QUASI-MONTE CARLO AND DISCONTINUOUS GALERKIN*

VESA KAARNIOJA† AND ANDREAS RUPP‡

Abstract. In this study, we consider the development of tailored quasi-Monte Carlo (QMC) cubatures for non-conforming discontinuous Galerkin (DG) approximations of elliptic partial differential equations (PDEs) with random coefficients. We consider both the affine and uniform and the lognormal models for the input random field and investigate the use of QMC cubatures to approximate the expected value of the PDE response subject to input uncertainty. In particular, we prove that the resulting QMC convergence rate for DG approximations behaves in the same way as if continuous finite elements were chosen. Notably, the parametric regularity bounds for DG, which are developed in this work, are also useful for other methods such as sparse grids. Numerical results underline our analytical findings.

Key words. diffusion equation, discontinuous Galerkin, quasi-Monte Carlo, random coefficient

AMS subject classifications. 65C05, 65N30

1. Introduction. In this paper, we consider the design of tailored quasi-Monte Carlo (QMC) rules for a class of non-conforming discontinuous Galerkin (DG) methods used to discretize an elliptic partial differential equation (PDE) with a random diffusion coefficient. More precisely, we investigate the class of interior penalty DG (IPDG) methods, and we show that there exist constructible rank-1 lattice rules satisfying rigorous error bounds independently of the stochastic dimension of the problem.

Let $D \subset \mathbb{R}^d$, $d \in \{1, 2, 3\}$, be a physical domain with Lipschitz boundary, and let $(\Omega, \Gamma, \mathbb{P})$ be a probability space. Let us consider the problem of finding $u \colon D \times \Omega \to \mathbb{R}$ which satisfies

$$(1.1) \qquad \begin{cases} -\nabla \cdot (a(\boldsymbol{x}, \omega) \nabla u(\boldsymbol{x}, \omega)) = f(\boldsymbol{x}), & \boldsymbol{x} \in D, \\ u(\boldsymbol{x}, \omega) = 0, & \boldsymbol{x} \in \partial D, \end{cases}$$

for almost all $\omega \in \Omega$. Many studies in uncertainty quantification typically consider one of the following two models for the input random field:

1. the affine and uniform model (cf., e.g., [9, 13, 15, 19, 39, 40, 49, 50])

$$a(\boldsymbol{x}, \omega) = a_0(\boldsymbol{x}) + \sum_{j=1}^{\infty} y_j(\omega) \psi_j(\boldsymbol{x}), \qquad \boldsymbol{x} \in D, \ \omega \in \Omega,$$

   where $y_1, y_2, \ldots$ are independently and identically distributed random variables with uniform distribution on $[-\frac{1}{2}, \frac{1}{2}]$;

2. the lognormal model (cf., e.g., [21, 22, 23, 24, 30, 38, 51])

$$a(\boldsymbol{x}, \omega) = a_0(\boldsymbol{x}) \exp\left(\sum_{j=1}^{\infty} y_j(\omega) \psi_j(\boldsymbol{x}) \right), \qquad \boldsymbol{x} \in D, \ \omega \in \Omega,$$

   where $y_1, y_2, \ldots$ are independently and identically distributed standard normal random variables and $a_0(\boldsymbol{x}) > 0$.

†Department of Mathematics and Computer Science, Free University of Berlin, Arnimallee 6, DE-14195 Berlin (vesa.kaarnioja@fu-berlin.de).

‡Department of Mathematics, Faculty of Mathematics and Computer Science, Saarland University, DE-66123 Saarbrücken (andreas.rupp@uni-saarland.de).

Here, $(\psi_j)_{j \geq 1}$ are assumed to be real-valued L^∞-functions on D such that

$$\sum_{j=1}^{\infty} \|\psi_j\|_{L^\infty} < \infty.$$

In the study of uncertainty quantification for PDEs, a natural quantity to investigate is the expected value

$$\mathbb{E}[u] := \int_\Omega u(\cdot, \omega) \, \mathrm{d}\mathbb{P}(\omega).$$

In practice, the problem needs to be discretized in several ways before it is possible to approximate this quantity numerically. The infinite-dimensional input random field is first replaced by a finite-dimensional one, meaning that we end up analyzing a dimensionally truncated PDE solution u_s. The dimensionally truncated solution of the PDE is replaced by a finite element solution $u_{s,h}$, and the high-dimensional integral needs to be approximated by an s-dimensional cubature rule with n cubature nodes $Q_{s,n}$. The overall error is then comprised of *dimension truncation error*, *finite element error*, and *cubature error* as

$$\|\mathbb{E}[u] - Q_{s,n}(u_{s,h})\| \leq \|\mathbb{E}[u - u_s]\| + \|\mathbb{E}[u_s - u_{s,h}]\| + \|\mathbb{E}[u_{s,h}] - Q_{s,n}(u_{s,h})\|,$$

for some appropriately chosen norm $\|\cdot\|$. We focus on the cubature error and discuss the spatial discretization error briefly in Section 5.3. We remark that the order of the last two error contributions—the finite element error and cubature error, respectively—can be flipped when the diffusion coefficient is represented using a sequence of bounded random variables (cf., e.g., [36, 37, 39]), but applying this tactic for the lognormal model would require making additional assumptions to ensure that the QMC approximation of $\mathbb{E}[\|u_{s,h}\|_{H^2}]$ is convergent. Hence, in this paper, we perform the QMC analysis for the dimensionally truncated, finite element solution of the parametric PDE.

In recent years, QMC methods have been demonstrated to be very effective at approximating the response statistics of PDE problems such as (1.1). The success of modern QMC theory for uncertainty quantification can largely be attributed to the introduction of weighted spaces (in the sense of Sloan and Woźniakowski [54] and Hickernell [32]) used to analyze the cubature error. While traditional QMC methods can also exhibit faster-than-Monte-Carlo convergence rates, as a general rule, the convergence rates of classical QMC methods have exponential dependence on the dimension of the integration problem in the unweighted setting [53]. However, QMC error bounds developed in weighted spaces for integrands satisfying certain moderate smoothness and anisotropy conditions can be shown to be independent of the stochastic dimension while retaining their faster-than-Monte-Carlo convergence rates. For a detailed survey of the development of QMC methods over the past few decades, see [14].

QMC methods are well suited for large scale computations since they are trivial to parallelize. Other advantages include the possibility to generate the cubature point sets for certain classes of QMC methods, such as randomly shifted rank-1 lattice rules or digital nets, on the fly. Randomization of QMC rules also enables the computation of practical error estimates. These features make QMC methods ideal for heavy-duty uncertainty quantification compared to regular Monte Carlo methods (slow convergence rate) or sparse grids (not easily parallelizable). QMC methods have been applied successfully to many problems such as Bayesian inverse problems [12, 29, 47, 48], spectral eigenvalue problems [20], optimization under uncertainty [25, 26, 27], the Schrödinger equation [56], the wave equation [17], problems arising in quantum physics [28], and various others.

QMC error bounds in the weighted space setting have the following generic form:

$$\text{root mean squared error} \leq C_{\boldsymbol{\gamma},n,s}\|G\|_{s,\boldsymbol{\gamma}},$$

where $G\colon [0,1]^s \to \mathbb{R}$ denotes the integrand, $\boldsymbol{\gamma} = (\gamma_{\mathfrak{u}})_{\mathfrak{u}\subseteq\{1,\dots,s\}}$ denotes a collection of positive weights, $C_{\boldsymbol{\gamma},n,s} > 0$ denotes the (*shift-averaged*) *worst case error* depending on the weights $\boldsymbol{\gamma}$, the number of QMC nodes n, and the dimension s, while $\|\cdot\|_{s,\boldsymbol{\gamma}}$ denotes the norm of an appropriately chosen weighted Sobolev space. The analysis of the cubature error for PDE uncertainty quantification typically consists of the following steps: first, the parametric regularity of the integrand needs to be estimated by deriving *a priori* upper bounds for the higher-order partial derivatives $\partial^{\boldsymbol{\nu}} G$ with respect to the (uncertain) variables. In consequence, these *a priori* upper bounds can then be used to estimate the weighted Sobolev norm $\|G\|_{s,\boldsymbol{\gamma}}$ of the integrand. Finally, the weights $\boldsymbol{\gamma}$ in the QMC error bound serve essentially as free parameters and can be chosen freely to optimize the cubature error bound. Typically, one chooses the weights in such a way that $C_{\boldsymbol{\gamma},n,s}\|G\|_{s,\boldsymbol{\gamma}} \leq C_{\boldsymbol{\gamma},n} < \infty$ for some constant $C_{\boldsymbol{\gamma},n} > 0$ not dependent on s, ensuring that the error rate of the resulting cubature rule can be bounded independently of the dimension s. From a practical point of view, the weights that minimize the error bound can be used as inputs to the so-called *component-by-component (CBC) algorithm* [10, 42], which produces a QMC rule satisfying the theoretically derived error bound. For a detailed description of this procedure, the reader is referred to the surveys [36, 37]. There also exist freely available software [6, 35] for the construction of QMC point sets.

A common feature of virtually all of the aforementioned QMC literature related to PDE uncertainty quantification is that the QMC rule is designed for the *non-discretized* PDE problem (1.1), whereas, in practical computations, one only has access to a discrete approximation of the PDE system. Of course, as long as one uses a conforming finite element (FE) method to perform the discretization of the PDE, the parametric regularity of the non-discretized PDE naturally transfers to the discretized PDE problem, and the theory remains sound even for the discretized PDE system. However, in many cases it is preferable to consider *non-conforming FE methods* such as DG methods (e.g., see the books [11, 16, 31, 33, 46] for a comprehensive overview and [2] for a unified analysis framework for elliptic PDEs using DG) to solve PDEs with uncertain coefficients. Using DG methods, the inter-element continuity constraint of conforming FE methods is dropped, and concatenations of arbitrary local polynomials with support in only one element can be used as test and trial functions. This property resembles the finite volume (FV) approach, in which the solution per element is approximated by a local constant. As a consequence, DG methods have similar parallelization properties as FV (but do not need any reconstructions to achieve higher order). Moreover, the problem of hanging nodes is intrinsically bypassed in the DG framework allowing for more general meshes than conforming FE; see [4]. Beyond that, DG directly supports hp refinement, where both the mesh and the local degree of approximating polynomials can be adapted locally; see [4, 5, 44]. Lastly, DG comes with an intuitive notion of local mass conservation, since it is based on inter-element fluxes (and mass conservative in an element itself). These advantages have made DG methods a popular tool for many applications. In particular, they have been used in the field of computational fluid dynamics (see [52] for a survey—notable mentions are DG methods for incompressible Navier–Stokes [7] and shallow-water equations [1]), multicomponent reactive transport [55], and many more. However, since discontinuous Galerkin methods are non-conforming, we cannot directly apply the existing QMC theory.

This paper tries to bridge this theoretical gap. It is structured as follows: Notations and preliminaries are introduced in Section 2. Section 3 describes randomly shifted lattice rules, and Section 4 gives a brief overview on the analysis of conforming FE methods. DG in

the QMC framework is presented in Section 5, and the corresponding parametric regularity analysis is considered in Section 6. Numerical results, which confirm our theoretical findings, are given in Section 7, while a short conclusion wraps up our exposition.

2. Notations, preliminaries, and assumptions. Let $\mathcal{F} := \{\boldsymbol{\nu} \in \mathbb{N}_0^{\mathbb{N}} \mid |\mathrm{supp}(\boldsymbol{\nu})| < \infty\}$ denote the set of all multi-indices with finite support $\mathrm{supp}(\boldsymbol{\nu}) := \{j \in \mathbb{N} \mid \nu_j \neq 0\}$. Moreover, we denote the order of a multi-index $\boldsymbol{\nu} = (\nu_1, \nu_2, \ldots) \in \mathcal{F}$ by

$$|\boldsymbol{\nu}| := \sum_{j \in \mathrm{supp}(\boldsymbol{\nu})} \nu_j.$$

We will also use the shorthand notation $\{1 : s\} := \{1, \ldots, s\}$.

The relevant function space for our PDE problem will be $V := H_0^1(D)$. Its dual space $V' = H^{-1}(D)$ is understood with respect to the pivot space $H := L^2(D)$, which we identify with its own dual. We define

$$\|v\|_V := \|\nabla v\|_H.$$

Let U be a space of parameters and for the moment not specified. We consider the variational formulation of the PDE problem (1.1): for all $\boldsymbol{y} \in U$, find $u(\cdot, \boldsymbol{y}) \in V$ such that

$$(2.1) \qquad \int_D a(\boldsymbol{x}, \boldsymbol{y}) \nabla u(\boldsymbol{x}, \boldsymbol{y}) \cdot \nabla v(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} = \langle f, v \rangle \qquad \text{for all } v \in V,$$

where $f \in V'$ and $\langle \cdot, \cdot \rangle := \langle \cdot, \cdot \rangle_{V', V}$ denotes the duality pairing of V and V'.

Uniform and affine setting. In the uniform and affine setting, we fix the set of parameters to be $U := [-\frac{1}{2}, \frac{1}{2}]^{\mathbb{N}}$, define

$$a(\boldsymbol{x}, \boldsymbol{y}) := a_0(\boldsymbol{x}) + \sum_{j=1}^{\infty} y_j \psi_j(\boldsymbol{x}), \qquad \boldsymbol{x} \in D, \ \boldsymbol{y} \in U,$$

and assume the following:

(U1) $a_0 \in L^{\infty}(D)$ and $\sum_{j=1}^{\infty} \|\psi_j\|_{L^{\infty}} < \infty$;

(U2) there exist $a_{\min}, a_{\max} > 0$ such that $a_{\min} \leq a(\boldsymbol{x}, \boldsymbol{y}) \leq a_{\max}$ for all $\boldsymbol{x} \in D$ and $\boldsymbol{y} \in U$;

(U3) $\sum_{j=1}^{\infty} \|\psi_j\|_{L^{\infty}}^p < \infty$ for some $0 < p < 1$;

(U4) $a_0 \in W^{1,\infty}(D)$ and $\sum_{j=1}^{\infty} \|\psi_j\|_{W^{1,\infty}} < \infty$, where

$$\|v\|_{W^{1,\infty}} := \max\{\|v\|_{L^{\infty}}, \|\nabla v\|_{L^{\infty}}\};$$

(U5) the spatial domain $D \subset \mathbb{R}^d$, $d \in \{1, 2, 3\}$, is a convex and bounded polyhedron.

We consider the expected value

$$\mathbb{E}[u] = \int_U u(\cdot, \boldsymbol{y}) \, \mathrm{d}\boldsymbol{y}.$$

Lognormal setting. In the lognormal setting, we fix the set of (admissible) parameters to be $U := U_{\boldsymbol{\beta}} := \{\boldsymbol{y} \in \mathbb{R}^{\mathbb{N}} : \sum_{j \geq 1} \beta_j |y_j| < \infty\}$, where $\boldsymbol{\beta} := (\beta_j)_{j \geq 1}$ and $\beta_j := \|\psi_j\|_{L^{\infty}}$. We model the input coefficient as

$$(2.2) \qquad a(\boldsymbol{x}, \boldsymbol{y}) := a_0(\boldsymbol{x}) \exp\left(\sum_{j=1}^{\infty} y_j \psi_j(\boldsymbol{x}) \right), \qquad \boldsymbol{x} \in D, \ \boldsymbol{y} \in U,$$

and assume the following:

(L1) $a_0 \in L^\infty(D)$ and $\sum_{j=1}^\infty \|\psi_j\|_{L^\infty} < \infty$;

(L2) there hold $\min_{\boldsymbol{x} \in \overline{D}} a_0(\boldsymbol{x}) > 0$ and $\max_{\boldsymbol{x} \in \overline{D}} a_0(\boldsymbol{x}) < \infty$, and we define

$$a_{\min}(\boldsymbol{y}) := \big(\min_{\boldsymbol{x} \in \overline{D}} a_0(\boldsymbol{x}) \big) \exp\bigg(-\sum_{j=1}^\infty |y_j| b_j \bigg),$$

$$a_{\max}(\boldsymbol{y}) := \big(\max_{\boldsymbol{x} \in \overline{D}} a_0(\boldsymbol{x}) \big) \exp\bigg(\sum_{j=1}^\infty |y_j| b_j \bigg),$$

for all $\boldsymbol{y} \in U_{\boldsymbol{\beta}}$.

(L3) $\sum_{j=1}^\infty \|\psi_j\|_{L^\infty}^p < \infty$ for some $0 < p < 1$;

(L4) $a_0 \in W^{1,\infty}(D)$ and $\sum_{j=1}^\infty \|\psi_j\|_{W^{1,\infty}} < \infty$;

(L5) the spatial domain $D \subset \mathbb{R}^d$, $d \in \{1, 2, 3\}$, is a convex and bounded polyhedron.

Since the set U has full measure with respect to the infinite-dimensional product Gaussian measure $\mu_G = \bigotimes_{j=1}^\infty \mathcal{N}(0,1)$, we can identify [50, Lem. 2.28]

$$\mathbb{E}[u] = \int_{\mathbb{R}^{\mathbb{N}}} u(\cdot, \boldsymbol{y}) \, \mu_G(\mathrm{d}\boldsymbol{y}) = \int_U u(\cdot, \boldsymbol{y}) \, \mu_G(\mathrm{d}\boldsymbol{y}).$$

Thus, it is sufficient to constrain our parametric regularity analysis of $u(\cdot, \boldsymbol{y})$ to the parameters $\boldsymbol{y} \in U$ for which the parametric PDE problem is well-defined.

REMARK 2.1.

