
ISSN: 2008-1898



Journal of Nonlinear Sciences and Applications



Journal Homepage: www.isr-publications.com/jnsa

On the convergence of adaptive gPC for non-linear random difference equations: Theoretical analysis and some practical recommendations



J. Calatayud, J.-C. Cortés*, M. Jornet

Instituto Universitario de Matemática Multidisciplinar, Universitat Politècnica de València, Valencia, Spain.

Communicated by R. Saadati

Abstract

In this paper, the application of adaptive generalized polynomial chaos (gPC) to quantify the uncertainty for non-linear random difference equations is analyzed. It is proved in detail that, under certain assumptions, the stochastic Galerkin projection technique converges algebraically in mean square to the solution process of the random recursive equation. The effect of the numerical errors on the convergence is also studied. A full numerical experiment illustrates our theoretical findings and gives useful insights to reduce the accumulation of numerical errors in practice.

Keywords: Adaptive gPC, stochastic Galerkin projection technique, non-linear random difference equations, uncertainty quantification, numerical analysis.

2010 MSC: 65Q10, 60H35, 39A50, 93E03.

©2018 All rights reserved.

1. Introduction

The generalized polynomial chaos (gPC) technique [11, 12] and general Galerkin methods have been widely used to quantify computationally the uncertainty for classical and fractional continuous stochastic systems (random ordinary and partial differential equations) [7–9, 13]. From the numerical experiments, it is usually stated that the convergence of the stochastic Galerkin projection is exponential or algebraic (spectral convergence). Some theoretical analyses have been carried out to prove this rate of convergence in continuous models [5, 10, 11]. To the best of our knowledge, the theoretical convergence analysis of the stochastic Galerkin projection in random difference equations has not been done. Moreover, we will work with a slight modification of the classical gPC, called adaptive gPC [2, 3]. By using a Gram-Schmidt procedure from the canonical bases of the random inputs, orthogonal polynomials not belonging to the

*Corresponding author

doi: 10.22436/jnsa.011.09.06

Received: 2018-04-23 Revised: 2018-05-30 Accepted: 2018-05-31

Email addresses: jucagre@alumni.uv.es (J. Calatayud), jccortes@imm.upv.es (J.-C. Cortés), marjorsa@doctor.upv.es (M. Jornet)

Askey-Wiener scheme are constructed. This gives more generality to the distributions of the random inputs. We will establish conditions under which the stochastic Galerkin projection constructed via adaptive gPC converges algebraically to the solution process of the random difference equation in the mean square sense. In addition, an analysis of how numerical errors could invalidate the theoretical results in practical experiments is performed.

The structure of the paper is the following. In Section 2, we describe the method of adaptive gPC and the application of the stochastic Galerkin projection technique to non-linear random difference equations. In Section 3, we prove a theoretical result on the algebraic convergence of the stochastic Galerkin projection to the solution process of the non-linear random difference equation. In addition, an analysis of how numerical errors could invalidate the convergence is done. In Section 4, a full illustrative example is included to support the theoretical results. Finally, in Section 5 conclusions are drawn.

2. Adaptive gPC

In this section we will revise the computational method of adaptive gPC from [2]. Using [5], we will recall when adaptive gPC converges for square integrable random vectors. Finally, we will formulate the problem of uncertainty quantification for non-linear random difference equations via the stochastic Galerkin projection technique.

2.1. Method

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be an underlying complete probability space, where Ω is the set of all outcomes ω equipped with a σ -algebra of events \mathcal{F} and a probability measure \mathbb{P} . We will work in the Hilbert space $(L^2(\Omega), \langle, \rangle)$ of random vectors $\mathfrak{u} : \Omega \to \mathbb{R}^q$ with finite second order moment: $\|\mathfrak{u}\|_{L^2(\Omega)} = \sqrt{\mathbb{E}[|\mathfrak{u}|^2]} < \infty$, where \mathbb{E} stands for the expectation operator.

Assume u can be written as a function of random input parameters: $u = g(\zeta_1, ..., \zeta_s)$, where $\zeta_1, ..., \zeta_s : \Omega \to \mathbb{R}$ are random variables and $g : \mathcal{D}_{(\zeta_1,...,\zeta_s)} \subseteq \mathbb{R}^s \to \mathbb{R}^q$ is a Borel measurable function on the support $\mathcal{D}_{(\zeta_1,...,\zeta_s)}$ of the random vector $\zeta = (\zeta_1, ..., \zeta_s)$. It is assumed that $\zeta_1, ..., \zeta_s$ are independent, absolutely continuous (not necessarily identically distributed), and with finite moments. Denote by f_{ζ_i} the density function of ζ_i , $1 \leq i \leq s$. Let $f_{\zeta} = f_{\zeta_1} \times \cdots \times f_{\zeta_s}$ be the joint density function of ζ .

