Advanced Modeling and Simulation in Engineering Sciences
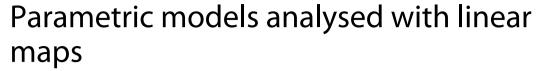
RESEARCH ARTICLE

Open Access

Parametric models analysed with linear maps

Hermann G. Matthies[1*] and Roger Ohayon[2]

*Correspondence:
wire@tu-bs.de
[1] Institute of Scientific
Computing, Technische
Universität Braunschweig, 38092
Brunswick, Germany
Full list of author information is
available at the end of the article

Abstract

Parametric entities appear in many contexts, be it in optimisation, control, modelling of random quantities, or uncertainty quantification. These are all fields where reduced order models (ROMs) have a place to alleviate the computational burden. Assuming that the parametric entity takes values in a linear space, we show how is is associated to a linear map or operator. This provides a general point of view on how to consider and analyse different representations of such entities. Analysis of the associated linear map in turn connects such representations with reproducing kernel Hilbert spaces and affine-/linear-representations in terms of tensor products. A generalised correlation operator is defined through the associated linear map, and its spectral analysis helps to shed light on the approximation properties of ROMs. This point of view thus unifies many such representations under a functional analytic roof, leading to a deeper understanding and making them available for appropriate analysis.

Keywords: Parametric models, Reduced order models, Reproducing kernel Hilbert space, Correlation, Factorisation, Spectral decomposition, Representation

Mathematics Subject Classification: 35B30, 37M99, 41A05, 41A45, 41A63, 60G20, 60G60, 65J99, 93A30

Introduction

Many mathematical and computational models depend on parameters. These may be quantities which have to be optimised during a design, or controlled in a real-time setting, or these parameters may be uncertain and represent uncertainty present in the model. Such parameter dependent models are usually specified in such a way that an input to the model, e.g. a process or a field, depends on these parameters. In an analogous fashion, the output or the "state" of the model will depend on those parameters. Any of these entities may be called a parametric model. To make things a bit more specific, we look at an example: Consider the parametric entities in the following equation:

$$A(\mu; u) = f(\mu). \tag{1}$$

Here $A(\zeta(\mu); \cdot) : \mathcal{V} \to \mathcal{V}$ is a possibly nonlinear opertor from the Hilbert space $\mathcal{U}$ into itself, dependent on $\zeta(\mu) \in \mathcal{Z}$—a vector in another Hilbert space $\mathcal{Z}$ used to specify the system— $u \in \mathcal{V}$ is the state of the system described by A, whereas $f(\mu) \in \mathcal{V}$ is the excitation resp.

action on the system. The *parameters* $\mu \in \mathcal{M}$ are elements of some admissible parameter set $\mathcal{M}$. Here $f(\mu)$ and $\zeta(\mu)$ are two examples of such parametric entities; and as the whole equation depends on μ, we assume that for each $\mu \in \mathcal{M}$ the system Eq. (1) will be well-posed and allow for the state $u(\mu)$ to also be a unique function of the parameters—another example of a parametric entity.

When one has to do computations with a system such as Eq. (1), one needs computational *representations* of the parametric entities such as the "inputs" $f(\mu)$, $\zeta(\mu)$, and also the to be determined *state* $u(\mu)$, the "output". Let us denote any of such generic entities as $r(\mu)$; then one seeks a computational expression to compute $r(\mu)$ for any given parameter $\mu \in \mathcal{M}$. The first question which has to be addressed is how to choose "good co-ordinates" on the parameter set $\mathcal{M}$. With this we mean scalar functions $\xi_m : \mathcal{M} \to \mathcal{R}$, so that the collection and specification of all $\{\xi_m(\mu)\}_{m=1,\dots,M}$ will on one hand specify the particular $\mu \in \mathcal{M}$ as regards the system Eq. (1), and on the other hand be a computational handle for the parametric entities $r(\mu)$, which now can be expressed as $r(\xi_1, \dots, \xi_M)$. Often the parameter set is already given as $\mathcal{M} \subseteq \mathbb{R}^d$, so that $\mu = [\eta_1, \dots, \eta_d] \in \mathbb{R}^d$ are directly given co-ordinates, and the co-ordinate functions η_k may directly serve as co-ordinates. But often, and not only, but especially, when $d \in \mathbb{N}$ is a large number, it may be advisable to choose other co-ordinates ξ_m, which should be free of possible constraints and be as "independent" as possible. This is usually part of finding a good computational representation for $r(\mu)$, and will be addressed as part of our analysis. One may term this as a *re-parametrisation* of the problem.

The second question to be addressed is the actual number of degrees-of-freedom needed to describe the behaviour of the system Eq. (1) through some finite-dimensional approximation or discretisation. Often the initial process of discretisation produces a first approximation with a large number of degrees-of-freedom; this initial computational model is often referred to as a full-scale or high-fidelity model. For many computational purposes it is necessary to reduce the number of degrees-of-freedom in the computational model in order to be able to carry out the computations involved in an optimisation or uncertainty quantification in a acceptable amount of time; such computational models are then termed reduced order models (ROMs). If the high-fidelity model is a parametric model, the same is required from the ROM $r_a(\mu) \approx r(\mu)$.

The question of how to produce ROMs for specific kinds of systems like Eq. (1) is an important one, and is the subject of extensive current research. For the general subject of model order reduction there is an excellent collection of recent work in [40] and survey in [11], as well as an introductory text in [39]; see also [20,34] for important contributions. Besides these general considerations, in the present case parametrised ROMs are of particular interest. The general survey [5] covers the literature up to 2015 very well, as well as the later one [10], which is concerned mainly with uncertainty quantification. Excellent collections on the topic of parametrised ROMs are contained in [4,6]. A recent systematic monograph is [27], and important recent contributions are e.g. [9,45,46]. Machine learning and so-called data-driven procedures have also been used in this context, see the recent contributions in [28,41–44], but this is at the very beginning.

Here a particular point of view will be taken for the analysis—not to be found in the recent literature just surveyed—namely the identification of a parametric entity with a linear mapping defined on the dual space, which is introduced in the "Parametric models and linear maps" section. This idea has been around for a long time, and has surfaced

mostly when the "strong" notion of a concept has to be replaced by a "weaker" one. In this sense one may see the present point of view as a generalisation of the view of distributions or generalised functions as linear mappings [21,23]. They were used to define weak notions of random quantities [24], and some of the present ideas are also contained in [33]. In some sense these ideas are already contained in [30]—see also the English translation [31]—and may most probably be found even earlier. The reason on why to approach the subject in this way is that for linear operators there is a host of methods which can be used for their analysis, and it puts all such parametric entities under one "roof".

It has to be pointed out though, that through the identification of a parametric object with a linear map, it is *linear* methods of analysis which are the main interest here, and they being used to study and analyse those parametric objects. It is one of the objectives of this note to show the power of linear methods in this kind of analysis. On the other hand, beyond the question of "generalising" the concept of a parametric objects, another interesting subject is their approximation—often termed as parametric reduced order models (parametric ROMs) as described to above. Although there are nonlinear methods to construct parametric ROMs such as those alluded to in the previous paragraph, here the emphasis is on linear methods and their use in constructing ROMs. But even though such a ROM may be constructed by nonlinear methods, this does not mean that linear methods have no rôle to play, as they can be used in the analysis of such a ROM and its approximation properties.

Here we want to explain the basic abstract framework and how it applies to ROMs. This present work is a continuation of [35–38]. The general theory was shown in [38], and here the purpose is primarily to give an introduction into this kind of analysis, which draws strongly on the spectral analysis of self-adjoint operators (e.g. [16,22,24]), and an overview on how to use it in the analysis of ROMs. This is the topic of "Correlation and representation" section. Coupled systems and their ROMs are the focus of [37], and [35] is a short note on how this is used for random fields and processes. In the "Structure preservation" section some examples of such refinements of the basic concept are given.

As will be seen, it is very natural to deal with tensor products in this topic of parametrised entities. In the form of the *proper generalised decomposition* (PGD) this idea has been explained and used in [2,11–13,17]. The topic of tensor approximations [26] turns out to be particularly relevant here, and recently new connections between such approximations and machine learning with deep networks have been exposed [14,32]. In "Conclusion" section we conclude with a recapitulation of the main ideas.

Parametric models and linear maps

This is a gentle introduction and short recap of the developments in [35,37,38], where the interested reader may find more detail. To start, and to take a simple motivating example, one could think of a scalar function $r(x, \mu)$, defined on some set $\mathcal{X}$, which depends on some parameters in a set $\mathcal{M}$—in other words a parametric function. In what follows, this function will be viewed as a mapping

$$r : \mathcal{M} \to \mathcal{U}, \tag{2}$$

so that for each value of $\mu \in \mathcal{M}$ the function $w := r(\cdot, \mu) \in \mathcal{U}$ is a scalar function $w(x)$ defined on the set $\mathcal{X}$.

