \left| D_i^h(\varepsilon N(y, \varphi_\varepsilon)) - \partial_i N(y, \varphi_\varepsilon) \right|^p \, dx \right)^{\frac{1}{p}} \leq \varepsilon^{\theta(\tau - 1)} \left( \int_{\Omega} \left| \nabla u_0 \right|^p \, dx \right)^{\frac{1}{p}} + \varepsilon^{1-(1-\alpha)\tau} \left( \int_{\Sigma_{4\varepsilon}} \left| \nabla u_0 \right|^{p-2} \left| D_i^h(\psi_{4\varepsilon} \nabla u_0) \right|^2 \, dx \right)^{\frac{1}{p}},
\]
where we use the facts (3.10), and
\[
D_i^h[S_\varepsilon(\psi_{4\varepsilon} \nabla u_0)] = S_\varepsilon(D_i^h(\psi_{4\varepsilon} \nabla u_0)).
\]
Then we turn to the second term in the right-hand side above, and
\[
\int_{\Sigma_{4\varepsilon}} \left| \nabla u_0 \right|^{p-2} \left| D_i^h(\psi_{4\varepsilon} \nabla u_0) \right|^2 \, dx = \int_{\Sigma_{4\varepsilon}} \left| \nabla u_0 \right|^{p-2} \left| D_i^h(\psi_{4\varepsilon} \nabla u_0 + \psi_{4\varepsilon}(x + \text{the}_i)D_i^h(\nabla u_0) \right|^2 \, dx
\]
\[
\leq \varepsilon^{-2} \int_{\Omega \setminus \Sigma_{9\varepsilon}} \left| \nabla u_0 \right|^p \, dx + \int_{\Sigma_{3\varepsilon}} \left( \int_{\Omega \setminus \Sigma_{9\varepsilon}} \left| \nabla u_0 \right|^q \, dx \right)^{\frac{1}{q}} \left| \nabla u_0 \right|^2 \, dx
\]
by noting that
\[
|D_i^h(\psi_{4\varepsilon})(x)| \leq \int_0^1 |\nabla \psi_{4\varepsilon}(x + \text{the}_i)| \, dt \leq C/\varepsilon, \quad \text{and} \quad \text{supp}(D_i^h(\psi_{4\varepsilon})) \subset \Sigma_{3\varepsilon} \setminus \Sigma_{9\varepsilon}.
\]
Using the same argument as in the proof of Lemma 3.9, we can easily carry out the proof this theorem, and we just mention that the estimates (3.28) and (3.29) are in fact proved by an analysis of certain difference quotients (see for example [13, Theorem 8.1]). In terms of $I_1$, it follows from Hölder’s inequality that
\[
I_1 \leq \varepsilon^{-2}|\Omega \setminus \Sigma_{9\varepsilon}|^{1 - \frac{q}{p}} \left( \int_{\Omega} \left| \nabla u_0 \right|^q \right)^{\frac{1}{q}} \leq \varepsilon^{-1 + \frac{2}{q}} \left( \int_{\Omega} \left| \nabla u_0 \right|^q \right)^{\frac{1}{q}},
\]
and we proceed to show
\[
I_2 \lesssim \int_{\Sigma_{3\varepsilon}} \int_{B(x, \delta(x)/4)} \left| \nabla u_0 \right|^{p-2} \left| D_i^h(\nabla u_0) \right|^2 \, dy \, dx \lesssim \int_{\Sigma_{3\varepsilon}} \left[ \delta(x) \right]^{-2} \int_{B(x, \delta(x)/2)} \left| \nabla u_0 \right|^p \, dy \, dx
\]
\[
\lesssim \int_{\Sigma_{3\varepsilon}} \left[ \delta(x) \right]^{-2} \left| \nabla u_0 \right|^p \, dx \lesssim \varepsilon^{-1 + \frac{2}{q}} \left( \int_{\Omega} \left| \nabla u_0 \right|^q \right)^{\frac{1}{q}}
\]
where the first and third inequalities are due to the fact (3.34), and the second one is based upon the estimate (3.28). Clearly, the above two estimates give
\[
\left( \int_{\Sigma_{4\varepsilon}} \left| \nabla u_0 \right|^{p-2} \left| D_i^h(\psi_{4\varepsilon} \nabla u_0) \right|^2 \, dx \right)^{\frac{1}{p}} \lesssim \varepsilon^{-\frac{1}{p} - \frac{2}{q}} \left( \int_{\Omega} \left| \nabla u_0 \right|^q \, dx \right)^{\frac{1}{q}}.
\]
Then by noting that $-(1/p + 1/q) = -\alpha + \beta$, we have proved
\[
\left( \int_{\Omega} \left| D_i^h(\varepsilon N(y, \varphi_\varepsilon)) - \partial_i N(y, \varphi_\varepsilon) \right|^p \, dx \right)^{\frac{1}{p}} \lesssim \left\{ \varepsilon^{\theta(\tau - 1)} + \varepsilon^{(1-\alpha)(1-\tau) + \beta} \right\} \left( \int_{\Omega} \left| \nabla u_0 \right|^q \, dx \right)^{\frac{1}{q}}.
\]
in the case of $2 \leq p < \infty$. This requires the condition $1 < \tau < (1 + \beta - \alpha)/(1 - \alpha)$. We proceed to show the second line of (3.35), and it follows from the estimate (2.18) that
\[
\left( \int_\Omega \left| D_i^h(\varepsilon N(y, \varphi_\varepsilon)) - \partial_i N(y, \varphi_\varepsilon) \right|^p dx \right)^{\frac{1}{p}} 
\leq \varepsilon^{\theta(\tau-1)} \left( \int_\Omega |\nabla u_0|^p dx \right)^{\frac{1}{p}} + \varepsilon^{1-(\alpha-\beta)} \left( \int_\Omega |\nabla u_0|^p dx \right)^{\frac{1}{p}} \left( \int_\Omega | D_i^h(\varepsilon N(y, \varphi_\varepsilon))|^p dx \right)^{\frac{1}{p}}.
\]

By a similar argument, there holds
\[
\left( \int_{\Sigma_{\delta}} | D_i^h(\varepsilon N(y, \varphi_\varepsilon)) - \partial_i N(y, \varphi_\varepsilon) |^p dx \right)^{\frac{1}{p}} \leq \left\{ \varepsilon^{-\frac{1}{q}} \left( \int_\Omega |\nabla u_0|^p dx \right)^{\frac{1}{p}} + \left( \int_{\Sigma_{\delta}} | D_i^h(\varepsilon N(y, \varphi_\varepsilon))|^p dx \right)^{\frac{1}{p}} \right\}^{\frac{1}{p^{(3-p)}}},
\]
where we use the estimate (3.29) in the second inequality. Noting that $1/p - 1/q = \alpha(\beta - 1)$, we consequently obtain
\[
\left( \int_\Omega \left| D_i^h(\varepsilon N(y, \varphi_\varepsilon)) - \partial_i N(y, \varphi_\varepsilon) \right|^p dx \right)^{\frac{1}{p}} \leq \left\{ \varepsilon^{\theta(\tau-1)} + \varepsilon^{(1-\alpha)(1-\tau)+\alpha\beta} \right\} \left( \int_\Omega |\nabla u_0|^q dx \right)^{\frac{1}{q}}
\]
with $1 < \tau < \frac{1-\alpha+\alpha\beta}{1-\alpha}$, where $\beta = 1/p - 1/q$. We have completed the proof.

