

On Stability and Monotonicity Requirements of Discretized Stochastic Conservation Laws with Random Viscosity

Per Pettersson^{a,b}, Alireza Doostan^c and Jan Nordström^d

^a Department of Information Technology,
Uppsala University, P.O. Box 337, SE-75105 Uppsala, Sweden.

^b Institute for Computational and Mathematical Engineering,
Stanford University, Stanford, CA 94305, USA.

^c Aerospace Engineering Science Department,
University of Colorado, Boulder, CO 80309, USA.

^d Department of Mathematics,
Linköping University, SE-58183 Linköping, Sweden.

Abstract

The stochastic Galerkin and collocation methods are used to solve an advection-diffusion equation with uncertain and spatially varying viscosity. We investigate well-posedness, monotonicity and stability for the extended system resulting from the Galerkin projection of the advection-diffusion equation onto the stochastic basis functions. High-order summation-by-parts operators and weak imposition of boundary conditions are used to prove stability of the semi-discrete system.

It is essential that the eigenvalues of the resulting viscosity matrix of the stochastic Galerkin system are positive and we investigate conditions for this to hold. When the viscosity matrix is diagonalizable, stochastic Galerkin and stochastic collocation are similar in terms of computational cost, and for some cases the accuracy is higher for stochastic Galerkin provided that monotonicity requirements are met. We also investigate the total spatial operator of the semi-discretized system and its impact on the convergence to steady-state.

1 Introduction

Partial differential equations (PDEs) subject to uncertain inputs require not only accurate representation of the stochastic input parameters, but also numerical solution methods that limit the numerical errors and converge to the exact solution. For *non-intrusive* methods where the solution statistics are based on multiple runs of deterministic codes, standard analysis techniques can be applied sample-wise to prove numerical properties of interest, such as stability and monotonicity requirements.

Intrusive methods are not based on sampling of deterministic codes, but rather rely on problem formulations that result in extended systems of equations. For the stochastic Galerkin method [12], one generally cannot rely directly on results derived for the deterministic systems evaluated at sample points in the stochastic space. Instead, numerical analysis must be applied directly to the stochastic Galerkin systems.

Generalized polynomial chaos (gPC) methods, for instance stochastic Galerkin, have previously been used in a number of fluid flow problems subject to uncertain viscosity. Ghanem and Dham [11] considered a lognormal diffusion coefficient in a multiphase porous medium problem. Le Maître et. al. investigated a set of Navier-Stokes problems, resulting in coupled sets of advection-diffusion equations with uncertain viscosity [18]. Wan et. al. investigated the advection-diffusion equation in two dimensions with random transport velocity [35], and the effect of long-term time integration of flow problems with gPC methods [34]. Xiu and Karniadakis studied the Navier-Stokes equations with various stochastic boundary conditions [40], as well as steady-state problems with random diffusivity [38].

An alternative to gPC methods is the class of stochastic collocation methods, i.e., sampling methods with interpolation in the stochastic space, c.f. [37]. Several investigations of the relative performance of stochastic Galerkin and collocation methods have been performed, c.f. [20, 3, 32]. The significant size of the stochastic Galerkin system may lead to inefficient direct implementations compared to collocation methods and preconditioned iterative Krylov subspace methods. However, the use of suitable techniques for large systems, such as preconditioners, may result in speedup for the solution of stochastic Galerkin systems compared to multiple collocation runs [32].

The aim of this paper is to present accurate and stable numerical schemes for the solution of a class of linear diffusive transport problems. The advection-diffusion equation subject to uncertain viscosity with known statistical description is represented by a spectral expansion in the stochastic dimension. The gPC framework and the stochastic Galerkin method are used to obtain an extended system which is analyzed to find discretization constraints on monotonicity, stiffness and stability. A comparison of stochastic collocation versus stochastic Galerkin is not the primary focus of this paper. However, we include a few examples on relative performance and numerical properties with respect to monotonicity requirements and convergence to steady-state, to motivate the use of stochastic Galerkin methods.

Special care is exercised to ensure that the stochastic Galerkin projection results in a system with positive semi-definite viscosity matrix. The sign of the eigenvalues of a pure advection problem is not a problem as long as the boundary conditions are properly adjusted to match the number of ingoing characteristics, as shown in [15].

Unlike the case of stochastic advection, the sign of the eigenvalues of the viscosity matrix of the advection-diffusion problem is crucial. A negative eigenvalue leads to the growth of the solution norm and hence numerical instability. The source of the growth is in the volume term, and no boundary condition can address that.

For the spatial discretization, we use high-order accurate summation-by-parts (SBP) operators [17], i.e., central finite differences in the interior of the computational domain complemented with boundary stencils chosen to satisfy the summation-by-parts property, the discrete analogue of integration-by-parts. The SBP operators are combined with weak imposition of boundary conditions with properties that lead to an energy estimate and conditions for numerical stability. Thus, we can prove that a certain order of accuracy will be obtained, provided that the stochastic representation is accurate enough.

The stochastic advection-diffusion equation is presented in Section 2, followed by an outline of the gPC framework and the Galerkin formulation of advection-diffusion in Section 3. Different basis functions and estimates of the eigenvalues of the viscosity matrix are given in Section 4. We prove well-posedness of the problem in Section 5 and monotonicity requirements for the solution are discussed in Section 6. In Section 7, we investigate the time-step limitations of the numerical schemes using von Neumann analysis for a periodic case and summation-by-parts operators for the nonperiodic case. We consider a spatially constant as well as a spatially varying viscosity. Section 7 also includes analysis regarding the convergence rate of the steady-state problem. Numerical results are then presented in Section 8.

2 Problem definition

Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a suitable probability space with the set of elementary events Ω and probability measure \mathcal{P} defined on the σ -algebra \mathcal{F} . Let $\xi(\omega)$, $\omega \in \Omega$, be a random variable defined on this space. Consider the following mixed hyperbolic-parabolic stochastic PDE defined on $(0, 1) \times [0, T]$ which holds \mathcal{P} -almost surely in Ω ,

$$\frac{\partial u}{\partial t} + v \frac{\partial u}{\partial x} = \frac{\partial}{\partial x} \left(\mu(x, \xi) \frac{\partial u}{\partial x} \right), \quad (1)$$

$$u(0, t, \xi) = g_0(t, \xi),$$

$$\frac{\partial u(x, t, \xi)}{\partial x} \Big|_{x=1} = g_1(t, \xi),$$

$$u(x, 0, \xi) = u_{init}(x, \xi). \quad (2)$$

Here the velocity $v > 0$ is a deterministic scalar and the viscosity $\mu(x, \xi) > \mu_0 > 0$ is a finite variance random field. As a special case of (1), we consider the case of $\mu(\xi)$ being spatially constant, but with the same initial and boundary conditions.

In what follows, we approximate the stochastic solution $u(x, t, \xi)$ using a gPC expansion in the random space. We use the stochastic Galerkin method and compare with the stochastic collocation method. Our objective is then to explore the stability, stiffness and monotonicity requirements associated with the numerical solution of the resulting coupled system of equations.

2.1 Uncertainty and solution procedure

We will consider the case where μ has a uniform marginal distribution and thus bounded range, and the case where μ takes a lognormal distribution, a common model in geophysics applications such as transport in porous media [8]. For other distributions, we assume that the viscosity coefficient $\mu(\xi)$ has the cumulative distribution function F . One may parameterize the uncertainty with a uniform random variable ξ , defined on the interval $[-1, 1]$ with constant probability density 0.5, denoted $\xi \sim \mathcal{U}[-1, 1]$. Then we get the expression

$$\mu(\xi) = F^{-1}\left(\frac{\xi + 1}{2}\right), \quad (3)$$

which holds for general distributions F as far as F^{-1} is defined. For the cases of interest here, F^{-1} is a linear function in the case of a uniform μ . In the case of a lognormal μ , we will alternatively use a different representation of μ in terms of the Hermite chaos expansion.

3 Generalized polynomial chaos expansion

The polynomial chaos method introduced in [12] and formulated in a generalized (gPC) framework in [39] is based on the theory developed in [36]. The framework allows for the representation of any finite-variance random variable $f(\xi)$ as a generalized Fourier expansion. Specifically, let $\langle \cdot, \cdot \rangle$ denote the inner product weighted by the probability measure of ξ , defined by

$$\langle f, g \rangle = \int_{\Omega} f(\xi)g(\xi)d\mathcal{P}(\xi) = \langle fg \rangle, \quad (4)$$

where $\langle \cdot \rangle$ denotes the mathematical expectation operator. The inner product in (4) induces a weighted L_2 -norm, denoted by $\|\cdot\|_{L_2(\Omega, \mathcal{P})}$. Let $\{\psi_k(\xi)\}_{k=0}^{\infty}$ be an orthonormal basis for $L_2(\Omega, \mathcal{P})$. Then, for any $f(\xi)$ satisfying $\|f\|_{L_2(\Omega, \mathcal{P})} < \infty$, there exists a gPC expansion

$$\tilde{f}(\xi) = \sum_{k=0}^{\infty} f_k \psi_k(\xi), \quad (5)$$

where the coefficients f_k are defined by the projections

$$f_k = \langle f, \psi_k \rangle = \langle f \psi_k \rangle.$$

In (5), $\tilde{f}(\xi)$ converges to $f(\xi)$ in the mean-squares sense by a generalization of the Cameron-Martin theorem [5] with some additional assumptions, c.f. [9]. For notational convenience (and with some abuse of notation), we will not distinguish between f and its gPC expansion \tilde{f} .

In the context of a stochastic Galerkin solution of Eq. (1), we expand the solution $u(x, t, \xi)$ with respect to a gPC basis $\{\psi_k(\xi)\}_{k=0}^{\infty}$. For optimal convergence, under

certain conditions [39], these polynomials are chosen to be orthonormal with respect to the probability measure \mathcal{P} ,

$$\langle \psi_j, \psi_k \rangle = \delta_{jk},$$

where δ_{jk} is the Kronecker δ . Possible choices of bases include Legendre polynomials that are orthogonal with respect to the uniform measure and Hermite polynomials that are orthogonal with respect to the Gaussian measure. These two sets of orthogonal polynomials are both used in the numerical experiments.

In the computations, we need to use a basis with finite cardinality. Hence, we truncate the gPC basis $\{\psi_k(\xi)\}_{k=0}^{\infty}$ to exactly represent polynomials up to order p ,

$$u^p(x, t, \xi) = \sum_{k=0}^p u_k(x, t) \psi_k(\xi), \quad (6)$$

where $\{\psi_k(\xi)\}_{k=0}^p$ is the set of gPC basis functions of maximum order p .

3.1 Stochastic Galerkin projection

The unknown coefficients $u_k(x, t)$ are then computed through a Galerkin projection onto the subspace spanned by the basis $\{\psi_k(\xi)\}_{k=0}^p$. Specifically, the truncated series (6) is inserted into (1) and multiplied by each one of the basis functions $\{\psi_k(\xi)\}_{k=0}^p$. The resulting expression is integrated with respect to the probability measure \mathcal{P} over the stochastic domain. This leads to a coupled linear system of deterministic PDE's of the form

$$\begin{aligned} \frac{\partial u_k}{\partial t} + v \frac{\partial u_k}{\partial x} &= \sum_{j=0}^p \frac{\partial}{\partial x} \left(\langle \mu \psi_j \psi_k \rangle \frac{\partial u_j}{\partial x} \right), \quad k = 0, \dots, p, \\ u_k(0, t) &= (g_0)_k, \quad k = 0, \dots, p, \\ \frac{\partial u_k(x, t)}{\partial x} \Big|_{x=1} &= (g_1)_k, \quad k = 0, \dots, p, \\ u_k(x, 0) &= (u_{init})_k, \quad k = 0, \dots, p, \end{aligned} \quad (7)$$

where the orthogonality of the basis functions $\{\psi_k(\xi)\}_{k=0}^p$ has been used to cancel terms. Here, $(g_0)_k$, $(g_1)_k$ and $(u_{init})_k$ are obtained by the projection of the left and right boundary data and the initial function on basis polynomial $\psi_k(\xi)$, $k = 0, \dots, p$. In the sequel we use a compact notation to represent the system (7). Let $\mathbf{u} \equiv (u_0 \ u_1 \ \dots \ u_p)^T$ be the vector of chaos coefficients in (6), then the system (7) can be equivalently written as

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{V} \frac{\partial \mathbf{u}}{\partial x} = \frac{\partial}{\partial x} \left(\mathbf{B}(x) \frac{\partial \mathbf{u}}{\partial x} \right), \quad (8)$$

$$\begin{aligned} \mathbf{u}(0, t) &= \mathbf{g}_0(t), \\ \frac{\partial \mathbf{u}(x, t)}{\partial x} \Big|_{x=1} &= \mathbf{g}_1(t), \\ \mathbf{u}(x, 0) &= \mathbf{u}_{init}(x), \end{aligned} \quad (9)$$

where $\mathbf{V} = \text{diag}(v)$ and the matrix \mathbf{B} is defined by

$$[\mathbf{B}(x)]_{jk} = \langle \mu(x) \psi_j \psi_k \rangle \quad j, k = 0, \dots, p. \quad (10)$$

We will frequently refer to the case of spatially independent $\mu(\xi)$. Then, (8) can be simplified to

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{V} \frac{\partial \mathbf{u}}{\partial x} = \mathbf{B} \frac{\partial^2 \mathbf{u}}{\partial x^2}. \quad (11)$$

With the gPC expansion of the viscosity, $\mu = \sum_{k=0}^{\infty} \mu_k \psi_k(\xi)$, (10) can be rewritten as

$$[\mathbf{B}]_{ij} = \langle \mu \psi_i \psi_j \rangle = \sum_{k=0}^{\infty} \mu_k(x) \langle \psi_i \psi_j \psi_k \rangle, \quad i, j = 0, \dots, p. \quad (12)$$

For the basis functions that will be used in this paper, all triple (inner) products $\langle \psi_i \psi_j \psi_k \rangle$ satisfy

$$\langle \psi_i \psi_j \psi_k \rangle = 0, \quad \text{for } k > 2p \text{ and } i, j \leq p. \quad (13)$$

Explicit formulas for $\langle \psi_i \psi_j \psi_k \rangle$ for Hermite and Legendre polynomials can be found in [1, 33]. Hence, using (13), (12) may be simplified to

$$[\mathbf{B}]_{ij} = \sum_{k=0}^{2p} \mu_k(x) \langle \psi_i \psi_j \psi_k \rangle, \quad i, j = 0, \dots, p. \quad (14)$$

The entries of \mathbf{B} can thus be evaluated as finite sums of triple products that can be computed exactly. Moreover, since $[\mathbf{B}]_{ij} = \langle \mu \psi_i \psi_j \rangle = \langle \mu \psi_j \psi_i \rangle = [\mathbf{B}]_{ji}$, it follows that \mathbf{B} is symmetric.

It is essential that the matrix \mathbf{B} be always positive definite when it is derived from a well-defined $\mu(\xi) > 0$. We prove that this holds in the following proposition.