We describe the procedure from [2] to computationally approximate u by means of orthonormal polynomials evaluated at ζ . Let $C_i^p = \{1, \zeta_i, \dots, \zeta_i^p\}$ be the canonical basis of the polynomials in ζ_i up to degree p. Using a Gram-Schmidt procedure, we obtain an orthonormal basis $\Xi_i^p = \{\varphi_0^i(\zeta_i), \varphi_1^i(\zeta_i), \dots, \varphi_p^i(\zeta_i)\}$ with respect to \langle , \rangle , for $1 \leq i \leq s$. The simple tensor product from [2] generates an orthonormal basis $\Xi^P = \{\varphi_1(\zeta), \dots, \varphi_p(\zeta)\}$ with respect to \langle , \rangle , where $\varphi_1 = 1$ and P = (p+s)!/(p!s!). If we increase p and P up to infinity, we obtain an orthonormal sequence of polynomials, $\{\varphi_i(\zeta)\}_{i=1}^{\infty}$, with respect to \langle , \rangle . Formally, $u = \sum_{i=1}^{\infty} \tilde{u}_i \varphi_i(\zeta)$ in $L^2(\Omega)$, where the Fourier coefficients are given by $\tilde{u}_i = \mathbb{E}[u\varphi_i(\zeta)]$.

In [5, Th. 3.6], it is proved that, if the moment problem is uniquely solvable for each ζ_i , $1 \le i \le s$, then the convergence of the series holds in $L^2(\Omega)$. For the sake of completeness, it is said that the moment problem is uniquely solvable if the distribution is determinate by the sequence of its moments. By [5, Th. 3.4], bounded random variables have a uniquely solvable moment problem.

2.2. Application to non-linear random difference equations

Consider a system of non-linear difference equations

$$\mathfrak{u}(\mathfrak{m}+1) = \mathsf{R}(\zeta_1, \dots, \zeta_s, \mathfrak{u}(\mathfrak{m})), \tag{2.1}$$

where $u(m) : \Omega \to \mathbb{R}^q$ is a random vector for each step m (i.e., a time-discrete stochastic process) and $R : \mathcal{D}_{(\zeta_1,...,\zeta_s)} \times \mathbb{R}^q \to \mathbb{R}^q$ is a Borel measurable function. The initial condition is $u(0) = u_0$, where $u_0 \in \mathbb{R}^q$ is a deterministic vector.

We describe a method based on adaptive gPC to intuitively approximate u(m) in $L^2(\Omega)$ for each

 $\mathfrak{m} \ge 0$. Formally,

$$\mathfrak{u}(\mathfrak{m}) = \sum_{\mathfrak{i}=1}^{\infty} \tilde{\mathfrak{u}}_{\mathfrak{i}}(\mathfrak{m}) \varphi_{\mathfrak{i}}(\zeta), \qquad (2.2)$$

where $\tilde{u}_i(m) = \mathbb{E}[u(m)\phi_i(\zeta)]$. Thus, for a fixed P, we look for a solution to (2.1) of the form

$$\hat{\mathbf{u}}^{\mathbf{P}}(\mathbf{m}) = \sum_{i=1}^{\mathbf{P}} \hat{\mathbf{u}}_{i}^{\mathbf{P}}(\mathbf{m}) \phi_{i}(\zeta), \qquad (2.3)$$

which, intuitively, should approximate u(m) in an $L^2(\Omega)$ sense. We impose $\hat{u}^P(m+1) = R(\zeta, \hat{u}^P(m))$, that is, $\sum_{i=1}^{P} \hat{u}_i^P(m+1)\phi_i(\zeta) = R\left(\zeta, \sum_{i=1}^{P} \hat{u}_i^P(m)\phi_i(\zeta)\right)$. We perform the inner product with each $\phi_k(\zeta)$, k = 1, ..., P, and use the orthonormality of Ξ^P (stochastic Galerkin projection technique):

$$\hat{\mathbf{u}}_{k}^{P}(m+1) = \langle \mathsf{R}\left(\zeta, \sum_{i=1}^{P} \hat{\mathbf{u}}_{i}^{P}(m)\phi_{i}(\zeta)\right), \phi_{k}(\zeta) \rangle.$$
(2.4)

This gives a deterministic system of difference equations for $\{\hat{u}_i^P(m)\}_{i=1}^P$, which can be solved numerically. Thus, the approximation $\hat{u}^P(m)$ is numerically computable.

3. Convergence of adaptive gPC for non-linear random difference equations

In this section, we will establish hypotheses on (2.1) under which (2.3) converges to u(m) when $P \rightarrow \infty$ in the mean square sense. We will also see how the numerical errors in the computer may affect the theoretical convergence.