To simplify further and make everything finite-dimensional, assume that we are only interested in four positions in $\mathcal{X}$, namely $x_1, x_2, x_3, x_4 \in \mathcal{X}$, or, alternatively and even simpler, that $\mathcal{X} = \{x_1, x_2, x_3, x_4\}$ has only four elements, and finally for the sake of simplicity, that the parameter set has only three elements $\mathcal{M} = \{\mu_1, \mu_2, \mu_3\}$. Then one can arrange all the possible values of $r(x, \mu)$ with the abbreviation $r_{i,j} = r(x_i, \mu_j)$, $(i = 1, \ldots, 3, j = 1, \ldots, 4)$ in the following matrix:

$$\boldsymbol{R} = [r_{i,j}]_{i=1,\ldots,3,j=1,\ldots,4} \in \mathbb{R}^{3 \times 4}.$$

It is obvious that knowing the function $r(x, \mu)$ is equivalent with knowing the matrix $\boldsymbol{R}$. As a matrix $\boldsymbol{R}$ obviously corresponds to a linear mapping from $\mathcal{U} = \mathbb{R}^4$ to $\mathcal{R} = \mathbb{R}^3$, and one has for any $\boldsymbol{u} = [u(x_1), u(x_2), u(x_3), u(x_4)]^\mathsf{T} = [u_1, u_2, u_3, u_4]^\mathsf{T} \in \mathcal{U}$ that

$$\boldsymbol{R}\boldsymbol{u} = [\phi(\mu_1), \phi(\mu_2), \phi(\mu_3)]^\mathsf{T} = \phi \in \mathbb{R}^3 = \mathcal{F} = \mathbb{R}^\mathcal{M}, \tag{3}$$

where $\phi(\mu_i) = \sum_{j=1}^4 r_j(\mu_i)u_j$—a *weighted* average of $r(\cdot, \mu_i)$—is a scalar function $\phi \in \mathcal{F} = \mathbb{R}^\mathcal{M}$ in the linear space $\mathcal{F}$ of scalar functions $\mathbb{R}^\mathcal{M}$ on the parameter set $\mathcal{M}$. If we denote the function of Eq. (2) in this case by $\boldsymbol{r}(\cdot)$, which for every $\mu \in \mathcal{M}$ is an element $\boldsymbol{w} := \boldsymbol{r}(\mu) \in \mathcal{U} = \mathbb{R}^4$, then the weighted average $\phi \in \mathcal{R} = \mathbb{R}^3$ in Eq. (3) obviously satisfies $\phi(\mu_i) = \boldsymbol{r}(\mu_i)^\mathsf{T}\boldsymbol{u}$, so that

$$\boldsymbol{R}\boldsymbol{u} = [\boldsymbol{r}(\mu_1)^\mathsf{T}\boldsymbol{u}, \boldsymbol{r}(\mu_2)^\mathsf{T}\boldsymbol{u}, \boldsymbol{r}(\mu_3)^\mathsf{T}\boldsymbol{u}]^\mathsf{T} = \langle \boldsymbol{r}(\cdot), \boldsymbol{u} \rangle_\mathcal{U} = [\phi(\mu_1), \phi(\mu_2), \phi(\mu_3)]^\mathsf{T} = \phi. \tag{4}$$

Obviously, knowing $\boldsymbol{R}$ is the same as knowing $\boldsymbol{R}\boldsymbol{u}$ for every $\boldsymbol{u} \in \mathcal{U} = \mathbb{R}^4$—actually a basis in $\mathcal{U}$ would suffice—which in turn is the same as knowing $\langle \boldsymbol{r}(\cdot), \boldsymbol{u} \rangle_\mathcal{U}$ for every $\boldsymbol{u} \in \mathcal{U}$.

The point to take away from this simple example is that the parametric function $r(x, \mu)$, where for each parameter value $\mu \in \mathcal{M}$ one has $r(\cdot, \mu) \in \mathcal{U}$ in some linear space $\mathcal{U}$—of functions on $\mathcal{X}$ in this case—is equivalent to a linear map

$$\boldsymbol{R} : \mathcal{U} \to \mathcal{F}$$

into a space $\mathcal{F} \subseteq \mathbb{R}^\mathcal{M}$ of scalar functions on the parameter set $\mathcal{M}$.

It is now easy to see how to generalise this further to cases where the set $\mathcal{X}$ or $\mathcal{M}$ or both have infinitely many values, and even further to a case where the vector space of functions $\mathcal{U}$ just has an inner product, say given by some integral, so that for $u, v \in \mathcal{U}$ one has

$$\langle u, v \rangle_\mathcal{U} = \int_\mathcal{X} u(x)v(x) \, \mathsf{m}(\mathrm{d}x)$$

with some measure m on $\mathcal{X}$. Then for each parameter $\mu \in \mathcal{M}$ one has $r(\cdot, \mu) \in \mathcal{U}$, a function on $\mathcal{X}$, or in other words an element of the linear space $\mathcal{U}$. In this case one defines the linear map

$$\tilde{R} : \mathcal{U} \to \mathcal{F} \subseteq \mathbb{R}^\mathcal{M}$$

as

$$\tilde{R} : u \mapsto \int_\mathcal{X} u(x)r(x, \mu) \, \mathsf{m}(\mathrm{d}x) = \langle r(\cdot, \mu, u \rangle_\mathcal{U} =: \phi(\mu) \in \mathcal{F} \subseteq \mathbb{R}^\mathcal{M},$$

which is a linear map from $\mathcal{U}$ onto a linear space of scalar functions $\phi \in \mathcal{F} \subseteq \mathbb{R}^\mathcal{M}$ on the parameter set $\mathcal{M}$.

This then is almost the general situation, where one views $r : \mathcal{M} \to \mathcal{V}$ as a map from the parameters $\mu \in \mathcal{M}$, where $\mathcal{M}$ may be some arbitrary set, into a topological vector space $\mathcal{V}$. One then defines a linear map

$$\tilde{R} : \mathcal{V}^* \to \mathcal{F} \subseteq \mathbb{R}^{\mathcal{M}}$$

from the dual space $\mathcal{V}^*$ onto a space of scalar functions $\mathcal{F}$ on $\mathcal{M}$ by

$$\mathcal{V}^* \ni u \mapsto \tilde{R}u = \langle r(\mu) \mid u \rangle_{(\mathcal{V},\mathcal{V}^*)} =: \phi(\mu) \in \mathcal{F} \subseteq \mathbb{R}^{\mathcal{M}},$$

where $\langle \cdot \mid \cdot \rangle_{(\mathcal{V},\mathcal{V}^*)}$ is the duality pairing between $\mathcal{V}$ and its dual space $\mathcal{V}^*$. For the following exposition of the main ideas we shall take a slightly less general situation by assuming for the sake of simplicity that the linear space $\mathcal{V}$ is in fact a separable Hilbert space with an inner product $\langle \cdot, \cdot \rangle_{\mathcal{U}}$, and use this in the usual manner to identify it with its dual.

Associated linear map

So with a vector-valued map $r : \mathcal{M} \to \mathcal{V}$, one defines the corresponding associated linear map $\tilde{R} : \mathcal{V} \to \mathcal{F}$ as

$$\forall v \in \mathcal{V} : \tilde{R}v = \langle r(\mu), v \rangle_{\mathcal{U}} =: \phi(\mu) \in \mathcal{F} \subseteq \mathbb{R}^{\mathcal{M}}. \tag{5}$$

Obviously only the Hilbert subspace $\mathcal{U} = \mathrm{cl}(\mathrm{span}\, r(\mathcal{M})) \subseteq \mathcal{V}$ actually reached by the map r is interesting, whereas $\mathcal{U}^\perp = \ker \tilde{R} \subseteq \mathcal{V}$ is not. Hence from now on we shall only look at $r : \mathcal{M} \to \mathcal{U}$, and additionally assume that $\mathcal{U} = \mathrm{cl}(\mathrm{span}\, r(\mathcal{M}))$, or in other words, that the vectors $\{r(\mu) \mid \mu \in \mathcal{M}\} = r(\mathcal{M})$ form a total set in $\mathcal{U}$. The map $\tilde{R}$ is thus formally redefined as

$$\tilde{R} : \mathcal{U} \to \mathbb{R}^{\mathcal{M}}. \tag{6}$$

Again, in the linear space $\mathbb{R}^{\mathcal{M}}$ of *all* scalar functions on $\mathcal{M}$, only the part $\mathcal{F} = \mathrm{im}\, \tilde{R} = \tilde{R}(\mathcal{U})$ is interesting.

Allow here a little digression, to point out similarities and analogies to other connected concepts. First, on the parameter set $\mathcal{M}$, where up to now no additional mathematical structure was used, we now have the linear space $\mathcal{F}$. This can be viewed as a first step to introduce some kind of "co-ordinates" on the set $\mathcal{M}$, and is in line with many other constructs where potentially complicated sets are characterised by algebraic constructs, such as groups or vector spaces for e.g. homology or cohomology. Even if from the outset the parameter set $\mathcal{M} \subseteq \mathbb{R}^m$ is given as some subset of some $\mathbb{R}^m$ and therefore has already coordinates, these may not be *good* ones, and as we shall see, it may be worthwhile to contemplate *re-parametrisations*, i.e. choosing some $\phi_k \in \mathcal{F}$ as "co-ordinates". These real valued functions are in general of course not "real co-ordinates", as they only distinguish what is being felt by the parametric object r.

Reproducing kernel Hilbert space

The second concept to touch on comes from the idea to use the function space $\mathcal{F}$ in place of span $r(\mathcal{M})$: As is easy to see, the map in Eq. (6) is injective, hence invertible on its image $\mathcal{F} = \mathrm{im}\tilde{R} = \tilde{R}(\mathcal{U})$, and this may be used to define an inner product on $\mathcal{F}$ as

$$\forall \phi, \psi \in \mathcal{F} \quad \langle \phi, \psi \rangle_{\mathcal{R}} := \langle \tilde{R}^{-1}\phi, \tilde{R}^{-1}\psi \rangle_{\mathcal{U}}, \tag{7}$$

and to denote the completion of $\mathcal{F}$ with this inner product by $\mathcal{R} = \mathrm{cl}\,\mathcal{F} \subseteq \mathbb{R}^{\mathcal{M}}$. One immediately obtains that $\tilde{R}^{-1}$ is a bijective isometry between $\mathrm{span}\,r$ and $\mathcal{F}$, hence extends to a *unitary* map $\bar{R}^{-1}$ between $\mathcal{U}$ and $\mathcal{R}$, and the same hold for $\tilde{R}$, the extension being denoted by $\bar{R}$.

Given the maps $r : \mathcal{M} \to \mathcal{U}$ and $\bar{R} : \mathcal{U} \to \mathcal{R}$, one may define the *reproducing kernel* [7,29] given by $\varkappa(\mu_1, \mu_2) := \langle r(\mu_1), r(\mu_2) \rangle_{\mathcal{U}}$. It is straightforward to verify that $\varkappa(\mu, \cdot) \in \mathcal{F} \subseteq \mathcal{R}$, and $\mathrm{span}\{\varkappa(\mu, \cdot) \mid \mu \in \mathcal{M}\} = \mathcal{F}$, as well as the reproducing property $\phi(\mu) = \langle \varkappa(\mu, \cdot), \phi \rangle_{\mathcal{R}}$ for all $\phi \in \mathcal{F}$. Another way of stating this reproducing property is to say that the linear map $\mathcal{R} \ni \phi(\mu_1) \mapsto \langle \varkappa(\mu_2, \cdot), \phi \rangle_{\mathcal{R}} = \phi(\mu_2) \in \mathcal{R}$ for all $\phi \in \mathcal{R}$ is the identity $I_{\mathcal{R}}$ on $\mathcal{R}$. An abstract way of putting this using the adjoint $\bar{R}^* = \bar{R}^{-1}$ of the unitary map $\bar{R}$ would be to note that that map is in fact $\bar{R}\bar{R}^* = \bar{R}\bar{R}^{-1} = I_{\mathcal{R}}$.

