

Approximation of Spatio-Temporal Random Processes Using Tensor Decomposition

Debraj Ghosh* and Anup Suryawanshi

*Department of Civil Engineering, Indian Institute of Science, Bangalore 560012,
India.*

Received 20 November 2012; Accepted (in revised version) 19 November 2013

Available online 31 March 2014

Abstract. A new representation of spatio-temporal random processes is proposed in this work. In practical applications, such processes are used to model velocity fields, temperature distributions, response of vibrating systems, to name a few. Finding an efficient representation for any random process leads to encapsulation of information which makes it more convenient for a practical implementations, for instance, in a computational mechanics problem. For a single-parameter process such as spatial or temporal process, the eigenvalue decomposition of the covariance matrix leads to the well-known Karhunen-Loève (KL) decomposition. However, for multiparameter processes such as a spatio-temporal process, the covariance function itself can be defined in multiple ways. Here the process is assumed to be measured at a finite set of spatial locations and a finite number of time instants. Then the spatial covariance matrix at different time instants are considered to define the covariance of the process. This set of square, symmetric, positive semi-definite matrices is then represented as a third-order tensor. A suitable decomposition of this tensor can identify the dominant components of the process, and these components are then used to define a closed-form representation of the process. The procedure is analogous to the KL decomposition for a single-parameter process, however, the decompositions and interpretations vary significantly. The tensor decompositions are successfully applied on (i) a heat conduction problem, (ii) a vibration problem, and (iii) a covariance function taken from the literature that was fitted to model a measured wind velocity data. It is observed that the proposed representation provides an efficient approximation to some processes. Furthermore, a comparison with KL decomposition showed that the proposed method is computationally cheaper than the KL, both in terms of computer memory and execution time.

AMS subject classifications: 60G12, 65F99, 65D15

Key words: Random process, spatio-temporal process, tensor decomposition, uncertainty quantification, probabilistic mechanics.

*Corresponding author. *Email addresses:* dghosh@civil.iisc.ernet.in, debrajghosh@gmail.com (D. Ghosh), suryawanshi.anup@gmail.com (A. Suryawanshi)

1 Introduction

In a probabilistic treatment of uncertainties in analyzing and designing physical systems, random processes are used to describe and model various parameters and phenomena. Sources of these uncertainties can be attributed to insufficient data, variability in manufacturing process, error incurred during mathematical idealization of the problem, to mention a few. Let $(\Omega, \mathcal{F}, \mu)$ denote a probability space where Ω denotes the set of elementary events θ , \mathcal{F} denotes a σ -algebra on this event set, and μ denotes the probability measure. Let $x \in \mathbb{R}^d$ denote a spatial location where $d = 1, 2$ or 3 , and $t \in \mathbb{R}^+$ denote the time. Then the heterogeneity of Young's modulus of a solid can be modeled as a spatial random process or field $u(x, \theta)$, a time-varying excitation can be modeled as a temporal random field $u(t, \theta)$. Similarly the parameters that are dependent upon both space and time — such as a velocity field of a fluid in motion, temperature field, dynamic response of a large structure — can be modeled as spatio-temporal process $u(x, t, \theta)$. In this work a spatially and temporally discrete version of the real-valued processes is considered, that is, the processes are measured or evaluated at spatial locations $x_i: i = 1, 2, \dots, N_s$ and time instants $t_j: j = 1, 2, \dots, N_t$. Therefore, the spatio-temporal processes can now be written in the following matrix form

$$\mathbf{U}(\theta) = \begin{bmatrix} u(x_1, t_1, \theta) & u(x_1, t_2, \theta) & \cdots & u(x_1, t_{N_t}, \theta) \\ u(x_2, t_1, \theta) & u(x_2, t_2, \theta) & \cdots & u(x_2, t_{N_t}, \theta) \\ \vdots & \vdots & \vdots & \vdots \\ u(x_{N_s}, t_1, \theta) & u(x_{N_s}, t_2, \theta) & \cdots & u(x_{N_s}, t_{N_t}, \theta) \end{bmatrix} \in \mathbb{R}^{(N_s \times N_t)}. \quad (1.1)$$

The spatial and temporal processes can accordingly be expressed in a vector form. However, this explicit form, which is often needed for computational purpose, is not known in most practical cases. Either a few realizations of the process or some information about the covariance is known. Therefore a representation of the process needs to be found using this available information. In the current work, it is assumed that the only available information are the mean and spatial covariance for a set of time instants.

Let the mathematical expectation operator $\int_{\Omega} \cdot d\mu(\theta)$ be denoted as $\mathbb{E}\{\cdot\}$. Then, for a single-parameter process such as a spatial process $u(x, \theta)$ the covariance between two spatial locations x_1 and x_2 is defined as

$$\begin{aligned} & Cov(u(x_1, \theta), u(x_2, \theta)) \\ &= \mathbb{E}\{(u(x_1, \theta) - \bar{u}(x_1))(u(x_2, \theta) - \bar{u}(x_2))\}, \end{aligned} \quad (1.2)$$

with $\bar{u} \in \mathbb{R}$ denoting the mean of the process. A few largest eigenvalues and corresponding eigenvectors of this $(N_s \times N_s)$ symmetric positive-semidefinite covariance matrix hold a significant amount of information about the process $u(x, \theta)$. These eigenvectors serve as the set of bases in an approximate representation of this process, known as the Karhunen-Loève (KL) decomposition [1–5]. Similarly, the covariance matrix for a spatio-temporal process can be constructed with the elements as $Cov(u(x_i, t_k, \theta), u(x_j, t_l, \theta))$, $i, j = 1, 2, \dots, N_s$

and $k, l = 1, 2, \dots, N_t$, computed in the similar way as in Eq. (1.2). In this definition of covariance, no distinction is made between the notion of space and time [6]. However, in this case the size of the covariance matrix is $(N_s N_t \times N_s N_t)$, which amounts to storing and processing $N_s^2 N_t^2$ scalar quantities — and it could lead to a serious computational overhead. A generalized form of this space-time covariance is used in some applications [7, 8], where the distinction between the notion of space and time are preserved, at least in the continuous form of the process. A numerical treatment of this covariance will again involve $N_s^2 N_t^2$ scalar quantities.

For some spatio-temporal processes, it can be expected that for a given pair of spatial locations, the covariance at two distant time-instants may not be significant for some spatio-temporal processes. Expressing mathematically, $Cov(u(x_i, t_k, \theta), u(x_j, t_l, \theta))$ may be small for $t_k \neq t_l$, especially when $|t_k - t_l|$ is large. For these processes an alternative definition of covariance is given here. Accordingly, for any given time instant t_k the spatio-temporal process reduces to a spatial process, and the $(N_s \times N_s)$ covariance matrix of this spatial process can be computed using Eq. (1.2). This procedure is followed for $k = 1, 2, \dots, N_t$, and finally these N_t number of spatial covariance matrices are stored in a third-order tensor of size $(N_s \times N_s \times N_t)$ — which is a significant reduction in the size compared to $(N_s N_t \times N_s N_t)$. This covariance tensor is then decomposed as a sum of outer products of a set of vectors. Using this set of vectors as the bases, a new representation of the spatio-temporal process is proposed.

Similar to matrix decompositions, there's an active interest in finding several useful decompositions of higher order tensors. An excellent review can be found in [9] with an exhaustive list of references. In the current work the most relevant decompositions are (i) Rank-1 [10], (ii) CANDECOMP/PARAFAC (CP) [11–14], and (iii) Tucker [10, 15–17] decompositions. In all these decompositions the goal is to decompose — or to approximate when an exact decomposition does not exist — a tensor in terms of outer products of a set of vectors. For the current work, the Rank-1 decomposition is directly used to find the dominant component of the spatio-temporal process. Furthermore, another decomposition derived as a constrained version of both CP and Tucker decompositions is found to model the covariance of a few processes better than the Rank-1 decomposition. The vectors in these tensor decompositions are used here as the bases for a new representation of spatio-temporal processes. It is observed that this technique reduces the computational demand significantly, especially on memory. To comment on accuracy, it is observed here that for a heat conduction problem the probability density functions are retrieved accurately, whereas for a vibration problem only the covariance is retrieved accurately.