Proposition 1. *The viscosity matrix \mathbf{B} given by (14) derived from any $\mu(\xi)$ satisfying $\mu(\xi) \geq 0$ \mathcal{P} -almost surely in Ω , has non-negative eigenvalues.*

Proof. For any order p of gPC expansion and any vector $\mathbf{u} \in \mathbb{R}^{p+1}$,

$$\begin{aligned} \mathbf{u}^T \mathbf{B} \mathbf{u} &= \sum_{i=0}^p \sum_{j=0}^p u_i u_j \sum_{k=0}^{2p} \langle \psi_i \psi_j \psi_k \rangle \mu_k = \sum_{i=0}^p \sum_{j=0}^p u_i u_j \langle \psi_i \psi_j \mu \rangle = \\ &= \int_{\Omega} \left(\sum_{i=0}^p u_i \psi_i \right)^2 \mu(\xi) d\mathcal{P}(\xi) \geq 0. \end{aligned} \quad (15)$$

□

Remark 1: The above proposition does not hold for the order p approximation $\tilde{\mu}(\xi) = \sum_{k=0}^p \mu_k \psi_k(\xi)$. The second equality of (15) relies on substituting the gPC expansion of μ of order $2p$ with the full gPC expansion of μ . This substitution is valid following (13), but it would not be valid for the order p gPC approximation of μ . In the latter case, the resulting \mathbf{B} may have negative eigenvalues, thus, ruining the stability of the

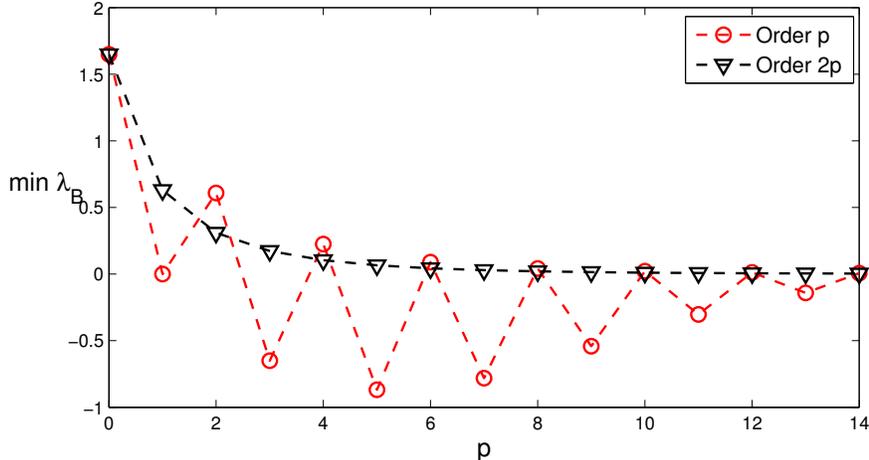


Figure 1: Minimum $\lambda_{\mathbf{B}}$ for $\mu = \exp(\xi)$. Here $[\mathbf{B}^{(p)}]_{ij} = \sum_{k=0}^p \langle \psi_i \psi_j \psi_k \rangle \mu_k$ and $[\mathbf{B}^{(2p)}]_{ij} = \sum_{k=0}^{2p} \langle \psi_i \psi_j \psi_k \rangle \mu_k$, respectively, and $\{\psi_k(\xi)\}$ are the Hermite polynomials.

discrete approximation of (8). Therefore, the $2p$ order of gPC expansion of μ is crucial. Figure 1 illustrates this for the case of a lognormal $\mu(\xi) = \exp(\xi)$ with $\xi \sim \mathcal{N}(0, 1)$.

The proof of Proposition 1 follows closely that of the positive- (negative-) definiteness of the advection matrix of Theorem 2.1 in [15]. However, here we also emphasize the importance of a suitable polynomial chaos approximation of \mathbf{B} , since in this case negative eigenvalues would lead to instability of the numerical method.

3.2 Diagonalization of the stochastic Galerkin system

In order to reduce the computational cost, it is advantageous to diagonalize the stochastic Galerkin systems whenever possible. If this is indeed the case, exact or numerical diagonalization can be done as a preprocessing step, followed by the numerical solution of $p + 1$ scalar advection-diffusion problems with different, but strictly positive, viscosity $\mu((\lambda_{\mathbf{B}})_j)$, where $(\lambda_{\mathbf{B}})_j$ are the eigenvalues of \mathbf{B} , $j = 0, \dots, p$. The system (8) can be diagonalized under certain conditions, which we elaborate on next. Assuming, for a moment, that $\mathbf{B}(x) = \mathbf{W} \mathbf{\Lambda}_B(x) \mathbf{W}^T$, i.e., that the eigenvectors \mathbf{W} of $\mathbf{B}(x)$ are not spatially dependent, then the system (8) can be diagonalized. Multiplying (8) from the left by \mathbf{W}^T and letting $\tilde{\mathbf{u}} = \mathbf{W}^T \mathbf{u}$, we get the diagonalized system

$$\frac{\partial \tilde{\mathbf{u}}}{\partial t} + \mathbf{V} \frac{\partial \tilde{\mathbf{u}}}{\partial x} = \frac{\partial}{\partial x} \left(\mathbf{\Lambda}_B(x) \frac{\partial \tilde{\mathbf{u}}}{\partial x} \right).$$

When the stochastic and space dependent components of $\mu(x, \xi)$ can be factorized or only occur in separate terms of a sum, $\mathbf{B}(x)$ can be diagonalized. That is, for general nonlinear functions f , g and h , and $\mu(x, \xi) = f(x)g(\xi) + h(x)$, we have

$$\mathbf{B}(x) = f(x) \mathbf{W} \mathbf{\Lambda}_g \mathbf{W}^T + h(x) \mathbf{I} = \mathbf{W} (f(x) \mathbf{\Lambda}_g + h(x) \mathbf{I}) \mathbf{W}^T = \mathbf{W} \mathbf{\Lambda}_B(x) \mathbf{W}^T,$$

where $\mathbf{\Lambda}_B(x) = f(x)\mathbf{\Lambda}_g + h(x)\mathbf{I}$, $\mathbf{\Lambda}_g$ is a diagonal matrix, \mathbf{W} is the eigenvector matrix of the eigenvalue decomposition of $[\mathbf{B}_g]_{ij} = \langle g\psi_i\psi_j \rangle$, and \mathbf{I} is the identity matrix. The only requirement on f, g , and h is that the resulting $\mu(x, \xi)$ is positive for all ξ , and bounded in the $L^2(\Omega, \mathcal{P})$ norm.

Notice that the form $\mu(x, \xi) = f(x)g(\xi) + h(x)$ has a given distribution throughout the domain, but not necessarily with the parameters of the distribution being constant. For instance, with $\mu = c_1(x) + c_2(x)\exp(\xi)$ and $\xi \sim \mathcal{N}(0, 1)$, the viscosity is lognormal for all x but with spatially varying statistics and diagonalization. However, for the general case $\mu(x, \xi_1, \dots, \xi_d) = \exp(G(x, \xi_1, \dots, \xi_d))$, with G being a multivariate Gaussian field, diagonalization is not possible.

For the general case of any empirical distribution with simultaneous spatial and stochastic variation diagonalization is not possible. Then we solve the full stochastic Galerkin system for which we perform analysis in the following sections. We also present results on the diagonalizable case, since this admits a very direct comparison to the stochastic collocation techniques, presented next.

3.3 Stochastic collocation

An alternative to the polynomial chaos approach with stochastic Galerkin projection is to use interpolation of solutions corresponding to some realizations of the stochastic inputs. Such *non-intrusive methods* do not require modification of existing codes but rely exclusively on repeated runs of the deterministic code which make them computationally attractive. Stochastic collocation takes a set of solutions $\{u^{(j)}\}$ evaluated at a set $\{\xi^{(j)}\}$ of values of the random input ξ , and constructs an interpolating polynomial from these solution realizations [19, 37, 2].

A common choice of interpolation polynomials is the set of Lagrange polynomials $\{\mathcal{L}_j^{(M)}(\xi)\}_{j=1}^M$, defined by M points $\{\xi^{(j)}\}_{j=1}^M$, for which the polynomial interpolant becomes

$$\mathcal{I}u = \sum_{j=1}^M u^{(j)} \mathcal{L}_j(\xi). \quad (16)$$

The distribution of the grid points $\{\xi^{(j)}\}_{j=1}^M$ is implied by the measure \mathcal{P} of ξ . For instance, we choose $\{\xi^{(j)}\}$ to be the set of Gauss-Legendre quadrature points for the case of uniformly distributed μ , and the set of Gauss-Hermite quadrature points for the case of lognormal μ . The integral statistics of interest, such as moments, may then be approximated by the corresponding quadrature rules. For instance, for some quantity of interest $\langle S(u) \rangle$, we have

$$\langle S(u) \rangle \approx \sum_{j=1}^M S(u^{(j)}) w_j, \quad (17)$$

where w_j is the weight corresponding to the quadrature point $\xi^{(j)}$. The quadrature points and weights can be computed through the Golub-Welsch algorithm [13]. Note that there is no need to find the Lagrange polynomials of (16) explicitly since $(\mathcal{I}u)(\xi^{(j)}) = u^{(j)}$ and we only need the values of $\mathcal{I}u$ at the quadrature points in (17).

Stochastic collocation is similar to other non-intrusive methods such as pseudospectral projection [26] and stochastic response surfaces (least-squares minimization) [4],

in that it relies on evaluating deterministic solutions associated with quadrature points in the stochastic space. The difference is the postprocessing step where quantities of interest are reconstructed by different means of numerical quadrature. Specifically, in stochastic collocation, quantities of interest are computed directly without representing the solutions as a gPC series. Pseudospectral projection, on the other hand, involves the computation of the polynomial chaos coefficients of u through numerical quadrature. Quantities of interest are then calculated as functions of the polynomial chaos coefficients.

We note that properties of monotonicity, stiffness and stability of the PDE's evaluated at the quadrature points is independent of the postprocessing step in which the solution statistics are reconstructed. Therefore, the comparison between stochastic Galerkin and stochastic collocation should be valid for a larger class of non-intrusive methods, but we restrict ourselves to stochastic collocation for the numerical experiments.

4 The eigenvalues of the diffusion matrix \mathbf{B}

In the analysis of the mathematical properties and the numerical scheme, e.g. well-posedness, monotonicity, stiffness and stability, we will need estimates of the eigenvalues of \mathbf{B} . We may express

$$\mathbf{B} = \sum_{k=0}^{\infty} \mu_k \mathbf{C}_k, \quad (18)$$

where μ_k 's are the polynomial chaos coefficients of $\mu(\xi)$ and $[\mathbf{C}_k]_{ij} = \langle \psi_i \psi_j \psi_k \rangle$.

4.1 General bounds on the eigenvalues of \mathbf{B}

Some eigenvalue estimates pertain to all gPC expansions, independent of the actual choice of stochastic basis functions. For example, in cases where $\mu(\xi)$ is bounded within an interval of the real line, the eigenvalues of the viscosity matrix \mathbf{B} can essentially be bounded from above and below by the upper and lower interval boundaries of possible values of μ , respectively. More generally, for any countable basis $\{\psi_k(\xi)\}_{k=0}^{\infty}$ of $L_2(\Omega, \mathcal{P})$, by Theorem 2 of [29], it follows that there is a bound on the set $\{(\lambda_{\mathbf{B}})_j\}_{j=0}^p$ of the eigenvalues of \mathbf{B} , given by

$$(\lambda_{\mathbf{B}})_j \in \text{conv}(\text{spect}(\mu(\xi))) = [\mu_{min}, \mu_{max}], \quad (19)$$

where *conv* denotes the convex hull, and the spectrum *spect* of $\mu(\xi)$ is the essential range, i.e., the set of all possible values (measurable) μ can attain. For a more general exposition and cases where μ is not confined to a convex region, we refer the interested reader to [29]. In this paper, we only consider μ in intervals of finite or infinite length (convex sets), and do not consider degenerate sets or single point values. Following (19), for bounded μ such as uniformly distributed viscosity, the eigenvalues $(\lambda_{\mathbf{B}})_j$ will be restricted to an interval for all orders p of gPC expansion. We expect that the order of chaos expansion has a limited impact on system properties such as monotonicity and stiffness for these cases, as demonstrated in Section 7.3. For unbounded μ (e.g.

lognormal distribution) there is no upper bound on the eigenvalues of \mathbf{B} and the system properties change with the order of gPC, also shown in Section 7.3.

4.2 Legendre polynomial representation

When the viscosity μ is given by $\mu = \mu_0 + \hat{\sigma}\xi$, $\xi \sim \mathcal{U}[-1, 1]$ and $\hat{\sigma}$ is a deterministic scaling factor, only the first two Legendre polynomials are needed to represent μ exactly, that is $\mu = \mu_0\psi_0 + \hat{\sigma}/\sqrt{3}\psi_1$. Then, the stochastic Galerkin projection yields a matrix \mathbf{B} of the form

$$[\mathbf{B}]_{jk} = \langle \mu\psi_j\psi_k \rangle = \mu_0\mathbf{I} + \mu_1\mathbf{C}_1, \quad j, k = 0, \dots, p,$$

where the eigenvalues of \mathbf{C}_1 are given by the Gauss-Legendre quadrature nodes scaled by $\sqrt{3}$. The scaling factor is due to the normalization performed to obtain unit-valued inner double products of the Legendre polynomials. This result follows from the fact that the eigenvalues of the matrix with (i, j) entries defined by $\langle \xi\psi_i\psi_j \rangle$ are the same as those of the Jacobi matrix corresponding to the three-term recurrence of the Legendre polynomials. Thus, they are equal to the Gauss-Legendre quadrature nodes, see e.g. [33, 14] for further details on this assertion.

The Gauss-Legendre nodes are located in the interval $[-1, 1]$, from which it follows that $(\lambda_{\mathbf{B}})_j \in [\mu_0 - \hat{\sigma}, \mu_0 + \hat{\sigma}]$. Note that this holds exactly only for a uniformly distributed μ ; for non-uniform μ the polynomial expansion would result in a matrix series representation of \mathbf{B} of the form (18), where the matrices \mathbf{C}_k are nonzero also for $k > 1$.

4.3 Hermite polynomial representation

Representing the uncertainty of the input parameters with an orthogonal polynomial basis whose weight function does not match the probability measure of the input parameters may lead to poor convergence rates [40]. However, problems where the inputs are functions of Gaussian variables may be represented by gPC expansions in the Hermite polynomials with a weight function matching the Gaussian measure. For instance, log-normal random processes can effectively be represented by Hermite polynomial chaos expansion, see e.g., [10]. Let

$$\mu(\xi) = c_1 + c_2e^\xi, \quad c_1, c_2 \geq 0, \xi \sim \mathcal{N}(0, 1). \quad (20)$$

Then, the Hermite polynomial chaos coefficients of μ are given by

$$\mu_j = \frac{c_2e^{1/2}}{\sqrt{j!}}, \quad j \geq 1. \quad (21)$$

The inner triple products of Hermite polynomials are given by

$$\langle \psi_i\psi_j\psi_k \rangle = \begin{cases} \frac{\sqrt{i!j!k!}}{(s-i)!(s-j)!(s-k)!} & s \text{ integer, } i, j, k \leq s \\ 0 & \text{otherwise,} \end{cases} \quad (22)$$

with $s = (i + j + k)/2$.