(i) Let $(\Omega, \Gamma, \mathbb{P})$ be a probability space. Every square-integrable Gaussian random field $Z : D \times \Omega \to \mathbb{R}$ with a continuous, symmetric, and positive definite covariance function admits a Karhunen–Loève expansion

$$Z(\boldsymbol{x}, \omega) = \sum_{j=1}^\infty \sqrt{\lambda_j} y_j(\omega) \psi_j(\boldsymbol{x}), \qquad y_j \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1),$$

where $\{\psi_j\}_{j=1}^\infty$ constitutes an orthonormal basis for $L^2(D)$ and $\lambda_1 \geq \lambda_2 \geq \cdots \geq 0$ with $\lim_{j \to \infty} \lambda_j = 0$. However, not all random fields with such expansions are Gaussian. Hence, we concentrate on a broader class of random fields characterized by expansions of the form (2.2), which we refer to as the "lognormal model" for brevity.

(ii) Assumption (L2) together with [21, Lem. 3.10] ensure that both $a_{\max}(\cdot)$ and $1/a_{\min}(\cdot)$ are integrable with respect to the Gaussian product measure $\bigotimes_{j \geq 1} \mathcal{N}(0,1)$.

3. Quasi-Monte Carlo cubature. Since QMC methods can only be applied to finite-dimensional integrals and the analysis of the dimension truncation error is independent of the chosen spatial discretization scheme (cf., e.g., [18, 36]), we restrict our analysis to the finite-dimensional setting in what follows.

3.1. Uniform and affine model. We are interested in solving an s-dimensional integration problem

$$I_s(G) := \int_{[-1/2, 1/2]^s} G(\boldsymbol{y}) \, \mathrm{d}\boldsymbol{y}.$$

As our QMC estimator of $I_s(G)$, we take

$$Q_{\mathrm{ran}}(G) := \frac{1}{nR} \sum_{i=1}^n \sum_{r=1}^R G(\{\boldsymbol{t}_i + \boldsymbol{\Delta}^{(r)}\} - \tfrac{1}{2}),$$

where $\boldsymbol{\Delta}^{(1)}, \dots, \boldsymbol{\Delta}^{(R)}$ are i.i.d. realizations of a random variable $\boldsymbol{\Delta} \sim U([0,1]^s)$, $\{\cdot\}$ denotes the componentwise fractional part, $\frac{1}{2} := [\frac{1}{2}, \dots, \frac{1}{2}]^{\mathrm{T}} \in \mathbb{R}^s$, and

$$(3.1) \qquad \boldsymbol{t}_i := \left\{ \frac{i\boldsymbol{z}}{n} \right\}, \qquad \text{for } i \in \{1, \dots, n\},$$

where $\boldsymbol{z} \in \{0, \dots, n-1\}^s$ is called the *generating vector*.

Let us assume that the integrand G belongs to a weighted, unanchored Sobolev space with bounded first-order mixed partial derivatives, the norm of which is given by

$$\|G\|_{s,\boldsymbol{\gamma}}^2 := \sum_{\mathfrak{u} \subseteq \{1:s\}} \frac{1}{\gamma_\mathfrak{u}} \int_{[-1/2,1/2]^{|\mathfrak{u}|}} \left(\int_{[-1/2,1/2]^{s-|\mathfrak{u}|}} \frac{\partial^{|\mathfrak{u}|}}{\partial \boldsymbol{y}_\mathfrak{u}} G(\boldsymbol{y}) \, \mathrm{d}\boldsymbol{y}_{-\mathfrak{u}} \right)^2 \mathrm{d}\boldsymbol{y}_\mathfrak{u}.$$

Here, $\boldsymbol{\gamma} := (\gamma_\mathfrak{u})_{\mathfrak{u} \subseteq \{1:s\}}$ is a collection of positive weights, $\mathrm{d}\boldsymbol{y}_\mathfrak{u} := \prod_{j \in \mathfrak{u}} \mathrm{d}y_j$, and $\mathrm{d}\boldsymbol{y}_{-\mathfrak{u}} := \prod_{j \in \{1:s\} \setminus \mathfrak{u}} \mathrm{d}y_j$, for $\mathfrak{u} \subseteq \{1:s\}$.

The following well-known result shows that it is possible to construct a generating vector satisfying a rigorous error bound using a CBC algorithm [42] (see also [14]).

LEMMA 3.1 (cf. [36, Thm. 5.1]). *Let G belong to the weighted unanchored Sobolev space over $[0,1]^s$ with weights $\boldsymbol{\gamma} = (\gamma_\mathfrak{u})_{\mathfrak{u} \subseteq \{1:s\}}$. A randomly shifted lattice rule with $n = 2^m$ points in s dimensions can be constructed by a CBC algorithm such that for R independent random shifts and for all $\lambda \in (1/2, 1]$, it holds that*

$$\sqrt{\mathbb{E}_{\boldsymbol{\Delta}}[|I_s(G) - Q_{\mathrm{ran}}(G)|^2]} \leq \frac{1}{\sqrt{R}} \left(\frac{2}{n} \sum_{\emptyset \neq \mathfrak{u} \subseteq \{1:s\}} \gamma_\mathfrak{u}^\lambda \varrho(\lambda)^{|\mathfrak{u}|} \right)^{1/(2\lambda)} \|G\|_{s,\boldsymbol{\gamma}},$$

where

$$\varrho(\lambda) := \frac{2\zeta(2\lambda)}{(2\pi^2)^\lambda}.$$

Here, $\zeta(x) := \sum_{k=1}^\infty k^{-x}$ is the Riemann zeta function *for $x > 1$, and $\mathbb{E}_{\boldsymbol{\Delta}}[\cdot]$ denotes the expected value with respect to uniformly distributed random shifts over $[0,1]^s$.*

3.2. Lognormal model. We are interested in solving an s-dimensional integration problem

$$I_s^\varphi(G) := \int_{\mathbb{R}^s} G(\boldsymbol{y}) \prod_{j=1}^s \varphi(y_j) \, \mathrm{d}\boldsymbol{y} = \int_{[0,1]^s} G(\Phi^{-1}(\boldsymbol{w})) \, \mathrm{d}\boldsymbol{w},$$

where $\varphi \colon \mathbb{R} \to \mathbb{R}_+$ is the probability density function of the standard normal distribution, i.e.,

$$(3.2) \qquad \varphi(x) := \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} x^2 \right), \qquad x \in \mathbb{R},$$

and Φ^{-1} denotes the inverse cumulative distribution function of $\prod_{j=1}^s \varphi(y_j)$.

As our QMC estimator of $I_s(G)$ we take

$$Q_{\mathrm{ran}}^\varphi(G) := \frac{1}{nR} \sum_{i=1}^n \sum_{r=1}^R G(\Phi^{-1}(\{\boldsymbol{t}_i + \boldsymbol{\Delta}^{(r)}\})),$$

where $\boldsymbol{\Delta}^{(1)}, \dots, \boldsymbol{\Delta}^{(R)}$ are i.i.d. realizations of a random variable $\boldsymbol{\Delta} \sim U([0,1]^s)$ and the nodes $\boldsymbol{t}_1, \dots, \boldsymbol{t}_n \in [0,1]^s$ are defined as in (3.1).

Let us assume that the integrand G belongs to a special weighted Sobolev space in $\mathbb{R}^s$ with bounded first-order mixed partial derivatives, the norm of which is given by

$$\|G\|_{s,\boldsymbol{\gamma}}^2 :=$$
$$\sum_{\mathfrak{u}\subseteq\{1:s\}} \frac{1}{\gamma_\mathfrak{u}} \int_{\mathbb{R}^{|\mathfrak{u}|}} \left(\int_{\mathbb{R}^{s-|\mathfrak{u}|}} \frac{\partial^{|\mathfrak{u}|}}{\partial \boldsymbol{y}_\mathfrak{u}} G(\boldsymbol{y}) \left(\prod_{j\in\{1:s\}\setminus\mathfrak{u}} \varphi(y_j) \right) \mathrm{d}\boldsymbol{y}_{-\mathfrak{u}} \right)^2 \left(\prod_{j\in\mathfrak{u}} \varpi_j^2(y_j) \right) \mathrm{d}\boldsymbol{y}_\mathfrak{u},$$

where $\boldsymbol{\gamma} := (\gamma_\mathfrak{u})_{\mathfrak{u}\subseteq\{1:s\}}$ is a collection of positive weights, and we define the weight functions

$$(3.3) \qquad \varpi_j(x) := \exp(-\alpha_j |x|), \qquad \alpha_j > 0, \quad j \in \mathbb{Z}_+.$$

The parameters $\alpha_j > 0$ in (3.3) are for the moment arbitrary, but we will end up fixing their values later on in order to obtain dimension-independent QMC convergence for the lognormal model problem.

In analogy to the affine and uniform setting, the following well-known result gives an error bound for a QMC rule based on a generating vector constructed using the CBC algorithm.

LEMMA 3.2 (cf. [41] and [22, Thm. 15]). *Let G belong to the weighted function space over $\mathbb{R}^s$ with weights $\boldsymbol{\gamma} = (\gamma_\mathfrak{u})_{\mathfrak{u}\subseteq\{1:s\}}$, let $\varphi : \mathbb{R} \to \mathbb{R}_+$ be the standard normal density defined by (3.2), and let the weight functions ϖ_j be defined by (3.3). A randomly shifted lattice rule with $n = 2^m$ points in s dimensions can be constructed by a CBC algorithm such that for R independent random shifts and for all $\lambda \in (1/2, 1]$, it holds that*

$$\sqrt{\mathbb{E}_{\boldsymbol{\Delta}}[|I_s^\varphi(G) - Q_{\mathrm{ran}}^\varphi(G)|^2]} \le \frac{1}{\sqrt{R}} \left(\frac{2}{n} \sum_{\emptyset\neq\mathfrak{u}\subseteq\{1:s\}} \gamma_\mathfrak{u}^\lambda \prod_{j\in\mathfrak{u}} \varrho_j(\lambda) \right)^{1/(2\lambda)} \|G\|_{s,\boldsymbol{\gamma}},$$

where

$$\varrho_j(\lambda) := 2\left(\frac{\sqrt{2\pi}\exp(\alpha_j^2/\eta_*)}{\pi^{2-2\eta_*}(1-\eta_*)\eta_*} \right)^\lambda \zeta(\lambda + \tfrac{1}{2}), \qquad \eta_* := \frac{2\lambda-1}{4\lambda}.$$

Here, $\zeta(x) := \sum_{k=1}^\infty k^{-x}$ is the Riemann zeta function *for $x > 1$, and $\mathbb{E}_{\boldsymbol{\Delta}}[\cdot]$ denotes the expected value with respect to uniformly distributed random shifts over $[0,1]^s$.*

4. Conforming FE methods. In the case of conforming finite element discretizations for the elliptic PDE problem, it is enough to analyze the parametric regularity of the continuous problem. The parametric regularity results are inherited by the conforming FE solution. Below, we briefly recap the main parametric regularity results for the affine and uniform as well as the lognormal model.

Let us consider the weak formulation (2.1): for all $\boldsymbol{y} \in U$, find $u(\cdot, \boldsymbol{y}) \in V$ such that

$$(4.1) \qquad \int_D a(\boldsymbol{x}, \boldsymbol{y})\nabla u(\boldsymbol{x}, \boldsymbol{y}) \cdot \nabla v(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} = \langle f, v \rangle \qquad \text{for all } v \in V.$$

If we use a conforming FE method to solve the equation above, the variational formulation stays the same:

$$\int_D a(\boldsymbol{x}, \boldsymbol{y})\nabla u_h(\boldsymbol{x}, \boldsymbol{y}) \cdot \nabla v_h(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} = \langle f, v_h \rangle \qquad \text{for all } v_h \in V_h,$$

where the test and trial space V_h is a finite-dimensional subspace of V.

Uniform and affine setting. It was shown in [9] that the solution to (4.1) satisfies the regularity bounds

$$(4.2) \qquad \|\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u(\cdot, \boldsymbol{y})\|_V \le |\boldsymbol{\nu}|! \, \boldsymbol{b}^{\boldsymbol{\nu}} \frac{\|f\|_{V'}}{a_{\min}}, \qquad \boldsymbol{b} := (b_j)_{j \ge 1}, \ b_j := \frac{\|\psi_j\|_{L^\infty}}{a_{\min}},$$

for all finitely supported multi-indices $\boldsymbol{\nu} \in \mathcal{F}$ and $\boldsymbol{y} \in U$. Since V_h is now a subspace of V, it follows that the same regularity bound holds for the dimensionally truncated FE approximation:

$$(4.3) \qquad \|\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_{s,h}(\cdot, \boldsymbol{y})\|_V \le |\boldsymbol{\nu}|! \, \boldsymbol{b}^{\boldsymbol{\nu}} \frac{\|f\|_{V'}}{a_{\min}}.$$

Lognormal setting. The corresponding regularity bound for the lognormal case is [22]

$$\|\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u(\cdot, \boldsymbol{y})\|_V \le \frac{|\boldsymbol{\nu}|!}{(\log 2)^{|\boldsymbol{\nu}|}} \boldsymbol{\beta}^{\boldsymbol{\nu}} \frac{\|f\|_{V'}}{a_{\min}(\boldsymbol{y})}, \qquad \boldsymbol{\beta} := (\beta_j)_{j \ge 1}, \ \beta_j := \|\psi_j\|_{L^\infty},$$

for all finitely supported multi-indices $\boldsymbol{\nu} \in \mathcal{F}$ and $\boldsymbol{y} \in U_{\boldsymbol{\beta}}$. Note that u can be again replaced by a dimensionally truncated, conforming FE approximation $u_{s,h}$ in the above inequality.

5. Discontinuous Galerkin FE methods. In this section, we analyze the DG method provided that $D \subset \mathbb{R}^d$ is a convex polyhedron, $\emptyset \ne U \subset \mathbb{R}^{\mathbb{N}}$, $f \in L^2(D)$, and that there are

$$0 < a_{\min}(\boldsymbol{y}) = \min_{x \in \bar{D}} a(\boldsymbol{x}, \boldsymbol{y}) \qquad \text{and} \qquad a_{\max}(\boldsymbol{y}) = \max_{x \in \bar{D}} a(\boldsymbol{x}, \boldsymbol{y}) < \infty.$$

This setup holds for both the uniform and affine and the lognormal cases.

In DG, the idea is to modify the variational formulation, and we can no longer exploit conformity to obtain regularity bounds for the DG solutions to (1.1). This means that the regularity analysis must be rewritten for the DG system, which will also affect the choice of the optimal QMC rule for the computation of the expectation of the stochastic response.

The most prominent DG method for the model problem (1.1) is given by the family of interior penalty (IPDG) methods. For these methods, we define

$$V_h = \{v_h \in L^2(D) \colon v_h|_T \in \mathcal{P}_k(T) \ \forall T \in \mathcal{T}_h\},$$

where $\mathcal{P}_k(T)$ denotes the space of (multivariate) polynomials of (total) degree at most k on T, and $\mathcal{T}_h$ is supposed to be a member of a shape and contact regular mesh sequence in the sense of [11, Def. 1.38]. Basically, this says that the mesh elements are neither distorted nor have arbitrary small angles. Purely for the sake of simplicity in the notation, we assume the mesh to be geometrically conforming (have no hanging nodes). Importantly, we do not require our mesh to be simplicial or quadrilateral, etc. In fact, DG methods can easily be used on significantly more general meshes. Note that V_h does not impose any inter-element continuity constraints. Thus, we also define the mean value $\{\![\cdot]\!\}$ and the jump $[\![\cdot]\!]$ on a common face F of the elements $T^+ \ne T^- \in \mathcal{T}_h$ by

$$\{\![v_h]\!\} = \frac{1}{2} (v_h|_{T^+} + v_h|_{T^-}) \qquad \text{and} \qquad [\![v_h]\!] = v_h|_{T^+} \boldsymbol{n}_{T^+} + v_h|_{T^-} \boldsymbol{n}_{T^-},$$

where $\boldsymbol{n}_T$ denotes the outward pointing unit normal with respect to T. Importantly, the jump turns a scalar function to a vector function. Likewise, the average of a vector quantity is defined component-wise, and if $F \subset \partial D$, then we set

$$\{\![v_h]\!\} = v_h|_T \qquad \text{and} \qquad [\![v_h]\!] = v_h|_T \boldsymbol{n},$$

where T refers to the unique $T \in \mathcal{T}_h$ with $F \subset \partial T$ and $\boldsymbol{n}$ is the outward pointing unit normal of D. The set of all faces is denoted by $\mathcal{F}$.

