ESAIM: M2AN 53 (2019) 1507–1552
https://doi.org/10.1051/m2an/2019016

ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

MULTILEVEL QUASI-MONTE CARLO INTEGRATION WITH PRODUCT WEIGHTS FOR ELLIPTIC PDES WITH LOGNORMAL COEFFICIENTS ⋆

LUKAS HERRMANN[1,*] AND CHRISTOPH SCHWAB[1]

Abstract. We analyze the convergence rate of a multilevel quasi-Monte Carlo (MLQMC) Finite Element Method (FEM) for a scalar diffusion equation with log-Gaussian, isotropic coefficients in a bounded, polytopal domain $D \subset \mathbb{R}^d$. The multilevel algorithm Q^*_L which we analyze here was first proposed, in the case of parametric PDEs with sequences of independent, uniformly distributed parameters in Kuo *et al.* (*Found. Comput. Math.* **15** (2015) 411–449). The random coefficient is assumed to admit a representation with *locally supported coefficient functions, as arise for example in spline- or multiresolution representations of the input random field.* The present analysis builds on and generalizes our single-level analysis in Herrmann and Schwab (*Numer. Math.* **141** (2019) 63–102). It also extends the MLQMC error analysis in Kuo *et al.* (*Math. Comput.* **86** (2017) 2827–2860), to locally supported basis functions in the representation of the Gaussian random field (GRF) in D, and to product weights in QMC integration. In particular, in polytopal domains $D \subset \mathbb{R}^d$, $d = 2, 3$, our analysis is based on weighted function spaces to describe solution regularity with respect to the spatial coordinates. These spaces allow GRFs and PDE solutions whose realizations become singular at edges and vertices of D. This allows for *non-stationary* GRFs whose covariance operators and associated precision operator are fractional powers of elliptic differential operators in D with boundary conditions on ∂D. In the weighted function spaces in D, first order, Lagrangian Finite Elements on regular, locally refined, simplicial triangulations of D yield optimal asymptotic convergence rates. Comparison of the ε-complexity for a class of Matérn-like GRF inputs indicates, for input GRFs with low sample regularity, superior performance of the present MLQMC-FEM with locally supported representation functions over alternative representations, *e.g.* of Karhunen–Loève type. Our analysis yields general bounds for the ε-complexity of the MLQMC algorithm, uniformly with respect to the dimension of the parameter space.

Mathematics Subject Classification. 65D30, 65N30, 60G60.

Received April 30, 2017. Accepted February 18, 2019.

1. INTRODUCTION

The numerical analysis of solution methods for partial differential equations (PDEs) and more general operator equations with random input data has received increasing attention in recent years, in particular with the development of computational uncertainty quantification and computational science and engineering. There,

Keywords and phrases. Quasi-Monte Carlo methods, multilevel quasi-Monte Carlo, uncertainty quantification, error estimates, high-dimensional quadrature, elliptic partial differential equations with lognormal input.

⋆ *This work was supported in part by the Swiss National Science Foundation under grant SNF 159940. The authors acknowledge the computational resources provided by the EULER cluster of ETH Zürich.*

[1] Seminar for Applied Mathematics, ETH Zürich, 8092 Zürich, Switzerland.
[*]Corresponding author: `lukas.herrmann@sam.math.ethz.ch`

 © EDP Sciences, SMAI 2019

particular models of randomness in the PDEs' input entail particular requirements to efficient computational uncertainty quantification algorithms. A basic case arises when there are a finite (possibly large) number s of random variables whose densities have bounded support and which parametrize uncertain input from function spaces, such as diffusion coefficients or source terms in the forward PDE model: computation of statistical moments of "responses" being (functionals of) solution families of these PDEs. Numerical Bayesian inversion then amounts to numerical integration over a parameter domain of finite parameter space dimension s, which itself is a discretization parameter. Statistical independence and scaling reduces this task to numerical integration over the unit cube $[0,1]^s$, against a product probability measure. In the context of PDEs, so-called *distributed random inputs* such as spatially heterogeneous diffusion coefficients, uncertain physical domains, etc. imply, via *uncertainty parametrizations* (such as Fourier-, Karhunen–Loève , B-spline or wavelet expansions) in physical domains D, a countably-infinite number of random parameters (being, for example, Fourier- or wavelet coefficients). This, in turn, renders the problem of numerical estimation of response statistics of PDE solutions a problem of infinite-dimensional numerical integration. Assuming statistical independence of the system of (countably many) random input parameters results in the problem of numerical integration against a product probability measure. The case of the uncertain PDE input being a Gaussian random field (GRF) is particularly important in applications, and the numerical analysis has received considerable attention in recent years. Here, the numerical estimation of statistical moments of PDE solutions amounts to integrating parametric PDE solutions against Gaussian measures on function spaces of admissible input data. Adopting uncertainty parametrizations of the input GRFs renders the domain Ω of integration a countable product of real lines $\mathbb{R}^{\mathbb{N}}$, endowed with the Gaussian product measure (GM) μ and with the product sigma algebra obtained by completing the finite dimensional cylinders of Borel sets on $\mathbb{R}$ (we refer to [10] for details on GMs on $\mathbb{R}^{\mathbb{N}}$).

Here, as in [26, 36] and the references there, we analyze the combined discretization by quasi-Monte Carlo (QMC) quadratures and by the Finite Element Method (FEM) of linear, second order elliptic PDEs in a bounded, polytopal domain $D \subset \mathbb{R}^d$, $d = 2, 3$. Unlike the applications in [26, 36] and the references there, and in [28], where stationarity enters the algorithms and the error analysis in an essential way, here we consider isotropic (*i.e.* scalar), log-Gaussian diffusion coefficient $a = \exp(Z)$, where Z is a *possibly non-stationary* GRF in D.

We place the present work in perspective with other recent work on the numerical analysis of PDEs with GRF inputs. In [26,36], an error analysis of single- and multilevel algorithms was developed for Karhunen–Loève type representations of the GRF Z. Except for rather special settings where Karhunen–Loève eigenfunctions are explicitly known (when D is a torus or a sphere [32]), Karhunen–Loève type representations of GRFs are not explicitly available but must be computed numerically. This entails the accurate numerical approximation of a large number of eigenpairs of the covariance operator of the GRF Z in the domain D, a significant computational overhead.

Moreover, the covariance eigenfunctions in Karhunen–Loève representations of GRFs in domains or manifolds D typically have *global support* in D. This was shown in [26, 36] to imply in the error analysis of QMC quadrature rules so-called *product-and-order dependent* (POD) weights. Constructing QMC points with POD weights introduces, via the corresponding fast component-by-component (CBC) algorithm, a *quadratic scaling* w.r. to the QMC integration dimension s of the construction cost for QMC rules, see [14, 41] and the references there. For this reason, the QMC construction cost is not explicitly accounted for in recent complexity estimates of QMC-Finite Element (FE) algorithms for PDEs.

To bypass the need for numerical Karhunen–Loève eigenfunction computation, under (strong) assumptions on stationarity of the GRF Z, fast Fourier transform (FFT)-based numerical methods have been proposed for efficient numerical realization of stationary GRFs. We refer to [16] and the references there for details. FFT based techniques have recently been used in conjunction with QMC and FEM for elliptic PDEs with coefficients given by (exponentials of) stationary GRFs Z in [28]. While allowing essentially linear scaling w.r. to the number of FE degrees of freedom in the domain D, *stationarity of the GRF Z* is a key condition for the applicability of FFT-based, so-called *"circulant embedding"* methods. Being essentially Fourier-based techniques, the QMC-FE error analysis in [28] involves QMC weight sequences with POD structure and, hence, quadratic w.r. to QMC integration dimension s scaling of the cost for QMC rule construction via the fast CBC construction (see [28], Eq. (3.16), Rem. 9).

In recent years, computational modeling of noisy spatial data has increasingly employed *non-stationary GRFs* in bounded domains D. We mention recently used random field models in spatial statistics (see [18, 38] and the references there), GRFs on manifolds such as the sphere (see [32] and the references there), and *deep Gaussian processes* (see [17] and the references there). As proposed in [38], rather general non-stationary GRFs Z in bounded domains or on manifolds D can be modeled and sampled as solutions of *stochastic (integro) PDEs* (SPDEs). A widely used equation which generalizes the classical Whittle–Matérn [39, 48] covariances of stationary Gaussian random fields reads as

$$(-\nabla \cdot (A(x)\nabla) + \kappa^2(x))^{\alpha/2} Z = \mathcal{W} \quad \text{in } D. \tag{1.1}$$

Here $\mathcal{W}$ denotes spatial white noise on D, and $\alpha > 0$ is suitably chosen. If $D = \mathbb{R}^d$, $A(x) \equiv \mathrm{Id}$, and $\kappa(x) \equiv$ const, the solution Z to (1.1) (we assume Z to be centered throughout this paper) is stationary with so-called *Matérn-type covariance*, *cf.* [47, 48]. For a variable coefficient matrix $A(x)$ and variable $\kappa(x)$, or in bounded domains D with homogeneous Dirichlet or Neumann boundary conditions, equation (1.1) results in non-stationary, "Matérn-like" Gaussian random fields. On bounded domains D, boundary conditions for Z are mandatory for the unique solvability of the SPDE (1.1). Imposition of boundary conditions on ∂D generally entails non-stationarity of the GRF Z, *cf.* Section A.4 of [38]. Then, FFT-based methods are generally not available, and computation of Karhunen–Loève eigenbases for (1.1) will entail, again, prohibitive cost. Alternative, *covariance independent* representations of GRFs via multiresolution systems in D allow us to circumvent the numerical solution of Karhunen–Loève eigenproblems, the classical example being the Brownian bridge in $D = (0, 1)$, going back to P. Lévy and Z. Ciesielski. The basis functions in corresponding representation systems are well-localized in D (either compactly supported or exponentially decaying) and allow for fast evaluation of the GRF in D, similar to FFTs. While retaining linear scaling w.r. to the spatial resolution of the approximate GRF in D, the hierarchical nature of multiresolution analyses (MRAs) naturally enables *multilevel QMC (MLQMC) algorithms* with a discretization level dependent resolution of GRF and QMC integration. In addition, as observed by us recently in [20, 31], the localization of the supports of the representation system in D allows us to use QMC quadrature with *product weights*. This, in turn, is known to afford *linear scaling* of the work with respect to the parameter dimension s to compute the QMC generating vectors (see [14, 42] and the references there). To provide a complete error *vs.* work analysis of a MLQMC-FE algorithm for the numerical solution of a linear, second order elliptic PDE with GRF input and locally supported basis functions in a bounded, polytopal domain $D \subset \mathbb{R}^d$, $d = 2, 3$, where the GRF satisfies (1.1) with suitable boundary conditions on ∂D is the purpose of the present paper. Recently, independent of the present work, in [33] a combined QMC and wavelet-based discretization of log-Gaussian random fields was proposed and error bounds were presented. The present results go in several respects beyond those in [33]. We consider, in particular, MLQMC-FE discretizations, and use sharper bounds than those in [33] on the error caused by truncating the expansion of the GRF, from our single-level analysis in [31]. We also generalize, based on [31], the QMC error analysis by admitting Gaussian type weight functions in the anisotropic QMC norms, as opposed to the exponential weights used in [26, 36]. We prove that this extends the summability range of Karhunen–Loève and wavelet expansions of the admissible GRFs for the applicability of QMC with product weights, in addition to obviating stationarity, as compared to [26, 36]. Furthermore, in the present paper, we provide a full regularity analysis of the PDE as required for MLQMC-FEM. As is well-known, this requires a form of "mixed regularity" analysis, with possibly sharp, quantitative bounds of the sensitivities of the parametric integrand functions with respect to the coordinates in the GRF Z, in weighted $H^2(D)$ norms. In the present paper, we also develop these norm bounds.

We confine QMC integration error analysis to first order, randomly shifted lattice rules proposed originally in [34], and to continuous, piecewise linear "Courant" FEM in D. We adopt the setting of our analysis [31] of the single-level QMC-FE algorithm: in a bounded, polytopal domain $D \subset \mathbb{R}^d$, $d = 2, 3$, we consider a model

Dirichlet problem

$$-\nabla \cdot (a\nabla u) = f, \quad u\Big|_{\partial D} = 0. \tag{1.2}$$

As in [31], we assume that the GRF $Z = \log(a) : \Omega \to L^\infty(D)$ is (formally) represented as

$$Z := \sum_{j\geq 1} y_j \psi_j, \tag{1.3}$$

where $(\psi_j)_{j\geq 1}$ is a sequence of real-valued, bounded, and measurable functions in D. In particular, with respect to the GM μ, the terms in the sequence $\boldsymbol{y} = (y_j)_{j\geq 1}$ in (1.3) are i.i.d standard normal, $y_j \sim \mathcal{N}(0,1)$, i.i.d. for $j \in \mathbb{N}$. The lognormal coefficient a in (1.2) is given by

$$a := \exp(Z). \tag{1.4}$$

Here, (1.3) converges in $L^q(\Omega; L^\infty(D))$, $q \in [1,\infty)$, under the assumption that there exists a positive sequence $(b_j)_{j\geq 1} \in \ell^p(\mathbb{N})$ for some $p \in (0,\infty)$ such that

$$K_0 := \left\| \sum_{j\geq 1} \frac{|\psi_j|}{b_j} \right\|_{L^\infty(D)} < \infty. \tag{\textbf{A1}}$$

In the setting of (1.3) and (1.4), the expectation with respect to the GM μ of the solution to (1.2) can be computed with QMC by randomly shifted lattice rules and product weights with dimension-independent convergence rates under the assumption (**A1**) with $p < 2$, *cf.* [31]. The assumption in (**A1**) can account for locality in the support of the functions ψ_j. This may also be achieved by exponentially decaying ψ_j, which are not compactly supported, see [8]. An assumption of the type of (**A1**) in the case of so called affine-parametric coefficients in conjunction with the application of QMC with product weights was already discussed by us in [20]. In the present work, we extend our analysis of [31] to a MLQMC-FE algorithm with log-Gaussian inputs to reduce the overall work. The perspective of MLQMC integration with product weights for random inputs represented by ψ_j with localized supports was originally introduced in [19] for the case of so-called affine-parametric coefficients. Multilevel QMC for elliptic PDEs with affine coefficients was first introduced in [35] (there for globally supported Karhunen–Loève eigenfunctions and with POD weights). As we showed there, localization of supports allows to obtain in certain cases estimates for the work of the evaluation of the MLQMC quadrature, which are asymptotically equal to the work to solve one instance of the corresponding deterministic PDE with the same error tolerance also in the case that the FE convergence rate is higher than 1/2 with respect to the number of FE degrees of freedom. The convergence rate of first order FEM is higher than $1/d$ if, for example, the spatial error is considered in a weaker Sobolev norm for $d = 2, 3$. In contrast to [36], the present complexity analysis does account for the cost of the CBC algorithm of [42, 43] for the computation of the generating vector of the QMC points. This is due to product weights affording CBC construction cost with linear scaling in terms of the dimension s of the integration domain. The work estimates of MLQMC quadrature obtained here are compared to previous results from [36] for the same underlying GRF. For locally supported representations of the random field inputs, MLQMC with product weights requires asymptotically less work to obtain a certain accuracy than global supports and POD weights. To facilitate comparison with [36] (where a Karhunen–Loève expansion is truncated to a fixed number of terms on all FE mesh levels and analyzed assuming the cost to evaluate one Karhunen–Loève function is $\mathcal{O}(1)$), in the present paper the error analysis from Section 5 of [36] is sharpened in Appendices A and B. There, novel parametric regularity and error *vs.* work estimates are proved that also cover variable Karhunen–Loève truncation dimensions and fast (*e.g.* FFT) methods to sample the GRF in the case of globally supported representation systems with QMC POD weights accounting for the cost of the CBC algorithm. As a byproduct, we show that MLQMC with global supports and POD weights requires in certain cases asymptotically the same work as the corresponding deterministic PDE, which constitutes an extension of the theory in [36] on MLQMC with global supports and POD weights.

The outline of this paper is as follows. In Section 2, we recapitulate known results on the well-posedness of problem (1.2)–(1.4) under assumption (**A1**), and on the integrability of random solution with respect to the GM. We also present bounds on the error incurred in the random solution when the expansion (1.3) is truncated to a finite number of s terms. As we combine QMC quadrature approximation of the GM with continuous, piecewise linear FE discretization of (1.2) of the random solution in polytopal domains $D \subset \mathbb{R}^d$, $d = 2, 3$, we also review in Section 2 elements of elliptic regularity theory and FE approximation theory in D; notably, handling corner and (in space dimension $d = 3$) edge singularities induced by D we review weighted Sobolev spaces in D in which (1.2) admits a full regularity shift. Corresponding weighted spaces also appear in our convergence rate analysis of the expansion (1.3) of the GRF. In Section 3, we review QMC convergence theory from [31,41]. Suitable (weighted) spaces on $\mathbb{R}^s$ of integrand functions with mixed first derivatives which ensure (nearly) first order convergence with dimension-independent constants are introduced. Section 4 presents the key mathematical results: parametric regularity analysis for the integrand functions which arise from the dimensionally truncated, FE discretized problem, generalizing the single-level QMC analysis in [31] by admitting locally supported functions ψ_j in the representation (1.3) of the GRF; while similar in spirit to the multilevel analysis in [26], there are significant technical differences due to accounting for local supports of ψ_j, analogous to the recent polynomial chaos N-term approximation rate analysis in [7]. The error bounds are then combined in Section 5 to a novel, MLQMC convergence rate bound in terms of the (sequences of) truncation dimensions $(s_\ell)_{\ell=0,\dots,L}$, numbers $(M_\ell)_{\ell\geq 0}$ of FE degrees of freedom and of QMC sample numbers $(N_\ell)_{\ell=0,\dots,L}$, where L denotes the number of discretization levels. Judicious choices of these parameters for concrete MLQMC-FE algorithms are derived in Section 6.1 by the "usual" error *vs.* work analysis through optimization, of the error bounds in Section 5, derived analogously to [35,36]. Several cases of these error *vs.* work estimates are discussed in Section 6.2 and compared to error *vs.* work estimates for multilevel QMC with global supports and POD weights in Section 6.3. Numerical experiments of this MLQMC algorithm for non-stationary GRF input represented by a multiresolution function system are presented in the univariate case in Section 7.

2. Well-posedness and spatial approximation

2.1. Well-posedness

We consider the variational formulation of the PDE (1.2) with lognormal coefficient $a = \exp(Z)$, *i.e.*, to find $u : \Omega \to V$ such that

$$\int_D a\nabla u \cdot \nabla v \mathrm{d}x = f(v), \quad v \in V, \tag{2.1}$$

where $V := H_0^1(D)$ with dual space V^*. Throughout, we identify $L^2(D)$ with its dual space $L^2(D)^*$, *i.e.*, $V \subset L^2(D) \simeq L^2(D)^* \subset V^*$. Under the assumption that for some $p_0 \in (0, \infty)$, $(b_j)_{j\geq 1} \in \ell^{p_0}(\mathbb{N})$ it holds that $Z \in L^q(\Omega; L^\infty(D))$ for every $q \in [1, \infty)$, *cf.* Theorem 2 of [31]. Hence, $0 < \operatorname{ess\,inf}_{x\in D}\{a(x)\} \leq \|a\|_{L^\infty(D)} < \infty$, μ-a.s.. As in previous works [26,31,36], in the ensuing error analysis, the quantities

$$a_{\min} := \operatorname*{ess\,inf}_{x\in D}\{a(x)\} \quad \text{and} \quad a_{\max} := \|a\|_{L^\infty(D)}$$

will play an important role. Under assumption (**A1**), $a_{\min}$ and $a_{\max}$ are random variables on the probability space $(\Omega, \bigotimes_{j\geq 1} \mathcal{B}(\mathbb{R}), \mu)$ (see *e.g.*, [10], Example 2.3.5). Therefore, continuity and coercivity of the random bilinear form $(w, v) \mapsto \int_D a\nabla w \cdot \nabla v \mathrm{d}x$ in (2.1) on $V \times V$ holds with coercivity constant $a_{\min}$ and continuity constant $a_{\max}$, μ-a.s. By the Lax–Milgram Lemma, a unique solution u to (2.1) exists μ-a.s. By Proposition 3 of [31] (see also [12]), for every $q \in [1, \infty)$,

$$\|u\|_{L^q(\Omega;V)} \leq \|1/a_{\min}\|_{L^q(\Omega)}\|f\|_{V^*} < \infty,$$

where the strong measurability of $u : \Omega \to V$ follows, since the V-valued random solution u of (2.1) depends continuously on the $L^\infty(D)$-valued coefficent a (via a Strang type argument).

Numerical approximation of (functionals of) the random solution by QMC quadratures requires a finite dimensional domain of integration. To this end, the expansion of the GRF Z in (1.3) is truncated to a finite number s of terms: the *s-term truncated lognormal random field* a^s is defined by $a^s := \exp(Z^s) = \exp(\sum_{j=1}^{s} y_j\psi_j)$, for every $s \in \mathbb{N}$. With a^s, we associate the random variables

$$a^s_{\min} := \operatorname*{ess\,inf}_{x\in D}\{a^s(x)\} \quad \text{and} \quad a^s_{\max} := \|a^s\|_{L^\infty(D)}.$$

By u^s we denote the solution of the variational formulation (2.1) with the s-term truncated, parametric coefficient a^s in place of a, *i.e.*,

$$u^s : \Omega \to V \text{ s.t. } \quad \int_D a^s \nabla u^s \cdot \nabla v \mathrm{d}x = f(v), \quad v \in V. \tag{2.2}$$

The *truncation error* can be controlled if the sequence $(b_j)_{j\geq 1}$ is p-summable. Specifically, if $(b_j)_{j\geq1} \in \ell^{p_0}(\mathbb{N})$ for some $p_0 \in (0,\infty)$, Proposition 7 of [31] implies that for every $\varepsilon > 0$ there exists a constant $C_\varepsilon > 0$ such that for every $G(\cdot) \in V^*$ and for every $s \in \mathbb{N}$

$$|\mathbb{E}(G(u)) - \mathbb{E}(G(u^s))| \leq C_\varepsilon \|G(\cdot)\|_{V^*}\|f\|_{V^*} \begin{cases} \sup_{j>s}\{b_j^{1-\varepsilon}\} & \text{if } p_0 > 2, \\ \sup_{j>s}\{b_j^{2-p_0/2}\} & \text{if } p_0 \leq 2. \end{cases} \tag{2.3}$$

2.2. Sample regularity in D. Weighted function spaces

Approximations of second order, elliptic PDEs with regular, simplicial FEs in a Lipschitz polytope $D \subset \mathbb{R}^d$, $d = 2, 3$, on regular, simplicial families of uniformly refined triangulations may produce suboptimal convergence rates, due to the occurrence of singularities in the parametric solutions u and u^s at vertices and, in space dimension $d = 3$, also at edges of ∂D. In such domains, linear elliptic PDEs admit regularity shifts in certain weighted Sobolev spaces, *cf.* [5,40], which we now recapitulate as we require the precise definition of the weighted norms in D in the ensuing QMC error analysis. We assume the polygon resp. polyhedron D to have straight edges resp. plane faces and J corners $\mathcal{C} := \{c_1, \ldots, c_J\} \subset \partial D$.

For $d = 2$, let $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_J)$ be a J-tuple of weight exponents. We define the *corner weight function*

$$\Phi_{\boldsymbol{\beta}}(x) := \prod_{i=1}^{J} |x - c_i|^{\beta_i}, \quad x \in D,$$

where $\beta_i \in [0,1)$, $i = 1, \ldots, J$. Here and in the following, the Euclidean norm in $\mathbb{R}^d$ is denoted by $|\cdot|$. The weighted function spaces $L^q_{\boldsymbol{\beta}}(D)$ and $H^2_{\boldsymbol{\beta}}(D)$ are defined as closures of $C^\infty(\overline{D})$ with respect to the norms

$$\|v\|_{L^q_{\boldsymbol{\beta}}(D)} := \|v\Phi_{\boldsymbol{\beta}}\|_{L^q(D)}, \quad q \in [1,\infty],$$

and

$$\|v\|^2_{H^2_{\boldsymbol{\beta}}(D)} := \|v\|^2_{H^1(D)} + \sum_{|\boldsymbol{\alpha}|=2} \||\partial_x^{\boldsymbol{\alpha}} v|\Phi_{\boldsymbol{\beta}}\|^2_{L^2(D)}.$$

For $d = 3$, let the polyhedron D have J' straight edges $\mathcal{E} := \{e_1, \ldots, e_{J'}\} \subset \partial D$ and define $\mathcal{X}_j := \{k : c_j \in \overline{e_k}\}$ as the index set of edges that meet at corner c_j, $j = 1, \ldots, J$. Let r_k denote the distance to the edge e_k and let ρ_j denote the distance to the corner c_j. Let $(V_j : j = 1, \ldots, J)$ be a finite, open covering of D such that

$$\overline{D} \subset \bigcup_{j=1}^{J} V_j, \quad c_i \notin \overline{V}_j, \text{ if } i \neq j, \quad \text{and} \quad \overline{V}_j \cap \overline{e}_k = \emptyset \text{ if } k \notin \mathcal{X}_j.$$

For a real-valued J-tuple $\boldsymbol{\beta} \in [0,1)^J$ and a real-valued J'-tuple $\boldsymbol{\delta} \in [0,1)^{J'}$, define the corner-edge weight function

$$\Phi_{(\boldsymbol{\beta},\boldsymbol{\delta})}(x) := \sum_{j=1}^{J} \rho_j^{\beta_j}(x) \prod_{k\in\mathcal{X}_j} \left(\frac{r_k(x)}{\rho_j(x)}\right)^{\delta_k} \mathbb{1}_{V_j}(x), \quad x \in D. \tag{2.4}$$

With this weight, we associate the weighted Sobolev spaces $L^2_{(\boldsymbol{\beta},\boldsymbol{\delta})}(D)$ and $H^2_{(\boldsymbol{\beta},\boldsymbol{\delta})}(D)$, *cf.* Section 4.1.2 of [40] as closures of $C_0^\infty(\overline{D}\backslash(\mathcal{C}\cup\mathcal{E}))$ with respect to the norms

$$\|v\|_{L^2_{(\boldsymbol{\beta},\boldsymbol{\delta})}(D)} := \|v\Phi_{\boldsymbol{\beta},\boldsymbol{\delta}}\|_{L^2(D)}$$

and for $\iota = 0,1,2$,

$$\|v\|_{H^\iota_{(\boldsymbol{\beta},\boldsymbol{\delta})}(D)} := \left(\sum_{j=1}^{J}\sum_{|\boldsymbol{\alpha}|\le\iota}\int_{D\cap V_j} \rho_j^{2(\beta_j-\iota+|\boldsymbol{\alpha}|)}(x) \prod_{k\in\mathcal{X}_j}\left(\frac{r_k(x)}{\rho_j(x)}\right)^{2(\delta_j-\iota+|\boldsymbol{\alpha}|)} |\partial_x^{\boldsymbol{\alpha}} v|^2 \mathrm{d}x\right)^{1/2}.$$

We note that the spaces $L^2_{(\boldsymbol{\beta},\boldsymbol{\delta})}(D)$ and $H^2_{(\boldsymbol{\beta},\boldsymbol{\delta})}(D)$ are isomorphic with equivalent norms: for every $x \in D$,

$$\sum_{j=1}^{J} \rho_j^{2\beta_j}(x) \prod_{k\in\mathcal{X}_j}\left(\frac{r_k(x)}{\rho_j(x)}\right)^{2\delta_k} \mathbb{1}_{V_j}(x) \le (\Phi_{(\boldsymbol{\beta},\boldsymbol{\delta})}(x))^2 \le J\sum_{j=1}^{J} \rho_j^{2\beta_j}(x) \prod_{k\in\mathcal{X}_j}\left(\frac{r_k(x)}{\rho_j(x)}\right)^{2\delta_k} \mathbb{1}_{V_j}(x),$$

Also, we define the weighted seminorm

$$|v|_{H^2_{(\boldsymbol{\beta},\boldsymbol{\delta})}(D)} := \left(\sum_{j=1}^{J}\sum_{|\boldsymbol{\alpha}|=2}\int_{D\cap V_j} \rho_j^{2\beta_j}(x) \prod_{k\in\mathcal{X}_j}\left(\frac{r_k(x)}{\rho_j(x)}\right)^{2\delta_j} |\partial_x^{\boldsymbol{\alpha}} v|^2 \mathrm{d}x\right)^{1/2}.$$

Lemma 2.1. *For a polygon D (i.e. in spatial dimension $d = 2$), there exists a constant $C > 0$ such that for every $f \in L^2_{\boldsymbol{\beta}}(D)$,*

$$\|f\|_{V^*} \le C\|f\|_{L^2_{\boldsymbol{\beta}}(D)}.$$

For a polyhedron D (i.e. in spatial dimension $d = 3$), there exists a constant $C > 0$ such that for every $f \in L^2_{(\boldsymbol{\beta},\boldsymbol{\delta})}(D)$,

$$\|f\|_{V^*} \le C\|f\|_{L^2_{(\boldsymbol{\beta},\boldsymbol{\delta})}(D)}.$$

Proof. The case $d = 2$ is proven in Lemma 1 from [30]. The case $d = 3$ follows by Lemma 4.1.4 of [40]. Specifically, in the notation of [40] the assertion of this lemma reads that the embedding $V^{0,2}_{\boldsymbol{\beta},\boldsymbol{\delta}}(D) \subset V^{-1,2}_{\mathbf{0},\mathbf{0}}(D)$ is continuous, if $\beta_j < 1$ and $\delta_k < 1$, $j = 1,\dots,J$, $k = 1,\dots,J'$. We note that here the space $V^{0,2}_{\boldsymbol{\beta},\boldsymbol{\delta}}(D)$ of [40] coincides with our spaces $H^0_{(\boldsymbol{\beta},\boldsymbol{\delta})}(D) = L^2_{(\boldsymbol{\beta},\boldsymbol{\delta})}(D)$ and the space $V^{-1,2}_{\mathbf{0},\mathbf{0}}(D)$ is isomorphic to V^*. In the definition of the weighted space $L^2_{(\boldsymbol{\beta},\boldsymbol{\delta})}(D) = H^0_{(\boldsymbol{\beta},\boldsymbol{\delta})}(D)$, it has been assumed that $\beta_j < 1$ and $\delta_k < 1$, $j = 1,\dots,J$, $k = 1,\dots,J'$. □

In polygons D in space dimension $d = 2$ and for functions in $H^2_{\boldsymbol{\beta}}(D)$, a full regularity shift for the Laplacian is available, *cf.* for example Theorem 3.2 of [5]: there exists a constant $C > 0$ such that for every $w \in V$ with $\Delta w \in L^2_{\boldsymbol{\beta}}(D)$,

$$\|w\|_{H^2_{\boldsymbol{\beta}}(D)} \le C\|\Delta w\|_{L^2_{\boldsymbol{\beta}}(D)}, \tag{2.5}$$

where we assume that the weight exponent sequence $\boldsymbol{\beta}$ satisfies $\max\{0, 1-\pi/\omega_i\} < \beta_i < 1$, $i = 1,\dots,J$. Here, ω_i denotes the interior angle of the polygon D at corner c_i, $i = 1,\dots,J$. Since [5] considers the Poisson boundary value problem with a zero order term, *i.e.*, $-\Delta u + u = f$, we note that Lemma 2.1 implies that there exists a constant C such that for every $w \in V \cap H^2_{\boldsymbol{\beta}}(D)$, $\|w\|_{L^2_{\boldsymbol{\beta}}(D)} \le C\|\Delta w\|_{L^2_{\boldsymbol{\beta}}(D)}$.

In space dimension $d = 3$, when D is a polyhedral domain with plane sides and for functions in $H^2_{(\boldsymbol{\beta},\boldsymbol{\delta})}(D) \cap V$, there holds a corresponding regularity shift of the Dirichlet Laplacian by Lemma 4.3.1 of [40] and by the inverse mapping theorem, *cf.* [13], Theorem 5.6.2: there exists a constant $C > 0$ such that for every $w \in H^2_{(\boldsymbol{\beta},\boldsymbol{\delta})}(D) \cap V$ holds

$$\|w\|_{H^2_{(\boldsymbol{\beta},\boldsymbol{\delta})}(D)} \leq C \|\Delta w\|_{L^2_{(\boldsymbol{\beta},\boldsymbol{\delta})}(D)}, \tag{2.6}$$

where we assume that

$$\frac{1}{2} - \lambda_j < \beta_j < 1, \quad j = 1, \ldots, J, \quad \text{and} \quad 1 - \frac{\pi}{\omega_k} < \delta_k < 1, \quad k = 1, \ldots, J',$$

where ω_k is the interior angle between two faces meeting at edge e_k and λ_j is given by

$$\lambda_j := -\frac{1}{2} + \sqrt{\Lambda_j + \frac{1}{4}},$$

where Λ_j is the smallest, strictly positive eigenvalue of the Dirichlet Laplace–Beltrami operator on the intersection of the unit sphere centered at c_j and the infinite, interior polyhedral tangent cone to ∂D with vertex c_j, *cf.* Section 4.3.1 of [40].

2.3. FE convergence theory

Let $\{\mathcal{T}_\ell\}_{\ell \geq 0}$ denote a sequence of regular, simplicial triangulations of D with proper mesh refinements near vertices and, if $d = 3$, also near edges of D. Let further $\mathbb{P}^1(K)$ denote the affine functions on a subset K of $\mathbb{R}^d$, *i.e.*, the polynomial degree $r = 1$. In FE spaces $V_\ell := \{v \in V : v|_K \in \mathbb{P}^1(K), K \in \mathcal{T}_\ell\}$ of continuous, piecewise affine functions on $\{\mathcal{T}_\ell\}_{\ell \geq 0}$, optimal asymptotic convergence rates are achievable, also in the presence of singularities. We state these for subsequent reference, recapitulating from [2,4,5,21] approximation properties in $H^1(D)$ of the FE spaces V_ℓ.