Applying Proposition 1 of Section 3.1 to the lognormal μ in (20), it follows that the eigenvalues of \mathbf{B} are bounded below by c_1 . The largest eigenvalue grows with the order p of gPC expansion. Since the entries of \mathbf{B} are non-negative due to (21) and (22), by the Gershgorin's circle theorem, the largest eigenvalue is bounded by the maximum row (column) sum of \mathbf{B} . This gives an estimate of the stiffness of the problem, where a problem is loosely defined as stiff when its numerical solution requires an excessively small time-step for stability.

5 Boundary conditions for well-posedness

A problem is *well-posed* if a solution exists, is unique and depends continuously on the problem data. Boundary conditions that lead to a bounded energy are necessary for well-posedness. For hyperbolic stochastic Galerkin systems, boundary conditions have been derived in [15] for the linear wave equation, and in [25] for the nonlinear case of Burgers' equation. Given the setting of (8) we derive the energy equation by multiplying \mathbf{u}^T with the first equation in (8) and integrating over the spatial extent of the problem. More specifically,

$$\int_0^1 \mathbf{u}^T \frac{\partial \mathbf{u}}{\partial t} dx + \int_0^1 \mathbf{u}^T \mathbf{V} \frac{\partial \mathbf{u}}{\partial x} dx = \int_0^1 \mathbf{u}^T \frac{\partial}{\partial x} \left(\mathbf{B}(x) \frac{\partial \mathbf{u}}{\partial x} \right) dx, \quad (23)$$

which can be compactly written as

$$\frac{\partial \|\mathbf{u}\|^2}{\partial t} + 2 \int_0^1 \frac{\partial \mathbf{u}^T}{\partial x} \mathbf{B}(x) \frac{\partial \mathbf{u}}{\partial x} dx = \left[\mathbf{u}^T \mathbf{V} \mathbf{u} - 2 \mathbf{u}^T \mathbf{B}(x) \frac{\partial \mathbf{u}}{\partial x} \right]_{x=0} - \left[\mathbf{u}^T \mathbf{V} \mathbf{u} - 2 \mathbf{u}^T \mathbf{B}(x) \frac{\partial \mathbf{u}}{\partial x} \right]_{x=1}. \quad (24)$$

Proposition 2. *The problem (8) is well-posed.*

Proof. We consider homogeneous boundary conditions, i.e., let $\mathbf{g}_0 = \mathbf{g}_1 = \mathbf{0}$ in (9). Notice that the right-hand-side of Eq. (24) is negative for the choice of boundary conditions in (9), hence leading to a bounded energy norm of solution \mathbf{u} in time. Uniqueness follows directly from the energy estimate by replacing the solution by the difference between two solutions \mathbf{u} and \mathbf{v} and noticing that the norm of the difference is non-increasing with time, thus $\mathbf{u} \equiv \mathbf{v}$. The problem is parabolic with full-rank \mathbf{B} and the correct number of boundary conditions. This implies the existence of the solution. Therefore, the problem (8) (and also (1)) is well-posed. \square

6 Monotonicity of the solution

In this section we use a *normal modal analysis* technique [16] to derive necessary conditions for the monotonicity of the steady-state solution of the system of equations (11) with spatially constant, but random, viscosity. We provide these conditions for second and fourth order discretization operators.

6.1 Second order operators

With standard second order central differences and a uniform grid, the semi-discrete representation of (11) for the steady-state limit reads

$$\mathbf{V} \frac{\mathbf{u}_{i+1} - \mathbf{u}_{i-1}}{2\Delta x} = \mathbf{B} \frac{\mathbf{u}_{i+1} - 2\mathbf{u}_i + \mathbf{u}_{i-1}}{\Delta x^2}, \quad (25)$$

where \mathbf{u}_i denotes the value of the vector of solution \mathbf{u} at the grid point i in space. This is a system of difference equations with a solution of the form

$$\mathbf{u}_i = \mathbf{y}\kappa^i, \quad (26)$$

for some scalar κ and vector $\mathbf{y} \in \mathbb{R}^{p+1}$ to be determined. By inserting (26) into (25) we arrive at the eigen-problem

$$\left[\frac{\Delta x(\kappa^2 - 1)}{2} \mathbf{V} - (\kappa - 1)^2 \mathbf{B} \right] \mathbf{y} = \mathbf{0} \quad (27)$$

whose non-trivial solution is obtained by requiring

$$\det \left(\frac{\Delta x(\kappa^2 - 1)}{2} \mathbf{V} - (\kappa - 1)^2 \mathbf{B} \right) = 0. \quad (28)$$

The spectral decomposition of the symmetric positive definite matrix \mathbf{B} , i.e., $\mathbf{B} = \mathbf{W} \boldsymbol{\Lambda}_{\mathbf{B}} \mathbf{W}^T$, inserted into (28) leads to

$$v\Delta x(\kappa_j^2 - 1) - (\lambda_{\mathbf{B}})_j(\kappa_j - 1)^2 = 0, \quad j = 0, \dots, p. \quad (29)$$

The solution to (29) is

$$\kappa_j = 1 \text{ or } \frac{2 + \theta_j}{2 - \theta_j}, \quad j = 0, \dots, p, \quad (30)$$

where $\theta_j = \frac{v\Delta x}{(\lambda_{\mathbf{B}})_j}$.

For a monotonic solution \mathbf{u} , we must have $\kappa_j \geq 0$ which demands a mesh such that

$$\text{Re}_{\text{mesh}} = \max_j \theta_j \leq 2. \quad (31)$$

In the case of stochastic collocation, each realization will have a different mesh Reynolds number Re_{mesh} based on the value of $\mu(\xi)$. In combination with the CFL restriction on the time-step Δt , this allows for larger time-steps for simulations corresponding to large values of $\mu(\xi)$, but forces small ones for small $\mu(\xi)$.

The importance of the mesh Reynolds number is illustrated in Figure 2. A step function initially located at $x = 0.2$ is transported to the right and is increasingly smeared by viscosity $\mu \sim \mathcal{U}[0.05, 0.15]$. The mean value is monotonically decreasing, but this property is clearly not preserved by numerical schemes not satisfying the mesh Reynolds number requirement. It also has the effect of erroneously predicting the location of the variance peaks.

When \mathbf{B} can be diagonalized, the solution statistics are functions of linear combinations of scalar advection-diffusion solutions with viscosity given by the eigenvalues

$(\lambda_{\mathbf{B}})_j$. Then there is a *local* mesh Reynolds number $(\text{Re}_{\text{mesh}})_j = \theta_j$ for each eigenvalue $(\lambda_{\mathbf{B}})_j$, and a *global* mesh Reynolds number Re_{mesh} defined by (31). Re_{mesh} is defined also for the cases when \mathbf{B} cannot be diagonalized. If the global mesh Reynolds number for the Galerkin system $\text{Re}_{\text{mesh}} > 2$, but the local mesh Reynolds number $(\text{Re}_{\text{mesh}})_j < 2$ for some instances of the scalar advection-diffusion equation after diagonalization, the lack of monotonicity may not be obvious in the statistics, since these are affected by averaging effects from all scalar solutions. Hence, the lack of monotonicity of the mean solution is more obvious if $(\text{Re}_{\text{mesh}})_j > 2$ for all $j = 0, \dots, p$. This is shown in Figure 3 with $\mu \sim \mathcal{U}[0.14, 0.16]$ for $\text{Re}_{\text{mesh}} = 3$ (and $(\text{Re}_{\text{mesh}})_j > 2, j = 0, \dots, p$) and $\text{Re}_{\text{mesh}} = 1$, respectively.

6.2 Fourth order operators

With fourth order central differences, the semi-discrete representation of (11) for the steady-state limit is given by

$$\mathbf{V} \frac{-\mathbf{u}_{i+2} + 8\mathbf{u}_{i+1} - 8\mathbf{u}_{i-1} + \mathbf{u}_{i-2}}{12\Delta x} = \mathbf{B} \frac{-\mathbf{u}_{i+2} + 16\mathbf{u}_{i+1} - 30\mathbf{u}_i + 16\mathbf{u}_{i-1} - \mathbf{u}_{i-2}}{12\Delta x^2}. \quad (32)$$

Following the procedure of monotonicity analysis used for the second order operators with the ansatz $\mathbf{u}_i = \mathbf{y}\kappa^i$ inserted in (32), we arrive at the eigen-problem

$$[(-\kappa^4 + 8\kappa^3 - 8\kappa + 1)\Delta x \mathbf{V} - (-\kappa^4 + 16\kappa^3 - 30\kappa^2 + 16\kappa - 1)\mathbf{B}] \mathbf{y} = \mathbf{0}. \quad (33)$$

One may verify that $\kappa = 1$ is a root of (33), just as in the case of second order central differences. Using the spectral decomposition of \mathbf{B} and factoring out $(\kappa - 1)$, we obtain the third order equation

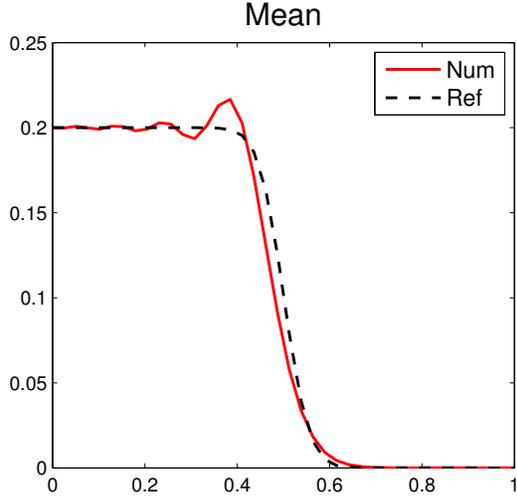
$$(1 - \theta_j) \kappa_j^3 - (15 - 7\theta_j) \kappa_j^2 + (15 + 7\theta_j) \kappa_j - (1 + \theta_j) = 0, \quad (34)$$

for $j = 0, \dots, p$. By Descartes' rule of signs, (34) has only positive roots $\kappa_j > 0$ for $0 < \theta_j < 1$. For $\theta_j > 1$, (34) has as at least one negative root. For $\theta_j = 1$, (34) reduces to a second order equation with two positive roots. Hence, the monotonicity condition $\kappa_j \geq 0$ for the fourth order operators is equivalent to the mesh Reynolds number bound

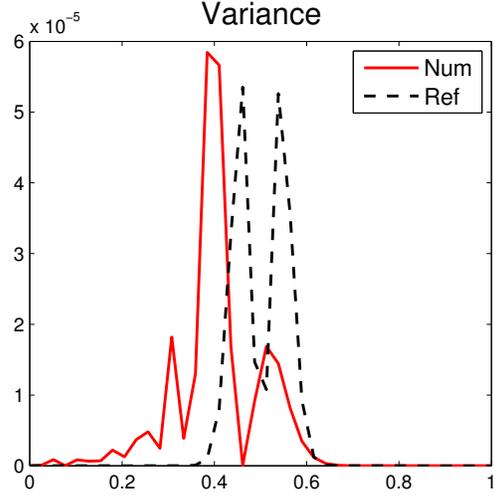
$$\text{Re}_{\text{mesh}} = \max_j \theta_j \leq 1. \quad (35)$$

Remark 2: The monotonicity analysis for sixth order operators can be performed by following the method used for the fourth order ones. The mesh Reynolds number monotonicity condition for sixth order operators is $\text{Re}_{\text{mesh}} \leq \frac{2}{3}$.

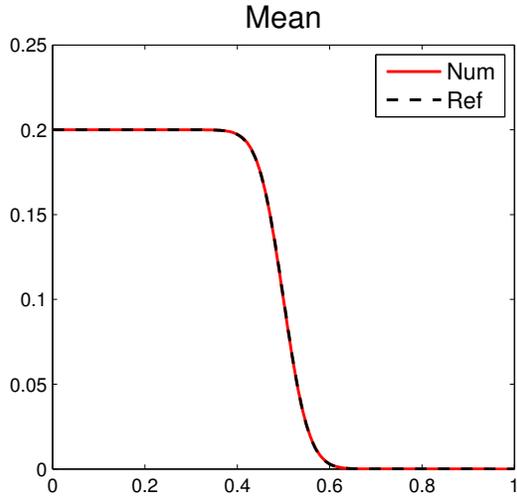
Figure 4 depicts an initial step function after 40 time steps, solved with second, fourth and sixth order operators, respectively. The undershoots of the solutions tend to increase with the order of the scheme which is in-line with the restriction on Re_{mesh} that becomes more severe for higher order operators.



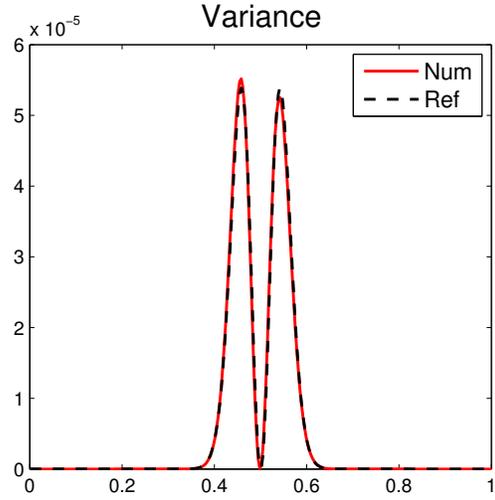
(a) $m = 40$, $Re_{\text{mesh}} = 14$.



(b) $m = 40$, $Re_{\text{mesh}} = 14$.

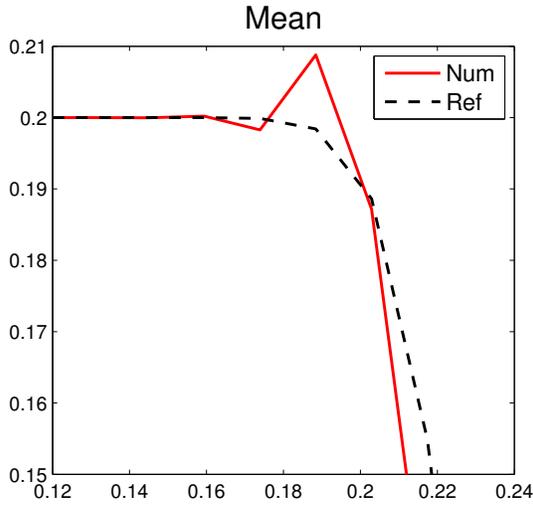


(c) $m = 300$, $Re_{\text{mesh}} = 1.9$.

