Finite Element Approximation of Elliptic Homogenization Problems in Nondivergence-Form

Timo Sprekeler

Keble College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

2021

This thesis is dedicated to
my parents,
Christa and Frank,
and to my brother,
Jan.

Annotation

This thesis contains original work taken from our papers/preprints listed below.

- [29] (with Y. Capdeboscq and E. Süli). Finite element approximation of elliptic homogenization problems in nondivergence-form. *ESAIM Math. Model. Numer. Anal.*, 54(4):1221–1257, 2020.
- [50] (with D. Gallistl and E. Süli). Mixed Finite Element Approximation of Periodic Hamilton–Jacobi–Bellman Problems With Application to Numerical Homogenization. *Multiscale Model. Simul.*, 19(2):1041–1065, 2021.
- [69] (with E. L. Kawecki). Discontinuous Galerkin and C^0-IP finite element approximation of periodic Hamilton–Jacobi–Bellman–Isaacs problems with application to numerical homogenization. arXiv:2104.14450 [math.NA], submitted.
- [90] (with H. V. Tran). Optimal Convergence Rates for Elliptic Homogenization Problems in Nondivergence-Form: Analysis and Numerical Illustrations. *Multiscale Model. Simul.*, 2021 (Forthcoming).

Chapter 2 is based on [29, 90]. Chapter 3 is based on [29]. Chapter 4 is based on [50]. Chapter 5 is based on [69].

Acknowledgements

First and foremost, I would like to thank my terrific team of advisors Prof. Endre Süli and Prof. Yves Capdeboscq for the many interesting mathematical conversations and their kind support. Secondly, I would like to express my thanks to my collaborators Prof. Dietmar Gallistl, Prof. Hung V. Tran and Dr Ellya L. Kawecki for many useful discussions and their support beyond our joint projects. Let me also thank the people who shaped me as a mathematician along my university education: A big thanks to Prof. Ben Schweizer and Prof. Stefan Turek for introducing me to the analysis and numerical analysis of PDEs and sharing their enthusiasm for the subject, and to Prof. Carola-Bibiane Schönlieb for some invaluable postgraduate research experience.

Thank you to my friends at Oxford who have made these few years truly unforgettable; let me mention in particular Katrin Harter, Christoph Höppke, Panas Kalayanamit, Fabian Laakmann, Yikun Qiao, Tommaso Seneci, and Jingmin Xia. Finally, thanks to my whole family for their limitless support.

My DPhil studies at Oxford have been funded by the Clarendon Fund and Keble College Sloane Robinson Scholarship for which I am very grateful.

Abstract

This thesis focuses on the construction of finite element numerical homogenization schemes for both linear and selected fully-nonlinear elliptic partial differential equations in nondivergence-form.

In the first part of the thesis, we study periodic homogenization problems of the form $A(x/\varepsilon) : D^2 u_\varepsilon = f$ subject to a homogeneous Dirichlet boundary condition. We provide a qualitative $W^{2,p}$ theory and obtain optimal gradient and Hessian bounds with correction terms taken into account in the L^p-norm. Consequently, we find that $(u_\varepsilon)_{\varepsilon>0}$ converges strongly in the $W^{1,p}$-norm to the solution of the corresponding effective problem, and that the optimal rate for this convergence is $\mathcal{O}(\varepsilon)$. Based on these quantitative homogenization results, we propose and rigorously analyze a finite element-type numerical homogenization scheme for the approximation of the solution to the effective problem and the solution u_ε to the original problem in the H^1 and H^2 Sobolev-norms. We extend the scheme to the framework of nonuniformly oscillating coefficients and provide a variety of numerical experiments illustrating the theoretical results.

In the second part of the thesis, we propose and rigorously analyze numerical homogenization schemes for the fully-nonlinear Hamilton–Jacobi–Bellman (HJB) and HJB–Isaacs (HJBI) equations. More precisely, we are interested in the approximation of the effective Hamiltonian which determines the effective equation. Our numerical schemes are based on finite element approximations for suitable corrector problems arising in the periodic homogenization of these equations. We present a mixed finite element scheme as well as discontinuous Galerkin and C^0 interior penalty finite element approaches. Several numerical experiments accompany the theoretical results and illustrate the performance of the numerical schemes.

Contents

Chapter 1

Introduction

This thesis is set within the areas of (periodic) homogenization and numerical homogenization. In many applications from physics and engineering, such as the study of composite materials, the underlying mathematical model is described by a partial differential equation (PDE) involving a fine microstructure. For a small parameter $\varepsilon > 0$, we consider (possibly nonlinear) second-order elliptic PDEs of the form

$$F\left(x, \frac{x}{\varepsilon}, u_\varepsilon, \nabla u_\varepsilon, D^2 u_\varepsilon\right) = 0 \qquad \text{in } \Omega$$

together with a periodicity assumption on the map $F = F(x, y, u, p, R)$ in the y-variable. We call x the slow variable and $y = x/\varepsilon$ the fast variable.

Their accurate computational resolution is very costly as classical numerical schemes such as finite element methods require a sufficiently fine discretization of the computational domain Ω to capture the oscillations at the microscale-level. In (numerical) homogenization, we try to overcome this difficulty by establishing an effective macroscopic model

$$F^{\text{eff}}\left(x, u_0, \nabla u_0, D^2 u_0\right) = 0 \qquad \text{in } \Omega$$

that is computationally cheap to solve with comparatively coarse discretizations. Typically, the construction of the effective operator F^{eff} relies on the solution of suitable local problems, which leads to a multiscale approach.

The development of numerical multiscale methods for divergence-form equations started around 1990 with the development of two of the most popular and effective approaches called the multiscale finite element method (MsFEM) and the heterogeneous multiscale method (HMM). More recently, in the past two decades the method of Localized Orthogonal Decomposition (LOD) has been developed, which is particularly interesting as it is not restricted to the structural assumptions of scale separation and periodicity. Today, there is a range of multiscale methods for divergence-form equations to choose from whereas the research for the nondivergence-form-case is lagging behind. We briefly explain the contributions of this thesis in the following paragraphs and provide an overview of the current literature.

1.1 Linear equations in nondivergence-form

In the first part of this thesis we consider the linear prototype problem of a second-order elliptic equation in nondivergence-form, that is,

$$A\left(\frac{\cdot}{\varepsilon}\right) : D^2 u_\varepsilon := \sum_{i,j=1}^{n} a_{ij}\left(\frac{\cdot}{\varepsilon}\right) \partial_{ij}^2 u_\varepsilon = f \quad \text{in } \Omega, \tag{1.1}$$

subject to the homogeneous Dirichlet boundary condition

$$u_\varepsilon = 0 \quad \text{on } \partial\Omega. \tag{1.2}$$

Here we assume that $\Omega \subset \mathbb{R}^n$ is a sufficiently regular bounded domain, $\varepsilon > 0$ is a small parameter and that $A : \mathbb{R}^n \to \mathbb{R}^{n\times n}$ is a symmetric, uniformly elliptic and Y-periodic matrix-valued function such that

$$A \in C^{0,\alpha}(\mathbb{R}^n; \mathbb{R}^{n\times n})$$

for some $\alpha \in (0, 1]$. Throughout this thesis, we use the notation

$$Y := (0, 1)^n \subset \mathbb{R}^n$$

to denote the unit cell in $\mathbb{R}^n$.

The main goal of Part I is to propose and rigorously analyze a finite element numerical homogenization scheme for (1.1), (1.2) that is based on novel quantitative homogenization results.

Periodic homogenization

We only discuss the nondivergence-form case and refer the reader interested in the divergence-form case to the books Allaire [8], Bensoussan, Lions, Papanicolaou [20], Cioranescu, Donato [30], Tartar [92], and the references therein.

Periodic homogenization is the study of the limiting behavior of the sequence of solutions $(u_\varepsilon)_{\varepsilon>0}$ to (1.1), (1.2) as the oscillation parameter ε tends to zero. It is well-known that $(u_\varepsilon)_{\varepsilon>0}$ converges uniformly on $\bar{\Omega}$ (that is, in the $L^\infty(\Omega)$-norm) to the solution u_0 of a constant-coefficient problem

$$\begin{cases} A^0 : D^2 u_0 = f & \text{in } \Omega, \\ u_0 = 0 & \text{on } \partial\Omega, \end{cases} \tag{1.3}$$

which we call the homogenized (or effective) problem; see e.g., Bensoussan, Lions, Papanicolaou [20], Jikov, Kozlov, Oleinik [65]. The effective coefficient $A^0 = (a^0_{ij})_{1\le i,j\le n} \in \mathbb{R}^{n\times n}$ is a constant symmetric positive definite matrix which can be obtained via integration against an invariant measure m, that is,

$$A^0 = \int_Y Am, \tag{1.4}$$

with $m : \mathbb{R}^n \to \mathbb{R}$ the solution to the periodic problem

$$\begin{cases} D^2 : (Am) = 0 \quad \text{in } Y, \\ m \text{ is } Y\text{-periodic}, \ m > 0, \ \int_Y m = 1; \end{cases}$$

see Avellaneda, Lin [15], Engquist, Souganidis [38]. Equivalently, the effective coefficient is characterized via corrector functions: For $i, j \in \{1, \dots, n\}$, the (i, j)-th entry $a^0_{ij} \in \mathbb{R}$ of A^0 is the unique value such that the periodic cell problem

$$\begin{cases} A : D^2\chi_{ij} = a^0_{ij} - a_{ij} \quad \text{in } Y, \\ \chi_{ij} \text{ is } Y\text{-periodic}, \ \int_Y \chi_{ij} = 0 \end{cases}$$

admits a unique periodic solution $\chi_{ij} : \mathbb{R}^n \to \mathbb{R}$, which we call a corrector function.

Heuristically, one may be tempted to believe that a formal two-scale expansion argument gives rise to the expansion

$$u_\varepsilon(x) \approx u_0(x) + \varepsilon^2 \sum_{i,j=1}^{n} \chi_{ij}\left(\frac{x}{\varepsilon}\right) \partial^2_{ij} u_0(x) + \text{higher order terms};$$

see [48], which suggests that $\|u_\varepsilon - u_0\|_{L^\infty(\Omega)} = \mathcal{O}(\varepsilon^2)$. It has just recently been shown in Guo, Tran, Yu [57] that this claim is incorrect: The optimal rate of convergence of $(u_\varepsilon)_{\varepsilon>0}$ to the homogenized solution u_0 in the L^∞-norm is generically only $\mathcal{O}(\varepsilon)$.

We briefly illustrate a formal two-scale asymptotic expansion argument, relying on an expansion of the form

$$u_\varepsilon(x) = \sum_{k=0}^{\infty} \varepsilon^k u_k\left(x, \frac{x}{\varepsilon}\right) \tag{1.5}$$

with functions $u_k = u_k(x, y)$ being Y-periodic in the y-variable for $k \geq 0$. Substituting (1.5) into the equation $[A(y) : D^2 u_\varepsilon(x)]|_{y=\frac{x}{\varepsilon}} = f(x)$ and comparing coefficients of powers of ε yields the equations

$$\begin{aligned}
&\text{order } \varepsilon^{-2} &&: \quad L_{yy}u_0 = 0 &&,\\
&\text{order } \varepsilon^{-1} &&: \quad L_{yy}u_1 = -L_{xy}u_0 &&,\\
&\text{order } \varepsilon^{0} &&: \quad L_{yy}u_2 = f - L_{xy}u_1 - L_{xx}u_0 &&,\\
&\text{order } \varepsilon^{k-2},\ k \geq 3 &&: \quad L_{yy}u_k = -L_{xy}u_{k-1} - L_{xx}u_{k-2}, &&
\end{aligned}$$

where $L_{yy} := A(y) : D_y^2$, $L_{xy} := 2A(y) : \nabla_x \nabla_y^{\mathrm{T}}$, and $L_{xx} := A(y) : D_x^2$. It is quickly seen that $u_0(x, y) \equiv u_0(x)$ and $u_1(x, y) \equiv u_1(x)$ are independent of y. The solvability condition for the equation at order ε^0 yields the homogenized equation $A^0 : D^2 u_0 = f$, and we obtain that $u_2(x, y) = \sum_{i,j=1}^n \chi_{ij}(y)\partial_{ij}^2 u_0(x) + \text{constant}$. At this point one might think that u_1 can be set to zero. However, the solvability condition for the equation at order ε^1 yields $A^0 : D^2 u_1 = -2\sum_{j,k,l=1}^n c_j^{kl}\partial_{jkl}^3 u_0$ with certain constants $c_j^{kl} = c_j^{kl}(A)$ (see (2.20)), which will generally prevent the first-order term u_1 from vanishing. Let us emphasize that this formal asymptotic expansion argument is purely heuristic and neglects the boundary condition, but it provides some useful intuition.

Chapter 2 contains novel qualitative and quantitative homogenization results from our papers Capdeboscq, Sprekeler, Süli [29] and Sprekeler, Tran [90], which add to the results of Guo, Tran, Yu [57] and Kim, Lee [70] on convergence rates. In addition to a qualitative $W^{2,p}$ theory (Section 2.2), the main quantitative results (Sections 2.3 and 2.4) include the L^∞-bound

$$\|u_\varepsilon - u_0 + 2\varepsilon z\|_{L^\infty(\Omega)} = \mathcal{O}(\varepsilon^2),$$

the $W^{1,p}$-bound

$$\left\| u_\varepsilon - u_0 + 2\varepsilon z - \varepsilon^2 \sum_{i,j=1}^n \chi_{ij}\left(\frac{\cdot}{\varepsilon}\right)\partial_{ij}^2 u_0 \right\|_{W^{1,p}(\Omega)} = \mathcal{O}(\varepsilon^{1+\frac{1}{p}}),$$

and the $W^{2,p}$-bound

$$\left\| u_\varepsilon - u_0 - \varepsilon^2 \sum_{i,j=1}^n \chi_{ij}\left(\frac{\cdot}{\varepsilon}\right)\partial_{ij}^2 u_0 \right\|_{W^{2,p}(\Omega)} = \mathcal{O}(\varepsilon^{\frac{1}{p}}),$$

which hold for any $p \in (1, \infty)$ under suitable assumptions (see Remark 2.5.1). Here, z is the solution to a constant-coefficient elliptic problem (see (2.25)). The obtained rates have been numerically demonstrated to be optimal. Let us note that the L^∞ bound was already shown in Guo, Tran, Yu [57] under stronger assumptions on the data, and that our main contribution here are the results in the higher-order $W^{1,p}$ and $W^{2,p}$ Sobolev norms.

In particular, we find that $u_\varepsilon \to u_0$ strongly in $W^{1,p}(\Omega)$ with generically optimal rate of convergence $\mathcal{O}(\varepsilon)$, i.e., we have

$$\|u_\varepsilon - u_0\|_{W^{1,p}(\Omega)} = \mathcal{O}(\varepsilon). \tag{1.6}$$

These quantitative homogenization results are extremely useful in the development of numerical homogenization schemes.

Numerical homogenization

The task of numerical homogenization (Chapter 3) concerns the numerical approximation of the effective coefficient A^0, the solution u_0 to the homogenized problem (1.3), and the solution u_ε to the original problem (1.1), (1.2).

Over the past decades, significant work has been done on periodic homogenization of elliptic problems in divergence-form; numerical homogenization for nondivergence-form problems is however less developed. In particular, we did not find finite element schemes for the numerical homogenization of nondivergence-form problems such as (1.1), (1.2) in the literature. A finite difference scheme has been considered in Froese, Oberman [48].

For divergence-form problems, various multiscale finite element methods (MsFEM) have been developed, which have the advantage over classical finite element methods of providing accurate approximations for very small values of ε even for moderate values of the grid size. For a detailed overview of these methods, we refer the reader to Efendiev, Hou [36], Efendiev, Wu [37], Hou, Wu [61], and the references therein.

The numerical method presented in this work has resemblances with the finite element heterogeneous multiscale method (HMM). The HMM has been introduced in E, Engquist [34] and has been successfully applied to many multiscale problems. For an overview of the field of finite element HMM, we refer to the articles [2, 3, 4, 5] by Abdulle and co-authors, to the review E, Engquist, Li, Ren, Vanden-Eijnden [35], and the references therein. An *a priori* error analysis for the fully discrete finite element HMM for elliptic homogenization problems in divergence-form can be found in the work Abdulle [1]. Concerning nondivergence-form problems, a finite difference HMM has recently been used for the numerical homogenization of second-order hyperbolic nondivergence-form problems in Arjmand, Kreiss [13].

As a third representative of a numerical scheme for divergence-form problems, in addition to MsFEM and HMM, let us mention the method of Localized Orthogonal Decomposition (LOD) introduced in Målqvist, Peterseim [76]. As the name suggests, LOD relies on an orthogonal decomposition of coarse and fine scales, and constitutes a powerful method even for problems beyond the periodic framework. An overview of recent developments can be found in the book Målqvist, Peterseim [77].

Let us briefly outline the finite element numerical homogenization scheme introduced in Capdeboscq, Sprekeler, Süli [29], which we present in Chapter 3.

The *first step* in the development of the proposed numerical homogenization scheme is the construction of a finite element method to obtain approximations $(m_h)_{h>0} \subset H^1_{\mathrm{per}}(Y)$ to the invariant measure with optimal order convergence rate

$$\|m - m_h\|_{L^2(Y)} + h\|m - m_h\|_{H^1(Y)} \lesssim h \inf_{\tilde{v}_h \in \tilde{M}_h} \|m - (\tilde{v}_h + 1)\|_{H^1(Y)},$$

where $\tilde{M}_h$ denotes the finite-dimensional subspace of $H^1_{\mathrm{per}}(Y)$ consisting of continuous Y-periodic piecewise linear functions on the triangulation with zero mean over Y; see Theorem 3.2.1.

The *second step* is to obtain approximations $(A^0_h)_{h>0} \subset \mathbb{R}^{n\times n}$ to the constant matrix A^0; see Lemma 3.2.1. To this end, the integrand in (1.4) is replaced by its continuous piecewise linear interpolant and the invariant measure m is replaced by the approximation m_h, i.e.,

$$A^0_h := \int_Y \mathcal{I}_h(Am_h),$$

which can be computed exactly using an appropriate quadrature rule.

The *third step* is to perform an $H^s(\Omega)$-conforming ($s \in \{1,2\}$) finite element approximation for the problem

$$\begin{cases} A^0_h : D^2 u^h_0 = f & \text{in } \Omega, \\ \qquad\quad u^h_0 = 0 & \text{on } \partial\Omega, \end{cases}$$

on a family of triangulations of the computational domain $\bar{\Omega}$, parametrized by a discretization parameter $H > 0$, measuring the granularity of the triangulation, to obtain $(u^{h,H}_0)_{h,H>0} \subset H^s(\Omega) \cap H^1_0(\Omega)$ with

$$\left\| u^h_0 - u^{h,H}_0 \right\|_{H^s(\Omega)} \lesssim H\|f\|_{H^{s-1}(\Omega)},$$

where the constant is independent of h; see Lemma 3.2.3. Note that for the sake of approximating u_0, an $H^1(\Omega)$-conforming finite element method is sufficient.

The approximation $(u_0^{h,H})_{h,H>0} \subset H^s(\Omega) \cap H_0^1(\Omega)$ obtained by this procedure approximates u_0, i.e., the solution to (1.3), with convergence rate

$$\left\| u_0 - u_0^{h,H} \right\|_{H^s(\Omega)} \lesssim (h + H)\|f\|_{H^{s-1}(\Omega)},$$

which can be improved to $\mathcal{O}(h^2 + H)$ for more regular A; see Theorem 3.2.2, Theorem 3.2.3 and Remark 3.2.3.

Concerning the approximation of u_ε, i.e., the solution to (1.1), (1.2), we show in Chapter 2 that under certain assumptions on the domain and the right-hand side, one has that

$$\left\| u_\varepsilon - u_0 - \varepsilon^2 \sum_{i,j=1}^{n} \chi_{ij}\left(\frac{\cdot}{\varepsilon}\right) \partial_{ij}^2 u_0 \right\|_{H^2(\Omega)} \lesssim \sqrt{\varepsilon}\, \|u_0\|_{W^{2,\infty}(\Omega)} + \varepsilon \|u_0\|_{H^4(\Omega)},$$

and we show in Section 3.3 how the above estimate, together with a finite element approximation scheme for the corrector functions, can be used to obtain approximations to $D^2 u_\varepsilon$. Note that in order to approximate u_ε in the $H^1(\Omega)$-norm, it is sufficient to approximate u_0 in the $H^1(\Omega)$-norm as we have (1.6). However, for an approximation of $D^2 u_\varepsilon$ based on the above corrector estimate, we need to approximate u_0 in the $H^2(\Omega)$-norm.

In Section 3.4, we extend our results to the case of nonuniformly oscillating coefficients, i.e., to problems of the form

$$\begin{cases} A\left(\cdot, \frac{\cdot}{\varepsilon}\right) : D^2 u_\varepsilon = f & \text{in } \Omega, \\ \qquad\qquad u_\varepsilon = 0 & \text{on } \partial\Omega, \end{cases}$$

where $A = A(x, y) : \Omega \times \mathbb{R}^n \to \mathbb{R}^{n \times n}$ is a sufficiently regular symmetric, uniformly elliptic matrix-valued function that is Y-periodic in y for fixed $x \in \Omega$.

1.2 Nonlinear equations in nondivergence-form

In the second part of this thesis we consider fully-nonlinear homogenization problems with the nonlinearity being of Hamilton–Jacobi–Bellman (HJB) or more generally of Hamilton–Jacobi–Bellman–Isaacs (HJBI) type. More precisely, we study the numerical homogenization of problems of the form

$$\begin{cases} u_\varepsilon + F\left[x, \frac{x}{\varepsilon}, \nabla u_\varepsilon, D^2 u_\varepsilon\right] = 0 & \text{in } \Omega, \\ \qquad\qquad u_\varepsilon = 0 & \text{on } \partial\Omega, \end{cases} \tag{1.7}$$

with $\Omega \subset \mathbb{R}^n$ being a convex domain in dimension $n \in \{2, 3\}$, a parameter $\varepsilon > 0$ considered to be small, and HJB-type nonlinearity

$$F^{\mathrm{HJB}}\left[x, y, \nabla w, D^2 w\right] := \sup_{\alpha \in \Lambda} \left\{-A(x, y, \alpha) : D^2 w - b(x, y, \alpha) \cdot \nabla w - f(x, y, \alpha)\right\},$$

or HJBI-type nonlinearity

$$\begin{aligned} &F^{\mathrm{HJBI}}\left[x, y, \nabla w, D^2 w\right] \\ &\quad := \inf_{\alpha \in \mathcal{A}} \sup_{\beta \in \mathcal{B}} \left\{-A(x, y, \alpha, \beta) : D^2 w - b(x, y, \alpha, \beta) \cdot \nabla w - f(x, y, \alpha, \beta)\right\}. \end{aligned}$$

Here, $\Lambda, \mathcal{A}, \mathcal{B}$ are compact metric spaces and the coefficients $A = (a_{ij})_{1 \leq i,j \leq n}, b = (b_i)_{1 \leq i \leq n}, f$ are assumed to be Y-periodic in $y \in \mathbb{R}^n$ with respect to their second arguments, and satisfy suitable regularity assumptions. Further, we assume that A is uniformly elliptic and that the coefficients satisfy a generalized Cordes condition, i.e., that there exist constants $\bar{\lambda} > 0$ and $\bar{\delta} \in (0, 1)$ such that

$$|A|^2 + \frac{|b|^2}{2\bar{\lambda}} + \frac{1}{\bar{\lambda}^2} \leq \frac{1}{n + \bar{\delta}} \left(\mathrm{tr}(A) + \frac{1}{\bar{\lambda}}\right)^2.$$

Under these assumptions, it is known that the problem (1.7) admits a unique strong solution $u_\varepsilon \in H^2(\Omega) \cap H_0^1(\Omega)$; see Smears, Süli [88].

It is well-known (see, e.g., Caffarelli, Souganidis, Wang [25], Evans [39, 40]) that the viscosity solution $u_\varepsilon \in C(\bar{\Omega})$ to (1.7) converges uniformly, as $\varepsilon \searrow 0$, to the viscosity solution $u_0 \in C(\bar{\Omega})$ of the homogenized problem

$$\begin{cases} u_0 + H(x, \nabla u_0, D^2 u_0) = 0 & \text{in } \Omega, \\ u_0 = 0 & \text{on } \partial\Omega, \end{cases}$$

for some function $H : \bar{\Omega} \times \mathbb{R}^n \times \mathbb{R}^{n \times n}_{\mathrm{sym}} \to \mathbb{R}$, the so-called effective Hamiltonian. The value of the effective Hamiltonian at a fixed point $(x, p, R) \in \bar{\Omega} \times \mathbb{R}^n \times \mathbb{R}^{n \times n}_{\mathrm{sym}}$ can be obtained as the uniform limit of the sequence $\{-\sigma v^\sigma\}_{\sigma > 0}$ as $\sigma \searrow 0$, where the so-called approximate corrector $v^\sigma = v^\sigma(\cdot\,; x, p, R)$ is the solution to the problem

$$\begin{cases} \sigma v^\sigma + F\left[x, y, p, R + D_y^2 v^\sigma\right] = 0 & \text{in } y \in Y, \\ y \mapsto v^\sigma(y; x, p, R) \text{ is } Y\text{-periodic}; \end{cases} \tag{1.8}$$

see e.g., Alvarez, Bardi [10, 11], Camilli, Marchi [28]. For further homogenization results we refer to Section 4.3. The main goal of this second part of the thesis is the efficient numerical approximation of the effective Hamiltonian.

The motivation for studying the fully nonlinear second-order HJB and HJBI equation comes from stochastic control theory for Markov diffusion processes and we refer the reader to Fleming, Soner [47]. Its study is a mathematically challenging task as there is no natural variational formulation and solvability has to be considered either in the sense of viscosity solutions (see Definition 4.3.1 and the user's guide Crandall, Ishii, Lions [33] for a comprehensive overview), or in the sense of strong solutions, i.e., functions admitting weak derivatives up to order two satisfying the equation pointwise almost everywhere.

The numerical homogenization of HJB/HJBI equations has not been studied a lot so far. For the case of second-order HJB equations, a finite difference scheme for the whole space problem has been proposed in Camilli, Marchi [28]. In Finlay, Oberman [45, 46], the effective Hamiltonian is computed exactly for HJB operators of certain types and some numerical simulations have been conducted. It seems that finite element schemes for the numerical homogenization of the problem (1.7) have not been constructed yet.

Let us note that there is a lot more work in the literature on the numerical approximation of the effective Hamiltonian arising in the homogenization of first-order Hamilton–Jacobi equations; see various authors [6, 41, 54, 55, 75, 80, 82, 83].

Our numerical schemes are based on mixed, discontinuous Galerkin, or C^0 interior penalty methods for the approximate corrector problems (1.8). The finite element approximation of periodic HJB/HJBI problems seems to have not been studied so far, the Dirichlet problem however, has been an active area of research in the past decades; see Feng, Glowinski, Neilan [43] and Neilan, Salgado, Zhang [78] for a survey on recent developments. The mixed finite element method presented in this work is a modified version of the mixed scheme for the Dirichlet problem with coefficients satisfying a generalized Cordes condition introduced in Gallistl, Süli [51], which allows the use of H^1-conforming finite elements. For further H^1-conforming finite element schemes, we refer to Camilli, Falcone [26], Camilli, Jakobsen [27], Jensen [63], and Jensen, Smears [64]. The first numerical scheme for HJB equations in the Cordes framework has been the discontinuous Galerkin finite element method in Smears, Süli [88, 89].

For the case of a HJBI-type nonlinearity, there is only little work in the numerical analysis literature. The well-posedness and the discontinuous Galerkin (DG) and C^0 interior penalty (C^0-IP) finite element approximation has only very recently been discussed in Kawecki, Smears [67, 68], paving the way to study numerical homogenization on nonsmooth domains in a Cordes framework. Our numerical homogenization scheme for Isaacs problems is based on a periodic adaptation of the method in Kawecki, Smears [67].

Numerical homogenization via mixed FEM for approximate correctors

In Chapter 4, which is a presentation of our paper Gallistl, Sprekeler, Süli [50], we propose and rigorously analyze a numerical homogenization scheme for HJB equations based on a mixed finite element approximation of approximate corrector problems.

The *first step* of the numerical scheme is the mixed finite element approximation of the solution to (1.8). More generally, we present the method for problems of the form

$$\begin{cases} \sup_{\alpha\in\Lambda}\{-A^\alpha : D^2u - b^\alpha\cdot\nabla u + c^\alpha u - f^\alpha\} = 0 \quad \text{in } Y, \\ u \text{ is } Y\text{-periodic}, \end{cases} \tag{1.9}$$

with uniformly continuous functions $a_{ij} = a_{ji}, b_i, c, f \in C(\mathbb{R}^n\times\Lambda)$ and positive zeroth-order coefficient $c > 0$. It is assumed that $A^\alpha, b^\alpha, c^\alpha, f^\alpha$ are Y-periodic on $\mathbb{R}^n$ and that the coefficients satisfy the Cordes condition, i.e., that there exist constants $\lambda > 0$ and $\delta \in (0,1)$ such that

$$|A^\alpha|^2 + \frac{|b^\alpha|^2}{2\lambda} + \frac{(c^\alpha)^2}{\lambda^2} \le \frac{1}{n+\delta}\left(\mathrm{tr}(A^\alpha) + \frac{c^\alpha}{\lambda}\right)^2 \tag{1.10}$$

holds in $\mathbb{R}^n$ for all $\alpha \in \Lambda$. Under these assumptions, the periodic HJB problem (1.9) admits a unique strong solution $u \in H^2_{\mathrm{per}}(Y)$; see Section 4.1.2.

The mixed formulation relies on the observation that we seek a pair $(w,u) = (\nabla u, u)$ such that

$$\begin{cases} F_\gamma[(w,u)] := \sup_{\alpha\in\Lambda}\{\gamma^\alpha\left(-A^\alpha : Dw - b^\alpha\cdot\nabla u + c^\alpha u - f^\alpha\right)\} = 0, \\ \nabla u - w = 0, \\ (w,u) \in X := W_{\mathrm{per}}(Y;\mathbb{R}^n)\times H^1_{\mathrm{per}}(Y), \end{cases}$$

where $\gamma = \gamma^\alpha(y) \in C(\mathbb{R}^n\times\Lambda)$ is a suitable positive renormalization function and we write $W_{\mathrm{per}}(Y)$ to denote the subspace of functions in $H^1_{\mathrm{per}}(Y)$ having mean zero over the unit cell Y. For a chosen $M \subset W_{\mathrm{per}}(Y)$, the mixed formulation seeks a pair $((w,u),m) \in X\times M$ such that for all $((w',u'),m') \in X\times M$ we have

$$\begin{aligned} \langle F_\gamma[(w,u)], L_\lambda(w',u')\rangle_{L^2(Y)} + S((w,u),(w',u')) + \langle \nabla m, \nabla u' - w'\rangle_{L^2(Y)} &= 0, \\ \langle \nabla m', \nabla u - w\rangle_{L^2(Y)} &= 0. \end{aligned}$$

Here, $S : X\times X \to \mathbb{R}$ is a bilinear form containing stabilization terms and our test functions are of the form $L_\lambda(w',u') := \lambda u' - \nabla\cdot w'$. The discrete mixed formulation is defined similarly using suitable closed linear subspaces $X_h \subset X$ and $M_h \subset M$, seeking a pair $((w_h, u_h), m_h) \in X_h \times M_h$ such that for all $((w_h', u_h'), m_h') \in X_h \times M_h$ we have

$$\langle F_\gamma[(w_h, u_h)], L_\lambda(w_h', u_h')\rangle_{L^2(Y)} + S((w_h, u_h), (w_h', u_h')) + \langle \nabla m_h, \nabla u_h' - w_h'\rangle_{L^2(Y)} = 0,$$
$$\langle \nabla m_h', \nabla u_h - w_h\rangle_{L^2(Y)} = 0.$$

We refer to Section 4.2 for a rigorous *a priori* and *a posteriori* analysis of the scheme.

The *second step* of the approximation scheme is the approximation of the effective Hamiltonian. To this end, we note that the approximate corrector problem (1.8) is of the form (1.9) and satisfies the Cordes condition (1.10) with Cordes parameters $\lambda = \sigma\bar{\lambda}$ and $\delta = \bar{\delta}$.

Using the mixed finite element method from the first step, we can obtain a numerical approximation $(w_h^\sigma(\cdot\,; x, p, R), v_h^\sigma(\cdot\,; x, p, R)) \in M_h \times X_h$ to the pair $(\nabla v^\sigma(\cdot\,; x, p, R), v^\sigma(\cdot\,; x, p, R))$ for any fixed point $(x, p, R) \in \bar{\Omega} \times \mathbb{R}^n \times \mathbb{R}^{n\times n}_{\mathrm{sym}}$, and we then define the approximated effective Hamiltonian as

$$H_{\sigma,h} : \bar{\Omega} \times \mathbb{R}^n \times \mathbb{R}^{n\times n}_{\mathrm{sym}} \to \mathbb{R}, \qquad H_{\sigma,h}(x, p, R) := -\sigma \int_Y v_h^\sigma(\cdot\,; x, p, R).$$

We prove that under suitable assumptions we have the error bound

$$|H_{\sigma,h}(x, p, R) - H(x, p, R)| \lesssim (h^r + \sigma)\,(1 + |p| + |R|)$$

for some $r > 0$ and we refer the reader to Sections 4.3.3 and 4.3.4 for a detailed analysis. Finally, the numerical experiments in Section 4.4 demonstrate the approximation scheme for the effective Hamiltonian as well as a least-squares scheme for the resulting effective equation.

Numerical homogenization via DG/C^0-IP FEM for approximate correctors

Chapter 5 presents the results of our paper Kawecki, Sprekeler [69], in which we propose and rigorously analyze a numerical homogenization scheme for HJBI equations based on a discontinuous Galerkin or C^0-IP finite element approximation of approximate corrector problems. The idea of the method is to approximate the effective Hamiltonian similarly as before via $H_\mathcal{T}^\sigma : \bar{\Omega} \times \mathbb{R}^n \times \mathbb{R}^{n\times n}_{\mathrm{sym}} \to \mathbb{R}$ given by

$$H_\mathcal{T}^\sigma(x, p, R) := -\sigma \int_Y v_\mathcal{T}^\sigma(\cdot\,; x, p, R), \tag{1.11}$$

with an approximation $v_\mathcal{T}^\sigma(\cdot\,; x, p, R)$ to the approximate corrector $v^\sigma(\cdot\,; x, p, R)$. Although the previous mixed approach can be easily extended to the HJBI framework, we would like to present a different approach by considering DG/C^0-IP approximations.

As a *first step* we look at periodic HJBI cell problems

$$\begin{cases} \inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} \left\{-A^{\alpha\beta} : D^2u - b^{\alpha\beta}\cdot\nabla u + c^{\alpha\beta}u - f^{\alpha\beta}\right\} = 0 \quad \text{in } Y, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad u \text{ is } Y\text{-periodic}, \end{cases} \tag{1.12}$$

with similar assumptions as in the HJB case, i.e., uniformly continuous functions $a_{ij} = a_{ji}, b_i, c, f \in C(\mathbb{R}^n \times \mathcal{A} \times \mathcal{B})$, a positive zeroth-order coefficient $c > 0$ and $A^{\alpha\beta}, b^{\alpha\beta}, c^{\alpha\beta}, f^{\alpha\beta}$ are Y-periodic on $\mathbb{R}^n$. We further assume that the coefficients satisfy the Cordes condition, i.e., that there exist constants $\lambda > 0$ and $\delta \in (0,1)$ such that

$$|A^{\alpha\beta}|^2 + \frac{|b^{\alpha\beta}|^2}{2\lambda} + \frac{(c^{\alpha\beta})^2}{\lambda^2} \le \frac{1}{n+\delta}\left(\operatorname{tr}(A^{\alpha\beta}) + \frac{c^{\alpha\beta}}{\lambda}\right)^2$$

holds in $\mathbb{R}^n$ for all $(\alpha,\beta) \in \mathcal{A}\times\mathcal{B}$. Under these assumptions, the periodic HJBI problem (1.12) admits a unique strong solution $u \in H^2_{\text{per}}(Y)$; see Section 5.1.

We are seeking an approximation to the solution in either the discontinuous Galerkin finite element space $V^0_{\mathcal{T}}$ or the C^0-IP finite element space $V^1_{\mathcal{T}}$, which are defined by

$$V^0_{\mathcal{T}} := \left\{v_{\mathcal{T}} \in L^2(Y) : v_{\mathcal{T}}|_K \in \mathbb{P}_{\bar{p}}\ \forall K \in \mathcal{T}\right\}, \qquad V^1_{\mathcal{T}} := V^0_{\mathcal{T}} \cap H^1_{\text{per}}(Y)$$

for a given $\bar{p} \ge 2$. Here, $\mathbb{P}_{\bar{p}}$ denotes the space of polynomials of degree at most $\bar{p}$ and $\mathcal{T}$ is a triangulation of the computational domain consistent with periodicity requirements; see Section 5.1.2. For a chosen $s \in \{0,1\}$, we consider an abstract numerical scheme of finding a function $u_{\mathcal{T}} \in V^s_{\mathcal{T}}$ such that

$$a_{\mathcal{T}}(u_{\mathcal{T}}, v_{\mathcal{T}}) = 0 \qquad \forall v_{\mathcal{T}} \in V^s_{\mathcal{T}}$$

with a map $a_{\mathcal{T}} : V^s_{\mathcal{T}} \times V^s_{\mathcal{T}} \to \mathbb{R}$ that satisfies some reasonable assumptions. We perform a rigorous abstract *a priori* and *a posteriori* analysis and provide an example of a family of numerical schemes. In particular, let us point out that the *a posteriori* analysis is completely independent of the chosen numerical scheme. The numerical method that we consider in this chapter is presented in Section 5.1.4 and we refer the reader to Chapter 5 for detailed results.

The *second step* of the approximation scheme is the approximation of the effective Hamiltonian via (1.11), where $v^\sigma_{\mathcal{T}}$ is a discontinuous Galerkin or C^0-IP finite element approximation to the approximate corrector obtained through the first step. Let us note that the error analysis is a little different from the HJB case as e.g., we do not have a rate for the uniform convergence of $(-\sigma v^\sigma)_{\sigma>0}$ to the effective Hamiltonian and we do not have as nice regularity properties of the approximate correctors which have previously been guaranteed from the convexity of HJB operators $F^{\mathrm{HJB}} = F^{\mathrm{HJB}}[x, y, p, R]$ in the R-variable. We refer the reader to Section 5.2.3 for the error analysis. Finally, numerical experiments demonstrate the theoretical results for both the discontinuous Galerkin and the C^0-IP method.

Part I

Linear elliptic equations in nondivergence-form

Chapter 2

Homogenization of linear nondivergence-form equations

In this chapter, we study the homogenization of elliptic problems in nondivergence-form with periodic coefficients. The outline of this chapter is as follows.

We provide the statement of the problem in Section 2.1, i.e., we define sets of assumptions for the domain, the coefficients and the right-hand side, ensuring well-posedness of the problem. In Section 2.2 (qualitative homogenization), we introduce the invariant measure and describe a known procedure for transforming the original nondivergence-form problem into a divergence-form problem. This is used in combination with uniform $W^{2,p}$ estimates to carry out the homogenization for the problem under consideration.

In Section 2.3 (quantitative homogenization), we introduce corrector functions and obtain several corrector estimates in the $W^{2,p}$ and $C^{1,\nu}$ norms. Corrector estimates are crucial in deriving the optimal convergence rates for the convergence to the homogenized solution in Section 2.4.

Finally, numerical illustrations demonstrating the optimality of the obtained results are provided in Section 2.5.

Annotation: Unless stated otherwise, this chapter contains novel results which have been obtained in Capdeboscq, Sprekeler, Süli [29] and Sprekeler, Tran [90]. The contribution of Y. Capdeboscq, E. Süli and H. V. Tran was of advisory nature.

2.1 Framework

Throughout this work, we denote the unit cell in $\mathbb{R}^n$ by

$$Y := (0,1)^n.$$

Let us consider a symmetric matrix-valued function $A : \mathbb{R}^n \to \mathbb{R}^{n\times n}_{\mathrm{sym}}$ with the following properties:

(i) Regularity: $A \in C^{0,\alpha}(\mathbb{R}^n; \mathbb{R}^{n\times n})$ for some $\alpha \in (0, 1]$;

(ii) Periodicity: A is Y-periodic;

(iii) Ellipticity: there are $\lambda, \Lambda > 0$ such that $\forall\, \xi, y \in \mathbb{R}^n :\ \lambda|\xi|^2 \leq A(y)\xi \cdot \xi \leq \Lambda|\xi|^2$.

Note that we can equivalently write

$$A \in C^{0,\alpha}(\mathbb{T}^n; \mathcal{S}^n_+) \text{ for some } \alpha \in (0, 1]$$

with $\mathbb{T}^n := \mathbb{R}^n/\mathbb{Z}^n$ being the n-dimensional flat torus and $\mathcal{S}^n_+ \subset \mathbb{R}^{n\times n}$ the set of symmetric positive definite $n \times n$ matrices.

For a small parameter $\varepsilon > 0$, we are then concerned with linear elliptic problems of the form

$$\begin{cases} A\left(\frac{\cdot}{\varepsilon}\right) : D^2 u_\varepsilon = f & \text{in } \Omega, \\ \qquad\qquad u_\varepsilon = 0 & \text{on } \partial\Omega, \end{cases} \tag{2.1}$$

where the triple (Ω, A, f) satisfies one of the following sets of assumptions.

Definition 2.1.1 (Sets of assumptions $\mathcal{G}^{m,p}$, $\mathcal{H}^m$)**.** *For $m \in \mathbb{N}_0$ and $p \in (1, \infty)$, we define the set of assumptions $\mathcal{G}^{m,p}$ as*

$$(\Omega, A, f) \in \mathcal{G}^{m,p} \quad \Longleftrightarrow \quad \begin{cases} \Omega \subset \mathbb{R}^n \textit{ is a bounded } C^{2,\gamma} \textit{ domain}, \gamma \in (0, 1), \\ A \in C^{0,\alpha}(\mathbb{T}^n; \mathcal{S}^n_+) \textit{ for some } \alpha \in (0, 1], \\ f \in W^{m,p}(\Omega), \end{cases}$$

and the set of assumptions $\mathcal{H}^m$ as

$$(\Omega, A, f) \in \mathcal{H}^m \quad \Longleftrightarrow \quad \begin{cases} \Omega \subset \mathbb{R}^n \textit{ is a bounded convex domain}, \\ A \in C^{0,\alpha}(\mathbb{T}^n; \mathcal{S}^n_+) \textit{ for some } \alpha \in (0, 1], \\ \exists\, \delta \in (0, 1] : \dfrac{|A|^2}{(\mathrm{tr}A)^2} \leq \dfrac{1}{n - 1 + \delta} \quad \textit{in } \mathbb{R}^n, \\ f \in H^m(\Omega). \end{cases}$$

Before we proceed, let us note that the Cordes condition (dating back to [31]), i.e., that there exists a constant $\delta \in (0, 1]$ such that

$$\frac{|A(y)|^2}{(\mathrm{tr}A(y))^2} := \frac{A(y) : A(y)}{(\mathrm{tr}A(y))^2} \leq \frac{1}{n - 1 + \delta} \qquad \forall\, y \in \mathbb{R}^n, \tag{2.2}$$

is a consequence of the uniform ellipticity condition in dimension $n = 2$. Indeed, let $A \in C^{0,\alpha}(\mathbb{T}^2; \mathcal{S}_+^2)$ for some $\alpha \in (0,1]$ with ellipticity constants $0 < \lambda \leq \Lambda$ such that $A(y)\frac{\xi}{|\xi|} \cdot \frac{\xi}{|\xi|} \in [\lambda, \Lambda]$ for all $\xi, y \in \mathbb{R}^2$ with $|\xi| \neq 0$. Then, introducing the eigenvalues $\lambda_{\min}, \lambda_{\max} : \mathbb{R}^2 \to [\lambda, \Lambda]$ given by

$$\lambda_{\min}(y) := \min_{\xi \in \mathbb{R}^2 \setminus \{0\}} \frac{A(y)\xi \cdot \xi}{|\xi|^2}, \qquad \lambda_{\max}(y) := \max_{\xi \in \mathbb{R}^2 \setminus \{0\}} \frac{A(y)\xi \cdot \xi}{|\xi|^2},$$

we find that the Cordes condition (2.2) holds with $\delta = \frac{\lambda}{\Lambda} \in (0,1]$ as we have that

$$\frac{|A|^2}{(\mathrm{tr}A)^2} = \frac{\lambda_{\max}^2 + \lambda_{\min}^2}{(\lambda_{\max} + \lambda_{\min})^2} = \frac{1 + \left(\frac{\lambda_{\min}}{\lambda_{\max}}\right)^2}{\left(1 + \frac{\lambda_{\min}}{\lambda_{\max}}\right)^2} \leq \frac{1}{1 + \frac{\lambda_{\min}}{\lambda_{\max}}} \leq \frac{1}{1 + \frac{\lambda}{\Lambda}} \quad \text{in } \mathbb{R}^2.$$

Here we have used that $\mathrm{tr}(A) = \lambda_{\max} + \lambda_{\min}$ and $\det(A) = \lambda_{\max}\lambda_{\min}$, as well as the fact that we can express the Frobenius norm of a symmetric 2×2 matrix as

$$|M| = \sqrt{(\mathrm{tr}M)^2 - 2\det(M)} \qquad \forall M \in \mathbb{R}^{2\times 2}_{\mathrm{sym}}.$$

Therefore, the set of assumptions $\mathcal{H}^m$ can be simplified when $n = 2$.

Remark 2.1.1. *When $n = 2$, the set $\mathcal{H}^m$ can be simplified to*

$$(\Omega, A, f) \in \mathcal{H}^m \qquad \Longleftrightarrow \qquad \begin{cases} \Omega \subset \mathbb{R}^2 \text{ is a bounded convex domain,} \\ A \in C^{0,\alpha}(\mathbb{T}^2; \mathcal{S}_+^2) \text{ for some } \alpha \in (0,1], \\ f \in H^m(\Omega). \end{cases}$$

The following theorem asserts well-posedness, i.e., the existence and uniqueness of strong solutions for the problem (2.1); see [53, 87]. We provide a proof for the case of convex domains for completeness.

Theorem 2.1.1 ([29, Theorem 2.3] Existence and uniqueness of strong solutions)**.** *Assume either that $(\Omega, A, f) \in \mathcal{G}^{0,p}$ for some $p \in (1,\infty)$, or that $(\Omega, A, f) \in \mathcal{H}^0$ and $p = 2$. Then, for any $\varepsilon > 0$, problem (2.1) admits a unique solution $u_\varepsilon \in W^{2,p}(\Omega) \cap W_0^{1,p}(\Omega)$.*

Proof. We note that the existence and uniqueness of strong solutions in the case $(\Omega, A, f) \in \mathcal{G}^{0,p}$ with $p \in (1,\infty)$ is a standard result from elliptic PDE theory, and we refer to the classical book Gilbarg, Trudinger [53, Theorem 9.15] and omit the proof for this case. We now prove the result for the case $(\Omega, A, f) \in \mathcal{H}^0$ and follow the proof of [87, Theorem 3].

Let us assume $(\Omega, A, f) \in \mathcal{H}^0$ and fix $\varepsilon > 0$. Noting that $\Omega \subset \mathbb{R}^n$ is a bounded convex domain, we recall the Miranda–Talenti estimates

$$\|D^2 v\|_{L^2(\Omega)} \leq \|\Delta v\|_{L^2(\Omega)}, \qquad \|v\|_{H^2(\Omega)} \leq C\|\Delta v\|_{L^2(\Omega)} \tag{2.3}$$

for $v \in H^2(\Omega) \cap H_0^1(\Omega)$, where $C = C(n, \operatorname{diam}(\Omega)) > 0$ is a constant only depending on n and $\operatorname{diam}(\Omega)$; see [87]. For simplicity, let us write $H := H^2(\Omega) \cap H_0^1(\Omega)$.

Step 1: We introduce the continuous function $\gamma : \mathbb{R}^n \to \mathbb{R}$ given by

$$\gamma(y) := \frac{\operatorname{tr}(A(y))}{|A(y)|^2}, \qquad y \in \mathbb{R}^n,$$

and note that $\gamma \in L^\infty(\mathbb{R}^n)$ is positive, i.e., $\inf_{\mathbb{R}^n} \gamma > 0$. We claim that

$$\left\| \gamma\left(\frac{\cdot}{\varepsilon}\right) A\left(\frac{\cdot}{\varepsilon}\right) : D^2 v - \Delta v \right\|_{L^2(\Omega)} \leq \sqrt{1-\delta}\|\Delta v\|_{L^2(\Omega)} \qquad \forall\, v \in H. \tag{2.4}$$

Indeed, a simple calculation yields that for any $v \in H$, we have

$$\left|\gamma\left(\frac{\cdot}{\varepsilon}\right) A\left(\frac{\cdot}{\varepsilon}\right) : D^2 v - \Delta v\right|^2 \leq \left|\gamma\left(\frac{\cdot}{\varepsilon}\right) A\left(\frac{\cdot}{\varepsilon}\right) - I\right|^2 |D^2 v|^2 = \left(n - \frac{(\operatorname{tr}(A))^2}{|A|^2}\right) |D^2 v|^2$$

almost everywhere in Ω and hence, (2.4) follows from the Cordes condition (2.2) and the Miranda–Talenti estimate (2.3).

Step 2: We claim that there exists a unique $u_\varepsilon \in H$ such that

$$a(u_\varepsilon, v) = l(v) \qquad \forall\, v \in H, \tag{2.5}$$

where a is the bilinear form

$$a : H \times H \to \mathbb{R}, \qquad a(u, v) := \int_\Omega \gamma\left(\frac{\cdot}{\varepsilon}\right) A\left(\frac{\cdot}{\varepsilon}\right) : D^2 u \,\Delta v$$

and l the linear functional

$$l : H \to \mathbb{R}, \qquad l(v) := \int_\Omega \gamma\left(\frac{\cdot}{\varepsilon}\right) f \,\Delta v.$$

Indeed, the claim follows from the Lax–Milgram theorem as a and l are bounded and the bilinear form a is coercive. The latter follows from the fact that (2.4) yields

$$a(v, v) = \|\Delta v\|_{L^2(\Omega)}^2 + \int_\Omega \left(\gamma\left(\frac{\cdot}{\varepsilon}\right) A\left(\frac{\cdot}{\varepsilon}\right) : D^2 v - \Delta v\right) \Delta v \geq (1 - \sqrt{1-\delta})\|\Delta v\|_{L^2(\Omega)}^2$$

and hence, by the Miranda–Talenti estimate (2.3),

$$\|v\|_{H^2(\Omega)}^2 \lesssim a(v, v) \tag{2.6}$$

for any $v \in H$, where the constant only depends on δ, n and $\mathrm{diam}(\Omega)$.

Step 3: We deduce from the result of the previous step (2.5) and the surjectivity of the operator $\Delta : H \to L^2(\Omega)$, that

$$\gamma\left(\frac{\cdot}{\varepsilon}\right) A\left(\frac{\cdot}{\varepsilon}\right) : D^2 u_\varepsilon = \gamma\left(\frac{\cdot}{\varepsilon}\right) f$$

almost everywhere in Ω. Therefore, as $\inf_{\mathbb{R}^n} \gamma > 0$, we conclude that $u_\varepsilon \in H$ is the unique strong solution to (2.1). $\square$

2.2 Qualitative homogenization: the convergence result

2.2.1 Transformation into divergence-form

We recall a well-known procedure to transform the problem (2.1) into divergence-form; see [15, 20]. We use the notation

$$W_{\mathrm{per}}(Y) := \left\{ v \in H^1_{\mathrm{per}}(Y) : \int_Y v = 0 \right\}.$$

Let us start by introducing the notion of invariant measure; see [20].

