SIAM J. SCI. COMPUT.
Vol. 43, No. 2, pp. A929–A952

HYPERBOLICITY-PRESERVING AND WELL-BALANCED STOCHASTIC GALERKIN METHOD FOR SHALLOW WATER EQUATIONS*

DIHAN DAI†, YEKATERINA EPSHTEYN†, AND AKIL NARAYAN†‡

Abstract. A stochastic Galerkin formulation for a stochastic system of balanced or conservation laws may fail to preserve the hyperbolicity of the original system. In this work, we develop a hyperbolicity-preserving stochastic Galerkin formulation for the one-dimensional shallow water equations by carefully selecting the polynomial chaos expansion of the nonlinear q^2/h term in terms of the polynomial chaos expansions of the conserved variables. In addition, in an arbitrary finite stochastic dimension, we establish a sufficient condition to guarantee the hyperbolicity of the stochastic Galerkin system through a finite number of conditions at stochastic quadrature points. Further, we develop a well-balanced central-upwind scheme for the stochastic shallow water model and derive the associated hyperbolicity-preserving CFL-type condition. The performance of the developed method is illustrated on a number of challenging numerical tests.

Key words. finite volume method, stochastic Galerkin method, shallow water equations, hyperbolic systems of conservation law and balance laws

AMS subject classifications. 35L65, 35Q35, 35R60, 65M60, 65M70

DOI. 10.1137/20M1360736

1. Introduction.

The classical one-dimensional deterministic Saint-Venant system of shallow water equations is

$$\begin{aligned} &(h)_t + (q)_x = 0, \\ &(q)_t + \left(\frac{q^2}{h} + \frac{1}{2}gh^2\right)_x = -ghB_x, \end{aligned} \tag{1.1}$$

where $h = h(x,t)$ is the water height, $q = q(x,t)$ is the water discharge, g is the gravitational constant, and $B = B(x)$ is the time-independent bottom topography. This system was first derived in [9] and since then has been widely used in modeling the flows whose horizontal scales are significantly larger than their vertical scales, such as water flows in rivers, lakes, and coastal areas. However, the accuracy and prediction capabilities of shallow water models depend strongly on the presence of various uncertainties that naturally arise in measuring or empirically approximating, e.g., the bottom topography data or initial and boundary conditions. Hence, it is important to consider a stochastic version of the shallow water equations (SWE). In this work, we focus on uncertainty that results in *parameterized* SWE, where parameters are modeled as random variables. In particular, we study the polynomial chaos expansions (PCE) strategy, which is very effective when quantities of interest vary smoothly with respect to the parameters.

*Submitted to the journal's Methods and Algorithms for Scientific Computing section August 18, 2020; accepted for publication (in revised form) December 21, 2020; published electronically March 11, 2021.

https://doi.org/10.1137/20M1360736

Funding: The work of the third author was partially supported by the NSF through grant DMS-1848508.

†Department of Mathematics, University of Utah, Salt Lake City, UT 84112 USA (dai@math.utah.edu, epshteyn@math.utah.edu).

‡Scientific Computing and Imaging (SCI) Institute, University of Utah, Salt Lake City, UT 84112 USA (akil@sci.utah.edu).

There are two widely used classes of methods for addressing uncertainty in (parameterized) partial differential equations using PCE. One class, of *nonintrusive* type methods, computes stochastic quantities by generating an ensemble of solution realizations, each of which may be treated as a deterministic problem. Statistical information is obtained from this ensemble by postprocessing the ensemble solutions. Examples of such methods include Monte Carlo–type methods that use randomly selected samples and stochastic collocation methods that use a priori preselected samples (see, e.g., [42, 31, 29]). Since they rely on multiple queries of existing deterministic solvers, nonintrusive methods are easy to implement and highly parallelizable, but they can result in less accurate approximations than the intrusive-type methods.

The other group of methods are *intrusive* methods. Such methods typically require a substantial rewrite of legacy code and solvers. In the context of PCE methods, the prototypical intrusive strategy is the stochastic Galerkin (SG) approach, wherein one replaces an underlying stochastic process with its truncated PCE [40, 43] and then forms a system of differential equations via Galerkin projection in stochastic space. As a consequence, one derives a new system of partial differential equations whose unknowns are (time- and space-varying) coefficients of the PCE. Intrusive methods are projection-based approximations, and thus their accuracy is near-optimal in an L^2 sense for static problems. Discussion on the existing convergence theory for SG methods can be found, for example, in [2, 27]. SG methods have been successfully employed for modeling uncertainty in diffusion models [44, 12], kinetic equations [17, 37], and conservation and balanced laws with symmetric Jacobian matrices [39].

For hyperbolic systems, such as the SWE, the associated SG system may not be hyperbolic in general [11, 18]. Thus, the intrusive SG formulation can result in a system of differential equations of a class different from the original deterministic system. There are currently several efforts to resolve this issue for more general types of equations and to preserve the hyperbolicity of the SG system. For quasi-linear hyperbolic systems, hyperbolicity can be ensured by multiplying the SG formulation of the system by the left eigenvector matrix of its flux Jacobian matrix [41]. Unfortunately, this transformation results in a nonconservative form and numerical solvers designed for conservative formulations cannot be applied directly. A recent operator-splitting-based approach has been developed for both the Euler equations [8] and the SWE [7], where the original systems are split into hyperbolic subsystems whose SG formulations remain hyperbolic. However, this may still lead to complex eigenvalues due to the mismatch in the hyperbolicity sets of the subsystems [36]. Another strategy to resolve the hyperbolicity issue of SG formulation is to introduce an appropriate change of variables. For example, the SG system of balanced/conservation laws in terms of entropic variables can be shown to be hyperbolic [35, 34]. In addition, an optimization-based method, called the intrusive polynomial moment method (IPMM), was proposed to calculate the PCE of entropic variables given the PCE of the conserved variables [11, 35, 34]. However, the optimization problem in the IPMM that must be solved for each cell and at each time step can be computationally expensive. There are also strategies that employ Roe variable formulations: In [33, 15, 14], the flux of the SG system is constructed using Roe variables and the conservative form of the system is preserved. It has been shown that both the SG formulations of the Euler equation [33] and the SWE [15] in terms of Roe variables are hyperbolic when using a Wiener–Haar expansion. The SG formulation of the isothermal Euler equations in terms of Roe variables is hyperbolic for any basis function under a positive definiteness condition [15]. However, it can still be expensive to implement the Roe formulation since the PCE of Roe variables needs to be calculated by solving both a

nonlinear equation and a linear equation.

The SG formulation of the SWE may not be hyperbolic due to the PCE of the nonlinear, nonpolynomial term q^2/h [11]. This issue can be partially resolved by using the Roe variables and the Wiener–Haar expansion [15, 14]. In this work, we develop hyperbolicity-preserving SG PCE formulation for the SWE by carefully selecting the PCE of the q^2/h term using only the PCE of the conserved variables. Further, we establish a connection between the hyperbolicity of the SG system and the original system. Namely, we show that preserving positivity of the water height at a finite number of stochastic quadrature points is sufficient to preserve the hyperbolicity of the SG formulation of the SWE. In addition, we will present the well-balanced discretization for our SG formulation of SWE, which preserves positivity of the water height at certain quadrature points in the stochastic domain. In this paper, we adopt the filter from [36] to ensure the positivity-preserving property of the algorithm at stochastic quadrature points, which is one ingredient for ensuring hyperbolicity. However, one can go further in filtering. For example, recent work [26] utilizes a more sophisticated Lasso-regression-based filter to reduce oscillations of the numerical solution at shocks in the spatial domain.

In this work, we consider central-upwind scheme as an example of the underlying numerical scheme for the stochastic SWE. However, the main ideas developed in this work are independent of the particular choice of the numerical solver for hyperbolic problems and can be employed with various choices of the numerical schemes for hyperbolic problems. The central Nessyahu–Tadmor schemes and their generalization into higher resolution central schemes and semidiscrete central-upwind schemes are a class of robust Godunov-type Riemann problem-free projection-evolution methods for hyperbolic systems. They were originally developed in [30, 25, 22]. The family of central-upwind schemes has been successfully applied to problems in science and engineering and, in particular, to deterministic SWE and related models. A second-order central-upwind scheme was first extended to SWE in [20]. However, the scheme did not simultaneously satisfy the positivity-preserving and well-balanced properties. It was improved in [23], where the developed method captures the "lake-at-rest" steady state and preserves the positivity of the water height. We refer the interested reader to [24, 21, 5, 6, 28, 19] for examples of other closely related works. The numerical scheme developed in this work is mainly based on further extension to stochastic SWE of the framework proposed in [22, 23].

This paper is organized as follows. In section 2, we introduce the stochastic SWE and the SG discretization of the system using a particular choice of the PCE for q^2/h. In section 3, we discuss the hyperbolicity of the SG system obtained in section 2 and present a sufficient condition to guarantee the hyperbolicity of the SG SWE system. In section 4, we present a well-balanced central-upwind scheme for the SG SWE model and derive a hyperbolicity-preserving CFL-type condition. In section 5, we illustrate the robustness of the developed numerical scheme with several challenging tests.

2. Modeling stochastic SWE.

This section sets up the stochastic SWE problem and introduces notation used in this article.

2.1. Stochastic modeling of the SWE.

We consider a complete probability space $(\Omega, \mathcal{F}, P)$ with event space Ω, σ-algebra $\mathcal{F}$, and probability measure P. For $\omega \in \Omega$, a stochastic version of (1.1) is

$$
\begin{aligned}
&(h(x,t,\omega))_t + (q(x,t,\omega))_x = 0,\\
&(q(x,t,\omega))_t + \left(\frac{q^2(x,t,\omega)}{h(x,t,\omega)} + \frac{1}{2}gh^2(x,t,\omega)\right)_x = -gh(x,t,\omega)B_x(x,\omega),
\end{aligned}
\tag{2.1}
$$

where uncertainty enters the equation through, e.g., a stochastic model of the initial conditions or of the bottom topography B. Here, we present a stochastic model of the bottom topography. However, all our results generalize to other models of uncertainty (e.g., in the initial conditions). We model B as a finite-dimensional random field,

$$B = B(x,\xi) = B_0(x) + \sum_{k=1}^{d} B_k(x)\xi_k,$$

where $\xi = (\xi_1, \ldots, \xi_d)$ is a d-dimensional random variable. Such a model can result, for example, from truncation of an infinite-dimensional Karhunen–Loève decomposition. Under this model, the stochastic SWE model (2.1) can be written as a function of ξ,

$$\begin{aligned} &(h(x,t,\xi))_t + (q(x,t,\xi))_x = 0, \\ &(q(x,t,\xi))_t + \left(\frac{q^2(x,t,\xi)}{h(x,t,\xi)} + \frac{1}{2}gh^2(x,t,\xi)\right)_x = -gh(x,t,\xi)B_x(x,\xi), \end{aligned} \tag{2.2}$$

which, for the purposes of this paper, forms the continuous model problem for which we seek to compute numerical solutions.

2.2. Polynomial chaos expansions. We assume that the random variable ξ has a Lebesgue density $\rho : \mathbb{R}^d \to \mathbb{R}$. Polynomial chaos expansions (PCE) seek to approximate dependence on ξ by a polynomial function of ξ. With $\nu = (\nu_1, \ldots, \nu_d) \in \mathbb{N}_0^d$ a multi-index, then for $\zeta \in \mathbb{R}^d$ we adopt the standard notation,

$$\zeta^\nu \coloneqq \prod_{j=1}^{d} \zeta_j^{\nu_j}, \qquad\qquad \zeta^0 = \zeta^{(0,0,\ldots,0)} = 1.$$

We let $\Lambda \subset \mathbb{N}_0^d$ denote any nonempty, size-K finite set of multi-indices. We will assume throughout that $0 = (0,0,\ldots,0) \in \Lambda$. Our PCE approximations will take place in a polynomial subspace defined by Λ:

$$P_\Lambda = \operatorname{span}\{\zeta^\nu \;\big|\; \nu \in \Lambda\}, \qquad\qquad \dim P_\Lambda = K \coloneqq |\Lambda|.$$

We will also need "powers" of this set, defined by r-fold products of P_Λ elements:

$$P_\Lambda^r \coloneqq \operatorname{span}\left\{\prod_{j=1}^{r} p_j \;\big|\; p_j \in P_\Lambda,\ j = 1,\ldots,r\right\}, \quad \dim P_\Lambda^r \le \left(\!\!\binom{K}{r}\!\!\right) = \binom{K+r-1}{r}, \tag{2.3}$$

where the dimension bound results from a combinatoric argument. Note that since $0 \in \Lambda$, then $P_\Lambda^r \subseteq P_\Lambda^s$ for any $r \le s$. We will later exercise the notation above for $r = 3$. If ρ has finite polynomial moments of all orders, then there is an $L^2_\rho(\mathbb{R}^d)$-orthonormal basis $\{\phi_k\}_{k=1}^\infty$ of P_Λ, i.e.,

$$\langle \phi_k, \phi_\ell \rangle_\rho \coloneqq \int_{\mathbb{R}} \phi_k(s)\phi_\ell(s)\rho(s)ds = \delta_{k\ell}, \qquad\qquad \phi_1(\xi) \equiv 1, \tag{2.4}$$

for all $k, \ell \in \{1,\ldots,K\}$, with the latter identification of ϕ_1 being an assumption we make without loss since $0 \in \Lambda$. If $y(x,t,\cdot) \in L^2_\rho(\mathbb{R})$, then under mild conditions on the probability measure ρ (see [13]) there exists a convergent expansion of y in these basis functions,

$$y(x,t,\cdot) \overset{L^2_\rho}{=} \sum_{k=1}^{\infty} \hat{y}_k(x,t)\phi_k(\cdot),$$

where $\hat{y}_k(x,t)$ are (stochastic) Fourier coefficients in the basis $\{\phi_k\}_{k\in\mathbb{N}}$, and $\{\phi_\ell\}_{\ell>K}$ are any $L^2_\rho(\mathbb{R}^d)$-orthonormal basis for the orthogonal complement of P_Λ in the space of all d-variate polynomials. A K-term P_Λ PCE *approximation* of the stochastic process y is then formed by truncating the summation above to terms in P_Λ:

$$y(x,t,\xi) \approx \sum_{k=1}^{K} \hat{y}_k(x,t)\phi_k(\xi) =: \mathcal{G}_\Lambda[y](x,t,\xi). \tag{2.5}$$

Above, we have defined the linear projection operator $\mathcal{G}_\Lambda : L^2_\rho \to P_\Lambda$.