There exist a few different approaches to model spatio-temporal processes in different application areas. In [18], targeted at meteorological applications, only the spatial covariance is considered and a space-time separable representation is used — which is apparently similar to our approach. However, there the bases in the spatial domain are chosen as the KL eigenvectors of the spatial covariance at an arbitrary time instant and the elements of the temporal bases are chosen to be independent. For a process that is stationary in time and homogeneous in space, spectral decompositions can be used for

both space and time [19]. A few more works on KL decomposition of spatio-temporal processes can be found in [20,21]. One other extreme is to treat a spatio-temporal process effectively as a pure spatial process by assuming that different time instants represent different realizations of this spatial process [22]. In [23], the temporal variation of the spatial covariance of the response of a structure was studied in a random vibration context. In [24] a spatio-temporal process was approximated using the polynomial chaos expansion (PCE), and applied to a temperature dataset obtained from an oceanographic experiment. More about theory and applications of spatio-temporal processes and their covariances can be found in [25–30]. The novelty of the current work lies in (i) using tensor decomposition techniques in approximating the covariance of a spatio-temporal process, and (ii) subsequently proposing closed-form expressions for this process using the bases found by tensor decompositions. A few possible applications of this work could be in (a) sampling realizations of a spatio-temporal process for a given covariance kernel, (b) using the proposed closed-form expression as an input to a subsequent analysis, in a way similar to the KL representation of a spatial field being used in solving stochastic mechanics problems [1,3], (iii) obtaining a model for a spatio-temporal process from real data by estimating the bases of the proposed representation.

Recently the tensor decomposition has drawn attraction in the stochastic mechanics community. For instance, in [31] the tensor decomposition was used in solving a stochastic elliptic partial differential equation (PDE). In that work, the solution is expressed in PCE, followed by a Galerkin projection. The tensor decomposition was used to accelerate the solution of a linear algebraic system arising from this Galerkin projection.

This paper is organized as follows. In the following section several definitions of covariance of spatio-temporal processes are presented. Then tensor decompositions are discussed in the next section. The proposed representations of the processes based upon these tensor decompositions are presented next. KL representation of the spatio-temporal processes is also presented in the same section. Then a numerical study is conducted where the covariance decompositions and proposed process representations are implemented and tested for three problems. A comparison with the KL representation is also performed. Finally the conclusions are drawn.

2 Covariance of spatio-temporal processes

For a one-parameter process, the covariance is uniquely defined as Eq. (1.2). For a multiparameter such as a spatio-temporal process, the covariance can be defined in multiple ways. Two distinct definitions are given in this section. In this work, the second definition will be used which has a better physical interpretation than the first one. Also it leads to a significant amount of computational saving in modeling the process.

2.1 Covariance in space-time

One possible way to treat a spatio-temporal process is to ignore the distinction between the physical interpretations of space and time. In other words, the covariance between

two spatial locations at a fixed time instant has the same interpretation as the covariance between two time instants for a fixed spatial location. According to this view, the covariance between the process values measured at two spatial locations x_i and x_j and at two time instants t_1 and t_2 is defined as

$$\begin{aligned} & Cov(u(x_i, t_1, \theta), u(x_j, t_2, \theta)) \\ &= \mathbb{E}\{[u(x_i, t_1, \theta) - \bar{u}(x_i, t_1)][u(x_j, t_2, \theta) - \bar{u}(x_j, t_2)]\}. \end{aligned} \quad (2.1)$$

In a discrete setting, where the process is evaluated at total N_s spatial locations and at total N_t spatial notations, this covariance is represented as an $(N_s N_t \times N_s N_t)$ matrix Cov_{uu} — which will be referred to here as the spatio-temporal covariance matrix. To discretize this process in terms of a finite set of random variables, the KL expansion can be directly used [6]. However, in this treatment, in addition to the lack of physical interpretation, the amount of information processed here is significantly high: $N_s^2 N_t^2$ real numbers.

2.2 Temporal variation of spatial covariance

One other way to treat a spatio-temporal process where a significant reduction of the storage requirement of the covariance will be achieved is to acknowledge the distinction between the notion of time and space. Here, the covariance between two spatial locations is considered for a fixed time instant t , and is expressed as

$$\begin{aligned} & Cov(u(x_i, t, \theta), u(x_j, t, \theta)) \\ &= \mathbb{E}\{[u(x_i, t, \theta) - \bar{u}(x_i, t)][u(x_j, t, \theta) - \bar{u}(x_j, t)]\}. \end{aligned} \quad (2.2)$$

Therefore, for a given time instant t , the covariance between N_s spatial locations is expressed as an $(N_s \times N_s)$ real symmetric matrix $Cov_{uu}(t)$, with the $(i, j)^{th}$ element defined as above. Considering all N_t time instants, the covariance of this process is completely characterized by N_t number of $(N_s \times N_s)$ spatial covariance matrices, therefore the total storage demand is $N_s^2 N_t$ real numbers — lower by a factor N_t from the previous definition of covariance. In this work, this temporal variation of the spatial matrix will be studied. This collection of matrices will be represented as a third order tensor \mathcal{C} of size $(N_s \times N_s \times N_t)$ and decomposed accordingly.

3 Tensor decompositions and computation

Let \mathcal{A} denote a third order tensor of size $(I_1 \times I_2 \times I_3)$, with its elements denoted as \mathcal{A}_{ijk} ; $i = 1, \dots, I_1$; $j = 1, \dots, I_2$; $k = 1, \dots, I_3$. A few relevant decompositions of this tensor are discussed in this section. In this work although only the third-order tensors are used, the following discussion is valid for higher order tensors as well.

3.1 Tensor decompositions

For matrices, several powerful decompositions such as singular value decompositions (SVD), spectral decompositions are found to be extremely useful in numerous applications. Therefore, naturally it has been of interest of various researchers to find similar decompositions for the higher-order tensors. This is an active research area with a number of open problems, some interesting and important works can be found in [9, 10, 16] and references therein. Extensions of the matrix decompositions to tensors are never straightforward, and in most cases do not carry the same properties and physical interpretation from matrices to tensors. However, from an approximation point of view, some tensor decompositions are found to be useful for practical applications. Some of these decompositions — that will be of most relevance for the current work — are described here. In the following discussion although the word *decomposition* is used, in practice an exact decomposition may not be attainable — even may not exist, in many cases only an *approximation* to a desired decomposition form is attainable.

3.1.1 Rank-1 decomposition

A rank-1 tensor is defined as

$$\hat{\mathcal{A}} = \lambda \cdot \mathbf{v}^{(1)} \circ \mathbf{v}^{(2)} \circ \mathbf{v}^{(3)}, \quad (3.1)$$

where $\lambda \in \mathbb{R}$, $\mathbf{v}^{(i)} \in \mathbb{R}^{I_i}$ with unit-norm, $i = 1, 2, 3$, and \circ denotes the outer product of vectors. Here the hat $\hat{\cdot}$ over \mathcal{A} will denote an approximation — described below, and this notation is used for tensors with any rank. Writing Eq. (3.1) element-wise,

$$\hat{a}_{i,j,k} = \lambda v_i^{(1)} v_j^{(2)} v_k^{(3)}, \quad i = 1, \dots, I_1, \quad j = 1, \dots, I_2, \quad k = 1, \dots, I_3, \quad (3.2)$$

where $\hat{a}_{i,j,k}$ denotes the (i, j, k) -th element of the tensor $\hat{\mathcal{A}}$, $v_i^{(1)}, v_j^{(2)}, v_k^{(3)}$ denote i -th, j -th, k -th element of the vectors $\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \mathbf{v}^{(3)}$, respectively. A rank-1 tensor $\hat{\mathcal{A}}$ is referred to as the *best rank-1 approximation* [9, 10] of a given tensor \mathcal{A} if the quantity

$$er_T = \frac{\|\mathcal{A} - \hat{\mathcal{A}}\|_{Fr}}{\|\mathcal{A}\|_{Fr}} \quad (3.3)$$

is minimized over all the rank-1 tensors of size $(I_1 \times I_2 \times I_3)$. Here $\|\cdot\|_{Fr}$ denotes the Frobenius norm of a tensor defined as

$$\|\mathcal{A}\|_{Fr} = \sqrt{\sum_{i=1}^{I_1} \sum_{j=1}^{I_2} \sum_{k=1}^{I_3} a_{i,j,k}^2}. \quad (3.4)$$

Therefore, to achieve this approximation, one needs to find the scalar λ and three vectors $\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \mathbf{v}^{(3)}$.