3.1. Theoretical convergence

Let $u = g(\zeta_1, ..., \zeta_s)$ be the random vector from Section 2. Let $\zeta_1, ..., \zeta_s$ be independent and absolutely continuous random variables with finite moments, not necessarily identically distributed. We will assume that $\zeta_1, ..., \zeta_s$ are bounded, with bounded density function f_{ζ_i} . Let $\tilde{u}^P = \sum_{i=1}^{P} \tilde{u}_i \phi_i(\zeta)$, where \tilde{u}_i is the Fourier coefficient defined by $\mathbb{E}[u\phi_i(\zeta)]$. As discussed in Section 2, $\tilde{u}^P \to u$ as $P \to \infty$ in $L^2(\Omega)$. In what follows, we analyze the rate of convergence of \tilde{u}^P to u depending on the smoothness of g, by using some ideas from [4, Lemma 2.1]. Concerning notation, we will denote by L^2_w the weighted L^2 space. The polynomials from Ξ^P are orthonormal in $L^2_{f_z}(\mathbb{R}^s)$.

Let $J \subseteq \mathbb{R}$ be a compact interval that contains the support of ζ_1, \ldots, ζ_s . Let $I = J^s$ be the multidimensional compact rectangle that contains the support of ζ . Consider the sequence of Legendre polynomials $\{L_i\}_{i=1}^{\infty}$ on J, where the degree of L_i is i-1. These polynomials are orthogonal in $L^2(J)$ (the weight is 1). Define the sequence of multivariate Legendre polynomials $\{\Psi_i\}_{i=1}^{\infty}$ on I by using a simple tensor product, which is orthogonal in $L^2(I)$.

Since $\tilde{u}^P \to u$ as $P \to \infty$ in $L^2(\Omega)$, we have $g(x) = \sum_{i=1}^{\infty} \tilde{u}_i \phi_i(x)$ in $L^2_{f_{\zeta}}(I)$. On the other hand, if $g \in L^2(I)$, then $g(x) = \sum_{i=1}^{\infty} \bar{u}_i \psi_i(x)$ in $L^2(I)$, where $\bar{u}_i = \int_I g(x) \psi_i(x) dx / \|\psi_i\|^2_{L^2(I)}$ is the Fourier coefficient associated to the multivariate Legendre basis.

Let $\tilde{\pi}_P g(x) = \sum_{i=1}^P \tilde{u}_i \phi_i(x)$, $\tilde{\pi}_P^{\perp} g(x) = \sum_{i=P+1}^\infty \tilde{u}_i \phi_i(x)$ (with convergence in $L^2_{f_{\zeta}}(I)$), $\bar{\pi}_P g(x) = \sum_{i=1}^P \tilde{u}_i \psi_i(x)$, and $\bar{\pi}_P^{\perp} g(x) = \sum_{i=P+1}^\infty \tilde{u}_i \psi_i(x)$ (with convergence in $L^2(I)$). We have $g = \tilde{\pi}_P g + \tilde{\pi}_P^{\perp} g = \tilde{\pi}_P g + \tilde{\pi}_P^{\perp} g$. Notice that $\|\cdot\|_{L^2_{f_{\zeta}}(I)} \leq C \|\cdot\|_{L^2(I)}$, for certain C > 0, by the assumed boundedness of f_{ζ} . Therefore,

$$C^{2}\|\bar{\pi}_{P}^{\perp}g\|_{L^{2}(I)}^{2} \geq \|\bar{\pi}_{P}^{\perp}g\|_{L_{f_{\zeta}}(I)}^{2} = \|(\tilde{\pi}_{P}g - \bar{\pi}_{P}g) + \tilde{\pi}_{P}^{\perp}g\|_{L_{f_{\zeta}}(I)}^{2} = \|\tilde{\pi}_{P}g - \bar{\pi}_{P}g\|_{L_{f_{\zeta}}(I)}^{2} + \|\tilde{\pi}_{P}^{\perp}g\|_{L_{f_{\zeta}}(I)}^{2} \geq \|\tilde{\pi}_{P}^{\perp}g\|_{L_{f_{\zeta}}(I)}^{2}$$

where Pythagoras's Theorem has been applied, because $\bar{\pi}_P g - \bar{\pi}_P g$ is a multivariate polynomial of degree P - 1, and $\phi_{P+1}, \phi_{P+2}, \ldots$ are orthogonal to the space of multivariate polynomials of degree P - 1 in