2.3. Operations on truncated PCE. Polynomial statistics of PCE expansions can be computed from a straightforward manipulation of their coefficients. For example,

$$\mathbb{E}[\mathcal{G}_\Lambda[y](x,t,\xi)] = \hat{y}_1(x,t), \quad \mathrm{Var}[\mathcal{G}_\Lambda[y](x,t,\xi)] = \sum_{k=2}^{K} \hat{y}_k^2(x,t), \tag{2.6}$$

where $\mathbb{E}$ is the expectation operator, and Var is the variance. In contrast, computing PCE of nonlinear expressions is more complicated. To calculate the P_Λ-truncated PCE of the product of two stochastic processes $y(x,t,\xi)$ and $z(x,t,\xi)$, we introduce the notation

$$\mathcal{G}_\Lambda[y,z] \coloneqq \mathcal{G}_\Lambda\left[\mathcal{G}_\Lambda[y]\,\mathcal{G}_\Lambda[z]\right] = \sum_{m=1}^{K}\left(\sum_{k,\ell=1}^{K} \hat{y}_k\hat{z}_\ell\langle\phi_k\phi_\ell,\phi_m\rangle_\rho\right)\phi_m(\xi). \tag{2.7}$$

The approximation above defines the *pseudospectral product*, which is a widely used strategy for computing PCE products (see, e.g., [10, 15]). The pseudospectral product is an exact projection onto P_Λ of the product of two P_Λ projections. Such an operation can be cast in linear algebraic terms by considering vectors comprised of the PCE coefficients. Given $y \in P_\Lambda$, we will hereafter let $\hat{y} \in \mathbb{R}^K$ denote its ϕ_k-expansion coefficients. We now introduce the linear operator $\mathcal{P} : \mathbb{R}^K \to \mathbb{R}^{K\times K}$,

$$\mathcal{P}(\hat{y}) \coloneqq \sum_{k=1}^{K} \hat{y}_k\mathcal{M}_k, \qquad \mathcal{M}_k \in \mathbb{R}^{K\times K}, \qquad (\mathcal{M}_k)_{\ell m} = \langle\phi_k,\phi_\ell\phi_m\rangle_\rho, \tag{2.8}$$

where $\mathcal{M}_k$ is a symmetric matrix for each k. The following properties hold:

$$\mathcal{P}(\hat{y}) = \left(\mathcal{M}_1\hat{y}|\mathcal{M}_2\hat{y}|\cdots|\mathcal{M}_K\hat{y}\right), \quad \mathcal{P}(\hat{y})\hat{z} = \mathcal{P}(\hat{z})\hat{y}, \quad \widehat{\mathcal{G}_\Lambda[y,z]} = \mathcal{P}(\hat{y})\hat{z}, \tag{2.9}$$

where the last property is due to (2.7) and allows us to conclude the following.

LEMMA 2.1. *Let $a(\xi), b(\xi), c(\xi) \in P_\Lambda$ have ϕ_j-expansion coefficients $\hat{a},\hat{b},\hat{c} \in \mathbb{R}^K$, respectively. Then $\langle a, b\,c\rangle_\rho = \hat{a}^T\mathcal{P}(\hat{b})\hat{c}$.*

Proof. Since $a \in P_\Lambda$, then

$$\langle a, b\,c\rangle_\rho = \langle b\,c, a\rangle_\rho = \langle\mathcal{G}_\Lambda[b,c], a\rangle_\rho = \hat{a}^T\widehat{\mathcal{G}_\Lambda[b,c]} \overset{(2.9)}{=} \hat{a}^T\mathcal{P}(\hat{b})\hat{c}. \qquad \square$$

We will also need to compute P_Λ truncations of ratios of processes (when for each (x,t) the denominator is a single-signed process with probability 1). We start by noting the following exact representation when y is a single-signed process:

$$\mathcal{G}_\Lambda\left[y\,\frac{z}{y}\right](x,t,\xi) = \mathcal{G}_\Lambda[z](x,t,\xi). \tag{2.10}$$

We then use this to motivate the assumption

$$\mathcal{G}_\Lambda\left[y, \frac{z}{y}\right] = \mathcal{G}_\Lambda[z] \iff \mathcal{P}(\hat{y})\widehat{\left(\frac{z}{y}\right)} = \hat{z}. \tag{2.11}$$

This expression motivates the following definition for a new operator $\mathcal{G}_\Lambda^\dagger\left[\frac{z}{y}\right]$:

$$\mathcal{G}_\Lambda^\dagger\left[\frac{z}{y}\right](\xi) \coloneqq \sum_{k=1}^{K} c_k \phi_k(\xi), \tag{2.12}$$

where c_i is the ith element of $\widehat{\left(\frac{z}{y}\right)}$ defined by (2.11), assuming $\mathcal{P}(\hat{y})$ is invertible.

2.4. SG formulation for SWE. We start with (2.2) and perform a standard Galerkin procedure in stochastic (ξ) space using polynomials from P_Λ. In other words, the first step is to replace h and q by the ansatz,

$$h \simeq h_\Lambda \coloneqq \sum_{k=1}^{K} \hat{h}_j(x,t)\phi_j(\xi), \qquad q \simeq q_\Lambda \coloneqq \sum_{k=1}^{K} \hat{q}_j(x,t)\phi_j(\xi), \tag{2.13}$$

respectively, and B by $\mathcal{G}_\Lambda[B]$. Following this, we apply the projection operator $\mathcal{G}_\Lambda$ to both sides of (2.2) and insist on equality. However, in addition we make the following crucial assumption about how we approximate the term q^2/h:

$$\frac{q^2}{h} = \frac{q}{h}\, q \quad \longrightarrow \quad \mathcal{G}_\Lambda\left[\frac{q_\Lambda^2}{h_\Lambda}\right] = \mathcal{G}_\Lambda\left[q_\Lambda\, \mathcal{G}_\Lambda^\dagger\left[\frac{q_\Lambda}{h_\Lambda}\right]\right].$$

Performing these steps on (2.2) results in the system

$$\frac{\partial}{\partial t}\begin{pmatrix}\hat{h}\\ \hat{q}\end{pmatrix} + \frac{\partial}{\partial x}\begin{pmatrix}\hat{q}\\ \frac{1}{2}g\mathcal{P}(\hat{h})\hat{h} + \mathcal{P}(\hat{q})\mathcal{P}^{-1}(\hat{h})\hat{q}\end{pmatrix} = \begin{pmatrix}0\\ -g\mathcal{P}(\hat{h})\widehat{B}_x\end{pmatrix}, \tag{2.14}$$

where $\hat{h}$ and $\hat{q}$ are each length-K vectors whose entries are the coefficients introduced in (2.13). With $\hat{U} \coloneqq (\hat{h}, \hat{q})^T$, and the flux and source terms

$$F(\hat{U}) = \begin{pmatrix}\hat{q}\\ \frac{1}{2}g\mathcal{P}(\hat{h})\hat{h} + \mathcal{P}(\hat{q})\mathcal{P}^{-1}(\hat{h})\hat{q}\end{pmatrix}, \qquad S(\hat{U}, \hat{B}) = \begin{pmatrix}0\\ -g\mathcal{P}(\hat{h})\widehat{B}_x\end{pmatrix}, \tag{2.15}$$

then the system (2.14) can be written in general conservation law form,

$$\hat{U}_t + (F(\hat{U}))_x = S(\hat{U}, \hat{B}), \tag{2.16}$$

with flux Jacobian

$$J(\hat{U}) \coloneqq \frac{\partial F}{\partial \hat{U}} = \begin{pmatrix}O & I\\ g\mathcal{P}(\hat{h}) - \mathcal{P}(\hat{q})\mathcal{P}^{-1}(\hat{h})\mathcal{P}(\hat{u}) & \mathcal{P}(\hat{u}) + \mathcal{P}(\hat{q})\mathcal{P}^{-1}(\hat{h})\end{pmatrix}, \tag{2.17}$$

where we have introduced

$$\hat{u} = \mathcal{P}^{-1}(\hat{h})\hat{q}, \tag{2.18}$$

which can be viewed as the PCE coefficient vector of the velocity $u \coloneqq \frac{q}{h}$. The computation that gives the expression (2.17) for the Jacobian uses the property (2.9). For more details, we refer interested readers to section 2.2 of [18].

We emphasize that (h, q) are the (x, t, ξ)-dependent solutions to the original stochastic SWE (2.2), whereas (h_Λ, q_Λ) are the (x, t, ξ)-dependent solutions to our SGSWE (2.16). In general, these two solutions are distinct. We first articulate sufficient conditions under which (2.16) is a well-posed hyperbolic system.

3. Hyperbolicity of the SG system. In this section, we show that the system (2.16) is hyperbolic under the condition that the matrix $\mathcal{P}(\hat{h})$ is positive definite. When there is no uncertainty, this condition reduces to $h > 0$, which ensures hyperbolicity for the deterministic SWE (1.1).

THEOREM 3.1. *If the matrix $\mathcal{P}(\hat{h})$ is strictly positive definite, the SG formulation* (2.16) *is hyperbolic.*

Proof. We will show that the Jacobian $\frac{\partial F}{\partial \hat{U}}$ is diagonalizable with real eigenvalues. Since $\mathcal{P}(\hat{h})$ is positive definite, then define

$$G \coloneqq \sqrt{g\mathcal{P}(\hat{h})}, \qquad A \coloneqq gG^{-1}\mathcal{P}(\hat{q})G^{-1}, \qquad B \coloneqq \mathcal{P}(\hat{u}), \tag{3.1}$$

where $\sqrt{M}$ is the (unique) symmetric positive definite square root of a symmetric positive definite matrix M. Using these matrices, define

$$P_1 \coloneqq \begin{pmatrix} I & I \\ B+G & B-G \end{pmatrix}, \qquad P_1^{-1} = \begin{pmatrix} -\frac{1}{2} \end{pmatrix} \begin{pmatrix} G^{-1}B - I & -G^{-1} \\ -G^{-1}B - I & G^{-1} \end{pmatrix},$$

where the formula for P_1^{-1} can be verified by direct computation. Then a calculation shows that

$$P_1^{-1}\frac{\partial F}{\partial \hat{U}}P_1 = -\frac{1}{2}\begin{pmatrix} -2G - B - A & A - B \\ A - B & 2G - B - A \end{pmatrix}, \tag{3.2}$$

which is symmetric. Thus, $\frac{\partial F}{\partial \hat{U}}$ is similar to a diagonalizable matrix with real eigenvalues, and so is itself real diagonalizable. ∎

Remark 3.2. In the deterministic case, all the PCE coefficients are zero, except possibly the very first coefficient, and the matrix P_1 in (3.2) reduces to the eigenmatrix that symmetrizes the deterministic Jacobian matrix, and the matrix on the right-hand side of (3.2) reduces to a diagonal matrix.

For the deterministic SWE (1.1), the velocity u is bounded between the smallest and the largest eigenvalues of the Jacobian of the deterministic SWE. For the SG formulation (2.14), we have an analogous relation.