Specifically, there exists $C > 0$ such that for every $w \in H^2_{\boldsymbol{\beta}}(D) \cap V$ for $d = 2$, resp. for every $w \in H^2_{(\boldsymbol{\beta},\boldsymbol{\delta})}(D) \cap V$ for $d = 3$, and for every $\ell \geq 0$ there exists $w_\ell \in V_\ell$ such that, with $M_\ell := \dim(V_\ell)$,

$$\|w - w_\ell\|_V \leq C M_\ell^{-1/d} \begin{cases} \|w\|_{H^2_{\boldsymbol{\beta}}(D)} & \text{if } d = 2, \\ \|w\|_{H^2_{(\boldsymbol{\beta},\boldsymbol{\delta})}(D)} & \text{if } d = 3. \end{cases} \tag{2.7}$$

For $d = 2$, the convergence rate bound (2.7) is due to Lemmas 4.1 and 4.5 of [5] for regular, graded simplicial meshes, resp. due to [21] for simplicial bisection tree meshes. In polyhedral domains D in space dimension $d = 3$, this estimate follows by Theorem 4.6 of [4] for every $w \in H^2_{(\boldsymbol{\beta},\boldsymbol{\delta})}(D)$ (in [4] denoted by $W^{2,2}_{\boldsymbol{\beta},\boldsymbol{\delta}}(D)$). Specifically, in the proof of Theorem 4.6 from [4] an interpolation error bound on $H^1(D)$ for functions in $H^2_{(\boldsymbol{\beta},\boldsymbol{\delta})}(D)$ is obtained for an interpolant Z_{h_ℓ} defined on page 1212 of [4] based on Lemma 4.4 of [4]. Inspecting the proof of Lemma 4.4 from [4], for every $v \in H^2_{(\boldsymbol{\beta},\boldsymbol{\delta})}(D)$, $v|_{\partial D} = 0$ implies $Z_{h_\ell}(v)|_{\partial D} = 0$, $Z_{h_\ell} : H^2_{(\boldsymbol{\beta},\boldsymbol{\delta})}(D) \cap V \to V_\ell$, *i.e.*, the interpolant in ([4], Lem. 4.4) preserves homogeneous boundary values. We also assume $\beta_j < 2/3$, $j = 1, \ldots, J$, and $\delta_k < 2/3$, $k = 1, \ldots, J'$. Since $1/2 - \lambda_j \leq 1/2$, $j = 1, \ldots, J$, and $1 - \pi/\omega_k \leq 1/2$, $k = 1, \ldots, J'$, this does not restrict generality. See also [3, 37] for further approximation results of FEM on anisotropically refined meshes in a polyhedral domain.

In the case of quasi-uniform mesh refinement, the convergence rates are well-known to be limited by the strongest singularity. For example for $d = 2$, the FE approximation w_h of the solution w to the diffusion equation $-\nabla \cdot (\exp(\widehat{Z}) \nabla w) = f$, $w|_{\partial D} = 0$, with deterministic $\widehat{Z} \in C^t(\overline{D})$ and $f \in C^\infty(D)$ satisfies

$$\|w - w_h\|_V \leq C M_h^{-\min\{\tau, r\}/d},$$

where $r \in \mathbb{N}$ is the polynomial degree of the FE space with dimension M_h and $\tau \in (0, \min\{t, \pi/\beta_{\max}\})$, $\beta_{\max} := \max\{\omega_1, \ldots, \omega_J\}$[2]. The Hölder spaces $C^t(\overline{D})$, $t \in [0, \infty)$, are sometimes also denoted by $C^{\lfloor t \rfloor, t - \lfloor t \rfloor}(\overline{D})$.

[2]Please note a misprint in Proposition 15 and Theorem 17 from [31], $\max\{t, \pi/\beta_{\max}\}$ should be replaced by $\min\{t, \pi/\beta_{\max}\}$ which is inconsequential for the preceding derivations and conclusions of [31].

2.4. Combined dimension truncation FE error bound

We now derive an error bound for the combined effect of truncating the GRF Z to a finite number of parameters s, and to FE discretization in D of the resulting s-parametric problem (2.2).

Let accordingly $u^{s,\mathcal{T}_\ell}:\Omega\to V_\ell$ denote the FE solution, *i.e.*,

$$\int_D a^s \nabla u^{s,\mathcal{T}_\ell}\cdot\nabla v \mathrm{d}x = f(v),\quad \forall v\in V_\ell. \tag{2.8}$$

For notational convenience, we introduce

$$\overline{\boldsymbol{\beta}} := \begin{cases} \boldsymbol{\beta} & \text{if } d=2,\\ (\boldsymbol{\beta},\boldsymbol{\delta}) & \text{if } d=3.\end{cases} \tag{2.9}$$

The Banach space $W^{1,\infty}_{\overline{\boldsymbol{\beta}}}(D)$ is the space of all measurable functions $v:D\to\mathbb{R}$ that have finite $W^{1,\infty}_{\overline{\boldsymbol{\beta}}}(D)$-norm, where

$$\|v\|_{W^{1,\infty}_{\overline{\boldsymbol{\beta}}}(D)} := \max\{\|v\|_{L^\infty(D)}, \||\nabla v|\Phi_{\overline{\boldsymbol{\beta}}}\|_{L^\infty(D)}\}.$$

In order for the multilevel algorithm Q^*_L to yield improved (w.r. to the single-level algorithm) error *vs.* work bounds, we require stronger assumptions than in the single-level analyses of [26, 31] on the function system $(\psi_j)_{j\ge1}$. This corresponds to what was found for uniform random parameters in [35] and in the lognormal case for ψ_j with global supports in [36]. The "local support" condition (**A1**) is, in the MLQMC-FE algorithm, strengthened as follows: there exist a constant $K_1>0$ and a positive sequence $(\bar{b}_j)_{j\ge1}$ such that

$$K_1 := \left\|\sum_{j\ge1}\frac{\max\{|\nabla\psi_j|\Phi_{\overline{\boldsymbol{\beta}}}, |\psi_j|\}}{\bar{b}_j}\right\|_{L^\infty(D)} < \infty. \tag{\textbf{A2}}$$

Remark 2.2. When the precision operator of Z is a positive power of a shifted Dirichlet Laplacian on D the Karhunen–Loève eigenfunctions v_j are, by the spectral mapping theorem, eigenfunctions of the Dirichlet Laplacian on D: $-\Delta v_j=\nu_j v_j$, $v_j|_{\partial D}=0$, $j\in\mathbb{N}$. Here, the eigenvalues ν_j are related to the ones appearing in the Karhunen–Loève expansion of the GRF Z by the spectral mapping theorem. Elliptic regularity shifts for the Dirichlet Laplacian are also known in certain weighted Hölder spaces in D: for $d=3$, Lemma 4.3.1.2 of [40], implies that $v_j\in W^{1,\infty}_{(\boldsymbol{\beta},\boldsymbol{\delta})}(D)$ provided that $1-\lambda_j<\beta_j<1$, $j=1,\dots,J$, and $1-\pi/\theta_k<\delta_k<1$, $k=1,\dots,J'$, where we used here that the weighted $C^{1+\varepsilon}(\overline{D})$-type space $N^{1,\varepsilon}_{\boldsymbol{\beta},\boldsymbol{\delta}}(D)$ (in the notation of [40], Sects. 4.2 and 4.3) embeds continuously into $W^{1,\infty}_{(\boldsymbol{\beta},\boldsymbol{\delta})}(D)$. Note that this condition on $\boldsymbol{\beta}$ for the KL eigenfunctions is stronger than in assumption (**A2**). Similar statements hold for $d=2$. Here, singularities at corners and (for $d=3$) along edges of the Karhunen–Loève eigenfunctions appear as a consequence of regularity shifts for the Dirichlet Laplacian in weighted Hölder spaces. The structure of the weight functions $\Phi_{\overline{\boldsymbol{\beta}}}$ (which depend only on D and on the (Dirichlet) Laplacian) in the assumption (**A2**) on the Karhunen–Loève eigenfunctions is identical to the weights in the elliptic regularity shift (2.6). In the case of Matérn-type covariance functions, as induced by solutions to (1.1), there is neither dependence of the functional form of the weight functions on the regularity nor on the positive correlation length of the respective GRF. Note, however, that in general, Karhunen–Loève eigenfunctions have global support in D.

Assumption (**A2**) implies $W^{1,\infty}_{\overline{\boldsymbol{\beta}}}(D)$-regularity of the GRF Z and strong approximation by its truncated expansion. This is made precise in the following proposition. Its proof is completely analogous to Theorem 2 of [31] and therefore not detailed.

Proposition 2.3. *Let the assumption in* (**A2**) *be satisfied for some sequence* $(\bar{b}_j)_{j\geq 1}$ *such that* $(\bar{b}_j)_{j\geq 1} \in \ell^{p_0}(\mathbb{N})$ *for some* $p_0 \in (0,\infty)$*. For every* $\varepsilon > 0$ *and* $q \in [1,\infty)$ *there exists a constant* $C > 0$ *such that for every* $s \in \mathbb{N}$*,*

$$\|Z - Z^s\|_{L^q(\Omega:W^{1,\infty}_{\overline{\boldsymbol{\beta}}}(D))} \leq C \sup_{j>s}\{\bar{b}_j^{1-\varepsilon}\}.$$

Since $(\nabla a)\Phi_{\overline{\boldsymbol{\beta}}} = (a\nabla Z)\Phi_{\overline{\boldsymbol{\beta}}}$ holds in $L^\infty(D)^d$, μ-a.s., Proposition 2.3 and Corollary 6 of [31] imply with the Cauchy–Schwarz inequality that for every $q \in [1,\infty)$ there exists a constant $C > 0$ such that for every $s \in \mathbb{N}$,

$$\|a\|_{L^q(\Omega;W^{1,\infty}_{\overline{\boldsymbol{\beta}}}(D))} < \infty \quad \text{and} \quad \|a^s\|_{L^q(\Omega;W^{1,\infty}_{\overline{\boldsymbol{\beta}}}(D))} \leq C < \infty. \tag{2.10}$$

To obtain an estimate of the Laplacian of u, we note that in any compact subset $\widetilde{D} \subset\subset D$ it holds $-a\Delta u = f - \nabla a \cdot \nabla u$, μ-a.s., where we assume that $f \in L^2_{\overline{\boldsymbol{\beta}}}(D)$. This equation may be tested with $-\Delta u \Phi^2_{\overline{\boldsymbol{\beta}}}/a$, which implies with Lemma 2.1

$$\|\Delta u\|_{L^2_{\overline{\boldsymbol{\beta}}}(D)} \leq \frac{\|f\|_{L^2_{\overline{\boldsymbol{\beta}}}(D)}}{a_{\min}} + \|Z\|_{W^{1,\infty}_{\overline{\boldsymbol{\beta}}}(D)}\|u\|_V \leq C\frac{\|f\|_{L^2_{\overline{\boldsymbol{\beta}}}(D)}}{a_{\min}}(1 + \|Z\|_{W^{1,\infty}_{\overline{\boldsymbol{\beta}}}(D)}). \tag{2.11}$$

An Aubin–Nitsche duality argument, (2.3), (2.5), (2.7), Proposition 2.3, (2.10), and (2.11) imply that for every $\varepsilon > 0$, there exists a constant $C > 0$ such that for every $s \in \mathbb{N}$, $\ell \in \mathbb{N}_0$

$$|\mathbb{E}(G(u)) - \mathbb{E}(G(u^{s,\mathcal{T}_\ell}))| \leq C\left(\sup_{j>s}\{b_j^t\} + M_\ell^{-2/d}\right)\|f\|_{L^2_{\overline{\boldsymbol{\beta}}}(D)}\|G\|_{L^2_{\overline{\boldsymbol{\beta}}}(D)}, \tag{2.12}$$

where $t = 2 - p_0/2$ if $p_0 \leq 2$ and $t = 1 - \varepsilon$ otherwise. Recall that $(b_j)_{j\geq 1} \in \ell^{p_0}(\mathbb{N})$ for some $p_0 \in (0,\infty)$. In this setting, $p_0 \in (0,2)$ is the range of applicability of QMC, *cf.* Theorem 11 of [31].

Remark 2.4. By interpolation, the error estimate in (2.12) extends to the case that $f \in (V^*, L^2_{\overline{\boldsymbol{\beta}}}(D))_{t,\infty}$ and $G(\cdot) \in (V^*, L^2_{\overline{\boldsymbol{\beta}}}(D))_{t',\infty}$ for some $t, t' \in [0,1]$. Then the estimate in (2.12) holds with the term $M_\ell^{-2/d}$ replaced by $M_\ell^{-(t+t')/d}$. To see this, we observe that the real method of interpolation can be applied to the regularity shifts in (2.5) and in (2.11). Specifically, the linear operator relating the solution $u \in V$ to its approximation error with a V-bounded, and quasioptimal projector $\Pi_\ell : V \to V_\ell$, where Π_ℓ is, for example, the $H^1_0(D)$-projection. From the approximation property in (2.7), the interpolation couple $L^2_{\overline{\boldsymbol{\beta}}}(D) \subset V^*$ then yields the fractional convergence order. Here and throughout what follows, interpolation spaces shall be understood with respect to the real method of interpolation; we refer to Chapter 1 of [45].

3. QMC integration

With convergence rate bounds on the dimension truncation and the FE discretization error at hand, we address the numerical approximation of the expectations in (2.12) with respect to the GM μ. Due to dimension truncation, we evaluate its s-variate section, *i.e.* we integrate w.r.t. to the GM on $\mathbb{R}^s$. As in [26], we approximate the s-variate integrals by so-called randomly shifted lattice rules proposed in [41]. Accordingly, we review QMC error estimates of randomly shifted lattice rules for high-dimensional integrals with respect to the s-variate normal distribution. The construction of generating vectors for such QMC rules in particular with respect to Gaussian and exponentially decaying weight functions with a fast CBC construction have been found in [41]. There, concrete error estimates of the resulting QMC rules in the mean-square sense (with respect to the random shift) have been derived, *cf.* Theorem 8 of [41]. See also Examples 4 and 5 of [34] for the estimation of constants appearing in the error bound of Theorem 8 from [41] for Gaussian and exponential weight functions, respectively.

The error analysis of randomly shifted lattice rules requires, for sequences of positive weights $\boldsymbol{\gamma} = (\gamma_{\mathfrak{u}})_{\mathfrak{u}\subset\mathbb{N}}$, indexed by all finite subsets $\mathfrak{u} \subset \mathbb{N}$, the weighted Sobolev space $\mathcal{W}_{\boldsymbol{\gamma}}(\mathbb{R}^s)$ of mixed first order derivatives, which is defined by the following norm

$$\|F\|_{\mathcal{W}_{\boldsymbol{\gamma}}(\mathbb{R}^s)} := \left(\sum_{\mathfrak{u}\subset\{1:s\}} \gamma_{\mathfrak{u}}^{-1} \int_{\mathbb{R}^{|\mathfrak{u}|}} \left| \int_{\mathbb{R}^{s-|\mathfrak{u}|}} \partial^{\mathfrak{u}} F(\boldsymbol{y}) \prod_{j\in\{1:s\}\setminus\mathfrak{u}} \phi(y_j) \mathrm{d}\boldsymbol{y}_{\{1:s\}\setminus\mathfrak{u}} \right|^2 \prod_{j\in\mathfrak{u}} w_j^2(y_j) \mathrm{d}\boldsymbol{y}_{\mathfrak{u}} \right)^{1/2}, \tag{3.1}$$

where $\{1 : s\} := \{1, \dots, s\}$.
Here, ϕ denotes the univariate normal density

$$\phi(y) := \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}, \quad y \in \mathbb{R}.$$

The norm in (3.1) is considered with respect to Gaussian and exponential weight functions

$$w_{\mathrm{g},j}^2(y) := e^{-\frac{y^2}{2\alpha_{\mathrm{g},j}}}, \quad y \in \mathbb{R}, j \in \mathbb{N}, \quad \text{and} \quad w_{\exp,j}^2(y) := e^{-\alpha_{\exp,j}|y|}, \quad y \in \mathbb{R}, j \in \mathbb{N}. \tag{3.2}$$

The parameters $\alpha_{\mathrm{g},j} > 1$ and $\alpha_{\exp,j} > 0$ will be determined in the ensuing error analysis. If the parameters $\alpha_{\mathrm{g},j}$ or $\alpha_{\exp,j}$ are constant with respect to j, we omit j for ease of presentation. In the following we consider the case $\alpha_{\mathrm{g},j} = \alpha_{\mathrm{g}} > 1$ and $\alpha_{\exp,j} = \alpha_{\exp} > 0$ for every $j \in \mathbb{N}$. In this work, we consider in (3.1) product weights $\boldsymbol{\gamma} = (\gamma_{\mathfrak{u}})_{\mathfrak{u}\subset\mathbb{N}}$, determined by a positive QMC weight sequence $(\gamma_j)_{j\geq1}$, *i.e.*,

$$\gamma_{\mathfrak{u}} = \prod_{j\in\mathfrak{u}} \gamma_j, \quad \mathfrak{u} \subset \mathbb{N}, |\mathfrak{u}| < \infty.$$

We will denote the QMC approximation in s dimensions with N points by $Q_{s,N}(\cdot)$. It shall approximate integrals with respect to the multivariate normal distribution which we denote for every integrand $F \in L^1(\mathbb{R}^s, \mu)$ by

$$I_s(F) := \int_{\mathbb{R}^s} F(\boldsymbol{y}) \prod_{j\in\{1:s\}} \phi(y_j) \mathrm{d}\boldsymbol{y}.$$

For a sequence of dimension truncations $(s_\ell)_{\ell=0,\dots,L}$ and a sequence $(N_\ell)_{\ell=0,\dots,L}$, $L \in \mathbb{N}_0$, the MLQMC quadrature algorithm of [35] is defined by

$$Q_L^*(G(u^L)) := \sum_{\ell=0}^{L} Q_{s_\ell,N_\ell}(G(u^\ell) - G(u^{\ell-1})), \quad L \geq 0, \tag{3.3}$$

with the understanding that $G(u^{-1}) := 0$. We used the notation that $u^\ell := u^{s_\ell,\mathcal{T}_\ell}$, $\ell \geq 0$. Multilevel QMC algorithms stemming from randomly shifted lattice rules have been considered in [35, 36]. The following error estimate (see [35], Eq. (23) or [36], Eq. (3.2)) holds due to the independence of the random shifts on the different levels

$$\mathbb{E}^{\boldsymbol{\Delta}}(|I_s(G(u^L)) - Q_L^*(G(u^L))|^2) = \sum_{\ell=0}^{L} \mathbb{E}^{\boldsymbol{\Delta}}(|I_s(G(u^\ell - u^{\ell-1})) - Q_{s_\ell,N_\ell}(G(u^\ell - u^{\ell-1}))|^2), \tag{3.4}$$

where we apply a randomly shifted lattice rule with respect to (possibly) a different QMC weight sequence on the PDE discretization level $\ell = 0$. The expectation with respect to the random shifts is denoted by $\mathbb{E}^{\boldsymbol{\Delta}}(\cdot)$.

In [31], convergence of randomly shifted lattice rules with product weights is investigated, which relies on parametric regularity estimates of a particular form. We summarize the QMC convergence theory in the following theorem.

Theorem 3.1. *Let $(\widetilde{b}_j)_{j\geq 1}$ be a positive sequence and $\kappa > 0$ such that for some $F : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$ there exists a constant $C > 0$ and a positive function $H(\boldsymbol{y})$ such that for every $\boldsymbol{y} \in \{\boldsymbol{y} \in \mathbb{R}^{\mathbb{N}} : \exists s \in \mathbb{N}, y_j = 0 : \forall j > s\}$,*

$$\sum_{\mathfrak{u}\subset\mathbb{N},|\mathfrak{u}|<\infty} |\partial^{\mathfrak{u}} F(\boldsymbol{y})|^2 \prod_{j\in\mathfrak{u}} \left(\frac{\kappa}{\widetilde{b}_j}\right)^2 \leq CH(\boldsymbol{y})^2.$$

(1) *Let $(\widetilde{b}_j)_{j\geq 1} \in \ell^p(\mathbb{N})$ for some $p \in (2/3, 2)$. For $\varepsilon \in (0, 3/4 - 1/(2p))$, set $p' = p/4 + 1/2 - \varepsilon p \in (0, 1)$. Consider the Gaussian weight functions $(w_{\mathrm{g},j})_{j\geq 1}$ with parameter α_{g} and QMC integration weight sequence*

$$\alpha_{\mathrm{g}} \in \left(\frac{p}{2(p-p')}, \frac{p}{p-2(1-p')}\right) \quad \text{and} \quad \gamma_j = \widetilde{b}_j^{2p'}, \quad j \geq 1.$$

Then, there exists a constant C (independent of F) such that for $q_0 = 2qq'/(q'-q)$, where $q = p/(2(1-p'))$ and $q' \in (q, \alpha_{\mathrm{g}}/(\alpha_{\mathrm{g}} - 1))$,

$$\sqrt{\mathbb{E}^{\boldsymbol{\Delta}}(|I_s(F) - Q_{s,N}(F)|^2)} \leq C(\varphi(N))^{-1/(2p)-1/4+\varepsilon} \|H\|_{L^{q_0}(\mathbb{R}^s,\mu)}.$$

(2) *Let $(\widetilde{b}_j)_{j\geq 1} \in \ell^p(\mathbb{N})$ for some $p \in (2/3, 1]$. Assume that $H(\boldsymbol{y}) \leq \eta_1 \exp(\eta_2 \sum_{j\geq 1} \widetilde{b}_j |y_j|)$ for some $\eta_1, \eta_2 > 0$. Set $p' = 1 - p/2$. Consider the exponential weight functions $(w_{\exp,j})_{j\geq 1}$ with parameter $\alpha_{\exp}$ and QMC integration weight sequence*

$$\alpha_{\exp} > 2\eta_2 \sup_{j\geq 1}\{\widetilde{b}_j\} \quad \text{and} \quad \gamma_j = \widetilde{b}_j^{2p'}, \quad j \geq 1.$$

Then, there exists a positive constant C (independent of η_1) such that

$$\sqrt{\mathbb{E}^{\boldsymbol{\Delta}}(|I_s(F) - Q_{s,N}(F)|^2)} \leq C\,(\varphi(N))^{-1/p+1/2}\,\eta_1.$$

The Euler totient function is denoted by $\varphi(\cdot)$.

This theorem was, in the case of Gaussian weight functions, obtained in Theorems 9 and 11 of [31] and in the case of exponential weight functions in Theorems 9 and 12 of [31]. The main ingredient of the proof of Theorem 9 from [31] is a parametric regularity estimate of the form assumed in Theorem 3.1. The parametric regularity estimates derived in [26, 36] for globally supported ψ_j afforded bounds for each partial derivative separately. In [31], we used the bound from Theorem 4.1 of [7] which does account for local supports and affords control of "bulk" sums of (norms of) solution derivatives with respect to the parameters y_j. We also note that in applications, the sequence $(\widetilde{b}_j)_{j\geq 1}$ may be arbitrarily scaled by a factor κ in order to satisfy such a regularity estimate.

4. Parametric regularity

In this section we derive parametric regularity estimates that allow to prove dimension independent convergence rates of MLQMC. We extend the argument that results in the estimate in Theorem 4.1 of [7] to dimensionally truncated and FE differences. The estimate in Theorem 4.1 of [7] was used in our single-level QMC analysis in [31] to prove dimension independent convergence rates of QMC with product weights in the case of local supports. In view of the parametric regularity estimate of an integrand F, which is the condition to apply Theorem 3.1, we seek to prove suitable estimates in the case that $F(\boldsymbol{y}) = G(u(\boldsymbol{y})) - G(u^s(\boldsymbol{y}))$ and $F(\boldsymbol{y}) = G(u(\boldsymbol{y})) - G(u^{\mathcal{T}_\ell}(\boldsymbol{y}))$. To be suitable for MLQMC, the quantitative approximation properties of $u^s(\boldsymbol{y}) \approx u(\boldsymbol{y})$ and $u^{\mathcal{T}_\ell}(\boldsymbol{y}) \approx u(\boldsymbol{y})$ shall be contained in these estimates; see ahead Theorems 4.3 and 4.10.

For every finite $s \in \mathbb{N}$, the truncated fields Z^s, a^s, and u^s, are well-defined regardless of assumption (**A1**). In particular, $Z^s = \sum_{j=1}^{s} y_j \psi_j$ is well-defined for every $\boldsymbol{y} \in \Omega = \mathbb{R}^{\mathbb{N}}$. We may therefore interpret Z^s as a mapping from $\mathbb{R}^s$ to $L^\infty(D)$ such that pointwise evaluation is well-defined for every $\boldsymbol{y} \in \mathbb{R}^s$. Similarly, a^s and u^s may be interpreted as mappings from $\mathbb{R}^s$ to $L^\infty(D)$ and to V, respectively. In the same way Z, a, and u are mappings with pointwise evaluation from the set

$$U := \{\boldsymbol{y} \in \Omega : \exists s \in \mathbb{N}, y_j = 0, j > s\}$$

to $L^\infty(D)$ and V, respectively. Note that $\mathbb{R}^s \times \{\mathbf{0}\} \subset U = \bigcup_{s\in\mathbb{N}} \mathbb{R}^s \times \{\mathbf{0}\}$ for every $s \in \mathbb{N}$, where $\mathbf{0} \in \mathbb{R}^{\mathbb{N}\setminus\{1:s\}}$. Hence, the set U of admissible parameters $\boldsymbol{y}$ is sufficiently rich for the ensuing QMC convergence analysis. The mappings Z^s, a^s, and u^s extend naturally to mappings from U to $L^\infty(D)$ and to V, respectively. For the ensuing analysis, let us introduce the parametric energy norm

$$\|v\|_{a(\boldsymbol{y})} := \left(\int_D a(\boldsymbol{y})|\nabla v|^2 \mathrm{d}x\right)^{1/2}, \quad \forall v \in V, \boldsymbol{y} \in U.$$

4.1. Dimension truncation

For every $\boldsymbol{y} \in U$, the difference $u(\boldsymbol{y}) - u^s(\boldsymbol{y})$ satisfies the variational formulation

$$\int_D a(\boldsymbol{y})\nabla(u(\boldsymbol{y}) - u^s(\boldsymbol{y})) \cdot \nabla v \mathrm{d}x = -\int_D (a(\boldsymbol{y}) - a^s(\boldsymbol{y}))\nabla u^s(\boldsymbol{y}) \cdot \nabla v \mathrm{d}x, \quad \forall v \in V. \tag{4.1}$$

We will mostly (in the proofs) omit the $\boldsymbol{y}$ dependence in the following. Set $\mathcal{F} := \{\boldsymbol{\tau} \in \mathbb{N}_0^{\mathbb{N}} : |\boldsymbol{\tau}| < \infty\}$. For every real-valued sequence $\rho = (\rho_j)_{j\geq 1}$ and $\boldsymbol{\tau} \in \mathcal{F}$, we shall use the notation $\rho^{\boldsymbol{\tau}} = \prod_{j\geq 1} \rho_j^{\tau_j}$. For every $\boldsymbol{\tau} \in \mathcal{F}$ and a positive sequence $(\rho_j)_{j\geq 1}$, let us define

$$\kappa_0(\boldsymbol{\tau}, \boldsymbol{\nu}) := \frac{\sqrt{\boldsymbol{\tau}!}}{\sqrt{\boldsymbol{\nu}!}} \frac{\rho^{\boldsymbol{\tau}-\boldsymbol{\nu}} |\psi|^{\boldsymbol{\tau}-\boldsymbol{\nu}}}{(\boldsymbol{\tau}-\boldsymbol{\nu})!}, \quad \boldsymbol{\nu} \leq \boldsymbol{\tau}.$$

Also, for given $k, r \in \mathbb{N}$ introduce the set $\Lambda_k := \{\boldsymbol{\tau} \in \mathcal{F} : |\boldsymbol{\tau}| = k, \|\boldsymbol{\tau}\|_{\ell^\infty} \leq r\}$ and for any integer $\ell \leq k-1$ and for $\boldsymbol{\nu} \in \Lambda_\ell$, introduce

$$R_{\boldsymbol{\nu},k} := \{\boldsymbol{\tau} \in \Lambda_k : \boldsymbol{\tau} \geq \boldsymbol{\nu}\},$$

where r denotes the maximal order of differentiability to be considered. The following lemma will be useful in the ensuing analysis.

Lemma 4.1. *Assume that there exists a positive sequence $(\rho_j)_{j\geq 1}$ such that, for some $r \in \mathbb{N}$,*

$$\widetilde{K}_0 := \left\| \sum_{j\geq 1} \rho_j |\psi_j| \right\|_{L^\infty(D)} < \frac{\log(2)}{\sqrt{r}} \cdot \tag{4.2}$$

Then, for every $\boldsymbol{\tau} \in \mathcal{F}$ such that $\|\boldsymbol{\tau}\|_{\ell^\infty} \leq r$,

$$\sum_{\boldsymbol{\nu}\leq\boldsymbol{\tau}, \boldsymbol{\nu}\neq\boldsymbol{\tau}} \kappa_0(\boldsymbol{\tau}, \boldsymbol{\nu}) \leq e^{\sqrt{r}\widetilde{K}_0} - 1 < 1$$

and for every positive integer $\ell \leq k-1$ and multi-index $\boldsymbol{\nu} \in \Lambda_\ell$,

$$\sum_{\boldsymbol{\tau}\in R_{\boldsymbol{\nu},k}} \kappa_0(\boldsymbol{\tau}, \boldsymbol{\nu}) \leq \frac{(\sqrt{r}\widetilde{K}_0)^{k-\ell}}{(k-\ell)!} \cdot$$

The estimates in this lemma are given in equations (4.12) and (4.14) from [7]. The second estimate of Lemma 4.1 holds even if the smallness assumption in (4.2) is not guaranteed. We note that the condition (4.2) is implied by (**A1**) with $\rho_j^{-1} = b_j\bar{K}\sqrt{r}/\log(2)$ provided that $\bar{K} > \|\sum_{j\geq 1}|\psi_j|/b_j\|_{L^\infty(D)}$. For every $s\in\mathbb{N}$, integers $\ell \leq k-1$, and $\boldsymbol{\nu}\in\Lambda_\ell$, introduce the set

$$R^s_{\boldsymbol{\nu},k} := \{\boldsymbol{\tau}\in R_{\boldsymbol{\nu},k} : \exists j > s \text{ such that } \tau_j > \nu_j\}.$$

Lemma 4.2. *Let the assumptions of Lemma 4.1 hold for positive weights* $(\rho_j)_{j\geq 1}$ *such that* $c := \|(\rho_j^{-1})_{j\geq 1}\|_{\ell^\infty(\mathbb{N})} < \infty$. *Further assume that for some* $\eta > 0$

$$\widetilde{K}_\eta := \left\| \sum_{j\geq 1} \rho_j^{1+\eta}|\psi_j| \right\|_{L^\infty(D)} < \infty.$$

Then, for $s\in\mathbb{N}$ *and every* $\boldsymbol{\tau}\in\mathcal{F}$ *such that* $\|\boldsymbol{\tau}\|_{\ell^\infty}\leq r$ *and* $\tau_j > 0$ *for some* $j > s$,

$$\sum_{\boldsymbol{\nu}\leq\boldsymbol{\tau},\nu_j=0\,\forall j>s} \kappa_0(\boldsymbol{\tau},\boldsymbol{\nu}) \leq 2(e^{\sqrt{r}\widetilde{K}_\eta c^\eta}-1)c^{-\eta}\sup_{j>s}\{\rho_j^{-\eta}\}.$$

For $s\in\mathbb{N}$, *positive integers* $\ell\leq k-1$, *and* $\boldsymbol{\nu}\in\Lambda_\ell$ *such that* $\nu_j = 0$, $j > s$,

$$\sum_{\boldsymbol{\tau}\in R^s_{\boldsymbol{\nu},k}} \kappa_0(\boldsymbol{\tau},\boldsymbol{\nu}) \leq \frac{(\sqrt{r}\widetilde{K}_\eta c^\eta)^{k-\ell}}{(k-\ell)!} c^{-\eta}\sup_{j>s}\{\rho_j^{-\eta}\}.$$

Proof. There is $j > s$ such that $\tau_j > 0$. Since κ_0 is a product, by Lemma 4.1,

$$\begin{aligned}\sum_{\boldsymbol{\nu}\leq\boldsymbol{\tau},\nu_j=0\,\forall j>s} \kappa_0(\boldsymbol{\tau},\boldsymbol{\nu}) &= \left(\sum_{\boldsymbol{\nu}_{\{1:s\}}\leq\boldsymbol{\tau}_{\{1:s\}}} \kappa_0(\boldsymbol{\tau}_{\{1:s\}},\boldsymbol{\nu}_{\{1:s\}})\right)\kappa_0(\boldsymbol{\tau}_{\mathbb{N}\setminus\{1:s\}},\mathbf{0}_{\mathbb{N}\setminus\{1:s\}})\\ &\leq 2\kappa_0(\boldsymbol{\tau}_{\mathbb{N}\setminus\{1:s\}},\mathbf{0}_{\mathbb{N}\setminus\{1:s\}}),\end{aligned}$$

where we used the notation that for every $\mathfrak{u}\subset\mathbb{N}$, $\tau_{\mathfrak{u}}$ is a multi-index that satisfies $(\tau_{\mathfrak{u}})_j = \tau_j$, $j\in\mathfrak{u}$, and $(\tau_{\mathfrak{u}})_j = 0$ otherwise. With $c = \|(\rho_j^{-1})_{j\geq 1}\|_{\ell^\infty(\mathbb{N})}$, we obtain

$$\begin{aligned}\kappa_0(\boldsymbol{\tau}_{\mathbb{N}\setminus\{1:s\}},\mathbf{0}_{\mathbb{N}\setminus\{1:s\}}) &\leq \frac{\rho^{\boldsymbol{\tau}_{\mathbb{N}\setminus\{1:s\}}}}{\sqrt{\boldsymbol{\tau}_{\mathbb{N}\setminus\{1:s\}}!}}|\psi|^{\boldsymbol{\tau}_{\mathbb{N}\setminus\{1:s\}}} \leq \exp\left(\sqrt{r}\sum_{j>s}\rho_j|\psi_j|\right)-1\\ &\leq (e^{\sqrt{r}\widetilde{K}_\eta c^\eta}-1)c^{-\eta}\sup_{j>s}\{\rho_j^{-\eta}\}.\end{aligned}$$

For the proof of the second inequality, we observe that

$$\sum_{\boldsymbol{\tau}\in R^s_{\boldsymbol{\nu},k}} \kappa_0(\boldsymbol{\tau},\boldsymbol{\nu}) \leq \sum_{\boldsymbol{\tau}\in R^s_{\boldsymbol{\nu},k}} \frac{\sqrt{\boldsymbol{\tau}!}}{\sqrt{\boldsymbol{\nu}!}}\frac{(\rho^{1+\eta}c^\eta)^{(\boldsymbol{\tau}-\boldsymbol{\nu})}|\psi|^{\boldsymbol{\tau}-\boldsymbol{\nu}}}{(\boldsymbol{\tau}-\boldsymbol{\nu})!}c^{-\eta}\sup_{j>s}\{\rho_j^{-\eta}\},$$

where we used that for every $\boldsymbol{\tau}\in R^s_{\boldsymbol{\nu},k}$ there exists $j > s$ such that $\tau_j - \nu_j > 0$ and that $\rho_j^{-1}/c\leq 1$, $j\geq 1$. By the first statement of Lemma 4.1,

$$\sum_{\boldsymbol{\tau}\in R^s_{\boldsymbol{\nu},k}} \frac{\sqrt{\boldsymbol{\tau}!}}{\sqrt{\boldsymbol{\nu}!}}\frac{(\rho^{1+\eta}c^\eta)^{(\boldsymbol{\tau}-\boldsymbol{\nu})}|\psi|^{\boldsymbol{\tau}-\boldsymbol{\nu}}}{(\boldsymbol{\tau}-\boldsymbol{\nu})!} \leq \sum_{\boldsymbol{\tau}\in R_{\boldsymbol{\nu},k}} \frac{\sqrt{\boldsymbol{\tau}!}}{\sqrt{\boldsymbol{\nu}!}}\frac{(\rho^{1+\eta}c^\eta)^{(\boldsymbol{\tau}-\boldsymbol{\nu})}|\psi|^{\boldsymbol{\tau}-\boldsymbol{\nu}}}{(\boldsymbol{\tau}-\boldsymbol{\nu})!} \leq \frac{(\sqrt{r}\widetilde{K}_\eta c^\eta)^{k-\ell}}{(k-\ell)!},$$

which implies the assertion of the lemma. ∎

Theorem 4.3. *Let the assumptions of Lemmas 4.1 and 4.2 be satisfied for a positive sequence* $(\rho_j)_{j\geq 1}$ *and* $\eta > 0$. *There exists a constant* $C > 0$ *such that for every* $s \in \mathbb{N}$ *and every* $\boldsymbol{y} \in U$

$$\sum_{\boldsymbol{\tau}\in\mathcal{F},\|\boldsymbol{\tau}\|_{\ell^\infty}\leq r} \frac{\rho^{2\boldsymbol{\tau}}}{\boldsymbol{\tau}!}\|\partial^{\boldsymbol{\tau}}(u(\boldsymbol{y}) - u^s(\boldsymbol{y}))\|^2_{a(\boldsymbol{y})} \leq C\left(\left\|\frac{a(\boldsymbol{y}) - a^s(\boldsymbol{y})}{a(\boldsymbol{y})}\right\|^2_{L^\infty(D)} + \sup_{j>s}\{\rho_j^{-2\eta}\}\right)\|u^s(\boldsymbol{y})\|^2_{a(\boldsymbol{y})}.$$

Proof. We divide the index set $\mathcal{F}_r := \{\boldsymbol{\tau} \in \mathcal{F} : \tau_j \leq r, j \in \mathbb{N}\}$ into $\mathcal{F}_1^s := \{\boldsymbol{\tau} \in \mathcal{F}_r : \tau_j = 0 \forall j > s\}$ and $\mathcal{F}_2^s := \{\boldsymbol{\tau} \in \mathcal{F}_r : \exists j > s \text{ s.t. } \tau_j > 0\}$. Obviously, $\mathcal{F}_r = \mathcal{F}_1^s \cup \mathcal{F}_2^s$.

