Breaking the curse of dimensionality for a class of PDEs with stochastic inputs

By A. Doostan AND G. Iaccarino

1. Motivation and objectives

Sampling methods have been utilized for a long time as a general purpose scheme for uncertainty propagation. There is an increasing recent interest in developing computational models for the analysis of uncertain systems that are more efficient compared to Monte Carlo techniques. Most notably *stochastic Galerkin schemes* (Ghanem & Spanos 2003; Deb, Babuška & Oden 2001; Xiu & Karniadakis 2002; Babuška, Tempone & Zouraris 2004) have been successfully applied to a variety of engineering problems and are extremely useful for situations in which the number of uncertain parameters is not large.

In many applications where a large number of uncertain parameters are required to characterize the uncertainty in the system, the methods mentioned above (with the exception of the Monte Carlo technique) suffer from the so-called *curse of dimensionality*: Their computational cost grows exponentially as a function of the number of random variables defining the underlying probability space of the problem. More specifically, the computational cost of the stochastic Galerkin schemes with global polynomials, e.g., Wiener Hermite chaos, (Ghanem & Spanos 2003), depends on the number of terms in the solution expansion. For the case of order p approximation in d independent random variables, the cardinality of the associated basis is $P = (p+d)!/p!d! - 1$ which increases exponentially with respect to p and d.

Stochastic collocation schemes (Xiu & Hesthaven 2005; Babuška, Nobile & Tempone 2007; Nobile, Tempone & Webster 2008, 2007) based on isotropic and anisotropic sparse grids reduce the problem of the curse of dimensionality normally associated with conventional tensor-product schemes. The computational cost of the stochastic collocation scheme using a tensor-product grid and isotropic sparse-grid constructed from M points in each direction of a d-dimensional space is $\mathcal{O}(M^d)$ and $\mathcal{O}(C^d M(\log M)^{d-1})$, respectively. Note that both estimates grow exponentially with respect to d.

Recently, we have successfully extended the alternating least-squares (ALS) approximation technique of (Beylkin & Mohlenkamp 2002, 2005) to obtain a low-rank separated representation of the solution to partial differential equations with high-dimensional random inputs, (Doostan, Iaccarino & Etemadi 2007; Doostan & Iaccarino 2008). In the present study, a novel alternative approach based on a Rayleigh-Ritz variational technique is investigated to further reduce the computational cost associate with low-rank approximation of a class of PDEs with high-dimensional random inputs. More specifically, when the random linear system of equations arising from spatial/temporal discretization of a PDE is symmetric positive-definite, the low-rank approximation of the solution can be obtained using an *alternating Rayleigh-Ritz* (ARR) scheme as opposed to the ALS. While the computational cost of the proposed technique, similar to that of the ALS algorithm, scales linearly with respect to the number of uncertain variables, the total computational cost is significantly smaller. The proposed approach is funda-

mentally different from stochastic Galerkin and collocation schemes. More specifically, it deviates from stochastic Galerkin schemes by not assuming any pre-determind basis along the stochastic dimension for the approximation. It is also different from stochastic collocation schemes as it does not solve the problem on the quadrature grid *exactly*.

The paper is organized as follows. Section 2.1 summarizes the separated representation of a $d-$dimensional function where a similar notation to that of (Beylkin & Mohlenkamp 2002, 2005) is adopted for the sake of consistency. In Section 2.2 it will be shown how one can cast the discrete analog of the governing stochastic PDE in a form that can be readily incorporated in the separated approximation framework. Finally, in Section 3, numerical experiments are performed on a 1-D unsteady Burgers' equation with random initial condition and viscosity to illustrate the performance of the proposed procedure.

2. Numerical method

2.1. *Separated representation*

Separation-of-variable techniques have been widely used to approximate high-dimensional functions using one-dimensional operations, thus virtually eliminating the curse of dimensionality, see (Pereyra & Scherer 1973; Kroonenberg & Leeuw 1980; Kolda 2001; Beylkin & Mohlenkamp 2002; Tyrtyshnikov 2004; Hackbusch & Khoromskij 2004; Beylkin & Mohlenkamp 2005) and references therein. Let u be a d-dimensional function. It can be approximated as

$$u(y_1, \cdots, y_d) \approx \phi_1(y_1) \cdots \phi_d(y_d). \tag{2.1}$$

The above approximation can be improved by introducing a series of such representations,

$$u(y_1, \cdots, y_d) = \sum_{l=1}^{r} s_l \phi_1^l(y_1) \cdots \phi_d^l(y_d) + \mathcal{O}(\epsilon), \tag{2.2}$$

which is called a *separated representation* with *separation rank* r. Given a target accuracy ϵ, the approximation Eq. (2.2) can be achieved by tailoring the unknown quantities $\{\phi_i^l(y_i)\}$, $\{s_l\}$, and an optimal separation rank r, for instance, through a non-linear optimization scheme. For general functions, the separated representation Eq. (2.2) is not unique; however, any separated approximation satisfying the accuracy ϵ is acceptable. The main advantage of adopting the separated representation Eq. (2.2) is that many algebraic operations in d dimensions can be performed using a series of one-dimensional operations. Therefore, in theory, the computational efforts increase only linearly with respect to d.