PROPOSITION 3.3. *The eigenvalues of the matrix $\mathcal{P}(\hat{u})$ are bounded between the smallest and the largest eigenvalues of the Jacobian matrix $J(\hat{U})$, i.e.,*

$$\lambda_{\max}(J(\hat{U})) \geq \lambda_{\max}\left(\mathcal{P}(\hat{u})\right) \geq \lambda_{\min}\left(\mathcal{P}(\hat{u})\right) \geq \lambda_{\min}(J(\hat{U})). \tag{3.3}$$

Proof. By the proof of Theorem 3.1, the matrix $J(\hat{U})$ is similar to the symmetric matrix $D \coloneqq P_1^{-1}\frac{\partial F}{\partial \hat{U}}P_1$ defined in (3.2). For an arbitrary unit vector $\hat{y} = (\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_K)^{\mathrm{T}} \in \mathbb{R}^K$, then $\hat{z} \coloneqq \frac{1}{\sqrt{2}}[\hat{y}^T, \hat{y}^T]^T \in \mathbb{R}^{2K}$ is also a unit vector. Then

$$\hat{z}^{\mathrm{T}} D \hat{z} = \hat{y}^{\mathrm{T}} \mathcal{P}(\hat{u}) \hat{y}. \tag{3.4}$$

From the above relation, and using properties of the Rayleigh quotient for $\mathcal{P}(\hat{u})$,

$$\lambda_{\max}(\mathcal{P}(\hat{u})) \geq \hat{z}^{\mathrm{T}} D \hat{z} \geq \lambda_{\min}(\mathcal{P}(\hat{u})),$$

where equalities can be achieved by proper selections of $\hat{y}$. Using similar Rayleigh quotient properties for D and noting that $\hat{z}$ ranges over a subset of $\mathbb{R}^{2K}$, then

$$\lambda_{\max}(D) \geq \lambda_{\max}\left(\mathcal{P}(\hat{u})\right) \geq \lambda_{\min}\left(\mathcal{P}(\hat{u})\right) \geq \lambda_{\min}(D). \tag{3.5}$$

The inequalities (3.3) follow since D is similar to $J(\hat{U})$. ∎

In the deterministic SWE, positivity of the water height h ensures the hyperbolicity of the PDE system. Theorem 3.1 shows that the stochastic variant of the positivity condition is that $\mathcal{P}(\hat{h})$ is positive definite. Much of the rest of this paper is devoted to deriving numerical procedures to guarantee this condition.

3.1. Positive definiteness of $\mathcal{P}(\hat{h})$. In this subsection, we present a computationally convenient sufficient condition that guarantees $\mathcal{P}(\hat{h}) > 0$ and hence guarantees hyperbolicity.

THEOREM 3.4. *Given Λ, let nodes ξ_m and weights τ_m satisfying $\{(\xi_m, \tau_m)\}_{m=1}^M \subset \mathbb{R}^d \times (0,\infty)$ represent any M-point positive quadrature rule that is exact on P_Λ^3, i.e.,*

$$\int_{\mathbb{R}^d} p(\xi)\rho(\xi)d\xi = \sum_{m=1}^M p(\xi_m)\tau_m, \qquad p \in P_\Lambda^3. \tag{3.6}$$

If

$$h_\Lambda(x,t,\xi_m) > 0 \quad \forall\, m = 1,\dots,M, \tag{3.7}$$

then the SGSWE system (2.16) *is hyperbolic.*

Proof. We will show that (3.7) implies $\mathcal{P}(\hat{h}) > 0$, which in turn ensures hyperbolicity from Theorem 3.1. Let $\hat{z} = (\hat{z}_k)_{k=1}^K$ be any nontrivial vector in $\mathbb{R}^K$, and define its associated P_Λ polynomial $z(\xi) \coloneqq \sum_{k=1}^K \hat{z}_j \phi_k(\xi) \neq 0$. Then $z(\xi)$ cannot vanish at all quadrature points simultaneously since if it did, we obtain the contradiction,

$$0 \neq \|\hat{z}\|^2 = \langle z, z\rangle_\rho \overset{(3.6)}{=} \sum_{j=1}^M z^2(\xi_j)\tau_j = 0,$$

where we have used the fact that $P_\Lambda^2 \subseteq P_\Lambda^3$ to utilize (3.6). Then, since the quadrature rule is positive and (3.7) holds, we have

$$0 < \sum_{j=1}^M h_\Lambda(x,t,\xi_j) z^2(\xi_j)\tau_j \overset{(3.6)}{=} \left\langle h_\Lambda(x,t,\xi), z^2(\xi)\right\rangle \overset{\text{Lemma 2.1}}{=} \hat{z}^T \mathcal{P}(\hat{h})\hat{z},$$

establishing that $\mathcal{P}(\hat{h})$ is positive definite. ∎

Thus, by guaranteeing the positivity of h_Λ at a finite number of points, we can ensure the hyperbolicity of the SGSWE system. For arbitrary stochastic dimension d and polynomial space P_Λ, there is a worst-case upper bound on the size of this finite set.

COROLLARY 3.5. *There is some $M \leq \dim P_\Lambda^3 \leq \frac{K(K+1)(K+2)}{6}$ such that the discrete pointwise positivity condition* (3.7) *guarantees the hyperbolicity of* (2.16).

We give the proof in Lemma B.2 in Appendix B. One might consider the somewhat simpler condition of restricting $\hat{h}_1 > 0$ for hyperbolicity since $\hat{h}_1$ is the expected value of h_Λ. This condition is actually implied by the condition in Theorem 3.4.

COROLLARY 3.6. *If the conditions of Theorem* 3.4 *are satisfied, then $\hat{h}_1 > 0$.*

Proof. Since $\tau_j > 0$ and $h_\Lambda > 0$ at the quadrature points, then

$$\hat{h}_1 = \int_{\mathbb{R}^d} h_\Lambda(x,t,\zeta)\rho(\zeta)d\zeta = \sum_{j=1}^M h_\Lambda(x,t,\xi_j)\tau_j > 0. \qquad ∎$$

A computable condition ensuring hyperbolicity therefore requires a positive quadrature rule that is exact on P_Λ^3. For general densities ρ over $\mathbb{R}^d$, computing such a quadrature rule is a very difficult task. But this is possible in specialized cases.

For example, if $d = 1$ and $\Lambda = \{0, 1, \dots, K-1\}$, then an optimal choice of positive quadrature is the ρ-Gaussian quadrature. Since $P_\Lambda^3 = \operatorname{span}\{1, \zeta, \dots, \zeta^{3K-3}\}$, then choosing the positive M-point Gaussian quadrature,

$$\{\xi_m\}_{m=1}^M = \phi_{M+1}^{-1}(0), \qquad \tau_m = \frac{1}{\sum_{j=1}^M \phi_j^2(\xi_m)},$$

with $M \geq \left\lceil \frac{3K}{2} \right\rceil - 1$ satisfies the conditions of Theorem 3.4 (and does so with substantially fewer points than the $\sim K^3/6$ worst-case bound from Corollary 3.5). Gaussian quadrature rules have real-valued nodes and positive weights [38].

In spaces with $d > 1$, if ρ is tensorial, then tensorizing Gauss quadrature rules achieves similar results. In other words, assume

$$\rho(\xi) = \prod_{J=1}^d \rho_J(\xi_J), \qquad \xi \in \mathbb{R}^d.$$

We can always enclose P_Λ within a tensor-product polynomial space:

$$P_\Lambda^3 \subseteq P_{3k,\infty} \coloneqq \left\{ \lambda \in \mathbb{N}_0^d \;\middle|\; \lambda_J \leq 3\kappa_J \text{ for } J = 1, \dots, d \right\}, \qquad \kappa_J \coloneqq \max_{\nu \in \Lambda} \nu_J.$$

For a fixed $J \in \{1, \dots, d\}$, let $\{(\xi_{m,M_J}^{(J)}, \tau_{m,M_J}^{(J)})\}_{m=1}^{M_J}$ denote the $M_J \coloneqq (\left\lceil \frac{3\kappa_J}{2} \right\rceil - 1)$-point ρ_J-Gaussian quadrature rule on $\mathbb{R}$. Then the tensorization of these d univariate quadrature rules results in an $M \coloneqq \left(\prod_{J=1}^d M_J\right)$-point positive quadrature rule that is exact on $P_{3k,\infty}$, and hence on P_Λ^3, and thus satisfies the conditions of Theorem 3.4.

4. Numerical scheme for stochastic SWE. In this section, we derive a well-balanced central-upwind scheme that preserves the hyperbolicity of the SG formulation (2.16) at every time step.

4.1. Central-upwind scheme for the SG system. We first introduce the central-upwind scheme for the SG system (2.16). Appendix A provides a brief summary of the second-order central-upwind schemes for balance laws. With $\{\mathcal{C}_i\}_{i=1}^N$ a partition of a bounded closed interval, let $x_{i\pm\frac{1}{2}}$ denote the partition boundaries, and define the cell average of the vector $\hat{U}$ over the ith cell $\mathcal{C}_i =: \left[x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}\right]$ as

$$\overline{\mathbf{U}}_i(t) \coloneqq \begin{pmatrix} \overline{\mathbf{h}}_i(t) \\ \overline{\mathbf{q}}_i(t) \end{pmatrix} \coloneqq \frac{1}{\Delta x} \int_{\mathcal{C}_i} \begin{pmatrix} \hat{h}(x,t) \\ \hat{q}(x,t) \end{pmatrix} dx \in \mathbb{R}^{2K}.$$

We have introduced notation for common quantities in finite volume–type schemes. While $\hat{U}_k$ is the kth component of the vector $\hat{U}$, the bold letter $\mathbf{U}$ with subscripts and superscripts is used here to introduce the cell averages and pointwise reconstructions, respectively, of the vector $\hat{U}(x,t)$. For instance, $\mathbf{U}_{i+\frac{1}{2}}^-$ is the approximated value of $\hat{U}$ at the left-hand side of spatial location $x = x_{i+\frac{1}{2}}$, which is reconstructed from the cell averages $\overline{\mathbf{U}}_i$. A similar reasoning applies to $(\mathbf{h}, \hat{h}, \hat{h}_k)$ and $(\mathbf{q}, \hat{q}, \hat{q}_k)$. To minimize clutter, we will notationally suppress t dependence from here onward. The possible discontinuities of the system (2.16) at the cell interface $x = x_{i+\frac{1}{2}}$, where

$\mathcal{C}_i = \left[x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}\right]$, propagates with left- and right-sided local speeds that can be estimated by

$$
\begin{aligned}
a^-_{i+\frac{1}{2}} &= \min\left\{\lambda_1\left(J(\mathbf{U}^-_{i+\frac{1}{2}})\right), \lambda_1\left(J(\mathbf{U}^+_{i+\frac{1}{2}})\right), 0\right\},\\
a^+_{i+\frac{1}{2}} &= \max\left\{\lambda_{2K}\left(J(\mathbf{U}^-_{i+\frac{1}{2}})\right), \lambda_{2K}\left(J(\mathbf{U}^+_{i+\frac{1}{2}})\right), 0\right\},
\end{aligned} \tag{4.1}
$$

where $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_{2K}$ are the eigenvalues of the $J(\cdot)$ in (2.17), and $\mathbf{U}^-_{i+\frac{1}{2}}$ and $\mathbf{U}^+_{i+\frac{1}{2}}$ are the left- and right-sided pointwise reconstructions in the ith cell. The semidiscrete form of the central-upwind scheme for the SG system (2.16) reads as

$$
\frac{d}{dt}\overline{\mathbf{U}}_i = -\frac{\mathcal{F}_{i+\frac{1}{2}} - \mathcal{F}_{i-\frac{1}{2}}}{\Delta x} + \overline{\mathbf{S}}_i, \qquad \overline{\mathbf{S}}_i \approx \frac{1}{\Delta x}\int_{\mathcal{C}_i} S(\mathbf{U},\mathbf{B})dx, \tag{4.2}
$$

with $\overline{\mathbf{S}}_i$ a well-balanced discretization of the source term, which we discuss below. With F the flux term in (2.15), the numerical flux $\mathcal{F}$ is given by

$$
\mathcal{F}_{i+\frac{1}{2}} \coloneqq \frac{a^+_{i+\frac{1}{2}}F(\mathbf{U}^-_{i+\frac{1}{2}}) - a^-_{i+\frac{1}{2}}F(\mathbf{U}^+_{i+\frac{1}{2}})}{a^+_{i+\frac{1}{2}} - a^-_{i+\frac{1}{2}}} + \frac{a^+_{i+\frac{1}{2}}a^-_{i+\frac{1}{2}}}{a^+_{i+\frac{1}{2}} - a^-_{i+\frac{1}{2}}}\left[\mathbf{U}^+_{i+\frac{1}{2}} - \mathbf{U}^-_{i+\frac{1}{2}}\right]. \tag{4.3}
$$