Let $\mathbf{0} \neq \boldsymbol{\tau} \in \mathcal{F}_1^s$ be arbitrary. We observe that for every $v \in V$,

$$\begin{aligned}\int_D a\nabla\partial^{\boldsymbol{\tau}}(u - u^s)\cdot\nabla v \mathrm{d}x = &-\sum_{\boldsymbol{\nu}\leq\boldsymbol{\tau},\boldsymbol{\nu}\neq\boldsymbol{\tau}}\binom{\boldsymbol{\tau}}{\boldsymbol{\nu}}\int_D \psi^{\boldsymbol{\tau}-\boldsymbol{\nu}} a\nabla\partial^{\boldsymbol{\nu}}(u - u^s)\cdot\nabla v\mathrm{d}x\\ &-\sum_{\boldsymbol{\nu}\leq\boldsymbol{\tau}}\binom{\boldsymbol{\tau}}{\boldsymbol{\nu}}\int_D \psi^{\boldsymbol{\tau}-\boldsymbol{\nu}}(a - a^s)\nabla\partial^{\boldsymbol{\nu}}u^s\cdot\nabla v\mathrm{d}x.\end{aligned}$$

Set

$$\sigma_k := \sum_{\boldsymbol{\tau}\in\Lambda_k\cap\mathcal{F}_1^s}\frac{\rho^{2\boldsymbol{\tau}}}{\boldsymbol{\tau}!}\|\partial^{\boldsymbol{\tau}}(u - u^s)\|_a^2$$

and take $v = \partial^{\boldsymbol{\tau}}(u - u^s)$. By a twofold application of the Cauchy–Schwarz inequality and by Lemma 4.1

$$\begin{aligned}\sigma_k &\leq \int_D\sum_{\boldsymbol{\tau}\in\Lambda_k\cap\mathcal{F}_1^s}\sum_{\boldsymbol{\nu}\leq\boldsymbol{\tau},\boldsymbol{\nu}\neq\boldsymbol{\tau}} a\kappa_0(\boldsymbol{\tau},\boldsymbol{\nu})\frac{\rho^{\boldsymbol{\nu}}}{\sqrt{\boldsymbol{\nu}!}}|\nabla\partial^{\boldsymbol{\nu}}(u - u^s)|\frac{\rho^{\boldsymbol{\tau}}}{\sqrt{\boldsymbol{\tau}!}}|\nabla\partial^{\boldsymbol{\tau}}(u - u^s)|\mathrm{d}x\\
&\quad+\int_D\sum_{\boldsymbol{\tau}\in\Lambda_k\cap\mathcal{F}_1^s}\sum_{\boldsymbol{\nu}\leq\boldsymbol{\tau}}|a - a^s|\kappa_0(\boldsymbol{\tau},\boldsymbol{\nu})\frac{\rho^{\boldsymbol{\nu}}}{\sqrt{\boldsymbol{\nu}!}}|\nabla\partial^{\boldsymbol{\nu}}u^s|\frac{\rho^{\boldsymbol{\tau}}}{\sqrt{\boldsymbol{\tau}!}}|\nabla\partial^{\boldsymbol{\tau}}(u - u^s)|\mathrm{d}x\\
&\leq\int_D\left(\sum_{\boldsymbol{\tau}\in\Lambda_k\cap\mathcal{F}_1^s}\sum_{\boldsymbol{\nu}\leq\boldsymbol{\tau},\boldsymbol{\nu}\neq\boldsymbol{\tau}} a\kappa_0(\boldsymbol{\tau},\boldsymbol{\nu})\frac{\rho^{2\boldsymbol{\nu}}}{\boldsymbol{\nu}!}|\nabla\partial^{\boldsymbol{\nu}}(u - u^s)|^2\right)^{1/2}\left(a\sum_{\boldsymbol{\tau}\in\Lambda_k\cap\mathcal{F}_1^s}\frac{\rho^{2\boldsymbol{\tau}}}{\boldsymbol{\nu}!}|\nabla\partial^{\boldsymbol{\tau}}(u - u^s)|^2\right)^{1/2}\mathrm{d}x\\
&\quad+\int_D\left(\sum_{\boldsymbol{\tau}\in\Lambda_k\cap\mathcal{F}_1^s}\sum_{\boldsymbol{\nu}\leq\boldsymbol{\tau}}|a - a^s|\kappa_0(\boldsymbol{\tau},\boldsymbol{\nu})\frac{\rho^{2\boldsymbol{\nu}}}{\boldsymbol{\nu}!}|\nabla\partial^{\boldsymbol{\nu}}u^s|^2\right)^{1/2}\left(2\sum_{\boldsymbol{\tau}\in\Lambda_k\cap\mathcal{F}_1^s}|a - a^s|\frac{\rho^{2\boldsymbol{\tau}}}{\boldsymbol{\tau}!}|\nabla\partial^{\boldsymbol{\tau}}(u - u^s)|^2\right)^{1/2}\mathrm{d}x\end{aligned}$$

Further, we apply the Cauchy–Schwarz inequality on the integral and obtain that

$$\begin{aligned}\sigma_k &\leq\left(\int_D\sum_{\boldsymbol{\tau}\in\Lambda_k\cap\mathcal{F}_1^s}\sum_{\boldsymbol{\nu}\leq\boldsymbol{\tau},\boldsymbol{\nu}\neq\boldsymbol{\tau}} a\kappa_0(\boldsymbol{\tau},\boldsymbol{\nu})\frac{\rho^{2\boldsymbol{\nu}}}{\boldsymbol{\nu}!}|\nabla\partial^{\boldsymbol{\nu}}(u - u^s)|^2\right)^{1/2}\sqrt{\sigma_k}\\
&\quad+\left(\int_D\sum_{\boldsymbol{\tau}\in\Lambda_k\cap\mathcal{F}_1^s}\sum_{\boldsymbol{\nu}\leq\boldsymbol{\tau}}|a - a^s|\kappa_0(\boldsymbol{\tau},\boldsymbol{\nu})\frac{\rho^{2\boldsymbol{\nu}}}{\boldsymbol{\nu}!}|\nabla\partial^{\boldsymbol{\nu}}u^s|^2\right)^{1/2}\sqrt{2\left\|\frac{a - a^s}{a}\right\|_{L^\infty(D)}}\sqrt{\sigma_k}.\end{aligned}$$

By equation (4.18) of [7] in the proof of Theorem 4.1 from [7], for any $\delta \in [\sqrt{r}\widetilde{K}_0/\log(2), 1)$ and for every $\ell \in \mathbb{N}$,

$$\sum_{\boldsymbol{\tau}\in\Lambda_\ell}\frac{\rho^{2\boldsymbol{\tau}}}{\boldsymbol{\tau}!}\|\partial^{\boldsymbol{\tau}}u^s\|_a^2 \leq \|u^s\|_a^2\,\delta^\ell. \tag{4.3}$$

We change the order of summation in order to apply the second estimate in Lemma 4.1 and insert (4.3) to obtain with Young's inequality that for any $\varepsilon > 0$

$$\begin{aligned}\sigma_k &\le (1+\varepsilon)\sum_{\ell=0}^{k-1}\frac{(\sqrt{r}\widetilde{K}_0)^{k-\ell}}{(k-\ell)!}\sigma_\ell + \left(1+\frac{1}{\varepsilon}\right)2\left\|\frac{a-a^s}{a}\right\|_{L^\infty(D)}^2 \sum_{\ell=0}^{k}\frac{(\sqrt{r}\widetilde{K}_0)^{k-\ell}}{(k-\ell)!}\sum_{\boldsymbol{\tau}\in\Lambda_\ell}\frac{\rho^{2\boldsymbol{\tau}}}{\boldsymbol{\tau}!}\|\partial^{\boldsymbol{\tau}}u^s\|_a^2 \\ &\le (1+\varepsilon)\sum_{\ell=0}^{k-1}\frac{(\sqrt{r}\widetilde{K}_0)^{k-\ell}}{(k-\ell)!}\sigma_\ell + \left(1+\frac{1}{\varepsilon}\right)2\left\|\frac{a-a^s}{a}\right\|_{L^\infty(D)}^2 \|u^s\|_a^2\sum_{\ell=0}^{k}\frac{(\sqrt{r}\widetilde{K}_0)^{k-\ell}}{(k-\ell)!}\delta^\ell \\ &\le (1+\varepsilon)\sum_{\ell=0}^{k-1}\frac{(\sqrt{r}\widetilde{K}_0)^{k-\ell}}{(k-\ell)!}\sigma_\ell + \left(1+\frac{1}{\varepsilon}\right)4\left\|\frac{a-a^s}{a}\right\|_{L^\infty(D)}^2 \|u^s\|_a^2\,\delta^k.\end{aligned}$$

By a change of the order of summation, we obtain that

$$\sum_{k\ge1}\sum_{\ell=0}^{k-1}\frac{(\sqrt{r}\widetilde{K}_0)^{k-\ell}}{(k-\ell)!}\sigma_\ell = \sum_{\ell\ge0}\left(\sum_{k=\ell+1}^{\infty}\frac{(\sqrt{r}\widetilde{K}_0)^{k-\ell}}{(k-\ell)!}\right)\sigma_\ell \le (e^{\sqrt{r}\widetilde{K}_0}-1)\sum_{\ell\ge0}\sigma_\ell. \tag{4.4}$$

Let us choose $\varepsilon > 0$ such that $\varepsilon < (2-e^{\sqrt{r}\widetilde{K}_0})/(e^{\sqrt{r}\widetilde{K}_0}-1)$, which implies that $(1+\varepsilon)(e^{\sqrt{r}\widetilde{K}_0}-1) < 1$. Denote $C^* := (1-(1+\varepsilon)(e^{\sqrt{r}\widetilde{K}_0}-1))^{-1}$. We sum σ_k over $k \ge 1$ and obtain that

$$\sum_{k\ge1}\sigma_k \le (1+\varepsilon)(e^{\sqrt{r}\widetilde{K}_0}-1)\sum_{\ell\ge0}\sigma_\ell + \left(1+\frac{1}{\varepsilon}\right)4\left\|\frac{a-a^s}{a}\right\|_{L^\infty(D)}^2\|u^s\|_a^2\frac{\delta}{1-\delta}.$$

Since $(1+\varepsilon)(e^{\sqrt{r}\widetilde{K}_0}-1) < 1$, we conclude that

$$\sum_{k\ge1}\sigma_k \le C^*\sigma_0 + C^*\left(1+\frac{1}{\varepsilon}\right)4\left\|\frac{a-a^s}{a}\right\|_{L^\infty(D)}^2\|u^s\|_a^2\frac{\delta}{1-\delta},$$

which implies

$$\sum_{\boldsymbol{\tau}\in\mathcal{F}_1^s}\frac{\rho^{2\boldsymbol{\tau}}}{\boldsymbol{\tau}!}\|\partial^{\boldsymbol{\tau}}(u-u^s)\|_a^2 \le C\left(\|u-u^s\|_a^2 + \left\|\frac{a-a^s}{a}\right\|_{L^\infty(D)}^2\|u^s\|_a^2\right).$$

In the other case $\boldsymbol{\tau}\in\mathcal{F}_2^s$, we observe that for arbitrary $\mathbf{0}\ne\boldsymbol{\tau}\in\mathcal{F}_2^s$,

$$\begin{aligned}\int_D a\nabla\partial^{\boldsymbol{\tau}}(u-u^s)\cdot\nabla v\mathrm{d}x = &-\sum_{\boldsymbol{\nu}\le\boldsymbol{\tau},\boldsymbol{\nu}\ne\boldsymbol{\tau}}\binom{\boldsymbol{\tau}}{\boldsymbol{\nu}}\int_D\psi^{\boldsymbol{\tau}-\boldsymbol{\nu}}a\nabla\partial^{\boldsymbol{\nu}}(u-u^s)\cdot\nabla v\mathrm{d}x \\ &-\sum_{\boldsymbol{\nu}\le\boldsymbol{\tau}}\binom{\boldsymbol{\tau}}{\boldsymbol{\nu}}\int_D\psi^{\boldsymbol{\tau}-\boldsymbol{\nu}}a\nabla\partial^{\boldsymbol{\nu}}u^s\cdot\nabla v\mathrm{d}x, \qquad \forall v\in V.\end{aligned} \tag{4.5}$$

We used that there is $j > s$ such that $\tau_j > 0$, which implies that for $\boldsymbol{\nu}\ne\boldsymbol{\tau}$ such that $\boldsymbol{\nu}\le\boldsymbol{\tau}$, either $\tau_j-\nu_j > 0$ yielding $\partial^{\boldsymbol{\tau}-\boldsymbol{\nu}}a^s = 0$ or $\tau_j = \nu_j > 0$ yielding $\partial^{\boldsymbol{\nu}}u^s = 0$. Moreover, in the second sum above, we can restrict the index set to those $\boldsymbol{\nu}$ satisfying $\nu_j = 0$ for every $j > s$. In particular, always $\boldsymbol{\nu}\ne\boldsymbol{\tau}$. The estimate of the sum over $\boldsymbol{\tau}\in\mathcal{F}_2^s$ follows with a similar argument using Lemma 4.1 for the first sum and Lemma 4.2 for the second sum of the right hand side of equality (4.5), where we crucially use that $\boldsymbol{\nu}\ne\boldsymbol{\tau}$, which yields that the sum runs only over $\ell\in\{0,\dots,k-1\}$. Specifically,

$$\sum_{\boldsymbol{\tau}\in\mathcal{F}_2^s}\frac{\rho^{2\boldsymbol{\tau}}}{\boldsymbol{\tau}!}\|\partial^{\boldsymbol{\tau}}(u-u^s)\|_a^2 \le C\left(\|u-u^s\|_a^2 + \max_{j>s}\{\rho_j^{-2\eta}\}\|u^s\|_a^2\right).$$

Since by (4.1) and by the Cauchy–Schwarz inequality

$$\|u-u^s\|_a \le \left\|\frac{a-a^s}{a}\right\|_{L^\infty(D)} \|u^s\|_a,$$

the assertion of the theorem follows. □

Remark 4.4. The statement of Theorem 4.3 also holds true for the FE solution $u^{\mathcal{T}_\ell}$ and $u^{s,\mathcal{T}_\ell}$ for every truncation dimension $s \in \mathbb{N}$.

4.2. Discretization

First we show parametric regularity estimates of the solution u. Thus, we bound weighted sums over sensitivities of u in the norm of the smoothness space. For every $\boldsymbol{\tau} \in \mathcal{F}$, we define the quantities

$$\kappa_1(\boldsymbol{\tau},\boldsymbol{\nu}) := \frac{\sqrt{\boldsymbol{\tau}!}}{\sqrt{\boldsymbol{\nu}!}} \frac{\boldsymbol{\rho}^{\boldsymbol{\tau}-\boldsymbol{\nu}} |\nabla \psi^{\boldsymbol{\tau}-\boldsymbol{\nu}}|\Phi_{\overline{\boldsymbol{\beta}}}}{(\boldsymbol{\tau}-\boldsymbol{\nu})!}, \quad \boldsymbol{\nu} \le \boldsymbol{\tau}.$$

Lemma 4.5. *Assume that for $r \in \mathbb{N}$*

$$\widetilde{K}_1 := \left\| \sum_{j\ge 1} \rho_j \max\{|\nabla\psi_j|\Phi_{\overline{\boldsymbol{\beta}}}, |\psi_j|\} \right\|_{L^\infty(D)} < C_r := \sup\left\{c>0 : \sqrt{r}c\, e^{\sqrt{r}c} \le 1\right\}. \tag{4.6}$$

Then for every $\boldsymbol{\tau} \in \mathbb{N}_0^{\mathbb{N}}$ such that $\|\boldsymbol{\tau}\|_{\ell^\infty} \le r$

$$\sum_{\boldsymbol{\nu}\le\boldsymbol{\tau},\boldsymbol{\nu}\ne\boldsymbol{\tau}} \kappa_1(\boldsymbol{\tau},\boldsymbol{\nu}) \le \sqrt{r}\widetilde{K}_1 e^{\sqrt{r}\widetilde{K}_1} < 1$$

and for every $\ell \le k-1$ and $\boldsymbol{\nu} \in \Lambda_\ell$,

$$\sum_{\boldsymbol{\tau}\in R_{\boldsymbol{\nu},k}} \kappa_1(\boldsymbol{\tau},\boldsymbol{\nu}) \le (k-\ell)\frac{(\sqrt{r}\widetilde{K}_1)^{k-\ell}}{(k-\ell)!}.$$

Proof. We set $k=|\boldsymbol{\tau}|$ and observe with the multinomial theorem

$$\begin{aligned}
\sum_{\boldsymbol{\nu}\le\boldsymbol{\tau},\boldsymbol{\nu}\ne\boldsymbol{\tau}} \kappa_1(\boldsymbol{\tau},\boldsymbol{\nu}) &= \sum_{\ell=1}^{k} \sum_{\boldsymbol{\nu}\le\boldsymbol{\tau},|\boldsymbol{\tau}-\boldsymbol{\nu}|=\ell} \kappa_1(\boldsymbol{\tau},\boldsymbol{\nu}) \\
&\le \sum_{\ell=1}^{k} r^{\ell/2} \sum_{\boldsymbol{\nu}\le\boldsymbol{\tau},|\boldsymbol{\tau}-\boldsymbol{\nu}|=\ell} \ell \frac{\boldsymbol{\rho}^{\boldsymbol{\tau}-\boldsymbol{\nu}} \max\{|\nabla\psi|\Phi_{\overline{\boldsymbol{\beta}}}, |\psi|\}^{\boldsymbol{\tau}-\boldsymbol{\nu}}}{(\boldsymbol{\tau}-\boldsymbol{\nu})!} \\
&\le \sum_{\ell=1}^{k} r^{\ell/2} \ell \sum_{|\boldsymbol{m}|=\ell} \frac{\boldsymbol{\rho}^{\boldsymbol{m}} \max\{|\nabla\psi|\Phi_{\overline{\boldsymbol{\beta}}}, |\psi|\}^{\boldsymbol{m}}}{\boldsymbol{m}!} \\
&= \sum_{\ell=1}^{k} \frac{r^{\ell/2}}{(\ell-1)!} \left(\sum_{j\ge1} \rho_j \max\{|\nabla\psi_j|\Phi_{\overline{\boldsymbol{\beta}}}, |\psi_j|\}\right)^{\ell} \le \sqrt{r}\widetilde{K}_1 e^{\sqrt{r}\widetilde{K}_1} < 1,
\end{aligned}$$

where we applied that

$$|\nabla\psi^{\boldsymbol{\tau}-\boldsymbol{\nu}}|\Phi_{\overline{\boldsymbol{\beta}}} \le \sum_{j\ge1} (\tau_j-\nu_j)|\psi_j|^{\tau_j-\nu_j-1}|\nabla\psi_j|\Phi_{\overline{\boldsymbol{\beta}}} \prod_{i\ne j} |\psi_i|^{\tau_i-\nu_i} \le |\boldsymbol{\tau}-\boldsymbol{\nu}| \max\{|\nabla\psi|\Phi_{\overline{\boldsymbol{\beta}}}, |\psi|\}^{\boldsymbol{\tau}-\boldsymbol{\nu}}.$$

The second estimate follows similarly. □

Theorem 4.6. *Let the assumption of Lemma 4.5 be satisfied for a positive sequence* $(\rho_j)_{j\geq 1}$*, and assume that* $r \in \mathbb{N}$ *and* $\widetilde{K}_1 < C_r$*. There exists a constant* $C > 0$ *such that for every* $\boldsymbol{y} \in U$

$$\sum_{\boldsymbol{\tau}\in\mathcal{F},\|\boldsymbol{\tau}\|_{\ell^\infty}\leq r} \frac{\rho^{2\boldsymbol{\tau}}}{\boldsymbol{\tau}!}\|\Delta\partial^{\boldsymbol{\tau}} u(\boldsymbol{y}))\|^2_{L^2_{\overline{\boldsymbol{\beta}}}(D)} \leq C\left(\frac{1}{a_{\min}(\boldsymbol{y})}(1+\||\nabla Z(\boldsymbol{y})|\Phi_{\overline{\boldsymbol{\beta}}}\|^2_{L^\infty(D)})\|u(\boldsymbol{y})\|^2_{a(\boldsymbol{y})} + \|\Delta u(\boldsymbol{y})\|^2_{L^2_{\overline{\boldsymbol{\beta}}}(D)}\right).$$

Proof. Let $\mathbf{0}\neq\boldsymbol{\tau}\in\mathcal{F}$ be given such that $\|\boldsymbol{\tau}\|_{\ell^\infty}\leq r$. We observe that for every $v\in C_0^\infty(D)$,

$$-\int_D av\Delta\partial^{\boldsymbol{\tau}} u \mathrm{d}x = \int_D\left(\nabla a\cdot\nabla\partial^{\boldsymbol{\tau}} u + \sum_{\boldsymbol{\nu}\leq\boldsymbol{\tau},\boldsymbol{\nu}\neq\boldsymbol{\tau}}\binom{\boldsymbol{\tau}}{\boldsymbol{\nu}}(\nabla\partial^{\boldsymbol{\tau}-\boldsymbol{\nu}}a\cdot\nabla\partial^{\boldsymbol{\nu}}u + \partial^{\boldsymbol{\tau}-\boldsymbol{\nu}}a\Delta\partial^{\boldsymbol{\nu}}u)\right)v\mathrm{d}x.$$

Using the density of $C_0^\infty(D)$ in $L^2_{\overline{\boldsymbol{\beta}}}(D)$, we choose the test function $v = -\Phi^2_{\overline{\boldsymbol{\beta}}}/a\Delta\partial^{\boldsymbol{\tau}}u$, multiply by $\rho^{2\boldsymbol{\tau}}/\boldsymbol{\tau}!$, and apply the Young inequality for arbitrary $\varepsilon > 0$ to obtain

$$\begin{aligned}\frac{\rho^{2\boldsymbol{\tau}}}{\boldsymbol{\tau}!}\|\Delta\partial^{\boldsymbol{\tau}}u\|^2_{L^2_{\overline{\boldsymbol{\beta}}}(D)} &= -\frac{\rho^{2\boldsymbol{\tau}}}{\boldsymbol{\tau}!}\int_D\left(\frac{\nabla a}{a}\cdot\nabla\partial^{\boldsymbol{\tau}}u + \sum_{\boldsymbol{\nu}\leq\boldsymbol{\tau},\boldsymbol{\nu}\neq\boldsymbol{\tau}}\binom{\boldsymbol{\tau}}{\boldsymbol{\nu}}\frac{\nabla\partial^{\boldsymbol{\tau}-\boldsymbol{\nu}}a}{a}\cdot\nabla\partial^{\boldsymbol{\nu}}u\right)\Delta\partial^{\boldsymbol{\tau}}u\Phi^2_{\overline{\boldsymbol{\beta}}}\mathrm{d}x\\ &\quad -\frac{\rho^{2\boldsymbol{\tau}}}{\boldsymbol{\tau}!}\int_D\left(\sum_{\boldsymbol{\nu}\leq\boldsymbol{\tau},\boldsymbol{\nu}\neq\boldsymbol{\tau}}\binom{\boldsymbol{\tau}}{\boldsymbol{\nu}}\frac{\partial^{\boldsymbol{\tau}-\boldsymbol{\nu}}a}{a}\Delta\partial^{\boldsymbol{\nu}}u\right)\Delta\partial^{\boldsymbol{\tau}}u\Phi^2_{\overline{\boldsymbol{\beta}}}\mathrm{d}x\\ &\leq \varepsilon\frac{\rho^{2\boldsymbol{\tau}}}{\boldsymbol{\tau}!}\|\Delta\partial^{\boldsymbol{\tau}}u\|^2_{L^2_{\overline{\boldsymbol{\beta}}}(D)} + \frac{1}{4\varepsilon}\int_D\left(|\nabla Z|\Phi_{\overline{\boldsymbol{\beta}}}\frac{\rho^{\boldsymbol{\tau}}|\nabla\partial^{\boldsymbol{\tau}}u|}{\sqrt{\boldsymbol{\tau}!}} + \sum_{\boldsymbol{\nu}\leq\boldsymbol{\tau},\boldsymbol{\nu}\neq\boldsymbol{\tau}}\kappa_1(\boldsymbol{\tau},\boldsymbol{\nu})\frac{\rho^{\boldsymbol{\nu}}|\nabla\partial^{\boldsymbol{\nu}}u|}{\sqrt{\boldsymbol{\nu}!}}\right)^2\mathrm{d}x\\ &\quad + \int_D\sum_{\boldsymbol{\nu}\leq\boldsymbol{\tau},\boldsymbol{\nu}\neq\boldsymbol{\tau}}\kappa_0(\boldsymbol{\tau},\boldsymbol{\nu})\frac{\rho^{\nu}|\Delta\partial^{\boldsymbol{\nu}}u|\Phi_{\overline{\boldsymbol{\beta}}}}{\sqrt{\boldsymbol{\nu}!}}\frac{\rho^{\boldsymbol{\tau}}|\Delta\partial^{\boldsymbol{\tau}}u|\Phi_{\overline{\boldsymbol{\beta}}}}{\sqrt{\boldsymbol{\tau}!}}\mathrm{d}x.\end{aligned} \tag{4.7}$$

Note that $\nabla((\partial^{\boldsymbol{\tau}-\boldsymbol{\nu}}a)/a) = \nabla\psi^{\boldsymbol{\tau}-\boldsymbol{\nu}}$. Note also the change of the order of summation: for any sequence $(\kappa'(\boldsymbol{\tau},\boldsymbol{\nu}))$ and for any $k\in\mathbb{N}$

$$\sum_{\boldsymbol{\tau}\in\Lambda_k}\sum_{\boldsymbol{\nu}\leq\boldsymbol{\tau},\boldsymbol{\nu}\neq\boldsymbol{\tau}}\kappa'(\boldsymbol{\tau},\boldsymbol{\nu}) = \sum_{\ell=0}^{k-1}\sum_{\boldsymbol{\nu}\in\Lambda_\ell}\sum_{\boldsymbol{\tau}\in R_{\boldsymbol{\nu},k}}\kappa'(\boldsymbol{\tau},\boldsymbol{\nu}), \tag{4.8}$$

which implies with Lemma 4.1 and with the elementary estimate $xy\leq(x^2+y^2)/2$, $x,y>0$, that for any $k\geq 1$,

$$\begin{aligned}&\sum_{\boldsymbol{\tau}\in\Lambda_k}\int_D\sum_{\boldsymbol{\nu}\leq\boldsymbol{\tau},\boldsymbol{\nu}\neq\boldsymbol{\tau}}\kappa_0(\boldsymbol{\tau},\boldsymbol{\nu})\frac{\rho^{\nu}|\Delta\partial^{\boldsymbol{\nu}}u|\Phi_{\overline{\boldsymbol{\beta}}}}{\sqrt{\boldsymbol{\nu}!}}\frac{\rho^{\boldsymbol{\tau}}|\Delta\partial^{\boldsymbol{\tau}}u|\Phi_{\overline{\boldsymbol{\beta}}}}{\sqrt{\boldsymbol{\tau}!}}\mathrm{d}x\\ &\leq\frac{1}{2}\sum_{\ell=0}^{k-1}\frac{(\sqrt{r}\widetilde{K}_1)^{k-\ell}}{(k-\ell)!}\sum_{\boldsymbol{\nu}\in\Lambda_\ell}\frac{\rho^{2\boldsymbol{\nu}}}{\boldsymbol{\nu}!}\|\Delta\partial^{\boldsymbol{\nu}}u\|^2_{L^2_{\overline{\boldsymbol{\beta}}}(D)} + \frac{1}{2}(e^{\sqrt{r}\widetilde{K}_1}-1)\sum_{\boldsymbol{\tau}\in\Lambda_k}\frac{\rho^{2\boldsymbol{\tau}}}{\boldsymbol{\tau}!}\|\Delta\partial^{\boldsymbol{\tau}}u\|^2_{L^2_{\overline{\boldsymbol{\beta}}}(D)}.\end{aligned} \tag{4.9}$$

Note that the assumptions of Lemma 4.1 are weaker than those of Lemma 4.5 and $\widetilde{K}_0 \leq \widetilde{K}_1$. Similarly, we obtain with Lemma 4.5

$$\begin{aligned}\sum_{\boldsymbol{\tau}\in\Lambda_k} \frac{1}{4\varepsilon}\int_D \left(|\nabla Z|\Phi_{\overline{\boldsymbol{\beta}}} \frac{\rho^{\boldsymbol{\tau}}|\nabla\partial^{\boldsymbol{\tau}}u|}{\sqrt{\boldsymbol{\tau}!}} + \sum_{\boldsymbol{\nu}\leq\boldsymbol{\tau},\boldsymbol{\nu}\neq\boldsymbol{\tau}} \kappa_1(\boldsymbol{\tau},\boldsymbol{\nu}) \frac{\rho^{\boldsymbol{\nu}}|\nabla\partial^{\boldsymbol{\nu}}u|}{\sqrt{\boldsymbol{\nu}!}} \right)^2 \mathrm{d}x \\ \leq \frac{1}{2\varepsilon}\frac{\||\nabla Z|\Phi_{\overline{\boldsymbol{\beta}}}\|^2_{L^\infty(D)}}{a_{\min}} \sum_{\boldsymbol{\tau}\in\Lambda_k}\frac{\rho^{2\boldsymbol{\tau}}}{\boldsymbol{\tau}!}\|\partial^{\boldsymbol{\tau}}u\|_a^2 + \frac{1}{2\varepsilon}\sum_{\boldsymbol{\tau}\in\Lambda_k}\int_D \left(\sum_{\boldsymbol{\nu}\leq\boldsymbol{\tau},\boldsymbol{\nu}\neq\boldsymbol{\tau}} \kappa_1(\boldsymbol{\tau},\boldsymbol{\nu}) \frac{\rho^{\boldsymbol{\nu}}|\nabla\partial^{\boldsymbol{\nu}}u|}{\sqrt{\boldsymbol{\nu}!}}\right)^2 \mathrm{d}x \qquad (4.10) \\ \leq \frac{1}{2\varepsilon}\frac{1}{a_{\min}}\left(\||\nabla Z|\Phi_{\overline{\boldsymbol{\beta}}}\|^2_{L^\infty(D)} \sum_{\boldsymbol{\tau}\in\Lambda_k}\frac{\rho^{2\boldsymbol{\tau}}}{\boldsymbol{\tau}!}\|\partial^{\boldsymbol{\tau}}u\|_a^2 + \sum_{\ell=0}^{k-1}\frac{(\sqrt{r}\widetilde{K}_1)^{k-\ell}}{(k-\ell-1)!}\sum_{\boldsymbol{\nu}\in\Lambda_\ell}\frac{\rho^{2\boldsymbol{\nu}}}{\boldsymbol{\nu}!}\|\partial^{\boldsymbol{\nu}}u\|_a^2\right).\end{aligned}$$