**Lemma 3.11.** Assume the same conditions as in Lemma 2.11. Let $\varphi_\varepsilon = S_\varepsilon(\psi_\varepsilon \nabla u_0)$, and $\beta = 1/p - 1/q$ with some $q > p$, then we have
\[
\left( \int_\Omega \left| D_j^{-h}(\varepsilon E_{ji}(y, \varphi_\varepsilon)) - \partial_j E_{ji}(y, \varphi_\varepsilon) \right|^p dx \right)^{\frac{1}{p}} \leq \left\{ \varepsilon^{\theta(\tau-1)} + \varepsilon^{(1-\gamma)(1-\tau)+\beta} \right\} \left( \int_\Omega |\nabla u_0|^q dx \right)^{\frac{1}{q}} \text{ if } 2 < p < \infty;
\]
\[
\left\{ \varepsilon^{\theta(\tau-1)} + \varepsilon^{(1-\gamma)(1-\tau)+\beta} \right\} \left( \int_\Omega |\nabla u_0|^q dx \right)^{\frac{1}{q}} \text{ if } 1 < p < 2,
\]
where $1 < \tau < (1 - \gamma + \beta)/(1 - \gamma)$ if $2 < p < \infty$, and $1 < \tau < (1 - \gamma + \gamma\beta)/(1 - \gamma)$ if $1 < p < 2$. Here the up to constant depends on $\mu_0, \mu_1, \mu_0, \mu_1, \mu_2, p, d$ and $r_0$.

**Proof.** The proof is similar to that given in the previous lemma. and we omit it.

**Remark 3.12.** Let $p = 2$, and $\beta = 1/2 - 1/q$ with some $q > 2$. Then we have
\[
\int_\Omega \left| \frac{\partial}{\partial \xi_k} E_{ji}(x/\varepsilon, \varphi_\varepsilon) \right| \left| \frac{\partial}{\partial x_j} (\varphi_\varepsilon) K_i \right| \frac{dx}{\varepsilon} \leq \varepsilon^{\frac{1}{q}(1+\frac{2}{q})} \left( \int_\Omega |\nabla u_0|^q dx \right)^{\frac{1}{q}} \left( \int_\Omega |K|^2 dx \right)^{\frac{1}{2}},
\]
where we just note that $|\nabla \xi_k E_{ji}(y, \cdot)| \leq C$ for a.e. $y \in \mathbb{R}^d$ by the estimate (2.11). The same result might be found in [29, Theorem 3.4].
3.6 Proof of Theorem 1.1

The estimate \((1.13)\) has already been proved in Lemma 2.2. Here we just focus on the error estimates stated in the theorem. For the ease of the statement. Let \(K = \nabla u_\varepsilon - V\) as in \((3.3)\), and \(\tilde{K} = K / \|K\|_{L^p(\Omega)}\). The proof is divided into two parts. One is to handle the singular case \((1 < p < 2)\). The other is to deal with the degenerated case \((2 < p < \infty)\).

Part 1 \((1 < p < 2)\). We show the estimate \((1.15)\), and the proof mainly consists of three steps according to Lemma 3.2. The first step is to give the estimate

\[
\int_\Omega |\hat{A}(\nabla u_0) - \hat{A}(\varphi_\varepsilon)| \tilde{K} \, dx
\lesssim |\Omega_{M,\varepsilon}|^{\beta(p-1)} \|\nabla u_0\|_{L^q(\Omega_{M,\varepsilon})}^{p-1} + \left\{ \varepsilon^{\beta(p-1)} + \varepsilon^{\beta \gamma} \right\} \|\nabla u_0\|_{L^p(\Omega)}^{p-1}
\lesssim |\Omega_{M,\varepsilon}|^{\beta(p-1)} \|\nabla u_0\|_{L^q(\Omega_{M,\varepsilon})}^{p-1} + \varepsilon^{\beta \gamma} \|\nabla u_0\|_{L^p(\Omega)}^{p-1},
\]

whose first inequality follows from the second line of \((3.8)\) coupled with the layer and co-layer estimates \((3.31)\) and \((3.33)\).

Then the second step is to show

\[
\left| \int_\Omega [\hat{A}(\varphi_\varepsilon) - A(y, P(y, \varphi_\varepsilon))] \cdot \tilde{K} \, dx \right| \lesssim \left\{ \varepsilon^{\rho(\tau-1)} + \varepsilon^{(1-\gamma)(1-\tau) + \beta \gamma} \right\} \|\nabla u_0\|_{L^p(\Omega)}^{p-1}
\lesssim \max \left\{ \varepsilon^{\rho(\tau-1)}, \varepsilon^{(1-\gamma)(1-\tau) + \beta \gamma} \right\} \|\nabla u_0\|_{L^p(\Omega)}^{p-1},
\]

where \(\rho = \theta(p-1)\) (defined in Lemma 2.9), and we employ the estimates \((3.15)\) and \((3.38)\).

In the third step, combining the estimates \((3.27)\) and \((3.35)\) together with \((3.31)\) and \((3.33)\) leads to

\[
\left| \int_\Omega [A(y, P(y, \varphi_\varepsilon)) - A(y, V)] \cdot \tilde{K} \, dx \right|
\lesssim \left\{ \varepsilon^{\rho(\tau-1)} + \varepsilon^{\beta(\tau-1)(p-1) + \beta \gamma} + \varepsilon^{(1-\alpha)(1-\tau)(p-1) + \beta \gamma} \right\} \|\nabla u_0\|_{L^p(\Omega)}^{p-1}
\lesssim \max \left\{ \varepsilon^{(1-\alpha)(1-\tau)(p-1) + \beta \gamma}, \varepsilon^{\rho(\tau-1)(p-1) + \beta \gamma} \right\} \|\nabla u_0\|_{L^p(\Omega)}^{p-1},
\]

in which the second inequality is due to the fact \(\beta(p-1) > (1-\alpha)(1-\tau)(p-1) + \gamma \beta\).

Hence, in view of the estimates \((3.41)\) and \((3.42)\), the exact range of \(\tau\) might be figured out by the following three cases:

(a) \(0 < \rho(\tau-1) < (1-\gamma)(1-\tau) + \beta \gamma < (1-\alpha)(1-\tau)(p-1) + \gamma \beta\);

(b) \(0 < (1-\gamma)(1-\tau) + \beta \gamma < \rho(\tau-1) < (1-\alpha)(1-\tau)(p-1) + \gamma \beta\);

(c) \(0 < (1-\gamma)(1-\tau) + \beta \gamma < (1-\alpha)(1-\tau)(p-1) + \gamma \beta < \rho(\tau-1)\).