 $L^2_{f_{\zeta}}(I)$. By [1, Th. 2.3], if $g \in H^{\sigma}(I)$ (i.e., g is in the Sobolev space of functions in $L^2(I)$ whose weak partial derivatives of order σ belong to $L^2(I)$), then $\|\bar{\pi}_P^{\perp}g\|_{L^2(I)} \leq C_{\sigma}P^{-\sigma}\|g\|_{H^{\sigma}(I)}$. As a consequence, if $g \in H^{\sigma}(I)$, then $\|u - \tilde{u}^P\|_{L^2(\Omega)} = \|\tilde{\pi}_P^{\perp}g\|_{L_{f_{\zeta}}(I)} \leq C \|\bar{\pi}_P^{\perp}g\|_{L^2(I)} \leq C_{\sigma}P^{-\sigma}\|g\|_{H^{\sigma}(I)}$. In particular, if $g \in C^{\infty}(I)$, then, for every $\sigma \ge 0$, there is a constant $C_{\sigma} > 0$ such that

$$\|\mathbf{u} - \tilde{\mathbf{u}}^{\mathsf{P}}\|_{\mathsf{L}^{2}(\Omega)} \leqslant \mathsf{C}_{\sigma}\mathsf{P}^{-\sigma} \tag{3.1}$$

for all $P \ge 1$, i.e., there is algebraic convergence.

Consider the non-linear system of random difference equations (2.1). Take the series from adaptive gPC (2.2). Denote the partial sums of the series in (2.2) by $\tilde{u}^{P}(m) = \sum_{i=1}^{P} \tilde{u}_{i}(m)\phi_{i}(\zeta)$. Consider the Galerkin projection (2.3). The goal is to demonstrate that (2.3) converges to u(m) when $P \to \infty$, under certain assumptions. Some ideas for the proof are taken from [10].

Theorem 3.1. Consider the system of random difference equations $u(m+1) = R(\zeta, u(m))$, where $\zeta = (\zeta_1, ..., \zeta_s)$: $\Omega \to \mathbb{R}^s$ is a random vector, $u(m) : \Omega \to \mathbb{R}^q$ is a random vector, R is a Borel measurable function, and the initial condition $u(0) = u_0$ is a deterministic constant in \mathbb{R}^q . Let $\zeta_1, ..., \zeta_s$ be independent and absolutely continuous random variables with finite moments, not necessarily identically distributed. Assume that $\zeta_1, ..., \zeta_s$ are bounded, with bounded density function f_{ζ_i} . Let $J \subseteq \mathbb{R}$ be a compact interval that contains the support of $\zeta_1, ..., \zeta_s$. Let $I = J^s$ be the multidimensional compact rectangle that contains the support of ζ . Let g_m be the solution of the system of difference equations for u(m), i.e., $u(m) = g_m(\zeta)$. Assume that $g_m \in C^{\infty}(I)$. Suppose that R is Lipschitz on $I \times \mathbb{R}^q$: there exists a constant K > 0 such that $\|R(\zeta, v_1) - R(\zeta, v_2)\|_{L^2(\Omega)} \leq K \|v_1 - v_2\|_{L^2(\Omega)}$ for every pair of random vectors $v_1, v_2 \in L^2(\Omega)$.

Then the Galerkin projection $\hat{u}^{P}(m)$ defined by (2.3) converges to u(m) as $P \to \infty$ in $L^{2}(\Omega)$. Moreover, the convergence is algebraic.

Proof. We have that (2.2) holds in $L^{2}(\Omega)$, and we put this equality in the recursive equation (2.1):

$$\sum_{i=1}^{\infty} \tilde{u}_{i}(m+1)\varphi_{i}(\zeta) = R\left(\zeta, \sum_{i=1}^{\infty} \tilde{u}_{i}(m)\varphi_{i}(\zeta)\right).$$

Fix $P \ge 1$. Performing the inner product with each $\phi_k(\zeta)$, k = 1, ..., P, and using the orthonormality of $\{\phi_i(\zeta)\}_{i=1}^{\infty}$, we obtain $\tilde{u}_k(m+1) = \langle R(\zeta, \sum_{i=1}^{\infty} \tilde{u}_i(m)\phi_i(\zeta)), \phi_k(\zeta) \rangle$. From (2.4), we derive that

$$\tilde{\mathbf{u}}_{k}(\mathfrak{m}+1) - \hat{\mathbf{u}}_{k}^{\mathsf{P}}(\mathfrak{m}+1) = \langle \mathsf{R}\left(\zeta, \sum_{i=1}^{\infty} \tilde{\mathbf{u}}_{i}(\mathfrak{m})\phi_{i}(\zeta)\right) - \mathsf{R}(\zeta, \hat{\mathbf{u}}^{\mathsf{P}}(\mathfrak{m})), \phi_{k}(\zeta) \rangle.$$