With the *reproducing kernel Hilbert space* (RKHS) $\mathcal{R}$ one can build a first representation and thus obtain a relevant "co-ordinate system" for $\mathcal{M}$. As $\mathcal{U}$ is separable, it has a Hilbert basis or complete orthonormal system (CONS) $\{y_k\}_{k \in \mathbb{N}}$. As $\bar{R}$ is unitary, the set $\{\varphi_k = \bar{R}y_k\}_{k \in \mathbb{N}}$ is a CONS in $\mathcal{R}$.

With this, the unitary operator $\bar{R}$, its adjoint or inverse $\bar{R}^* = \bar{R}^{-1}$, and the parametric element $r(\mu)$ become [38]

$$\bar{R} = \sum_k \varphi_k \otimes y_k; \quad \text{i.e.} \quad \bar{R}(u)(\cdot) = \sum_k \langle y_k \mid u \rangle_{\mathcal{U}} \varphi_k(\cdot), \tag{8}$$

$$\bar{R}^* = \bar{R}^{-1} = \sum_k y_k \otimes \varphi_k; \quad r(\mu) = \sum_k \varphi_k(\mu) y_k = \sum_k \varphi_k(\mu)\,\bar{R}^*\varphi_k. \tag{9}$$

Observe that the relations Eqs. (8), (9) exhibit the tensorial nature of the representation mapping. One sees that *model reductions* may be achieved by choosing only subspaces of $\mathcal{R}$, i.e. spanned by a—typically finite—subset of the CONS $\{\varphi_k\}_k$. Furthermore, the representation of $r(\mu)$ in Eq. (9) is *linear* in the new "parameters" φ_k.

Coherent states

The third concept one should mention in this context is the one of *coherent states*, e.g. see [1,3]. In this development from quantum theory, these quantum states were initially introduced as eigenstates of certain operators, and the name refers originally to their high coherence, minimum uncertainty, and quasi classical behaviour. What is important here is that the idea has been abstracted, and represents overcomplete sets of vectors or frames $\{r(\mu) \mid \mu \in \mathcal{M}\}$ in a Hilbert space $\mathcal{U}$, which depend on a parameter $\mu \in \mathcal{M}$ from a locally compact measure space. This space often has more structure, e.g. a Lie group, and the coherent states are connected with group representations in the unitary group of $\mathcal{U}$, i.e. if $\mu \mapsto U(\mu) \in \mathscr{L}(\mathcal{U})$ is a unitary representation, the coherent states may be defined by $r(\mu) = U(\mu)w$ for some $w \in \mathcal{U}$. There are usually further requirements like weak continuity for the map $\mathcal{M} \ni \mu \mapsto r(\mu) \in \mathcal{U}$, and that these coherent states form a *resolution of the identity*, in that one has (weakly)

$$I_{\mathcal{U}} = \int_{\mathcal{M}} r(\mu) \otimes r(\mu)\,\varpi(\mathrm{d}\mu),$$

where ϖ is a measure on $\mathcal{M}$—naturally defined on some σ-algebra of subsets of $\mathcal{M}$, a detail which needs no further mention here. We shall leave this topic here, and come back to similar representations later, but note in passing the tensor product structure under

the integral. The above requirement of the resolution of the identity may sometimes be too strong, and one often falls back to the case of RKHS discussed above.

Reduced models

As was noted in the introduction, the emphasis is here on linear methods of analysis. The nonlinear construction of ROMs is outside the scope of this paper. Accordingly, in the following "Correlation and representation" section, only *linear* methods of constructing ROMs $r_a(\mu)$ which offer an approximate version of the full parametric objects $r_a(\mu) \approx r(\mu)$ will be touched upon. One possibility of producing such a ROM was already mentioned above by letting the sum in Eq. (9) run over fewer terms.

To actually construct a ROM, usually auxiliary information on what is important in the approximation will be used, see "Correlation and representation" section, where a second inner product will be introduced on the space $\mathbb{R}^{\mathcal{M}}$ of scalar functions of $\mathcal{M}$ for this purpose. Then the essentially same associated linear map R of a parametric object $r(\mu)$ is defined formally in Eq. (10). But whatever the method of construction for the ROM, assume now that $\mathcal{M} \ni \mu \mapsto r_a(\mu) \in \mathcal{U}$ is such an approximate or *reduced order model* (ROM) of $r(\mu)$. The linear methods to be explained later still can be used in the *analysis* of the approximation properties of the ROM. As any parametric map ($\mathcal{M} \to \mathcal{U}$), the ROM $r_a(\mu)$ thus has an associated linear map R_a, defined similarly as in Eq. (5), or rather Eq. (10).

As the associated linear maps carry all the relevant information, the analysis of both the original parametric object $r(\mu)$, and the comparison and analysis of accuracy of the approximation $r_a(\mu)$, can be carried out in terms of the associated linear maps R and R_a. Thus, to analyse the difference between $r(\mu)$ and $r_a(\mu)$, one may look at the difference between the the associated linear maps, R and R_a.

A slightly different but related way is to look at the error $r_\delta(\mu) = r(\mu) - r_a(\mu)$. This error is again a parametric map ($\mathcal{M} \to \mathcal{U}$), and thus has an associated linear map R_δ defined analogously as in Eq. (10). One may then use linear methods such as the ones presented here to analyse the linear map R_δ. A moments thought shows the close relation between these two approaches, as obviously $R_\delta = R - R_a$.

Possible generalisations

It is generally assumed here that the parametric object $r : \mathcal{M} \to \mathcal{U}$ is a *mapping*, i.e. there is only *one* value $r(\mu) \in \mathcal{U}$ for each $\mu \in \mathcal{M}$. This assumption is in line with assuming that in Eq. (1) the problem is *well-posed* for each $\mu \in \mathcal{M}$, this makes the state or solution $u(\mu)$ again a function of $\mu \in \mathcal{M}$. There are several interesting situations where this assumption may no longer hold, and some of these are briefly sketched here.

One such situation is the branching of solutions, like e.g. buckling or other instabilities, or spontaneous symmetry breaking. One way to look at such a situation is to realise that the set of parameters $\mathcal{M}$ is not "good" in the vicinity of such occurrences. In the description of these parameters one then has an unfolding in the sense of *catastrophe theory*, in that the solution $u(\mu)$ of e.g. Eq. (1) lies in this case actually on a manifold, and the parameter set is not "good" in the sense that it is *not* a valid coordinate system here, and one needs a re-parametrisation. The usual way of attack for such a situation is again through the analysis of *linear* maps, like e.g. the derivative $D_\mu u(\mu)$ in the case of differentiable dependence.

Instead of the simple vector space $\mathcal{U}$ one would have to consider a manifold where $\mathcal{U}$ or a subspace could be a tangent space.

Another situation where there is an apparent loss of uniqueness is when phenomena like irreversible behaviour—a typical simple case of this is plasticity—or hysteresis like the magnetisation of ferro-magnetic materials occur. To take plasticity as a simple but typical example, the model described by Eq. (1) would be the equilibrium of a mechanical system, $f(\mu)$ would be the loading, and $u(\mu)$ would be the displacement. The parameter $\mu \in \mathcal{M}$ could be envisaged like a time-like variable describing the loading program. It is then well known that the displacement $u(\mu)$ alone does not describe the complete state of the system, and that one may need additional so-called phenomenological or internal variables $q(\mu)$ to describe the state of the system, and in addition an evolution law for these internal variables. Similar remarks apply in the case of hysteresis, like The system state would then better be described by $r(\mu) = (u(\mu), q(\mu))$, and this would be the parametric object to be analysed, and where one would like to produce ROMs for.

One topic which becomes important in these situations, and which is touched on briefly in the "Structure preservation" section, is structure preservation. This means that the ROM system should in principle display certain salient features of the full model. In the case of solution branching and catastrophes described earlier that could mean that the ROM should be a able to reproduce these kind of behaviours. Similarly, in the case of irreversibility and hysteresis the ROM would be required to in principle reproduce such behaviour which involves energy storage and dissipation. The detailed treatment of such behaviours is beyond the scope of the present work, but the general principles laid out here are obviously equally applicable.

Correlation and representation

In what was detailed up to now in the previous "Parametric models and linear maps" section with regard to the RKHS, was that the structure of the Hilbert space was carried reproduced on the subspace $\mathcal{R} \subseteq \mathbb{R}^{\mathcal{M}}$ of the full function space. In the remarks about coherent states one could already see an additional structure, namely a measure ϖ on $\mathcal{M}$. This measure structure can be used to define the subspace $\mathcal{A} := \mathrm{L}_0(\mathcal{M}, \varpi)$ of measurable functions, as well as its Hilbert subspace of square-integrable functions $\mathcal{Q} := \mathrm{L}_2(\mathcal{M}, \varpi)$ with associated inner product

$$\langle \phi, \psi \rangle_{\mathcal{Q}} := \int_{\mathcal{M}} \phi(\mu)\psi(\mu)\, \varpi(\mathrm{d}\mu).$$

We shall simply assume here that there is a Hilbert space $\mathcal{Q} \subseteq \mathbb{R}^{\mathcal{M}}$ of functions with inner product $\langle \cdot, \cdot \rangle_{\mathcal{Q}}$, which may or may not come from an underlying measure space. The associated linear map $\tilde{R} : \mathcal{U} \to \mathcal{R}$, essentially defined in Eq. (5) with range the RKHS $\mathcal{R}$, will now be seen as a map $R : \mathcal{U} \to \mathcal{Q}$ into the Hilbert space $\mathcal{Q}$, i.e. with a different range with different inner product $\langle \cdot, \cdot \rangle_{\mathcal{Q}}$ from the RKHS inner product $\langle \cdot, \cdot \rangle_{\mathcal{R}}$ on $\mathcal{R}$. One may view this inner product as a way to tell what is important in the parameter set $\mathcal{M}$: functions ϕ with large $\mathcal{Q}$-norm are considered more important than those where this norm is small. The map $R : \mathcal{U} \to \mathcal{Q}$ is thus generally not unitary any more, but for the sake of simplicity, we shall assume that it is a densely defined closed operator, see e.g. [16]. As it may be only densely defined, it is sometimes a good idea to define R through a densely defined bilinear

form in $\mathcal{U} \otimes \mathcal{Q}$:

$$\forall u \in \mathrm{dom}R, \phi \in \mathcal{Q} : \langle Ru, \phi\rangle_{\mathcal{Q}} := \langle\langle r(\cdot), u\rangle_{\mathcal{U}}, \phi\rangle_{\mathcal{Q}}. \tag{10}$$

Following [33,35,37,38], one now obtains a densely defined map C in $\mathcal{U}$ through the densely defined bilinear form, in line with Eq. (10):

$$\forall u, v : \quad \langle Cu, v\rangle_{\mathcal{U}} := \langle Ru, Rv\rangle_{\mathcal{Q}}. \tag{11}$$

The map $C = R^*R$—observe that now the adjoint is w.r.t. the $\mathcal{Q}$-inner product—may be called the *"correlation"* operator, and is by construction self-adjoint and positive, and if R is bounded resp. continuous, so is C.