In the case (a), it follows from \((3.40), (3.41), (3.42)\) and \((3.4)\) that

\[
\|K\|_{L^p(\Omega)}^{p-1} \leq |\Omega_{M,\varepsilon}|^{\beta(p-1)} \|\nabla u_0\|_{L^q(\Omega_{M,\varepsilon})}^{p-1} + \varepsilon^{\beta(\tau-1)(p-1)} \|\nabla u_0\|_{L^p(\Omega)}^{p-1},
\]

and this gives the stated estimate \((1.15)\). Meanwhile, it is not hard to see that the range \((1.16)\) follows from the relationship (a). By the same token, we can derive the estimate \((1.17)\) on account of
the cases (b) and (c). However, as stated in (1.18), we have to pay more careful on the computations for the corresponding range of \( \tau \). From (b), it follows that

\[
\frac{1 - \gamma + \gamma \beta + \rho}{1 - \gamma + \rho} < \tau \leq \min\left\{ \frac{1 - \alpha + \alpha \beta + \theta}{1 - \alpha + \theta}, \frac{1 - \gamma + \gamma \beta}{1 - \gamma} \right\},
\]

and keep the above inequality for a while. In terms of (c), we may have

\[
\begin{cases}
1 < \tau < \frac{1 - \gamma + \gamma \beta}{1 - \gamma} ; \\
1 - \alpha + \alpha \beta + \theta < \tau < \frac{1 - \alpha + \alpha \beta}{1 - \alpha},
\end{cases}
\]

under the condition \( \frac{2 + \delta}{1 + \delta} < p < 2 \), since such the condition guarantees the following inequality

\[
\frac{1 - \alpha + \alpha \beta + \theta}{1 - \alpha + \theta} < \tau < \frac{1 - \gamma + \gamma \beta}{1 - \gamma}.
\]

Recalling (3.43), we have the range

\[
\frac{1 - \gamma + \gamma \beta + \rho}{1 - \gamma + \rho} < \tau < \frac{1 - \alpha + \alpha \beta + \theta}{1 - \alpha + \theta}.
\]

under the same condition, and this together with (3.44) gives (1.18). For the case \( 1 < p \leq \frac{2 + \delta}{1 + \delta} \), the stated range (1.18) also follows from (3.43), and in fact (c) can not happen in such the case. The proof of Part 1 is done.

Part 2. In this part \( (2 < p < \infty) \), we will address the estimate (1.19) and the range (1.20) of \( \tau \). We first show that

\[
\int_{\Omega} |\hat{A}(\nabla u_{0}) - \hat{A}(\varphi_{\varepsilon}) - \hat{A}(y, P(y, \varphi_{\varepsilon}))| \cdot K \, dx \lesssim |\Omega_{M, \varepsilon}|^{\beta(p-1)} \| \nabla u_{0} \|_{L^{\beta}(\Omega_{M, \varepsilon})}^{p-1} + \left\{ \varepsilon^{\beta(p-1)} + \varepsilon^{\beta/\alpha} \right\} \| \nabla u_{0} \|_{L^{\beta}(\Omega)}^{p-1}
\]

where we use the first line of (3.8) coupled with the layer and co-layer type estimates (3.31), (3.32). Then, we employ the first line of (3.15) and (3.38) to obtain

\[
\left| \int_{\Omega} \left[ \hat{A}(\varphi_{\varepsilon}) - A(y, P(y, \varphi_{\varepsilon})) \right] \cdot K \, dx \right| \lesssim \varepsilon^{\beta(\tau-1)} \left( \int_{\Omega} |\nabla u_{0}|^{q} \, dx \right)^{\frac{p-1}{q}} + \varepsilon^{(1-\gamma)(1-\tau)+\beta} \left( \int_{\Omega} |\nabla u_{0}|^{q} \, dx \right)^{\frac{p-1}{q}}
\]

under the condition (3.48), where we note that \( \gamma = \alpha \) in the case of \( p > 2 \). Furthermore, let

\[
R_{M, \varepsilon} := |\Omega_{M, \varepsilon}|^{\beta(p-1)} \| \nabla u_{0} \|_{L^{\beta}(\Omega_{M, \varepsilon})}^{p-1},
\]

and it follows from the estimates (3.19) and (3.35) that

\[
\left| \int_{\Omega} \left[ A(y, P(y, \varphi_{\varepsilon})) - A(y, V) \right] \cdot K \, dx \right| \lesssim \left\{ \varepsilon^{\beta/\alpha} + \varepsilon^{1-\alpha+\beta} + \varepsilon^{(p-2)(\tau-1)} + \varepsilon^{(p-2)(1-\alpha)(1-\tau)+\beta} \right\} \| \nabla u_{0} \|_{L^{\beta}(\Omega)}^{p-1} + R_{M, \varepsilon}
\]

\[
\lesssim \left\{ \varepsilon^{\beta/\alpha} + \varepsilon^{1-\alpha+\beta} + \varepsilon^{(p-2)(1-\alpha)(1-\tau)+\beta} \right\} \| \nabla u_{0} \|_{L^{\beta}(\Omega)}^{p-1} + R_{M, \varepsilon}
\]

\[
\lesssim \varepsilon^{(1-\alpha)(1-\tau)+\beta} \| \nabla u_{0} \|_{L^{\beta}(\Omega)}^{p-1} + R_{M, \varepsilon},
\]
where we also use the estimates (3.31), (3.32) in the first inequality, and the fact that

\[
0 < (1 - \alpha)(1 - \tau) + \beta < \theta(\tau - 1) \iff \frac{1 - \alpha + \beta + \theta}{1 - \alpha + \theta} < \tau < \frac{1 - \alpha + \beta}{1 - \alpha}
\]

(3.48)
in the second one. The last inequality of (3.47) is due to

\[
[(1 - \alpha)(1 - \tau) + \beta](p - 2) < \beta/\alpha < 1 - \alpha + \beta
\]

whose second inequality follows from the fact \(0 < \beta < \alpha\). Here we also point out that the estimate (3.47) merely holds for the case of \(2 < p < 3\). Thus, the estimate (3.4) together with the estimates (3.45), (3.46) and (3.47) implies that

\[
\|K\|_{L^p(\Omega)} \leq \varepsilon^{\frac{(p-2)(1-\alpha)(1-\tau)+\beta}{p-1}} \|\nabla u_0\|_{L^q(\Omega)} + |\Omega_{M, \varepsilon}|^\beta \|\nabla u_0\|_{L^q(\Omega_{M, \varepsilon})},
\]

which is exactly the first line of (1.19). Concerning the case \(p \geq 3\), we merely make a few modifications on the estimate (3.47), which is

\[
\left| \int_\Omega \left[ A(y, P(y, \varphi_\varepsilon)) - A(y, V) \right] \cdot \tilde{K} \, dx \right| \lesssim \left\{ \varepsilon^{\beta/\alpha} + \varepsilon^{1-\alpha+\beta} + \varepsilon^{\theta(\tau-1)} + \varepsilon^{(1-\alpha)(1-\tau)+\beta} \right\} \|\nabla u_0\|_{L^q(\Omega)}^{p-1} + R_{M, \varepsilon}
\]

(3.49)

where we use the estimates (3.19) and (3.35) in the case of \(p \geq 3\), coupled with the layer and co-layer type estimates (3.31) and (3.32). Also, the condition (3.48) and the fact \(0 < \beta < \alpha\) guarantee the second inequality of (3.49). Thus, combining the estimates (3.4), (3.49) and (3.45) leads to the second line of the stated estimate (1.19). We end this part by mention that (3.48) is exactly the stated range (1.20).