Lemma 2.2.1 (Invariant measure and solvability condition)**.** *Let $A \in C^{0,\alpha}(\mathbb{T}^n; \mathcal{S}^n_+)$ for some $\alpha \in (0,1]$. Then, there exists a unique solution $m : \mathbb{R}^n \to \mathbb{R}$ to the problem*

$$D^2 : (Am) = 0 \quad \text{in } Y, \qquad m \text{ is } Y\text{-periodic}, \qquad \int_Y m = 1.$$

This function m, called the invariant measure corresponding to the coefficient A, is Hölder continuous (see [22, 23]) and satisfies $\inf_{\mathbb{R}^n} m > 0$*. Moreover, for a Y-periodic function $g \in L^2_{\mathrm{per}}(Y)$, the adjoint problem*

$$A : D^2 v = g \quad \text{in } Y, \qquad v \text{ is } Y\text{-periodic}, \qquad \int_Y v = 0,$$

admits a (unique) solution $v \in W_{\mathrm{per}}(Y)$ if and only if

$$\langle g, m \rangle_{L^2(Y)} = 0. \tag{2.7}$$

With the invariant measure m at hand, we can easily convert the problem into divergence-form as follows. We define a matrix-valued function $B = (b_{ij})_{1 \le i,j \le n} : \mathbb{R}^n \to \mathbb{R}^{n \times n}$ by

$$b_{ij} := \partial_i v_j - \partial_j v_i, \quad (1 \le i, j \le n),$$

with $v_l \in W_{\mathrm{per}}(Y)$ denoting the solution to

$$-\Delta v_l = \operatorname{div}(Am) \cdot e_l \quad \text{in } Y, \qquad v_l \text{ is } Y\text{-periodic}, \qquad \int_Y v_l = 0,$$

for $1 \leq l \leq n$. Since the coefficient A and the invariant measure m are Hölder continuous, so is the matrix-valued function B by elliptic regularity. Further, we observe that B is skew-symmetric, Y-periodic with zero mean over Y, and that

$$\operatorname{div}(B) = -\operatorname{div}(Am) \quad \text{a.e. on } \mathbb{R}^n.$$

Now we define the matrix-valued function

$$A^{\mathrm{div}} := Am + B,$$

which is again Hölder continuous. Then, since

$$\operatorname{div}(A^{\mathrm{div}}) = 0,$$

and using the fact that B is skew-symmetric, we obtain

$$\nabla \cdot \left(A^{\mathrm{div}} \left(\frac{\cdot}{\varepsilon} \right) \nabla u_\varepsilon \right) = A^{\mathrm{div}} \left(\frac{\cdot}{\varepsilon} \right) : D^2 u_\varepsilon = (Am) \left(\frac{\cdot}{\varepsilon} \right) : D^2 u_\varepsilon,$$

that is, we have converted (2.1) into divergence-form:

$$\begin{cases} \nabla \cdot \left(A^{\mathrm{div}} \left(\frac{\cdot}{\varepsilon} \right) \nabla u_\varepsilon \right) = f\, m \left(\frac{\cdot}{\varepsilon} \right) & \text{in } \Omega, \\ u_\varepsilon = 0 & \text{on } \partial\Omega, \end{cases} \tag{2.8}$$

and it is straightforward to check that A^{div} is Y-periodic, Hölder continuous on $\mathbb{R}^n$ and uniformly elliptic.

Remark 2.2.1. *If the coefficient A admits Sobolev regularity $A \in W^{1,q}(Y; \mathbb{R}^{n \times n})$ for some $q > n$, then so does the corresponding invariant measure $m \in W^{1,q}(Y)$, see [22, 23], and the coefficient $A^{\mathrm{div}} \in W^{1,q}(Y; \mathbb{R}^{n \times n})$.*

2.2.2 Uniform estimates in the $W^{2,p}$-norm

The transformation described in the previous section can be used to obtain uniform $W^{2,p}(\Omega)$ *a priori* estimates for the solution of (2.1), which are crucial in deriving homogenization results.

Theorem 2.2.1 ([29, Theorem 2.5] Uniform $W^{2,p}$ *a priori* estimates)**.** *Assume either that* $(\Omega, A, f) \in \mathcal{G}^{0,p}$ *for some* $p \in (1,\infty)$*, or that* $(\Omega, A, f) \in \mathcal{H}^0$ *and* $p = 2$*. Then, for* $\varepsilon \in (0,1]$*, the solution* $u_\varepsilon \in W^{2,p}(\Omega) \cap W_0^{1,p}(\Omega)$ *to* (2.1)*, whose existence and uniqueness are guaranteed by Theorem 2.1.1, satisfies*

$$\|u_\varepsilon\|_{W^{2,p}(\Omega)} \lesssim \|f\|_{L^p(\Omega)},$$

with the constant absorbed into the notation $\lesssim$ *being independent of* ε*.*

Proof. Let us first assume that $(\Omega, A, f) \in \mathcal{G}^{0,p}$ for some $p \in (1,\infty)$. We showed in the previous section that we can transform problem (2.1) into the divergence-form problem (2.8), where $A^{\text{div}} : \mathbb{R}^n \to \mathbb{R}^{n\times n}$ is a Y-periodic, Hölder continuous, and uniformly elliptic matrix-valued function satisfying

$$\operatorname{div}(A^{\text{div}}) = 0.$$

Therefore, we can apply [16, Theorem D] to problem (2.8) to obtain

$$\|u_\varepsilon\|_{W^{2,p}(\Omega)} \lesssim \left\| f\, m\left(\frac{\cdot}{\varepsilon}\right)\right\|_{L^p(\Omega)} \lesssim \|f\|_{L^p(\Omega)}$$

with constants independent of ε.

Let us now assume that $(\Omega, A, f) \in \mathcal{H}^0$. The proof of Theorem 2.1.1, more precisely (2.5) and (2.6), yields the estimate

$$\|u_\varepsilon\|^2_{H^2(\Omega)} \lesssim a(u_\varepsilon, u_\varepsilon) = l(u_\varepsilon) \leq \|\gamma\|_{L^\infty(\mathbb{R}^n)}\|f\|_{L^2(\Omega)}\|\Delta u_\varepsilon\|_{L^2(\Omega)} \lesssim \|f\|_{L^2(\Omega)}\|u_\varepsilon\|_{H^2(\Omega)}$$

with constants only depending on $\delta, n, \operatorname{diam}(\Omega)$ and $\|\gamma\|_{L^\infty(\mathbb{R}^n)}$. In particular,

$$\|u_\varepsilon\|_{H^2(\Omega)} \lesssim \|f\|_{L^2(\Omega)}$$

with constant independent of ε. □

2.2.3 The convergence result

This leads to a simple proof of the homogenization theorem for problem (2.1), using the compactness of the embedding $W^{2,p}(\Omega) \hookrightarrow W^{1,p}(\Omega)$ and the fact that we can rewrite the problem as (2.8).

Theorem 2.2.2 ([29, Theorem 2.6] Homogenization theorem)**.** *Assume either that* $(\Omega, A, f) \in \mathcal{G}^{0,p}$ *for some* $p \in (1,\infty)$*, or that* $(\Omega, A, f) \in \mathcal{H}^0$ *and* $p = 2$*. Then the* *solution* $u_\varepsilon \in W^{2,p}(\Omega) \cap W_0^{1,p}(\Omega)$ *to* (2.1) *converges weakly in* $W^{2,p}(\Omega)$ *to the solution* $u_0 \in W^{2,p}(\Omega) \cap W_0^{1,p}(\Omega)$ *of the homogenized problem*

$$\begin{cases} A^0 : D^2 u_0 = f & \text{in } \Omega, \\ u_0 = 0 & \text{on } \partial\Omega, \end{cases} \tag{2.9}$$

with the effective coefficient $A^0 \in \mathbb{R}^{n \times n}$ *being the constant matrix given by*

$$A^0 := \int_Y Am, \tag{2.10}$$

where m *is the invariant measure corresponding to* A.

Proof. By Theorem 2.2.1, the reflexivity of $W^{2,p}(\Omega)$ for $p \in (1, \infty)$, the compactness of the embedding $W^{2,p}(\Omega) \hookrightarrow W^{1,p}(\Omega)$, and the properties of the trace operator, there exists a $u_0 \in W^{2,p}(\Omega) \cap W_0^{1,p}(\Omega)$ such that (for a subsequence, not indicated,)

$$\begin{aligned} u_\varepsilon &\rightharpoonup u_0 \quad \text{weakly in } W^{2,p}(\Omega), \text{ and} \\ u_\varepsilon &\to u_0 \quad \text{strongly in } W^{1,p}(\Omega). \end{aligned}$$

We can transform (2.1) as in Section 2.2.1 into the divergence-form problem (2.8) with

$$A^{\text{div}} = Am + B$$

being Y-periodic, Hölder continuous and uniformly elliptic on $\mathbb{R}^n$. Recalling that B is of mean zero over Y, we have

$$A^{\text{div}}\left(\frac{\cdot}{\varepsilon}\right) \overset{*}{\rightharpoonup} \int_Y Am = A^0 \quad \text{weakly-}* \text{ in } L^\infty(\Omega).$$

Since we have that

$$\nabla u_\varepsilon \to \nabla u_0 \quad \text{strongly in } L^p(\Omega),$$

we can pass to the limit in the weak formulation of (2.8) to obtain that $u_0 \in W^{2,p}(\Omega) \cap W_0^{1,p}(\Omega)$ solves (2.9). We conclude the proof by noting that (2.9) admits a unique strong solution in $W^{2,p}(\Omega) \cap W_0^{1,p}(\Omega)$. □

Let us note that the effective coefficient A^0 from (2.10) can be equivalently characterized via so-called correctors: For $i, j \in \{1, \dots, n\}$, the (i,j)-th entry $a_{ij}^0 \in \mathbb{R}$ of A^0 is the unique value such that the periodic cell problem

$$A : D^2 \chi_{ij} = a_{ij}^0 - a_{ij} \quad \text{in } Y, \qquad \chi_{ij} \text{ is } Y\text{-periodic}, \qquad \int_Y \chi_{ij} = 0 \tag{2.11}$$

admits a unique solution $\chi_{ij} : \mathbb{R}^n \to \mathbb{R}$, called a corrector function. Indeed, this is a consequence of the solvability condition from Lemma 2.2.1.

2.3 Quantitative homogenization: corrector estimates

Let us introduce the matrix of corrector functions $V = (\chi_{ij})_{1\leq i,j\leq n}$ given by (2.11) and a boundary corrector θ_ε as the solution to the problem

$$\begin{cases} A\left(\frac{\cdot}{\varepsilon}\right) : D^2\theta_\varepsilon = 0 & \text{in } \Omega, \\ \theta_\varepsilon = -V\left(\frac{\cdot}{\varepsilon}\right) : D^2 u_0 & \text{on } \partial\Omega. \end{cases} \tag{2.12}$$

Then, defining the function

$$\phi_\varepsilon := \varepsilon^2 \left[V\left(\frac{\cdot}{\varepsilon}\right) : D^2 u_0 + \theta_\varepsilon\right], \tag{2.13}$$

we have that the function $u_\varepsilon - u_0 - \phi_\varepsilon$ satisfies the problem

$$\begin{cases} A\left(\frac{\cdot}{\varepsilon}\right) : D^2(u_\varepsilon - u_0 - \phi_\varepsilon) = -\varepsilon F^\varepsilon & \text{in } \Omega, \\ u_\varepsilon - u_0 - \phi_\varepsilon = 0 & \text{on } \partial\Omega, \end{cases}$$

with F^ε given by

$$F^\varepsilon := \sum_{i,j,k,l=1}^{n} a_{ij}\left(\frac{\cdot}{\varepsilon}\right)\left[2\partial_i\chi_{kl}\left(\frac{\cdot}{\varepsilon}\right)\partial^3_{jkl}u_0 + \varepsilon\chi_{kl}\left(\frac{\cdot}{\varepsilon}\right)\partial^4_{ijkl}u_0\right].$$

If the sequence $(F^\varepsilon)_\varepsilon$ is uniformly bounded in a suitable function space, we can apply uniform estimates to this problem to deduce results of the form

$$\|u_\varepsilon - u_0 - \phi_\varepsilon\|_X = \mathcal{O}(\varepsilon)$$

in a suitable norm $\|\cdot\|_X$.

2.3.1 Corrector estimate in the $W^{2,p}$-norm

If $(F^\varepsilon)_\varepsilon$ is uniformly bounded in $L^p(\Omega)$, we can apply the uniform $W^{2,p}$ estimate to obtain a convergence result:

Theorem 2.3.1 ([29, Theorem 2.7] $W^{2,p}$ corrector estimate I)**.** *Assume either that $(\Omega, A, f) \in \mathcal{G}^{2,p}$ for some $p \in (1,\infty)$, or that $(\Omega, A, f) \in \mathcal{H}^2$ and $p = 2$. Let $\varepsilon \in (0,1]$ and assume $u_0 \in W^{4,p}(\Omega)$. Then, we have the bound*

$$\left\|u_\varepsilon - u_0 - \varepsilon^2\left(V\left(\frac{\cdot}{\varepsilon}\right) : D^2 u_0 + \theta_\varepsilon\right)\right\|_{W^{2,p}(\Omega)} \lesssim \varepsilon\|u_0\|_{W^{4,p}(\Omega)}. \tag{2.14}$$

Proof. First, we note that $\chi_{ij} \in C^{2,\alpha}(\mathbb{R}^n)$ for any $1 \leq i, j \leq n$ since $A \in C^{0,\alpha}(\mathbb{T}^n; \mathcal{S}^n_+)$ by elliptic regularity theory. Since $u_0 \in W^{4,p}(\Omega)$, the sequence $(F^\varepsilon)_\varepsilon$ is uniformly bounded in $L^p(\Omega)$ and we conclude using the uniform $W^{2,p}$ estimate from Theorem 2.2.1 that

$$\|u_\varepsilon - u_0 - \phi_\varepsilon\|_{W^{2,p}(\Omega)} \lesssim \varepsilon \|F^\varepsilon\|_{L^p(\Omega)} \lesssim \varepsilon \|u_0\|_{W^{4,p}(\Omega)},$$

which yields (2.14) by definition of the function ϕ_ε. □

The following theorem shows that if $u_0 \in W^{4,p}(\Omega) \cap W^{2,\infty}(\Omega)$, then we can absorb the term involving the boundary corrector into the right-hand side at the cost of powers of ε.

Theorem 2.3.2 ([29, Theorem 2.8] $W^{2,p}$ corrector estimate II)**.** *Assume either that $(\Omega, A, f) \in \mathcal{G}^{2,p}$ for some $p \in (1,\infty)$, or that $(\Omega, A, f) \in \mathcal{H}^2$ and $p = 2$. Let $\varepsilon \in (0,1]$ and assume $u_0 \in W^{4,p}(\Omega) \cap W^{2,\infty}(\Omega)$. Then we have*

$$\varepsilon^2 \|\theta_\varepsilon\|_{W^{2,p}(\Omega)} \lesssim \varepsilon^{\frac{1}{p}} \|u_0\|_{W^{2,\infty}(\Omega)} + \varepsilon \|u_0\|_{W^{4,p}(\Omega)},$$

and therefore, in view of Theorem 2.3.1, we have

$$\left\| u_\varepsilon - u_0 - \varepsilon^2 V\left(\frac{\cdot}{\varepsilon}\right) : D^2 u_0 \right\|_{W^{2,p}(\Omega)} \lesssim \varepsilon^{\frac{1}{p}} \|u_0\|_{W^{2,\infty}(\Omega)} + \varepsilon \|u_0\|_{W^{4,p}(\Omega)}.$$

Proof. Let $\eta \in C_c^\infty(\mathbb{R}^n)$ be a cut-off function with $0 \leq \eta \leq 1$,

$$\begin{aligned} \eta &\equiv 1 \quad \text{in } \left\{ x \in \Omega : \operatorname{dist}(x, \partial\Omega) < \frac{\varepsilon}{2} \right\}, \\ \eta &\equiv 0 \quad \text{in } \left\{ x \in \Omega : \operatorname{dist}(x, \partial\Omega) \geq \varepsilon \right\}, \end{aligned}$$

and let η satisfy

$$|\nabla \eta| + \varepsilon |D^2 \eta| \lesssim \frac{1}{\varepsilon} \quad \text{in } \Omega.$$

We introduce the function

$$\tilde{\theta}_\varepsilon := \theta_\varepsilon + \eta\, V\left(\frac{\cdot}{\varepsilon}\right) : D^2 u_0,$$

and verify that it satisfies

$$\begin{cases} A\left(\frac{\cdot}{\varepsilon}\right) : D^2 \tilde{\theta}_\varepsilon = \frac{1}{\varepsilon^2} S_1 + \frac{1}{\varepsilon} S_2 + S_3 & \text{in } \Omega, \\ \tilde{\theta}_\varepsilon = 0 & \text{on } \partial\Omega, \end{cases}$$

where S_1, S_2 and S_3 are given by

$$S_1 := \sum_{i,j,k,l=1}^{n} a_{ij}\left(\frac{\cdot}{\varepsilon}\right) \eta\, \partial^2_{ij}\chi_{kl}\left(\frac{\cdot}{\varepsilon}\right) \partial^2_{kl}u_0,$$

$$S_2 := 2\sum_{i,j,k,l=1}^{n} a_{ij}\left(\frac{\cdot}{\varepsilon}\right) \left(\partial_i\eta\, \partial_j\chi_{kl}\left(\frac{\cdot}{\varepsilon}\right) \partial^2_{kl}u_0 + \eta\, \partial_i\chi_{kl}\left(\frac{\cdot}{\varepsilon}\right) \partial^3_{jkl}u_0\right),$$

$$S_3 := \sum_{i,j,k,l=1}^{n} a_{ij}\left(\frac{\cdot}{\varepsilon}\right) \left(\partial^2_{ij}\eta\, \chi_{kl}\left(\frac{\cdot}{\varepsilon}\right) \partial^2_{kl}u_0 + 2\,\partial_i\eta\, \chi_{kl}\left(\frac{\cdot}{\varepsilon}\right) \partial^3_{jkl}u_0 + \eta\, \chi_{kl}\left(\frac{\cdot}{\varepsilon}\right) \partial^4_{ijkl}u_0\right).$$

Due to the fact that $u_0 \in W^{4,p}(\Omega) \cap W^{2,\infty}(\Omega)$, the boundedness of A and $\chi_{ij} \in W^{2,\infty}(\mathbb{R}^n)$, the right-hand side belongs to $L^p(\Omega)$, and we have by Theorem 2.2.1 that

$$\left\|\tilde{\theta}_\varepsilon\right\|_{W^{2,p}(\Omega)} \lesssim \frac{1}{\varepsilon^2}\|S_1\|_{L^p(\Omega)} + \frac{1}{\varepsilon}\|S_2\|_{L^p(\Omega)} + \|S_3\|_{L^p(\Omega)}.$$

We bound the terms on the right-hand side separately using the following bounds on the cut-off:

$$\|\eta\|_{L^p(\Omega)} + \varepsilon\|\nabla\eta\|_{L^p(\Omega)} + \varepsilon^2\|D^2\eta\|_{L^p(\Omega)} \lesssim \varepsilon^{\frac{1}{p}}.$$

For S_1, we have

$$\|S_1\|_{L^p(\Omega)} \lesssim \|\eta\|_{L^p(\Omega)}\|u_0\|_{W^{2,\infty}(\Omega)} \lesssim \varepsilon^{\frac{1}{p}}\|u_0\|_{W^{2,\infty}(\Omega)}.$$

For S_2, we obtain similarly that

$$\begin{aligned}\|S_2\|_{L^p(\Omega)} &\lesssim \|\nabla\eta\|_{L^p(\Omega)}\|u_0\|_{W^{2,\infty}(\Omega)} + \|\eta\|_{L^\infty(\Omega)}\|u_0\|_{W^{4,p}(\Omega)}\\ &\lesssim \varepsilon^{\frac{1}{p}-1}\|u_0\|_{W^{2,\infty}(\Omega)} + \|u_0\|_{W^{4,p}(\Omega)}.\end{aligned}$$

Finally, for S_3, we have that

$$\begin{aligned}\|S_3\|_{L^p(\Omega)} &\lesssim \|D^2\eta\|_{L^p(\Omega)}\|u_0\|_{W^{2,\infty}(\Omega)} + \left(\|\nabla\eta\|_{L^\infty(\Omega)} + \|\eta\|_{L^\infty(\Omega)}\right)\|u_0\|_{W^{4,p}(\Omega)}\\ &\lesssim \varepsilon^{\frac{1}{p}-2}\|u_0\|_{W^{2,\infty}(\Omega)} + \varepsilon^{-1}\|u_0\|_{W^{4,p}(\Omega)}.\end{aligned}$$

Altogether, we have shown that

$$\left\|\tilde{\theta}_\varepsilon\right\|_{W^{2,p}(\Omega)} \lesssim \varepsilon^{\frac{1}{p}-2}\|u_0\|_{W^{2,\infty}(\Omega)} + \varepsilon^{-1}\|u_0\|_{W^{4,p}(\Omega)}.$$

A direct computation, using the bounds on the cut-off, yields the estimate

$$\left\|\eta\, V\left(\frac{\cdot}{\varepsilon}\right) : D^2u_0\right\|_{W^{2,p}(\Omega)} \lesssim \varepsilon^{\frac{1}{p}-2}\|u_0\|_{W^{2,\infty}(\Omega)} + \varepsilon^{-1}\|u_0\|_{W^{4,p}(\Omega)}$$

and therefore, using the triangle inequality, we obtain

$$\|\theta_\varepsilon\|_{W^{2,p}(\Omega)} \lesssim \varepsilon^{\frac{1}{p}-2}\|u_0\|_{W^{2,\infty}(\Omega)} + \varepsilon^{-1}\|u_0\|_{W^{4,p}(\Omega)},$$

which yields the claim. □

Let us remark that $W^{4,p}(\Omega) \hookrightarrow W^{2,\infty}(\Omega)$ for $p > \frac{n}{2}$, i.e., the assumption $u_0 \in W^{4,p}(\Omega) \cap W^{2,\infty}(\Omega)$ is for $p > \frac{n}{2}$ a consequence of $u_0 \in W^{4,p}(\Omega)$; in particular, for dimensions $n \in \{2, 3\}$ and $p = 2$, one can replace the condition $u_0 \in W^{4,p}(\Omega) \cap W^{2,\infty}(\Omega)$ by the sufficient condition $u_0 \in H^4(\Omega)$.

Let us recall that u_0 is the solution to the elliptic constant-coefficient problem (2.9). For bounded convex polygonal domains ($n = 2$), $u_0 \in H^4(\Omega)$ can be ensured by assuming that $f \in H^2(\Omega)$ satisfies certain compatibility conditions at the corners of the domain. In the case of Poisson's equation on $\Omega = (0,1)^2$, a necessary and sufficient condition for $u_0 \in H^4(\Omega) \cap H_0^1(\Omega)$ is that $f \in H^2(\Omega)$ and $f = 0$ at the corners of Ω; see [59]. We note that these conditions are satisfied for functions $f \in H^2(\Omega)$ such that $\mathrm{supp}(f) \Subset \Omega$; see [56].

2.3.2 Corrector estimate in the $C^{1,\nu}$-norm and $W^{1,p}$ rate

If $(F^\varepsilon)_\varepsilon$ is uniformly bounded in $L^q(\Omega)$ for some $q > n$, we can make use of a uniform $C^{1,\nu}$ estimate from [15] which we recall below:

Lemma 2.3.1 (Uniform $C^{1,\nu}$ estimate)**.** *Let $(\Omega, A, f) \in \mathcal{G}^{0,q}$ for some $q > n$. For $\varepsilon > 0$, let u_ε be the solution to the problem (2.1). Then there exists $\nu \in (0,1]$ such that there holds*

$$\|u_\varepsilon\|_{C^{1,\nu}(\Omega)} \leq C\|f\|_{L^q(\Omega)}$$

with a constant $C > 0$ independent of ε.

Note that this result holds for $C^{2,\gamma}$ domains and we will discuss the case of non-smooth domains later. Using Lemma 2.3.1, we obtain a corrector estimate in the $C^{1,\nu}$-norm:

Lemma 2.3.2 ($C^{1,\nu}$ corrector estimate)**.** *Assume that $(\Omega, A, f) \in \mathcal{G}^{2,q}$ and $u_0 \in W^{4,q}(\Omega)$ for some $q > n$. Then, there exists $\nu \in (0,1]$ such that we have*

$$\left\| u_\varepsilon - u_0 - \varepsilon^2 \left(V\left(\frac{\cdot}{\varepsilon}\right) : D^2 u_0 + \theta_\varepsilon \right) \right\|_{C^{1,\nu}(\Omega)} \lesssim \varepsilon \|u_0\|_{W^{4,q}(\Omega)}. \tag{2.15}$$

Proof. Since $u_0 \in W^{4,q}(\Omega)$, the sequence $(F^\varepsilon)_\varepsilon$ is uniformly bounded in $L^q(\Omega)$ and we conclude using the uniform estimate from Lemma 2.3.1 that there exists $\nu \in (0,1]$ such that we have

$$\|u_\varepsilon - u_0 - \phi_\varepsilon\|_{C^{1,\nu}(\Omega)} \lesssim \varepsilon \|F^\varepsilon\|_{L^q(\Omega)} \lesssim \varepsilon \|u_0\|_{W^{4,q}(\Omega)},$$

which yields (2.15) by definition of the function ϕ_ε. □

We are going to consider the weaker $W^{1,p}$-norm and eliminate the boundary corrector from this estimate. It turns out to be useful to transform the problem (2.12) into the divergence-form problem

$$\begin{cases} \nabla \cdot \left(A^{\mathrm{div}} \left(\frac{\cdot}{\varepsilon} \right) \nabla \theta_\varepsilon \right) = 0 & \text{in } \Omega, \\ \theta_\varepsilon = -V \left(\frac{\cdot}{\varepsilon} \right) : D^2 u_0 & \text{on } \partial\Omega, \end{cases}$$

with a coefficient $A^{\mathrm{div}} \in C^{0,\alpha}(\mathbb{T}^n; \mathbb{R}^{n\times n})$ for some $\alpha \in (0,1]$ that is uniformly elliptic, see Section 2.2.1.

We will then make use of the uniform $W^{1,p}$ estimate from [16] for divergence-form homogenization problems.

Lemma 2.3.3 (Uniform $W^{1,p}$ estimate for divergence-form equations)**.** *Let $\Omega \subset \mathbb{R}^n$ be a bounded $C^{2,\gamma}$ domain. Assume that $A^{\mathrm{div}} \in C^{0,\alpha}(\mathbb{T}^n; \mathbb{R}^{n\times n})$ for some $\alpha \in (0,1]$ is a uniformly elliptic coefficient, $F \in L^p(\Omega)$ and $G \in W^{1,p}(\Omega)$ for some $p \in (1,\infty)$. For $\varepsilon \in (0,1]$, let $\rho_\varepsilon \in W^{1,p}(\Omega)$ be the solution to the problem*

$$\begin{cases} -\nabla \cdot \left(A^{\mathrm{div}} \left(\frac{\cdot}{\varepsilon} \right) \nabla \rho_\varepsilon \right) = -\nabla \cdot F & \text{in } \Omega, \\ \rho_\varepsilon = G & \text{on } \partial\Omega. \end{cases}$$

Then we have the estimate

$$\|\rho_\varepsilon\|_{W^{1,p}(\Omega)} \leq C \left(\|F\|_{L^p(\Omega)} + \|G\|_{W^{1,p}(\Omega)} \right)$$

with a constant $C > 0$ independent of ε.

Note that symmetry of the coefficient is not required in the above result. This is important as the procedure of transforming nondivergence-form equations into divergence-form usually yields a nonsymmetric coefficient A^{div}.

We then have the following result on the asymptotic behavior of the boundary corrector θ_ε in the $W^{1,p}$-norm:

Lemma 2.3.4 (Boundary corrector $W^{1,p}$ bound)**.** *Assume that $(\Omega, A, f) \in \mathcal{G}^{2,q}$ and $u_0 \in W^{4,q}(\Omega)$ for some $q > n$. Then, for all $p \in (1,\infty)$, we have that*

$$\varepsilon \, \|\theta_\varepsilon\|_{W^{1,p}(\Omega)} = \mathcal{O}(\varepsilon^{\frac{1}{p}}). \tag{2.16}$$

Proof. Firstly note that as $u_0 \in W^{4,q}(\Omega)$ for some $q > n$ we have $u_0 \in W^{3,\infty}(\Omega)$. We further note that, as $A \in C^{0,\alpha}(\mathbb{R}^n; \mathbb{R}^{n\times n})$ for some $\alpha \in (0,1]$, we have $V \in C^{2,\alpha}(\mathbb{R}^n; \mathbb{R}^{n\times n})$ by elliptic regularity theory [53].

We let $\eta \in C_c^\infty(\mathbb{R}^n)$ be a cut-off function with the properties $0 \leq \eta \leq 1$,

$$\begin{aligned} \eta \equiv 1 \quad &\text{in} \quad \left\{x \in \Omega : \operatorname{dist}(x, \partial\Omega) < \frac{\varepsilon}{2}\right\}, \\ \eta \equiv 0 \quad &\text{in} \quad \left\{x \in \Omega : \operatorname{dist}(x, \partial\Omega) \geq \varepsilon\right\}, \end{aligned}$$

and $|\nabla \eta| = \mathcal{O}(\varepsilon^{-1})$. Note that this implies that

$$\|\eta\|_{L^p(\Omega)} + \varepsilon \|\nabla \eta\|_{L^p(\Omega)} = \mathcal{O}(\varepsilon^{\frac{1}{p}}) \tag{2.17}$$

for any $p \in (1, \infty)$. We then define the function

$$\tilde{\theta}_\varepsilon := \theta_\varepsilon + \eta\, V\left(\frac{\cdot}{\varepsilon}\right) : D^2 u_0$$

and note that it is the solution to the problem

$$\begin{cases} \nabla \cdot \left(A^{\text{div}}\left(\frac{\cdot}{\varepsilon}\right) \nabla \tilde{\theta}_\varepsilon\right) = \nabla \cdot F_1^\varepsilon & \text{in } \Omega, \\ \tilde{\theta}_\varepsilon = 0 & \text{on } \partial\Omega, \end{cases}$$

with F_1^ε given by

$$F_1^\varepsilon := A^{\text{div}}\left(\frac{\cdot}{\varepsilon}\right) \nabla \left[\eta\, V\left(\frac{\cdot}{\varepsilon}\right) : D^2 u_0\right].$$

Using the uniform $W^{1,p}$ estimate from Lemma 2.3.3, we find that for $\varepsilon \in (0, 1]$ and any $p \in (1, \infty)$, we have

$$\|\tilde{\theta}_\varepsilon\|_{W^{1,p}(\Omega)} \leq C \|F_1^\varepsilon\|_{L^p(\Omega)} \leq C \left\|\eta\, V\left(\frac{\cdot}{\varepsilon}\right) : D^2 u_0\right\|_{W^{1,p}(\Omega)}.$$

Therefore, by the triangle inequality, we obtain the estimate

$$\|\theta_\varepsilon\|_{W^{1,p}(\Omega)} \leq C \left\|\eta\, V\left(\frac{\cdot}{\varepsilon}\right) : D^2 u_0\right\|_{W^{1,p}(\Omega)}. \tag{2.18}$$

As we have the bound

$$\left\|V\left(\frac{\cdot}{\varepsilon}\right) : D^2 u_0\right\|_{L^\infty(\Omega)} + \varepsilon \left\|\nabla \left[V\left(\frac{\cdot}{\varepsilon}\right) : D^2 u_0\right]\right\|_{L^\infty(\Omega)} = \mathcal{O}(1),$$

and the asymptotic behavior of the cut-off (2.17), we deduce from (2.18) that there holds

$$\varepsilon \|\theta_\varepsilon\|_{W^{1,p}(\Omega)} \leq C\varepsilon \left(\|\nabla \eta\|_{L^p(\Omega)} + \left(\varepsilon^{-1} + 1\right) \|\eta\|_{L^p(\Omega)}\right) = \mathcal{O}(\varepsilon^{\frac{1}{p}}),$$

which is precisely the claimed bound (2.16). $\square$

Let us remark that the estimate (2.16) for $p = 2$ has already been shown in the context of divergence-form homogenization; see [9, 81]. It is worth noting here that we only obtain $W^{1,p}$ and $W^{2,p}$ bounds for the boundary corrector θ_ε, and we do not study qualitative and quantitative homogenization of (2.12) as in, e.g., [7, 14, 42, 52].

Combining the results of Lemma 2.3.2 and Lemma 2.3.4, we obtain a convergence rate for the convergence of u_ε to the homogenized solution u_0 in the $W^{1,p}$-norm.

Theorem 2.3.3 ([90, Theorem 1.5(i)] $W^{1,p}$ convergence rate $\mathcal{O}(\varepsilon)$)**.** *Assume that $(\Omega, A, f) \in \mathcal{G}^{2,q}$ and $u_0 \in W^{4,q}(\Omega)$ for some $q > n$. Then, for any $p \in (1, \infty)$, we have*

$$\|u_\varepsilon - u_0\|_{W^{1,p}(\Omega)} = \mathcal{O}(\varepsilon).$$

Proof. The result follows from Lemma 2.3.2 and Lemma 2.3.4. □

Convergence rate in H^1 for nonsmooth domains

We briefly discuss an extension of this result for $p = 2$ to the case of nonsmooth domains. To this end, let us consider a triple $(\Omega, A, f) \in \mathcal{H}^2$ and assume that $u_0 \in H^4(\Omega) \cap W^{2,\infty}(\Omega)$. Recall from Theorem 2.3.1 that

$$\left\| u_\varepsilon - u_0 - \varepsilon^2 \left(V\left(\frac{\cdot}{\varepsilon}\right) : D^2 u_0 + \theta_\varepsilon \right) \right\|_{H^2(\Omega)} = \mathcal{O}(\varepsilon).$$

Noting that the result of Lemma 2.3.3 for $p = 2$ still holds in this case by standard arguments, we find that

$$\varepsilon \|\theta_\varepsilon\|_{H^1(\Omega)} \leq C\varepsilon \left\| \eta\, V\left(\frac{\cdot}{\varepsilon}\right) : D^2 u_0 \right\|_{H^1(\Omega)} = \mathcal{O}(\sqrt{\varepsilon})$$

analogously to the proof of Lemma 2.3.4 and we obtain the following result:

Theorem 2.3.4 (H^1 rate for convex domains)**.** *Assume that $(\Omega, A, f) \in \mathcal{H}^2$ and $u_0 \in H^4(\Omega) \cap W^{2,\infty}(\Omega)$. Then, we have that*

$$\|u_\varepsilon - u_0\|_{H^1(\Omega)} = \mathcal{O}(\varepsilon).$$

Remark 2.3.1. *The result also holds for triples $(\Omega, A, f) \in \mathcal{G}^{2,2}$ which are chosen such that $u_0 \in H^4(\Omega) \cap W^{2,\infty}(\Omega)$.*

2.4 Optimal rates of convergence

In this section we would like to derive the optimal convergence rate for the convergence of the solution u_ε to (2.1) to the solution u_0 of the homogenized problem (2.9) in the $W^{1,p}$-norm for $p \in (1, \infty)$. The result is for the case of $C^{2,\gamma}$ domains and we discuss extensions to nonsmooth domains in the end.

2.4.1 Higher order corrector estimate in the $C^{1,\nu}$-norm

As a first step we would like to derive a $\mathcal{O}(\varepsilon^2)$ corrector estimate in the $C^{1,\nu}$-norm. To this end, we need to introduce some auxiliary functions:

Let us introduce the function z_ε to be the solution to the problem

$$\begin{cases} A\left(\frac{\cdot}{\varepsilon}\right) : D^2 z_\varepsilon = g & \text{in } \Omega, \\ z_\varepsilon = 0 & \text{on } \partial\Omega, \end{cases}$$

where the right-hand side is defined to be

$$g := \sum_{j,k,l=1}^{n} c_j^{kl} \partial^3_{jkl} u_0 \tag{2.19}$$

with the values $c_j^{kl} \in \mathbb{R}$, $j,k,l \in \{1,\dots,n\}$, given by

$$c_j^{kl} = c_j^{kl}(A) := \int_Y A e_j \cdot \nabla \chi_{kl} \, m. \tag{2.20}$$

We further introduce the functions ξ_{jkl}, $j,k,l \in \{1,\dots,n\}$ to be the solutions to the periodic problems

$$A : D^2 \xi_{jkl} = c_j^{kl} - A e_j \cdot \nabla \chi_{kl} \quad \text{in } Y, \quad \xi_{jkl} \text{ is } Y\text{-periodic}, \quad \int_Y \xi_{jkl} = 0.$$

Note that the functions ξ_{jkl} are well-defined as by definition (2.20) of c_j^{kl}, the right-hand side integrated against the invariant measure equals zero, i.e., there holds

$$\int_Y \left(c_j^{kl} - A e_j \cdot \nabla \chi_{kl}\right) m = 0.$$

We also introduce a corresponding boundary corrector θ_ξ^ε to be the solution to the following problem:

$$\begin{cases} A\left(\frac{\cdot}{\varepsilon}\right) : D^2 \theta_\xi^\varepsilon = 0 & \text{in } \Omega, \\ \theta_\xi^\varepsilon = -\sum_{j,k,l=1}^{n} \xi_{jkl}\left(\frac{\cdot}{\varepsilon}\right) \partial^3_{jkl} u_0 & \text{on } \partial\Omega. \end{cases} \tag{2.21}$$

Similarly to the definition of ϕ_ε from (2.13), we then define the function

$$\psi_\varepsilon := \varepsilon^2 \left[\sum_{j,k,l=1}^{n} \xi_{jkl}\left(\frac{\cdot}{\varepsilon}\right) \partial^3_{jkl} u_0 + \theta_\xi^\varepsilon \right]. \tag{2.22}$$

We then have the following higher order corrector estimate:

Theorem 2.4.1 (Higher order $C^{1,\nu}$ corrector estimate)**.** *Assume that* $(\Omega, A, f) \in \mathcal{G}^{3,q}$ *and* $u_0 \in W^{5,q}(\Omega)$ *for some* $q > n$. *Then, there exists a* $\nu \in (0,1]$ *such that for all* $p \in (1,\infty)$, *we have that*

$$\|u_\varepsilon - u_0 + 2\varepsilon z_\varepsilon - \phi_\varepsilon - 2\varepsilon\psi_\varepsilon\|_{C^{1,\nu}(\Omega)} = \mathcal{O}(\varepsilon^2).$$

Proof. Let w_ε be the solution to the problem

$$\begin{cases} A\left(\frac{\cdot}{\varepsilon}\right) : D^2 w_\varepsilon = -\displaystyle\sum_{i,j,k,l=1}^{n} a_{ij}\left(\frac{\cdot}{\varepsilon}\right) \partial_i \chi_{kl}\left(\frac{\cdot}{\varepsilon}\right) \partial^3_{jkl} u_0 & \text{in } \Omega, \\ w_\varepsilon = 0 & \text{on } \partial\Omega. \end{cases}$$

Then we have that the function $u_\varepsilon - u_0 - \phi_\varepsilon - 2\varepsilon w_\varepsilon$ satisfies the problem

$$\begin{cases} A\left(\frac{\cdot}{\varepsilon}\right) : D^2(u_\varepsilon - u_0 - \phi_\varepsilon - 2\varepsilon w_\varepsilon) = -\varepsilon^2 F_2^\varepsilon & \text{in } \Omega, \\ u_\varepsilon - u_0 - \phi_\varepsilon - 2\varepsilon w_\varepsilon = 0 & \text{on } \partial\Omega, \end{cases}$$

with F_2^ε given by

$$F_2^\varepsilon := \sum_{i,j,k,l=1}^{n} a_{ij}\left(\frac{\cdot}{\varepsilon}\right) \chi_{kl}\left(\frac{\cdot}{\varepsilon}\right) \partial^4_{ijkl} u_0.$$

As $u_0 \in W^{5,q}(\Omega)$ for $q > n$, we have that F_2^ε is uniformly bounded in $L^q(\Omega)$ and hence, by the uniform estimate from Lemma 2.3.1, we find

$$\|u_\varepsilon - u_0 - \phi_\varepsilon - 2\varepsilon w_\varepsilon\|_{C^{1,\nu}(\Omega)} \leq C\varepsilon^2 \|F_2^\varepsilon\|_{L^q(\Omega)} = \mathcal{O}(\varepsilon^2). \tag{2.23}$$

We also have that the function $w_\varepsilon + z_\varepsilon - \psi_\varepsilon$ satisfies the problem

$$\begin{cases} A\left(\frac{\cdot}{\varepsilon}\right) : D^2(w_\varepsilon + z_\varepsilon - \psi_\varepsilon) = -\varepsilon F_3^\varepsilon & \text{in } \Omega, \\ w_\varepsilon + z_\varepsilon - \psi_\varepsilon = 0 & \text{on } \partial\Omega, \end{cases}$$

with F_3^ε given by

$$F_3^\varepsilon := \sum_{d,i,j,k,l=1}^{n} a_{ij}\left(\frac{\cdot}{\varepsilon}\right) \left[2\,\partial_i \xi_{dkl}\left(\frac{\cdot}{\varepsilon}\right) \partial^4_{djkl} u_0 + \varepsilon \xi_{dkl}\left(\frac{\cdot}{\varepsilon}\right) \partial^5_{dijkl} u_0\right].$$

As $u_0 \in W^{5,q}(\Omega)$ for some $q > n$, we have that F_3^ε is uniformly bounded in $L^q(\Omega)$ and hence, by the uniform estimate from Lemma 2.3.1, we find

$$\|w_\varepsilon + z_\varepsilon - \psi_\varepsilon\|_{C^{1,\nu}(\Omega)} \leq C\varepsilon \|F_3^\varepsilon\|_{L^q(\Omega)} = \mathcal{O}(\varepsilon). \tag{2.24}$$

Combining the bounds (2.23) and (2.24), we obtain

$$\|u_\varepsilon - u_0 + 2\varepsilon z_\varepsilon - \phi_\varepsilon - 2\varepsilon\psi_\varepsilon\|_{C^{1,\nu}(\Omega)} = \mathcal{O}(\varepsilon^2),$$

which is the claimed result. ◻

2.4.2 Optimal convergence rate in the $W^{1,p}$-norm

Before we state the main result on the optimal $W^{1,p}$ convergence rate, let us introduce the function z as the solution to the problem

$$\begin{cases} A^0 : D^2 z = g & \text{in } \Omega, \\ z = 0 & \text{on } \partial\Omega, \end{cases} \tag{2.25}$$

where g is defined in (2.19). Observe that the function z is precisely the homogenized solution corresponding to $(z_\varepsilon)_{\varepsilon>0}$. We note that, assuming $u_0 \in W^{5,q}(\Omega)$ for some $q > n$, we have $g \in W^{2,q}(\Omega)$ and therefore, we can apply Theorem 2.3.3 to find that for any $p \in (1,\infty)$, there holds

$$\|z_\varepsilon - z\|_{W^{1,p}(\Omega)} = \mathcal{O}(\varepsilon). \tag{2.26}$$

Let us also note that we can transform the problem (2.21) into the divergence-form problem

$$\begin{cases} \nabla \cdot \left(A^{\text{div}} \left(\frac{\cdot}{\varepsilon}\right) \nabla \theta_\xi^\varepsilon \right) = 0 & \text{in } \Omega, \\ \theta_\xi^\varepsilon = - \displaystyle\sum_{j,k,l=1}^n \xi_{jkl} \left(\frac{\cdot}{\varepsilon}\right) \partial^3_{jkl} u_0 & \text{on } \partial\Omega, \end{cases}$$

with a coefficient $A^{\text{div}} \in C^{0,\alpha}(\mathbb{T}^n; \mathbb{R}^{n\times n})$ for some $\alpha \in (0,1]$ that is uniformly elliptic. Let us note that since $u_0 \in W^{5,q}(\Omega)$ for some $q > n$ and $\xi_{jkl} \in C^{2,\beta}(\mathbb{R}^n)$ for some $\beta \in (0,1]$ by elliptic regularity theory [53], we can apply Lemma 2.3.3 to find the bound

$$\varepsilon \|\theta_\xi^\varepsilon\|_{W^{1,p}(\Omega)} \le C\varepsilon \sum_{j,k,l=1}^n \left\| \xi_{jkl} \left(\frac{\cdot}{\varepsilon}\right) \partial^3_{jkl} u_0 \right\|_{W^{1,\infty}(\Omega)} = \mathcal{O}(1) \tag{2.27}$$

for any $p \in (1,\infty)$.

We then have the following theorem on the optimal rate for the convergence of u_ε to the homogenized solution u_0 in the $W^{1,p}(\Omega)$-norm:

Theorem 2.4.2 ([90, Theorem 1.5(ii)] $W^{1,p}$ estimate and optimal rate)**.** *Assume that $(\Omega, A, f) \in \mathcal{G}^{3,q}$ and $u_0 \in W^{5,q}(\Omega)$ for some $q > n$. Then, for all $p \in (1,\infty)$, we have that*

$$\left\| u_\varepsilon - u_0 + 2\varepsilon z - \varepsilon^2 V\left(\frac{\cdot}{\varepsilon}\right) : D^2 u_0 \right\|_{W^{1,p}(\Omega)} = \mathcal{O}(\varepsilon^{1+\frac{1}{p}}).$$

In particular, for all $p \in (1,\infty)$, we have

$$\|u_\varepsilon - u_0\|_{W^{1,p}(\Omega)} = \mathcal{O}(\varepsilon),$$

and this rate of convergence $\mathcal{O}(\varepsilon)$ is optimal in general.

Proof. From Theorem 2.4.1, using the definitions of ϕ_ε and ψ_ε from (2.13) and (2.22), we have that

$$\left\| u_\varepsilon - u_0 + 2\varepsilon z_\varepsilon - \varepsilon^2 V\left(\frac{\cdot}{\varepsilon}\right) : D^2 u_0 - \varepsilon^2 \theta_\varepsilon - 2\varepsilon^3 \theta_\xi^\varepsilon \right\|_{W^{1,\infty}(\Omega)} = \mathcal{O}(\varepsilon^2). \qquad (2.28)$$

Finally, using the rate of convergence of z_ε to z given by (2.26), and Lemma 2.3.4 and the estimate (2.27) to bound the boundary correctors, we conclude that

$$\left\| u_\varepsilon - u_0 + 2\varepsilon z - \varepsilon^2 V\left(\frac{\cdot}{\varepsilon}\right) : D^2 u_0 \right\|_{W^{1,p}(\Omega)} = \mathcal{O}(\varepsilon^{1+\frac{1}{p}})$$

for any $p \in (1, \infty)$. □

Nonsmooth domains

We would like to briefly discuss the case of nonsmooth convex domains. To this end, let us consider $(\Omega, A, f) \in \mathcal{H}^0$ and assume that the homogenized solution is of regularity $u_0 \in W^{5,q}(\Omega)$ for some $q > n$. By the uniform H^2 estimate from Theorem 2.2.1 and the Sobolev embedding, we have the uniform $W^{1,p}$ estimate

$$\|u_\varepsilon\|_{W^{1,p}(\Omega)} \leq C\|u_\varepsilon\|_{H^2(\Omega)} \leq C\|f\|_{L^2(\Omega)}$$

for any $p < 2^*$ with constants independent of ε. Here, we write $2^* := \frac{2n}{n-2}$ to denote the critical Sobolev exponent (with the convention that $2^* := \infty$ if $n = 2$). This uniform estimate replaces the need for the uniform $C^{1,\nu}$ estimate from Lemma 2.3.1.

Finally, in order to estimate the boundary corrector, we have previously transformed the problem into divergence-form and used that for problems of the form

$$\begin{cases} -\nabla \cdot \left(A^{\text{div}}\left(\frac{\cdot}{\varepsilon}\right) \nabla \rho_\varepsilon\right) = -\nabla \cdot F & \text{in } \Omega, \\ \rho_\varepsilon = G & \text{on } \partial\Omega, \end{cases}$$

we have (Lemma 2.3.3) the uniform $W^{1,p}$ estimate

$$\|\rho_\varepsilon\|_{W^{1,p}(\Omega)} \leq C\left(\|F\|_{L^p(\Omega)} + \|G\|_{W^{1,p}(\Omega)}\right) \qquad (2.29)$$

with a constant $C > 0$ independent of ε, assuming that $A^{\text{div}} \in C^{0,\alpha}(\mathbb{T}^n; \mathbb{R}^{n\times n})$ for some $\alpha \in (0, 1]$ is uniformly elliptic and that Ω is sufficiently smooth.

Now as Ω is merely assumed to be convex, we still have (2.29) for $p = 2$ by standard arguments and hence, we find that the result of Theorem 2.4.2 remains true for $p = 2$ under the assumptions made in this section. Uniform $W^{1,p}$ estimates for divergence-form problems for a wider range of values p require a more sophisticated approach. With a symmetry assumption on A^{div}, uniform $W^{1,p}$ estimates for divergence-form problems on Lipschitz domains (recall that bounded convex domains are Lipschitz [56]) have been obtained in [86] for values of p in a certain range around $p = 2$.

2.4.3 Optimal convergence rate in the L^∞-norm

Let us note that an inspection of the proof of Theorem 2.4.2, see (2.28) and (2.26), yields that

$$\|u_\varepsilon - u_0 + 2\varepsilon z\|_{L^\infty(\Omega)} = \mathcal{O}(\varepsilon^2),$$

i.e., the optimal rate for the convergence of u_ε to u_0 in the L^∞-norm is $\mathcal{O}(\varepsilon)$ generically. This result has first been obtained in [57] under stronger assumptions:

Theorem 2.4.3 (Theorem 1.2 in [57]). *Assume that $A \in C^2(\mathbb{T}^n; \mathcal{S}^n_+)$ and $f \in C^3(\bar{\Omega})$. Let u_ε, u_0 and z be the solutions to* (2.1), (2.9) *and* (2.25) *respectively. Then we have*

$$\|u_\varepsilon - u_0 + 2\varepsilon z\|_{L^\infty(\Omega)} = \mathcal{O}(\varepsilon^2).$$

In particular, with g given by (2.19), *the following assertions hold:*

(i) If $g \equiv 0$, then $\|u_\varepsilon - u_0\|_{L^\infty(\Omega)} = \mathcal{O}(\varepsilon^2)$ and this rate of convergence is optimal.

(ii) If $g \not\equiv 0$, then $\|u_\varepsilon - u_0\|_{L^\infty(\Omega)} = \mathcal{O}(\varepsilon)$ and this rate of convergence is optimal.

Remark 2.4.1. *There is a typo in Theorem 1.2 in [57], which uses the opposite sign for the $\mathcal{O}(\varepsilon)$-term.*

As a consequence of Theorem 2.4.3, we can classify coefficients $A \in C^2(\mathbb{T}^n; \mathcal{S}^n_+)$ into those that give optimal rate of convergence $\mathcal{O}(\varepsilon^2)$, called the c-good coefficients, and those that give optimal rate of convergence $\mathcal{O}(\varepsilon)$, called the c-bad coefficients.

Corollary 2.4.1 (c-good and c-bad matrices). *Let $A \in C^2(\mathbb{T}^n; \mathcal{S}^n_+)$. Then, with $\{c_j^{kl}\}_{1\le j,k,l\le n}$ given by* (2.20), *the following assertions hold:*

(i) If $c_j^{kl}(A) = 0$ for all $j,k,l \in \{1,\dots,n\}$, then the situation (i) of Theorem 2.4.3 occurs for any choice of f. We then say that A is c-good.

(ii) If $c_j^{kl}(A) \neq 0$ for some $j,k,l \in \{1,\dots,n\}$, then there exists an f such that the situation (ii) of Theorem 2.4.3 occurs. We then say that A is c-bad.

It has further been shown in [57, Theorem 1.4] that the set of c-bad matrices is open and dense in $C^2(\mathbb{T}^n; \mathcal{S}^n_+)$ for dimensions $n \geq 2$. Therefore, we have generically that the optimal rate is $\mathcal{O}(\varepsilon)$ in $L^\infty(\Omega)$.

An explicit c-bad matrix

For the numerical illustrations we use an explicit c-bad matrix (recall Corollary 2.4.1 for the definition of c-bad) and consider a homogenization problem of the form (2.1) with $z \not\equiv 0$. This is the first direct proof of the existence of a c-bad matrix.

Theorem 2.4.4 ([90, Theorem 1.10] Explicit c-bad matrix)**.** *The matrix-valued function $A : \mathbb{R}^2 \to \mathbb{R}^{2\times 2}$ given by*

$$A(y_1, y_2) := \frac{1}{m(y_1,y_2)} \begin{pmatrix} 1 - \frac{1}{2}\sin(2\pi y_1)\sin(2\pi y_2) & 0 \\ 0 & 1 + \frac{1}{2}\sin(2\pi y_1)\sin(2\pi y_2) \end{pmatrix} \tag{2.30}$$

with $m : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$m(y_1, y_2) := 1 + \frac{1}{4}(\cos(2\pi y_1) - 2\sin(2\pi y_1))\sin(2\pi y_2) \tag{2.31}$$

is c-bad. More precisely, there holds $c_1^{11} = c_1^{22} = -\frac{1}{128\pi}$ and $c_j^{kl} = 0$ otherwise.

Before we prove the theorem, we observe the following:

Remark 2.4.2. *The function $m : \mathbb{R}^2 \to \mathbb{R}$ given by* (2.31) *is the invariant measure of $A : \mathbb{R}^2 \to \mathbb{R}^{2\times 2}$ given by* (2.30)*. Further note that the problem* (2.1) *can then be transformed into the divergence-form problem*

$$\begin{cases} \nabla \cdot \left(A^{\mathrm{div}}\left(\frac{\cdot}{\varepsilon}\right)\nabla u_\varepsilon\right) = m\left(\frac{\cdot}{\varepsilon}\right) f & \text{in } \Omega, \\ u_\varepsilon = 0 & \text{on } \partial\Omega, \end{cases} \tag{2.32}$$

with the matrix-valued function $A^{\mathrm{div}} : \mathbb{R}^n \to \mathbb{R}^{n\times n}$ given by

$$A^{\mathrm{div}}(y) := \begin{pmatrix} 1 - \frac{1}{2}\sin(2\pi y_1)\sin(2\pi y_2) & \frac{1}{2}\cos(2\pi y_1)\cos(2\pi y_2) \\ -\frac{1}{2}\cos(2\pi y_1)\cos(2\pi y_2) & 1 + \frac{1}{2}\sin(2\pi y_1)\sin(2\pi y_2) \end{pmatrix}.$$

We can check that A is c-bad by explicitly computing the matrix of corrector functions $V = (\chi_{ij})_{1\le i,j\le 2}$ given by (2.11) and computing the values $\{c_j^{kl}\}_{1\le j,k,l\le 2}$ given by (2.20).

