QUANTIZED TENSOR FEM FOR MULTISCALE PROBLEMS: DIFFUSION PROBLEMS IN TWO AND THREE DIMENSIONS.

V. KAZEEV †, I. OSELEDETS ‡, M. RAKHUBA §, AND CH. SCHWAB §

Abstract. Homogenization in terms of multiscale limits transforms a multiscale problem with \( n + 1 \) asymptotically separated microscales posed on a physical domain \( D \subset \mathbb{R}^d \) into a one-scale problem posed on a product domain of dimension \( (n + 1)d \) by introducing \( n \) so-called “fast variables”. This procedure allows to convert \( n + 1 \) scales in \( d \) physical dimensions into a single-scale structure in \( (n + 1)d \) dimensions. We prove here that both the original, physical multiscale problem and the corresponding high-dimensional, one-scale limiting problem can be efficiently treated numerically with the recently developed quantized tensor-train finite-element method (QTT-FEM).

The method is based on restricting computation to sequences of nested subspaces of low dimensions (which are called tensor ranks) within a vast but generic “virtual” (background) discretization space. In the course of computation, these subspaces are computed iteratively and data-adaptively at runtime, bypassing any “offline precomputation”. For the purpose of theoretical analysis, such low-dimensional subspaces are constructed analytically so as to bound the tensor ranks vs. error tolerance \( \tau > 0 \).

We consider a model linear elliptic multiscale problem in several physical dimensions and show, theoretically and experimentally, that both (i) the solution of the associated high-dimensional one-scale problem and (ii) the corresponding approximation to the solution of the multiscale problem admit efficient approximation by the QTT-FEM. These problems can therefore be numerically solved in a scale-robust fashion by standard (low-order) PDE discretizations combined with state-of-the-art general-purpose solvers for tensor-structured linear systems. We prove scale-robust exponential convergence, i.e., that QTT-FEM achieves accuracy \( \tau \) with the number of effective degrees of freedom scaling polynomially in \( \log \tau \).

1. Introduction. The efficient numerical solution of mathematical models of physical processes with multiple scales has undergone a rapid development during recent years. Several classes of computational approaches have been put forward which try, usually through selective and sparing access of the microscopic structure of the problem, to correctly numerically approximate the “effective”, macroscopic or “homogenized” features of the solution. In the context of Finite Element discretizations, these methodologies are referred to as multiscale FEM (MsFEM). In a broader context, such computational approaches for the numerical approximation of multiscale differential equation models (ordinary or partial) have been referred to as hierarchic multiscale methods (HMM). We refer to [19, 1] and the references therein for a comprehensive discussion.

In these approaches, the solution of the correct macroscopic, or “upscaled” mathematical model is numerically approximated by selective, localized access to the microscopic information. This can be achieved by the mentioned methods in (essentially optimal) numerical complexity that is independent of the microscopic length scale of the problem. Additionally, postprocessing techniques allow for localized numerical recovery of the microscopic structure of the physical solution, at extra computational costs.

An alternative computational approach aims at the simultaneous numerical

---

* Work initiated while VK and IO visited the Institute for Mathematical Research (FIM) of ETH Zürich in 2017 (IO) and 2018 (VK), and performed in part while the authors participated in the thematic programme “Numerical Analysis of Complex PDE Models in the Sciences” of the Erwin Schrödinger International Institute for Mathematics and Physics (ESI), Vienna, Austria, in June–August 2018.

† Faculty of Mathematics, University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria. vladimir.kazeev@univie.ac.at.

‡ Skolkovo Institute of Science and Technology, Novaya Ul. 100, 143025 Skolkovo, Moscow Region, Russia and Institute of Numerical Mathematics, Russian Academy of Sciences, Gubkina St. 8, 119333, Moscow, Russia. i.oseledets@skoltech.ru

§ Seminar for Applied Mathematics, ETH Zurich, Rämistrasse 101, 8092 Zurich, Switzerland. {maksim.rakhuba,christoph.schwab}@sam.math.ethz.ch
approximation of the macroscopic, homogenized solution and at the numerical approximation of the microscopic structure of the physical solution, throughout the physical domain, at computational work which is independent of the physical length scale of data. This is feasible, in general, under additional assumptions on the microstructure, such as (locally) periodicity or ergodicity. Under such assumptions, it is known that for linear, second order elliptic PDEs the physical solution and the interaction of all scales can be described by certain two- and \((n+1)\)-scale limits \([35, 3, 2]\). These limits take the form of solutions of *high-dimensional, elliptic boundary value problems*, which are independent of the scale parameters and posed on a Cartesian product of the physical domain \(D\) and of the \(n\) “unit-cells” \(Y_i, i = 1,...,n\) that describe the structure of the fast scales of the multiscale solution. As a result, \((n+1)\)-scale limits trade scale-resolving requirements for high-dimensionality \([42]\). Loosely speaking, scale-resolution is traded for the “curse of dimensionality”: once efficient numerical approximations for such high-dimensional \((n+1)\)-scale limiting problems are available, robust, scale-independent discretizations of multiscale models can be derived. This idea, put forward in \([42]\), has been developed in the context of sparse tensor FEM multiscale diffusion problems in \([18]\) and, subsequently, for elasticity and electromagnetics \([17, 46, 47, 48]\). In particular, *algebraic convergence rates independent of the scale parameter with weak or no dependence on the number of \(n\) of fast variables* were established. The implementation of these sparse tensor FEM discretizations of the high-dimensional limits requires, however, *explicit derivation of the PDEs which describe the \((n+1)\)-scale limits*. This may, in particular for nonlinear multiscale problems, not be feasible, even though the existence of \((n+1)\)-scale limits is mathematically assured.

1.1. Contributions. We analyze the novel, tensor-structured numerical approximation of the solution of a linear second-order elliptic PDE whose diffusion tensor depends on \(n+1\) separated scales, i.e., in the classical setting of \((n+1)\)-scale homogenization. Specifically, following earlier work \([28, 25, 22]\) we consider the *quantized tensor-train finite-element method* (QTT-FEM), combining adaptive low-rank tensor approximation with *quantization* \([37, 24]\) to efficiently represent the multiscale structure of data.

In the present paper, we first prove that the QTT-FEM allows for *exponentially convergent* numerical approximations to the scale-interaction functions involved in the \((n+1)\)-scale limits and, as a consequence, to the homogenized solutions. Specifically, we construct “by hand” certain approximations that, with respect to the discretization parameter, are sufficiently accurate and have sufficiently low tensor ranks.

The idea of approximating the multiscale problem by reapproximating the homogenized problem (derived by \((n+1)\)-scale convergence \([2, 12]\), proposed for elliptic multiscale problems in \([42]\), was exploited in the context of sparse grid approximations \([18, 17, 16]\). However, our present perspective extends further, as the motivation for considering approximations based on homogenization. In practice, the QTT-FEM can completely bypass the homogenization procedure and operate entirely on the physical domain, adaptively accessing the fine-scale information of the PDE. Naturally, the numerical approximations found by this approach are better adapted to the data and are more efficient than any particular approximations we construct “by hand” through the re-approximation of the corresponding homogenized problem. In Section 5 we report numerical results obtained by such a practical computational multiscale QTT-FEM algorithm, built upon the TT Toolbox \([39]\).

1.2. Structure of the present paper. In Section 2 we describe the \(n\)-scale homogenization problem, and present in particular the QTT discretization of this problem in the physical domain in Section 2.1. The emphasis in Section 2 is to present the \(n\)-scale problem and its quantized, tensor-formatted discretization entirely in the physical domain. Section 3 presents the asymptotic analysis of the
an $n$-scale solution by the so-called unfolding method: the asymptotic limit of the physical problem is described by a high-dimensional one-scale problem. To this end, we recapitulate results from [33, 32, 29] on reiterated homogenization for linear, elliptic multiscale problems, which are required in the ensuing numerical analysis of the QTT-FE approach.

Section 4 will develop novel approximation rate results for the solution of the $(n+1)$-scale limit which are, subsequently, used to obtain quantized tensor-rank bounds for the physical, $(n+1)$-scale solution.

Section 5 then will present numerical experiments which model multiscale problems where the QTT-ranks of the numerical solutions are explicitly estimated numerically.

Finally, Section 6 and the Appendix contain a discussion of the results and a few proofs postponed due to their technicality.

2. Model elliptic multiscale problem. We consider a bounded “physical” domain $D \subset \mathbb{R}^d$ (with which, for notational convenience, we associate the macroscopic scale parameter $\varepsilon_0 = 1$) and a moderate number $n \in \mathbb{N}$ of microscales $\varepsilon_1, \ldots, \varepsilon_n$, which we assume to be positive functions of a scale parameter $\varepsilon$ such that $\lim_{\varepsilon \to 0} \varepsilon_i = 0$ for all $i \in \{1, \ldots, n\}$. We additionally assume asymptotic scale separation:

$$\lim_{\varepsilon \to 0} \varepsilon_{i+1}/\varepsilon_i = 0$$

for $i \in \{1, \ldots, n-1\}$.

Further, we assume that there exist $n$ unit cells $Y_1, \ldots, Y_n$ such that $D$ is partitioned into a union of translations of $\varepsilon_1 Y_1$ and each $Y_{i-1}$ with $i \in \{2, \ldots, n\}$ is partitioned into a union of translations of $\varepsilon_i Y_i$. Specifically, we deal with the case of $Y_1, \ldots, Y_n = (0,1)^d$ in the present paper, while more sophisticated constructions may be used to model, e.g., perforated media. For notational convenience, we set $Y_0 = \{0\}$ and $Y_i = Y_1 \times \cdots \times Y_i$ for each $i \in \{1, \ldots, n\}$.

To formulate a multiscale diffusion problem on $D$, we consider a matrix function $A$ defined on $D \times Y_n$, which therefore depends on a macroscopic (“slow”) variable and on $n$ microscopic (“fast”) variables. We will consider multiscale diffusion coefficients $A^\varepsilon$ induced by functions satisfying the following assumption.

**Assumption 2.1.** $A \in L^\infty(D; C_\#(Y_n; \mathbb{R}^{d \times d}))$ is essentially bounded and uniformly positive definite with constants $\Gamma$ and $\gamma$: $\gamma \leq \xi^T A(x,y_n) \xi \leq \Gamma$ for every unit vector $\xi \in \mathbb{R}^d$, a.e. $x \in D$ and all $y_n \in Y_n$.

Here and throughout, by $C_\#(Y_n)$ we denote the space of functions that are continuous on $Y_n$ and $Y_i$-periodic with respect to the $i$th variable for each $i \in \{1, \ldots, n\}$.

For every $\varepsilon > 0$, a function $A$ satisfying Assumption 2.1 induces a multiscale coefficient $A^\varepsilon \in L^\infty(D)$ as follows:

$$A^\varepsilon(x) = A\left(x, \frac{x}{\varepsilon_1}, \ldots, \frac{x}{\varepsilon_n}\right) \text{ for all } x \in D. \quad (2.2)$$

With such a coefficient, we consider the following model variational problem on $V = H^1_0(D)$:

$$\text{find } u^\varepsilon \in V \text{ such that } \int_D (\nabla v)^T A^\varepsilon \nabla u^\varepsilon = \int_D f v \text{ for all } v \in V, \quad (2.3)$$

where $f \in L^2(D)$ is a forcing term. Assumption 2.1 and the Lax–Milgram theorem guarantee that this problem has a unique solution, which satisfies the stability bound

$$|u^\varepsilon|_{H^1(D)} \leq \gamma^{-1} \sup_{v \in V \setminus \{0\}} \frac{|f(v)|}{|v|_{H^1(D)}} \leq C \gamma^{-1} \sup_{v \in V \setminus \{0\}} \frac{|f(v)|}{|v|_{L^2(D)}} \leq C \gamma^{-1} \|f\|_{L^2(D)},$$

where $C$ is the classical Poincaré constant for $D$. 

3
Although the forcing term $f$ is assumed to be independent of the scale parameter $\varepsilon$ for simplicity, we hasten to add that all results that follow admit a straightforward generalization to the case when $f$ exhibits a microscale structure analogous to the one expressed by (2.2).

2.1. Low-rank tensor multilevel discretization. In this section, we give an explicit construction of the low-rank tensor multilevel discretization of the problem (2.3) for the case when $D = Y_1 = \cdots = Y_n = (0, 1)^d$. We start with defining the underlying virtual grid and the associated finite-element spaces.

2.1.1. Virtual grids and low order finite-element spaces. Let $L \in \mathbb{N}_0$ be arbitrary. We define index sets $\mathcal{I}^L = \{1, \ldots, 2^L\}$, and $\mathcal{J}^L = \{1, \ldots, 2^L - 1\}$, select the meshwidth at mesh level $L$ as $h_L = 2^{-L}$ and consider a set of equispaced points on $(0, 1)$:

$$t^L_j = jh_L \quad \text{with} \quad j \in \{0\} \cup \mathcal{I}^L.$$  (2.4)

The corresponding continuous piecewise-linear functions $\varphi^L_j$, $j \in \{0\} \cup \mathcal{I}^L$ are given by the condition $\varphi^L_j(t^L_j) = \delta_{jj'}$ for all $j, j' \in \{0\} \cup \mathcal{I}^L$. These functions form a basis in the finite-element space $\bar{U}^L = \text{span}\{\varphi^L_j : j \in \{0\} \cup \mathcal{I}^L\}$, whose subspace $U^L = \text{span}\{\varphi^L_j : j \in \mathcal{J}^L\}$ allows to explicitly impose the boundary conditions of the problem (2.3). Similarly, the corresponding space of piecewise-constant functions is $\bar{U}^L = \text{span}\{\varphi^L_i : i \in \mathcal{I}^L\}$ with $\varphi^L_i$ with $i \in \mathcal{I}^L$ given by the condition $\varphi^L_i(t^L_j) = \delta_{ii'}$ for all $i, i' \in \mathcal{I}^L$.

To obtain coefficients of finite-element approximations with respect to these bases, we will use the analysis operators $\Phi^L : H^1(0, 1) \to \mathbb{C}^d$ and $\tilde{\Phi}^L : L^2(0, 1) \to \mathbb{C}^d$ defined as follows: for all $v \in H^1(0, 1)$, $w \in L^2(0, 1)$ and $i \in \mathcal{I}^L$, we set

$$(\Phi^L v)_i = v(t^L_i) \quad \text{and} \quad (\tilde{\Phi}^L w)_i = 2^L \int_{t^L_{i-1}}^{t^L_i} w.$$  (2.5)

Tensorizing the univariate basis functions defined above, we obtain $d$-variate basis functions that span the corresponding finite-element spaces:

$$V^L = \bigotimes_{k=1}^d U^L \subset V \quad \text{and} \quad \bar{V}^L = \bigotimes_{k=1}^d \bar{U}^L \subset L^2(D).$$  (2.6)

Classical approximation bounds (see, e.g., [11]) give

$$\inf_{v^L \in V^L} \|v - v^L\|_{H^t(D)} \leq C 2^{-tL} \|v\|_{H^{t+1}(D)} \quad \text{for all} \quad v \in H^{t+1}(D),$$  (2.7)

where $t > 0$ is a fractional order of Sobolev smoothness and $C > 0$ is a coefficient that depends on $t$ but not on $L$.

Since the solution $u^\varepsilon$ of (2.3) may exhibit algebraic singularities at the boundary of $D$ due to a combination of the domain’s geometry, boundary conditions and diffusion coefficient, $u^\varepsilon \in H^{1+t}(D)$ may hold only for $t$ significantly less than one. To efficiently approximate such solutions in low-rank form, we will follow [21, 23, 31] in using the multilevel QTT format for the low-rank separation of the indices associated with different levels and, for example, not different physical variables. This consists in applying the isomorphism

$$\bigotimes_{k=1}^d \bigotimes_{\ell=1}^L \mathbb{C}^2 \simeq \bigotimes_{\ell=1}^\ell \left\{ \bigotimes_{k=1}^d \mathbb{C}^2 \right\}$$  (2.8)

so that the $2^{dL}$ degrees of freedom in $V^L$ in (2.6) are represented by $d$-indices corresponding to the $L$ levels of discretization, each taking $2^d$ values that enumerate the
elements of the corresponding factor marked by curly brackets in (2.8). To refer to this isomorphism explicitly, we define \( \Pi^L \) with \( L \in \mathbb{N} \) as the permutation matrix of order \( 2^{dL} \) satisfying
\[
(\Pi^L)_{i_1,1 \ldots, i_d,1; i_1,l \ldots, i_d,l} = 1
\]
for all \( i_{k\ell} \in \{1,2\} \) with \( k \in \{1,\ldots,d\} \) and \( \ell \in \{1,\ldots,L\} \).