As before by the proof of Theorem 4.1 and equation (4.18) of [7], for any $\delta \in [\sqrt{r}\widetilde{K}_1/\log(2), 1)$ and for every $\ell \in \mathbb{N}_0$,

$$\sum_{\boldsymbol{\nu}\in\Lambda_\ell}\frac{\rho^{2\boldsymbol{\nu}}}{\boldsymbol{\nu}!}\|\partial^{\boldsymbol{\nu}}u\|_a^2 \leq \delta^\ell \|u\|_a^2. \qquad (4.11)$$

Hence,

$$\begin{aligned}\sum_{\ell=0}^{k-1}\frac{(\sqrt{r}\widetilde{K}_1)^{k-\ell}}{(k-\ell-1)!}\sum_{\boldsymbol{\nu}\in\Lambda_\ell}\frac{\rho^{2\boldsymbol{\nu}}}{\boldsymbol{\nu}!}\|\partial^{\boldsymbol{\nu}}u\|_a^2 \leq \sum_{\ell=0}^{k-1}\frac{(\sqrt{r}\widetilde{K}_1)^{k-\ell}}{(k-\ell-1)!}\delta^\ell\|u\|_a^2 \leq \delta^k\sum_{\ell=0}^{k-1}\log(2)\frac{(\log(2))^{k-\ell-1}}{(k-\ell-1)!}\|u\|_a^2 \qquad (4.12) \\ \leq \delta^k\log(2)2\|u\|_a^2 = \delta^k\log(4)\|u\|_a^2.\end{aligned}$$

We choose $0 < \varepsilon < 1 - e^{\sqrt{r}\widetilde{K}_1}/2$, which implies that $C_\varepsilon := (1-\varepsilon-(e^{\sqrt{r}\widetilde{K}_1}-1)/2)^{-1} < 2$. This allows us to subtract $\Delta\partial^{\boldsymbol{\tau}}u$-terms summed over Λ_k in (4.7) and (4.9) while obtaining a constant $C_\varepsilon^{-1} > 1/2$ which is shifted to the left hand side, *i.e.*,

$$\begin{aligned}\sum_{\boldsymbol{\tau}\in\Lambda_k}\frac{\rho^{2\boldsymbol{\tau}}}{\boldsymbol{\tau}!}\|\Delta\partial^{\boldsymbol{\tau}}u\|^2_{L^2_{\overline{\boldsymbol{\beta}}}(D)} \leq \frac{C_\varepsilon}{2\varepsilon}\frac{1}{a_{\min}}\left(\||\nabla Z|\Phi_{\overline{\boldsymbol{\beta}}}\|^2_{L^\infty(D)} + \log(4)\right)\delta^k\|u\|_a^2 \\ + \frac{C_\varepsilon}{2}\sum_{\ell=0}^{k-1}\frac{(\sqrt{r}\widetilde{K}_1)^{k-\ell}}{(k-\ell)!}\sum_{\boldsymbol{\nu}\in\Lambda_\ell}\frac{\rho^{2\boldsymbol{\nu}}}{\boldsymbol{\nu}!}\|\Delta\partial^{\boldsymbol{\nu}}u\|^2_{L^2_{\overline{\boldsymbol{\beta}}}(D)},\end{aligned}$$

where we have also inserted (4.10), (4.11) and (4.12). We sum over $k \geq 1$ and obtain with (4.4)

$$\begin{aligned}\sum_{k\geq 1}\sum_{\boldsymbol{\tau}\in\Lambda_k}\frac{\rho^{2\boldsymbol{\tau}}}{\boldsymbol{\tau}!}\|\Delta\partial^{\boldsymbol{\tau}}u\|^2_{L^2_{\overline{\boldsymbol{\beta}}}(D)} \leq \frac{C_\varepsilon}{2\varepsilon}\frac{1}{a_{\min}}\left(\||\nabla Z|\Phi_{\overline{\boldsymbol{\beta}}}\|^2_{L^\infty(D)} + \log(4)\right)\frac{\delta}{1-\delta}\|u\|_a^2 \\ + \frac{C_\varepsilon}{2}(e^{\sqrt{r}\widetilde{K}_1}-1)\sum_{\ell\geq 0}\sum_{\boldsymbol{\nu}\in\Lambda_\ell}\frac{\rho^{2\nu}}{\boldsymbol{\nu}!}\|\Delta\partial^{\boldsymbol{\nu}}u\|^2_{L^2_{\overline{\boldsymbol{\beta}}}(D)},\end{aligned}$$

which implies the assertion as at the end of the proof of Theorem 4.3, since $(C_\varepsilon/2)(e^{\sqrt{r}\widetilde{K}_1}-1) < 1$. □

We remark that a related estimate to Theorem 4.6 has been derived in Theorem 6.1 of [6] without taking into account the spatial weight function.

Theorem 4.7. *Let the assumption of Lemma 4.5 be satisfied for a positive sequence $(\rho_j)_{j\geq 1}$, $r \in \mathbb{N}$, and $\widetilde{K}_1 < C_r$. There exists a constant $C > 0$ such that for every $\boldsymbol{y} \in U$*

$$\sum_{\boldsymbol{\tau}\in\mathcal{F},\|\boldsymbol{\tau}\|_{\ell^\infty}\leq r}\frac{\rho^{2\boldsymbol{\tau}}}{\boldsymbol{\tau}!}\|\partial^{\boldsymbol{\tau}}(u(\boldsymbol{y}) - u^{\mathcal{T}_\ell}(\boldsymbol{y}))\|^2_{a(\boldsymbol{y})} \leq C\left(\frac{(a_{\max}(\boldsymbol{y}))}{(a_{\min}(\boldsymbol{y}))^2}(1 + \|\nabla Z(\boldsymbol{y})|\Phi_{\overline{\boldsymbol{\beta}}}\|^2_{L^\infty(D)})\right)\|f\|^2_{L^2_{\overline{\boldsymbol{\beta}}}(D)} M_\ell^{-2/d}.$$

Proof. To simplify notation, we do not indicate in this proof the dependence of quantities on the parameters $\boldsymbol{y}$. Define the Galerkin projection $\mathcal{P}_h : V \to V_\ell$ for every $w \in V$ by

$$\int_D a\nabla(w - \mathcal{P}_h w) \cdot \nabla v \mathrm{d}x = 0, \quad \forall v \in V_\ell.$$

Since $(\mathcal{I} - \mathcal{P}_h)v = 0$ for every $v \in V_\ell$, it holds that for every $\boldsymbol{\tau} \in \mathcal{F}$,

$$\|\partial^{\boldsymbol{\tau}}(u - u^{\mathcal{T}_\ell})\|_a \le \|\mathcal{P}_h \partial^{\boldsymbol{\tau}}(u - u^{\mathcal{T}_\ell})\|_a + \|(\mathcal{I} - \mathcal{P}_h)\partial^{\boldsymbol{\tau}} u\|_a. \tag{4.13}$$

Let $\boldsymbol{\tau} \in \mathcal{F}$ be such that $\|\boldsymbol{\tau}\|_{\ell^\infty(\mathbb{N})} \le r$ and $|\boldsymbol{\tau}| = k$ for some $k \in \mathbb{N}$. We observe that

$$\frac{\rho^{2\boldsymbol{\tau}}}{\boldsymbol{\tau}!}\int_D a|\nabla \mathcal{P}_h \partial^{\boldsymbol{\tau}}(u - u^{\mathcal{T}_\ell})|^2 \mathrm{d}x \le \int_D \sum_{\boldsymbol{\nu}\le\boldsymbol{\tau},\boldsymbol{\nu}\ne\boldsymbol{\tau}} \kappa_0(\boldsymbol{\tau},\boldsymbol{\nu}) a \frac{\rho^{\boldsymbol{\nu}}|\nabla\partial^{\boldsymbol{\nu}}(u - u^{\mathcal{T}_\ell})|}{\sqrt{\boldsymbol{\nu}!}} \frac{\rho^{\boldsymbol{\tau}}|\nabla \mathcal{P}_h \partial^{\boldsymbol{\tau}}(u - u^{\mathcal{T}_\ell})|}{\sqrt{\boldsymbol{\tau}!}} \mathrm{d}x.$$

A twofold application of the Cauchy–Schwarz inequality using that by the first estimate of Lemma 4.1 for fixed $\boldsymbol{\tau} \in \mathcal{F}$ such that $\|\boldsymbol{\tau}\|_{\ell^\infty(\mathbb{N})} \le r$ the sum of $(\kappa_0(\boldsymbol{\tau},\boldsymbol{\nu}))_{\boldsymbol{\nu}\le\boldsymbol{\tau},\boldsymbol{\nu}\ne\boldsymbol{\tau}}$ is less than one implies with the change of the order of summation in (4.8) and the second estimate in Lemma 4.1 the bound

$$\sum_{|\boldsymbol{\tau}|=k} \frac{\rho^{2\boldsymbol{\tau}}}{\boldsymbol{\tau}!}\|\mathcal{P}_h \partial^{\boldsymbol{\tau}}(u - u^{\mathcal{T}_\ell})\|_a^2 \le \sum_{\ell=0}^{k-1} \frac{(\sqrt{r}\widetilde{K}_1)^{k-\ell}}{(k-\ell)!} \sum_{|\boldsymbol{\nu}|=\ell} \frac{\rho^{2\boldsymbol{\nu}}}{\boldsymbol{\nu}!}\|\partial^{\boldsymbol{\nu}}(u - u^{\mathcal{T}_\ell})\|_a^2. \tag{4.14}$$

By the approximation property in (2.7), by (4.13), (4.14), the Young inequality for any $\varepsilon > 0$, and by the change of the order of summation that implied (4.4)

$$\begin{aligned}
&\sum_{k\ge1}\sum_{|\boldsymbol{\tau}|=k} \frac{\rho^{2\boldsymbol{\tau}}}{\boldsymbol{\tau}!}\|\partial^{\boldsymbol{\tau}}(u - u^{\mathcal{T}_\ell})\|_a^2 \\
&\quad\le (1+\varepsilon)\sum_{k\ge1}\sum_{\ell=0}^{k-1} \frac{(\sqrt{r}\widetilde{K}_1)^{k-\ell}}{(k-\ell)!} \sum_{|\boldsymbol{\nu}|=\ell} \frac{\rho^{2\boldsymbol{\nu}}}{\boldsymbol{\nu}!}\|\partial^{\boldsymbol{\nu}}(u - u^{\mathcal{T}_\ell})\|_a^2 + \left(1 + \frac{1}{\varepsilon}\right)\sum_{k\ge1}\sum_{|\boldsymbol{\tau}|=k} \frac{\rho^{2\boldsymbol{\tau}}}{\boldsymbol{\tau}!}\|(\mathcal{I} - \mathcal{P}_h)\partial^{\boldsymbol{\tau}} u\|_a^2 \\
&\quad\le (1+\varepsilon)(e^{\sqrt{r}\widetilde{K}_1} - 1)\sum_{\ell\ge0}\sum_{|\boldsymbol{\nu}|=\ell} \frac{\rho^{2\boldsymbol{\nu}}}{\boldsymbol{\nu}!}\|\partial^{\boldsymbol{\nu}}(u - u^{\mathcal{T}_\ell})\|_a^2 + \left(1 + \frac{1}{\varepsilon}\right) C\|a\|_{L^\infty(D)} M_\ell^{-2/d} \sum_{k\ge1}\sum_{|\boldsymbol{\tau}|=k} \frac{\rho^{2\boldsymbol{\tau}}}{\boldsymbol{\tau}!}\|\Delta\partial^{\boldsymbol{\tau}} u\|_{L^2_{\overline{\boldsymbol{\beta}}}(D)}^2.
\end{aligned}$$

Hence, we choose $\varepsilon < (2 - e^{\sqrt{r}\widetilde{K}_1})/(e^{\sqrt{r}\widetilde{K}_1} - 1)$ and conclude with Theorem 4.6 and (2.11) that there exists a constant $C > 0$ such that

$$\sum_{\boldsymbol{\tau}\in\mathcal{F},\|\boldsymbol{\tau}\|_{\ell^\infty}\le r} \frac{\rho^{2\boldsymbol{\tau}}}{\boldsymbol{\tau}!}\|\partial^{\boldsymbol{\tau}}(u - u^{\mathcal{T}_\ell})\|_a^2 \le C\left(\frac{(a_{\max})}{(a_{\min})^2}(1 + \|\nabla Z|\Phi_{\overline{\boldsymbol{\beta}}}\|_{L^\infty(D)}^2)\right)\|f\|_{L^2_{\overline{\boldsymbol{\beta}}}(D)}^2 M_\ell^{-2/d}.$$

□

Remark 4.8. The parametric regularity estimate in Theorem 4.7 also holds if $f \in (V^*, L^2_{\overline{\boldsymbol{\beta}}}(D))_{t,\infty}$ for some $t \in [0,1]$ with the FE error bounded by an absolute constant times $M_\ell^{-2t/d}$. This can be shown by interpolation applied in the last and next to last step of the proof of Theorem 4.7, see also Remark 2.4.

Let $G(\cdot) \in L^2_{\overline{\boldsymbol{\beta}}}(D)$ denote a solution functional of interest which is deterministic, *i.e.*, which does not depend on $\boldsymbol{y}$. To analyze the parametric regularity of $G(u - u^{\mathcal{T}_\ell})$, we introduce v_G and $v_G^{\mathcal{T}_\ell}$ to be the solution and respective FE solution to the adjoint problem with right hand side $G(\cdot)$. It holds that

$$G(u - u^{\mathcal{T}_\ell}) = \int_D a\nabla(u - u^{\mathcal{T}_\ell}) \cdot \nabla(v_G - v_G^{\mathcal{T}_\ell})\mathrm{d}x.$$

Proposition 4.9. *For $(\widetilde{\rho})_{j\geq 1}$ defined by $\widetilde{\rho}_j := \sqrt{2}\rho_j$, $j \in \mathbb{N}$, assume that $(\widetilde{\rho})_{j\geq 1}$ satisfies the sparsity assumption in (4.2) of Lemma 4.1. Then, for every $\boldsymbol{y} \in U$, there holds*

$$\sum_{\boldsymbol{\tau}\in\mathcal{F},\|\boldsymbol{\tau}\|_{\ell^\infty}\leq r} \frac{\rho^{2\boldsymbol{\tau}}}{\tau!}|\partial^{\boldsymbol{\tau}} G(u(\boldsymbol{y}) - u^{\mathcal{T}_\ell}(\boldsymbol{y}))|^2 \leq 4\left(\sum_{\|\boldsymbol{\tau}\|_{\ell^\infty}\leq r} \frac{\widetilde{\rho}^{2\boldsymbol{\tau}}}{\tau!}\|\partial^{\boldsymbol{\tau}}(u(\boldsymbol{y}) - u^{\mathcal{T}_\ell}(\boldsymbol{y}))\|^2_{a(\boldsymbol{y})}\right)$$
$$\times\left(\sum_{\|\boldsymbol{\tau}\|_{\ell^\infty}\leq r} \frac{\widetilde{\rho}^{2\boldsymbol{\tau}}}{\tau!}\|\partial^{\boldsymbol{\tau}}(v_G(\boldsymbol{y}) - v_G^{\mathcal{T}_\ell}(\boldsymbol{y}))\|^2_{a(\boldsymbol{y})}\right).$$

Proof. We observe that for every $\boldsymbol{\tau} \in \mathcal{F}$

$$\frac{\rho^{\boldsymbol{\tau}}}{\sqrt{\boldsymbol{\tau}!}}\partial^{\boldsymbol{\tau}} G(u - u^{\mathcal{T}_\ell}) = \frac{\rho^{\boldsymbol{\tau}}}{\sqrt{\boldsymbol{\tau}!}}\int_D \sum_{\boldsymbol{\nu}\leq\boldsymbol{\tau}} \binom{\boldsymbol{\tau}}{\boldsymbol{\nu}}\left[\sum_{\boldsymbol{m}\leq\boldsymbol{\nu}}\binom{\boldsymbol{\nu}}{\boldsymbol{m}}\psi^{\boldsymbol{\nu}-\boldsymbol{m}}(\sqrt{a}\nabla\partial^{\boldsymbol{m}}(u - u^{\mathcal{T}_\ell}))\right](\sqrt{a}\nabla\partial^{\boldsymbol{\tau}-\boldsymbol{\nu}}(v_G - v_G^{\mathcal{T}_\ell}))\mathrm{d}x$$
$$= \int_D \sum_{\boldsymbol{\nu}\leq\boldsymbol{\tau}}\sqrt{\binom{\boldsymbol{\tau}}{\boldsymbol{\nu}}}\left[\sum_{\boldsymbol{m}\leq\boldsymbol{\nu}}\kappa_0(\boldsymbol{\nu},\boldsymbol{m})\left(\frac{\rho^{\boldsymbol{m}}}{\sqrt{\boldsymbol{m}!}}\sqrt{a}\nabla\partial^{\boldsymbol{m}}(u - u^{\mathcal{T}_\ell})\right)\right]$$
$$\times\left(\frac{\rho^{\boldsymbol{\tau}-\boldsymbol{\nu}}}{\sqrt{(\boldsymbol{\tau}-\boldsymbol{\nu})!}}\sqrt{a}\nabla\partial^{\boldsymbol{\tau}-\boldsymbol{\nu}}(v_G - v_G^{\mathcal{T}_\ell})\right)\mathrm{d}x.$$

It holds that $\sum_{\boldsymbol{\nu}\leq\boldsymbol{\tau}}\binom{\boldsymbol{\tau}}{\boldsymbol{\nu}} = 2^{\boldsymbol{\tau}}$. By a twofold application of the Cauchy–Schwarz inequality

$$\frac{\rho^{2\boldsymbol{\tau}}}{\boldsymbol{\tau}!}|\partial^{\boldsymbol{\tau}} G(u - u^{\mathcal{T}_\ell})|^2 \leq \left(\sum_{\boldsymbol{\nu}\leq\boldsymbol{\tau}}\sqrt{\binom{\boldsymbol{\tau}}{\boldsymbol{\nu}}}\|[\dots]\|_{L^2(D)}\frac{\rho^{\boldsymbol{\tau}-\boldsymbol{\nu}}}{\sqrt{(\boldsymbol{\tau}-\boldsymbol{\nu})!}}\|\partial^{\boldsymbol{\tau}-\boldsymbol{\nu}}(v_G - v_G^{\mathcal{T}_\ell})\|_a\right)^2$$
$$\leq 2^{\boldsymbol{\tau}}\sum_{\boldsymbol{\nu}\leq\boldsymbol{\tau}}\|[\dots]\|^2_{L^2(D)}\frac{\rho^{2(\boldsymbol{\tau}-\boldsymbol{\nu})}}{(\boldsymbol{\tau}-\boldsymbol{\nu})!}\|\partial^{\boldsymbol{\tau}-\boldsymbol{\nu}}(v_G - v_G^{\mathcal{T}_\ell})\|^2_a.$$

We define the sequence $(\widetilde{\rho})_{j\geq 1}$ by $\widetilde{\rho}_j := \sqrt{2}\rho_j$, $j \in \mathbb{N}$. By a change of the order of summation

$$\sum_{\|\boldsymbol{\tau}\|_{\ell^\infty}\leq r}\frac{\rho^{2\boldsymbol{\tau}}}{\boldsymbol{\tau}!}|\partial^{\boldsymbol{\tau}} G(u - u^{\mathcal{T}_\ell})|^2 \leq \sum_{\|\boldsymbol{\nu}\|_{\ell^\infty}\leq r} 2^{\boldsymbol{\nu}}\|[\dots]\|^2_{L^2(D)}\sum_{\|\boldsymbol{\tau}\|_{\ell^\infty}\leq r,\boldsymbol{\tau}\geq\boldsymbol{\nu}}\frac{\widetilde{\rho}^{2(\boldsymbol{\tau}-\boldsymbol{\nu})}}{(\boldsymbol{\tau}-\boldsymbol{\nu})!}\|\partial^{\boldsymbol{\tau}-\boldsymbol{\nu}}(v_G - v_G^{\mathcal{T}_\ell})\|^2_a.$$

Since $\sum_{\boldsymbol{m}\leq\boldsymbol{\nu}}\widetilde{\kappa}_0(\boldsymbol{\nu},\boldsymbol{m}) \leq 2$ due to Lemma 4.1, where $\widetilde{\kappa}_0$ is w.r. to $(\widetilde{b}_j)_{j\geq 1}$, by the Cauchy–Schwarz inequality and (4.8)

$$\sum_{k\geq 0}\sum_{\boldsymbol{\nu}\in\Lambda_k}\int_D\left(\sum_{\boldsymbol{m}\leq\boldsymbol{\nu}}\widetilde{\kappa}_0(\boldsymbol{\nu},\boldsymbol{m})\frac{\widetilde{\rho}^{\boldsymbol{m}}}{\sqrt{\boldsymbol{m}!}}\sqrt{a}|\nabla\partial^{\boldsymbol{m}}(u - u^{\mathcal{T}_\ell})|\right)^2\mathrm{d}x \leq 2\sum_{k\geq 0}\sum_{\boldsymbol{\nu}\in\Lambda_k}\sum_{\boldsymbol{m}\leq\boldsymbol{\nu}}\widetilde{\kappa}_0(\boldsymbol{\nu},\boldsymbol{m})\frac{\widetilde{\rho}^{2\boldsymbol{m}}}{\boldsymbol{m}!}\|\partial^{\boldsymbol{m}}(u - u^{\mathcal{T}_\ell})\|^2_a$$
$$\leq 2\sum_{k\geq 0}\sum_{\ell=0}^{k}\frac{(\sqrt{r}\widetilde{K}_0)^{k-\ell}}{(k-\ell)!}\sum_{\boldsymbol{m}\in\Lambda_\ell}\frac{\widetilde{\rho}^{2\boldsymbol{m}}}{\boldsymbol{m}!}\|\partial^{\boldsymbol{m}}(u - u^{\mathcal{T}_\ell})\|^2_a$$
$$= 2\sum_{\ell\geq 0}\sum_{k\geq\ell}\frac{(\sqrt{r}\widetilde{K}_0)^{k-\ell}}{(k-\ell)!}\sum_{\boldsymbol{m}\in\Lambda_\ell}\frac{\widetilde{\rho}^{2\boldsymbol{m}}}{\boldsymbol{m}!}\|\partial^{\boldsymbol{m}}(u - u^{\mathcal{T}_\ell})\|^2_a$$
$$\leq 4\sum_{\|\boldsymbol{m}\|_{\ell^\infty}\leq r}\frac{\widetilde{\rho}^{2\boldsymbol{m}}}{\boldsymbol{m}!}\|\partial^{\boldsymbol{m}}(u - u^{\mathcal{T}_\ell})\|^2_a,$$

which proves the assertion together with the previous inequality. ∎

The following theorem is directly implied by Theorem 4.7 and Proposition 4.9.

Theorem 4.10. *Let the assumption of Lemma 4.5 be satisfied for a positive sequence* $(\rho_j)_{j\geq 1}$*, and let* $r \in \mathbb{N}$ *and assume that* $\widetilde{K}_1 < C_r/\sqrt{2}$*. Then there exists a constant* $C > 0$ *such that for every* $\boldsymbol{y} \in U$

$$\sum_{\boldsymbol{\tau}\in\mathcal{F},\|\boldsymbol{\tau}\|_{\ell^\infty}\leq r} \frac{\rho^{2\boldsymbol{\tau}}}{\boldsymbol{\tau}!}|\partial^{\boldsymbol{\tau}} G(u(\boldsymbol{y}) - u^{\mathcal{T}_\ell}(\boldsymbol{y}))|^2 \leq C\left(\frac{(a_{\max}(\boldsymbol{y}))}{(a_{\min}(\boldsymbol{y}))^2}(1 + \||\nabla Z(\boldsymbol{y})|\Phi_{\overline{\boldsymbol{\beta}}}\|^2_{L^\infty(D)})\right)^2 M_\ell^{-4/d}\|f\|^2_{L^2_{\overline{\boldsymbol{\beta}}}(D)}\|G\|^2_{L^2_{\overline{\boldsymbol{\beta}}}(D)}.$$

Remark 4.11. The statement of Theorem 4.10 also holds true for the dimensionally truncated solutions u^s and $u^{s,\mathcal{T}_\ell}$, for every truncation dimension $s \in \mathbb{N}$. In particular, the constant C which appears in the error bound is independent of s.

Remark 4.12. The parametric regularity estimate in Theorem 4.10 also holds if $f \in (V^*, L^2_{\overline{\boldsymbol{\beta}}}(D))_{t,\infty}$ and $G(\cdot) \in (V^*, L^2_{\overline{\boldsymbol{\beta}}}(D))_{t',\infty}$ for $t, t' \in [0,1]$. Then, the FE discretization error contribution to the overall error is bounded by a constant times $M_\ell^{-2(t+t')/d}$. This follows from Remark 4.8.

5. Multilevel QMC convergence analysis

The sequences $(b_j)_{j\geq1}$ and $(\bar{b}_j)_{j\geq1}$ in the assumptions in (**A1**) and (**A2**) will be the input for the QMC weight sequence $(\gamma_j)_{j\geq1}$ of product weights. In the MLQMC quadrature algorithm Q_L^* in (3.3), we apply a randomly shifted lattice rule on level $\ell = 0$ with respect to the QMC weight sequence

$$\gamma_j = b_j^{2p'}, \quad j \geq 1, \tag{5.1}$$

for some $p' \in (0,1)$ and on the levels $\ell = 1, \ldots, L$ with respect to the QMC weight sequence

$$\bar{\gamma}_j = (b_j^{1-\theta} \vee \bar{b}_j)^{2\bar{p}'}, \quad j \geq 1 \tag{5.2}$$

for some $p' \in (0,1)$ and some $\theta \in (0,1)$. Here, for $c_1, c_2 \in \mathbb{R}$, $c_1 \vee c_2 := \max\{c_1, c_2\}$.

Theorem 5.1. *For every* $L \in \mathbb{N}_0$ *and sequences* $(s_\ell)_{\ell=0,\ldots,L}$ *and* $(N_\ell)_{\ell=0,\ldots,L}$*, the ensuing error estimate holds under the following conditions:*

(1) *Gaussian weight functions:* $(b_j)_{j\geq1} \in \ell^p(\mathbb{N})$ *for some* $p \in (2/3, 2)$ *and* $(b_j^{1-\theta} \vee \bar{b}_j)_{j\geq1} \in \ell^{\bar{p}}(\mathbb{N})$ *for some* $\bar{p} \in [p, 2)$ *with* $\chi = 1/(2p) + 1/4 - \varepsilon$ *and* $\bar{\chi} = 1/(2\bar{p}) + 1/4 - \bar{\varepsilon}$*. The QMC weight sequence in (5.1) is applied with* $p' = p/4 + 1/2 - \varepsilon p$ *on the level* $\ell = 0$ *for* $\varepsilon \in (0, 3/4 - 1/(2p))$*. The QMC weight sequence in (5.2) is applied with* $\bar{p}' = \bar{p}/4 + 1/2 - \bar{\varepsilon}\bar{p}$ *on the levels* $\ell = 1, \ldots, L$ *for* $\bar{\varepsilon} \in (0, 3/4 - 1/(2\bar{p}))$*.*
(2) *Exponential weight functions:* $(b_j)_{j\geq1} \in \ell^p(\mathbb{N})$ *for some* $p \in (2/3, 1]$ *and for* $(b_j^{1-\theta} \vee \bar{b}_j)_{j\geq1} \in \ell^{\bar{p}}(\mathbb{N})$ *for some* $\bar{p} \in [p, 1]$ *with* $\chi = 1/p - 1/2$ *and* $\bar{\chi} = 1/\bar{p} - 1/2$*. The QMC weight sequence in (5.1) is applied with* $p' = 1 - p/2$ *on the level* $\ell = 0$*. The QMC weight sequence in (5.2) is applied with* $\bar{p}' = 1 - \bar{p}/2$ *on the levels* $\ell = 1, \ldots, L$*.*

There exists a constant $C > 0$ *that is in particular independent of* $(M_\ell)_{\ell\geq0}$*,* $(s_\ell)_{\ell=0,\ldots,L}$*,* $(N_\ell)_{\ell=0,\ldots,L}$*, and of* $L \in \mathbb{N}_0$*, such that*

$$\sqrt{\mathbb{E}^{\boldsymbol{\Delta}}(|\mathbb{E}(G(u)) - Q_L^*(G(u^L))|^2)} \leq C\Bigg(\sup_{j>s_L}\{b_j^{4-p}\} + M_L^{-4/d} + (\varphi(N_0))^{-2\chi} + \sum_{\ell=1}^{L}(\varphi(N_\ell))^{-2\bar{\chi}}\left(\xi_{\ell,\ell-1}\sup_{j>s_{\ell-1}}\{b_j^{2\theta}\} + M_{\ell-1}^{-4/d}\right)\Bigg)^{1/2},$$

where $\xi_{\ell,\ell-1} := 0$ *if* $s_\ell = s_{\ell-1}$ *and* $\xi_{\ell,\ell-1} := 1$ *otherwise.*

Proof. By the triangle inequality, for $\ell = 1, \dots, L$,

$$
\begin{aligned}
&|(I_{s_\ell} - Q_{s_\ell,N_\ell})(G(u^\ell) - G(u^{\ell-1}))| \\
&\quad \le |(I_{s_\ell} - Q_{s_\ell,N_\ell})(G(u^{s_\ell,\mathcal{T}_\ell}) - G(u^{s_\ell,\mathcal{T}_{\ell-1}}))| + |(I_{s_\ell} - Q_{s_\ell,N_\ell})(G(u^{s_\ell,\mathcal{T}_{\ell-1}}) - G(u^{s_{\ell-1},\mathcal{T}_{\ell-1}}))|
\end{aligned}
$$

and

$$
\begin{aligned}
&|(I_{s_\ell} - Q_{s_\ell,N_\ell})(G(u^{s_\ell,\mathcal{T}_\ell}) - G(u^{s_\ell,\mathcal{T}_{\ell-1}}))| \\
&\quad \le |(I_{s_\ell} - Q_{s_\ell,N_\ell})(G(u^{s_\ell}) - G(u^{s_\ell,\mathcal{T}_\ell}))| + |(I_{s_\ell} - Q_{s_\ell,N_\ell})(G(u^{s_\ell}) - G(u^{s_\ell,\mathcal{T}_{\ell-1}}))|,
\end{aligned}
$$

where we recall that $u^\ell := u^{s_\ell,\mathcal{T}_\ell}$, $\ell = 0, \dots, L$. We wish to show the conditions of Theorem 3.1 for integrands $\boldsymbol{y} \mapsto G(u^{s_\ell}(\boldsymbol{y})) - G(u^{s_\ell,\mathcal{T}_\ell}(\boldsymbol{y}))$ and $\boldsymbol{y} \mapsto G(u^{s_\ell,\mathcal{T}_{\ell-1}}(\boldsymbol{y})) - G(u^{s_{\ell-1},\mathcal{T}_{\ell-1}}(\boldsymbol{y}))$.

Recall

$$
K_1 := \left\| \sum_{j\ge1} \frac{\max\{|\nabla\psi_j|\Phi_{\overline{\boldsymbol{\beta}}}, |\psi_j|\}}{\bar{b}_j} \right\|_{L^\infty(D)} < \infty, \tag{5.3}
$$

the conditions of Theorem 3.1 are satisfied for the integrand $\boldsymbol{y} \mapsto G(u^{s_\ell}(\boldsymbol{y})) - G(u^{s_\ell,\mathcal{T}_\ell}(\boldsymbol{y}))$ with the sequence $(\bar{b}_j)_{j\ge1}$ and $\kappa < C_r/(\sqrt{2}K_1)$ by Theorem 4.10 and Remark 4.11 with $r = 1$. Specifically, we apply Theorem 4.10 and Remark 4.11 with $\rho_j = \kappa/\bar{b}_j$, $j \ge 1$.

For the integrand $\boldsymbol{y} \mapsto G(u^{s_\ell,\mathcal{T}_{\ell-1}}(\boldsymbol{y})) - G(u^{s_{\ell-1},\mathcal{T}_{\ell-1}}(\boldsymbol{y}))$, we apply Theorem 4.3 with $\rho_j = \kappa/b_j^{1-\theta}$, $j \ge 1$. Then, the condition of Theorem 4.3 is satisfied for $\eta = \theta/(1-\theta)$ and $\kappa < \log(2)/K_0$, where K_0 is as in assumption (**A1**). Hence, the conditions of Theorem 3.1 are satisfied for the integrand $\boldsymbol{y} \mapsto G(u^{s_\ell,\mathcal{T}_{\ell-1}}(\boldsymbol{y})) - G(u^{s_{\ell-1},\mathcal{T}_{\ell-1}}(\boldsymbol{y}))$. Since the sequence $(b_j^{1-\theta} \vee \bar{b}_j)_{j\ge1}$ dominates $(b_j^{1-\theta})_{j\ge1}$ and $(\bar{b}_j)_{j\ge1}$, Theorem 3.1 can be applied with $\tilde{b}_j = b_j^{1-\theta} \vee \bar{b}_j$, $j \ge 1$. For the exponential weight functions, we note that $\eta_1 = C(\max_{j>s_{\ell-1}}\{b_j^\theta\} + M_{\ell-1}^{-2/d})$ for a constant $C > 0$ (independent of ℓ), and with $\eta_2 = 5$ in the notation of the second item of Theorem 3.1.