Proof of Theorem 2.4.4. The effective coefficient $A^0 \in \mathcal{S}_+^2$ is given by

$$A^0 = \int_Y Am = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{2.33}$$

and it is a straightforward calculation to check that the matrix of corrector functions $V = (\chi_{ij})_{1\le i,j\le 2} : \mathbb{R}^2 \to \mathbb{R}^{2\times 2}$ is given by

$$V(y) = -\frac{\sin(2\pi y_2)}{32\pi^2} \begin{pmatrix} \cos(2\pi y_1) & 0 \\ 0 & \cos(2\pi y_1) - 4\sin(2\pi y_1) \end{pmatrix}.$$

Computation of the values c_j^{kl} for $j,k,l \in \{1,2\}$ given by (2.20) yields that

$$c_1^{11} = \int_Y m a_{11} \partial_1 \chi_{11} = -\frac{1}{128\pi} = \int_Y m a_{11} \partial_1 \chi_{22} = c_1^{22}$$

for the values of c_1^{11}, c_1^{22}, and that

$$c_2^{11} = \int_Y m a_{22} \partial_2 \chi_{11} = 0 = \int_Y m a_{22} \partial_2 \chi_{22} = c_2^{22}$$

for the values of c_2^{11}, c_2^{22}. Clearly we have that $c_j^{kl} = 0$ for any $(j,k,l) \in \{1,2\}^3$ with $k \neq l$. □

Let us note that the effective coefficient (2.33) is the identity matrix and hence, the homogenized problem for this c-bad matrix is the Poisson problem

$$\begin{cases} \Delta u_0 = f & \text{in } \Omega, \\ u_0 = 0 & \text{on } \partial\Omega. \end{cases} \tag{2.34}$$

Further, we have that the function z defined by (2.25) is given as the solution to the Poisson problem

$$\begin{cases} \Delta z = \dfrac{\partial_1 f}{128\pi} & \text{in } \Omega, \\ z = 0 & \text{on } \partial\Omega. \end{cases} \tag{2.35}$$

An explicit c-good matrix

Finally, let us note that the factor $\frac{1}{m}$ in the definition of the c-bad matrix (2.30) is crucial for c-badness. Indeed, removing this factor we obtain a c-good matrix:

Remark 2.4.3. *The matrix-valued function $A : \mathbb{R}^2 \to \mathbb{R}^{2\times 2}$ given by*

$$A(y) := \begin{pmatrix} 1 - \frac{1}{2}\sin(2\pi y_1)\sin(2\pi y_2) & 0 \\ 0 & 1 + \frac{1}{2}\sin(2\pi y_1)\sin(2\pi y_2) \end{pmatrix}$$

is c-good.

Proof. The invariant measure is the constant function $m \equiv 1$ and hence, the effective coefficient $A^0 \in \mathcal{S}_+^2$ is given by

$$A^0 = \int_Y A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

It is a straightforward calculation to check that the matrix of corrector functions $V = (\chi_{ij})_{1\le i,j\le 2} : \mathbb{R}^2 \to \mathbb{R}^{2\times 2}$ is given by

$$V(y) = -\frac{\sin(2\pi y_1)\sin(2\pi y_2)}{16\pi^2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Computation of the values c_j^{kl} given by (2.20) yields $c_j^{kl} = 0$ for all $j,k,l \in \{1,2\}$. □

Note that the effective problem for this c-good matrix is again the Poisson problem (2.34), i.e., the homogenized solution coincides with the one from the c-bad problem.

Generalization

Let us note that the aforementioned c-bad matrix from Theorem 2.4.4 and the c-good matrix from Remark 2.4.3 fit into a more general framework outlined below.

Let $a : \mathbb{R}^2 \to (-1,1)$ and $m : \mathbb{R}^2 \to (0,\infty)$ be Y-periodic functions of the form

$$\begin{aligned} a(y_1,y_2) &:= f_1(y_1+y_2) - f_1(y_1-y_2), \\ m(y_1,y_2) &:= 1 + f_2(y_1+y_2) - f_2(y_1-y_2) \end{aligned}$$

for some 1-periodic functions $f_1, f_2 \in C^2(\mathbb{R})$ with $\int_0^1 f_1(s)\mathrm{d}s = \int_0^1 f_2(s)\mathrm{d}s = 0$. We then consider the matrix-valued function $A : \mathbb{R}^2 \to \mathbb{R}^{2\times 2}$ given by

$$A = \frac{1}{m}\begin{pmatrix} 1-a & 0 \\ 0 & 1+a \end{pmatrix}.$$

Observe that the function m is the invariant measure corresponding to A and hence, noting that $\int_Y a = 0$, the effective coefficient is the identity matrix

$$A^0 = \int_Y Am = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

It is straightforward to check that the matrix of corrector functions $V = (\chi_{ij})_{1\le i,j\le 2} : \mathbb{R}^2 \to \mathbb{R}^{2\times 2}$ is given by

$$V(y_1,y_2) = \begin{pmatrix} \frac{F_+(y_1+y_2)-F_+(y_1-y_2)}{2} & 0 \\ 0 & \frac{F_-(y_1+y_2)-F_-(y_1-y_2)}{2} \end{pmatrix}, \qquad F_\pm := F_2 \pm F_1,$$

where $F_i \in C^2(\mathbb{R})$, $i \in \{1,2\}$, denotes the unique solution to

$$F_i'' = f_i \quad \text{in } (0,1), \quad F_i \text{ is 1-periodic}, \quad \int_0^1 F_i(s)\,\mathrm{d}s = 0.$$

Finally, we leave it to the reader to show that

$$\begin{aligned} c_1^{11} &= \int_Y (1-a)\partial_1\chi_{11} = \int_0^1\int_0^1 a(y_1,y_2)F_2'(y_1-y_2)\,\mathrm{d}y_1\,\mathrm{d}y_2 = \int_Y (1-a)\partial_1\chi_{22} = c_1^{22}, \\ c_2^{11} &= \int_Y (1+a)\partial_2\chi_{11} = 0 = \int_Y (1+a)\partial_2\chi_{22} = c_2^{22}, \end{aligned}$$

and $c_j^{kl} = 0$ for any $(j,k,l) \in \{1,2\}^3$ with $k \ne l$. Therefore, A is c-bad if and only if

$$Q := \int_0^1\int_0^1 a(y_1,y_2)F_2'(y_1-y_2)\,\mathrm{d}y_1\,\mathrm{d}y_2 \ne 0.$$

Remark 2.4.4. *(i) The c-bad matrix from Theorem 2.4.4 fits into the above framework with* $f_1, f_2 : \mathbb{R} \to \mathbb{R}$ *given by*

$$f_1(s) := -\frac{1}{4}\cos(2\pi s), \quad f_2(s) := \frac{1}{8}\left(\sin(2\pi s) + 2\cos(2\pi s)\right)$$

for $s \in \mathbb{R}$*. Indeed, with these choices we have that*

$$\begin{aligned} a(y) &:= \frac{1}{2}\sin(2\pi y_1)\sin(2\pi y_2) &&= f_1(y_1+y_2) - f_1(y_1-y_2), \\ m(y) - 1 &:= \frac{1}{4}\left(\cos(2\pi y_1) - 2\sin(2\pi y_1)\right)\sin(2\pi y_2) &&= f_2(y_1+y_2) - f_2(y_1-y_2) \end{aligned}$$

for any $y = (y_1, y_2) \in \mathbb{R}^2$*. Noting that the function* $F_2 : \mathbb{R} \to \mathbb{R}$ *is given by* $F_2(s) := -\frac{1}{32\pi^2}\left(\sin(2\pi s) + 2\cos(2\pi s)\right)$ *for* $s \in \mathbb{R}$*, we find that*

$$Q = \int_0^1 \int_0^1 a(y_1, y_2) F_2'(y_1 - y_2)\, \mathrm{d}y_1\, \mathrm{d}y_2 = -\frac{1}{128\pi} \neq 0,$$

and recover the result from Theorem 2.4.4.

(ii) The c-good matrix from Remark 2.4.3 fits into the above framework with $f_2 \equiv 0$ *(observe* $m \equiv 1$*) and* $f_1 : \mathbb{R} \to \mathbb{R}$ *as in (i). Therefore, noting that* $f_2 \equiv 0$ *implies* $F_2 \equiv 0$*, we find that* $Q = 0$ *and recover the result from Remark 2.4.3.*

2.5 Numerical illustrations

Finally, we demonstrate through numerical experiments that the obtained rates in the previously stated results cannot be improved in general. We illustrate the optimality of the obtained results.

Remark 2.5.1 (L^∞ estimate, gradient estimate and Hessian estimate)**.** *In the situation of Theorem 2.4.2, the following assertions hold.*

(i) L^∞ *bound: We have*

$$\|u_\varepsilon - u_0 + 2\varepsilon z\|_{L^\infty(\Omega)} = \mathcal{O}(\varepsilon^2). \tag{2.36}$$

(ii) Gradient bound: For all $p \in (1, \infty)$*, we have*

$$\left\| \nabla u_\varepsilon - \nabla u_0 + 2\varepsilon \nabla z - \varepsilon \sum_{i,j=1}^n \nabla \chi_{ij}\left(\frac{\cdot}{\varepsilon}\right) \partial^2_{ij} u_0 \right\|_{L^p(\Omega)} = \mathcal{O}(\varepsilon^{1+\frac{1}{p}}). \tag{2.37}$$

(iii) Hessian bound: In view of Theorem 2.3.2, for all $p \in (1,\infty)$, there holds

$$\left\| D^2 u_\varepsilon - D^2 u_0 - \sum_{i,j=1}^{n} D^2\chi_{ij}\left(\frac{\cdot}{\varepsilon}\right)\partial_{ij}^2 u_0 \right\|_{L^p(\Omega)} = \mathcal{O}(\varepsilon^{\frac{1}{p}}). \tag{2.38}$$

The rate $\mathcal{O}(\varepsilon^2)$ in the $L^\infty(\Omega)$ estimate (2.36)*, the rate $\mathcal{O}(\varepsilon^{1+\frac{1}{p}})$ in the gradient estimate* (2.37) *and the rate $\mathcal{O}(\varepsilon^{\frac{1}{p}})$ in the Hessian estimate* (2.38) *are optimal in general.*

Let us revisit the bounds in Remark 2.5.1 and note that, for values $p \geq 2$, the gradient bound (2.37) follows from the L^∞ bound (2.36) and the Hessian bound (2.38) via the Gagliardo–Nirenberg interpolation inequality [79] applied to the function

$$\varphi_\varepsilon := u_\varepsilon - u_0 + 2\varepsilon z - \varepsilon^2\, V\left(\frac{\cdot}{\varepsilon}\right) : D^2 u_0.$$

Indeed, assuming that $\|\varphi_\varepsilon\|_{L^\infty(\Omega)} = \mathcal{O}(\varepsilon^2)$ and $\|D^2\varphi_\varepsilon\|_{L^p(\Omega)} = \mathcal{O}(\varepsilon^{\frac{1}{p}})$ for any $p \in (1,\infty)$, an application of the Gagliardo–Nirenberg inequality yields

$$\|\nabla\varphi_\varepsilon\|_{L^p(\Omega)} \leq C\left(\|D^2\varphi_\varepsilon\|_{L^{\frac{p}{2}}(\Omega)}^{\frac{1}{2}} \|\varphi_\varepsilon\|_{L^\infty(\Omega)}^{\frac{1}{2}} + \|\varphi_\varepsilon\|_{L^\infty(\Omega)} \right) = \mathcal{O}(\varepsilon^{1+\frac{1}{p}})$$

for any $p \geq 2$. This shows that the optimality of the bounds (2.36)–(2.38) is natural to expect.

2.5.1 Numerical illustration of the L^∞ rate

We consider the problem (2.1) with the c-bad coefficient matrix A from Theorem 2.4.4, the domain $\Omega := (0,1)^2$ and the right-hand side

$$f : \bar{\Omega} \to \mathbb{R}, \quad f(x_1,x_2) := -8\pi^2 \sin(2\pi x_1)\sin(2\pi x_2).$$

Then, the solution to the homogenized problem (2.34) is given by

$$u_0 : \bar{\Omega} \to \mathbb{R}, \quad u_0(x_1,x_2) = \sin(2\pi x_1)\sin(2\pi x_2),$$

and the solution z to the problem (2.35) is given by

$$z : \bar{\Omega} \to \mathbb{R}, \quad z(x) = \frac{1}{64}\left(\frac{\cosh(2\pi x_1 - \pi)}{\cosh(\pi)} - \cos(2\pi x_1) \right)\sin(2\pi x_2).$$

Figure 2.1 illustrates the estimate (2.36) from Remark 2.5.1, i.e., for several values of ε, we plot

$$E_{0,\infty}^{\varepsilon} := \|u_\varepsilon - u_0 + 2\varepsilon z\|_{L^\infty(\Omega)}. \tag{2.39}$$

We approximate the solution u_ε to (2.1) with $\mathbb{P}_1$ finite elements on a fine mesh, based on the natural variational formulation of the divergence-form problem (2.32). We observe the rate $E_{0,\infty}^{\varepsilon} = \mathcal{O}(\varepsilon^2)$ as ε tends to zero, as expected from Remark 2.5.1.

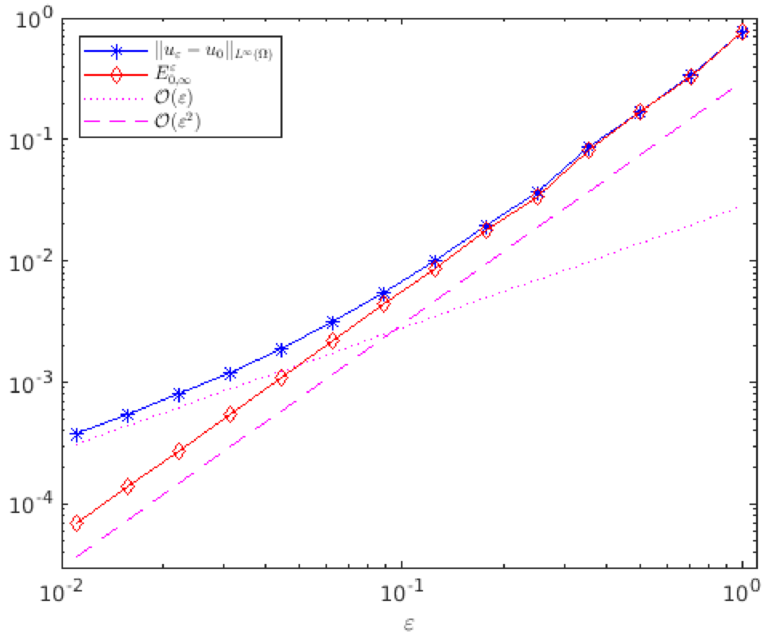


Figure 2.1: *blue*: Plot of $\|u_\varepsilon - u_0\|_{L^\infty(\Omega)}$, *red*: Plot of $E^\varepsilon_{0,\infty}$ (see (2.39)). We observe $\|u_\varepsilon - u_0\|_{L^\infty(\Omega)} = \mathcal{O}(\varepsilon)$ and $E^\varepsilon_{0,\infty} = \mathcal{O}(\varepsilon^2)$ as expected from Remark 2.5.1.

2.5.2 Numerical illustration of the $W^{1,p}$ and $W^{2,p}$ rates

We consider the problem (2.1) with the c-bad coefficient matrix A from Theorem 2.4.4, the domain $\Omega := (0,1)^2$ and the right-hand side

$$f : \bar{\Omega} \to \mathbb{R}, \quad f(x_1, x_2) := -x_1(1-x_1) - x_2(1-x_2). \tag{2.40}$$

Then, the solution of the homogenized problem (2.34) is given by

$$u_0 : \bar{\Omega} \to \mathbb{R}, \quad u_0(x_1, x_2) = \frac{1}{2} x_1(1-x_1)x_2(1-x_2). \tag{2.41}$$

Figure 2.2 (left) illustrates the estimate (2.37) from Remark 2.5.1, i.e., for several values of ε, we plot

$$E^\varepsilon_{1,p} := \left\| \nabla u_\varepsilon - \nabla u_0 + 2\varepsilon \nabla z - \varepsilon \sum_{i,j=1}^{n} \nabla \chi_{ij} \left(\frac{\cdot}{\varepsilon}\right) \partial^2_{ij} u_0 \right\|_{L^p(\Omega)}$$

for the values $p = 2, 3, 4, 5$. We approximate the solution u_ε to (2.1) and the solution z to (2.35) with $\mathbb{P}_2$ finite elements on a fine mesh, based on the natural variational formulation of the divergence-form problems (2.32) and (2.35). We observe the rate $E^\varepsilon_{1,p} = \mathcal{O}(\varepsilon^{1+\frac{1}{p}})$ as ε tends to zero, as expected from Remark 2.5.1.

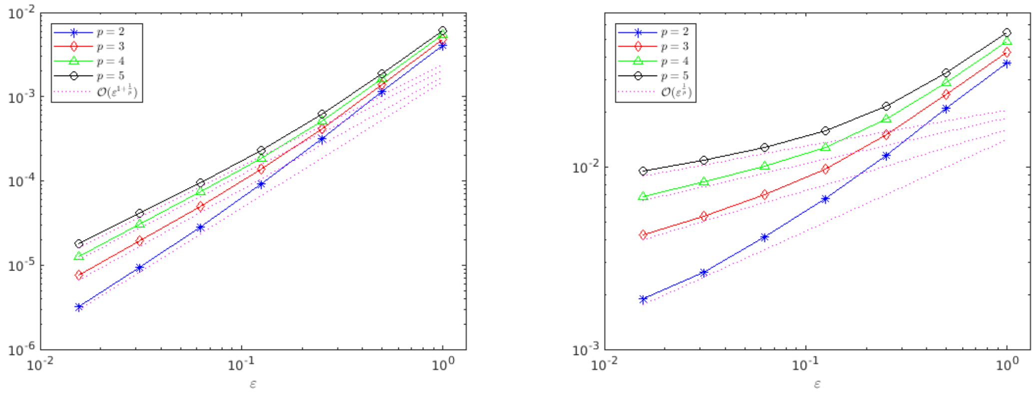


Figure 2.2: Illustration of the optimality of the gradient bound $E^\varepsilon_{1,p} = \mathcal{O}(\varepsilon^{1+\frac{1}{p}})$ (left) and the Hessian bound $E^\varepsilon_{2,p} = \mathcal{O}(\varepsilon^{\frac{1}{p}})$ (right) for $p = 2, 3, 4, 5$.

Figure 2.2 (right) illustrates the estimate (2.38) from Remark 2.5.1, i.e., for several values of ε, we plot

$$E^\varepsilon_{2,p} := \left\| D^2 u_\varepsilon - D^2 u_0 - \sum_{i,j=1}^{n} D^2 \chi_{ij}\left(\frac{\cdot}{\varepsilon}\right) \partial^2_{ij} u_0 \right\|_{L^p(\Omega)}$$

for the values $p = 2, 3, 4, 5$. We approximate the solution u_ε to (2.1) with an H^2 conforming finite element method on a fine mesh, using the HCT element in FreeFem++ [58]. We multiply the equation (2.1) by the invariant measure and use the variational formulation from the framework of linear nondivergence-form equations with coefficients satisfying the Cordes condition (see (2.5)): The solution u_ε to (2.1) is the unique function in $H := H^2(\Omega) \cap H^1_0(\Omega)$ such that there holds

$$\int_\Omega \frac{\mathrm{tr}\left([mA]\left(\frac{\cdot}{\varepsilon}\right)\right)}{\left|[mA]\left(\frac{\cdot}{\varepsilon}\right)\right|^2} \left([mA]\left(\frac{\cdot}{\varepsilon}\right) : D^2 u_\varepsilon\right) \Delta v = \int_\Omega \frac{\mathrm{tr}\left([mA]\left(\frac{\cdot}{\varepsilon}\right)\right)}{\left|[mA]\left(\frac{\cdot}{\varepsilon}\right)\right|^2} m\left(\frac{\cdot}{\varepsilon}\right) f \, \Delta v$$

for any $v \in H$. We observe the rate $E^\varepsilon_{2,p} = \mathcal{O}(\varepsilon^{\frac{1}{p}})$ as ε tends to zero, as expected from Remark 2.5.1.

2.5.3 Comparison of c-bad and c-good problems

We refer to the problem (2.1) with the c-bad coefficient matrix from Theorem 2.4.4 as the c-bad problem and to the problem (2.1) with the c-good coefficient matrix from Remark 2.4.3 as the c-good problem. We perform experiments for these two problems with two different choices of right-hand sides, one with known homogenized solution u_0 and one with unknown homogenized solution u_0. All experiments are performed on the domain $\Omega := (0, 1)^2$.

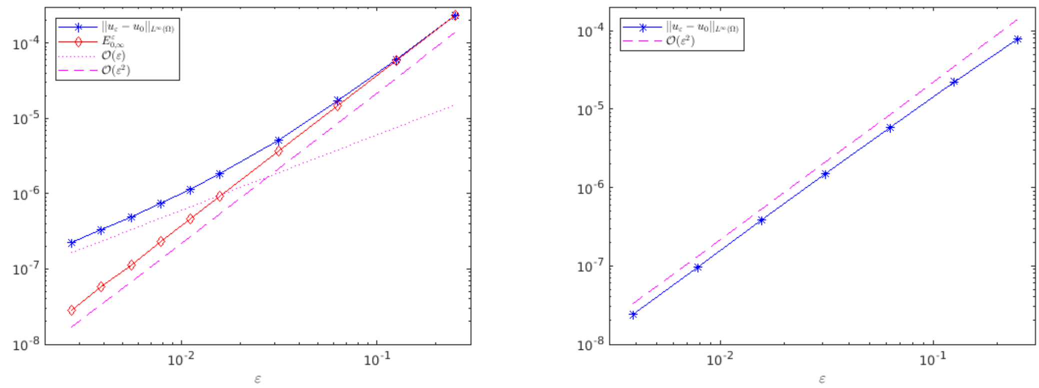


Figure 2.3: Illustration of the L^∞-rates $\|u_\varepsilon - u_0\|_{L^\infty(\Omega)} = \mathcal{O}(\varepsilon)$ and $E^\varepsilon_{0,\infty} = \mathcal{O}(\varepsilon^2)$ for the c-bad problem (left), and $\|u_\varepsilon - u_0\|_{L^\infty(\Omega)} = \mathcal{O}(\varepsilon^2)$ for the c-good problem (right) with the right-hand side (2.40).

Let us recall that the homogenized problems corresponding to the c-bad and the c-good problem coincide and that the homogenized solution u_0 is the solution to the Poisson problem (2.34).

c-bad and c-good problems with known (common) homogenized function

We consider the right-hand side f given by (2.40). Then, the solution u_0 of the homogenized problem is known and given by (2.41).

Figure 2.3 illustrates the L^∞ convergence rate $\mathcal{O}(\varepsilon)$ for the c-bad problem and the convergence rate $\mathcal{O}(\varepsilon^2)$ for the c-good problem. We also illustrate the corrected L^∞ bound $E^\varepsilon_{0,\infty} = \mathcal{O}(\varepsilon^2)$ for the c-bad problem. We approximate the solution u_ε to (2.1) and the solution z to (2.35) with $\mathbb{P}_2$ finite elements on a fine mesh, based on the natural variational formulation of the divergence-form problems (2.32) (note $m \equiv 1$ for the c-good problem) and (2.35).

c-bad and c-good problems with unknown (common) homogenized function

We consider the right-hand side f given by

$$f : \bar{\Omega} \to \mathbb{R}, \quad f(x) := -x_1^3(1-x_1)^3 \sin(2\pi(x_1 - 2x_2)). \tag{2.42}$$

Let us note that we do not know the homogenized solution u_0 exactly, we have however that $u_0 \in H^6(\Omega) \cap H^1_0(\Omega)$ as the right-hand side $f \in H^4(\Omega)$ satisfies the compatibility conditions $f = 0$ and $\partial_1^2 f - \partial_2^2 f = 0$ at the corners of the square $(0,1)^2 = \Omega$; see [59].

Figure 2.4 illustrates the L^∞ convergence rate $\mathcal{O}(\varepsilon)$ for the c-bad problem and the convergence rate $\mathcal{O}(\varepsilon^2)$ for the c-good problem. We also illustrate the corrected

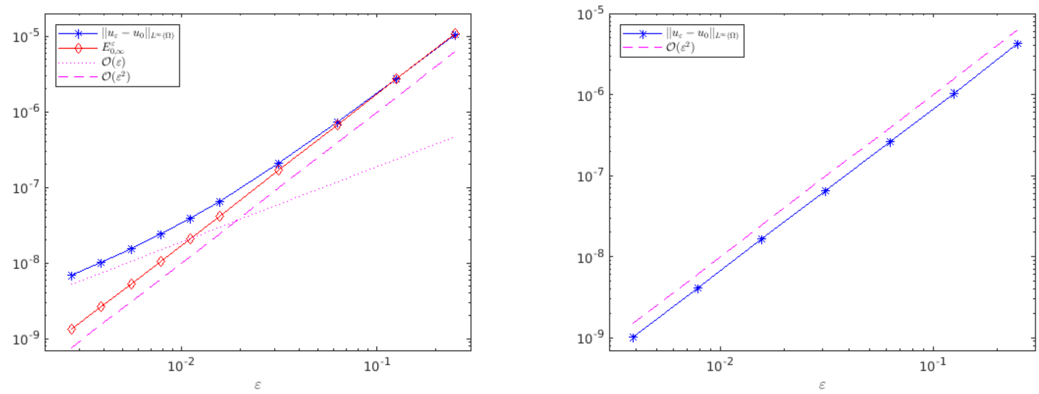


Figure 2.4: Illustration of the L^∞-rates $\|u_\varepsilon - u_0\|_{L^\infty(\Omega)} = \mathcal{O}(\varepsilon)$ and $E^\varepsilon_{0,\infty} = \mathcal{O}(\varepsilon^2)$ for the c-bad problem (left), and $\|u_\varepsilon - u_0\|_{L^\infty(\Omega)} = \mathcal{O}(\varepsilon^2)$ for the c-good problem (right) with the right-hand side (2.42).

L^∞ bound $E^\varepsilon_{0,\infty} = \mathcal{O}(\varepsilon^2)$ for the c-bad problem. We approximate the functions u_ε, u_0 and z with $\mathbb{P}_2$ finite elements as before.

Chapter 3

Numerical homogenization of linear nondivergence-form equations

In this chapter, we present and rigorously analyze the proposed numerical scheme. The outline of this chapter is as follows.

After providing the framework in Section 3.1, we discuss the numerical homogenization in Section 3.2, which is divided into three parts. In the first part, we approximate the invariant measure by a finite element method and provide a convergence result for the approximation. This is then used in the second part to obtain an approximation to the effective coefficient, i.e., to the constant matrix A^0. In the third part, we use a finite element method to discretize the homogenized problem and show convergence results for the approximation of the homogenized solution in $H^1(\Omega)$ and $H^2(\Omega)$, using the approximated effective coefficient, a comparison result, and two technical lemmata. Improvements to the convergence rates are given, provided more regularity on the coefficients is assumed.

In Section 3.3.1, we address the approximation of the corrector functions, presenting a method of successively approximating higher derivatives. We then use the homogenization results obtained in Chapter 2 and the approximations of the homogenized solution and the corrector functions from the previous subsections to approximate the original solution u_ε in Section 3.3.2.

Finally, we study the case of nonuniformly oscillating coefficients in Section 3.4, derive homogenization results similar to the case of periodic coefficients and discuss the numerical homogenization for this case. Numerical experiments demonstrating the theoretical results are provided in Section 3.5.

Annotation: Unless stated otherwise, this chapter contains novel results which have been obtained in Capdeboscq, Sprekeler, Süli [29]. The contribution of Y. Capdeboscq and E. Süli was of advisory nature.

3.1 Framework

The framework is the one considered in the previous chapter (Section 2.1) with a slight modification concerning the regularity of the coefficient A: We consider a symmetric matrix-valued function $A : \mathbb{R}^n \to \mathbb{R}^{n\times n}_{\mathrm{sym}}$ with the properties

$$\begin{cases} A \in W^{1,q}(Y) \text{ for some } q \in (n,\infty], \\ A \text{ is } Y\text{-periodic}, \\ \exists\, \lambda, \Lambda > 0 : \; \lambda|\xi|^2 \leq A(y)\xi\cdot\xi \leq \Lambda|\xi|^2 \quad \forall\, \xi, y \in \mathbb{R}^n. \end{cases} \tag{3.1}$$

Note that by the Sobolev embedding, we then have that

$$A \in C^{0,\alpha}(\mathbb{R}^n) \text{ for some } 0 < \alpha \leq 1.$$

Let us recall that under these assumptions on the coefficient, the corresponding invariant measure is of regularity $m \in W^{1,q}(Y)$ as well; see Remark 2.2.1. Note however that the function m is only in $W^{1,q}(Y)$ in general, and in particular it does not belong to $H^2(Y)$, as can be seen from the example chosen in the numerical experiments in Section 3.5.

As in the previous chapter, we are concerned with the problem

$$\begin{cases} A\left(\frac{\cdot}{\varepsilon}\right) : D^2 u_\varepsilon = f & \text{in } \Omega, \\ u_\varepsilon = 0 & \text{on } \partial\Omega, \end{cases} \tag{3.2}$$

with a small parameter $\varepsilon > 0$ and the triple (Ω, A, f) satisfying one of the following sets of assumptions:

For $m \in \mathbb{N}_0$ and $p \in (1,\infty)$, we define the set of assumptions $\mathcal{G}^{m,p}$ as

$$(\Omega, A, f) \in \mathcal{G}^{m,p} \quad \Longleftrightarrow \quad \begin{cases} \Omega \subset \mathbb{R}^n \text{ is a bounded } C^{2,\gamma} \text{ domain}, \gamma \in (0,1), \\ A : \mathbb{R}^n \to \mathbb{R}^{n\times n}_{\mathrm{sym}} \text{ satisfies (3.1)}, \\ f \in W^{m,p}(\Omega), \end{cases}$$

and the set of assumptions $\mathcal{H}^m$ as

$$(\Omega, A, f) \in \mathcal{H}^m \quad \Longleftrightarrow \quad \begin{cases} \Omega \subset \mathbb{R}^n \text{ is a bounded convex domain}, \\ A : \mathbb{R}^n \to \mathbb{R}^{n\times n}_{\mathrm{sym}} \text{ satisfies (3.1)}, \\ \exists\, \delta \in (0,1] : \dfrac{|A|^2}{(\mathrm{tr} A)^2} \leq \dfrac{1}{n-1+\delta} \text{ in } \mathbb{R}^n, \\ f \in H^m(\Omega). \end{cases}$$

We present a scheme for the numerical homogenization of the problem (3.2) that is based on finite element approximations.

3.2 Numerical homogenization via finite element approximations

The presentation of this section is divided into three parts, namely, the approximation of the invariant measure m, the approximation of the effective coefficient A^0, and the approximation of the homogenized solution u_0.

We start by discussing the finite element approximation of the invariant measure.

3.2.1 Approximation of the invariant measure

Let us recall from Lemma 2.2.1 that the invariant measure $m \in H^1_{\mathrm{per}}(Y)$ is the unique solution to the problem

$$\begin{cases} D^2 : (Am) = 0 \quad \text{in } Y, \\ m \text{ is } Y\text{-periodic}, \ \int_Y m = 1. \end{cases} \tag{3.3}$$

We have the positivity property $\inf_{\mathbb{R}^n} m > 0$ and the regularity $m \in W^{1,q}(Y)$; see Remark 2.2.1. Our approximation scheme will be based on the observation that

$$\tilde{m} := m - 1 \in W_{\mathrm{per}}(Y) = \left\{ v \in H^1_{\mathrm{per}}(Y) : \int_Y v = 0 \right\}$$

is the unique solution to the problem

$$\begin{cases} -\nabla \cdot (A\nabla \tilde{m} + \tilde{m}\, \mathrm{div} A) = \nabla \cdot (\mathrm{div} A) \quad \text{in } Y, \\ \tilde{m} \text{ is } Y\text{-periodic}, \ \int_Y \tilde{m} = 0, \end{cases}$$

that is, it satisfies

$$\int_Y (A\nabla \tilde{m} + \tilde{m}\, \mathrm{div} A) \cdot \nabla \varphi = -\int_Y (\mathrm{div} A) \cdot \nabla \varphi \qquad \forall\, \varphi \in W_{\mathrm{per}}(Y).$$

For the approximation of the invariant measure m, we consider a shape-regular triangulation of $\bar{Y}$ into triangles with longest edge $h > 0$ and let

$$\tilde{M}_h \subset W_{\mathrm{per}}(Y)$$

be the finite-dimensional subspace of $W_{\mathrm{per}}(Y)$ consisting of continuous Y-periodic piecewise linear functions on the triangulation with zero mean over Y. We assume that

$$\overline{\bigcup_{h>0} \tilde{M}_h} = W_{\mathrm{per}}(Y).$$

Then we have the following approximation result for m.

Theorem 3.2.1 ([29, Theorem 3.1] Approximation of the invariant measure)**.** *Let* $A : \mathbb{R}^n \to \mathbb{R}^{n\times n}_{\mathrm{sym}}$ *satisfy* (3.1)*. Then, for* $h > 0$ *sufficiently small, there exists a unique* $\tilde{m}_h \in \tilde{M}_h$ *such that*

$$\int_Y (A\nabla\tilde{m}_h + \tilde{m}_h \,\mathrm{div}A)\cdot\nabla\varphi_h = -\int_Y (\mathrm{div}A)\cdot\nabla\varphi_h \qquad \forall\,\varphi_h \in \tilde{M}_h, \tag{3.4}$$

and writing $m_h := \tilde{m}_h + 1$*, we have the error bound*

$$\|m - m_h\|_{L^2(Y)} + h\|m - m_h\|_{H^1(Y)} \lesssim h \inf_{\tilde{v}_h\in\tilde{M}_h} \|m - (\tilde{v}_h + 1)\|_{H^1(Y)},$$

where m *is the invariant measure given by* (3.3)*.*

Proof. We define the bilinear form

$$a : W_{\mathrm{per}}(Y) \times W_{\mathrm{per}}(Y) \longrightarrow \mathbb{R}, \quad a(u,v) := \int_Y A\nabla u\cdot\nabla v + \int_Y u(\mathrm{div}A)\cdot\nabla v,$$

and observe that $\tilde{m} := m - 1 \in W_{\mathrm{per}}(Y)$ satisfies

$$a(\tilde{m},\varphi) = -\int_Y (\mathrm{div}A)\cdot\nabla\varphi \qquad \forall\,\varphi \in W_{\mathrm{per}}(Y).$$

We further observe that (3.4) is equivalent to finding $\tilde{m}_h \in \tilde{M}_h$ such that

$$a(\tilde{m}_h,\varphi_h) = -\int_Y (\mathrm{div}A)\cdot\nabla\varphi_h \qquad \forall\,\varphi_h \in \tilde{M}_h. \tag{3.5}$$

We start by showing boundedness of a and a Gårding-type inequality. We claim that there exist constants $C_b, C_g > 0$ such that

$$|a(u,v)| \le C_b\|u\|_{H^1(Y)}\|v\|_{H^1(Y)} \qquad \forall\,u,v \in W_{\mathrm{per}}(Y), \tag{3.6}$$

and

$$a(u,u) \ge \frac{\lambda}{2}\|u\|^2_{H^1(Y)} - C_g\|u\|^2_{L^2(Y)} \qquad \forall\,u \in W_{\mathrm{per}}(Y). \tag{3.7}$$

Let us first show (3.6). For $u, v \in W_{\mathrm{per}}(Y)$, by Hölder's inequality and Sobolev embeddings (note that, according to (3.1), $q > n$), we have that

$$\left|\int_Y u(\mathrm{div}A)\cdot\nabla v\right| \le \|\mathrm{div}A\|_{L^q(Y)}\|u\|_{L^{\frac{2q}{q-2}}(Y)}\|\nabla v\|_{L^2(Y)} \lesssim \|u\|_{H^1(Y)}\|v\|_{H^1(Y)}.$$

Using the fact that $A \in W^{1,q}(Y) \hookrightarrow L^\infty(Y)$ since $q > n$, we obtain the bound

$$|a(u,v)| \le \left|\int_Y A\nabla u\cdot\nabla v\right| + \left|\int_Y u(\mathrm{div}A)\cdot\nabla v\right| \lesssim \|u\|_{H^1(Y)}\|v\|_{H^1(Y)}$$

for any $u, v \in W_{\text{per}}(Y)$, i.e., (3.6) holds.

Let us now show the estimate (3.7). For $u \in W_{\text{per}}(Y)$, by ellipticity and Hölder's inequality, we have

$$\begin{aligned} a(u,u) &= \int_Y A\nabla u \cdot \nabla u + \int_Y u(\text{div}A) \cdot \nabla u \\ &\geq \lambda \|\nabla u\|^2_{L^2(Y)} - \|\text{div}A\|_{L^q(Y)} \|u\|_{L^{\frac{2q}{q-2}}(Y)} \|\nabla u\|_{L^2(Y)}. \end{aligned}$$

For the second term we use the Gagliardo–Nirenberg inequality and Young's inequality to obtain

$$\begin{aligned} \|\text{div}A\|_{L^q(Y)} \|u\|_{L^{\frac{2q}{q-2}}(Y)} \|\nabla u\|_{L^2(Y)} &\leq C(q,n) \|\text{div}A\|_{L^q(Y)} \|u\|^{1-\frac{n}{q}}_{L^2(Y)} \|\nabla u\|^{1+\frac{n}{q}}_{L^2(Y)} \\ &\leq \frac{\lambda}{2} \|\nabla u\|^2_{L^2(Y)} + C(q,n,\lambda, \|\text{div}A\|_{L^q(Y)}) \|u\|^2_{L^2(Y)}. \end{aligned}$$

Therefore, we have

$$\begin{aligned} a(u,u) &\geq \frac{\lambda}{2} \|\nabla u\|^2_{L^2(Y)} - C(q,n,\lambda, \|\text{div}A\|_{L^q(Y)}) \|u\|^2_{L^2(Y)} \\ &= \frac{\lambda}{2} \|u\|^2_{H^1(Y)} - \left(\frac{\lambda}{2} + C(q,n,\lambda, \|\text{div}A\|_{L^q(Y)}) \right) \|u\|^2_{L^2(Y)} \end{aligned}$$

for any $u \in W_{\text{per}}(Y)$, i.e., (3.7) holds with

$$C_g := \frac{\lambda}{2} + C(q,n,\lambda, \|\text{div}A\|_{L^q(Y)}).$$

We use Schatz's method to derive an *a priori* estimate; see [85].

From our Gårding-type inequality (3.7) we see that (note that $\tilde{m} - \tilde{m}_h \in W_{\text{per}}(Y)$)

$$\begin{aligned} \|\tilde{m} - \tilde{m}_h\|_{H^1(Y)} - \frac{2C_g}{\lambda} \|\tilde{m} - \tilde{m}_h\|_{L^2(Y)} &\leq \|\tilde{m} - \tilde{m}_h\|_{H^1(Y)} - \frac{2C_g}{\lambda} \frac{\|\tilde{m} - \tilde{m}_h\|^2_{L^2(Y)}}{\|\tilde{m} - \tilde{m}_h\|_{H^1(Y)}} \\ &\leq \frac{2}{\lambda} \frac{a(\tilde{m} - \tilde{m}_h, \tilde{m} - \tilde{m}_h)}{\|\tilde{m} - \tilde{m}_h\|_{H^1(Y)}}. \end{aligned} \tag{3.8}$$

By Galerkin-orthogonality and boundedness, we have for any $\tilde{v}_h \in \tilde{M}_h$ that

$$\frac{a(\tilde{m} - \tilde{m}_h, \tilde{m} - \tilde{m}_h)}{\|\tilde{m} - \tilde{m}_h\|_{H^1(Y)}} = \frac{a(\tilde{m} - \tilde{m}_h, \tilde{m} - \tilde{v}_h)}{\|\tilde{m} - \tilde{m}_h\|_{H^1(Y)}} \leq C_b \|\tilde{m} - \tilde{v}_h\|_{H^1(Y)},$$

and taking the infimum over all $\tilde{v}_h \in \tilde{M}_h$, we find that

$$\frac{a(\tilde{m} - \tilde{m}_h, \tilde{m} - \tilde{m}_h)}{\|\tilde{m} - \tilde{m}_h\|_{H^1(Y)}} \leq C_b \inf_{\tilde{v}_h \in \tilde{M}_h} \|\tilde{m} - \tilde{v}_h\|_{H^1(Y)}.$$

Combining this estimate with (3.8) yields

$$\|\tilde{m} - \tilde{m}_h\|_{H^1(Y)} - \frac{2C_g}{\lambda}\|\tilde{m} - \tilde{m}_h\|_{L^2(Y)} \leq \frac{2C_b}{\lambda} \inf_{\tilde{v}_h \in \tilde{M}_h} \|\tilde{m} - \tilde{v}_h\|_{H^1(Y)}. \tag{3.9}$$

Next, we use an Aubin–Nitsche-type duality argument.

Let $\phi \in W_{\mathrm{per}}(Y)$ be the unique solution to

$$\begin{cases} -\nabla \cdot (A\nabla\phi) + (\mathrm{div} A) \cdot \nabla\phi = \frac{\tilde{m} - \tilde{m}_h}{m} & \text{in } Y, \\ \phi \text{ is } Y\text{-periodic}, \ \int_Y \phi = 0. \end{cases} \tag{3.10}$$

We note that the solvability condition (2.7) is satisfied:

$$\int_Y \frac{\tilde{m} - \tilde{m}_h}{m} m = \int_Y (\tilde{m} - \tilde{m}_h) = 0.$$

We have, using the bounds on the invariant measure, the weak formulation of (3.10) and the symmetry of A, that

$$\begin{aligned} \|\tilde{m} - \tilde{m}_h\|^2_{L^2(Y)} &\lesssim \int_Y \frac{\tilde{m} - \tilde{m}_h}{m}(\tilde{m} - \tilde{m}_h) \\ &\lesssim \int_Y A\nabla\phi \cdot \nabla(\tilde{m} - \tilde{m}_h) + \int_Y (\mathrm{div} A) \cdot \nabla\phi \, (\tilde{m} - \tilde{m}_h) \\ &\lesssim \int_Y A\nabla(\tilde{m} - \tilde{m}_h) \cdot \nabla\phi + \int_Y (\tilde{m} - \tilde{m}_h)(\mathrm{div} A) \cdot \nabla\phi. \end{aligned}$$

Next, we use Galerkin orthogonality, the boundedness (3.6) and an interpolation inequality to obtain

$$\begin{aligned} \|\tilde{m} - \tilde{m}_h\|^2_{L^2(Y)} &\lesssim a(\tilde{m} - \tilde{m}_h, \phi) \\ &\lesssim a(\tilde{m} - \tilde{m}_h, \phi - \mathcal{I}_h\phi) \\ &\lesssim \|\tilde{m} - \tilde{m}_h\|_{H^1(Y)} \|\phi - \mathcal{I}_h\phi\|_{H^1(Y)} \\ &\lesssim h\|\tilde{m} - \tilde{m}_h\|_{H^1(Y)} \|\phi\|_{H^2(Y)}, \end{aligned}$$

where $\mathcal{I}_h\phi$ denotes the continuous piecewise linear interpolant (for $n \leq 3$ and quasi-interpolant for $n \geq 4$) of ϕ on the triangulation. Finally, by a regularity estimate for ϕ and the bounds on the invariant measure, we arrive at the bound

$$\|\phi\|_{H^2(Y)} \lesssim \left\| \frac{\tilde{m} - \tilde{m}_h}{m} \right\|_{L^2(Y)} \lesssim \|\tilde{m} - \tilde{m}_h\|_{L^2(Y)},$$

which provides us with the estimate

$$\|\tilde{m} - \tilde{m}_h\|_{L^2(Y)} \leq C_0 h \|\tilde{m} - \tilde{m}_h\|_{H^1(Y)}$$

for some $C_0 > 0$. Combining this with (3.9) we have

$$\left(1 - \frac{2C_g C_0}{\lambda} h\right) \|\tilde{m} - \tilde{m}_h\|_{H^1(Y)} \leq \|\tilde{m} - \tilde{m}_h\|_{H^1(Y)} - \frac{2C_g}{\lambda} \|\tilde{m} - \tilde{m}_h\|_{L^2(Y)}$$
$$\leq \frac{2C_b}{\lambda} \inf_{\tilde{v}_h \in \tilde{M}_h} \|\tilde{m} - \tilde{v}_h\|_{H^1(Y)}.$$

Therefore, for h sufficiently small, we arrive at the bounds

$$\|\tilde{m} - \tilde{m}_h\|_{H^1(Y)} \lesssim \inf_{\tilde{v}_h \in \tilde{M}_h} \|\tilde{m} - \tilde{v}_h\|_{H^1(Y)},$$

and

$$\|\tilde{m} - \tilde{m}_h\|_{L^2(Y)} \leq C_0 h \|\tilde{m} - \tilde{m}_h\|_{H^1(Y)} \lesssim h \inf_{\tilde{v}_h \in \tilde{M}_h} \|\tilde{m} - \tilde{v}_h\|_{H^1(Y)}.$$

We have thus established the *a priori* estimate

$$\|\tilde{m} - \tilde{m}_h\|_{L^2(Y)} + h\|\tilde{m} - \tilde{m}_h\|_{H^1(Y)} \lesssim h \inf_{\tilde{v}_h \in \tilde{M}_h} \|\tilde{m} - \tilde{v}_h\|_{H^1(Y)},$$

which immediately implies existence and uniqueness of solutions to (3.5).

Finally, using that $m = \tilde{m} + 1$ and $m_h = \tilde{m}_h + 1$, we conclude that

$$\|m - m_h\|_{L^2(Y)} + h\|m - m_h\|_{H^1(Y)} \lesssim h \inf_{\tilde{v}_h \in \tilde{M}_h} \|m - (\tilde{v}_h + 1)\|_{H^1(Y)}.$$

□

Remark 3.2.1. *In particular, since*

$$\inf_{\tilde{v}_h \in \tilde{M}_h} \|m - (\tilde{v}_h + 1)\|_{H^1(Y)} = o(1),$$

we have that $m_h \to m$ in $H^1(Y)$ as h tends to zero.

3.2.2 Approximation of the effective coefficient

We use this finite element approximation of the invariant measure to obtain an approximation to the effective coefficient. We recall that the effective coefficient is the constant matrix $A^0 \in \mathbb{R}^{n \times n}_{\text{sym}}$ given by

$$A^0 = \int_Y Am. \tag{3.11}$$

To this end, we first replace the invariant measure m by the approximation m_h from Theorem 3.2.1, and then replace the integrand by its piecewise linear interpolant,

$$A^0_h := \int_Y \mathcal{I}_h(Am_h).$$

This integral can be computed exactly using an appropriate quadrature rule. The following lemma gives an error estimate for this approximation.

Lemma 3.2.1 (Approximation of A^0)**.** *Let $A : \mathbb{R}^n \to \mathbb{R}^{n\times n}_{\mathrm{sym}}$ satisfy* (3.1)*. Further, let $A^0 = (a^0_{ij})_{1\le i,j\le n} \in \mathbb{R}^{n\times n}$ be the effective coefficient given by* (3.11)*, let m_h be the approximation to the invariant measure given by Theorem 3.2.1, and let $A^0_h = (a^0_{ij,h})_{1\le i,j\le n} \in \mathbb{R}^{n\times n}$ be the matrix given by*

$$a^0_{ij,h} := \int_Y \mathcal{I}_h(a_{ij}m_h), \qquad 1 \le i,j \le n.$$

Then, for $h > 0$ sufficiently small, A^0_h is elliptic and we have the error bound

$$\max_{1\le i,j\le n} \left|a^0_{ij} - a^0_{ij,h}\right| \lesssim h.$$

Proof. Fix $1 \le i,j \le n$. Using the definition of $A^0 = (a^0_{ij})$, i.e.,

$$a^0_{ij} = \int_Y a_{ij}m,$$

we obtain the estimate

$$|a^0_{ij} - a^0_{ij,h}| \le \|a_{ij}(m - m_h)\|_{L^1(Y)} + \|a_{ij}m_h - \mathcal{I}_h(a_{ij}m_h)\|_{L^1(Y)}.$$

For the first term, we have

$$\|a_{ij}(m - m_h)\|_{L^1(Y)} \lesssim \|m - m_h\|_{L^1(Y)} \lesssim \|m - m_h\|_{L^2(Y)}.$$

For the second term, let us first note that using $a_{ij} \in W^{1,q}(Y)$ with $q > n$ and Sobolev embeddings, we have

$$\begin{aligned}|a_{ij}m_h|_{H^1(Y)} &\le \|\nabla a_{ij}\|_{L^q(Y)}\|m_h\|_{L^{\frac{2q}{q-2}}(Y)} + \|a_{ij}\|_{L^\infty(Y)}\|\nabla m_h\|_{L^2(Y)} \\ &\lesssim \|a_{ij}\|_{W^{1,q}(Y)}\|m_h\|_{H^1(Y)}.\end{aligned}$$

Therefore, using a standard interpolation error bound, we obtain

$$\begin{aligned}\|a_{ij}m_h - \mathcal{I}_h(a_{ij}m_h)\|_{L^1(Y)} &\lesssim \|a_{ij}m_h - \mathcal{I}_h(a_{ij}m_h)\|_{L^2(Y)} \\ &\lesssim h|a_{ij}m_h|_{H^1(Y)} \\ &\lesssim h\|a_{ij}\|_{W^{1,q}(Y)}\|m_h\|_{H^1(Y)}.\end{aligned}$$

By Theorem 3.2.1, for $h > 0$ sufficiently small, we have that

$$\begin{aligned}|a^0_{ij} - a^0_{ij,h}| &\lesssim \|m - m_h\|_{L^2(Y)} + h\|m_h\|_{H^1(Y)} \\ &\lesssim \|m - m_h\|_{L^2(Y)} + h\|m - m_h\|_{H^1(Y)} + h\|m\|_{H^1(Y)} \\ &\lesssim h \inf_{\tilde{v}_h\in\tilde{M}_h} \|m - (\tilde{v}_h + 1)\|_{H^1(Y)} + h\|m\|_{H^1(Y)} \\ &\lesssim h\|m - 1\|_{H^1(Y)} + h\|m\|_{H^1(Y)} \\ &\lesssim h.\end{aligned}$$

Finally, we note that this implies that for $h > 0$ sufficiently small, A^0_h is elliptic. $\square$

3.2.3 Approximation of the homogenized solution

In this section we discuss the approximation of the homogenized solution u_0, that is, the solution $u_0 \in H^2(\Omega) \cap H_0^1(\Omega)$ to the effective problem

$$\begin{cases} A^0 : D^2 u_0 = f & \text{in } \Omega, \\ u_0 = 0 & \text{on } \partial\Omega. \end{cases} \tag{3.12}$$

For its numerical approximation, we use the following comparison result for the error committed when replacing A^0 by A_h^0.

Lemma 3.2.2 (Comparison result). *Assume either that $(\Omega, A, f) \in \mathcal{G}^{0,2}$ or that $(\Omega, A, f) \in \mathcal{H}^0$. Let $A_h^0 \in \mathbb{R}^{n\times n}$ be the approximation to A^0 as in Lemma 3.2.1. Then, for $h > 0$ sufficiently small, we have that*

$$\|u_0 - u_0^h\|_{H^2(\Omega)} \lesssim h \|f\|_{L^2(\Omega)},$$

where $u_0^h \in H^2(\Omega) \cap H_0^1(\Omega)$ is the solution to the problem

$$\begin{cases} A_h^0 : D^2 u_0^h = f & \textit{in } \Omega, \\ u_0^h = 0 & \textit{on } \partial\Omega, \end{cases} \tag{3.13}$$

and $u_0 \in H^2(\Omega) \cap H_0^1(\Omega)$ is the solution to the homogenized problem (3.12).