3.1.2 CANDECOMP/PARAFAC (CP) decomposition

A generalization of the aforementioned rank-1 decomposition is arrived when the tensor is decomposed as the sum of R (> 1) rank-1 tensors, thereby leading to the expression [11–14]

$$\hat{\mathcal{A}} = \sum_{r=1}^R \lambda_r \cdot \mathbf{v}_r^{(1)} \circ \mathbf{v}_r^{(2)} \circ \mathbf{v}_r^{(3)}, \quad (3.5)$$

where $\lambda_r \in \mathbb{R}$, $\mathbf{v}_r^{(i)} \in \mathbb{R}^{I_i}$ with unit-norm, $r = 1, \dots, R$, $i = 1, 2, 3$. Writing element-wise,

$$\hat{a}_{i,j,k} = \sum_{r=1}^R \lambda_r v_{r,i}^{(1)} v_{r,j}^{(2)} v_{r,k}^{(3)}, \quad i = 1, \dots, I_1, \quad j = 1, \dots, I_2, \quad k = 1, \dots, I_3, \quad (3.6)$$

where $v_{r,i}^{(1)}, v_{r,j}^{(2)}, v_{r,k}^{(3)}$ denote the i -th, j -th, k -th element of the vectors $\mathbf{v}_r^{(1)}, \mathbf{v}_r^{(2)}, \mathbf{v}_r^{(3)}$, respectively. This decomposition was first proposed in [11], and is also often referred to as the CP decomposition, derived from its other names canonical decomposition (CANDECOMP) [13] and parallel factors (PARAFAC) [14]. The integer R is called as a rank of the tensor — this is one of the several types of tensor ranks. For a given tensor \mathcal{A} and an integer R , an approximation $\hat{\mathcal{A}}$ is sought so that the error defined in Eq. (3.3) is minimized. The basic approach of an algorithm for finding this approximation is presented in Section 3.2.

CP decomposition is unique only under certain conditions [9, 12]. Furthermore, existence of a CP decomposition of any given rank is not guaranteed. One other important aspect of this decomposition is that a truncation of a rank- R decomposition to a lower rank $R' < R$ with $|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_{R'}| \geq \dots \geq |\lambda_R|$ does not provide the best rank- R' approximation — unlike matrix SVD. Therefore care must be taken in choosing the rank R . However, when such a truncation is used the approximation error reduces monotonically with increasing R . That is, er_T for a rank R decomposition is less than er_T for a rank R' decomposition when $R > R'$.

3.1.3 Tucker decomposition

To explain the Tucker decomposition, the concept of the n -mode product of a tensor and a matrix should be defined first. Let \mathcal{A} denote an l -th order tensor, that is, $\mathcal{A} \in \mathbb{R}^{(I_1 \times I_2 \times \dots \times I_l)}$. The n -mode product \times_n , where $n = 1, 2, \dots, l$ of the tensor \mathcal{A} and a $(J \times I_n)$ matrix \mathbf{V} is denoted as $\mathcal{A} \times_n \mathbf{V}$, and is defined as a tensor of size $(I_1 \times \dots \times I_{n-1} \times J \times I_{n+1} \times \dots \times I_l)$ whose elements are

$$(\mathcal{A} \times_n \mathbf{V})_{i_1, \dots, i_{n-1}, j, i_{n+1}, \dots, i_l} = \sum_{i_n=1}^{I_n} a_{i_1, \dots, i_{n-1}, i_n, i_{n+1}, \dots, i_l} v_{j, i_n}, \quad j = 1, \dots, J, \quad (3.7)$$

where v_{j, i_n} denotes the $(j, i_n)^{th}$ element of the matrix \mathbf{V} . For instance, when \mathcal{A} is a third order tensor, its product \times_1 with a $(J \times I_1)$ matrix \mathbf{V} is a tensor of size $(J \times I_2 \times I_3)$ with the

elements

$$(\mathcal{A} \times_1 \mathbf{V})_{j,i_2,i_3} = \sum_{i_1=1}^{I_1} a_{i_1,i_2,i_3} v_{j,i_1}, \quad j=1, \dots, J. \quad (3.8)$$

Given this definition, the Tucker decomposition of an $(I_1 \times I_2 \times I_3)$ tensor \mathcal{A} is given as [15, 16]

$$\begin{aligned} \mathcal{A} &= \mathcal{G} \times_1 \mathbf{V}^{(1)} \times_2 \mathbf{V}^{(2)} \times_3 \mathbf{V}^{(3)} \\ &= \sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \sum_{i_3=1}^{I_3} g_{i_1,i_2,i_3} \cdot \mathbf{v}_{i_1}^{(1)} \circ \mathbf{v}_{i_2}^{(2)} \circ \mathbf{v}_{i_3}^{(3)}, \end{aligned} \quad (3.9)$$

where \mathcal{G} is a tensor of size $(I_1 \times I_2 \times I_3)$ — referred to as the *core tensor*, and $\mathbf{V}^{(1)}, \mathbf{V}^{(2)}, \mathbf{V}^{(3)}$ are unitary matrices of size $(I_1 \times I_1), (I_2 \times I_2), (I_3 \times I_3)$, respectively; $\mathbf{v}_j^{(i)}$ denotes the j^{th} column of the matrix $\mathbf{V}^{(i)}$. Although this decomposition — also referred to as the higher order singular value decomposition (HOSVD) — exists for any tensor [17], it may be very expensive to compute for large tensors. An alternative is to estimate an approximation $\hat{\mathcal{A}}$ as

$$\hat{\mathcal{A}} = \mathcal{G}_{red} \times_1 \mathbf{V}_{red}^{(1)} \times_2 \mathbf{V}_{red}^{(2)} \times_3 \mathbf{V}_{red}^{(3)}, \quad (3.10)$$

where the size of the core tensor \mathcal{G}_{red} is $(R_1 \times R_2 \times R_3)$, the matrices $\mathbf{V}_{red}^{(1)}, \mathbf{V}_{red}^{(2)}, \mathbf{V}_{red}^{(3)}$ are of size $(I_1 \times R_1), (I_2 \times R_2), (I_3 \times R_3)$, respectively. The integers R_i are chosen by the user and normally are much lower than I_i for all i . The approximation is sought for minimizing the error defined in Eq. (3.3).

Once again, as opposed to the matrix SVD, for a given R_1, R_2, R_3 any truncation of the expression (3.9) does not guarantee to give the best approximation in the form of Eq. (3.10). Therefore to achieve the best approximation one must compute directly Eq. (3.10). An algorithm for this computation is proposed in [10], where it is called as the rank- (R_1, R_2, \dots, R_N) decomposition because $R_i = \text{rank}_i(\hat{\mathcal{A}})$ for all i . Here $\text{rank}_i(\hat{\mathcal{A}})$ is defined as the rank of a matrix created using the vectors obtained by varying only the index i of $\hat{\mathcal{A}}$, keeping other indices fixed.

An important observation is that both rank-1 and CP decompositions can be seen as special structural forms of the Tucker decomposition. The CP decomposition follows when the core tensor \mathcal{G} is superdiagonal, that is, $g_{ijk} = 0$ except $i = j = k$. However, in CP, the matrices $\mathbf{V}^{(i)}$ constructed using the vector set $\{\mathbf{v}_r^{(i)}\}_{r=1}^R$ no longer need to be unitary. Moreover, in general the superdiagonalized form (that is, the exact CP decomposition) does not exist. The rank-1 decomposition follows from the CP decomposition when only one term is used.