In the end, we address the special case \(p = 2\), and proceed as in the proof of Part 1. The first step is given by

\[
\int_\Omega |\tilde{A}(\nabla u_0) - \hat{A}(\varphi_\varepsilon)\|\tilde{K}\| dx \lesssim \varepsilon^{\beta} \|\nabla u_0\|_{L^q(\Omega)},
\]

where we use the first line of the estimate (3.8) coupled with the estimates (3.31) and (3.32). Then, it follows from the estimates (3.15) and (3.39) that

\[
\left| \int_\Omega \left[ \hat{A}(\varphi_\varepsilon) - A(y, P(y, \varphi_\varepsilon)) \right] \cdot \tilde{K} \, dx \right| \lesssim \varepsilon^{\beta} \|\nabla u_0\|_{L^q(\Omega)}.
\]

Also, in view of the estimates (3.27) and (2.22), we consequently have

\[
\left| \int_\Omega \left[ A(y, P(y, \varphi_\varepsilon)) - A(y, V) \right] \cdot \tilde{K} \, dx \right| \lesssim \varepsilon^{\beta} \|\nabla u_0\|_{L^q(\Omega)}^{p-1},
\]

and the above three estimates together with Lemma 3.2 gives the desired estimate. We end the proof by mention that, in such the case \(p = 2\), the size of \(h\) might be independent of \(\varepsilon\). \(\square\)
4 Large scale estimates

4.1 An approximating argument

**Lemma 4.1** (Caccioppoli’s inequality). Let $1 < p < \infty$. Suppose $A$ satisfies (1.1) and (1.2), and $u_\varepsilon \in W^{1,p}_{\text{loc}}(\Omega)$ is a weak solution of $L_\varepsilon u_\varepsilon = 0$ in $\Omega$, then for any $B(x,r) \subset \Omega$, there holds

$$
\left( \int_{B(x,r)} |\nabla u_\varepsilon|^p \right)^{1/p} \leq \frac{C}{r} \left( \int_{B(x,2r)} |u_\varepsilon - c|^p \right)^{1/p}
$$

(4.1)

for any constant $c \in \mathbb{R}$, where $C$ depends on $\mu_0, \mu_1, p$ and $d$.

**Proof.** On account of the assumption (1.1) and (1.2), there hold

$$
\langle A(y, \xi), \xi \rangle \geq \mu_0 |\xi|^p; \quad |A(y, \xi)| \leq \mu_1 |\xi|^{p-1} \quad \forall y, \xi \in \mathbb{R}^d.
$$

(4.2)

By definition of the weak solution, we have

$$
\int_{\Omega} A(x/\varepsilon, \nabla u_\varepsilon) \nabla \varphi dx = 0 \quad \forall \varphi \in W^{1,p}_0(\Omega).
$$

Let $\varphi = \psi^p_r(u_\varepsilon - c)$, where $\psi_r \in C^1(\Omega)$ is a cut-off function satisfying $\psi_r = 1$ in $B(x, r)$ and $\psi_r = 0$ outside $B(x, 2r)$ with $|\nabla \psi_r| \leq C/r$. Thus, a routine computation leads to

$$
0 = \int_{\Omega} \psi^p_r A(x, \nabla u_\varepsilon) \nabla u_\varepsilon dx + p \int_{\Omega} \psi^{p-1}_r (u_\varepsilon - c) A(x, \nabla u_\varepsilon) \nabla \psi_r dx
$$

$$
\geq \mu_0 \int_{\Omega} \psi^p_r |\nabla u_\varepsilon|^p dx - \mu_1 p \left( \int_{\Omega} \psi^p_r |\nabla u_\varepsilon|^p dx \right)^{1/p} \left( \int_{\Omega} |\nabla \psi_r|^p |u_\varepsilon - c|^p dx \right)^{1/p}
$$

$$
\geq \frac{\mu_0}{2} \int_{\Omega} \psi^p_r |\nabla u_\varepsilon|^p dx - C(\mu_0, \mu_1, p, d) \int_{\Omega} |\nabla \psi_r|^p |u_\varepsilon - c|^p dx,
$$

where we use Young’s inequality in the last step. This implies the stated estimate (4.1) and we may end the proof here. \qed

**Remark 4.2.** The above estimate in fact implies the so-called Meyer’s estimate. There exists a small number $\delta \in (0, 1)$, such that

$$
\left( \int_{B(x,r)} |\nabla u_\varepsilon|^{p(1+\delta)} \right)^{1/(1+\delta)} \lesssim \left( \int_{B(x,2r)} |\nabla u_\varepsilon|^p \right)^{1/p}.
$$

(4.3)

It is also known as a self-improvement property, and relies on the reverse Hölder inequality developed by M. Giaquinta and G. Modica (see for example [18, Theorem 6.38]). We usually choose $q = (1+\delta)p$ in the paper.

In particular, it has been known from Lemma 2.4 that $\hat{A}$ admits the same property as (4.2). Given $g \in W^{1-1/q,q}(\partial \Omega)$, suppose that $u_0 \in W^{1,p}(\Omega)$ is a weak solution of $L_0 u_0 = 0$ in $\Omega$ and $u_0 = g$ on $\partial \Omega$. Then one may have $|\nabla u_0| \in L^q(\Omega)$, and

$$
\|\nabla u_0\|_{L^q(\Omega)} \leq C \|g\|_{W^{1-1/q,q}(\partial \Omega)},
$$

(4.4)

where $C$ depends on $\mu_0, \mu_1, p, d$ and $\Omega$. We refer the reader to [29, Theorem 2.13] or [11, Lemma 4.1] for an analogy argument for the proof, which reduces the problem to establish the corresponding boundary Caccioppoli’s inequality, and the details is left to the reader.
Lemma 4.3 (approximating lemma). Let $1 < p < \infty$, and $B = B(0, R) \subset \Omega$ with $\varepsilon \leq R \leq 1$. Assume the same conditions as in Theorem 1.2. Suppose that $u_{\varepsilon} \in W^{1,p}(2B)$ satisfies $\mathcal{L}_u u_{\varepsilon} = 0$ in $2B$. Then there exist a function $v_R \in W^{1,p}(B)$ satisfying $\|\nabla v_R\|_{L^\infty(B)} \leq M$, and $\eta \in (0, 1)$, such that

$$
\left( \int_{B} |u_{\varepsilon} - v_R|^p \right)^{1/p} \lesssim \left\{ \left( \frac{\varepsilon}{R} \right)^{\eta} + \left\{ x \in B : |\nabla v_R| \leq M(\varepsilon/R)^\beta \right\} \left( \int_{2B} |u_{\varepsilon}|^p \right)^{1/p},
\right.
$$

where the up-to-constant is independent of $\varepsilon$ and $R$.