4.2. Well-balanced property. In applications of the deterministic SWE, simulations should accurately capture the so-called *lake-at-rest* steady state solution or small perturbations of the lake-at-rest steady state. A *well-balanced* numerical scheme for the SWE captures the lake-at-rest solution exactly at discrete level. An analogous lake-at-rest state for the stochastic SWE (2.14) is

$$
q_\Lambda(x,t,\xi) \equiv 0, \quad h_\Lambda + \mathcal{G}_\Lambda[B](x,t,\xi) \equiv C(\xi), \tag{4.4}
$$

where $C(\xi)$ depends only on ξ. This solution corresponds to still water with a flat stochastic water surface. Equation (4.4) can be rewritten in the vector form

$$
\hat{q} \equiv \mathbf{0}, \quad \hat{h} + \hat{B} \equiv \hat{C}. \tag{4.5}
$$

In order to derive a well-balanced central upwind scheme for the SGSWE, we first replace the original bottom function $\hat{B}$ by its continuous linear interpolant. At every time step, we compute the PCE vector for the cell averages of the water surface by $\overline{\mathbf{w}}_i \coloneqq \overline{\mathbf{h}}_i + \overline{\mathbf{B}}_i$ and the pointwise reconstructions of the water surface by $\mathbf{w}^\pm_{i+\frac{1}{2}}$ using a generalized minmod limiter (see Appendix A). The pointwise reconstructions of the water height are then computed by

$$
\mathbf{h}^\pm_{i+\frac{1}{2}} \coloneqq \mathbf{w}^\pm_{i+\frac{1}{2}} - \mathbf{B}_{i+\frac{1}{2}}, \tag{4.6}
$$

where $\mathbf{B}_{i+\frac{1}{2}}$ is the PCE vector for $\mathcal{G}_\Lambda\left[B(x_{i+\frac{1}{2}},t,\xi)\right]$. The numerical fluxes $\{\mathcal{F}_{i+\frac{1}{2}}\}_{i=1}^N$ are subsequently computed using the reconstructed PCE of the water height defined in (4.6). After that, the well-balanced property of the scheme is ensured by a special choice of the source term $\overline{\mathbf{S}}_i$.

LEMMA 4.1. *With $\mathbf{B}_{i\pm\frac{1}{2}}$ the PCE vectors for $\mathcal{G}_\Lambda\left[B(x_{i\pm\frac{1}{2}},t,\xi)\right]$, if we choose*

$$
\overline{\mathbf{S}}_i \coloneqq \begin{pmatrix} \mathbf{0} \\ -\frac{1}{\Delta x} g\mathcal{P}(\overline{\mathbf{h}}_i)\left(\mathbf{B}_{i+\frac{1}{2}} - \mathbf{B}_{i-\frac{1}{2}}\right) \end{pmatrix}, \tag{4.7}
$$

then the central-upwind scheme (4.2) *satisfies the well-balanced property.*

Proof. We have $\overline{\mathbf{B}}_i = (\mathbf{B}_{i+\frac{1}{2}} + \mathbf{B}_{i-\frac{1}{2}})/2$, and the cell average PCE vector of the water surface is $\overline{\mathbf{w}}_i := \overline{\mathbf{h}}_i + \overline{\mathbf{B}}_i$. Let the pointwise reconstructions for water surface be $\mathbf{w}^{\pm}_{i+\frac{1}{2}}$. Assume that at time t, the stochastic water surface is flat and the water is still, i.e., $\overline{\mathbf{w}}_i \equiv \mathbf{w}^*$ is a constant vector for all i, and $\overline{\mathbf{q}}_i \equiv \mathbf{0}$. Then a second-order piecewise linear reconstruction procedure produces $\mathbf{w}^{\pm}_{i+\frac{1}{2}} \equiv \mathbf{w}^*$ and $\mathbf{q}^{\pm}_{i+\frac{1}{2}} \equiv \mathbf{0}$. Hence, the numerical flux defined in (4.3) becomes

$$\mathcal{F}_{i+\frac{1}{2}} = \begin{pmatrix} \mathbf{0} \\ \frac{g}{2}\mathcal{P}(\mathbf{w}^* - \mathbf{B}_{i+\frac{1}{2}})(\mathbf{w}^* - \mathbf{B}_{i+\frac{1}{2}}) \end{pmatrix} =: \begin{pmatrix} \mathcal{F}^{\hat{h}}_{i+\frac{1}{2}} \\ \mathcal{F}^{\hat{q}}_{i+\frac{1}{2}} \end{pmatrix}. \tag{4.8}$$

Then, with $\overline{\mathbf{S}}_i = \left(\overline{\mathbf{S}}^T_{i,1}, \overline{\mathbf{S}}^T_{i,2}\right)^T$, the corresponding semidiscrete form is

$$\begin{aligned} &\frac{d}{dt}\overline{\mathbf{h}}_i = \overline{\mathbf{S}}_{i,1}, \\ &\frac{d}{dt}\overline{\mathbf{q}}_i = -\frac{1}{\Delta x}\frac{g}{2}\left[\mathcal{P}(\mathbf{w}^* - \mathbf{B}_{i+\frac{1}{2}})(\mathbf{w}^* - \mathbf{B}_{i+\frac{1}{2}}) - \mathcal{P}(\mathbf{w}^* - \mathbf{B}_{i-\frac{1}{2}})(\mathbf{w}^* - \mathbf{B}_{i-\frac{1}{2}})\right] + \overline{\mathbf{S}}_{i,2}. \end{aligned} \tag{4.9}$$

To balance these equations, we choose $\overline{\mathbf{S}}_{i,1}$ and $\overline{\mathbf{S}}_{i,2}$ so that the right-hand side vanishes. Clearly, we need $\overline{\mathbf{S}}_{i,1} \equiv \mathbf{0}$. To simplify the computation for $\overline{\mathbf{S}}_{i,2}$, let $\Delta\mathbf{B}_i = \mathbf{B}_{i+\frac{1}{2}} - \mathbf{B}_{i-\frac{1}{2}}$; then $\overline{\mathbf{B}}_i = \mathbf{B}_{i+\frac{1}{2}} - \frac{1}{2}\Delta\mathbf{B}_i = \mathbf{B}_{i-\frac{1}{2}} + \frac{1}{2}\Delta\mathbf{B}_i$. By linearity of the operator $\mathcal{P}$ and the property (2.9),

$$\begin{aligned} \overline{\mathbf{S}}_{i,2} &= \frac{1}{\Delta x}\frac{g}{2}\left[\mathcal{P}(\mathbf{w}^* - \mathbf{B}_{i+\frac{1}{2}})(\mathbf{w}^* - \mathbf{B}_{i+\frac{1}{2}}) - \mathcal{P}(\mathbf{w}^* - \mathbf{B}_{i-\frac{1}{2}})(\mathbf{w}^* - \mathbf{B}_{i-\frac{1}{2}})\right] \\ &= \frac{1}{\Delta x}\frac{g}{2}\left[\mathcal{P}\left(\mathbf{w}^* - \overline{\mathbf{B}}_i - \frac{1}{2}\Delta\mathbf{B}_i\right)\left(\mathbf{w}^* - \overline{\mathbf{B}}_i - \frac{1}{2}\Delta\mathbf{B}_i\right)\right. \\ &\qquad \left. - \mathcal{P}\left(\mathbf{w}^* - \overline{\mathbf{B}}_i + \frac{1}{2}\Delta\mathbf{B}_i\right)\left(\mathbf{w}^* - \overline{\mathbf{B}}_i + \frac{1}{2}\Delta\mathbf{B}_i\right)\right] \\ &= \frac{1}{\Delta x}\frac{g}{2}\left[\mathcal{P}(\mathbf{w}^* - \overline{\mathbf{B}}_i)\left(-\Delta\mathbf{B}_i\right) - \mathcal{P}\left(\frac{\Delta\mathbf{B}_i}{2}\right)\left(2\mathbf{w}^* - 2\overline{\mathbf{B}}_i\right)\right] \\ &= -g\mathcal{P}(\mathbf{w}^* - \overline{\mathbf{B}}_i)\left(\frac{\mathbf{B}_{i+\frac{1}{2}} - \mathbf{B}_{i-\frac{1}{2}}}{\Delta x}\right) = -g\mathcal{P}(\overline{\mathbf{h}}_i)\left(\frac{\mathbf{B}_{i+\frac{1}{2}} - \mathbf{B}_{i-\frac{1}{2}}}{\Delta x}\right). \end{aligned} \tag{4.10}$$

This completes the proof. ∎

In the meantime, (4.7) reduces to the deterministic well-balanced quadrature approximation when there is no uncertainty. The deterministic formula is obtained by applying the midpoint quadrature rule to the cell averages (4.2) with the derivative term $\mathbf{B}_x(x_i)$ approximated by the finite difference $\left(\mathbf{B}_{i+\frac{1}{2}} - \mathbf{B}_{i-\frac{1}{2}}\right)/\Delta x$ [23].

4.3. Hyperbolicity-preserving CFL-type conditions. In order to determine hyperbolicity-preserving CFL-type conditions, we focus on the first K equations in (4.2) which prescribe evolution of $\overline{\mathbf{h}}_i$,

$$\frac{d}{dt}\overline{\mathbf{h}}_i = -\frac{1}{\Delta x}\left[\mathcal{F}^{\hat{h}}_{i+\frac{1}{2}}(t) - \mathcal{F}^{\hat{h}}_{i-\frac{1}{2}}(t)\right], \tag{4.11}$$

where

$$\mathcal{F}^{\hat{h}}_{i+\frac{1}{2}} = \frac{a^+_{i+\frac{1}{2}}\mathbf{q}^-_{i+\frac{1}{2}} - a^-_{i+\frac{1}{2}}\mathbf{q}^+_{i+\frac{1}{2}}}{a^+_{i+\frac{1}{2}} - a^-_{i+\frac{1}{2}}} + \frac{a^+_{i+\frac{1}{2}}a^-_{i+\frac{1}{2}}}{a^+_{i+\frac{1}{2}} - a^-_{i+\frac{1}{2}}}\left[\mathbf{h}^+_{i+\frac{1}{2}} - \mathbf{h}^-_{i+\frac{1}{2}}\right]. \tag{4.12}$$

A fully discrete version of (4.11) computes the unknowns at fixed values of time, t^n, $n \in \mathbb{N}_0$, with $t^n < t^{n+1}$. For example, with $\overline{\mathbf{h}}_i^n$ the numerical approximation to $\overline{\mathbf{h}}_i(t^n)$, and $\Delta t^n := t^{n+1} - t^n$, the forward Euler discretization of (4.11) reads as

$$\overline{\mathbf{h}}_i^{n+1} = \overline{\mathbf{h}}_i^n - \lambda_i^n \left[\mathcal{F}^{\hat{h}}_{i+\frac{1}{2}}(t^n) - \mathcal{F}^{\hat{h}}_{i-\frac{1}{2}}(t^n) \right], \qquad \lambda_i^n := \frac{\Delta t^n}{\Delta x_i}. \tag{4.13}$$

The following CFL condition guarantees the hyperbolicity of the system (4.13) at $t = t^{n+1}$ for all cell averages by enforcing the positivity condition prescribed in Theorem 3.4.