By Cauchy-Schwarz inequality, orthonormality of $\{\phi_k(\zeta)\}_{k=1}^p$, the Lipschitz condition, the triangular inequality, and the well-known inequality $(a + b)^2 \leq 2(a^2 + b^2)$, we derive

$$\begin{split} (\tilde{u}_{k}(\mathfrak{m}+1)-\hat{u}_{k}^{P}(\mathfrak{m}+1))^{2} &\leqslant \left\| R\left(\zeta,\sum_{i=1}^{\infty}\tilde{u}_{i}(\mathfrak{m})\varphi_{i}(\zeta)\right) - R(\zeta,\hat{u}^{P}(\mathfrak{m})) \right\|_{L^{2}(\Omega)}^{2} \leqslant K^{2} \left\| \sum_{i=1}^{\infty}\tilde{u}_{i}(\mathfrak{m})\varphi_{i}(\zeta) - \hat{u}^{P}(\mathfrak{m}) \right\|_{L^{2}(\Omega)}^{2} \\ &\leqslant K^{2} \left(\left\| \tilde{u}^{P}(\mathfrak{m}) - \hat{u}^{P}(\mathfrak{m}) \right\|_{L^{2}(\Omega)}^{2} + \left\| \sum_{i=P+1}^{\infty}\tilde{u}_{i}(\mathfrak{m})\varphi_{i}(\zeta) \right\|_{L^{2}(\Omega)}^{2} \right)^{2} \\ &\leqslant 2K^{2} \left\| \tilde{u}^{P}(\mathfrak{m}) - \hat{u}^{P}(\mathfrak{m}) \right\|_{L^{2}(\Omega)}^{2} + 2K^{2} \left\| \sum_{i=P+1}^{\infty}\tilde{u}_{i}(\mathfrak{m})\varphi_{i}(\zeta) \right\|_{L^{2}(\Omega)}^{2}. \end{split}$$

Summing at both sides for $1 \leq k \leq P$ and considering that

$$\sum_{k=1}^{P} (\tilde{u}_{k}(m+1) - \hat{u}_{k}^{P}(m+1))^{2} = \|\tilde{u}^{P}(m+1) - \hat{u}^{P}(m+1)\|_{L^{2}(\Omega)}^{2}$$

by orthonormality of $\{\phi_i(\zeta)\}_{i=1}^{P}$ and Pythagoras's Theorem, we deduce that

$$\left\|\tilde{u}^{P}(\mathfrak{m}+1)-\hat{u}^{P}(\mathfrak{m}+1)\right\|_{L^{2}(\Omega)}^{2} \leqslant 2\mathsf{P}\mathsf{K}^{2}\left\|\tilde{u}^{P}(\mathfrak{m})-\hat{u}^{P}(\mathfrak{m})\right\|_{L^{2}(\Omega)}^{2}+2\mathsf{P}\mathsf{K}^{2}\left\|\sum_{\mathfrak{i}=\mathsf{P}+1}^{\infty}\tilde{u}_{\mathfrak{i}}(\mathfrak{m})\varphi_{\mathfrak{i}}(\zeta)\right\|_{L^{2}(\Omega)}^{2}$$

This may be viewed as a recurrence (with inequality) for $\|\tilde{u}^{P}(m) - \hat{u}^{P}(m)\|_{L^{2}(\Omega)}^{2}$, with initial condition given by $\|\tilde{u}^{P}(0) - \hat{u}^{P}(0)\|_{L^{2}(\Omega)}^{2} = \|u_{0} - u_{0}\|_{L^{2}(\Omega)}^{2} = 0$. Going backwards in the recursive expression, one gets that

$$\begin{split} \|\tilde{u}^{P}(m+1) - \hat{u}^{P}(m+1)\|_{L^{2}(\Omega)}^{2} &\leq \sum_{j=0}^{m} \left\| \sum_{i=P+1}^{\infty} \tilde{u}_{i}(m-j)\varphi_{i}(\zeta) \right\|_{L^{2}(\Omega)}^{2} (2PK^{2})^{j+1} \\ &= \sum_{j=0}^{m} \left\| u(m-j) - \tilde{u}^{P}(m-j) \right\|_{L^{2}(\Omega)}^{2} (2PK^{2})^{j+1} \end{split}$$

Let $\sigma > (m+1)/2$. Since $g_m \in C^{\infty}(I)$, by (3.1) there exists a constant $C_{m,\sigma} > 0$ such that $\|u(m-j) - \tilde{u}^P(m-j)\|_{L^2(\Omega)} \leq C_{m,\sigma}P^{-\sigma}$ for all $j = 0, \ldots, m$. Thereby,

$$\|\tilde{\mathfrak{u}}^{\mathsf{P}}(\mathfrak{m}+1) - \hat{\mathfrak{u}}^{\mathsf{P}}(\mathfrak{m}+1)\|_{L^{2}(\Omega)}^{2} \leqslant \sum_{j=0}^{\mathfrak{m}} \frac{C_{\mathfrak{m},\sigma}^{2}}{\mathsf{P}^{2\sigma}} (2\mathsf{P}\mathsf{K}^{2})^{j+1} \leqslant (\mathfrak{m}+1) \frac{C_{\mathfrak{m},\sigma}^{2}}{\mathsf{P}^{2\sigma}} (2\mathsf{P}\mathsf{K}^{2})^{\mathfrak{m}+1}.$$