Using these definitions, we can define the DG bilinear form $B_h \colon V_h \times V_h \to \mathbb{R}$ as

$$B_h(\boldsymbol{y}; u_h(\cdot, \boldsymbol{y}), v_h)$$

$$= \sum_{T \in \mathcal{T}_h} \int_T a(\boldsymbol{x}, \boldsymbol{y}) \nabla u_h(\boldsymbol{x}, \boldsymbol{y}) \cdot \nabla v_h(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}$$

$$+ \sum_{F \in \mathcal{F}} \int_F \left[\theta \{ [a(\boldsymbol{x}, \boldsymbol{y}) \nabla v_h(\boldsymbol{x})] \} \cdot [\![u_h(\boldsymbol{x}, \boldsymbol{y})]\!] \right.$$

$$- \{ [a(\boldsymbol{x}, \boldsymbol{y}) \nabla u_h(\boldsymbol{x}, \boldsymbol{y})] \} \cdot [\![v_h(\boldsymbol{x})]\!]$$

$$\left. + \frac{\eta(\boldsymbol{y})}{h_F} [\![u_h(\boldsymbol{x}, \boldsymbol{y})]\!] \cdot [\![v_h(x)]\!] \right] \mathrm{d}\sigma$$

(5.1)
$$= \sum_{T \in \mathcal{T}_h} \int_T a(\boldsymbol{x}, \boldsymbol{y}) \nabla u_h(\boldsymbol{x}, \boldsymbol{y}) \cdot \nabla v_h(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}$$

$$+ \sum_{T \in \mathcal{T}_h} \int_{\partial T} \left[\theta \{ [a(\boldsymbol{x}, \boldsymbol{y}) \nabla v_h(\boldsymbol{x})] \} \cdot u_h(\boldsymbol{x}, \boldsymbol{y}) \boldsymbol{n}_T(\boldsymbol{x}) \right.$$

$$- \frac{1}{2} a(\boldsymbol{x}, \boldsymbol{y}) \nabla u_h(\boldsymbol{x}, \boldsymbol{y}) \cdot [\![v_h(\boldsymbol{x})]\!]$$

$$\left. + \frac{\eta(\boldsymbol{y})}{h_F} u_h(\boldsymbol{x}, \boldsymbol{y}) \boldsymbol{n}_T(\boldsymbol{x}) \cdot [\![v_h(x)]\!] \right] \mathrm{d}\sigma$$

and the corresponding linear form $\mathcal{L} \colon V_h \to \mathbb{R}$ as

(5.2)
$$\mathcal{L}(v_h) = \int_D f(\boldsymbol{x}) v_h(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}.$$

Here, h_F denotes the diameter of the considered face. The parameter $\theta \in \{-1, 0, +1\}$ is chosen to be equal to 1 for the non-symmetric IP method (NIPDG), equal to 0 for the incomplete IP method (IIPDG), and equal to -1 for the symmetric IP method (SIPDG). The role of the parameter $\eta(\boldsymbol{y})$ is to penalize jumps in u_h across interfaces and thereby stabilize the method. It is well known that the NIPDG method is stable for parameters $\eta(\boldsymbol{y}) \geq 0$, while IIPDG and SIPDG methods are stable if $\eta(\boldsymbol{y}) \geq \bar{\eta}(\boldsymbol{y}) > 0$. Here, $\bar{\eta}(\boldsymbol{y})$ depends, among others, on the infimum and supremum of $a(\cdot, \boldsymbol{y})$; see Lemma 5.3. Thus, we cannot expect to choose η independent of $\boldsymbol{y}$ for the IIPDG and SIPDG methods in the lognormal case. The NIPDG method for $\eta(\boldsymbol{y}) = 0$ is attributed to Oden, Babuška, and Baumann [43], but we exclude it from our considerations (since its analysis is more involved) and assume that $\eta(\boldsymbol{y}) > 0$ is used. For a detailed discussion of these methods, the reader is referred to the books [11, 16, 31, 33, 46] and the references therein.

5.1. Preliminaries. Notably, $V_h \not\subset H^1(D)$, and thus we have to define a norm that allows for a generalization of Poincaré's inequality. We choose

$$\|v_h\|_{V_h}^2 = \sum_{T \in \mathcal{T}_h} \|\sqrt{a(\cdot, \boldsymbol{y})} \nabla v_h(\cdot)\|_{L^2(T)}^2 + \sum_{F \in \mathcal{F}} \frac{\eta(\boldsymbol{y})}{h_F} \|[\![v_h]\!]\|_{L^2(F)}^2.$$

The norm $\| \cdot \|_{V_h}$ actually also depends on $\boldsymbol{y}$, which is suppressed in the notation. In order to deduce Poincaré's inequality, we cite [11, Cor. 5.4]:

LEMMA 5.1. *Let $\mathcal{T}_h$ be a shape- and contact-regular mesh sequence member. For the norm*

$$\|v_h\|_{\mathrm{dG}} = \left(\sum_{T \in \mathcal{T}_h} \|\nabla v_h\|_{L^2(T)}^2 + \sum_{F \in \mathcal{F}} \frac{1}{h_F} \|[\![v_h]\!]\|_{L^2(F)}^2 \right)^{1/2},$$

there is a constant $\sigma > 0$, independent of h, such that

$$\forall v_h \in V_h : \qquad \|v_h\|_{L^2(D)} \leq \sigma \|v_h\|_{\mathrm{dG}}$$

holds.

Here, the notion "shape and contact regular mesh sequence" refers to [11, Def. 1.38] and requires that each mesh in the sequence admits a simplicial submesh without hanging nodes or deteriorating simplices.

This can be further adapted to suit our purposes by observing that

(5.3) $$\|v_h\|_{L^2(D)} \leq \sigma \|v_h\|_{\mathrm{dG}} \leq \underbrace{\frac{\sigma}{\min\{\sqrt{a_{\min}(\boldsymbol{y})}, \sqrt{\eta(\boldsymbol{y})}\}}}_{=C_{\mathrm{Poin}}^{V_h}(\boldsymbol{y})} \|v_h\|_{V_h} \qquad \forall v_h \in V_h.$$

The Poincaré constant $C_{\mathrm{Poin}}^{V_h}(\boldsymbol{y})$ can be chosen independently of $\boldsymbol{y}$ in the uniform and affine setting (there are uniform upper and lower bounds for the equivalence constants between norms), which is not possible (by our proof techniques—the equivalence constants between norms depend on $\boldsymbol{y}$) in the lognormal setting. Together with assumptions (U1)–(U2) and (L1)–(L2), the following lemmas describe how $\eta(\boldsymbol{y})$ should be chosen in order to ensure the integrability of $C_{\mathrm{Poin}}^{V_h}(\boldsymbol{y})$:

LEMMA 5.2 ([11, Lem. 1.46]). *Let $\mathcal{T}_h$ be a member of a shape- and contact-regular mesh sequence. Then, for all $v_h \in V_h$, all $T \in \mathcal{T}_h$, all $F \subset \partial T$, and $h_T = \mathrm{diam}(T)$,*

$$h_T^{1/2} \|v_h\|_{L^2(F)} \leq C_{\mathrm{tr}} \|v_h\|_{L^2(T)},$$

where C_{tr} does not depend on h.

This allows for a stability estimate of the approximate u_h.

LEMMA 5.3. *Assume that there is $\tau > 1$ (independent of $\boldsymbol{y}$ and h) such that*

(5.4) $$\eta(\boldsymbol{y}) \geq \tau \frac{a_{\max}^2(\boldsymbol{y}) C_{\mathrm{tr}}^2 N_\partial (\theta - 1)^2}{4 a_{\min}(\boldsymbol{y})},$$

where C_{tr} denotes the constant of Lemma 5.2 and N_∂ is the maximum amount of faces of a mesh element (for simplicial meshes, $N_\partial = d + 1$). There is $\alpha > 0$ (independent of h) such that

$$\alpha \|u_h\|_{V_h}^2 \overset{\text{Ⓐ}}{\leq} B_h(\boldsymbol{y}; u_h(\cdot, \boldsymbol{y}), u_h(\cdot, \boldsymbol{y})) = F(u_h(\cdot, \boldsymbol{y}))$$
$$\leq \|f\|_{L^2(D)} \|u_h\|_{L^2(D)} \leq C_{\mathrm{Poin}}^{V_h}(\boldsymbol{y}) \|f\|_{L^2(D)} \|u_h\|_{V_h},$$

which gives the stability estimate

(5.5) $$\|u_h\|_{V_h} \leq \frac{C_{\mathrm{Poin}}^{V_h}(\boldsymbol{y})}{\alpha} \|f\|_{L^2(D)} \overset{\text{Ⓑ}}{=} \frac{\sigma}{\alpha \sqrt{a_{\min}(\boldsymbol{y})}} \|f\|_{L^2(D)},$$

where Ⓑ *holds if* $\eta(\boldsymbol{y}) \geq a_{\min}(\boldsymbol{y})$. *Additionally, for the NIPDG method* ($\theta = 1$), *inequality* Ⓐ *holds with equality and* $\alpha = 1$ *independent of* $\eta(\boldsymbol{y}) > 0$.

Proof. This is a direct consequence of [33, Lemma 7.30] except for the identity Ⓑ. However, Ⓑ follows directly from combining (5.3) and $\eta(\boldsymbol{y}) \geq a_{\min}(\boldsymbol{y})$. □

Additionally, we define the norm

$$\|v_h\|_{V_h^*}^2 = \|v_h\|_{V_h}^2 + \sum_{F \in \mathcal{F}} \frac{h_F}{\eta(\boldsymbol{y})} \|\{[a(\cdot, \boldsymbol{y}) \nabla v_h]\}\|_{L^2(F)}^2.$$

LEMMA 5.4. *We have* $\|v_h\|_{V_h} \leq \|v_h\|_{V_h^*}$. *Moreover, if there is a* $\tilde{\tau} > 0$ *(independent of* $\boldsymbol{y}$ *and* h*) such that*

$$(5.6) \qquad\qquad \eta(\boldsymbol{y}) \geq \tilde{\tau} \frac{a_{\max}(\boldsymbol{y})^2}{a_{\min}(\boldsymbol{y})} \qquad \text{for all } \boldsymbol{y},$$

then we have $C_{\mathrm{equiv}} > 0$ *independent of* $\boldsymbol{y}$ *and* h *such that*

$$\|v_h\|_{V_h^*} \leq C_{\mathrm{equiv}} \|v_h\|_{V_h} \qquad \text{for all } v_h \in V_h.$$

Proof. The first inequality $\|v_h\|_{V_h} \leq \|v_h\|_{V_h^*}$ is obvious. For the second inequality, we observe that we need to bound the additional term in the definition of $\| \cdot \|_{V_h^*}$ using

$$\begin{aligned}
\sum_{F \in \mathcal{F}} \frac{h_F}{\eta(\boldsymbol{y})} \|\{[a(\cdot, \boldsymbol{y}) \nabla v_h]\}\|_{L^2(F)}^2 &\leq \sum_{F \in \mathcal{F}} \frac{a_{\max}(\boldsymbol{y})^2 h_F}{\eta(\boldsymbol{y})} \|\{[\nabla v_h]\}\|_{L^2(F)}^2 \\
&\leq \sum_{F \in \mathcal{F}} \frac{a_{\max}(\boldsymbol{y})^2 C_{\mathrm{tr}}^2}{\eta(\boldsymbol{y})} \sum_{T \in \mathcal{T}_h : \partial T \supset F} \|\nabla v_h\|_{L^2(T)}^2 \\
&\leq \sum_{T \in \mathcal{T}_h} \frac{a_{\max}(\boldsymbol{y})^2 C_{\mathrm{tr}}^2 N_\partial}{\eta(\boldsymbol{y})} \|\nabla v_h\|_{L^2(T)}^2 \\
&\leq \sum_{T \in \mathcal{T}_h} \frac{a_{\max}(\boldsymbol{y})^2 C_{\mathrm{tr}}^2 N_\partial}{a_{\min}(\boldsymbol{y}) \eta(\boldsymbol{y})} \|\sqrt{a(\cdot, \boldsymbol{y})} \nabla v_h\|_{L^2(T)}^2.
\end{aligned}$$

Here, the second inequality is the discrete trace inequality in Lemma 5.2 using that $h_F = \mathrm{diam}(F) \leq \mathrm{diam}(T) = h_T$, for $F \subset T$. The constant becomes independent of $\boldsymbol{y}$ if (5.6) holds. □

REMARK 5.5. In the lognormal setting, the condition of Lemma 5.3 can be satisfied by using the NIPG method corresponding to $\theta = 1$. Lemma 5.4 suggests that we can define

$$\eta(\boldsymbol{y}) := \frac{a_{\max}(\boldsymbol{y})^2}{a_{\min}(\boldsymbol{y})} \qquad \text{for all } \boldsymbol{y}.$$

Analogously, in the affine and uniform case, we can choose

$$\eta(\boldsymbol{y}) := \eta = \frac{a_{\max}^2}{a_{\min}} \qquad \text{for all } \boldsymbol{y}.$$

Later on, we will choose $\eta(\boldsymbol{y})$ precisely in this way. Although the choice of $\eta(\boldsymbol{y})$ is more subtle for the SIPG and IIPG methods, we experienced that a too large $\eta(\boldsymbol{y})$ does not have a visual effect on the results in Section 7 while a too small $\eta(\boldsymbol{y})$ can lead to wrong approximations. In principle, increasing η adds inter-element diffusion to the method. Thus, enlarging η stabilizes the method while its scaling ensures that the convergence rates in h do not deteriorate. Too large η-values have an adversarial effect on the constants of DG methods, but as long as integrability (concerning $\boldsymbol{y}$) is ensured, the QMC methods will work.

5.2. Bounds for the derivatives with respect to the random variable. The ultimate goal of our analysis is to derive regularity bounds in the flavor of (4.2) and (4.3) for the diffusion problem that is discretized with an IPDG method instead of continuous finite elements. To this end, we start with some preliminary considerations before we prove Theorem 5.6 stating that

$$\|\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\cdot, \boldsymbol{y})\|_{V_h}$$
$$\leq \underbrace{C_{\mathrm{equiv}}^2 C_{\mathrm{DG}}'}_{=C_{\mathrm{DG}}} \sum_{\boldsymbol{0} \neq \boldsymbol{m} \leq \boldsymbol{\nu}} \binom{\boldsymbol{\nu}}{\boldsymbol{m}} \left[\left\| \frac{\partial_{\boldsymbol{y}}^{\boldsymbol{m}} a(\cdot, \boldsymbol{y})}{a(\cdot, \boldsymbol{y})} \right\|_{L^\infty(D)} + \left| \frac{\partial_{\boldsymbol{y}}^{\boldsymbol{m}} \eta(\boldsymbol{y})}{\eta(\boldsymbol{y})} \right| \right] \|\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}-\boldsymbol{m}} u_h(\cdot, \boldsymbol{y})\|_{V_h}$$

under certain conditions. Notably, the derivative $\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\cdot, \boldsymbol{y})$ is an element of V_h. This can be seen by similar arguments as conducted in [9] and is a well-known observation when approximating time-dependent equations by DG methods where the derivative concerns a unknown time instead of a random variable; see [8] and [11, Sect. 3.1].

To derive this regularity bound for the IPDG methods, we first observe that $\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} \mathcal{L}(v_h) = 0$ (see (5.2)) and that $a(\boldsymbol{x}, \boldsymbol{y})$ is continuous with respect to $\boldsymbol{x}$ since a_0 and all ψ_j have been assumed to be Lipschitz continuous. Thus,

(5.7)
$$B_h(\boldsymbol{y}; u_h(\cdot, \boldsymbol{y}), v_h)$$
$$= \sum_{T \in \mathcal{T}_h} \int_T a(\boldsymbol{x}, \boldsymbol{y}) \nabla u_h(\boldsymbol{x}, \boldsymbol{y}) \cdot \nabla v_h(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}$$
$$+ \sum_{T \in \mathcal{T}_h} \int_{\partial T} \left[\theta a(\boldsymbol{x}, \boldsymbol{y}) u_h(\boldsymbol{x}, \boldsymbol{y}) \boldsymbol{n}_T(\boldsymbol{x}) \cdot \{\!\{\nabla v_h(\boldsymbol{x})\}\!\} \right.$$
$$- \frac{1}{2} a(\boldsymbol{x}, \boldsymbol{y}) \nabla u_h(\boldsymbol{x}, \boldsymbol{y}) \cdot [\![v_h(\boldsymbol{x})]\!]$$
$$+ \frac{\eta(\boldsymbol{y})}{h_F} u_h(\boldsymbol{x}, \boldsymbol{y}) \boldsymbol{n}_T(\boldsymbol{x}) \cdot [\![v_h(\boldsymbol{x})]\!] \Big] \mathrm{d}\sigma,$$

and we can deduce by $B_h(\boldsymbol{y}; u_h(\cdot, \boldsymbol{y}), v_h) = \mathcal{L}(v_h)$ and $\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} \mathcal{L}(v_h) = 0$ that

$$0 = \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} B_h(\boldsymbol{y}; u_h(\cdot, \boldsymbol{y}), v_h)$$
$$\overset{(5.7)}{=} \sum_{T \in \mathcal{T}_h} \int_T \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} [a(\boldsymbol{x}, \boldsymbol{y}) \nabla u_h(\boldsymbol{x}, \boldsymbol{y})] \cdot \nabla v_h(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}$$
$$+ \sum_{T \in \mathcal{T}_h} \int_{\partial T} \left[\theta \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} [a(\boldsymbol{x}, \boldsymbol{y}) u_h(\boldsymbol{x}, \boldsymbol{y})] \boldsymbol{n}_T(\boldsymbol{x}) \cdot \{\!\{\nabla v_h(\boldsymbol{x})\}\!\} \right.$$
$$- \frac{1}{2} \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} [a(\boldsymbol{x}, \boldsymbol{y}) \nabla u_h(\boldsymbol{x}, \boldsymbol{y})] \cdot [\![v_h(\boldsymbol{x})]\!]$$
$$+ h_F^{-1} \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} [\eta(\boldsymbol{y}) u_h(\boldsymbol{x}, \boldsymbol{y})] \boldsymbol{n}_T(\boldsymbol{x}) \cdot [\![v_h(\boldsymbol{x})]\!] \Big] \mathrm{d}\sigma$$
$$= \sum_{T \in \mathcal{T}_h} \int_T \sum_{\boldsymbol{m} \leq \boldsymbol{\nu}} \binom{\boldsymbol{\nu}}{\boldsymbol{m}} \partial_{\boldsymbol{y}}^{\boldsymbol{m}} a(\boldsymbol{x}, \boldsymbol{y}) \nabla \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}-\boldsymbol{m}} u_h(\boldsymbol{x}, \boldsymbol{y}) \cdot \nabla v_h(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}$$
$$+ \sum_{T \in \mathcal{T}_h} \int_{\partial T} \left[\theta \sum_{\boldsymbol{m} \leq \boldsymbol{\nu}} \binom{\boldsymbol{\nu}}{\boldsymbol{m}} \partial_{\boldsymbol{y}}^{\boldsymbol{m}} a(\boldsymbol{x}, \boldsymbol{y}) \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}-\boldsymbol{m}} u_h(\boldsymbol{x}, \boldsymbol{y}) \boldsymbol{n}_T(\boldsymbol{x}) \cdot \{\!\{\nabla v_h(\boldsymbol{x})\}\!\} \right.$$
$$- \frac{1}{2} \sum_{\boldsymbol{m} \leq \boldsymbol{\nu}} \binom{\boldsymbol{\nu}}{\boldsymbol{m}} \partial_{\boldsymbol{y}}^{\boldsymbol{m}} a(\boldsymbol{x}, \boldsymbol{y}) \nabla \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}-\boldsymbol{m}} u_h(\boldsymbol{x}, \boldsymbol{y}) \cdot [\![v_h(\boldsymbol{x})]\!]$$