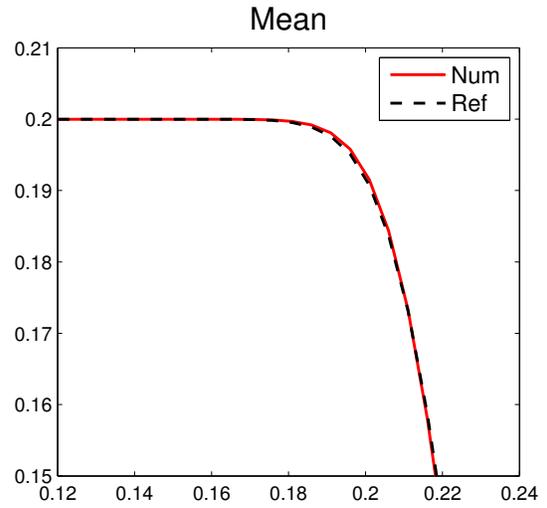


(d) $m = 300$, $Re_{\text{mesh}} = 1.9$.

Figure 2: Solution statistics at $t = 0.01$ using stochastic Galerkin with $p = 4$ for diffusion of a moving step function, $u(x, t, \xi) = \rho_0 \text{erfc} \left(\frac{(x - (x_0 + v(t + \tau))) / \sqrt{(4\mu(\xi)(t + \tau))}} \right)$, $\mu(\xi) \sim \mathcal{U}[0.05, 0.15]$, $\rho_0 = 0.1$, $\tau = 0.005$, $x_0 = 0.2$, and $v = 1$. Here, m denotes the number of spatial grid points.

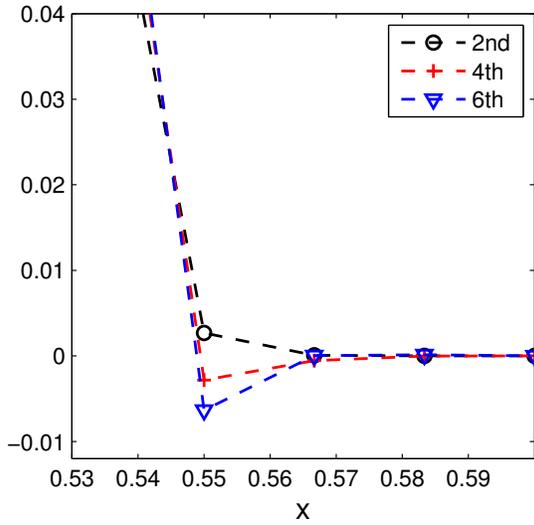


(a) $Re_{\text{mesh}} = 3, m = 70$.

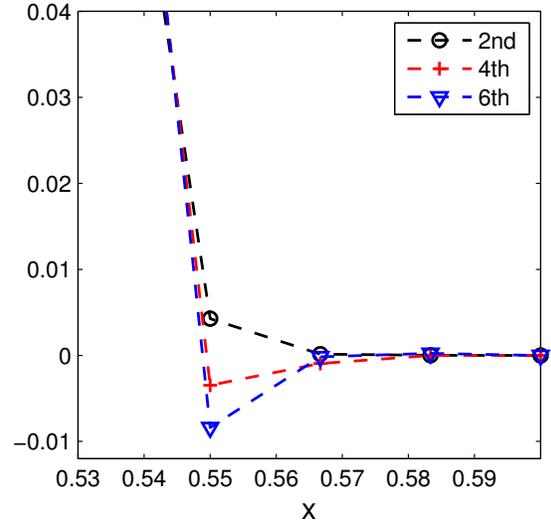


(b) $Re_{\text{mesh}} = 1, m = 200$.

Figure 3: Mean solution at $t = 0.001$ for diffusion of a moving step function, $p = 4$.



(a) $Re_{\text{mesh}} = 0.90$.



(b) $Re_{\text{mesh}} = 1.90$.

Figure 4: Mean solution for diffusion of a moving step function after 40 time steps, $p = 4$, $\mu \sim \mathcal{U}[0.0095, 0.0195]$, $m = 61$ spatial points and two different Re_{mesh} . The undershoot grows with the order of the operators.

7 Stability of the semi-discretized problem

A numerical scheme is *stable* if the semi-discrete problem with homogeneous boundary conditions leads to a bounded energy norm. A stable and consistent scheme converges by the Lax equivalence theorem. Our primary interest is the general case of non-periodic boundary conditions, but the well-known periodic case with spatially constant viscosity $\mu(\xi)$ is also included for comparison.

7.1 The initial value problem: von Neumann analysis

We consider the cases of second and fourth order accurate periodic versions of the central finite difference operators in [22], and show that the amplification factors have negative real parts, describing ellipses in the negative half-plane of the complex plane. The generalization to higher order operators is straightforward.

7.1.1 Second order operators

Assuming spatially constant μ and diagonalizing (11), and using the standard central difference discretization, for all $k = 0, \dots, p$, we get

$$\frac{\partial \tilde{u}_j}{\partial t} + v \frac{\tilde{u}_{j+1} - \tilde{u}_{j-1}}{2\Delta x} = \lambda_k \frac{\tilde{u}_{j+1} - 2\tilde{u}_j + \tilde{u}_{j-1}}{(\Delta x)^2}. \quad (36)$$

We assume periodic boundary conditions and use the Fourier ansatz $\tilde{u}_j = \hat{u} e^{i\alpha \Delta x j}$, where α is the Fourier parameter. Then, with $\theta_k = v\Delta x / (2(\lambda_{\mathbf{B}})_k)$, (36) becomes

$$\begin{aligned} \frac{\partial \hat{u}}{\partial t} &= -i \frac{v}{\Delta x} \frac{e^{i\alpha \Delta x} - e^{-i\alpha \Delta x}}{2i} \hat{u} + \lambda_k \frac{e^{i\alpha \Delta x} - 2 + e^{-i\alpha \Delta x}}{(\Delta x)^2} \hat{u} = \\ &= -\frac{v}{\Delta x} \left[\sin(\alpha \Delta x) i + \frac{2}{\theta_k} (1 - \cos(\alpha \Delta x)) \right] \hat{u}. \end{aligned} \quad (37)$$

The coefficient of \hat{u} in the right hand side of (37) is an expression of the form $f(\omega) = c_1 \cos(\omega) + ic_2 \sin(\omega) + c_3$, i.e., the parametrization of an ellipse in the complex plane. The real part is always non-positive due to the additive constant, so the spectrum is an ellipse in the negative half-plane.

7.1.2 Fourth order operators

The fourth order semi-discretization is given by

$$\frac{\partial \tilde{u}_j}{\partial t} + v \frac{-\tilde{u}_{j+2} + 8\tilde{u}_{j+1} - 8\tilde{u}_{j-1} + \tilde{u}_{j-2}}{12\Delta x} = \lambda_k \frac{-\tilde{u}_{j+2} + 16\tilde{u}_{j+1} - 30\tilde{u}_j + 16\tilde{u}_{j-1} - \tilde{u}_{j-2}}{12(\Delta x)^2}. \quad (38)$$

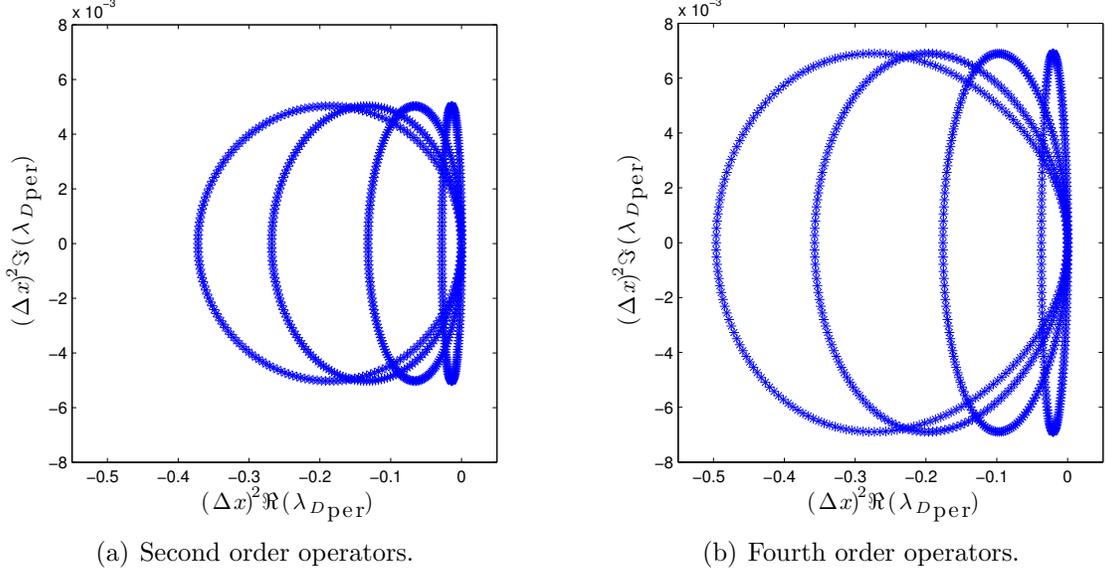


Figure 5: Eigenvalues for order $p = 3$ Legendre polynomial chaos with 200 grid points, $\mu(\xi) \sim \mathcal{U}[0, 0.1]$, $v = 1$.

Again using the Fourier ansatz, we have

$$\begin{aligned}
\frac{\partial \hat{u}}{\partial t} &= i \frac{v}{6\Delta x} \left[\frac{e^{i2\alpha\Delta x} - e^{-i2\alpha\Delta x}}{2i} - 8 \frac{e^{i\alpha\Delta x} - e^{-i\alpha\Delta x}}{2i} \right] \hat{u} + \\
&\quad + \frac{\lambda_k}{6(\Delta x)^2} \left[-\frac{e^{i2\alpha\Delta x} + e^{-i2\alpha\Delta x}}{2} + 16 \frac{e^{i\alpha\Delta x} + e^{-i\alpha\Delta x}}{2} - 15 \right] \hat{u} = \\
&= \left[i \frac{v}{6\Delta x} [\sin(2\alpha\Delta x) - 8 \sin(\alpha\Delta x)] - \frac{\lambda_k}{3(\Delta x)^2} [\cos^2(2\alpha\Delta x) + 8(1 - \cos(\alpha\Delta x))] \right] \hat{u},
\end{aligned} \tag{39}$$

which again is an ellipse in the negative half-plane. This is illustrated in Figure 5, showing the eigenvalues of the second and fourth order periodic spatial discretization matrices \mathbf{D}_{per} . Since \mathbf{D}_{per} is applied to periodic functions, no special boundary treatment is needed. Therefore, the entries of \mathbf{D}_{per} are completely determined by the first and second derivative approximations of (36) and (38), respectively. In Figure 5, the real part of the eigenvalues is denoted by \Re , and the complex part by \Im . Each of the eigenvalues $(\lambda_{\mathbf{B}})_k$, $k = 0, 1, 2, 3$, of \mathbf{B} corresponds to one of the ellipses. For uniformly distributed μ , the range of the eigenvalues is bounded, and increasing the order of gPC does not increase the maximal eigenvalue significantly. Therefore, the order of gPC expansion has a negligible impact on the time-step restriction in this case.

7.2 The initial boundary value problem

In order to obtain stability of the semi-discretized problem for various orders of accuracy and non-periodic boundary conditions, we use discrete operators satisfying a

7.2.1 Spatially varying viscosity

Consider the case of a spatially varying $\mu = \mu(x, \xi)$, given by (8). Since μ depends on x , we cannot write the semi-discretized version of \mathbf{B} as a Kronecker product. Instead, we introduce the block diagonal matrix

$$\hat{\mathbf{B}} = \text{diag}(\mathbf{B}(x_1), \mathbf{B}(x_2), \dots, \mathbf{B}(x_m)).$$

Note that $\hat{\mathbf{B}}$ and the matrix $(\mathbf{P}^{-1} \otimes \mathbf{I})$ commute, i.e.,

$$(\mathbf{P}^{-1} \otimes \mathbf{I})\hat{\mathbf{B}} = \hat{\mathbf{B}}(\mathbf{P}^{-1} \otimes \mathbf{I}). \quad (41)$$

Additionally, $\hat{\mathbf{B}}$ is symmetric, positive definite, and block diagonal. The matrix $(\mathbf{P}^{-1} \otimes \mathbf{I})\hat{\mathbf{B}}$ is a scaling of each diagonal block $\mathbf{B}(x_j)$ of $\hat{\mathbf{B}}$ with the factor $p_{jj}^{-1} > 0$. Thus, $(\mathbf{P}^{-1} \otimes \mathbf{I})\hat{\mathbf{B}}$ is symmetric and positive definite. The numerical approximation of (8) using SBP operators is given by

$$\begin{aligned} \frac{\partial \mathbf{u}}{\partial t} + (\mathbf{P}^{-1}\mathbf{Q} \otimes \mathbf{V})\mathbf{u} &= (\mathbf{P}^{-1}\mathbf{Q} \otimes \mathbf{I})\hat{\mathbf{B}}(\mathbf{P}^{-1}\mathbf{Q} \otimes \mathbf{I})\mathbf{u} \\ &+ (\mathbf{P}^{-1} \otimes \mathbf{I})(\mathbf{E}_0 \otimes \Sigma_0^I)(\mathbf{u} - \mathbf{0}) + (\mathbf{P}^{-1} \otimes \mathbf{I})(\mathbf{Q}^T \mathbf{P}^{-1} \otimes \mathbf{I})(\mathbf{E}_0 \otimes \Sigma_0^V)(\mathbf{u} - \mathbf{0}) \\ &+ (\mathbf{P}^{-1} \otimes \mathbf{I})(\mathbf{E}_N \otimes \Sigma_N^V)((\mathbf{P}^{-1}\mathbf{Q} \otimes \mathbf{I})\mathbf{u} - \mathbf{0}), \end{aligned} \quad (42)$$

where the first line corresponds to the discretization of the PDE, and the second and third lines enforce the homogeneous boundary conditions weakly, here expressed as $(\mathbf{u} - \mathbf{0})$. Although the numerical experiments are performed with nonzero boundary conditions, it is sufficient to consider the homogeneous case in the analysis of stability.