In the following, the goal is to obtain a low-rank separated representation of the solution of a linear system of equations arising from spatial/temporal discretization of stochastic PDEs using only one-dimensional operations. To achieve this objective some preliminary notations and definitions are introduced.

Notation 1. A scalar function $u(y_1, \cdots, y_d)$ in d dimensions is a mapping from the Euclidean space $\mathbb{R}^d$ to real line $\mathbb{R}$, $u : \mathbb{R}^d \to \mathbb{R}$. A vector $\mathcal{U} \equiv u(j_1, \cdots, j_d)$ with $j_k = 1, \cdots, M_k$ is a discrete representation of function u on a d-dimensional tensor-product grid $\mathcal{G}$ with $\prod_{k=1}^{d} M_k$ nodes. Without loss of generality it is assumed that the grid is isotropic, i.e., $M_k = M$ for $k = 1, \cdots, d$.

Definition 1 (Beylkin & Mohlenkamp 2002, 2005) (Separated representation of a vector).

Let $\otimes$ denote the Kronecker product. For a given accuracy, ϵ, a vector $\mathcal{U} = u(j_1, \cdots, j_d)$ in d dimension is approximated by

$$\sum_{l=1}^{r} s_l u_1^l(j_1) u_2^l(j_2) \cdots u_d^l(j_d) \equiv \sum_{l=1}^{r} s_l \mathbf{u}_1^l \otimes \mathbf{u}_2^l \otimes \cdots \otimes \mathbf{u}_d^l, \tag{2.3}$$

with $s_l \in \mathbb{R}$ being a scalar and $\mathbf{u}_k^l \in \mathbb{R}^M$ a one-dimensional vector with entries $u_k^l(j_k)$ and unit Euclidean norm. The approximation error is required to be less than ϵ, i.e.,

$$\|\mathcal{U} - \sum_{l=1}^{r} s_l \mathbf{u}_1^l \otimes \mathbf{u}_2^l \otimes \cdots \otimes \mathbf{u}_d^l\| \leq \epsilon, \tag{2.4}$$

where $\|\cdot\|$ denotes the Frobenius norm over $\mathcal{G}$.

In the following, the objective is to compute a low-rank separated approximation to the solution of a stochastic PDE when its spatial-temporal discretization leads to a symmetric positive-definite random linear system of equations.

2.2. *Solution of a stochastic PDE*

Consider a linear partial differential equation with stochastic operator and forcing, and let $u(x, t, \omega)$ be the solution in $\bar{\mathcal{D}} \times [0, T] \times \Omega \to \mathbb{R}$, such that the following equation holds almost surely in Ω,

$$\begin{aligned} \mathcal{L}(x, t, \omega; u) &= f(x, t, \omega) \quad (x, t) \in \mathcal{D} \times [0, T] \\ u(x, t, \omega; u) &= g(x, t) \quad (x, t) \in \partial\mathcal{D} \times [0, T] \\ u(x, 0, \omega) &= h(x, \omega) \quad x \in \mathcal{D}, \end{aligned} \tag{2.5}$$

where Ω is the set of elementary events and $\omega \in \Omega$, $\mathcal{D}$ and T denote the physical extent and the time interval of the problem, respectively. The randomness in the problem is induced by the uncertainty in the underlying parameters of the corresponding physical system, e.g., heat conductivity, viscosity, initial conditions, etc., and is assumed to be a function of a finite, possibly very large, number of random variables. Such representation can be obtained through, for instance, spectral decomposition of the covariance kernel of the underlying (second-order) random fields/processes which is referred to as Karhunen-Loeve expansion (Loeve 2003). Therefore, the random differential operator of Eq. (2.5) can be represented as:

$$\mathcal{L}(x, t, \omega; u) = \mathcal{L}\left(x, t, y_1(\omega), \cdots, y_{d_O}(\omega); u\right), \tag{2.6}$$

where $\{y_k\}_{k=1}^{d_O}$ are random variables, e.g., Karhunen-Loeve expansion random variables. Similarly, one can rewrite:

$f(x, t, \omega) = f\left(x, t, y_1(\omega), \cdots, y_{d_f}(\omega)\right)$ and $h(x, \omega) = h\left(x, y_1(\omega), \cdots, y_{d_h}(\omega)\right)$.

In the present work, it is assumed that the random variables $\{y_k\}_{k=1}^{d}$ with $d \equiv d_O + d_f + d_h$ are independent with known probability density functions $\{\rho_k\}_{k=1}^{d}$, respectively. Let $\Gamma_k \equiv y_k(\Omega)$ be the image of the random variable $y_k(\omega)$; the underlying probability space Γ is the product of images of random variables $y_k(\omega)$, i.e., $\Gamma \equiv \prod_{k=1}^{d} \Gamma_k$.

The solution of Eq. (2.5) is a mapping of $\{x, t, y_1(\omega), \cdots, y_d(\omega)\}$, i.e.,

$$u(x,t,\omega) := u(x,t,y_1(\omega),\cdots,y_d(\omega)), \tag{2.7}$$

which is in general non-linear.

Considering a spatial-temporal discretization scheme; the semi-discrete equivalent form of Eq. (2.5) typically simplifies to solution of a random linear system of equations written as

$$\mathbf{A}_n(\omega)\mathbf{u}_n(\omega) = \mathbf{f}_n(\omega) \quad \forall n, \tag{2.8}$$

where n denotes the index associated with the time integration scheme and is henceforth omitted for the sake of a simpler notation.