LEMMA 4.2. *Let $\{\xi_j\}_{j=1}^M$ be the nodes of a quadrature rule satisfying the conditions of Theorem* 3.4. *Assume that $\overline{\mathbf{h}}_i^n(\xi_j) > 0$ for $1 \le j \le M$. If Δt^n satisfies*

$$\Delta t^n < \Delta t_h^n := \min_{\substack{1 \le j \le M \\ i}} \left\{ \Delta x_i \left| \frac{(\overline{\mathbf{h}}_i^n)^{\mathrm{T}} \mathbf{\Phi}(\xi_j)}{\left[\mathcal{F}^{\hat{h}}_{i+\frac{1}{2}}(t_n) - \mathcal{F}^{\hat{h}}_{i-\frac{1}{2}}(t_n) \right]^{\mathrm{T}} \mathbf{\Phi}(\xi_j)} \right| \right\}, \tag{4.14}$$

then the flux Jacobian (2.17), $J\big(\overline{\mathbf{U}}_i^{n+1}\big)$, *is diagonalizable with real eigenvalues.*

Proof. Theorem 3.4 guarantees the conclusion if $\overline{\mathbf{h}}_i^{n+1}(\xi_j) > 0$, for $1 \le j \le M$, so we proceed to show this latter property. For each j, the inequality

$$0 < (\overline{\mathbf{h}}_i^{n+1})^{\mathrm{T}} \mathbf{\Phi}(\xi_j) = (\overline{\mathbf{h}}_i^n)^{\mathrm{T}} \mathbf{\Phi}(\xi_j) - \lambda_i^n \left[\mathcal{F}^{\hat{h}}_{i+\frac{1}{2}}(t_n) - \mathcal{F}^{\hat{h}}_{i-\frac{1}{2}}(t_n) \right]^{\mathrm{T}} \mathbf{\Phi}(\xi_j) \tag{4.15}$$

holds if we choose

$$\frac{\Delta t^n}{\Delta x_i} = \lambda_i^n < \min_{1 \le j \le M} \left\{ \left| \frac{(\overline{\mathbf{h}}_i^n)^{\mathrm{T}} \mathbf{\Phi}(\xi_j)}{\left[\mathcal{F}^{\hat{h}}_{i+\frac{1}{2}}(t_n) - \mathcal{F}^{\hat{h}}_{i-\frac{1}{2}}(t_n) \right]^{\mathrm{T}} \mathbf{\Phi}(\xi_j)} \right| \right\}.$$

Multiplying both sides by Δx_i and minimizing over i yields the conclusion. ∎

The condition (4.14) ensures the positivity of the water height, but we also need to adhere to standard wavespeed-based CFL stability conditions. Thus, we will choose

$$\Delta t^n = 0.9 \min \left\{ \Delta t_h^n, \min_i \frac{\Delta x_i}{\max\{a^+_{i+\frac{1}{2}}, -a^-_{i+\frac{1}{2}}\}} \right\}. \tag{4.16}$$

To extend these conditions to hold for higher-order schemes, we use strong stability-preserving Runge–Kutta schemes [16] to solve the semidiscrete system (4.2). The analysis above for the condition (4.14) still holds for this solver since the ODE solver can be written as a convex combination of several forward Euler steps. However, an adaptive time-step control needs to be adopted to determine the time step [6, 19]. The analysis above can also be naturally extended to any other finite volume solvers.

Remark 4.3. The CFL condition (4.14) can be relaxed if the signs of the fluxes are taken into account in the inequality (4.15). In implementation, this can be used to reduce the simulation time.

It is important to note that the CFL-type condition provided above is limited to the cell averages. For the second-order (or higher-order) central-upwind scheme, additional correction is required for the pointwise reconstructions $\mathbf{U}^{\pm}_{i+\frac{1}{2}}$ to ensure the hyperbolicity of (4.13). Similarly, special correction is needed for the near-dry states, where the matrices $\mathcal{P}(\mathbf{h}^{\pm}_{i+\frac{1}{2}})$ are close to singular, to ensure hyperbolicity.

4.3.1. Hyperbolicity-preserving correction to the reconstruction. Assuming $(\overline{\mathbf{h}}_i^n)^{\mathrm{T}}\mathbf{\Phi}(\xi_j) > 0$, we are able to enforce $(\overline{\mathbf{h}}_i^{n+1})^{\mathrm{T}}\mathbf{\Phi}(\xi_j) > 0$ for $j = 1, \dots, M$ under the CFL-type condition (4.16); see Lemma 4.2. However, the one-sided propagation speeds (4.1) in the central-upwind scheme (4.13) are estimated by the eigenvalues of the Jacobian $\frac{\partial F}{\partial \hat{U}}$ using the pointwise values at the cell interfaces. Thus, computation of these wave speeds requires the positivity of the pointwise reconstructions at quadrature points, i.e., $(\mathbf{h}^{\pm}_{i+\frac{1}{2}})^T\mathbf{\Phi}(\xi_j) > 0$, which is not guaranteed by $(\overline{\mathbf{h}}_i^n)^{\mathrm{T}}\mathbf{\Phi}(\xi_j) > 0$. To resolve this problem, we use the filtering strategy proposed in [36] to filter $\mathbf{h}^{\pm}_{i+\frac{1}{2}}$.

Given a polynomial $p_{\hat{y}}(\xi) = \sum_{k=1}^{K} \hat{y}_k \phi_k(\xi)$ with positive moment $\hat{y}_1$, we find the smallest possible weight μ' such that the weighted averages of the polynomial $p_{\hat{y}}(\xi)$ and the moment $\hat{y}_1$ are nonnegative at given quadrature points $\{\xi_j\}_{j=1}^M$, i.e.,

$$\mu'\hat{y}_1 + (1-\mu')p_{\hat{y}}(\xi) \geq 0 \Leftrightarrow \hat{y}_1 + \sum_{k=2}^{K}(1-\mu')\hat{y}_k\phi_k(\xi_j) \geq 0, \quad j = 1, \dots, M, \tag{4.17}$$

and the coefficients of the polynomial are filtered by

$$\hat{\mathsf{y}}_1 = \hat{y}_1, \qquad \hat{\mathsf{y}}_k = (1-\mu)\hat{y}_k, k = 2, \dots, K, \tag{4.18}$$

where $\mu = \min\{\mu' + \delta, 1\}$, and we select $\delta = 10^{-10}$ in our scheme. Hence, the filtered polynomial $p_{\hat{\mathsf{y}}}(\xi) = \sum_{k=1}^{K} \hat{\mathsf{y}}_k \phi(\xi)$ is positive at given quadrature points $\{\xi_j\}_{j=1}^M$. We filter $p_{\hat{y}}(\xi) = \sum_{k=1}^{K} \hat{y}_k \phi_k(\xi)$ and $p_{\hat{z}}(\xi) = \sum_{k=1}^{K} \hat{z}_k \phi_k(\xi)$ simultaneously by calculating the individual filtering parameters $\mu'_{\hat{y}}$ and $\mu'_{\hat{z}}$ for $p_{\hat{y}}(\xi)$ and $p_{\hat{z}}(\xi)$, respectively, through (4.17). Then the simultaneous filtering parameter is set to $\mu = \min\{\mu'_{\hat{y}} + \delta, \mu'_{\hat{z}} + \delta, 1\}$.

We will exercise the filtering strategy (4.17)–(4.18) for pointwise reconstructions. We compute the filtering parameter μ_i^n at time $t = t^n$ for the ith cell for $(\mathbf{h}^{\pm}_{i\mp\frac{1}{2}})^{\mathrm{T}}\mathbf{\Phi}(\xi)$ according to (4.17). The pointwise reconstructions $\mathbf{h}^{\pm}_{i\mp\frac{1}{2}}$ are then filtered by

$$\left(\mathsf{h}^{\pm}_{i\mp\frac{1}{2}}\right)_1 = \left(\mathbf{h}^{\pm}_{i\mp\frac{1}{2}}\right)_1, \quad \left(\mathbf{h}^{\pm}_{i\mp\frac{1}{2}}\right)_k = (1-\mu_i^n)\left(\mathsf{h}^{\pm}_{i\mp\frac{1}{2}}\right)_k, \quad k = 2, \dots, K. \tag{4.19}$$

The corresponding cell average is adjusted accordingly in order to remain consistent:

$$\overline{\mathsf{h}}_i^n = \frac{1}{2}\left(\mathsf{h}^{+}_{i-\frac{1}{2}} + \mathsf{h}^{-}_{i+\frac{1}{2}}\right). \tag{4.20}$$

Remark 4.4. To reduce oscillations in $q_\Lambda(x,t,\xi)$, we can also filter the discharge reconstructions $\mathbf{q}^{\pm}_{i-\frac{1}{2}}$. The corresponding cell average needs to be adjusted similarly to (4.20). In subsection 5.3, when $(\alpha, \beta) = (1, 3)$, we adopt this filtering approach to reduce oscillations in the discharge.

As an alternative to the filtering above, one can use a convex-optimization-based method [4] to enforce the positivity of $(\mathbf{h}^{\pm}_{i\mp\frac{1}{2}})^{\mathrm{T}}\mathbf{\Phi}(\xi)$ at quadrature points $\{\xi_j\}_{j=1}^M$.

4.3.2. Near-dry state correction. When the polynomial $(\overline{\mathbf{h}}_i^n)^{\mathrm{T}}\mathbf{\Phi}(\xi) \sim 0$, two issues related to the dry state may occur. One is that the first moments of the polynomials $(\mathbf{h}^{\pm}_{i\mp\frac{1}{2}})^{\mathrm{T}}\mathbf{\Phi}(\xi)$ may become nonpositive. This can happen even when the system is deterministic [23]. Nonpositive first moments may lead to the failure of the

filtering correction (4.17)–(4.18). In our scheme, we adopt the following correction for nonpositive first moments. Denote the first moments of $\mathbf{h}^{\pm}_{i\mp\frac{1}{2}}$ by $\left(\mathbf{h}^{\pm}_{i\mp\frac{1}{2}}\right)_1$; then

$$\text{if } \left(\mathbf{h}^{\pm}_{i\mp\frac{1}{2}}\right)_1 \le 0, \text{ then take } \mathbf{h}^{\pm}_{i\mp\frac{1}{2}} = \mathbf{0}, \;\; \mathbf{h}^{\mp}_{i\pm\frac{1}{2}} = 2\overline{\mathbf{h}}^n_i. \tag{4.21}$$

Note that this strategy reduces to a similar correction in the central-upwind scheme for the deterministic SWE [23].

Another issue may happen when the matrix $\mathcal{P}(\mathbf{h}^{+}_{i+\frac{1}{2}})$ or $\mathcal{P}(\mathbf{h}^{-}_{i+\frac{1}{2}})$ is ill-conditioned, which may lead to problems with round-off errors when solving the corresponding linear system (2.18). To resolve this issue, we extend to the stochastic model the desingularization process for the deterministic problem [23, 19]. We demonstrate our correction using the matrix $\mathcal{P}(\mathbf{h}^{-}_{i+\frac{1}{2}})$ as an example. Let

$$\mathcal{P}(\mathbf{h}^{-}_{i+\frac{1}{2}}) = Q^{\mathrm{T}}\Pi Q$$

be the eigenvalue decomposition for $\mathcal{P}(\mathbf{h}^{-}_{i+\frac{1}{2}})$, where $\Pi = \operatorname{diag}(\lambda_1, \dots, \lambda_K)$. For $k = 1, \dots, K$ and a given $\epsilon > 0$, define

$$\Pi^{\mathrm{cor}} = \operatorname{diag}(\lambda_1^{\mathrm{cor}}, \dots, \lambda_K^{\mathrm{cor}}), \qquad \lambda_k^{\mathrm{cor}} = \frac{\sqrt{2}\lambda_k}{\sqrt{\lambda_k^4 + \max\{\lambda_k^4, \epsilon^4\}}}. \tag{4.22}$$

In our scheme, we choose $\epsilon = \Delta x$. Then the corrected PCE coefficient vector for the velocity $\mathbf{u}^{-}_{i+\frac{1}{2}}$ is given by

$$\mathbf{u}^{-}_{i+\frac{1}{2}} = Q^{\mathrm{T}}\Pi^{\mathrm{cor}} Q \mathbf{q}^{-}_{i+\frac{1}{2}}. \tag{4.23}$$

For well-conditioned $\mathcal{P}(\mathbf{h}^{-}_{i+\frac{1}{2}})$, the correction (4.23) reduces to the system (2.18), but when $\mathcal{P}(\mathbf{h}^{-}_{i+\frac{1}{2}})$ is near singular, the discharge needs to be recomputed,

$$\mathbf{q}^{-}_{i+\frac{1}{2}} = \mathcal{P}(\mathbf{h}^{-}_{i+\frac{1}{2}})\mathbf{u}^{-}_{i+\frac{1}{2}}, \tag{4.24}$$

in order to keep the scheme consistent.

Remark 4.5. If there is no uncertainty, the correction (4.22)–(4.23) reduces to the deterministic velocity desingularization in [23, 19].