Proposition 3. *The scheme in (42) with $\Sigma_N^V = -\mathbf{B}(x_N)$, $\Sigma_0^V = \mathbf{B}(x_1)$, and $\Sigma_0^I \leq -\mathbf{V}/2$ is stable.*

Proof. Multiplying (42) by $\mathbf{u}^T(\mathbf{P} \otimes \mathbf{I})$ and replacing $\mathbf{Q} = \mathbf{E}_N - \mathbf{E}_0 - \mathbf{Q}^T$ in the first term of the right-hand-side, we obtain

$$\begin{aligned} \mathbf{u}^T(\mathbf{P} \otimes \mathbf{I})\frac{\partial \mathbf{u}}{\partial t} + \underbrace{\mathbf{u}^T(\mathbf{Q} \otimes \mathbf{V})\mathbf{u}}_{\text{Advective term}} &= \underbrace{\mathbf{u}^T(\mathbf{E}_N \otimes \mathbf{I})\hat{\mathbf{B}}(\mathbf{P}^{-1}\mathbf{Q} \otimes \mathbf{I})\mathbf{u}}_{\text{Viscous terms from PDE}} \\ - \underbrace{\mathbf{u}^T(\mathbf{E}_0 \otimes \mathbf{I})\hat{\mathbf{B}}(\mathbf{P}^{-1}\mathbf{Q} \otimes \mathbf{I})\mathbf{u} - \mathbf{u}^T(\mathbf{Q}^T \otimes \mathbf{I})\hat{\mathbf{B}}(\mathbf{P}^{-1} \otimes \mathbf{I})(\mathbf{Q} \otimes \mathbf{I})\mathbf{u}}_{\text{Viscous terms from PDE}} &+ \underbrace{\mathbf{u}^T(\mathbf{E}_0 \otimes \Sigma_0^I)\mathbf{u}}_{\text{Adv. penalty term}} \\ &+ \underbrace{\mathbf{u}^T(\mathbf{Q}^T \mathbf{P}^{-1} \otimes \mathbf{I})(\mathbf{E}_0 \otimes \Sigma_0^V)\mathbf{u}}_{\text{Left viscous penalty term}} + \underbrace{\mathbf{u}^T(\mathbf{E}_N \otimes \Sigma_N^V)(\mathbf{P}^{-1}\mathbf{Q} \otimes \mathbf{I})\mathbf{u}}_{\text{Right viscous penalty term}}. \end{aligned} \quad (43)$$

The right viscous penalty term and the first viscous term from the PDE cancel if we set $\Sigma_N^V = -\mathbf{B}(x_N)$. Adding the transpose of the remaining terms of (43) to themselves and using (41), we arrive at the energy equation

$$\begin{aligned}
& \frac{\partial}{\partial t} \|\mathbf{u}\|_{P \otimes I}^2 + \overbrace{\mathbf{u}^T (\mathbf{E}_N \otimes \mathbf{V}) \mathbf{u} - \mathbf{u}^T (\mathbf{E}_0 \otimes \mathbf{V}) \mathbf{u}}^{\text{Advective boundary terms}} = \\
& = \underbrace{-\mathbf{u}^T (\mathbf{E}_0 \otimes \mathbf{I}) \hat{\mathbf{B}} (\mathbf{P}^{-1} \mathbf{Q} \otimes \mathbf{I}) \mathbf{u} - \mathbf{u}^T (\mathbf{Q}^T \mathbf{P}^{-1} \otimes \mathbf{I}) \hat{\mathbf{B}} (\mathbf{E}_0 \otimes \mathbf{I}) \mathbf{u}}_{\text{Viscous terms from PDE}} \\
& \quad - 2\mathbf{u}^T (\mathbf{Q}^T \otimes \mathbf{I}) \hat{\mathbf{B}} (\mathbf{P}^{-1} \otimes \mathbf{I}) (\mathbf{Q} \otimes \mathbf{I}) \mathbf{u} + \underbrace{2\mathbf{u}^T (\mathbf{E}_0 \otimes \boldsymbol{\Sigma}_0^I) \mathbf{u}}_{\text{Adv. penalty term}} \\
& \quad + \underbrace{\mathbf{u}^T (\mathbf{Q}^T \mathbf{P}^{-1} \otimes \mathbf{I}) (\mathbf{E}_0 \otimes \boldsymbol{\Sigma}_0^V) \mathbf{u} + \mathbf{u}^T (\mathbf{E}_0 \otimes \boldsymbol{\Sigma}_0^V) (\mathbf{P}^{-1} \mathbf{Q} \otimes \mathbf{I}) \mathbf{u}}_{\text{Left viscous penalty terms}}. \quad (44)
\end{aligned}$$

The viscous terms from the PDE and the left viscous penalty terms cancel if we set $\boldsymbol{\Sigma}_0^V = \mathbf{B}(x_1)$. Pairing the second advective boundary term with the advective penalty term for stability and choosing $\boldsymbol{\Sigma}_0^I = -\delta \mathbf{V}$ where $\delta \in \mathbb{R}$ leads to

$$\frac{\partial}{\partial t} \|\mathbf{u}\|_{P \otimes I}^2 = u_0^T (1 - 2\delta) v u_0 - u_N^T v u_N - 2 [(\mathbf{Q} \otimes \mathbf{I}) \mathbf{u}]^T \hat{\mathbf{B}} (\mathbf{P}^{-1} \otimes \mathbf{I}) [(\mathbf{Q} \otimes \mathbf{I}) \mathbf{u}]. \quad (45)$$

For $\delta \geq 1/2$, i.e., $\boldsymbol{\Sigma}_0^I \leq -\mathbf{V}/2$, the energy rate (45) shows that the scheme (42) with variable \mathbf{B} is stable, as the norm of \mathbf{u} decays with time. \square

7.2.2 Spatially constant viscosity

For the case of spatially constant viscosity $\mu(\xi)$, we use compact second-derivative SBP operators. We show that the choice of penalty matrices is similar to the case of spatially varying viscosity $\mu(x, \xi)$ presented in the preceding section. The scheme is given by

$$\begin{aligned}
\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{P}^{-1} \mathbf{Q} \otimes \mathbf{V}) \mathbf{u} &= (\mathbf{P}^{-1} (-\mathbf{M} + \tilde{\mathbf{B}} \mathbf{D}) \otimes \mathbf{B}) \mathbf{u} \\
&+ (\mathbf{P}^{-1} \otimes \mathbf{I}) (\mathbf{E}_0 \otimes \boldsymbol{\Sigma}_0^I) (\mathbf{u} - \mathbf{0}) + (\mathbf{P}^{-1} \otimes \mathbf{I}) (\mathbf{D}^T \otimes \mathbf{I}) (\mathbf{E}_0 \otimes \boldsymbol{\Sigma}_0^V) (\mathbf{u} - \mathbf{0}) \\
&+ (\mathbf{P}^{-1} \otimes \mathbf{I}) (\mathbf{E}_N \otimes \boldsymbol{\Sigma}_N^V) ((\mathbf{D} \otimes \mathbf{I}) \mathbf{u} - \mathbf{0}), \quad (46)
\end{aligned}$$

Proposition 4. *The scheme in (46) with the parameters $\boldsymbol{\Sigma}_0^V = \mathbf{B}$, $\boldsymbol{\Sigma}_N^V = -\mathbf{B}$, and $\boldsymbol{\Sigma}_0^I \leq -\mathbf{V}/2$ is stable.*

The proof of Proposition 4 is given in Appendix A.

7.3 Eigenvalues of the total system matrix

The semi-discrete scheme (46) is an ODE system of the form

$$\frac{\partial \mathbf{u}}{\partial t} = \mathbf{D}_{tot} \mathbf{u},$$

whose properties are determined by the complex-valued eigenvalues of the total system matrix \mathbf{D}_{tot} . The eigenvalues of \mathbf{D}_{tot} must all have negative real parts for stability.

The utmost right lying eigenvalue determines the slowest decay rate, and thus the speed of convergence to steady-state, see [23, 24]. The total spatial operator defined by the scheme (46) with $\Sigma_0^V = \mathbf{B}$, $\Sigma_N^V = -\mathbf{B}$, and $\Sigma_0^I = -\mathbf{V}/2$ is given by the matrix

$$\mathbf{D}_{tot} = (\mathbf{P}^{-1} \otimes \mathbf{I}) (-(\mathbf{Q} + \mathbf{E}_0/2) \otimes \mathbf{V} + (\mathbf{D}^T \mathbf{E}_0 - \mathbf{E}_0 \mathbf{D} - \mathbf{M}) \otimes \mathbf{B}). \quad (47)$$

The location in the complex plane of the eigenvalues of \mathbf{D}_{tot} depends on the distribution of μ , the spatial step Δx , and the ratio between viscosity and advective speed.

Figure 6 depicts the eigenvalues of \mathbf{D}_{tot} for uniform $\mu(\xi) \sim \mathcal{U}[0, 0.04]$, $v = 1$, different orders of polynomial chaos, and number of spatial grid points. The fourth order SBP operators have been used and penalty coefficients are chosen according to the stability analysis above. The eigenvalues all have negative real parts, showing that the discretizations are indeed stable. Note that for an order of gPC expansion p , there will be $p + 1$ eigenvalues for each single eigenvalue of the corresponding deterministic system matrix. The groups of $p + 1$ eigenvalues are clustered around the corresponding eigenvalue of the deterministic system matrix. When the range of possible viscosity values (uncertainty) is increased, the spreading of the eigenvalues within each cluster increases. When the mean of the viscosity is increased, the eigenvalues with non-zero complex part will move away farther from the origin.

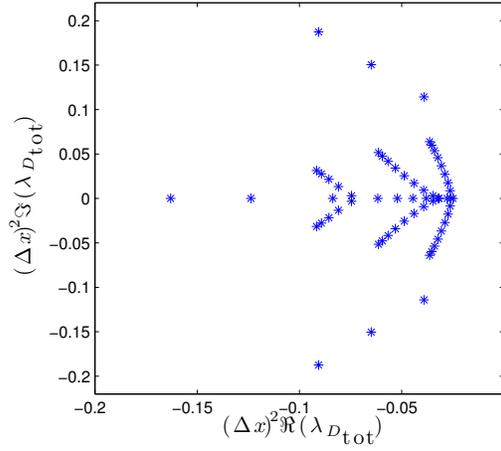
The change of location of the eigenvalues with increasing order of gPC expansion gives an idea how the time-step restriction changes with the order of gPC. Figure 7 shows the eigenvalues of the total system matrix for uniform and lognormal μ for first order (left) and fourth order (right) gPC. For the random viscosities to be comparable, the coefficients are chosen such that the first and second moments of the uniform and the lognormal μ match each other. For low-order polynomial chaos expansions, the eigenvalues are close to each other and the systems are similar in terms of stiffness. As the order of gPC expansion is increased, the scattering of the eigenvalues of \mathbf{B} resulting from lognormal μ increases (μ is unbounded). Hence, the stochastic Galerkin system becomes stiffer with increasing order of gPC. The time-step restriction for the uniform viscosity does not change significantly with the order of gPC. The fourth order operators are a factor of approximately 1.5 stiffer than the second-order operators. Here, we calculate stiffness as

$$\rho_{\text{stiff}} = \frac{\max |\lambda_{\mathbf{D}_{tot}}|}{\min |\lambda_{\mathbf{D}_{tot}}|},$$

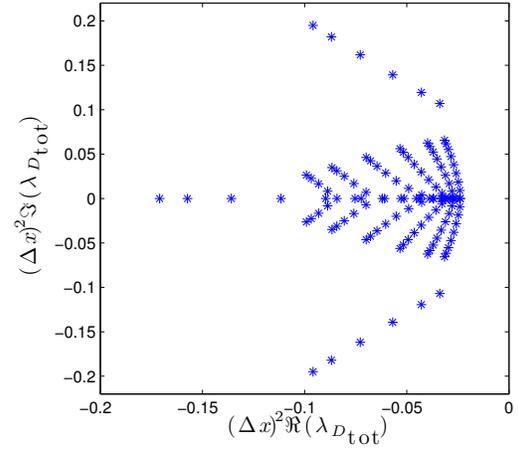
where $|\lambda_{\mathbf{D}_{tot}}|$ denote the absolute values of the complex eigenvalues of the total spatial operator \mathbf{D}_{tot} .

7.4 Convergence to steady-state

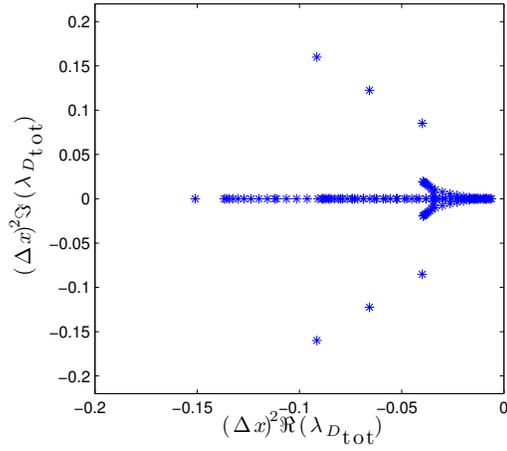
As we let $t \rightarrow \infty$, the problem (8) with $\mathbf{B}(x) > 0$ will reach steady-state, i.e., it will satisfy $\partial \mathbf{u} / \partial t = 0$. This situation can be formulated as a time-independent problem



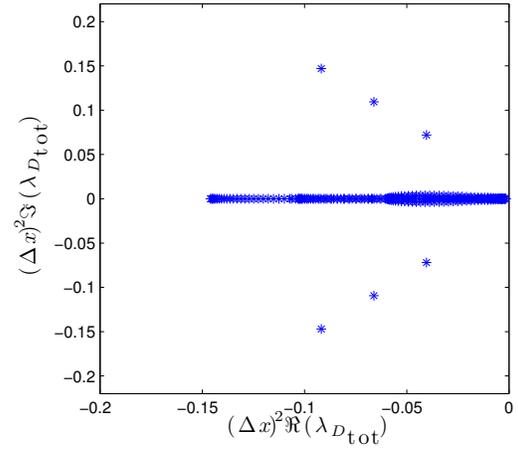
(a) $p = 2, m = 20$.



(b) $p = 5, m = 20$.



(c) $p = 2, m = 40$.



(d) $p = 2, m = 80$.

Figure 6: Eigenvalues of the total operator \mathbf{D}_{tot} (including force terms). Comparison of different orders of gPC (a) and (b), and different grid sizes (a), (c), and (d).