In the above case that the $\mathcal{Q}$-inner product comes from a measure, one has from Eq. (11)

$$\langle Cu, v\rangle_{\mathcal{U}} = \int_{\mathcal{M}} \langle r(\mu), u\rangle_{\mathcal{U}}\langle r(\mu), v\rangle_{\mathcal{U}} \, \varpi(\mathrm{d}\mu), \text{ i.e. } C = R^*R = \int_{\mathcal{M}} r(\mu) \otimes r(\mu) \, \varpi(\mathrm{d}\mu).$$

This is reminiscent of what was required for coherent states. But it also shows that if ϖ were a probability measure—i.e. $\varpi(\mathcal{M}) = 1$—with the usual expectation operator

$$\mathbb{E}(\phi) := \int_{\mathcal{M}} \phi(\mu) \, \varpi(\mathrm{d}\mu),$$

then the above would be really the familiar correlation operator [33,35] $\mathbb{E}(r \otimes r)$ of the $\mathcal{U}$-valued *random variable* (RV) r, therefore from now on we shall simply refer to C as the correlation operator, even in the general case not based on a probability measure.

The fact that the correlation operator is self-adjoint and positive implies that its spectrum $\sigma(C) \subseteq \mathbb{R}^+$ is real and non-negative. This will be used when analysing it with any of the versions of the spectral theorem for self-adjoint operators (e.g. [16]). The easiest and best known version of this is for finite dimensional maps.

Finite dimensional beginnings

So let us return to the simple example at the beginning of the "Parametric models and linear maps" section where the associated linear map can be represented by a matrix $\boldsymbol{R}$. If we remember the each row $\boldsymbol{r}^{\mathsf{T}}(\mu_j)$ is the value for the vector $\boldsymbol{r}(\mu)$ for one particular $\mu \in \mathcal{M}$, we see that the matrix can be written as

$$\boldsymbol{R} = [\boldsymbol{r}(\mu_1), \ldots, \boldsymbol{r}(\mu_j), \ldots]^{\mathsf{T}},$$

and that the rows are just "snapshots" for different values μ_j. What is commonly done now is the so-called method of proper orthogonal decomposition (POD) to produce a ROM.

The matrix $\boldsymbol{R}$—to generalise a bit, assume it of size $m \times n$—can be decomposed according to its *singular value decomposition* (SVD)

$$\boldsymbol{R} = \boldsymbol{\Phi}\boldsymbol{\Sigma}V^{\mathsf{T}} = \sum_{k=1}^{\min(m,n)} \varsigma_k \, \phi_k \otimes \boldsymbol{v}_k, \tag{12}$$

where the matrices $\boldsymbol{\Phi} = [\phi_k]$ and $V = [\boldsymbol{v}_k]$ are orthogonal with unit length orthogonal columns—right and left singular vectors—ϕ_k resp. $\boldsymbol{v}_k$, and $\boldsymbol{\Sigma} = \mathrm{diag}(\varsigma_k)$ is diagonal with non-negative diagonal elements ς_k, the singular values. For clarity, we arrange the

singular values in a decreasing sequence, $\varsigma_1 \geq \varsigma_2 \geq \cdots \geq 0$. It is well known that this decomposition is connected with the eigenvalue or spectral decomposition of the correlation

$$C = R^{\mathsf{T}}R = V\Sigma\Phi^{\mathsf{T}}\Phi\Sigma V^{\mathsf{T}} = V\Sigma^2 V^{\mathsf{T}} = \sum_{k=1}^{\min(m,n)} \varsigma_k^2\, v_k \otimes v_k, \tag{13}$$

with eigenvalues ς_k^2, eigenvectors v_k, and its companion

$$C_{\mathcal{Q}} := RR^{\mathsf{T}} = \Phi\Sigma^2\Phi^{\mathsf{T}} = \sum_{k=1}^{\min(m,n)} \varsigma_k^2\, \phi_k \otimes \phi_k, \tag{14}$$

with the same eigenvalues, but eigenvectors ϕ_k. The representation is based on R^{T}, and its accompanying POD or *Karhunen–Loève* decomposition:

$$R^{\mathsf{T}} = \sum_{k=1}^{\min(m,n)} \varsigma_k\, v_k \otimes \phi_k, \qquad r(\mu_j) = \sum_{k=1}^{\min(m,n)} \varsigma_k\, v_k \otimes \phi_k(\mu_j), \tag{15}$$

where $\phi_k(\mu_j) = \phi_k^j$, and $\phi_k = [\phi_k^1, \ldots, \phi_k^j, \ldots]^{\mathsf{T}}$.

The second expression in Eq. (15) is a representation for $r(\mu)$, and that is the purpose of the whole exercise. Similar expressions may be used as approximations. It clearly exhibits the tensorial nature of the representation, which is also evident in the expressions Eqs. (12), (13), and (14). One sees here that this is just the j-th column of R^{T}, so that with the canonical basis in $\mathcal{Q} = \mathbb{R}^m$, $e_j^{(m)} = [\delta_{ij}]^{\mathsf{T}}$ with the Kronecker-δ, that expression becomes just

$$r(\mu_j) = R^{\mathsf{T}}e_j^{(m)}; \quad \text{and} \quad r \approx r_a = R^{\mathsf{T}}\psi \tag{16}$$

by taking other vectors ψ in $\mathcal{Q} = \mathbb{R}^m$ to give weighted averages or interpolations.

The general picture which emerges is that the matrix R is a kind of "square root"—or more precisely factorisation—of the correlation $C = R^{\mathsf{T}}R$, and that the left part of this factorisation is used for reconstruction resp. representation. In any other factorisation like

$$C = B^{\mathsf{T}}B, \quad \text{with} \quad B : \mathcal{U} \to \mathcal{H}, \tag{17}$$

where B maps into some other space $\mathcal{H}$; the map B will necessarily have essentially the same singular values ς_k and right singular vectors v_k as R, and can now be used to have a representation or reconstruction of r on $\mathcal{H}$ via

$$r \approx B^{\mathsf{T}}h \quad \text{for some} \quad h \in \mathcal{H}. \tag{18}$$

A popular choice is to use the Choleski-factorisation $C = LL^{\mathsf{T}}$ of the correlation into two triangular matrices, and then take $B^{\mathsf{T}} = L$ for the reconstruction.

As we have introduced the correlation's spectral factorisation in Eq. (13), some other factorisations come to mind, although they may be mostly of theoretical value:

$$C = B^{\mathsf{T}}B = (V\Sigma)(V\Sigma)^{\mathsf{T}} = (V\Sigma V^{\mathsf{T}})(V\Sigma V^{\mathsf{T}})^{\mathsf{T}}, \tag{19}$$

where then the reconstruction map is $B^\mathsf{T} = (V\Sigma)$ or $B^\mathsf{T} = (V\Sigma V^\mathsf{T})$. Obviously, in the second case the reconstruction map is symmetric $B^\mathsf{T} = B = C^{1/2}$, and is actually the true square root of the correlation C.

Other factorisation can come from looking at the companion $C_\mathcal{Q}$ in Eq. (14). Any factorisation $F : \mathcal{Z} \to \mathcal{Q}$ or approximate factorisation F_a of

$$C_\mathcal{Q} = FF^\mathsf{T} \approx F_a F_a^\mathsf{T} \tag{20}$$

is naturally a factorisation or approximate factorisation of the correlation

$$C = W^\mathsf{T}W \approx W_a^\mathsf{T}W_a, \quad \text{with} \quad W = F^\mathsf{T}\Phi V^\mathsf{T} \text{ and } W_a = F_a^\mathsf{T}\Phi V^\mathsf{T}, \tag{21}$$

where V and Φ are the left and right singular vectors—see Eq. (12)—of the associated map R resp. the eigenvectors of the correlation C in Eq. (13) and its companion $C_\mathcal{Q}$ in Eq. (14). A new ROM representation can now be found for $z \in \mathcal{Z}$ via

$$r \approx r_a = W_a^\mathsf{T}z = V\Phi^\mathsf{T}F_a z. \tag{22}$$

One last observation here is important: the expressions for r resp. one of its ROMs r_a are *linear* in the newly introduced parameters or "co-ordinates" ϕ_k in Eq. (15), resp. ψ in Eq. (16), resp. h in Eq. (18) and Eq. (25), as well as z in Eq. (22); which is an important requirement in many numerical methods.

Reduced order models—ROMs

As has become clear now, and was mentioned before, that approximations or ROMs $r_a(\mu)$ to the full model $r(\mu) \approx r_a(\mu)$ produce associated maps R_a, which are approximate factorisations of the correlation:

$$C \approx R_a^\mathsf{T}R_a.$$

This introduces different ways of judging how good an approximation is. If one looks at the difference between the full model $r(\mu)$ and ist approximation $r_a(\mu)$ as a residual, and computes weighted versions of it

$$\langle r(\cdot) - r_a(\cdot), u\rangle_\mathcal{U} = (R - R_a)u = Ru - R_a u, \tag{23}$$

then this is just the difference linear map $R - R_a$ applied to the weighting vector u. In Eq. (15) is was shown that $r(\cdot) = \sum_{k=1}^{\min(m,n)} \varsigma_k v_k \otimes \phi_k(\cdot)$ is a representation. As usual, one may now approximate such an expressions by leaving out terms with small or vanishing singular values, say using only $\varsigma_1, \ldots, \varsigma_\ell$, getting an approximation of rank ℓ—this also means that the associated linear map R_a in Eq. (15) has rank ℓ. As is well known [26], this is the best ℓ-term approximation in the norms of $\mathcal{U}$ and $\mathcal{Q}$. But from Eq. (23) one may gather that the error can also be described through the difference $R - R_a$. As error measure one may take the norm of that difference, and, depending on which norm one chooses, the error is then in—this example approximation —$\varsigma_{\ell+1}$ in the operator norm, or $\sum_{k=\ell+1}^{\min(m,n)} \varsigma_k$ in the trace- resp. nuclear norm, or $\sqrt{\sum_{k=\ell+1}^{\min(m,n)} \varsigma_k^2}$ in the Frobenius- resp. Hilbert–Schmidt norm.