Proof. We let $w_h := u_0 - u_0^h \in H^2(\Omega) \cap H_0^1(\Omega)$ and note that w_h is the unique solution to the boundary-value problem

$$\begin{cases} A^0 : D^2 w_h = (A_h^0 - A^0) : D^2 u_0^h & \text{in } \Omega, \\ w_h = 0 & \text{on } \partial\Omega. \end{cases}$$

We recall that $A^0 \in \mathbb{R}^{n\times n}$ is an elliptic constant matrix. For $h > 0$ sufficiently small, by an H^2 *a priori* estimate, the Cauchy–Schwarz inequality and Lemma 3.2.1,

$$\begin{aligned} \|w_h\|_{H^2(\Omega)} &\lesssim \|(A_h^0 - A^0) : D^2 u_0^h\|_{L^2(\Omega)} \\ &\lesssim \left(\int_\Omega \left| \sum_{i,j=1}^n (a_{ij,h}^0 - a_{ij}^0) \partial_{ij}^2 u_0^h \right|^2 \right)^{\frac{1}{2}} \\ &\lesssim \left(\int_\Omega \left(\sum_{i,j=1}^n |a_{ij,h}^0 - a_{ij}^0|^2 \right) \left(\sum_{i,j=1}^n |\partial_{ij}^2 u_0^h|^2 \right) \right)^{\frac{1}{2}} \\ &\lesssim h \, |u_0^h|_{H^2(\Omega)}. \end{aligned}$$

Finally, we show that for $h > 0$ sufficiently small, we have

$$\|u_0^h\|_{H^2(\Omega)} \lesssim \|f\|_{L^2(\Omega)} \tag{3.14}$$

with the constant being independent of h. This can be seen by rewriting (3.13) as

$$\begin{cases} A^0 : D^2 u_0^h = f + (A^0 - A_h^0) : D^2 u_0^h & \text{in } \Omega, \\ u_0^h = 0 & \text{on } \partial\Omega. \end{cases} \tag{3.15}$$

Then, again by an H^2 *a priori* estimate and Lemma 3.2.1,

$$\|u_0^h\|_{H^2(\Omega)} \lesssim \|f + (A^0 - A_h^0) : D^2 u_0^h\|_{L^2(\Omega)} \lesssim \|f\|_{L^2(\Omega)} + h\|u_0^h\|_{H^2(\Omega)}$$

with constants independent of h, i.e., for $h > 0$ sufficiently small, (3.14) holds with the constant being independent of h. □

Finally, we can use an $H_0^1(\Omega)$-conforming finite element approximation $u_0^{h,H}$ to the solution u_0^h of (3.13), satisfying the error bound

$$\left\|u_0^h - u_0^{h,H}\right\|_{H^1(\Omega)} \lesssim H\|u_0^h\|_{H^2(\Omega)} \lesssim H\|f\|_{L^2(\Omega)}$$

with constants independent of h. By the triangle inequality and the results obtained in this section, we have the following approximation result for u_0.

Theorem 3.2.2 ([29, Theorem 3.5] H^1-norm approximation of u_0)**.** *Assume either that $(\Omega, A, f) \in \mathcal{G}^{0,2}$, or that $(\Omega, A, f) \in \mathcal{H}^0$. Then, the approximation $u_0^{h,H}$ obtained by the procedure described above satisfies the error bound*

$$\left\|u_0 - u_0^{h,H}\right\|_{H^1(\Omega)} \lesssim (h + H)\|f\|_{L^2(\Omega)}.$$

Let us now assume either that $(\Omega, A, f) \in \mathcal{G}^{1,2}$ or that $(\Omega, A, f) \in \mathcal{H}^1$. Further, assume that for $h > 0$ sufficiently small, we have that $u_0^h \in H^3(\Omega)$ with

$$\left\|u_0^h\right\|_{H^3(\Omega)} \lesssim \|f\|_{H^1(\Omega)}, \tag{3.16}$$

where the constant is independent of h. The following lemma provides two situations where this is satisfied.

Lemma 3.2.3. *Let (Ω, A, f) be such that*

(i) $(\Omega, A, f) \in \mathcal{G}^{1,2}$ with $\partial\Omega \in C^3$, or

(ii) $(\Omega, A, f) \in \mathcal{H}^1$ with $\Omega \subset \mathbb{R}^2$ being a polygon and $f \in H_0^1(\Omega)$.

Then, for $h > 0$ sufficiently small, (3.16) *holds.*

Proof. We start with the case (i). To this end, let $(\Omega, A, f) \in \mathcal{G}^{1,2}$ with $\partial\Omega \in C^3$. Then, by elliptic regularity theory, we have $u_0^h \in H^3(\Omega)$. Using elliptic regularity for problem (3.15) yields

$$\|u_0^h\|_{H^3(\Omega)} \lesssim \|f + (A^0 - A_h^0) : D^2 u_0^h\|_{H^1(\Omega)} \lesssim \|f\|_{H^1(\Omega)} + h\|u_0^h\|_{H^3(\Omega)}$$

with constants independent of h, i.e., for $h > 0$ sufficiently small, (3.16) holds with the constant being independent of h.

Let us now show the claim for the case (ii). To this end, let $(\Omega, A, f) \in \mathcal{H}^1$ with $\Omega \subset \mathbb{R}^2$ being a polygon and $f \in H_0^1(\Omega)$. Since

$$A_h^0 = A^0 + \left(A_h^0 - A^0\right) =: A^0 + B_h$$

is symmetric and elliptic for $h > 0$ sufficiently small, there exists an orthogonal matrix $Q_h \in \mathbb{R}^{2\times 2}$ with $Q_h Q_h^{\mathrm{T}} = Q_h^{\mathrm{T}} Q_h = I_2$ such that

$$Q_h \left(A^0 + B_h\right) Q_h^{\mathrm{T}} = \operatorname{diag}(\lambda_h^+, \lambda_h^-) =: \Lambda_h,$$

where $\lambda_h^\pm > 0$ are given by

$$2\lambda_h^\pm = \operatorname{tr}\left(A^0 + B_h\right) \pm \left(\left(\operatorname{tr}\left(A^0 + B_h\right)\right)^2 - 4\det\left(A^0 + B_h\right)\right)^{\frac{1}{2}}.$$

We note that, by Lemma 3.2.1, the entries of $B_h = (b_{ij}^h)_{1\le i,j\le 2}$ satisfy $b_{ij}^h \lesssim h$, and therefore, for $h > 0$ sufficiently small, we have $0 < \lambda_h^\pm + (\lambda_h^\pm)^{-1} \lesssim 1$.

The problem (3.13) in the new coordinates reads

$$\begin{cases} \Delta U_h = F_h & \text{in } P_h, \\ U_h = 0 & \text{on } \partial P_h, \end{cases} \tag{3.17}$$

where $U_h := u_0^h\left(Q_h^{\mathrm{T}} \Lambda_h^{\frac{1}{2}} \cdot\right)$, $F_h := f\left(Q_h^{\mathrm{T}} \Lambda_h^{\frac{1}{2}} \cdot\right)$, and $P_h := \Lambda_h^{-\frac{1}{2}} Q_h \Omega$. Note that P_h is still a bounded convex polygonal domain and that $F_h \in H_0^1(P_h)$. By the change of variables formula and the orthogonality of Q_h,

$$\begin{aligned} \|f\|_{H^1(\Omega)}^2 = \int_\Omega \left(|f|^2 + |\nabla f|^2\right) &= \det \Lambda_h^{\frac{1}{2}} \int_{P_h} \left(\left|f\left(Q_h^{\mathrm{T}} \Lambda_h^{\frac{1}{2}} \cdot\right)\right|^2 + \left|\nabla f\left(Q_h^{\mathrm{T}} \Lambda_h^{\frac{1}{2}} \cdot\right)\right|^2\right) \\ &= \det \Lambda_h^{\frac{1}{2}} \int_{P_h} \left(|F_h|^2 + \left|Q_h^{\mathrm{T}} \Lambda_h^{-\frac{1}{2}} \nabla F_h\right|^2\right) \\ &= \det \Lambda_h^{\frac{1}{2}} \int_{P_h} \left(|F_h|^2 + \left|\Lambda_h^{-\frac{1}{2}} \nabla F_h\right|^2\right) \\ &\gtrsim \int_{P_h} \left(|F_h|^2 + |\nabla F_h|^2\right) = \|F_h\|_{H^1(P_h)}^2. \end{aligned}$$

Using Lemma 3.2.4, we have that, for $h > 0$ sufficiently small, the solution to (3.17) satisfies

$$\|U_h\|_{H^3(P_h)} \lesssim \|F_h\|_{H^1(P_h)} \lesssim \|f\|_{H^1(\Omega)}$$

with constants independent of h. It remains to show the bound

$$\|u_0^h\|_{H^3(\Omega)} \lesssim \|U_h\|_{H^3(P_h)}. \tag{3.18}$$

By the change of variables formula and the orthogonality of Q_h, we obtain similarly as before,

$$\begin{aligned}
\|u_0^h\|_{H^3(\Omega)}^2 &= \int_\Omega \left(|u_0^h|^2 + \left|\nabla u_0^h\right|^2 + \left|D^2 u_0^h\right|^2\right) + \sum_{i=1}^2 \int_\Omega \left|D^2 \partial_i u_0^h\right|^2 \\
&= \det \Lambda_h^{\frac{1}{2}} \int_{P_h} \left(|U_h|^2 + \left|Q_h^{\mathrm{T}} \Lambda_h^{-\frac{1}{2}} \nabla U_h\right|^2 + \left|Q_h^{\mathrm{T}} \Lambda_h^{-\frac{1}{2}}\, D^2 U_h\, \Lambda_h^{-\frac{1}{2}} Q_h\right|^2\right) \\
&\quad + \sum_{i=1}^2 \det \Lambda_h^{\frac{1}{2}} \int_{P_h} \left|\sum_{j=1}^2 \frac{(Q_h)_{ji}}{\sqrt{(\Lambda_h)_{jj}}} Q_h^{\mathrm{T}} \Lambda_h^{-\frac{1}{2}}\, D^2 \partial_j U_h\, \Lambda_h^{-\frac{1}{2}} Q_h\right|^2 \\
&= \det \Lambda_h^{\frac{1}{2}} \int_{P_h} \left(|U_h|^2 + \left|\Lambda_h^{-\frac{1}{2}} \nabla U_h\right|^2 + \left|\Lambda_h^{-\frac{1}{2}}\, D^2 U_h\, \Lambda_h^{-\frac{1}{2}}\right|^2\right) \\
&\quad + \sum_{i=1}^2 \frac{\det \Lambda_h^{\frac{1}{2}}}{(\Lambda_h)_{ii}} \int_{P_h} \left|\Lambda_h^{-\frac{1}{2}}\, D^2 \partial_i U_h\, \Lambda_h^{-\frac{1}{2}}\right|^2 \\
&\lesssim \int_{P_h} \left(|U_h|^2 + |\nabla U_h|^2 + \left|D^2 U_h\right|^2\right) + \sum_{i=1}^2 \int_{P_h} \left|D^2 \partial_i U_h\right|^2 = \|U_h\|_{H^3(P_h)}^2,
\end{aligned}$$

i.e., we have established the bound (3.18). We conclude that, for $h > 0$ sufficiently small, we have (3.16), i.e.,

$$\|u_0^h\|_{H^3(\Omega)} \lesssim \|f\|_{H^1(\Omega)},$$

where the constant is independent of h. □

In the proof of Lemma 3.2.3, we have used the following result on the regularity of solutions to Poisson's problem on convex polygons; see also [56, 59, 60, 71].

Lemma 3.2.4. *Let $\Omega \subset \mathbb{R}^2$ be a convex polygonal domain and $f \in H_0^1(\Omega)$. Then the solution $u \in H_0^1(\Omega)$ to the problem*

$$\begin{cases} \Delta u = f & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega, \end{cases}$$

satisfies the bound

$$\|u\|_{H^3(\Omega)} \lesssim \|f\|_{H^1(\Omega)}. \tag{3.19}$$

Proof. First, note that since $\Omega \subset \mathbb{R}^2$ is a convex polygonal domain, we have $u \in H^2(\Omega) \cap H_0^1(\Omega)$ with $\|u\|_{H^2(\Omega)} \lesssim \|f\|_{L^2(\Omega)}$; see [56]. Since $f \in H_0^1(\Omega)$, there exists a sequence of smooth functions with compact support $(f_m)_m \subset C_c^\infty(\Omega)$ such that $f_m \to f$ in $H^1(\Omega)$. Let $(u_m)_m \subset H_0^1(\Omega)$ be the sequence of solutions in $H_0^1(\Omega)$ to $\Delta u_m = f_m$ in Ω, and note that $(u_m)_m \subset C^\infty(\bar{\Omega})$ since the functions f_m satisfy compatibility conditions of any order; see [56, Sec. 5.1]. Again we use the H^2-regularity result for solutions of Poisson's problem on convex polygons to obtain

$$\|u_m - u\|_{H^2(\Omega)} \lesssim \|f_m - f\|_{L^2(\Omega)} \to 0,$$

i.e., $u_m \to u$ in $H^2(\Omega)$.

Next, we shall use the fact that

$$|v|_{H^3(\Omega)} = \|\nabla(\Delta v)\|_{L^2(\Omega)} \quad \forall\, v \in \left\{w \in H_0^1(\Omega) : \Delta w \in H_0^1(\Omega)\right\} \cap C^\infty(\bar{\Omega}); \tag{3.20}$$

see [71]. We apply (3.20) to the difference of two elements of the sequence $(u_m)_m$ to find that $(u_m)_m$ is a Cauchy sequence in $H^3(\Omega)$, using that $f_m \to f$ in $H^1(\Omega)$. Thus, $u_m \to u$ in $H^3(\Omega)$ and passing to the limit in (3.20) applied to the functions u_m yields

$$|u|_{H^3(\Omega)} = \|\nabla f\|_{L^2(\Omega)}.$$

Since $\|u\|_{H^2(\Omega)} \lesssim \|f\|_{L^2(\Omega)}$, we conclude the bound (3.19). □

Remark 3.2.2. *The assumption $f \in H_0^1(\Omega)$ in Lemma 3.2.4 can be weakened provided f satisfies certain compatibility conditions; see [56, Theorem 5.1.2.4].*

Then an $H^2(\Omega) \cap H_0^1(\Omega)$-conforming finite element approximation $u_0^{h,H}$ to the solution u_0^h of (3.13), that satisfies the error bound

$$\left\|u_0^h - u_0^{h,H}\right\|_{H^2(\Omega)} \lesssim H \left\|u_0^h\right\|_{H^3(\Omega)} \lesssim H \left\|f\right\|_{H^1(\Omega)}, \tag{3.21}$$

provides by Lemma 3.2.2 and the triangle inequality an approximation to u_0.

Theorem 3.2.3 ([29, Theorem 3.9] H^2-norm approximation of u_0)**.** *Assume either that $(\Omega, A, f) \in \mathcal{G}^{1,2}$ or that $(\Omega, A, f) \in \mathcal{H}^1$, and assume* (3.16)*. Then, the approximation $u_0^{h,H}$ obtained by the procedure described above satisfies*

$$\left\|u_0 - u_0^{h,H}\right\|_{H^2(\Omega)} \lesssim (h + H)\|f\|_{H^1(\Omega)}.$$

Remark 3.2.3 (Improvements)**.** *We note that if we assume that $A \in W^{2,\infty}(Y)$, then we have the following improved results.*

(i) *Approximation of m: In this case, $m \in H^2(Y)$ and we have that*

$$\inf_{\tilde{v}_h \in \tilde{M}_h} \|m - (\tilde{v}_h + 1)\|_{H^1(Y)} \leq \left\| m - \mathcal{I}_h m - \int_Y (m - \mathcal{I}_h m) \right\|_{H^1(Y)} \lesssim h \|m\|_{H^2(Y)},$$

by choosing $\tilde{v}_h = \mathcal{I}_h m - \int_Y \mathcal{I}_h m$, and using an interpolation error bound. Therefore, Theorem 3.2.1 yields

$$\|m - m_h\|_{L^2(Y)} + h\|m - m_h\|_{H^1(Y)} \lesssim h^2 \|m\|_{H^2(Y)}.$$

(ii) *Approximation of A^0: By an interpolation error bound and the fact that m_h is piecewise linear, one has*

$$\|a_{ij} m_h - \mathcal{I}_h(a_{ij} m_h)\|_{L^1(Y)} \lesssim h^2 \|a_{ij}\|_{W^{2,\infty}(Y)} \|m_h\|_{H^1(Y)}.$$

Therefore, the proof of Lemma 3.2.1 yields

$$\max_{1 \leq i,j \leq n} \left| a^0_{ij} - a^0_{ij,h} \right| \lesssim h^2 \|A\|_{W^{2,\infty}(Y)} \|m\|_{H^2(Y)} \lesssim h^2 \|A\|_{W^{2,\infty}(Y)}.$$

(iii) *Approximation of u_0: It follows that the results of Lemma 3.2.2, Theorem 3.2.2 and Theorem 3.2.3 can be improved to second-order convergence in h, i.e.,*

$$\left\| u_0 - u_0^{h,H} \right\|_{H^s(\Omega)} \lesssim (h^2 \|A\|_{W^{2,\infty}(Y)} + H) \|f\|_{H^{s-1}(\Omega)} = \mathcal{O}(h^2 + H),$$

for $s = 1, 2$, respectively.

We note that second-order convergence with respect to h could not have been obtained by using a piecewise constant approximation of $a_{ij} m_h$ instead of the piecewise linear approximation considered here. For the approximation of derivatives of u_0 of higher than second order, the post-processing method of Babuška in [17] can be used to obtain error bounds in norms involving derivatives of higher order than the energy norm (the norm natural to the problem).

For bounded convex polygonal domains $\Omega \subset \mathbb{R}^2$, an H^2-conforming approximation to the solution of (3.13) can be obtained as follows. Assume that $f \in H^1_0(\Omega)$ so that (3.16) holds. Consider a shape-regular triangulation of Ω into triangles with longest edge $H > 0$, and let

$$V_H \subset H^2(\Omega) \cap H^1_0(\Omega)$$

be an appropriate finite element space. In practice, the Hsieh–Clough–Tocher element and the Argyris element can be used as H^2-conforming elements. Then, for $h > 0$

sufficiently small, standard finite element analysis can be used to show that there is a unique function $u_0^{h,H} \in V_H$ such that

$$\int_\Omega \left(A_h^0 : D^2 u_0^{h,H}\right)\left(A_h^0 : D^2\varphi_H\right) = \int_\Omega f\left(A_h^0 : D^2\varphi_H\right) \qquad \forall\, \varphi_H \in V_H, \tag{3.22}$$

and that the error bound (3.21) holds.

Further finite element approaches for approximating the solution of problems in nondivergence-form include the conforming method [72] that makes use of a finite element Hessian, the discontinuous *hp*-Galerkin method [87, 88], the primal method [44] similar to an interior penalty discontinuous Galerkin method, the mixed finite element method [51], and the variational formulations presented in [49].

3.3 Approximation of u_ε via correctors

We assume either that $(\Omega, A, f) \in \mathcal{G}^{2,2}$ or that $(\Omega, A, f) \in \mathcal{H}^2$. Let $n \in \{2,3\}$, $\varepsilon \in (0,1]$, and assume that

$$u_0 \in H^4(\Omega).$$

Then we know that $u_0 \in W^{2,\infty}(\Omega)$ and by Theorem 2.3.2 we have that

$$\left\| u_\varepsilon - u_0 - \varepsilon^2 \sum_{i,j=1}^n \chi_{ij}\left(\frac{\cdot}{\varepsilon}\right)\partial_{ij}^2 u_0 \right\|_{H^2(\Omega)} \lesssim \sqrt{\varepsilon}\,\|u_0\|_{W^{2,\infty}(\Omega)} + \varepsilon\|u_0\|_{H^4(\Omega)}, \tag{3.23}$$

where u_0 is the solution to the homogenized problem, and χ_{ij} are the corrector functions given as the solutions to the periodic cell problems

$$A : D^2\chi_{ij} = a_{ij}^0 - a_{ij} \quad \text{in } Y, \qquad \chi_{ij} \text{ is } Y\text{-periodic}, \qquad \int_Y \chi_{ij} = 0, \tag{3.24}$$

as introduced in (2.11).

3.3.1 Approximation of the corrector

We now address problem (3.24) and present a method for $A \in W^{2,\infty}(Y)$. To simplify the notation and the arguments, we assume that we know the invariant measure m and the matrix $A^0 = (a_{ij}^0)_{1\le i,j\le n}$ exactly instead of working with our approximation A_h^0.

For a given Y-periodic right-hand side $g \in W^{2,\infty}(Y)$, we consider the problem

$$\begin{cases} -\nabla\cdot(A\nabla\chi) + (\mathrm{div}A)\cdot\nabla\chi = -g & \text{in } Y,\\ \chi \text{ is } Y\text{-periodic}, \ \int_Y \chi = 0. \end{cases}$$

Obtaining an approximation for second-order derivatives via finite elements is not straightforward since the natural solution space is $W_{\text{per}}(Y)$. We present a method of successively approximating higher derivatives.

Let χ_h be a $W_{\text{per}}(Y)$-conforming finite element approximation to χ, i.e.,

$$\chi_h \in V_h, \quad \int_Y A\nabla\chi_h \cdot \nabla\varphi + \int_Y \varphi\,(\text{div}A)\cdot\nabla\chi_h = -\int_Y g\varphi \qquad \forall\,\varphi \in V_h,$$

with $V_h \subset W_{\text{per}}(Y)$ finite-dimensional, and satisfying the error estimate

$$\|\chi_h - \chi\|_{H^1(Y)} \lesssim h.$$

Let $r \in \{1,\dots,n\}$ and write $\xi_r := \partial_r\chi$. Then, using the equation

$$-\nabla\cdot(A\nabla\chi) + (\text{div}A)\cdot\nabla\chi = -g \quad \text{in } Y,$$

we find that in a weak sense, one has

$$-\nabla\cdot(A\nabla\xi_r) + (\text{div}A)\cdot\nabla\xi_r = -\partial_r g + \nabla\cdot(\partial_r A\,\nabla\chi) - (\text{div}(\partial_r A))\cdot\nabla\chi \quad \text{in } Y.$$

Further, we claim that $\xi_r \in W_{\text{per}}(Y)$. Indeed, this follows from the regularity and periodicity of χ and

$$\int_Y \partial_r\chi = \int_{\partial Y} \chi\nu\cdot e_r = 0.$$

Therefore, $\xi_r \in W_{\text{per}}(Y)$ satisfies

$$\begin{cases} -\nabla\cdot(A\nabla\xi_r) + (\text{div}A)\cdot\nabla\xi_r = -\partial_r g + \nabla\cdot(\partial_r A\,\nabla\chi) - (\text{div}(\partial_r A))\cdot\nabla\chi \quad \text{in } Y, \\ \xi_r \text{ is } Y\text{-periodic}, \ \int_Y \xi_r = 0. \end{cases}$$

Now we use our H^1-conforming approximation for χ for the right-hand side and use a $W_{\text{per}}(Y)$-conforming finite element method for approximating the solution $v \in W_{\text{per}}(Y)$ to the following problem:

$$\begin{cases} -\nabla\cdot(A\nabla v) + (\text{div}A)\cdot\nabla v = -\partial_r g + \nabla\cdot(\partial_r A\,\nabla\chi_h) - (\text{div}(\partial_r A))\cdot\nabla\chi_h - c, \\ v \text{ is } Y\text{-periodic}, \ \int_Y v = 0, \end{cases} \tag{3.25}$$

where c is such that this problem admits a unique solution (such that the solvability condition (2.7) is satisfied). By looking at the problem for $v - \xi_r$, one obtains the comparison result

$$\begin{aligned} \|v - \xi_r\|_{H^1(Y)} &\lesssim \|\nabla\cdot(\partial_r A\,\nabla(\chi_h - \chi))\|_{W_{\text{per}}(Y)'} + \|\,(\text{div}(\partial_r A))\cdot\nabla(\chi_h - \chi)\|_{W_{\text{per}}(Y)'} \\ &\lesssim \|A\|_{W^{2,\infty}(Y)}\|\chi_h - \chi\|_{H^1(Y)} \\ &\lesssim h\|A\|_{W^{2,\infty}(Y)} = \mathcal{O}(h). \end{aligned}$$

Let v_h be a $W_{\mathrm{per}}(Y)$-conforming finite element approximation to the solution v of (3.25) satisfying

$$\|v_h - v\|_{H^1(Y)} \leq Ch$$

for some constant $C = C(\|A\|_{W^{2,\infty}(Y)}) > 0$. Then, using the triangle inequality, we obtain

$$\|v_h - \xi_r\|_{H^1(Y)} \leq Ch$$

for some constant $C = C(\|A\|_{W^{2,\infty}(Y)}) > 0$. Using this procedure for $r = 1, \ldots, n$, we eventually obtain approximations to derivatives of order up to two of χ.

3.3.2 Approximation of u_ε

The H^2 corrector estimate (3.23) can be used to construct an approximation of u_ε, i.e., to the solution of problem (3.2) for small ε. We note that (3.23) implies that

$$\begin{aligned} &\|u_\varepsilon - u_0\|_{H^1(\Omega)} + \sum_{k,l=1}^{n} \left\| \partial^2_{kl} u_\varepsilon - \left(\partial^2_{kl} u_0 + \sum_{i,j=1}^{n} \left(\partial^2_{kl}\chi_{ij}\right)\left(\frac{\cdot}{\varepsilon}\right) \partial^2_{ij} u_0 \right) \right\|_{L^2(\Omega)} \\ &\quad \lesssim \sqrt{\varepsilon}\, \|u_0\|_{W^{2,\infty}(\Omega)} + \varepsilon \|u_0\|_{H^4(\Omega)}. \end{aligned} \tag{3.26}$$

This leads to the following approximation result for u_ε.

Theorem 3.3.1 ([29, Theorem 3.11] Approximation of u_ε). *In the situation described above, let $(u_{0,h})_{h>0} \subset H^2(\Omega)$ be a family of H^2-conforming approximations for u_0 satisfying the error bound*

$$\|u_0 - u_{0,h}\|_{H^2(\Omega)} \lesssim h\|f\|_{H^1(\Omega)},$$

and for $1 \leq i,j,k,l \leq n$, let $(z^{kl}_{ij,h})_{h>0} \subset L^2_{\mathrm{per}}(Y)$ be a family of L^2 approximations for $\partial^2_{kl}\chi_{ij}$ satisfying the error bound

$$\|\partial^2_{kl}\chi_{ij} - z^{kl}_{ij,h}\|_{L^2(Y)} \lesssim h.$$

Then, by writing

$$u^{kl}_{\varepsilon,h} := \partial^2_{kl} u_{0,h} + \sum_{i,j=1}^{n} z^{kl}_{ij,h}\left(\frac{\cdot}{\varepsilon}\right) \partial^2_{ij} u_{0,h},$$

we have that

$$\begin{aligned} &\|u_\varepsilon - u_{0,h}\|_{H^1(\Omega)} + \sum_{k,l=1}^{n} \left\|\partial^2_{kl} u_\varepsilon - u^{kl}_{\varepsilon,h}\right\|_{L^1(\Omega)} \\ &\quad \lesssim \left(\sqrt{\varepsilon} + h\right) \|u_0\|_{W^{2,\infty}(\Omega)} + \varepsilon\|u_0\|_{H^4(\Omega)} + h\|f\|_{H^1(\Omega)}. \end{aligned}$$

Proof. We use (3.26) and the triangle inequality to obtain

$$\begin{aligned}\|u_\varepsilon - u_{0,h}\|_{H^1(\Omega)} &\le \|u_\varepsilon - u_0\|_{H^1(\Omega)} + \|u_0 - u_{0,h}\|_{H^1(\Omega)} \\ &\lesssim \sqrt{\varepsilon}\,\|u_0\|_{W^{2,\infty}(\Omega)} + \varepsilon\|u_0\|_{H^4(\Omega)} + h\|f\|_{H^1(\Omega)},\end{aligned}$$

and for $1 \le k, l \le n$,

$$\begin{aligned}\left\|\partial^2_{kl} u_\varepsilon - u^{kl}_{\varepsilon,h}\right\|_{L^1(\Omega)} &\lesssim \sqrt{\varepsilon}\,\|u_0\|_{W^{2,\infty}(\Omega)} + \varepsilon\|u_0\|_{H^4(\Omega)} + h\|f\|_{H^1(\Omega)} \\ &\quad + \sum_{i,j=1}^{n} \left\|\left(\partial^2_{kl}\chi_{ij}\right)\left(\frac{\cdot}{\varepsilon}\right)\partial^2_{ij}u_0 - z^{kl}_{ij,h}\left(\frac{\cdot}{\varepsilon}\right)\partial^2_{ij}u_{0,h}\right\|_{L^1(\Omega)}.\end{aligned}$$

It remains to study the last term on the right-hand side of the above inequality. For fixed $1 \le i, j \le n$, we use again the triangle inequality to obtain

$$\begin{aligned}&\left\|\left(\partial^2_{kl}\chi_{ij}\right)\left(\frac{\cdot}{\varepsilon}\right)\partial^2_{ij}u_0 - z^{kl}_{ij,h}\left(\frac{\cdot}{\varepsilon}\right)\partial^2_{ij}u_{0,h}\right\|_{L^1(\Omega)} \\ &\le \left\|z^{kl}_{ij,h}\left(\frac{\cdot}{\varepsilon}\right)\left(\partial^2_{ij}u_0 - \partial^2_{ij}u_{0,h}\right)\right\|_{L^1(\Omega)} + \left\|\left(\partial^2_{kl}\chi_{ij} - z^{kl}_{ij,h}\right)\left(\frac{\cdot}{\varepsilon}\right)\partial^2_{ij}u_0\right\|_{L^1(\Omega)} \\ &\lesssim \left\|z^{kl}_{ij,h}\left(\frac{\cdot}{\varepsilon}\right)\right\|_{L^2(\Omega)}\|u_0 - u_{0,h}\|_{H^2(\Omega)} + \left\|\left(\partial^2_{kl}\chi_{ij} - z^{kl}_{ij,h}\right)\left(\frac{\cdot}{\varepsilon}\right)\right\|_{L^2(\Omega)}\|u_0\|_{W^{2,\infty}(\Omega)} \\ &\lesssim h\left(\left\|z^{kl}_{ij,h}\left(\frac{\cdot}{\varepsilon}\right)\right\|_{L^2(\Omega)}\|f\|_{H^1(\Omega)} + \|u_0\|_{W^{2,\infty}(\Omega)}\right).\end{aligned}$$

In the last step, we used that by the transformation formula and periodicity (cover Ω/ε by $\mathcal{O}(\varepsilon^{-n})$ many cells of unit length), there holds

$$\left\|\left(\partial^2_{kl}\chi_{ij} - z^{kl}_{ij,h}\right)\left(\frac{\cdot}{\varepsilon}\right)\right\|_{L^2(\Omega)} \lesssim \left\|\partial^2_{kl}\chi_{ij} - z^{kl}_{ij,h}\right\|_{L^2(Y)} \lesssim h. \tag{3.27}$$

We claim that

$$\left\|z^{kl}_{ij,h}\left(\frac{\cdot}{\varepsilon}\right)\right\|_{L^2(\Omega)} \lesssim h + 1.$$

Indeed, we use the triangle inequality, (3.27) and the fact that $\chi_{ij} \in W^{2,\infty}(Y)$ to obtain

$$\left\|z^{kl}_{ij,h}\left(\frac{\cdot}{\varepsilon}\right)\right\|_{L^2(\Omega)} \le \left\|\left(\partial^2_{kl}\chi_{ij} - z^{kl}_{ij,h}\right)\left(\frac{\cdot}{\varepsilon}\right)\right\|_{L^2(\Omega)} + \left\|\partial^2_{kl}\chi_{ij}\right\|_{L^\infty(Y)} \lesssim h + 1.$$

□

In connection with the previously described approximation of the homogenized solution u_0 and the corrector functions χ_{ij}, note that Theorem 3.2.3 provides an $H^2(\Omega)$-conforming approximation to u_0 and the method presented in Section 3.3.1

provides $L^2_{\mathrm{per}}(Y)$ approximations for the second-order partial derivatives of χ_{ij}, as required for the setting of Theorem 3.3.1.

Let us conclude this section by remarking that if the second derivatives of the corrector functions are approximated in the space $L^\infty(Y)$ or if the solution to the homogenized problem is approximated in the space $W^{2,\infty}(\Omega)$, then one obtains by a similar proof an approximation result for the second derivatives of u_ε in $L^2(\Omega)$.

Remark 3.3.1. *If $(z^{kl}_{ij,h})_{h>0} \subset L^\infty_{\mathrm{per}}(Y)$ is a family of L^∞ approximations for $\partial^2_{kl}\chi_{ij}$ satisfying the error bound*

$$\|\partial^2_{kl}\chi_{ij} - z^{kl}_{ij,h}\|_{L^\infty(Y)} = \mathcal{O}(h),$$

and $(u_{0,h})_{h>0}$ is as in Theorem 3.3.1, then we have that

$$\|u_\varepsilon - u_{0,h}\|_{H^1(\Omega)} + \sum_{k,l=1}^{n} \left\|\partial^2_{kl}u_\varepsilon - u^{kl}_{\varepsilon,h}\right\|_{L^2(\Omega)} = \mathcal{O}(\sqrt{\varepsilon} + h).$$

The same holds true when $(u_{0,h})_{h>0} \subset W^{2,\infty}(\Omega)$ is a family of $W^{2,\infty}$-conforming approximations for u_0 satisfying the error bound

$$\|u_0 - u_{0,h}\|_{W^{2,\infty}(\Omega)} = \mathcal{O}(h),$$

and $(z^{kl}_{ij,h})_{h>0}$ is as in Theorem 3.3.1.

3.4 Extension to nonuniformly oscillating coefficients

In this section, we discuss the case of nonuniformly oscillating coefficients, i.e., coefficients depending on x and $\frac{x}{\varepsilon}$. We consider the problem

$$\begin{cases} A\left(\cdot, \frac{\cdot}{\varepsilon}\right) : D^2 u_\varepsilon = f & \text{in } \Omega, \\ u_\varepsilon = 0 & \text{on } \partial\Omega, \end{cases} \tag{3.28}$$

where the triple (Ω, A, f) satisfies one of the following sets of assumptions.

Definition 3.4.1 (Sets of assumptions $\mathcal{G}, \mathcal{H}$)**.** *We write*

(i) $(\Omega, A, f) \in \mathcal{G}$ if and only if $\Omega \subset \mathbb{R}^n$ is a bounded $C^{2,\gamma}$ domain, $f \in L^2(\Omega)$, and $A : \Omega \times \mathbb{R}^n \to \mathbb{R}^{n\times n}_{\mathrm{sym}}$ satisfies

$$\begin{cases} A = A(x,y) \in W^{2,\infty}(\Omega; W^{1,q}(Y)) \text{ for some } q \in (n,\infty], \\ A(x,\cdot) \text{ is } Y\text{-periodic}, \\ \exists\, \lambda, \Lambda > 0 : \ \lambda|\xi|^2 \le A(x,y)\xi\cdot\xi \le \Lambda|\xi|^2 \quad \forall\, \xi, y \in \mathbb{R}^n,\, x \in \Omega. \end{cases} \tag{3.29}$$

(ii) $(\Omega, A, f) \in \mathcal{H}$ *if and only if* $\Omega \subset \mathbb{R}^n$ *is a bounded convex domain,* $f \in L^2(\Omega)$*, and* $A : \Omega \times \mathbb{R}^n \to \mathbb{R}^{n\times n}_{\mathrm{sym}}$ *satisfies* (3.29) *and*

$$\exists\, \delta \in (0,1] : \quad \frac{|A(x,y)|^2}{(\mathrm{tr} A(x,y))^2} \leq \frac{1}{n-1+\delta} \quad \forall\, (x,y) \in \Omega \times \mathbb{R}^n. \tag{3.30}$$

In view of Remark 2.1.1, we see that the Cordes condition (3.30) is always satisfied when $n = 2$. Well-posedness of the problem (3.28) is guaranteed by the following theorem; see [53, Theorem 9.15] and [87, Theorem 3].

Theorem 3.4.1 ([29, Theorem 3.14] Existence and uniqueness of strong solutions)**.** *Assume either that* $(\Omega, A, f) \in \mathcal{G}$*, or that* $(\Omega, A, f) \in \mathcal{H}$*. Then, for any* $\varepsilon > 0$*, the problem* (3.28) *admits a unique solution* $u_\varepsilon \in H^2(\Omega) \cap H^1_0(\Omega)$*.*

3.4.1 Homogenization results

As in Chapter 2, uniform *a priori* estimates for the solution to (3.28) allow passage to the limit in equation (3.28); see [19, 20]. The coefficient matrix of the homogenized problem now depends on the slow variable x, and is obtained by integrating against an invariant measure. Corrector results can then be shown as before.

Theorem 3.4.2 ([29, Theorem 3.15] Nonuniformly oscillating coefficients)**.** *Assume that* $\varepsilon \in (0,1]$ *and that either* $(\Omega, A, f) \in \mathcal{G}$ *or* $(\Omega, A, f) \in \mathcal{H}$*. Then, the following assertions hold:*

(i) Uniform a priori estimate: The solution $u_\varepsilon \in H^2(\Omega) \cap H^1_0(\Omega)$ *to* (3.28) *satisfies*

$$\|u_\varepsilon\|_{H^2(\Omega)} \lesssim \|f\|_{L^2(\Omega)}.$$

(ii) Homogenization: The solution $u_\varepsilon \in H^2(\Omega) \cap H^1_0(\Omega)$ *to* (3.28) *converges weakly in* $H^2(\Omega)$ *to the solution* $u_0 \in H^2(\Omega) \cap H^1_0(\Omega)$ *of the homogenized problem*

$$\begin{cases} A^0 : D^2 u_0 = f & \text{in } \Omega, \\ u_0 = 0 & \text{on } \partial\Omega, \end{cases} \tag{3.31}$$

with $A^0 : \Omega \to \mathbb{R}^{n\times n}$ *given by*

$$A^0(x) := \int_Y A(x,\cdot)\, m(x,\cdot),$$

where $m = m(x,y)$ *is the unique function* $m : \Omega \times \mathbb{R}^n \to \mathbb{R}$ *with* $m \in C(\bar{\Omega} \times \mathbb{R}^n)$*,* $0 < \bar{m} \leq m \leq M$ *for some constants* $\bar{m}, M > 0$*, such that*

$$\begin{cases} D^2 : (A(x,\cdot)\, m(x,\cdot)) = 0 & \text{in } Y, \\ m(x,\cdot) \text{ is } Y\text{-periodic}, \ \int_Y m(x,\cdot) = 1, \end{cases}$$

for any fixed $x \in \Omega$*. The function* m *is called the invariant measure.*

(iii) Corrector estimate: Assume that $f \in H^2(\Omega)$ and $u_0 \in H^4(\Omega) \cap W^{2,\infty}(\Omega)$. Introducing the corrector function χ_{ij}, $1 \leq i, j \leq n$, as the solution to

$$\begin{cases} A(x,y) : D_y^2 \chi_{ij}(x,y) = a_{ij}^0(x) - a_{ij}(x,y), & (x,y) \in \Omega \times Y, \\ \chi_{ij}(x,\cdot) \text{ is } Y\text{-periodic}, \ \int_Y \chi_{ij}(x,\cdot) = 0, \end{cases}$$

we have that

$$\left\| u_\varepsilon - u_0 - \varepsilon^2 \sum_{i,j=1}^n \chi_{ij}\left(\cdot, \frac{\cdot}{\varepsilon}\right) \partial_{ij}^2 u_0 \right\|_{H^2(\Omega)} \lesssim \sqrt{\varepsilon}\, \|u_0\|_{W^{2,\infty}(\Omega)} + \varepsilon \|u_0\|_{H^4(\Omega)}.$$

Proof. (i) For $(\Omega, A, f) \in \mathcal{H}$, one shows similarly to the proof of [87, Theorem 3] and Theorem 2.2.1 that

$$\|u_\varepsilon\|_{H^2(\Omega)} \lesssim \left\| \frac{\mathrm{tr} A\left(\cdot, \frac{\cdot}{\varepsilon}\right)}{|A\left(\cdot, \frac{\cdot}{\varepsilon}\right)|^2} \right\|_{L^\infty(\Omega)} \|f\|_{L^2(\Omega)} \lesssim \|f\|_{L^2(\Omega)}.$$

For $(\Omega, A, f) \in \mathcal{G}$, the claim follows from the method of freezing coefficients, using the uniform estimate from Theorem 2.2.1 for the operators $L_{x_0} := A\left(x_0, \frac{\cdot}{\varepsilon}\right) : D^2$ for fixed $x_0 \in \Omega$.

(ii) The uniform estimate from (i) yields weak convergence in $H^2(\Omega)$ and strong convergence in $H^1(\Omega)$ for a subsequence of $(u_\varepsilon)_{\varepsilon>0}$ to some limit function $u_0 \in H^2(\Omega) \cap H_0^1(\Omega)$. We multiply (3.28) by $m\left(\cdot, \frac{\cdot}{\varepsilon}\right)$ and follow the transformation performed in [19] to find that the equality

$$m^\varepsilon f = 2\nabla \cdot \left(\tilde{A}^\varepsilon \nabla u_\varepsilon + \left[\mathrm{div}_x \tilde{A}\right]^\varepsilon u_\varepsilon\right) - 2\left[\mathrm{div}_x \tilde{A}\right]^\varepsilon \cdot \nabla u_\varepsilon - \left[D_x^2 : \tilde{A}\right]^\varepsilon u_\varepsilon - D^2 : (\tilde{A}^\varepsilon u_\varepsilon)$$

holds weakly, where $\tilde{A} := Am$ and v^ε denotes $v\left(\cdot, \frac{\cdot}{\varepsilon}\right)$. Passing to the limit, we obtain that u_0 is a weak solution of (3.31). We conclude the proof by noting that (3.31) admits a unique strong solution, since A^0 is uniformly elliptic and Lipschitz continuous on $\bar{\Omega}$; see [53, 56].

(iii) This can be proved similarly to Theorem 2.3.1 and Theorem 2.3.2, using that, by the assumptions made on A and elliptic regularity, we have

$$\chi_{kl}^\varepsilon,\ [\partial_{x_i}\chi_{kl}]^\varepsilon,\ [\partial_{y_i}\chi_{kl}]^\varepsilon,\ \left[\partial_{x_i y_j}^2 \chi_{kl}\right]^\varepsilon,\ \left[\partial_{x_i x_j}^2 \chi_{kl}\right]^\varepsilon \in L^\infty(\Omega)$$

for any $1 \leq i, j, k, l \leq n$. □

3.4.2 Numerical scheme

Let us explain how the numerical scheme from Section 3.2 can be used for the numerical homogenization of (3.28).

First, we consider a shape-regular triangulation $\mathcal{T}_H$ on $\bar{\Omega}$ consisting of nodes $\{x_i\}_{i\in I}$ with grid size $H > 0$, and a shape-regular triangulation $\mathcal{T}_h$ on Y with grid size $h > 0$. Then, for any $i \in I$, we can use the scheme from Section 3.2 (see Theorem 3.2.1) to obtain an approximation $m_h^i \in H^1(Y)$ to $m_{x_i} = m(x_i, \cdot)$ such that

$$\|m_{x_i} - m_h^i\|_{L^2(Y)} + h\|m_{x_i} - m_h^i\|_{H^1(Y)} \lesssim h \inf_{\tilde{v}_h \in \tilde{M}_h} \|m_{x_i} - (\tilde{v}_h + 1)\|_{H^1(Y)}.$$

Further, we obtain that

$$A_h^{0,i} := \int_Y \mathcal{I}_h \left(A(x_i, \cdot)\, m_h^i \right)$$

is an approximation to $A^0(x_i)$ (see Lemma 3.2.1),

$$\left| A^0(x_i) - A_h^{0,i} \right| \lesssim h. \tag{3.32}$$

Now we define $A^0_{h,H}$ to be a continuous piecewise linear function on the triangulation $\mathcal{T}_H$ such that

$$A^0_{h,H}(x_i) = A_h^{0,i}$$

for all $i \in I$. Then, using (3.32) and denoting the continuous piecewise linear interpolant of a function ϕ on the triangulation $\mathcal{T}_H$ by $\mathcal{I}_H \phi$, we have

$$\begin{aligned} \|A^0 - A^0_{h,H}\|_{L^\infty(\Omega)} &\leq \|A^0 - \mathcal{I}_H A^0\|_{L^\infty(\Omega)} + \|\mathcal{I}_H A^0 - A^0_{h,H}\|_{L^\infty(\Omega)} \\ &\lesssim \|A^0 - \mathcal{I}_H A^0\|_{L^\infty(\Omega)} + h. \end{aligned} \tag{3.33}$$

We observe that, similarly to the proof of Lemma 3.2.2, we obtain that the solution $u_0^{h,H} \in H^2(\Omega) \cap H^1_0(\Omega)$ to

$$\begin{cases} A^0_{h,H} : D^2 u_0^{h,H} = f & \text{in } \Omega, \\ \qquad\qquad u_0^{h,H} = 0 & \text{on } \partial\Omega, \end{cases} \tag{3.34}$$

satisfies, for $h, H > 0$ sufficiently small,

$$\|u_0 - u_0^{h,H}\|_{H^2(\Omega)} \lesssim \|A^0 - A^0_{h,H}\|_{L^\infty(\Omega)} \|f\|_{L^2(\Omega)},$$

and in view of (3.33),

$$\|u_0 - u_0^{h,H}\|_{H^2(\Omega)} \lesssim \left(\|A^0 - \mathcal{I}_H A^0\|_{L^\infty(\Omega)} + h \right) \|f\|_{L^2(\Omega)} = \mathcal{O}(H^2 + h),$$

where u_0 is the solution to the homogenized problem (3.31). Here we have used the interpolation estimate

$$\|A^0 - \mathcal{I}_H A^0\|_{L^\infty(\Omega)} \lesssim H^2 \|A^0\|_{W^{2,\infty}(\Omega)},$$

which follows from $A^0 \in W^{2,\infty}(\Omega)$ and standard interpolation theory.

Remark 3.4.1. *For problems in divergence-form, similar results have been derived previously using heterogeneous multiscale methods; see e.g. [1].*

At this point, let us note that in contrast with our procedure of approximating the effective coefficient A^0 at the nodes of the coarse triangulation $\mathcal{T}_H$ and interpolating linearly, in the framework of the finite element heterogeneous multiscale method A^0 is typically approximated at the coarse integration nodes; see e.g. [1, 2]. The use of piecewise linear interpolation allows us to obtain second-order convergence. Assuming more regularity on the coefficient $A(x, y)$ in y, as in Remark 3.2.3, the error in the approximation of the homogenized solution u_0 can be improved to order $\mathcal{O}(H^2 + h^2)$. Finally, the solution to (3.34) can be approximated by a standard finite element method on the triangulation $\mathcal{T}_H$, which yields an approximation $u_{0,h,H} \in H^2(\Omega) \cap H_0^1(\Omega)$ to u_0 in the $H^2(\Omega)$-norm.

The approximation of u_ε can be obtained based on the corrector estimate from Theorem 3.4.2 analogously as in Section 3.3.2.

3.5 Numerical experiments

In this section, we illustrate the theoretical results through numerical experiments. We provide an example for the case of periodic coefficients in Section 3.5.1, and one for the case of nonuniformly oscillating coefficients in Section 3.5.2. In both cases, we provide not only an example with an unknown u_0, but also a set-up with a known u_0 in order to test the approximation scheme for the homogenized solution.

The experiments demonstrate the performance of the scheme for the approximation of the invariant measure m, the effective coefficient A^0, the homogenized solution u_0, as well as the approximation of the solution u_ε to the original problem for a fixed value of ε.

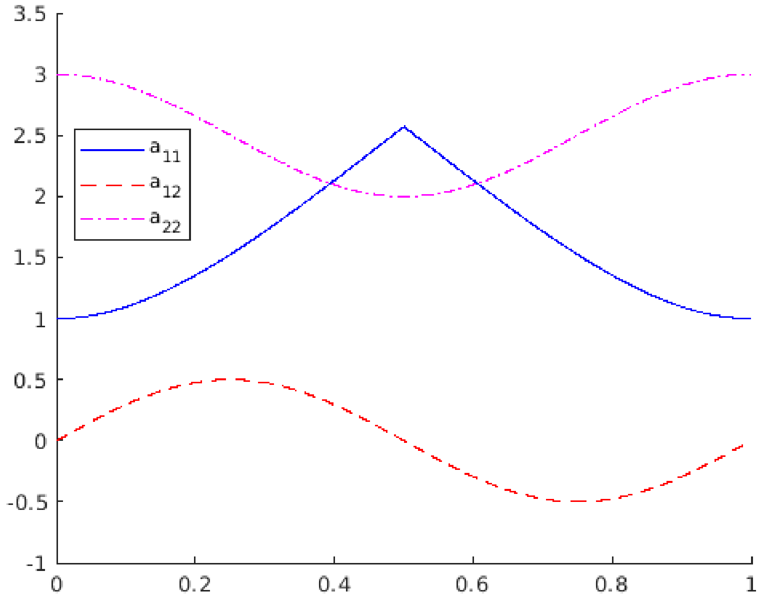


Figure 3.1: The functions $y_1 \mapsto a_{ij}(y_1)$ plotted on the interval $(0,1)$.

3.5.1 Periodic coefficients

We consider the homogenization problem

$$\begin{cases} A\left(\frac{\cdot}{\varepsilon}\right) : D^2 u_\varepsilon = f & \text{in } \Omega, \\ u_\varepsilon = 0 & \text{on } \partial\Omega, \end{cases} \tag{3.35}$$

on the domain

$$\Omega := Y = (0,1)^2,$$

with the matrix-valued map

$$A : \mathbb{R}^2 \to \mathbb{R}^{2\times 2}, \quad A(y_1, y_2) := \begin{pmatrix} 1 + \arcsin\left(\sin^2(\pi y_1)\right) & \sin(\pi y_1)\cos(\pi y_1) \\ \sin(\pi y_1)\cos(\pi y_1) & 2 + \cos^2(\pi y_1) \end{pmatrix},$$

and the right-hand side $f : \Omega \to \mathbb{R}$ to be specified below. We observe that the matrix-valued function A satisfies (3.1) with $q = \infty$. Further, note that

$$A(y) = (a_{ij}(y_1))_{1\le i,j\le 2}$$

depends only on the first coordinate of $y = (y_1, y_2) \in \mathbb{R}^2$; see Figure 3.1.

In this case we know that the homogenized problem is given by

$$\begin{cases} A^0 : D^2 u_0 = f & \text{in } \Omega, \\ u_0 = 0 & \text{on } \partial\Omega, \end{cases} \tag{3.36}$$

where $A^0 \in \mathbb{R}^{2\times 2}$ denotes the constant matrix $A^0 = \int_Y Am$ with m being the invariant measure

$$m : \mathbb{R}^2 \to \mathbb{R}, \quad m(y_1, y_2) = \left(\int_0^1 \frac{\mathrm{d}t}{a_{11}(t)} \right)^{-1} \frac{1}{a_{11}(y_1)};$$

see [48]. Explicit computation yields that

$$A^0 = \begin{pmatrix} \left(\int_0^1 \frac{\mathrm{d}t}{a_{11}(t)} \right)^{-1} & \left(\int_0^1 \frac{\mathrm{d}t}{a_{11}(t)} \right)^{-1} \int_0^1 \frac{a_{12}(t)}{a_{11}(t)} \,\mathrm{d}t \\ \left(\int_0^1 \frac{\mathrm{d}t}{a_{11}(t)} \right)^{-1} \int_0^1 \frac{a_{12}(t)}{a_{11}(t)} \,\mathrm{d}t & \left(\int_0^1 \frac{\mathrm{d}t}{a_{11}(t)} \right)^{-1} \int_0^1 \frac{a_{22}(t)}{a_{11}(t)} \,\mathrm{d}t \end{pmatrix} \approx \begin{pmatrix} 1.4684 & 0 \\ 0 & 2.6037 \end{pmatrix}.$$

We also note that for the corrector functions χ_{ij} $(1 \leq i, j \leq 2)$, i.e., the solutions to

$$\begin{cases} A : D^2 \chi_{ij} = a_{ij}^0 - a_{ij} & \text{in } Y, \\ \chi_{ij} \text{ is } Y\text{-periodic}, \ \int_Y \chi_{ij} = 0, \end{cases}$$

we have that

$$\partial_{kl}^2 \chi_{ij}(y_1, y_2) = \begin{cases} \dfrac{a_{ij}^0 - a_{ij}(y_1)}{a_{11}(y_1)} & \text{if } k = l = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Figure 3.2 shows the error in the approximation of m and A^0 by the scheme presented in Section 3.2.