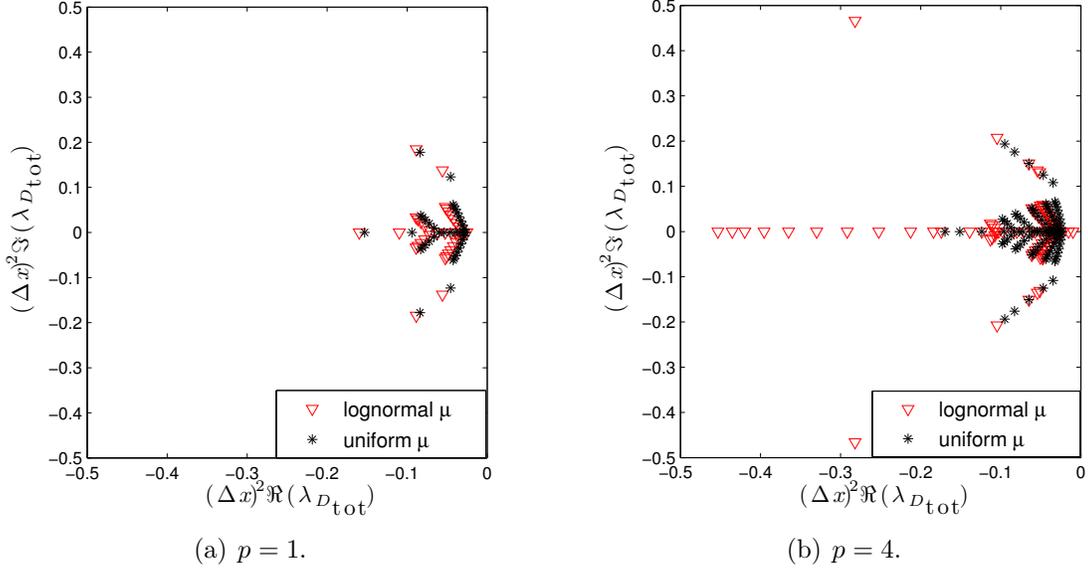


Figure 7: Eigenvalues of the total operator for $m = 20$ and different orders of polynomial chaos. Here the viscosity μ has mean $\langle \mu \rangle = 0.02$ and variance $\text{Var}(\mu) = 3.33e - 5$, and has uniform and lognormal distributions.

with solution $\tilde{\mathbf{u}}$, that satisfies

$$\begin{aligned} \mathbf{V} \frac{\partial \tilde{\mathbf{u}}}{\partial x} &= \frac{\partial}{\partial x} \left(\mathbf{B}(x) \frac{\partial \tilde{\mathbf{u}}}{\partial x} \right), \\ \tilde{\mathbf{u}}(x=0) &= \tilde{\mathbf{g}}_0, \\ \frac{\partial \tilde{\mathbf{u}}(x)}{\partial x} \Big|_{x=1} &= \tilde{\mathbf{g}}_1. \end{aligned} \quad (48)$$

By subtracting (48) from (8), we get the initial boundary value problem for the deviation $\mathbf{e} = \mathbf{u} - \tilde{\mathbf{u}}$ from steady-state,

$$\begin{aligned} \frac{\partial \mathbf{e}}{\partial t} + \mathbf{V} \frac{\partial \mathbf{e}}{\partial x} &= \frac{\partial}{\partial x} \left(\mathbf{B}(x) \frac{\partial \mathbf{e}}{\partial x} \right), \\ \mathbf{e}(0, t) &= 0, \\ \frac{\partial \mathbf{e}(x, t)}{\partial x} \Big|_{x=1} &= 0, \\ \mathbf{e}(x, 0) &= \mathbf{u}_{init}(x), \end{aligned} \quad (49)$$

where it has been used that as $t \rightarrow \infty$, the boundary data must be independent of time and vanish. The problem (49) can be semi-discretized analogous to the numerical schemes presented in Section 7.2. Thus, with \mathbf{D}_{tot} defined in (47), the aim is to solve the initial value problem

$$\frac{\partial \mathbf{e}}{\partial t} = \mathbf{D}_{tot} \mathbf{e}, \quad t > 0, \quad (51)$$

$$\mathbf{e} = \mathbf{u}_{init}(x), \quad t = 0, \quad (52)$$

with the solution $\mathbf{e}(x, t) = \mathbf{u}_{init}(x) \exp(\mathbf{D}_{tot}t)$. The largest real component of the eigenvalues of \mathbf{D}_{tot} , denoted by $\max \Re(\lambda_{\mathbf{D}_{tot}})$, must be negative; otherwise, the solution will not converge to steady-state. The more negative $\max \Re(\lambda_{\mathbf{D}_{tot}})$ is, the faster the convergence to steady-state.

Although the boundary conditions may be altered in different ways to accelerate the convergence to steady-state [24], we use the weak imposition of boundary conditions described in Section 7.2 and compare the convergence to steady-state for a diagonalizable stochastic Galerkin system with that of the stochastic collocation. The number of iterations to reach convergence to steady-state depends on the size of the time-step and the exponential decay of the solution, governed by the rightmost lying eigenvalue of the total system matrix, $\max \Re(\lambda_{\mathbf{D}_{tot}})$. For each stochastic quadrature point of the advection-diffusion equation, there is a maximal time-step as well as a maximal eigenvalue of the total system matrix. For stochastic Galerkin, each scalar instance of the advection-diffusion equation corresponds to one of the eigenvalues of \mathbf{B} , and for stochastic collocation each instance corresponds to μ evaluated at a stochastic quadrature point.

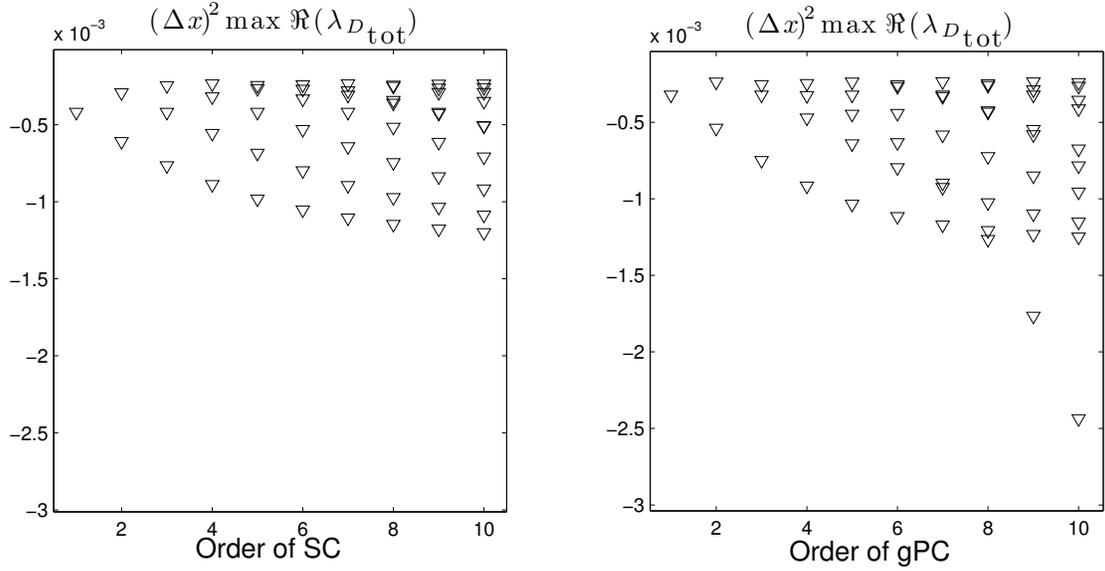
Figure 8 depicts the maximum time-step and the maximum eigenvalue of \mathbf{D}_{tot} for each one of the instances of advection-diffusion equations for different approximation orders of the gPC and stochastic collocation. If diagonalization is possible, and for sufficiently high orders of stochastic Galerkin, the scalar instances of the continuous advection-diffusion equation with the most negative $\max \Re(\lambda_{\mathbf{D}_{tot}})$ converge to steady-state faster than the corresponding instances of stochastic collocation. However, the severe time-step limit of stochastic Galerkin implies that a large number of time-steps are needed to reach steady-state numerically with explicit time-stepping. It is not clear from Figure 8 alone whether stochastic Galerkin or stochastic collocation reaches steady-state numerically in the smaller number of time-steps. This will be investigated in Section 8.2.

For non-diagonalizable stochastic Galerkin, the local bounds of Figure 8 on time-steps and maximum eigenvalues no longer apply. Instead, the most severe local time-step limit and eigenvalue will dominate the entire stochastic Galerkin system, with deteriorating performance as a consequence. Stochastic collocation is still subject to local time-step restrictions and local maximum eigenvalues, and is expected to converge faster to steady-state than stochastic Galerkin.

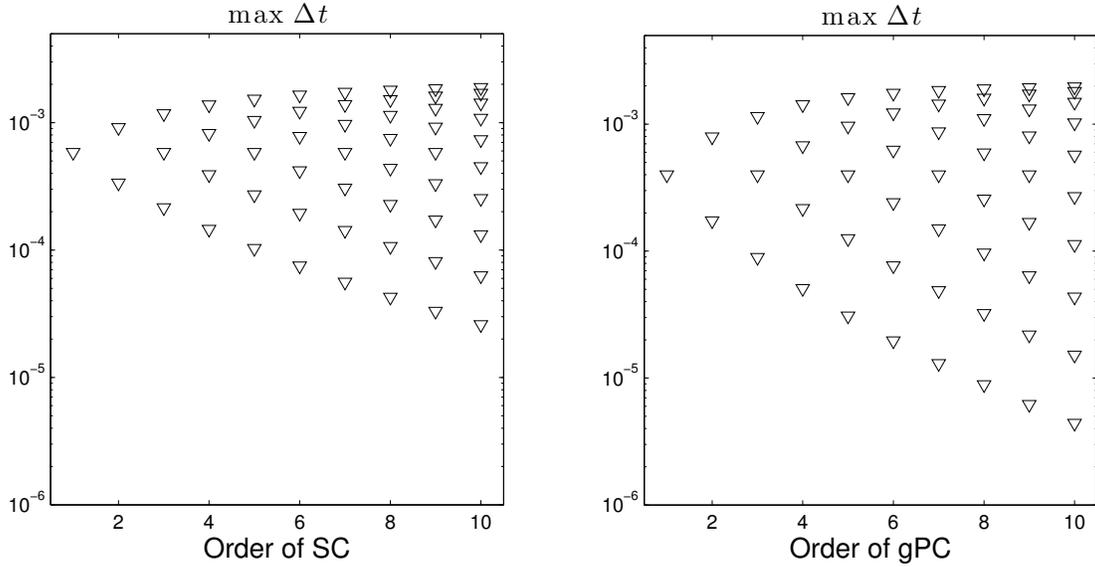
8 Numerical results

In the numerical examples of this Section, we use a fourth order Runge-Kutta method for the time integration and the fourth order accurate SBP-SAT scheme in space. The matrix operators can be found in [22]. The problem (1) with spatially independent μ is solved for the initial function

$$u_0(x, \xi) = \frac{\rho_0}{\sqrt{4\pi\mu(\xi)\tau}} \exp\left(-\frac{(x - (x_0 + v\tau))^2}{4\mu(\xi)\tau}\right), \quad \rho_0 > 0, \quad x_0 \in [0, 1], \quad \tau > 0,$$



(a) $\max \Re(\lambda_{\mathbf{D}_{tot}})$ for each quadrature point as a function of the order of stochastic collocation. (b) $\max \Re(\lambda_{\mathbf{D}_{tot}})$ for the scalar advection-diffusion equations (one for each eigenvalue ($\lambda_{\mathbf{B}}$)) for different orders of stochastic Galerkin.



(c) Time-step limit for each quadrature point of stochastic collocation. (d) Time-step limit for each scalar advection diffusion equation (diagonalizable system) of stochastic Galerkin.

Figure 8: Convergence to steady-state depends on the limit on Δt and $\max \Re(\lambda_{\mathbf{D}_{tot}})$. These quantities are plotted for lognormal viscosity $\mu(\xi) = 0.02 + 0.05 \exp(\xi)$, $\xi \sim \mathcal{N}(0, 1)$. Stochastic collocation (left) and stochastic Galerkin (right).

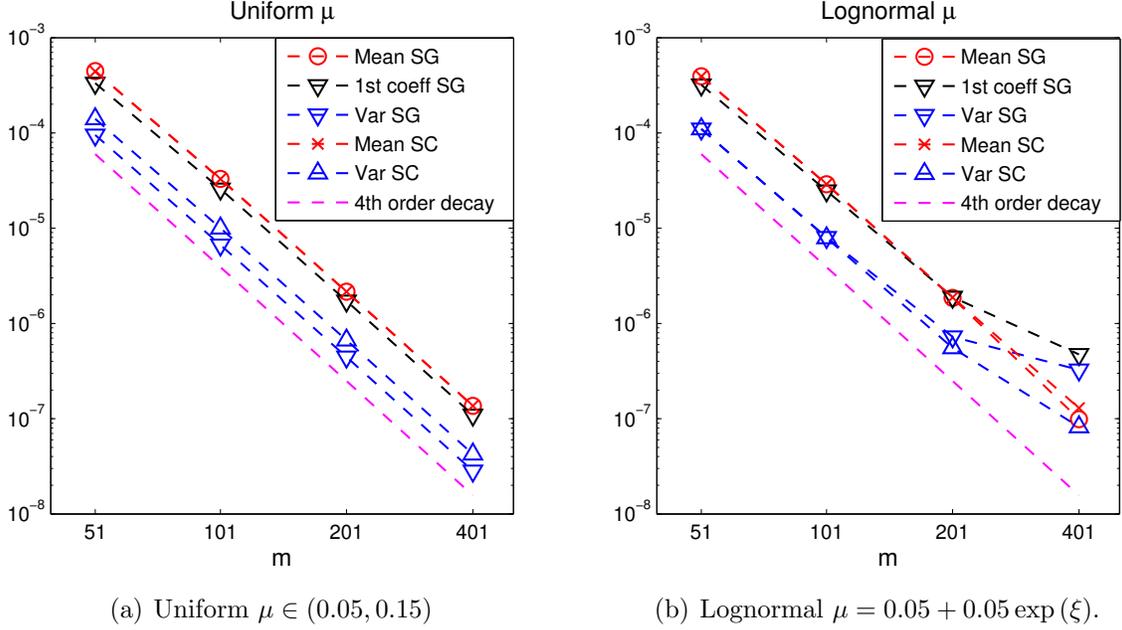


Figure 9: Convergence with respect to the spatial discretization using stochastic Galerkin (SG) and stochastic collocation (SC). Plotted are norms of the absolute errors in mean, first coefficient and variance with $p = 12$ order of generalized Legendre/Hermite chaos, and $N = 13$ quadrature points for stochastic collocation.

for which the analytical solution at time t is given by

$$u(x, t, \xi) = \frac{\rho_0}{\sqrt{4\pi\mu(\xi)(t + \tau)}} \exp\left(-\frac{(x - (x_0 + v(t + \tau)))^2}{4\mu(\xi)(t + \tau)}\right). \quad (53)$$

For the spatially varying $\mu(x, \xi)$, we employ the method of manufactured solutions [27, 28] where we get the same solution as in the case of spatially constant $\mu(\xi)$ with the aid of an appropriate source function $s(x, t, \xi)$ in (1).

The stochastic reference solution (53) is projected onto the gPC basis functions using a high-order numerical quadrature. The order N of the quadrature is chosen sufficiently large so that the difference between two successive reference solutions of order $N - 1$ and N are several orders of magnitude smaller than the difference between the solution from the numerical scheme and the reference solution.

Figure 9 illustrates the convergence as the spatial grid is refined for constant order of gPC, $p = 12$ and $N = 13$ collocation points. For this high-order stochastic representation, the theoretical fourth order convergence rate is attained for the mean using stochastic Galerkin and stochastic collocation. For the variance, the stochastic truncation error becomes visible for fine spatial meshes with lognormal μ , see Figure 9 (b). There is no significant difference in performance between stochastic collocation and stochastic Galerkin for this test case.