On the other hand, different approximations or ROMs can now be obtained by starting with an approximate factorisation

$$C \approx B_a^\mathsf{T} B_a, \tag{24}$$

and introducing a ROM via

$$r \approx r_a = B_a^\mathsf{T} h. \tag{25}$$

Such a representing linear map B, may, e.g. via its SVD, be written as a sum of tensor products, and approximations B_a are often lower rank expressions, directly reflected in a reduced sum for the tensor products. As will become clearer at the end of this section, the bilinear forms Eq. (10) resp. Eq. (11) can sometimes split into multi-linear forms, thus enabling the further approximation of B_a through *hierarchical tensor products* [26].

Infinite dimensional continuation—discrete spectrum

For the cases where both $\mathcal{U}$ and $\mathcal{Q}$ are infinite dimensional, the operators R and C live on infinite dimensional spaces, and the spectral theory gets a bit more complicated. We shall distinguish some simple cases. After finite dimensional resp. finite rank operators just treated in matrix form, the next simplest case is certainly the case when the associated linear map R and the correlation operator $C = R^*R$ has a discrete spectrum, e.g. if C is compact, or a function of a compact operator, like for example its inverse. In this case the spectrum is discrete (e.g. [16]), and in the case of a compact operator the non-negative eigenvalues λ_k of C may be arranged as a decreasing sequence $\infty > \lambda_1 \geq \lambda_2 \geq \cdots \geq 0$ with only possible accumulation point the origin. It is not uncommon when dealing with random fields that C is a nuclear or trace-class operator, i.e. an operator which satisfies the stronger requirement $\sum_k \lambda_k < \infty$. The spectral theorem for an operator with purely discrete spectrum takes the form

$$C = R^*R = \sum_{k=1}^{\infty} \lambda_k \, (v_k \otimes v_k), \tag{26}$$

where the eigenvectors $\{v_k\}_k \subset \mathcal{U}$ form a CONS in $\mathcal{U}$. Defining a new corresponding CONS $\{s_k\}_k$ in $\mathcal{Q}$ via $\lambda_k^{1/2} s_k := R v_k$, one obtains the singular value decomposition of R and R^* with singular values $\varsigma_k = \lambda_k^{1/2}$:

$$R = \sum_{k=1}^{\infty} \varsigma_k (s_k \otimes v_k); \quad \text{i.e.} \quad R(u)(\cdot) = \sum_{k=1}^{\infty} \varsigma_k \langle v_k, u \rangle_{\mathcal{U}} s_k(\cdot), \quad R^* = \sum_{k=1}^{\infty} \varsigma_k (v_k \otimes s_k);$$

$$r(\mu) = \sum_{k=1}^{\infty} \varsigma_k \, s_k(\mu) v_k = \sum_{k=1}^{\infty} s_k(\mu) R^* s_k, \text{ as } R^* s_k = \varsigma_k v_k. \tag{27}$$

It is not necessary to repeat in this setting of compact maps all the different factorisations considered in the preceding paragraphs, and especially their approximations, which will be usually finite dimensional as they are made to be used for actual computations, e.g. the approximations will usually involve only finite portions of the infinite series in Eqs. (26) and (27), which means that the induced linear maps have finite rank and essentially become finite dimensional, so that the preceding paragraphs apply practically verbatim.

But one consideration is worth to follow up further. In infinite dimensional Hilbert spaces, self-adjoint operators may have a continuous spectrum, e.g. [16]; this is what is usually the case when homogeneous random fields or stationary stochastic processes have to be represented, This means that the expressions developed for purely discrete spectra in Eqs. (26) and (27) are not general enough. These expressions are really generalisations of the last equalities in Eqs. (13) and (12); but is is possible to give meaning to the matrix equalities in those equations, which simultaneously cover the case of a continuous spectrum.

In infinite dimensions—non-discrete spectrum

To this end we introduce the so called multiplication operator: Let $L_2(\mathcal{T})$ be the usual Hilbert space on some locally compact measure space $\mathcal{T}$, and let $\gamma \in L_\infty(\mathcal{T})$ be an essentially bounded function. Then the map

$$M_\gamma : L_2(\mathcal{T}) \to L_2(\mathcal{T}); \qquad M_\gamma : \xi(t) \to \gamma(t)\xi(t)$$

for $\xi \in L_2(\mathcal{T})$ is a bounded operator $M_\gamma \in \mathscr{L}(\mathcal{X})$ on $L_2(\mathcal{T})$. Such a multiplication operator is the direct analogue of a diagonal matrix in finite dimensions.

Using such a multiplication operator, one may introduce a formulation of the spectral decomposition different from Eq. (26) which does not require C to be compact [16], C resp. R do not even have to be continuous resp. bounded:

$$C = R^*R = VM_\gamma V^*, \tag{28}$$

where $V : L_2(\mathcal{T}) \to \mathcal{U}$ is unitary between some $L_2(\mathcal{T})$ on a measure space $\mathcal{T}$ and $\mathcal{U}$. In case C is continuous resp. bounded, one has $\gamma \in L_\infty(\mathcal{T})$. As C is positive, the function γ is non-negative ($\gamma(t) \geq 0$ a.e. for $t \in \mathcal{T}$). This covers the previous case of operators with purely discrete spectrum if the function γ is a step function and takes only a discrete (countable) set of values—the eigenvalues. This theorem is actually quite well known in the special case that C is the correlation operator of a stationary stochastic process—an integral operator where the kernel is the correlation function; in this case V is the *Fourier transform*, and γ is known as the *power spectrum*.

General factorisations

To investigate the analogues of further factorisations of R, $C = R^*R$, and its companion $C_\mathcal{Q} = RR^*$, we need the SVD of R and R^*. They derive generally in the same manner as for the finite dimensional case from the spectral factorisations of C in Eq. (28) and a corresponding one for its companion

$$C_\mathcal{Q} = RR^* = \Phi M_\gamma \Phi^* \tag{29}$$

with a unitary $\Phi : L_2(\mathcal{T}_*) \to \mathcal{Q}$ between some $L_2(\mathcal{T}_*)$ on a measure space $\mathcal{T}_*$ and $\mathcal{Q}$. Here in Eq. (29), and in Eq. (28), the multiplication operator M_γ plays the role of the diagonal matrix $\boldsymbol{\Sigma}^2$ in Eqs. (13) and (14). For the SVD of R one needs its square root, and as γ is non-negative, this is simply given by $M_\gamma^{1/2} = M_{\sqrt{\gamma}}$, i.e. multiplication by $\sqrt{\gamma}$. Hence the SVD of R and R^* is given by

$$R = \Phi M_{\sqrt{\gamma}} V^*, \quad R^* = V M_{\sqrt{\gamma}} \Phi^*. \tag{30}$$

These are all examples of a general factorisation $C = B^*B$, where $B : \mathcal{U} \to \mathcal{H}$ is a map to a Hilbert space $\mathcal{H}$ with all the properties demanded from R—see the beginning of this section. It can be shown [38] that any two such factorisations $B_1 : \mathcal{U} \to \mathcal{H}_1$ and $B_2 : \mathcal{U} \to \mathcal{H}_2$ with $C = B_1^*B_1 = B_2^*B_2$ are *unitarily equivalent* in that there is a unitary map $X_{21} : \mathcal{H}_1 \to \mathcal{H}_2$ such that $B_2 = X_{21}B_1$. Equivalently, each such factorisation is unitarily equivalent to R, i.e. there is a unitary $X : \mathcal{H} \to \mathcal{Q}$ such that $R = XB$.

Analogues of the factorisations considered in Eq. (19) are

$$C = B^*B = (VM_{\sqrt{\gamma}})(VM_{\sqrt{\gamma}})^* = (VM_{\sqrt{\gamma}}V^*)(VM_{\sqrt{\gamma}}^*V)^*, \tag{31}$$

where again $C^{1/2} = VM_{\sqrt{\gamma}}V^*$ is the square root of C.

And just as in the case of the factorisations of $C_{\mathcal{Q}}$ considered in Eq. (20) and the resulting factorisation of C in Eq. (21), it is also here possible to consider factorisations of $C_{\mathcal{Q}}$ in Eq. (29), such as

$$C_{\mathcal{Q}} = FF^* \approx F_aF_a^*, \quad \text{with} \quad F, F_a : \mathcal{E} \to \mathcal{U} \tag{32}$$

with some Hilbert space $\mathcal{E}$, which lead again to factorisations of

$$C = W^*W \approx W_a^*W_a, \quad \text{with} \quad W = F^*\Phi V^* \text{ and } W_a = F_a^*\Phi V^*, \tag{33}$$

and representation on the space $\mathcal{E}$; with the representing linear maps given by $W^* = V\Phi^*F$ resp. $W_a^* = V\Phi^*F_a$.

Coming back to the situation where C has a purely discrete spectrum and a CONS of eigenvectors $\{v_m\}_m$ in $\mathcal{U}$, the map B from the decomposition $C = B^*B$ can be used to define a CONS $\{h_m\}_m$ in $\mathcal{H}$: $h_m := BC^{-1/2}v_m$, which is an eigenvector CONS of the operator $C_{\mathcal{H}} := BB^* : \mathcal{H} \to \mathcal{H}$, with $C_{\mathcal{H}}h_m := \lambda_m h_m$, see [38]. From this follows a SVD of B and B^* in a manner analogous to Eq. (27). The main result is [38] that in the case of a nuclear C with necessarily purely discrete spectrum every factorisation leads to a separated representation in terms of a series, and vice versa. In case C is not nuclear, the representation of a "parametric object" via a linear map is actually more general [35,38] and allows to the rigorous and uniform treatment of also "idealised" objects, like for example Gaussian white noise on a Hilbert space.