The elements of \( V^L \) and \( \tilde{V}^L \) can be parametrized by their coefficients extracted using the analysis operators
\[
\psi^L = H^L \bigotimes_{k=1}^{d} \phi^L : H^1(0,1) \to \mathbb{C}^{2^{dL}} \quad \text{and} \quad \tilde{\psi}^L = H^L \bigotimes_{k=1}^{d} \tilde{\phi}^L : L^2(D) \to \mathbb{C}^{2^{dL}}.
\]
(2.10)

Note that the restriction of \( \tilde{\psi}^L \) to \( V^L \) is not surjective. This lack of surjectivity stems from that we choose to use nested finite-element spaces \( V^L \) with \( L \in \mathbb{N} \) given by (2.6) but represent every function from \( V^L \) with \( L \in \mathbb{N} \) by \( 2^{dL} \) values instead of \( (2^L - 1)^d \), the extra values, in agreement with the boundary conditions of the problem (2.3), being zero.

2.1.2. Discrete multiscale problem and low-rank tensor parametrization. For every \( L \in \mathbb{N} \), we consider the following discretization of the problem (2.3):
\[
\text{find } u^\varepsilon;L \in V^L \text{ such that } \int_D (\nabla v^L)^T A^\varepsilon \nabla u^\varepsilon;L = \int_D f v^L \quad \text{for all } v^L \in V^L.
\]
(2.11)

As for the original problem, Assumption 2.1 and the Lax–Milgram theorem guarantee that the discrete solution has a unique solution. By the Cea’s lemma, the discrete solution is quasi optimal: \( \|u^\varepsilon - u^\varepsilon;L\|_{H^1(D)} \leq C \gamma^{-1} \Gamma 2^{-L} \|u^\varepsilon\|_{H^{1+\varepsilon}(D)} \), where \( C \) is the constant appearing in the approximation bound (2.7).

For a tensor \( u \in \mathbb{C}^{n_1 \times \cdots \times n_L} \) with \( L \in \mathbb{N} \) dimensions and mode sizes \( n_1, \ldots, n_L \in \mathbb{N} \), a representation
\[
u_{i_1,\ldots, i_L} = \sum_{\alpha_1}^{r_1} \cdots \sum_{\alpha_{L-1}}^{r_{L-1}} U_1(1, i_1, \alpha_1) \cdot U_2(\alpha_1, i_2, \alpha_2) \cdot \cdots \cdot U_{L-1}(\alpha_{L-2}, i_{L-1}, \alpha_{L-1}) \cdot U_L(\alpha_{L-1}, i_L, 1)
\]
(2.12)
in terms of arrays \( U_\ell \in \mathbb{C}^{r_{\ell-1} \times n_{\ell} \times r_\ell} \) with \( \ell \in \{1,\ldots,L\} \), where we use \( r_0 = 1 = r_L \) for convenience, is referred to as a tensor train (TT) decomposition \cite{20, 35}, or, alternatively, as a matrix-product state (MPS) representation \cite{45, 14, 41}. The arrays \( U_\ell \) with \( \ell \in \{1,\ldots,L\} \) are called cores, and the parameters \( r_1, \ldots, r_{L-1} \), governing the number of entries of the cores, are called ranks. In the present paper, we use the TT-MPS representation as a multilevel tensor decomposition \cite{43}, by which we mean that the indices of a tensor represented as in (2.12) represent the scales (not the physical dimensions) of the data. In the context of the TT-MPS decomposition, this has been known in the literature as the quantized tensor train decomposition \cite{37, 26, 24, 27}.

3. Reiterated homogenization and high-dimensional one-scale limit.

For analysis, instead of the original multiscale problem (2.3), we consider a one-scale high-dimensional limit problem posed in (3.6) in this section. The limit problem is obtained from the original multiscale problem (2.3) by homogenization, analyzed for \( n = 1 \), i.e., for a single microscale in \cite{10, 9, 20, 33, 35, 4}, and for \( n > 1 \) fast scales by iteration in \cite{2}. For a general discussion, we refer to \cite{14}.

3.1. One-scale high-dimensional limit problem. To formulate reiterated homogenization, we consider the following assumption, of which Assumption 2.1 is a particular case with \( i = n \) and \( A_n = A \).
Using the following result, the physical solution of the solution of the one-scale high-dimensional limit problem.

\[(2.3)\]

and consider the Cartesian-product space
\[V_i = V \times V_1 \times \cdots \times V_i\]
endowed with the inner product \(\langle \cdot, \cdot \rangle_{V_i}\) given by
\[
\langle \psi, \phi \rangle_{V_i} = \sum_{|\alpha|=1} \langle \partial^{\alpha} \psi_0, \partial^{\alpha} \phi_0 \rangle_{L^2(D)} + \sum_{j=1}^i \sum_{|\alpha|=1} \langle \partial_{j}^{\alpha} \psi_j, \partial_{j}^{\alpha} \phi_j \rangle_{L^2(D \times Y_j)}
\]
for all \(\psi = (\psi_0, \psi_1, \ldots, \psi_i), \phi = (\phi_0, \phi_1, \ldots, \phi_i) \in V_i\). We denote the norm induced by \(\langle \cdot, \cdot \rangle_{V_i}\) with \(||\cdot||_{V_i}\). Here and throughout, the symbol \(\partial^{\alpha}\) with \(\alpha \in \mathbb{N}_0^n\) denotes the differentiation of functions with respect to the first \(d\) scalar variables indicated by the multi-index \(\alpha\), whereas \(\partial_{j}^{\alpha}\) with \(i \in \{1, \ldots, n\}\) and \(\alpha \in \mathbb{N}_0^n\) denotes the differentiation of functions with respect to the scalar variables \(i d + 1, \ldots, (i + 1) d\) according to the multi-index \(\alpha\). Further, we define a bilinear form \(B_i : V_i \times V_i \to \mathbb{R}\):
\[
B_i(\psi, \phi) = \int_{D \times Y_i} \left( \nabla \psi_0 + \sum_{j=1}^i \nabla_{j} \psi_j \right)^T A_i \left( \nabla \phi_0 + \sum_{j=1}^i \nabla_{j} \phi_j \right)
\]
for all \(\psi = (\psi_0, \psi_1, \ldots, \psi_i), \phi = (\phi_0, \phi_1, \ldots, \phi_i) \in V_i\), where \(A_i\) is a matrix function satisfying Assumption 3.1 with \(i\) microscales and with positive constants \(\gamma\) and \(\Gamma\). Then the bilinear form \(B_i\) is continuous and coercive: the inequalities
\[
\gamma \|\phi\|_{V_i}^2 \leq B_i(\phi, \phi) \quad \text{and} \quad B_i(\psi, \phi) \leq \Gamma \|\psi\|_{V_i} \|\phi\|_{V_i}
\]
hold for all \(\psi, \phi \in V_i\). Then, since \(f \in L^2(D)\), the problem of finding \(u \in V_i\) such that
\[
B_i(u, \phi) = \int_D f \phi_0 \quad \text{for all} \quad \phi = (\phi_0, \phi_1, \ldots, \phi_i) \in V_i
\]
has a unique solution \(u = (u_0, u_1, \ldots, u_i)\) (by the Lax–Milgram theorem). For notational convenience, we introduce
\[
v_i = \sum_{j=0}^i \nabla u_j \quad \text{with} \quad i = 1, \ldots, n.
\]
We remark that the bilinear forms \(B_i, i = 1, \ldots, n,\) in \((3.6)\) satisfy property \((3.5)\) with constants uniform with respect to the scale parameter \(\varepsilon_i\).

The problem \((3.6)\) with \(i = n\) microscales, representing the result of \(n\) iterations of homogenization applied to the original multiscale problem \((2.3)\), approximates the multiscale problem in the following sense.

**Theorem 3.2** (Theorem 2.11 and equation (2.9) in [2]). The solution \(u^\varepsilon\) of the problem \((2.3)\) converges weakly to \(u_0\) in \(H^1_0(D)\), and \(\nabla u^\varepsilon\) \((n+1)\)-scale converges to \(v_n\).

Using the following result, the physical solution \(u^\varepsilon\), including the oscillations induced by the multiscale structure of the diffusion coefficient \((2.2)\), can be approximated in terms of the solution of the one-scale high-dimensional limit problem.
Theorem 3.3 (Theorem 2.14 in [2]). Assume that the solution \((u, u_1, \ldots, u_n)\) of problem (2.3) is sufficiently smooth, say \(u \in C^1(D)\) and \(u_i \in C^1(D, C^1(Y_i))\) for all \(i \in \{1, \ldots, n\}\). Then, as \(\varepsilon \to 0\),

\[
u_\varepsilon(x) \to u_0(x) + \sum_{i=1}^n \varepsilon_i u_i \left( x, \frac{x}{\varepsilon_1}, \ldots, \frac{x}{\varepsilon_i} \right) \quad \text{in} \quad H^1(D).
\]

3.2. Convergence in physical variables for multiple scales. Unfolding and averaging operators. For problems with \(n+1 > 2\) scales, an error estimate in the form (3.10) appears not to be available. We still base the rank bounds for the QT discretization on the structure of the one-scale limiting problem. To this end, generalizing [13] Definitions 2.1 and 2.16 to the case of multiple microscales, we introduce unfolding and averaging operators.

Definition 3.4. For all \(i \in \{1, \ldots, n\}\), the operators \(T^\varepsilon_i : L^2(D \times Y_{i+1} \times \cdots \times Y_n) \to L^2(D \times Y_i \times \cdots \times Y_n)\) and \(U^\varepsilon_i : L^2(D \times Y_i \times \cdots \times Y_n) \to L^2(D \times Y_{i+1} \times \cdots \times Y_n)\) of unfolding and averaging with respect to the \(i\)th microscale are defined by

\[
(T^\varepsilon_i \phi)(x, y_i, y_{i+1}, \ldots, y_n) = \phi \left( \varepsilon_i \left[ \frac{x}{\varepsilon_i} \right] + \varepsilon_i y_i, y_{i+1}, \ldots, y_i \right)
\]

for a.e. \((x, y_i, y_{i+1}, \ldots, y_n) \in D \times Y_i \times \cdots \times Y_n\) and all \(\phi \in L^2(D \times Y_i \times Y_{i+1} \times \cdots \times Y_n)\), where \(\phi\) is extended by zero outside its domain, and

\[
(U^\varepsilon_i \phi)(x, y_{i+1}, \ldots, y_n) = |Y_i|^{-1} \int_{Y_i} \phi \left( \varepsilon_i \left[ \frac{x}{\varepsilon_i} \right] + \varepsilon_i z, \left\{ \frac{x}{\varepsilon_i} \right\}, y_{i+1}, \ldots, y_n \right) dz
\]

for a.e. \((x, y_{i+1}, \ldots, y_n) \in D \times Y_{i+1} \times \cdots \times Y_n\) and all \(\phi \in L^2(D \times Y_i)\).

For every \(i \in \{1, \ldots, n\}\), the \(n\)-microscale unfolding and averaging operators are defined as \(T^\varepsilon = T^\varepsilon_n \circ \cdots \circ T^\varepsilon_1 : L^2(D) \to L^2(D \times Y_n)\) and \(U^\varepsilon = U^\varepsilon_n \circ \cdots \circ U^\varepsilon_1 : L^2(D \times Y_n) \to L^2(D)\).

In the case of one microscale, certain basic properties of the unfolding and averaging operators are analyzed in [13]. In particular, by [13] Proposition 2.17, the operator \(U^\varepsilon_i : L^2(D \times Y_i \times \cdots \times Y_n) \to L^2(D \times Y_{i+1} \times \cdots \times Y_n)\) is continuous and has norm \(|Y_i|^{-1/2}\) for all \(i \in \{1, \ldots, n\}\). This implies

\[
\|U^\varepsilon_i(\Phi - \bar{\Phi})\|_{L^2(D \times Y_n)} \leq \|\Phi - \bar{\Phi}\|_{L^2(D \times Y_n)}
\]

for all \(\Phi, \bar{\Phi} \in L^2(D \times Y_n)\).

As in [13], one can show that the solution \(u^\varepsilon\) of the multiscale problem (2.3) under the scale-separation condition (2.1) satisfies

\[
T^\varepsilon \nabla u^\varepsilon \to v_n \quad \text{strongly in} \quad L^2(D \times Y_n) \quad \text{as} \quad \varepsilon \to 0.
\]

Using the folding operator \(U^\varepsilon\), we can state an analog of (3.10) for several microscales, showing that the scale-interaction functions \(u_1, \ldots, u_n\) in (3.9) describe to leading order the oscillations of the functions \(u^\varepsilon\) with \(\varepsilon > 0\) as they approach the weak limit \(u^0\).

Lemma 3.5. Under the scale-separation condition (2.1), for the multiscale problem (2.3) we have \(\nabla u^\varepsilon - U^\varepsilon v_n \to 0\) strongly in \(L^2(D)\) as \(\varepsilon \to 0\), the averaging operators being applied componentwise.

For a proof, we refer to [13] Theorem 6.1 for the case \(n = 1\) of a single microscale and [13] Remark 7.5 regarding the case of \(n > 1\) microscales.

Remark 3.6. When the unfolded solution \((u_0, u_1, \ldots, u_n)\) consists of infinitely differentiable functions of all variables, this result can be inferred from the corrector result in Theorem 3.3.
By the Lax–Milgram theorem, the problem of finding \( f \) posed and has a unique solution.

Assume moreover that the physical domain \( D \) has a smooth boundary and that \( f \in L^2(D) \). Then the solution \((u_0, u_1, \ldots, u_n)\) of the limit problem \((3.6)\) satisfies \( u_0 \in H^2(D) \).

### 3.3. Convergence in physical variables for two scales.

We estimate the error between the solution \( u^\varepsilon \) of the physical problem \((2.3)\) in terms of the FE approximations of the limit problem \((3.6)\). We base this on an explicit error estimate between \( u^\varepsilon \) and the correctors for the two scale case (\( n = 1 \)).

**Proposition 3.8.** Assume that \( A \in C^\infty(D, C^\#_\#(Y_1))^{d \times d} \) and that the homogenized solution \( u_0 \) belongs to \( H^2(D) \). Then
\[
\left\| u^\varepsilon - \left( u_0(x) + \varepsilon u_1 \left( x, \frac{x}{\varepsilon} \right) \right) \right\|_{H^1(D)} \leq C \varepsilon^{\frac{1}{2}}.
\]

The constant \( C \) is independent of \( \varepsilon \) but depends on \( u_0 \) and \( u_1 \).

### 3.4. Recurrence for scale-interaction functions.

Let \( i \in \{1, \ldots, n\} \) and assume that \( A_i \) is a matrix function satisfying Assumption 3.1 with \( i \) microscales and positive constants \( \gamma \) and \( \Gamma \). Then the limit problem \((3.6)\), posed on \( D \times Y_i \), is well posed and has a unique solution.

Assume that \( \xi \in \mathbb{R}^d \) is a unit vector. For a.e. \((x, y_{i-1}) \in D \times Y_{i-1} \), define a bilinear form \( b_i(x, y_{i-1}, \cdot, \cdot) : H^1_{sym}(Y) / \mathbb{R} \times H^1_{sym}(Y) / \mathbb{R} \rightarrow \mathbb{R} \) and a linear form \( f_i(x, y_{i-1}, \xi, \cdot) : H^1_{sym}(Y) / \mathbb{R} \rightarrow \mathbb{R} \) as follows:
\[
\begin{align*}
b_i(x, y_{i-1}, \psi, \phi) &= \int_{Y_i} \langle \nabla \psi \rangle^T A_i(x, y_{i-1}, \cdot) \nabla \phi, \\
f_i(x, y_{i-1}, \xi, \phi) &= - \int_{Y_i} \xi^T A_i(x, y_{i-1}, \cdot) \nabla \phi
\end{align*}
\]
for all \( \psi, \phi \in H^1_{sym}(Y) / \mathbb{R} \). Then the following holds for a.e. \((x, y_{i-1}) \in D \times Y_{i-1} \).