5. Numerical results. In this section, we summarize numerical tests to illustrate the robustness of the proposed schemes for the SGSWE system (2.16) with different uncertainty models and parametric distributions. For simplicity, we consider only one-dimensional stochastic spaces ($d = 1$) associated to a Beta density over $[-1, 1]$,

$$\rho(\xi) \coloneqq \rho^{(\alpha,\beta)}(\xi) = C(\alpha,\beta)(1-\xi)^{\alpha}(1+\xi)^{\beta}, \quad C(\alpha,\beta)^{-1} = 2^{\alpha+\beta+1}\mathcal{B}(\beta+1, \alpha+1),$$

where $\mathcal{B}(\cdot,\cdot)$ is the Beta function, and the parameters $\alpha, \beta > -1$ can be chosen freely and control how mass concentrates at $\xi = 1$ and $\xi = -1$, respectively. In particular, $\alpha = \beta = 0$ corresponds to the uniform distribution on $[-1, 1]$. The numerical examples in the coming sections consist of the following numerical experiments:

- Subsection 5.1: Stochastic bottom topography model comparing the SGSWE solution (2.16) with $K = 9$ and $K = 17$ with the uniform density, $\alpha = \beta = 0$. The results are compared against a $K = 9$ stochastic collocation solution computed with $\mathcal{S} = 100$ stochastic points. The stochastic collocation solution for, e.g., the water height h, is computed via quadrature,

$$h_{SC}(x,t,\xi) \coloneqq \sum_{j=1}^{K} \hat{h}_{SC,j}(x,t)\phi_k(\xi), \quad \hat{h}_{SC,j}(x,t) \coloneqq \sum_{s=1}^{\mathcal{S}} h(x,t,\zeta_s)\phi_j(\zeta_s)z_s,$$

  where $\{\zeta_s, z_s\}_{s=1}^{\mathcal{S}}$ is the $\mathcal{S}$-point ρ-Gaussian quadrature rule, and $h(x,t,\zeta_s)$ is a numerical solution to a deterministic specialization of the SWE (2.2) obtained by setting $\xi = \zeta_s$ and numerically solved using a deterministic central-upwind scheme.
- Subsection 5.2: Stochastic water surface model testing the well-balanced property of the scheme with $\alpha = \beta = 0$.
- Subsection 5.3: Stochastic discontinuous bottom topography model investigating the effects of different values of M used to enforce $\mathcal{P}(\hat{h}) > 0$. This example also investigates different distributions with $(\alpha, \beta) = (3, 1)$ and $(\alpha, \beta) = (1, 3)$.

The parameter θ in the generalized minmod limiter is set to $\theta = 1.3$ for the first two examples and $\theta = 1$ for the third example. The gravitational constant g is set to $g = 1$ for the first two examples and $g = 2$ for the last example. We filter only the water heights h_Λ except in the very last numerical test. In the third numerical example, when $(\alpha, \beta) = (1, 3)$, we filter both the water heights and the discharges of the water. In all examples, the CFL condition we use in our simulation is (4.16). However, we observe that in practice, a relaxed time step $c\Delta t^n (c > 1)$ will not result in loss of hyperbolicity and the plots are similar visually to the results obtained from the condition (4.16). We believe this is because condition (3.7) is only a sufficient but not necessary condition for the hyperbolicity of SGSWE.

Our numerical results will report quantile regions indicating the range of behavior for solutions. These quantile regions are computed empirically by computing the corresponding PCE on 10^5 randomly sampled points from the density ρ on $[-1, 1]$.

For a fixed spatial grid, the computational cost depends on the dimension K of the chosen polynomial subspace P_Λ. In order to compute the propagation speeds (4.1), the eigenvalues of the $2K \times 2K$ Jacobian $J(\mathbf{U})$ matrix must be computed, making this cost increase as K increases. In addition, to preserve hyperbolicity, we need to ensure the positivity of the water height at all the quadrature points for every spatial-temporal point (Theorem 3.1). Therefore, the cost for preserving the hyperbolicity is at most of order $O(K^3)$ per cell per time step (Corollary 3.5). These relations are formally independent of the dimension d of the stochastic space, but in practice K can grow considerably as d is increased. For example, one may choose P_Λ to be the space of the polynomials with degree up to L. In this case, $K = \binom{L+d}{d}$. When $L \geq d$, as d increases, K increases and also therefore does the computational cost. In this paper, we only consider numerically the case $d = 1$. We plan to investigate higher dimensional stochastic space in a future work. However, note that the developed theory in sections 2 and 3 extends to $d > 1$.

5.1. Stochastic bottom topography. We consider the shallow water system with deterministic initial conditions

$$w(x,0) = \begin{cases} 1, & x < 0, \\ 0.5, & x > 0, \end{cases} \quad q(x,0) = 0, \tag{5.1}$$

and with a stochastic bottom topography

$$B(x,\xi) = \begin{cases} 0.125(\cos(5\pi x)+2) + 0.125\xi, & |x| < 0.2, \\ 0.125 + 0.125\xi & \text{otherwise.} \end{cases} \tag{5.2}$$

In this example, we model ξ as a uniform random variable ($\alpha = \beta = 0$). The corresponding orthonormal basis functions ϕ_j are the orthonormal Legendre polynomials on $[-1,1]$ with density $\rho(\xi) = \frac{1}{2}$. Initially, the highest possible bottom barely touches the initial water surface at $x = 0.5$. In Figures 1 and 2, we use a uniform grid size Δx over the physical domain $x \in [-1,1]$ and compute up to terminal time $t = 0.8$. We present the numerical solutions for $K = 9$ and $K = 17$ using $M = 17$ and $M = 33$-point Gaussian quadrature nodes, respectively, to enforce the positivity condition (3.7).

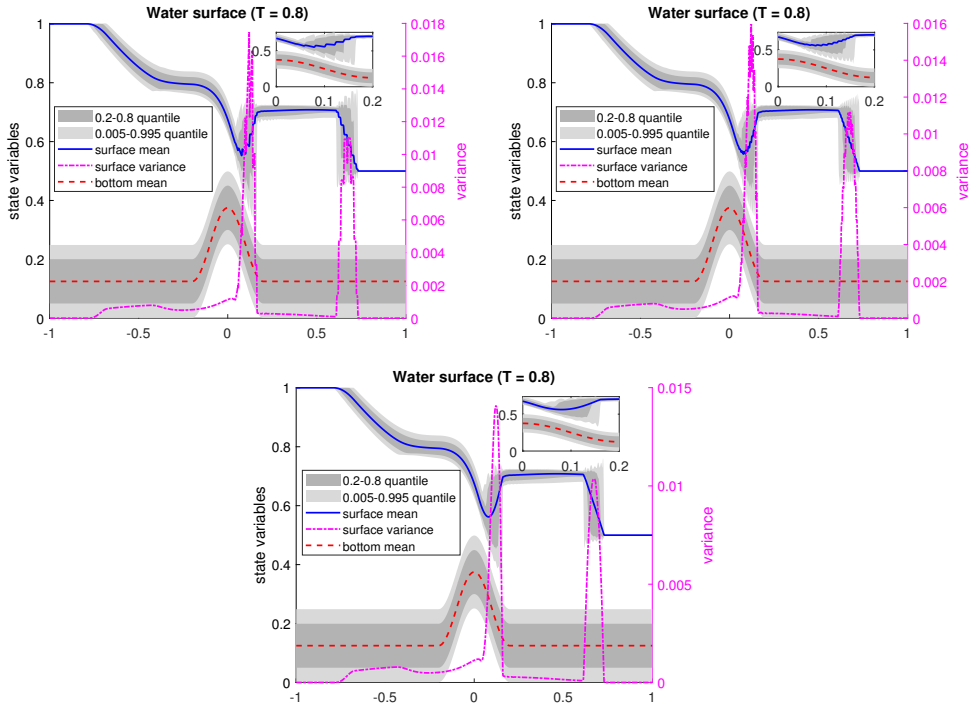

FIG. 1. *Results for subsection* 5.1, *water surfaces. Top left: SG,* $K = 9$, $\Delta x = 1/800$. *Top right: SG,* $K = 17$, $\Delta x = 1/800$. *Bottom: SC,* $K = 9$, $\Delta x = 1/800$.

The 99% confidence region of the water surface stays above the 99% confidence region of the bottom function in the first three (top left, top right, bottom) subfigures in Figure 1.

For reference and comparison, a solution obtained by the stochastic collocation method (100 quadrature points, $K = 9$-term PCE as explained in section 5) is computed. Results for water surface and discharge are shown in the bottom subfigures of Figures 1 and 2, respectively. We note that the stochastic collocation (SC) solution is a different PDE model, so we do not necessarily expect the numerical results from the SG and SC solvers to be identical for a fixed, finite K. In particular, we do not expect "convergence" of one model to the other as, say, $\mathcal{S} \uparrow \infty$ and/or $\Delta x \downarrow 0$. However,

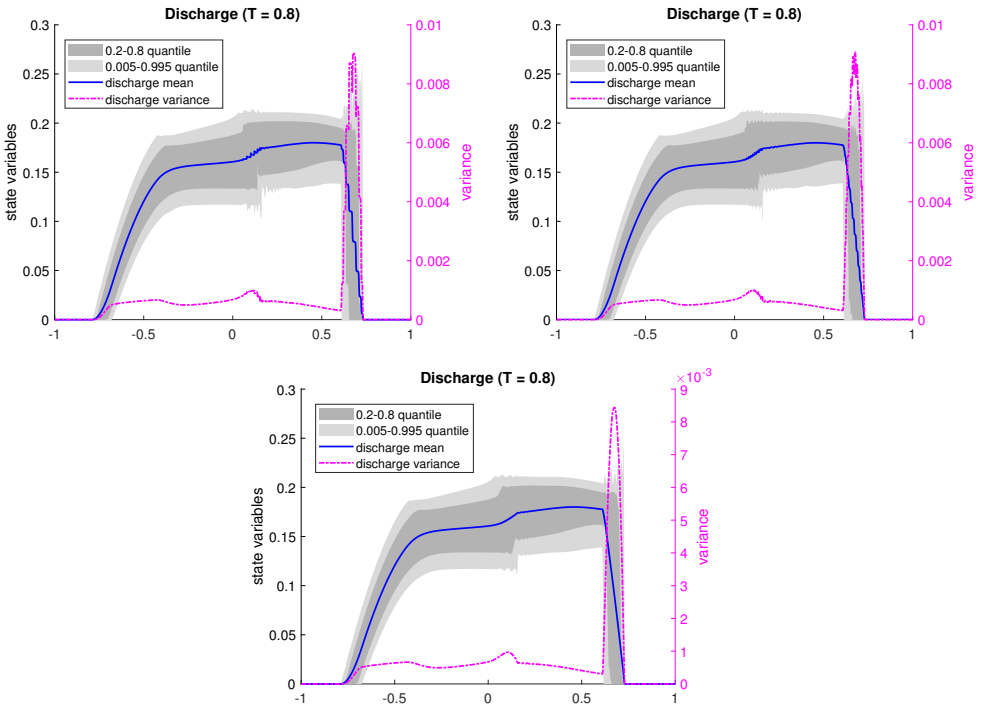

FIG. 2. *Results for subsection* 5.1, *discharges. Top left: SG,* $K = 9$, $\Delta x = 1/800$. *Top right: SG,* $K = 17$, $\Delta x = 1/800$. *Bottom: SC,* $K = 9$, $\Delta x = 1/800$.

the results in the figures do show substantial similarity between these solutions. The numerical solution obtained from the collocation method is less oscillatory near sharp gradients of water surface and discharges.

We observe small oscillations near sharp gradients of the water surface and discharge in all of the figures. We investigate the oscillations for the discharge more carefully in Figure 3. We observe that both higher resolution and larger K can reduce the magnitude of the oscillations that appear in quantiles.

5.2. Stochastic water surface. Consider a stochastic shallow water system with a deterministic bottom function,

$$B(x,\xi) = \begin{cases} 10(x-0.3), & 0.3 \le x \le 0.4, \\ 1 - 0.0025\sin^2(25(\pi(x-0.4))), & 0.4 \le x \le 0.6, \\ -10(x-0.7), & 0.6 \le x \le 0.7, \\ 0 & \text{otherwise}, \end{cases} \tag{5.3}$$

and a stochastic water surface,

$$w(x,0,\xi) = \begin{cases} 1.001 + 0.001\xi, & 0.1 < x < 0.2, \\ 1 & \text{otherwise}, \end{cases} \qquad q(x,0,\xi) \equiv 0. \tag{5.4}$$

We again model ξ as a uniform random variable ($\alpha = \beta = 0$) with $K = 9$. A small uncertain region was originally at $0.1 \le x \le 0.2$, where the water surface is slightly

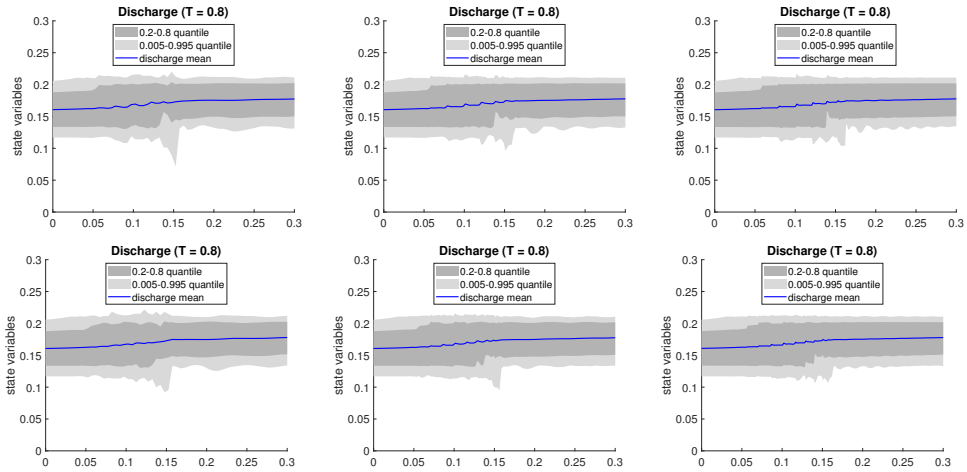

FIG. 3. *Results for subsection* 5.1*, discharges on* $[0, 0.3]$ *for different values of* K *and* Δx*, zoom view. Top:* $K = 9$*; bottom:* $K = 13$*. Left:* $\Delta x = 1/200$*; middle:* $\Delta x = 1/400$*; right* $\Delta x = 1/800$.

perturbed. The 17-point ρ-Gaussian quadrature rule is used to enforce the condition (3.7) to guarantee hyperbolicity. We compute the cell averages of the vector of PCE coefficients for water surface and discharges at terminal time $t = 1.0$ on the physical domain $[-1, 1]$ with uniform grid size $\Delta x = 1/400$. We observe from the top right of Figure 4 that the perturbed water surface with uncertainties propagates along different directions. The right-moving wave interacts with the nonflat bottom and gets partially reflected. The magnitude of the uncertainties doesn't seem to exceed the magnitude of the initial uncertainties, which illustrates the well-balanced property of our scheme.