$$+ h_F^{-1} \sum_{\boldsymbol{m} \leq \boldsymbol{\nu}} \binom{\boldsymbol{\nu}}{\boldsymbol{m}} \partial_{\boldsymbol{y}}^{\boldsymbol{m}} \eta(\boldsymbol{y}) \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}-\boldsymbol{m}} u_h(\boldsymbol{x}, \boldsymbol{y}) \boldsymbol{n}_T(\boldsymbol{x}) \cdot [\![v_h(\boldsymbol{x})]\!] \Big] \mathrm{d}\sigma$$

$$= \sum_{T \in \mathcal{T}_h} \int_T \sum_{\boldsymbol{m} \leq \boldsymbol{\nu}} \binom{\boldsymbol{\nu}}{\boldsymbol{m}} \partial_{\boldsymbol{y}}^{\boldsymbol{m}} a(\boldsymbol{x}, \boldsymbol{y}) \nabla \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}-\boldsymbol{m}} u_h(\boldsymbol{x}, \boldsymbol{y}) \cdot \nabla v_h(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}$$

$$+ \sum_{F \in \mathcal{F}} \int_F \Big[\theta \sum_{\boldsymbol{m} \leq \boldsymbol{\nu}} \binom{\boldsymbol{\nu}}{\boldsymbol{m}} \partial_{\boldsymbol{y}}^{\boldsymbol{m}} a(\boldsymbol{x}, \boldsymbol{y}) [\![\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}-\boldsymbol{m}} u_h(\boldsymbol{x}, \boldsymbol{y})]\!] \cdot \{\!\{\nabla v_h(\boldsymbol{x})\}\!\}$$

$$- \sum_{\boldsymbol{m} \leq \boldsymbol{\nu}} \binom{\boldsymbol{\nu}}{\boldsymbol{m}} \partial_{\boldsymbol{y}}^{\boldsymbol{m}} a(\boldsymbol{x}, \boldsymbol{y}) \{\!\{\nabla \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}-\boldsymbol{m}} u_h(\boldsymbol{x}, \boldsymbol{y})\}\!\} \cdot [\![v_h(\boldsymbol{x})]\!]$$

$$+ h_F^{-1} \sum_{\boldsymbol{m} \leq \boldsymbol{\nu}} \binom{\boldsymbol{\nu}}{\boldsymbol{m}} \partial_{\boldsymbol{y}}^{\boldsymbol{m}} \eta(\boldsymbol{y}) [\![\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}-\boldsymbol{m}} u_h(\boldsymbol{x}, \boldsymbol{y})]\!] \cdot [\![v_h(\boldsymbol{x})]\!] \Big] \mathrm{d}\sigma,$$

where the third equality is a consequence of the Leibniz product rule and the fourth equality follows from the definitions of $\{\!\{\cdot\}\!\}$ and $[\![\cdot]\!]$ analogously to (5.1).

Next, we set $v_h = \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\cdot, \boldsymbol{y})$ and do some algebraic manipulations to obtain

$$\|\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\cdot, \boldsymbol{y})\|_{V_h}^2$$

$$= - \sum_{T \in \mathcal{T}_h} \int_T \sum_{\boldsymbol{0} \neq \boldsymbol{m} \leq \boldsymbol{\nu}} \binom{\boldsymbol{\nu}}{\boldsymbol{m}} \partial_{\boldsymbol{y}}^{\boldsymbol{m}} a(\boldsymbol{x}, \boldsymbol{y}) \nabla \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}-\boldsymbol{m}} u_h(\boldsymbol{x}, \boldsymbol{y}) \cdot \nabla \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\boldsymbol{x}, \boldsymbol{y}) \mathrm{d}\boldsymbol{x}$$

$$- \sum_{F \in \mathcal{F}} \int_F \Big[\theta \sum_{\boldsymbol{m} \leq \boldsymbol{\nu}} \binom{\boldsymbol{\nu}}{\boldsymbol{m}} \partial_{\boldsymbol{y}}^{\boldsymbol{m}} a(\boldsymbol{x}, \boldsymbol{y}) [\![\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}-\boldsymbol{m}} u_h(\boldsymbol{x}, \boldsymbol{y})]\!] \cdot \{\!\{\nabla \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\boldsymbol{x}, \boldsymbol{y})\}\!\}$$

$$- \sum_{\boldsymbol{m} \leq \boldsymbol{\nu}} \binom{\boldsymbol{\nu}}{\boldsymbol{m}} \partial_{\boldsymbol{y}}^{\boldsymbol{m}} a(\boldsymbol{x}, \boldsymbol{y}) \{\!\{\nabla \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}-\boldsymbol{m}} u_h(\boldsymbol{x}, \boldsymbol{y})\}\!\} \cdot [\![\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\boldsymbol{x}, \boldsymbol{y})]\!]$$

$$+ h_F^{-1} \sum_{\boldsymbol{0} \neq \boldsymbol{m} \leq \boldsymbol{\nu}} \binom{\boldsymbol{\nu}}{\boldsymbol{m}} \partial_{\boldsymbol{y}}^{\boldsymbol{m}} \eta(\boldsymbol{y}) [\![\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}-\boldsymbol{m}} u_h(\boldsymbol{x}, \boldsymbol{y})]\!] \cdot [\![\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\boldsymbol{x}, \boldsymbol{y})]\!] \Big] \mathrm{d}\sigma$$

$$= - \sum_{T \in \mathcal{T}_h} \int_T \sum_{\boldsymbol{0} \neq \boldsymbol{m} \leq \boldsymbol{\nu}} \binom{\boldsymbol{\nu}}{\boldsymbol{m}} \partial_{\boldsymbol{y}}^{\boldsymbol{m}} a(\boldsymbol{x}, \boldsymbol{y}) \nabla \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}-\boldsymbol{m}} u_h(\boldsymbol{x}, \boldsymbol{y}) \cdot \nabla \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\boldsymbol{x}, \boldsymbol{y}) \mathrm{d}\boldsymbol{x}$$

$$- \sum_{F \in \mathcal{F}} \int_F \Big[\theta \sum_{\boldsymbol{0} \neq \boldsymbol{m} \leq \boldsymbol{\nu}} \binom{\boldsymbol{\nu}}{\boldsymbol{m}} \partial_{\boldsymbol{y}}^{\boldsymbol{m}} a(\boldsymbol{x}, \boldsymbol{y}) [\![\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}-\boldsymbol{m}} u_h(\boldsymbol{x}, \boldsymbol{y})]\!] \cdot \{\!\{\nabla \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\boldsymbol{x}, \boldsymbol{y})\}\!\}$$

$$- \sum_{\boldsymbol{0} \neq \boldsymbol{m} \leq \boldsymbol{\nu}} \binom{\boldsymbol{\nu}}{\boldsymbol{m}} \partial_{\boldsymbol{y}}^{\boldsymbol{m}} a(\boldsymbol{x}, \boldsymbol{y}) \{\!\{\nabla \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}-\boldsymbol{m}} u_h(\boldsymbol{x}, \boldsymbol{y})\}\!\} \cdot [\![\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\boldsymbol{x}, \boldsymbol{y})]\!]$$

$$+ h_F^{-1} \sum_{\boldsymbol{0} \neq \boldsymbol{m} \leq \boldsymbol{\nu}} \binom{\boldsymbol{\nu}}{\boldsymbol{m}} \partial_{\boldsymbol{y}}^{\boldsymbol{m}} \eta(\boldsymbol{y}) [\![\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}-\boldsymbol{m}} u_h(\boldsymbol{x}, \boldsymbol{y})]\!] \cdot [\![\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\boldsymbol{x}, \boldsymbol{y})]\!] \Big] \mathrm{d}\sigma$$

$$\underbrace{- \sum_{F \in \mathcal{F}} \int_F (\theta - 1) a(\boldsymbol{x}, \boldsymbol{y}) [\![\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\boldsymbol{x}, \boldsymbol{y})]\!] \cdot \{\!\{\nabla \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\boldsymbol{x}, \boldsymbol{y})\}\!\} \mathrm{d}\sigma}_{=: \Psi_h(\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h)} .$$

For the second equality we separated $\boldsymbol{m} = \boldsymbol{0}$ for face integrals from the rest.

Interestingly, if $\theta = 1$, then the highest-order terms ($\boldsymbol{m} = \boldsymbol{0}$) with respect to the faces cancel, and $\Psi_h(\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h) = 0$ for all $\eta(\boldsymbol{y}) > 0$. In any case, Ψ_h can be bounded by first using Hölder's inequality, second absorbing a into the constants (twice), third using Lemma 5.2,

fourth incorporating $a_{\min}$ into the volume norms, and finally using Young's inequality. To wit:

$$
\begin{aligned}
|\Psi_h(\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h)| &\leq \sum_{F \in \mathcal{F}} \sqrt{\frac{a_{\min}(\boldsymbol{y}) h_F}{2 a_{\max}(\boldsymbol{y}) C_{\mathrm{tr}}^2 N_\partial}} \| \sqrt{a(\cdot, \boldsymbol{y})} \{\!\!\{ \nabla \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\cdot, \boldsymbol{y}) \}\!\!\} \|_{L^2(F)} \\
&\qquad \times |\theta - 1| \sqrt{\frac{2 a_{\max}(\boldsymbol{y}) C_{\mathrm{tr}}^2 N_\partial}{a_{\min}(\boldsymbol{y}) h_F}} \| [\![\sqrt{a(\cdot, \boldsymbol{y})} \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\cdot, \boldsymbol{y})]\!] \|_{L^2(F)} \\
&\leq \sum_{F \in \mathcal{F}} \sqrt{\frac{a_{\min}(\boldsymbol{y}) h_F}{2 C_{\mathrm{tr}}^2 N_\partial}} \| \{\!\!\{ \nabla \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\cdot, \boldsymbol{y}) \}\!\!\} \|_{L^2(F)} \\
&\qquad \times \sqrt{\frac{2 a_{\max}^2(\boldsymbol{y}) C_{\mathrm{tr}}^2 N_\partial (\theta - 1)^2}{a_{\min}(\boldsymbol{y}) h_F}} \| [\![\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\cdot, \boldsymbol{y})]\!] \|_{L^2(F)} \\
&\leq \left[\sum_{F \in \mathcal{F}} \frac{a_{\min}(\boldsymbol{y})}{2 N_\partial} \left(\sum_{T \in \mathcal{T}_h : \partial T \supset F} \| \nabla \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\cdot, \boldsymbol{y}) \|_{L^2(T)} \right)^2 \right]^{1/2} \\
&\qquad \times \left[\sum_{F \in \mathcal{F}} \frac{2 a_{\max}^2(\boldsymbol{y}) C_{\mathrm{tr}}^2 N_\partial (\theta - 1)^2}{a_{\min}(\boldsymbol{y}) h_F} \| [\![\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\cdot, \boldsymbol{y})]\!] \|_{L^2(F)}^2 \right]^{1/2} \\
&\leq \left[\sum_{T \in \mathcal{T}_h} \| \sqrt{a(\cdot, \boldsymbol{y})} \nabla \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\cdot, \boldsymbol{y}) \|_{L^2(T)}^2 \right]^{1/2} \\
&\qquad \times \sqrt{\frac{2 a_{\max}^2(\boldsymbol{y}) C_{\mathrm{tr}}^2 N_\partial (\theta - 1)^2}{a_{\min}(\boldsymbol{y})}} \left[\sum_{F \in \mathcal{F}} \frac{1}{h_F} \| [\![\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\cdot, \boldsymbol{y})]\!] \|_{L^2(F)}^2 \right]^{1/2} \\
&\leq \delta \sum_{T \in \mathcal{T}_h} \| \sqrt{a(\cdot, \boldsymbol{y})} \nabla \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\cdot, \boldsymbol{y}) \|_{L^2(T)}^2 \\
&\qquad + \frac{a_{\max}^2(\boldsymbol{y}) C_{\mathrm{tr}}^2 N_\partial (\theta - 1)^2}{2 \delta a_{\min}(\boldsymbol{y})} \sum_{F \in \mathcal{F}} \frac{1}{h_F} \| [\![\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\cdot, \boldsymbol{y})]\!] \|_{L^2(F)}^2,
\end{aligned}
$$

and $\delta \in (0, 1)$ can be chosen arbitrarily. Notably, the fourth inequality uses that

$$
\begin{aligned}
\sum_{F \in \mathcal{F}} \frac{a_{\min}(\boldsymbol{y})}{2 N_\partial} &\left(\sum_{T \in \mathcal{T}_h : \partial T \supset F} \| \nabla \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\cdot, \boldsymbol{y}) \|_{L^2(T)} \right)^2 \\
&\leq \frac{1}{2 N_\partial} \sum_{F \in \mathcal{F}} \left(\sum_{T \in \mathcal{T}_h : \partial T \supset F} \| \sqrt{a(\cdot, \boldsymbol{y})} \nabla \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\cdot, \boldsymbol{y}) \|_{L^2(T)} \right)^2,
\end{aligned}
$$

and the fact that each element has at most N_∂ faces. Hence, each element appears at most N_∂ times on the right-hand side. Moreover, we have that

$$
\begin{aligned}
\left(\sum_{T \in \mathcal{T}_h : \partial T \supset F} \| \sqrt{a(\cdot, \boldsymbol{y})} \nabla \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\cdot, \boldsymbol{y}) \|_{L^2(T)} \right)^2 \\
\leq 2 \sum_{T \in \mathcal{T}_h : \partial T \supset F} \| \sqrt{a(\cdot, \boldsymbol{y})} \nabla \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\cdot, \boldsymbol{y}) \|_{L^2(T)}^2
\end{aligned}
$$

as each face has at most two neighbors. Thus, if (5.4) holds, we can choose $\delta = \frac{2}{\tau+1}$, and $\Psi_h(\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h) < \|\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\cdot, \boldsymbol{y})\|_{V_h}^2$ can be incorporated into the norm on the left-hand side producing a constant that is independent of $\boldsymbol{y}$:

$$
\|\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\cdot, \boldsymbol{y})\|_{V_h}^2
$$

$$
\overset{\text{\textcircled{A}}}{\lesssim} - \sum_{T \in \mathcal{T}_h} \int_T \sum_{\boldsymbol{0} \neq \boldsymbol{m} \leq \boldsymbol{\nu}} \binom{\boldsymbol{\nu}}{\boldsymbol{m}} \partial_{\boldsymbol{y}}^{\boldsymbol{m}} a(\boldsymbol{x}, \boldsymbol{y}) \nabla \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}-\boldsymbol{m}} u_h(\boldsymbol{x}, \boldsymbol{y}) \cdot \nabla \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\boldsymbol{x}, \boldsymbol{y}) \mathrm{d}\boldsymbol{x}
$$

$$
- \sum_{F \in \mathcal{F}} \int_F \Bigg[\theta \sum_{\boldsymbol{0} \neq \boldsymbol{m} \leq \boldsymbol{\nu}} \binom{\boldsymbol{\nu}}{\boldsymbol{m}} \partial_{\boldsymbol{y}}^{\boldsymbol{m}} a(\boldsymbol{x}, \boldsymbol{y}) \llbracket \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}-\boldsymbol{m}} u_h(\boldsymbol{x}, \boldsymbol{y}) \rrbracket \cdot \{\!\!\{ \nabla \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\boldsymbol{x}, \boldsymbol{y}) \}\!\!\}
$$

$$
- \sum_{\boldsymbol{0} \neq \boldsymbol{m} \leq \boldsymbol{\nu}} \binom{\boldsymbol{\nu}}{\boldsymbol{m}} \partial_{\boldsymbol{y}}^{\boldsymbol{m}} a(\boldsymbol{x}, \boldsymbol{y}) \{\!\!\{ \nabla \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}-\boldsymbol{m}} u_h(\boldsymbol{x}, \boldsymbol{y}) \}\!\!\} \cdot \llbracket \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\boldsymbol{x}, \boldsymbol{y}) \rrbracket
$$

$$
+ h_F^{-1} \sum_{\boldsymbol{0} \neq \boldsymbol{m} \leq \boldsymbol{\nu}} \binom{\boldsymbol{\nu}}{\boldsymbol{m}} \partial_{\boldsymbol{y}}^{\boldsymbol{m}} \eta(\boldsymbol{y}) \llbracket \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}-\boldsymbol{m}} u_h(\boldsymbol{x}, \boldsymbol{y}) \rrbracket \cdot \llbracket \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\boldsymbol{x}, \boldsymbol{y}) \rrbracket \Bigg] \mathrm{d}\sigma
$$