Assumption 1. In the present study, it is assumed that the random matrix $\mathbf{A}(\omega)$ is symmetric positive-definite for all $\omega \in \Omega$. This clearly restricts the use of proceeding procedures to particular classes of PDEs Eq. (2.5) and semi-discretizations schemes. Furthermore, it is assumed that $\mathbf{A}(\omega)$ and the random vector $\mathbf{f}(\omega)$ in Eq. (2.8) have separated representations with respect to their coordinates $\{y_1(\omega),\cdots,y_d(\omega)\}$. For the case of linear problem Eq. (2.5), the latter assumption readily holds once the corresponding uncertain parameters are characterized in a separated representation, for instance, through a procedure similar to Karhunen-Loeve expansion (here with independent random variables) or a polynomial chaos expansion (possibly with large separation rank that can further be reduced using the ALS algorithm of (Beylkin & Mohlenkamp 2002, 2005; Doostan, Iaccarino & Etemadi 2007; Doostan & Iaccarino 2008)).

The discrete representation of Eq. (2.8) on a tensor-product grid $\mathcal{G}$ consisting of M nodes along each direction y_k (total of M^d nodes) reads,

$$\mathcal{A}\mathcal{U} = \mathcal{F}, \tag{2.9}$$

with

$$\mathcal{A} \equiv \sum_{l=1}^{r_{\mathbf{A}}} \mathbf{A}_0^l \otimes \mathbf{A}_1^l \otimes \cdots \otimes \mathbf{A}_d^l \tag{2.10}$$

and

$$\mathcal{F} \equiv \sum_{l=1}^{r_{\mathbf{f}}} \mathbf{f}_0^l \otimes \mathbf{f}_1^l \otimes \cdots \otimes \mathbf{f}_d^l. \tag{2.11}$$

Here $\mathbf{A}_0^l \in \mathbb{R}^{N,N}$ and $\mathbf{f}_0^l \in \mathbb{R}^N$ are deterministic (sparse) matrices and vectors, respectively, whose size, N, is determined by the spatial discretization scheme and are obtained from the discretization of deterministic modes in the representation of underlying stochastic fields. For $k = 1,\cdots,d$, the diagonal matrices $\mathbf{A}_k^l \in \mathbb{R}^{M,M}$ and vectors $\mathbf{f}_k^l \in \mathbb{R}^M$ hold quadrature values in the finite-dimensional representation of corresponding stochastic fields. Finally, $\mathcal{U}$ is the tensor-product representation of the solution to be calculated. In general, the construction of separated representations of the coefficient tensor $\mathcal{A}$ and the right-hand side tensor $\mathcal{F}$ in Eq. (2.9) along with their separation ranks $r_{\mathbf{A}}$ and $r_{\mathbf{f}}$, respectively, depend on the type of the stochastic PDE, the spatial-temporal discretizations, and how the finite-dimensional uncertainty representations are obtained.

Remark 1. Notice that the representation Eq. (2.9) is in fact a compact notation for the tensor-product stochastic collocation approximation of Eq. (2.8) on $\mathcal{G}$. The solution $\mathcal{U}$ in Eq. (2.9) is therefore the solution of the tensor-product stochastic collocation applied to Eq. (2.8). The procedure proposed here, however, departs from the collocation approach by approximating the solution $\mathcal{U}$ on $\mathcal{G}$, within the class of separated representations, and not evaluating it exactly at the quadrature points. As is subsequently described in greater details, such approximation allows one to break the issue of the curse of dimensionality associated with the tensor-product collocation approach.

2.2.1. *Separated approximation of Eq. (2.9)*

The exact solution, $\mathcal{U}$, of linear system Eq. (2.9) minimizes the functional

$$F(\mathcal{V}) = \langle \mathcal{A}\mathcal{V}, \mathcal{V} \rangle - 2\langle \mathcal{F}, \mathcal{V} \rangle, \tag{2.12}$$

where $\langle \mathcal{U}, \mathcal{V} \rangle$ denotes the inner product of two tensors $\mathcal{U}$ and $\mathcal{V}$. Given that $\mathcal{A}$ is symmetric, the functional F is equivalently written as

$$F(\mathcal{V}) = \langle \mathcal{A}\left(\mathcal{U} - \mathcal{V}\right), \mathcal{U} - \mathcal{V} \rangle - \langle \mathcal{F}, \mathcal{U} \rangle. \tag{2.13}$$

Furthermore, since $\mathcal{A}$ is positive definite, minimization of the functional $F(\mathcal{V})$ leads to minimization of $\|\mathcal{U} - \mathcal{V}\|$ as

$$\|\mathcal{U} - \mathcal{V}\|^2 \leq \langle \mathcal{A}\left(\mathcal{U} - \mathcal{V}\right), \mathcal{U} - \mathcal{V} \rangle. \tag{2.14}$$