5.3. Stochastic discontinuous bottom. For our last example, consider the shallow water system with deterministic initial conditions,

$$w(x, 0, \xi) = \begin{cases} 5.0, & x \le 0.5, \\ 1.6, & x > 0.5, \end{cases} \qquad u(x, 0, \xi) = \begin{cases} 1.0, & x \le 0.5, \\ -2.0, & x > 0.5, \end{cases} \tag{5.5}$$

and a stochastic discontinuous bottom,

$$B(x, \xi) = \begin{cases} 1.5 + 0.1\xi, & x \le 0.5, \\ 1.1 + 0.1\xi, & x > 0.5, \end{cases} \tag{5.6}$$

where initially we model ξ as a random variable with Beta density defined by $(\alpha, \beta) = (3, 1)$, which is more concentrated toward $\xi = -1$, and hence the bottom topography has a higher probability of having smaller values. At time $t = 0$, the highest possible bottom barely touches the initial water height at $x = 0.5$. We compute the numerical solutions of a $K = 9$-term PCE with an $M = 17$-point ρ-Gaussian quadrature to enforce the condition (3.7). We compute on a physical domain $x \in [0, 1]$ with uniform cell size $\Delta x = 1/400$ up to terminal time $t = 0.15$.

In this example, we observe over- and undershoots in the neighborhood of the bottom discontinuity for both the water surface w and the discharge q (see Figure 5). This phenomenon also occurs in deterministic version of (5.5)–(5.6) when numerical

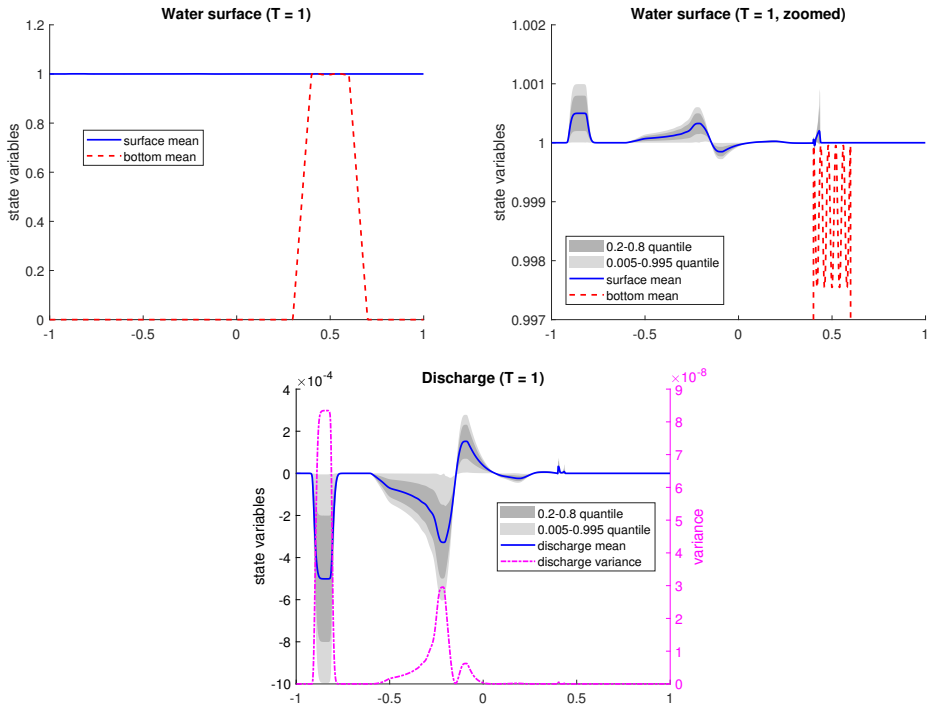

FIG. 4. *Results for subsection* 5.2*: water surface (top left), zoomed water surface (top right), and discharge (bottom) at* $t = 1$ *for* (5.3)–(5.4)*,* $K = 9$.

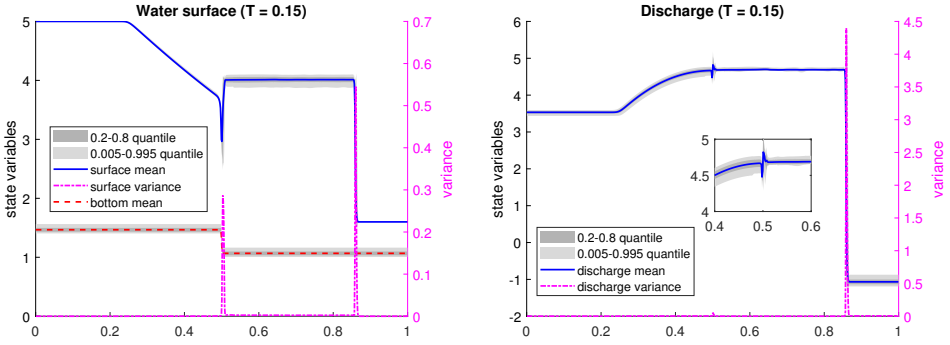

FIG. 5. *Results for subsection* 5.3*:* $K = 9$*,* $t = 0.15$*,* $(\alpha, \beta) = (3, 1)$*. Left figure: water surface and bottom. Right figure: discharge.*

solutions are computed using the schemes from [1, 32]. In addition, we observe in this example a numerical artifact resulting from our enforcement of the positivity of the water height (3.7) at only a finite number of points: although the 99% quantile region of water heights lies above 0, the ξ-global minimum of the water height in some cells can still be negative. Since $\mathcal{P}(\hat{h}) > 0$ only requires the positivity of h_Λ at a finite number of points, there are (low-probability) regions of the domain where the height can be negative. Note, however, that the SGSWE system is still hyperbolic and simulation can continue, despite the low probability of negative water height.

Nevertheless, the existence of negative water heights imposes doubts about the applicability of the SGSWE model. Fortunately, this situation can be mitigated by increasing the number of points M where positivity of h_Λ is enforced. We observe that if the positivity of the water height is enforced at more points, the stochastic region of negative height shrinks. We demonstrate this with results in Table 1. In particular, we observe that (a) the negative region occurs on a subinterval containing ξ values greater than the maximum quadrature point, and (b) the probability of ξ lying in this region is quite small.

In a separate experiment, we also compute the numerical results when ξ is modeled as random according to a $(\alpha, \beta) = (1, 3)$ distribution, which is more concentrated toward $\xi = 1$. Figure 6 shows that at the terminal time the "pressure" from the stochastic bottom that skews positively causes more oscillations on the water surface and the discharge compared to Figure 5. In this experiment, we filter both the water heights and the discharges.

Appendix A. The semidiscrete second-order central-upwind scheme. We briefly describe the central-upwind schemes for one-dimensional balance laws. For a complete description and derivation, we refer the reader to [22]. Consider the balance law,

$$\mathbf{U}_t + (F(\mathbf{U}))_x = S(\mathbf{U}). \tag{A.1}$$

For a uniform mesh with cells $\mathcal{C}_i := \left[x_{i-1/2}, x_{i+1/2}\right]$ of size $|\mathcal{C}_i| \equiv \Delta x$, centered at $x_i = (x_{i-1/2} + x_{i+1/2})/2$, assume that at a certain time level, the cell averages

$$\overline{\mathbf{U}}_i^n \approx \frac{1}{\Delta x}\int_{\mathcal{C}_i} \mathbf{U}(x, t^n)dx \tag{A.2}$$

are available. The cell averages are then used to construct a nonoscillatory second-order linear piecewise reconstruction,

$$\widetilde{\mathbf{U}}_i^n(x) = \overline{\mathbf{U}}_i^n + (\mathbf{U}_x)_i(x - x_i), \quad x \in \mathcal{C}_i, \tag{A.3}$$

whose slopes $(\mathbf{U}_x)_i$ are obtained by generalized minmod limiter,

$$(\mathbf{U}_x)_i = \text{minmod}\left(\theta\frac{\overline{\mathbf{U}}_{i+1}^n - \overline{\mathbf{U}}_i^n}{\Delta x}, \frac{\overline{\mathbf{U}}_{i+1}^n - \overline{\mathbf{U}}_{i-1}^n}{2\Delta x}, \theta\frac{\overline{\mathbf{U}}_i^n - \overline{\mathbf{U}}_{i-1}^n}{\Delta x}\right), \tag{A.4}$$

where the minmod function is defined to be

$$\text{minmod}(z_1, z_2, \dots) := \begin{cases} \min\{z_1, z_2, \dots\} & \text{if } z_i > 0 \ \forall i, \\ \max\{z_1, z_2, \dots\} & \text{if } z_i < 0 \ \forall i, \\ 0 & \text{otherwise,} \end{cases}$$

and the parameter $\theta \in [1, 2]$ controls the amount of numerical dissipation. The left- and right-sided reconstructions at the endpoints of $\mathcal{C}_i$ are

$$\mathbf{U}^+_{i-\frac{1}{2}} = \overline{\mathbf{U}}_i^n - \frac{\Delta x}{2}(\mathbf{U}_x)_i, \quad \mathbf{U}^-_{i+\frac{1}{2}} = \overline{\mathbf{U}}_i^n + \frac{\Delta x}{2}(\mathbf{U}_x)_i. \tag{A.5}$$

The semidiscrete form of the central-upwind scheme is then given by

$$\frac{d}{dt}\overline{\mathbf{U}}_i(t) = -\frac{\mathcal{F}_{i+\frac{1}{2}} - \mathcal{F}_{i-\frac{1}{2}}}{\Delta x} + \overline{\mathbf{S}}_i, \tag{A.6}$$

TABLE 1
Numerical study of the ξ-region and associated probabilities where the water height is negative.

M	$\max_m \xi_m$	Negative region N_M	$\Pr[\xi \in N_M]$
15	0.934077	$[0.934079, 1]$	5.75×10^{-6}
17	0.946839	$[0.946899, 1]$	2.43×10^{-6}
19	0.956205	$[0.956320, 1]$	1.12×10^{-6}
21	0.963310	$[0.963980, 1]$	5.18×10^{-7}

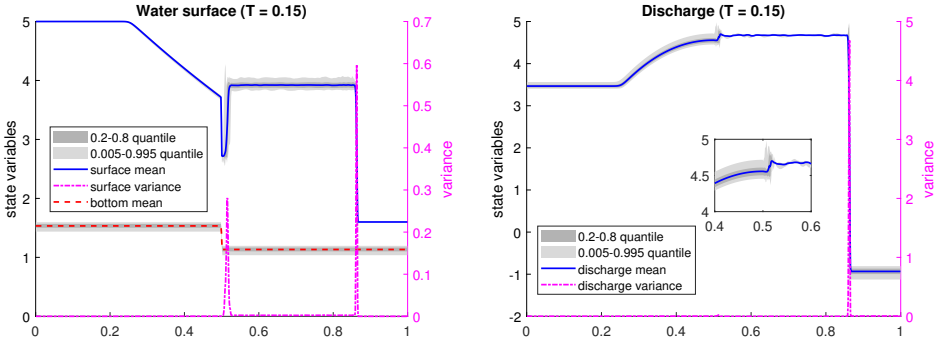

FIG. 6. *Numerical results with $(\alpha, \beta) = (1, 3)$, $K = 9$, $t = 0.15$. Left figure: water surface and bottom. Right figure: discharge.*

where the numerical flux $\mathcal{F}$ and the source term $\overline{\mathbf{S}}_i$ are given in (4.3) and (4.2), respectively.

Appendix B. Proof of Corollary 3.5. The corollary is immediate from the following lemma.