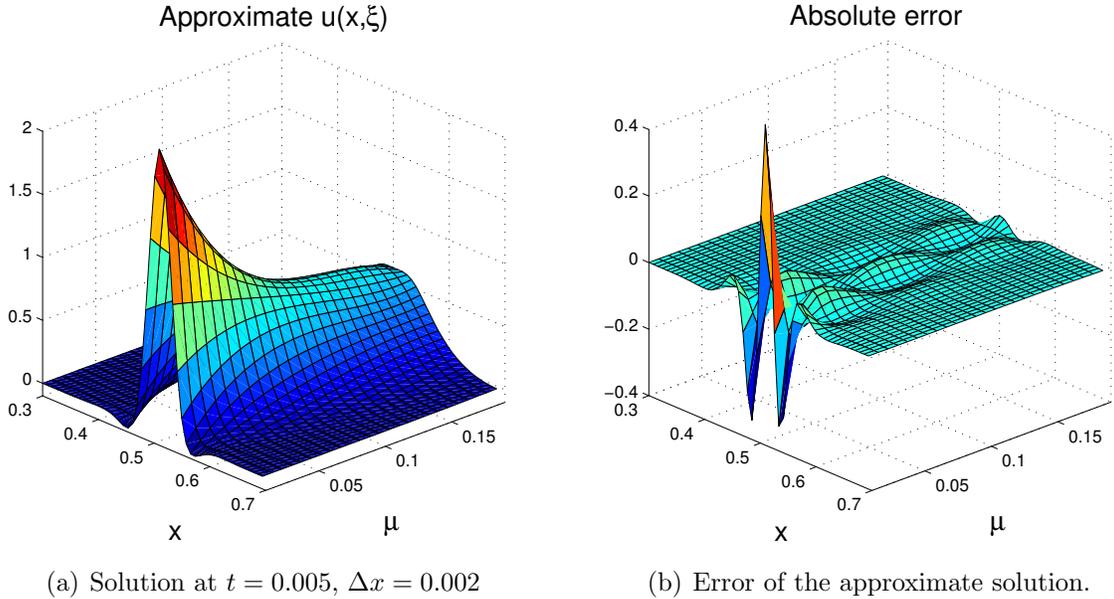


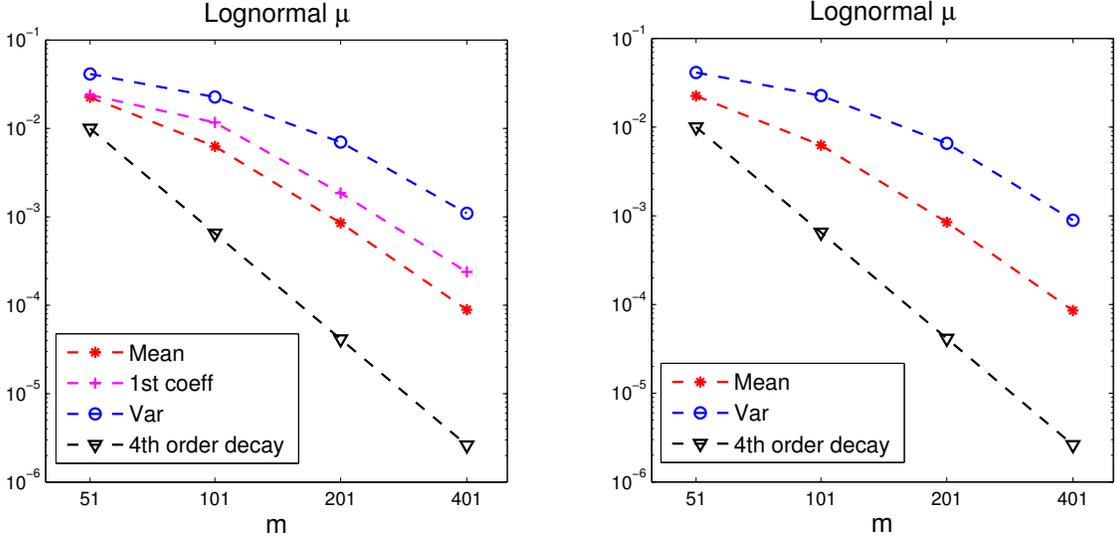
Figure 10: Approximate solution with $p = 3$ order of Legendre chaos.

8.1 The inviscid limit

The theoretical results for the advection-diffusion problem are based on $\mu > 0$. When μ is arbitrarily close to 0 (but non-negative), the problem becomes nearly hyperbolic. In the stochastic setting, this happens with non-zero probability whenever $\mu(\xi) \in [0, c]$, $c > 0$. For small μ the mesh must be very fine, otherwise the mesh Reynolds number requirement discussed in Section 6 will be violated. This is illustrated in Figure 10 (numerical solution left and error right) for results obtained with fourth order SBP operators, $v = 1$ and $\mu \sim \mathcal{U}[0.01, 0.19]$. Note that the error is maximal close to the inviscid limit of $\mu = 0.01$. The solution (53) is a Gaussian in space for any fixed value of ξ and t and varies exponentially in x with the inverse of μ . Thus, spatial convergence requires a fine mesh for small μ . Deterioration of the convergence properties for small μ is a well-known phenomenon for other problems with a parabolic term, c.f. the Navier-Stokes equations in the inviscid limit.

Figure 11 shows the convergence in space for $\mu = 0.02 \exp(\xi)$, $\xi \sim \mathcal{N}(0, 1)$, using stochastic Galerkin (left) and stochastic collocation (right). When μ approaches zero, the gradients become steeper, which requires finer meshes. The fourth order convergence rate is clearly seen. As long as the stochastic basis is rich enough to represent the uncertainty, the choice of stochastic collocation versus stochastic Galerkin has neither any significant effect on the rate of spatial convergence, nor on the actual error. However, the number of stochastic basis functions needed for a certain level of resolution increases as μ goes to zero; therefore, a simultaneous increase in spatial and stochastic resolution is necessary for convergence in the inviscid limit.

The performance of stochastic Galerkin versus stochastic collocation depends on the proximity to the inviscid limit. Figure 12 shows the convergence in the order of gPC expansion (stochastic Galerkin) and the number of quadrature points (stochastic



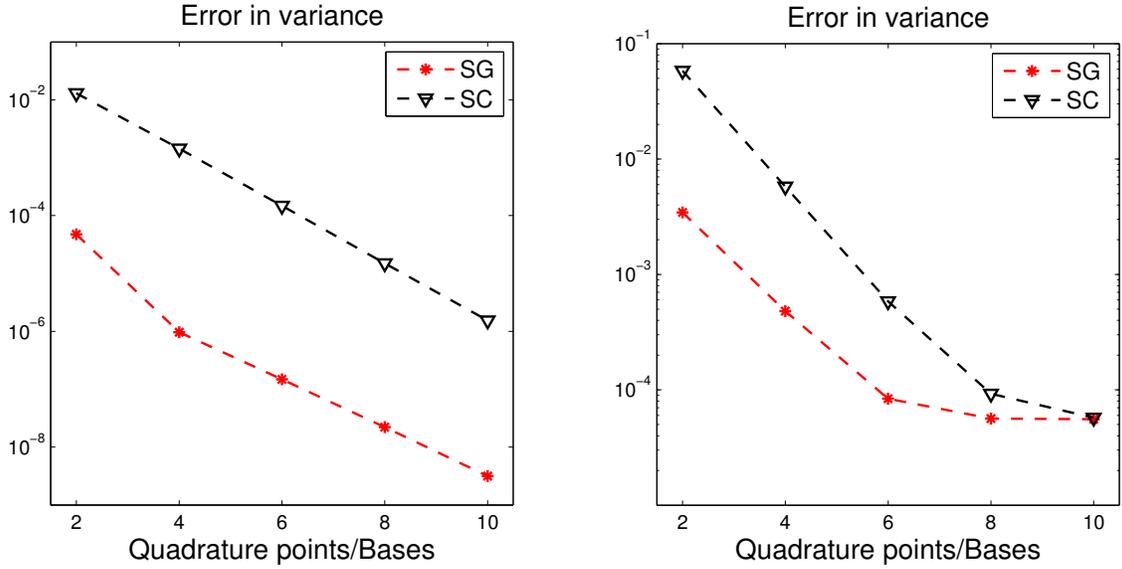
(a) Stochastic Galerkin with $p = 9$.

(b) Stochastic collocation with 10 quadrature points.

Figure 11: Lognormal viscosity, $\mu = 0.02 \exp(\xi)$. $T = 0.001$ and $\tau = 0.005$.

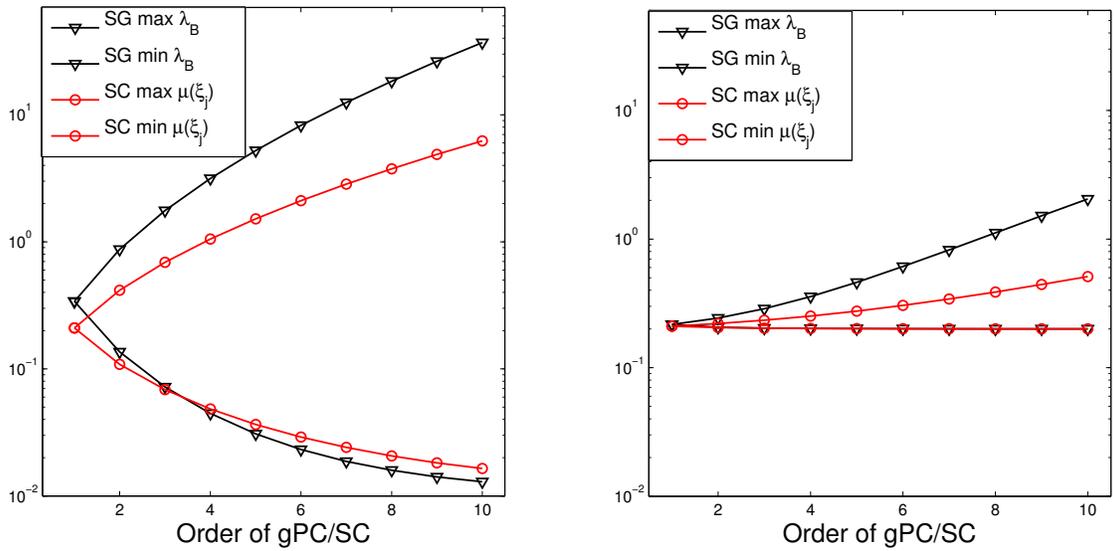
collocation) for a fixed spatial grid. Two cases of lognormal μ are compared; one with $\mu_{min} = 0.2$ and one with $\mu_{min} = 0.01$. For these cases, the stochastic Galerkin system can be diagonalized, so the cost for stochastic Galerkin with an expansion order $p - 1$ is equivalent to the cost of stochastic collocation with p quadrature points. If the problem is highly diffusive, stochastic Galerkin is the more efficient method. If the viscosity is close to zero with some non-zero probability, the difference in performance decreases. Low viscosity sharpens the solution features. The effect of this on the spatial convergence is seen in the low-viscosity case ($\mu_{min} = 0.01$) in Figure 12 (b), where the spatial truncation error becomes visible for high-order polynomial chaos expansions. Due to the fixed number of spatial grid points, the convergence rate decreases for high-order stochastic representations. With a sufficiently fine mesh, one could show exponential convergence rate for any given order of stochastic representation.

Both the stochastic collocation and diagonalizable stochastic Galerkin rely on a set of scalar advection-diffusion problems, with the difference between the methods lying in the choice of stochastic viscosity point values and the postprocessing used to obtain statistics of interest. Figure 13 displays the difference in the range of the effective values of μ for stochastic collocation and stochastic Galerkin (this corresponds to the range of eigenvalues of \mathbf{B} for stochastic Galerkin). From a purely numerical point of view, stochastic Galerkin poses an additional challenge compared to stochastic collocation in that a wider range of scales of diffusion must be handled simultaneously, as shown in Figure 13. If we were to choose the eigenvalues of the matrix \mathbf{B} as the collocation points, the two methods would only differ in the postprocessing.



(a) Lognormal viscosity, $\mu = 0.2 + 0.01 \exp(\xi)$. (b) Lognormal viscosity, $\mu = 0.01 + 0.01 \exp(\xi)$.

Figure 12: Stochastic Galerkin (SG) and stochastic collocation (SC) as a function of the order of gPC/number of quadrature points. Fixed mesh of 201 spatial points.



(a) Lognormal viscosity, $\mu = 0.01 + 0.2 \exp(\xi)$. (b) Lognormal viscosity, $\mu = 0.2 + 0.01 \exp(\xi)$.

Figure 13: Minimum and maximum viscosity for different orders of stochastic Galerkin (SG) and stochastic collocation (SC) for two different distributions of μ . This corresponds to the minimum and maximum λ_B for SG and to the minimum and maximum $\mu(\xi)$ for SC.

8.2 Steady-state calculations

Let the time of numerical convergence to steady-state be defined as the time T_{ss} when the discretized residual \mathbf{e} satisfies $\|\mathbf{e}\|_{2,\Delta x} = (\Delta x \sum_{i=1}^m (e(x_i, T_{ss}))^2)^{1/2} < tol$, where tol is a numerical tolerance to be chosen a priori. When μ is sufficiently large so that diffusion is the dominating feature compared to advection, larger values of the range of μ implies that T_{ss} decreases, and steady-state is reached sooner. The number of iterations to steady-state (i.e. the number of time-steps $T_{ss}/\Delta t$) is inversely proportional to Δt . On the other hand, the limit on Δt decreases with μ . Hence, there is a trade-off in the number of iterations to steady-state between the size of the time-step and the eigenvalues or quadrature point values of μ . In Figure 14, this issue is explored for a lognormal $\mu = c_1 + c_2 \exp(\xi)$ with different choices of c_1 and c_2 .

From the previous analysis and Figures 8 and 13, we have observed how the eigenvalues grow with the order p of gPC. For the most advection dominated case, Figure 14 (a), the number of iterations grows superlinearly with the number of quadrature points in the stochastic collocation approach. The same holds for up to order $p = 8$ of stochastic Galerkin. For this case, the fastest convergence to steady-state is obtained for stochastic collocation, where the range of possible μ values is narrower than in the case of stochastic Galerkin, see Figure 13. In the more diffusive case, i.e., Figure 14 (b), the relative performance for stochastic collocation versus stochastic Galerkin is less pronounced. In the most diffusive case considered here, Figure 14 (c), the number of stochastic Galerkin iterations required to steady-state is a sublinear function of the order of gPC. In this case, the largest eigenvalues of $\lambda_{\mathbf{B}}$ yield advection-diffusion equations that converge within a relatively short time T_{ss} , which compensates for a severe time-step restriction. For these diffusive cases, stochastic Galerkin is more efficient than stochastic collocation. The stochastic Galerkin problem has been diagonalized to make the computational cost per iteration similar. In summary, Figure 14 shows that stochastic collocation converges faster than stochastic Galerkin to steady-state for problems that are advection dominated or moderately diffusive. For diffusion dominated flows, stochastic Galerkin converges faster to steady-state compared to stochastic collocation.

9 Summary and conclusions

Summation-by-parts operators and weak boundary treatment have been applied to a stochastic Galerkin formulation of the advection-diffusion equation. We have presented conditions for monotonicity for stochastic, but spatially constant, viscosity, and stable schemes for the more general case of spatially varying and stochastic viscosity.

Violation of the derived upper bound on the mesh Reynolds number may lead to spurious oscillations, but it may also result in less obviously recognizable errors that are visible in different ways, e.g. as incorrect predictions of regions of large variation. The limit on the mesh Reynolds number gets more severe for higher order operators.