In this work, the domain of $\mathcal{A}$ is restricted to separated tensors and F is minimized by defining

$$\mathcal{U}_\epsilon^M \equiv \sum_{l=1}^{r_\mathbf{u}} s_l^\mathbf{u} \mathbf{u}_0^l \otimes \mathbf{u}_1^l \otimes \cdots \otimes \mathbf{u}_d^l, \tag{2.15}$$

which satisfies

$$\|\mathcal{F} - \mathcal{A}\mathcal{U}_\epsilon^M\| \leq \epsilon \|\mathcal{F}\|, \tag{2.16}$$

to a desired accuracy ϵ. This is a non-linear optimization problem as $r_\mathbf{u}$, $\{s_l^\mathbf{u}\}_{l=1}^{r_\mathbf{u}}$, and $\{\mathbf{u}_k^l\}_{l=1}^{r_\mathbf{u}}$ for $k = 0, 1, \cdots, d$ are to be computed. Due to the already separated form of $\mathcal{U}_\epsilon^M$, $\mathcal{A}$, and $\mathcal{F}$, such minimization can be performed in the form of a sequence of linear one-dimensional Rayleigh-Ritz problems which is, here, referred to as *alternating Rayleigh-Ritz* (ARR) algorithm.

Alternating Rayleigh-Ritz (ARR) *algorithm.* For a fixed $r_\mathbf{u}$, an initial guess for $\mathcal{U}_\epsilon^M$ is constructed by randomly initializing $\{\mathbf{u}_i^l\}$. The optimization steps are then as follows:

- Loop over dimensions $k = 0, 1, \cdots, d$
  - Loop over grid points $j_k = 1, \cdots, J$ ($J = N$ when $k = 0$ and $J = M$ otherwise) in direction k
    - Fix $\{\mathbf{u}_i^{\tilde{l}}\}_{i \neq k}$ and solve the following linear system of equations associated with the Rayleigh-Ritz problem:

$$\mathbf{B}\mathbf{c} = \mathbf{b}, \tag{2.17}$$

where $\mathbf{B}$ contains $r_{\mathbf{u}} \times r_{\mathbf{u}}$ blocks with size $J \times J$. The $(\hat{l}, \tilde{l})$-th block is obtained from

$$\mathbf{B}_{(\hat{l},\tilde{l})}[i^{'}, j^{'}] = \sum_{l=1}^{r_{\mathbf{A}}} \mathbf{A}_k^l[i^{'}, j^{'}] \prod_{i \neq k} \langle \mathbf{A}_i^l \mathbf{u}_i^{\tilde{l}}, \mathbf{u}_i^{\hat{l}} \rangle. \tag{2.18}$$

Moreover $\mathbf{b}$ contains $r_{\mathbf{u}}$ vectors each with size J where the $\hat{l}$-th component is computed as

$$\mathbf{b}_{\hat{l}}(i^{'}) = \sum_{l=1}^{r_{\mathbf{f}}} \mathbf{f}_k^l(i^{'}) \prod_{i \neq k} \langle \mathbf{f}_i^l, \mathbf{u}_i^{\hat{l}} \rangle. \tag{2.19}$$

- Update $s_{\tilde{l}}^{\mathbf{u}} = \left(\sum_{j_k} \mathbf{c}_{\tilde{l}}^2(j_k) \right)^{1/2}$ and $u_k^{\tilde{l}}(j_k) = \mathbf{c}_{\tilde{l}}(j_k)/s_{\tilde{l}}^{\mathbf{u}}$ for all $\tilde{l} = 1, \cdots, r_{\mathbf{u}}$.

Here $\langle \cdot, \cdot \rangle$ denotes the usual inner product of two vectors. The above algorithm monotonically reduces the functional F in Eq. (2.12), hence $\|\mathcal{U} - \mathcal{U}_\epsilon^M\|$, until the rate of the reduction is small. If the desired accuracy, based on Eq. (2.16), has not been achieved, the rank $r_{\mathbf{u}}$ must be increased to reduce F further. It is in general not possible to determine the optimal rank $r_{\mathbf{u}}$ *a priori*. However, in order to achieve a near-optimal rank $r_{\mathbf{u}}$, it is proposed to start from a low separation rank, e.g., $r_{\mathbf{u}} = 1$, and reduce the representation error using the ARR algorithm and, if necessary, increase $r_{\mathbf{u}}$ until Eq. (2.16) is attained. The overall procedure is then summarized as:

Separated approximation algorithm:
- Set $r_{\mathbf{u}} = 1$; (randomly) initialize $\{\mathbf{u}_i^1\}$ and $s_1^{\mathbf{u}}$
- Loop while $\|\mathcal{F} - \mathcal{A}\mathcal{U}_\epsilon^M\| > \epsilon \|\mathcal{F}\|$
  - Perform the ARR algorithm until the decrease in $F(\mathcal{U}_\epsilon^M)$ in successive iterations is smaller than a prescribed tolerance.
  - Set $r_{\mathbf{u}} = r_{\mathbf{u}} + 1$; (randomly) initialize $\{\mathbf{u}_i^{r_{\mathbf{u}}}\}$ and $s_{r_{\mathbf{u}}}^{\mathbf{u}}$.

Note that for updates along random directions, $k = 1, \cdots, d$, the linear system Eq. (2.17) decouples into M smaller linear systems each with size $r_{\mathbf{u}} \times r_{\mathbf{u}}$ for each nodal point j_k. However, when $k = 0$ the linear system Eq. (2.17) is coupled according to the spatial discretization and thus has the size $(N \cdot r_{\mathbf{u}}) \times (N \cdot r_{\mathbf{u}})$.