3.2 Computation of tensor decompositions

As mentioned earlier, finding efficient methods for computing the tensor decompositions is an active area of research. In this work, methods based on alternating least-square

approach (ALS) are used. For CP decomposition the approach can be outlined as follows. Starting from an initial iterate, and freezing the sets $\{\mathbf{v}_r^{(2)}\}_{r=1}^R$ and $\{\mathbf{v}_r^{(3)}\}_{r=1}^R$, the set $\{\mathbf{v}_r^{(1)}\}_{r=1}^R$ is estimated by solving a least-square minimization problem. Next, $\{\mathbf{v}_r^{(2)}\}_{r=1}^R$ is similarly computed by freezing the other two sets, and then $\{\mathbf{v}_r^{(3)}\}_{r=1}^R$ is computed by freezing $\{\mathbf{v}_r^{(1)}\}_{r=1}^R$ and $\{\mathbf{v}_r^{(2)}\}_{r=1}^R$. This completes one iteration, the iterations are continued till the desired accuracy is attained. The λ_r -s are also updated within the iterations. Analogous methods are used for the Rank-1 and Tucker decompositions.

4 Representation of spatio-temporal processes based upon tensor decompositions

In the current work, the covariance is defined as the temporal variation of the spatial covariance, as described in Section 2.2. This covariance, as defined in Eq. (2.2), is treated as a third order tensor \mathcal{A} of size $(N_s \times N_s \times N_t)$ with the $(i, j, k)^{th}$ element as

$$\mathbb{E}\{(u(\mathbf{x}_i, t_k, \theta) - \bar{u}(\mathbf{x}_i, t_k, \theta))(u(\mathbf{x}_j, t_k, \theta) - \bar{u}(\mathbf{x}_j, t_k, \theta))\}. \quad (4.1)$$

As mentioned in the previous section, a suitable decomposition or an approximation $\hat{\mathcal{A}}$ will be used to find the dominant components of this covariance tensor \mathcal{A} — analogous to the eigenvalue or singular value decomposition of a matrix. The random process can then be represented using the bases of this decomposition. To arrive at this representation, first the process mentioned in Eq. (1.1) is written as

$$\mathbf{U}(\theta) = \bar{\mathbf{U}} + \mathbf{U}_{rand}(\theta), \quad (4.2)$$

where $\bar{\mathbf{U}}$ and $\mathbf{U}_{rand}(\theta)$ denote the mean and the zero-mean random parts, respectively. Then the proposed representation of the spatio-temporal process has the form

$$\mathbf{U}(\theta) = \bar{\mathbf{U}} + \mathbf{U}_{rand}(\theta) = \bar{\mathbf{U}} + \sum_{i=1}^{i=N} \sqrt{\lambda_i} \phi_i \psi_i^T \eta_i, \quad \phi_i \in \mathbb{R}^{N_s}, \quad \psi_i \in \mathbb{R}^{N_t}. \quad (4.3)$$

Here ϕ_i and ψ_j are the basis vectors corresponding to spatial and temporal components, respectively, and to be estimated using a decomposition of the covariance tensor. The zero-mean random variables η_i are functions of θ , and a method is needed to estimate these random variables. Note that if the tensor decomposition is not exact, then $\mathbf{U}(\theta)$ and $\mathbf{U}_{rand}(\theta)$ should be replaced by $\hat{\mathbf{U}}(\theta)$ and $\hat{\mathbf{U}}_{rand}(\theta)$, respectively. However, this notational strictness is relaxed to keep the discussion simpler. It will be shown later in this section that the process in Eq. (4.3) yields the same covariance tensor \mathcal{A} . Depending upon the type of the tensor decomposition used, these bases ϕ_i and ψ_j may vary. Here two decompositions are used, details of which is given later in this section.

Difference with the KL decomposition: In a KL decomposition, the covariance is expressed as a $(N_s N_t \times N_s N_t)$ matrix \mathbf{Cov}_{uu} with the elements defined in Eq. (2.1). Let

us construct an $N_s N_t$ dimensional vector $\mathbf{u}(\theta)$ by arranging all the columns of $\mathbf{U}(\theta)$ in Eq. (1.1) in a single column. That is,

$$\mathbf{u}(\theta)^T = \{u(\mathbf{x}_1, t_1, \theta) u(\mathbf{x}_2, t_1, \theta) \cdots u(\mathbf{x}_{N_s}, t_1, \theta) u(\mathbf{x}_1, t_2, \theta) \cdots u(\mathbf{x}_{N_s}, t_{N_t}, \theta)\}^T. \quad (4.4)$$

Then

$$\mathbf{Cov}_{uu} = \mathbb{E}\{\mathbf{u}(\theta)\mathbf{u}(\theta)^T\}. \quad (4.5)$$

Let $\lambda_i^{KL}, \phi_i^{KL}$ denote the i^{th} eigenpair of this covariance matrix, where the eigenvalues are ordered in numerically ascending way. Accordingly, the KL representation of the process $\mathbf{u}(\theta)$ is written as

$$\mathbf{u}(\theta) = \bar{\mathbf{u}} + \sum_{i=1}^N \sqrt{\lambda_i^{KL}} \phi_i^{KL} \eta_i, \quad (4.6)$$

where $\bar{\mathbf{u}}$ denotes the mean of the process and η_i are independent, zero-mean random variables, N denotes the number of terms retained. Distributions of η_i are found by using orthogonality of the eigenvectors ϕ_i^{KL} . Therefore the main difference between the tensor decomposition based approach and KL approach lies in treating the covariance, which is a tensor in the first case and a matrix in the second case. Note that the covariance matrix (in KL) contains N_t times more elements than the covariance tensor (in the tensor decomposition). As will be seen later in this section, unlike in KL, the choice of the random variables η_i is not unique in the tensor decomposition.

4.1 Rank-1 tensor decomposition

First the spatio-temporal covariance tensor \mathcal{C} is decomposed into its rank-1 decomposition, as defined in Eq. (3.1). For a given time instant t , the spatial covariance $\mathbf{Cov}_{uu}(t)$ is a symmetric matrix of size $(N_s \times N_s)$. Due to this symmetry, the rank-1 decomposition of the covariance tensor \mathcal{C} yields $\mathbf{v}^{(1)} = \mathbf{v}^{(2)}$, where $\mathbf{v}^{(1)}, \mathbf{v}^{(2)}$ are defined in Eq. (3.1). When the random process is expressed as Eq. (4.3) based on this decomposition, the natural choice of the spatial basis becomes $\phi = \mathbf{v}^{(1)}$. The temporal basis can be chosen as $\psi = \sqrt{\mathbf{v}^{(3)}}$, where the square-root is calculated term-wise. To avoid ambiguity in the sign convention here only the *principal square root* is taken, which means that only the positive roots are considered. According to this choice of bases the expression of the random process becomes

$$\mathbf{U}(\theta) = \bar{\mathbf{U}} + \sqrt{\lambda} \mathbf{v}^{(1)} \sqrt{\mathbf{v}^{(3)}}^T \boldsymbol{\eta}, \quad \mathbf{v}^{(1)} \in \mathbb{R}^{N_s}, \quad \mathbf{v}^{(3)} \in \mathbb{R}^{N_t}, \quad \mathbb{E}\{\boldsymbol{\eta}\} = 0. \quad (4.7)$$

The covariance of the process can be retrieved from this expression as follows. Recall from Eq. (1.1) that the $(i, k)^{th}$ element of the matrix $\mathbf{U}(\theta)$ is $u(\mathbf{x}_i, t_k, \theta)$. Accordingly, the spatial covariance between the points \mathbf{x}_i and \mathbf{x}_j at a given time instant t_k is estimated from Eq. (4.7) as

$$\begin{aligned} & \mathbb{E}\{(u(\mathbf{x}_i, t_k, \theta) - \bar{u}(\mathbf{x}_i, t_k, \theta))(u(\mathbf{x}_j, t_k, \theta) - \bar{u}(\mathbf{x}_j, t_k, \theta))\} \\ & = \lambda v_i^{(1)} v_j^{(1)} v_k^{(3)} \mathbb{E}\{\eta^2\}. \end{aligned} \quad (4.8)$$