For the approximation of the invariant measure we observe convergence of order

$$\|m - m_h\|_{L^2(Y)} = \mathcal{O}(h^{\frac{3}{2}}), \tag{3.37}$$

and superconvergence of order $\mathcal{O}(h^2)$ for $h > 0$ when grid points fall on the line $\{y_1 = \frac{1}{2}\}$, which is the set along which $\partial_1 m$ possesses a jump. The observed rate of convergence (3.37) is consistent with Theorem 3.2.1. Indeed, we have $m \in H^{\frac{3}{2}-\tilde{\varepsilon}}(Y)$ for any $\tilde{\varepsilon} > 0$, and Theorem 3.2.1 yields

$$\begin{aligned} \|m - m_h\|_{L^2(Y)} + h\|m - m_h\|_{H^1(Y)} &\lesssim h \inf_{\tilde{v}_h \in \tilde{M}_h} \|m - (\tilde{v}_h + 1)\|_{H^1(Y)} \\ &\lesssim h \left\| m - \mathcal{I}_h m - \int_Y (m - \mathcal{I}_h m) \right\|_{H^1(Y)} \\ &\lesssim h^{\frac{3}{2}-\tilde{\varepsilon}} \|m\|_{H^{\frac{3}{2}-\tilde{\varepsilon}}(Y)}, \end{aligned}$$

by making the choice $\tilde{v}_h = \mathcal{I}_h m - \int_Y \mathcal{I}_h m$, and using an interpolation error bound. In connection with the superconvergence we note that $m|_{(0,\frac{1}{2})\times(0,1)} \in H^2((0, \frac{1}{2}) \times (0, 1))$ and $m|_{(\frac{1}{2},1)\times(0,1)} \in H^2((\frac{1}{2}, 1) \times (0, 1))$. For the approximation of the matrix A^0, we observe second-order convergence.

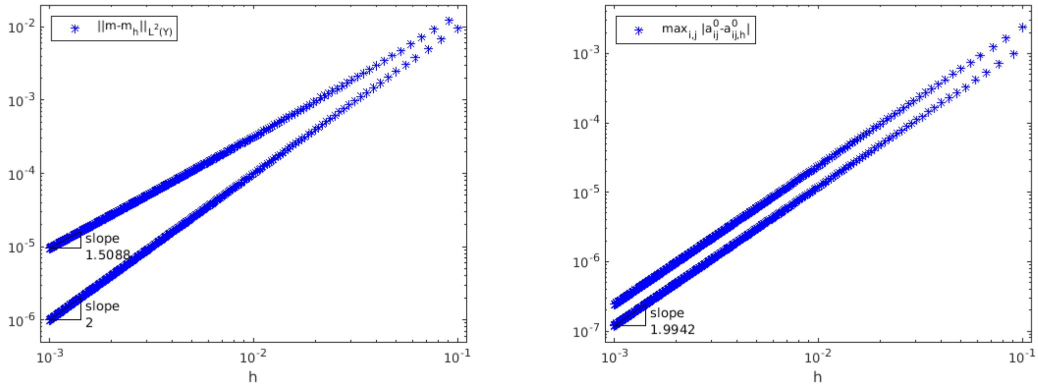


Figure 3.2: Approximation error for the invariant measure m (left) and the matrix A^0 (right). Two curves are observed, corresponding to whether or not grid points fall on the line $\{y_1 = \frac{1}{2}\}$, i.e., the set along which $\partial_1 m$ exhibits a jump.

Problem with a known u_0

We consider the right-hand side given by

$$f : \Omega \to \mathbb{R}, \quad f(x_1, x_2) := a^0_{22} x_1(x_1 - 1) + a^0_{11} x_2(x_2 - 1).$$

Then it is straightforward to check that the exact solution $u_0 \in H^2(\Omega) \cap H^1_0(\Omega)$ to the homogenized problem (3.36) is given by

$$u_0 : \Omega \to \mathbb{R}, \quad u_0(x_1, x_2) = \frac{1}{2} x_1(x_1 - 1) x_2(x_2 - 1).$$

Note that we are in the situation $(\Omega, A, f) \in \mathcal{H}^2$, that $f = 0$ at the corners of Ω and that $u_0 \in H^4(\Omega)$.

We use the scheme presented in Section 3.2 to approximate the homogenized solution u_0, where we use the same mesh for approximating m and u_0. The Hsieh–Clough–Tocher (HCT) element in FreeFem++ is used in the formulation (3.22) for the H^2 approximation of u_0; see [58]. The gradient on the boundary is set to be the gradient of an H^1 approximation by $\mathbb{P}_2$ elements on a fine mesh.

Concerning the approximation of u_ε, from Chapter 2 and Section 3.3.2 we obtain that

$$E_\varepsilon := \|u_\varepsilon - u_0\|^2_{H^1(\Omega)} + \sum_{k,l=1}^{2} \left\| \partial^2_{kl} u_\varepsilon - \left(\partial^2_{kl} u_0 + \sum_{i,j=1}^{2} \left(\partial^2_{kl}\chi_{ij}\right) \left(\frac{\cdot}{\varepsilon}\right) \partial^2_{ij} u_0 \right) \right\|^2_{L^2(\Omega)} = \mathcal{O}(\varepsilon).$$

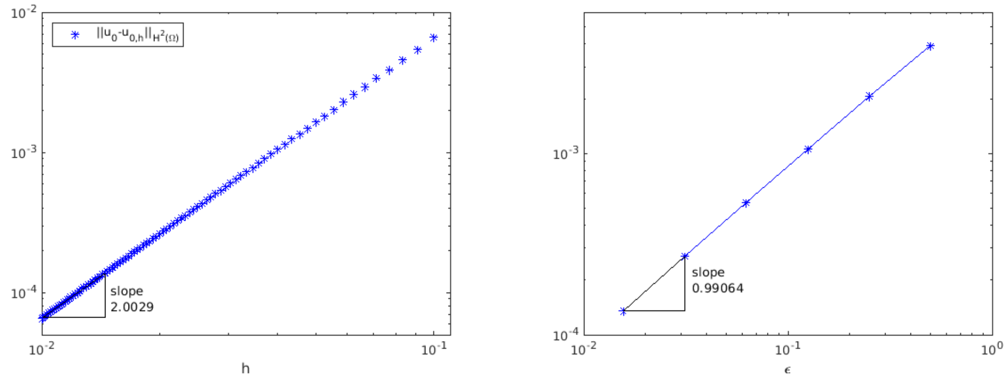


Figure 3.3: Approximation error for u_0 (left) and the error E_ε in the approximation of u_ε for different values of ε (right).

For the numerical approximation, we replace u_ε by an H^2-conforming finite element approximation on a fine mesh, based on the formulation

$$\text{Find } u_\varepsilon \in V: \quad \int_\Omega \frac{\operatorname{tr} A\left(\frac{\cdot}{\varepsilon}\right)}{|A\left(\frac{\cdot}{\varepsilon}\right)|^2} A\left(\frac{\cdot}{\varepsilon}\right) : D^2 u_\varepsilon \, \Delta v = \int_\Omega \frac{\operatorname{tr} A\left(\frac{\cdot}{\varepsilon}\right)}{|A\left(\frac{\cdot}{\varepsilon}\right)|^2} f \Delta v \qquad \forall\, v \in V,$$

where $V := H^2(\Omega) \cap H_0^1(\Omega)$. To this end, we use again the HCT element and set the gradient on the boundary to be the gradient of an H^1 approximation by $\mathbb{P}_2$ elements on a fine mesh.

Figure 3.3 shows the error in the approximation of u_0 and we observe second-order convergence. Further, with the exact u_0 being available, we can compute the error E_ε for different values of ε; see Figure 3.3. We observe first-order convergence as ε tends to zero, as expected.

Problem with an unknown u_0

Next, let us consider the problem (3.35) with the same domain Ω and matrix-valued function A as before, but with the right-hand side given by

$$f : \Omega \to \mathbb{R}, \quad f(x_1, x_2) := \exp\left(- \frac{1}{\frac{1}{2} - \left(x_1 - \frac{1}{2}\right)^2 - \left(x_2 - \frac{1}{2}\right)^2} \right).$$

Note that now we are in the situation $(\Omega, A, f) \in \mathcal{H}^2$. Further, since the right-hand side $f \in H^2(\Omega)$ of the homogenized problem (3.36) satisfies $f = 0$ at the corners of Ω, the solution u_0 to (3.36) belongs to the class $H^4(\Omega)$; see [60, Prop. 2.6].

As before, we use the scheme presented in Section 3.2 to approximate m, A^0 and u_0. Using the second-order $H^2(\Omega)$-conforming approximation $u_{0,h}$ to u_0 obtained as

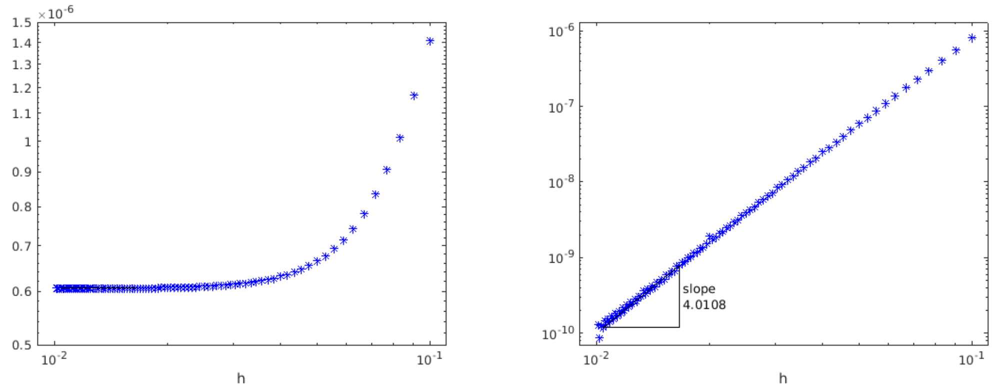


Figure 3.4: The error $E^h_{0.01}$ in the approximation of u_ε for a fixed value, $\varepsilon = \frac{1}{100}$, (left) and the error after subtraction of $6.0657 \cdot 10^{-7}$ (right), which is approximately the limit of $E^h_{0.01}$ in the figure on the left for this fixed value of ε as h tends to zero.

previously described,

$$\|u_0 - u_{0,h}\|_{H^2(\Omega)} = \mathcal{O}(h^2),$$

we have that

$$E^h_\varepsilon := \|u_\varepsilon - u_{0,h}\|^2_{H^1(\Omega)} + \sum_{k,l=1}^{2} \left\| \partial^2_{kl} u_\varepsilon - \left(\partial^2_{kl} u_{0,h} + \sum_{i,j=1}^{2} \left(\partial^2_{kl}\chi_{ij}\right)\left(\frac{\cdot}{\varepsilon}\right) \partial^2_{ij} u_{0,h} \right) \right\|^2_{L^2(\Omega)}$$
$$= \mathcal{O}(\varepsilon + h^4).$$

Figure 3.4 shows the error $E^h_{0.01}$ of the approximation of u_ε for different grid sizes and $\varepsilon = \frac{1}{100}$ fixed. We observe fourth-order convergence in h for the error as expected.

3.5.2 Nonuniformly oscillating coefficients

We consider the homogenization problem

$$\begin{cases} A\left(\cdot, \frac{\cdot}{\varepsilon}\right) : D^2 u_\varepsilon = f & \text{in } \Omega, \\ u_\varepsilon = 0 & \text{on } \partial\Omega, \end{cases} \tag{3.38}$$

on the domain

$$\Omega := Y = (0,1)^2,$$

with the matrix-valued map $A : \Omega \times \mathbb{R}^2 \to \mathbb{R}^{2\times 2}$,

$$((x_1,x_2),(y_1,y_2)) \mapsto \begin{pmatrix} \mathrm{e}^{x_1 x_2} + \frac{1}{4}|x|^2 \arcsin\left(\sin^2(\pi y_1)\right) & 0 \\ 0 & 2 + x_2\cos(2\pi y_2 + x_1) \end{pmatrix},$$

and the right-hand side $f : \Omega \to \mathbb{R}$ to be specified below. We observe that the matrix-valued function A satisfies (3.29) with $q = \infty$. Further, note that it is of the form

$$A(x,y) = \operatorname{diag}\left(a_{11}(x,y_1), a_{22}(x,y_2)\right).$$

In this case we know that the homogenized problem is given by

$$\begin{cases} A^0 : D^2 u_0 = f & \text{in } \Omega, \\ u_0 = 0 & \text{on } \partial\Omega, \end{cases} \tag{3.39}$$

where $A^0 : \Omega \to \mathbb{R}^{2\times 2}$ is given by

$$A^0(x) = \int_Y A(x,\cdot)\, m(x,\cdot),$$

with m being the invariant measure

$$m : \Omega \times \mathbb{R}^2 \to \mathbb{R}, \quad m(x,y) = \left(\int_0^1 \int_0^1 \frac{\mathrm{d}s\, \mathrm{d}t}{a_{11}(x,s)\, a_{22}(x,t)} \right)^{-1} \frac{1}{a_{11}(x,y_1)\, a_{22}(x,y_2)};$$

see [48]. Therefore, we have

$$a^0_{ij}(x) = \delta_{ij} \left(\int_0^1 \frac{\mathrm{d}t}{a_{ij}(x,t)} \right)^{-1}, \quad 1 \leq i,j \leq 2.$$

We also note that for the corrector functions χ_{ij} ($1 \leq i,j \leq 2$), i.e., the solutions to

$$\begin{cases} A(x,y) : D^2_y \chi_{ij}(x,y) = a^0_{ij}(x) - a_{ij}(x,y), & (x,y) \in \Omega \times Y, \\ \chi_{ij}(x,\cdot) \text{ is } Y\text{-periodic}, \ \int_Y \chi_{ij}(x,\cdot) = 0, \end{cases}$$

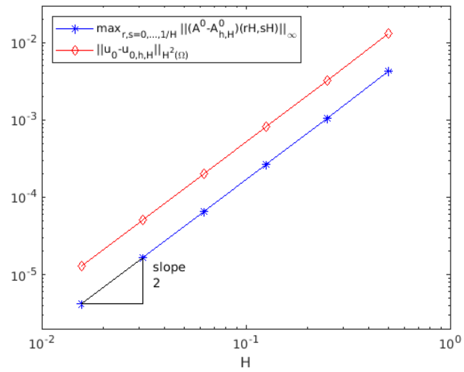
we have that

$$\partial^2_{y_k y_l} \chi_{ij}(x,y) = \begin{cases} \dfrac{a^0_{11}(x) - a_{11}(x,y_1)}{a_{11}(x,y_1)} & \text{if } i=j=k=l=1, \\ \dfrac{a^0_{22}(x) - a_{22}(x,y_2)}{a_{22}(x,y_2)} & \text{if } i=j=k=l=2, \\ 0 & \text{otherwise.} \end{cases}$$

Problem with a known u_0

We consider the right-hand side given by

$$f : \Omega \to \mathbb{R}, \quad x = (x_1,x_2) \mapsto f(x) := a^0_{22}(x)\, x_1(x_1 - 1) + a^0_{11}(x)\, x_2(x_2 - 1).$$



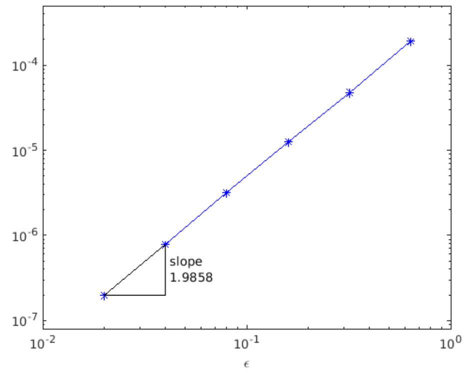


Figure 3.5: Approximation error for A^0 and u_0 for different values of H, using $h = \frac{H}{4}$, (left) and the error E_ε in the approximation of u_ε for different values of ε (right).

Then it is straightforward to check that the exact solution $u_0 \in H^2(\Omega) \cap H_0^1(\Omega)$ to the homogenized problem (3.39) is given by

$$u_0 : \Omega \to \mathbb{R}, \quad u_0(x_1, x_2) = \frac{1}{2} x_1(x_1 - 1) x_2(x_2 - 1).$$

Note that the assumptions of Theorem 3.4.2 (iii) are satisfied.

For $H > 0$ such that $\frac{1}{H} \in \mathbb{N}$, we take a triangulation $\mathcal{T}_H$ on $\bar{\Omega}$ consisting of nodes $\{(sH, rH)\}_{s,r=0,\dots,1/H}$, and a triangulation $\mathcal{T}_h$ on Y with grid size $h = \frac{H}{4}$. We use the scheme presented in Section 3.4 to approximate A^0 and u_0, and we observe second-order convergence; see Figure 3.5.

For the approximation of u_ε, Theorem 3.4.2 yields

$$E_\varepsilon := \|u_\varepsilon - u_0\|_{H^1(\Omega)}^2 + \sum_{k,l=1}^{2} \left\| \partial_{kl}^2 u_\varepsilon - \left(\partial_{kl}^2 u_0 + \sum_{i,j=1}^{2} \left(\partial_{y_k y_l}^2 \chi_{ij} \right) \left(\cdot, \frac{\cdot}{\varepsilon} \right) \partial_{ij}^2 u_0 \right) \right\|_{L^2(\Omega)}^2$$
$$= \mathcal{O}(\varepsilon).$$

For the numerical approximation, we replace u_ε by an H^2-conforming finite element method on a fine mesh, based on the formulation

$$\text{Find } u_\varepsilon \in V : \quad \int_\Omega \frac{\operatorname{tr} A\left(\cdot, \frac{\cdot}{\varepsilon}\right)}{|A\left(\cdot, \frac{\cdot}{\varepsilon}\right)|^2} A\left(\cdot, \frac{\cdot}{\varepsilon}\right) : D^2 u_\varepsilon \, \Delta v = \int_\Omega \frac{\operatorname{tr} A\left(\cdot, \frac{\cdot}{\varepsilon}\right)}{|A\left(\cdot, \frac{\cdot}{\varepsilon}\right)|^2} f \Delta v \qquad \forall v \in V,$$

where $V := H^2(\Omega) \cap H_0^1(\Omega)$. To this end, we use again the HCT element and set the gradient on the boundary to be the gradient of an H^1-conforming approximation by $\mathbb{P}_2$ elements on a fine mesh.

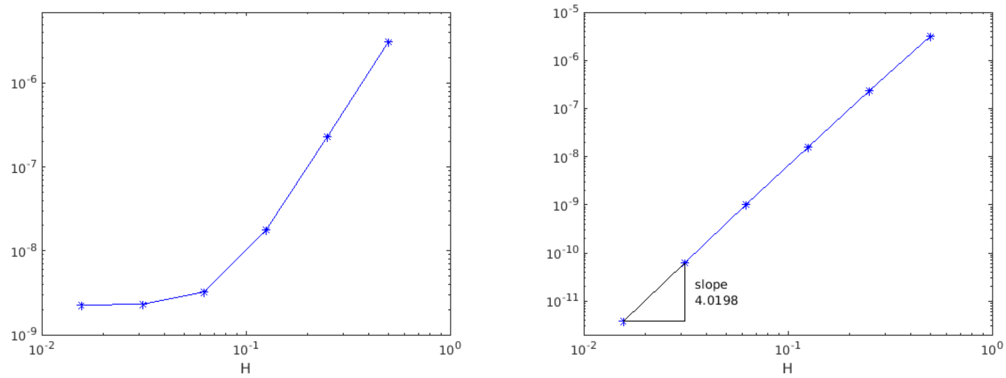


Figure 3.6: The error $E^H_{0.02}$ in the approximation of u_ε for a fixed value, $\varepsilon = \frac{1}{50}$, (left) and the error after subtraction of $2.2653 \cdot 10^{-9}$ (right), which is approximately the limit of $E^H_{0.02}$ in the figure on the left for this fixed value of ε as H tends to zero.

Problem with an unknown u_0

Finally, let us consider the problem (3.38) with the same domain Ω and matrix-valued function A as before, but with the right-hand side given by

$$f : \Omega \to \mathbb{R}, \quad f(x_1, x_2) := \exp\left(-\frac{1}{\frac{1}{2} - \left(x_1 - \frac{1}{2}\right)^2 - \left(x_2 - \frac{1}{2}\right)^2}\right).$$

Note that we are in the situation $(\Omega, A, f) \in \mathcal{H}$. Using the second-order H^2-conforming approximation $u_{0,H}$ to u_0 obtained as previously described (again with $h = \frac{H}{4}$),

$$\|u_0 - u_{0,H}\|_{H^2(\Omega)} = \mathcal{O}(H^2),$$

we have that

$$E^H_\varepsilon := \|u_\varepsilon - u_{0,H}\|^2_{H^1(\Omega)} + \sum_{k,l=1}^{2} \left\| \partial^2_{kl} u_\varepsilon - \partial^2_{kl} u_{0,H} - \sum_{i,j=1}^{2} \left(\partial^2_{y_k y_l} \chi_{ij}\right)\left(\cdot, \frac{\cdot}{\varepsilon}\right) \partial^2_{ij} u_{0,H} \right\|^2_{L^2(\Omega)}$$
$$= \mathcal{O}(\varepsilon + H^4).$$

Figure 3.6 shows the error $E^H_{0.02}$ of the approximation of u_ε for different grid sizes and $\varepsilon = \frac{1}{50}$ fixed. We observe fourth-order convergence in H for the error as expected.

Part II

Nonlinear elliptic equations in nondivergence-form

Chapter 4

Numerical homogenization of Hamilton–Jacobi–Bellman equations

This chapter discusses the numerical homogenization of Hamilton–Jacobi–Bellman (HJB) equations based on a mixed finite element method for the approximate corrector problems and is structured as follows:

We consider periodic HJB cell problems in Section 4.1 and prove the existence and uniqueness of a periodic strong solution in a suitable Cordes framework. These periodic cell problems arise naturally in the homogenization of HJB equations.

In Section 4.2, we propose and rigorously analyze a mixed finite element method for the approximation of the periodic solution to the HJB equation (1.9). We prove *a priori* (see Theorem 4.2.2) as well as *a posteriori* (see Theorem 4.2.3, Remark 4.2.3) error bounds with explicit error constants.

In Section 4.3, we discuss the numerical homogenization of problems of the form (1.7). We provide the framework and theoretical homogenization results in Sections 4.3.1 and 4.3.2 respectively. We then analyze the approximation of the approximate corrector (1.8) by the mixed finite element scheme from Section 4.2, and present a scheme for the approximation of the effective Hamiltonian in Sections 4.3.3 and 4.3.4 respectively.

In Section 4.4, we present numerical experiments for the approximate corrector problems and the homogenized effective equation.

Annotation: Unless stated otherwise, this chapter contains novel results which have been obtained in Gallistl, Sprekeler, Süli [50] (see also Kawecki, Sprekeler [69] for Section 4.1.2). D. Gallistl implemented the numerical homogenization scheme and generated the numerical data for our numerical experiments. The contribution of E.

Süli was of advisory nature. I would like to thank Y. Capdeboscq for some very useful discussions.

4.1 Periodic HJB cell problems

4.1.1 Setting

We let Λ be a compact metric space and write $Y := (0,1)^n$ for the unit cell in $\mathbb{R}^n$. We work in dimension $n \in \{2,3\}$. In order to simplify the presentation in this chapter, we use the notation

$$\varphi^\alpha(y) := \varphi(y,\alpha), \qquad (y,\alpha) \in \mathbb{R}^n \times \Lambda$$

for functions $\varphi : \mathbb{R}^n \times \Lambda \to \mathcal{R}$ with $\mathcal{R} \in \{\mathbb{R}, \mathbb{R}^n, \mathbb{R}^{n\times n}_{\mathrm{sym}}\}$. We study the periodic Hamilton–Jacobi–Bellman (HJB) problem

$$\begin{cases} F[u] := \sup_{\alpha\in\Lambda} \left\{ -A^\alpha : D^2u - b^\alpha \cdot \nabla u + c^\alpha u - f^\alpha \right\} = 0 \quad \text{in } Y, \\ \hfill u \text{ is } Y\text{-periodic,} \end{cases} \tag{4.1}$$

with given uniformly continuous coefficient functions

$$A : \mathbb{R}^n \times \Lambda \to \mathbb{R}^{n\times n}_{\mathrm{sym}}, \quad b : \mathbb{R}^n \times \Lambda \to \mathbb{R}^n, \quad c, f : \mathbb{R}^n \times \Lambda \to \mathbb{R}.$$

We assume that $A^\alpha, b^\alpha, c^\alpha, f^\alpha$ are Y-periodic in $\mathbb{R}^n$ for fixed $\alpha \in \Lambda$ and that $c > 0$ in $\mathbb{R}^n \times \Lambda$. Finally, we assume uniform ellipticity, i.e.,

$$\exists\, \zeta_1, \zeta_2 > 0 : \qquad \zeta_1 |\xi|^2 \leq A^\alpha(y)\xi \cdot \xi \leq \zeta_2 |\xi|^2 \quad \forall y, \xi \in \mathbb{R}^n, \alpha \in \Lambda,$$

and that the (generalized) Cordes condition (see [88]) is satisfied, that is,

$$\frac{|A|^2 + \frac{|b|^2}{2\lambda} + \frac{c^2}{\lambda^2}}{\left(\mathrm{tr}(A) + \frac{c}{\lambda}\right)^2} \leq \frac{1}{n+\delta} \tag{4.2}$$

holds in $\mathbb{R}^n \times \Lambda$ for some constants $\lambda > 0$ and $\delta \in (0,1)$. Let us point out the connection of this Cordes condition to the condition (2.2).

Remark 4.1.1. *Note that the Cordes condition* (4.2) *is equivalent to*

$$\frac{|\tilde{A}|^2}{(\mathrm{tr}\tilde{A})^2} \leq \frac{1}{\tilde{n} - 1 + \delta}$$

with $\tilde{n} := n+1$ *and* $\tilde{A} : \mathbb{R}^n \times \Lambda \to \mathbb{R}^{\tilde{n}\times\tilde{n}}_{\mathrm{sym}}$ *given by*

$$\tilde{A} := \left(\begin{array}{c|c} A & \frac{b}{2\sqrt{\lambda}} \\ \hline \frac{b^{\mathrm{T}}}{2\sqrt{\lambda}} & \frac{c}{\lambda} \end{array} \right).$$

The Cordes condition arises naturally for stochastic control problems as can be seen from the following example, which is taken from [88, Example 1].

Remark 4.1.2 (Cordes condition in practice; an explicit example taken from [88]). *Let us consider control variables $(\theta, Q) \in [0, \frac{\pi}{3}] \times \mathrm{SO}(3) =: \Lambda$ corresponding to angle and orientation between independent Wiener diffusions along directions Ge_i, $i \in \{1,2\}$, with*

$$G : \Lambda \to \mathbb{R}^{2\times 2}, \qquad (\theta, Q) \mapsto Q^{\mathrm{T}} \begin{pmatrix} 1 & \sin(\theta) \\ 0 & \cos(\theta) \end{pmatrix},$$

and look at the coefficient $A : \Lambda \to \mathbb{R}^{2\times 2}_{\mathrm{sym}}$ defined via $A := \frac{1}{2} G G^{\mathrm{T}}$. Typically, one has a constant zeroth-order coefficient $c \equiv \mathrm{constant} =: \bar{c} > 0$ for these stochastic control problems. Then, if $b \equiv 0$, we have that

$$\max_{(\theta,Q)\in\Lambda} \frac{|A(\theta,Q)|^2 + \frac{c^2}{\lambda^2}}{\left(\mathrm{tr}(A(\theta,Q)) + \frac{c}{\lambda}\right)^2} = \max_{(\theta,Q)\in\Lambda} \frac{\frac{1}{2}(1+\sin^2(\theta)) + \left(\frac{\bar{c}}{\lambda}\right)^2}{(1+\frac{\bar{c}}{\lambda})^2} = \frac{\frac{7}{8} + \left(\frac{\bar{c}}{\lambda}\right)^2}{(1+\frac{\bar{c}}{\lambda})^2}$$

is minimized at $\lambda = \frac{8}{7}\bar{c}$ with minimum value $\frac{7}{15} = (2 + \frac{1}{7})^{-1}$ and therefore, the Cordes condition (4.2) holds for any $\delta \in (0, \frac{1}{7}]$. For a non-vanishing first-order coefficient $b : \Lambda \to \mathbb{R}^2$, the Cordes condition (4.2) will still be satisfied if $\frac{|b|^2}{\bar{c}}$ is sufficiently small, which is essentially a coercivity assumption.

4.1.2 Existence and uniqueness of periodic strong solutions

In this section, we show that the periodic HJB problem (4.1) is well-posed in the sense that there exists a unique periodic strong solution, i.e., a unique function $u \in H^2_{\mathrm{per}}(Y)$ satisfying $F[u] = 0$ almost everywhere in Y. Recall that the space $H^2_{\mathrm{per}}(Y) \subset H^2(Y)$ is defined as the closure of

$$C^\infty_{\mathrm{per}}(Y) := \{ v|_Y : v \in C^\infty(\mathbb{R}^n) \text{ is } Y\text{-periodic}\}$$

with respect to the H^2-norm.

The renormalized problem

Let us introduce the function $\gamma = \gamma(y,\alpha) \in C(\mathbb{R}^n \times \Lambda)$ defined by

$$\gamma := \frac{\mathrm{tr}(A) + \frac{c}{\lambda}}{|A|^2 + \frac{|b|^2}{2\lambda} + \frac{c^2}{\lambda^2}} \tag{4.3}$$

and note that, by the assumptions on the coefficients from Section 4.1.1, we have

$$\inf_{\mathbb{R}^n\times\Lambda}\gamma > 0. \tag{4.4}$$

We then consider the renormalized HJB equation

$$\begin{cases} F_\gamma[u] := \sup_{\alpha\in\Lambda}\left\{\gamma^\alpha\left(-A^\alpha : D^2u - b^\alpha\cdot\nabla u + c^\alpha u - f^\alpha\right)\right\} = 0 \quad \text{in } Y, \\ \hfill u \text{ is } Y\text{-periodic.} \end{cases} \tag{4.5}$$

It is easily checked that the renormalized problem (4.5) is equivalent to the original problem (4.1) in the sense that they have the same set of periodic strong solutions. More precisely, we can characterize strong solutions to (4.1) as follows:

Remark 4.1.3. *For $u\in H^2_{\mathrm{per}}(Y)$, the following assertions are equivalent:*

(i) $F[u]=0$ *a.e. in Y, i.e., u is a periodic strong solution to the HJB problem* (4.1)*.*

(ii) $F_\gamma[u]=0$ *a.e. in Y, i.e., u is a periodic strong solution to the renormalized problem* (4.5)*.*

(iii) There holds

$$\int_Y F_\gamma[u]L_\lambda v = 0 \qquad \forall v\in H^2_{\mathrm{per}}(Y),$$

where $L_\lambda v := \lambda v - \Delta v$ for functions $v\in H^2_{\mathrm{per}}(Y)$.

Indeed, the equivalence $(i)\Leftrightarrow(ii)$ follows from (4.4) and the compactness of the metric space Λ, and $(ii)\Leftrightarrow(iii)$ is a consequence of the surjectivity of the linear differential operator

$$L_\lambda : H^2_{\mathrm{per}}(Y)\to L^2(Y), \qquad L_\lambda v := \lambda v - \Delta v.$$

Consequences of the Cordes condition

We point out a crucial Lipschitz-type estimate for the nonlinear operator F_γ. This is a direct consequence of the Cordes condition (4.2) and can be found in [67, 88]. A short proof is provided for demonstrating how the Cordes condition comes into play.

Lemma 4.1.1. *Let $\omega\subset Y$ be an open set. For any $u_1,u_2\in H^2(\omega)$, writing $\delta_u := u_1-u_2$, we have that*

$$|F_\gamma[u_1]-F_\gamma[u_2]-L_\lambda\delta_u| \le \sqrt{1-\delta}\sqrt{|D^2\delta_u|^2+2\lambda|\nabla\delta_u|^2+\lambda^2\delta_u^2} \tag{4.6}$$

almost everywhere in ω.

Proof. Let $u_1, u_2 \in H^2(\omega)$ and set $\delta_u := u_1 - u_2$. Note that for any bounded sets $\{x^\alpha\}_{\alpha\in\Lambda} \subset \mathbb{R}$ and $\{y^\alpha\}_{\alpha\in\Lambda} \subset \mathbb{R}$ we have that

$$|\sup_{\alpha\in\Lambda} x^\alpha - \sup_{\alpha\in\Lambda} y^\alpha| \leq \sup_{\alpha\in\Lambda}|x^\alpha - y^\alpha|.$$

This yields

$$\begin{aligned}
&|F_\gamma[u_1] - F_\gamma[u_2] - L_\lambda\delta_u|^2 \\
&\leq \sup_{\alpha\in\Lambda}\left|\gamma^\alpha\left(-A^\alpha : D^2\delta_u - b^\alpha\cdot\nabla\delta_u + c^\alpha\delta_u\right) + \Delta\delta_u - \lambda\delta_u\right|^2 \\
&\leq \left(|D^2\delta_u|^2 + 2\lambda\,|\nabla\delta_u|^2 + \lambda^2\delta_u^2\right)\sup_{\alpha\in\Lambda}\left\{|-\gamma^\alpha A^\alpha + I|^2 + \frac{|\gamma^\alpha b^\alpha|^2}{2\lambda} + \frac{|\gamma^\alpha c^\alpha - \lambda|^2}{\lambda^2}\right\} \\
&= \left(|D^2\delta_u|^2 + 2\lambda\,|\nabla\delta_u|^2 + \lambda^2\delta_u^2\right)\sup_{\alpha\in\Lambda}\left\{n + 1 - \frac{\left(\mathrm{tr}(A^\alpha) + \frac{c^\alpha}{\lambda}\right)^2}{|A^\alpha|^2 + \frac{|b^\alpha|^2}{2\lambda} + \frac{|c^\alpha|^2}{\lambda^2}}\right\} \\
&\leq (1-\delta)\left(|D^2\delta_u|^2 + 2\lambda\,|\nabla\delta_u|^2 + \lambda^2\delta_u^2\right)
\end{aligned}$$

almost everywhere in ω, where we used the Cauchy–Schwarz inequality, calculation and the Cordes condition (4.2). □

Observe that, using the triangle and Cauchy-Schwarz inequalities, we can eliminate the term $L_\lambda\delta_u$ from the left-hand side of (4.6). We thus find that, in the situation of Lemma 4.1.1, we have

$$|F_\gamma[u_1] - F_\gamma[u_2]| \leq \left(\sqrt{1-\delta} + \sqrt{n+1}\right)\sqrt{|D^2\delta_u|^2 + 2\lambda|\nabla\delta_u|^2 + \lambda^2\delta_u^2} \tag{4.7}$$

almost everywhere in ω.

Existence and uniqueness of solutions

We are now in a position to prove the existence and uniqueness of periodic strong solutions to the HJB problem (4.1). In view of Remark 4.1.3, let us define

$$B : H^2_{\mathrm{per}}(Y) \times H^2_{\mathrm{per}}(Y) \to \mathbb{R}, \qquad B(u,v) := \int_Y F_\gamma[u]L_\lambda v.$$

Let us note that integration by parts and a density argument yields that

$$\|\Delta v\|_{L^2(Y)} = \|D^2 v\|_{L^2(Y)} \qquad \forall v \in H^2_{\mathrm{per}}(Y)$$

and we have

$$\begin{aligned}
\forall v \in H^2_{\mathrm{per}}(Y) : \quad \|L_\lambda v\|^2_{L^2(Y)} &= \|D^2 v\|^2_{L^2(Y)} + 2\lambda\|\nabla v\|^2_{L^2(Y)} + \lambda^2\|v\|^2_{L^2(Y)} \\
&\geq C_\lambda\|v\|^2_{H^2(Y)}.
\end{aligned} \tag{4.8}$$

We can now proceed similarly to [67, 88] in showing that the Browder–Minty theorem applies and we obtain the following theorem:

Theorem 4.1.1 ([50, Theorem 2.1] Well-posedness)**.** *In the situation of Section 4.1.1, there exists a unique periodic strong solution* $u \in H^2_{\mathrm{per}}(Y)$ *to the HJB problem* (4.1)*.*

Proof. Note that it is enough to show that B satisfies the Lipschitz property

$$|B(u_1, v) - B(u_2, v)| \lesssim \|u_1 - u_2\|_{H^2(Y)} \|v\|_{H^2(Y)} \qquad \forall u_1, u_2, v \in H^2_{\mathrm{per}}(Y), \tag{4.9}$$

and strong monotonicity, i.e.,

$$\|u_1 - u_2\|^2_{H^2(Y)} \lesssim B(u_1, u_1 - u_2) - B(u_2, u_1 - u_2) \qquad \forall u_1, u_2 \in H^2_{\mathrm{per}}(Y). \tag{4.10}$$

The Browder–Minty theorem then yields that there exists a unique $u \in H^2_{\mathrm{per}}(Y)$ such that

$$B(u, v) = 0 \qquad \forall v \in H^2_{\mathrm{per}}(Y),$$

which proves the theorem in view of Remark 4.1.3.

The Lipschitz property (4.9) now immediately follows from (4.7) and it remains to show strong monotonicity. To this end, let $u_1, u_2 \in H^2_{\mathrm{per}}(Y)$ and write $\delta_u := u_1 - u_2$. Using Lemma 4.1.1, we find

$$\begin{aligned} B(u_1, \delta_u) - B(u_2, \delta_u) &= \|L_\lambda \delta_u\|^2_{L^2(Y)} + \int_Y (F_\gamma[u_1] - F_\gamma[u_2] - L_\lambda \delta_u)\, L_\lambda \delta_u \\ &\geq (1 - \sqrt{1-\delta}) \|L_\lambda \delta_u\|^2_{L^2(Y)} \end{aligned}$$

and hence, by (4.8), there holds (4.10) and the claim is proved. □

Remark 4.1.4. *For the unique periodic strong solution* $u \in H^2_{\mathrm{per}}(Y)$ *to the HJB problem* (4.1)*, we have the bound*

$$\|L_\lambda u\|_{L^2(Y)} = \left(\|D^2 u\|^2_{L^2(Y)} + 2\lambda \|\nabla u\|^2_{L^2(Y)} + \lambda^2 \|u\|^2_{L^2(Y)} \right)^{\frac{1}{2}} \leq \frac{\|F_\gamma[0]\|_{L^2(Y)}}{1 - \sqrt{1-\delta}}.$$

Proof. Note that we have already obtained the first equality (see (4.8)). We use Lemma 4.1.1 and the solution property $F_\gamma[u] = 0$ to find

$$\begin{aligned} (1 - \sqrt{1-\delta}) \|L_\lambda u\|^2_{L^2(Y)} &\leq \|L_\lambda u\|^2_{L^2(Y)} + \int_Y (F_\gamma[u] - F_\gamma[0] - L_\lambda u)\, L_\lambda u \\ &= - \int_Y F_\gamma[0] L_\lambda u. \end{aligned}$$

We conclude the proof by using Hölder's inequality to obtain

$$\|L_\lambda u\|^2_{L^2(Y)} \leq - \frac{1}{1 - \sqrt{1-\delta}} \int_Y F_\gamma[0] L_\lambda u \leq \frac{\|F_\gamma[0]\|_{L^2(Y)}}{1 - \sqrt{1-\delta}} \|L_\lambda u\|_{L^2(Y)},$$

which yields the desired bound. □

4.2 Mixed FEM for periodic HJB cell problems

4.2.1 Mixed formulation

We construct a mixed finite element method for the numerical approximation of the strong periodic solution to (4.1) similarly to the scheme presented in [51]. The mixed formulation relies on rewriting the problem (4.5) as

$$\sup_{\alpha\in\Lambda}\{\gamma^\alpha(-A^\alpha : Dw - b^\alpha\cdot\nabla u + c^\alpha u - f^\alpha)\} = 0 \tag{4.11}$$

with the coupling

$$\nabla u - w = 0.$$

Let us recall the notation

$$W_{\mathrm{per}}(Y) := \left\{ v\in H^1_{\mathrm{per}}(Y) : \int_Y v = 0\right\},\qquad W_{\mathrm{per}}(Y;\mathbb{R}^n) := (W_{\mathrm{per}}(Y))^n.$$

Noting that a solution $u\in H^2_{\mathrm{per}}(Y)$ to (4.11) satisfies $w=\nabla u\in W_{\mathrm{per}}(Y;\mathbb{R}^n)$, we define the function space

$$X := W_{\mathrm{per}}(Y;\mathbb{R}^n)\times H^1_{\mathrm{per}}(Y)$$

and let $M\subset W_{\mathrm{per}}(Y)$ be a closed linear subspace. Admissible choices include $M = \{0\}$ and $M = W_{\mathrm{per}}(Y)$.

The mixed formulation

The mixed formulation is defined as the following problem: Find $m\in M$ and $(w,u)\in X$ such that

$$\begin{cases} a((w,u),(w',u')) + b(m,(w',u')) = 0 & \forall\,(w',u')\in X,\\ b(m',(w,u)) = 0 & \forall\, m'\in M,\end{cases} \tag{4.12}$$

where the semilinear form $a : X\times X\to\mathbb{R}$ is given by

$$\begin{aligned} &a((w,u),(w',u'))\\ &:= \int_Y F_\gamma[(w,u)]\,L_\lambda(w',u') + \sigma_1\int_Y \mathrm{rot}(w)\cdot\mathrm{rot}(w') + \sigma_2\int_Y(\nabla u - w)\cdot(\nabla u' - w'),\end{aligned}$$

and the bilinear form $b : M\times X\to\mathbb{R}$ is given by

$$b(m,(w,u)) := \int_Y \nabla m\cdot(\nabla u - w)$$

for $(w,u),(w',u')\in X$ and $m\in M$. Here, we have used the operators

$$F_\gamma[(w,u)] := \sup_{\alpha\in\Lambda}\left\{\gamma^\alpha\left(-A^\alpha : Dw - b^\alpha\cdot\nabla u + c^\alpha u - f^\alpha\right)\right\},$$
$$L_\lambda(w,u) := \lambda u - \nabla\cdot w,$$

acting on $(w,u)\in X$, and the positive constants

$$\sigma_1 := \sigma_1(\delta) := 1 - \frac{1}{2}\sqrt{1-\delta},$$
$$\sigma_2 := \lambda\,\tilde\sigma_2(\delta) := \lambda\left(\frac{1-\sqrt{1-\delta}}{2} + \frac{1}{4\left(1-\sqrt{1-\delta}\right)}\right).$$

We proceed by showing well-posedness of this mixed formulation.

Well-posedness of the mixed formulation

We define a norm on the space $X = W_{\mathrm{per}}(Y;\mathbb{R}^n)\times H^1_{\mathrm{per}}(Y)$ by

$$|||(w,u)|||_\lambda^2 := \|Dw\|^2_{L^2(Y)} + 2\lambda\|\nabla u\|^2_{L^2(Y)} + \lambda^2\|u\|^2_{L^2(Y)}, \qquad (w,u)\in X.$$

It is easy to verify that this does indeed define a norm on X and we observe that there holds

$$\|Dw\|^2_{L^2(Y)} = \|\mathrm{rot}(w)\|^2_{L^2(Y)} + \|\nabla\cdot w\|^2_{L^2(Y)} \qquad \forall w\in H^1_{\mathrm{per}}(Y;\mathbb{R}^n), \tag{4.13}$$

which follows from the formal calculation (using integration by parts twice)

$$\int_Y |Dw|^2 - \int_Y |\mathrm{rot}(w)|^2 = \sum_{i,j=1}^n \int_Y \partial_i w_j\,\partial_j w_i = \sum_{i,j=1}^n \int_Y \partial_i w_i\,\partial_j w_j = \int_Y |\nabla\cdot w|^2,$$

and a density argument. Note that compared to the usual Maxwell-type inequality [32], we obtain an equality in (4.13) thanks to periodicity. We obtain two preliminary estimates.

Lemma 4.2.1 (Preliminary estimates)**.** *Let $(w,u),(w',u')\in X$ and $\rho\in(0,2)$. Then, writing $(\delta_w,\delta_u) := (w-w',u-u')$, there holds*

$$\|F_\gamma[(w,u)] - F_\gamma[(w',u')] - L_\lambda(\delta_w,\delta_u)\|_{L^2(Y)} \le \sqrt{1-\delta}\,|||(\delta_w,\delta_u)|||_\lambda, \tag{4.14}$$

and we have the Miranda–Talenti-type estimate

$$\frac{2-\rho}{2}|||(w,u)|||_\lambda^2 \le \|\mathrm{rot}(w)\|^2_{L^2(Y)} + \|L_\lambda(w,u)\|^2_{L^2(Y)} + \frac{\lambda}{\rho}\|\nabla u - w\|^2_{L^2(Y)}. \tag{4.15}$$

Proof. The first part of the Lemma, i.e., the estimate (4.14), is shown analogously to Lemma 4.1.1. For the second part, we use (4.13), integration by parts and Young's inequality to find

$$\begin{aligned}
|||(w,u)|||_\lambda^2 &= \|\mathrm{rot}(w)\|^2_{L^2(Y)} + \|\nabla\cdot w\|^2_{L^2(Y)} + 2\lambda\|\nabla u\|^2_{L^2(Y)} + \lambda^2\|u\|^2_{L^2(Y)} \\
&= \|\mathrm{rot}(w)\|^2_{L^2(Y)} + \|-\nabla\cdot w + \lambda u\|^2_{L^2(Y)} + 2\lambda\int_Y (\nabla u - w)\cdot\nabla u \\
&\leq \|\mathrm{rot}(w)\|^2_{L^2(Y)} + \|L_\lambda(w,u)\|^2_{L^2(Y)} + \frac{\lambda}{\rho}\|\nabla u - w\|^2_{L^2(Y)} + \lambda\rho\|\nabla u\|^2_{L^2(Y)} \\
&\leq \|\mathrm{rot}(w)\|^2_{L^2(Y)} + \|L_\lambda(w,u)\|^2_{L^2(Y)} + \frac{\lambda}{\rho}\|\nabla u - w\|^2_{L^2(Y)} + \frac{\rho}{2}|||(w,u)|||_\lambda^2,
\end{aligned}$$

which yields the Miranda–Talenti-type estimate (4.15). □

With these estimates at hand, we can proceed with showing essential properties of the maps a and b, namely monotonicity, Lipschitz continuity and an inf-sup condition, which will allow us to show well-posedness of the mixed formulation. We will use that we have the Poincaré inequality (see [18, Theorem 3.2]) for scalar functions,

$$\|v\|_{L^2(Y)} \leq \frac{\sqrt{n}}{\pi}\|\nabla v\|_{L^2(Y)} \qquad \forall\, v \in W_{\mathrm{per}}(Y), \tag{4.16}$$

and the corresponding inequality for vector-valued functions,

$$\|w\|_{L^2(Y)} \leq \frac{\sqrt{n}}{\pi}\|Dw\|_{L^2(Y)} \qquad \forall\, w \in W_{\mathrm{per}}(Y;\mathbb{R}^n). \tag{4.17}$$

We then have the following result:

Lemma 4.2.2 (Monotonicity, Lipschitz continuity and inf-sup condition)**.** *We have the following properties:*

(i) Monotonicity: For any $(w,u),(w',u') \in X$, writing $(\delta_w,\delta_u) := (w-w', u-u')$, we have

$$C_M|||(\delta_w,\delta_u)|||_\lambda^2 \leq a\left((w,u),(\delta_w,\delta_u)\right) - a\left((w',u'),(\delta_w,\delta_u)\right)$$

with the monotonicity constant $C_M := \frac{1}{4}\left(1-\sqrt{1-\delta}\right) > 0$.

(ii) Lipschitz continuity: For any $(w,u),(w',u'),(z,v) \in X$, writing $(\delta_w,\delta_u) := (w-w', u-u')$, we have

$$|a\left((w,u),(z,v)\right) - a\left((w',u'),(z,v)\right)| \leq C_L|||(\delta_w,\delta_u)|||_\lambda |||(z,v)|||_\lambda \tag{4.18}$$

with the Lipschitz constant $C_L := 2 + \sqrt{2}\sqrt{1-\delta} + \sigma_1(\delta) + \tilde{\sigma}_2(\delta)\left(\frac{1}{2} + \frac{n}{\pi^2}\lambda\right) > 0$.

(iii) Inf-sup condition: We have

$$\inf_{m'\in M\setminus\{0\}} \sup_{(w',u')\in X\setminus\{0\}} \frac{b(m',(w',u'))}{\|\nabla m'\|_{L^2(Y)}\,|||(w',u')|||_\lambda} \geq c_b \tag{4.19}$$

with the inf-sup constant $c_b := \lambda^{-\frac{1}{2}}\left(2+\frac{n}{\pi^2}\lambda\right)^{-\frac{1}{2}} > 0$.

Proof. We are going to prove the claimed results (i), (ii), (iii) separately.

(i) By (4.14), Young's inequality and the Miranda–Talenti-type estimate (4.15) with $\rho = 2-2\sqrt{1-\delta}$, we find that

$$\begin{aligned}
&a\left((w,u),(\delta_w,\delta_u)\right) - a\left((w',u'),(\delta_w,\delta_u)\right) - \sigma_1\|\mathrm{rot}(\delta_w)\|^2_{L^2(Y)} - \sigma_2\|\nabla\delta_u - \delta_w\|^2_{L^2(Y)}\\
&= \int_Y \left(F_\gamma[(w,u)] - F_\gamma[(w',u')]\right) L_\lambda(\delta_w,\delta_u)\\
&\geq \|L_\lambda(\delta_w,\delta_u)\|^2_{L^2(Y)} - \sqrt{1-\delta}\,|||(\delta_w,\delta_u)|||_\lambda \|L_\lambda(\delta_w,\delta_u)\|_{L^2(Y)}\\
&\geq \frac{2-\sqrt{1-\delta}}{2}\|L_\lambda(\delta_w,\delta_u)\|^2_{L^2(Y)} - \frac{\sqrt{1-\delta}}{2}|||(\delta_w,\delta_u)|||^2_\lambda\\
&\geq \frac{1-\sqrt{1-\delta}}{2}\|L_\lambda(\delta_w,\delta_u)\|^2_{L^2(Y)} - \frac{1}{2}\|\mathrm{rot}(\delta_w)\|^2_{L^2(Y)} - \frac{\lambda}{4-4\sqrt{1-\delta}}\|\nabla\delta_u - \delta_w\|^2_{L^2(Y)}.
\end{aligned}$$

Therefore, by definition of the constants σ_1, σ_2 and the Miranda–Talenti-type estimate (4.15) with the choice $\rho = 1$, we conclude that

$$\begin{aligned}
&a\left((w,u),(\delta_w,\delta_u)\right) - a\left((w',u'),(\delta_w,\delta_u)\right)\\
&\geq \frac{1-\sqrt{1-\delta}}{2}\left(\|L_\lambda(\delta_w,\delta_u)\|^2_{L^2(Y)} + \|\mathrm{rot}(\delta_w)\|^2_{L^2(Y)} + \lambda\|\nabla\delta_u - \delta_w\|^2_{L^2(Y)}\right)\\
&\geq \frac{1-\sqrt{1-\delta}}{4}|||(\delta_w,\delta_u)|||^2_\lambda,
\end{aligned}$$

which is the claimed inequality.

(ii) We note that we have

$$\|L_\lambda(w,u)\|_{L^2(Y)} \leq \sqrt{2}\,|||(w,u)|||_\lambda \qquad \forall\,(w,u)\in X, \tag{4.20}$$

as there holds $\|\nabla\cdot w\|_{L^2(Y)} \leq \|Dw\|_{L^2(Y)}$ for any $w \in W_{\mathrm{per}}(Y;\mathbb{R}^n)$ by (4.13). We bound the terms arising in the quantity on the left-hand side of (4.18) separately. For the term involving the nonlinearity, using (4.14), we have

$$\begin{aligned}
&\left|\int_Y \left(F_\gamma[(w,u)] - F_\gamma[(w',u')]\right) L_\lambda(z,v)\right|\\
&\leq \|L_\lambda(z,v)\|_{L^2(Y)}\left(\|L_\lambda(\delta_w,\delta_u)\|_{L^2(Y)} + \sqrt{1-\delta}\,|||(\delta_w,\delta_u)|||_\lambda\right)\\
&\leq \sqrt{2}\left(\sqrt{2}+\sqrt{1-\delta}\right)|||(\delta_w,\delta_u)|||_\lambda |||(z,v)|||_\lambda.
\end{aligned}$$

For the term multiplying the constant σ_1, we have

$$\left|\sigma_1 \int_Y \mathrm{rot}(\delta_w) \cdot \mathrm{rot}(z)\right| \leq \sigma_1 \|D\delta_w\|_{L^2(Y)} \|Dz\|_{L^2(Y)} \leq \sigma_1 |||(\delta_w, \delta_u)|||_\lambda |||(z, v)|||_\lambda,$$

as there holds $\|\mathrm{rot}(w)\|_{L^2(Y)} \leq \|Dw\|_{L^2(Y)}$ for any $w \in W_{\mathrm{per}}(Y; \mathbb{R}^n)$ by (4.13). For the term multiplying the constant σ_2, we have by the triangle, Poincaré (4.17) and Cauchy–Schwarz inequalities that

$$\begin{aligned}
&\left|\sigma_2 \int_Y (\nabla \delta_u - \delta_w) \cdot (\nabla v - z)\right| \\
&\leq \sigma_2 \left(\|\nabla \delta_u\|_{L^2(Y)} + \frac{\sqrt{n}}{\pi} \|D\delta_w\|_{L^2(Y)}\right) \left(\|\nabla v\|_{L^2(Y)} + \frac{\sqrt{n}}{\pi} \|Dz\|_{L^2(Y)}\right) \\
&\leq \sigma_2 \left(\frac{1}{2\lambda} + \frac{n}{\pi^2}\right) |||(\delta_w, \delta_u)|||_\lambda |||(z, v)|||_\lambda.
\end{aligned}$$

Altogether, we obtain the claimed inequality (4.18) with the constant

$$C_L = 2 + \sqrt{2}\sqrt{1 - \delta} + \sigma_1 + \sigma_2 \left(\frac{1}{2\lambda} + \frac{n}{\pi^2}\right),$$

which is identical to the one given in Lemma 4.2.2 (ii) using that $\sigma_2 = \lambda\, \tilde{\sigma}_2$.