By (3.1) again, $\|u(m+1) - \tilde{u}^P(m+1)\|_{L^2(\Omega)}^2 \leq D_{m,\sigma}^2 P^{-(2\sigma-m-1)}$, for certain constant $D_{m,\sigma} > 0$. By the triangular inequality,

$$\begin{split} \|u(m+1) - \hat{u}^{P}(m+1)\|_{L^{2}(\Omega)}^{2} &\leqslant (\|u(m+1) - \tilde{u}^{P}(m+1)\|_{L^{2}(\Omega)} + \|\tilde{u}^{P}(m+1) - \hat{u}^{P}(m+1)\|_{L^{2}(\Omega)})^{2} \\ &\leqslant 2\|u(m+1) - \tilde{u}^{P}(m+1)\|_{L^{2}(\Omega)}^{2} + 2\|\tilde{u}^{P}(m+1) - \hat{u}^{P}(m+1)\|_{L^{2}(\Omega)}^{2} \\ &\leqslant 2\frac{D_{m,\sigma}^{2}}{p^{2\sigma-m-1}} + 2(m+1)\frac{C_{m,\sigma}^{2}}{p^{2\sigma}}(2\mathsf{P}\mathsf{K}^{2})^{m+1} = \frac{\mathsf{E}_{m,\sigma}^{2}}{p^{2\sigma-m-1}}. \end{split}$$
(3.2)

Since $\sigma > (m+1)/2$ is arbitrary, we conclude that the Galerkin projection $\hat{u}^{P}(m)$ converges algebraically to u(m).

The hypothesis of boundedness for ζ_1, \ldots, ζ_s allowed comparing (2.2) with the multivariate Legendre basis via [1, Th. 2.3]. From a mathematical standpoint, this condition is a restriction. However, in practical applications, this assumption is not so strict. If a random input has an unbounded distribution, then one may truncate it to adapt to Theorem 3.1. In this regard, it is important to point out that applying Chebyshev-Markov inequalities or the inverse cumulative distribution function (quantile function) allows us to determine appropriate finite intervals where the truncated distribution captures most of the probabilistic information of the corresponding unbounded random variable. Moreover, in some models, for instance epidemic models, we deal with random proportions, which are bounded in [0, 1] (see the numerical example from Section 4).

3.2. Numerical convergence

We have proved that, under certain hypotheses, the Galerkin projection $\hat{u}^{P}(m)$ converges algebraically to u(m) as $P \to \infty$ in $L^{2}(\Omega)$. In practice, one must take into account the computational errors associated to Ξ^{P} and $\hat{u}_{i}^{P}(m)$, i = 1, ..., P. Thus, in reality, one works with $\hat{u}^{P,\text{error}}(m) = \sum_{i=1}^{P} \hat{u}_{i}^{P,\text{error}}(m) \varphi_{i}^{\text{error}}(\zeta)$, where $\hat{u}_{i}^{P,\text{error}}(m) = \hat{u}_{i}^{P}(m) + \varepsilon_{i}^{P}(m)$ takes into account the errors when computing the coefficients, and

 $\|\varphi_i^{error}(\zeta) - \varphi_i(\zeta)\|_{L^2(\Omega)} \leq \eta_i$ is the error from the Gram-Schmidt procedure (which affects the error $\varepsilon_i^p(m)$). The rate of convergence (3.2) becomes

$$\begin{aligned} \|\mathfrak{u}(\mathfrak{m}) - \hat{\mathfrak{u}}^{\mathsf{P},\mathsf{error}}(\mathfrak{m})\|_{L^{2}(\Omega)} &\leq \|\mathfrak{u}(\mathfrak{m}) - \hat{\mathfrak{u}}^{\mathsf{P}}(\mathfrak{m})\|_{L^{2}(\Omega)} + \|\hat{\mathfrak{u}}^{\mathsf{P}}(\mathfrak{m}) - \hat{\mathfrak{u}}^{\mathsf{P},\mathsf{error}}(\mathfrak{m})\|_{L^{2}(\Omega)} \\ &\leq C_{\sigma,\mathfrak{m}}\mathsf{P}^{-\sigma} + \|\hat{\mathfrak{u}}^{\mathsf{P}}(\mathfrak{m}) - \hat{\mathfrak{u}}^{\mathsf{P},\mathsf{error}}(\mathfrak{m})\|_{L^{2}(\Omega)}. \end{aligned}$$

If the numerical error $\|\hat{u}^{P}(m) - \hat{u}^{P,error}(m)\|_{L^{2}(\Omega)}$ is large, then there might be no algebraic convergence.