In practice, the linear system Eq. (2.17) is not formed explicitly, particularly when updates along spatial direction x, i.e., $k = 0$, are performed. Depending on the choice of the solver, the linear system Eq. (2.17) can be solved by using only matrix-vector multiplications and vector-vector dot-products. In the present study, a preconditioned conjugate gradient solver is employed; an incomplete LU decomposition of the mean of the random matrix $\mathbf{A}(\omega)$, i.e., $\mathbf{A}_0^1$, is used as the preconditioner.

Remark 2. Similar to tensor-product stochastic collocation schemes, the choice of grid points j_k affects the convergence and the accuracy of the approximation. In the present work, based on the distribution of the random variables y_k, quadrature rules are used to distribute these abscissas along each direction y_k.

2.2.2. *Computational cost*

Operations such as summation and product of two tensors having separated form incure a computational cost that scales linearly with respect to dimension d; incidentally, the

memory storage requirements also scale linearly. In theory, as far as $r_{\mathbf{u}}$ is finite and does not depend on d, it can be shown that one full ARR iteration, as described above, to update quantities from $k = 0$ to $k = d$ requires only linear increase of computational cost with respect to d. More precisely, for the interesting case of $N \gg d \cdot M$, the number of operations needed to compute vectors $\mathbf{b}$ and matrices $\mathbf{B}$ for updates along $k = 0$ to $k = d$ is $\mathcal{O}\left(d \cdot r_{\mathbf{u}} \cdot r_{\mathbf{f}}.N\right)$ and $\mathcal{O}\left(d \cdot r_{\mathbf{A}} \cdot r_{\mathbf{u}}^2.N^2\right)$, respectively. Finally the solution of linear systems Eq. (2.17) requires $\mathcal{O}(r_{\mathbf{u}}^3.N^3)$ operating using a direct solver. However, in practice, $r_{\mathbf{u}}$ may depend mildly on d; therefore the computational cost of the above ARR scheme is nearly linear with respect to d.

2.2.3. *Regularization*

In general, the problem of funding a separated solution $\mathcal{U}_\epsilon^M$ that minimizes the functional F in Eq. (2.12) is ill-conditioned due to loss of precision and thus requires regularization. A simple *Tikhonov* regularization technique has been proposed in (Beylkin & Mohlenkamp 2005) which aims at controlling the *condition number*,

$$\kappa \equiv \frac{\left(\sum_{l=1}^{r_{\mathbf{u}}} \left(s_l^{\mathbf{u}}\right)^2\right)^{1/2}}{\|\mathcal{U}_\epsilon^M\|}, \tag{2.20}$$

of the separated representation $\mathcal{U}_\epsilon^M$. More specifically, to prevent the loss of significant digits and thus to achieve the approximation Eq. (2.16), one has to satisfy $\kappa\mu\|\mathcal{U}_\epsilon^M\| \leq \epsilon$, where μ is the machine precision. Such regularization can be implemented for the case of ARR scheme by simply replacing $\mathbf{B}$ with $\mathbf{B} + \lambda\mathbf{I}$ in Eq. (2.17) and choosing the scalar λ slightly larger than μ, (Beylkin & Mohlenkamp 2005). Here, $\mathbf{I}$ denotes the identity matrix of the same size as that of $\mathbf{B}$.

2.2.4. *Response statistics*

The computation of response statistics for the proposed technique follows that of a tensor-product stochastic collocation technique (Xiu & Hesthaven 2005; Babuška, Nobile & Tempone 2007) except that the computational cost grows linearly with respect to dimension d. Given the discrete solution $\mathcal{U}_\epsilon^M$ in Eq. (2.15), one can estimate the desirable statistics, e.g., moments, of u based on 1) sampling from interpolating surface of $\mathcal{U}_\epsilon^M$ or 2) numerical integration using quadrature rules for its moments, e.g., mean and variance. These are described in greater detail below.

Statistics based on interpolation. A multi-dimensional interpolation technique with, for instance, Lagrange polynomials can be employed to construct the solution response surface based on the nodal values of u along the random dimensions. Given the separated solution $\mathcal{U}_\epsilon^M$, such interpolation is cast as a sequence of one-dimensional interpolations, hence, resulting in a total computational cost that grows linearly with respect to d. More specifically, let $\mathcal{I}\left(\mathbf{u}_\epsilon^M\right)(y_1, \cdots, y_d) : \mathbb{R}^N \times \mathcal{G} \to \mathbb{R}^N \times \Gamma$ denote the interpolation of the solution $\mathcal{U}_\epsilon^M$ in Γ, then

$$\mathcal{I}\left(\mathbf{u}_\epsilon^M\right)(y_1, \cdots, y_d) = \sum_{l=1}^{r_{\mathbf{u}}} s_l^{\mathbf{u}} \mathbf{u}_0^l \prod_{k=1}^{d} \mathcal{I}(u_k^l)(y_k), \tag{2.21}$$

where

$$\mathcal{I}(u_k^l)(y_k) = \sum_{j_k=1}^{M} u_k^l\,(j_k)\,L_{j_k}(y_k), \quad \forall k, l, \tag{2.22}$$

is the one-dimensional interpolation along direction y_k and $L_{j_k}(y_k)$ is the Lagrange polynomial corresponding to the node j_k. The desired solution statistics are then computed by sampling from $\mathcal{I}\left(\mathbf{u}_\epsilon^M\right)(y_1, \cdots, y_d)$.