In this equation the left hand side is defined in Eq. (2.2), and if $\mathbb{E}\{\eta^2\} = 1$ then the right hand side equals the term defined in Eq. (3.2) — which is nothing but the $(i,j,k)^{th}$ element in the decomposed tensor $\hat{\mathcal{A}}$. Recall that the error er_T incurred in approximating the covariance tensor is measured by Eq. (3.3). This implies that if η is chosen as a random variable with mean as zero and variance as one, then the covariance of the random process in Eq. (4.7) converges to the covariance defined in Eq. (2.2) as the approximation error er_T reduces. A method to estimate this variable is now needed, which can be done as follows. Multiplying both sides of Eq. (4.7) by $\mathbf{v}^{(1)T}$ results in

$$\mathbf{v}^{(1)T} \mathbf{U}_{rand}(\theta) = \sqrt{\lambda} \mathbf{v}^{(1)T} \mathbf{v}^{(1)} \sqrt{\mathbf{v}^{(3)T}} \eta, \quad \mathbf{v}^{(1)} \in \mathbb{R}^{N_s}, \quad \mathbf{v}^{(3)} \in \mathbb{R}^{N_t}. \quad (4.9)$$

From this expression, the random coefficient η is computed as

$$\eta = \frac{\mathbf{v}^{(1)T} \mathbf{U}_{rand}(\theta) \mathbf{1}}{\sqrt{\lambda} \sqrt{\mathbf{v}^{(3)T}} \mathbf{1}}, \quad (4.10)$$

where $\mathbf{1}$ denotes a vector of length N_t with all elements as 1. The choice of the estimator (4.10) is not unique. For instance, Eq. (4.9) could be post multiplied by any other N_t dimensional vector than $\mathbf{1}$ which would yield in a different estimator.

From Eq. (4.7) it is clear that for the process to be real-valued, it is necessary for each term of the vector $\lambda \mathbf{v}^{(3)}$ to be non-negative. As was mentioned in Section 2.2, for any given time instant, the spatial covariance is positive definite. Furthermore, the matrix $\mathbf{v}^{(1)} \mathbf{v}^{(1)T}$ is symmetric and positive-definite for any choice of the real vector $\mathbf{v}^{(1)}$. Therefore each element of the product $\lambda \mathbf{v}^{(3)}$ — which corresponds to a time instant — must be non-negative. This argument also guarantees that the denominator in Eq. (4.10) is non-zero (except the trivial case when the either λ or $\mathbf{v}^{(3)}$ is zero).

4.2 A constrained decomposition derived from the CP and Tucker

When the CP decomposition is used for finding the dominant components of the covariance tensor, the following issues were faced in numerical studies:

1. The bases are not orthogonal in general.
2. All terms of the vector $\lambda_r \mathbf{v}_r^{(3)}$ are not positive, it will lead to a complex-valued process.
3. $\mathbf{v}_r^{(1)} \neq \mathbf{v}_r^{(2)}$, that is, the symmetry is not preserved.

Whereas in the Tucker decomposition although the orthogonality and symmetry of the bases are maintained, the challenges remaining are:

1. For a given i_3 the vectors $\mathbf{v}_{i_3}^{(3)}$ in Eq. (3.9) may contain both positive and negative entries, leading to a complex valued process.

2. Finding an expression for the process itself is a complicated task due to the non-zero off-diagonal terms in the core tensor \mathcal{G} .

To address these issues a constrained version of both Tucker and CP decompositions is considered here as

$$\hat{\mathcal{C}} = \left[\sum_{r=1}^R \lambda_r \cdot \mathbf{v}_r^{(1)} \circ \mathbf{v}_r^{(2)} \right] \circ \mathbf{v}^{(3)}, \quad (4.11)$$

with an added constraint that the vectors in the set $\{\mathbf{v}_r^{(1)}\}_{r=1}^R$ are orthogonal. This decomposition can either be viewed as (i) the CP decomposition (3.5) with the added orthogonality constraint and $\mathbf{v}_r^{(3)}$ being unaltered for any r , or (ii) as the Tucker decomposition (3.10) with $R_3 = 1$. Computationally, the Tucker decomposition algorithm is used with the constraint $R_3 = 1$. Note that the choice of $R_3 = 1$ may not lead to the exact decomposition of the covariance tensor, that is, $er_T = 0$ may not always hold. However, this choice led to a simplification in choosing an expression for the random process, as proposed in Eq. (4.12). Note that the spatial covariance matrices are symmetric. For such tensors, it is proved in [17] (Theorem 3) that $\mathbf{v}_r^{(1)} = \mathbf{v}_r^{(2)}$ for all r . Taking advantage of this equality, the following expression of the random process is proposed:

$$\mathbf{U}(\theta) = \bar{\mathbf{U}} + \sum_{r=1}^R \sqrt{\lambda_r} \mathbf{v}_r^{(1)} \sqrt{\mathbf{v}^{(3)T}} \eta_r, \quad \mathbf{v}_r^{(1)} \in \mathbb{R}^{N_s}, \quad \mathbf{v}^{(3)} \in \mathbb{R}^{N_t}, \quad (4.12)$$

where η_r are independent, zero-mean random variables. Similar to the Rank-1 tensor decomposition described earlier in this section, the covariance of the process can be retrieved from this expression as follows. Now the spatial covariance between the points \mathbf{x}_i and \mathbf{x}_j at a given time instant t_k is estimated from Eq. (4.12) as

$$\begin{aligned} & \mathbb{E}\{(u(\mathbf{x}_i, t_k, \theta) - \bar{u}(\mathbf{x}_i, t_k, \theta))(u(\mathbf{x}_j, t_k, \theta) - \bar{u}(\mathbf{x}_j, t_k, \theta))\} \\ &= \sum_{r=1}^R \lambda_r v_{r,i}^{(1)} v_{r,j}^{(1)} v_{r,k}^{(3)} \mathbb{E}\{\eta_r^2\}. \end{aligned} \quad (4.13)$$

If the random variables η_r chosen such as $\mathbb{E}\{\eta_r^2\} = 1$ for all r , then the right hand side of this equation equals the $(i, j, k)^{th}$ element in the decomposed tensor in Eq. (4.11). Therefore, once again, the covariance of the random process in Eq. (4.12) converges to the covariance defined in Eq. (2.2) as the approximation error er_T reduces.

Using the orthogonality among the bases $\{\mathbf{v}_r^{(1)}\}_{r=1}^R$ the random coefficients η_r are computed as

$$\eta_r = \frac{\mathbf{v}_r^{(1)T} \mathbf{U}_{rand}(\theta) \mathbf{1}}{\sqrt{\lambda_r} \sqrt{\mathbf{v}^{(3)T}} \mathbf{1}}. \quad (4.14)$$

Once again, this choice is not unique as other conditions can be used to estimate these coefficients.

5 Numerical study

The tensor decompositions of the covariance and associated proposed representations are numerically tested in this section. Three problems are considered here: a one-dimensional heat equation, a vibration problem, and a readily available covariance function for wind velocity. A comparison with the KL decomposition is performed for accuracy and computational cost. The tensor decomposition related computations are performed using the MATLAB Tensor Toolbox [32].