On discretization level $\ell = 0$, the parametric integrand is $\boldsymbol{y} \mapsto G(u^{s_0,\mathcal{T}_0})$. The conditions of Theorem 3.1 are satisfied with $\tilde{b}_j = b_j$, $j \ge 1$ (see also [31], Thms. 11 and 13). The assertion follows with (2.12) and (3.4). ☐

Remark 5.2. If $f \in (V^*, L^2_{\overline{\boldsymbol{\beta}}}(D))_{t,\infty}$ and $G(\cdot) \in (V^*, L^2_{\overline{\boldsymbol{\beta}}}(D))_{t',\infty}$ for some $t, t' \in [0,1]$, then the error estimate in Theorem 5.1 also holds with an error bounded by an absolute multiple of $M_\ell^{-2(t+t')/d}$ on mesh level ℓ.

Remark 5.3. When the GRF Z is stationary in $D \subset \mathbb{R}^d$, the *covariance kernel* $k(x, x') := \mathbb{E}(Z(x)Z(x'))$ of Z depends only on $x - x'$, *cf.* [1]. A widely used parametric family of covariances for stationary GRFs was proposed by B. Matérn [39]. Here, the covariance kernel depends on two parameters $\nu, \lambda > 0$, where λ is referred to as correlation length and $Z \in C^t(\overline{D})$, μ-a.s., for every positive real number $t < \nu$. Wavelet type function systems exist which allow to represent the GRF Z in terms of a sequence $(y_j)_{j\ge1}$ of independent, standard normally distributed y_j, that satisfy assumption (**A1**) with $b_j \sim j^{-\widehat{\beta}/d}$, $j \ge 1$, for every $\widehat{\beta} < \nu$, *cf. e.g.* Corollary 4.3 of [8]. In [8], the random field Z in D is constructed by restriction of a GRF defined on suitable product domain that depends on the correlation length λ and which is a superset of D. For a constructive approach to obtain function systems of expansions with i.i.d. coefficients, we refer for example to [23] and the references there. For a discussion of the Hölder regularity and $L^q(\Omega)$ integrability of GRFs expanded in generic wavelets, we refer to Section 9 of [31]. There, also if $C^t(\overline{D})$-regularity of the respective GRF Z holds as an implication by Proposition 18 of [31], the generic wavelets satisfy assumption (**A1**) with $b_j \sim j^{-\widehat{\beta}/d}$, $j \ge 1$, for every $\widehat{\beta} < t$.

Remark 5.4. In the case of single-level QMC, also fractional Hölder regularity of the lognormal coefficent $a = \exp(Z)$ is covered by our theory in [31]. The GRF of the model function system of generic wavelets discussed in Section 9 from [31] is for $d = 1$ and for wavelets that are scaled to decay as $\|\psi_j\|_{L^\infty(D)} \sim j^{-1/2-\varepsilon}$, $j \ge 1$, a member of $L^q(\Omega; C^{1/2+\varepsilon'}(\overline{D}))$, for every $q \in [1,\infty)$ and for every $\varepsilon > \varepsilon' > 0$, *cf.* Proposition 19 of [31]. The sequence $(b_j)_{j\ge1}$ may be chosen such that $b_j \sim j^{-1/2-\varepsilon'}$, for every $j \ge 1$ and for some $\varepsilon' \in (0,\varepsilon)$. For every $p > 2/(1+2\varepsilon')$, this sequence $(b_j)_{j\ge1} \in \ell^p(\mathbb{N})$ is admissible with Gaussian weight functions for every

$\varepsilon' > 0$, *cf.* Theorem 11 of [31] and therefore QMC with Gaussian weight functions and product weights is applicable for every $\varepsilon > 0$. However, for $1/2 > \varepsilon > 0$, the convergence theory for QMC with product weights in Theorem 13 from [31] does not seem to be applicable with exponential weight functions in this case. Numerical experiments in Section 11 from [31] for Z being the Brownian bridge and Gaussian QMC weight functions reported convergence rates slightly larger than $1/2$. The Brownian bridge Z is a borderline case of our theory in [31]; then, $Z \in C^{1/2-\varepsilon}(\overline{D})$ μ-a.s. for every $0 < \varepsilon \le 1/2$ and $b_j \sim j^{-1/2}$ using the Lévy–Ciesielski decomposition of Z.

6. Error *vs.* work analysis

We discuss concrete choices of algorithmic steering parameters in the preceding error bounds to obtain asymptotic error *vs.* work estimates. We elaborate the widely used case of GRFs Z with Matérn-like covariances, and compare the present results to previous work [26,36] and the references there.

6.1. Error *vs.* work for local supports and product weights

In the estimate of Theorem 5.1, the error contributions of the QMC quadrature and the spatial approximation by FE and dimension truncation are coupled on the different levels. The number N_ℓ of QMC points at level $\ell = 0, 1, \ldots, L$ should minimize the error estimate subject to a prescribed work measure. We consider functions $(\psi_j)_{j\ge1}$ which are compactly supported in D, as for example certain MRA. Note that this will only affect the choice of the work measure for the assembly of stiffness matrices.

Let us assume that the MRA $(\psi_\lambda)_{\lambda\in\triangledown}$ results from a finite number of generating (or "mother") wavelets by scaling and translation, *i.e.*,

$$\psi_\lambda(x) := \psi(2^{|\lambda|}x - k), \quad k \in \triangledown_{|\lambda|}, x \in D. \tag{6.1}$$

We use notation that is standard for MRA, *i.e.*, the function system is indexed by $\lambda = (|\lambda|, k) \in \triangledown$, where $|\lambda| \in \mathbb{N}_0$ refers to the level and $k \in \triangledown_{|\lambda|}$ to the translation. The index set $\triangledown_\ell$ has cardinality $|\triangledown_\ell| = \mathcal{O}(2^{d\ell})$, $\ell \in \mathbb{N}_0$. Let $j : \triangledown \to \mathbb{N}$ be a suitable enumeration. The overlap on every level $|\lambda| = \ell \in \mathbb{N}_0$ is assumed to be uniformly bounded, *i.e.*, there exists $K > 0$ such that for every $\ell \in \mathbb{N}_0$ and every $x \in D$,

$$|\{\lambda \in \triangledown : |\lambda| = \ell, \psi_\lambda(x) \neq 0\}| \le K.$$

Additionally, for constants $\sigma, \widehat{\alpha} > 0$ we introduce the scaling

$$\|\psi_\lambda\|_{L^\infty(D)} \le \sigma 2^{-\widehat{\alpha}|\lambda|}, \quad \lambda \in \triangledown. \tag{6.2}$$

Under this assumption, the work to assemble one sample of the stiffness matrix (*i.e.* for one QMC point) on discretization level $\ell \in \mathbb{N}_0$ scales for large ℓ as $\mathcal{O}(M_\ell |j^{-1}(s_\ell)|) = \mathcal{O}(M_\ell \log(s_\ell))$.

Proposition 6.1. *For $d = 1$, the work to solve the linear system that corresponds to (2.8) for one sample is $\mathcal{O}(M_\ell)$, $\ell \in \mathbb{N}_0$.*

Proof. The parametric stiffness matrix is tridiagonal and symmetric, positive definite with probability one. Therefore both, Cholesky decomposition and backsubstitution, can be performed in $\mathcal{O}(M_\ell)$ work and memory (see, *e.g.*, [24], Chap. 4.3.6). □

Due to Proposition 6.1 and Remark 6.2, we stipulate availability of a PDE solver with work

$$\text{work}_{\text{PDEsolve}} = \mathcal{O}(M_\ell^{1+\eta}) \tag{A3}$$

for some $\eta \ge 0$ with implied constants independent of $\ell \in \mathbb{N}_0$ *and, in particular, of the realization of the PDE coefficients*. For $D = (0,1)^2$ and a sparse direct solver based on *nested dissection* it is known that $\eta = 1/2$, *cf.* [22]. Note that $\eta = 0$ corresponds to linear complexity as is afforded by multigrid or multilevel preconditioned iterative solvers for elliptic PDEs in the deterministic setting; see, *e.g.*, [11,49]. The results in [29] on convergence of these methods for log-Gaussian coefficients in the $L^q(\Omega; V)$-norm, $q \in [1, \infty)$, and $d = 2, 3$ suggest that $\eta = 0$ may not be admissible for MLQMC and $d = 2, 3$.

Remark 6.2. The uniformity w.r. to the coefficient realizations of the work estimate (**A3**) is, for the presently considered log-Gaussian diffusion coefficient models, by no means to be taken for granted [29]. Since for $d = 2, 3$ stiffness matrices will not be tridiagonal, usually iterative solvers are used. In [29], strong convergence (in the $L^q(\Omega; V)$-norm) for iterative methods is shown for every $\eta > 0$ in the general framework of [49], which is nearly optimal complexity (w.r. to the degrees of freedom) of a PDE solver. This is sufficient for single-level QMC and multilevel Monte Carlo (MLMC). Applicability to MLQMC does not seem to be a direct consequence. In practice also direct solvers have been used for $d = 2$ with observed $\eta < 1/2$ using different sparse direct solvers than in [22], *e.g.*, in Figure 5 from [36] for $D = (0, 1)^2$, $\eta_{\text{observed}} \approx 0$ and in Figure 3 from [32], $\eta_{\text{observed}} \approx 0.3$ for $D = \mathbb{S}^2$ (the two dimensional sphere).

Under (**A3**) the model for the computational work for the MLQMC quadrature reads, for every $L \in \mathbb{N}_0$, as

$$\text{work}_L = \mathcal{O}\left(\sum_{\ell=0}^{L} s_\ell N_\ell \log(N_\ell) + \sum_{\ell=0}^{L} N_\ell (M_\ell \log(s_\ell) + M_\ell^{1+\eta})\right), \tag{6.3}$$

where the first sum in (6.3) is the work of the generation of the QMC points which includes the work to obtain the generating vectors by the fast CBC construction, *cf.* [42,43]. The work model in (6.3) depends on the choices for $(s_\ell)_{\ell=0,\dots,L}$, $(N_\ell)_{\ell=0,\dots,L}$, and $(M_\ell)_{\ell\geq 0}$, which we shall not indicate explicitly in our notation and simply write "work_L". The second sum in (6.3) is the work of the evaluation of the MLQMC quadrature. The sequence

$$b_{j(\lambda)} = b_\lambda := c2^{-\widehat{\beta}|\lambda|}, \quad \lambda \in \triangledown, \tag{6.4}$$

together with $(\psi_\lambda)_{\lambda\in\triangledown}$ defined in (6.1) and (6.2) satisfies the assumption in (**A1**) for $\max\{1, d/2\} < \widehat{\beta} < \widehat{\alpha}$ and some $c > 0$. Since $\||\nabla\psi_\lambda|\|_{L^\infty(D)} \leq \sigma 2^{-(\widehat{\alpha}-1)|\lambda|}\||\nabla\psi|\|_{L^\infty(D)}$, (assuming $\||\nabla\psi|\|_{L^\infty(D)} < \infty$) $\lambda \in \triangledown$, the sequence

$$\bar{b}_j := b_j^{(\widehat{\beta}-1)/\widehat{\beta}}, \quad j \in \mathbb{N}, \tag{6.5}$$

and $(\psi_j)_{j\geq 1}$ defined in (6.1) and (6.2) satisfy the assumption (**A2**). In this section we assume that

$$f \in (V^*, L^2_{\boldsymbol{\beta}}(D))_{t,\infty} \quad \text{and} \quad G(\cdot) \in (V^*, L^2_{\boldsymbol{\beta}}(D))_{t',\infty}, \quad t, t' \in [0, 1], \tag{\textbf{A4}}$$

and define $\tau := t + t'$. In the following, we assume that

$$M_\ell \sim 2^{d\ell}, \quad \ell \geq 0. \tag{\textbf{A5}}$$

The ensuing analysis is inspired by [35], Section 3.7 (see also [19, 36]). We will restrict the analysis to one QMC rule with respect to the QMC weight sequence (5.2) on all levels $\ell = 0, \dots, L$, but remark that in some cases it might be beneficial to use a second one with respect to the QMC weight sequence (5.1) on the level $\ell = 0$. The MLQMC quadrature depends on the algorithmic steering parameters $(N_\ell)_{\ell=0,\dots,L}$, $(s_\ell)_{\ell=0,\dots,L}$, $(M_\ell)_{\ell\geq 0}$, and also on $\theta \in (0, 1)$. The number of degrees of freedom $(M_\ell)_{\ell\geq 0}$ of the FE discretization in D are assumed to be given. The parameter $\theta \in (0, 1)$ is for now left arbitrary. According to the estimate in Theorem 5.1, θ can be used to balance the truncation error with the FE error on the levels $\ell = 0, \dots, L$. We will use this feature to discuss two possible strategies to choose the truncation dimensions $(s_\ell)_{\ell=0,\dots,L}$.

Strategy 1: The contributions in the QMC weight sequence in (5.2) are equilibrated, *i.e.*, we choose $\theta = 1/\widehat{\beta}$, which implies that $b^{1-\theta} = \bar{b}_j$, $j \in \mathbb{N}$. The truncation dimension s_L is also chosen to equilibrate the respective truncation and FE error in the estimate of Theorem 5.1. We choose

$$s_L \sim 2^{d\lceil L\tau/\widetilde{\beta}\rceil}$$

for some

$$1 < \widetilde{\beta} < 2\widehat{\beta} - \frac{d}{2} \tag{6.6}$$

close to $2\widehat{\beta} - d/2$, where we use that $M_\ell = \mathcal{O}(2^{d\ell})$, $\ell = 0, \dots, L$, and $(b_j)_{j\geq 1} \in \ell^p(\mathbb{N})$ for every $p > d/\widehat{\beta}$. On the levels $\ell = 0, \dots, L-1$, we either increase s_ℓ or leave it constant. We choose

$$s_\ell \sim \min\{2^{d\lceil \tau\ell \rceil}, s_L\}, \quad \ell = 0, \dots, L-1.$$

Strategy 2: For particular $(\psi_\lambda)_{\lambda\in\bigtriangledown}$ and meshes, it may be interesting to align the level structure $(\psi_\lambda)_{\lambda\in\bigtriangledown}$ and the used FE meshes. Therefore, we choose

$$s_\ell \sim M_\ell, \quad \ell = 0, \dots, L.$$

The choice $\theta = \tau/\widehat{\beta}$ equilibrates the truncation and FE error in the estimate of Theorem 5.1 on the levels $\ell = 0, \dots, L$ assuming that $\widehat{\beta} > \tau$. Then, $(b_j^{1-\theta} \vee \bar{b}_j)_{j\geq 1} \in \ell^{\bar{p}}(\mathbb{N})$ for every $\bar{p} > d/(\min\{\widehat{\beta} - \tau, \widehat{\beta} - 1\})$.

For either of the strategies and for every $L \in \mathbb{N}_0$, by Theorem 5.1 we obtain the error estimate

$$\text{error}_L^2 = \mathcal{O}\left(M_L^{-2\tau/d} + \sum_{\ell=0}^{L} (\varphi(N_\ell))^{-2\bar{\chi}} M_\ell^{-2\tau/d} \right). \tag{6.7}$$

Since the Euler totient function satisfies that $(\varphi(N))^{-1} \leq N^{-1}(e^{\widehat{\gamma}} \log\log N + 3/\log\log(N))$ for every $N \geq 3$, where $\widehat{\gamma} \approx 0.5772$ is the Euler–Mascheroni constant, $(\varphi(N))^{-1} \leq 9/N$ for every $N = 3, \dots, 10^{30}$. We will for simplicity restrict in our analysis the range of N to $N \leq 10^{30}$ and use the bound $(\varphi(N))^{-1} \leq 9/N$. In Strategies 1 and 2, the $\bar{p}$-summability of the sequence $(b_j^{1-\theta} \vee \bar{b}_j)_{j\geq 1}$ holds with a strict inequality condition on $\bar{p}$, *i.e.*, $(b_j^{1-\theta} \vee \bar{b}_j)_{j\geq 1} \in \ell^{\bar{p}}(\mathbb{N})$, for every $\bar{p} > d/(\widehat{\beta} - 1)$ in the case of Strategy 1 and for every $\bar{p} > d/\min\{\widehat{\beta} - \tau, \widehat{\beta} - 1\}$ in the case of Strategy 2. Since the QMC convergence rate $\bar{\chi}$ depends on the exponent $\bar{p}$, there exists $\varepsilon > 0$ such that $\bar{\chi}(1+\varepsilon)$ is also admissible in (6.7) due to Theorem 5.1. Using $\log(N) \leq N^\varepsilon/(\varepsilon e)$ for every $N \in \mathbb{N}$, the factor $N_\ell \log(N_\ell)$ in (6.3) may be estimated by $N_\ell^{1+\varepsilon}$. Since $N_\ell^{1+\varepsilon}$ appears then in the estimate of the work (6.3) and in the error estimate (6.7), it can be substituted by N_ℓ, using the strict inequalities in the above bounds for the admissible indices, and choosing $\varepsilon > 0$ sufficiently small.

We obtain with the choices for $(s_\ell)_{\ell=0,\dots,L}$ in Strategies 1 and 2

$$\text{work}_L = \begin{cases} \mathcal{O}\left(\sum_{\ell=0}^{L} N_\ell (M_\ell \log(M_\ell) + \max\{M_\ell^{1+\eta}, \min\{M_\ell^\tau, M_L^{\tau/\widetilde{\beta}}\}\}) \right), & \text{for Strategy 1,} \\ \mathcal{O}\left(\sum_{\ell=0}^{L} N_\ell (M_\ell \log(M_\ell) + M_\ell^{1+\eta}) \right), & \text{for Strategy 2.} \end{cases}$$

and

$$\text{error}_L^2 = \mathcal{O}\left(M_L^{-2\tau/d} + \sum_{\ell=0}^{L} N_\ell^{-2\bar{\chi}} M_\ell^{-2\tau/d} \right).$$

We will distinguish between the cases that $\eta = 0$ and $\eta > 0$ in (**A3**). We treat Strategy 2 and the case $\eta > 0$ first. As above, $\log(M) \leq M^\eta/(\eta e)$ for every $M \in \mathbb{N}$. To obtain optimal choices for the sample numbers $(N_\ell)_{\ell=0,\dots,L}$, we search for a stationary point of the function

$$g(\xi) := M_L^{-2\tau/d} + \sum_{\ell=0}^{L} N_\ell^{-2\bar{\chi}} M_\ell^{-2\tau/d} + \xi \sum_{\ell=0}^{L} N_\ell M_\ell^{1+\eta}$$

with respect to N_ℓ, *i.e.*, we solve the first order necessary condition $\partial g/\partial N_\ell = 0$ (see also [35], Sect. 3.7). This gives

$$N_\ell = \left\lceil N_0 M_\ell^{-(2\tau/d+1+\eta)/(1+2\bar{\chi})} \right\rceil, \quad \ell = 1, \dots, L, \tag{6.8}$$

and with setting $E_\ell = M_\ell^{(1+\eta-\tau/(d\bar{\chi}))2\bar{\chi}/(1+2\bar{\chi})}$, $\ell = 0, \dots, L$,

$$\text{error}_L^2 = \mathcal{O}\left(M_L^{-2\tau/d} + N_0^{-2\bar{\chi}} \sum_{\ell=0}^{L} E_\ell\right) \quad \text{and} \quad \text{work} = \mathcal{O}\left(N_0 \sum_{\ell=0}^{L} E_\ell\right), \tag{6.9}$$

where

$$\sum_{\ell=0}^{L} E_\ell = \begin{cases} \mathcal{O}(1) & \text{if } 1+\eta < \tau/(d\bar{\chi}), \\ \mathcal{O}(L) & \text{if } 1+\eta = \tau/(d\bar{\chi}), \\ \mathcal{O}(2^{(2\bar{\chi}d(1+\eta)-2\tau)L/(1+2\bar{\chi})}) & \text{if } 1+\eta > \tau/(d\bar{\chi}). \end{cases} \tag{6.10}$$

The parameter N_0 is chosen to balance the error contributions, *i.e.*, $N_0^{-2\bar{\chi}} \sum_{\ell=0}^{L} E_\ell = \mathcal{O}(M_L^{-2\tau/d})$, which implies

$$N_0 = \begin{cases} \lceil 2^{\tau L/\bar{\chi}} \rceil & \text{if } 1+\eta < \tau/(d\bar{\chi}), \\ \lceil 2^{\tau L/\bar{\chi}} L^{1/(2\bar{\chi})} \rceil & \text{if } 1+\eta = \tau/(d\bar{\chi}), \\ \lceil 2^{(2\tau+d(1+\eta))L/(1+2\bar{\chi})} \rceil & \text{if } 1+\eta > \tau/(d\bar{\chi}). \end{cases} \tag{6.11}$$

We conclude that $\text{error}_L^2 = \mathcal{O}(M_L^{-2\tau/d})$ can be achieved with

$$\text{work}_L = \begin{cases} \mathcal{O}(2^{\tau L/\bar{\chi}}) & \text{if } 1+\eta < \tau/(d\bar{\chi}), \\ \mathcal{O}(2^{\tau L/\bar{\chi}} L^{(1+2\bar{\chi})/(2\bar{\chi})}) & \text{if } 1+\eta = \tau/(d\bar{\chi}), \\ \mathcal{O}(2^{dL(1+\eta)}) & \text{if } 1+\eta > \tau/(d\bar{\chi}). \end{cases}$$

In the case that $\eta = 0$, the resulting work measure is considered in Section 3.7 from [35]. In particular, we obtain by equations (74) and (77) of [35]

$$N_\ell = \left\lceil N_0 \left(M_\ell^{-1-2\tau/d} \log(s_\ell)^{-1}\right)^{1/(1+2\bar{\chi})} \right\rceil, \quad \ell = 1, \dots, L, \tag{6.12}$$

and

$$N_0 = \begin{cases} \lceil 2^{\tau L/\bar{\chi}} \rceil & \text{if } d < \tau/\bar{\chi}, \\ \lceil 2^{\tau L/\bar{\chi}} L^{(1+4\bar{\chi})/(\bar{\chi}(2+4\bar{\chi}))} \rceil & \text{if } d = \tau/\bar{\chi}, \\ \lceil 2^{(d+2\tau)L/(1+2\bar{\chi})} L^{1/(1+2\bar{\chi})} \rceil & \text{if } d > \tau/\bar{\chi}. \end{cases} \tag{6.13}$$

Note that the corresponding work estimates are given on page 443 of [35]. We summarize this analysis as ε-complexity bounds in the following theorem.

Theorem 6.3 (Error *vs.* work for Strategy 2)**.** *Let the truncation dimensions $(s_\ell)_{\ell=0,\dots,L}$ be chosen according to Strategy 2 assuming $\widehat{\beta} > \max\{\tau, 1\}$. Let the assumptions (**A5**) and (**A3**) be satisfied for $\eta \geq 0$. If $\eta > 0$, the sample numbers for $Q_L^*(\cdot)$ are given by (6.11) and (6.8), $L \in \mathbb{N}_0$. If $\eta = 0$, the sample numbers for $Q_L^*(\cdot)$ are given by (6.13) and (6.12), $L \in \mathbb{N}_0$. Let f and $G(\cdot)$ satisfy (**A4**).*

1. *Gaussian weight functions: for $\bar{p} \in (\max\{2/3, d/(\widehat{\beta} - \tau), d/(\widehat{\beta} - 1)\}, 2)$, $\bar{\chi} = 1/(2\bar{p}) + 1/4 - \varepsilon'$ for $\varepsilon' > 0$ sufficiently small assuming $d/\min\{\widehat{\beta} - \tau, \widehat{\beta} - 1\} < 2$.*
2. *Exponential weight functions: for $\bar{p} \in (\max\{2/3, d/(\widehat{\beta}-\tau), d/(\widehat{\beta}-1)\}, 1]$, $\bar{\chi} = 1/\bar{p}-1/2$ assuming $d/\min\{\widehat{\beta}-\tau, \widehat{\beta} - 1\} < 1$.*

For an error threshold $1 > \varepsilon > 0$, we obtain

$$\sqrt{\mathbb{E}^{\boldsymbol{\Delta}}(|\mathbb{E}(G(u)) - Q_L^*(G(u^L))|^2)} = \mathcal{O}(\varepsilon)$$

is achieved with

$$\text{work}_L = \begin{cases} \mathcal{O}(\varepsilon^{-1/\bar{\chi}}) & \text{if } 1+\eta < \tau/(d\bar{\chi}), \\ \mathcal{O}(\varepsilon^{-1/\bar{\chi}} \log(\varepsilon^{-1})^{(1+2\bar{\chi})/(2\bar{\chi})}) & \text{if } 1+\eta = \tau/(d\bar{\chi}), \eta > 0, \\ \mathcal{O}(\varepsilon^{-1/\bar{\chi}} \log(\varepsilon^{-1})^{(1+4\bar{\chi})/(2\bar{\chi})}) & \text{if } d = \tau/\bar{\chi}, \eta = 0, \\ \mathcal{O}(\varepsilon^{-d/\tau(1+\eta)}) & \text{if } 1+\eta > \tau/(d\bar{\chi}), \eta > 0, \\ \mathcal{O}(\varepsilon^{-d/\tau} \log(\varepsilon^{-1})) & \text{if } d > \tau/\bar{\chi}, \eta = 0. \end{cases}$$

Here, the implied constants are independent of L, $(s_\ell)_{\ell=0,\dots,L}$, $(N_\ell)_{\ell=0,\dots,L}$, *and of* $(M_\ell)_{\ell\geq 0}$.

Remark 6.4. In Strategy 2, there is one parameter respectively one dimension of integration, per spatial degree of freedom, so that $s_\ell \sim M_\ell$, $\ell \geq 0$. This coupling occurs, for example, when circulant embedding is applied to evaluate a GRF on uniformly spaced spatial grid points such that each element of the FE mesh contains at least one of these points to perform a one point quadrature for computing the stiffness matrix. Numerical experiments with a QMC rule using a circulant embedding are presented in [25] and the references there.

For Strategy 1, we may restrict the analysis to the case $\tau > 1$, since for $\tau \leq 1$ the additional restriction $\widehat{\beta} > \tau$ for Strategy 2 is redundant and Strategy 2 can be applied. For $\eta > 0$, we obtain following the same line of argument as applied in the analysis of Strategy 2

$$N_\ell = \left\lceil N_0 \left(M_\ell^{2\tau/d} \max\{M_\ell^{1+\eta}, \min\{M_\ell^\tau, M_L^{\tau/\widetilde{\beta}}\}\} \right)^{-1/(1+2\bar{\chi})} \right\rceil, \quad \ell = 1,\dots,L, \tag{6.14}$$

where also (6.9) holds with

$$E_\ell = \left(M_\ell^{-\tau/(d\bar{\chi})} \max\{M_\ell^{1+\eta}, \min\{M_\ell^\tau, M_L^{\tau/\widetilde{\beta}}\}\} \right)^{2\bar{\chi}/(1+2\bar{\chi})}, \quad \ell = 0,\dots,L.$$

We observe that

$$\sum_{\ell=0}^{L} \left(M_\ell^{-\tau/(d\bar{\chi})} \max\{M_\ell^{1+\eta}, M_L^{\tau/\widetilde{\beta}}\} \right)^{2\bar{\chi}/(1+2\bar{\chi})} = \begin{cases} \mathcal{O}(2^{dL(\tau/\widetilde{\beta})2\bar{\chi}/(1+2\bar{\chi})}) & \text{if } 1+\eta \leq \tau/(d\bar{\chi}), \\ \mathcal{O}(2^{dL\max\{1+\eta-\tau/(d\bar{\chi}),\tau/\widetilde{\beta}\}2\bar{\chi}/(1+2\bar{\chi})}) & \text{if } 1+\eta > \tau/(d\bar{\chi}), \end{cases}$$

where we used that $\max\{x,y\} \leq x+y$ for every $x,y \in [0,\infty)$. The respective estimate for the sum over $(M_\ell^{-\tau/(d\bar{\chi})} \max\{M_\ell^{1+\eta}, M_\ell^\tau\})^{2\bar{\chi}/(1+2\bar{\chi})}$ is given in (6.10) with $\max\{1+\eta,\tau\}$ in place of $1+\eta$ (also in the conditions of the three cases). To estimate $\sum_{\ell=0}^L E_\ell$, we use the identity that $\max\{x,\min\{y,z\}\} = \min\{\max\{x,y\},\max\{x,z\}\}$ for every $x,y,z \in \mathbb{R}$, and apply the superadditivity of the minimum to obtain that

$$\sum_{\ell=0}^{L} E_\ell = \begin{cases} \mathcal{O}(1) & \text{if } \max\{\tau,1+\eta\} < \tau/(d\bar{\chi}), \\ \mathcal{O}(L) & \text{if } \max\{\tau,1+\eta\} = \tau/(d\bar{\chi}), \\ \mathcal{O}(2^{dL(1+\eta-\tau/(d\bar{\chi}))2\bar{\chi}/(1+2\bar{\chi})}) & \text{if } 1+\eta \geq \tau, 1+\eta > \tau/(d\bar{\chi}), \\ \mathcal{O}(2^{dL\min\{\tau-\tau/(d\bar{\chi}),\max\{(1+\eta-\tau/(d\bar{\chi})),\tau/\widetilde{\beta}\}\}2\bar{\chi}/(1+2\bar{\chi})}) & \text{if } 1+\eta < \tau, 1 < d\bar{\chi}. \end{cases}$$

As above, N_0 is chosen to balance the error, *i.e.*, $N_0 \sim M_L^{\tau/(d\bar{\chi})} (\sum_{\ell=0}^L E_\ell)^{1/(2\bar{\chi})}$. Specifically,

$$N_0 = \begin{cases} \lceil 2^{L\tau/\bar{\chi}} \rceil & \text{if } \max\{\tau,1+\eta\} < \tau/(d\bar{\chi}), \\ \lceil 2^{L\tau/\bar{\chi}} L^{1/(2\bar{\chi})} \rceil & \text{if } \max\{\tau,1+\eta\} = \tau/(d\bar{\chi}), \\ \lceil 2^{(2\tau+d(1+\eta))L/(1+2\bar{\chi})} \rceil & \text{if } 1+\eta \geq \tau, 1+\eta > \tau/(d\bar{\chi}), \\ \lceil 2^{dL\min\{\tau-\tau/(d\bar{\chi}),\max\{(1+\eta-\tau/(d\bar{\chi})),\tau/\widetilde{\beta}\}\}/(1+2\bar{\chi})+L\tau/\bar{\chi}} \rceil & \text{if } 1+\eta < \tau, 1 < d\bar{\chi}. \end{cases} \tag{6.15}$$

For $\eta = 0$,

$$N_\ell = \left\lceil N_0 \left(M_\ell^{2\tau/d} \max\{M_\ell \log(M_\ell), \min\{M_\ell^\tau, M_L^{\tau/\widetilde{\beta}}\}\} \right)^{-1/(1+2\bar{\chi})} \right\rceil, \quad \ell = 1,\dots,L, \tag{6.16}$$

and $E_\ell = (M_\ell^{-\tau/(d\bar\chi)} \max\{M_\ell \log(M_\ell), \min\{M_\ell^\tau, M_L^{\tau/\widetilde\beta}\}\})^{2\bar\chi/(1+2\bar\chi)}$, $\ell = 0,\dots,L$. We obtain similarly using $\tau > 1$,

$$\sum_{\ell=0}^L E_\ell = \begin{cases} \mathcal{O}(1) & \text{if } 1 < 1/(d\bar\chi), \\ \mathcal{O}(L) & \text{if } 1 = 1/(d\bar\chi), \\ \mathcal{O}(2^{(d-\tau/\bar\chi)L2\bar\chi/(1+2\bar\chi)} L^{2\bar\chi/(1+2\bar\chi)}) & \text{if } d-\tau/\bar\chi \geq d\tau/\widetilde\beta, \\ \mathcal{O}(2^{d(\tau/\widetilde\beta)L2\bar\chi/(1+2\bar\chi)}) & \text{if } d-\tau/\bar\chi < d\tau/\widetilde\beta < \tau(d-1/\bar\chi), \\ \mathcal{O}(2^{\tau(d-1/\bar\chi)L2\bar\chi/(1+2\bar\chi)}) & \text{if } d-\tau/\bar\chi < d\tau/\widetilde\beta, \tau(d-1/\bar\chi) \leq d\tau/\widetilde\beta. \end{cases}$$

Again by $N_0 \sim M_L^{\tau/(d\bar\chi)} (\sum_{\ell=0}^L E_\ell)^{1/(2\bar\chi)}$,

$$N_0 = \begin{cases} \lceil 2^{L\tau/\bar\chi} \rceil & \text{if } 1 < 1/(d\bar\chi), \\ \lceil 2^{L\tau/\bar\chi} L^{1/(2\bar\chi)} \rceil & \text{if } 1 = 1/(d\bar\chi), \\ \lceil 2^{(2\tau+d)L/(1+2\bar\chi)} L^{1/(1+2\bar\chi)} \rceil & \text{if } d-\tau/\bar\chi \geq d\tau/\widetilde\beta, \\ \lceil 2^{(\tau/\bar\chi + d\tau/(\widetilde\beta(1+2\bar\chi))L} \rceil & \text{if } 0 < d-\tau/\bar\chi < d\tau/\widetilde\beta < \tau(d-1/\bar\chi), \\ \lceil 2^{(2\tau+d\tau)L/(1+2\bar\chi)} \rceil & \text{if } 0 < d-\tau/\bar\chi < d\tau/\widetilde\beta, \tau(d-1/\bar\chi) \leq d\tau/\widetilde\beta. \end{cases} \tag{6.17}$$

Explicit error *vs.* work estimates are summarized as ε-complexity bounds in the following theorem, where we recall that work $= N_0 \sum_{\ell=0}^L E_\ell = M_L^{\tau/(d\bar\chi)} (\sum_{\ell=0}^L E_\ell)^{(1+2\bar\chi)/(2\bar\chi)}$.