LEMMA B.1. *For some $M \leq \dim P_\Lambda^3$, there is an M-point positive quadrature rule that is exact on P_Λ^3.*

The veracity of this lemma immediately yields $M \leq \dim P_\Lambda^3$ in Corollary 3.5. The second bound in that corollary results from chaining this with the dimension bound in (2.3). Thus, we need only prove the above lemma, which in turn is a simple consequence of Tchakaloff's theorem.

LEMMA B.2 (Tchakaloff's theorem [3]). *Let $P_{T,\ell}$ denote the space of polynomials of degree up to ℓ on $\mathbb{R}^d$:*

$$P_{T,\ell} := \operatorname{span}\left\{ \zeta^\nu \;\middle|\; \sum_{J=1}^{d} \nu_J \leq \ell \right\}.$$

Then, for some $M \leq \dim P_{T,\ell}$, there exists a set of quadrature nodes $\{\zeta_m\}_{m=1}^M$ and positive weights $\{\tau_m\}_{m=1}^M$ such that

$$\int_{\mathbb{R}^d} p(\zeta)\rho(\zeta)d\zeta = \sum_{m=1}^{M} p(\zeta_m)\tau_m, \qquad p \in P_{T,\ell}.$$

Now given P_Λ^3, let ℓ^* denote the maximum polynomial degree of any element in

P_Λ^3:

$$\ell^* \coloneqq \sup_{p \in P_\Lambda^3} \deg p = \max_{k=1,\dots,K} \deg \phi_k,$$

which is finite. Then clearly we have $P_\Lambda^3 \subseteq P_{T,\ell^*}$. By Lemma B.2, there is some $M^* \leq \dim P_{T,\ell^*}$ such that $\{\zeta_m^*\}_{m=1}^{M^*}$ and $\{\tau_m^*\}_{m=1}^{M^*}$ are nodes and (positive) weights, respectively, corresponding to a quadrature rule that is exact on P_Λ (since it's exact on the larger set P_{T,ℓ^*}). Note that if $M^* \leq \dim P_\Lambda^3 =: Q$, then the result of Lemma B.1 is immediate, so we assume otherwise. Let $\{\psi_k\}_{k=1}^Q$ denote any basis for P_Λ^3, and define

$$\mathbf{\Psi}(\zeta) \coloneqq [\psi_1(\zeta), \ \ \psi_2(\zeta), \dots, \psi_Q(\zeta)]^T \in \mathbb{R}^Q.$$

Then exactness of the quadrature rule on P_Λ^3 implies the vector-valued equality,

$$\sum_{m=1}^{M^*} \tau_m^* \mathbf{\Psi}(\zeta_m^*) = \mathbf{e}, \qquad\qquad (e)_k \coloneqq \int_{\mathbb{R}^d} \psi_k(\zeta)\rho(\zeta)d\zeta.$$

In other words, $\mathbf{e} \in \mathbb{R}^Q$ lies in the convex hull of $\{\mathbf{\Psi}(\zeta_m^*)\}_{m=1}^{M^*}$. By Carathéodory's theorem, there must be a size-Q subset of nodes $\{\zeta_m\}_{m=1}^Q \subset \{\zeta_m^*\}_{m=1}^{M^*}$, with positive weights $\{\tau_m\}_{m=1}^Q$, such that $\sum_{m=1}^Q \tau_m \mathbf{\Psi}(\zeta_m) = \mathbf{e}$, which proves Lemma B.1.

REFERENCES

[1] E. Audusse, F. Bouchut, M.-O. Bristeau, R. Klein, and B. Perthame, *A fast and stable well-balanced scheme with hydrostatic reconstruction for shallow water flows*, SIAM J. Sci. Comput., 25 (2004), pp. 2050–2065, https://doi.org/10.1137/S1064827503431090.

[2] I. Babuška, R. Tempone, and G. E. Zouraris, *Galerkin finite element approximations of stochastic elliptic partial differential equations*, SIAM J. Numer. Anal., 42 (2004), pp. 800–825, https://doi.org/10.1137/S0036142902418680.

[3] C. Bayer and J. Teichmann, *The proof of Tchakaloff's theorem*, Proc. Amer. Math. Soc., 134 (2006), pp. 3035–3040.

[4] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, UK, 2004.

[5] S. Bryson, Y. Epshteyn, A. Kurganov, and G. Petrova, *Well-balanced positivity preserving central-upwind scheme on triangular grids for the Saint-Venant system*, ESAIM Math. Model. Numer. Anal., 45 (2011), pp. 423–446.

[6] A. Chertock, S. Cui, A. Kurganov, and T. Wu, *Well-balanced positivity preserving central-upwind scheme for the shallow water system with friction terms*, Internat. J. Numer. Methods Fluids, 78 (2015), pp. 355–383.

[7] A. Chertock, S. Jin, and A. Kurganov, *A Well-Balanced Operator Splitting Based Stochastic Galerkin Method for the One-Dimensional Saint-Venant System with Uncertainty*, preprint, 2015.

[8] A. Chertock, S. Jin, and A. Kurganov, *An Operator Splitting Based Stochastic Galerkin Method for the One-Dimensional Compressible Euler Equations with Uncertainty*, preprint, 2015.

[9] A. J.-C. B. de Saint-Venant, *Théorie du mouvement non-permanent des eaux, avec application aux crues des rivières et à l'introduction des marées dans leur lit*, C. R. Acad. Sci. Paris, 73 (1871), pp. 148–154, 237–240.

[10] B. J. Debusschere, H. N. Najm, P. P. Pébay, O. M. Knio, R. G. Ghanem, and O. P. Le Maître, *Numerical challenges in the use of polynomial chaos representations for stochastic processes*, SIAM J. Sci. Comput., 26 (2004), pp. 698–719, https://doi.org/10.1137/S1064827503427741.

[11] B. Després, G. Poëtte, and D. Lucor, *Robust uncertainty propagation in systems of conservation laws with the entropy closure method*, in Uncertainty Quantification in Computational Fluid Dynamics, Springer, Heidelberg, 2013, pp. 105–149.

[12] M. Eigel, C. J. Gittelson, C. Schwab, and E. Zander, *Adaptive stochastic Galerkin FEM*, Comput. Methods Appl. Mech. Engrg., 270 (2014), pp. 247–269.
[13] O. G. Ernst, A. Mugler, H.-J. Starkloff, and E. Ullmann, *On the convergence of generalized polynomial chaos expansions*, ESAIM Math. Model. Numer. Anal., 46 (2012), pp. 317–339.
[14] S. Gerster and M. Herty, *Entropies and symmetrization of hyperbolic stochastic Galerkin formulations*, Commun. Comput. Phys., 27 (2020), pp. 639–671.
[15] S. Gerster, M. Herty, and A. Sikstel, *Hyperbolic stochastic Galerkin formulation for the p-system*, J. Comput. Phys., 395 (2019), pp. 186–204.
[16] S. Gottlieb, C.-W. Shu, and E. Tadmor, *Strong stability-preserving high-order time discretization methods*, SIAM Rev., 43 (2001), pp. 89–112, https://doi.org/10.1137/S003614450036757X.
[17] J. Hu and S. Jin, *A stochastic Galerkin method for the Boltzmann equation with uncertainty*, J. Comput. Phys., 315 (2016), pp. 150–168.
[18] S. Jin and R. Shu, *A study of hyperbolicity of kinetic stochastic Galerkin system for the isentropic Euler equations with uncertainty*, Chin. Ann. Math. Ser. B, 40 (2019), pp. 765–780.
[19] A. Kurganov, *Finite-volume schemes for shallow-water equations*, Acta Numer., 27 (2018), pp. 289–351.
[20] A. Kurganov and D. Levy, *Central-upwind schemes for the Saint-Venant system*, ESAIM Math. Model. Numer. Anal., 36 (2002), pp. 397–425.
[21] A. Kurganov and C.-T. Lin, *On the reduction of numerical dissipation in central-upwind schemes*, Commun. Comput. Phys, 2 (2007), pp. 141–163.
[22] A. Kurganov, S. Noelle, and G. Petrova, *Semidiscrete central-upwind schemes for hyperbolic conservation laws and Hamilton–Jacobi equations*, SIAM J. Sci. Comput., 23 (2001), pp. 707–740, https://doi.org/10.1137/S1064827500373413.
[23] A. Kurganov and G. Petrova, *A second-order well-balanced positivity preserving central-upwind scheme for the Saint-Venant system*, Commun. Math. Sci., 5 (2007), pp. 133–160.
[24] A. Kurganov, G. Petrova, and B. Popov, *Adaptive semidiscrete central-upwind schemes for nonconvex hyperbolic conservation laws*, SIAM J. Sci. Comput., 29 (2007), pp. 2381–2401, https://doi.org/10.1137/040614189.
[25] A. Kurganov and E. Tadmor, *New high-resolution central schemes for nonlinear conservation laws and convection–diffusion equations*, J. Comput. Phys., 160 (2000), pp. 241–282.
[26] J. Kusch, R. G. McClarren, and M. Frank, *Filtered stochastic Galerkin methods for hyperbolic equations*, J. Comput. Phys., 403 (2020), 109073.
[27] O. Le Maître and O. M. Knio, *Spectral Methods for Uncertainty Quantification: With Applications to Computational Fluid Dynamics*, Springer, Dordrecht, The Netherlands, 2010.
[28] X. Liu, J. Albright, Y. Epshteyn, and A. Kurganov, *Well-balanced positivity preserving central-upwind scheme with a novel wet/dry reconstruction on triangular grids for the Saint-Venant system*, J. Comput. Phys., 374 (2018), pp. 213–236.
[29] S. Mishra, Ch. Schwab, and J. Šukys, *Multilevel Monte Carlo finite volume methods for shallow water equations with uncertain topography in multi-dimensions*, SIAM J. Sci. Comput., 34 (2012), pp. B761–B784, https://doi.org/10.1137/110857295.
[30] H. Nessyahu and E. Tadmor, *Non-oscillatory central differencing for hyperbolic conservation laws*, J. Comput. Phys., 87 (1990), pp. 408–463.
[31] F. Nobile, R. Tempone, and C. G. Webster, *A sparse grid stochastic collocation method for partial differential equations with random input data*, SIAM J. Numer. Anal., 46 (2008), pp. 2309–2345, https://doi.org/10.1137/060663660.
[32] B. Perthame and C. Simeoni, *A kinetic scheme for the Saint-Venant system with a source term*, Calcolo, 38 (2001), pp. 201–231.
[33] P. Pettersson, G. Iaccarino, and J. Nordström, *A stochastic Galerkin method for the Euler equations with Roe variable transformation*, J. Comput. Phys., 257 (2014), pp. 481–500.
[34] G. Poëtte, *Contribution to the Mathematical and Numerical Analysis of Uncertain Systems of Conservation Laws and of the Linear and Nonlinear Boltzmann Equation*, Ph.D. thesis, 2019.
[35] G. Poëtte, B. Després, and D. Lucor, *Uncertainty quantification for systems of conservation laws*, J. Comput. Phys., 228 (2009), pp. 2443–2467.
[36] L. Schlachter and F. Schneider, *A hyperbolicity-preserving stochastic Galerkin approximation for uncertain hyperbolic systems of equations*, J. Comput. Phys., 375 (2018), pp. 80–98.
[37] R. Shu, J. Hu, and S. Jin, *A stochastic Galerkin method for the Boltzmann equation with multi-dimensional random inputs using sparse wavelet bases*, Numer. Math. Theory Meth-

ods Appl., 10 (2017), pp. 465–488.
[38] G. Szegö, *Orthogonal Polynomials*, 4th ed., American Mathematical Society, Providence, RI, 1975.
[39] J. Tryoen, O. Le Maître, M. Ndjinga, and A. Ern, *Intrusive Galerkin methods with upwinding for uncertain nonlinear hyperbolic systems*, J. Comput. Phys., 229 (2010), pp. 6485–6511.
[40] N. Wiener, *The homogeneous chaos*, Amer. J. Math., 60 (1938), pp. 897–936.
[41] K. Wu, H. Tang, and D. Xiu, *A stochastic Galerkin method for first-order quasilinear hyperbolic systems with uncertainty*, J. Comput. Phys., 345 (2017), pp. 224–244.
[42] D. Xiu and J. S. Hesthaven, *High-order collocation methods for differential equations with random inputs*, SIAM J. Sci. Comput., 27 (2005), pp. 1118–1139, https://doi.org/10.1137/040615201.
[43] D. Xiu and G. E. Karniadakis, *The Wiener–Askey polynomial chaos for stochastic differential equations*, SIAM J. Sci. Comput., 24 (2002), pp. 619–644, https://doi.org/10.1137/S1064827501387826.
[44] D. Xiu and J. Shen, *Efficient stochastic Galerkin methods for random diffusion equations*, J. Comput. Phys., 228 (2009), pp. 266–281.