5.1 Heat equation

The first problem considered is the heat equation in one-dimension, stated as

$$\frac{\partial u(x,t)}{\partial t} = \kappa \frac{\partial^2 u(x,t)}{\partial x^2}, \quad 0 \leq x \leq 2\pi, \quad t \geq 0 \quad (5.1)$$

with the initial condition as $u(x,0) = \sin(\vartheta x)$ and periodic boundary conditions. Here $u(x,t)$ denotes the temperature and κ denotes the thermal diffusivity. Two sources of randomness are considered: (i) the initial condition is random with ϑ being a random variable represented as $\vartheta = 3 + 0.3 \zeta_1$ and (ii) the thermal diffusivity κ is another random variable represented as $\kappa = 0.5 + 0.1 \zeta_2$, where ζ_1 and ζ_2 are two uniformly distributed independent random variables varying between 0 and 1. For this problem, Eq. (5.1) is first solved numerically. The spatial domain is discretized using 101 equally spaced grid points, and 200 time-steps of 0.001 seconds each are considered. This resolution implies $N_s = 101$ and $N_t = 201$. The covariance is estimated using a Monte Carlo simulation with 10^4 realizations. Accordingly, for each realization a deterministic heat equation is solved and the solution is stored. Finally the covariance defined in Eq. (2.2) is estimated by statistical averaging. The mean temperature is plotted in Fig. 1.

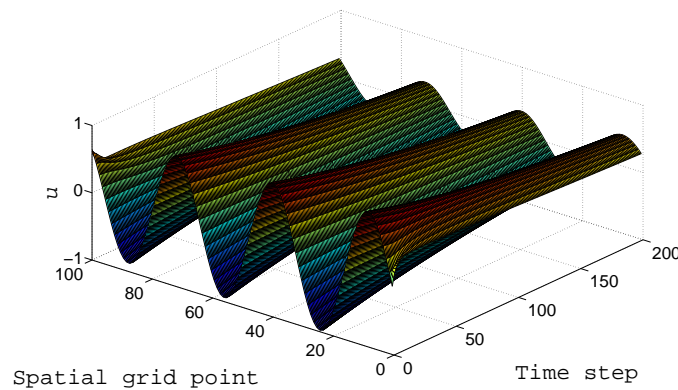
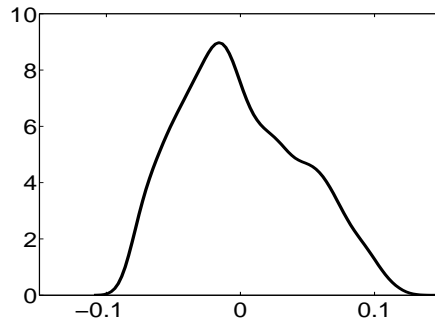
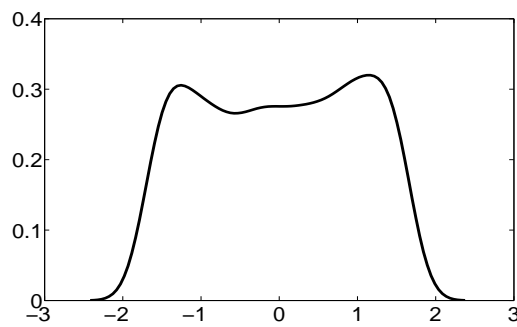


Figure 1: Heat problem: Mean of the field $u(x,t)$.

Figure 2: Heat problem: Probability density function of η_1 .Figure 3: Heat problem: Probability density function of η_2 .

Next this $(101 \times 101 \times 201)$ covariance tensor is decomposed using the constrained decomposition (4.11). Therefore $\mathbf{v}_r^{(1)}, \mathbf{v}_r^{(2)} \in \mathbb{R}^{101}$, $\mathbf{v}^{(3)} \in \mathbb{R}^{201}$. The approximation error er_T defined in Eq. (3.3) is found to be 0.13 for $R=1$, and 0.077 for $R=2$. The error remains almost constant thereafter as R grows. For a three-term approximation of the process using Eq. (4.12) with $R=2$, the random variables η_1 and η_2 are estimated using Eq. (4.14). The probability density functions (PDF) of these two variables are plotted in Figs. 2 and 3. One interesting observation here is that although the random variables ϑ and κ are uniformly distributed, distribution of the variable η_1 departs significantly from a uniform distribution. To check the accuracy of the approximation, for arbitrarily chosen spatial location and time instant the random temperature is sampled using Eq. (4.12) and the PDF is plotted in Fig. 4. A comparison with the PDF estimated using the original process (found during the Monte Carlo simulation involving the Heat equation) shows that the proposed tensor decomposition based representation offers a good approximation of the original process. Also the PDF is estimated using KL decomposition. For this purpose the covariance matrix given in Eq. (4.5) is first estimated using a Monte Carlo simulation. Upon finding the dominant eigenvalues and eigenvectors, the process is regenerated using Eq. (4.6), with $N=2$. The PDF of the temperature thus estimated is also plotted in Fig. 4 and a good match is observed.

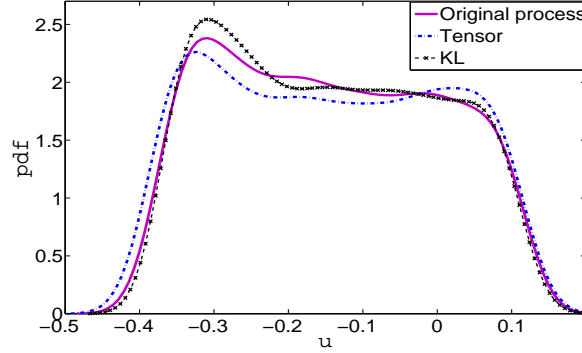


Figure 4: Heat problem: Probability density function of u at arbitrarily chosen point and time instant. The regenerated processes using both tensor decomposition and KL decomposition show a fair agreement with the original process.

However, the main difference between the proposed tensor decomposition based approximation and KL decomposition lies in the computational cost, which is presented in Table 1. Here only the cost of decomposition is reported since in a practical application the covariance is not intended to be estimated from a simulation. One such application will be given in the third example in this paper, where the covariance is evaluated from an analytical expression. From Table 1 it is observed that the KL decomposition is computationally more expensive than the tensor decomposition, especially the memory requirement becomes prohibitive.

Table 1: Comparison of computational cost for the heat problem.

Discretization		Tensor		KL	
Spatial (N_s)	Temporal (N_t)	Memory	CPU time	Memory	CPU time
101	201	16 MB	2.2 sec	3.3 GB	10 sec

5.2 Vibration of a plate

The next problem considered is a forced linear vibration problem of a square plate of size $3m \times 3m$ with a uniform thickness of $5mm$. Total seven random variables are used to model the uncertainty of the system: (i) Young's modulus, (ii) mass density, (iii) three spatial dimensions, (iv) Poisson's ratio, and (v) loading. Except the excitation all other random parameters are modeled as positive-valued random variables. For instance, the Young's modulus is modeled as

$$E(\xi_1) = \bar{E} + \frac{\sigma_E}{\sqrt{2}}(\xi_1^2 - 1), \quad (5.2)$$

where \bar{E} denotes the mean value, σ_E denotes the standard deviation, and ξ_1 denotes a standard normal random variable. To ensure the positivity of the Young's modulus

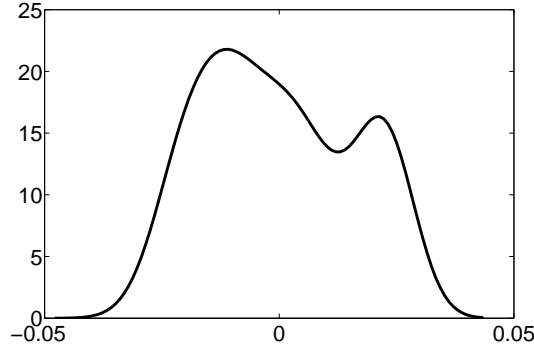


Figure 5: Vibration problem: Probability density function of η .