First, the assumption regarding \( A_i \) results in the continuity and ellipticity of \( b_i(x, y_{i-1}, \cdot, \cdot) \): for all \( \psi, \phi \in H^1_{sym}(Y) / \mathbb{R} \),
\[
b_i(x, y_{i-1}, \psi, \phi) \leq \Gamma \| \psi \|_{H^1(Y_i)} \| \phi \|_{H^1(Y_i)},
\]
and
\[
b_i(x, y_{i-1}, \phi, \phi) \geq \gamma \| \phi \|_{H^1_{sym}(Y_i)}^2.
\]
Second, by the same argument, the linear form \( f_i(x, y_{i-1}, \xi, \cdot) \) is continuous:
\[
|f_i(x, y_{i-1}, \xi, \phi)| \leq \Gamma \| \phi \|_{H^1_{sym}(Y_i)} \quad \text{for all } \phi \in H^1_{sym}(Y_i) / \mathbb{R}.
\]

By the Lax–Milgram theorem, the problem of finding \( w_\xi(x, y_{i-1}, \cdot) \in H^1_{sym}(Y_i) / \mathbb{R} \) such that
\[
b_i(x, y_{i-1}, w_\xi(x, y_{i-1}, \cdot), \phi) = f_i(x, y_{i-1}, \xi, \phi) \quad \text{for all } \phi \in H^1_{sym}(Y_i) / \mathbb{R},
\]
admits a unique solution, which satisfies \( \| w_\xi(x, y_{i-1}, \cdot) \|_{H^1_{sym}(Y_i) / \mathbb{R}} \leq \gamma^{-1} \Gamma \).

Let \( \xi_1, \ldots, \xi_d \) be the columns of the identity matrix \( I \) of order \( d \). Being valid for a.e. \((x, y_{i-1}) \in D \times Y_{i-1} \) and every unit vector \( \xi \in \mathbb{R}^d \), the above argument defines \( w_i \in W^d_i \) whose components \( w_{ik} \in W_i \) with \( k \in \{1, \ldots, d\} \) are given by
\[
w_{ik}(x, y_{i-1}, y_k) = w_{\xi_k}(x, y_{i-1}, y_k) \quad \text{for a.e. } (x, y_{i-1}, y_k) \in D \times Y_{i-1} \times Y_k \text{ and for each } k \in \{1, \ldots, d\}.
\]
Note that \( w_i \) is also an element of \( V^d_i \). Furthermore, it is the only element of \( V^d_i \) such that
\[
\int_{D \times Y_i} (I + J_i w_i) A_i \nabla \phi = 0
\]
for all \( \phi \in V_i \). Here, \( J_i \) denotes the differential operator returning the Jacobi matrix with respect to the last variable (varying in \( Y_i \)), as a function of all variables (taking values in \( D \times Y_i \)).
Since \( A_i \in L^\infty(D; C_\#(Y_i; \mathbb{R}^{d \times d})) \), one can define \( A_{i-1} \in L^\infty(D; C_\#(Y_{i-1}; \mathbb{R}^{d \times d})) \) by setting
\[
A_{i-1}(x, y_{i-1}) = \int_{Y_i} (I + J_i w_i(x, y_{i-1}, \cdot)) A_i(x, y_{i-1}, \cdot) (I + J_i w_i(x, y_{i-1}, \cdot))^T
\]

\[
= \int_{Y_i} (I + J_i w_i(x, y_{i-1}, \cdot)) A_i(x, y_{i-1}, \cdot)
\]

(3.14)

for a.e. \( x \in D \) and for all \( y_{i-1} \in Y_{i-1} \). By [10, Theorem 3.9], the matrix function \( A_{i-1} \), which is called an \textit{upcaled coefficient}, satisfies Assumption 3.1 with \( i \) microscales and with the identical positive constants \( \gamma \) and \( \Gamma \). The corresponding problem (3.6), involving \( i \) variables, is therefore well posed and has a unique solution \((u_0, \ldots, u_{i-1}) \in V_{i-1}\)

Since \( u_{i-1} \in V_{i-1} \), we have \( \nabla_{i-1} u_{i-1} \in L^2(D \times Y_{i-1})^d \). On the other hand, we have noted that \( w_i \in W_i^d \), so we can define \( u_i \in V_i \) by setting
\[
u_i(x, y_{i-1}, \cdot) = (w_i(x, y_{i-1}, \cdot))^T \nabla_{i-1} u_{i-1}(x, y_{i-1}) \quad \text{in} \quad H_\#(Y_i)/\mathbb{R}
\]

(3.15)

for a.e. \( x \in D \) and \( y_{i-1} \in Y_{i-1} \). Indeed, this entails that \( u_i(x, y_{i-1}, \cdot) \) has the gradient
\[
\nabla_i u_i(x, y_{i-1}, \cdot) = J_i w_i(x, y_i) \nabla_{i-1} u_{i-1}(x, y_{i-1}) \quad \text{in} \quad L^2(Y_i)
\]

(3.16)

for a.e. \( x \in D \) and \( y_{i-1} \in Y_{i-1} \), so that the bound \( \|u_i\|_V^2 \lesssim \|w_i\|_{W_i^d}^2 \|u_{i-1}\|_{V_{i-1}}^2 \) holds with an equivalence constant determined by the choice of a norm for \( W_i^d \). This implies that \((u_0, \ldots, u_{i-1}, u_i) \in V_i \) and, as one verifies using (3.13) and (3.14), also that this tuple solves the problem (3.6) with \( i + 1 \) variables.

Applying the above argument iteratively, we obtain the “effective” macroscopic diffusion coefficient \( A_0 \in L^\infty(D; \mathbb{R}^{d \times d}) \):
\[
A_0 = \int_{Y_1} \cdots \int_{Y_n} (I + J_1 w_1) \cdots (I + J_n w_n) A, \quad \text{(3.17)}
\]

which satisfies Assumption 3.1 with zero microscales and with the identical constants \( \gamma \) and \( \Gamma \). The “effective” problem for the homogenized limit \( u_0 \) reads: find \( u_0 \in V_0 \) such that for every \( \phi \in V_0 \)
\[
\int_D (\nabla \phi)^T A_0 \nabla u_0 = \int_D f \phi. \quad \text{(3.18)}
\]

Then the solution \((u_0, \ldots, u_n) \in V_n \) of the limit problem (3.6) with \( n + 1 \) variables can be solved using the recursion (3.15), so that the scale-interaction functions \( u_i \) and the sums of their gradients given by (3.7) satisfy
\[
u_i = w_i^T v_{i-1} \quad \text{and} \quad v_i = (I + J_i w_i)^T v_{i-1} = (I + J_i w_i)^T (I + J_i w_i)^T \nabla u_0 \quad \text{(3.19)}
\]
in \( V_i \) and \( L^2(D \times Y_i)^d \) respectively.

#### 3.5. Approximate recurrence for scale-interaction functions.
In order to obtain low-rank tensor-structured approximations of \((u_0, u_1, \ldots, u_n) \in V_n \), we use the following approximation scheme with a discretization parameter \( L \in \mathbb{N} \). For every \( i = 1, \ldots, n \), we approximate \( w_i \) and \( J_i w_i \) by \( w_i^L \) and \( J_i^L \) in \( W_i^d \) and \( L^\infty(D \times Y_{i-1}, L^\infty(Y_i))^{d \times d} \) respectively. Assuming that \( u_0 \) and \( \nabla u_0 \) are approximated by \( u_0^L \) and \( v_0^L \) in \( V \) and \( L^2(D)^d \) respectively, we follow (3.19) to define the corresponding approximations \( u_i^L \) and \( v_i^L \) to \( u_i \) and \( v_i \) with \( i \in \{1, \ldots, n\} \): in \( V_i \) and \( L^2(D \times Y_i)^d \) respectively, we set
\[
u_i^L = (w_i^L)^T v_{i-1}^L \quad \text{and} \quad v_i^L = (I + J_i^L)^T v_{i-1}^L = (I + J_i^L)^T (I + J_i^L)^T v_0^L. \quad \text{(3.20)}
\]
The associated errors can be represented by telescoping sums: for example,

\[ v_i - v_i^L = (I + \mathcal{F}_i w_i) \cdots (I + \mathcal{F}_1 w_1) (v_0 - v_0^L) + \sum_{j=1}^i \left\{ \prod_{m=j+1}^i (I + \mathcal{F}_m w_m) \right\} \left( \mathcal{F}_j w_j - J_j^L \right) \right\} v_0^L \]

for every \( i \in \{1, \ldots, n\} \), where sums and products over empty ranges are to be omitted. Assuming that the errors \( w_i - w_i^L, \mathcal{F}_i w_i - J_i^L \) and \( v_0 - v_0^L \) are bounded, respectively, in \( W^d, L^\infty(D \times Y_{i-1}, L^2(Y_i))^d \times d \) and \( L^2(D)^d \) uniformly with respect to \( L \in \mathbb{N} \) and \( i \in \{1, \ldots, n\} \), we obtain, with a positive equivalence constant independent of the discretization parameter \( L \in \mathbb{N} \), the bounds

\[ \|v_i - v_i^L\|_{L^2(D \times Y_i)^d} \lesssim \|v_0 - v_0^L\|_{L^2(D)^d} + \sum_{j=1}^i \|\mathcal{F}_j w_j - J_j^L\|_{L^\infty(D \times Y_{i-1}, L^2(Y_i))} \]

(3.21)

and

\[ \|u_i - u_i^L\|_{v_i} \lesssim \|v_0 - v_0^L\|_{L^2(D)^d} + \sum_{j=1}^i \|\mathcal{F}_j w_j - J_j^L\|_{L^\infty(D \times Y_{i-1}, L^2(Y_i))} + \|w_i - w_i^L\|_{W_i} \]

(3.22)

for \( i \in \{1, \ldots, n\} \). In Section 4.3.3 we construct particular approximations \( w_i^L, J_i^L, u_i^L \) and \( v_i^L \) with \( i \in \{1, \ldots, n\} \) and \( L \in \mathbb{N} \) in the finite-element spaces specified in Section 4.1.1.

4. Approximability under the assumption of analyticity. In the present section, we investigate regularity and approximability of \( u_0, u_1, \ldots, u_n \). With the aim of establishing convergence rates and (quantized) tensor rank bounds which are independent of the scales, we impose additional assumptions on the data \( D, A \) and \( f \). Specifically, we consider a tensor-product physical domain and analytic data.

The first set of additional assumptions consists in the following.

ASSUMPTION 4.1. For every \( \varepsilon \) and \( i \in \{0, 1, \ldots, n\} \), we have \( \varepsilon_i = 2^{-\lambda_i} \) with \( \lambda_i \in \mathbb{N} \) depending on \( \varepsilon \) (we set \( \lambda_0 = 0 \) for notational convenience). For the physical domain and the unit cells, we have \( D = Y_1 = \cdots = Y_n = (0,1)^d \). The diffusion coefficient \( A \) is analytic and one-periodic with respect to each of the last \( nd \) scalar variables on \( D \times Y_n \). The right-hand side \( f \) is analytic on \( D \times Y_n \).

Assumption 4.1 allows to prove that the solution of the one-scale high-dimensional limiting problem can be approximated by finite-element functions of tensor ranks that are logarithmic in accuracy. This implies that the solution of the one-scale high-dimensional limiting problem admits an infinite sequence of approximations that converge exponentially with respect to the number of parameters used to represent them.

4.1. Low-order finite-element approximation. In this section, we extend the construction of finite-element spaces given in Section 2.1.1 to address the boundary conditions of the high-dimensional problem (3.6) and establish main approximation results. As stated in Assumption 4.1, we consider the case \( D = Y_1 = \cdots = Y_n = (0,1)^d \).

4.1.1. Low-order approximation on an interval. For \( L \in \mathbb{N} \), in order to accommodate the periodic boundary conditions of the high-dimensional problem (3.6), we consider the following subspace of \( U^L \):

\[ U^L_{\#} = \text{span}\{ \varphi_{\# j}: j \in \mathcal{J}^L \} , \]

where \( \varphi_{\# j} = \varphi_j^L \) for every \( j \in \mathcal{J}^L \) and \( \varphi_{\# 2 L}^L = \varphi_0^L + \varphi_{2 L}^L \).
We will use the analysis operators introduced in [2.5] to extract the coefficients of finite-element approximations in $U^L$, $\bar{U}^L$ and $U^L_\#$. To construct such approximations, we will use the following projection operators, $\pi^L: H^1(0,1) \to \bar{U}^L$ and $\bar{\pi}^L: L^2(0,1) \to \bar{U}^L$. The first we define as the operator of continuous, piecewise-linear Lagrange interpolation at the nodes given in [2.4], in the basis of $\varphi_j^L$ with $j \in \{0\} \cup \mathcal{I}^L$. The second operator we define as the operator of piecewise-constant $L^2$ approximation associated with the basis functions $\bar{\varphi}_i^L$ with $i \in \mathcal{I}^L$, which are defined in Section 2.1.1. Note that $(\pi^L v)' = \bar{\pi}^L v'$ for every $v \in H^1(0,1)$. Finally, both the projection operators can be expressed in terms of the analysis operators defined in [2.5], for all $u \in H^1_0(0,1)$, $v \in H^1_0(0,1)$ and $w \in L^2(0,1)$, we have

$$\pi^L u = \sum_{j \in \mathcal{I}^L} (\Phi^L u)_j \varphi_j^L, \quad \pi^L v = \sum_{j \in \mathcal{I}^L} (\Phi^L v)_j \varphi_j^L \quad \text{and} \quad \bar{\pi}^L w = \sum_{i \in \mathcal{I}^L} (\Phi^L w)_i \bar{\varphi}_i^L. \quad \tag{4.1}$$

In the following proposition, we summarize classical bounds for the projection operators $\pi^L$ and $\bar{\pi}^L$ for $L \in \mathbb{N}$.

**Proposition 4.2.** For all $v \in C[0,1] \cap C^2(0,1)$, $w \in C[0,1] \cap C^1(0,1)$ and $L \in \mathbb{N}$, the projections $\pi^L v$ and $\bar{\pi}^L w$ satisfy the error bounds

$$\|v - \pi^L v\|_{L^\infty(0,1)} \leq 2^{-2L-3} \|v''\|_{L^\infty(0,1)}, \quad \|v - \pi^L v\|_{L^\infty(0,1)} \leq 2^{-L} \|v''\|_{L^\infty(0,1)},$$

and the stability bounds

$$\|\pi^L v\|_{L^\infty(0,1)} \leq \|v\|_{L^\infty(0,1)}, \quad \|\pi^L v\|_{L^\infty(0,1)} \leq \|v\|_{L^\infty(0,1)}, \quad \bar{\pi}^L w\|_{L^\infty(0,1)} \leq \|w\|_{L^\infty(0,1)}.$$

**4.1.2. Low-order approximation on $D \times Y_\#$.** From the univariate bases defined above, we obtain by tensorization $d$-vari ate bases which span the corresponding finite-element spaces:

$$\tilde{V}^L = \bigotimes_{k=1}^d \bar{U}^L = \tilde{V}^L \subset H^1(D), \quad V^L = \bigotimes_{k=1}^d U^L = \bar{V}^L \subset H^1_0(D), \quad \tilde{V}^L_\# = \bigotimes_{k=1}^d U^L_\# = \tilde{V}^L \cap H^1_\#(Y) \quad \text{and} \quad \tilde{V}^L = \bigotimes_{k=1}^d \bar{U}^L \subset L^2(D) = L^2(Y) \quad \tag{4.2}$$

with $L \in \mathbb{N}$.

Using the spaces of $d$-vari ate finite-element functions specified above, define

$$\tilde{V}_i^L = (\tilde{V}^L)^{\otimes (i+1)}, \quad \tilde{V}_i^L = \tilde{V}_{i-1}^L \otimes \tilde{V}^L \quad \text{and} \quad V^L_{\#i} = \tilde{V}_{i-1}^L \otimes V^L_{\#} \quad \tag{4.3}$$

for all $i \in \{0, \ldots, n\}$ and $L \in \mathbb{N}$.

Further, for all $i \in \{0, \ldots, n\}$ and $L \in \mathbb{N}$, to construct approximations by finite-element functions from $V^L_{\#i}$, we will use the operators $\bar{\Pi}_i^L: L^2(D \times Y_1) \to \tilde{V}_i^L$ and $\Pi_i^L: L^2(D \times Y_{i-1}) \otimes \bar{H}^1_\#(0,1) \otimes \cdots \to V^L_{\#i}$ given by

$$\bar{\Pi}_i^L = \bigotimes_{j=0}^i \bar{\pi}^L \quad \text{and} \quad \Pi_i^L = \tilde{\Pi}_{i-1}^L \otimes \bigotimes_{k=1}^d \pi^L \quad \tag{4.4}$$

The following accuracy bounds for $\bar{\Pi}_i^L$ and $\Pi_i^L$ with $i \in \{1, \ldots, n\}$ and $L \in \mathbb{N}$ can be derived from Proposition 4.2.