Proof. The proof is based upon the estimate (1.17), and there are two steps to complete the whole argument. Here we set $\eta = \alpha \beta + \frac{(1-\alpha)(1-\beta)}{p-1}$, and construct $v \in W^{1,p}(B)$ satisfying $\mathcal{L}_v v = \mathcal{L}_u u_{\varepsilon} = 0$ in $B$ with $v_R = u_{\varepsilon}$ on $\partial B$. Since $u_{\varepsilon} \in C^{1,\theta}_{loc}(2B)$, it is fine to assume that $\|\nabla v_R\|_{L^\infty(B)} \leq M$.

Step 1. In the special case of $R = 1$, we may show the estimate (4.5). For any fixed $i = 1, \cdots, d$, there holds

$$
\|u_{\varepsilon} - v_1\|_{L^p(B)} = \|\nabla (u_{\varepsilon} - v_1)\|_{W^{-1,p}(B)} \leq \|\nabla u_{\varepsilon} - \nabla v_1 - D_i^{h}(\varepsilon N(\cdot, \varphi_{\varepsilon}))\|_{W^{-1,p}(B)} + \|D_i^{h}(\varepsilon N(y, \varphi_{\varepsilon}))\|_{W^{-1,p}(B)},
$$

where $h = \varepsilon^\tau$ with $\tau$ fixed in (1.18), and $\varphi_{\varepsilon} = S_{\varepsilon}(\psi_{\varepsilon} \nabla v_1)$ with $\psi_{\varepsilon}$ being the corresponding cut-off function. For the ease of the statement, we denote by

$$
\Omega_{M,\varepsilon}^R := \{ x \in B(0, R) : |\nabla v_R| \leq M\varepsilon \}
$$

Thus, it follows from the estimate (1.17) that

$$
I_1 \leq \|\nabla u_{\varepsilon} - \nabla v_1 - D_i^{h}(\varepsilon N(\cdot, \varphi_{\varepsilon}))\|_{L^p(B)} \lesssim \left\{ \varepsilon^{\eta} + |\Omega^1_{M,\varepsilon}| \right\} \|\nabla v_1\|_{L^q(B)} \lesssim \left\{ \varepsilon^{\eta} + |\Omega^1_{M,\varepsilon}| \right\} \|\nabla u_{\varepsilon}\|_{L^q(B)},
$$

where we employ the $L^q$ estimate (4.4) coupled with the trace theorem in the third line. Note that $u_{\varepsilon}$ up to a constant is still a solution to $\mathcal{L}_u u_{\varepsilon} = 0$, which leads to the estimate in the brace. The fourth line comes from Meyer’s estimate (4.3). Then we turn to show a computation on $I_2$. By definition,

$$
I_2 = \sup_{\|\phi\|_{W^{1,p'}(B)} \leq 1} \left| \int_B \varepsilon N(\cdot, \varphi_{\varepsilon}) D^{-h}(\phi) dx \right| \leq C \varepsilon \left( \int_B \|N(\cdot, \varphi_{\varepsilon})\|_p dx \right)^{1/p} \|\nabla \phi\|_{L^{p'}(B)} \lesssim \varepsilon \|\nabla v_1\|_{L^p(B)} \lesssim \varepsilon \|\nabla u_{\varepsilon}\|_{L^p(B)},
$$

in which we note that $|D_i^{h}(\varepsilon N(\cdot, \varphi_{\varepsilon}))|$ is an integrable function due to the estimate (2.13) under our assumption on the size of $h$. The second inequality might be derived from the estimate (3.7).

Hence, combining the above estimates, we have

$$
\|u_{\varepsilon} - v_1\|_{L^p(B)} \lesssim \left\{ \varepsilon^{\eta} + |\Omega^1_{M,\varepsilon}| + \varepsilon \right\} \|\nabla u_{\varepsilon}\|_{L^p(\Omega_{M,\varepsilon}^B)} \lesssim \left\{ \varepsilon^{\eta} + |\Omega^1_{M,\varepsilon}| \right\} \|u_{\varepsilon}\|_{L^p(2B)}
$$

where the second step follows from Caccioppoli’s inequality (4.1).
Step 2. We show that the estimate (4.5) is also valid for \( \varepsilon \leq R < 1 \), by a rescaling argument. Let \( \tilde{u}_\varepsilon(y) = u_\varepsilon(Ry) = u_\varepsilon(x) \), where \( y \in B(0, 2) \) and \( x \in B(0, 2R) \). On account of the assumption (1.1), we have

\[
\text{div}_y A(y/\varepsilon', \nabla_y \tilde{u}_\varepsilon) = 0 \quad \text{in} \quad B(0, 2), \quad \varepsilon' = \varepsilon/R.
\]

We set \( v_R(x) = v_1(x/R) \) for any \( x \in B(0, R) \), and denote by \( \hat{B}(\xi) = R^{p-1}\hat{A}(R^{-1}\xi) \) for any \( \xi \in \mathbb{R}^d \). Obviously, the operator \( \hat{B} \) still satisfies the structure (1.7). Then we have

\[
\text{div}_y \hat{B}(\nabla_y v_1) = \text{div}_y A(y/\varepsilon', \nabla_y \tilde{u}_\varepsilon) = 0 \quad \text{in} \quad B(0, 1) \quad v_1 = \tilde{u}_\varepsilon \quad \text{on} \partial B(0, 1),
\]

with the condition \( \|\nabla_y v_1\|_{L^\infty(B(0,1))} \leq RM \). By last step, it follows from the estimate (4.7) that

\[
\|\tilde{u}_\varepsilon - v_1\|_{L^p(B(0,1))} \lesssim \left\{ (\varepsilon')^\eta + |\Omega_{RM, \varepsilon'}^1|^\beta \right\} \|\tilde{u}_\varepsilon\|_{L^p(B(0,2))},
\]

and this implies

\[
\|u_\varepsilon - v_R\|_{L^p(B(0,R))} \lesssim \left\{ (R/\varepsilon)^\eta + |\Omega_{M, (\varepsilon/R)}^R|^\beta \right\} \|u_\varepsilon\|_{L^p(B(0,2R))},
\]

where we just mention that \( \Omega_{M, (\varepsilon/R)}^R = \{x \in B(0, R) : |\nabla v_R| \leq M(\varepsilon/R)^\theta \} \) by definition. Multiplying \( R^{-\frac{d}{p}} \) on the both sides above leads to the stated estimate (4.5), and we have completed the proof. \( \square \)