In this instance of a discrete spectrum and a nuclear C and hence nuclear $C_{\mathcal{Q}}$, the abstract equation $C_{\mathcal{Q}} = \sum_k \lambda_k s_k \otimes s_k$ can be written in a more familiar form in the case when the inner product on $\mathcal{Q}$ is given by a measure ϖ on $\mathcal{M}$. It becomes for all $\varphi, \psi \in \mathcal{Q}$:

$$
\begin{aligned}
\langle C_{\mathcal{Q}}\varphi, \psi \rangle_{\mathcal{Q}} &= \sum_k \lambda_k \langle \varphi, s_k \rangle_{\mathcal{Q}} \langle s_k, \psi \rangle_{\mathcal{Q}} = \langle R^*\varphi, R^*\psi \rangle_{\mathcal{U}} \\
&= \iint_{\mathcal{M}\times\mathcal{M}} \varphi(\mu_1)\langle r(\mu_1), r(\mu_2)\rangle_{\mathcal{U}}\psi(\mu_2)\, \varpi(\mathrm{d}\mu_1)\varpi(\mathrm{d}\mu_2) \\
&= \iint_{\mathcal{M}\times\mathcal{M}} \varphi(\mu_1)\varkappa(\mu_1, \mu_2)\psi(\mu_2)\, \varpi(\mathrm{d}\mu_1)\varpi(\mathrm{d}\mu_2) \\
&= \iint_{\mathcal{M}\times\mathcal{M}} \varphi(\mu_1)\left(\sum_k \lambda_k s_k(\mu_1)s_k(\mu_2)\right)\psi(\mu_2)\, \varpi(\mathrm{d}\mu_1)\varpi(\mathrm{d}\mu_2).
\end{aligned}
$$

This shows that C_Q is really a Fredholm integral operator, and its spectral decomposition is nothing but the familiar theorem of Mercer [15] for the kernel

$$\varkappa(\mu_1, \mu_2) = \sum_k \lambda_k s_k(\mu_1) s_k(\mu_2). \tag{34}$$

Factorisations of C_Q are then usually expressed as factorisations of the kernel $\varkappa(\mu_1, \mu_2)$, which may involve integral transforms already envisioned in [30]—see also the English translation [31]:

$$\varkappa(\mu_1, \mu_2) = \int_{\mathcal{Y}} \rho(\mu_1, y) \rho(\mu_2, y) \, \mathsf{n}(\mathrm{d}y),$$

where the "factors" $\rho(\mu, y)$ are measurable functions on the measure space $(\mathcal{Y}, \mathsf{n})$. This is the classical analogue of the general "kernel theorem" [24].

Connections to tensor products

Although not as obvious as for the case of a discrete spectrum in Eqs. (12), (13), and (14); and Eqs. (26), (27), such a connection is also possible in the general case of a non-discrete spectrum. But as the spectral values in the continuous part have no corresponding eigenvectors, one has to use the concept of generalised eigenvectors [16,22,24,38]. Then it is possible to formulate the spectral theorem in the following way:

$$\langle Cu, w \rangle_{\mathcal{U}} = \int_{\mathbb{R}^+} \lambda \, \langle u, v_\lambda \rangle \langle w, v_\lambda \rangle \, \nu(\mathrm{d}\lambda), \quad \text{or in a weak sense} \tag{35}$$

$$C = \int_{\mathbb{R}^+} \lambda \, (v_\lambda \otimes v_\lambda) \, \nu(\mathrm{d}\lambda), \tag{36}$$

with the spectral measure ν on $\mathbb{R}^+$. Observe the analogy, especially of Eq. (36), with Eq. (26), where the sum now has been generalised to an integral to account for the continuous spectrum. Equation (35) is for the case of a simple spectrum; in the more general case of spectral multiplicity large than one, the Hilbert space $\mathcal{U} = \bigoplus_m \mathcal{U}_m$ can be written as an orthogonal sum [16,22,24] of Hilbert subspaces $\mathcal{U}_m$, each invariant under the operator C, on which an expression like Eq. (35) holds, and on which the spectrum is simple. For the sake of brevity we shall only consider the case of a simple spectrum now, and avoid writing the sums over m. The difficulty in going from Eqs. (26) to (36) is that the values λ in the truly continuous spectrum have no corresponding eigenvector, i.e. $v_\lambda \notin \mathcal{U}$, but it has to be found in a generally larger space. The possibility of writing an expression like Eq. (35) rests on the concept of a "rigged" resp. "equipped" Hilbert space or Gel'fand triple. This means that one can find [24] a nuclear space $\mathcal{K} \hookrightarrow \mathcal{U}$ densely embedded in the Hilbert space $\mathcal{U}$, such that Eq. (35) holds for all $u, v \in \mathcal{K}$. This also means that the generalised eigenvectors should be seen as linear functionals on $\mathcal{K}$. As the subspace $\mathcal{K}$ is densely embedded in $\mathcal{U}$, it also holds that $\mathcal{U} \hookrightarrow \mathcal{K}^*$ is densely embedded in the topological dual $\mathcal{K}^*$ of $\mathcal{K}$, i.e. one has the Gel'fand triple

$$\mathcal{K} \hookrightarrow \mathcal{U} \hookrightarrow \mathcal{K}^*. \tag{37}$$

The generalised eigenvectors can now be seen as elements of the dual, $v_\lambda \in \mathcal{K}^*$, where the generalised eigenvalue equation $Cv_\lambda = \lambda v_\lambda$ holds after an appropriate extension of C.

If an expressions such as Eqs. (35) or (36) have to be approximated numerically, it becomes necessary to evaluate the integral in an approximate way. The integral is really only over the spectrum of $\sigma(C)$ of C, as outside of $\sigma(C)$ the spectral measure ν vanishes. Obviously, one would first split the spectrum $\sigma(C) = \sigma_d(C) \cup \sigma_c(C)$ into a discrete $\sigma_d(C)$ and a continuous part $\sigma_c(C)$. On the discrete part, the integral is just a sum as shown before. On the continuous part, the integral has to be evaluated by a quadrature formula. Choosing quadrature points $\lambda_z \in \sigma_c(C)$ and appropriate integration weights $w_z \in \mathbb{R}$, the integral can be approximated by

$$\int_{\sigma_c(C)} \lambda \, \langle u, v_\lambda \rangle \langle w, v_\lambda \rangle \, \nu(\mathrm{d}\lambda) \approx \sum_z w_z \lambda_z \langle u, v_{\lambda_z} \rangle \langle w, v_{\lambda_z} \rangle,$$

an expression very similar to the ones used in case of discrete spectra.

Further tensor products

Essentially, the constructions we have been investigating could be seen as elements of the tensor product $\mathcal{U} \otimes \mathcal{Q}$, or extensions thereof as in the preceding paragraph. Often one, or both of the spaces $\mathcal{U}$ or $\mathcal{Q}$, can be further so factored in tensor products, say without loss of generality that $\mathcal{Q} = \mathcal{Q}_I \otimes \mathcal{Q}_{II}$. This is for example the case for the white-noise modelling of random fields [29,33,35,36], where one has $\mathcal{Q} = \bigotimes_{m=1}^{\infty} \mathcal{S}_m$. We just want to indicate how this structure can be used for further approximation.

It essentially means that the whole foregoing is applied, instead on $\mathcal{U} \otimes \mathcal{Q}$, on the tensor product $\mathcal{Q}_I \otimes \mathcal{Q}_{II}$. Combined with the upper level decomposition on $\mathcal{U} \otimes \mathcal{Q}$, one sees that this becomes one on $\mathcal{U} \otimes (\mathcal{Q}_I \otimes \mathcal{Q}_{II})$. The bilinear forms Eqs. (10) and (11) can thus be written as tri-linear forms, making a direct connection to tensor products and multi-linear forms [26]. Like in the just cited example of random fields [35,36], often this can be extended to higher order tensor products in a tree-like manner—by splitting $\mathcal{U}$, or $\mathcal{Q}_I$ resp. $\mathcal{Q}_{II}$. This leads to a hierarchical structure encoded in this binary tree, with the top product $\mathcal{U} \otimes \mathcal{Q}$ the root of the tree, and the individual factors as leaves of the tree. The higher the order of the tensor product, the better it is possible to exploit dependencies in low-rank formats [25,26]. This has been recently also pointed out in the tight connections between *deep neural networks* [14,32] and such tensor decompositions, which come in different *formats* or *representations* [26]. The indicated binary tree leads to what is known as a *hierarchical Tucker-* or HT-format; but obviously the multi-factor tensor product can be split also in a non-binary fashion, leading to more general tree-based tensor formats [19]. A completely flat tree structure with only root and leaves corresponds to the well known *canonic polyadic-* or CP-decomposition or format, the original *proper generalised decomposition* (PGD) falls into this category [2,11–13,17].

Structure preservation

The foregoing development for a parametric model $r : \mathcal{M} \to \mathcal{U}$ did not assume anything more than that $\mathcal{U}$ is a Hilbert space. In "Parametric models and linear maps" section it was already indicated on how to proceed if $\mathcal{U}$ is not a Hilbert space, but a more general topological vector space. The treatment so far preserves the linear structure of the space $\mathcal{U}$, and the approximations are using this linear structure as well. The tensor based representations using tensors of certain rank already have a more difficult geometric structure [18], indeed a manifold structure [8].

But here the concern is about the structure of the image set of the parametric object $r(\mu)$ and its preservation under the approximations or ROMs $r_a(\mu)$. In case the image set—here $\mathcal{U}$—is not a vector space, but say a differential manifold, things are bound to get more complicated; one possible route of attack seems to use the previous linear methods like the ones described here to map into the tangent spaces. One instance of this, which seems to be more accessible, is the case when the image set is a Lie group $\mathcal{G}$. Then everything can be done in the tangent space at the group identity, the Lie algebra $\mathfrak{g}$ of the Lie group $\mathcal{G}$. The Lie algebra is a linear space, and one may take $\mathcal{U} := \mathfrak{g}$. One then has to map further from $\mathfrak{g}$ to $\mathcal{G}$, but this can be achieved by the canonical exponential map $\exp : \mathfrak{g} \to \mathcal{G}$. A representation or ROM then would have the form

$$\mathcal{M} \xrightarrow{r,r_a} \mathcal{U} = \mathfrak{g} \xrightarrow{\exp} \mathcal{G}.$$

This has the added advantage that interpolations along straight lines in $\mathfrak{g}$, which like in any Euclidean or unitary space are also geodetics, is mapped into interpolations along geodetics on the Riemannian manifold structure on $\mathcal{G}$. We shall come back to a somewhat similar situation later in this section.