**Lemma 4.3.** Let $i \in \{0, \ldots, n\}$ and $\|\cdot\|_{L^\infty(D \times Y_1)}$ denote $\|\cdot\|_{L^\infty(D \times Y_1)}$. Assume that $v \in C^1(D \times Y_1)$ and $w \in C^0(D \times Y_1)$. Then the following error bounds hold for all
\[ L \in \mathbb{N} \text{ and } k \in \{1, \ldots, d\}\):

\[ \|v - \Pi^L_i v\|_{\infty} \leq 2^{-L} \sum_{j'=0}^{i} \sum_{k'=1}^{d} \|\partial^{j'k'} w\|_{\infty}, \]

\[ \|w - \Pi^L_i w\|_{\infty} \leq 2^{-L} \sum_{j'=0}^{i-1} \sum_{k'=1}^{d} \|\partial^{j'k'} w\|_{\infty} + 2^{-2L-3} \sum_{k'=1}^{d} \|\partial_{ik}^2 w\|_{\infty}, \]

\[ \|\partial_{ik}(w - \Pi^L_i w)\|_{\infty} \leq 2^{-L} \sum_{j'=0}^{i-1} \sum_{k'=1}^{d} \|\partial^{j'k'} \partial_{ik} w\|_{\infty} + 2^{-L} \sum_{k'=1}^{d} \|\partial_{ik}^2 \partial_{ik} w\|_{\infty}. \]

We give a proof of Lemma 4.3 in the Appendix.

For all \( i \in \{0, \ldots, n\} \) and \( L \in \mathbb{N} \), the projections produced by the operators \( \Pi^L_i \) and \( \Pi^L_i \), defined by (4.4), can be parametrized by the coefficients extracted using the analysis operators \( \Psi^L_i : L^2(D \times Y_i) \to C^{2(1+1)4L} \) and \( \Phi^L_i : L^2(D \times Y_{i-1}) \otimes (H^2(0,1))^d \to C^{2(1+1)4L} \), given by

\[ \Psi^L_i = \bigotimes_{j=0}^{i} \Pi^L_j \otimes \Phi^L_j \quad \text{and} \quad \Phi^L_i = \bigotimes_{k=1}^{d} \Phi^L_k \quad \text{for } i = 0, \ldots, n. \]

Note that while the restrictions of \( \Phi^L_i \) and \( \Phi^L_i \) to \( V^L_i \) and \( \tilde{V}^L_i \) are bijective, that of \( \Psi^L_i \) to \( V^L_i \) is injective. The lack of bijectivity stems from the fact that we use nested finite-element spaces (4.2)–(4.3) and represent every function from \( V^L_i \) by \( 2^{(1+1)4L} \) values associated with a uniform tensor-product grid. We take into account this lack of bijectivity in the design of our numerical method.

4.2. High-order approximation.

4.2.1. High-order approximation on an interval. By \( \tilde{T}_\alpha \) with \( \alpha \in \mathbb{N}_0 \), we denote the Chebyshev polynomials of the first kind orthogonal on \( (0,1) \):

\[ \tilde{T}_\alpha(x) = \cos \left\{ n \arccos(2x - 1) \right\} \quad \text{for all } x \in (0,1) \text{ and } \alpha \in \mathbb{N}_0, \]

so that the orthogonality property holds with respect to the weight function \( \omega \) given by

\[ \omega(x) = \frac{1}{\sqrt{x(1-x)}} \quad \text{for all } x \in (0,1). \]

Specifically, we have

\[ \langle \tilde{T}_\alpha, \tilde{T}_{\alpha'} \rangle_{L^2_\omega(0,1)} = \int_0^1 \omega T_\alpha T_{\alpha'} = \delta_{\alpha\alpha'} \|T_\alpha\|_{L^2_\omega(0,1)}^2 \quad \text{for all } \alpha, \alpha' \in \mathbb{N}_0, \]

where \( \|T_0\|_{L^2_\omega(0,1)}^2 = \pi \) and \( \|T_\alpha\|_{L^2_\omega(0,1)}^2 = \frac{\pi}{2} \) for all \( \alpha \in \mathbb{N} \).

Further, we consider the complex exponentials \( \hat{T}_\alpha \) with \( \alpha \in \mathbb{Z} \) defined as follows:

\[ \hat{T}_\alpha(x) = \exp(2\pi i \alpha x) \quad \text{for all } x \in (0,1) \text{ and } \alpha \in \mathbb{Z}. \]

These are also orthogonal on \( (0,1) \):

\[ \langle \hat{T}_\alpha, \hat{T}_{\alpha'} \rangle_{L^2(0,1)} = \int_0^1 \hat{T}_\alpha \hat{T}_{\alpha'} = \delta_{\alpha\alpha'} \quad \text{for all } \alpha, \alpha' \in \mathbb{Z}. \]

We will use the following notation for the spaces of univariate algebraic and trigonometric polynomials of degree at most \( p \in \mathbb{N}_0 \):

\[ \mathcal{P}_p = \text{span}\{\tilde{T}_\alpha\}_{\alpha=0}^p \quad \text{and} \quad \mathcal{P}_p = \text{span}\{\hat{T}_\alpha\}_{\alpha=-p}^p. \]
where the span is meant with respect to the field \( \mathbb{C} \).

We will use polynomial approximations obtained by the following orthogonal projections onto \( \mathcal{P}_p \) and \( \mathcal{P}_{\#p} \) with \( p \in \mathbb{N}_0^d \):

\[
\pi_p = \frac{1}{\pi} \mathcal{T}_0 (\mathcal{T}_0, \cdot)_{L^2(0,1)} + \frac{2}{\pi} \sum_{\alpha=1}^{p} \mathcal{T}_\alpha (\mathcal{T}_\alpha, \cdot)_{L^2(0,1)} : L^2(0,1) \to \mathcal{P}_p,
\]

\[
\pi_{\#p} = \mathcal{T}_0 (\mathcal{T}_0, \cdot)_{L^2(0,1)} + \sum_{\alpha=1}^{p} \mathcal{T}_\alpha (\mathcal{T}_\alpha, \cdot)_{L^2(0,1)} : L^2(0,1) \to \mathcal{P}_{\#p}.
\]

4.2.2. High-order approximation on \( D \times Y_i \). For every \( i \in \{1, \ldots, n\} \), denoting by \( \text{id} \) the identity transformation of \( \mathbb{C}^d \), let us define the following tensor-product operators:

\[
\Pi_{i,p} = \left( \bigotimes_{k=1}^{d} \pi_p \right) \otimes \left( \bigotimes_{j=1}^{d} \pi_{\#p} \right) : L^2_{\omega \otimes \cdot \otimes \text{id}}(D \times Y_i) \to \left( \bigotimes_{k=1}^{d} \mathcal{P}_p \right) \otimes \left( \bigotimes_{j=1}^{d} \mathcal{P}_{\#p} \right)
\]

for all \( p \in \mathbb{N}_0^d \). Here, \( \omega \) denotes the weight function in (4.7). The following lemma verifies that, when applied to analytic functions, these operators yield approximations that converge exponentially with respect to \( p \in \mathbb{N}_0^d \).

**Lemma 4.4.** Assume that \( i \in \{1, \ldots, n\} \) and \( w \in V_i \) is analytic and one-periodic with respect to each of the last id scalar variables on \( D \times Y_i \). Let \( \epsilon_0 > 0 \). Then there exist positive constants \( C \) and \( c \) such that, for any \( \epsilon \in (0, \epsilon_0) \) and for \( p = \lceil c \log \epsilon^{-1} \rceil \), the following bounds hold for every \( k \in \{1, \ldots, d\} \) and \( j \in \{1, \ldots, i\} \):

\[
\|w - \Pi_{i,p} w\|_{L^\infty(D \times Y_i)} \leq C \epsilon, \quad \|\partial_k (w - \Pi_{i,p} w)\|_{L^\infty(D \times Y_i)} \leq C \epsilon p^2, \quad \|\partial_{j,k} (w - \Pi_{i,p} w)\|_{L^\infty(D \times Y_i)} \leq C \epsilon p.
\]

The result is classical; for completeness, we provide a proof of Lemma 4.4 in the Appendix.

**Lemma 4.5.** Let the assumptions of Lemma 4.4 hold and \( \|\cdot\|_{\infty} \) denote \( \|\cdot\|_{L^\infty(D \times Y_i)} \). Then there exist positive constants \( C \) and \( c \) such that, for any \( L \in \mathbb{N} \) and for \( p = \lceil c L \rceil \), the following bounds hold for every \( k \in \{1, \ldots, d\} \):

\[
\|\partial_{i,k} (w - \Pi_{i,p}^L w)\|_{\infty} \leq C p^2 2^{-L} \quad \text{and} \quad \|\partial_{i,k} (w - \Pi_{i,p}^L w)\|_{\infty} \leq C p^2 2^{-L}.
\]

We give a proof of Lemma 4.5 in the Appendix.

4.3. Low-rank tensor approximation. In this section, for \( i \in \{1, \ldots, n\} \) and \( L \in \mathbb{N} \), we consider

\[
w_{i}^{L} = (\Pi_{i}^{L} \Pi_{i,p} w_{ik})_{k=1}^{d} \in V_{i}^{L} \quad \text{and} \quad J_{i}^{L} = (\Pi_{i}^{L} \partial_{i,k} \Pi_{i,p} w_{ik})_{k'=1}^{d} \in (\bar{V}_{i}^{L})^{d \times d}
\]

with a suitable \( p_L \in \mathbb{N} \) as approximations to \( w_i \) and \( f_i w_i \), where \( w_i \) is the solution of (3.13). Then the approximation scheme (3.20) produces \( u_i \in V_i^L \) and \( v_i \in (\bar{V}_i^L)^d \).

Section 4.3.1 relates the error of the approximation scheme (3.20), bounded by (3.21)–(3.22), to the error of \( w_i^L \) and \( J_i^L \) as approximations to \( w_i \) and \( f_i w_i \) for all \( i \in \{1, \ldots, n\} \) and \( L \in \mathbb{N} \).

In Section 4.3.2, the error bounds proved in Section 4.3.1 are followed by a quantized tensor-rank analysis, which is based on auxiliary definitions and rank bounds which are also provided in Section 4.3.2.

The analysis is based on the following assumption regarding the approximation \( v_0^L \in \bar{V}_0^L \) to \( v_0 \).

**Assumption 4.6.** For all \( L \in \mathbb{N} \), the subspace \( S^L = \text{span}\{\bar{V}_0^L, v_0^L\}_{k=1}^{d} \) satisfies the following with some rank \( r_L \in \mathbb{N} \): for every \( \ell \in \{1, \ldots, L-1\} \), there exist subspaces \( \mathcal{S}_\ell^L \subset \mathbb{R}^{2^\ell} \) and \( \mathcal{M}_\ell^L \subset \mathbb{R}^{2^d(2^{d} - \ell)} \) of dimensions at most \( r_L \) such that \( S^L \subset \mathcal{S}^L \otimes \mathcal{M}_\ell^L \).
The purpose of Sections 4.3.3 and 4.3.4 is to bound the tensor ranks of the coefficients of \( u_1^n, \ldots, u_n^n \) and \( U^n U^n \) in terms of both \( L \in \mathbb{N} \) and \( r_L \in \mathbb{N} \). Finally, in Section 4.3.6 we restrict the setting to the case of \( d = 2 \) physical dimensions and invoke a result from [21, 23]. We remark that corresponding results in space dimension \( d = 3 \) are also available in [24]. This gives simultaneous bounds on \( r_L \) and on the errors \( \| u_0 - u_1^n \|_{L^2(D)} \) and \( \| v_0 - v_1^n \|_{L^2(D)^d} \) for every \( L \in \mathbb{N} \), which lead to an analogous result for \( \Psi_L u_1^n U^n \).

### 4.3.1. Accuracy of the approximation scheme

Under Assumption 4.1 differentiating the equation expressing the cell problem (3.13) in the strong form iteratively for \( i = n, \ldots, 1 \), one verifies that the solutions \( w_i \in V_i^d \) with \( i \in \{1, \ldots, n\} \) satisfy the assumption of Lemma (4.4). This gives that, with a positive constant \( c \), for any \( L \in \mathbb{N}_0 \) and for

\[
\rho_L = [cL],
\]

the approximations \( u_i^L \) and \( J_i^L \) defined by (4.13) satisfy the error bounds

\[
\| w_i - w_i^L \|_{W_i} \lesssim L^2 2^{-L} \quad \text{and} \quad \| f_i w_i - J_i^L \|_{L^\infty(D \times Y_{i-1} \times Y_i)} \lesssim L^2 2^{-L}
\]

The equivalence holds with a positive constant that is independent of \( L \in \mathbb{N}_0 \) and \( i \in \{1, \ldots, n\} \). Then the bounds (3.21)–(3.22) for the approximation scheme (3.20) show that the resulting approximations \( u_i^L \in V_i \) and \( v_i^L \in L^2(D \times Y_i)^d \) satisfy the bounds

\[
\| u_i - u_i^L \|_{V_i} \leq \| v_i - v_i^L \|_{L^2(D \times Y_i)^d} \lesssim \| v_0 - v_0^L \|_{L^2(D)^d} + L^2 2^{-L}
\]

with a positive equivalence constant independent of \( L \in \mathbb{N} \) and \( i \in \{1, \ldots, n\} \).

### 4.3.2. Auxiliary subspaces and results

For all \( L \in \mathbb{N}_0 \) and \( p \in \mathbb{N}_0 \), we will use the following notation for the sets of tensors obtained by evaluating \( d \)-variate algebraic and trigonometric polynomials of maximum degree at most \( p \) on a uniform tensor-product grid with \( 2^L \) nodes in each variable:

\[
\mathcal{P}_p^{L,d} = \bigotimes_{k=1}^d \Psi^L \mathcal{P}_p \subset \mathbb{C}^{2^dL} \quad \text{and} \quad \mathcal{P}_{\#p}^{L,d} = \bigotimes_{k=1}^d \Psi^L \mathcal{P}_{\#p} \subset \mathbb{C}^{2^dL}.
\]

Let us extend (4.11) and (4.16) by introducing, for all \( p \in \mathbb{N}_0 \), \( L \in \mathbb{N} \) and \( \lambda \in \mathbb{Z} \),

\[
\mathcal{P}_{p,\lambda} = \text{span} \{ \tilde{T}_\alpha(2^L \cdot) \}_{\alpha=-p}^p \quad \text{and} \quad \mathcal{P}_{\#p,\lambda}^{L,d} = \bigotimes_{k=1}^d \Psi^L \mathcal{P}_{p,\lambda} \subset \mathbb{C}^{2^dL}.
\]

We will use several results, stated below, to analyze the low-rank structure of the approximations \( u_i^L \) and \( v_i^L \) in \( V_i \) and \( L^2(D \times Y_i)^d \) with \( i \in \{1, \ldots, n\} \) and \( L \in \mathbb{N} \), defined by (4.20) and (4.13), as elements of the respective spaces \( Q_i^L \) with \( i \in \{1, \ldots, n\} \) and \( L \in \mathbb{N} \), given by (4.19).

**Proposition 4.7.** For all \( p \in \mathbb{N}_0 \), \( L \in \mathbb{N} \) and \( \ell \in \{1, \ldots, L - 1\} \), we have

\[
\mathcal{P}_{p}^{L,d} \subset \mathcal{P}_{p}^{\ell,d} \otimes \mathcal{P}_{p}^{L-\ell,d}.
\]

The embedding stated in Proposition 4.7 means the following: for every \( u \in \mathcal{P}_{p}^{L,d} \subset \mathbb{C}^{2^dL} \), there exist \( u' \in \mathcal{P}_{p}^{\ell,d} \subset \mathbb{C}^{2^dL} \) and \( u'' \in \mathcal{P}_{p}^{L-\ell,d} \subset \mathbb{C}^{2^dL} \) such that \( u = u' \otimes u'' \) in the sense of the Kronecker product of vectors (tensors). The proof follows trivially from the binomial formula applied to the standard basis of monomials.

An immediate consequence of Proposition 4.7 is that the tensor of the values of any \( d \)-variate polynomial of maximum degree at most \( p \in \mathbb{N} \) at any tensor-product uniform grid with \( 2^L \) entries in each dimension can be represented in the multilevel TT-MPS format with the transposition (2.9) with ranks not exceeding \((p + 1) d\). This was originally shown, in the case of \( d = 1 \), in [15, Corollary 13]. The language of
space factorization, which we adopt for Assumption 4.6, Proposition 4.7 and for the whole section, is different from that of \[36\, 37\, 24\, 15\], we use it here to mostly avoid lengthy expressions with numerous indices associated with nodes of tensor-product grids.

The additional notation \[4.17\] allows to state the following analog of Proposition 1.7 for trigonometric polynomials, which is an immediate consequence of the separability of the exponential function.