Theorem 6.5 (Error *vs.* work for Strategy 1)**.** *Let the truncation dimension $(s_\ell)_{\ell\geq 1}$ be chosen according to Strategy 1 assuming $\widehat\beta > 1$ and $\tau > 1$. Let the assumptions (**A5**) and (**A3**) be satisfied for $\eta \geq 0$. The sample numbers for $Q_L^*(\cdot)$ are given by (6.15) and (6.14) for $\eta > 0$ and by (6.17) and (6.16) for $\eta = 0$, $L \in \mathbb{N}_0$. Let f and $G(\cdot)$ satisfy (**A4**).*

1. *Gaussian weight functions: for $\bar p \in (\max\{2/3, d/(\widehat\beta - 1)\}, 2)$, $\bar\chi = 1/(2\bar p) + 1/4 - \varepsilon'$ for $\varepsilon' > 0$ sufficiently small assuming $d/(\widehat\beta - 1) < 2$.*
2. *Exponential weight functions: for $\bar p \in (\max\{2/3, d/(\widehat\beta - 1)\}, 1]$, $\bar\chi = 1/\bar p - 1/2$ assuming $d/(\widehat\beta - 1) < 1$.*

For an error threshold $\varepsilon > 0$, we obtain

$$\sqrt{\mathbb{E}^{\boldsymbol{\Delta}}(|\mathbb{E}(G(u)) - Q_L^*(G(u^L))|^2)} = \mathcal{O}(\varepsilon)$$

is achieved with

$$\text{work}_L = \begin{cases} \mathcal{O}(\varepsilon^{-1/\bar\chi}) & \text{if } \max\{\tau, 1+\eta\} < \tau/(d\bar\chi), \\ \mathcal{O}(\varepsilon^{-1/\bar\chi} \log(\varepsilon^{-1})^{(1+2\bar\chi)/(2\bar\chi)}) & \text{if } \max\{\tau, 1+\eta\} = \tau/(d\bar\chi), \\ \mathcal{O}(\varepsilon^{-d/\tau(1+\eta)}) & \text{if } 1+\eta \geq \tau, 1+\eta > \tau/(d\bar\chi), \eta > 0, \\ \mathcal{O}(\varepsilon^{-d/\tau} \log(\varepsilon^{-1})) & \text{if } d-\tau/\bar\chi \geq d\tau/\widetilde\beta, \eta = 0, \\ \mathcal{O}(\varepsilon^{-d\min\{1, \max\{(1+\eta)/\tau, 1/\widetilde\beta + 1/(d\bar\chi)\}\}}) & \text{if } 1+\eta < \tau, 1 < d\bar\chi, \eta > 0, \\ \mathcal{O}(\varepsilon^{-(1/\bar\chi + d/\widetilde\beta)}) & \text{if } d-\tau/\bar\chi < d\tau/\widetilde\beta < \tau(d-1/\bar\chi), \eta = 0, \\ \mathcal{O}(\varepsilon^{-d}) & \text{if } d-\tau/\bar\chi < d\tau/\widetilde\beta, \tau(d-1/\bar\chi) \leq d\tau/\widetilde\beta, \eta = 0. \end{cases}$$

Here, $\widetilde\beta$ is as in (6.6) chosen close to $0 < 2\widehat\beta - d/2$ such that $\widetilde\beta < 2\widehat\beta - d/2$ and all implied constants are independent of L, $(s_\ell)_{\ell=0,\dots,L}$, $(N_\ell)_{\ell=0,\dots,L}$, and $(M_\ell)_{\ell\geq 0}$.

6.2. Application of error *vs.* work estimates: Matérn-like covariance

In practical applications, GRFs may be parametrized by MRAs, such that assumptions (**A1**) and (**A2**) are satisfied. A class of such GRFs represented by plain wavelets has been discussed in Section 6.1. The assumption (**A2**) is on the gradients of the function system and implies first order differentiability of the GRF. However, if the function system is sufficiently regular, also higher spatial regularity of realizations of the GRF may be deduced, which has been analyzed by the authors in Section 9 from [31]. Specifically, in the setting of Section 6.1, if the function system is able to characterize Besov norms (*i.e.*, [31], assumption (**A4**) holds), then by Proposition 18 of [31], the realizations of the GRF are μ-a.s. Hölder regular with exponent $\nu-\varepsilon>0$ for any $\nu>\varepsilon>0$ and $\nu=\widehat{\alpha}$. In this section, we aim to discuss the error *vs.* work estimates of Theorems 6.3 and 6.5 in dependence of the parameter ν (the smoothness of the GRF) or respectively the parameter $\widehat{\alpha}$ from Section 6.1 and d,η for a larger class of GRFs.

Solutions Z to the SPDE (1.1) may be represented by various function systems. If it is posed on D and also boundary conditions are prescribed certain wavelet bases may be used to solve (1.1). The function system that results by rearranging terms suitably for an expansion of Z with i.i.d. coefficients is generally not compactly supported, but decays exponentially, which means it is well-localized and may satisfy the bound (6.2) with $\widehat{\alpha}\approx\nu=\alpha-d/2$ (recall the parameter α from (1.1)), see ahead Section 7 for a particular choice and in particular Figures 1a and 1b. Wavelet bases on polytopal domains are available which also satisfy boundary conditions (see, *e.g.* [44] and the references there). If (1.1) is posed on $\mathbb{R}^d$ with $A(x)=\mathrm{Id}$ and $\kappa(x)=\bar{\kappa}$, $x\in\mathbb{R}^d$, then function systems to represent the GRF Z have been proposed in [8] by studying the covariance operator of Z. In this case it is well-known that the GRF Z has so called Matérn covariance [38] with smoothness parameter $\nu=\alpha-d/2$. The function system proposed in [8] is based on so-called *Meyer* wavelets and is globally supported, but well-localized in D. By Corollary 4.3 of [8], it satisfies assumption (**A1**) with $(b_j)_{j\geq1}\in\ell^p(\mathbb{N})$ for every $p>d/\nu$. The statement of Corollary 4.3 from [8] may be extended to gradients of the function system, which would imply that assumption (**A2**) could be satisfied with $(\bar{b}_j)_{j\geq1}\in\ell^{\bar{p}}(\mathbb{N})$ for every $\bar{p}>d/(\nu-1)$. We suppose that $\tau>1$ and let assumption (**A3**) be satisfied for $\eta>0$. The error *vs.* work estimates for the truncation Strategy 2 given in Theorem 6.3 may be applied using the borderline cases $\varepsilon'=0$ and $\bar{p}=d/(\nu-\tau)$ in our error bounds. It follows that accuracy $\varepsilon>0$ may be achieved using Gaussian weight functions with

$$\text{work}=\begin{cases}\mathcal{O}(\varepsilon^{-2d/(\nu-\tau+d/2)-\delta}) & \text{if } 1+\eta<2\tau/(\nu-\tau+d/2),\\ \mathcal{O}(\varepsilon^{-d/\tau(1+\eta)}) & \text{if } 1+\eta>2\tau/(\nu-\tau+d/2),\end{cases}$$

for any $\delta>0$, provided that

$$\frac{d}{2}+\tau<\nu<\frac{3d}{2}+\tau.$$

In case that $\nu\geq 3d/2+\tau$, the work is also $\mathcal{O}(\varepsilon^{-d/\tau(1+\eta)})$, which is the complexity of the Poisson problem under the assumption that (**A3**) is satisfied for $\eta>0$. Note that the case $1+\eta=2\tau/(\nu-\tau+d/2)$ is not considered, since the formal value $\bar{p}=d/(\nu-\tau)$ was used instead of $\bar{p}+\varepsilon'=d/(\nu-\tau)$ for some $0<\varepsilon'\ll1$, which renders this case unimportant.

Suppose that $1+\eta<\tau$. The error *vs.* work estimates for the truncation Strategy 1 given in Theorem 6.5 are applied using the borderline case $\varepsilon'=0$ and $\bar{p}=d/(\nu-1)$ and $\widetilde{\beta}=2\nu-d/2$. It follows that accuracy $\varepsilon>0$ may be achieved (based on QMC error bounds with the norm (3.1) and with Gaussian weight functions $w_{\mathrm{g},j}$ in (3.2)) with

$$\text{work}=\begin{cases}\mathcal{O}(\varepsilon^{-2d/(\nu-1+d/2)-\delta}) & \text{if } \nu+d/2<3,\\ \mathcal{O}(\varepsilon^{-d\min\{1,1/(2\nu-d/2)+2/(\nu-1+d/2)\}-\delta}) & \text{if } \nu+d/2>3, (1+\eta)/\tau<1/(2\nu-d/2)+2/(\nu-1+d/2),\\ \mathcal{O}(\varepsilon^{-d/\tau(1+\eta)}) & \text{if } \nu+d/2>3, (1+\eta)/\tau>1/(2\nu-d/2)+2/(\nu-1+d/2),\end{cases}$$

for any $\delta>0$, provided that the Matérn parameter ν satisfies

$$\frac{d}{2}+1<\nu<\frac{3d}{2}+1. \tag{6.18}$$

For $3d/2+1\leq\nu<3d/2+\tau$, intermediate cases hold according to Theorem 6.5 with $\bar{\chi}\approx1$. If $\nu\geq 3d/2+\tau$, Strategy 2 is applicable as mentioned above.

We discuss the cases $\nu \in \{d+1, 3d/2+1\}$, which allow error bounds with both, Strategy 1 and Strategy 2. We suppose maximal regularity $\tau = 2$, which is for example the case if $f, G(\cdot) \in L^2(D)$. For $\nu = 3d/2+1$, accuracy $\varepsilon > 0$ is achievable with Strategy 1 with work $\mathcal{O}(\varepsilon^{-9/7-\delta})$ (for $d = 2$) and $\mathcal{O}(\varepsilon^{-d/\tau(1+\eta)})$ (for $d = 3$). With Strategy 2, work $\mathcal{O}(\varepsilon^{-4/3-\delta})$ (for $d = 2$) and $\mathcal{O}(\varepsilon^{-d/\tau(1+\eta)})$ (for $d = 3$) is needed. In the case $\nu = d+1$, accuracy $\varepsilon > 0$ is achievable using Strategy 1 with work $\mathcal{O}(\varepsilon^{-26/15-\delta})$ (for $d = 2$) and $\mathcal{O}(\varepsilon^{-210/117-\delta})$ (for $d = 3$). With Strategy 2, work $\mathcal{O}(\varepsilon^{-2-\delta})$ (for $d = 2$) and $\mathcal{O}(\varepsilon^{-12/7-\delta})$ (for $d = 3$) is needed. Note that $1.71 \approx 12/7 < 210/117 \approx 1.79$. We assumed that $\eta > 0$ in (**A3**) is sufficiently small (the theory in [29] implies that $\eta > 0$ may be chosen arbitrarily small for MLMC-FEM).

6.3. Local supports and product weights *vs.* global supports and POD weights

We suppose that the GRF Z can be represented with a Karhunen–Loève expansion given by the eigenpairs $(\lambda_j, \psi_j^{\mathrm{KL}})_{j\geq 1}$ of the covariance operator of Z normalized in $L^2(D)$. Thus, $Z = \sum_{j\geq 1} y_j \psi_j^{\mathrm{gl}}$, where $\psi_j^{\mathrm{gl}} = \sqrt{\lambda_j}\psi_j^{\mathrm{KL}}$, $j \geq 1$. In the case that Z has Matérn covariance, Corollary 5 of [26] implies that there exists $c > 0$ such that $\sqrt{\lambda_j} \leq c j^{-(1/2+\nu/d)}$, $j \geq 1$. By the proof of Proposition 9 from [26], generally in the Matérn case for any $\delta > 0$ there is $c > 0$ such that $\|\psi_j^{\mathrm{KL}}\|_{L^\infty(D)} \leq c j^{1/2+\delta}$ and $\||\nabla \psi_j^{\mathrm{KL}}|\|_{L^\infty(D)} \leq c j^{1/2+1/d+\delta}$, $j \geq 1$. As observed in Remark 10 from [36], this implies $(\||\nabla\psi_j^{\mathrm{gl}}|\|_{L^\infty(D)})_{j\geq 1} \in \ell^{\bar{p}_{\mathrm{POD}}}(\mathbb{N})$ for any $\bar{p}_{\mathrm{POD}} > d/(\nu-1)$. Also, $(\|\psi_j^{\mathrm{gl}}\|_{L^\infty(D)})_{j\geq 1} \in \ell^{p_{\mathrm{POD}}}(\mathbb{N})$ for any $p_{\mathrm{POD}} > d/\nu$. For the MLQMC error analysis in [36] to be applicable, it is required that $\bar{p}_{\mathrm{POD}} < 1$. The regime that $\bar{p}_{\mathrm{POD}} \in (2/3, 1)$ is equivalent to

$$d + 1 < \nu < \frac{3d}{2} + 1, \tag{6.19}$$

which is a more restrictive condition than what was required in (6.18) for Strategy 1 of the local support theory with product weights to be applicable. Under this assumption, Corollary 2 of [36] implies that an accuracy $\varepsilon > 0$ may be achieved (*excluding* the computational cost of the CBC algorithm) using MLQMC with POD weights with

$$\text{work} = \begin{cases} \mathcal{O}(\varepsilon^{-d/(\nu-1-d/2)-1/\alpha'-\delta}) & \text{if } \tau > \nu - 1 - d/2, \\ \mathcal{O}(\varepsilon^{-d/\tau-1/\alpha'}) & \text{if } \tau < \nu - 1 - d/2, \end{cases} \tag{6.20}$$

for any $\delta > 0$. If $\nu \geq 3d/2+1$, the work is also $\mathcal{O}(\varepsilon^{-d/\tau-1/\alpha'})$. The value of the parameter α' has been shown in Proposition 9 from [26] to be at least $\nu/d - 1/2 - \delta$ for any $\delta > 0$. In the setting of Corollary 2 from [36] $s_\ell = s_L$ for all levels $\ell = 0, \ldots, L-1$, where $L \in \mathbb{N}$ is the maximal level.

Remark 6.6. Consider a convex polygon $D \subset \mathbb{R}^2$, *i.e.*, $d = 2$, and $f, G(\cdot) \in L^2(D)$. This implies that $\tau = 2$. In the borderline case $\nu = 3d/2 + 1 = 4$ of the estimate (6.20), accuracy $\varepsilon > 0$ may be achieved using MLQMC-FEM with $(\psi_j^{\mathrm{gl}})_{j\geq 1}$ and POD weights with work $\mathcal{O}(\varepsilon^{-5/3-\delta})$. In the same situation accuracy $\varepsilon > 0$ may be achieved using Strategy 1 of MLQMC-FEM with locally supported ψ_j and QMC with product weights in work $\mathcal{O}(\varepsilon^{-9/7-\delta})$. In this comparison we assumed that $\eta > 0$ in (**A3**) is sufficiently small (the theory in [29] implies that $\eta > 0$ may be chosen arbitrarily small for MLMC-FEM). Multilevel QMC with local supports and product weights requires here less work to achieve target accuracy $\varepsilon > 0$ than MLQMC-FEM with globally supported ψ_j and POD QMC weights as presented in [36], even when the cost of the CBC construction of the QMC generating vector for POD weights is excluded from the work estimates as in [36], where this (excluded) work depends quadratically on the truncation dimension. Also, the work estimates from [36] assumed the cost of evaluating ψ_j^{KL} to be $\mathcal{O}(1)$ for every $j \geq 1$.

Remark 6.7. The parameter α' increases linearly with respect to the parameter ν. The work estimate in (6.20) includes always the term $1/\alpha'$. This is a consequence of the work model that the cost of assembling instances of stiffness matrices has cost proportional to the number of FE degrees of freedom multiplied by the truncation dimension s_L. If $\nu \gg 1$, then $1/\alpha' \ll 1$ and this term becomes insignificant and since $s_L \sim h_L^{-\tau/\alpha'}$ the uncertainty quantification problem is then of moderate effective dimension. The resulting work would still be $\mathcal{O}(\varepsilon^{-d/\tau-1/\alpha'})$

and τ is at most two, since first order FE is used in the MLQMC algorithm. Another restriction is the rate of the QMC quadrature, which is of essentially first order for randomly shifted lattice rules. In the case that a higher order QMC rule were available (which is the case for bounded parameter vectors, *cf.* [15]) and is used (for simplicity) as single-level QMC in combination with the function system $(\psi_j^{\mathrm{gl}})_{j\geq 1}$ stemming from the Karhunen–Loève expansion and QMC with POD weights, then the order of convergence that could be achieved is only restricted by the value of ν given that proper mesh refinement is available. Approximation rates of Z by function systems based on spline wavelets are generally restricted by their maximal order, which limits corresponding complexity estimates of single-level QMC with local supports.

In certain cases the Karhunen–Loève functions may have stronger properties concerning their decay in the sup-norm and the computational cost to compute linear combinations of them efficiently. One of these cases should also be discussed in order to provide a more thorough comparison of locally supported ψ_j and QMC product weights to globally supported and QMC POD weights. As discussed in Section 10 from [31], if the GRF Z is computed on a product domain D and $A(x) = \mathrm{Id}$, $\kappa(x) = \text{const} > 0$ in (1.1) (which is solved on $\mathbb{R}^d$ in this case), the Karhunen–Loève functions may take the form of products of trigonometric functions and are thus uniformly bounded in the sup-norm. I.e., there exists $c > 0$ such that for all $j \geq 1$ holds $\| |\nabla\psi_j^{\mathrm{KL}}| \|_{L^\infty(D)} \leq c j^{1/d}$. Thus, $(\| |\nabla\psi_j^{\mathrm{gl}}| \|_{L^\infty(D)})_{j\geq 1} \in \ell^{\bar{p}_{\mathrm{POD}}}(\mathbb{N})$ for any $\bar{p}_{\mathrm{POD}} > d/(\nu-1+d/2)$. Also the truncated GRF may be computed efficiently with FFT techniques. To also accommodate these cases with variable truncation dimension (s_ℓ not necessarily equal to s_L), in Appendix A the MLQMC error analysis from [36] is suitably extended. In Appendix B we take into account the cost of the CBC algorithm, which in the POD weight case grows quadratically with respect to the truncation dimensions. The error analysis there applies to Matérn GRFs with Matérn parameter $\nu > d/2+1$. To illustrate the results, consider a convex polygon or polyhedron $D \subset \mathbb{R}^d$ with $d = 2, 3$. Then, the Dirichlet Laplacian is boundedly invertible from $H^2(D) \cap V$ to $L^2(D)$ and $f, G(\cdot) \in L^2(D)$. This implies the FE convergence rate $\tau = 2$ for FE on quasiuniform meshes in D. For $\nu = 3d/2+1$, Theorem B.1 may be applied with the borderline cases $p_{\mathrm{POD}} = (2 + 1/d)^{-1}$ and $\bar{p}_{\mathrm{POD}} = 3/2$. Note that the parameter $\alpha = (1/p_{\mathrm{POD}} - 1/\bar{p}_{\mathrm{POD}})^{-1}$ in Theorem B.1 is then $\alpha = 2d/(2+d)$. The accuracy $\varepsilon > 0$ may be achieved using MLQMC-FEM with $(\psi_j^{\mathrm{gl}})_{j\geq 1}$ and POD weights with work $\mathcal{O}(\varepsilon^{-4d/(3d+2)-1-\delta})$ taking into account the cost of the CBC algorithm. Note that $4d/(3d + 2) + 1 = 2$ for $d = 2$ and $4d/(3d + 2) + 1 = 1 + 12/11$ for $d = 3$. In the same situation, MLQMC-FEM with local support and product weights achieves accuracy ε with work $\mathcal{O}(\varepsilon^{-9/7-\delta})$ for $d = 2$ and Strategy 1 and $\mathcal{O}(\varepsilon^{-4/3-\delta})$ for $d = 2$ and Strategy 2. For $d = 3$, the corresponding required work for MLQMC-FEM with local support and product weights is $\mathcal{O}(\varepsilon^{-d/\tau(1+\eta)})$ for either of Strategy 1 and 2. In this comparison we assumed the larger value $\alpha' \approx \nu/d$ for the truncation error with respect to the globally supported function system $(\psi_j^{\mathrm{gl}})_{j\geq 1}$, which was shown in this case in Proposition 20 from [31]. In Figure 1 of [36], the value α' has been investigated empirically for $d = 1$ and $\nu = 1$, the functional $G(\cdot)$ being point evaluation. The shown empirical data suggests the higher value of $\approx 2\nu/d$ in this case. Assuming this stronger decay of the truncation error with respect to s_L and $(\psi_j^{\mathrm{gl}})_{j\geq 1}$, the required work for MLQMC with $(\psi_j^{\mathrm{gl}})_{j\geq 1}$ and POD weights is $\mathcal{O}(\varepsilon^{-3/2-\delta})$ for $d = 2$ and $\mathcal{O}(\varepsilon^{-1-6/11-\delta})$ for $d = 3$. Throughout this paragraph, $\delta > 0$ was an arbitrarily small number and the constants hidden in the $\mathcal{O}(\cdot)$ depend on δ. In these cases MLQMC with local support and product weights achieves prescribed accuracy ε with an asymptotically smaller work compared to MLQMC with $(\psi_j^{\mathrm{gl}})_{j\geq 1}$ and POD weights.

In the case that ν is sufficiently large also MLQMC with $(\psi_j^{\mathrm{gl}})_{j\geq 1}$ and POD weights is able to achieve accuracy $\varepsilon > 0$ with work $\mathcal{O}(\varepsilon^{-d/\tau(1+\eta)})$. This is the work to solve one instance of the FE method on the finest mesh level. For example for $d = 2$, $\tau = 2$, and $\nu = 6$, Theorem B.1 is applicable with the borderline values $p_{\mathrm{POD}} = 2/7$, $\bar{p}_{\mathrm{POD}} = 2/3$, and $\alpha = 1/2$ in the case that the Karhunen–Loève functions are uniformly bounded. Due to the term $1/\alpha'$ in the work estimate in Corollary 2 of [36], this situation was not achievable with the theory presented in [36]. Note again that we assumed that $\eta > 0$ in (**A3**) is sufficiently small (the theory in [29] implies that $\eta > 0$ may be chosen arbitrarily small for MLMC-FEM).

7. Numerical experiments

We illustrate the complexity estimates and algorithmic details on GRF generation in locally supported representation systems with numerical tests. To this end, we consider the following class of GRFs. We admit GRFs Z which are sample-wise, weak solutions to the SPDE (1.1), where $\mathcal{W}$ denotes spatial white noise on D. See *e.g.* [1] for details on this. In (1.1), we assume that $A(x) \in \mathbb{R}^{d\times d}$ is symmetric for a.e. $x \in D$ and there exists $\overline{A} > 0$ such that

$$\operatorname*{ess\,inf}_{x\in D} \xi^\top A(x)\xi \geq \overline{A}\xi^\top \xi, \quad \forall \xi \in \mathbb{R}^d,$$

and $\operatorname{ess\,inf}_{x\in D} \kappa(x) > 0$.

We recall from Section 1 that if $D = \mathbb{R}^d$, $A(x) \equiv \mathrm{Id}$, and if also $\kappa(x) \equiv$ const, then the stationary solution Z to (1.1) is well-known to have Matérn covariance. As proposed in [38], the SPDE (1.1) can be used to define and numerically sample non-stationary GRFs in bounded domains and for general coefficients κ and A, which accommodates non-stationary GRFs Z. As in [38], both in stationary and non-stationary cases (see [38], Sect. 3.2), we shall refer to solutions to (1.1) as *Matérn fields*.

We choose $D = (0, 1)$ with periodic boundary conditions, which can be identified with the one-dimensional sphere or the one-dimensional torus $\mathbb{T}^1$. To obtain a series expansion of Z with i.i.d. standard normally distributed coefficients, *i.e.*, the form of (1.3), we discretize (1.1) by biorthogonal and continuous, piecewise linear spline prewavelets as in [46]. Let $(V_\ell)_{\ell\geq 0}$ be a sequence of FE spaces of piecewise affine functions on uniformly refined meshes with mesh width $h_\ell = 2^{-\ell-2}$ and $\dim(V_\ell) = 2^{\ell+2}$. Each FE space is spanned by continuous, piecewise affine functions, *i.e.*, $V_\ell = \operatorname{span}\{\varphi^\ell_1, \ldots, \varphi^\ell_{2^{\ell+2}}\}$, $\ell \geq 0$, where $\varphi^\ell_1, \ldots, \varphi^\ell_{2^{\ell+2}}$ are the "hat" function basis. We shall use the following representation system for the GRF in D: for every $\ell \in \mathbb{N}$, define the spline-prewavelets as in equation (4.7) from [46] by

$$\phi_{\ell,k} := \sum_{\iota=1}^{5} a_\iota \varphi^\ell_{2k-4+\iota}, \quad k = 1, \ldots, N_\ell, a = (1/2, -3, 5, -3, 1/2), \tag{7.1}$$

where $N_\ell = 2^{\ell+1}$ and subscript indices of φ^ℓ_k are taken modulo $2^{\ell+2}$ plus 1. The barycenter of the support of the prewavelet $\phi_{\ell,k}$ is $x_{\ell,k} = (2k-1)2^{-\ell-2}$. We define the wavelet spaces $W^\ell := \operatorname{span}\{\phi_{\ell,1}, \ldots, \phi_{\ell,N_\ell}\}$, $\ell \geq 1$, with the understanding that $W^0 := V_0$. For $\ell = 0$, we define $\phi_{\ell,k} = \varphi^\ell_k$, $k = 1, \ldots, N_0 := 4$. These spaces are L^2-orthogonal across levels, *i.e.*, $\int_D w_1 w_2 \mathrm{d}x = 0$ for all $w_1 \in W^{\ell_1}$, $w_2 \in W^{\ell_2}$ s.t. $\ell_1, \ell_2 \in \mathbb{N}_0$ and $\ell_1 \neq \ell_2$. Hence, we obtain the multilevel splitting

$$V^\ell = W^0 \oplus W^1 \oplus \ldots \oplus W^\ell, \quad \ell \geq 1.$$

We note that $\{\phi_{\ell,k} : \ell \geq 0, k = 1, \ldots, N_\ell\}$ is a Riesz basis of $L^2(\mathbb{T}^1)$ and of $H^1(\mathbb{T}^1)$. Upon proper scaling, there hold stable norm equivalences in scale of spaces $H^t(\mathbb{T}^1)$ for $t \in [0, 3/2)$, *cf.* Proposition 4.1 of [46]. There are constants C_1, C_2 such that for every $L \geq 0$ and for every $v = \sum_{\ell=0}^{L} \sum_{k=1}^{N_\ell} v_{\ell,k}\phi_{\ell,k}$

$$C_1 \sum_{\ell=0}^{L} 2^{2(t-1/2)\ell} \sum_{k=1}^{N_\ell} |v_{\ell,k}|^2 \leq \|v\|^2_{H^t(\mathbb{T}^1)} \leq C_2 \sum_{\ell=0}^{L} 2^{2(t-1/2)\ell} \sum_{k=1}^{N_\ell} |v_{\ell,k}|^2.$$

Let us define the sequence space $\ell^2_1 := \{\boldsymbol{c} \in \mathbb{R}^{\mathbb{N}} : \sum_{\ell\geq 0} 2^\ell \sum_{k=1}^{N_\ell} |c_{\ell,k}|^2 < \infty\}$ that corresponds to $H^1(\mathbb{T}^1)$. The white noise $\mathcal{W}$, applied to the prewavelets (7.1), results in a random vector $\boldsymbol{W}$ with components that are normally distributed with zero mean and covariance determined by the "mass matrix" $\boldsymbol{M}$, *i.e.*, $\operatorname{cov}(\mathcal{W}(\phi_{\ell_2,k_2}), \mathcal{W}(\phi_{\ell_1,k_1})) = \int_D \phi_{\ell_2,k_2}\phi_{\ell_1,k_1} \mathrm{d}x =: \boldsymbol{M}_{\ell_1,k_1,\ell_2,k_2}$, $\ell_1, \ell_2 \geq 0$, $k_1 \in \{1, \ldots, N_{\ell_1}\}$, $k_2 \in \{1, \ldots, N_{\ell_2}\}$. Due to the orthogonality of the prewavelets across levels the bi-infinite mass matrix $\boldsymbol{M}$ is block diagonal with diagonal blocks given by the mass matrices of W^ℓ: there is no correlation between the different levels. This will be convenient in sampling realizations of $\boldsymbol{W}$ with a block Cholesky algorithm; its complexity is $\mathcal{O}(N_\ell)$ on every block. Note that the random vectors $\boldsymbol{W}$ and $\boldsymbol{L}\boldsymbol{y}$ have the same distribution for an i.i.d. standard normally distributed $\boldsymbol{y}$ and any operator $\boldsymbol{L}$ that satisfies $\boldsymbol{L}\boldsymbol{L}^\top = \boldsymbol{M}$. For example, $\boldsymbol{L}$ can be the operator that results by applying the Cholesky algorithm to every block of $\boldsymbol{M}$. For $\alpha = 2$, the operator in (1.1) is local and reads $(-\mathrm{div}(A(x)\nabla) + \kappa^2(x))$ It can be equivalently represented as a bi-infinite matrix in a prewavelet basis. We denote the resulting (bi-infinite) matrix by $\boldsymbol{A}$, where $\boldsymbol{A}_{\ell_1,k_1,\ell_2,k_2} = \int_D (\nabla\phi_{\ell_2,k_2})^\top A \nabla\phi_{\ell_1,k_1} + \kappa^2\phi_{\ell_2,k_2}\phi_{\ell_1,k_1}\mathrm{d}x$, $\ell_1,\ell_2 \geq 0$, $k_1 \in \{1,\dots,N_{\ell_1}\}$, $k_2 \in \{1,\dots,N_{\ell_2}\}$.

For $\alpha = 2$, the variational formulation of (1.1) with respect to this prewavelet basis is: for a given parameter vector $\boldsymbol{y}$, find $\boldsymbol{Z}(\boldsymbol{y}) \in \ell_1^2$ such that

$$\boldsymbol{A}\boldsymbol{Z}(\boldsymbol{y}) = \boldsymbol{L}\boldsymbol{y},$$

where the parametric coefficients $\boldsymbol{Z}(\boldsymbol{y})$ and the GRF Z are related by

$$Z(\boldsymbol{y}) = \sum_{\ell\geq 0}\sum_{k=1}^{N_\ell} \boldsymbol{Z}(\boldsymbol{y})_{\ell,k}\phi_{\ell,k}$$

Let us define $\boldsymbol{e}(\ell,k)$ by $\boldsymbol{e}(\ell,k)_{\ell',k'} := 1$ if $\ell = \ell'$ and $j = k'$ and zero otherwise. A piecewise linear multiresolution representation of the Matérn field Z with i.i.d. coefficients and a function system $(\psi_{\ell,k}^{\alpha/2})_{\ell\geq 0,k=1,\dots,N_\ell}$ (here $\alpha/2 = 1$) is now obtained by

$$Z(\boldsymbol{y}) = \sum_{\ell\geq 0}\sum_{k=1}^{N_\ell} y_{\ell,k}\psi_{\ell,k}^1, \quad \text{with} \quad \psi_{\ell,k}^1 = \sum_{\ell'\geq 0}\sum_{k'=1}^{N_{\ell'}} (\boldsymbol{A}^{-1}\boldsymbol{L}\boldsymbol{e}(\ell,k))_{\ell',k'}\phi_{\ell',k'}, \tag{7.2}$$

where $\ell \geq 0$ is the level or dilation index and k the translation. For $\alpha > 2$, let $\{\alpha/2\}$ be the fractional part of $\alpha/2$ and $\lfloor\alpha/2\rfloor = \alpha/2 - \{\alpha/2\}$. The SPDE (1.1) is rewritten recursively as

$$\begin{aligned} (-\mathrm{div}(A(x)\nabla) + \kappa^2(x))^{\{\alpha/2\}} Z &= Z_i \\ (-\mathrm{div}(A(x)\nabla) + \kappa^2(x)) Z_i &= Z_{i-1}, \quad i = 1,\dots,\lfloor\alpha/2\rfloor, \end{aligned} \tag{7.3}$$

where $Z_0 := \mathcal{W}$. If $\{\alpha/2\} = 0$, $(-\mathrm{div}(A(x)\nabla) + \kappa^2(x))^0$ is understood as the identity operator. For $\alpha \in 2\mathbb{N}$, there is no fractional PDE to be solved in (7.3) and by standard Galerkin techniques

$$\psi_{\ell,k}^{\alpha/2} = \sum_{\ell'\geq 0}\sum_{k'=1}^{N_{\ell'}} ((\boldsymbol{A}^{-1}\boldsymbol{M})^{\alpha/2-1}\boldsymbol{A}^{-1}\boldsymbol{L}\boldsymbol{e}(\ell,k))_{\ell',k'}\phi_{\ell',k'}, \quad \ell \geq 0, k = 1,\dots,N_\ell. \tag{7.4}$$

Generally, for $\alpha \geq 2$,

$$\psi_{\ell,k}^{\alpha/2} = \boldsymbol{A}_{\{\alpha/2\}}^{-1}\boldsymbol{M}\psi_{\ell,k}^{\lfloor\alpha/2\rfloor} \quad \ell \geq 0, k = 1,\dots,N_\ell,$$

where $\boldsymbol{A}_{\{\alpha/2\}}$ is the wavelet representation of the fractional operator $(-\mathrm{div}(A(x)\nabla) + \kappa^2(x))^{\{\alpha/2\}}$. In the case that $\{\alpha/2\} > 0$, a fractional PDE with a non-local operator needs to be discretized. Efficient FE methods for the numerical solution of this problem have been recently analyzed in [9]. We remark that the application of first order prewavelets is sufficient here, since the convergence rate of randomly shifted lattice rules is limited by one and also first order FE is considered to discretize (1.2) in space.