$$
= - \sum_{\boldsymbol{0} \neq \boldsymbol{m} \leq \boldsymbol{\nu}} \binom{\boldsymbol{\nu}}{\boldsymbol{m}} \Bigg[\sum_{T \in \mathcal{T}_h} \int_T \frac{\partial_{\boldsymbol{y}}^{\boldsymbol{m}} a(\boldsymbol{x}, \boldsymbol{y})}{a(\boldsymbol{x}, \boldsymbol{y})} a(\boldsymbol{x}, \boldsymbol{y}) \nabla \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}-\boldsymbol{m}} u_h(\boldsymbol{x}, \boldsymbol{y}) \cdot \nabla \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\boldsymbol{x}, \boldsymbol{y}) \mathrm{d}\boldsymbol{x}
$$

$$
+ \sum_{F \in \mathcal{F}} \int_F \frac{\partial_{\boldsymbol{y}}^{\boldsymbol{m}} a(\boldsymbol{x}, \boldsymbol{y})}{a(\boldsymbol{x}, \boldsymbol{y})} \Bigg[\theta \llbracket \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}-\boldsymbol{m}} u_h(\boldsymbol{x}, \boldsymbol{y}) \rrbracket \cdot \{\!\!\{ a(\boldsymbol{x}, \boldsymbol{y}) \nabla \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\boldsymbol{x}, \boldsymbol{y}) \}\!\!\}
$$

$$
- \{\!\!\{ a(\boldsymbol{x}, \boldsymbol{y}) \nabla \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}-\boldsymbol{m}} u_h(\boldsymbol{x}, \boldsymbol{y}) \}\!\!\} \cdot \llbracket \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\boldsymbol{x}, \boldsymbol{y}) \rrbracket \Bigg] \mathrm{d}\sigma
$$

$$
+ \sum_{F \in \mathcal{F}} \int_F \frac{\partial_{\boldsymbol{y}}^{\boldsymbol{m}} \eta(\boldsymbol{y})}{\eta(\boldsymbol{y})} \frac{\eta(\boldsymbol{y})}{h_F} \llbracket \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}-\boldsymbol{m}} u_h(\boldsymbol{x}, \boldsymbol{y}) \rrbracket \cdot \llbracket \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\boldsymbol{x}, \boldsymbol{y}) \rrbracket \mathrm{d}\sigma \Bigg]
$$

$$
= (\bigstar),
$$

where '$\lesssim$' is to be interpreted as smaller than or equal to up to a positive, multiplicative constant, which does not depend on h, $\boldsymbol{\nu}$, or $\boldsymbol{y}$. Importantly, for the non-symmetric IP method ($\theta = 1$), the symbol '$\lesssim$' at $\text{\textcircled{A}}$ can be replaced by '$=$' independently of τ. Otherwise $\tau \searrow 1$ yields an unstable method.

Our next task consists in estimating $|(\bigstar)|$. Hölder's inequality yields

$$
|(\bigstar)| \leq \sum_{\boldsymbol{0} \neq \boldsymbol{m} \leq \boldsymbol{\nu}} \binom{\boldsymbol{\nu}}{\boldsymbol{m}} \left\| \frac{\partial_{\boldsymbol{y}}^{\boldsymbol{m}} a(\cdot, \boldsymbol{y})}{a(\cdot, \boldsymbol{y})} \right\|_{L^\infty(D)}
$$

$$
(5.8a) \qquad \times \Bigg[\sum_{T \in \mathcal{T}_h} \|\sqrt{a(\cdot, \boldsymbol{y})} \nabla \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}-\boldsymbol{m}} u_h(\cdot, \boldsymbol{y})\|_{L^2(T)} \|\sqrt{a(\cdot, \boldsymbol{y})} \nabla \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\cdot, \boldsymbol{y})\|_{L^2(D)}
$$

$$
(5.8b) \qquad + \sum_{F \in \mathcal{F}} \int_F \left| \theta \llbracket \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}-\boldsymbol{m}} u_h(\boldsymbol{x}, \boldsymbol{y}) \rrbracket \cdot \{\!\!\{ a(\boldsymbol{x}, \boldsymbol{y}) \nabla \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\boldsymbol{x}, \boldsymbol{y}) \}\!\!\} \right| \mathrm{d}\sigma
$$

$$
(5.8c) \qquad + \sum_{F \in \mathcal{F}} \int_F \left| \{\!\!\{ a(\boldsymbol{x}, \boldsymbol{y}) \nabla \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}-\boldsymbol{m}} u_h(\boldsymbol{x}, \boldsymbol{y}) \}\!\!\} \cdot \llbracket \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\boldsymbol{x}, \boldsymbol{y}) \rrbracket \right| \mathrm{d}\sigma \Bigg]
$$

$$
(5.8d) \qquad + \sum_{\boldsymbol{0} \neq \boldsymbol{m} \leq \boldsymbol{\nu}} \binom{\boldsymbol{\nu}}{\boldsymbol{m}} \left| \frac{\partial_{\boldsymbol{y}}^{\boldsymbol{m}} \eta(\boldsymbol{y})}{\eta(\boldsymbol{y})} \right| \sum_{F \in \mathcal{F}} \int_F \frac{\eta(\boldsymbol{y})}{h_F} \llbracket \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}-\boldsymbol{m}} u_h(\boldsymbol{x}, \boldsymbol{y}) \rrbracket \cdot \llbracket \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\boldsymbol{x}, \boldsymbol{y}) \rrbracket \mathrm{d}\sigma.
$$

The former three lines can be bounded separately using $|\theta| \leq 1$:

$$|(5.8\text{a})| \leq \left[\sum_{T \in \mathcal{T}_h} \|\sqrt{a(\cdot, \boldsymbol{y})}\nabla \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\cdot, \boldsymbol{y})\|_{L^2(D)}^2 \right]^{\frac{1}{2}}$$
$$\times \left[\sum_{T \in \mathcal{T}_h} \|\sqrt{a(\cdot, \boldsymbol{y})}\nabla \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}-\boldsymbol{m}} u_h(\cdot, \boldsymbol{y})\|_{L^2(T)}^2 \right]^{\frac{1}{2}},$$

$$|(5.8\text{b})| \leq \left[\sum_{F \in \mathcal{F}} \frac{\eta(\boldsymbol{y})}{h_F} \|[\![\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}-\boldsymbol{m}} u_h(\cdot, \boldsymbol{y})]\!] \|_{L^2(F)}^2 \right]^{\frac{1}{2}}$$
$$\times \left[\sum_{F \in \mathcal{F}} \frac{h_F}{\eta(\boldsymbol{y})} \|\{\!\{ a(\cdot, \boldsymbol{y})\nabla \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\cdot, \boldsymbol{y}) \}\!\} \|_{L^2(F)}^2 \right]^{\frac{1}{2}},$$

$$|(5.8\text{c})| \leq \left[\sum_{F \in \mathcal{F}} \frac{\eta(\boldsymbol{y})}{h_F} \|[\![\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\cdot, \boldsymbol{y})]\!] \|_{L^2(F)}^2 \right]^{\frac{1}{2}}$$
$$\times \left[\sum_{F \in \mathcal{F}} \frac{h_F}{\eta(\boldsymbol{y})} \|\{\!\{ a(\cdot, \boldsymbol{y})\nabla \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}-\boldsymbol{m}} u_h(\cdot, \boldsymbol{y}) \}\!\} \|_{L^2(F)}^2 \right]^{\frac{1}{2}},$$

and the last line can be estimated via

$$|(5.8\text{d})| \leq \left[\sum_{F \in \mathcal{F}} \frac{\eta(\boldsymbol{y})}{h_F} \|[\![\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}-\boldsymbol{m}} u_h(\cdot, \boldsymbol{y})]\!] \|_{L^2(F)}^2 \right]^{\frac{1}{2}} \left[\sum_{F \in \mathcal{F}} \frac{\eta(\boldsymbol{y})}{h_F} \|[\![\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\cdot, \boldsymbol{y})]\!] \|_{L^2(F)}^2 \right]^{\frac{1}{2}}.$$

As a consequence, we obtain by the Cauchy–Schwarz inequality that

$$|(\bigstar)| \leq \sum_{\boldsymbol{0} \neq \boldsymbol{m} \leq \boldsymbol{\nu}} \binom{\boldsymbol{\nu}}{\boldsymbol{m}} \left[\left\| \frac{\partial_{\boldsymbol{y}}^{\boldsymbol{m}} a(\cdot, \boldsymbol{y})}{a(\cdot, \boldsymbol{y})} \right\|_{L^\infty(D)} + \left| \frac{\partial_{\boldsymbol{y}}^{\boldsymbol{m}} \eta(\boldsymbol{y})}{\eta(\boldsymbol{y})} \right| \right]$$
$$\times \|\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\cdot, \boldsymbol{y})\|_{V_h^*} \|\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}-\boldsymbol{m}} u_h(\cdot, \boldsymbol{y})\|_{V_h^*}.$$

Thus, considering Lemma 5.4 and division by $\|\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\cdot, \boldsymbol{y})\|_{V_h}$ yield the following:

THEOREM 5.6. *If* (5.4) *and* (5.6) *hold, then there are constants* $C_{\text{DG}}', C_{\text{DG}} > 0$ *(independent of $\boldsymbol{y}$ and h) such that*

$$\|\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\cdot, \boldsymbol{y})\|_{V_h}$$
$$(5.9) \qquad \leq \underbrace{C_{\text{equiv}}^2 C_{\text{DG}}'}_{=C_{\text{DG}}} \sum_{\boldsymbol{0} \neq \boldsymbol{m} \leq \boldsymbol{\nu}} \binom{\boldsymbol{\nu}}{\boldsymbol{m}} \left[\left\| \frac{\partial_{\boldsymbol{y}}^{\boldsymbol{m}} a(\cdot, \boldsymbol{y})}{a(\cdot, \boldsymbol{y})} \right\|_{L^\infty(D)} + \left| \frac{\partial_{\boldsymbol{y}}^{\boldsymbol{m}} \eta(\boldsymbol{y})}{\eta(\boldsymbol{y})} \right| \right] \|\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}-\boldsymbol{m}} u_h(\cdot, \boldsymbol{y})\|_{V_h}.$$

For NIPDG ($\theta = 1$), one may choose $C_{\text{DG}}' = 1$.

5.3. A brief note on error estimates. The error estimate does not significantly deviate from standard DG error estimates. Thus, we keep it short and refer to the already mentioned references [11, 33] for the details. We start by introducing the broken H^2-space

$$H^2(\mathcal{T}_h) = \{v \in L^2(D) \colon v|_T \in H^2(T)\}$$

and observe the following:

COROLLARY 5.7. *Under the assumptions of Lemma 5.4, there is a constant M (independent of $\boldsymbol{y}$ and h) such that for all $w_h \in H^2(\mathcal{T}_h)$ and all $v_h \in V_h$, we have*

$$B_h(w_h, v_h) \le M\|w_h\|_{V_h^*}\|v_h\|_{V_h}.$$

This follows directly from the Cauchy–Schwarz inequality and Lemma 5.4. Thus, Strang's lemma [33, Rem. 4.9] implies that for any $w_h \in V_h$

$$\|u - u_h\|_{V_h^*} \le \left(1 + \frac{MC_{\text{equiv}}}{\alpha}\right)\|u - w_h\|_{V_h^*}.$$

We choose w_h to be the local L^2-projection Πu of u and will need to estimate the respective terms of $\|u - \Pi u\|_{V_h^*}$. Exploiting the standard scaling arguments of DG analysis, we arrive at the following:

THEOREM 5.8. *Let $u(\cdot, \boldsymbol{y}) \in H^{k+1}(D)$ and the assumptions of Lemmas 5.3 and 5.4 hold. Then we have*

$$\frac{1}{C_{\text{Poin}}^{V_h}(\boldsymbol{y})}\|u(\cdot, \boldsymbol{y}) - u_h(\cdot, \boldsymbol{y})\|_{L^2(D)} \le \|u(\cdot, \boldsymbol{y}) - u_h(\cdot, \boldsymbol{y})\|_{V_h^*}$$

$$\lesssim \max\left\{\sqrt{\eta(\boldsymbol{y})}, a_{\max}(\boldsymbol{y}), \sqrt{\tfrac{1}{\eta(\boldsymbol{y})}}\right\} h^k |u(\cdot, \boldsymbol{y})|_{H^{k+1}(D)}.$$

Notably, the convergence order for SIPG can be increased by one using a duality argument as in [11, Sect. 4.2.4] if we have elliptic regularity. Importantly, the FE error in the QMC setting will be integrated with respect to $\boldsymbol{y}$. That is why we additionally assume that $|u(\cdot, \boldsymbol{y})|_{H^{k+1}(D)}$ is integrable with respect to $\boldsymbol{y}$. By standard elliptic regularity theory, this holds for example when D is a bounded, convex domain, the diffusion coefficient a is Lipschitz continuous, and the source term $f \in L^2(D)$. For $k = 1$, the integrability of the seminorm is shown in [57].

6. Parametric regularity analysis.

6.1. Uniform and affine setting. In this setting, the regularity analysis is completely analogous to the conforming FE setting. Specifically, the choice of $\eta(\boldsymbol{y})$ is not as critical as in the lognormal case to be discussed later on: $\eta(\boldsymbol{y})$ can be chosen independently of $\boldsymbol{y}$ and large enough to satisfy Lemmas 5.3 and 5.4.

LEMMA 6.1 (Regularity bound in the uniform and affine setting). *Let assumptions* (U1)–(U5) *hold with η sufficiently large (independently of $\boldsymbol{y}$). Let $\boldsymbol{b} = (b_j)_{j \ge 1}$ with $b_j := \frac{C_{\text{DG}}\|\psi_j\|_{L^\infty(D)}}{\alpha a_{\min}}$. Then for all $\boldsymbol{\nu} \in \mathcal{F}$ and $\boldsymbol{y} \in U$, it holds that*

$$\|\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\cdot, \boldsymbol{y})\|_{V_h} \le |\boldsymbol{\nu}|!\boldsymbol{b}^{\boldsymbol{\nu}} \frac{C_{\text{Poin}}^{V_h}}{\alpha}\|f\|_{L^2(D)}.$$

Proof. This is an immediate consequence of [36, Lem. 9.1] (in fact, it already appears in [9]), but since the proof is short, we present it for completeness. The proof is carried out by induction with respect to the order of the multi-indices $\boldsymbol{\nu} \in \mathcal{F}$. In the affine setting, we make use of the fact that

$$\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} a(\boldsymbol{x}, \boldsymbol{y}) = \begin{cases} a(\boldsymbol{x}, \boldsymbol{y}) & \text{if } \boldsymbol{\nu} = \boldsymbol{0}, \\ \psi_j(\boldsymbol{x}) & \text{if } \boldsymbol{\nu} = \boldsymbol{e}_j, \ j \geq 1, \\ 0 & \text{otherwise}, \end{cases}$$

where $\boldsymbol{e}_j$ denotes the jth Euclidean standard unit vector.

The base of the induction is resolved by observing that the claim is equivalent to the *a priori* bound when $\boldsymbol{\nu} = \boldsymbol{0}$; see Lemma 5.3.