$E(\xi_1)$ for all values of ξ_1 , one constraint needs to be added as $\frac{\sigma_E}{\sqrt{2}} < \bar{E}$. Similarly the other five random parameters are modeled using a set of independent standard normal random variables $\{\xi_i\}_{i=2}^{i=6}$. The assumed mean values are as follows. Young's modulus: 2×10^5 MPa, mass density: 7860 Kg/m^3 , Poisson's ratio: 0.3. The standard deviations of these parameters are chosen as 5% of their mean. The plate is modeled using four-noded bicubic rectangular conforming elements with sixteen degrees-of-freedom (DOF) per element. Total sixteen elements are used to discretize the plate. The random excitation is modeled as

$$f(t) = f_1 \sin(\omega_{f_1} t) + f_2 \sin(\omega_{f_2} t) + \xi_7 f_3 \sin(\omega_{f_3} t), \quad (5.3)$$

where ξ_7 denotes a standard normal random variable independent of the set $\{\xi_i\}_{i=1}^{i=6}$, f_1, f_2, f_3 , denote arbitrarily chosen amplitude vectors of loading components, the frequencies $\omega_{f_1}, \omega_{f_2}, \omega_{f_3}$ are chosen arbitrarily but remain within the range of first four natural frequencies of the system.

For this problem, the covariance tensor of the response is estimated using the a Monte Carlo simulation and the tensor decompositions are computed. It is observed that the Rank-1 decomposition defined in Eq. (3.1) gave a good approximate with an error of 0.003. Therefore, the expansion in Eq. (4.3) with $N = 1$ can be expected to yield a good approximation of the spatio-temporal response process. However, when the η was estimated using Eq. (4.10), the recomputed process showed a different variance from the original process. The PDF of η is plotted in Fig. 5, the variance of theta is far less than one, whereas it should be equal to one as was discussed in Section 4.1. Therefore it is concluded from this example that although a tensor decomposition may yield a good approximation of the covariance tensor, the regenerated process using Eq. (4.3) with the η estimated using Eq. (4.10) may not yield a good approximation of the process.

Further investigation on the results from the Monte Carlo simulation shows that the for a given spatial location, the type of distribution of the response varies with time. However, the representation (4.3) with $N = 1$ is unable to capture the change in the type of distribution with time, as it depends only upon one random variable η . Furthermore,

the formula (4.10) yields a time-averaged random variable. This also could be a reason for η not being of unit variance (as the average of a set of zero-mean unit-variance random variables is not always a variable with unit-variance). However, further detailed exploration on suitability of either the representation (4.3) or the formula (4.10) is needed and modifications are required for more robust applicability.

5.3 Sampling from a readily available covariance function

In the previous two examples the target was to demonstrate and test the validity of tensor decompositions and the proposed process representations. Now an example is considered which could be a potential application of the present work. Here the realizations of a random process are sampled from the known covariance function. Similar sampling problem occurs in stochastic mechanics examples, where samples of random material properties are drawn from a KL expansion starting from a spatial covariance function. However, here the goal is to sample from a spatio-temporal process.

A readily available analytical covariance model from the literature [8] is used for this sampling example. A generic form of this covariance function is

$$C(\mathbf{x}_{dist}; t_{diff}) = \frac{\sigma^2}{f(|t_{diff}|^2)^{d/2}} \times g\left(\frac{\|\mathbf{x}_{dist}\|^2}{f(|t_{diff}|^2)}\right), \quad (5.4)$$

where \mathbf{x}_{dist} denotes a vector joining two spatial locations in the d -dimensional spatial domain, t_{diff} denotes the lag between two time instants, f and g are two functions determining the actual form of the covariance kernel, σ^2 is the variance of the spatio-temporal process. Based on this generic form, using an available dataset on Irish wind [33], the following correlation model of the velocity (technically a *velocity-measure*, derived after performing a series of pre-processing steps on the actual velocity, including a square-root transformation) was developed in [8]:

$$C(\mathbf{x}_{dist}; t_{diff}) = \begin{cases} (0.901|t_{diff}|^{1.544} + 1)^{-1} & \text{if } \mathbf{x}_{dist} = \mathbf{0}, \\ 0.968(0.901|t_{diff}|^{1.544} + 1)^{-1} \\ \quad \times \exp\left(-\frac{0.00134\|\mathbf{x}_{dist}\|}{(0.901|t_{diff}|^{1.544} + 1)^{0.305}}\right) & \text{otherwise.} \end{cases} \quad (5.5)$$

This model is used with a spatial domain being a circular area of radius 300 kilometers over a 24 hours timespan. The spatial covariance from this model at an arbitrarily chosen time instant is plotted in Fig. 6. For $N_s=50$ and $N_t=10$, first the elements of the covariance matrix Cov_{uu} and the covariance tensor \mathcal{C} are evaluated directly from Eq. (5.5). Therefore, unlike the previous two examples, a Monte Carlo simulation is not needed to estimate the covariance matrix and tensor. The constrained decomposition given in Eq. (4.11) is then carried out for the covariance tensor. For different values of R , the approximation error er_T is plotted in Fig. 7. To compare the results with KL decomposition, an eigenvalue decomposition of the covariance matrix is then carried out. For a given truncation level

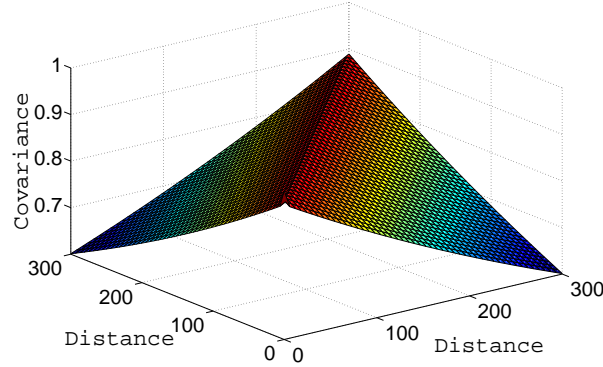


Figure 6: Spatial covariance of the wind velocity covariance model at an arbitrarily chosen time instant.

of the KL expansion in Eq. (4.6), the relative error between the actual covariance and the covariance of the generated process is measured in a similar way as er_T by replacing the tensor norm by a matrix norm. Decay of this error with respect to the levels of truncation is also plotted in Fig. 7. From this figure it is observed that both the decompositions follow a similar pattern of convergence. However, the error in the tensor decomposition is lower than the error in the KL decomposition when the level of truncation is low.

The PDF of the velocity at arbitrarily chosen spatial location and time instant is estimated using the KL expansion in Eq. (4.6) and the proposed representation in Eq. (4.12), and are plotted as in Fig. 8. In the KL expansion 10 terms are included whereas in the proposed representation only 5 terms are included as they show equal level of accuracy in Fig. 7. The distribution is chosen as Gaussian due to the lack of further information. Also the mean part is left out in this plot, that is, only the random fluctuation is considered. Inclusion of the mean will shift both the PDFs rightward by an equal amount. The PDFs estimated from these two methods are found to be in good agreement. The relative computational cost is presented in Table 2. Here the computational time includes both computing the covariance matrix or tensor, and their decompositions. Once again it is observed that the proposed tensor-based representation is computationally cheaper. Note that with an increase in N_t , the increase in the cost in KL decomposition is significantly higher than the tensor decomposition. This difference is due to the fact that the number of elements in the covariance matrix varies quadratically with N_t whereas for covariance tensor it varies linearly.

Table 2: Comparison of computational cost for the wind velocity covariance model.

Discretization		Tensor		KL	
Spatial (N_s)	Temporal (N_t)	Memory	CPU time	Memory	CPU time
50	10	0.2 MB	less than 1 sec	2 MB	~3 sec
50	50	1 MB	~ 2 sec	50 MB	70 sec

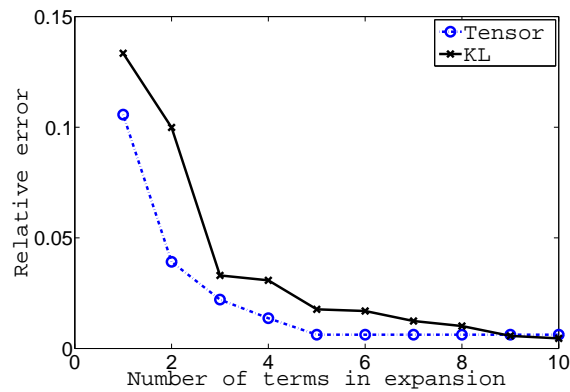


Figure 7: Convergence of KL and tensor approximations for the covariance model of the wind data.