**Proposition 4.8.** For all \(p \in \mathbb{N}_0\), \(L \in \mathbb{N}\) and \(\ell \in \{1, \ldots, L - 1\}\), we have

\[
p_{\#p}^{L,d} \subset p_{\#p}^{\ell,d} \otimes p_{\#p,-\ell}^{L-\ell,d}.
\]

### 4.3.3. Approximation of the one-scale high-dimensional limit problem.

Apart from realizing arbitrary accuracy, the approximations given by (3.20) and (4.13) are also structured in the sense that

\[
\Psi^L_{i} u^L_{i} \in P^{L,d}_{pL} \otimes i \bigotimes_{j=1} v^L_{i} \in Q^L_{i} \otimes i \bigotimes_{j=1} p^{L,d}_{#(i+1-j)pL} \subset C^{2d+iL}_{\ell d}.
\]

Here, the operation "\(\otimes\)" between two spaces denotes taking the span of the pointwise products of all pairs of elements from the respective spaces.

**Theorem 4.9.** Let Assumptions 2.1, 4.1, and 4.10 hold and \((u_0, u_1, \ldots, u_n) \in V^L_{\alpha}\) be the solution of (3.6). Consider \(v^L_0 \in V^L_n\) with \(L \in \mathbb{N}\) satisfying Assumption 4.6 and such that \(\|v^L_0 - v^L_0\|_{L(\Omega)^d} \leq C_0 L^{\gamma_0} 2^{-\alpha L}\) for all \(L \in \mathbb{N}\) with \(\alpha \in (0, 1]\) and with positive constants \(C_0\) and \(\gamma_0\). Then the approximations \(u^L_{i} \in V^L_{i}\) with \(L \in \mathbb{N}\) and \(i \in \{1, \ldots, n\}\) given by (3.20) and (4.13) satisfy the following with positive constants \(\bar{C}\) and \(\bar{c}\):

For all \(L \in \mathbb{N}\) and \(i \in \{1, \ldots, n\}\), the bound \(\|u^L_{i} - u^L_{i}\|_{V^L_{i}} \leq \bar{C} L^{\gamma_0} 2^{-\alpha L}\) holds and the coefficient tensor \(\Psi^L_{i} u^L_{i}\) admits a decomposition of the form (2.12) with \((n + 1)L\) levels and ranks bounded from above by \(R^L = \bar{c} L^{n+1} d r^L\).

**Proof.** Consider \(i \in \{1, \ldots, n\}\) and \(L \in \mathbb{N}\). The claimed accuracy bound follows from (4.15).

To bound the first \(L\) ranks, we consider \(\ell \in \{1, \ldots, L\}\) and factorize \(Q^L_{i}\) so that the first factor is a subspace of \(C^{2d_i}\). Then this subspace corresponds to the \(\ell\) coarsest levels of the macroscale, and its dimension majorates the corresponding rank of \(\Psi^L_{i} u^L_{i}\). First, we obtain from Proposition 4.7 that \(\mathcal{P}^{\ell,d}_{(i+1)pL} \subset C^{2d_{(i+1)}}_{\ell d}\), where the dimension of \(\mathcal{P}^{\ell,d}_{(i+1)pL} \subset C^{2d(L-i)}\) is \((i + 1)pL + 1)\). On the other hand, by Assumption 4.6, we have \(\mathcal{S}^L \subset L^d_{\ell} \otimes C^{2d_{(L-i)}}\), where the dimension of \(L^d_{\ell} \subset C^{2d}\) does not exceed \(r^L\). This results in \(Q^L_{i} \subset L^d_{\ell} \otimes C^{2d_{(L-i)}} \otimes i \bigotimes_{j=1} p^{d}_{#(i+1-j)pL}\), where \(\tilde{L}^d_{\ell} = L^d_{\ell} \otimes p^{d}_{#(i+1-j)pL}\). We note then that \(\dim\tilde{L}^d_{\ell} \leq \dim L^d_{\ell} \cdot \dim p^{d}_{#(i+1-j)pL} \leq ((i + 1)pL + 1) d r^L \leq \bar{c} L^d r^L = R^L\) for a suitable positive constant \(\bar{c}\) independent of \(L\), due to the linear dependence (4.14) of \(p^L_{\ell}\) on \(L\).

To bound the other ranks, we now consider \(j \in \{1, \ldots, i\}\) and \(\ell \in \{0, \ldots, L\}\) and factorize \(Q^L_{i}\) so that the last factor is a subspace of \(C^{2d_{(L-i)+j+1)}}_{\ell d}\). This subspace corresponds to the \(L - \ell\) finest levels of the \(j\)th microscale and all levels of all finer microscales. The dimension of this subspace majorates the corresponding rank of \(u^L_{i} \in Q^L_{i}\.\)
Applying Proposition 4.8 we obtain $\mathcal{P}_{#(i+1-j)p_L}^{L,d} \subset \mathcal{P}_{#(i+1-j)p_L}^{L',d} \otimes \mathcal{P}_{#(i+1-j)p_L}^{L''-d}$, where the dimension of both the factors is $(2(i+1-j)p_L+1)^d$. Then we have

$$\mathcal{G}_L^i \subset \left\{ \mathcal{S}_L \otimes \mathcal{P}_{(i+1)p_L}^{L,d} \right\} \otimes \bigotimes_{m=1}^{j-1} \mathcal{P}_{#(i+1-m)p_L}^{L,d} \otimes \mathcal{P}_{#(i+1-j)p_L}^{L',d} \otimes \tilde{\mathcal{M}}_{j,t}^L$$

with $\tilde{\mathcal{M}}_{j,t}^L = \mathcal{P}_{#(i+1-j)p_L}^{L',d} \otimes \bigotimes_{m=j+1}^{L} \mathcal{P}_{#(i+1-m)p_L}^{L,d}$. For the last factor, we find that

$$\dim \tilde{\mathcal{M}}_{j,t}^L \leq \dim \mathcal{P}_{#(i+1-j)p_L}^{L',d} \otimes \bigotimes_{m=j+1}^{L} \mathcal{P}_{#(i+1-m)p_L}^{L,d} = \prod_{m=j}^{L} (2(i+1-m)p_L+1)^d \leq cL(i+1)^d = R_L^i.$$

As above, the latter inequality holds with a suitable positive constant $c$ independent of $L$ due to the linear dependence \([1.14]\) of $p_L$ on $L$. \(\square\)

### 4.3.4. Approximation of the multiscale problem.

In this section, we analyze the low-rank structure and accuracy of $\hat{U}^{\epsilon}_{V_n}$ as an approximation to $U^{\epsilon}v_n$. As in Section 4.3.3 we develop our analysis here under Assumptions 2.1, 4.1 and 4.6. Additionally, we make the following technical assumption, which simplifies the analysis of tensor structure in the present section.

**Assumption 4.10.** For every $i \in \{1, \ldots, n\}$, we have $L \geq \lambda_i - \lambda_{i-1}$.

We start with defining finite-element subspaces in which we will consider averaged approximations. To this end, we set

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \xi = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

and, for every $L \in \mathbb{N}$ and $i \in \{1, \ldots, n\}$, consider the matrices

$$\tilde{M}_{L-1}^i = I \otimes (\lambda_i - \lambda_{i-1}) \otimes \left( \frac{1}{2} \xi^T \right) \otimes (L - \lambda_i + \lambda_{i-1}) \quad \text{and} \quad M_{i-1}^L = I \otimes d_{\lambda_i - 1} \otimes \tilde{M}_{L-1}^i. \quad (4.20)$$

The action of the matrix $M_{L-1}^i$ on the coefficient tensor of a piecewise-constant function subordinate to a uniform partition with $2d(L+\lambda_{i-1})$ elements of linear size $2^{-(L+\lambda_{i-1})}$ consists in averaging the function over the $2d\lambda_i$ cells of scale $\varepsilon_i = 2^{-\lambda_i}$. The matrix $M_{L-1}^i$, on the other hand, represents the same averaging operation on subtensors corresponding to single cells of scale $\varepsilon_{i-1} = 2^{-\lambda_{i-1}}$. The order of the factors in (4.20) reflects the use of transposition (2.9) for the coefficient tensors, see (4.5).

Using the matrices introduced in (4.20), we iteratively define the following spaces:

$$\mathcal{S}_0^L = \mathcal{S}_L \otimes \mathcal{P}_{(n+1)p_L}^{L,d} \subset \mathbb{C}^{2^{dL}}$$

and

$$\mathcal{S}_i^L = (\tilde{M}_{L-1}^i \mathcal{S}_{i-1}^L) \otimes \mathcal{P}_{#(n+1-i)p_L}^{L,d} \subset \mathbb{C}^{2^{d(L_i + L)}} \quad (4.21)$$

for all $i \in \{1, \ldots, n\}$ and $L \in \mathbb{N}$. Eventually, we are interested the subspaces with index $i = n$, which are relevant due to the following result. It is a corollary of auxiliary technical Lemmas 6.1 and 0.2, which we present in the Appendix.

**Lemma 4.11.** For all $L \in \mathbb{N}$ and $v \in V_n^L$ such that $\tilde{\Psi}_n^L v \in \mathcal{G}_n^L$, we have $\hat{U}^{\epsilon}_v \in \tilde{\Psi}_n^L$.

Iterating (4.21) under Assumption 4.10 we arrive at

$$\mathcal{S}_i^L = (\tilde{M}_0^L \mathcal{S}_0^L) \otimes \left( \bigotimes_{j=1}^{i-1} \tilde{M}_j^L \mathcal{P}_{#(n+1-j)p_L}^{L,d} \right) \otimes \mathcal{P}_{#p_L}^{L,d} = (\tilde{M}_0^L \mathcal{S}_0^L) \otimes \left( \bigotimes_{j=1}^{i-1} \mathcal{P}_{#(n+1-j)p_L}^{\lambda_{i+1-j} - \lambda_{i-1}} \right) \otimes \mathcal{P}_{#p_L}^{L,d}, \quad (4.22)$$

16
where the second equality holds due to that \( \tilde{M}_j^L \mathcal{P}_{\#(n+1-j)^{j}P_L} = \mathcal{P}_{\#(n+1-j)^{j}P_L}^{\lambda_j-\lambda_{j-1},d} \) for any \( j \in \{1, \ldots, n\} \) by Proposition 4.8.

**THEOREM 4.12.** Let Assumptions 2.1, 4.1, and 4.10 hold and \( (v_0, u_1, \ldots, u_n) \in V_n \) be the solution of (3.6). Consider \( v_0 = \nabla u_0, v_n = \nabla u_0 + \nabla_1 u_1 + \cdots + \nabla_n u_n \) and \( v_m^L \in V^L \) with \( L \in \mathbb{N} \) satisfying Assumption 4.6 and such that \( \|v_0 - v_m^L\|_{L^2(D)^d} \leq C_0 L^n \gamma^{2-\alpha L} \) for all \( L \in \mathbb{N} \) with \( \alpha \in [0, 1] \) and with positive constants \( C_0 \) and \( \gamma_0 \). Then the approximations \( \mathcal{U}^L \mathcal{V}_n = \mathcal{U}^L v^L_n \in (V_n^{1+L})^d \) with \( L \in \mathbb{N} \) defined by (3.20) and (4.13) satisfy the following with positive constants \( \bar{C} \) and \( \bar{c} \) and \( \bar{\gamma} = \max\{2, \gamma_0\} \).

For all \( L \in \mathbb{N} \) sufficiently large (to satisfy Assumption 4.10), the error bound \( \|\mathcal{U}^L v_n - \mathcal{U}^L v^L_n\|_{L^2(D)^d} \leq \bar{C} L^n \gamma^{2-\alpha L} \) holds for \( \bar{\gamma} \) and with positive constants \( \bar{c} \) and \( \bar{\gamma} \). Theorem 4.12.

Proof. The claimed accuracy bound follows from (4.15) combined with (3.8).

The claimed rank bounds follows from the fact that \( \bar{\Psi}^{\lambda_n+L}_0 \mathcal{U}^L v^L_n \in \mathcal{S}^L \) by Lemma 4.11 combined with (4.22), Assumption 4.6 and Propositions 4.7 and 4.8.

First, let us consider \( \ell \in \{1, \ldots, \lambda_n - 1\} \). From Assumption 4.6 and Proposition 4.7, we obtain \( \mathcal{S}^L \subset \mathcal{D}^L \otimes \mathbb{C}^{2^d(L-\ell)} \) and \( \mathcal{P}^{L,d}_{(i+1)p_L} \subset \mathcal{P}^{L,d}_{(i+1)p_L} \otimes \mathbb{C}^{2^d(L-\ell)} \), where the dimensions of the first factors are bounded form above by \( r_L \) and \( ((i+1)p_L + 1)^d \) respectively. This implies the inclusion \( \mathcal{M}^L_0 \mathcal{S}^L \subset \mathcal{D}^L \otimes \mathbb{C}^{2^d(L-\ell)} \) with a subspace \( \mathcal{D}^L_\ell \subset \mathbb{C}^{2^d} \) of dimension at most \( r_L ((i+1)p_L + 1)^d \) and hence, by (4.22), also \( \mathcal{S}^L_\ell \subset \mathcal{D}^L_\ell \otimes \mathbb{C}^{2^d(L-\ell)} \). The first \( \lambda_n - 1 \) ranks of \( \bar{\Psi}^{\lambda_n+L}_0 \mathcal{U}^L v^L_n \) therefore do not exceed \( r_L ((i+1)p_L + 1)^d \).

To obtain a bound for all the remaining ranks at once, let us set \( \lambda_{n+1} = \lambda_n + L \) for notational convenience and consider \( \ell \in \{\lambda_n, \ldots, \lambda_{n+1} - 1\} \) with \( k \in \{1, \ldots, n\} \). Inasmuch as the corresponding factors indexed by \( j \in \{k, \ldots, n\} \) in (4.19) and (4.21) with \( i = n \) are completely analogous, the second part of the proof of Theorem 4.9 applies herein upon replacing \( L \) with \( \lambda_j + 1 - \lambda_j \) in superscript in the \( j \)-th term for every \( j \in \{k, \ldots, n\} \).

**4.3.5. The case of a separable scalar diffusion coefficient.** Improved bounds can be obtained under additional scale-separability assumptions on the diffusion coefficient \( A^c \). For example, let us consider the case where the function \( A \) is of separable form

\[ A = (a_0 \otimes a_1 \otimes \cdots \otimes a_n) K, \]

where \( K_i \) is a symmetric positive-definite matrix with spectrum in \([\gamma, \Gamma]\) for some positive constants \( \gamma \) and \( \Gamma \), \( a_0 \in L^\infty(D), a_i \in C_0(Y_i) \) for every \( i \in \{1, \ldots, n\} \) and \( J \) is the identity matrix of order \( d \). Let us also assume that \( a_0(x) > 1 \) for a.e. \( x \in D \) and \( a_i(y) > 1 \) for a.e. \( y \in Y_i \) and every \( i \in \{1, \ldots, n\} \).

Let us consider the following modification of Assumption 3.1.

**ASSUMPTION 4.13** (on a coefficient \( A_i \) with \( i \in \{0, \ldots, n\} \) microscales, with positive constants \( \gamma \) and \( \Gamma \)). The coefficient \( A_i \) is of the form \( A_i = (a_0 \otimes a_1 \otimes \cdots \otimes a_{i-1} \otimes a_i) K_i \), where \( K_i \) is a symmetric positive-definite matrix with spectrum in \([\gamma, \Gamma]\).

Note that under the conditions imposed on \( A \) in this section, Assumption 4.13 holds for \( i = n \) with \( K_n = K \).

For every \( i \in \{1, \ldots, n\} \), under Assumption 4.13 the problem (3.11) becomes

\[ G_i(x, y_1, \ldots, y_{i-1}, \psi, \phi) = a_0(x) a_1(y_1) \cdots a_{i-1}(y_{i-1}) \int_{Y_i} a_i(y) \nabla^T K_i \nabla \phi, \]

\[ F_i(x, y_1, \ldots, y_{i-1}, \xi, \phi) = -a_0(x) a_1(y_1) \cdots a_{i-1}(y_{i-1}) \int_{Y_i} a_i \xi \nabla K_i \nabla \phi \]

(4.23)

for all \( \psi, \phi \in H^1_\#(Y_i)/\mathbb{R} \) and for a.e. \( (x, y_1, \ldots, y_{i-1}) \in D \times Y_{i-1} \). As a result, the solution of (3.12) is independent of \( x \in D \) and \( y_{i-1} \in Y_{i-1} \), so that \( w_i \in (H^1_\#(Y_i)/\mathbb{R})^d \).
Then the upscaled coefficient $A_{i-1}$, given by (3.14), satisfies Assumption 4.13 with $i-1$ microscales and

$$K_{i-1} = \int_{Y_i} a_i (I + J w_i) K_i (I + J w_i)^T = \int_{Y_i} a_i (I + J w_i) K_i > 0,$$

where the integrand is a function of a single microscale variable taking values in $Y_i$.