(iii) For any $m' \in M \backslash \{0\}$ we have $(0, m') \in X$ and hence

$$\begin{aligned}
\sup_{(w', u') \in X \backslash \{0\}} \frac{b(m', (w', u'))}{|||(w', u')|||_\lambda} &\geq \frac{b(m', (0, m'))}{|||(0, m')|||_\lambda} \\
&= \frac{\|\nabla m'\|^2_{L^2(Y)}}{\sqrt{2\lambda \|\nabla m'\|^2_{L^2(Y)} + \lambda^2 \|m'\|^2_{L^2(Y)}}} \\
&\geq \frac{\|\nabla m'\|_{L^2(Y)}}{\sqrt{2\lambda + \frac{n}{\pi^2}\lambda^2}}
\end{aligned}$$

by Poincaré's inequality (4.16) (recall that $M \subset W_{\mathrm{per}}(Y)$), which yields the claimed result (4.19). □

Remark 4.2.1 (Local Lipschitz estimate)**.** *Similarly, one obtains the local Lipschitz estimate*

$$\begin{aligned}
&|a_I((w, u), (z, v)) - a_I((w', u'), (z, v))| \\
&\leq C_L' \left[|||(w - w', u - u')|||_{\lambda, I} + \|w - w'\|_{L^2(I)}\right] \left[|||(z, v)|||_{\lambda, I} + \|z\|_{L^2(I)}\right]
\end{aligned}$$

for all $(w, u), (w', u'), (z, v) \in X$ and any open $I \subset Y$ with a constant $C_L' = C_L'(\delta, \lambda, n) > 0$. Here, the subscript I in a_I and $|||\cdot|||_{\lambda, I}$ denotes that integrals in the corresponding definitions are taken over the set I.

Now we are in a position to show well-posedness of the mixed formulation, i.e., the existence and uniqueness of a solution $(m,(w,u)) \in M \times X$ to (4.12).

Theorem 4.2.1 ([50, Theorem 2.5] Well-posedness of the mixed formulation)**.** *There exists a unique solution $(m,(w,u)) \in M \times X$ to (4.12). Further, $m = 0$, $u \in H^2_{\mathrm{per}}(Y)$ with $\nabla u = w$ and u is the solution to (4.1).*

Proof. The existence of a unique solution $(m,(w,u)) \in M \times X$ to (4.12) follows from the Brezzi-splitting; see [21] and [51, Proposition 2.5], as we have the monotonicity and Lipschitz continuity for a and an inf-sup condition from Lemma 4.2.2. For the second part of the claim, i.e., that $m = 0$, $u \in H^2_{\mathrm{per}}(Y)$ with $w = \nabla u$ and u is the solution to (4.1), we note that L_λ is surjective from the set $X_g := \{(w', u') \in X : w' = \nabla u'\}$ onto $L^2(Y)$. We first test the mixed formulation (4.12) with pairs (w', u') from X_g to obtain $F_\gamma[(w,u)] = 0$ almost everywhere and then with the solution pair (w,u) to find $w = \nabla u$ and thus $u \in H^2_{\mathrm{per}}(Y)$. We conclude the proof by noting that this implies that u is the solution to (4.5) (and hence to (4.1) by Theorem 4.1.1) and that $m = 0$. $\square$

4.2.2 Discrete mixed formulation and error analysis

We take closed linear subspaces $W_h \subset W_{\mathrm{per}}(Y;\mathbb{R}^n)$, $U_h \subset H^1_{\mathrm{per}}(Y)$, $M_h \subset U_h \cap M$ (recall that $M \subset W_{\mathrm{per}}(Y)$), and define

$$X_h := W_h \times U_h \subset X.$$

We then define the discrete mixed formulation as the following problem: Find $m_h \in M_h$ and $(w_h, u_h) \in X_h$ such that

$$\begin{cases} a\left((w_h,u_h),(w_h',u_h')\right) + b(m_h,(w_h',u_h')) = 0 & \forall\, (w_h',u_h') \in X_h, \\ b(m_h',(w_h,u_h)) = 0 & \forall\, m_h' \in M_h. \end{cases} \tag{4.21}$$

We note that we have boundedness of b and a discrete inf-sup condition.

Lemma 4.2.3 (Boundedness of b and the discrete inf-sup condition)**.** *For any $(m',(w',u')) \in M \times X$, we have*

$$b(m',(w',u')) \leq C_b \|\nabla m'\|_{L^2(Y)} |||(w',u')|||_\lambda$$

with the constant $C_b := \lambda^{-\frac{1}{2}}\left(\frac{1}{2} + \frac{n}{\pi^2}\lambda\right)^{\frac{1}{2}} > 0$. Further, the discrete inf-sup condition

$$\inf_{m_h' \in M_h \setminus \{0\}} \sup_{(w_h',u_h') \in X_h \setminus \{0\}} \frac{b(m_h',(w_h',u_h'))}{\|\nabla m_h'\|_{L^2(Y)}\, |||(w_h',u_h')|||_\lambda} \geq c_b \tag{4.22}$$

holds with $c_b > 0$ as in Lemma 4.2.2 (iii).

Proof. We use the triangle, Poincaré (4.17) and the Cauchy–Schwarz inequalities to obtain

$$\begin{aligned} b(m', (w', u')) &\leq \|\nabla m'\|_{L^2(Y)} \left(\|\nabla u'\|_{L^2(Y)} + \frac{\sqrt{n}}{\pi} \|Dw'\|_{L^2(Y)} \right) \\ &\leq \sqrt{\frac{1}{2\lambda} + \frac{n}{\pi^2}} \|\nabla m'\|_{L^2(Y)} |||(w', u')|||_\lambda \end{aligned}$$

for all $(m', (w', u')) \in M \times X$.

The discrete inf-sup condition holds as for $m_h' \in M_h \backslash \{0\}$ we have $(0, m_h') \in X_h \backslash \{0\}$ since $M_h \subset U_h \cap M$, and hence

$$\begin{aligned} \sup_{(w_h', u_h') \in X_h \backslash \{0\}} \frac{b(m_h', (w_h', u_h'))}{|||(w_h', u_h')|||_\lambda} &\geq \frac{b(m_h', (0, m_h'))}{|||(0, m_h')|||_\lambda} \\ &= \frac{\|\nabla m_h'\|_{L^2(Y)}^2}{\sqrt{2\lambda \|\nabla m_h'\|_{L^2(Y)}^2 + \lambda^2 \|m_h'\|_{L^2(Y)}^2}} \\ &\geq \frac{\|\nabla m_h'\|_{L^2(Y)}}{\sqrt{2\lambda + \frac{n}{\pi^2}\lambda^2}} \end{aligned}$$

by Poincaré's inequality (4.16) (recall that $M \subset W_{\mathrm{per}}(Y)$), which yields the claimed result (4.22). □

It follows that we have well-posedness of the discrete mixed formulation analogously to Theorem 4.2.1. We also obtain an error bound.

Theorem 4.2.2 ([50, Theorem 2.7] Well-posedness and error bound)**.** *There exists a unique solution* $(m_h, (w_h, u_h)) \in M_h \times X_h$ *to the discrete mixed formulation* (4.21)*. Further, we have*

$$|||(w - w_h, u - u_h)|||_\lambda \leq C_e \inf_{(w_h', u_h') \in X_h} |||(w - w_h', u - u_h')|||_\lambda,$$

where $(m, (w, u)) \in M \times X$ *denotes the solution to* (4.12) *and* $C_e = C_e(\delta, \lambda, n) > 0$ *is the constant*

$$C_e := 2 \frac{C_L}{C_M} \left(1 + \frac{C_b}{c_b} \right)$$

with $C_L, C_M, C_b, c_b > 0$ *from Lemmata 4.2.2 and 4.2.3.*

Proof. We only show the error bound as the existence and uniqueness of solutions for (4.21) follows from Lemma 4.2.2 and Lemma 4.2.3 in a standard way; see [51, Proposition 3.1].

Step 1: We introduce the discrete kernel

$$Z_h := \{(w'_h, u'_h) \in X_h : b\,(m'_h, (w'_h, u'_h)) = 0 \quad \forall\, m'_h \in M_h\}$$

and claim that there holds

$$|||(w - w_h, u - u_h)|||_\lambda \leq \frac{C_L}{C_M} \inf_{(w'_h, u'_h) \in Z_h} |||(w - w'_h, u - u'_h)|||_\lambda. \tag{4.23}$$

Indeed, we use successively the monotonicity from Lemma 4.2.2 (i), the solution property of (w, u) from Theorem 4.2.1 and the fact that (w_h, u_h) solves the discrete problem (4.21), and the Lipschitz estimate from Lemma 4.2.2 (ii) to find that

$$\begin{aligned}
&C_M |||(w - w_h, u - u_h)|||_\lambda^2 \\
&\quad \leq a\,((w, u), (w - w_h, u - u_h)) - a\,((w_h, u_h), (w - w_h, u - u_h)) \\
&\quad = -a\,((w_h, u_h), (w, u)) \\
&\quad = -a\,((w_h, u_h), (w - w'_h, u - u'_h)) \\
&\quad = a\,((w, u), (w - w'_h, u - u'_h)) - a\,((w_h, u_h), (w - w'_h, u - u'_h)) \\
&\quad \leq C_L |||(w - w_h, u - u_h)|||_\lambda |||(w - w'_h, u - u'_h)|||_\lambda
\end{aligned}$$

for any $(w'_h, u'_h) \in Z_h$, which implies the desired estimate (4.23).

Step 2: We let $(w_*, u_*) \in X_h$ denote the best-approximation to (w, u) from Z_h, i.e.,

$$|||(w - w_*, u - u_*)|||_\lambda = \inf_{(w'_h, u'_h) \in Z_h} |||(w - w'_h, u - u'_h)|||_\lambda, \tag{4.24}$$

and we derive a linear mixed problem for (w_*, u_*).

By the discrete inf-sup condition (4.22), there exists $m_* \in M_h$ such that

$$\begin{cases} \langle (w_*, u_*), (w'_h, u'_h) \rangle_\lambda + b(m_*, (w'_h, u'_h)) = \langle (w, u), (w'_h, u'_h) \rangle_\lambda & \forall\, (w'_h, u'_h) \in X_h, \\ \qquad\qquad\qquad\qquad b(m'_h, (w_*, u_*)) = 0 & \forall\, m'_h \in M_h, \end{cases}$$

where $\langle \cdot, \cdot \rangle_\lambda : X \times X \to \mathbb{R}$ is the inner product given by

$$\langle (w', u'), (w'', u'') \rangle_\lambda := \int_Y Dw' : Dw'' + 2\lambda \int_Y \nabla u' \cdot \nabla u'' + \lambda^2 \int_Y u' u''.$$

We also note that the solution pair $((w, u), m)$ satisfies the similar system (recall that $m = 0$)

$$\begin{cases} \langle (w, u), (w', u') \rangle_\lambda + b(m, (w', u')) = \langle (w, u), (w', u') \rangle_\lambda & \forall\, (w', u') \in X, \\ \qquad\qquad\qquad\qquad b(m', (w, u)) = 0 & \forall\, m' \in M. \end{cases}$$

Step 3: We derive an error bound for $(w - w_*, u - u_*)$ in the $|||\cdot|||_\lambda$ norm using the classical linear mixed finite element theory.

Note that for any $(w', u'), (w'', u'') \in X$, we have

$$|\langle (w', u'), (w'', u'')\rangle_\lambda| \leq |||(w', u')|||_\lambda |||(w'', u'')|||_\lambda, \quad \langle (w', u'), (w', u')\rangle_\lambda = |||(w', u')|||_\lambda^2.$$

In particular, we have boundedness and coercivity on the whole space, i.e.,

$$\begin{aligned} |\langle (w', u'), (w'', u'')\rangle_\lambda| &\leq C_a |||(w', u')|||_\lambda |||(w'', u'')|||_\lambda, \\ \langle (w', u'), (w', u')\rangle_\lambda &\geq c_a |||(w', u')|||_\lambda^2 \end{aligned}$$

for all $(w', u'), (w'', u'') \in X$ with constants $C_a := c_a := 1$. Further, from Lemma 4.2.3 we have the discrete inf-sup condition (4.22) with constant c_b and boundedness of b with constant C_b. Then, by linear mixed finite element theory (see [91]), we obtain

$$\begin{aligned} |||(w - w_*, u - u_*)|||_\lambda &\leq \left(1 + \frac{C_a}{c_a}\right)\left(1 + \frac{C_b}{c_b}\right) \inf_{(w_h', u_h') \in X_h} |||(w - w_h', u - u_h')|||_\lambda \\ &\qquad + \frac{C_b}{c_a} \inf_{m_h' \in M_h} \|\nabla (m - m_h')\|_{L^2(Y)} \qquad (4.25) \\ &= 2\left(1 + \frac{C_b}{c_b}\right) \inf_{(w_h', u_h') \in X_h} |||(w - w_h', u - u_h')|||_\lambda, \end{aligned}$$

where we used $m = 0$ and $C_a = c_a = 1$ in the last line.

Step 4: We conclude by combining (4.23), (4.24) and (4.25):

$$\begin{aligned} |||(w - w_h, u - u_h)|||_\lambda &\leq \frac{C_L}{C_M} |||(w - w_*, u - u_*)|||_\lambda \\ &\leq 2\frac{C_L}{C_M}\left(1 + \frac{C_b}{c_b}\right) \inf_{(w_h', u_h') \in X_h} |||(w - w_h', u - u_h')|||_\lambda, \end{aligned}$$

which is the desired error bound. □

Remark 4.2.2. *Note that the error constant $C_e = C_e(\delta, \lambda, n)$ is monotonically increasing in λ.*

Besides this *a priori* error bound, the monotonicity property from Lemma 4.2.2 allows us to obtain an *a posteriori* error bound.

Theorem 4.2.3 ([50, Theorem 2.9] *a posteriori* error bound and efficiency estimate)**.** *For the solution $(m, (w, u)) \in M \times X$ to the mixed formulation* (4.12) *and the solution*

$(m_h,(w_h,u_h)) \in M_h \times X_h$ to the discrete mixed formulation (4.21), writing $e_h := (w - w_h, u - u_h)$, we have the error bound

$$\begin{aligned}&|||e_h|||_\lambda\\&\le \sqrt{2}\, C_M^{-\frac{1}{2}} \left(C_M^{-1} \left\| F_\gamma[(w_h,u_h)] \right\|_{L^2(Y)}^2 + \sigma_1 \left\| \mathrm{rot}(w_h) \right\|_{L^2(Y)}^2 + \sigma_2 \left\| w_h - \nabla u_h \right\|_{L^2(Y)}^2 \right)^{\frac{1}{2}}\end{aligned}$$

and the efficiency estimate

$$\begin{aligned}\frac{1}{2} \left\| F_\gamma[(w_h,u_h)] \right\|_{L^2(Y)}^2 + \sigma_1 \left\| \mathrm{rot}(w_h) \right\|_{L^2(Y)}^2 + \sigma_2 \left\| w_h - \nabla u_h \right\|_{L^2(Y)}^2&\\ \le \left(C_L + \frac{1-\delta}{2} \right) |||e_h|||_\lambda^2&,\end{aligned}$$

where $C_M, C_L > 0$ are the constants from Lemma 4.2.2.

Proof. We use successively the monotonicity from Lemma 4.2.2 (i), the solution property of (w,u) from Theorem 4.2.1, the Cauchy–Schwarz inequality (note $w = \nabla u$), the bound (4.20), and Young's inequality:

$$\begin{aligned}&C_M |||e_h|||_\lambda^2\\&\le a\left((w,u),e_h\right) - a\left((w_h,u_h),e_h\right)\\&= -a\left((w_h,u_h),e_h\right)\\&\le \left\| F_\gamma[(w_h,u_h)] \right\|_{L^2(Y)} \left\| L_\lambda e_h \right\|_{L^2(Y)} + \sigma_1 \left\| \mathrm{rot}(w_h) \right\|_{L^2(Y)}^2 + \sigma_2 \left\| w_h - \nabla u_h \right\|_{L^2(Y)}^2\\&\le \sqrt{2} \left\| F_\gamma[(w_h,u_h)] \right\|_{L^2(Y)} |||e_h|||_\lambda + \sigma_1 \left\| \mathrm{rot}(w_h) \right\|_{L^2(Y)}^2 + \sigma_2 \left\| w_h - \nabla u_h \right\|_{L^2(Y)}^2\\&\le C_M^{-1} \left\| F_\gamma[(w_h,u_h)] \right\|_{L^2(Y)}^2 + \frac{C_M}{2} |||e_h|||_\lambda^2 + \sigma_1 \left\| \mathrm{rot}(w_h) \right\|_{L^2(Y)}^2 + \sigma_2 \left\| w_h - \nabla u_h \right\|_{L^2(Y)}^2 .\end{aligned}$$

Upon rearranging, we find the claimed *a posteriori* estimate.

For the efficiency estimate, recall the solution property of (w,u) from Theorem 4.2.1, in particular $w = \nabla u$ and $F_\gamma[(w,u)] = 0$ almost everywhere. With the Lipschitz property from Lemma 4.2.2 (ii) and with Lemma 4.2.1, we then obtain

$$\begin{aligned}&C_L |||e_h|||_\lambda^2 - \sigma_1 \left\| \mathrm{rot}(w_h) \right\|_{L^2(Y)}^2 - \sigma_2 \left\| w_h - \nabla u_h \right\|_{L^2(Y)}^2\\&\ge a\left((w,u),e_h\right) - a\left((w_h,u_h),e_h\right) - \sigma_1 \left\| \mathrm{rot}(w_h) \right\|_{L^2(Y)}^2 - \sigma_2 \left\| w_h - \nabla u_h \right\|_{L^2(Y)}^2\\&= -a\left((w_h,u_h),e_h\right) - \sigma_1 \left\| \mathrm{rot}(w_h) \right\|_{L^2(Y)}^2 - \sigma_2 \left\| w_h - \nabla u_h \right\|_{L^2(Y)}^2\\&= \left\| F_\gamma[(w_h,u_h)] \right\|_{L^2(Y)}^2 + \int_Y F_\gamma[(w_h,u_h)] \left(F_\gamma[(w,u)] - F_\gamma[(w_h,u_h)] - L_\lambda e_h \right)\\&\ge \left\| F_\gamma[(w_h,u_h)] \right\|_{L^2(Y)}^2 - \sqrt{1-\delta} \left\| F_\gamma[(w_h,u_h)] \right\|_{L^2(Y)} |||e_h|||_\lambda\\&\ge \frac{1}{2} \left\| F_\gamma[(w_h,u_h)] \right\|_{L^2(Y)}^2 - \frac{1-\delta}{2} |||e_h|||_\lambda^2,\end{aligned}$$

which yields the efficiency estimate upon rearranging. □

Remark 4.2.3 (Local efficiency)**.** *Similarly, one obtains the local efficiency estimate*

$$\frac{1}{2}\|F_\gamma[(w_h,u_h)]\|_{L^2(I)}^2+\sigma_1\|\mathrm{rot}(w_h)\|_{L^2(I)}^2+\sigma_2\|w_h-\nabla u_h\|_{L^2(I)}^2 \\ \leq\left(2C_L'+\frac{1-\delta}{2}\right)\left(|||(w-w_h,u-u_h)|||_{\lambda,I}^2+\|w-w_h\|_{L^2(I)}^2\right)$$

for any open $I\subset Y$, where $C_L'>0$ is the constant from Remark 4.2.1.

4.3 Numerical homogenization of HJB equations

4.3.1 Framework

Let $\Omega\subset\mathbb{R}^n$ be a bounded convex domain in dimension $n\in\{2,3\}$ and let Λ be a compact metric space. For $\varepsilon>0$ small, we consider problems of the form

$$\begin{cases}\sup_{\alpha\in\Lambda}\left\{-A^\alpha\left(\cdot,\frac{\cdot}{\varepsilon}\right):D^2u_\varepsilon-b^\alpha\left(\cdot,\frac{\cdot}{\varepsilon}\right)\cdot\nabla u_\varepsilon+u_\varepsilon-f^\alpha\left(\cdot,\frac{\cdot}{\varepsilon}\right)\right\}=0 & \text{in }\Omega,\\ u_\varepsilon=0 & \text{on }\partial\Omega,\end{cases} \tag{4.26}$$

where we assume that the functions

$$\begin{aligned} A=(a_{ij})_{1\leq i,j\leq n}&:\bar\Omega\times\mathbb{R}^n\times\Lambda\to\mathbb{R}^{n\times n}_{\mathrm{sym}}, && (x,y,\alpha)\mapsto A(x,y,\alpha)=:A^\alpha(x,y),\\ b=(b_i)_{1\leq i\leq n}&:\bar\Omega\times\mathbb{R}^n\times\Lambda\to\mathbb{R}^n, && (x,y,\alpha)\mapsto b(x,y,\alpha)=:b^\alpha(x,y),\\ f&:\bar\Omega\times\mathbb{R}^n\times\Lambda\to\mathbb{R}, && (x,y,\alpha)\mapsto f(x,y,\alpha)=:f^\alpha(x,y)\end{aligned}$$

satisfy the following assumptions:

(i) Continuity: A, b, f are continuous on $\bar\Omega\times\mathbb{R}^n\times\Lambda$,

(ii) Periodicity: $A^\alpha(x,\cdot)$, $b^\alpha(x,\cdot)$, $f^\alpha(x,\cdot)$ are Y-periodic for fixed $\alpha\in\Lambda$ and $x\in\bar\Omega$,

(iii) Regularity: A^α, b^α, f^α are Lipschitz on $\bar\Omega\times\mathbb{R}^n$ uniformly in $\alpha\in\Lambda$,

(iv) Ellipticity: There exist $\zeta_1,\zeta_2>0$ such that $\zeta_1|\xi|^2\leq A\xi\cdot\xi\leq\zeta_2|\xi|^2$ in $\bar\Omega\times\mathbb{R}^n\times\Lambda$ for all $\xi\in\mathbb{R}^n$.

Further, it is assumed that the (generalized) Cordes condition

$$\frac{|A|^2+\frac{|b|^2}{2\lambda}+\frac{1}{\lambda^2}}{\left(\mathrm{tr}(A)+\frac{1}{\lambda}\right)^2}\leq\frac{1}{n+\delta}\qquad\text{in }\bar\Omega\times\mathbb{R}^n\times\Lambda \tag{4.27}$$

holds for some constants $\lambda>0$ and $\delta\in(0,1)$. Then we have well-posedness in the sense of strong solutions; see [88].

Theorem 4.3.1 (Existence and uniqueness of strong solutions)**.** *In this situation, for any given $\varepsilon > 0$, there exists a unique strong solution $u_\varepsilon \in H^2(\Omega) \cap H^1_0(\Omega)$ to (4.26).*

Remark 4.3.1. *Problems involving a non-constant zeroth-order coefficient, i.e., problems of the form*

$$\sup_{\alpha\in\Lambda}\left\{-A^\alpha\left(\cdot,\frac{\cdot}{\varepsilon}\right):D^2v_\varepsilon - b^\alpha\left(\cdot,\frac{\cdot}{\varepsilon}\right)\cdot\nabla v_\varepsilon + c^\alpha\left(\cdot,\frac{\cdot}{\varepsilon}\right)v_\varepsilon - f^\alpha\left(\cdot,\frac{\cdot}{\varepsilon}\right)\right\} = 0 \quad \text{in } \Omega,$$
$$v_\varepsilon = 0 \quad \text{on } \partial\Omega,$$

with c^α satisfying the same assumptions as the components of b^α and additionally $\inf_{\bar\Omega\times\mathbb{R}^n\times\Lambda} c > 0$, can be reduced to a problem of the form (4.26). This is due to the fact that division by $c^\alpha(x, x/\varepsilon)$ inside the argument of the supremum does not change the sets of strong and viscosity solutions; see [67, Remark 2.2].

4.3.2 Homogenization

In this section, we briefly recall known homogenization results from the literature. Let us start by recalling one of the several equivalent definitions of a viscosity solution; see [73].

Definition 4.3.1 (Viscosity solution)**.** *Let $\Omega \subset \mathbb{R}^n$ be open and $F : \Omega \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^{n\times n}_{\mathrm{sym}} \to \mathbb{R}$ be continuous. A continuous function $u : \Omega \to \mathbb{R}$, $u \in C(\bar\Omega)$, is called a viscosity solution to the equation*

$$F(x, u, \nabla u, D^2 u) = 0 \quad \text{in } \Omega,$$

if for any $\phi \in C^2(\Omega)$ there holds

$$x_0 \in \Omega \text{ local maximum point of } u - \phi \implies F(x_0, u(x_0), \nabla\phi(x_0), D^2\phi(x_0)) \leq 0,$$
$$x_0 \in \Omega \text{ local minimum point of } u - \phi \implies F(x_0, u(x_0), \nabla\phi(x_0), D^2\phi(x_0)) \geq 0.$$

For an overview of the theory of viscosity solutions for second-order equations we refer the reader to [33]. We note that the strong solution $u_\varepsilon \in H^2(\Omega) \cap H^1_0(\Omega)$ to (4.26) belongs to $C(\bar\Omega)$ (recall $n \in \{2, 3\}$) and a natural question to ask is whether u_ε is a viscosity solution. If the strong solution u_ε is such that $u_\varepsilon \in W^{2,n}_{\mathrm{loc}}(\Omega)$, then it is a viscosity solution to (4.26); see [24, 73, 74]. We also note that the viscosity solution to (4.26) is unique; see [62]. We then have the following result; see [84]:

Remark 4.3.2 (Regularity)**.** *Let $u_\varepsilon \in H^2(\Omega) \cap H^1_0(\Omega)$ be the unique strong solution to* (4.26) *given by Theorem 4.3.1. Then*

$$u_\varepsilon \in C^{2,\tilde{\alpha}}(\Omega) \cap C(\bar{\Omega})$$

for some $\tilde{\alpha} > 0$ and u_ε is the unique viscosity solution to (4.26)*. Further, if $\partial\Omega \in C^{2,\beta}$ for some $\beta > 0$, then $u_\varepsilon \in C^{2,\tilde{\alpha}}(\bar{\Omega})$ for some $\tilde{\alpha} > 0$.*

With this observation at hand, we can use the well-known homogenization results for viscosity solutions; see [25, 39, 40].

Theorem 4.3.2 (Homogenization of HJB problems)**.** *The solution u_ε to* (4.26) *converges uniformly on $\bar{\Omega}$ to the viscosity solution $u_0 \in C(\bar{\Omega})$ of*

$$\begin{cases} u_0 + H(x, \nabla u_0, D^2 u_0) = 0 & \text{in } \Omega, \\ u_0 = 0 & \text{on } \partial\Omega, \end{cases} \tag{4.28}$$

with an effective Hamiltonian $H : \bar{\Omega} \times \mathbb{R}^n \times \mathbb{R}^{n\times n}_{\mathrm{sym}} \to \mathbb{R}$ defined as follows: For given $(x, p, R) \in \bar{\Omega} \times \mathbb{R}^n \times \mathbb{R}^{n\times n}_{\mathrm{sym}}$ we define $H(x, p, R) \in \mathbb{R}$ to be the unique real number such that there exists a function $v \in C(\mathbb{R}^n)$, a so-called corrector, that is a viscosity solution to

$$\begin{cases} \sup_{\alpha\in\Lambda} \left\{-A^\alpha_x : D^2 v - g^\alpha_{x,p,R}\right\} = H(x, p, R) & \text{in } \mathbb{R}^n, \\ v \text{ is } Y\text{-periodic}, \end{cases} \tag{4.29}$$

where $A^\alpha_x(y) := A^\alpha(x, y)$ and $g^\alpha_{x,p,R}(y) := A^\alpha(x, y) : R + b^\alpha(x, y) \cdot p + f^\alpha(x, y)$ for $y \in \mathbb{R}^n$, $\alpha \in \Lambda$.

Let us note that rates for the convergence of u_ε to the homogenized solution u_0 have been derived for the whole space problem in [28].

The effective Hamiltonian can also be obtained through a limit of ergodic approximations, the so-called approximate correctors; see [11] and the references therein. For $(x, p, R) \in \bar{\Omega} \times \mathbb{R}^n \times \mathbb{R}^{n\times n}_{\mathrm{sym}}$ and $\sigma > 0$, the approximate corrector $v^\sigma = v^\sigma(\cdot\,; x, p, R) \in C(\mathbb{R}^n)$ is defined to be the viscosity solution to

$$\begin{cases} \sigma v^\sigma + \sup_{\alpha\in\Lambda} \left\{-A^\alpha_x : D^2 v^\sigma - g^\alpha_{x,p,R}\right\} = 0 & \text{in } \mathbb{R}^n, \\ v^\sigma \text{ is } Y\text{-periodic.} \end{cases} \tag{4.30}$$

Remark 4.3.3 (Regularity of approximate correctors)**.** *The viscosity solution $v^\sigma = v^\sigma(\cdot\,; x, p, R) \in C(\mathbb{R}^n)$ to* (4.30) *is in fact a classical solution $v^\sigma \in C^2(\mathbb{R}^n)$. Further, there exists an $\tilde{\alpha} \in (0, 1)$ such that*

$$\|\sigma v^\sigma(\cdot\,; x, p, R)\|_\infty + \|v^\sigma(\cdot\,; x, p, R) - v^\sigma(0\,; x, p, R)\|_{C^{2,\tilde{\alpha}}(\mathbb{R}^n)} \lesssim 1 + |p| + |R|$$

for all $(x, p, R) \in \bar{\Omega} \times \mathbb{R}^n \times \mathbb{R}^{n\times n}_{\mathrm{sym}}$; see [10, 28].

The value $H(x,p,R) \in \mathbb{R}$ for the effective Hamiltonian at the point (x,p,R) is then the uniform limit of the sequence $\{-\sigma v^\sigma\}_{\sigma>0}$ as $\sigma \to 0$; see [28].

Lemma 4.3.1 (Properties of the effective Hamiltonian)**.** *The following holds true.*

(i) The sequence $\{-\sigma v^\sigma(\cdot\,; x,p,R)\}_{\sigma>0}$ *converges uniformly to the constant value* $H(x,p,R)$ *with*

$$\|-\sigma v^\sigma(\cdot\,; x,p,R) - H(x,p,R)\|_\infty \lesssim \sigma\,(1+|p|+|R|)$$

for all $(x,p,R) \in \bar{\Omega} \times \mathbb{R}^n \times \mathbb{R}^{n\times n}_{\mathrm{sym}}$ *and* $\sigma > 0$ *sufficiently small.*

(ii) The effective Hamiltonian $H = H(x,p,R)$ *is uniformly elliptic, it is convex in* R*, and we have*

$$\begin{aligned} |H(x,p_1,R_1) - H(x,p_2,R_2)| &\lesssim |p_1-p_2| + |R_1-R_2|, \\ |H(x_1,p,R) - H(x_2,p,R)| &\lesssim |x_1-x_2|\,(1+|p|+|R|)\,, \end{aligned}$$

for any $x, x_1, x_2 \in \bar{\Omega}$*,* $p, p_1, p_2 \in \mathbb{R}^n$ *and* $R, R_1, R_2 \in \mathbb{R}^{n\times n}_{\mathrm{sym}}$*.*

Note that the properties of the approximate correctors from Remark 4.3.3 and Lemma 4.3.1 (i) allow passage to the limit $\sigma \to 0$ in (4.30) and guarantee the existence of a corrector $v \in C^2(\mathbb{R}^n)$ (i.e., a classical solution to (4.29)). We also note that the properties of the effective Hamiltonian from Lemma 4.3.1 (ii) yield a regularity result for the homogenized solution as it is of the type of problems studied in [84].

Remark 4.3.4 (Regularity of the homogenized solution)**.** *The viscosity solution* $u_0 \in C(\bar{\Omega})$ *to the homogenized problem* (4.28) *satisfies*

$$u_0 \in C^{2,\tilde{\alpha}}(\Omega) \cap C(\bar{\Omega})$$

for some $\tilde{\alpha} > 0$*. Further, if* $\partial\Omega \in C^{2,\beta}$ *for some* $\beta > 0$*, then* $u_0 \in C^{2,\tilde{\alpha}}(\bar{\Omega})$ *for some* $\tilde{\alpha} > 0$*.*

4.3.3 Approximation of the approximate corrector

We construct a mixed finite element method for the numerical approximation of the approximate corrector for fixed $(x,p,R) \in \bar{\Omega} \times \mathbb{R}^n \times \mathbb{R}^{n\times n}_{\mathrm{sym}}$. For $\sigma \in (0,1)$ we consider the problem (4.30), i.e., the problem of finding a strong solution v^σ to

$$\begin{cases} \sup_{\alpha\in\Lambda} \{-A_x^\alpha : D^2 v^\sigma + \sigma v^\sigma - g^\alpha_{x,p,R}\} = 0 \quad \text{in } Y, \\ \qquad\qquad\qquad\qquad v^\sigma \text{ is } Y\text{-periodic.} \end{cases} \tag{4.31}$$

Recall the notation $A_x^\alpha(y) := A_x(y, \alpha) := A(x, y, \alpha)$ and

$$g_{x,p,R}^\alpha(y) := g_{x,p,R}(y, \alpha) := A^\alpha(x, y) : R + b^\alpha(x, y) \cdot p + f^\alpha(x, y)$$

for $y \in \mathbb{R}^n$ and $\alpha \in \Lambda$ from Theorem 4.3.2.

Note that $g_{x,p,R} : \mathbb{R}^n \times \Lambda \to \mathbb{R}$ is continuous, and that $g_{x,p,R}^\alpha$ is Y-periodic for fixed $\alpha \in \Lambda$ and Lipschitz on $\mathbb{R}^n$ uniformly in α. We also note that we have the Cordes condition (4.27), which yields

$$\frac{|A_x|^2 + \frac{\sigma^2}{\lambda_\sigma^2}}{(\mathrm{tr}(A_x) + \frac{\sigma}{\lambda_\sigma})^2} \leq \frac{1}{n + \delta} \qquad \text{in } \mathbb{R}^n \times \Lambda, \tag{4.32}$$

where $\lambda_\sigma > 0$ is given by

$$\lambda_\sigma := \sigma\lambda.$$

The corresponding scaling function $\gamma^\alpha(y) := \gamma(y, \alpha)$ is given by (compare to (4.3))

$$\gamma := \frac{\mathrm{tr}(A_x) + \frac{\sigma}{\lambda_\sigma}}{|A_x|^2 + \frac{\sigma^2}{\lambda_\sigma^2}} = \frac{\mathrm{tr}(A_x) + \frac{1}{\lambda}}{|A_x|^2 + \frac{1}{\lambda^2}}.$$

Observe that (4.32) is the Cordes condition (4.2) for the problem (4.31) with Cordes constants δ and λ_σ. Therefore, Theorem 4.1.1 ensures the well-posedness of the problem (4.31), i.e., existence and uniqueness of a strong periodic solution. We apply the mixed finite element method from Section 4.2 to problem (4.31) to obtain an approximation.

The scheme from Section 4.2 applied to the problem (4.31) yields an approximation $(m_h^\sigma, (w_h^\sigma, v_h^\sigma)) \in M_h \times X_h$, whose existence and uniqueness are guaranteed by Theorem 4.2.2, satisfying the error bound

$$|||(\nabla v^\sigma - w_h^\sigma, v^\sigma - v_h^\sigma)|||_{\lambda_\sigma} \leq C_e(\delta, \lambda_\sigma, n) \inf_{(w_h', u_h') \in X_h} |||(\nabla v^\sigma - w_h', v^\sigma - u_h')|||_{\lambda_\sigma}, \tag{4.33}$$

and we have that $C_e(\delta, \lambda_\sigma, n) \leq C_e(\delta, \lambda, n)$ for all $\sigma \in (0, 1)$. In particular, in view of Remark 4.3.3, we have boundedness of the sequence of numerical approximations in the sense that

$$|||(w_h^\sigma, v_h^\sigma)|||_{\lambda_\sigma} \leq C(\delta, \lambda, n) \sup_{\sigma \in (0,1)} |||(\nabla v^\sigma, v^\sigma)|||_{\lambda_\sigma},$$

uniformly with respect to h and σ.

For a shape-regular triangulation $\mathcal{T}_h$ on Y, denoting the Lagrange finite element space of degree $q \in \mathbb{N}$ over the triangulation by $\mathcal{S}^q(\mathcal{T}_h)$, we obtain the following approximation result:

Theorem 4.3.3 ([50, Theorem 3.9] Error bound for approximate corrector)**.** *For $\sigma \in (0,1)$, if we have $v^\sigma \in H^{2+r}(Y)$ for some $r \geq 0$ and the choice*

$$X_h := (\mathcal{S}^q(\mathcal{T}_h;\mathbb{R}^n) \cap W_{\text{per}}(Y;\mathbb{R}^n)) \times \left(\mathcal{S}^l(\mathcal{T}_h) \cap H^1_{\text{per}}(Y)\right)$$

for some $q, l \in \mathbb{N}$ and a shape-regular triangulation $\mathcal{T}_h$ on Y (consistent with the requirement of periodicity), we find that

$$|||(\nabla v^\sigma - w_h^\sigma, v^\sigma - v_h^\sigma)|||_{\lambda_\sigma} \leq C h^{\min\{r,q,l\}} \|\nabla v^\sigma\|_{H^{1+r}(Y)}$$

for $h > 0$ sufficiently small, with a constant $C > 0$ only depending on δ, λ, n and interpolation constants.

Proof. Using the definition of the $|||\cdot|||_{\lambda_\sigma}$ norm and interpolation inequalities, denoting the interpolation operators on the finite element spaces by $\mathcal{I}_h^{\mathcal{S}^q}, \mathcal{I}_h^{\mathcal{S}^l}$, we find

$$\begin{aligned}
&\inf_{(w_h', u_h') \in X_h} |||(\nabla v^\sigma - w_h', v^\sigma - u_h')|||_{\lambda_\sigma} \\
&\leq \left|\left|\left| \left(\nabla v^\sigma - \left(\mathcal{I}_h^{\mathcal{S}^q}(\nabla v^\sigma) - \int_Y \mathcal{I}_h^{\mathcal{S}^q}(\nabla v^\sigma) \right), v^\sigma - \mathcal{I}_h^{\mathcal{S}^l}(v^\sigma) \right) \right|\right|\right|_{\lambda_\sigma} \\
&= \left[\left\| D\left(\nabla v^\sigma - \mathcal{I}_h^{\mathcal{S}^q}(\nabla v^\sigma)\right) \right\|_{L^2(Y)}^2 + 2\lambda_\sigma |v^\sigma - \mathcal{I}_h^{\mathcal{S}^l}(v^\sigma)|_{H^1(Y)}^2 + \lambda_\sigma^2 \|v^\sigma - \mathcal{I}_h^{\mathcal{S}^l}(v^\sigma)\|_{L^2(Y)}^2 \right]^{\frac{1}{2}} \\
&\leq C_i \left(h^{2\min\{r,q\}} + 2\lambda_\sigma h^{2\min\{1+r,l\}} + \lambda_\sigma^2 h^{2\min\{2+r,l\}} \right)^{\frac{1}{2}} \|\nabla v^\sigma\|_{H^{1+r}(Y)} \\
&\leq C_i \left(1 + 2\lambda_\sigma + \lambda_\sigma^2\right)^{\frac{1}{2}} h^{\min\{r,q,l\}} \|\nabla v^\sigma\|_{H^{1+r}(Y)}
\end{aligned}$$

for $h > 0$ sufficiently small, where $C_i > 0$ is the constant arising in applying the interpolation inequalities. The claimed result now follows from (4.33), i.e.,

$$\begin{aligned}
|||(\nabla v^\sigma - w_h^\sigma, v^\sigma - v_h^\sigma)|||_{\lambda_\sigma} &\leq C_e(\delta, \lambda_\sigma, n) \inf_{(w_h', u_h') \in X_h} |||(\nabla v^\sigma - w_h', v^\sigma - u_h')|||_{\lambda_\sigma} \\
&\leq C_e(\delta, \lambda_\sigma, n) C_i \left(1 + \lambda_\sigma\right) h^{\min\{r,q,l\}} \|\nabla v^\sigma\|_{H^{1+r}(Y)} \\
&\leq C_e(\delta, \lambda, n) C_i \left(1 + \lambda\right) h^{\min\{r,q,l\}} \|\nabla v^\sigma\|_{H^{1+r}(Y)},
\end{aligned}$$

where we used $\lambda_\sigma \leq \lambda$ and Remark 4.2.2. □

Remark 4.3.5. *The proof yields that the error constant can be taken to be*

$$C := C_e(\delta, \lambda, n) C_i (1 + \lambda),$$

where C_i is a constant arising from interpolation inequalities.

4.3.4 Approximation of the effective Hamiltonian

The approximation of the approximate corrector from the previous section allows us to obtain an approximation to the effective Hamiltonian as follows.

First, we note that with $\tilde{\alpha} \in (0,1)$ from Remark 4.3.3 we have that, for any $r \in [0, \tilde{\alpha})$, there holds

$$\sup_{\sigma\in(0,1)} \|\nabla v^\sigma(\cdot\,; x,p,R)\|_{H^{1+r}(Y)} \lesssim \sup_{\sigma\in(0,1)} \|\nabla v^\sigma(\cdot\,; x,p,R)\|_{C^{1,\tilde{\alpha}}(\mathbb{R}^n)} \lesssim 1 + |p| + |R|,$$

uniformly in σ. Using the error bound from Theorem 4.3.3, we deduce that

$$|||(\nabla v^\sigma - w_h^\sigma, v^\sigma - v_h^\sigma)|||_{\lambda_\sigma} \lesssim h^{\min\{r,q,l\}} \|\nabla v^\sigma\|_{H^{1+r}(Y)} \lesssim h^{\min\{r,q,l\}}(1 + |p| + |R|)$$

with constants independent of σ and the choice of (x,p,R). In particular, by definition of $|||\cdot|||_{\lambda_\sigma}$, we have

$$\|\sigma v^\sigma - \sigma v_h^\sigma\|_{L^2(Y)} \lesssim h^{\min\{r,q,l\}}(1 + |p| + |R|). \tag{4.34}$$

We then define the approximated effective Hamiltonian as

$$H_{\sigma,h} : \bar{\Omega} \times \mathbb{R}^n \times \mathbb{R}^{n\times n}_{\text{sym}} \to \mathbb{R}, \qquad H_{\sigma,h}(x,p,R) := -\sigma \int_Y v_h^\sigma(\cdot\,; x,p,R). \tag{4.35}$$

Then, the following approximation result holds:

Theorem 4.3.4 ([50, Theorem 3.11] Approximation of the effective Hamiltonian)**.** *Let $\sigma \in (0,1)$ and $(w_h^\sigma, v_h^\sigma) \in X_h$ as in Theorem 4.3.3. Further let $H_{\sigma,h}$ be defined as in (4.35). Then, for $(x,p,R) \in \bar{\Omega} \times \mathbb{R}^n \times \mathbb{R}^{n\times n}_{\text{sym}}$, we have the error bound*

$$|H_{\sigma,h}(x,p,R) - H(x,p,R)| \lesssim (h^r + \sigma)\,(1 + |p| + |R|)$$

for any $r \in [0, \tilde{\alpha})$ with $\tilde{\alpha} \in (0,1)$ from Remark 4.3.3 and $\sigma, h > 0$ sufficiently small. More generally, for fixed $(x,p,R) \in \bar{\Omega} \times \mathbb{R}^n \times \mathbb{R}^{n\times n}_{\text{sym}}$, we have

$$|H_{\sigma,h}(x,p,R) - H(x,p,R)| = \mathcal{O}\left(h^{\min\{r,q,l\}} + \sigma\right)$$

for any $r \geq 0$ such that $\{\|\nabla v^\sigma(\cdot\,; x,p,R)\|_{H^{1+r}(Y)}\}_{\sigma\in(0,1)}$ is uniformly bounded.

Proof. We use Hölder and triangle inequalities, Lemma 4.3.1 and the error bound (4.34) to obtain

$$\begin{aligned}\left|\int_Y (-\sigma v_h^\sigma(\cdot\,; x,p,R)) - H(x,p,R)\right| &= \left|\int_Y (-\sigma v_h^\sigma(\cdot\,; x,p,R) - H(x,p,R))\right| \\ &\leq \|-\sigma v_h^\sigma(\cdot\,; x,p,R) - H(x,p,R)\|_{L^2(Y)} \\ &\lesssim \sigma\|v_h^\sigma(\cdot\,; x,p,R) - v^\sigma(\cdot\,; x,p,R)\|_{L^2(Y)} \\ &\qquad + \sigma(1 + |p| + |R|) \\ &\lesssim \left(h^{\min\{r,q,l\}} + \sigma\right)(1 + |p| + |R|).\end{aligned}$$

The second part of the claim can be shown analogously. ∎

4.4 Numerical experiments

4.4.1 Set-up

We consider the problem of approximating the solution u_ε to the HJB equation

$$\begin{cases} \sup_{\alpha\in\Lambda} \left\{ -A^\alpha\left(\frac{\cdot}{\varepsilon}\right) : D^2 u_\varepsilon + u_\varepsilon - 1 \right\} = 0 & \text{in } \Omega, \\ u_\varepsilon = 0 & \text{on } \partial\Omega, \end{cases}$$

where $\Omega := (0,1)^2 \subseteq \mathbb{R}^2$ is the unit square and $\Lambda := [0,1]$. The coefficient A has the structure

$$A : \mathbb{R}^2 \times \Lambda \to \mathbb{R}^{2\times 2}_{\mathrm{sym}}, \qquad A(y,\alpha) := A^\alpha(y) := (a_0(y) + \alpha a_1(y))\, B$$

for Y-periodic functions $a_0, a_1 : \mathbb{R}^2 \to (0,\infty)$ and a symmetric positive definite matrix $B \in \mathbb{R}^{2\times 2}_{\mathrm{sym}}$. The homogenized problem (4.28) is then given by

$$\begin{cases} u_0 + H(D^2 u_0) = 0 & \text{in } \Omega, \\ u_0 = 0 & \text{on } \partial\Omega, \end{cases}$$

and an explicit expression for the effective Hamiltonian $H : \mathbb{R}^{2\times 2}_{\mathrm{sym}} \to \mathbb{R}$ according to [45] is given by

$$H(R) = \max\left\{ -\left(\int_Y \frac{1}{a_0}\right)^{-1} B : R, -\left(\int_Y \frac{1}{a_0 + a_1}\right)^{-1} B : R \right\} - 1. \tag{4.36}$$

Explicitly, we choose in our numerical experiments

$$B := \begin{pmatrix} 2 & -1 \\ -1 & 4 \end{pmatrix}, \quad a_0 \equiv 1, \quad a_1(y_1, y_2) := \sin^2(2\pi y_1)\cos^2(2\pi y_2) + 1.$$

4.4.2 Approximation of the effective Hamiltonian in a point

Our objective in the first numerical experiment is to investigate the approximation of the effective Hamiltonian $H(R)$ by the numerically computed approximate Hamiltonian $H_{\sigma,h}(R)$ at some given point $R \in \mathbb{R}^{2\times 2}_{\mathrm{sym}}$. We choose

$$R := \begin{pmatrix} -2 & 1 \\ 1 & -3 \end{pmatrix}$$

as a negative definite matrix so that the maximum in (4.36) is realized by the term involving the harmonic mean of $a_0 + a_1$ (i.e., the term involving $\left[\int_Y (a_0 + a_1)^{-1}\right]^{-1}$). For our discretization, we choose a first-order discretization with $q = l = 1$ and

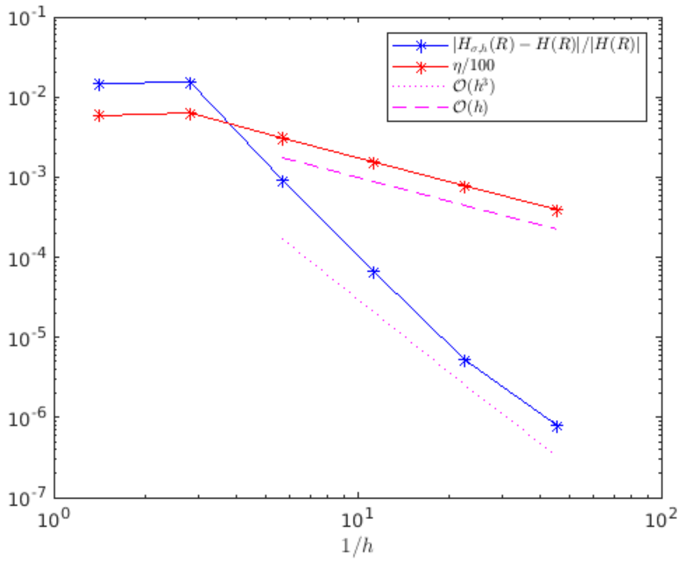


Figure 4.1: Approximation of $H(R)$ by $H_{\sigma,h}(R)$ under mesh refinement with fixed $\sigma = 0.01$.

$M_h := \{0\}$. In order to compare the experimental results with the theoretical bound of Theorem 4.3.4, we consider convergence in h and σ separately. We test convergence with respect to h by fixing a (sufficiently small) value $\sigma = 0.01$ and choosing uniform mesh-refinement of the periodicity cell $Y = (0,1)^2$. Since the error bound for the approximate corrector from Theorem 4.3.3 is given in the norm $|||\cdot|||_{\lambda_\sigma}$, we first numerically test the convergence rate predicted by Theorem 4.3.3. The exact approximate corrector v^σ is unknown, and thus we instead compute the *a posteriori* error estimator

$$\eta(h) := \|F_\gamma[(w_h, u_h)]\|^2_{L^2(Y)} + \sigma_1 \|\mathrm{rot}(w_h)\|^2_{L^2(Y)} + \sigma_2 \|w_h - \nabla u_h\|^2_{L^2(Y)},$$

which is, up to a constant factor, equivalent to the error in Theorem 4.3.3; see Theorem 4.2.3 and Remark 4.2.3. The convergence histories of $\frac{\eta}{100}$ and the relative error

$$\frac{|H_{\sigma,h}(R) - H(R)|}{|H(R)|}$$

are displayed in Figure 4.1. As we are mainly interested in the rate of convergence, we plot $\frac{\eta}{100}$ so that both error quantities can be shown in the same diagram.

As expected from Theorem 4.3.3, the error estimator is of order $\mathcal{O}(h)$, whereas we observe cubic convergence $\mathcal{O}(h^3)$ for the relative error of the effective Hamiltonian at the point R. This rate is higher than predicted by Theorem 4.3.4 , which is based on an error estimate in the norm $|||\cdot|||_{\lambda_\sigma}$ and is therefore indeed expected to overestimate

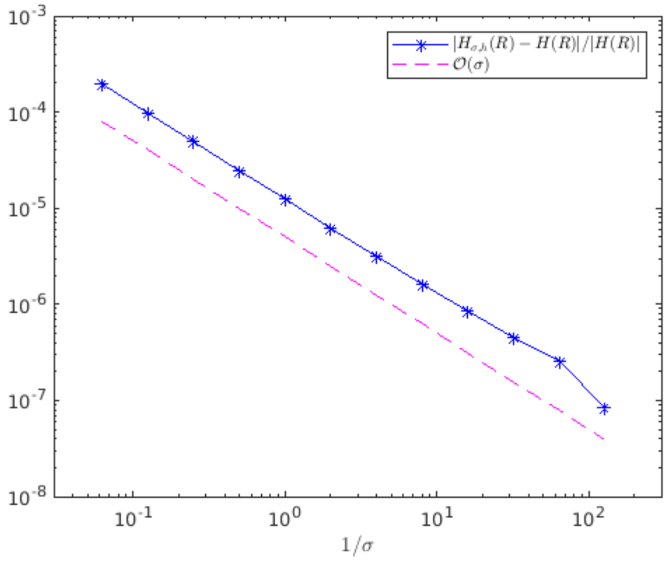


Figure 4.2: Approximation of $H(R)$ by $H_{\sigma,h}(R)$ for varying σ with fixed mesh size $h = \sqrt{2} \times 2^{-7}$.

the actual error between $H_{\sigma,h}(R)$ and $H(R)$ related to the weaker integral functional from (4.35).

Next, we test convergence with respect to σ by fixing a fine mesh size $h = \sqrt{2} \times 2^{-7}$ and letting σ vary from 2^4 to 2^{-7}. The convergence history of the relative error is displayed in Figure 4.2. We observe linear convergence with respect to σ, which indicates that the bound in Theorem 4.3.4 is sharp in σ.