Vector fields

One of the probably simplest situations is when the image set has the structure of $\mathcal{V} = \mathcal{U} \otimes \mathcal{E}$, where $\mathcal{E}$ is a *finite-dimensional* inner-product (Hilbert) space [33]:

$$\mathsf{r} : \mathcal{M} \to \mathcal{V} = \mathcal{U} \otimes \mathcal{E}; \quad \mathsf{r}(\mu) = \sum_k r_k(\mu) \boldsymbol{r}_k, \tag{38}$$

and the r_k are maps $r_k(\mu)\mathcal{M} \to \mathcal{U}$ as before in the "Parametric models and linear maps" and "Correlation and representation" section, whereas the $\boldsymbol{r}_k$ are typically linearly independent vectors in $\mathcal{E}$. Often one wants to preserve the structure $\mathcal{V} = \mathcal{U} \otimes \mathcal{E}$; one can think of this in the following way: $\mathcal{U}$ is a space of scalar functions, on some domain in Euclidean space, and $\mathcal{E}$ are vectors from the associated vector space. Hence one could call this a vector field in some sense. The associated linear map is then defined a bit differently, namely as

$$R_\mathcal{E} : \mathcal{U} \to \mathcal{Q} \otimes \mathcal{E}; \quad R_\mathcal{E} : u \mapsto \sum_k (R_k(\mu)u) \boldsymbol{r}_k,$$

where the maps $R_k : \mathcal{U} \to \mathcal{Q}$ are defined as before in Eq. (10).

The "correlation" can now be given by a bilinear form; namely the densely defined map $C_\mathcal{E}$ in $\mathcal{V} = \mathcal{U} \otimes \mathcal{E}$ is defined on elementary tensors $\mathsf{u} = u \otimes \boldsymbol{u}, \mathsf{v} = v \otimes \boldsymbol{v} \in \mathcal{V} = \mathcal{U} \otimes \mathcal{E}$ as

$$\langle C_\mathcal{E} \mathsf{u}, \mathsf{v} \rangle_\sqsubseteq := \sum_{k,j} \langle R_k(u), R_j(v) \rangle_\mathcal{Q} (\boldsymbol{u}^\mathsf{T} \boldsymbol{r}_k)(\boldsymbol{r}_j^\mathsf{T} \boldsymbol{v}) \tag{39}$$

and extended by linearity, where each $R_k : \mathcal{U} \to \mathcal{Q}$ is the map associated to $r_k(\mu)$ as before for just a single map $r(\mu)$. It may be called the *"vector correlation"*. By construction it is self-adjoint and positive. The corresponding kernel is not scalar, but has values in $\mathcal{E} \otimes \mathcal{E}$:

$$\varkappa_\mathcal{E}(\mu_1, \mu_2) = \sum_{k,j} \langle r_k(\mu_1), r_j(\mu_2) \rangle_\mathcal{U} \, \boldsymbol{r}_k \otimes \boldsymbol{r}_j. \tag{40}$$

The eigenvalue problem on for an integral operator with such a kernel—representing the companion map—is on $\mathcal{W} = \mathcal{Q} \otimes \mathcal{E}$.

Coupled systems

A in some way similar situation is when the state space $\mathcal{U} = \mathcal{U}_1 \times \mathcal{U}_2$ comes from a combined or *coupled* system [37], and one wants to conserve this information or structure. The state is represented as $\boldsymbol{u} = (u_1, u_2)$, and the natural inner product on such a normal product space is

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle_{\mathcal{U}} = \langle (u_1, u_2), (v_1, v_2) \rangle_{\mathcal{U}} = \langle u_1, v_1 \rangle_{\mathcal{U}_1} + \langle u_2, v_2 \rangle_{\mathcal{U}_2}$$

for $\boldsymbol{u} = (u_1, u_2), \boldsymbol{v} = (v_1, v_2) \in \mathcal{U}$. This is for two coupled systems, labelled as '1' and '2'. The parametric map is

$$\boldsymbol{r} : \mathcal{M} \to \mathcal{U} = \mathcal{U}_1 \times \mathcal{U}_2; \quad \boldsymbol{r}(\mu) = (r_1(\mu), r_2(\mu)). \tag{41}$$

The associated linear map is

$$\boldsymbol{R}_c : \mathcal{U} \to \mathcal{Q}^2 = \mathcal{Q} \times \mathcal{Q}; \quad (\boldsymbol{R}_c(\boldsymbol{u}))(\mu) = (\langle u_1, r_1(\mu) \rangle_{\mathcal{U}_1}, \langle u_2, r_2(\mu) \rangle_{\mathcal{U}_2}). \tag{42}$$

As before, these $\mathbb{R}^2$ valued functions on $\mathcal{M}$ are like two problem-adapted co-ordinate systems on the joint parameter set, one for each sub-system. From this one obtains the "coupling correlation" $\boldsymbol{C}_c$, again defined through a bilinear form

$$\langle \boldsymbol{C}_c \boldsymbol{u}, \boldsymbol{v} \rangle_{\mathcal{U}} := \sum_{j=1}^{2} \langle R_j(u_j), R_j(v_j) \rangle_{\mathcal{Q}}. \tag{43}$$

The kernel is then a 2×2 matrix valued function in an integral operator on $\mathcal{W} = \mathcal{Q} \times \mathcal{Q}$:

$$\varkappa_c(\mu_1, \mu_2) = \mathrm{diag}(\langle r_k(\mu_1), r_k(\mu_2) \rangle_{\mathcal{U}_k}). \tag{44}$$

Other variations regarding coupled systems are possible, see [37], like when the parameter set $\mathcal{M} = \mathcal{M}_1 \times \mathcal{M}_2$ is a product. The parametric map can then defined as

$$\boldsymbol{r} : \mathcal{M} = \mathcal{M}_1 \times \mathcal{M}_2 \to \mathcal{U} = \mathcal{U}_1 \times \mathcal{U}_2; \quad \boldsymbol{r}((\mu_1, \mu_2)) = (r_1(\mu_1), r_2(\mu_2)), \tag{45}$$

with the associated linear map

$$\boldsymbol{R} : \mathcal{U} \to \mathcal{Q} = \mathcal{Q}_1 \times \mathcal{Q}_2; \quad (\boldsymbol{R}(\boldsymbol{u}))(\mu) = (\langle u_1, r_1(\mu_1) \rangle_{\mathcal{U}_1}, \langle u_2, r_2(\mu_2) \rangle_{\mathcal{U}_2}). \tag{46}$$

The correlation may be defined as before in Eq. (43), and also the kernel on $\mathcal{Q} = \mathcal{Q}_1 \times \mathcal{Q}_2$ is as in Eq. (44), but now the first diagonal entry is a function on $\mathcal{M}_1 \times \mathcal{M}_1$ only, and analogous for the second diagonal entry.

Tensor fields

This is similar to the case of vector fields in that the state space is $\mathcal{W} = \mathcal{U} \otimes \mathcal{A}$, where $\mathcal{U}$ is a space of scalar valued functions on some set; and $\mathcal{A} \subset \mathcal{B} = \mathcal{E} \otimes \mathcal{E}$, where $\mathcal{E}$ is a finite-dimensional vector space [38], and $\mathcal{A}$ is a manifold of tensors in the full tensor product $\mathcal{B}$ of tensors of even degree. If $\mathcal{A}$ were the full tensor product $\mathcal{B}$, which is a linear finite-dimensional space, there would be no difference to the case of vector fields. But in case of tensors of even degree are often used in more special situations. Obviously, such tensors may be identified with linear maps [38] $\mathcal{L}(\mathcal{E}) \cong \mathcal{B}$, which will be done here.

Therefore one may speak of e.g. the manifold of *special orthogonal* tensors, say $\mathsf{SO}(\mathcal{E})$, and of the manifold of *symmetric positive definite* tensors $\mathsf{Sym}^+(\mathcal{E})$.

We shall consider only these two mentioned examples. The special orthogonal tensors are a Lie group $\mathcal{A} := \mathsf{SO}(\mathcal{E})$ with Lie algebra $\mathfrak{a} := \mathfrak{so}(\mathcal{E})$, the *skew-symmetric* tensors, a free linear space. For $S \in \mathfrak{so}(\mathcal{E})$, the exponential map carries it onto $\exp(S) \in \mathsf{SO}(\mathcal{E})$. Therefore a parametric element in $\mathcal{W} = \mathcal{U} \otimes \mathcal{A}$ can first be represented as a parametric element in the linear space $\mathcal{Z} = \mathcal{U} \otimes \mathfrak{a}$, where all the preceding statements on vector fields apply. It is on this intermediary representation that one can define ROMs. Such a representation is then further mapped through exponentiation:

$$\exp_1 : \mathcal{Z} = \mathcal{U} \otimes \mathfrak{a} \ni u \otimes S \mapsto u \otimes \exp(S) \in \mathcal{U} \otimes \mathcal{A} = \mathcal{W}.$$

The associated linear map goes from $\mathcal{Z} = \mathcal{U} \otimes \mathfrak{a}$ to the linear space $\mathcal{Y} = \mathcal{Q} \otimes \mathfrak{a}$; and again from here one would use an analogue of the above exponential to map on $\mathcal{Q} \otimes \mathcal{A}$.

The positive definite tensors $\mathcal{A} := \mathsf{Sym}^+(\mathcal{E})$ are not a classical Lie algebra under multiplication, i.e. concatenation of linear maps, but rather only a Riemannian manifold, geometrically a convex cone. But there still is an exponential map, carrying the free linear space of symmetric tensors $\mathfrak{a} := \mathfrak{sym}(\mathcal{E})$ onto $\mathcal{A} := \mathsf{Sym}^+(\mathcal{E})$. In fact, for a $H \in \mathfrak{sym}(\mathcal{E})$, the exponential maps it onto $\exp(H) \in \mathsf{Sym}^+(\mathcal{E})$. Thus we have recovered formally the same situation as for orthogonal tensors just described, and the same procedures may be followed.