Remark 7.1. The application of the sparse operator $\boldsymbol{L}$ introduces a weight sequence $(2^{-(1/2)\ell})_{\ell\geq 0,j=1,\dots,N_\ell}$ and the application of the inverse of $\boldsymbol{A}$ introduces an additional weight sequence $(2^{-\ell})_{\ell\geq 0,k=1,\dots,N_\ell}$ on the parameter vector $\boldsymbol{y}$ or $e(\ell,j)$. The theory of pseudodifferential operators and wavelet compression suggests that $(\psi_{\ell,k}^{\alpha/2})_{\ell\geq 0,k=1,\dots,N_\ell}$ satisfies assumptions (**A1**) and (**A2**) with $b_{\ell,k} = 2^{-\widehat{\beta}\ell}$ and $\bar{b}_{\ell,k} = 2^{-(\widehat{\beta}-1)\ell}$, $\ell \geq 1$, $k = 1,\dots,N_\ell$, for all $1 < \widehat{\beta} < \alpha - 1/2$. See Figures 1a and 1b ahead for an illustration of this property. In practical implementations the infinite series in (7.2) needs to be truncated. A detailed analysis of these aspects in a more general setting will be presented in a forthcoming work.

For a scaling parameter $\theta > 0$ to be specified, the parametric PDE (1.2) with log-Matérn input $a = \exp(\theta Z)$ and right hand side $f \in L^2(\mathbb{T}^1)$ such that $\int_{\mathbb{T}^1} f \mathrm{d}x = 0$ is discretized by the FE spaces $(V_\ell)_{\ell \geq 0}$, which are spanned by the standard hat functions. We recall the variational formulation: for all $\ell \geq 0$, find $u^{h_\ell} : \Omega \to V_\ell$ such that

$$\int_{\mathbb{T}^1} a \, (u^{h_\ell})' v' \mathrm{d}x = \int_{\mathbb{T}^1} f v \mathrm{d}x, \quad \forall v \in V_\ell, \quad \text{and} \quad \int_{\mathbb{T}^1} u^{h_\ell} \mathrm{d}x = 0. \tag{7.5}$$

Due to periodicity, no essential boundary conditions enter the variational formulation (7.5). The vanishing mean condition on the solution is sufficient to ensure well-posedness, since the kernel of the differential operator $[w \mapsto -(aw')']$ comprises exactly the constant functions.

We will adopt Strategy 2 from Section 6.1, which means that on every discretization level $\ell \geq 0$ we set $s_\ell = M_\ell$. We also truncate the expansion of the $\psi^1_{\ell,k}$'s in (7.2) and (7.4) and the *infinite* matrices $\boldsymbol{A}$ to M_ℓ terms. Note that the *bi-infinite* matrices $\boldsymbol{M}$ and $\boldsymbol{L}$ do not need to be truncated as they are block-diagonal. The work to compute $Z^{s_\ell}(\boldsymbol{y})$ for a given $\boldsymbol{y}$ in a continuous, piecewise linear representation is $\mathcal{O}(M_\ell \log(M_\ell))$, since it amounts to the solution of a PDE discretized by prewavelets and the application of the Cholesky algorithm to sparse band matrices. The work for the approximate solution of the PDE discretized with the hat functions on mesh-level ℓ scales as $\mathcal{O}(M_\ell)$.

In our numerical tests, $D = (0, 1)$, $f(x) = \sin(2\pi x)$ with periodic boundary conditions and

$$\kappa^2(x) = \overline{\kappa}^2 \left(1 + \frac{1}{2} \sin(2\pi x)\right), \quad A(x) = \mathrm{Id}, \quad \alpha \in \{2, 4\}. \tag{7.6}$$

The parameter ν such that $\alpha = \nu + d/2$ determines the sample regularity of realizations of the Matérn field. We test the cases $\nu \in \{3/2, 7/2\}$ with the *correlation length* scale parameter

$$\lambda = \frac{2\sqrt{\nu}}{\overline{\kappa}},$$

which then determines the value of $\bar{\kappa}$ in (7.6). For $\kappa(x) \equiv \overline{\kappa}$ this corresponds to the correlation length parameter used in numerical experiments in Section 4 from [36]. We aim at testing for different values of λ without greatly affecting the variance of Z. Thus, we set the scaling θ of the (non-stationary) GRF Z to

$$\theta = \frac{\sigma_0}{\sigma(\alpha, \overline{\kappa})}$$

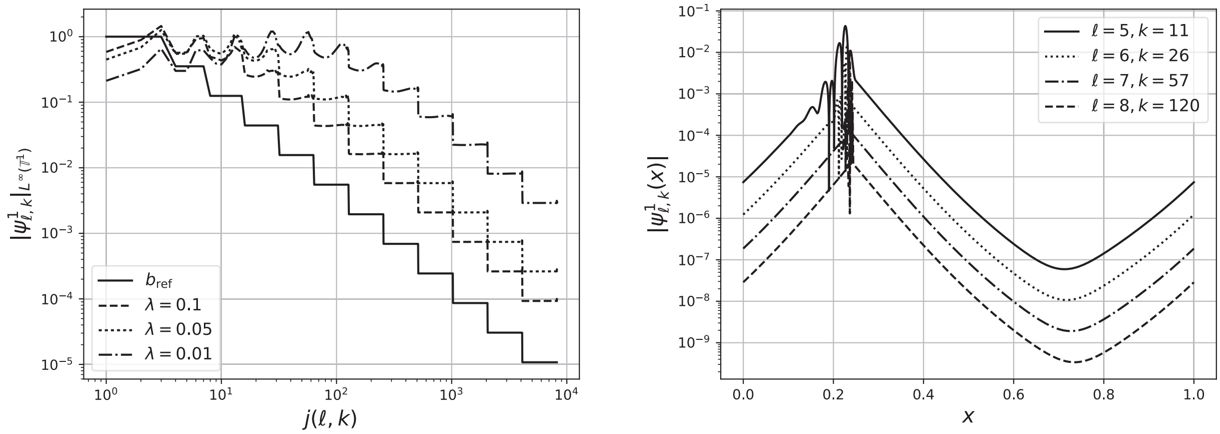
with $\sigma_0 > 0$ still at our disposal. In the stationary ("Matérn") case, *i.e.*, when $A(x) \equiv \mathrm{Id}$ and $\kappa(x) \equiv \overline{\kappa} > 0$, elementary Fourier analysis reveals that the marginal variance is given by

$$\sigma^2(\alpha, \overline{\kappa}) := \frac{1}{\overline{\kappa}^{2\alpha}} + \sum_{i \geq 1} \frac{2}{((2\pi i)^2 + \overline{\kappa}^2)^\alpha}.$$

The function systems $(\psi^{\alpha/2}_{\ell,k} : \ell \geq 0, k = 1, \ldots, N_\ell)$ are well localized and satisfy a decay condition which suits our MLQMC analysis with product weights. To illustrate this numerically, we define

$$b_{\mathrm{ref}} := (2^{-(3/2)\ell})_{\ell \geq 0, k=1,\ldots,N_\ell}.$$

In Figure 1a, we plot this reference sequence and $\|\psi^1_{\ell,k}\|_{L^\infty(\mathbb{T}^1)}$ for several choices of the correlation length $\lambda > 0$, indexed by $j(\ell, k)$. Figure 1b shows plots of $|\psi_{\ell,k}|$ for several values of ℓ and k in log-scale. For the illustrations in both figures, the expansion in (7.2) has been truncated to the maximal level $L = 11$. For $\alpha = 2$, Figures 1a and 1b suggest that $(\psi^1_{\ell,k})_{\ell \geq 0, k=1,\ldots,N_\ell}$ satisfies assumptions (**A1**) and (**A2**) with weight sequence $b^{1-\varepsilon}_{\mathrm{ref}}$ for every $\varepsilon > 0$ and $(\bar{b}_{\ell,k})_{\ell \geq 0, k=1,\ldots,N_\ell}$ defined according to (6.5) with $\widehat{\beta} \approx 3/2$, respectively.



(a) $\sigma_0 = 1.0$ (b) $\sigma_0 = 1.0$, $\lambda = 0.1$

FIGURE 1. Quantitative properties of $(\psi^{\alpha/2}_{\ell,k})_{\ell\geq 0,k=1,\dots,N_\ell}$ for $\sigma_0 = 1.0$, $\lambda \in \{0.1, 0.05, 0.01\}$, $\alpha = 2$ ($\nu = 3/2$). *Left panel*: decay of $\|\psi^1_{\ell,k}\|_{L^\infty(\mathbb{T}^1)}$ *vs.* j. Onset of asymptotic decay appears for $j \sim \lambda^{-1}$. *Right panel*: localization of $\psi^1_{\ell,k}$. Semi-logarithmic plot of the modulus of $\psi^1_{\ell,k}(x)$. Exponential decay away from barycenter $x_{\ell,k}$ of $\phi_{\ell,k}$.

Remark 7.2. For stationary GRFs, translation invariance implies that the Karhunen–Loève basis is trigonometric, and QMC error bounds from [26,36] with globally supported Karhunen–Loève basis functions and QMC integration with POD weights are applicable. When $A(x) \equiv \mathrm{Id}$ and $\kappa(x) \equiv \mathrm{const}$ and periodic boundary conditions are imposed on ∂D, the GRF Z is stationary. The SPDE (1.1) can be numerically solved by Fourier methods. We refer to [16,27] for details on this. Unlike the product weights for QMC integration which were derived in the present work, the appearance of QMC weights with POD structure in [26,36] implies that the construction cost for these QMC integration methods scales as $O(s^2 N + sN\log(N))$ [41].

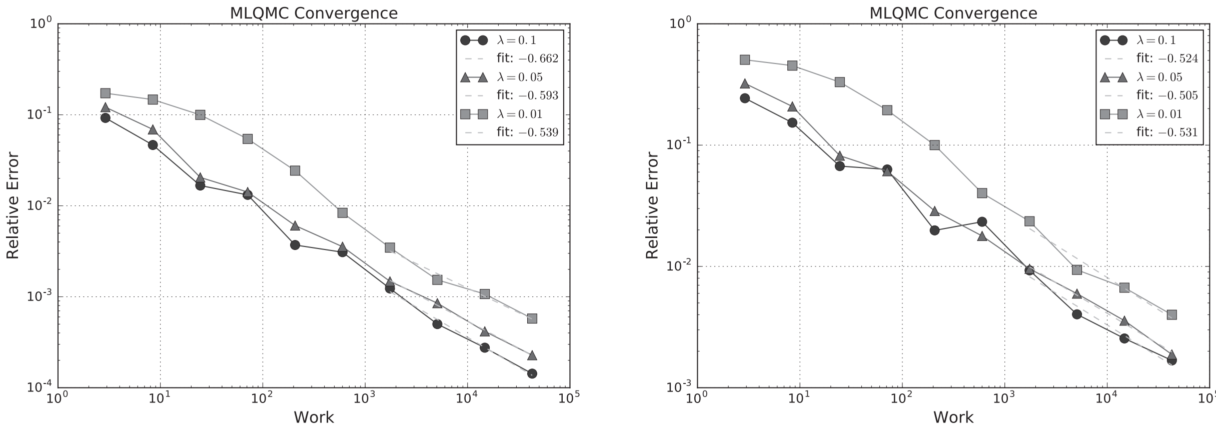
In our numerical tests, we consider the functional $G(v) := v(x_0)$ with $x_0 = 0.7$, which is not a node in any of our FE meshes for all levels $\ell \geq 0$. Note that $G(\cdot) \in H^{-1/2-\varepsilon}(\mathbb{T}^1)$ for every $\varepsilon > 0$. Since the QMC rate $\bar{\chi}$ is restricted to $[1/2, 1)$ the complexity estimate in Theorem 6.3 in the regime $1 < \tau/\bar{\chi}$ (here $d = 1$, $\eta = 0$) does not seem to benefit from $\tau > 1$. Thus, QMC sample numbers are chosen according to (6.12) and (6.13) with $\tau = 1$ and $\bar{\chi}$ to be specified, *i.e.*,

$$N_\ell = \left\lceil 2^{(\tau/\bar{\chi})(L+1)} \left(2^{-(2\tau+1)(\ell+1)}(\ell+1)^{-1}\right)^{1/(2\bar{\chi}+1)} \right\rceil, \quad \ell = 1,\dots,L, \quad N_0 = \left\lceil 2^{(\tau/\bar{\chi})(L+1)} \right\rceil. \tag{7.7}$$

They are rounded up to the next odd prime number. We compute the generating vectors by the fast CBC algorithm, Gaussian weight functions, and product weights according to Theorem 3.1 and the sequence

$$\bar{b}^{\alpha/2}_{\mathrm{ref}} = (c2^{-(\nu-1)\ell})_{\ell\geq 0,k=1,\dots,N_\ell}$$

for some $c > 0$ (see Rem. 7.1 and Fig. 1a). According to Theorem 3.1, also the value p' is needed for the product weights, which in turn are an input in the CBC algorithm to construct the generating vectors for the QMC lattice points. For $\alpha = 2$, we use the borderline values $p' = 1$ and $\varepsilon = 0$ in Theorem 3.1. We note that $\nu = 3/2$ is the borderline parameter for our MLQMC convergence theory in Theorem 6.3. We expect a convergence rate $\bar{\chi} \approx 1/2$. In our single-level QMC experiments, we observed in Figure 1 from [31] that the QMC rate for a borderline case of applicability (*i.e.* $(b_j)_{j\geq 1} \notin \ell^p(\mathbb{N})$ for every $p \in (0,2)$, but $(b_j)_{j\geq 1} \in \ell^2(\mathbb{N})$) was always larger than 0.65. So for $\alpha = 2$ ($\nu = 3/2$), we use QMC sample numbers (7.7) with $\bar{\chi} = 0.65$. For $\alpha = 4$, $\bar{b}^{\alpha/2}_{\mathrm{ref}} \in \ell^{2/3}(\mathbb{N})$.



(a) $\sigma_0 = 0.5$ (b) $\sigma_0 = 1.0$

FIGURE 2. MLQMC convergence with $\nu = 3/2$, $\alpha_{\mathrm{g}} = 1.05$, $c = 0.1$, $\lambda \in \{0.1, 0.05, 0.01\}$. *Left panel*: $\sigma_0 = 0.5$. *Right panel*: $\sigma_0 = 1.0$.

Thus, we use $p' = 2/3 = p$ and the boundary value $\varepsilon = 0$ in Theorem 3.1. For $\alpha = 4$, the sample numbers are chosen with $\bar{\chi} = 0.9$.

In Figures 2a and 2b, we display error *vs.* work for $\alpha = 2$, *i.e.*, $\nu = 3/2$, $L = 2, \dots, 11$ with reference solution on level $L_{\mathrm{ref}} = 12$. For $L = 11$, there are $s_L = 8192$ stochastic parameters on the finest mesh level L and for the reference solution the highest occurring dimension was $s_{L_{\mathrm{ref}}} = 16384$. Here, and in the following, the asymptotic work model is according to Theorem 6.3

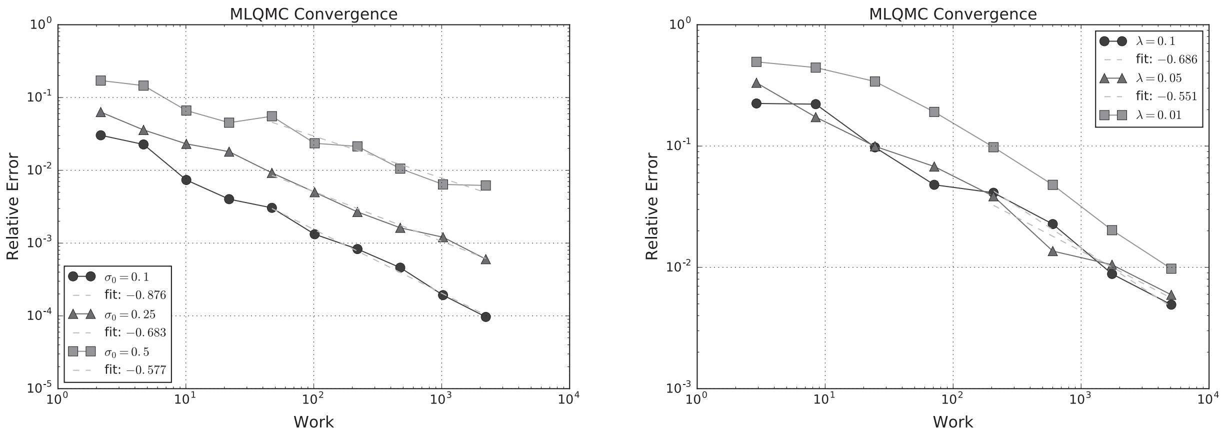
$$\mathrm{work}_L = 2^{(\tau/\bar{\chi})L}.$$

In all numerical tests, the mean square error is approximated by the empirical variance of R samples of Q_j corresponding to R i.i.d. realizations of the random shift, with the unbiased estimator

$$\sqrt{\frac{1}{R-1}\sum_{j=1}^{R}(Q_j - \bar{Q})^2} \approx \sqrt{\mathbb{E}^{\boldsymbol{\Delta}}(|\mathbb{E}(G(u)) - Q_L^*(G(u^L))|^2)}.$$

The reference value $\bar{Q}$ is the average over R i.i.d. random shifts of $Q^*_{L_{\mathrm{ref}}}(G(u^{L_{\mathrm{ref}}}))$ with $L_{\mathrm{ref}} = 12$. In all numerical tests we use $R = 20$. For $\alpha = 2$, this is the borderline case $\bar{p} = 2$ of the error bounds in Theorem 6.3 and we expect a convergence rate of the error as a function of the work of $\approx 1/2$. The empirically observed rate is a least squares fit taking into account the four data pairs corresponding to finer resolution. The MLQMC algorithms converge even for very small correlation length $\lambda > 0$, which is presented in Figures 2a and 2b for two choices of σ_0. Specifically, we observe that for correlation length $\lambda \in \{0.1, 0.05, 0.01\}$, there seems to be a pre-asymptotic regime until the (non-dimensional) correlation length λ can be resolved by the FE discretization in D.

In Figure 3a, we study the error *vs.* the variance of the Matérn field Z. We control this variance by the parameter σ_0, and monitor the convergence rate of the error as a function of the work. The test is carried out for $\alpha = 4$, *i.e.*, $\nu = 7/2$, fixed correlation length $\lambda = 0.1$, and $L = 2, \dots, 11$ with reference solution on level $L_{\mathrm{ref}} = 12$. Thus, $s_{11} = 8192$ dimensions on the highest considered level and $s_{L_{\mathrm{ref}}} = 16384$ dimensions of the reference solution. The empirically observed rate is a least squares fit taking into account the six data pairs corresponding to higher resolution. We observe that the convergence rate seems to be influenced by the variance of Z, the size of the fluctuations. This was also observed in previous numerical experiments in [26,30,31,36].



(a) $\alpha_{\mathrm{g}} = 5.05$, $\nu = 7/2$, $\lambda = 0.1$ (b) $\alpha_{\mathrm{g}} = 1.05$, $\nu = 3/2$, $\sigma_0 = 1.0$

FIGURE 3. MLQMC convergence with $\nu \in \{3/2, 7/2\}$ and $c = 0.1$. *Left panel*: $\nu = 7/2$, $\lambda = 0.1$, and $\sigma_0 \in \{0.1, 0.25, 0.5\}$. *Right panel*: $\nu = 3/2$, $\sigma_0 = 1.0$, and $\lambda \in \{0.1, 0.05, 0.01\}$ with *informed* QMC generating vectors.

All numerical tests in Figures 2a and 2b were performed with generating vectors that do not depend on the correlation length. However, Figure 1a suggests that the pre-asymptotic spatial decay of the functions $\psi^1_{\ell,k}$ depends on the correlation length $\lambda > 0$, which shall be reflected in the decay of the respective QMC integration weight sequence corresponding to the coefficients $(\bar{b}_{\ell,k})_{\ell\geq 0, k=1,\dots,N_\ell}$ to be defined below.

In Figure 3b, we present numerical results with QMC generating vectors that are *informed* by the pre-asymptotic decay of the $\bar{b}_{\ell,k}$. The test is carried out for $\alpha = 2$, *i.e.*, $\nu = 3/2$, $\sigma_0 = 1.0$, and $L = 2, \dots, 9$ with reference solution on level $L_{\mathrm{ref}} = 10$. The empirically observed rate is a least squares fit taking into account the four data pairs corresponding to higher resolution. Specifically, we compute the QMC generating vector with the sequence

$$\bar{b}_{\ell,k} = c b_{\ell,k}^{(\widehat{\beta}-1)/\widehat{\beta}}.$$

Here, the value of c is (as specified in the caption of the figure) $c = 0.1$ (according to (6.5)) for $\widehat{\beta} = 3/2$ and with

$$b_{\ell,k} = \|\psi_{\ell,k}\|_{L^\infty(D)}, \quad \ell \geq 0, k = 1, \dots, N_\ell.$$

In Figure 3b, we observe a very similar behavior compared to Figure 2b. We thus conclude that the observed pre-asymptotic regime is indeed due to the inability of the FE method to resolve the small correlation length until the MLQMC algorithm accesses discretization levels which resolve the spatial correlation length.

8. Conclusions

For linear, second order diffusion equations (1.2) in a polygonal or polyhedral domain D, and with diffusion coefficient $a = \exp(Z)$, where the GRF Z in D is represented in terms of a series expansion in a representation system with supports which are "localized in D" in the sense that (**A1**) and (**A2**) hold and with the GRF Z taking values in weighted Hölder spaces in D, we extended the convergence rate and error *vs.* work analysis of combined QMC quadratures and multilevel FE approximation from [26,36] in several directions. We considered randomly shifted lattice QMC rules introduced in [41] for numerical integration of PDE outputs against a dimension-truncated Gaussian measure. The present work extends previous results to possibly non-stationary GRFs Z, accounts explicitly for a discretization-level dependent truncation of the representation of the GRF, and accounts for possibly low sample regularity of Z and of the random solution u in a polytopal physical domain

D. In particular, Z and u are admitted in weighted Hölder and Sobolev spaces in the polytope $D \subset \mathbb{R}^d$, the weights allowing singularities in realizations of Z and of u due to corners (and edges in dimension $d = 3$) of D. This allows, in particular, GRFs Z whose covariance is non-stationary, with associated *Matérn SPDE* (1.1) in D with Dirichlet or Neumann boundary conditions, as proposed recently in [38]. Whereas in [26,35,36], globally (in D) supported ϕ_j's were admitted (implying QMC quadratures with so-called "POD" weights), the present analysis shows that for multilevel representation systems $(\psi_j)_{j\geq1}$ with localized supports, QMC quadratures with product weights are admissible and, in a sense, natural. We also provided a novel QMC error analysis with Gaussian weighted function spaces for QMC integration on the unbounded integration domain. It extends the range of summability exponents from $p \in (0, 1]$ obtained by exponential weight function as considered in [26, 36] to $p \in (2/3, 2)$ (Thm. 3.1, item 1.), while still retaining dimension independent convergence rate up to $1/(2p)+1/4$. For GRFs Z whose spatial variation is parametrized by a representation system $(\psi_j)_{j\geq1}$ of functions $\psi_j(x)$ defined in D with "localized supports" we proved that QMC combined with continuous, piecewise linear FE in D on families of regular, simplicial triangulations of D with suitable mesh refinement near vertices and (in space dimension $d = 3$) edges of D allows for parameter-dimension independent error *vs.* work bounds. Full elliptic regularity in function spaces in D without spatial weights and uniform truncation dimension, *i.e.* $s_\ell = s_L$, $\ell = 0, \ldots, L$, as considered in Corollary 2 and Section 5 from [36] is a particular case of our results. Here, we admitted bounded, polytopal domains D where ∂D consists of straight lines (in space dimension $d = 2$) or of plane faces (in space dimension $d = 3$) which require weighted spaces for the spatial coordinate, and we consider truncation dimensions s_ℓ of the input GRF that depend on the discretization level. The discretization level dependent truncations of the GRF allow, for elliptic PDEs with log-Gaussian coefficients, in certain cases an ε-complexity of MLQMC-FEM with product weights that is asymptotically equivalent to the ε-complexity of QMC in the case that integrand evaluation would be available at unit cost. The parametric regularity results in weighted function spaces in D hold also for polytopal D with piecewise smoothly curved boundaries as considered in [40].

Since the assumed localization of the supports of the ψ_j in D was shown to allow for QMC integration rules with so-called product weights, the present model of the computational work (6.3) includes the cost of the generation of the QMC points. This cost is dominated by the cost of the fast CBC construction of QMC generating vectors. It was considered an "off-line", pre-computation in [35, 36] and the (quadratic w.r.t. the parameter dimensions s_ℓ) work count for the (precomputed) CBC construction for globally supported ψ_j (as *e.g.* in Karhunen–Loève expansions) was omitted from the work counts in [35, 36]. We also note that the same generating vectors can be used for different right hand sides f. However, if the representation system $(\psi_j)_{j\geq1}$ of the GRF is altered due to modeling considerations of the lognormal diffusion coefficent, then the QMC generating vectors need in general to be recomputed. In the present QMC error analysis, being based on product weights, the work of the fast CBC construction of generating vectors due to Nuyens and Cools [42] and the generation of QMC points scales linearly with respect to the parameter dimensions s_ℓ. We conclude in certain cases the same asymptotic error *vs.* work bounds of the presently proposed MLQMC algorithm as for the numerical solution of (one instance of) the respective deterministic, elliptic PDE. The error *vs.* work estimates for local supports and product weights derived in this paper are in certain cases superior to corresponding error *vs.* work estimates for representations of the GRFs with global supports (such as Karhunen–Loève expansions) and QMC with POD weights as discussed in Section 6.3.

We considered only homogeneous Dirichlet boundary conditions on all of ∂D in (1.2) and, in the numerical experiments section, only even integer order precision operators. This was for ease of exposition only: the parametric regularity analysis of Section 4 and the elliptic regularity results in Section 2.2 remain valid verbatim for problems with Neumann or mixed boundary conditions provided that suitable regularity shifts of the Laplacian with these boundary conditions are available as well as FE spaces with suitable interpolants. In particular, an analogous structure of the corner- and edge-weights in (2.4) can be used to characterize elliptic regularity shifts in scales of weighted Sobolev- and Hölder spaces in D for these boundary conditions. Precision operators of fractional and odd integer order in (1.1), *i.e.* when $\alpha \in \mathbb{R}\backslash(2\mathbb{N})$, can be treated in exactly the same fashion, using recently developed methods for the efficient numerical solution of the SPDE (1.1). We refer to [9] and the references there for details of the corresponding algorithms.

Appendix A. Error estimates of multilevel QMC with level dependent truncation dimensions: global supports and POD weights

In this appendix, we augment the parametric regularity estimates in Section 5 from [36] in order to allow to truncate the dimension of the MLQMC algorithm depending on the FE mesh level ℓ also for globally supported functions ψ_j^{gl} in (1.3) as arises *e.g.* in Karhunen–Loève expansions; MLQMC-FEM was considered in [36], with fixed truncation s_L on all mesh levels $\ell = 0, 1, \dots, L$.

Suppose that the GRF Z is represented as in (1.3) with ψ_j^{gl} and define the sequence $(\upsilon_j)_{j\geq 1}$ by $\upsilon_j := \|\psi_j^{\mathrm{gl}}\|_{L^\infty}$. Assume the summability $(\upsilon_j)_{j\geq 1} \in \ell^{p_{\mathrm{POD}}}(\mathbb{N})$ for some $p_{\mathrm{POD}} \in (0,1)$ and that $(\upsilon_j)_{j\geq 1}$ is decreasing. The corresponding parametric solutions are denoted by u_{gl}^s, $u_{\mathrm{gl}}^{s,\mathcal{T}}$, and $u_{\mathrm{gl}}^{\ell} = u_{\mathrm{gl}}^{s_\ell,\mathcal{T}_\ell}$, respectively.

Proposition A.1. *Let $s \in \mathbb{N}$. For every $\boldsymbol{\tau} \in \mathcal{F}$, such that $\tau_j = 0$ for every $j > s$, and for every $\boldsymbol{y} \in U$,*

$$\|\partial^{\boldsymbol{\tau}}(u(\boldsymbol{y}) - u_{\mathrm{gl}}^s(\boldsymbol{y}))\|_{a(\boldsymbol{y})} \leq \frac{2^{|\boldsymbol{\tau}|+1}|\boldsymbol{\tau}|!}{\log(2)^{|\boldsymbol{\tau}|}}\upsilon^{\boldsymbol{\tau}} \left\|\frac{a(\boldsymbol{y}) - a_{\mathrm{gl}}^s(\boldsymbol{y})}{a(\boldsymbol{y})}\right\|_{L^\infty(D)} \|u_{\mathrm{gl}}^s(\boldsymbol{y})\|_{a(\boldsymbol{y})}.$$

For every $\boldsymbol{\tau} \in \mathcal{F}$ such that there exists $j > s$ with $\tau_j > 0$, and every $\boldsymbol{y} \in U$, there holds

$$\|\partial^{\boldsymbol{\tau}}(u(\boldsymbol{y}) - u_{\mathrm{gl}}^s(\boldsymbol{y}))\|_{a(\boldsymbol{y})} \leq \frac{|\boldsymbol{\tau}|!}{\log(2)^{|\boldsymbol{\tau}|}}\|u(\boldsymbol{y})\|_{a(\boldsymbol{y})}.$$

Proof. The dependence on the parameter vector $\boldsymbol{y}$ is omitted in the proof for notational convenience. Introduce the index sets $\mathcal{F}_1 := \{\boldsymbol{\tau} \in \mathcal{F} : \forall j > s, \tau_j = 0\}$ and $\mathcal{F}_2 := \{\boldsymbol{\tau} \in \mathcal{F} : \exists j > s, \tau_j > 0\}$. For any $\boldsymbol{\tau} \in \mathcal{F}_1$,

$$\begin{aligned}\int_D a\nabla\partial^{\boldsymbol{\tau}}(u - u_{\mathrm{gl}}^s)\cdot\nabla v \mathrm{d}x = &- \sum_{\boldsymbol{\nu}\leq\boldsymbol{\tau},\boldsymbol{\nu}\neq\boldsymbol{\tau}} \binom{\boldsymbol{\tau}}{\boldsymbol{\nu}} \int_D \psi^{\boldsymbol{\tau}-\boldsymbol{\nu}} a\nabla\partial^{\boldsymbol{\nu}}(u - u_{\mathrm{gl}}^s)\cdot\nabla v\mathrm{d}x \\ &- \sum_{\boldsymbol{\nu}\leq\boldsymbol{\tau}} \binom{\boldsymbol{\tau}}{\boldsymbol{\nu}} \int_D \psi^{\boldsymbol{\tau}-\boldsymbol{\nu}}(a - a_{\mathrm{gl}}^s)\nabla\partial^{\boldsymbol{\nu}} u_{\mathrm{gl}}^s\cdot\nabla v\mathrm{d}x.\end{aligned}$$

The choice $v = \partial^{\boldsymbol{\tau}}(u - u_{\mathrm{gl}}^s)$ implies with the Cauchy–Schwarz inequality and equation (3.10) of [26]

$$\begin{aligned}\|\partial^{\boldsymbol{\tau}}(u - u_{\mathrm{gl}}^s)\|_a &\leq \sum_{\boldsymbol{\nu}\leq\boldsymbol{\tau},\boldsymbol{\nu}\neq\boldsymbol{\tau}} \binom{\boldsymbol{\tau}}{\boldsymbol{\nu}} \upsilon^{\boldsymbol{\tau}-\boldsymbol{\nu}}\|\partial^{\boldsymbol{\nu}}(u - u_{\mathrm{gl}}^s)\|_a + \sum_{\boldsymbol{\nu}\leq\boldsymbol{\tau}} \binom{\boldsymbol{\tau}}{\boldsymbol{\nu}} \upsilon^{\boldsymbol{\tau}-\boldsymbol{\nu}} \left\|\frac{a - a_{\mathrm{gl}}^s}{a}\right\|_{L^\infty(D)} \|\partial^{\boldsymbol{\nu}} u_{\mathrm{gl}}^s\|_a \\ &\leq \sum_{\boldsymbol{\nu}\leq\boldsymbol{\tau},\boldsymbol{\nu}\neq\boldsymbol{\tau}} \binom{\boldsymbol{\tau}}{\boldsymbol{\nu}} \upsilon^{\boldsymbol{\tau}-\boldsymbol{\nu}}\|\partial^{\boldsymbol{\nu}}(u - u_{\mathrm{gl}}^s)\|_a + \sum_{\boldsymbol{\nu}\leq\boldsymbol{\tau}} \binom{\boldsymbol{\tau}}{\boldsymbol{\nu}} \upsilon^{\boldsymbol{\tau}} \frac{|\boldsymbol{\nu}|!}{\log(2)^{|\boldsymbol{\nu}|}} \left\|\frac{a - a_{\mathrm{gl}}^s}{a}\right\|_{L^\infty(D)} \|u_{\mathrm{gl}}^s\|_a.\end{aligned}$$

Since the previous estimate holds for every $\boldsymbol{\tau} \in \mathcal{F}_1$, Lemma 5 of [36] is applicable and thus for every $\boldsymbol{\tau} \in \mathcal{F}_1$,

$$\|\partial^{\boldsymbol{\tau}}(u - u_{\mathrm{gl}}^s)\|_a \leq \sum_{\boldsymbol{\nu}\leq\boldsymbol{\tau}} \binom{\boldsymbol{\tau}}{\boldsymbol{\nu}} \frac{|\boldsymbol{\nu}|!}{\log(2)^{|\boldsymbol{\nu}|}} \frac{2^{|\boldsymbol{\tau}-\boldsymbol{\nu}|}|\boldsymbol{\tau}-\boldsymbol{\nu}|!}{\log(2)^{|\boldsymbol{\tau}-\boldsymbol{\nu}|}} \upsilon^{\boldsymbol{\tau}} \left\|\frac{a - a_{\mathrm{gl}}^s}{a}\right\|_{L^\infty(D)} \|u_{\mathrm{gl}}^s\|_a.$$

From the identity $\sum_{\boldsymbol{\nu}\leq\boldsymbol{\tau},|\boldsymbol{\nu}|=\ell} \binom{\boldsymbol{\tau}}{\boldsymbol{\nu}} = \binom{|\boldsymbol{\tau}|}{\ell}$ (see *e.g.* [36], Eq. (5.25)),

$$\sum_{\boldsymbol{\nu}\leq\boldsymbol{\tau}} \binom{\boldsymbol{\tau}}{\boldsymbol{\nu}} |\boldsymbol{\nu}|!|\boldsymbol{\tau}-\boldsymbol{\nu}|!2^{|\boldsymbol{\tau}-\boldsymbol{\nu}|} = \sum_{\ell=0}^{|\boldsymbol{\tau}|} \sum_{\boldsymbol{\nu}\leq\boldsymbol{\tau},|\boldsymbol{\nu}|=\ell} \binom{\boldsymbol{\tau}}{\boldsymbol{\nu}} |\boldsymbol{\nu}|!|\boldsymbol{\tau}-\boldsymbol{\nu}|!2^{|\boldsymbol{\tau}-\boldsymbol{\nu}|} = \sum_{\ell=0}^{|\boldsymbol{\tau}|} |\boldsymbol{\tau}|!2^{|\boldsymbol{\tau}|-\ell} \leq 2^{|\boldsymbol{\tau}|+1}|\boldsymbol{\tau}|!.$$

Thus, for every $\boldsymbol{\tau} \in \mathcal{F}_1$,

$$\|\partial^{\boldsymbol{\tau}}(u - u_{\mathrm{gl}}^s)\|_a \le \frac{2^{|\boldsymbol{\tau}|+1}|\boldsymbol{\tau}|!}{\log(2)^{|\boldsymbol{\tau}|}} \upsilon^{\boldsymbol{\tau}} \left\| \frac{a - a_{\mathrm{gl}}^s}{a} \right\|_{L^\infty(D)} \|u_{\mathrm{gl}}^s\|_a.$$

The second assertion follows by equation (3.10) of [26] since for every $\boldsymbol{\tau} \in \mathcal{F}_2$ holds $\partial^{\boldsymbol{\tau}} u_{\mathrm{gl}}^s = 0$. □

Consider the space $\mathcal{W}_{\boldsymbol{\gamma}}(\mathbb{R}^s)$ defined in (3.1) with respect to exponential weight functions $w_{\exp,j}^2(y) = e^{-\alpha_{\exp}|y|}$, $j \ge 1$. For $L \in \mathbb{N}$, let the sequence $(s_\ell)_{\ell=0,\dots,L} \subset \mathbb{N}$ of truncation dimensions be non-decreasing and the sequence $(N_\ell)_{\ell=0,\dots,L} \subset \mathbb{N}$ of numbers of QMC points at discretization level ℓ be non-increasing.