To analyze a little bit further the error term $\|\hat{u}^{P}(m) - \hat{u}^{P,error}(m)\|_{L^{2}(\Omega)}$ from an intuitive point of view, let us assume for simplicity that $\eta_{i} = 0$ and that $\varepsilon_{i}^{P}(m) = \varepsilon$ for i = 1, ..., P. Then

$$\|\hat{\mathbf{u}}^{\mathsf{P}}(\mathbf{m}) - \hat{\mathbf{u}}^{\mathsf{P}, \text{error}}(\mathbf{m})\|_{L^{2}(\Omega)}^{2} = \epsilon^{2}\mathsf{P},$$

by Pythagoras's Theorem. Thus,

$$\|u(\mathfrak{m})-\hat{\mathfrak{u}}^{P,error}(\mathfrak{m})\|_{L^2(\Omega)}^2\leqslant 2C_{\sigma,\mathfrak{m}}^2P^{-2\sigma}+2\varepsilon^2P=:b(P).$$

This bound b(P) takes its minimum at $P = (2\sigma C_{\sigma,m}^2/\epsilon^2)^{1/(2\sigma+1)}$. Hence, if there is a large error ϵ in the computation of the coefficients of the Galerkin projection, to ensure a small error it is better not to take a large P.

To summarize, from a theoretical point of view it is better to take a large P, so that the approximation done by the Galerkin projection $\hat{u}^{P}(m)$ resembles the real solution u(m); however, a excessively large P may increase the computational errors and invalidate the results obtained.

4. Numerical experiments

We will show a numerical example to illustrate the theoretical results. Consider the non-linear random difference equation $u(m + 1) = \zeta u(m)(T - u(m))$, where T > 0 is a constant, ζ is a random variable and u(m) is a random variable for each time step m. This recursive equation is called logistic.

From a modeling point of view, T could represent the total population and u(m) the number of infected individuals at time m. It is assumed that infections last for one time step. The number of infections u(m+1) depends on the number of contacts between infected individuals u(m) and susceptible individuals T - u(m), at a random rate ζ . Notice that, for this model to make sense, we must ensure that $u(m) \in [0,T]$ for every $m \ge 0$. If we suppose that $0 \le \zeta \le 4/T$, then $\zeta u(m)(T - u(m)) \le \zeta((u(m) + T - u(m))/2)^2 = \zeta T^2/4 = T$ (by the AM-GM inequality), so that the number of infected persons lies in [0,T].

Let us see that the hypotheses of Theorem 3.1 hold. We have q = s = 1. Take ζ as an absolutely continuous random variable with finite moments, bounded and with bounded density function f_{ζ} . We take I = [0, 4/T]. If we denote the solution of the recursive equation by $u(m) = g_m(\zeta)$, then clearly $g_m \in C^{\infty}(I)$, because when we go backwards in the logistic equation, the computations consist of sums and products. Let $R(\zeta, u) = \zeta u(T - u)1_{[0,T]}(u)$, for $u \in \mathbb{R}$, which is Lipschitz on \mathbb{R} . By Theorem 3.1, the Galerkin projection $\hat{u}^P(m)$ tends algebraically to u(m) as $P \to \infty$ in $L^2(\Omega)$.

For the numerical example, we use the software Mathematica[®]. We take T = 100 and $u_0 = 3$. We consider the following two distributions for the random rate: $\zeta \sim \text{Triangular}(0.02, 0.03)$ and $\zeta \sim \text{Exponential}(3)|_{[0.031,0.035]}$. In both cases, $\zeta \leq 0.04 = 4/T$. In Figure 1, we show the numerical results obtained for $0 \leq m \leq 30$ and P = 4, 5, 6. We have estimated $\mathbb{E}[\hat{u}^{P}(m)]$ (average number of infected individuals) and a confidence interval with the rule $\mathbb{E}[\hat{u}^{P}(m)] \pm \sqrt{\mathbb{V}[\hat{u}^{P}(m)]}$, by utilizing $\mathbb{E}[\hat{u}^{P}(m)] = \hat{u}_{1}^{P}(m)$ and $\mathbb{V}[\hat{u}^{P}(m)] = \sum_{i=2}^{P} (\hat{u}_{i}^{P}(m))^{2}$ [11, p. 67]. It is curious that the results obtained for P = 4 and P = 5 are similar, which indicates that the convergence starts being achieved, but for P = 6 disastrous errors appear (from the time step m > 8). We have analyzed numerically where those catastrophic errors come from: when performing the orthonormalization procedure in Mathematica with the built-in function Orthogonalize, there is a loss of orthogonality [6], which increases as the length P of the basis grows. This example shows that one has to choose the correct order P, in the sense that a large P is theoretically better, but could make the numerical errors accumulate. In this example, the best choice for uncertainty quantification is P = 5.