Statistics based on quadrature integration. The integral-form statistics, e.g., moments, of the response quantities of interest can be obtained as a sequence of one-dimensional quadrature integrations once their separated representation is available. For instance, the estimate of the mean and second moment of the nodal solution vector $\mathbf{u}$ is obtained as

$$\begin{aligned}\mathbb{E}\left[\mathbf{u}_\epsilon^M\right] &= \int_\Gamma \mathbf{u}_\epsilon^M(y_1, \cdots, y_d) \prod_{k=1}^{d} (\rho_k dy_k) \\ &= \sum_{l=1}^{r_\mathbf{u}} s_l^\mathbf{u}\mathbf{u}_0^l \prod_{k=1}^{d} \left(\int_{\Gamma_k} u_k^l(y_k)\rho_k dy_k \right) = \sum_{l=1}^{r_\mathbf{u}} s_l^\mathbf{u}\mathbf{u}_0^l \prod_{k=1}^{d} \left(\sum_{j_k=1}^{M} u_k^l\,(j_k)\,w_{j_k} \right)\end{aligned} \tag{2.23}$$

and similarly

$$\begin{aligned}\mathbb{E}\left[(\mathbf{u}_\epsilon^M)^2\right] &\equiv \int_\Gamma \left(\mathbf{u}_\epsilon^M \circ \mathbf{u}_\epsilon^M\right)(y_1, \cdots, y_d) \prod_{k=1}^{d} (\rho_k dy_k) \\ &= \sum_{l=1}^{r_\mathbf{u}} \sum_{\check{l}=1}^{r_\mathbf{u}} s_l^\mathbf{u} s_{\check{l}}^\mathbf{u} (\mathbf{u}_0^l \circ \mathbf{u}_0^{\check{l}}) \prod_{k=1}^{d} \left(\sum_{j_k=1}^{M} u_k^l\,(j_k)\,u_k^{\check{l}}\,(j_k)\,w_{j_k} \right),\end{aligned} \tag{2.24}$$

respectively. In Eqs. 2.23 and 2.24, $\mathbb{E}$ is the mathematical expectation operator, $\circ$ denotes the Hadamard product of two vectors, and w_{j_k} is the weight associated with the quadrature point j_k.

3. Results

A 1-D unsteady viscous Burgers' equation with stochastic initial condition and diffusion fields is considered as a verification test example. The mathematical model for the random velocity field $u(x, t, \omega)$ reads,

$$\begin{aligned}&\frac{\partial u(x,t,\omega)}{\partial t} + u(x,t,\omega)\frac{\partial u(x,t,\omega)}{\partial x} = \frac{\partial}{\partial x}\left(\nu(x,\omega)\frac{\partial u(x,t,\omega)}{\partial x}\right) \quad (x,t) \in \mathcal{D}\times[0,T] \\ &u(0,t,\omega) = u(1,t,\omega) = 0 \quad (x,t) \in \partial\mathcal{D}\times[0,T] \\ &u(x,0,\omega) = g(x,\omega) \quad x \in \mathcal{D},\end{aligned} \tag{3.1}$$

where the uncertain viscosity $\nu(x, \omega)$ and initial condition $g(x, \omega)$ are characterized as

$$\nu(x,\omega) = \bar{\nu} + \sigma_\nu \sum_{k=1}^{d_\nu} \frac{1}{k}\cos(2\pi k x) y_k(\omega) \tag{3.2}$$

and

$$g(x,\omega) = \sin(\pi x) + \sigma_g \sum_{l=2}^{d_g} \frac{1}{l} \sin(\pi l x) y_l(\omega), \tag{3.3}$$

respectively. The parameters of representations Eq. (3.2) and Eq. (3.3) are $\bar{\nu} = 1 \times 10^{-2}$, $\sigma_\nu = 9.0634 \times 10^{-4}$, $d_\nu = 15$, $\sigma_g = 1.2840 \times 10^{-1}$, and $d_g = 16$. Random variables $\{y_k\}_{k=1}^{d_\nu + d_g - 1}$ are *i.i.d.* uniformly distributed on $[-1, +1]$, therefore the solution $u(x,t,\omega)$ is 30 dimensional with respect to the random space.

The problem Eq. (3.1) is discretized with the central difference of size $\Delta x = 6.250 \times 10^{-3}$ in space and the semi-implicit two-step Adam-Bashforth with $\Delta t = 3 \times 10^{-4}$ in time to the final analysis time $T = 1.5$. In particular, the convection and diffusion terms are treated explicitly and implicitly, respectively. Finally, a tensor-product grid based on Clenshaw-Curtis rules of size $M = 8$ along each direction y_k is considered for discretization of Eq. (3.1) in the random space.