Let $\boldsymbol{\nu} \in \mathcal{F}$, and suppose that the claim has already been proved for all multi-indices with order less than $|\boldsymbol{\nu}|$. Then it is a consequence of Theorem 5.6 and the fact that $\partial^{\boldsymbol{m}} \eta(\boldsymbol{y}) = 0$, for $|\boldsymbol{m}| \geq 1$, that

$$\|\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\cdot, \boldsymbol{y})\|_{V_h} \leq C_{\mathrm{DG}} \sum_{\boldsymbol{0} \neq \boldsymbol{m} \leq \boldsymbol{\nu}} \binom{\boldsymbol{\nu}}{\boldsymbol{m}} \left\| \frac{\partial_{\boldsymbol{y}}^{\boldsymbol{m}} a(\cdot, \boldsymbol{y})}{a(\cdot, \boldsymbol{y})} \right\|_{L^\infty(D)} \|\partial_{\boldsymbol{y}}^{\boldsymbol{\nu} - \boldsymbol{m}} u_h(\cdot, \boldsymbol{y})\|_{V_h}$$

$$\leq \sum_{j \in \mathrm{supp}(\boldsymbol{\nu})} \nu_j b_j |\boldsymbol{\nu} - \boldsymbol{e}_j|! \boldsymbol{b}^{\boldsymbol{\nu} - \boldsymbol{e}_j} \frac{C_{\mathrm{Poin}}^{V_h}}{\alpha} \|f\|_{L^2(D)}$$

$$= \boldsymbol{b}^{\boldsymbol{\nu}} (|\boldsymbol{\nu}| - 1)! \left(\sum_{j \in \mathrm{supp}(\boldsymbol{\nu})} \nu_j \right) \frac{C_{\mathrm{Poin}}^{V_h}}{\alpha} \|f\|_{L^2(D)}$$

$$= \boldsymbol{b}^{\boldsymbol{\nu}} |\boldsymbol{\nu}|! \frac{C_{\mathrm{Poin}}^{V_h}}{\alpha} \|f\|_{L^2(D)},$$

as desired. $\square$

6.2. Derivation of the QMC error in the uniform and affine setting. In what follows, we define the dimensionally truncated solution by setting

$$u_{s,h}(\cdot, \boldsymbol{y}) := u_h(\cdot, (y_1, \ldots, y_s, 0, 0, \ldots)), \qquad \boldsymbol{y} \in [-\tfrac{1}{2}, \tfrac{1}{2}]^s.$$

THEOREM 6.2. *Let assumptions* (U1)–(U5) *hold, and suppose that* $n = 2^m$, $m \in \mathbb{N}$. *There exists a generating vector constructed by the CBC algorithm such that*

$$\sqrt{\mathbb{E}_{\boldsymbol{\Delta}} \left\| \int_{[-1/2, 1/2]^s} u_{s,h}(\cdot, \boldsymbol{y}) \, \mathrm{d}\boldsymbol{y} - Q_{\mathrm{ran}}(u_{s,h}) \right\|_{L^2(D)}^2}$$

$$= \begin{cases} \mathcal{O}(n^{-1/p+1/2}) & \text{if } p \in (2/3, 1), \\ \mathcal{O}(n^{-1+\varepsilon}) & \text{for arbitrary } \varepsilon \in (0, 1/2) & \text{if } p \in (0, 2/3], \end{cases}$$

where the implied coefficient is independent of s *in both cases when the weights* $(\gamma_{\mathfrak{u}})_{\mathfrak{u} \subseteq \{1:s\}}$ *in Lemma* 3.1 *are chosen to be*

$$\gamma_{\mathfrak{u}} = \left(|\mathfrak{u}|! \prod_{j \in \mathfrak{u}} \frac{b_j}{\sqrt{\varrho(\lambda)}} \right)^{2/(1+\lambda)} \qquad \text{for } \mathfrak{u} \subseteq \{1 : s\}$$

and

$$\lambda = \begin{cases} \frac{p}{2-p} & \text{if } p \in (2/3, 1), \\ \frac{1}{2-2\varepsilon} & \text{for arbitrary } \varepsilon \in (0, 1/2) & \text{if } p \in (0, 2/3]. \end{cases}$$

Proof. In the uniform and affine setting, $C_{\text{Poin}}^{V_h}$ is independent of $\boldsymbol{y}$, and so we can proceed with the usual derivation of the QMC weights. Let the assumptions of Lemma 3.1 be in effect. Furthermore, suppose that assumptions (U1)–(U3) and (U5) hold. Then for almost all $\boldsymbol{x} \in D$, we have from Theorem 3.1 that

$$
\mathbb{E}_{\boldsymbol{\Delta}} \left| \int_{[-1/2,1/2]^s} u_{s,h}(\boldsymbol{x},\boldsymbol{y}) \, \mathrm{d}\boldsymbol{y} - Q_{\text{ran}}(u_{s,h}(\boldsymbol{x},\cdot)) \right|^2
$$
$$
\leq \left(\frac{2}{n} \right)^{1/\lambda} \frac{1}{R} \left(\sum_{\emptyset \neq \mathfrak{u} \subseteq \{1:s\}} \gamma_{\mathfrak{u}}^{\lambda} \varrho(\lambda)^{|\mathfrak{u}|} \right)^{1/\lambda}
$$
$$
\times \sum_{\mathfrak{u} \subseteq \{1:s\}} \frac{1}{\gamma_{\mathfrak{u}}} \int_{[-1/2,1/2]^{|\mathfrak{u}|}} \left(\int_{[-1/2,1/2]^{s-|\mathfrak{u}|}} \frac{\partial^{|\mathfrak{u}|}}{\partial \boldsymbol{y}_{\mathfrak{u}}} u_{s,h}(\boldsymbol{x},\boldsymbol{y}) \, \mathrm{d}\boldsymbol{y}_{-\mathfrak{u}} \right)^2 \mathrm{d}\boldsymbol{y}_{\mathfrak{u}}
$$
$$
\leq \left(\frac{2}{n} \right)^{1/\lambda} \frac{1}{R} \left(\sum_{\emptyset \neq \mathfrak{u} \subseteq \{1:s\}} \gamma_{\mathfrak{u}}^{\lambda} \varrho(\lambda)^{|\mathfrak{u}|} \right)^{1/\lambda} \sum_{\mathfrak{u} \subseteq \{1:s\}} \frac{1}{\gamma_{\mathfrak{u}}} \int_{[-1/2,1/2]^s} \left| \frac{\partial^{|\mathfrak{u}|}}{\partial \boldsymbol{y}_{\mathfrak{u}}} u_{s,h}(\boldsymbol{x},\boldsymbol{y}) \right|^2 \mathrm{d}\boldsymbol{y}.
$$

Integrating over $\boldsymbol{x} \in D$ on both sides and using Fubini's theorem, we obtain

$$
\mathbb{E}_{\boldsymbol{\Delta}} \left\| \int_{[-1/2,1/2]^s} u_{s,h}(\cdot,\boldsymbol{y}) \, \mathrm{d}\boldsymbol{y} - Q_{\text{ran}}(u_{s,h}) \right\|_{L^2(D)}^2
$$
$$
\leq \left(\frac{2}{n} \right)^{1/\lambda} \frac{1}{R} \left(\sum_{\emptyset \neq \mathfrak{u} \subseteq \{1:s\}} \gamma_{\mathfrak{u}}^{\lambda} \varrho(\lambda)^{|\mathfrak{u}|} \right)^{1/\lambda}
$$
$$
\times \sum_{\mathfrak{u} \subseteq \{1:s\}} \frac{1}{\gamma_{\mathfrak{u}}} \int_{[-1/2,1/2]^s} \left\| \frac{\partial^{|\mathfrak{u}|}}{\partial \boldsymbol{y}_{\mathfrak{u}}} u_{s,h}(\cdot,\boldsymbol{y}) \right\|_{L^2(D)}^2 \mathrm{d}\boldsymbol{y}
$$
$$
\leq \left(\frac{2}{n} \right)^{1/\lambda} \frac{(C_{\text{Poin}}^{V_h})^2}{R} \left(\sum_{\emptyset \neq \mathfrak{u} \subseteq \{1:s\}} \gamma_{\mathfrak{u}}^{\lambda} \varrho(\lambda)^{|\mathfrak{u}|} \right)^{1/\lambda}
$$
$$
\times \sum_{\mathfrak{u} \subseteq \{1:s\}} \frac{1}{\gamma_{\mathfrak{u}}} \int_{[-1/2,1/2]^s} \left\| \frac{\partial^{|\mathfrak{u}|}}{\partial \boldsymbol{y}_{\mathfrak{u}}} u_{s,h}(\cdot,\boldsymbol{y}) \right\|_{V_h}^2 \mathrm{d}\boldsymbol{y},
$$

where we used the discrete Poincaré inequality (5.3). (Recall that in the uniform and affine setting, the Poincaré constant can be bounded independently of $\boldsymbol{y}$.) We can now apply Lemma 6.1 to obtain

$$
\mathbb{E}_{\boldsymbol{\Delta}} \left\| \int_{[-1/2,1/2]^s} u_{s,h}(\cdot,\boldsymbol{y}) \, \mathrm{d}\boldsymbol{y} - Q_{\text{ran}}(u_{s,h}) \right\|_{L^2(D)}^2
$$
$$
\leq \left(\frac{2}{n} \right)^{1/\lambda} \frac{(C_{\text{Poin}}^{V_h})^4 \|f\|_{L^2(D)}^2}{\alpha^2 R} \left(\sum_{\emptyset \neq \mathfrak{u} \subseteq \{1:s\}} \gamma_{\mathfrak{u}}^{\lambda} \varrho(\lambda)^{|\mathfrak{u}|} \right)^{1/\lambda} \sum_{\mathfrak{u} \subseteq \{1:s\}} \frac{1}{\gamma_{\mathfrak{u}}} (|\mathfrak{u}|!)^2 \prod_{j \in \mathfrak{u}} b_j^2.
$$

It is easy to verify that the upper bound is minimized by choosing (cf., e.g., [39, Lem. 6.2])

$$
(6.1) \qquad \gamma_{\mathfrak{u}} = \left(|\mathfrak{u}|! \prod_{j \in \mathfrak{u}} \frac{b_j}{\sqrt{\varrho(\lambda)}} \right)^{2/(1+\lambda)} \qquad \text{for } \mathfrak{u} \subseteq \{1:s\}.
$$

By plugging in these weights into the error bound, we obtain

$$\sqrt{\mathbb{E}_{\boldsymbol{\Delta}}\bigg\|\int_{[-1/2,1/2]^s} u_{s,h}(\cdot,\boldsymbol{y})\,\mathrm{d}\boldsymbol{y} - Q_{\mathrm{ran}}(u_{s,h})\bigg\|_{L^2(D)}^2}$$

$$\leq \left(\frac{2}{n}\right)^{1/(2\lambda)} \frac{(C_{\mathrm{Poin}}^{V_h})^2\|f\|_{L^2(D)}}{\alpha\sqrt{R}} \underbrace{\left(\sum_{\mathfrak{u}\subseteq\{1:s\}} \left(|\mathfrak{u}|!\prod_{j\in\mathfrak{u}} \frac{b_j}{\sqrt{\varrho(\lambda)}}\right)^{\frac{2\lambda}{1+\lambda}} \varrho(\lambda)^{|\mathfrak{u}|}\right)^{\frac{\lambda+1}{2\lambda}}}_{=:C(s,\lambda)}.$$

The term $C(s,\lambda)$ can be bounded independently of s:

$$C(s,\lambda) = \sum_{\ell=0}^{s} (\ell!)^{2\lambda/(1+\lambda)} \varrho(\lambda)^{\frac{1}{1+\lambda}\ell} \sum_{\substack{\mathfrak{u}\subseteq\{1:s\}\\|\mathfrak{u}|=\ell}} \prod_{j\in\mathfrak{u}} b_j^{2\lambda/(1+\lambda)}$$

$$\leq \sum_{\ell=0}^{\infty} (\ell!)^{2\lambda/(1+\lambda)} \varrho(\lambda)^{\frac{1}{1+\lambda}\ell} \frac{1}{\ell!} \left(\sum_{j=1}^{\infty} b_j^{2\lambda/(1+\lambda)}\right)^{\ell},$$

where the series can be shown to converge by the d'Alembert ratio test provided that

$$\frac{2\lambda}{1+\lambda} \geq p \qquad \text{and} \qquad \frac{1}{2} < \lambda \leq 1.$$

We can ensure that both conditions are satisfied by choosing

$$(6.2) \qquad \lambda = \begin{cases} \frac{p}{2-p} & \text{if } p \in (2/3,1), \\ \frac{1}{2-2\varepsilon} & \text{for arbitrary } \varepsilon \in (0,1/2) \quad \text{if } p \in (0,2/3]. \end{cases}$$

We conclude that by choosing the weights (6.1) and the parameter (6.2), we obtain the QMC convergence rate

$$\sqrt{\mathbb{E}_{\boldsymbol{\Delta}}\bigg\|\int_{[-1/2,1/2]^s} u_{s,h}(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\boldsymbol{y} - Q_{\mathrm{ran}}(u_{s,h}(\boldsymbol{x},\cdot))\bigg\|_{L^2(D)}^2}$$

$$= \begin{cases} \mathcal{O}(n^{-1/p+1/2}) & \text{if } p \in (2/3,1), \\ \mathcal{O}(n^{-1+\varepsilon}) & \text{for arbitrary } \varepsilon \in (0,1/2) \quad \text{if } p \in (0,2/3], \end{cases}$$

where the implied coefficient is independent of s in both cases. This concludes the proof. $\square$

REMARK 6.3. Let $\mathcal{G}\colon V_h \to \mathbb{R}$ be a bounded linear functional. The same proof technique can be applied *mutatis mutandis* to show that there exists a generating vector constructed by the CBC algorithm using the weights specified in Theorem 6.2 such that

$$\sqrt{\mathbb{E}_{\boldsymbol{\Delta}}\bigg|\int_{[-1/2,1/2]^s} \mathcal{G}(u_{s,h}(\cdot,\boldsymbol{y}))\,\mathrm{d}\boldsymbol{y} - Q_{\mathrm{ran}}(\mathcal{G}(u_{s,h}))\bigg|^2} = \mathcal{O}(n^{\max\{-1/p+1/2,-1+\varepsilon\}}),$$

where the implied coefficient is independent of the dimension s with arbitrary $\varepsilon \in (0,1/2)$.

We note that the QMC convergence rate is at best essentially linear, and it is always better than $\mathcal{O}(n^{-1/2})$.

6.3. Lognormal setting. In the lognormal setting, we set

$$(6.3) \qquad \eta(\boldsymbol{y}) := \frac{(\max_{\boldsymbol{x} \in \overline{D}} a_0(\boldsymbol{x}))^2}{\min_{\boldsymbol{x} \in \overline{D}} a_0(\boldsymbol{x})} \exp\bigg(\sum_{j=1}^{\infty} 3\beta_j |y_j| \bigg),$$

where $\beta_j = \|\psi_j\|_{L^\infty}$ is defined as in Section 2. Then we have for all multi-indices $\boldsymbol{m}$ with $|\boldsymbol{m}|_\infty := \sup_{j \geq 1} m_j \leq 1$ that

$$\left\| \frac{\partial_{\boldsymbol{y}}^{\boldsymbol{m}} a(\cdot, \boldsymbol{y})}{a(\cdot, \boldsymbol{y})} \right\|_{L^\infty(D)} \leq \boldsymbol{\beta}^{\boldsymbol{m}} \qquad \text{for all } \boldsymbol{y} \in U_{\boldsymbol{\beta}},$$

$$\left| \frac{\partial_{\boldsymbol{y}}^{\boldsymbol{m}} \eta(\boldsymbol{y})}{\eta(\boldsymbol{y})} \right| \leq 3^{|\boldsymbol{m}|} \boldsymbol{\beta}^{\boldsymbol{m}} \quad \text{for a.e. } \boldsymbol{y} \in U_{\boldsymbol{\beta}}$$

(note that the second inequality holds almost everywhere since the derivatives of (6.3) are discontinuous on a set of measure zero), and the recurrence relation (5.9) takes the form

$$(6.4) \qquad \|\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\cdot, \boldsymbol{y})\|_{V_h} \leq C_{\mathrm{DG}} \sum_{\boldsymbol{0} \neq \boldsymbol{m} \leq \boldsymbol{\nu}} \binom{\boldsymbol{\nu}}{\boldsymbol{m}} (1 + 3^{|\boldsymbol{m}|}) \boldsymbol{\beta}^{\boldsymbol{m}} \|\partial_{\boldsymbol{y}}^{\boldsymbol{\nu} - \boldsymbol{m}} u_h(\cdot, \boldsymbol{y})\|_{V_h},$$

where we used Theorem 5.6 with $C_{\mathrm{DG}} > 0$ which is independent of $\boldsymbol{y}$. In this case, the assumptions of Lemma 5.3 are satisfied since we use NIPG while the assumptions of Lemma 5.4 are satisfied by construction with $\tilde{\tau} = 1$. In the ensuing analysis, we only consider the first-order parametric regularity for simplicity.

LEMMA 6.4. *Let assumptions* (L1)–(L5) *hold. It holds for all $\boldsymbol{\nu} \in \mathcal{F}$ with $0 \neq |\boldsymbol{\nu}|_\infty \leq 1$ and a.e. $\boldsymbol{y} \in U_{\boldsymbol{\beta}}$ that*

$$(6.5) \qquad \|\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\cdot, \boldsymbol{y})\|_{V_h} \leq C_{\mathrm{DG}}^{|\boldsymbol{\nu}|} 4^{|\boldsymbol{\nu}|} \Lambda_{|\boldsymbol{\nu}|} \boldsymbol{\beta}^{\boldsymbol{\nu}} \|u_h(\cdot, \boldsymbol{y})\|_{V_h},$$

where the sequence of ordered Bell numbers $(\Lambda_k)_{k=0}^{\infty}$ *is defined by the recursion*

$$\Lambda_0 := 1 \qquad and \qquad \Lambda_k := \sum_{\ell=1}^{k} \binom{k}{\ell} \Lambda_{k-\ell}, \quad k \geq 1.$$

Proof. We prove the claim by induction with respect to the modulus of $\boldsymbol{\nu}$. The base step is resolved by observing that, for any $\boldsymbol{\nu} = \boldsymbol{e}_k$, it follows from (6.4) that

$$\|\partial_{\boldsymbol{y}}^{\boldsymbol{e}_k} u_h(\cdot, \boldsymbol{y})\|_{V_h} \leq C_{\mathrm{DG}} (1+3) \boldsymbol{\beta}^{\boldsymbol{e}_k} \|u_h(\cdot, \boldsymbol{y})\|_{V_h} = C_{\mathrm{DG}} 4^{|\boldsymbol{e}_k|} \Lambda_{|\boldsymbol{e}_k|} \boldsymbol{\beta}^{\boldsymbol{e}_k} \|u_h(\cdot, \boldsymbol{y})\|_{V_h},$$

as desired.