In the case of spatially independent viscosity as well as spatially varying viscosity, the advection-diffusion Galerkin system can be diagonalized under some conditions. This results in a number of uncoupled systems and the numerical cost and perfor-

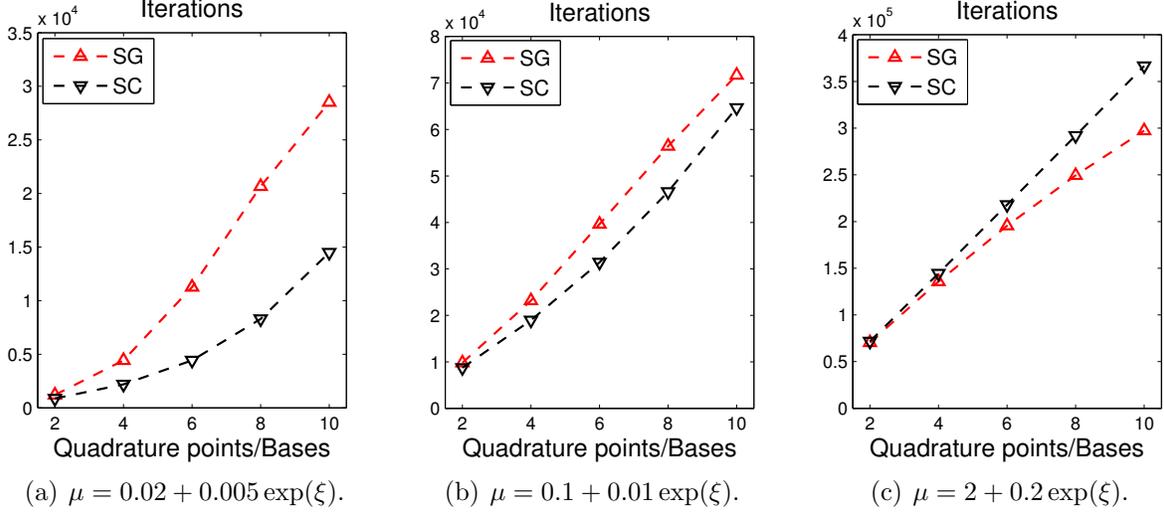


Figure 14: Number of iterations to steady-state for different lognormal viscosity $\mu = c_1 + c_2 \exp(\xi)$ using stochastic Galerkin and stochastic collocation. Here $tol = 10^{-6}$.

mance is very similar to non-intrusive methods such as pseudospectral projection and stochastic collocation.

For diffusive problems, the stochastic Galerkin formulation leads to better accuracy compared to stochastic collocation. For steady-state calculations, stochastic collocation is faster for advection dominated cases and stochastic Galerkin is faster for diffusive cases. When diagonalization of the viscosity matrix \mathbf{B} is possible, the problem should be solved in a non-intrusive way to reduce the computational cost.

SBP operators are suitable for smooth problems like the advection-diffusion equation investigated here, but many real-world flow problems contain regions of sharp gradients or discontinuities. For these problems, one may use hybrid schemes consisting of shock capturing methods in regions of strong variation, coupled through weak interfaces with SBP schemes in smooth regions.

Appendix A Proof of Proposition 4

Proof. Multiplying (46) by $\mathbf{u}^T(\mathbf{P} \otimes \mathbf{I})$ and using $\tilde{\mathbf{B}} = -\mathbf{E}_0 + \mathbf{E}_N$, we get

$$\begin{aligned}
 \mathbf{u}^T(\mathbf{P} \otimes \mathbf{I}) \frac{\partial \mathbf{u}}{\partial t} + \underbrace{\mathbf{u}^T(\mathbf{Q} \otimes \mathbf{V})\mathbf{u}}_{\text{Advective term}} &= \underbrace{-\mathbf{u}^T(\mathbf{M} \otimes \mathbf{B})\mathbf{u} - \mathbf{u}^T(\mathbf{E}_0 \mathbf{D} \otimes \mathbf{B})\mathbf{u} + \mathbf{u}^T(\mathbf{E}_N \mathbf{D} \otimes \mathbf{B})\mathbf{u}}_{\text{Viscous terms from PDE}} \\
 + \underbrace{\mathbf{u}^T(\mathbf{E}_0 \otimes \Sigma_0^I)\mathbf{u}}_{\text{Inviscid penalty term}} &+ \underbrace{\mathbf{u}^T(\mathbf{D}^T \otimes \mathbf{I})(\mathbf{E}_0 \otimes \Sigma_0^V)\mathbf{u}}_{\text{Left viscous penalty term}} + \underbrace{\mathbf{u}^T(\mathbf{E}_N \otimes \Sigma_N^V)(\mathbf{D} \otimes \mathbf{I})\mathbf{u}}_{\text{Right viscous penalty term}},
 \end{aligned} \tag{54}$$

As in the case of variable viscosity, setting $\Sigma_0^V = \mathbf{B}$ and $\Sigma_N^V = -\mathbf{B}$ cancels the viscous boundary terms of the ODE. Using the relation

$$\begin{aligned} \mathbf{u}^T(\mathbf{Q} \otimes \mathbf{V})\mathbf{u} &= \mathbf{u}^T \left(\frac{1}{2}(\mathbf{Q} + \mathbf{Q}^T) \otimes \mathbf{V} \right) \mathbf{u} + \underbrace{\mathbf{u}^T \left(\frac{1}{2}(\mathbf{Q} - \mathbf{Q}^T) \otimes \mathbf{V} \right) \mathbf{u}}_{=0} = \\ &= \frac{1}{2} \mathbf{u}^T ((-\mathbf{E}_0 + \mathbf{E}_N) \otimes \mathbf{V})\mathbf{u}, \end{aligned} \quad (55)$$

we arrive at

$$\mathbf{u}^T(\mathbf{P} \otimes \mathbf{I}) \frac{\partial \mathbf{u}}{\partial t} = -\mathbf{u}^T(\mathbf{M} \otimes \mathbf{B})\mathbf{u} + \mathbf{u}^T(\mathbf{E}_0 \otimes (\mathbf{V}/2 + \Sigma_0^I))\mathbf{u} - \mathbf{u}^T(\mathbf{E}_N \otimes \mathbf{V}/2)\mathbf{u} \quad (56)$$

Finally, setting $\Sigma_0^I = -\delta \mathbf{V}$ as in Section 7.2.1 and adding the transpose of (56) to itself, we get the energy estimate

$$\frac{\partial}{\partial t} \|\mathbf{u}\|_{(P \otimes I)}^2 = u_0^T(1 - 2\delta)vu_0 - u_N^Tvu_N - \mathbf{u}^T((\mathbf{M} + \mathbf{M}^T) \otimes \mathbf{B})\mathbf{u}. \quad (57)$$

Since $\mathbf{M} + \mathbf{M}^T$ and \mathbf{B} are positive definite, the relation (57) with $\delta \geq 1/2$, i.e. $\Sigma_0^I \leq -\mathbf{V}/2$, proves that the scheme in (46) is stable. \square

Acknowledgements

The first author gratefully acknowledges funding from King Abdullah University of Science and Technology (KAUST), Saudi Arabia. The second author gratefully acknowledges the support of the Department of Energy under grant DE-SC0006402.

References

- [1] Ruth Azor, J. Gillis, and J. D. Victor. Combinatorial applications of Hermite polynomials. *SIAM J. Math. Anal.*, 13(5):879–890, 1982.
- [2] Ivo Babuška, Fabio Nobile, and Raul Tempone. A stochastic collocation method for elliptic partial differential equations with random input data. *SIAM Journal on Numerical Analysis*, 45(3):1005–1034, 2007.
- [3] Bäck, J., Nobile, F., Tamellini, L., and Tempone, R. Implementation of optimal Galerkin and collocation approximations of PDEs with random coefficients. *ESAIM: Proc.*, 33:10–21, 2011.
- [4] Marc Berveiller, Bruno Sudret, and Maurice Lemaire. Stochastic finite element: a non intrusive approach by regression. *Eur. J. Comput. Mech.*, 15:81–92, 2006.
- [5] R.H. Cameron and W.T. Martin. The orthogonal development of non-linear functionals in series of Fourier-Hermite functionals. *The Annals of Mathematics*, 48(2):385–392, 1947.

- [6] Mark H. Carpenter, David Gottlieb, and Saul Abarbanel. Time-stable boundary conditions for finite-difference schemes solving hyperbolic systems: methodology and application to high-order compact schemes. *J. Comput. Phys.*, 111(2):220–236, 1994.
- [7] Mark H. Carpenter, Jan Nordström, and David Gottlieb. A stable and conservative interface treatment of arbitrary spatial accuracy. *J. Comput. Phys.*, 148:341–365, January 1999.
- [8] Frank A. Coutelieiris and J. M. P. Q. Delgado. *Transport Processes in Porous Media*. Springer-Verlag, 2012.
- [9] Oliver G. Ernst, Antje Mugler, Hans-Jörg Starkloff, and Elisabeth Ullmann. On the convergence of generalized polynomial chaos expansions. *DFGSP 1324*, Preprint(2):317–339, 2010.
- [10] Roger Ghanem. Stochastic finite elements for heterogeneous media with multiple random non-Gaussian properties. *ASCE J. Eng. Mech.*, 125:26–40, 1997.
- [11] Roger Ghanem and S. Dham. Stochastic finite element analysis for multiphase flow in heterogeneous porous media. *Porous Media*, 32:239–262, 1998.
- [12] Roger G. Ghanem and Pol D. Spanos. *Stochastic finite elements: a spectral approach*. Springer-Verlag, New York, 1991.
- [13] Gene H. Golub and John H. Welsch. Calculation of Gauss quadrature rules. Technical report, Stanford University, Stanford, CA, USA, 1967.
- [14] Gene H. Golub and John H. Welsch. Calculation of Gauss quadrature rules. *Mathematics of Computation*, 23(106):221–230, 1969.
- [15] David Gottlieb and Dongbin Xiu. Galerkin method for wave equations with uncertain coefficients. *Comm. Comput. Phys.*, 3(2):505–518, 2008.
- [16] Bertil Gustafsson, Heinz-Otto Kreiss, and Arne Sundström. Stability theory of difference approximations for mixed initial boundary value problems. II. *Mathematics of Computation*, 26(119):649–686, 1972.
- [17] Heinz-Otto Kreiss and Godela Scherer. Finite element and finite difference methods for hyperbolic partial differential equations. In *Mathematical Aspects of Finite Elements in Partial Differential Equations*, pages 179–183. Academic Press, New York, 1974.
- [18] Olivier P. Le Maître, Omar M. Knio, Matthew Reagan, Habib Najm, and Roger Ghanem. A stochastic projection method for fluid flow. I: Basic formulation. *J. Comp. Phys.*, 173:481–511, 2001.
- [19] Lionel Mathelin and M. Yousuff Hussaini. A stochastic collocation algorithm for uncertainty analysis. Technical Report 2003-212153, NASA Langley Research Center, 2003.
- [20] Lionel Mathelin, M. Yousuff Hussaini, Thomas A. Zang, and Françoise Bataille. Uncertainty propagation for turbulent, compressible flow in a quasi-1d nozzle using stochastic methods. In *AIAA-2003-4240, 16TH AIAA CFD CONFERENCE*, pages 23–26, 2003.

- [21] Ken Mattsson. Summation by parts operators for finite difference approximations of second-derivatives with variable coefficients. *Journal of Scientific Computing*, 51:650–682, 2012.
- [22] Ken Mattsson and Jan Nordström. Summation by parts operators for finite difference approximations of second derivatives. *J. Comput. Phys.*, 199(2):503–540, September 2004.
- [23] Jan Nordström. The influence of open boundary conditions on the convergence to steady state for the Navier-Stokes equations. *J. Comput. Phys.*, 85(1):210–244, 1989.
- [24] Jan Nordström, Sofia Eriksson, and Peter Eliasson. Weak and strong wall boundary procedures and convergence to steady-state of the Navier-Stokes equations. *J. Comput. Phys.*, 231(14):4867–4884, 2012.
- [25] Per Pettersson, Gianluca Iaccarino, and Jan Nordström. Numerical analysis of the Burgers’ equation in the presence of uncertainty. *J. Comput. Phys.*, 228:8394–8412, December 2009.
- [26] Matthew T. Reagan, Habib N. Najm, Roger G. Ghanem, and Omar M. Knio. Uncertainty quantification in reacting-flow simulations through non-intrusive spectral projection. *Combustion and Flame*, 132(3):545–555, 2003.
- [27] Patrick J. Roache. Verification of codes and calculations. *AIAA Journal*, 36(5):696–702, May 1988.
- [28] Lee Shunn, Frank E. Ham, and Parviz Moin. Verification of variable-density flow solvers using manufactured solutions. *J. Comput. Phys.*, 231(9):3801–3827, 2012.
- [29] Benjamin E. Sonday, Robert D. Berry, Habib N. Najm, and Bert J. Deusschere. Eigenvalues of the Jacobian of a Galerkin-projected uncertain ODE system. *SIAM J. Sci. Comput.*, 33:1212–1233, Jun 2011.
- [30] Bo Strand. Summation by parts for finite difference approximations for d/dx . *J. Comput. Phys.*, 110(1):47–67, January 1994.
- [31] Magnus Svärd and Jan Nordström. On the order of accuracy for difference approximations of initial-boundary value problems. *J. Comput. Phys.*, 218(1):333–352, October 2006.
- [32] Raymond S. Tuminaro, Eric T. Phipps, Christopher W. Miller, and Howard C. Elman. Assessment of collocation and Galerkin approaches to linear diffusion equations with random data. *International Journal for Uncertainty Quantification*, 1(1):19–33, 2011.
- [33] Elisabeth Ullmann. *Solution Strategies for Stochastic Finite Element Discretizations*. PhD thesis, Technische Universität Bergakademie Freiberg, Germany, 2008.
- [34] Xiaoliang Wan and George Em Karniadakis. Long-term behavior of polynomial chaos in stochastic flow simulations. *Computer methods in applied mechanics and engineering*, 195:5582–5596, 2006.
- [35] Xiaoliang Wan, Dongbin Xiu, George, and Em Karniadakis. Stochastic solutions for the two-dimensional advection-diffusion equation. *SIAM J. Sci. Computing*, 26:578–590, 2005.

- [36] Norbert Wiener. The homogeneous chaos. *American Journal of Mathematics*, 60(4):897–936, 1938.
- [37] Dongbin Xiu and Jan S. Hesthaven. High-order collocation methods for differential equations with random inputs. *SIAM J. Sci. Comput.*, 27:1118–1139, October 2005.
- [38] Dongbin Xiu and George Em. Karniadakis. Modeling uncertainty in steady state diffusion problems via generalized polynomial chaos. *Comput. Methods Appl. Math. Eng.*, 191:4927–4948, 2002.
- [39] Dongbin Xiu and George Em. Karniadakis. The Wiener–Askey polynomial chaos for stochastic differential equations. *SIAM J. Sci. Comput.*, 24(2):619–644, 2002.
- [40] Dongbin Xiu and George Em. Karniadakis. Modeling uncertainty in flow simulations via generalized polynomial chaos. *J. Comput. Phys.*, 187(1):137–167, May 2003.