Figure 1 compares the performance of the ARR algorithm in estimating the time-dependent solution mean and standard deviation with those of a Monte Carlo simulation using $50,000$ samples. We also observe that the accuracy of the scheme for estimating mean and standard deviation deteriorates for longer times. This issue is known as the "long-time integration" of unsteady stochastic systems and is typical to Stochastic Galerkin and collocation techniques as well. In Fig. 2 the evolution of the solution rank $r_{\mathbf{u}}$ for the case of $\epsilon = 2 \times 10^{-3}$, $\epsilon = 5 \times 10^{-3}$, and $\epsilon = 1 \times 10^{-2}$ is illustrated. As is observed from this plot, the solution rank decreases as a function of time, thus leading to decrease in computational cost over time. For the fist analysis time step, the initial approximation rank $r_{\mathbf{u}} = 1$ is selected and the algorithm is executed to achieve the target accuracy. For the subsequent time steps, the solution is initialized based on the solution of the previous time step with a rank that is minimum of 1 and $r_{\mathbf{u}} - 3$. This significantly reduces the overall computational cost as opposed to resetting the initial rank to $r_{\mathbf{u}} = 1$ for all time steps. Notice that the overall rank reduction of the solution is expected in this case because of the choice of the uncertainty. The details of the initial conditions and the small scale dynamics are successively reduced by the effect of the viscosity.

4. Conclusion and future directions

The present study addresses the issue of the curse of dimensionality associated with the approximation of high-dimensional stochastic partial differential equations whose spatial/temporal discretization leads to symmetric positive-definite linear systems. The proposed method is based on the low-rank separated representation of multi-variate functions and an alternating Rayleigh-Ritz (ARR) variational algorithm that minimizes the difference between the quantity to be estimated and the corresponding separated representation. The number of required operations, hence the computational cost (and also the memory requirements) for the proposed scheme is formally linear with respect to the dimension of the underlying probability space in which the solution exists. This is, in general, a significant advantage over the widely used stochastic Galerkin and collocation schemes when solving problems with a large number of random variables.

The proposed ARR algorithm is verified through its application to the 1-D unsteady Burgers' equation with random initial condition and viscosity. The existence of a low-rank separated solution for a non-linear PDE, namely the Burgers' equation, with a high-dimensional probability space has been observed. Furthermore, the ability of the

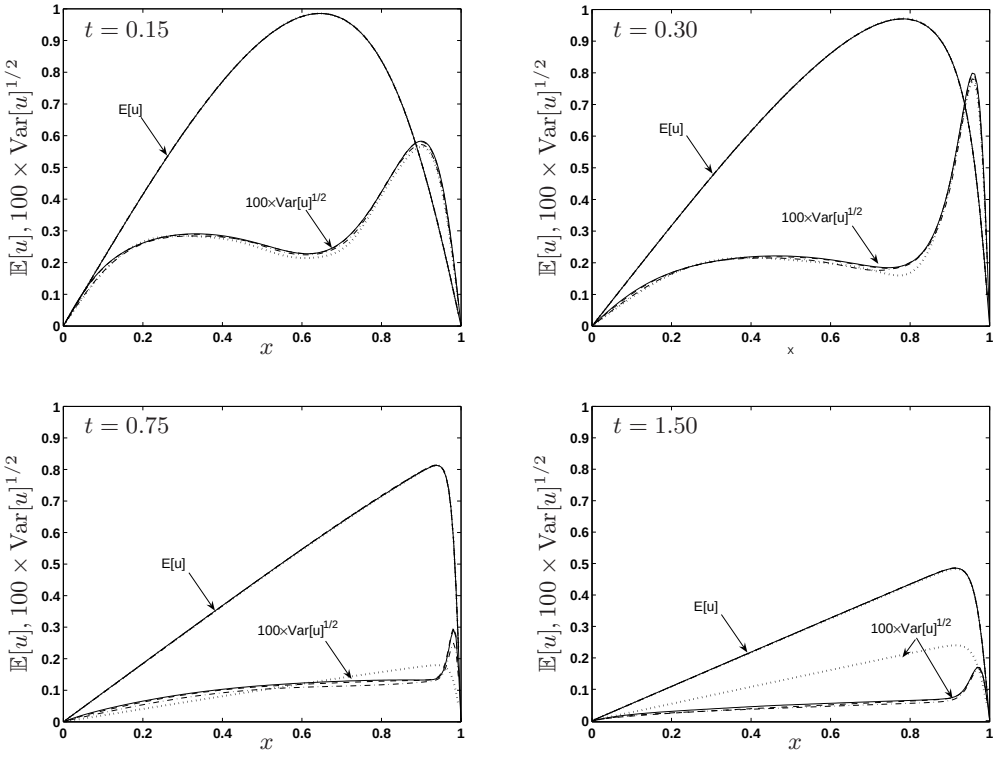

FIGURE 1. Comparison of the solution mean, $\mathbb{E}[u]$, and standard deviation, $\mathrm{Var}[u]^{1/2}$, at different analysis times, $t = 0.15, 0.30, 0.75, 1.50$, based on ARR algorithm and $50,000$ Monte Carlo samples. (Monte Carlo simulation ——— ; ARR algorithm with $M = 8, \epsilon = 2 \times 10^{-3}$ – – – – ; ARR algorithm with $M = 8, \epsilon = 5 \times 10^{-3}$ —·— ; ARR algorithm with $M = 8, \epsilon = 1 \times 10^{-2}$ $\cdots$).