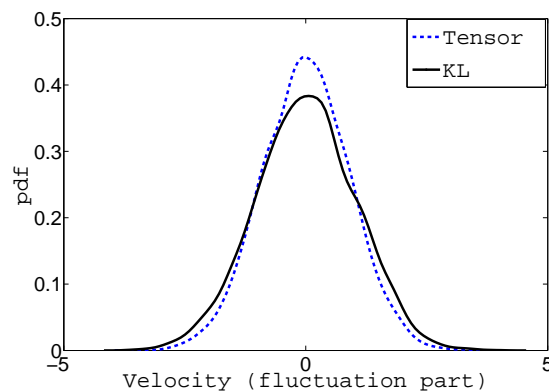


Figure 8: Comparison of PDFs generated by the tensor decomposition and KL expansion of the wind velocity covariance model, only the fluctuating part is considered. The time instant and the spatial location are chosen arbitrarily.

6 Concluding remarks

It is observed that for some random spatio-temporal processes, if the covariance is defined as the temporal variation of the spatial covariance, the resulting tensor can be decomposed into a low-rank tensor decomposition. This property allows representing the processes in a proposed series form, analogous to the KL expansion for a one-parameter process. In this work, for the heat equation, both the tensor decomposition and the proposed series representation of the process yielded results with sufficient accuracy. Also this approach is found to be computationally cheaper than the KL approach. However, for the vibration problem, only the tensor decomposition worked and not the representation of the process. To demonstrate a practical application of the proposed method, an example is presented where the realizations of a spatio-temporal process are successfully sampled from a readily available covariance function. These successful implemen-

tations of this representation opens up avenues of developing approximations to spatio-temporal processes in a computationally efficient way. Further studies are needed to identify the characteristics of the processes or physical phenomena that will make a process suitable for this tensor decomposition and series representation. Estimation of the scalar-valued random coefficients in this series representation also needs further exploration as the choice of this coefficient is not unique.

Acknowledgments

The authors would like to thank Indian Institute of Science and the Board of Research in Nuclear Sciences (BRNS) grant no. 2011/36/41-BRNS/1977 for their financial support.

References

- [1] R. Ghanem and P. D. Spanos. *Stochastic Finite Elements: A Spectral Approach*. Revised Edition, Dover Publications, 2003.
- [2] H. L. Van Trees. *Detection, Estimation and Modulation Theory, Part I*. Wiley, 2001.
- [3] D. Xiu and G. E. Karniadakis. A new stochastic approach to transient heat conduction modeling with uncertainty. *International Journal of Heat and Mass Transfer*, 46:4681–4693, 2003.
- [4] H. Sang and J. Z. Huang. A full scale approximation of covariance functions for large spatial data sets. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 74(1):111–132, 2011.
- [5] J. Yang, D. Zhang, and Z. Lu. Stochastic analysis of saturated-unsaturated flow in heterogeneous media by combining Karhunen-Loève expansion and perturbation method. *Journal of Hydrology*, 294:18–38, 2004.
- [6] C. A. Schenk and G. I. Schueller. *Uncertainty Assessment of Large Finite Element Systems*. Springer Berlin/Heidelberg, 2005.
- [7] M. L. Stein. Space-time covariance functions. *Journal of the American Statistical Association*, 100(469):310–321, 2005.
- [8] T. Gneiting. Nonseparable, stationary covariance functions for space-time data. *Journal of the American Statistical Association*, 97(458):590–600, 2002.
- [9] T. G. Kolda and B. W. Bader. Tensor decompositions and applications. *SIAM Review*, 51(3):455–500, 2009.
- [10] L. De Lathauwer, B. De Moor, and J. Vandewalle. On the best rank-1 and rank- (R_1, R_2, \dots, R_n) approximation of higher-order tensors. *SIAM Journal on Matrix Analysis and Applications*, 21(4):1324–1342, 2000.
- [11] F. L. Hitchcock. The expression of a tensor or a polyadic as a sum of products. *Journal of Mathematical Physics*, 6:164–189, 1927.
- [12] J. B. Kruskal. Three-way arrays: Rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics. *Linear Algebra and its Applications*, 18:95–138, 1977.
- [13] J. D. Carroll and J. J. Chang. Analysis of individual differences in multidimensional scaling via an N-way generalization of Eckart-Young decomposition. *Psychometrika*, 35(3):283–319, 1970.

- [14] R. A. Harshman. Foundations of the PARAFAC procedure: Models and conditions for an explanatory multi-modal factor analysis. UCLA Working Papers in Phonetics, Available at <http://publish.uwo.ca/~harshman/wpppfac0.pdf>, 16:1–84, 1970.
- [15] L. R. Tucker. Implications of factor analysis of three-way matrices for measurement of change. Problems in Measuring Change, Edited by Harris CW, University of Wisconsin Press, pages 122–137, 1963.
- [16] L. R. Tucker. Some mathematical notes on three-mode factor analysis. Psychometrika, 31(3):279–311, 1966.
- [17] L. De Lathauwer, B. De Moor, and J. Vandewalle. A multilinear singular value decomposition. SIAM Journal on Matrix Analysis and Applications, 21(4):1258–1278, 2000.
- [18] C. K. Wikle. Spatio-temporal methods in climatology. Encyclopedia of Life Support Systems (EOLSS), 2003.
- [19] H. L. Van Trees. Detection, Estimation and Modulation Theory, Part IV. Wiley, 2002.
- [20] M. D. Graham and I. G. Kevrikedis. Alternative approaches to the Karhunen-Loève decomposition for model reduction and data analysis. Computers & Chemical Engineering, 20(5):495–506, 1996.
- [21] S. Bellizzi and R. Sampaio. Poms analysis of randomly vibrating systems obtained from Karhunen-Loève expansion. Journal of Sound and Vibration, 297:774–793, 2006.
- [22] D. J. Nott and W. T. M. Dunsmuir. Estimation of nonstationary spatial covariance structure. Biometrika, 89(4):819–829, 2002.
- [23] C. W. S. To. Time-dependent variance and covariance of responses of structures to non-stationary random excitations. Journal of Sound and Vibration, 93(1):135–156, 1984.
- [24] S. Das, R. Ghanem, and S. Finette. Polynomial chaos representation of spatio-temporal random fields from experimental measurements. Journal of Computational Physics, 228:8726–8751, 2009.
- [25] C. Ma. Recent developments on the construction of spatio-temporal covariance models. Stochastic Environmental Research and Risk Assessment, 22:39–47, 2008.
- [26] F. Bruno, P. Guttorp, P. D. Sampson, and D. Cocchi. A simple non-separable, non-stationary spatiotemporal model for ozone. Environ Ecol Stat, 16:515–529, 2009.
- [27] S. D. Iaco, D. Posa, and D. E. Myers. Characteristics of some classes of space-time covariance functions. Journal of Statistical Planning and Inference, 143:2002–2015, 2013.
- [28] C. Ma. Families of spatio-temporal stationary covariance models. Journal of Statistical Planning and Inference, 116:489–501, 2003.
- [29] C. Ma. Linear combinations of space-time covariance functions and variograms. IEEE Transactions on Signal Processing, 53(3):857–864, 2005.
- [30] A. Kolovos, G. Christakos, D. T. Hristopulos, and M. L. Serre. Methods for generating non-separable spatiotemporal covariance models with potential environmental applications. Advances in Water Resources, 27:815V830, 2004.
- [31] B. N. Khoromskij and C. Schwab. Tensor-structured Galerkin approximation of parametric and stochastic elliptic PDEs. Research Report No. 2010-04, Swiss Federal Institute of Technology Zurich, 2010.
- [32] B. W. Bader and T. G. Kolda. MATLAB Tensor Toolbox. <http://csmr.ca.sandia.gov/~tgkolda/TensorToolbox/>.
- [33] J. Haslett and A. E. Raftery. Space-time modelling with long-memory dependence: Assessing Ireland’s wind-power resource. Applied Statistics, 38:1–50, 1989.