### 4.2 Proof of Theorem 1.2

The main idea may be found in [26, Theorem 5.2] or [29, Theorem 5.4], and we make a modification of the original proof to fit the situation here. To do so, we set \( \sigma \in (0, 1) \), and

\[
H(r, \sigma, v) = r^{-\sigma} \left\{ \left( \int_{B(\theta r, r)} |v|^p \right)^{1/p} \right\}.
\]

For each \( r \in [\varepsilon, 1/2] \), let \( v_r \) be the function given in Lemma 4.3. Then, in terms of the interior Lipschitz estimate (3.30), there holds

\[
H(\theta r, \sigma, v_r) \leq C \theta^{1-\sigma} H(r, \sigma, v_r) \quad (4.8)
\]

for any \( \theta \in (0, 1/4) \), where we just use Poincaré’s inequality and Caccioppoli’s inequality in the computation. Recall the notation \( \Omega_{M, (\varepsilon/r)}^r \) in (4.6), and we note that

\[
|\Omega_{M, (\varepsilon/r)}^r| \lesssim r^d.
\]

Then, by using the estimate (4.5) we acquire that

\[
H(\theta r, \sigma, u_\varepsilon) \leq (\theta r)^{-\sigma} \left( \int_{B(\theta r, r)} |u_\varepsilon - v_r|^p dx \right)^{1/p} + H(\theta r, \sigma, v_r)
\]

\[
\leq C_1 \theta^{-\sigma - \frac{d}{p}} \left\{ (\varepsilon/r)^\eta + r^d \right\} H(2r, \sigma, u_\varepsilon) + C_2 \theta^{1-\sigma} H(r, \sigma, v_r)
\]

\[
\leq \max \left\{ C_1 C_2 \theta^{-\sigma - \frac{d}{p}} (\varepsilon/r)^\eta, C_1 C_2 \theta^{-\sigma - \frac{d}{p}} r^d, C_2 \theta^{1-\sigma} \right\} H(2r, \sigma, u_\varepsilon),
\]

(4.9)
where we use the estimate (4.8) in the second step. Now, we first fix \( \theta \in (0, 1/4) \) such that \( C_2 \theta^\sigma \leq 1/4 \), and then let \( N \varepsilon \leq r < r_0 \). By choosing \( N > 1 \) large and \( r_0 \in (0, 1) \) sufficiently small, we also obtain

\[
\max \left\{ C_1 C_2 \theta^{-\sigma - \frac{d}{p}} N^{-\eta}, C_1 C_2 \theta^{-\sigma - \frac{d}{p}} r_0^{d/2} \right\} \leq 1/4.
\]

Obviously, we have

\[
H(\theta r, \sigma, u_\varepsilon) \leq \frac{1}{4} H(2r, \sigma, u_\varepsilon). \tag{4.10}
\]

Moreover, multiplying \( r^{-1} \) on the both sides of (4.10) and then integrating with respect to \( r \) from \( N \varepsilon \) to \( r_0 \), we have

\[
\int_{\theta N \varepsilon}^{\theta r_0} H(r, \sigma, u_\varepsilon) \frac{dr}{r} \leq \frac{1}{4} \int_{2N \varepsilon}^{2r_0} H(r, \sigma, u_\varepsilon) \frac{dr}{r}
\]

and this implies that

\[
\int_{\theta N \varepsilon}^{2r_0} H(s, \sigma, u_\varepsilon) \frac{ds}{s} \leq \frac{4}{3} \int_{\theta r_0}^{2r_0} H(s, \sigma, u_\varepsilon) \frac{ds}{s} \leq CH(2r_0, \sigma, u_\varepsilon).
\]

Thus we deduce from the above estimate that

\[
H(r, \sigma, u_\varepsilon) \leq CH(r_0, \sigma, u_\varepsilon) \tag{4.11}
\]

holds for any \( \varepsilon \leq r < r_0 \). We mention that due to the relationship \( N \gtrsim \theta^{-\frac{\sigma}{\eta} - \frac{d}{p}} \), it is trivial to prove the stated result (4.11) in the case of \( \varepsilon \leq r \leq \theta N \varepsilon \), as well as, of \( r_0 \leq r \leq 1 \). Then it is clear to see that

\[
\left( \int_{B(0,r)} |\nabla u_\varepsilon|^p \right)^{\frac{1}{p}} \lesssim r^{\sigma-1} H(2r, \alpha, u_\varepsilon)
\]

\[
\lesssim r^{\sigma-1} H(R, \alpha, u_\varepsilon) \lesssim \left( \frac{r}{R} \right)^{\sigma-1} \left( \int_{B(0,R)} |\nabla u_\varepsilon|^p \right)^{\frac{1}{p}}
\]

where we note that \( u_\varepsilon \) up to a constant is still a solution to \( \mathcal{L}_\varepsilon u_\varepsilon = 0 \) in \( B(0,1) \). The first step follows from Caccioppoli’s inequality (4.1), and second one comes from the estimate (4.11), while we use Poincaré’s inequality in the last step. Then, by setting \( \kappa = p(1 - \sigma) \), the above estimate consequently leads to the desired result (1.23), and we have completed the whole proof.

\[ \square \]

## 5 Examples on the assumption (1.7)

We take \( A(y, \xi) = a(y)|\xi|^{p-2}\xi \) as an example, with the assumption that

\[
a(y + z) = a(y) \quad \text{and} \quad \mu_0 \leq a(y) \leq \mu_1 \quad \forall y \in \mathbb{R}^d, z \in \mathbb{Z}^d,
\]

which is known as weighted p-Laplace equation. In the case of \( d = 1 \), the homogenized equation could be figured out, and

\[
\widehat{A}(\xi) = \widehat{a}|\xi|^{p-2}\xi
\]

with

\[
\widehat{a} = \left( \int_0^1 [a(y)]^{\frac{1}{1-p}} dy \right)^{1-p}
\]