Iterating this argument, one finds that Assumption 4.13 holds for every $i \in \{0, \ldots, n-1\}$ with the same constants as for $i = n$ and that the scale-interaction functions and their gradients, given by (3.19), are separable. Consequently, the factors in the right-hand sides of equalities (3.19), depending on variables corresponding to only a single scale each, can be approximated independently. This allows to consider, instead of the spaces $\mathcal{Q}^l$ and $\mathcal{S}^l$ with $i \in \{1, \ldots, n\}$, specified in (4.19) and (4.21), spaces of separable tensors, and to thereby avoid the dependence on $n$ of the exponent in the rank bounds for $\Psi^n u_{\ell n}^L, \Psi^L u_{nm}^L$ and $\Psi^L u_{nm}^L$ in Theorems 4.9 and 4.12.

4.3.6. The case of two dimensions: approximation of functions with corner singularities. In the remainder of this section, we consider the case of $d = 2$. We will use spaces of functions defined on a polygonal domain that are analytic on the closure of the domain except a number of points where algebraic singularities of certain order may occur.

With any set $\Theta$ of a finite number of distinct points in $\mathbb{R}^2$, we associate the weight function $\chi_\Theta$ given by

$$\chi_\Theta(x) = \prod_{\theta \in \Theta} \|x - \theta\|_2 \quad \text{for all} \quad x \in \mathbb{R}^2,$$

where $\| \cdot \|_2$ denotes the Euclidean norm on $\mathbb{R}^2$.

To quantify the analytic regularity of solutions to the high-dimensional one-scale problem, we use weighted Sobolev spaces and associated countably normed classes as introduced in [30, 29, 6, 5, 7] and denoted here by $H^m_{\Theta, \beta}(\Omega)$ and $C^\ell_{\Theta, \beta}(\Omega)$ with $\ell \in \{1, 2\}$, $m \in \{0, 1, \ldots, \ell \}$ and $\beta \in [0, 1)$, where $\Omega \subset \mathbb{R}^2$ is a polygonal domain and $\Theta$ is a set of $S \in \mathbb{N}$ distinct points in $\Omega$.

Specifically, we will use the following weighted Sobolev spaces:

$$H^m_{\Theta, \beta}(\Omega) = \{ u : \Omega \to \mathbb{R} : \chi_\Theta^{\beta + |\alpha|} \partial^\alpha u \in L^2(\Omega) \quad \text{if} \quad 0 \leq |\alpha| \leq m \}$$

for all $\ell \geq 0$ and

$$H^m_{\Theta, \beta}(\Omega) = \{ u \in H^{\ell-1}(\Omega) : \chi_\Theta^{\beta + |\alpha| - \ell} \partial^\alpha u \in L^2(\Omega) \quad \text{if} \quad 0 \leq |\alpha| \leq m \}$$

for all $m, \ell \geq 1$, where the differentiation is understood in the weak sense. By setting

$$|u|_{H^m_{\Theta, \beta}(\Omega)}^2 = \sum_{|\alpha| = m} \| \chi_\Theta^{\beta + m - \ell} \partial^\alpha u \|_{L^2(\Omega)}^2 \quad \text{for all} \quad u \in H^m_{\Theta, \beta}(\Omega),$$

we introduce $|\cdot|_{H^m_{\Theta, \beta}(\Omega)}$, a seminorm on $H^m_{\Theta, \beta}(\Omega)$. Also, by setting

$$\|u\|_{H^m_{\Theta, \beta}(\Omega)}^2 = \sum_{k=0}^m |u|_{H^m_{\Theta, \beta}(\Omega)}^2 \quad \text{for} \quad u \in H^m_{\Theta, \beta}(\Omega), \quad m \geq 0,$$

$$\|u\|_{H^m_{\Theta, \beta}(\Omega)}^2 = \|u\|_{H^{\ell-1}(\Omega)}^2 + \sum_{k=\ell}^m |u|_{H^m_{\Theta, \beta}(\Omega)}^2 \quad \text{for} \quad u \in H^m_{\Theta, \beta}(\Omega), \quad m \geq \ell \geq 1,$$

we define $\| \cdot \|_{H^m_{\Theta, \beta}(\Omega)}$ a norm on $H^m_{\Theta, \beta}(\Omega)$ for any $\ell, m \in \mathbb{N}_0$ such that $m \geq \ell$.

DEFINITION 4.14 (analyticity of a function with point algebraic singularities, with positive constants $M$ and $\rho$). Let $\beta \in [0, 1)$, $\Omega \subset \mathbb{R}^2$ be a polygonal domain, $\Theta$ be a
finite set of distinct points in \( \overline{\Omega} \) and \( \ell \in \{1, 2\} \). Then \( u \in C^\ell_{0,\beta}(\Omega) \) if \( u \in H^\ell_{0,\beta}(\Omega) \) and there exist positive constants \( M \) and \( \rho \) such that, for all \( \alpha \in \mathbb{N}_0^2 \) with \( |\alpha| \geq \ell - 1 \),

\[
\sup_{x \in \Omega} \chi_\Theta^{\beta+|\alpha|} \frac{|\partial^\alpha u(x)|}{|x^\alpha|!} \leq M \rho^{|\alpha|} |\alpha|!.
\]

The following result is a consequence of [5, Theorems 3.4–3.5] for the iterated-homogenization scheme of [3, 13, 16].

**Proposition 4.15.** Assume that \( \Theta \) is the set of vertices of the unit square \( \Omega = (0, 1)^2 \). Let Assumptions 2.1 and 4.1 hold. Then the solution \( u_0 \) of the homogenized problem (3.18) satisfies \( u_0 \in C^2_{0,\beta}(\Omega) \) with some \( \beta \in [0, 1) \).

Note that this statement remains valid for an arbitrary curvilinear polygon with an analytic boundary [5]. However, even in the setting of Assumption 4.1, the exponent \( \beta \) depends on the transformation diagonalizing the diffusion coefficient at the vertices of \( \Omega \) and can be estimated in terms of the spectral bounds \( \gamma \) and \( \Gamma \).

We will now combine the weighted-analyticity statement of Proposition 4.15 with rank bounds for the QTT-FE approximation of functions from \( C^2_{0,\beta}(\Omega) \). Let \( u_0 \in V \cap C^2_{0,\beta}(\Omega) \). Then the following holds with positive constants \( C \) and \( c \).

For every \( L \in \mathbb{N} \), there exist \( v_0^L \in V^L \) and \( v_0^L \in (V^L)^d \) satisfying Assumption 4.16 with rank bounds for the QTT-FE approximation of functions from \( C^2_{0,\beta}(\Omega) \) in [21, 23].

**Theorem 4.16.** Assume that \( \beta \in (0, 1) \) and \( \Theta \) is the set of the vertices of \( \Omega \). Let \( u_0 \in V \cap C^2_{0,\beta}(\Omega) \). Then the solution \( u_0 \) follows immediately from either [21, Theorem 5.3.7] or [23, Theorem 5.16]. In particular, for all \( L \in \mathbb{N} \) and \( \ell \in \{0, 1, \ldots, L\} \), there exist subspaces \( \mathcal{L}^L \subset \mathbb{R}^{2^{\ell+1}} \) and \( \mathcal{M}^L \subset \mathbb{R}^{2^{d(\ell-1)}} \), both of dimension at most \( r_L = \lfloor c L^2 \rfloor \), where \( c \) is a positive constant independent of \( L \), such that \( \Psi_0^L u_0^L \in \mathcal{L}^L \oplus \mathcal{M}^L \).

To obtain the statement regarding \( v_0^L \) with \( L \in \mathbb{N} \), we consider \( v_0^L \) with \( L \in \mathbb{N} \), we consider \( v_0^L \) with \( L \in \mathbb{N} \), we consider \( v_0^L \) with \( L \in \mathbb{N} \). Let \( u_0 \in V^L \) for all \( k \in \{1, \ldots, d\} \) and \( L \in \mathbb{N} \). Bound the piecewise-polynomial structure used to establish rank bounds in [21, Lemma 4.6.1 and Corollary 4.6.2] and in [23, Lemma 5.13 and Corollary 5.14].

Inspecting those proofs, one concludes that the rank analysis given there applies to\( v_0^L \) with \( k \in \{1, \ldots, d\} \) and \( L \in \mathbb{N} \): for every \( \ell \in \{0, 1, \ldots, L\} \), we have \( \Psi_0^L u_0^L \in \mathcal{L}^L \oplus \mathcal{M}^L \), where the subspaces \( \mathcal{L}^L \) and \( \mathcal{M}^L \) are identical to those constructed in the same proofs for \( \Psi_0^L u_0^L \). This shows that \( u_0^L \) and \( v_0^L \) satisfy Assumption 4.16 with \( r_L = \lfloor c L^2 \rfloor \). \( \square \)

The following is a corollary of Theorems 4.9, 4.12, and 4.16 and Proposition 4.15.

**Corollary 4.17.** Assume that \( D = (0, 1)^2 \). Let Assumptions 2.1, 4.4, and 4.10 hold and \( (u_0, u_1, \ldots, u_n) \in V_n \) be the solution of (3.6) and \( v_n = \nabla u_0 + \nabla u_1 + \cdots + \nabla u_n \). Then the approximations \( u_0^L \in V^L \) with \( L \in \mathbb{N} \) and \( \ell \in \{1, \ldots, n\} \) and \( \mathcal{U}^L u_0^L = (\mathcal{V}^L)^d \) with \( L \in \mathbb{N} \), defined by (3.20) and (4.13), satisfy the following with \( \beta \in [0, 1) \) and with positive constants \( C \) and \( c \).

For all \( L \in \mathbb{N} \) sufficiently large (to satisfy Assumption 4.10), the error bound \( \sum_{i=0}^n \|u_i - u_i^L\|_V + \|\mathcal{U}^L u_0 - \mathcal{U}^L u_0^L\|_{L^2(\Omega)^d} \leq C L^2 \delta \|L^{-\frac{1}{2}}\|_L \) holds and each of the coefficient tensors \( \Psi_0^L u_0^L \) with \( i \in \{0, \ldots, n\} \) and \( \Psi_0^L u_0^L \) admit decompositions of the form (2.12) with ranks bounded from above by \( R_L = \tilde{c} L^{2(n+2)} \).

**5. Numerical results.** We implement two approaches for the approximate numerical solution of the multiscale problem (2.3).

The first approach consists in immediately solving the discretization (2.11) of the multiscale problem (2.3), seeking the solution in the form of the multilevel TT-
MPS decomposition \((2.12)\). The implementation is based on the recent result \([8]\) on the preconditioning of elliptic second-order operators, which allows to avoid the ill-conditioning and numerical instability associated with the use of fine discretizations (large \(L\)) and with using the multilinear decomposition \((2.12)\) instead of storing all the entries of the coefficient tensor independently. We modified the BPX preconditioner developed in \([8]\), following the original derivations, so as to accommodate the Dirichlet boundary conditions of \((2.3)\), imposed on the whole of the boundary.

The second approach consists in solving the high-dimensional one-scale limit problem \((3.6)\) in a form analogous to \((2.12)\) and computing \(u^L_0\) and \(U^\varepsilon v_n^L\) that approximate \(u^\varepsilon\) and \(\nabla u^\varepsilon\) in the respective \(L^2\) norms.

We emphasize that the first approach bypasses the limit problem \((3.6)\) and aims at solving directly discretizations of the multiscale problem \((2.3)\). The second approach, on the contrary, explicitly involves the limit problem as an auxiliary computational problem. Neither approach requires the computation of effective (or “homogenized”, “upscaled”) coefficients. We cover the second approach only in some of the experiments, for reference and comparison. We did not incorporate the BPX preconditioner developed in \([8]\) in the second approach, so it can be used only for relatively coarse virtual grids (up to \(L = 15\) when \(d = 1\)). The source code developed for our numerical experiments is publicly available \(\footnote{\url{https://bitbucket.org/rakhuba/msqtt2d_numexp}}\).

5.1. Two scale problem, \(n = 1, d = 1\). We start with an instance of the problem \((2.2)–(2.3)\) with two scales, \(D = Y = (0, 1)\), \(d = 1\) and

\[
\frac{d}{dx} \left( A^\varepsilon(x) \frac{du^\varepsilon}{dx} \right) = 1 \quad \text{in} \quad D, \quad u^\varepsilon(0) = u^\varepsilon(1) = 0,
\]

where

\[
A(x, y) = \frac{2}{3}(1 + x)(1 + \cos^2(2\pi y)) \quad \text{for all} \quad x \in D \quad \text{and} \quad y \in Y, \tag{5.2}
\]

see \([18, \text{Section 6.1}]\), where this problem was solved with a sparse-grid FEM approach. The two-scale limiting equation has the exact (homogenized) solution \(u_0\) given by

\[
u_0(x) = \frac{3}{2\sqrt{2}} \left( x - \frac{\log(1 + x)}{\log 2} \right) \tag{5.3}
\]

for all \(x \in D\) and \(y \in Y\) and the scale interaction term \(u_1\) is given by

\[
u_1(x, y) = \frac{3}{2\sqrt{2}} \left( 1 - \frac{1}{(1 + x)\log 2} \right) \left( \frac{1}{2\pi} \tan^{-1} \left( \frac{\tan 2\pi y}{\sqrt{2}} \right) - y + \phi(y) + C \right) \tag{5.4}
\]

for all \(x \in D\) and \(y \in Y\), where \(\phi\) is chosen to ensure continuity of \(u_1\):

\[
\phi(y) = \begin{cases} 
0, & y \in [0, 1/4] \\
\frac{1}{2}, & y \in (1/4, 3/4] \\
1, & y \in (3/4, 1]. 
\end{cases} \tag{5.5}
\]

We consider two approaches to approximate solution of the problem \((2.2)–(2.3)\) with \((5.2)\): QTT-FEM discretization of the multiscale problem \((2.2)–(2.3)\) with \((5.2)\) and the QTT-FEM discretization of the corresponding high-dimensional limit problem \((3.6)\). For the first one we introduce nested grids with \(2^\ell - 1\) interior points and the corresponding FE discretization using piecewise-linear hat functions. For every \(\ell\) the Galerkin solution is parametrized by a \(2^\ell\)-component vector \(u^{\varepsilon, \ell}\), including zero coefficient of the basis function at corresponding to node 1.

The multidimensional limiting one-scale problem is discretized using tensor product basis functions with \(2^\ell\) basis functions both for the physical variable \(x\) and
for fast variables $y_i$. This discretization produces coefficient tensors $u_i^\ell \in \mathbb{R}^{2^{(i+1)\ell}}$ with $i \in \{0, 1, \ldots, n\}$.

The goal is to find QTT approximations $u^\epsilon_qtt$ and $u^\epsilon_qtt$ with $i = 0, \ldots, n$ of $u^\epsilon$ of the multiscale problem and $u_i^\ell$ with $i = 0, \ldots, n$ of the one-scale limit problem respectively. We denote the $|\cdot|_{H^1(D)}$ error corresponding to $u^\epsilon_qtt$ as follows: $\delta^\text{exact}_\epsilon = |u^\epsilon_qtt - u^\epsilon|_{H^1(D)}$. Since the exact solution $u^\epsilon$ is not available, we use instead the extrapolated solution

$$u^\epsilon_{\text{ext}} = 2u^\epsilon_{\text{qtt}} - u^\epsilon_{\text{qtt}}$$

with $L = 50$. In numerical experiments we therefore measure the following error:

$$|u^\epsilon_qtt - u^\epsilon_{\text{ext}}|_{H^1(D)} \approx |u^\epsilon_qtt - u^\epsilon|_{H^1(D)}.$$  

As for the one-scale limit problem corresponding to the problem $[2.2],[2.3]$ with $[5.2]$, we have its exact solution $(u_0, u_1)$ available through $[5.3]$ and $[5.4]$. So errors can be exactly computed as $\delta^\ell = |\cdot|_{H^1(D)}$, where $|\cdot|_{H^1(D)} = \|\nabla u_0\|_{L^2(D)} + \sum_{i=1}^{n} \|\nabla_i u_1\|_{L^2(D \times Y_1 \times \cdots \times Y_i)}$.

To find QTT approximations $u^\epsilon_qtt$ and $u^\epsilon_qtt$ with $i \in \{0, 1, \ldots, n\}$, we take the two approaches described in the beginning of Section 5. Figures 5.1a and 5.1b illustrate convergence with respect to the virtual grid level $\ell$ for each of them. In the both cases as anticipated we observe first order convergence.