Figure 1: Mean value estimates and confidence intervals (in dashed lines). Up: Triangular distribution. Down: Truncated exponential distribution. P = 4 is denoted by \bigcirc in green color, P = 5 is denoted by \square in blue color, and P = 6 is denoted by \Diamond in red color. Notice that similar results are presented for P = 4 and P = 5, which agrees with the theoretical convergence. However, for P = 6, the accumulation of numerical errors invalidates the results (from m > 8).

5. Conclusions

The gPC technique has been widely used to quantify the uncertainty for continuous stochastic systems (random ordinary and partial differential equations). However, applications of gPC to random difference equations have not been performed with such an emphasis. In this paper, we have studied the application of a novel variation of gPC, called adaptive gPC, to non-linear random difference equations. We have proved in detail that, under certain assumptions, the stochastic Galerkin projection technique converges algebraically in mean square to the solution process of the random recursive equation. We have also analyzed how numerical errors could invalidate in practice the theoretical approximations. An example has been included in the end to illustrate the theoretical aspects of the paper.

Acknowledgment

This work has been supported by Spanish Ministerio de Economía y Competitividad grant MTM2017– 89664–P. Marc Jornet acknowledges the doctorate scholarship granted by Programa de Ayudas de Investigación y Desarrollo (PAID), Universitat Politècnica de València.

References

- C. Canuto, A. Quarteroni, Approximation results for orthogonal polynomials in Sobolev spaces, Math. Comp., 38 (1982), 67–86. 3.1, 3.1
- [2] B.-M. Chen-Charpentier, J.-C. Cortés, J.-A. Licea, J.-V. Romero, M.-D. Roselló, F.-J. Santonja, R.-J. Villanueva, Constructing adaptive generalized polynomial chaos method to measure the uncertainty in continuous models: A computational approach, Math. Comput. Simulation, 109 (2015), 113–129. 1, 2, 2.1
- [3] J.-C. Cortés, J.-V. Romero, M.-D. Roselló, R.-J. Villanueva, Improving adaptive generalized polynomial chaos method to solve nonlinear random differential equations by the random variable transformation technique, Commun. Nonlinear Sci. Numer. Simul., 50 (2017), 1–15. 1
- [4] A. Ditkowski, R. Kats, On Spectral approximations with nonstandard weight functions and their implementations to generalized chaos expansions, preprint arXiv, (2017), 36 pages. 3.1
- [5] O. G. Ernst, A. Mugler, H.-J. Starkloff, E. Ullmann, On the convergence of generalized polynomial chaos expansions, ESAIM Math. Model. Numer. Anal., 46 (2012), 317–339. 1, 2, 2.1
- [6] L. Giraud, J. Langou, M. Rozloznik, *The loss of orthogonality in the Gram-Schmidt orthogonalization process*, Comput. Math. Appl., 50 (2005), 1069–1075.
- [7] G. González-Parra, B. M. Chen-Charpentier, A. J. Arenas, *Polynomial Chaos for random fractional order differential equations*, Appl. Math. Comput., **226** (2014), 123–130. 1
- [8] D. Lucor, C.-H. Su, G. E. Karniadakis, Generalized polynomial chaos and random oscillators, Internat. J. Numer. Methods Engrg., 60 (2004), 571–596.
- [9] F. Santonja, B. Chen-Charpentier, *Uncertainty quantification in simulations of epidemics using polynomial chaos*, Comput. Math. Methods Med., **2012** (2012), 8 pages. 1
- [10] W. J. Shi, C. J. Zhang, Error analysis of generalized polynomial chaos for nonlinear random ordinary differential equations, Appl. Numer. Math., 62 (2012), 1954–1964. 1, 3.1
- [11] D. B. Xiu, Numerical Methods for Stochastic Computations, Princeton University Press, Princeton, (2010). 1, 4
- [12] D. B. Xiu, G. E. Karniadakis, The Wiener-Askey polynomial chaos for stochastic differential equations, SIAM J. Sci. Comput., 24 (2002), 619–644. 1
- [13] G. Zou, A. Atangana, Y. Zhou, Error estimates of a semidiscrete finite element method for fractional stochastic diffusionwave equations, Numer Methods Partial Differ. Equ., 2018 (2018), 1–15. 1