4.4.3 Approximation of the homogenized problem

The second numerical experiment is devoted to the approximation of the effective problem (4.28). We first note that the discretization on the scales Ω and Y leads to a two-scale approach. We denote the triangulation of Ω by $\mathcal{T}_h^\Omega$ with mesh size h_Ω and the triangulation of Y by $\mathcal{T}_h^Y$ with mesh size h_Y. In view of the regularity result from Remark 4.3.2, we discretize the solution u_0 of this fully nonlinear equation by a least-squares approach, which is explained in the following. We discretize functions over Ω with continuous piecewise affine finite elements $\mathcal{S}_0^1(\mathcal{T}_h^\Omega)$ satisfying a homogeneous Dirichlet boundary condition, and their gradients by vector-valued continuous piecewise affine finite elements $\mathcal{S}^1(\mathcal{T}_h^\Omega;\mathbb{R}^2)$.

Given $w_h^\Omega \in \mathcal{S}^1(\mathcal{T}_h^\Omega;\mathbb{R}^2)$, we say that Dw_h^Ω is the discrete Hessian of some $u_h^\Omega \in$

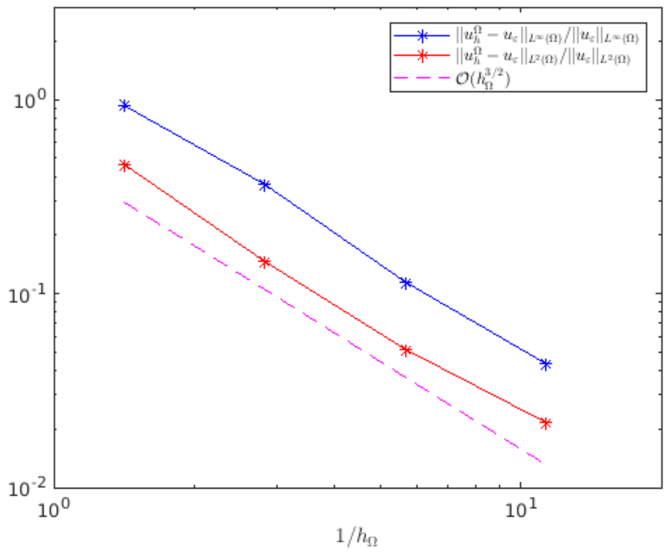


Figure 4.3: Convergence history under mesh-refinement of Ω for the approximation of the solution u_0 to the effective equation. The reference solution u_ϵ is computed for $\varepsilon = 0.1$. The cell problem is solved with $h_Y = \sqrt{2} \times 2^{-2}$ and $\sigma = 0.1$.

$\mathcal{S}_0^1(\mathcal{T}_h^\Omega)$ if it satisfies

$$\int_\Omega w_h^\Omega \cdot v = \int_\Omega \nabla u_h^\Omega \cdot v \qquad \forall v \in \mathcal{S}^1(\mathcal{T}_h^\Omega; \mathbb{R}^2)$$

and write $D_h^2 u_h^\Omega := D w_h^\Omega$. The discrete Hessian $D_h^2 u_h^\Omega$ is expected to be discontinuous across the element boundaries. In order to define a function that represents the evaluation of the discretized approximate Hamiltonian H_{σ,h_Y} at $D_h^2 u_h^\Omega$, we define the continuous and piecewise affine function $\tilde{H}_{\sigma,h_Y}(D^2 u_h^\Omega)$ by nodal averaging of the piecewise constant function

$$x \mapsto H_{\sigma,h_Y}(\mathrm{mid}(T)) \quad \text{for } T \in \mathcal{T}_h^\Omega \text{ with } x \in T$$

(defined a.e. in Ω) where $\mathrm{mid}(T)$ denotes the barycenter of T. We then define the numerical approximation $u_h^\Omega = u_h^\Omega(h_\Omega, \sigma, h_Y)$ as a minimizer of the following least-squares functional

$$u_h^\Omega \in \operatorname*{arg\,min}_{v_h^\Omega \in \mathcal{S}_0^1(\mathcal{T}_h^\Omega)} \|v_h^\Omega + \tilde{H}_{\sigma,h_Y}(D_h^2 v_h^\Omega)\|_{L^2(\Omega)}^2.$$

We choose $\sigma = 0.1$ and $h_Y = \sqrt{2} \times 2^{-2}$ fixed and consider a sequence of uniformly refined triangulations of Ω with mesh sizes $h^\Omega \in \sqrt{2} \times 2^{-\{1,2,3,4\}}$. For the error computation, we use as reference solution the approximation of u_ε with $\varepsilon = 0.1$ on a triangulation with mesh-size $\sqrt{2} \times 2^{-7}$. The convergence history of the errors in the L^∞ and L^2 norms is displayed in Figure 4.3. For both error norms we observe a convergence order of $\mathcal{O}(h_\Omega^{3/2})$, which indicates that the effective problem with the chosen data is possibly more regular than predicted in Remark 4.3.2.

Chapter 5

Numerical homogenization of HJB–Isaacs equations

This chapter discusses a numerical homogenization scheme for Hamilton–Jacobi–Bellman–Isaacs (HJBI) equations based on discontinuous Galerkin (DG) and C^0 interior penalty (C^0-IP) finite element approximations of the approximate corrector problems. This chapter is structured as follows.

In Section 5.1 we study the DG and C^0-IP finite element approximation of periodic HJBI cell problems. We start by proving well-posedness of the problem in a suitable Cordes framework and proceed by discussing discretization aspects. We perform an *a posteriori* analysis which is independent of the choice of numerical scheme and relies on what we refer to as the periodic enrichment of finite element functions. Afterwards an abstract *a priori* error analysis is given and applied to a particular family of numerical schemes.

In Section 5.2 we present the approximation scheme for the effective Hamiltonian based on finite element approximations of approximate correctors. After defining the effective Hamiltonian corresponding to ergodic HJBI operators, we prove an error bound for the DG/C^0-IP approximation of the approximate corrector and perform a rigorous error analysis for the numerical effective Hamiltonian.

Finally, in Section 5.3 we illustrate the theoretical results through numerical experiments.

Annotation: Unless stated otherwise, this chapter contains novel results which have been obtained in Kawecki, Sprekeler [69]. The presented theory was co-developed by E. L. Kawecki. I would like to thank D. Gallistl for some very useful discussions.

5.1 Periodic HJBI cell problems: DG and C^0-IP schemes

5.1.1 Framework

The framework is the natural generalization of the one for HJB equations presented in Section 4.1.

We let $\mathcal{A}, \mathcal{B}$ be compact metric spaces and write $Y := (0,1)^n$ for the unit cell in $\mathbb{R}^n$. We work in dimension $n \in \{2,3\}$ and write

$$\varphi^{\alpha\beta}(y) := \varphi(y,\alpha,\beta), \qquad (y,\alpha,\beta) \in \mathbb{R}^n \times \mathcal{A} \times \mathcal{B}$$

for functions $\varphi : \mathbb{R}^n \times \mathcal{A} \times \mathcal{B} \to \mathcal{R}$ with $\mathcal{R} \in \{\mathbb{R}, \mathbb{R}^n, \mathbb{R}^{n\times n}_{\mathrm{sym}}\}$. We study the periodic Hamilton–Jacobi–Bellman–Isaacs (HJBI) problem

$$\begin{cases} F[u] := \inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} \left\{ -A^{\alpha\beta} : D^2 u - b^{\alpha\beta} \cdot \nabla u + c^{\alpha\beta} u - f^{\alpha\beta} \right\} = 0 \quad \text{in } Y, \\ \hfill u \text{ is } Y\text{-periodic}, \end{cases} \tag{5.1}$$

with given uniformly continuous coefficient functions

$$A : \mathbb{R}^n \times \mathcal{A} \times \mathcal{B} \to \mathbb{R}^{n\times n}_{\mathrm{sym}}, \quad b : \mathbb{R}^n \times \mathcal{A} \times \mathcal{B} \to \mathbb{R}^n, \quad c, f : \mathbb{R}^n \times \mathcal{A} \times \mathcal{B} \to \mathbb{R}.$$

We assume that $A^{\alpha\beta}, b^{\alpha\beta}, c^{\alpha\beta}, f^{\alpha\beta}$ are Y-periodic in $\mathbb{R}^n$ for fixed $(\alpha,\beta) \in \mathcal{A} \times \mathcal{B}$ and that $c > 0$ in $\mathbb{R}^n \times \mathcal{A} \times \mathcal{B}$. Finally, we assume uniform ellipticity, i.e.,

$$\exists\, \zeta_1, \zeta_2 > 0 : \qquad \zeta_1 |\xi|^2 \le A^{\alpha\beta}(y)\xi \cdot \xi \le \zeta_2 |\xi|^2 \quad \forall y, \xi \in \mathbb{R}^n, (\alpha,\beta) \in \mathcal{A} \times \mathcal{B},$$

and that the (generalized) Cordes condition

$$\frac{|A|^2 + \frac{|b|^2}{2\lambda} + \frac{c^2}{\lambda^2}}{\left(\mathrm{tr}(A) + \frac{c}{\lambda}\right)^2} \le \frac{1}{n+\delta}$$

holds in $\mathbb{R}^n \times \mathcal{A} \times \mathcal{B}$ for some constants $\lambda > 0$ and $\delta \in (0,1)$.

The renormalized problem

As for HJB equations, let us introduce the (positive) function $\gamma = \gamma(y,\alpha,\beta) \in C(\mathbb{R}^n \times \mathcal{A} \times \mathcal{B})$ defined by

$$\gamma := \frac{\mathrm{tr}(A) + \frac{c}{\lambda}}{|A|^2 + \frac{|b|^2}{2\lambda} + \frac{c^2}{\lambda^2}}$$

and consider the renormalized HJBI equation

$$\begin{cases} F_\gamma[u] := \inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} \left\{\gamma^{\alpha\beta}\left(-A^{\alpha\beta} : D^2u - b^{\alpha\beta}\cdot\nabla u + c^{\alpha\beta}u - f^{\alpha\beta}\right)\right\} = 0 \quad \text{in } Y, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad u \text{ is } Y\text{-periodic}, \end{cases} \tag{5.2}$$

which is equivalent to the original problem (5.1) in the sense that the analogue of Remark 4.1.3 holds. Noting that

$$\left|\inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} x^{\alpha\beta} - \inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} y^{\alpha\beta}\right| \leq \sup_{(\alpha,\beta)\in\mathcal{A}\times\mathcal{B}} |x^{\alpha\beta} - y^{\alpha\beta}|$$

for any bounded sets $\{x^{\alpha\beta}\}_{(\alpha,\beta)\in\mathcal{A}\times\mathcal{B}} \subset \mathbb{R}$ and $\{y^{\alpha\beta}\}_{(\alpha,\beta)\in\mathcal{A}\times\mathcal{B}} \subset \mathbb{R}$, we make the following key observation (see [67]):

Remark 5.1.1. *The result of Lemma 4.1.1 also holds in this framework, i.e., for any open subset $\omega \subset Y$ and any $u_1, u_2 \in H^2(\omega)$, writing $\delta_u := u_1 - u_2$, we have that*

$$|F_\gamma[u_1] - F_\gamma[u_2] - L_\lambda\delta_u| \leq \sqrt{1-\delta}\sqrt{|D^2\delta_u|^2 + 2\lambda|\nabla\delta_u|^2 + \lambda^2\delta_u^2}$$

almost everywhere in ω, where $L_\lambda u := \lambda u - \Delta u$.

Well-posedness

Therefore, we have existence and uniqueness of strong solutions to the periodic HJBI problem (5.1) analogously to Section 4.1.2:

Theorem 5.1.1 (Well-posedness)**.** *In the situation described above, there exists a unique periodic strong solution $u \in H^2_{\mathrm{per}}(Y)$ to the HJBI problem* (5.1)*. Further, we have the bound*

$$\|L_\lambda u\|_{L^2(Y)} = \left(\|D^2u\|^2_{L^2(Y)} + 2\lambda\|\nabla u\|^2_{L^2(Y)} + \lambda^2\|u\|^2_{L^2(Y)}\right)^{\frac{1}{2}} \leq \frac{\|F_\gamma[0]\|_{L^2(Y)}}{1-\sqrt{1-\delta}}.$$

It is easily seen that all results from Section 4.2 on the mixed finite element approximation of periodic HJB cell problems can be extended to the HJBI cell problems considered here.

In this section we would like to present a different approach, namely discontinuous Galerkin and C^0 interior penalty methods for the periodic HJBI cell problem (5.1). The method presented here is the periodic adaptation of the one proposed in [67] with a novel *a posteriori* analysis.

5.1.2 Discretization

The partition $\mathcal{T}$

We consider a finite conforming partition $\mathcal{T}$ of the closed unit cell $\bar{Y}$ consisting of closed simplices that can be periodically extended in a Y-periodic fashion to $\mathbb{R}^n$, i.e., we require the discretization to be consistent with the identification of opposite faces by periodicity. We introduce the following mathematical objects associated with the partition $\mathcal{T}$:

(i) Set of faces $\mathcal{F}$ and associated unit normal n_F:
We let $\mathcal{F} := \mathcal{F}^I \cup \mathcal{F}^{BP}$ denote the set of $(n-1)$-dimensional faces, where $\mathcal{F}^I$ is the set of all interior faces of $\mathcal{T}$, and $\mathcal{F}^{BP}$ the set of all boundary face-pairs of $\mathcal{T}$, i.e., the boundary faces upon a periodic identification of opposite faces. For each face $F \in \mathcal{F}$, we associate a fixed choice of unit normal n_F, where we often only write n for simplicity; see Figure 5.1.

(ii) Shape-regularity parameter $\theta_{\mathcal{T}}$ and mesh-size function $h_{\mathcal{T}}$:
We let $\theta_T := \max\{\rho_K^{-1}\mathrm{diam}(K) : K \in \mathcal{T}\}$ with ρ_K being the diameter of the largest ball that can be inscribed in the element $K \in \mathcal{T}$. Further introduce $h_{\mathcal{T}} : \bar{Y} \to \mathbb{R}$ defined via $h_{\mathcal{T}}|_{\mathrm{int}(K)} := h_K := (\mathcal{L}^n(K))^{\frac{1}{n}}$ for all $K \in \mathcal{T}$ and $h_{\mathcal{T}}|_F := h_F := (\mathcal{H}^{n-1}(F))^{\frac{1}{n-1}}$ for all $F \in \mathcal{F}$.

Finite element spaces $V_{\mathcal{T}}^s$

For fixed $\bar{p} \geq 2$, we define the discontinuous Galerkin finite element space $V_{\mathcal{T}}^0$ and the C^0-IP finite element space $V_{\mathcal{T}}^1$ by

$$V_{\mathcal{T}}^0 := \left\{ v_{\mathcal{T}} \in L^2(Y) : v_{\mathcal{T}}|_K \in \mathbb{P}_{\bar{p}} \; \forall K \in \mathcal{T} \right\} \quad \text{and} \quad V_{\mathcal{T}}^1 := V_{\mathcal{T}}^0 \cap H^1_{\mathrm{per}}(Y),$$

where $\mathbb{P}_{\bar{p}}$ denotes the space of polynomials of degree at most $\bar{p}$.

Let us make some comments about the derivatives of functions in the finite element spaces. For a function $v \in V_{\mathcal{T}}^0 \subset \mathrm{BV}(Y)$, we define $\nabla v \in L^1(Y;\mathbb{R}^n)$ to be its approximate derivative, i.e., the density of the absolutely continuous part of its distributional derivative and it can be checked that ∇v coincides with the piecewise gradient over the elements of the partition. For $w \in \mathrm{BV}(Y;\mathbb{R}^n)$, we set $Dw := \nabla w \in L^1(Y;\mathbb{R}^{n\times n})$, i.e., we write Dw to denote the approximate derivative of w.

In particular, as $\nabla v \in \mathrm{BV}(Y;\mathbb{R}^n)$ for $v \in V_{\mathcal{T}}^0 \subset \mathrm{BV}(Y)$, we set $D^2 v := D(\nabla v) \in L^1(Y;\mathbb{R}^{n\times n})$ and observe that this coincides with the piecewise Hessian over the elements of the partition. Let us further define $\Delta v := \mathrm{tr}(D^2 v) \in L^1(Y)$.

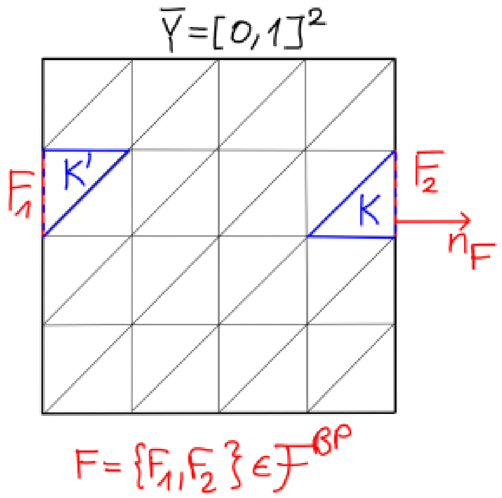


Figure 5.1: Illustration of a boundary face-pair $F \in \mathcal{F}^{BP}$ (left) and an interior face $F \in \mathcal{F}^I$ (right) in dimension $n = 2$.

We then equip the spaces $V_{\mathcal{T}}^s$, $s \in \{0, 1\}$, with the norm

$$\|v_{\mathcal{T}}\|_{\mathcal{T},\lambda}^2 := \int_Y \left(|D^2 v_{\mathcal{T}}|^2 + 2\lambda|\nabla v_{\mathcal{T}}|^2 + \lambda^2 v_{\mathcal{T}}^2\right) + |v_{\mathcal{T}}|_{J,\mathcal{T}}^2,$$
$$|v_{\mathcal{T}}|_{J,\mathcal{T}}^2 := \int_{\mathcal{F}} \left(h_{\mathcal{T}}^{-1}|[\![\nabla v_{\mathcal{T}}]\!]|^2 + h_{\mathcal{T}}^{-3}|[\![v_{\mathcal{T}}]\!]|^2\right)$$

for functions $v_{\mathcal{T}} \in V_{\mathcal{T}}^s$. In order to simplify the presentation, we write $\int_{\mathcal{E}} := \sum_{K\in\mathcal{E}} \int_K$ for collections $\mathcal{E} \subset \mathcal{T}$ of elements and $\int_{\mathcal{G}} := \sum_{F\in\mathcal{G}} \int_F$ for collections $\mathcal{G} \subset \mathcal{F}$ of faces. The jump operator $[\![\cdot]\!]$ is defined in the following paragraph.

Jump and average operators

For elements $K \in \mathcal{T}$, we write $\tau_{\partial K} : \mathrm{BV}(K) \to L^1(\partial K)$ to denote the trace operator. Further, for $v \in \mathrm{BV}(Y)$ we define $\tau_{\partial K} v := \tau_{\partial K}(v|_K)$ for elements $K \in \mathcal{T}$. We then introduce the jump $[\![v]\!]_F$ and the average $\{v\}_F$ of a function $v \in \mathrm{BV}(Y)$ over a face $F = \partial K \cap \partial K' \in \mathcal{F}$ shared by the elements $K, K' \in \mathcal{T}$ by

$$[\![v]\!]_F := \tau_{\partial K} v|_F - \tau_{\partial K'} v|_F \in L^1(F) \quad \text{and} \quad \{v\}_F := \frac{\tau_{\partial K} v|_F + \tau_{\partial K'} v|_F}{2} \in L^1(F),$$

where K, K' are labeled such that the unit normal n_F is the outward normal to K on the face F; see Figure 5.1. To simplify the presentation, we will often simply write $[\![\cdot]\!]$ and $\{\cdot\}$, and drop the subscript.

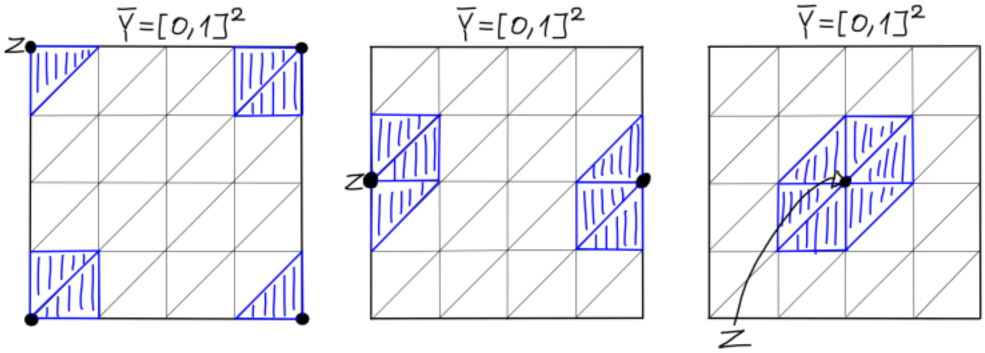


Figure 5.2: Illustration of the periodic neighborhood $N(z) \subset \mathcal{T}$ in dimension $n=2$. *Left*: $z \in \mathcal{Z} \cap \partial Y$ corner point, *middle*: $z \in \mathcal{Z} \cap \partial Y$ non-corner boundary point, *right*: $z \in \mathcal{Z} \cap Y$ interior point.

5.1.3 *A posteriori* analysis via periodic enrichment

Periodic enrichment operators

We let $\mathcal{Z}$ be the set of points in $\bar{Y}$ corresponding to the Lagrange degrees of freedom for the function space $V_{\mathcal{T}}^1 = V_{\mathcal{T}}^0 \cap H^1_{\mathrm{per}}(Y)$, where boundary nodes on ∂Y are identified with all their Y-periodic counterparts. For $z \in \mathcal{Z}$, we then define the *periodic neighborhood* $N(z) \subset \mathcal{T}$ to be the set of all elements $K \in \mathcal{T}$ that contain z or any periodically identical point to z; see Figure 5.2.

Let us introduce an operator

$$E_1 : V_{\mathcal{T}}^0 \to V_{\mathcal{T}}^0 \cap H^1_{\mathrm{per}}(Y),$$

which we call the H^1_{per}-enrichment operator, defined through averaging of the function values in periodic neighborhoods of points in $\mathcal{Z}$. That is, for $v_{\mathcal{T}} \in V_{\mathcal{T}}^0$, we define the function $E_1 v_{\mathcal{T}} \in V_{\mathcal{T}}^1$ by prescribing

$$E_1 v_{\mathcal{T}}(z) := \frac{1}{|N(z)|} \sum_{K \in N(z)} v_{\mathcal{T}}|_K (z)$$

at points $z \in \mathcal{Z}$ (here, $|N(z)|$ denotes the cardinality of the set $N(z)$). Denoting the collection of interior faces and boundary face-pairs neighboring an element $K \in \mathcal{T}$ by $\mathcal{F}_K := \{F \in \mathcal{F} : F \cap K \neq \emptyset\}$, we then have the bound

$$\begin{aligned} &\int_K \left|D^2(v_{\mathcal{T}} - E_1 v_{\mathcal{T}})\right|^2 + \int_K h_{\mathcal{T}}^{-2} \left|\nabla(v_{\mathcal{T}} - E_1 v_{\mathcal{T}})\right|^2 + \int_K h_{\mathcal{T}}^{-4} \left|v_{\mathcal{T}} - E_1 v_{\mathcal{T}}\right|^2 \\ &\lesssim \int_{\mathcal{F}_K} h_{\mathcal{T}}^{-3} |[\![v_{\mathcal{T}}]\!]|^2 \qquad \forall K \in \mathcal{T} \end{aligned} \tag{5.3}$$

for all $v_{\mathcal{T}} \in V_{\mathcal{T}}^0$. This follows from the arguments in [66].

Let us also discuss the periodic enrichment of vector fields. To this end, we define the space containing potential gradients of functions in the finite element spaces by

$$W_{\mathcal{T}} := \{v_{\mathcal{T}} \in L^2(Y;\mathbb{R}^n) : v_{\mathcal{T}}|_K \in \mathbb{P}^n_{\bar{p}-1}\ \forall K \in \mathcal{T}\}.$$

Indeed, observe that $\nabla v_{\mathcal{T}} \in W_{\mathcal{T}}$ for any $v_{\mathcal{T}} \in V^s_{\mathcal{T}}$, $s \in \{0,1\}$. Analogously to E_1, we can then construct a linear operator $E_1^g : W_{\mathcal{T}} \to W_{\mathcal{T}} \cap H^1_{\mathrm{per}}(Y;\mathbb{R}^n)$ satisfying

$$\int_K |D(w_{\mathcal{T}} - E_1^g w_{\mathcal{T}})|^2 + \int_K h_{\mathcal{T}}^{-2}\,|w_{\mathcal{T}} - E_1^g w_{\mathcal{T}}|^2 \lesssim \int_{\mathcal{F}_K} h_{\mathcal{T}}^{-1} |[\![w_{\mathcal{T}}]\!]|^2 \quad \forall K \in \mathcal{T} \tag{5.4}$$

for all $w_{\mathcal{T}} \in W_{\mathcal{T}}$. With the enrichment operators at hand we can proceed with the *a posteriori* analysis, independent of the choice of the numerical scheme.

***A posteriori* analysis**

Let $u \in H^2_{\mathrm{per}}(Y)$ denote the unique solution to the HJBI problem (5.1) and let $v_{\mathcal{T}} \in V^0_{\mathcal{T}}$ be arbitrary. The goal of this section is to estimate the $\|\cdot\|_{\mathcal{T},\lambda}$-distance, i.e.,

$$\|u - v_{\mathcal{T}}\|^2_{\mathcal{T},\lambda} = \int_Y \left(|D^2(u - v_{\mathcal{T}})|^2 + 2\lambda|\nabla(u - v_{\mathcal{T}})|^2 + \lambda^2|u - v_{\mathcal{T}}|^2\right) + |u - v_{\mathcal{T}}|^2_{J,\mathcal{T}}$$

in terms of a computable quantity not depending on the solution u.

It will be useful to introduce some notation from the mixed finite element theory from Chapter 4. Let us consider the function space

$$X := W_{\mathrm{per}}(Y;\mathbb{R}^n) \times H^1_{\mathrm{per}}(Y),$$

which we equip with the $|\!|\!|\cdot|\!|\!|_\lambda$-norm given by

$$|\!|\!|(w',u')|\!|\!|^2_\lambda := \|Dw'\|^2_{L^2(Y)} + 2\lambda\|\nabla u'\|^2_{L^2(Y)} + \lambda^2\|u'\|^2_{L^2(Y)}, \qquad (w',v') \in X.$$

We further define the mixed analogue F_γ^M to the nonlinear operator F_γ by

$$F_\gamma^M[(w',u')] := \inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} \left\{\gamma^{\alpha\beta}\left(-A^{\alpha\beta} : Dw' - b^{\alpha\beta}\cdot\nabla u' + c^{\alpha\beta}u' - f^{\alpha\beta}\right)\right\}$$

for pairs $(w',u') \in X$, and observe that the solution $u \in H^2_{\mathrm{per}}(Y)$ to (5.1) satisfies

$$F_\gamma^M[(\nabla u, u)] = F_\gamma[u] = 0 \quad \text{a.e. in } Y.$$

We have the *a posteriori* bound from Theorem 4.2.3 (identical proof for HJB and HJBI) on the $|\!|\!|\cdot|\!|\!|_\lambda$-distance between the solution pair $(\nabla u, u)$ and an arbitrary pair $(w',u') \in X$:

Lemma 5.1.1 (Mixed *a posteriori* bound). *Let $u \in H^2_{\mathrm{per}}(Y)$ denote the unique solution to the HJBI problem* (5.1)*. Then we have*

$$|||(\nabla u - w', u - u')|||_\lambda^2 \lesssim \|F_\gamma^M[(w',u')]\|_{L^2(Y)}^2 + \|\mathrm{rot}(w')\|_{L^2(Y)}^2 + \|\nabla u' - w'\|_{L^2(Y)}^2$$

for all $(w',u') \in X$, where the constant absorbed in $\lesssim$ only depends on the Cordes parameters δ and λ.

We can use Lemma 5.1.1 and the H^1_{per}-enrichment operators to prove the following *a posteriori* error bound:

Theorem 5.1.2 (*a posteriori* error bound). *Let $u \in H^2_{\mathrm{per}}(Y)$ denote the unique solution to the HJBI problem* (5.1)*. Then there holds*

$$\|u - v_{\mathcal{T}}\|_{\mathcal{T},\lambda}^2 \lesssim \int_Y |F_\gamma[v_{\mathcal{T}}]|^2 + |v_{\mathcal{T}}|_{J,\mathcal{T}}^2 \qquad \forall v_{\mathcal{T}} \in V_{\mathcal{T}}^0$$

with the constant absorbed in $\lesssim$ only depending on $n, \theta_{\mathcal{T}}, \bar{p}$ and the Cordes parameters δ, λ.

Proof. Let $v_{\mathcal{T}} \in V_{\mathcal{T}}^0$ be arbitrary and set

$$\begin{aligned} v &:= E_1 v_{\mathcal{T}} \in V_{\mathcal{T}}^0 \cap H^1_{\mathrm{per}}(Y), \\ w &:= E_1^g(\nabla v_{\mathcal{T}}) - \int_Y E_1^g(\nabla v_{\mathcal{T}}) \in W_{\mathcal{T}} \cap W_{\mathrm{per}}(Y;\mathbb{R}^n). \end{aligned}$$

By the triangle inequality, we have

$$\begin{aligned} \|u - v_{\mathcal{T}}\|_{\mathcal{T},\lambda}^2 \lesssim\ & |||(\nabla u - w, u - v)|||_\lambda^2 \\ & + \int_Y \left(|D(w - \nabla v_{\mathcal{T}})|^2 + 2\lambda|\nabla(v - v_{\mathcal{T}})|^2 + \lambda^2(v - v_{\mathcal{T}})^2\right) + |v_{\mathcal{T}}|_{J,\mathcal{T}}^2, \end{aligned}$$

which we can further bound, using the properties of the enrichment operators (5.3) and (5.4), to obtain that

$$\|u - v_{\mathcal{T}}\|_{\mathcal{T},\lambda}^2 \lesssim |||(\nabla u - w, u - v)|||_\lambda^2 + |v_{\mathcal{T}}|_{J,\mathcal{T}}^2.$$

We can apply Lemma 5.1.1 to find

$$\|u - v_{\mathcal{T}}\|_{\mathcal{T},\lambda}^2 \lesssim \|F_\gamma^M[(w,v)]\|_{L^2(Y)}^2 + \|\mathrm{rot}(w)\|_{L^2(Y)}^2 + \|\nabla v - w\|_{L^2(Y)}^2 + |v_{\mathcal{T}}|_{J,\mathcal{T}}^2. \quad (5.5)$$

Note that, using the triangle and Hölder inequalities, and the enrichment bounds (5.3) and (5.4), we have

$$\|\mathrm{rot}(w)\|_{L^2(Y)}^2 \lesssim \int_Y |\mathrm{rot}(w - \nabla v_{\mathcal{T}})|^2 \lesssim |v_{\mathcal{T}}|_{J,\mathcal{T}}^2$$

for the second term on the right-hand side of (5.5), and

$$\begin{aligned}
\|\nabla v - w\|_{L^2(Y)}^2 &\lesssim \left\| \nabla v - E_1^g(\nabla v_{\mathcal{T}}) - \int_Y (\nabla v - E_1^g(\nabla v_{\mathcal{T}})) \right\|_{L^2(Y)}^2 \\
&\lesssim \|\nabla v - E_1^g(\nabla v_{\mathcal{T}})\|_{L^2(Y)}^2 \\
&\lesssim \int_Y |\nabla(v - v_{\mathcal{T}})|^2 + \int_Y |\nabla v_{\mathcal{T}} - E_1^g(\nabla v_{\mathcal{T}})|^2 \\
&\lesssim |v_{\mathcal{T}}|_{J,\mathcal{T}}^2
\end{aligned}$$

for the third term on the right-hand side of (5.5). Finally, for the first term on the right-hand side of (5.5), we successively use the triangle inequality together with $F_\gamma[v_{\mathcal{T}}] = F_\gamma^M[(\nabla v_{\mathcal{T}}, v_{\mathcal{T}})]$, a Lipschitz property of F_γ^M which is shown analogously to (4.7), and the enrichment bounds (5.3) and (5.4) to obtain

$$\begin{aligned}
&\|F_\gamma^M[(w, v)]\|_{L^2(Y)}^2 \\
&\lesssim \int_Y |F_\gamma[v_{\mathcal{T}}]|^2 + \int_Y \left|F_\gamma^M[(w, v)] - F_\gamma^M[(\nabla v_{\mathcal{T}}, v_{\mathcal{T}})]\right|^2 \\
&\lesssim \int_Y |F_\gamma[v_{\mathcal{T}}]|^2 + \int_Y \left(|D(w - \nabla v_{\mathcal{T}})|^2 + 2\lambda|\nabla(v - v_{\mathcal{T}})|^2 + \lambda^2|v - v_{\mathcal{T}}|^2\right) \\
&\lesssim \int_Y |F_\gamma[v_{\mathcal{T}}]|^2 + |v_{\mathcal{T}}|_{J,\mathcal{T}}^2.
\end{aligned}$$

Altogether, in view of (5.5), we have proved the desired estimate. □

This concludes the *a posteriori* analysis and we proceed with an abstract *a priori* analysis for a wide class of numerical schemes in the next section.

5.1.4 Numerical scheme and *a priori* analysis

Let us consider an abstract numerical scheme written in the following form: For chosen $s \in \{0, 1\}$, find a function $u_{\mathcal{T}} \in V_{\mathcal{T}}^s$ satisfying

$$a_{\mathcal{T}}(u_{\mathcal{T}}, v_{\mathcal{T}}) = 0 \qquad \forall v_{\mathcal{T}} \in V_{\mathcal{T}}^s. \tag{5.6}$$

Abstract *a priori* analysis

Here, we assume that the nonlinear form $a_{\mathcal{T}} : V_{\mathcal{T}}^s \times V_{\mathcal{T}}^s \to \mathbb{R}$ satisfies the assumptions listed below:

(A1) Linearity in second argument: $a_{\mathcal{T}}(w_{\mathcal{T}}, \cdot) : V_{\mathcal{T}}^s \to \mathbb{R}$ is linear for any fixed $w_{\mathcal{T}} \in V_{\mathcal{T}}^s$.

(A2) Strong monotonicity: There exists a constant $C_M > 0$ such that

$$\|w_\mathcal{T} - v_\mathcal{T}\|_{\mathcal{T},\lambda}^2 \leq C_M \left(a_\mathcal{T}(w_\mathcal{T}, w_\mathcal{T} - v_\mathcal{T}) - a_\mathcal{T}(v_\mathcal{T}, w_\mathcal{T} - v_\mathcal{T})\right) \quad \forall w_\mathcal{T}, v_\mathcal{T} \in V_\mathcal{T}^s.$$

(A3) Lipschitz continuity: There exists a constant $C_L > 0$ such that

$$|a_\mathcal{T}(w_\mathcal{T}, v_\mathcal{T}) - a_\mathcal{T}(w'_\mathcal{T}, v_\mathcal{T})| \leq C_L \|w_\mathcal{T} - w'_\mathcal{T}\|_{\mathcal{T},\lambda} \|v_\mathcal{T}\|_{\mathcal{T},\lambda} \qquad \forall w_\mathcal{T}, w'_\mathcal{T}, v_\mathcal{T} \in V_\mathcal{T}^s.$$

(A4) Discrete consistency: There exists a linear operator $L_\mathcal{T} : V_\mathcal{T}^s \to L^2(Y)$ such that, for some constant $C_1 > 0$, we have

$$\|L_\mathcal{T} v_\mathcal{T}\|_{L^2(Y)} \leq C_1 \|v_\mathcal{T}\|_{\mathcal{T},\lambda} \qquad \forall v_\mathcal{T} \in V_\mathcal{T}^s,$$

and, for some constant $C_2 > 0$, we have

$$\left| a(w_\mathcal{T}, v_\mathcal{T}) - \int_Y F_\gamma[w_\mathcal{T}] L_\mathcal{T} v_\mathcal{T} \right| \leq C_2 |w_\mathcal{T}|_{J,\mathcal{T}} \|v_\mathcal{T}\|_{\mathcal{T},\lambda} \qquad \forall w_\mathcal{T}, v_\mathcal{T} \in V_\mathcal{T}^s.$$

Observe that the assumptions (A1)–(A4) guarantee the well-posedness of the numerical scheme, i.e., there exists a unique solution $u_\mathcal{T} \in V_\mathcal{T}^s$ satisfying (5.6). We can show an *a priori* bound in this general setting analogously to [67]. A proof is provided for completeness.

Theorem 5.1.3 (*a priori* error bound)**.** *For chosen $s \in \{0, 1\}$, let $a_\mathcal{T} : V_\mathcal{T}^s \times V_\mathcal{T}^s \to \mathbb{R}$ be a nonlinear form satisfying the assumptions (A1)–(A4). Further, let $u \in H^2_{\mathrm{per}}(Y)$ denote the unique solution to the HJBI problem* (5.1)*. Then, there exists a unique solution $u_\mathcal{T} \in V_\mathcal{T}^s$ to* (5.6) *and we have the near-best approximation bound*

$$\|u - u_\mathcal{T}\|_{\mathcal{T},\lambda} \leq C_e \inf_{v_\mathcal{T} \in V_\mathcal{T}^s} \|u - v_\mathcal{T}\|_{\mathcal{T},\lambda}, \tag{5.7}$$

where the constant $C_e > 0$ is given by

$$C_e := 1 + C_M \left(C_1 \left(\sqrt{1 - \delta} + \sqrt{n + 1} \right) + C_2 \right). \tag{5.8}$$

Proof. As we have already noted, the existence and uniqueness of a solution $u_\mathcal{T} \in V_\mathcal{T}^s$ to (5.6) follows from the assumptions on the nonlinear form $a_\mathcal{T}$, and it only remains to show the near-best approximation bound (5.7). To this end, let $v_\mathcal{T} \in V_\mathcal{T}^s$ be arbitrary and observe that

$$\begin{aligned} \|v_\mathcal{T} - u_\mathcal{T}\|_{\mathcal{T},\lambda}^2 &\leq C_M \left(a_\mathcal{T}(v_\mathcal{T}, v_\mathcal{T} - u_\mathcal{T}) - a_\mathcal{T}(u_\mathcal{T}, v_\mathcal{T} - u_\mathcal{T})\right) \\ &= C_M \, a_\mathcal{T}(v_\mathcal{T}, v_\mathcal{T} - u_\mathcal{T}) \end{aligned} \tag{5.9}$$

by strong monotonicity (A2) and the solution property (5.6) of $u_{\mathcal{T}}$. In order to further bound the right-hand side, we successively use the discrete consistency (A4), the solution property and regularity of u, and the Lipschitz property (4.7) of F_γ to obtain

$$\begin{aligned} a_{\mathcal{T}}(v_{\mathcal{T}}, v_{\mathcal{T}} - u_{\mathcal{T}}) &\le \left| \int_Y F_\gamma[v_{\mathcal{T}}] L_{\mathcal{T}}(v_{\mathcal{T}} - u_{\mathcal{T}}) \right| + C_2 |v_{\mathcal{T}}|_{J,\mathcal{T}} \|v_{\mathcal{T}} - u_{\mathcal{T}}\|_{\mathcal{T},\lambda} \\ &\le \left(C_1 \|F_\gamma[v_{\mathcal{T}}] - F_\gamma[u]\|_{L^2(Y)} + C_2 |v_{\mathcal{T}} - u|_{J,\mathcal{T}} \right) \|v_{\mathcal{T}} - u_{\mathcal{T}}\|_{\mathcal{T},\lambda} \\ &\le \left(C_1 \left(\sqrt{1-\delta} + \sqrt{n+1} \right) + C_2 \right) \|v_{\mathcal{T}} - u\|_{\mathcal{T},\lambda} \|v_{\mathcal{T}} - u_{\mathcal{T}}\|_{\mathcal{T},\lambda}. \end{aligned}$$

Combination with the previous estimate (5.9) yields

$$\|v_{\mathcal{T}} - u_{\mathcal{T}}\|_{\mathcal{T},\lambda} \le C_M \left(C_1 \left(\sqrt{1-\delta} + \sqrt{n+1} \right) + C_2 \right) \|u - v_{\mathcal{T}}\|_{\mathcal{T},\lambda},$$

which in turn implies

$$\|u - u_{\mathcal{T}}\|_{\mathcal{T},\lambda} \le \|u - v_{\mathcal{T}}\|_{\mathcal{T},\lambda} + \|v_{\mathcal{T}} - u_{\mathcal{T}}\|_{\mathcal{T},\lambda} \le C_e \|u - v_{\mathcal{T}}\|_{\mathcal{T},\lambda}$$

with $C_e > 0$ given by (5.8). We conclude the proof by taking the infimum over $v_{\mathcal{T}} \in V_{\mathcal{T}}^s$. □

We conclude this section by noting that Theorem 5.1.3 implies convergence of the numerical approximation under mesh-refinement. While convergence together with optimal rates follows immediately from standard approximation arguments in the case that the exact solution satisfies additional regularity assumptions, it is not that clear when we only have a minimal regularity solution $u \in H^2_{\mathrm{per}}(Y)$. For the latter case, we can argue as in [67, Corollary 4.4].

Remark 5.1.2 (Convergence of the numerical approximation)**.** *For a sequence of conforming simplicial meshes $\{\mathcal{T}_k\}_k$ with $\max_{K \in \mathcal{T}_k} h_K \to 0$ as $k \to \infty$, we have that*

$$\inf_{v_{\mathcal{T}_k} \in V^s_{\mathcal{T}_k}} \|u - v_{\mathcal{T}_k}\|_{\mathcal{T}_k,\lambda} \underset{k\to\infty}{\longrightarrow} 0.$$

In particular, in view of (5.7), given $a_{\mathcal{T}_k} : V^s_{\mathcal{T}_k} \times V^s_{\mathcal{T}_k} \to \mathbb{R}$ satisfying (A1)–(A4) with constants uniformly bounded in k, we have that

$$\|u - u_{\mathcal{T}_k}\|_{\mathcal{T}_k,\lambda} \underset{k\to\infty}{\longrightarrow} 0$$

for the sequence of numerical approximations $\{u_{\mathcal{T}_k}\}_k \subset V^s_{\mathcal{T}_k}$.

The numerical scheme

For chosen $s \in \{0,1\}$ and a parameter $\theta \in [0,1]$, we now consider the numerical scheme of finding $u_{\mathcal{T}} \in V^s_{\mathcal{T}}$ satisfying (5.6) with

$$a_{\mathcal{T}} : V^s_{\mathcal{T}} \times V^s_{\mathcal{T}} \to \mathbb{R}, \quad a_{\mathcal{T}}(w_{\mathcal{T}}, v_{\mathcal{T}}) := \int_Y F_\gamma[w_{\mathcal{T}}] L_{\lambda,\mathcal{T}} v_{\mathcal{T}} + \theta S_{\mathcal{T}}(w_{\mathcal{T}}, v_{\mathcal{T}}) + J_{\mathcal{T}}(w_{\mathcal{T}}, v_{\mathcal{T}}),$$

where we define the linear operator $L_{\lambda,\mathcal{T}} v_{\mathcal{T}} := \lambda v_{\mathcal{T}} - \Delta v_{\mathcal{T}}$ for $v_{\mathcal{T}} \in V^s_{\mathcal{T}}$, the stabilization bilinear form $S_{\mathcal{T}} : V^s_{\mathcal{T}} \times V^s_{\mathcal{T}} \to \mathbb{R}$ via

$$\begin{aligned} &S_{\mathcal{T}}(w_{\mathcal{T}}, v_{\mathcal{T}}) \\ &:= \int_Y \left(D^2 w_{\mathcal{T}} : D^2 v_{\mathcal{T}} - \Delta w_{\mathcal{T}} \Delta v_{\mathcal{T}}\right) + \int_{\mathcal{F}} \left(\{\Delta_T w_{\mathcal{T}}\} [\![\nabla v_{\mathcal{T}} \cdot n]\!] + \{\Delta_T v_{\mathcal{T}}\} [\![\nabla w_{\mathcal{T}} \cdot n]\!]\right) \\ &- \int_{\mathcal{F}} \left(\nabla_T \{\nabla w_{\mathcal{T}} \cdot n\} \cdot [\![\nabla_T v_{\mathcal{T}}]\!] + \nabla_T \{\nabla v_{\mathcal{T}} \cdot n\} \cdot [\![\nabla_T w_{\mathcal{T}}]\!]\right), \end{aligned}$$

and, for chosen parameters $\eta_1, \eta_2 > 0$, the jump penalization form $J_{\mathcal{T}} : V^s_{\mathcal{T}} \times V^s_{\mathcal{T}} \to \mathbb{R}$ via

$$J_{\mathcal{T}}(w_{\mathcal{T}}, v_{\mathcal{T}}) := \eta_1 \int_{\mathcal{F}} h_{\mathcal{T}}^{-1} [\![\nabla w_{\mathcal{T}}]\!] \cdot [\![\nabla v_{\mathcal{T}}]\!] + \eta_2 \int_{\mathcal{F}} h_{\mathcal{T}}^{-3} [\![w_{\mathcal{T}}]\!] [\![v_{\mathcal{T}}]\!].$$

Here, the tangential gradient and Laplacian on mesh faces are denoted by ∇_T and Δ_T.

This scheme is an adaptation of the method presented in [67] for the homogeneous Dirichlet problem. The analysis of this method, i.e., the verification of the assumptions (A1)–(A4), is analogous to [67] and hence omitted. The main result is the following:

Theorem 5.1.4. *There exist constants $\bar{\eta}_1, \bar{\eta}_2 > 0$, depending only on $n, \theta_{\mathcal{T}}, \bar{p}$ and the Cordes parameters δ, λ, such that, for any $\theta \in [0,1]$, if $\eta_1 \geq \bar{\eta}_1$ and $\eta_2 \geq \bar{\eta}_2$, the properties (A1)–(A4) are satisfied and Theorem 5.1.3 applies.*

Remark 5.1.3. *The constants $\bar{\eta}_1, \bar{\eta}_2$ and the constant C_e in the near-best approximation bound* (5.7) *remain bounded as $\lambda \searrow 0$.*

5.2 Approximation of effective Hamiltonians to HJBI operators

5.2.1 The effective Hamiltonian

We start by recalling the definition of the effective Hamiltonian based on the *cell σ-problem*; see [10, 11, 12].

Let us consider an HJBI operator $F : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{n\times n}_{\mathrm{sym}} \to \mathbb{R}$ given by

$$F(x,y,p,R) := \inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} \left\{ -A^{\alpha\beta}(y) : R - b^{\alpha\beta}(x,y)\cdot p - f^{\alpha\beta}(x,y) \right\} \tag{5.10}$$

with $\mathcal{A}$ and $\mathcal{B}$ denoting compact metric spaces, and functions

$$\begin{aligned} A = (a_{ij})_{1\le i,j\le n} : \mathbb{R}^n \times \mathcal{A} \times \mathcal{B} \to \mathbb{R}^{n\times n}_{\mathrm{sym}}, \qquad & (y,\alpha,\beta) \mapsto A(y,\alpha,\beta) =: A^{\alpha\beta}(y), \\ b = (b_i)_{1\le i\le n} : \mathbb{R}^n \times \mathbb{R}^n \times \mathcal{A} \times \mathcal{B} \to \mathbb{R}^n, \quad & (x,y,\alpha,\beta) \mapsto b(x,y,\alpha,\beta) =: b^{\alpha\beta}(x,y), \\ f : \mathbb{R}^n \times \mathbb{R}^n \times \mathcal{A} \times \mathcal{B} \to \mathbb{R}, \quad & (x,y,\alpha,\beta) \mapsto f(x,y,\alpha,\beta) =: f^{\alpha\beta}(x,y) \end{aligned}$$

satisfying the assumptions stated in the paragraph at the end of this subsection.

To the HJBI operator (5.10), we associate the corresponding *cell σ-problem*: For fixed $(x,p,R) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{n\times n}_{\mathrm{sym}}$ and a positive parameter $\sigma > 0$, seek the unique viscosity solution $v^\sigma = v^\sigma(\cdot\,; x,p,R) \in C(\mathbb{R}^n)$ to the problem

$$\begin{cases} \sigma v^\sigma + F(x,y,p,R + D_y^2 v^\sigma) = 0 \quad \text{for } y \in Y, \\ \qquad\qquad y \mapsto v^\sigma(y;x,p,R) \text{ is } Y\text{-periodic.} \end{cases} \tag{5.11}$$

The function $v^\sigma(\cdot\,;x,p,R)$ is called an *approximate corrector*.

Definition 5.2.1 (Ergodicity and effective Hamiltonian)**.** *Let $F : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{n\times n}_{\mathrm{sym}} \to \mathbb{R}$ be an HJBI operator of the form* (5.10)*.*

(i) We say F is ergodic (in the y-variable) at a point $(x,p,R) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{n\times n}_{\mathrm{sym}}$ if there exists a constant $H(x,p,R) \in \mathbb{R}$ such that

$$-\sigma v^\sigma(\cdot\,;x,p,R) \underset{\sigma\searrow 0}{\longrightarrow} H(x,p,R) \quad \textit{uniformly.} \tag{5.12}$$

Further, we call F ergodic if it is ergodic at every $(x,p,R) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{n\times n}_{\mathrm{sym}}$.

(ii) If F is ergodic, we call the function

$$H : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{n\times n}_{\mathrm{sym}} \to \mathbb{R}, \qquad (x,p,R) \mapsto H(x,p,R)$$

defined via (5.12) *the effective Hamiltonian corresponding to F.*

The assumptions on the coefficients in the following paragraph are such that the HJBI operator (5.10) fits into the framework considered in [12], which guarantees ergodicity. The corresponding effective Hamiltonian $H : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{n\times n}_{\mathrm{sym}} \to \mathbb{R}$ is automatically continuous and degenerate elliptic, that is,

$$R_1 - R_2 \ge 0 \quad \Longrightarrow \quad H(x,p,R_1) \le H(x,p,R_2).$$

for any $x,p \in \mathbb{R}^n$, $R_1, R_2 \in \mathbb{R}^{n\times n}_{\mathrm{sym}}$.

Remark 5.2.1. *In the periodic homogenization of elliptic and parabolic HJBI equations*

$$L^{\text{elliptic}} u_\varepsilon^{\text{e}} := u_\varepsilon^{\text{e}} + F\left(x, \frac{x}{\varepsilon}, \nabla u_\varepsilon^{\text{e}}, D^2 u_\varepsilon^{\text{e}}\right) = 0,$$
$$L^{\text{parabolic}} u_\varepsilon^{\text{p}} := \partial_t u_\varepsilon^{\text{p}} + F\left(x, \frac{x}{\varepsilon}, \nabla_x u_\varepsilon^{\text{p}}, D_x^2 u_\varepsilon^{\text{p}}\right) = 0,$$

posed in a suitable Dirichlet/Cauchy setting, the effective Hamiltonian determines the homogenized equation

$$\bar{L}^{\text{elliptic}} u_0^{\text{e}} := u_0^{\text{e}} + H\left(x, \nabla u_0^{\text{e}}, D^2 u_0^{\text{e}}\right) = 0,$$
$$\bar{L}^{\text{parabolic}} u_0^{\text{p}} := \partial_t u_0^{\text{p}} + H\left(x, \nabla_x u_0^{\text{p}}, D_x^2 u_0^{\text{p}}\right) = 0;$$

see [12, 39, 40] and Section 4.3.

In this setting, having $A = A(y, \alpha, \beta)$ being independent of the state variable x, it can be shown that

$$|H(x_1, p, R) - H(x_2, p, R)| \leq C|x_1 - x_2|(1 + |p|) + \omega(|x_1 - x_2|)$$

for all $x_1, x_2, p \in \mathbb{R}^n$ and $R \in \mathbb{R}^{n\times n}_{\text{sym}}$, for some constant $C > 0$ and modulus of continuity ω, which guarantees a comparison principle for the effective problem and implies homogenization; see [12].

Assumptions on the coefficients

We assume $A = \frac{1}{2}GG^{\text{T}} \in C(\mathbb{R}^n \times \mathcal{A} \times \mathcal{B}; \mathbb{R}^{n\times n})$, $b \in C(\mathbb{R}^n \times \mathbb{R}^n \times \mathcal{A} \times \mathcal{B}; \mathbb{R}^n)$ and $f \in C(\mathbb{R}^n \times \mathbb{R}^n \times \mathcal{A} \times \mathcal{B}; \mathbb{R})$ satisfy the assumptions listed below. Note that these fit into the framework of [12].