Next we investigate the QTT rank dependence of $u^\epsilon_qtt$ of the QTT-FE solution of the multiscale problem $[2.2],[2.3]$ with $[5.2]$. To this end, we first approximate $u^\epsilon$ by calculating $u^\epsilon_qtt$ with $10^{-12}$ tolerance of QTT arithmetic and amen_solve, which is utilized to solve arising linear systems. Then we calculate the error $\delta^\ell$

$$\delta^\ell = |u^\epsilon_qtt - u^\epsilon_{\text{ext}}|_{H^1(D)}.$$ (5.7)

Finally, we calculate a sequence of truncated representations $\text{round}(u^\epsilon_qtt, \text{tol})$ for different tolerance values $\text{tol}$. We introduce notation $u^\epsilon_qtt(\tau^\ell) = \text{round}(u^\epsilon_qtt, \tau^\ell)$. The goal is to find the largest $\tau^\ell$ so that the following inequality holds:

$$|u^\epsilon_qtt - u^\epsilon_{\text{ext}}|_{H^1(D)} \leq 2 |u^\epsilon_qtt(\tau^\ell) - u^\epsilon_{\text{ext}}|_{H^1(D)}.$$ (5.8)
where $u_{\text{qtt}}^{\varepsilon,\ell}[\tau_{\ell}]$ is the FE interpolant:

$$u_{\text{qtt}}^{\varepsilon,\ell}[\tau_{\ell}] = \sum_{j \in I_{\ell}} u_{\text{qtt}}^{\varepsilon,\ell}[\tau_{\ell}] j \varphi_j$$

Figure 5.1a presents the dependence of the rank of $\text{round}(u^{\varepsilon,\ell}, \tau_{\ell})$ against the $H^1$ error $\delta_{\ell}$:

$$\delta_{\ell} = \left| u_{\text{qtt}}^{\varepsilon,\ell}[\tau_{\ell}] - u_{\text{ext}}^{\varepsilon} \right|_{H^1(D)}$$  \hspace{1cm} (5.9)

Next we investigate the QTT rank dependence of the QTT-FEM solution $u_{\text{lim-qtt}}^{\ell}$ of the high-dimensional, one-scale limiting problem: $u_{\text{lim}}^{\ell} = u_0^{\ell} + u^{\varepsilon} u_1^{\ell}$. We set $\varepsilon = 2^{-\ell_0}$ thus obtaining solution given by coefficient tensor $u_{\text{lim}}^{\ell}$ of length $2^{\ell_0+\ell_0}$, which is approximated in QTT format by $u_{\text{lim-qtt}}^{\ell}$.

For the both cases we observe polylogarithmic scaling of the effective QTT-rank of both $u_{\text{qtt}}^{\varepsilon}$ and $u_{\text{lim-qtt}}^{\ell}$ with respect to the error in $|\cdot|_{H^1(D)}$ or with respect to the truncation parameter $\tilde{\delta}$:

$$r = O(\log^\kappa \tilde{\delta}^{-1}).$$  \hspace{1cm} (5.10)

In Figures 5.3a and 5.3b we fit the parameter $\kappa$. Figure 5.3a illustrates that $\kappa$ barely depends on $\varepsilon$.

5.2. $(n + 1)$-scale problem. In this section, we consider the problem (2.2)–(2.3) with $n + 1$ scales, $D = Y = (0, 1)$ and

$$A(x,y_1,\ldots,y_n) = \left(\frac{2}{3}\right)^n (1 + x) \prod_{i=1}^n (1 + \cos^2(2\pi y_i))$$  \hspace{1cm} (5.11)

for all $x \in D$ and $y_1,\ldots,y_n \in Y$. We discretize the problem using QTT-FEM with number of virtual grid levels $L = 50$. We fix the finest scale parameter to be $\varepsilon_n = 2^{-20} \approx 10^{-6}$ and then select the remaining scale parameters as follows

$$\varepsilon_k = 2^{2(n-k)}\varepsilon_n, \hspace{0.5cm} k = 1,\ldots,n-1.$$  

In Figure 5.4 the effective rank values (obtained for the fixed truncation threshold $10^{-8}$) against the number of scales are presented. In this plot, we observe superlinear
Figure 5.3: Multiscale multiscale problem \((2.2)\)–\((2.3)\) with the coefficient given by \((5.2)\). (a) QTT-FEM for the multiscale problem: effective rank \(r\) vs. \(|H^{1}(D)|\)-error for different \(\varepsilon\). (b) QTT-FEM for the corresponding one-scale limit problem: effective rank \(r\) vs. rounding parameter \(\tilde{\delta}\) for \(\ell = 10\) and \(\ell_{\varepsilon} = 17\).

Figure 5.4: QTT-FEM for the multiscale problem \((2.2)\)–\((2.3)\) with the coefficient given by \((5.11)\): effective rank \(r\) vs. number \(n + 1\) of scales.

The growth of the effective rank in the given range of the number of scales. In absolute values, the effective rank increased approximately from 2.2 for \(n = 1\) to 3.8 for \(n = 9\), which only leads to a moderate increase of the total amount of work to solve the problem.

5.3. Two scale problem in two physical dimensions. In this section, we consider the problem \((2.2)\)–\((2.3)\) with two scales, \(D = Y = (0,1)^2\) and \(A = aI\), where \(I\) is the identity matrix of order two and

\[
a(x, y) = \left(1 + \cos^2(2\pi y_1)\right) \left(1 + \cos^2(2\pi y_2)\right)
\]

\((5.12)\)

for all \(x \in D\) and \(y = (y_1, y_2) \in Y\).

Similarly to the one-dimensional case, we introduce nested tensor-product grids with \((2^\ell - 1)^2\) interior points (see Sections \(2.1.1\) and \(4.1\)). On this grid we introduce FE basis functions that are tensor product of one-dimensional piecewise-linear hat functions. Then for every \(\ell\) the Galerkin solution is parametrized by the \(2^{2\ell}\)-component vector \(u^{\ell}\). The error and ranks are measured as described in Section \(5.1\). In Figure 5.5a we plot the error w.r.t. the extrapolated solution \((5.6)\).
against virtual grid level $\ell$. As anticipated we observe first-order convergence with respect to the meshwidth $h_\ell = 2^{-\ell}$.

Figure 5.5 presents effective numerical rank of $\text{round}(u^{\epsilon,\ell}, \tau_\ell)$ with $\tau_\ell$ being the smallest positive value satisfying (5.8). We fit the effective numerical rank versus $\delta_\ell$ defined in (5.9) using $r = O(|\log \delta_\ell|)$. As for the case with one physical dimension, the fitted values of $\kappa$ hardly depend on the scale parameter $\epsilon$.

6. Conclusions and Generalizations. The present analysis and numerical experiments is focused on the model linear elliptic multiscale problem (2.3). Here, the physical length scales are assumed to be asymptotically separated, and the dependence of the diffusion coefficient $A^\epsilon$ on the fast variables $y_1, ..., y_n$ is assumed to be periodic. Similar structure and results hold for other types of PDEs (e.g. [46, 48] and the references there). The corresponding development of QTT-FE approaches for these problem classes is a natural extension of the present analysis.

The assumptions allow to consider, instead of the original $d$-dimensional multiscale problem, a one-scale limit problem which is high-dimensional. Analogous high-dimensional one-scale limit problems are obtained for perforated materials, and for so-called reticulated structures, as well as so-called lattice-materials; we refer to the survey [13] and to the references there. Additionally, we point out that high-dimensional one-scale limit problems with the same, tensorized structure as those considered here arise also for certain non-periodic multiscale problems, which fall into the class of the so-called homogenization structures, as proposed by Nguetseng in [34]. We also emphasize that analogous results are available for nonlinear problems with multiple scales; we refer to [17] and the references there for further details. The results of the present paper indicate that the resulting (nonlinear) one-scale high-dimensional limit problems can also be solved efficiently by QTT-FE discretization, combined with a nonlinear solver.

We obtained the QTT rank bounds of the solution of the high-dimensional, one-scale limit problem under strong (analyticity) assumptions on the data which implied, as we showed, the corresponding analyticity of the solutions $u_i(x, y_1, ..., y_i)$; this, in turn, allowed us to prove bounds on the TT-rank of the solution that are logarithmic in accuracy. This naturally leads to the question whether analogous results can be expected in the case that we do not have analyticity. Consider, for example, the case where the unit cells $Y_i = (0, 1)^d$ have ‘holes’, i.e. $Y_i = (0, 1)^d \setminus O_i$, where $O_i \subset \subset Y_i$ is polyhedral, e.g. a cube centered at the point $(1/2, ..., 1/2)$ with edge length $1/2$. The corresponding generalization of unfolding homogenization is given in [12]. In this case, the gradient $v_n$, given by (3.7), exhibits singularities on $\partial O_i$ with respect to the...
ith microscopic variable, for each \( i = 1, \ldots, n \), so that analyticity of \( v_i \) with respect to \( y_i \in \overline{Y_i} \setminus \overline{O_i} \) can not be expected anymore. Regularity results for the parametric unit-cell problems in countably normed spaces are available (for \( n = 1 \) microscale and \( d = 2 \) space dimensions) in [32]. When combined with the QTT-FE approximations from [23] (in space dimension \( d = 2 \)), also in this case, QTT-FE approximation rate and rank bounds completely analogous to the results in the present note can be obtained.
Applying Proposition 4.2 to these representations, we obtain the claimed bounds.

where, due to (4.8) and (4.10), the coefficients satisfy

\[
\kappa_\alpha = \kappa_{\alpha_1} \cdots \kappa_{\alpha_d}
\]

for all \( \alpha \in \mathbb{N}_d \) and \( \beta_1, \ldots, \beta_i \in \mathbb{Z}^d \) with \( \kappa_0 = 1, \kappa_{\pm \alpha} = 2 (-1)^\alpha \) for each \( \alpha \in \mathbb{N} \) and \( \kappa_\alpha = \kappa_{\alpha_1} \cdots \kappa_{\alpha_d} \) for every \( \alpha \in \mathbb{N}^d \).

The entire function \( z : \mathbb{C} \rightarrow C \) given by \( z(\zeta) = (1 - \cos 2\pi \zeta) / 2 \) for all \( \zeta \in \mathbb{C} \) bijectively maps each of the intervals \((0, 1/2)\) and \((1/2, 1)\) onto \((0, 1)\). Then, introducing \( \mathcal{X} = \zeta^{\otimes d} \otimes \text{id} : \mathbb{C}^{(i+1)d} \rightarrow \mathbb{C}^{(i+1)d} \), we can substitute \( \mathcal{X} \) in (6.2) to express the coefficients of \( w \) as follows:

\[
c_{\alpha, \beta_1, \ldots, \beta_i} = 2^{-d} \sum_{\sigma \in \{\pm\}^d} \tilde{c}_{\sigma, \alpha, \beta_1, \ldots, \beta_i}
\]
for all \( \alpha \in \mathbb{N}_0^d \) and \( \beta_1, \ldots, \beta_i \in \mathbb{Z}^d \), where \( \sigma \odot \alpha = (\sigma_1 \alpha_1, \ldots, \sigma_d \alpha_d) \) for any \( \sigma \in \{\pm 1\}^d \) and \( \alpha \in \mathbb{N}_0^d \) and

\[
\hat{c}_{\beta_0, \beta_1, \ldots, \beta_i} = k_{\alpha} \left\langle \left( \bigotimes_{k=1}^d \hat{T}_{j \beta_k} \right) \otimes \left( \bigotimes_{j=1}^i \bigotimes_{k=1}^d \hat{T}_{j \beta_k} \right), w \circ \mathcal{I} \right\rangle_{L^2(D \times Y_i)} (6.4)
\]

for all \( \beta_0, \beta_1, \ldots, \beta_i \in \mathbb{Z}^d \).

For every \( \delta > 0 \), the function \( z \) bijectively maps \( \mathcal{E}_\delta = \{ \xi - i \eta: \xi \in (0, 1), \eta \in (0, \delta) \} \subset \mathbb{C} \) onto \( \mathcal{E}_\delta = \{ (1 - a_\eta \cos 2\pi \xi)/2 - i (b_\eta \sin 2\pi \xi)/2: \xi \in (0, 1), \eta \in (0, \delta) \}, \) where \( a_\eta = \cosh 2\pi \eta \) and \( b_\eta = \sinh 2\pi \eta \) for every \( \eta > 0 \). Note that \( \mathcal{E}_\delta \cup ((1 - a_\delta)/2, 1] \) is the image of the standard open Bernstein ellipse with parameter \( \rho = e^{2 \pi \delta} \) (with foci \( \pm 1 \) and semi-axes \( a_\delta \) and \( b_\delta \)) under the affine mapping \( \mathbb{C} \ni z \mapsto (1 - z)/2 \in \mathbb{C} \). Since the function \( w \) is analytic on \( D \times Y_i \) by assumption, it admits extension by analytic continuation to an open neighborhood of \( D \times Y_i \). Specifically, for some \( \delta_{i0}, \delta_{i1}, \ldots, \delta_{ii} > 0 \), it has an extension to \( \mathcal{G}_i \), where

\[
\mathcal{G}_i = \left\{ \bigotimes_{k=1}^d \mathcal{E}_{\delta_{k0}} \right\} \times \left\{ \bigotimes_{j=1}^i \bigotimes_{k=1}^d \mathcal{E}_{\delta_{kj}} \right\},
\]

that is holomorphic on \( \mathcal{G}_i \) and continuous on \( \mathcal{G}_i \). We identify the original function \( w \) with this (unique) extension and set \( M_i = \sup_{z \in \mathcal{G}_i} |w(z)| \). For the domain

\[
\mathcal{D}_i = \bigotimes_{j=0}^{i} \bigotimes_{k=1}^d \mathcal{E}_{\delta_{kj}},
\]

we have \( \mathcal{G}_i = \mathcal{I}(\mathcal{D}_i) \) and \( \sup_{\zeta \in \mathcal{D}_i} |(w \circ \mathcal{I})(\zeta)| = M_i \). Furthermore, \( w \circ \mathcal{I} \) is holomorphic on \( \mathcal{D}_i \), continuous on \( \mathcal{D}_i \), \( \mathcal{D}_i = \mathcal{I}(\mathcal{D}_i) \) and one-periodic with respect to each of its \( (i + 1)d \) variables. Using these properties and applying the Cauchy–Goursat theorem for the domain \( \mathcal{D}_i \), we obtain

\[
\hat{c}_{\beta_0, \beta_1, \ldots, \beta_i} = k_{\alpha} \left( \int_{x_j = \iota \beta_j}^{x_j = \iota \beta_j + \iota \beta_{j+1}} \left( \bigotimes_{j=0}^i \bigotimes_{k=1}^d \hat{T}_{j \beta_k} \right) (w \circ \mathcal{I}) \right)
\]

and hence

\[
|\hat{c}_{\beta_0, \beta_1, \ldots, \beta_i}| \leq M_i k_{\alpha} \exp \left( - \sum_{j=0}^i 2\pi \delta_{kj} |\beta_j| \right) (6.5)
\]

for all \( \beta_0, \beta_1, \ldots, \beta_i \in \mathbb{N}_0^d \). Then (6.3) gives

\[
|c_{\alpha, \beta_1, \ldots, \beta_i}| \leq M_i k_{\alpha} \exp \left( - \sum_{j=0}^i 2\pi \delta_{kj} |\beta_j| \right) (6.6)
\]

for all \( \alpha \in \mathbb{N}_0^d \) and \( \beta_1, \ldots, \beta_i \in \mathbb{Z}^d \).

Now we set \( \delta_0 = \min\{\delta_{00}, \delta_{11}, \ldots, \delta_{ii}\} \) and verify the claimed bounds for \( c = (2\pi \delta_0)^{-1}, p = e \log e^{-1} \) and a suitable positive constant \( C \). Let \( \mathcal{J}_0 = \{0, 1, \ldots, p - 1\}, \mathcal{J}_d = \{0, \pm 1, \ldots, \pm (p - 1)\} \) and \( \mathcal{J}_1 = \mathbb{Z} \setminus \mathcal{J}_0 \). Using the product index sets \( \mathcal{I}_\mu = \mathcal{I}_{\mu_1} \times \cdots \times \mathcal{I}_{\mu_d} \) and \( \mathcal{J}_\mu = \mathcal{J}_{\mu_1} \times \cdots \times \mathcal{J}_{\mu_d} \) with \( \mu \in \{0, 1\}^d \), we can recast the expansion (6.1) in \( L^2_{\omega \odot \delta \odot \delta_0 d}(D \times Y_i) \) as follows:

\[
w = \sum_{m=0}^{(i+1)d} \sum_{\mu, \nu, \iota \in \{0, 1\}^d: |\mu| + \sum_{j=1}^i |\nu_j| = m} \sum_{\alpha \in \mathcal{I}_\mu} \sum_{\beta \in \mathcal{J}_\nu} \sum_{\beta \in \mathcal{J}_\iota} c_{\alpha, \beta_1, \ldots, \beta_i} \left( \bigotimes_{k=1}^d \hat{T}_{j \alpha_k} \right) \otimes \left( \bigotimes_{j=1}^i \bigotimes_{k=1}^d \hat{T}_{j \beta_k} \right). (6.7)
\]
In the right-hand side of (6.7), the term of the outer sum corresponding to \( m = 0 \) is \( \Pi_{i,p} w \), and the remainder can be bounded using (6.6):

\[
\|w - \Pi_{i,p} w\|_{L^\infty(D \times Y_t)} \leq \sum_{m=1}^{(i+1)d} \sum_{\substack{\mu, \nu_i, \ldots, \nu_i \in \{0,1\}^d; \\ |\mu| + \sum_{j=1}^d \nu_j = m}} \sum_{\substack{\alpha \in S_p \setminus \tilde{S}_p; \\ \beta_i \in \tilde{S}_v_i}} |c_{\alpha, \beta_1, \ldots, \beta_i}| \leq M_i 2^{d-1} 2^{i+1} \sum_{m=1}^{(i+1)d} \epsilon^m \left( \frac{(i+1)d}{m} \right) \leq C_0 \epsilon, \quad (6.8)
\]

where \( \lambda = e^{-2\pi \delta} \in (0,1) \) and \( C_0 = M_i (i+1) d^2 \sum_{m=1}^{(i+1)d} \epsilon^m \left( \frac{(i+1)d}{m} \right) \geq 0 \). This gives the first of the bounds (4.12) with any constant \( C \geq C_0 \), selected independently of \( \epsilon \).