(see [24, Example 25.4]).
Then we turn to the case $d \geq 2$. For the higher dimensions, we assume that
\[
\frac{\partial}{\partial y_i} [a(y)] = \frac{\partial}{\partial y_j} [a(y)], \quad i \neq j \text{ and } i, j = 1, 2, \cdots, d.
\] (5.1)
In such the case, one may construct the first order corrector in the form of
\[
N(y, \xi) = \left[ \chi(y) + \sum_{k=1}^{d} y_k \right] \sum_{k=1}^{d} \xi_k - y \cdot \xi, \quad y \in Y,
\] (5.2)
where the corrector $\chi$ satisfies
\[
\begin{cases}
\text{div} \left[ a(y) |\nabla \chi + 1|^p - 2 (\nabla \chi + 1) \right] = 0 \quad \text{in } Y, \\
\chi \in W^{1,p}_{\text{per}}(Y), \quad \int_Y \chi = 0.
\end{cases}
\] (5.3)
It is known that there exists the unique solution of the above cell problem, and this implies that $N(\cdot, \xi) \in W^{1,p}_{\text{per}}(Y)$ satisfies the equation (1.6). The homogenized equation is given by
\[
- \hat{a} \cdot \nabla \left( |\sum_{k=1}^{d} \nabla_k u_0|^p - 2 \sum_{k=1}^{d} \nabla_k u_0 \right) = F \quad \text{in } \Omega,
\] (5.4)
where $\hat{a} \in \mathbb{R}^d$ is a vector, and its component is given by
\[
\hat{a}_i = \int_Y \left[ a(y) |\nabla \chi(y) + 1|^p \nabla_i \chi(y) + 1 \right] dy.
\]
Note that the equation (5.4) is solvable according to $\hat{a}$ direction, which means if we take $F(x) = F(\hat{a} \cdot x)$ and $u_0(x) = u_0(\hat{a} \cdot x)$ for any $x \in \Omega$, then the equation (5.4) is transformed into
\[
- \frac{\partial}{\partial r} \left( |\sum_{k=1}^{d} \hat{a}_k|^p - 2 \sum_{k=1}^{d} \hat{a}_k \right) F(r) \quad \text{and } r = \hat{a} \cdot x,
\]
provided
\[
\sum_{k=1}^{d} \hat{a}_k \neq 0.
\] (5.5)
Thus, we go back to the equation (5.3), and it is possible to find a radial solution (i.e., $\chi(y) = \chi(|y|)$) to meet the condition (5.5), since it is fine to assume $a(y) = a(|y|)$ under the assumption (5.1). In such the case, we may have $\hat{a}_i = \hat{a}_j$ for any $i \neq j$.

**Remark 5.1.** In fact, the term $T_1$ of (5.2) leads to the equation (5.3), while the form of $T_2$ of (5.2) determines the formula of homogenized operator in (5.4), separately. Thus, if the corrector $\chi$ is a radial solution, then we may further simplify $T_2$ into $\sum_{k=1}^{d} \xi_k = d \xi_1$ for example, which gives
\[
|\sum_{k=1}^{d} \xi_k | \geq |\xi_1| / \sqrt{d}.
\] (5.6)
We mention that the above inequality will be true if there exists an integer $i_0 \in [1, d]$ such that $\max_{i \neq i_0} |\xi_i | \leq |\xi_j |$ and $\sum_{i \neq i_0} \xi_i \geq 0$. In fact, to verify the conditions (2.9) and (2.10), the inequality (5.6) will be used in later computations.
Then we plan to show the reasonableness of (5.2), at least, in the case of \( p = 2 \). According to the asymptotic expansion \( u_\varepsilon = u_0(x) + \varepsilon u_1(x, y) + \varepsilon^2 u_2(x, y) + \cdots \), one may derive that
\[
\frac{\partial}{\partial y_i} \left( a(y) \frac{\partial u_1}{\partial y_i} \right) = -\frac{\partial}{\partial y_i} \left[ a(y) \sum_{k=1}^{d} \frac{\partial u_0}{\partial x_k} \right],
\]
(5.7)
(see for example [26]), where \( y = x/\varepsilon \). On the one hand, let \( u_1(x, y) = N(y, \nabla u_0) \), and it is not hard to verify that \( N(y, \nabla u_0) \) in the form of (5.2) in fact satisfies the above equation under the condition (5.1) coupled with the equation (5.3). Firstly, in view of (5.2),
\[
\nabla_y N(y, \nabla u_0) = \left( \nabla_y \chi + e_i \right) \left( \sum_{k=1}^{d} \nabla_{x_k} u_0 \right) - \nabla_x u_0.
\]
Then
\[
\nabla_y \left\{ a(y) \nabla_y N(y, \nabla u_0) \right\} = \nabla_y \left\{ a(y) (\nabla_y \chi + 1) \right\} \left( \sum_{k=1}^{d} \nabla_{x_k} u_0 \right) - \nabla_y \left[ a(y) \right] \nabla_x u_0,
\]
where we use the equation (5.3) in the case \( p = 2 \). In this sense, the form of (5.2) is acceptable. On the other hand, we say that the form of (5.2) leads to the same equation that the corrector \( \chi \) ought to satisfy. From (5.7), we have
\[
\frac{\partial}{\partial y_i} \left( a(y) \frac{\partial u_1}{\partial y_i} \right) = -\frac{\partial}{\partial y_i} \left[ a(y) \sum_{k=1}^{d} \frac{\partial u_0}{\partial x_k} \right],
\]
that is due to the assumption (5.1). Let \( u_1(x, y) = \chi(y)(\sum_{k=1}^{d} \partial u_0 / \partial x_k) \), it is clear to see that \( \chi \) admits a solution of
\[
\nabla_y \left\{ a(y) (\nabla_y \chi + 1) \right\} = 0 \quad \text{in} \quad Y,
\]
which is exactly the form of the equation (5.3) in the case of \( p = 2 \). Thus, in terms of the equation that the corrector obeys, the form of (5.2) plays the same role as the above decomposition of \( u_1 \) in the linear case.

Now, we can show the desired estimates (2.9) and (2.10). Set
\[
P(y, \xi) = \left( \nabla \chi + 1 \right) \sum_{k=1}^{d} \xi_k
\]
(5.8)
with the quantity \( \sum_{k=1}^{d} \xi_k \) admitting the inequality (5.6). Then, it follows that
\[
|P(y, \xi)| \geq |\xi| |\nabla \chi + 1| / \sqrt{d},
\]
\[
|P(y, \xi) - P(y, \xi')| \geq |\xi - \xi'| |\nabla \chi + 1| / \sqrt{d} - 1,
\]
\[
|P(y, \xi) - P(y, \xi')| \leq |\xi - \xi'| (|\nabla \chi| + 2).
\]
(5.9)
In the case \( 2 < p < \infty \), we have
\[
\int_Y |P(y, \xi)|^{p-2} |P(y, \xi) - P(\xi', \xi')|^2 dy \\
\geq |\xi|^{p-2} |\xi - \xi'|^2 \int_Y |\nabla \chi + 1|^{p-2} |\nabla \chi + 1| / \sqrt{d} - 1|^2 dy,
\]
and this together with
\[
\langle \hat{A}(\xi) - \hat{A}(\xi'), \xi - \xi' \rangle \geq \mu_0 \int_Y |P(y, \xi) - P(y, \xi')|^2 (|P(y, \xi)| + |P(y, \xi')|)^{p-2} dy \\
\geq \int_Y |P(y, \xi) - P(y, \xi')|^2 (|P(y, \xi)|^{p-2} + |P(y, \xi')|^{p-2}) dy
\]
gives the desired estimate (2.9). Then we proceed to show (2.10), and on account of the assumption (1.2) in the case of \(1 < p < 2\), we arrive at
\[
|\hat{A}(\xi) - \hat{A}(\xi')| \leq \mu_1 \int_Y |P(y, \xi) - P(y, \xi')|(|P(y, \xi)| + |P(y, \xi')|)^{p-2} dy \\
\lesssim |\xi - \xi'|(|\xi| + |\xi'|)^{p-2} \int_Y |\nabla \chi + 2||\nabla \chi + 1|^{p-2} dy \\
\lesssim |\xi - \xi'|(|\xi| + |\xi'|)^{p-2}
\]
and this gives (2.10). In the above estimate, we use (5.9) in the second inequality. Here the up to constant also depend on the corrector \(\chi\) given by (5.3). In the end, we mention that, by the analogy idea one may construct other examples, such as \(A_i(y, \xi) = a_{ij}(y)|\xi|^{p-2}\xi_j\) under a similar assumption to (5.1), and we omit these details here.