Theorem A.2. *Suppose that* $(\upsilon_j)_{j\ge1} \in \ell^{p_{\mathrm{POD}}}(\mathbb{N})$ *for some* $0 < p_{\mathrm{POD}} < 1$*. Let* $q \in (p_{\mathrm{POD}}, 1]$*. Let* $\alpha_{\exp} > 2\sup_{j\ge1}\{\upsilon_j\}$*. For any* $n \in \mathbb{N}_0$*,* $\mathfrak{u} \subset \mathbb{N}$ *with* $|\mathfrak{u}| < \infty$ *and* $c_1, c_2 > 0$*, define*

$$\gamma_{\mathfrak{u}}^{\mathrm{POD}} := \left(c_1(|\mathfrak{u}| + n)! \prod_{j\in\mathfrak{u}} \frac{\upsilon_j^{p_{\mathrm{POD}}/q}}{c_2} \right)^{2-q}. \tag{A.1}$$

There exists a constant $C > 0$ *which does not depend on* $\mathcal{T}_\ell$*,* s_ℓ*,* $\ell \ge 1$*, such that for every* $\ell \ge 0$

$$\|G(u_{\mathrm{gl}}^{s_\ell,\mathcal{T}_\ell}) - G(u_{\mathrm{gl}}^{s_{\ell-1},\mathcal{T}_\ell})\|_{\mathcal{W}_{\boldsymbol{\gamma}^{\mathrm{POD}}}(\mathbb{R}^{s_\ell})} \le C\|G\|_{V^*}\|f\|_{V^*} s_{\ell-1}^{-1/p_{\mathrm{POD}}+1/q}.$$

Proof. We observe that $\|u_{\mathrm{gl}}^{s_{\ell-1}}\|_V \le \|f\|_{V^*}\exp(\sum_{j=1}^{s_{\ell-1}} |y_j|\upsilon_j)$ and that

$$\begin{aligned}\left\| \frac{a_{\mathrm{gl}}^{s_\ell} - a_{\mathrm{gl}}^{s_{\ell-1}}}{a_{\mathrm{gl}}^{s_\ell}} \right\|_{L^\infty(D)} &\le \left\| 1 + \exp\left(\sum_{j=s_{\ell-1}}^{s_\ell} -y_j\psi_j \right) \right\|_{L^\infty(D)} \left\| \sum_{j=s_{\ell-1}}^{s_\ell} y_j\psi_j \right\|_{L^\infty(D)} \\ &\le 2\exp\left(\sum_{j=s_{\ell-1}}^{s_\ell} |y_j|\upsilon_j \right) \sum_{j=s_{\ell-1}}^{s_\ell} |y_j|\upsilon_j\end{aligned}$$

depends only on y_j, $j \ge s_{\ell-1}$. By Proposition A.1 and the Cauchy–Schwarz inequality for any $\mathfrak{u} \subset \{1 : s_{\ell-1}\}$,

$$\begin{aligned}&\int_{\mathbb{R}^{s_\ell-|\mathfrak{u}|}} |\partial^{\mathfrak{u}} G(u_{\mathrm{gl}}^{s_\ell,\mathcal{T}_\ell}(\boldsymbol{y}) - u_{\mathrm{gl}}^{s_{\ell-1},\mathcal{T}_\ell}(\boldsymbol{y}))| \prod_{j\in\{1:s_\ell\}\setminus\mathfrak{u}} \phi(y_j)\mathrm{d}y_j \\ &\quad \le 2\|G\|_{V^*}\|f\|_{V^*} \frac{2^{|\mathfrak{u}|+1}|\mathfrak{u}|!}{\log(2)^{|\mathfrak{u}|}} \upsilon^{\mathfrak{u}} \exp\left(\sum_{j\in\mathfrak{u}} |y_j|\upsilon_j \right) \\ &\qquad \times \int_{\mathbb{R}^{s_\ell-|\mathfrak{u}|}} \exp\left(2 \sum_{j\in\{1:s_\ell\}\setminus\mathfrak{u}} |y_j|\upsilon_j \right) \sum_{j=s_{\ell-1}}^{s_\ell} |y_j|\upsilon_j \prod_{j\in\{1:s_\ell\}\setminus\mathfrak{u}} \phi(y_j)\mathrm{d}y_j \\ &\quad \le 2\|G\|_{V^*}\|f\|_{V^*} \frac{2^{|\mathfrak{u}|+1}|\mathfrak{u}|!}{\log(2)^{|\mathfrak{u}|}} \upsilon^{\mathfrak{u}} \exp\left(\sum_{j\in\mathfrak{u}} |y_j|\upsilon_j + \sum_{j\in\{1:s_\ell\}\setminus\mathfrak{u}} \left(4\upsilon_j^2 + \frac{4\upsilon_j}{\sqrt{2\pi}} \right) \right) \sum_{j=s_{\ell-1}}^{s_\ell} \upsilon_j,\end{aligned}$$

where we used that for any $c > 0$, $\int_{\mathbb{R}} e^{c|y|}\phi(y)\mathrm{d}y \leq \exp(c^2/2 + 2c/\sqrt{2\pi})$, *cf.* Equation (4.15) of [26], and $\int_{-\infty}^{c} \phi(y)\mathrm{d}y \leq \exp(2c/\sqrt{2\pi})/2$ (see *e.g.* [26], p. 355). Thus,

$$\begin{aligned}\int_{\mathbb{R}^{|\mathfrak{u}|}} &\left(\int_{\mathbb{R}^{s_\ell - |\mathfrak{u}|}} \partial^{\mathfrak{u}} G(u_{\mathrm{gl}}^{s_\ell,\mathcal{T}_\ell}(\boldsymbol{y}) - u_{\mathrm{gl}}^{s_{\ell-1},\mathcal{T}_\ell}(\boldsymbol{y})) \prod_{j\in\{1:s_\ell\}\setminus\mathfrak{u}} \phi(y_j)\mathrm{d}y_j \right)^2 \prod_{j\in\mathfrak{u}} w_{\exp,j}^2(y_j)\mathrm{d}y_j \\ &\leq 4\|G\|_{V^*}^2\|f\|_{V^*}^2 \left(\frac{2^{|\mathfrak{u}|+1}|\mathfrak{u}|!}{\log(2)^{|\mathfrak{u}|}}\right)^2 \prod_{j\in\mathfrak{u}} \frac{8\upsilon_j^2}{\alpha_{\exp} - 2\upsilon_j} \exp\left(16 \sum_{j\in\{1:s_\ell\}\setminus\mathfrak{u}} \left(\upsilon_j^2 + \frac{\upsilon_j}{\sqrt{2\pi}}\right)\right) \left(\sum_{j=s_{\ell-1}}^{s_\ell} \upsilon_j\right)^2.\end{aligned} \tag{A.2}$$

Similarly, we obtain using Proposition A.1 that for $\mathfrak{u} \subset \{1 : s_\ell\}$ such that $\mathfrak{u} \cap \{s_{\ell-1}+1 : s_\ell\} \neq \emptyset$,

$$\begin{aligned}\int_{\mathbb{R}^{|\mathfrak{u}|}} &\left(\int_{\mathbb{R}^{s_\ell - |\mathfrak{u}|}} \partial^{\mathfrak{u}} G(u_{\mathrm{gl}}^{s_\ell,\mathcal{T}_\ell}(\boldsymbol{y}) - u_{\mathrm{gl}}^{s_{\ell-1},\mathcal{T}_\ell}(\boldsymbol{y})) \prod_{j\in\{1:s_\ell\}\setminus\mathfrak{u}} \phi(y_j)\mathrm{d}y_j \right)^2 \prod_{j\in\mathfrak{u}} w_{\exp,j}^2(y_j)\mathrm{d}y_j \\ &\leq \|G\|_{V^*}^2\|f\|_{V^*}^2 \left(\frac{|\mathfrak{u}|!}{\log(2)^{|\mathfrak{u}|}}\right)^2 \prod_{j\in\mathfrak{u}} \frac{2\upsilon_j^2}{\alpha_{\exp} - 2\upsilon_j} \exp\left(\sum_{j\in\{1:s_\ell\}\setminus\mathfrak{u}} \upsilon_j^2 + \frac{4\upsilon_j}{\sqrt{2\pi}}\right).\end{aligned} \tag{A.3}$$

By Theorem 11 of [35], there exists a constant $C > 0$ such that

$$\sum_{\mathfrak{u}\subset\{1:s_\ell\},\mathfrak{u}\cap\{s_{\ell-1}+1:s_\ell\}\neq\emptyset} \frac{(|\mathfrak{u}|!)^2 \prod_{j\in\mathfrak{u}} \upsilon_j^2}{\gamma_{\mathfrak{u}}^{\mathrm{POD}}} \leq C s_{\ell-1}^{-2/p_{\mathrm{POD}}+2/q} \sum_{\mathfrak{u}\subset\{1:s_\ell\}} \frac{((|\mathfrak{u}|+n)!)^2}{\gamma_{\mathfrak{u}}^{\mathrm{POD}}}.$$

Since $\sum_{j>s_{\ell-1}} \upsilon_j \leq \min\{p_{\mathrm{POD}}/(1-p_{\mathrm{POD}}), 1\} \|(\upsilon_j)_{j\geq1}\|_{\ell^{p_{\mathrm{POD}}}(\mathbb{N})} s_{\ell-1}^{-1/p_{\mathrm{POD}}+1}$ (see *e.g.* [35], Eq. (14)), (A.2) and (A.3) imply

$$\|G(u_{\mathrm{gl}}^{s_\ell,\mathcal{T}_\ell}) - G(u_{\mathrm{gl}}^{s_{\ell-1},\mathcal{T}_\ell})\|_{\mathcal{W}_{\boldsymbol{\gamma}^{\mathrm{POD}}}(\mathbb{R}^{s_\ell})}^2 \leq \mathcal{C} s_{\ell-1}^{-2/p_{\mathrm{POD}}+2/q} \sum_{\mathfrak{u}\subset\{1:s_\ell\}} \prod_{j\in\mathfrak{u}} \frac{2\upsilon_j^2}{\alpha_{\exp} - 2\upsilon_j} \frac{((|\mathfrak{u}|+n)!)^2}{\gamma_{\mathfrak{u}}^{\mathrm{POD}} \log(2)^{2\mathfrak{u}}},$$

where

$$\mathcal{C} = C\|G\|_{V^*}^2\|f\|_{V^*}^2 \left(1 + \|(\upsilon_j)_{j\geq1}\|_{\ell^{p_{\mathrm{POD}}}(\mathbb{N})}\right) \exp\left(16 \sum_{j\geq1} \left(\upsilon_j^2 + \frac{\upsilon_j}{\sqrt{2\pi}}\right)\right) < \infty$$

and C is independent of f, and the parameter dimension. Boundedness of the quantity

$$\sum_{\mathfrak{u}\subset\{1:s_\ell\}} \prod_{j\in\mathfrak{u}} \frac{2\upsilon_j^2}{\alpha_{\exp} - 2\upsilon_j} \frac{((|\mathfrak{u}|+n)!)^2}{\gamma_{\mathfrak{u}}^{\mathrm{POD}} \log(2)^{2\mathfrak{u}}}$$

independently of the parameter dimension s_ℓ may be checked following the same arguments as in the proofs of Theorem 20 and Corollary 21 from [26]. □

To state the error estimate of the MLQMC algorithm with POD weights and global supports, we assume that $f, G(\cdot) \in L^2(D)$, and that the Dirichlet Laplacian is boundedly invertible from $H^2(D) \cap V$ to $L^2(D)$. In this case the FE spaces V_ℓ, $\ell \geq 0$, result by uniformly refining an initial triangulation. They have mesh width h_ℓ and dimension $\mathcal{O}(h_\ell^{-d})$, $\ell \geq 0$. Define

$$\bar{\upsilon}_j := \max\{\upsilon_j, \||\nabla\psi_j|\|_{L^\infty(D)}\}, \quad j \geq 1.$$

The following corollary generalizes Theorem 9 of [36] to the case that also the truncation dimensions s_ℓ may differ from s_L.

Corollary A.3. *Suppose that* $(\upsilon_j)_{j\geq 1} \in \ell^{p_{\rm POD}}(\mathbb{N})$ *and* $(\bar{\upsilon}_j)_{j\geq 1} \in \ell^{\bar{p}_{\rm POD}}(\mathbb{N})$ *for* $p_{\rm POD} \in (0,1)$ *and* $\bar{p}_{\rm POD} \in (\max\{2/3, p_{\rm POD}\}, 1)$*. Let* $q \in (p_{\rm POD}, 1)$*. Consider the POD weights in* (A.1) *with* $\upsilon_j^{p_{\rm POD}/q}$ *replaced by* $\beta_j := \max\{\bar{\upsilon}_j, \upsilon_j^{p_{\rm POD}/q}\}$*,* $j \geq 1$*, and* $n = 5$*. Let* $\alpha_{\exp} > 2\sup_{j\geq 1}\{2\upsilon_j, 9\bar{\upsilon}_j\}$*.*

Then, QMC randomly shifted lattice rules with N_ℓ *points in dimension* s_ℓ *may be constructed in* $\mathcal{O}(s_\ell^2 N_\ell + s_\ell N_\ell \log(N_\ell))$*,* $\ell = 0, \dots, L$*, with the fast CBC algorithm from Section 5.2 of [41], so that the MLQMC algorithm* Q_L^* *in* (3.3) *satisfies, for* $\bar{\chi} = 1/\max\{q, \bar{p}_{\rm POD}\} - 1/2$*,*

$$\sqrt{\mathbb{E}^{\boldsymbol{\Delta}}(|I_{s_L}(G(u_{\rm gl}^L)) - Q_L^*(G(u_{\rm gl}^L))|^2)} \leq C\left(\sum_{\ell=0}^{L}\varphi(N_\ell)^{-2\bar{\chi}}(\xi_{\ell,\ell-1}s_{\ell-1}^{-2/p_{\rm POD}+2/q} + h_{\ell-1}^4)\right)^{1/2},$$

where $\xi_{\ell,\ell-1} := 0$ *if* $s_\ell = s_{\ell-1}$ *and* $\xi_{\ell,\ell-1} := 1$ *otherwise.*

Proof. The assertion follows by Theorem A.2, (3.4), and Theorem 9 and Corollary 8 of [36]. ☐

Remark A.4. Standard error bounds for FE discretizations imply that Theorem 9 of [36] may be extended to $f \in H^{-1+t}(D)$, $G(\cdot) \in H^{-1+t'}(D)$, $t, t' \in [0,1]$. Applying interpolation in the appropriate places in the proof of Theorem 9 from [36]. Corollary A.3 remains valid with $h_{\ell-1}^{2(t+t')}$ in place of $h_{\ell-1}^4$.

Appendix B. Error *vs.* work analysis: global supports and POD weights

For the discussion of the required work to obtain a target accuracy in the case of global supports and POD weights, we suppose that there are fast methods available for the evaluation of $\sum_{j=1}^{s_\ell} y_j\psi_j^{\rm gl}(x_k)$ for nodes x_k, which is necessary to assemble the stiffness matrix for a given QMC point $\boldsymbol{y}$. Assume that the computational work for mesh width $h_\ell \in (0,1]$ is $\mathcal{O}(h_\ell^{-d}\log(h_\ell^{-1}) + s_\ell\log(s_\ell))$, where $\#\mathcal{T} = \mathcal{O}(h_\ell^{-d})$. Then, for $L \in \mathbb{N}$, the computational work of $Q_L^*(G(u_{\rm gl}^L))$ is

$$\text{work}_L = \mathcal{O}\left(\sum_{\ell=0}^{L} N_\ell(s_\ell^2 + s_\ell\log(N_\ell) + h_\ell^{-d(1+\eta)} + s_\ell\log(s_\ell))\right) = \mathcal{O}\left(\sum_{\ell=0}^{L} N_\ell(s_\ell^2 + h_\ell^{-d(1+\eta)})\right),$$

where we assume that s_ℓ asymptotically dominates $\log(N_\ell)$ and that the PDE may be solved in $\mathcal{O}(h_\ell^{-d(1+\eta)})$ for some $\eta > 0$ (see the assumption in (**A3**)). In contrast to the discussion in Section 3 from [36], we also take into account the computational cost of the CBC construction, which is required to realize the QMC points. As in Theorem 1 from [36], we suppose that for $\alpha' > 0$, there exists a constant $C > 0$ such that for every $L \in \mathbb{N}$

$$|\mathbb{E}(G(u)) - \mathbb{E}(G(u_{\rm gl}^L))| \leq C(h_L^\tau + s_L^{-\alpha'}).$$

As in Remark A.4, we assume $f \in H^{-1+t}(D)$ and $G(\cdot) \in H^{-1+t'}(D)$, $t, t' \in [0,1]$, and set $\tau = t + t'$. Set $s_L \sim h_L^{-\tau/\alpha'}$ to equilibrate the error contributions. Furthermore, we choose

$$s_\ell \sim \min\{ch_l^{-\tau p_{\rm POD}\bar{p}_{\rm POD}/(\bar{p}_{\rm POD}-p_{\rm POD})}, s_L\}, \quad \ell = 0, \dots, L-1.$$

As a result by Corollary A.3 (with $q = \bar{p}_{\rm POD}$) and Remark A.4 the following error estimate holds

$$\text{error}_L^2 = \mathcal{O}\left(h_L^\tau + \sum_{\ell=0}^{L} N_\ell^{-2\bar{\chi}} h_{\ell-1}^\tau\right),$$

where we used that the Euler totient function satisfies that $\varphi(N)^{-1} \leq 9/N$ for every $N \leq 10^{30}$. Corresponding work estimates may be obtained along the line of the error *vs.* work analysis from Section 6.1 in the case of Strategy 1. The QMC sample numbers are given by

$$N_\ell = \left\lceil N_0\left(h_\ell^{-2\tau}\max\{h_\ell^{-d(1+\eta)}, \min\{h_\ell^{-2\tau\alpha}, h_L^{-2\tau/\alpha'}\}\}\right)^{-1/(1+2\bar{\chi})}\right\rceil, \quad \ell = 1, \dots, L, \tag{B.1}$$

and

$$N_0 = \begin{cases} \lceil 2^{L\tau/\bar{\chi}} \rceil & \text{if } \max\{2\tau\alpha/d, 1+\eta\} < \tau/(d\bar{\chi}), \\ \lceil 2^{L\tau/\bar{\chi}} L^{1/(2\bar{\chi})} \rceil & \text{if } \max\{2\tau\alpha/d, 1+\eta\} = \tau/(d\bar{\chi}), \\ \lceil 2^{(2\tau+d(1+\eta))L/(1+2\bar{\chi})} \rceil & \text{if } 1+\eta \geq 2\tau\alpha/d, 1+\eta > \tau/(d\bar{\chi}), \\ \lceil 2^{L\min\{2\tau\alpha-\tau/\bar{\chi}, \max\{(d(1+\eta)-\tau/\bar{\chi}), 2\tau/\alpha'\}\}/(1+2\bar{\chi})+L\tau/\bar{\chi}} \rceil & \text{if } 1+\eta < 2\tau\alpha/d, 2\alpha > 1/\bar{\chi}, \end{cases} \tag{B.2}$$

where we have set $\alpha = p_{\mathrm{POD}}\bar{p}_{\mathrm{POD}}/(\bar{p}_{\mathrm{POD}} - p_{\mathrm{POD}})$.

Theorem B.1. *Let the QMC sample numbers for* $Q_L^*(\cdot)$ *be given by* (B.2) *and* (B.1). *Suppose that* $(\upsilon_j)_{j\geq 1} \in \ell^{p_{\mathrm{POD}}}(\mathbb{N})$ *and* $(\bar{\upsilon}_j)_{j\geq 1} \in \ell^{\bar{p}_{\mathrm{POD}}}(\mathbb{N})$ *for* $p_{\mathrm{POD}} \in (0,1)$ *and* $\bar{p}_{\mathrm{POD}} \in (\max\{2/3, p_{\mathrm{POD}}\}, 1)$. *Set* $\alpha = p_{\mathrm{POD}}\bar{p}_{\mathrm{POD}}/(\bar{p}_{\mathrm{POD}} - p_{\mathrm{POD}})$.

Then the error threshold $\varepsilon > 0$, *i.e.,*

$$\sqrt{\mathbb{E}^{\mathbf{\Delta}}(|\mathbb{E}(G(u)) - Q_L^*(G(u_{\mathrm{gl}}^L))|^2)} = \mathcal{O}(\varepsilon),$$

may be achieved with

$$\mathrm{work}_L = \begin{cases} \mathcal{O}(\varepsilon^{-1/\bar{\chi}}) & \text{if } \max\{2\tau\alpha/d, 1+\eta\} < \tau/(d\bar{\chi}), \\ \mathcal{O}(\varepsilon^{-1/\bar{\chi}} \log(\varepsilon^{-1})^{(1+2\bar{\chi})/(2\bar{\chi})}) & \text{if } \max\{2\tau\alpha/d, 1+\eta\} = \tau/(d\bar{\chi}), \\ \mathcal{O}(\varepsilon^{-d/\tau(1+\eta)}) & \text{if } 1+\eta \geq 2\tau\alpha/d, 1+\eta > \tau/(d\bar{\chi}), \\ \mathcal{O}(\varepsilon^{-d\min\{2\alpha/d, \max\{(1+\eta)/\tau, 2/(d\alpha')+1/(d\bar{\chi}\}\}}) & \text{if } 1+\eta < 2\tau\alpha/d, 2\alpha > 1/\bar{\chi}, \end{cases}$$

where $\bar{\chi} = 1/\bar{p}_{\mathrm{POD}} - 1/2$.

Acknowledgements. The authors acknowledge the help of Magdalena Keller, MSc ETH Applied Mathematics, for permission to use her C++ implementation of the fast CBC algorithm.

References

[1] R.J. Adler, The geometry of random fields. In: *Wiley Series in Probability and Mathematical Statistics*. John Wiley & Sons, Ltd., Chichester (1981).

[2] J.H. Adler and V. Nistor, Graded mesh approximation in weighted Sobolev spaces and elliptic equations in 2D. *Math. Comput.* **84** (2015) 2191–2220.

[3] T. Apel, A.L. Lombardi and M. Winkler, Anisotropic mesh refinement in polyhedral domains: error estimates with data in $L^2(\Omega)$. *ESAIM: M2AN* **48** (2014) 1117–1145.

[4] T. Apel, J. Pfefferer and M. Winkler, Local mesh refinement for the discretization of Neumann boundary control problems on polyhedra. *Math. Methods Appl. Sci.* **39** (2016) 1206–1232.

[5] I. Babuška, R. Bruce Kellogg and J. Pitkäranta, Direct and inverse error estimates for finite elements with mesh refinements. *Numer. Math.* **33** (1979) 447–471.

[6] M. Bachmayr, A. Cohen, D. Dũng and C. Schwab, Fully discrete approximation of parametric and stochastic elliptic PDEs. *SIAM J. Numer. Anal.* **55** (2017) 2151–2186.

[7] M. Bachmayr, A. Cohen, R. DeVore and G. Migliorati. Sparse polynomial approximation of parametric elliptic PDEs. part II: lognormal coefficients. *ESAIM: M2AN* **51** (2017) 341–363.

[8] M. Bachmayr, A. Cohen and G. Migliorati, Representations of Gaussian random fields and approximation of elliptic PDEs with lognormal coefficients. *J. Fourier Anal. Appl.* **24** (2018) 621–649.

[9] L. Banjai, J.M. Melenk, R.H. Nochetto, E. Otarola, A.J. Salgado and C. Schwab, Tensor FEM for spectral fractional diffusion. *Found. Comput. Math.* DOI:10.1007/s10208-018-9402-3 (2018).

[10] V.I. Bogachev, Gaussian Measures. Vol. 62 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI (1998).

[11] J.H. Bramble, Multigrid methods. In: Vol. 294 of *Pitman Research Notes in Mathematics Series*. Longman Scientific & Technical, Harlow; copublished in the United States with John Wiley & Sons, Inc., New York, NY (1993).

[12] J. Charrier, Strong and weak error estimates for elliptic partial differential equations with random coefficients. *SIAM J. Numer. Anal.* **50** (2012) 216–246.

[13] P.G. Ciarlet, Linear and Nonlinear Functional Analysis With Applications. Society for Industrial and Applied Mathematics, Philadelphia, PA (2013).

[14] J. Dick, F.Y. Kuo and I.H. Sloan, High-dimensional integration: the quasi-Monte Carlo way. *Acta Numer.* **22** (2013) 133–288.
[15] J. Dick, F.Y. Kuo, Q.T. Le Gia, D. Nuyens and C. Schwab, Higher order QMC Petrov-Galerkin discretization for affine parametric operator equations with random field inputs. *SIAM J. Numer. Anal.* **52** (2014) 2676–2702.
[16] C.R. Dietrich and G.N. Newsam, Fast and exact simulation of stationary Gaussian processes through circulant embedding of the covariance matrix. *SIAM J. Sci. Comput.* **18** (1997) 1088–1107.
[17] M.M. Dunlop, M.A. Girolami, A.M. Stuart and A.L. Teckentrup. How deep are deep Gaussian processes? *J. Mach. Learn. Res.* **19** (2018) 54.
[18] G.-A. Fuglstad, F. Lindgren, D. Simpson and H. Rue, Exploring a new class of non-stationary spatial Gaussian random fields with varying local anisotropy. *Statist. Sinica* **25** (2015) 115–133.
[19] R.N. Gantner, L. Herrmann and C. Schwab, Multilevel QMC with product weights for affine-parametric, elliptic PDEs, edited by J. Dick, F.Y. Kuo and H. Woźniakowski. In: Contemporary Computational Mathematics – A Celebration of the 80th Birthday of Ian Sloan. Springer International Publishing, Cham (2018) 373–405.
[20] R.N. Gantner, L. Herrmann and C. Schwab, Quasi-Monte Carlo integration for affine-parametric, elliptic PDEs: local supports and product weights. *SIAM J. Numer. Anal.* **56** (2018) 111–135.
[21] F.D. Gaspoz and P. Morin, Convergence rates for adaptive finite elements. *IMA J. Numer. Anal.* **29** (2009) 917–936.
[22] A. George, Nested dissection of a regular finite element mesh. *SIAM J. Numer. Anal.* **10** (1973) 345–363.
[23] C.J. Gittelson, Representation of Gaussian fields in series with independent coefficients. *IMA J. Numer. Anal.* **32** (2012) 294–319.
[24] G.H. Golub and C.F. Van Loan, Matrix computations, 3rd edition. In: Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, Baltimore, MD (1996).
[25] I.G. Graham, F.Y. Kuo, D. Nuyens, R. Scheichl and I.H. Sloan, Quasi-Monte Carlo methods for elliptic PDEs with random coefficients and applications. *J. Comput. Phys.* **230** (2011) 3668–3694.
[26] I.G. Graham, F.Y. Kuo, J.A. Nichols, R. Scheichl, C. Schwab and I.H. Sloan, Quasi-Monte Carlo finite element methods for elliptic PDEs with lognormal random coefficients. *Numer. Math.* **131** (2015) 329–368.
[27] I.G. Graham, F.Y. Kuo, D. Nuyens, R. Scheichl and I.H. Sloan, Analysis of circulant embedding methods for sampling stationary random fields. *SIAM J. Numer. Anal.* **56** (2018) 1871–1895.
[28] I.G. Graham, F.Y. Kuo, D. Nuyens, R. Scheichl and I.H. Sloan, Circulant embedding with QMC: analysis for elliptic PDE with lognormal coefficients. *Numer. Math.* **140** (2018) 479–511.
[29] L. Herrmann, Strong convergence analysis of iterative solvers for random operator equations. Technical Report 2017-35, Seminar for Applied Mathematics, ETH Zürich (2017).
[30] L. Herrmann and C. Schwab, QMC algorithms with product weights for lognormal-parametric, elliptic PDEs. In: *Monte Carlo and Quasi-Monte Carlo Methods: MCQMC 2016, Stanford, CA, August 14–19.* Vol. 241 of *Springer Proc. Math. Stat.* Springer, Cham (2018) 313–330.
[31] L. Herrmann and C. Schwab, QMC integration for lognormal-parametric, elliptic PDEs: local supports and product weights. *Numer. Math.* **141** (2019) 63–102.
[32] L. Herrmann, A. Lang and C. Schwab, Numerical analysis of lognormal diffusions on the sphere. *Stochastics Partial Differ. Equ.: Anal. Comput.* **6** (2018) 1–44.
[33] Y. Kazashi, Quasi-Monte Carlo integration with product weights for elliptic PDEs with log-normal coefficients. *IMA J. Numer. Anal.* **39** (2019) 1563–1593.
[34] F.Y. Kuo, I.H. Sloan, G.W. Wasilkowski and B.J. Waterhouse, Randomly shifted lattice rules with the optimal rate of convergence for unbounded integrands. *J. Complexity* **26** (2010) 135–160.
[35] F.Y. Kuo, C. Schwab and I.H. Sloan, Multi-level quasi-Monte Carlo finite element methods for a class of elliptic PDEs with random coefficients. *Found. Comput. Math.* **15** (2015) 411–449.
[36] F.Y. Kuo, R. Scheichl, C. Schwab, I.H. Sloan and E. Ullmann, Multilevel quasi-Monte Carlo methods for lognormal diffusion problems. *Math. Comput.* **86** (2017) 2827–2860.
[37] H. Li, An anisotropic finite element method on polyhedral domains: interpolation error analysis. *Math. Comput.* **87** (2018) 1567–1600.
[38] F. Lindgren, H. Rue and J. Lindström, An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **73** (2011) 423–498. With discussion and a reply by the authors.
[39] B. Matérn, Spatial variation, 2nd edition. With a Swedish summary. Vol. 36 of *Lecture Notes in Statistics*. Springer-Verlag, Berlin (1986).
[40] V. Maz'ya and J. Rossmann, Elliptic equations in polyhedral domains. Vol. 162 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI (2010).
[41] J.A. Nichols and F.Y. Kuo, Fast CBC construction of randomly shifted lattice rules achieving $\mathcal{O}(n^{-1+\delta})$ convergence for unbounded integrands over $\mathbb{R}^s$ in weighted spaces with POD weights. *J. Complexity* **30** (2014) 444–468.
[42] D. Nuyens and R. Cools, Fast algorithms for component-by-component construction of rank-1 lattice rules in shift-invariant reproducing kernel Hilbert spaces. *Math. Comput.* **75** (2006) 903–920.
[43] D. Nuyens and R. Cools, Fast component-by-component construction of rank-1 lattice rules with a non-prime number of points. *J. Complexity* **22** (2006) 4–28.
[44] N. Rekatsinas and R. Stevenson, A quadratic finite element wavelet Riesz basis. *Int. J. Wavelets Multiresolut. Inf. Process.* **16** (2018) 1850033.

[45] H. Triebel, Interpolation Theory, Function Spaces, Differential Operators, 2nd edition. Johann Ambrosius Barth, Heidelberg (1995).

[46] T. von Petersdorff and C. Schwab, Wavelet approximations for first kind boundary integral equations on polygons. *Numer. Math.* **74** (1996) 479–516.

[47] P. Whittle, On stationary processes in the plane. *Biometrika* **41** (1954) 434–449.

[48] P. Whittle, Stochastic processes in several dimensions. *Bull. Inst. Internat. Statist.* **40** (1963) 974–994.

[49] J. Xu, Iterative methods by space decomposition and subspace correction. *SIAM Rev.* **34** (1992) 581–613.