For derivatives of the shifted Chebyshev polynomials and exponentials, we have \( \|\tilde{T}_\alpha\|_{L^\infty(0,1)} = 2\alpha^2 \) for all \( \alpha \in \mathbb{N}_0 \) and \( \|\tilde{T}_\beta\|_{L^\infty(0,1)} = 2\pi |\beta| \) for all \( \beta \in \mathbb{Z} \). Note that there exist positive constants \( \gamma_1 \) and \( \gamma_2 \) such that \( \sum_{\beta=\delta}^\infty \beta \lambda^\beta \leq \gamma_1 (1-\lambda)^{-1} r \sum_{\beta=\delta}^\infty \lambda^\beta \) and \( \sum_{\beta=\delta}^\infty \beta^2 \lambda^\beta \leq \gamma_2 (1-\lambda)^{-2} r \sum_{\beta=\delta}^\infty \lambda^\beta \) for any \( r \in \mathbb{N}_0 \). Using this, we obtain, as in (6.8), the following inequalities:

\[
\|\partial_k (w - \Pi_{i,p} w)\|_{L^\infty(D \times Y_t)} \leq \frac{2\gamma_2 C_0 \epsilon^p \lambda^2}{(1-\lambda)^2}, \quad \|\partial_j k (w - \Pi_{i,p} w)\|_{L^\infty(D \times Y_t)} \leq \frac{2\gamma_1 C_0 \epsilon^p}{1-\lambda}.
\]

for all \( k \in \{1, \ldots, d\} \) and \( j \in \{1, \ldots, i\} \). This proves the last two of the bounds (4.12) with a suitable positive constant \( C \), which can be chosen independently of \( \epsilon \in (0, \epsilon_0) \).

\[ \square \]

6.1. Proof of Lemma 4.5

Proof. Let \( L \in \mathbb{N} \). Using the triangle inequality, we bound the errors as follows:

\[
\|\partial_k (w - \Pi_{i,p}^L w)\|_{\infty} \leq \|\partial_k (id - \Pi_{i,p}) w\|_{\infty} + \|\partial_k (id - \Pi_{i,p}^L) \Pi_{i,p} w\|_{\infty}, \\
\|\partial_i k w - \Pi_{i,p}^L \partial_i k \Pi_{i,p} w\|_{\infty} \leq \|\partial_i (id - \Pi_{i,p}) w\|_{\infty} + \|(id - \Pi_{i,p}^L) \partial_i k \Pi_{i,p} w\|_{\infty}
\]

(6.9)

for every \( k \in \{1, \ldots, d\} \). By Lemma 4.4 there exist positive constants \( C_0 \) and \( c \) such that, for \( p = [cL] \), we have

\[
\|\partial_k (id - \Pi_{i,p}) w\|_{\infty} \leq C_0 p^2 2^{-L}, \quad \|\partial_j k (id - \Pi_{i,p}) w\|_{\infty} \leq C_0 p 2^{-L}.
\]

(6.10)

Certain derivatives of \( \Pi_{i,p} w \) can be bounded in terms of first-order derivatives of \( \Pi_{i,p} w \) using the Bernstein’s inequality for trigonometric polynomials. Applying it together with the bounds (6.10) and Lemma 4.3 we obtain

\[
\|\partial_k (id - \Pi_{i,p}^L) \Pi_{i,p} w\|_{\infty} \leq 2^{-L} \sum_{k'=1}^d \|\partial_k \partial_{k'} \Pi_{i,p} w\|_{\infty} \\
+ 2^{-L} \sum_{j'=1}^{i-1} \sum_{k'=1}^d \|\partial_{j'} k' \partial_{k'} \Pi_{i,p} w\|_{\infty} + 2^{-L} \sum_{k'=1}^d \|\partial_{k'} \partial_{k'} \Pi_{i,p} w\|_{\infty} \\
\leq 2^{-L} \sum_{k'=1}^d 2\pi p \|\partial_{k'} \Pi_{i,p} w\|_{\infty} + 2^{-L} \sum_{j'=1}^{i-1} \sum_{k'=1}^d 2\pi p \|\partial_{j'} k' \Pi_{i,p} w\|_{\infty} \\
+ 2^{-L} \sum_{k'=1}^d (2\pi p)^2 \|\partial_{k'} \Pi_{i,p} w\|_{\infty} \leq 2^{-L} \sum_{k'=1}^d 2\pi p \left\{ \|\partial_{k'} w\|_{\infty} + C_0 p^2 2^{-L} \right\} \\
+ 2^{-L} \sum_{j'=1}^i (2\pi p)^2 \left\{ \|\partial_{j'} k' w\|_{\infty} + C_0 p^2 2^{-L} \right\} \leq C_1 p^2 2^{-L} \quad (6.11)
\]
for every \( k \in \{1, \ldots, d\} \) with a positive constant \( C_1 \) independent of \( L \). The same approach leads to the bound

\[
\| (\text{id} - \tilde{P}_L^k) \partial_{ik} \Pi_{i,p} w \|_\infty \leq 2^{-L} \sum_{k'=1}^d \| \partial_{ik'} \Pi_{i,p} w \|_\infty + 2^{-L} \sum_{j'=1}^i \sum_{k'=1}^d \| \partial_{j'k'} \Pi_{i,p} w \|_\infty
\]

\[
\leq 2^{-L} \sum_{k'=1}^d 2p \| \partial_{ik'} \Pi_{i,p} w \|_\infty + 2^{-L} \sum_{j'=1}^i \sum_{k'=1}^d 2p \| \partial_{j'k'} \Pi_{i,p} w \|_\infty
\]

\[
\leq 2^{-L} \sum_{k'=1}^d 2p \left\{ \| \partial_{ik'} w \|_\infty + C_0 p^2 2^{-L} \right\} + 2^{-L} \sum_{j'=1}^i \sum_{k'=1}^d 2p \left\{ \| \partial_{j'k'} w \|_\infty + C_0 p 2^{-L} \right\}
\]

\[
\leq C_2 p^2 2^{-L} \quad (6.12)
\]

for every \( k \in \{1, \ldots, d\} \) with a positive constant \( C_2 \) independent of \( L \).

Combining inequalities (6.11) and (6.12) with (6.10) and (6.9), we obtain the claimed error bounds with \( C = C_0 + \max\{C_1, C_2\} \). \( \square \)

**6.2. Auxiliary results for Lemma 4.11**

Lemma 4.11 is based on the following auxiliary Lemmas 6.1 and 6.2. We formulate these lemmas in terms of intermediate, starred finite element spaces with corresponding analysis operators and low-rank subspaces, which reflect the iterative averaging of all the \( n \) microscales, as defined in Definition 3.4.

First, similarly to as in (4.33) using the functions \( \lambda_0, \ldots, \lambda_n \) of \( \varepsilon \) from Assumption 4.1, we define the space

\[
\tilde{V}_i^L = \tilde{V}_i^{\lambda_i + L} \otimes \bigotimes_{j=i+1}^n \tilde{V}^L \quad (6.13)
\]

for every \( i \in \{0, 1, \ldots, n\} \), so that \( \tilde{V}_0^L = \tilde{V}_n^L \) and \( \tilde{V}_i^L = \tilde{V}_0^{\lambda_i + L} \). Further, as in (4.15), we define an analysis operator \( \tilde{\Psi}_i^L : L^2(D \times Y_i) \to C_{2d(\lambda_i + L) + \nu} \) by setting

\[
\tilde{\Psi}_i^L = \tilde{\Psi}_i^{\lambda_i + L} \otimes \bigotimes_{j=i+1}^n \tilde{\Psi}^L \quad (6.14)
\]

for every \( i \in \{0, 1, \ldots, n\} \). Then \( \tilde{\Psi}_i^L \) and \( \tilde{\Psi}_i^L \) are identical to \( \Psi_i^L \) and \( \Psi_i^L \) respectively.

Note that the starred finite-element spaces are introduced in (6.13) so as to ensure that averaging an element of each of these spaces (except the last) produces an element from the next space. Indeed, the following embedding property follows from Definition 3.4 and equality (6.13).

**Lemma 6.1.** For all \( L \in \mathbb{N} \) and \( i \in \{1, \ldots, n\} \), we have \( \mathcal{U}_i^L \tilde{V}_{i-1}^L \subset \tilde{V}_i^L \).

In order to analyze how the structure of functions from \( \tilde{V}_i^L \) with coefficients from \( \mathcal{Q}_i^L \) is transformed under averaging, we define, for every \( i \in \{0, 1, \ldots, n\} \),

\[
\mathcal{Q}_{i+1}^L = \mathcal{Q}_i^L \otimes \bigotimes_{j=i+1}^n \mathcal{P}_{(n+1-j)p_L} \subset C_{2d(\lambda_i + L) + \nu} \quad (6.15)
\]

In particular, the so defined subspaces \( \mathcal{Q}_0^L \) and \( \mathcal{Q}_n^L \) coincide with \( \mathcal{Q}_L^L \) and \( \mathcal{S}_L \), given by (4.19) and (6.15) respectively. These intermediate subspaces satisfy the following relation.

**Lemma 6.2.** For all \( L \in \mathbb{N} \), \( i \in \{1, \ldots, n\} \) and \( v \in \tilde{V}_{i-1}^L \) such that \( \tilde{\Psi}_{i-1}^L v \in \mathcal{Q}_{i-1}^L \), we have \( \tilde{\Psi}_{i-1}^L \mathcal{U}_i^L v \in \mathcal{Q}_i^L \).

**Proof.** Let us consider a function \( v \in W_i^L \) such that \( \tilde{\Psi}_{i-1}^L v = \kappa \otimes \mu \otimes \nu \) with

\[
\kappa \in \mathcal{S}_{i-1}^L, \quad \mu \in \mathcal{P}_{(n+1-i)p_L}, \quad \nu \in \bigotimes_{j=i+1}^n \mathcal{P}_{(n+1-j)p_L}
\]
and show that \( \tilde{\psi}_{\varepsilon} L \mathcal{U}_L^v \in \Omega_{\varepsilon}^L \). Due to the linearity and tensor-product structure of \( \delta_{L-1}^L \), defined by \( (6.15) \), this is sufficient to verify the claim.

Applying Definition [3.4] and Lemma [6.1] to \( v \), we obtain

\[
(\tilde{\psi}_{\varepsilon} L \mathcal{U}_L^v)(j-1)2^L + j_1, j_1+1, \ldots, j_n = (\mathcal{U}_L^v) \left( \frac{j - 1}{2^\lambda_i} + \frac{j_1 - \frac{1}{2}}{2^\lambda_i+L}, \frac{j_1+1 - \frac{1}{2}}{2^\lambda_i+L}, \ldots, \frac{j_n - \frac{1}{2}}{2^\lambda_i+L} \right)
\]

\[
= \tilde{\kappa} j, \mu, v_{j_1+1, \ldots, j_n} \tag{6.16}
\]

for all \( j \in \mathcal{J}^{\lambda_i, d} \) and \( j_1, j_1+1, \ldots, j_n \in \mathcal{J}^{L, d} \), where \( \tilde{\kappa} = M_L^L \kappa \) is the coefficient tensor of the component of \( v \) with respect to the first variable averaged over scale \( \varepsilon_i \):

\[
\tilde{\kappa}_j = 2^{-d(L+\lambda_i-1) - \lambda_i} \sum_{j' \in \mathcal{J}^{L+\lambda_i-1, -\lambda_i, d}} \kappa_{(j-1)2^L+\lambda_i-1, -\lambda_i+j'}
\]

for every \( j \in \mathcal{J}^{\lambda_i, d} \). With this notation, relation \( (6.16) \) implies

\[
\tilde{\psi}_{\varepsilon} L \mathcal{U}_L^v = (M_L^L \kappa) \otimes \mu \otimes \nu.
\]

Since, according to \( (4.21) \), \( (M_L^L \kappa) \otimes \mu \in \delta_{L}^L \), the claimed inclusion \( \tilde{\psi}_{\varepsilon} L \mathcal{U}_L^v \in \Omega_{\varepsilon}^L \) follows immediately from \( (6.15) \). \( \square \)

REFERENCES

[1] Assyr Abdulle, Weinan E, Björn Engquist, and Eric Vanden-Eijnden. The heterogeneous multiscale method. Acta Numer., 21:1–87, 2012.
[2] G. Allaire and M. Briane. Multiscale convergence and reiterated homogenisation. Proceedings of the Royal Society of Edinburgh: Section A. Mathematics, 126(2):297–342, 1996.
[3] Grégoire Allaire. Homogenization and two-scale convergence. SIAM J. Math. Anal., 23(6):1482–1518, 1992.
[4] Grégoire Allaire. Homogenization and two-scale convergence. SIAM Journal on Mathematical Analysis, 23(6):1482–1518, 1992.
[5] I. Babuška and B. Guo. The h-p version of the finite element method for domains with curved boundaries. SIAM Journal on Numerical Analysis, 25(4):837–861, 1988.
[6] I. Babuška and B. Guo. Regularity of the solution of elliptic problems with piecewise analytic data. Part I. boundary value problems for linear elliptic equation of second order. SIAM Journal on Mathematical Analysis, 19(1):172–203, 1988.
[7] I. Babuška and B. Guo. Regularity of the solution of elliptic problems with piecewise analytic data. II: The trace spaces and application to the boundary value problems with nonhomogeneous boundary conditions. SIAM Journal on Mathematical Analysis, 20(4):763–781, 1989.
[8] Markus Bachmayr and Vladimir Kazeev. Stability and preconditioning of elliptic PDEs with low-rank multilevel structure. Foundations of Computational Mathematics, 2020.
[9] N. Bakhvalov and G. Panasenko. Homogenisation: Averaging Processes in Periodic Media, volume 36 of Mathematics and its Applications. Springer Netherlands, 1989.
[10] Alain Bensoussan, Jacques-Louis Lions, and George Papanicolaou. Asymptotic analysis for periodic structures, volume 5 of Studies in Mathematics and its Applications. North-Holland Publishing Co., Amsterdam-New York, 1978.
[11] Philippe G. Ciarlet. The finite element method for elliptic problems. North-Holland, Amsterdam, New York, Oxford, 1978.
[12] D. Cioranescu, A. Dalibard, A. Donato, G. Griso, and R. Zaki. The periodic unfolding method in domains with holes. SIAM J. Math. Anal., 44(2):718–760, 2012.
[13] Doina Cioranescu, Alain Dalibard, and Georges Griso. The periodic unfolding method in homogenization. SIAM Journal on Mathematical Analysis, 40(4):1585–1620, 2008.
[14] B. Engquist and P. E. Souganidis. Asymptotic and numerical homogenization. Acta Numerica, 17:147–190, 2008.
[15] Lars Grasedyck. Polynomial approximation in Hierarchical Tucker format by vector-tensorization. Preprint 308, Institut für Geometrie und Praktische Mathematik, RWTH Aachen, April 2010.
[16] Helmut Harbrecht and Christoph Schwab. Sparse tensor finite elements for elliptic multiple scale problems. Comput. Methods Appl. Mech. Engrg., 200(45-46):3100–3110, 2011.
[17] Viet Ha Hoang. Sparse finite element method for periodic multiscale nonlinear monotone problems. Multiscale Model. Simul., 7(3):1042–1072, 2008.
[18] Viet Ha Hoang and Christoph Schwab. High-dimensional finite elements for elliptic problems with multiple scales. Multiscale Model. Simul., 3(1):168–194, 2004/05.