Acknowledgements

The second author appreciated Prof. Zhongwei Shen and Prof. Felix Otto for their illuminating instructions and encourages as he studied the homogenization theory. The first and third authors were supported by the National Natural Science Foundation of China (Grant NO. 11471147).

References

[1] S. Armstrong, C. Smart, Quantitative stochastic homogenization of convex integral functionals, Ann. Sci. Éc. Norm. Supér. (4)49 (2016), no.2, 423-481.
[2] S. Armstrong, Z. Shen, Lipschitz estimates in almost-periodic homogenization, Comm. Pure Appl. Math. 69 (2016), no.10, 1882-1923.
[3] M. Avellaneda, F. Lin, Compactness methods in the theory of homogenization, Comm. Pure Appl. Math. 40(1987), no.6, 803-847.
[4] A. Bensoussan, J.-L. Lions, and G.C. Papanicolaou, Asympotic Analysis for Periodic Structures, Studies in Mathematics and its Applications, North Holland, 1978.
[5] L. Caffarelli, A note on nonlinear homogenization, Comm. Pure Appl. Math. 52 (1999), no.7, 829-838.
[6] L. Caffarelli, I. Peral, On \(W^{1,p}\) estimates for elliptic equations in divergence form, Comm. Pure Appl. Math. 51(1998), no.1, 1-21.
[7] V. Chiadó Piat, G. Dal Maso, A. Defranceschi, G-convergence of monotone operators, Ann. Inst. H. Poincaré Anal. Non Linéaire 7(1990), no.3, 123-160.
[8] G. Dal Maso, A. Defranceschi, Correctors for the homogenization of monotone operators, Differential and Integral Equations, 3(1990), no.6, 1151-1166.

[9] E. DiBenedetto, $C^{1+\alpha}$ local regularity of weak solutions of degenerate elliptic equations, Nonlinear Anal. 7(1983), no.8, 827-850.

[10] W. E, P. Ming, P. Zhang, Analysis of the heterogeneous multiscale method for elliptic homogenization problems, J. Amer. Math. Soc. 18(2005), no.1, 121-156.

[11] X. Fan, Global $C^{1,\alpha}$ regularity for variable exponent elliptic equations in divergence form, J. Differential Equations 235(2007), no.2, 397-417.

[12] N. Fusco, G. Moscariello, On the homogenization of quasilinear divergence structure operators, Ann. Mat. Pura. Appl., 146(1987), 1-13.

[13] E. Giusti, Direct Methods in the Calculus of Variations, World Scientific Publishing Co., Inc., River Edge, NJ, 2003.

[14] A. Gloria, An analytical framework for the numerical homogenization of monotone elliptic operators and quasiconvex energies, Multiscale Model. Simul. 5(2006), no.3, 996-1043.

[15] A. Gloria, S. Neukamm, F. Otto, A regularity theory for random elliptic operators, arXiv:1409.2678v3, (2014).

[16] A. Gloria, S. Neukamm, F. Otto, Quantification of ergodicity in stochastic homogenization: optimal bounds via spectral gap on Glauber dynamics, Invent. Math. 199(2015), no.2, 455-515.

[17] A. Gloria, F. Otto, An optimal error estimate in stochastic homogenization of discrete elliptic equations, Ann. Appl. Probab. 22(2012), no.1, 1-28.

[18] M. Giaquinta, L. Martinazzi, An Introduction to the Regularity Theory for Elliptic Systems, Harmonic Maps and Minimal Graphs, Edizioni della Normale, Pisa, 2012.

[19] V. Jikov, S. Kozlov, O. Oleinik, Homogenization of Differential Operators and Integral Functionals, Springer-Verlag, Berlin, 1994.

[20] C. Kenig, F. Lin, Z. Shen, Homogenization of elliptic systems with Neumann boundary conditions, J. Amer. Math. Soc. 26(2013), no.4, 901-937.

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

[22] J. Malý, W. Ziemer, Fine Regularity of Solutions of Elliptic Partial Differential Equations, Mathematical Surveys and Monographs, 51, American Mathematical Society, Providence, RI, 1997.

[23] A. Pankov, G-Convergence and Homogenization of Nonlinear Partial Differential Operators, Mathematics and its Applications, 422, Kluwer Academic Publishers, Dordrecht, 1997.

[24] S. Pastukhova, Operator estimates in nonlinear problems of reiterated homogenization, (Russian) Tr. Mat. Inst. Steklova 261(2008), Differ. Uravn. i Din. Sist., 220-233; translation in Proc. Steklov Inst. Math. 261(2008), no.1, 214-228.
[25] Z. Shen, Lectures on Periodic Homogenization of Elliptic Systems, arXiv:1710.11257v1 (2017).

[26] Z. Shen, Boundary estimates in elliptic homogenization, Anal. PDE 10(2017), 653-694.

[27] T. Suslina, Homogenization of the Neumann problem for elliptic systems with periodic coefficients, SIAM J. Math. Anal. 45(2013), no.6, 3453-3493.

[28] Q. Xu, Convergence rates for general elliptic homogenization problems in Lipschitz domains, SIAM J. Math. Anal. 48(2016), no.6, 3742-3788.

[29] L. Wang, Q. Xu, P. Zhao, Quantitative estimates on periodic homogenization of nonlinear elliptic operators, arXiv:1807.10865v1 (2018).

[30] E. Zeidler, Nonlinear Functional Analysis and its Applications (II/B), nonlinear monotone operators. Translated from the German by the author and Leo F. Boron, Springer-Verlag, New York, 1990.

[31] V. Zhikov, S. Pastukhova, On operator estimates for some problems in homogenization theory, Russ. J. Math. Phys. 12(2005), no.4, 515-524.

[32] V. Zhikov, S. Pastukhova, On operator estimates in homogenization theory, (Russian) Uspekhi Mat. Nauk 71(2016), no.3(429), 27-122; translation in Russian Math. Surveys 71(2016), no.3, 417-511.