Next, let $\boldsymbol{\nu} \in \mathcal{F}$ be such that $0 \neq |\boldsymbol{\nu}|_\infty \leq 1$, and suppose that the claim holds for all multi-indices with order $< |\boldsymbol{\nu}|$. Then (6.4) implies that

$$\|\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_h(\cdot, \boldsymbol{y})\|_{V_h}$$

$$\leq C_{\mathrm{DG}} \boldsymbol{\beta}^{\boldsymbol{\nu}} \|u_h(\cdot, \boldsymbol{y})\|_{V_h} \sum_{\boldsymbol{0} \neq \boldsymbol{m} \leq \boldsymbol{\nu}} \binom{\boldsymbol{\nu}}{\boldsymbol{m}} (1 + 3^{|\boldsymbol{m}|}) C_{\mathrm{DG}}^{|\boldsymbol{\nu} - \boldsymbol{m}|} 4^{|\boldsymbol{\nu} - \boldsymbol{m}|} \Lambda_{|\boldsymbol{\nu} - \boldsymbol{m}|}$$

$$\leq C_{\mathrm{DG}}^{|\boldsymbol{\nu}|} 4^{|\boldsymbol{\nu}|} \boldsymbol{\beta}^{\boldsymbol{\nu}} \|u_h(\cdot, \boldsymbol{y})\|_{V_h} \sum_{\boldsymbol{0} \neq \boldsymbol{m} \leq \boldsymbol{\nu}} \binom{\boldsymbol{\nu}}{\boldsymbol{m}} \Lambda_{|\boldsymbol{\nu} - \boldsymbol{m}|},$$

where we used $C_{\mathrm{DG}} \geq 1$ and the inequality $(1 + 3^\ell)4^{n-\ell} \leq 4^n$ for all $1 \leq \ell \leq n$. The claim follows by simplifying the remaining sum using the generalized Vandermonde identity:

$$\sum_{\mathbf{0} \neq \mathbf{m} \leq \boldsymbol{\nu}} \binom{\boldsymbol{\nu}}{\mathbf{m}} \Lambda_{|\boldsymbol{\nu}-\mathbf{m}|} = \sum_{\ell=1}^{|\boldsymbol{\nu}|} \Lambda_{|\boldsymbol{\nu}|-\ell} \sum_{|\mathbf{m}|=\ell} \binom{\boldsymbol{\nu}}{\mathbf{m}} = \sum_{\ell=1}^{|\boldsymbol{\nu}|} \Lambda_{|\boldsymbol{\nu}|-\ell} \binom{|\boldsymbol{\nu}|}{\ell} = \Lambda_{|\boldsymbol{\nu}|}.$$

This concludes the proof. $\quad\square$

COROLLARY 6.5. *Let assumptions* (L1)–(L5) *hold. It holds for all* $\boldsymbol{\nu} \in \mathcal{F}$ *with* $0 \neq |\boldsymbol{\nu}|_\infty \leq 1$ *and a.e.* $\mathbf{y} \in U_{\boldsymbol{\beta}}$ *that*

$$\|\partial_{\mathbf{y}}^{\boldsymbol{\nu}} u_h(\cdot, \mathbf{y})\|_{V_h} \leq C_{\mathrm{DG}}^{|\boldsymbol{\nu}|} 4^{|\boldsymbol{\nu}|} \frac{|\boldsymbol{\nu}|!}{(\log 2)^{|\boldsymbol{\nu}|}} \boldsymbol{\beta}^{\boldsymbol{\nu}} \frac{\sigma}{\alpha a_{\min}(\mathbf{y})^{1/2}} \|f\|_{L^2}.$$

Proof. The claim follows by plugging the *a priori* bound (5.5) into (6.5) and using the well-known estimate [3, Lem. A.3]:

$$\Lambda_{|\boldsymbol{\nu}|} \leq \frac{|\boldsymbol{\nu}|!}{(\log 2)^{|\boldsymbol{\nu}|}},$$

which yields the assertion. $\quad\square$

6.4. Derivation of the QMC error in the lognormal setting. In analogy to Section 6.2, the dimensionally truncated solution is given by

$$u_{s,h}(\cdot, \mathbf{y}) := u_h(\cdot, (y_1, \ldots, y_s, 0, 0, \ldots)), \qquad \mathbf{y} \in \mathbb{R}^s.$$

THEOREM 6.6. *Let assumptions* (L1)–(L5) *hold, and let* $\eta(\mathbf{y})$ *be chosen as in* (6.3). *There exists a generating vector constructed by the CBC algorithm such that*

$$\sqrt{\mathbb{E}_{\boldsymbol{\Delta}} \left\| \int_{\mathbb{R}^s} u_{s,h}(\cdot, \mathbf{y}) \prod_{j=1}^{s} \varphi(y_j) \, \mathrm{d}\mathbf{y} - Q_{\mathrm{ran}}^\varphi(u_{s,h}) \right\|_{L^2(D)}^2}$$
$$= \begin{cases} \mathcal{O}(n^{-1/p+1/2}) & \text{if } p \in (2/3, 1), \\ \mathcal{O}(n^{-1+\varepsilon}) & \text{for arbitrary } \varepsilon \in (0, 1/2) \quad \text{if } p \in (0, 2/3], \end{cases}$$

where the implied coefficient is independent of s *in both cases when the weights* $(\gamma_{\mathfrak{u}})_{\mathfrak{u} \subseteq \{1:s\}}$ *in Lemma* 3.1 *are chosen to be*

$$\gamma_{\mathfrak{u}} = \left(|\mathfrak{u}|! \prod_{j \in \mathfrak{u}} \frac{\beta_j}{2(\log 2) \exp(\beta_j^2/2) \Phi(\beta_j) \sqrt{(\alpha_j - \beta_j) \varrho_j(\lambda)}} \right)^{2/(1+\lambda)},$$

where

$$\alpha_j = \frac{1}{2} \left(\beta_j + \sqrt{\beta_j^2 + 1 - \frac{1}{2\lambda}} \right)$$

and

$$\lambda = \begin{cases} \frac{p}{2-p} & \text{if } p \in (2/3, 1), \\ \frac{1}{2-2\varepsilon} & \text{for arbitrary } \varepsilon \in (0, 1/2) \quad \text{if } p \in (0, 2/3]. \end{cases}$$

Proof. We note that

$$\frac{1}{a_{\min}(\boldsymbol{y})^{1/2}} \leq \frac{1}{\min_{\boldsymbol{x}\in\overline{D}} a_0(\boldsymbol{x})^{1/2}} \prod_{j\geq 1} \exp(\tfrac{1}{2}\beta_j|y_j|) \qquad \text{for all } \boldsymbol{y} \in U_{\boldsymbol{\beta}},$$

which yields the upper bound

$$\left\| \frac{\partial^{|\mathfrak{u}|}}{\partial \boldsymbol{y}_{\mathfrak{u}}} u_{s,h}(\cdot,\boldsymbol{y}) \right\|_{V_h}^2 \lesssim \widehat{C}_{\mathrm{DG}}^{2|\mathfrak{u}|} \frac{(|\mathfrak{u}|!)^2}{(\log 2)^{2|\mathfrak{u}|}} \left(\prod_{j\in\mathfrak{u}} \beta_j^2 \right) \prod_{j\geq 1} \exp(\beta_j|y_j|),$$

where $\widehat{C}_{\mathrm{DG}} := 4C_{\mathrm{DG}}$.

Since our definition of $\eta(\boldsymbol{y})$ ensures that $\eta(\boldsymbol{y}) \geq a_{\min}(\boldsymbol{y})$, we obtain

$$\int_D \|u_{s,h}(\boldsymbol{x},\cdot)\|_{s,\boldsymbol{\gamma}}^2 \, \mathrm{d}\boldsymbol{x}$$

$$= \sum_{\mathfrak{u}\subseteq\{1:s\}} \frac{1}{\gamma_{\mathfrak{u}}} \int_{\mathbb{R}^{|\mathfrak{u}|}} \int_D \left(\int_{\mathbb{R}^{s-|\mathfrak{u}|}} \frac{\partial^{|\mathfrak{u}|}}{\partial \boldsymbol{y}_{\mathfrak{u}}} u(\boldsymbol{x},\boldsymbol{y}) \prod_{j\notin\mathfrak{u}} \varphi(y_j) \, \mathrm{d}\boldsymbol{y}_{-\mathfrak{u}} \right)^2 \prod_{j\in\mathfrak{u}} \varpi_j^2(y_j) \, \mathrm{d}\boldsymbol{y}_{\mathfrak{u}}$$

$$\leq \sum_{\mathfrak{u}\subseteq\{1:s\}} \frac{1}{\gamma_{\mathfrak{u}}} \int_{\mathbb{R}^s} \left\| \frac{\partial^{|\mathfrak{u}|}}{\partial \boldsymbol{y}_{\mathfrak{u}}} u(\cdot,\boldsymbol{y}) \right\|_{L^2(D)}^2 \prod_{j\notin\mathfrak{u}} \varphi(y_j) \prod_{j\in\mathfrak{u}} \varpi_j^2(y_j) \, \mathrm{d}\boldsymbol{y}$$

$$\leq \sum_{\mathfrak{u}\subseteq\{1:s\}} \frac{1}{\gamma_{\mathfrak{u}}} \int_{\mathbb{R}^s} C_{\mathrm{Poin}}^{V_h}(\boldsymbol{y})^2 \left\| \frac{\partial^{|\mathfrak{u}|}}{\partial \boldsymbol{y}_{\mathfrak{u}}} u(\cdot,\boldsymbol{y}) \right\|_{V_h}^2 \prod_{j\notin\mathfrak{u}} \varphi(y_j) \prod_{j\in\mathfrak{u}} \varpi_j^2(y_j) \, \mathrm{d}\boldsymbol{y}$$

$$\lesssim \sum_{\mathfrak{u}\subseteq\{1:s\}} \frac{\widehat{C}_{\mathrm{DG}}^{2|\mathfrak{u}|}}{\gamma_{\mathfrak{u}}} (|\mathfrak{u}|!)^2 \left(\prod_{j\in\mathfrak{u}} \frac{\beta_j}{\log 2} \right)^2$$

$$\times \int_{\mathbb{R}^s} \prod_{j=1}^s \exp\left(2\beta_j|y_j|\right) \prod_{j\notin\mathfrak{u}} \varphi(y_j) \prod_{j\in\mathfrak{u}} \varpi_j^2(y_j) \, \mathrm{d}\boldsymbol{y},$$

where we used $C_{\mathrm{Poin}}^{V_h}(\boldsymbol{y}) \leq \frac{\sigma}{\sqrt{a_{\min}(\boldsymbol{y})}}$. Thus,

$$\int_D \|u_{s,h}(\boldsymbol{x},\cdot)\|_{s,\boldsymbol{\gamma}}^2 \, \mathrm{d}\boldsymbol{x}$$

$$\leq \sum_{\mathfrak{u}\subseteq\{1:s\}} \frac{\widehat{C}_{\mathrm{DG}}^{2|\mathfrak{u}|}}{\gamma_{\mathfrak{u}}} (|\mathfrak{u}|!)^2 \left(\prod_{j\in\mathfrak{u}} \frac{\beta_j}{\log 2} \right)^2$$

$$\times \left(\prod_{j\notin\mathfrak{u}} \underbrace{\int_{\mathbb{R}} \exp(2\beta_j|y_j|)\varphi(y_j) \, \mathrm{d}y_j}_{\leq 2\exp(2\beta_j^2)\Phi(2\beta_j)} \right) \left(\prod_{j\in\mathfrak{u}} \int_{\mathbb{R}} \exp(2\beta_j|y_j|)\varpi_j^2(y_j) \, \mathrm{d}y_j \right)$$

$$\leq \sum_{\mathfrak{u}\subseteq\{1:s\}} \frac{\widehat{C}_{\mathrm{DG}}^{2|\mathfrak{u}|}}{\gamma_{\mathfrak{u}}} (|\mathfrak{u}|!)^2 \left(\prod_{j=1}^s 2\exp(2\beta_j^2)\Phi(2\beta_j) \right)$$

$$\times \left(\prod_{j\in\mathfrak{u}} \frac{\beta_j^2}{2(\log 2)^2 \exp(2\beta_j^2)\Phi(2\beta_j)} \int_{\mathbb{R}} \exp(2\beta_j|y_j|)\varpi_j^2(y_j) \, \mathrm{d}y_j \right).$$

Recalling that $\varpi_j(y) = \exp(-\alpha_j|y|)$, the remaining integral is bounded as long as we choose

$$(6.6) \qquad\qquad\qquad\qquad \alpha_j > \beta_j.$$

In fact, if (6.6) holds, then

$$\int_{\mathbb{R}} \exp(2\beta_j |y_j|) \varpi_j^2(y_j) \, \mathrm{d}y_j = \frac{1}{\alpha_j - \beta_j},$$

and we obtain

$$\int_D \|u_{s,h}(\boldsymbol{x}, \cdot)\|_{s,\boldsymbol{\gamma}}^2 \, \mathrm{d}\boldsymbol{x} \leq \sum_{\mathfrak{u} \subseteq \{1:s\}} \frac{\widehat{C}_{\mathrm{DG}}^{2|\mathfrak{u}|}}{\gamma_{\mathfrak{u}}} (|\mathfrak{u}|!)^2 \left(\prod_{j=1}^s 2 \exp(2\beta_j^2) \Phi(2\beta_j) \right)$$
$$\times \left(\prod_{j \in \mathfrak{u}} \frac{\beta_j^2}{2(\log 2)^2 \exp(2\beta_j^2) \Phi(2\beta_j)(\alpha_j - \beta_j)} \right).$$

The remainder of the argument is completely analogous to the derivation presented for the continuous setting in [22]: the weights $\gamma_{\mathfrak{u}}$ enter the expression for the upper bound in the same manner as in the continuous setting, leading us to conclude that the error criterion is minimized by setting

$$\alpha_j = \frac{1}{2}\left(\beta_j + \sqrt{\beta_j^2 + 1 - \frac{1}{2\lambda}} \right)$$

and choosing the weights

$$\gamma_{\mathfrak{u}} = \left(|\mathfrak{u}|! \prod_{j \in \mathfrak{u}} \frac{\beta_j}{2(\log 2) \exp(\beta_j^2/2) \Phi(\beta_j) \sqrt{(\alpha_j - \beta_j)\varrho_j(\lambda)}} \right)^{2/(1+\lambda)},$$

with λ chosen as in (6.2) and p as in (L3).   □

REMARK 6.7. Let $\mathcal{G}: V_h \to \mathbb{R}$ be a bounded linear functional. Similarly to the affine and uniform setting, it can be shown that there exists a generating vector constructed by the CBC algorithm using the weights specified in Theorem 6.6 such that

$$\sqrt{\mathbb{E}_{\boldsymbol{\Delta}} \left| \int_{\mathbb{R}^s} \mathcal{G}(u_{s,h}(\cdot, \boldsymbol{y})) \prod_{j=1}^s \varphi(y_j) \, \mathrm{d}\boldsymbol{y} - Q_{\mathrm{ran}}(\mathcal{G}(u_{s,h})) \right|^2} = \mathcal{O}(n^{\max\{-1/p+1/2, -1+\varepsilon\}}),$$

where the implied coefficient is independent of the dimension s with arbitrary $\varepsilon \in (0, 1/2)$.

7. Numerical results. We consider (1.1) with $f(\boldsymbol{x}) = x_1$ in $D = (0,1)^2$ and investigate the errors in the means of the numerical approximations of the unknown u. For the affine case, we set $U = [-\frac{1}{2}, \frac{1}{2}]^{\mathbb{N}}$ and truncate the series expansion for the input random coefficient into $s = 100$ terms, i.e.,

$$a_{\mathrm{affine}}(\boldsymbol{x}, \boldsymbol{y}) = 5 + \sum_{j=1}^{100} \frac{y_j}{(k_j^2 + \ell_j^2)^{1.3}} \sin(k_j \pi x_1) \sin(\ell_j \pi x_2),$$

where $(k_j, \ell_j)_{j \geq 1}$ is an ordering of elements of $\mathbb{Z}_+ \times \mathbb{Z}_+$ such that the sequence $(\|\psi_j\|_{L^\infty(D)})_{j \geq 1}$ is not increasing. In the lognormal case, we define $U = \mathbb{R}^{\mathbb{N}}$ and consider the dimensionally truncated coefficient with $s = 100$ terms

$$a_{\mathrm{lognormal}}(\boldsymbol{x}, \boldsymbol{y}) = \exp\left(\sum_{j=1}^{100} \frac{y_j}{(k_j^2 + \ell_j^2)^{1.3}} \sin(k_j \pi x_1) \sin(\ell_j \pi x_2) \right)$$
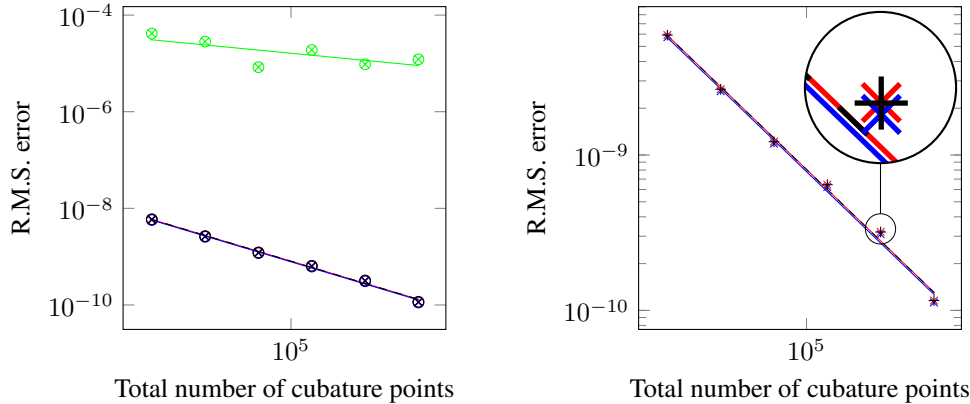
FIG. 7.1. *Root mean squared error for the affine case for first-order polynomial DG methods. The NIPG method with $\eta = 10$ is depited blue, while the SIPG method with $\eta = 100$ is depicted red, and the SIPG method with $\eta = 10$ is depicted green, and the conforming finite element solution is black. The green graph is omitted in the right picture.*

with an analogously defined sequence $(k_j, \ell_j)_{j \geq 1}$. In this case, we have $\|\psi_j\|_{L^\infty(D)} \sim j^{-1.3}$, and the expected convergence independent of the dimension is $\mathcal{O}(n^{-0.8+\varepsilon}), \varepsilon > 0$.