- G, b, f are bounded continuous functions on their respective domains.
- $G = G(y, \alpha, \beta), b = b(x, y, \alpha, \beta)$ are Lipschitz continuous in (x, y), uniformly in (α, β).
- $f = f(x, y, \alpha, \beta)$ is uniformly continuous in (x, y), uniformly in (α, β).
- G, b, f are Y-periodic in the fast variable y.
- Uniform ellipticity: $\exists\, \zeta_1, \zeta_2 > 0 : \quad \zeta_1|\xi|^2 \leq A(y, \alpha, \beta)\xi \cdot \xi \leq \zeta_2|\xi|^2 \quad \forall y, \xi \in \mathbb{R}^n$, $(\alpha, \beta) \in \mathcal{A} \times \mathcal{B}$.
- (Generalized) Cordes condition: There exist $\lambda > 0$ and $\delta \in (0, 1)$ such that
$$|A(y, \alpha, \beta)|^2 + \frac{|b(x, y, \alpha, \beta)|^2}{2\lambda} + \frac{1}{\lambda^2} \leq \frac{1}{n + \delta}\left(\text{tr}(A(y, \alpha, \beta)) + \frac{1}{\lambda}\right)^2 \tag{5.13}$$
for all $(x, y, \alpha, \beta) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathcal{A} \times \mathcal{B}$.

5.2.2 Approximation of the cell σ-problem

For fixed $(x,p,R) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{n\times n}_{\mathrm{sym}}$ and a positive parameter $\sigma \in (0,1)$, let us consider the cell σ-problem (5.11) in the rewritten form

$$\begin{cases} \inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} \left\{ -A^{\alpha\beta} : D^2 v^\sigma + \sigma v^\sigma - g^{\alpha\beta}_{x,p,R} \right\} = 0 \quad \text{in } Y, \\ \qquad\qquad y \mapsto v^\sigma(y;x,p,R) \text{ is } Y\text{-periodic}, \end{cases} \tag{5.14}$$

where $g^{\alpha\beta}_{x,p,R} : \mathbb{R}^n \to \mathbb{R}$ is the Y-periodic function given by

$$g^{\alpha\beta}_{x,p,R}(y) := g_{x,p,R}(y,\alpha,\beta) := A^{\alpha\beta}(y) : R + b^{\alpha\beta}(x,y)\cdot p + f^{\alpha\beta}(x,y)$$

for $y \in \mathbb{R}^n$ and $(\alpha,\beta) \in \mathcal{A}\times\mathcal{B}$. The following Lemma shows that, for any $\sigma > 0$, the problem (5.14) admits a unique strong solution $v^\sigma \in H^2_{\mathrm{per}}(Y)$ and that we have a uniform bound on $|v^\sigma|_{H^2(Y)}$.

Lemma 5.2.1. *Let $(x,p,R) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{n\times n}_{\mathrm{sym}}$ be fixed. Then, for any $\sigma > 0$, there exists a unique periodic strong solution $v^\sigma \in H^2_{\mathrm{per}}(Y)$ to the cell σ-problem (5.14). Further, we have the bound*

$$|v^\sigma|_{H^2(Y)} \le C \tag{5.15}$$

with $C > 0$ independent of σ.

Proof. It is straightforward to check that all assumptions of Theorem 4.1.1 are satisfied. In particular, the problem (5.14) satisfies the Cordes condition

$$|A|^2 + \frac{\sigma^2}{\lambda_\sigma^2} \le \frac{1}{n+\delta}\left(\mathrm{tr}(A) + \frac{\sigma}{\lambda_\sigma}\right)^2 \quad \text{in } \mathbb{R}^n \times \mathcal{A}\times\mathcal{B},$$

where $\lambda_\sigma > 0$ is defined by $\lambda_\sigma := \sigma\lambda$. Therefore, we find that there exists a unique periodic strong solution $v^\sigma \in H^2_{\mathrm{per}}(Y)$ to (5.14). Note that the corresponding renormalization function $\gamma^\sigma \in C(\mathbb{R}^n \times \mathcal{A}\times\mathcal{B})$ (see (4.3)) is given by

$$\gamma^\sigma := \frac{\mathrm{tr}(A) + \frac{\sigma}{\lambda_\sigma}}{|A|^2 + \frac{\sigma^2}{\lambda_\sigma^2}} = \frac{\mathrm{tr}(A) + \frac{1}{\lambda}}{|A|^2 + \frac{1}{\lambda^2}}$$

and hence, $\gamma := \gamma^\sigma$ is independent of σ. The uniform bound (5.15) now follows from Remark 4.1.4. □

Let us make the technical assumption that $v^\sigma \in W^{2,n}_{\mathrm{loc}}(\mathbb{R}^n)$, so that the strong solution coincides with the unique viscosity solution to (5.14); see [24, 73, 74]. This is no restriction when $n = 2$ or when we have an HJB problem as in Section 4.3.2.

The discontinuous Galerkin ($s = 0$) or the C^0-IP ($s = 1$) finite element method from Section 5.1.4 yields an approximation $v_{\mathcal{T}}^\sigma \in V_{\mathcal{T}}^s$ to the problem (5.14) satisfying

$$\|v^\sigma - v_{\mathcal{T}}^\sigma\|_{\mathcal{T},\lambda_\sigma} \leq C \inf_{z_{\mathcal{T}} \in V_{\mathcal{T}}^s} \|v^\sigma - z_{\mathcal{T}}\|_{\mathcal{T},\lambda_\sigma} \leq C \inf_{z_{\mathcal{T}} \in V_{\mathcal{T}}^s} \|v^\sigma - z_{\mathcal{T}}\|_{\mathcal{T},\lambda}, \tag{5.16}$$

where the constant $C > 0$ can be chosen to be independent of $\sigma \in (0, 1)$; see Section 5.1.4.

Lemma 5.2.2 (Approximation of the approximate corrector)**.** *In the situation described above, additionally assuming that the periodic strong solution* $v^\sigma = v^\sigma(\cdot\,; x, p, R) \in H^2_{\mathrm{per}}(Y)$ *to* (5.14) *satisfies* $v^\sigma \in H^{2+r_K}(K)$ *with* $r_K \geq 0$ *for all* $K \in \mathcal{T}$*, we have the error bound*

$$\|v^\sigma - v_{\mathcal{T}}^\sigma\|_{\mathcal{T},\lambda_\sigma} \lesssim \inf_{z_{\mathcal{T}} \in V_{\mathcal{T}}^s} \|v^\sigma - z_{\mathcal{T}}\|_{\mathcal{T},\lambda} \lesssim \left(\sum_{K \in \mathcal{T}} h_K^{2(\min\{2+r_K, \bar{p}+1\}-2)} \|\nabla v^\sigma\|^2_{H^{1+r_K}(K)} \right)^{\frac{1}{2}}$$

with constants independent of σ *and the choice of* (x, p, R)*.*

The proof is omitted as the first inequality is already obtained in (5.16), while the second estimate is a consequence of standard approximation arguments.

We observe that without any additional regularity assumptions on v^σ, we have that $\|\nabla v^\sigma\|_{H^1(Y)} \leq C$ is uniformly bounded in σ. Indeed, this follows from (5.15) and Poincaré's inequality.

5.2.3 Approximation of the effective Hamiltonian

Let us define the approximated effective Hamiltonian $H_{\mathcal{T}}^\sigma$ for $\sigma > 0$ via

$$H_{\mathcal{T}}^\sigma : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{n\times n}_{\mathrm{sym}} \to \mathbb{R}, \qquad H_{\mathcal{T}}^\sigma(x, p, R) := -\sigma \int_Y v_{\mathcal{T}}^\sigma(\cdot\,; x, p, R). \tag{5.17}$$

We note that this definition is quite natural as we have that

$$Q^\sigma_{x,p,R} := \|-\sigma v^\sigma(\cdot\,; x, p, R) - H(x, p, R)\|_{L^\infty(Y)} \underset{\sigma \searrow 0}{\longrightarrow} 0$$

for any $(x, p, R) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{n\times n}_{\mathrm{sym}}$.

Theorem 5.2.1 (Approximation of the effective Hamiltonian)**.** *Let* $H : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{n\times n}_{\mathrm{sym}} \to \mathbb{R}$ *denote the effective Hamiltonian given by* (5.12) *and* $H_{\mathcal{T}}^\sigma : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{n\times n}_{\mathrm{sym}} \to \mathbb{R}$ *its numerical approximation* (5.17)*. Then, for* $\sigma \in (0, 1)$ *and* $(x, p, R) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{n\times n}_{\mathrm{sym}}$*, we have the error bound*

$$|H_{\mathcal{T}}^\sigma(x, p, R) - H(x, p, R)| \lesssim Q^\sigma_{x,p,R} + \inf_{z_{\mathcal{T}} \in V_{\mathcal{T}}^s} \|v^\sigma(\cdot\,; x, p, R) - z_{\mathcal{T}}\|_{\mathcal{T},\lambda}. \tag{5.18}$$

In particular, we have the following assertions:

(i) If there exist non-negative numbers $\{r_K\}_{K\in\mathcal{T}} \subset [0,\infty)$ *such that the bound* $\sup_{K\in\mathcal{T}} \|\nabla v^\sigma(\cdot\,;x,p,R)\|_{H^{1+r_K}(K)} \le C_{x,p,R}|K|^{\frac{1}{2}}$ *holds uniformly in* σ*, then we have that*

$$|H^\sigma_{\mathcal{T}}(x,p,R) - H(x,p,R)| \lesssim Q^\sigma_{x,p,R} + C_{x,p,R}\left(\sum_{K\in\mathcal{T}} h_K^{2\min\{r_K,\bar{p}-1\}}|K|\right)^{\frac{1}{2}}. \quad (5.19)$$

(ii) Let us denote $h := \max_{K\in\mathcal{T}} h_K$ *and assume that there exists* $r \ge 0$ *such that* $\sup_{\sigma\in(0,1)} \|\nabla v^\sigma(\cdot\,;x,p,R)\|_{H^{1+r}(Y)} \le C_{x,p,R}$*. Then we have that*

$$|H^\sigma_{\mathcal{T}}(x,p,R) - H(x,p,R)| \lesssim Q^\sigma_{x,p,R} + C_{x,p,R}\, h^{\min\{r,\bar{p}-1\}}. \quad (5.20)$$

The constants absorbed in $\lesssim$ *are independent of* σ *and* (x,p,R)*.*

Proof. Let $\sigma \in (0,1)$ and $(x,p,R) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{n\times n}_{\mathrm{sym}}$. We observe that by Lemma 5.2.2, and recalling $\lambda_\sigma = \sigma\lambda$, we have

$$\begin{aligned}\|\sigma v^\sigma(\cdot\,;x,p,R) - \sigma v^\sigma_{\mathcal{T}}(\cdot\,;x,p,R)\|_{L^2(Y)} &\lesssim \|v^\sigma(\cdot\,;x,p,R) - v^\sigma_{\mathcal{T}}(\cdot\,;x,p,R)\|_{\mathcal{T},\lambda_\sigma} \\ &\lesssim \inf_{z_{\mathcal{T}}\in V^s_{\mathcal{T}}} \|v^\sigma(\cdot\,;x,p,R) - z_{\mathcal{T}}\|_{\mathcal{T},\lambda}\end{aligned} \quad (5.21)$$

with constants independent of σ and (x,p,R). Further, we note that

$$\|-\sigma v^\sigma(\cdot\,;x,p,R) - H(x,p,R)\|_{L^2(Y)} \le Q_{x,p,R}. \quad (5.22)$$

We can now conclude, using Hölder and triangle inequalities together with (5.21) and (5.22), that we have

$$\begin{aligned}|H^\sigma_{\mathcal{T}}(x,p,R) - H(x,p,R)| &= \left|-\sigma\int_Y v^\sigma_{\mathcal{T}}(\cdot\,;x,p,R) - H(x,p,R)\right| \\ &= \left|\int_Y (-\sigma v^\sigma_{\mathcal{T}}(\cdot\,;x,p,R) - H(x,p,R))\right| \\ &\le \|-\sigma v^\sigma_{\mathcal{T}}(\cdot\,;x,p,R) - H(x,p,R)\|_{L^2(Y)} \\ &\lesssim Q_{x,p,R} + \inf_{z_{\mathcal{T}}\in V^s_{\mathcal{T}}} \|v^\sigma(\cdot\,;x,p,R) - z_{\mathcal{T}}\|_{\mathcal{T},\lambda},\end{aligned}$$

where the constant absorbed in $\lesssim$ is independent of σ and (x,p,R). This completes the proof of (5.18). In view of Lemma 5.2.2, the bounds (5.19) and (5.20) are an immediate consequence of (5.18). □

Remark 5.2.2 (Improvement for HJB operators)**.** *Let us assume that the coefficients* A, b, f *from the HJBI operator* (5.10) *are such that the operator simplifies to an HJB operator*

$$F(x,y,p,R) := \sup_{\beta\in\mathcal{B}} \{-A(y,\beta) : R - b(x,y,\beta)\cdot p - f(x,y,\beta)\}$$

fitting into the framework of Section 4.3.1. We then have (see Section 4.3) *(i) the convergence rate* $Q^{\sigma}_{x,p,R} = \mathcal{O}\left(\sigma\left(1+|p|+|R|\right)\right)$ *as* $\sigma \searrow 0$*, and*

(ii) the uniform bound $\sup_{\sigma\in(0,1)} \|\nabla v^{\sigma}\|_{H^{1+r}(Y)} \leq C(1+|p|+|R|)$ *for some* $r>0$*.*

Therefore, by Theorem 5.2.1 (ii), we have the error bound

$$|H^{\sigma}_{\mathcal{T}}(x,p,R) - H(x,p,R)| \lesssim \left(\sigma + h^{\min\{r,\bar{p}-1\}}\right)\left(1+|p|+|R|\right),$$

where the constant absorbed in $\lesssim$ *is independent of* σ *and* (x,p,R)*.*

5.3 Numerical Experiments

5.3.1 Numerical solution of a periodic HJBI problem

In this numerical experiment, we consider the periodic HJBI problem

$$\begin{cases} \inf_{\alpha\in[0,\frac{1}{2}]} \sup_{\beta\in[0,2\pi]} \left\{-A^{\alpha\beta} : D^2u + c^{\alpha\beta}u - f^{\alpha\beta}\right\} = 0 \quad \text{in } Y, \\ u \text{ is } Y\text{-periodic}, \end{cases} \tag{5.23}$$

where we define the diffusion coefficient by

$$A^{\alpha\beta} := Q(\beta)\begin{pmatrix} \frac{\cos(\alpha)+\sin(\alpha)}{\sqrt{2}} & 0 \\ 0 & \frac{\cos(\alpha)-\sin(\alpha)}{\sqrt{2}} \end{pmatrix} Q(\beta)^{\mathrm{T}}, \qquad Q(\beta) := \begin{pmatrix} \cos(\beta) & -\sin(\beta) \\ \sin(\beta) & \cos(\beta) \end{pmatrix},$$

and set $c^{\alpha\beta} := \frac{\sec(\alpha)}{\sqrt{2}}$ and $f^{\alpha\beta} := \frac{\sec(\alpha)}{\sqrt{2}}\tilde{f}$ for $(\alpha,\beta) \in [0,\frac{1}{2}]\times[0,2\pi]$. Here, we choose $\tilde{f} \in C_{\mathrm{per}}(Y)$ such that the solution to (5.23) is given by

$$u : [0,1]^2 \to \mathbb{R}, \qquad u(y_1,y_2) = \cos(2\pi y_1)\cos(2\pi y_2).$$

We leave it to the reader to check that this problem fits into the setting of Section 5.1. In particular, we have that the Cordes condition (4.2) holds with $\lambda = 1$ and $\delta = \frac{1}{2}$.

Remark 5.3.1. *The renormalized HJBI problem* (5.2) *corresponding to* (5.23) *is given by*

$$\begin{cases} \inf_{\alpha\in[0,\frac{1}{2}]} \sup_{\beta\in[0,2\pi]} \left\{-\gamma^{\alpha\beta}A^{\alpha\beta} : D^2u + u\right\} = \tilde{f} \quad \textit{in } Y, \\ u \textit{ is } Y\textit{-periodic}, \end{cases}$$

where $\gamma^{\alpha\beta} := \sqrt{2}\cos(\alpha)$ *for* $(\alpha,\beta) \in [0,\frac{1}{2}]\times[0,2\pi]$*.*

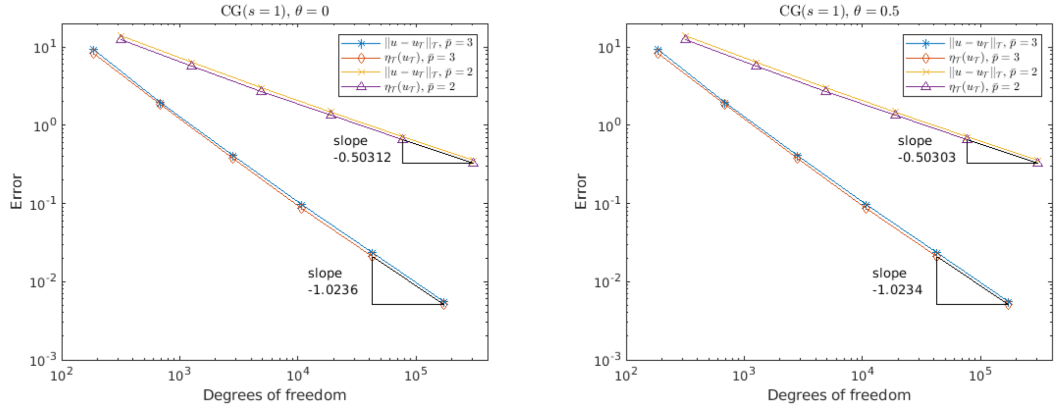


Figure 5.3: Approximation of the solution u to the HJBI problem (5.23) via the C^0-IP method with uniform mesh-refinement. We use polynomial degrees $\bar{p} \in \{2,3\}$ and set $\theta = 0$ (left) and $\theta = \frac{1}{2}$ (right). The plots illustrate the error (5.24) and the *a posteriori* error estimator (5.25) for the approximation $u_\mathcal{T} \in V^1_\mathcal{T}$.

We apply the C^0-IP and discontinuous Galerkin finite element schemes from Section 5.1.4 to the HJBI problem (5.23). Under uniform mesh-refinement, we illustrate the behavior of the error

$$\|u - u_\mathcal{T}\|_\mathcal{T} := \left(\int_Y \left(|D^2(u - u_\mathcal{T})|^2 + 2|\nabla(u - u_\mathcal{T})|^2 + (u - u_\mathcal{T})^2 \right) + |u - u_\mathcal{T}|^2_{J,\mathcal{T}} \right)^{\frac{1}{2}} \tag{5.24}$$

and of the *a posteriori* error estimator (see Theorem 5.1.2), i.e.,

$$\eta_\mathcal{T}(u_\mathcal{T}) := \left(\int_Y |F_\gamma[u_\mathcal{T}]|^2 + |u_\mathcal{T}|^2_{J,\mathcal{T}} \right)^{\frac{1}{2}} \tag{5.25}$$

for the numerical approximation $u_\mathcal{T} \in V^s_\mathcal{T}$. Figure 5.3 presents the C^0-IP method ($s = 1$) using the parameters $\theta \in \{0, \frac{1}{2}\}$ and the polynomial degrees $\bar{p} \in \{2,3\}$. Figure 5.4 presents the discontinuous Galerkin method ($s = 0$) using the parameters $\theta \in \{0, \frac{1}{2}\}$ and the polynomial degrees $\bar{p} \in \{2,3\}$. We observe optimal rates of convergence for both schemes: Denoting the number of degrees of freedom by N, we observe order $\mathcal{O}(N^{-\frac{1}{2}})$ for $\bar{p} = 2$ and order $\mathcal{O}(N^{-1})$ for $\bar{p} = 3$.

5.3.2 Numerical approximation of the effective Hamiltonian

In this numerical experiment, we demonstrate the numerical scheme for the approximation of the effective Hamiltonian corresponding to the HJBI operator

$$F : \mathbb{R}^2 \times \mathbb{R}^{2\times 2}_{\text{sym}} \to \mathbb{R}, \qquad F(y,R) := \inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} \left\{ -A^{\alpha\beta}(y) : R - 1 \right\} \tag{5.26}$$

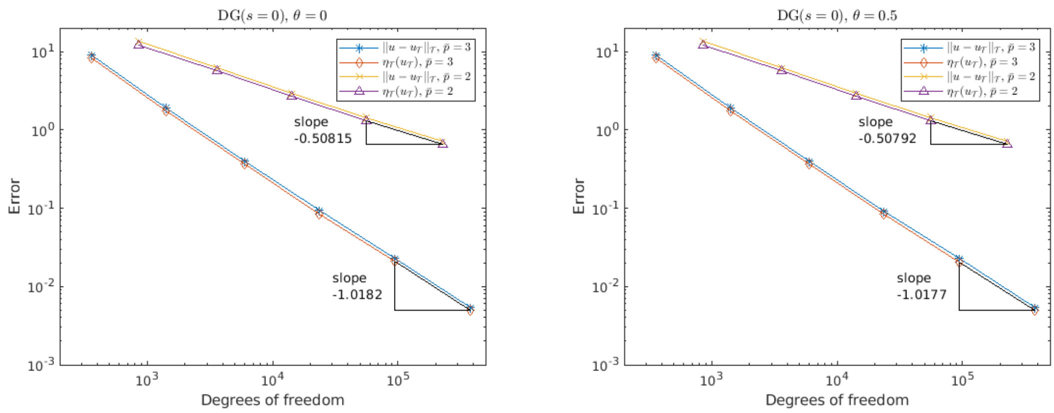


Figure 5.4: Approximation of the solution u to the HJBI problem (5.23) via the DG method with uniform mesh-refinement. We use polynomial degrees $\bar{p} \in \{2, 3\}$ and set $\theta = 0$ (left) and $\theta = \frac{1}{2}$ (right). The plots illustrate the error (5.24) and the *a posteriori* error estimator (5.25) for the approximation $u_{\mathcal{T}} \in V_{\mathcal{T}}^0$.

with $\mathcal{A} := [1, 2]$, $\mathcal{B} := [0, 1]$, and the coefficient $A = A(y, \alpha, \beta) : \mathbb{R}^2 \times \mathcal{A} \times \mathcal{B} \to \mathbb{R}^{2\times 2}_{\mathrm{sym}}$ given by

$$A^{\alpha\beta}(y) := (a_0(y) + \alpha\beta a_1(y))\, B,$$

where we choose positive scalar functions $a_0, a_1 : \mathbb{R}^2 \to (0, \infty)$ and a symmetric positive definite matrix $B \in \mathbb{R}^{2\times 2}_{\mathrm{sym}}$ defined by

$$B := \begin{pmatrix} 2 & -1 \\ -1 & 4 \end{pmatrix}, \qquad a_0 \equiv 1, \qquad a_1(y) := \sin^2(2\pi y_1)\cos^2(2\pi y_2) + 1.$$

It is straightforward to check that this problem fits into the framework of Section 5.2.1 and in particular we have that the Cordes condition (5.13) holds with $\lambda = \frac{1}{4}$. This HJBI operator is chosen so that we know the effective Hamiltonian explicitly:

Remark 5.3.2. *It can be checked that the HJBI operator* (5.26) *can be rewritten as HJB operator*

$$F(y, R) = \sup_{\beta \in [0,1]} \left\{ -\left(a_0(y) + \beta a_1(y)\right) B : R - 1 \right\}, \qquad (y, R) \in \mathbb{R}^2 \times \mathbb{R}^{2\times 2}_{\mathrm{sym}},$$

for which the effective Hamiltonian $H : \mathbb{R}^{2\times 2}_{\mathrm{sym}} \to \mathbb{R}$ *is known explicitly and given by*

$$H(R) := \max\left\{ -\left(\int_Y \frac{1}{a_0} \right)^{-1} B : R - 1, -\left(\int_Y \frac{1}{a_0 + a_1} \right)^{-1} B : R - 1 \right\}$$

for $R \in \mathbb{R}^{2\times 2}_{\mathrm{sym}}$*; see [45].*

We make it our goal to approximate the effective Hamiltonian $H(R)$ at the point

$$R := \begin{pmatrix} -2 & 1 \\ 1 & -3 \end{pmatrix},$$

noting that the same problem was already used for the numerical experiments in Chapter 4. As we have $B : R = -18 < 0$, the true effective Hamiltonian at this chosen point can be computed as

$$H(R) = -\left(\int_Y \frac{1}{a_0 + a_1}\right)^{-1} B : R - 1 = \frac{9\sqrt{6}\pi}{K(\frac{1}{3})} - 1 \approx 38.94291272989, \tag{5.27}$$

where K denotes the complete elliptic integral of the first kind.

In our numerical experiments, we approximate the true value of the effective Hamiltonian $H(R)$ from (5.27) by $H^\sigma_{\mathcal{T}}(R)$ as defined in (5.17), where we use the C^0-IP finite element method ($s = 1$) with $\theta = \frac{1}{2}$ to obtain the approximation $v^\sigma_{\mathcal{T}}(\cdot\,; R)$ to the solution $v^\sigma(\cdot\,; R)$ of the cell σ-problem as described in Section 5.2.2. We denote the relative approximation error by

$$E^\sigma_{\mathcal{T}} := \frac{|H^\sigma_{\mathcal{T}}(R) - H(R)|}{|H(R)|}, \qquad H^\sigma_{\mathcal{T}}(R) := -\sigma \int_Y v^\sigma_{\mathcal{T}}(\cdot\,; R)$$

and further write

$$E^\sigma := \frac{|H^\sigma(R) - H(R)|}{|H(R)|}, \qquad H^\sigma(R) := -\sigma \int_Y v^\sigma(\cdot\,; R).$$

Let us point out that the approximate corrector $v^\sigma(\cdot\,; R)$ is not known exactly, but we know that $E^\sigma = \mathcal{O}(\sigma)$ from Remark 5.2.2.

Figure 5.5 shows the behavior of the relative approximation error $E^\sigma_{\mathcal{T}}$ under uniform mesh refinement for fixed σ, and the corresponding *a posteriori* estimator (5.25) (re-scaled with a multiplicative constant C_σ for illustration purposes) using polynomial degree $\bar{p} = 3$. We observe that $E^\sigma_{\mathcal{T}}$ converges to a constant, namely E^σ, and that the *a posteriori* estimator is of order $\mathcal{O}(N^{-1})$ as expected, where N denotes the degrees of freedom. In particular, let us emphasize that this is the expected behavior and that the relative error for large numbers of degrees of freedom is entirely dominated by the σ-error E^σ.

Figure 5.6 shows the behavior of E^σ for varying values of σ, and we observe $E^\sigma = \mathcal{O}(\sigma)$ as σ tends to zero. The values E^σ have been obtained using a fine approximation with a high polynomial degree. For fixed values of σ, we further illustrate that the convergence rate for the convergence of $E^\sigma_{\mathcal{T}}$ to the value E^σ is of order $\mathcal{O}(N^{-\frac{3}{2}})$. This rate is higher than predicted by Remark 5.2.2, which is based on

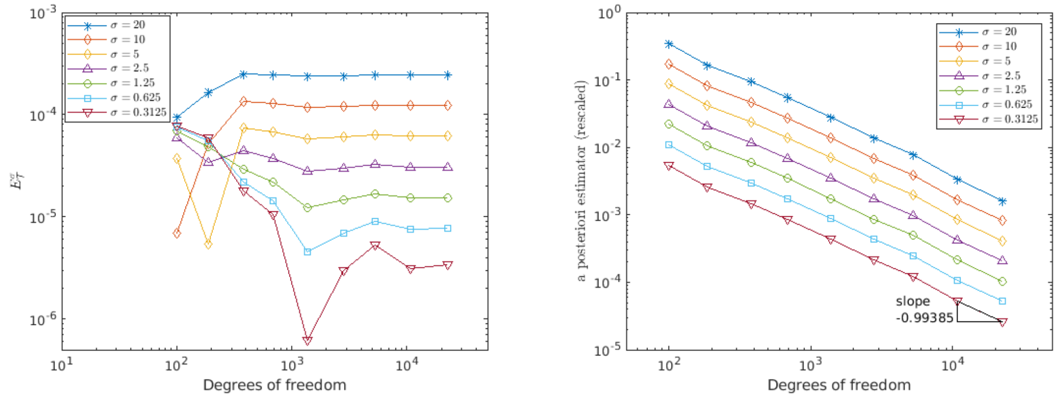


Figure 5.5: Relative error $E^\sigma_\mathcal{T}$ (left) and rescaled *a posteriori* error estimator (right) under uniform mesh refinement for fixed σ, using polynomial degree $\bar{p} = 3$.

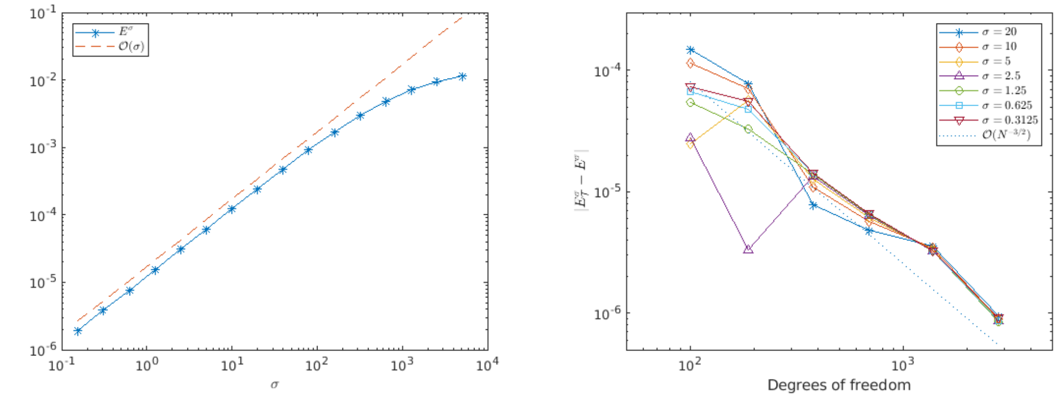


Figure 5.6: Accurate approximations to E^σ obtained via fine discretizations (left) and illustration of $|E^\sigma_\mathcal{T} - E^\sigma|$ under uniform mesh refinement with $\bar{p} = 3$ (right).

an error estimate in the $\|\cdot\|_{\mathcal{T},\lambda_\sigma}$-norm and is therefore indeed expected to overestimate the error between $H^\sigma_\mathcal{T}(R)$ and $H(R)$ related to the weaker integral functional from (5.17).

Chapter 6

Conclusion

6.1 Summary

In the first chapter of this work (Chapter 2) we studied the qualitative and quantitative homogenization of linear elliptic quations in nondivergence-form on $C^{2,\gamma}$ and polygonal domains. The convergence result is obtained via a transformation of the problem into divergence-form and uniform $W^{2,p}$ estimates. Quantitatively, through corrector results we obtained that the optimal rate of convergence of u_ε to the homogenized solution in the $W^{1,p}$-norm and the L^∞-norm is $\mathcal{O}(\varepsilon)$. Moreover, we obtained optimal estimates for the gradient and the Hessian of the solution with correction terms taken into account in the L^p-norm. In the final part of the chapter, we provided examples of an explicit c-good/bad matrix and presented several numerical experiments matching the theoretical results and illustrating the optimality of the obtained rates.

In Chapter 3 we introduced a scheme for the numerical approximation of such problems, which is based on a $W^{2,p}$ corrector estimate derived in the Chapter 2. We proved an optimal-order error bound for a finite element approximation of the corresponding invariant measure using continuous Y-periodic piecewise linear basis functions on a shape-regular triangulation of the unit cell Y under weak regularity assumptions on the coefficients. The coefficients are integrated against the so obtained approximation of the invariant measure after piecewise linear interpolation on the mesh to obtain an approximation of the constant coefficient-matrix of the homogenized problem. Using an H^2 comparison result for the solution of this perturbed problem, we eventually obtained an approximation of the solution u_0 to the homogenized problem in the H^2-norm. In the case of a polygonal domain in two space dimensions, we made use of compatibility conditions for the source term to ensure sufficiently high Sobolev-regularity of u_0.

We obtained an approximation to the solution u_ε of the original problem, i.e., the problem with oscillating coefficients, by making use of the H^2 approximation of u_0 and corrector functions, as well as an H^2 corrector result. A method of successively approximating higher derivatives for the approximation of corrector functions was provided and analyzed. The corrector functions are necessary in order to obtain an approximation of $D^2 u_\varepsilon$ whereas the task of approximating u_ε in the H^1-norm can be achieved using only an H^1 approximation of u_0.

Furthermore, we generalized our results to the case of nonuniformly oscillating coefficients, i.e., we derived an analogous corrector result and studied the approximation of the solution u_0 to the homogenized problem and the solution u_ε of the ε-dependent problem in this case.

In the final part of the chapter, we presented numerical experiments matching the theoretical results for problems with both known and unknown u_0, as well as problems with nonuniformly oscillating coefficients. We illustrated the performance of the scheme for the approximation of the invariant measure, the solution u_0 to the homogenized problem and the solution u_ε to the problem involving oscillating coefficients for a fixed value of ε.

In Chapter 4 we introduced a scheme for the numerical homogenization of the fully nonlinear second-order Hamilton–Jacobi–Bellman equation with coefficients satisfying a generalized Cordes condition, based on a mixed finite element method for the periodic corrector problems.

The focus of the first part of the chapter was the construction and the rigorous analysis of mixed finite element approximations to the periodic solution of the HJB equation. We derived a mixed formulation for the problem and proved well-posedness as well as *a priori* and *a posteriori* error bounds. Explicit formulas for the error constants were provided, showing the asymptotic behavior of the constants in the Cordes parameters.

In the second part of the chapter we focused on the numerical homogenization of HJB equations with locally periodic coefficients. Theoretical homogenization results were provided and used in the analysis of the numerical homogenization scheme. We presented and rigorously analyzed a method for the approximation of the effective Hamiltonian based on mixed finite element approximations of the periodic cell problem for the approximate corrector from the first part.

Finally, we presented numerical experiments illustrating the theoretical results. The experiments demonstrated the approximation of the effective Hamiltonian in a point as well as the approximation of the solution to the homogenized problem.

In Chapter 5 we presented a numerical homogenization scheme for Hamilton–Jacobi–Bellman–Isaacs equations with coefficients satisfying a generalized Cordes condition, which was based on a discontinuous Galerkin or C^0 interior penalty finite element approximation for periodic corrector problems.

The first part of the chapter was focused on periodic HJBI cell problems and provided a rigorous *a posteriori* and *a priori* analysis for a wide class of numerical schemes. In particular, the *a posteriori* analysis was independent of the choice of numerical scheme and used a periodic enrichment of finite element functions. We provided a family of numerical schemes which fits into this abstract framework.

The second part of the chapter was focused on the approximation of the effective Hamiltonian corresponding to ergodic HJBI operators. An approximation scheme for the effective Hamiltonian via a DG/C^0-IP approximation to approximate correctors was presented and rigorously analyzed.

Finally, we illustrated our theoretical results demonstrating the performance of the numerical scheme in numerical experiments.

6.2 Future work

This work seems to be the first systematic study of finite element schemes for the numerical homogenization of nondivergence-form equations and might open up new frontiers in this research area. In particular, it would be extremely interesting to obtain alternative numerical schemes (e.g., of LOD or MsFEM type) for the aforementioned problems. There is also scope for future work on the analysis of such homogenization problems to achieve better understanding of the class of c-bad matrices from Section 2. Future research will further include checking whether the generalized Cordes condition leads to a threshold in the behavior of solutions.

This work only discusses the periodic (or locally periodic) framework and it would be interesting to develop schemes for settings beyond the periodic framework for such problems. More generally speaking, we hope that this work will initiate further research on problems of nondivergence structure in regards of numerical homogenization.

References

[1] A. Abdulle. On a priori error analysis of fully discrete heterogeneous multiscale FEM. *Multiscale Model. Simul.*, 4(2):447–459, 2005.

[2] A. Abdulle. The finite element heterogeneous multiscale method: a computational strategy for multiscale PDEs. In *Multiple scales problems in biomathematics, mechanics, physics and numerics*, volume 31 of *GAKUTO Internat. Ser. Math. Sci. Appl.*, pages 133–181. Gakkotosho, Tokyo, 2009.

[3] A. Abdulle, W. E, B. Engquist, and E. Vanden-Eijnden. The heterogeneous multiscale method. *Acta Numer.*, 21:1–87, 2012.

[4] A. Abdulle and M. E. Huber. Finite element heterogeneous multiscale method for nonlinear monotone parabolic homogenization problems. *ESAIM Math. Model. Numer. Anal.*, 50(6):1659–1697, 2016.

[5] A. Abdulle and T. Pouchon. A priori error analysis of the finite element heterogeneous multiscale method for the wave equation over long time. *SIAM J. Numer. Anal.*, 54(3):1507–1534, 2016.

[6] Y. Achdou, F. Camilli, and I. Capuzzo Dolcetta. Homogenization of Hamilton-Jacobi equations: numerical methods. *Math. Models Methods Appl. Sci.*, 18(7):1115–1143, 2008.

[7] H. Aleksanyan, H. Shahgholian, and P. Sjölin. Applications of Fourier analysis in homogenization of the Dirichlet problem: L^p estimates. *Arch. Ration. Mech. Anal.*, 215(1):65–87, 2015.

[8] G. Allaire. *Shape optimization by the homogenization method*, volume 146 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 2002.

[9] G. Allaire and M. Amar. Boundary layer tails in periodic homogenization. *ESAIM Control Optim. Calc. Var.*, 4:209–243, 1999.

[10] O. Alvarez and M. Bardi. Viscosity solutions methods for singular perturbations in deterministic and stochastic control. *SIAM J. Control Optim.*, 40(4):1159–1188, 2001.

[11] O. Alvarez and M. Bardi. Ergodicity, stabilization, and singular perturbations for Bellman-Isaacs equations. *Mem. Amer. Math. Soc.*, 204(960):vi+77, 2010.

[12] O. Alvarez, M. Bardi, and C. Marchi. Multiscale problems and homogenization for second-order Hamilton-Jacobi equations. *J. Differential Equations*, 243(2):349–387, 2007.

[13] D. Arjmand and G. Kreiss. An equation-free approach for second order multiscale hyperbolic problems in non-divergence form. *Commun. Math. Sci.*, 16(8):2317–2343, 2018.

[14] S. Armstrong, T. Kuusi, J.-C. Mourrat, and C. Prange. Quantitative analysis of boundary layers in periodic homogenization. *Arch. Ration. Mech. Anal.*, 226(2):695–741, 2017.

[15] M. Avellaneda and F.-H. Lin. Compactness methods in the theory of homogenization. II. Equations in nondivergence form. *Comm. Pure Appl. Math.*, 42(2):139–172, 1989.

[16] M. Avellaneda and F.-H. Lin. L^p bounds on singular integrals in homogenization. *Comm. Pure Appl. Math.*, 44(8-9):897–910, 1991.

[17] I. Babuška. Computation of derivatives in the finite element method. *Comment. Math. Univ. Carolinae*, 11:545–558, 1970.

[18] M. Bebendorf. A note on the Poincaré inequality for convex domains. *Z. Anal. Anwendungen*, 22(4):751–756, 2003.

[19] A. Bensoussan, L. Boccardo, and F. Murat. Homogenization of elliptic equations with principal part not in divergence form and Hamiltonian with quadratic growth. *Comm. Pure Appl. Math.*, 39(6):769–805, 1986.

[20] A. Bensoussan, J.-L. Lions, and G. Papanicolaou. *Asymptotic analysis for periodic structures*. AMS Chelsea Publishing, Providence, RI, 2011. Corrected reprint of the 1978 original.

[21] D. Boffi, F. Brezzi, and M. Fortin. *Mixed finite element methods and applications*, volume 44 of *Springer Series in Computational Mathematics*. Springer, Heidelberg, 2013.

[22] V. I. Bogachev, N. V. Krylov, and M. Röckner. On regularity of transition probabilities and invariant measures of singular diffusions under minimal conditions. *Comm. Partial Differential Equations*, 26(11-12):2037–2080, 2001.

[23] V. I. Bogachev and S. V. Shaposhnikov. Integrability and continuity of solutions to double divergence form equations. *Ann. Mat. Pura Appl. (4)*, 196(5):1609–1635, 2017.

[24] L. Caffarelli, M. G. Crandall, M. Kocan, and A. Swiech. On viscosity solutions of fully nonlinear equations with measurable ingredients. *Comm. Pure Appl. Math.*, 49(4):365–397, 1996.

[25] L. A. Caffarelli, P. E. Souganidis, and L. Wang. Homogenization of fully nonlinear, uniformly elliptic and parabolic partial differential equations in stationary ergodic media. *Comm. Pure Appl. Math.*, 58(3):319–361, 2005.

[26] F. Camilli and M. Falcone. An approximation scheme for the optimal control of diffusion processes. *RAIRO Modél. Math. Anal. Numér.*, 29(1):97–122, 1995.

[27] F. Camilli and E. R. Jakobsen. A finite element like scheme for integro-partial differential Hamilton-Jacobi-Bellman equations. *SIAM J. Numer. Anal.*, 47(4):2407–2431, 2009.

[28] F. Camilli and C. Marchi. Rates of convergence in periodic homogenization of fully nonlinear uniformly elliptic PDEs. *Nonlinearity*, 22(6):1481–1498, 2009.

[29] Y. Capdeboscq, T. Sprekeler, and E. Süli. Finite element approximation of elliptic homogenization problems in nondivergence-form. *ESAIM Math. Model. Numer. Anal.*, 54(4):1221–1257, 2020.

[30] D. Cioranescu and P. Donato. *An introduction to homogenization*, volume 17 of *Oxford Lecture Series in Mathematics and its Applications*. The Clarendon Press, Oxford University Press, New York, 1999.

[31] H. O. Cordes. Über die erste Randwertaufgabe bei quasilinearen Differentialgleichungen zweiter Ordnung in mehr als zwei Variablen. *Math. Ann.*, 131:278–312, 1956.

[32] M. Costabel and M. Dauge. Maxwell and Lamé eigenvalues on polyhedra. *Math. Methods Appl. Sci.*, 22(3):243–258, 1999.

[33] M. G. Crandall, H. Ishii, and P.-L. Lions. User's guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc. (N.S.)*, 27(1):1–67, 1992.

[34] W. E and B. Engquist. The heterogeneous multiscale methods. *Commun. Math. Sci.*, 1(1):87–132, 2003.

[35] W. E, B. Engquist, X. Li, W. Ren, and E. Vanden-Eijnden. Heterogeneous multiscale methods: a review. *Commun. Comput. Phys.*, 2(3):367–450, 2007.

[36] Y. Efendiev and T. Y. Hou. *Multiscale finite element methods*, volume 4 of *Surveys and Tutorials in the Applied Mathematical Sciences*. Springer, New York, 2009. Theory and applications.

[37] Y. R. Efendiev and X.-H. Wu. Multiscale finite element for problems with highly oscillatory coefficients. *Numer. Math.*, 90(3):459–486, 2002.

[38] B. Engquist and P. E. Souganidis. Asymptotic and numerical homogenization. *Acta Numer.*, 17:147–190, 2008.

[39] L. C. Evans. The perturbed test function method for viscosity solutions of nonlinear PDE. *Proc. Roy. Soc. Edinburgh Sect. A*, 111(3-4):359–375, 1989.

[40] L. C. Evans. Periodic homogenisation of certain fully nonlinear partial differential equations. *Proc. Roy. Soc. Edinburgh Sect. A*, 120(3-4):245–265, 1992.

[41] M. Falcone and M. Rorro. On a variational approximation of the effective Hamiltonian. In *Numerical mathematics and advanced applications*, pages 719–726. Springer, Berlin, 2008.

[42] W. M. Feldman and I. C. Kim. Continuity and discontinuity of the boundary layer tail. *Ann. Sci. Éc. Norm. Supér. (4)*, 50(4):1017–1064, 2017.

[43] X. Feng, R. Glowinski, and M. Neilan. Recent developments in numerical methods for fully nonlinear second order partial differential equations. *SIAM Rev.*, 55(2):205–267, 2013.

[44] X. Feng, L. Hennings, and M. Neilan. Finite element methods for second order linear elliptic partial differential equations in non-divergence form. *Math. Comp.*, 86(307):2025–2051, 2017.

[45] C. Finlay and A. M. Oberman. Approximate homogenization of convex nonlinear elliptic PDEs. *Commun. Math. Sci.*, 16(7):1895–1906, 2018.

[46] C. Finlay and A. M. Oberman. Approximate homogenization of fully nonlinear elliptic PDEs: estimates and numerical results for Pucci type equations. *J. Sci. Comput.*, 77(2):936–949, 2018.

[47] W. H. Fleming and H. M. Soner. *Controlled Markov processes and viscosity solutions*, volume 25 of *Stochastic Modelling and Applied Probability*. Springer, New York, second edition, 2006.

[48] B. D. Froese and A. M. Oberman. Numerical averaging of non-divergence structure elliptic operators. *Commun. Math. Sci.*, 7(4):785–804, 2009.

[49] D. Gallistl. Variational formulation and numerical analysis of linear elliptic equations in nondivergence form with Cordes coefficients. *SIAM J. Numer. Anal.*, 55(2):737–757, 2017.

[50] D. Gallistl, T. Sprekeler, and E. Süli. Mixed Finite Element Approximation of Periodic Hamilton–Jacobi–Bellman Problems With Application to Numerical Homogenization. *Multiscale Model. Simul.*, 19(2):1041–1065, 2021.

[51] D. Gallistl and E. Süli. Mixed finite element approximation of the Hamilton-Jacobi-Bellman equation with Cordes coefficients. *SIAM J. Numer. Anal.*, 57(2):592–614, 2019.

[52] D. Gérard-Varet and N. Masmoudi. Homogenization and boundary layers. *Acta Math.*, 209(1):133–178, 2012.

[53] D. Gilbarg and N. S. Trudinger. *Elliptic partial differential equations of second order*. Classics in Mathematics. Springer-Verlag, Berlin, 2001. Reprint of the 1998 edition.

[54] R. Glowinski, S. Leung, and J. Qian. A simple explicit operator-splitting method for effective Hamiltonians. *SIAM J. Sci. Comput.*, 40(1):A484–A503, 2018.

[55] D. A. Gomes and A. M. Oberman. Computing the effective Hamiltonian using a variational approach. *SIAM J. Control Optim.*, 43(3):792–812, 2004.

[56] P. Grisvard. *Elliptic problems in nonsmooth domains*, volume 69 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2011. Reprint of the 1985 original.

[57] X. Guo, H. V. Tran, and Y. Yu. Remarks on optimal rates of convergence in periodic homogenization of linear elliptic equations in non-divergence form. *SN Partial Differ. Equ. Appl.*, 1(15), 2020.

[58] F. Hecht. New development in freefem++. *J. Numer. Math.*, 20(3-4):251–265, 2012.

[59] T. Hell and A. Ostermann. Compatibility conditions for Dirichlet and Neumann problems of Poisson's equation on a rectangle. *J. Math. Anal. Appl.*, 420(2):1005–1023, 2014.

[60] T. Hell, A. Ostermann, and M. Sandbichler. Modification of dimension-splitting methods—overcoming the order reduction due to corner singularities. *IMA J. Numer. Anal.*, 35(3):1078–1091, 2015.

[61] T. Y. Hou and X.-H. Wu. A multiscale finite element method for elliptic problems in composite materials and porous media. *J. Comput. Phys.*, 134(1):169–189, 1997.

[62] H. Ishii. On uniqueness and existence of viscosity solutions of fully nonlinear second-order elliptic PDEs. *Comm. Pure Appl. Math.*, 42(1):15–45, 1989.

[63] M. Jensen. $L^2(H^1_\gamma)$ finite element convergence for degenerate isotropic Hamilton-Jacobi-Bellman equations. *IMA J. Numer. Anal.*, 37(3):1300–1316, 2017.

[64] M. Jensen and I. Smears. On the convergence of finite element methods for Hamilton-Jacobi-Bellman equations. *SIAM J. Numer. Anal.*, 51(1):137–162, 2013.

[65] V. V. Jikov, S. M. Kozlov, and O. A. Oleĭnik. *Homogenization of differential operators and integral functionals*. Springer-Verlag, Berlin, 1994. Translated from the Russian.

[66] O. A. Karakashian and F. Pascal. A posteriori error estimates for a discontinuous Galerkin approximation of second-order elliptic problems. *SIAM J. Numer. Anal.*, 41(6):2374–2399, 2003.

[67] E. L. Kawecki and I. Smears. Unified analysis of discontinuous Galerkin and C^0-interior penalty finite element methods for Hamilton-Jacobi-Bellman and Isaacs equations. *ESAIM Math. Model. Numer. Anal.*, 55(2):449–478, 2021.

[68] E. L. Kawecki and I. Smears. Convergence of adaptive discontinuous Galerkin and C^0-interior penalty finite element methods for Hamilton-Jacobi-Bellman and Isaacs equations. *Found. Comput. Math.*, 2021. doi:10.1007/s10208-021-09493-0.

[69] E. L. Kawecki and T. Sprekeler. Discontinuous Galerkin and C^0-IP finite element approximation of periodic Hamilton–Jacobi–Bellman–Isaacs problems with application to numerical homogenization, arXiv:2104.14450 [math.NA].

[70] S. Kim and K.-A. Lee. Higher order convergence rates in theory of homogenization: equations of non-divergence form. *Arch. Ration. Mech. Anal.*, 219(3):1273–1304, 2016.

[71] T. Kinoshita, Y. Watanabe, N. Yamamoto, and M. T. Nakao. Some remarks on a priori estimates of highly regular solutions for the Poisson equation in polygonal domains. *Jpn. J. Ind. Appl. Math.*, 33(3):629–636, 2016.

[72] O. Lakkis and T. Pryer. A finite element method for second order nonvariational elliptic problems. *SIAM J. Sci. Comput.*, 33(2):786–801, 2011.

[73] P.-L. Lions. Optimal control of diffusion processes and Hamilton-Jacobi-Bellman equations. II. Viscosity solutions and uniqueness. *Comm. Partial Differential Equations*, 8(11):1229–1276, 1983.

[74] P.-L. Lions. A remark on Bony maximum principle. *Proc. Amer. Math. Soc.*, 88(3):503–508, 1983.

[75] S. Luo, Y. Yu, and H. Zhao. A new approximation for effective Hamiltonians for homogenization of a class of Hamilton-Jacobi equations. *Multiscale Model. Simul.*, 9(2):711–734, 2011.

[76] A. Målqvist and D. Peterseim. Localization of elliptic multiscale problems. *Math. Comp.*, 83(290):2583–2603, 2014.

[77] A. Målqvist and D. Peterseim. *Numerical homogenization by localized orthogonal decomposition*, volume 5 of *SIAM Spotlights*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, [2021] ©2021.

[78] M. Neilan, A. J. Salgado, and W. Zhang. Numerical analysis of strongly nonlinear PDEs. *Acta Numer.*, 26:137–303, 2017.

[79] L. Nirenberg. On elliptic partial differential equations. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (3)*, 13:115–162, 1959.

[80] A. M. Oberman, R. Takei, and A. Vladimirsky. Homogenization of metric Hamilton-Jacobi equations. *Multiscale Model. Simul.*, 8(1):269–295, 2009.

[81] D. Onofrei and B. Vernescu. Asymptotic analysis of second-order boundary layer correctors. *Appl. Anal.*, 91(6):1097–1110, 2012.

[82] J. Qian. Two approximations for effective Hamiltonians arising from homogenization of Hamilton-Jacobi equations. UCLA CAM report 03–39, 2003.

[83] J. Qian, H. V. Tran, and Y. Yu. Min-max formulas and other properties of certain classes of nonconvex effective Hamiltonians. *Math. Ann.*, 372(1-2):91–123, 2018.

[84] M. V. Safonov. Classical solution of second-order nonlinear elliptic equations. *Izv. Akad. Nauk SSSR Ser. Mat.*, 52(6):1272–1287, 1328, 1988.

[85] A. H. Schatz. An observation concerning Ritz-Galerkin methods with indefinite bilinear forms. *Math. Comp.*, 28:959–962, 1974.

[86] Z. Shen. $W^{1,p}$ estimates for elliptic homogenization problems in nonsmooth domains. *Indiana Univ. Math. J.*, 57(5):2283–2298, 2008.

[87] I. Smears and E. Süli. Discontinuous Galerkin finite element approximation of nondivergence form elliptic equations with Cordès coefficients. *SIAM J. Numer. Anal.*, 51(4):2088–2106, 2013.

[88] I. Smears and E. Süli. Discontinuous Galerkin finite element approximation of Hamilton-Jacobi-Bellman equations with Cordes coefficients. *SIAM J. Numer. Anal.*, 52(2):993–1016, 2014.

[89] I. Smears and E. Süli. Discontinuous Galerkin finite element methods for time-dependent Hamilton-Jacobi-Bellman equations with Cordes coefficients. *Numer. Math.*, 133(1):141–176, 2016.

[90] T. Sprekeler and H. V. Tran. Optimal Convergence Rates for Elliptic Homogenization Problems in Nondivergence-Form: Analysis and Numerical Illustrations. *Multiscale Model. Simul.*, 2021 (Forthcoming article).

[91] E. Süli. A brief excursion into the mathematical theory of mixed finite element methods. *Lecture Notes University of Oxford*, 2013.

[92] L. Tartar. *The general theory of homogenization*, volume 7 of *Lecture Notes of the Unione Matematica Italiana*. Springer-Verlag, Berlin; UMI, Bologna, 2009. A personalized introduction.