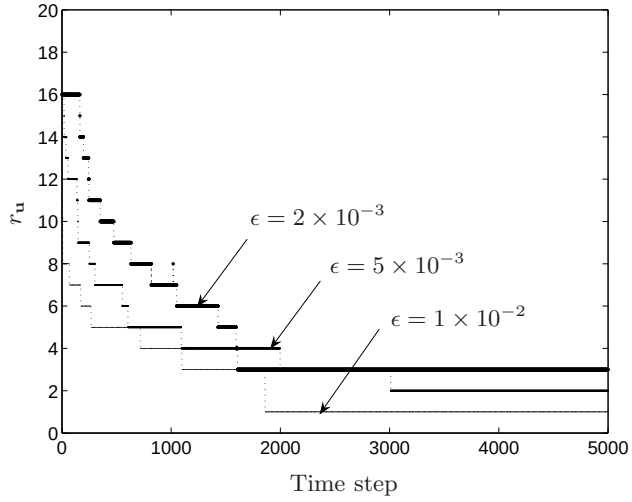

FIGURE 2. Evolution of the solution rank as a function of time and accuracy ϵ.

ARR scheme to capture the low-rank separated approximate solution is verified in the numerical experiment.

The effect of long-time integration has been investigated briefly in a numerical experiment. The target accuracy, ϵ, and the number of collocation points, M, along the range of each random variable in the ARR algorithm can be adjusted *a priori* to achieve a certain accuracy at the final analysis time. However, this might require a relatively small ϵ (large separation rank) to begin with, thus making the proposed procedures not as efficient. Similar to the case of stochastic Galerkin and collocation schemes, further studies are needed to improve the efficiency of the proposed approach for long-time integrations of unsteady problems. Furthermore, we plan to derive *a proiri* error estimates for the proposed scheme in order to investigate the convergence of the method with respect to M and ϵ.

Acknowledgments

This work was supported by the United States Department of Energy under the Predictive Science Academic Alliance Program (PSAAP) at Stanford University.

REFERENCES

Babuška, I., Nobile, F. & Tempone, R. 2007 A stochastic collocation method for elliptic partial differential equations with random input data. *SIAM J. Numer. Anal.* **45**(3), 1005–1034.

Babuška, I., Tempone, R. & Zouraris, G. 2004 Galerkin finite element approximations of stochastic elliptic partial differential equations. *SIAM J. Numer. Anal.* **42**(2), 800–825.

Beylkin, G. & Mohlenkamp, M. 2002 Numerical operator calculus in higher dimensions. *Proceedings of the National Academy of Science.* **99**(16), 10246–10251.

Beylkin, G. & Mohlenkamp, M. 2005 Algorithms for numerical analysis in high dimensions. *SIAM J. Sci. Comput.* **26**(16), 2133–2159.

Deb, M., Babuška, I. & Oden, J. 2007 Solution of stochastic partial differential equations using Galerkin finite element techniques. *Computer Methods in Applied Mechanics and Engineering.* **190**(48), 6359–6372.

Doostan, A., Iaccarino, G. & Etemadi, N. 2007 A least-squares approximation of high-dimensional uncertain systems. *Annual Research Briefs*, Stanford, California: Center for Turbulence Research, 121–132.

Doostan, A. & Iaccarino, G. 2008 A least-squares approximation of partial differential equations with high-dimensional random inputs. *submitted to Journal of Computational Physics.*

Ghanem, R. & Spanos, P. 2003 *Stochastic Finite Elements: A Spectral Approach.* Mineola, New York: Dover.

Hackbusch, W. & Khorminskij, B. 2004 Kronecker tensor-product approximation to certain matrix-valued functions in higher dimensions. *Technical Report Preprint 16*, Bonn, Germany: Max-Planck-Institut für Mathematik in den Naturwissenschaften.

Kolda, T. 2001 Orthogonal tensor decompositions. *SIAM Journal on Matrix Analysis and Applications.* **23**(1), 243–255.

Kroonenberg, P. & Leeuw, R. 1980 Principal component analysis of three-mode data by means of alternating least squares algorithms. *Psychometrika.* **45**(1), 69–97.

Loeve, M. 1977 *Probability Theory,* 4^{th} *edition.* Springer-Verlag, NewYork.

Mathelin, L. & Hussaini, M. 2003 A stochastic collocation algorithm for uncertainty analysis. *Technical Report NAS 1.26:212153; NASA/CR-2003-212153*, Langley, Virginia: NASA Langley Research Center.

Nobile, F., Tempone, R. & Webster, C. 2008 A sparse grid stochastic collocation method for partial differential equations with random input data. *SIAM J. Numer. Anal.* **46**(5), 2309–2345.

Nobile, F., Tempone, R. & Webster, C. 2007 An anisotropic sparse grid stochastic collocation method for partial differential equations with random input data. *SIAM J. Numer. Anal.* **46**(5), 2411–2442.

Pereyra, V. & Scherer, G. 1973 Efficient computer manipulation of tensor products with applications to multidimensional approximation. *Math. Comput.* **27**(123), 595–605.

Tyrtyshnikov, E. 2004 Kronecker-product approximations for some function-related matrices. *Linear Algebra and its Applications.* **379**, 423–437.

Xiu, D. & Hesthaven, J. 2005 High-order collocation methods for differential equations with random inputs. *SIAM J. Sci. Comput.* **27**(3), 1118–1139.

Xiu, D. & Karniadakis, G. 2002 The Wiener-Askey polynomial chaos for stochastic differential equations. *SIAM J. Sci. Comput.* **24**(2), 619–644.