We use in all experiments an off-the-shelf generating vector [34, lattice-39101-1024-1048576.3600] using a total amount of 16384, 32768, 65536, 131072, 262144, and 524288 cubature points with $R = 16$ random shifts. Although this generating vector has not been obtained using the CBC algorithm with the weights derived in Theorems 6.2 and 6.6, we found this off-the-shelf generating vector to perform well yielding the optimal rate of convergence. Thus, this confirms our analytical findings without excluding that there might be even better lattice rules. In practice, the off-the-shelf lattice rule has comparable performance to tailored lattice rules. The finite element discretization uses discontinuous Galerkin methods (the implementation of which is based on the 'Finite Element Simulation Toolbox for UNstructured Grids: FESTUNG' [45]) on a grid of mesh size $\frac{1}{16}$. Additionally, the DG methods are compared to a conforming finite element implementation on 30×30 elements.

Thus, the spatial grid of the discontinuous Galerkin method is significantly coarser than the conforming finite element grid. We do so to underline that it is generally considered 'unfair' to compare discontinuous Galerkin and conforming finite elements on the same grid since DG usually has many more degrees of freedom. There are several ways to compensate this imbalance (count degrees of freedom, number non-zero entries in the stiffness matrix, . . .). However, we will see that the accuracy of the DG and conforming element methods is of minor importance in the numerical experiments. Thus, we skip a detailed discussion about a fair comparison.

Figure 7.1 displays the numerical results for the affine case and locally linear DG or conforming finite element approximation. We recognize that the green plot for the SIPG method with $\eta = 10$ does not show the expected convergence behavior (with respect to the number of cubature points). This nicely resembles our analysis, which highlights that SIPG needs a large enough η to work stably, while the NIPG method is stable for all $\eta > 0$. We clearly see that SIPG (with $\eta = 100$, red), NIPG (with $\eta = 10$, blue), and conforming linear finite elements (black) all show a very similar convergence rate (between 1.098 and 1.1) in the number of cubature points indicating that all three methods work similarly well in the affine case.
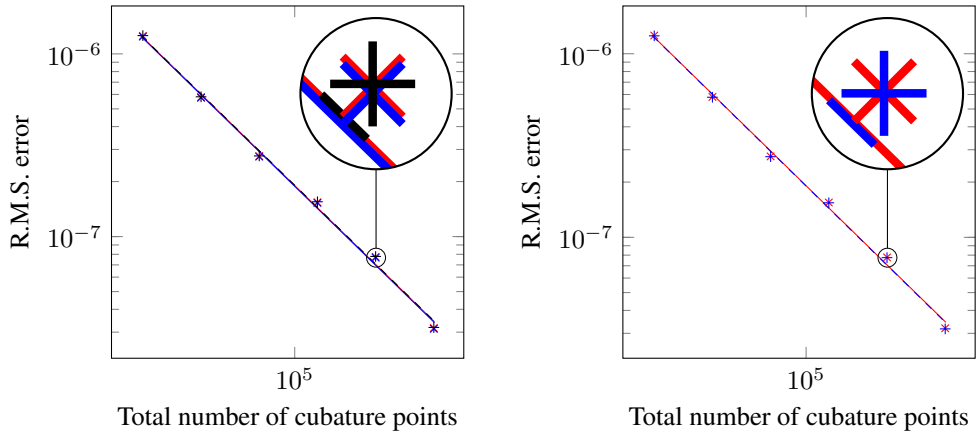
614   V. KAARNIOJA AND A. RUPP

FIG. 7.2. *Root mean squared error for the lognormal case. The NIPG method with $\eta = 10$ is depicted in red, the SIPG method with $\eta = 10$ is depicted in blue, and the conforming finite element solution is black. The left picture illustrates DG and conforming finite elements for locally linear polynomial approximation, while the second picture shows only DG for a locally quadratic approximation.*

In Figures 7.1 and 7.2 we recognize that the error plots of all stable methods are almost identical. Thus, we conclude that the cubature error dominates the discretization error, and hence the particular choice of the discretization appears to be almost irrelevant.

Figure 7.2 deals with the lognormal case. The left picture uses linear approximations of SIPG, NIPG, and the conforming finite element method, while the right picture displays the results for second-order SIPG and NIPG. In the left picture, we see that, again, all three methods work fine and similarly well. Their convergence rates are approximately 1.03 with only insignificant differences. Interestingly, we see similar convergence rates and absolute errors for the second-order SIPG and NIPG approximations. Thus, we deduce that the influence of the accuracy of the conforming finite element or DG approximation of the PDE does not significantly influence the convergence behavior (not even the constants) of the QMC method in our cases.

Notably, the NIPG method is robust concerning η, and we do not see visual differences in the error plots if we choose $\eta(\boldsymbol{y})$ as described in the remark in Section 5.1.

8. Conclusion. We have designed QMC cubatures using DG discretizations of an elliptic PDE with a random coefficient. Our analytical and numerical findings are consistent with the results derived from the conforming approximations in the literature.

In this study, we assumed the data to allow for solutions $u \in H^2$ with respect to the spatial variable. At the cost of some technical but standard extensions of the DG analysis, we can extend our results to the case that $u \in H^{3/2+\epsilon}$ in space. Moreover, using discrete gradients and lifting operators, we believe our results can be transferred to the case of u having minimal spatial regularity. We dropped extensive discussions on spatial regularity as detailed in the DG literature. For QMC methods, the advantages of DG are its more regular execution patterns and the additional local mass conservation. In that sense, it preserves more physics than continuous finite element methods. However, we are not aware of any indications that DG works under milder regularity assumptions than continuous finite elements.

REFERENCES

[1] V. AIZINGER AND C. DAWSON, *The local discontinuous Galerkin method for three-dimensional shallow water flow*, Comput. Methods Appl. Mech. Engrg., 196 (2007), pp. 734–746.

[2] D. N. ARNOLD, F. BREZZI, B. COCKBURN, AND L. D. MARINI, *Unified analysis of discontinuous Galerkin methods for elliptic problems*, SIAM J. Numer. Anal., 39 (2001/02), pp. 1749–1779.

[3] J. BECK, R. TEMPONE, F. NOBILE, AND L. TAMELLINI, *On the optimal polynomial approximation of stochastic PDEs by Galerkin and collocation methods*, Math. Models Methods Appl. Sci., 22 (2012), Paper No. 1250023, 33 pages.

[4] A. CANGIANI, E. H. GEORGOULIS, AND P. HOUSTON, *hp-version discontinuous Galerkin methods on polygonal and polyhedral meshes*, Math. Models Methods Appl. Sci., 24 (2014), pp. 2009–2041.

[5] P. CASTILLO, B. COCKBURN, D. SCHÖTZAU, AND C. SCHWAB, *Optimal a priori error estimates for the hp-version of the local discontinuous Galerkin method for convection-diffusion problems*, Math. Comp., 71 (2002), pp. 455–478.

[6] S.-C. T. CHOI, F. J. HICKERNELL, R. JAGADEESWARAN, M. J. MCCOURT, AND A. G. SOROKIN, *Quasi-Monte Carlo software*, in Monte Carlo and Quasi-Monte Carlo Methods, A. Keller, ed., vol. 387 of Springer Proc. Math. Stat., Springer, Cham, 2022, pp. 23–47.

[7] B. COCKBURN, G. KANSCHAT, AND D. SCHÖTZAU, *A locally conservative LDG method for the incompressible Navier-Stokes equations*, Math. Comp., 74 (2005), pp. 1067–1095.

[8] B. COCKBURN AND C.-W. SHU, *The local discontinuous Galerkin method for time-dependent convection-diffusion systems*, SIAM J. Numer. Anal., 35 (1998), pp. 2440–2463.

[9] A. COHEN, R. DEVORE, AND C. SCHWAB, *Convergence rates of best N-term Galerkin approximations for a class of elliptic sPDEs*, Found. Comput. Math., 10 (2010), pp. 615–646.

[10] R. COOLS, F. Y. KUO, AND D. NUYENS, *Constructing embedded lattice rules for multivariable integration*, SIAM J. Sci. Comput., 28 (2006), pp. 2162–2188.

[11] D. A. DI PIETRO AND A. ERN, *Mathematical Aspects of Discontinuous Galerkin Methods*, Springer, Heidelberg, 2012.

[12] J. DICK, R. N. GANTNER, Q. T. LE GIA, AND C. SCHWAB, *Higher order quasi-Monte Carlo integration for Bayesian PDE inversion*, Comput. Math. Appl., 77 (2019), pp. 144–172.

[13] 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), pp. 2676–2702.

[14] J. DICK, F. Y. KUO, AND I. H. SLOAN, *High-dimensional integration: the quasi-Monte Carlo way*, Acta Numer., 22 (2013), pp. 133–288.

[15] J. DICK, Q. T. LE GIA, AND C. SCHWAB, *Higher order quasi-Monte Carlo integration for holomorphic, parametric operator equations*, SIAM/ASA J. Uncertain. Quantif., 4 (2016), pp. 48–79.

[16] V. DOLEJŠÍ AND M. FEISTAUER, *Discontinuous Galerkin Method*, Springer, Cham, 2015.

[17] M. GANESH, F. Y. KUO, AND I. H. SLOAN, *Quasi-Monte Carlo finite element analysis for wave propagation in heterogeneous random media*, SIAM/ASA J. Uncertain. Quantif., 9 (2021), pp. 106–134.

[18] R. N. GANTNER, *Dimension truncation in QMC for affine-parametric operator equations*, in Monte Carlo and Quasi–Monte Carlo Methods, A. B. Owen and P. W. Glynn, eds., vol. 241 of Springer Proc. Math. Stat., Springer, Cham, 2018, pp. 249–264.

[19] R. N. GANTNER, L. HERRMANN, AND C. SCHWAB, *Multilevel QMC with product weights for affine-parametric, elliptic PDEs*, in Contemporary Computational Mathematics—A Celebration of the 80th Birthday of Ian Sloan. Vol. 1, 2, J. Dick, F. Y. Kuo, and H. Woźniakowski, eds., Springer, Cham, 2018, pp. 373–405.

[20] A. D. GILBERT, I. G. GRAHAM, F. Y. KUO, R. SCHEICHL, AND I. H. SLOAN, *Analysis of quasi-Monte Carlo methods for elliptic eigenvalue problems with stochastic coefficients*, Numer. Math., 142 (2019), pp. 863–915.

[21] C. J. GITTELSON, *Stochastic Galerkin discretization of the log-normal isotropic diffusion problem*, Math. Models Methods Appl. Sci., 20 (2010), pp. 237–263.

[22] 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), pp. 329–368.

[23] 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), pp. 3668–3694.

[24] 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), pp. 479–511.

[25] P. A. GUTH, V. KAARNIOJA, F. Y. KUO, C. SCHILLINGS, AND I. H. SLOAN, *A quasi-Monte Carlo method for optimal control under uncertainty*, SIAM/ASA J. Uncertain. Quantif., 9 (2021), pp. 354–383.

[26] ———, *Parabolic PDE-constrained optimal control under uncertainty with entropic risk measure using quasi-Monte Carlo integration*, Numer. Math., 156 (2024), pp. 565–608.

[27] P. A. GUTH AND A. VAN BAREL, *Multilevel quasi–Monte Carlo for optimization under uncertainty*, Numer. Math., 154 (2023), pp. 443–484.

[28] T. HARTUNG, K. JANSEN, F. Y. KUO, H. LEÖVEY, D. NUYENS, AND I. H. SLOAN, *Lattice meets lattice: application of lattice cubature to models in lattice gauge theory*, J. Comput. Phys., 443 (2021), Paper No. 110527, 24 pages.

[29] L. HERRMANN, M. KELLER, AND C. SCHWAB, *Quasi-Monte Carlo Bayesian estimation under Besov priors in elliptic inverse problems*, Math. Comp., 90 (2021), pp. 1831–1860.

[30] L. HERRMANN AND C. SCHWAB, *QMC integration for lognormal-parametric, elliptic PDEs: local supports and product weights*, Numer. Math., 141 (2019), pp. 63–102.

[31] J. S. HESTHAVEN AND T. WARBURTON, *Nodal Discontinuous Galerkin Methods*, Springer, New York, 2008.

[32] F. J. HICKERNELL, *Quadrature error bounds with applications to lattice rules*, SIAM J. Numer. Anal., 33 (1996), pp. 1995–2016.

[33] P. KNABNER AND L. ANGERMANN, *Numerical Methods for Elliptic and Parabolic Partial Differential Equations*, extended ed., Springer, Cham, 2021.

[34] F. Y. KUO, *Lattice rule generating vectors*, website.
https://web.maths.unsw.edu.au/~fkuo/lattice/index.html

[35] F. Y. KUO AND D. NUYENS, *QMC4PDE software*, software package.
https://people.cs.kuleuven.be/~dirk.nuyens/qmc4pde/

[36] ———, *Application of quasi-Monte Carlo methods to elliptic PDEs with random diffusion coefficients: a survey of analysis and implementation*, Found. Comput. Math., 16 (2016), pp. 1631–1696.

[37] ———, *Application of quasi–Monte Carlo methods to PDEs with random coefficients—an overview and tutorial*, in Monte Carlo and Quasi–Monte Carlo Methods, A. B. Owen and P. W. Glynn, eds., vol. 241 of Springer Proc. Math. Stat., Springer, Cham, 2018, pp. 53–71.

[38] F. Y. KUO, R. SCHEICHL, C. SCHWAB, I. H. SLOAN, AND E. ULLMANN, *Multilevel quasi-Monte Carlo methods for lognormal diffusion problems*, Math. Comp., 86 (2017), pp. 2827–2860.

[39] F. Y. KUO, C. SCHWAB, AND I. H. SLOAN, *Quasi-Monte Carlo finite element methods for a class of elliptic partial differential equations with random coefficients*, SIAM J. Numer. Anal., 50 (2012), pp. 3351–3374.

[40] ———, *Multi-level quasi-Monte Carlo finite element methods for a class of elliptic PDEs with random coefficients*, Found. Comput. Math., 15 (2015), pp. 411–449.

[41] J. A. NICHOLS AND F. Y. KUO, *Fast CBC construction of randomly shifted lattice rules achieving $O(n^{-1+\delta})$ convergence for unbounded integrands over $\mathbb{R}^s$ in weighted spaces with POD weights*, J. Complexity, 30 (2014), pp. 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. Comp., 75 (2006), pp. 903–920.

[43] J. T. ODEN, I. BABUŠKA, AND C. E. BAUMANN, *A discontinuous hp finite element method for diffusion problems*, J. Comput. Phys., 146 (1998), pp. 491–519.

[44] I. PERUGIA AND D. SCHÖTZAU, *An hp-analysis of the local discontinuous Galerkin method for diffusion problems*, J. Sci. Comput. 17 (2002), pp. 561–571.

[45] B. REUTER, H. HAJDUK, A. RUPP, F. FRANK, V. AIZINGER, AND P. KNABNER, *FESTUNG 1.0: overview, usage, and example applications of the MATLAB/GNU Octave toolbox for discontinuous Galerkin methods*, Comput. Math. Appl., 81 (2021), pp. 3–41.

[46] B. RIVIÈRE, *Discontinuous Galerkin Methods for Solving Elliptic and Parabolic Equations*, SIAM, Philadelphia, 2008.

[47] R. SCHEICHL, A. M. STUART, AND A. L. TECKENTRUP, *Quasi-Monte Carlo and multilevel Monte Carlo methods for computing posterior expectations in elliptic inverse problems*, SIAM/ASA J. Uncertain. Quantif., 5 (2017), pp. 493–518.

[48] C. SCHILLINGS, B. SPRUNGK, AND P. WACKER, *On the convergence of the Laplace approximation and noise-level-robustness of Laplace-based Monte Carlo methods for Bayesian inverse problems*, Numer. Math., 145 (2020), pp. 915–971.

[49] C. SCHWAB, *QMC Galerkin discretization of parametric operator equations*, in Monte Carlo and Quasi-Monte Carlo Methods 2012, J. Dick, F. Y. Kuo, G. W. Peters, and I. H. Sloan, eds., vol. 65 of Springer Proc. Math. Stat., Springer, Heidelberg, 2013, pp. 613–629.

[50] C. SCHWAB AND C. J. GITTELSON, *Sparse tensor discretizations of high-dimensional parametric and stochastic PDEs*, Acta Numer., 20 (2011), pp. 291–467.

[51] C. SCHWAB AND R. A. TODOR, *Karhunen-Loève approximation of random fields by generalized fast multipole methods*, J. Comput. Phys., 217 (2006), pp. 100–122.

[52] C. W. SHU, *A brief survey on discontinuous Galerkin methods in computational fluid dynamics*, Adv. Mech., 43 (2013), pp. 541–553.

[53] I. H. SLOAN AND S. JOE, *Lattice Methods for Multiple Integration*, Oxford University Press, New York, 1994.

[54] I. H. SLOAN AND H. WOŹNIAKOWSKI, *When are quasi-Monte Carlo algorithms efficient for high-dimensional integrals?*, J. Complexity, 14 (1998), pp. 1–33.

[55] S. SUN AND M. F. WHEELER, *Analysis of discontinuous Galerkin methods for multicomponent reactive transport problems*, Comput. Math. Appl., 52 (2006), pp. 637–650.

[56] Y. SUZUKI AND D. NUYENS, *Rank-1 lattices and higher-order exponential splitting for the time-dependent Schrödinger equation*, in Monte Carlo and Quasi-Monte Carlo Methods, B. Tuffin and P. L'Ecuyer, eds., vol. 324 of Springer Proc. Math. Stat., Springer, Cham, 2020, pp. 485–502.

[57] A. L. TECKENTRUP, R. SCHEICHL, M. B. GILES, AND E. ULLMANN, *Further analysis of multilevel Monte Carlo methods for elliptic PDEs with random coefficients*, Numer. Math., 125 (2013), pp. 569–600.