Conclusion

Parametric mappings $r : \mathcal{M} \to \mathcal{U}$ have been analysed with in a variety of settings via the associated linear map $R : \mathcal{U} \to \mathcal{Q} \subseteq \mathbb{R}^{\mathcal{M}}$. It was shown that the associated linear map contains the full information present in the parametric entity. It is actually a mathematically more general concept which allows one to define extreme or idealised such entities; this is particularly relevant in the field of uncertainty quantification when one has to deal with stochastic processes and random fields [35].

So instead of analysing a parametric entity $r(\mu)$ and ist approximations or ROMs $r_a(\mu)$ directly, one may take the cues on how to do this from considering the associated linear maps R and R_a. One has to say that in practical situations the associated linear maps are typically not simply available explicitly, but they provide a conceptual framework on how to deal with the situation. And even though they are not directly available, the desired quantities needed in such analyses are all in principle computable.

Very closely related to such an associated linear map R is the so-called "correlation operator" $C = R^*R$ and its companion $C_\mathcal{Q} = RR^*$, both self-adjoint and positive definite. Their spectral analysis turns out to be very helpful in understanding the nature of such parametric entities, as well as possible ROMs. The very general nature and mathematical embedding of parametric entities, which also incorporates random fields, is shown in the different spectral properties of the correlation operator. Such generalised parametric entities may yield correlation operators with continuous spectra—as it typically occurs for homogeneous random fields—and thus this needs the full generality of the spectral analysis in rigged Hilbert spaces for understanding the spectral analysis in terms of generalised tensor products. Other factorisations of the correlation, such as $C = B^*B$, induce other representations for the parametric entities, and any other representation or re-parametrisation may be understood in these terms.

Preservation of certain structural properties is often very desirable. Examples are given to show how the general idea can be refined to reflect some linear structures in the representation. This even extends to non-linear manifolds if they can be easily parametrised by linear spaces. Lie groups with their associated Lie algebras are one such example which is mentioned in a bit more detail. This last point is especially relevant to the representation of spatially varying or even random material properties, which are typically fields of symmetric positive tensors. A similar comment applies to "orientation fields", which are spatially varying and possibly random fields of orthogonal tensors.

Additionally it is explained how representations in tensor product spaces arise naturally in such situations, and how this process can be cascaded to produce a tree like structure for the analysis. Low-rank tensor approximations can thus be used as ROMs, and this certainly offers fresh new impulses. The same applies to machine learning and data-driven approaches, which obviously can also be analysed with the proposed framework. These deep learning methods have recently been shown to be closely connected with low-rank tensor approximations, offering some insights and avenues for their analysis. With the proposed framework of analysing such parametric entities via linear maps, we hope to introduce a fresh point of view which may lead to new ideas on how to construct and analyse ROMs.

Authors' contributions
The topic arose from joint discussion of HGM and RO about the topic of parametrised reduced order models. HGM did most of the actual writing, and RO discussed the results, proofread the manuscript, and confirmed its findings. All authors read and approved the final manuscript.

Availability of data and materials
This is a theoretical work, no data are used.

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Institute of Scientific Computing, Technische Universität Braunschweig, 38092 Brunswick, Germany, [2]Structural Mechanics and Coupled Systems Laboratory, Conservatoire National des Arts et Métiers (CNAM), 75141 Paris Cedex 03, France.

References
1. Ali ST, Antoine J-P, Gazeau J-P. Coherent states, wavelets, and their generalizations. 2nd ed. Berlin: Springer; 2014. https://doi.org/10.1007/978-1-4614-8535-3.
2. Ammar A, Chinesta F, Falcó A. On the convergence of a greedy rank-one update algorithm for a class of linear systems. Arch Comput Methods Eng. 2010;17:473–86. https://doi.org/10.1007/s11831-010-9048-z.
3. Antoine J-P, Bagarello F, Gazeau J-P, editors. Coherent states and their applications–a contemporary panorama, vol. 205., Springer proceedings in physicsBerlin: Springer; 2018. https://doi.org/10.1007/978-3-319-76732-1.
4. Benner P, Cohen A, Ohlberger M, Willcox K, editors. Model reduction and approximation: theory and algorithms, vol. 15. Philadelphia: SIAM; 2017.

5.   Benner P, Gugercin S, Willcox K. A survey of projection-based model reduction methods for parametric dynamical systems. SIAM Rev. 2015;57:483–531. https://doi.org/10.1137/130932715.
6.   Benner P, Ohlberger M, Patera AT, Rozza G, Urban K, editors. Model reduction of parametrized Systems, MS&A–modeling, simulation & applications, vol. 17. Berlin: Springer; 2017. https://doi.org/10.1007/978-3-319-58786-8.
7.   Berlinet A, Thomas-Agnan C. Reproducing kernel Hilbert spaces in probability and statistics. Berlin: Springer; 2004. https://doi.org/10.1007/978-1-4419-9096-9.
8.   Billaud-Friess M, Falcó A, Nouy A. Principal bundle structure of matrix manifolds. arXiv: 1705.04093 [math.DG]. 2017. http://arxiv.org/1705.04093.
9.   Buffa A, Maday Y, Patera AT, Prud'homme C, Turinic G. A priori convergence of the greedy algorithm for the parametrized reduced basis method. ESAIM: Math Model Numer Anal (M2AN). 2012;46:595–603. https://doi.org/10.1051/m2an/2011056.
10. Chen P, Schwab C. Model order reduction methods in computational uncertainty quantification. In: Ghanem R, Higdon D, Owhadi H, editors. Handbook of Uncertainty Quantification. Berlin: Springer; 2017. p. 937–90. https://doi.org/10.1007/978-3-319-12385-1.
11. Chinesta F, Huerta A, Rozza G, Willcox K. Model reduction methods, Encyclopaedia of computational mechanics. In: Stein E, de Borst R, Hughes TJR, editors. Part 1. Fundamentals. Encyclopaedia of computational mechanics, vol. 1. 2nd ed. Chichester: Wiley; 2017. https://doi.org/10.1002/9781119176817.ecm2110.
12. Chinesta F, Keunings R, Leygue A. The proper generalized decomposition for advanced numerical simulations. Berlin: Springer; 2014. https://doi.org/10.1007/978-3-319-02865-1.
13. Chinesta F, Ladevèze P, Cueto E. A short review on model order reduction based on proper generalized decomposition. Arch Comput Methods Eng. 2011;18:395–404. https://doi.org/10.1007/s11831-011-9064-7.
14. Cohen N, Sharri O, Shashua A. On the expressive power of deep learning: a tensor analysis. arXiv: 1509.05009 [cs.NE]. 2016. http://arxiv.org/abs/1509.05009.
15. Courant R, Hilbert D. Methods of mathematical physics. Chichester: Wiley; 1989. https://doi.org/10.1002/9783527617234.
16. Dautray R, Lions J-L. Spectral theory and applications, Mathematical analysis and numerical methods for science and technology, vol. 3. Berlin: Springer; 1990. https://doi.org/10.1007/978-3-642-61529-0.
17. Falcó A, Nouy A. Proper generalized decomposition for nonlinear convex problems in tensor Banach spaces. Numerische Mathematik. 2012;121:503–30. https://doi.org/10.1007/s00211-011-0437-5.
18. Falcó A, Hackbusch W, Nouy A. Geometric structures in tensor representations. arXiv: 1505.03027 [math.NA]. 2015. http://arxiv.org/1505.03027.
19. Falcó A, Hackbusch W, Nouy A. Tree-based tensor formats. arXiv: 1810.01262 [math.NA]. 2019. http://arxiv.org/1810.01262.
20. Fick L, Maday Y, Patera AT, Taddei T. A reduced basis technique for long-time unsteady turbulent flows. arXiv: 1710.03569 [math.NA]. 2017. http://arxiv.org/abs/1710.03569.
21. Gel'fand IM, Shilov GE. Properties and operations, Generalized functions, vol. 1. New York: Academic Press; 1964.
22. Gel'fand IM, Shilov GE. Theory of differential equations, Generalized functions, vol. 3. New York: Academic Press; 1967.
23. Gel'fand IM, Shilov GE. Spaces of fundamental and generalized functions, Generalized functions, vol. 2. New York: Academic Press; 1968.
24. Gel'fand IM, Vilenkin NY. Applications of harmonic analysis, Generalized Functions, vol. 4. New York: Academic Press; 1964.
25. Grasedyck L, Kressner D, Tobler C. A literature survey of low-rank tensor approximation techniques. GAMM-Mitteilungen. 2013;36:53–78. https://doi.org/10.1002/gamm.201310004.
26. Hackbusch W. Tensor spaces and numerical tensor calculus. Berlin: Springer; 2012. https://doi.org/10.1007/978-3-642-28027-6.
27. Hesthaven JS, Rozza G, Stamm B. Certified reduced basis methods for parametrized partial differential equations. Berlin: Springer; 2016. https://doi.org/10.1007/978-3-319-22470-1.
28. Hijazi S, Stabile G, Mola A, Rozza G. Data-driven POD-Galerkin reduced order model for turbulent flows. arXiv: 1907.09909 [math.NA]. 2019. http://arxiv.org/abs/1907.09909.
29. Janson S. Gaussian Hilbert spaces, Cambridge tracts in mathematics, vol. 129. Cambridge: Cambridge University Press; 1997. https://doi.org/10.1017/CBO9780511526169.
30. Karhunen K. Über lineare Methoden in der Wahrscheinlichkeitsrechnung. Ann Acad Sci Fennicae Ser A I Math Phys. 1947;37:1–79.
31. Karhunen K, Selin I. (transl.), On linear methods in probability theory—Über lineare Methoden in der Wahrscheinlichkeitsrechnung—1947, U.S. Air Force—Project RAND T-131, The RAND Corporation, St Monica, CA, USA, August 1960, Englisch Translation. https://www.rand.org/pubs/translations/T131.html.
32. Khrulkov V, Novikov A, Oseledets I. Eexpressive power of recurrent neural networks. arXiv: 1711.00811 [cs.LG]. 2018. http://arxiv.org/abs/1711.00811.
33. Krée P, Soize C. Mathematics of random phenomena—random vibrations of mechanical structures. Dordrecht: D. Reidel; 1986. https://doi.org/10.1007/978-94-009-4770-2.
34. Lam R, Zahm O, Marzouk Y, Willcox K. Multifidelity dimension reduction via active subspaces. arXiv: 1809.05567 [math.NA]. 2018. http://arxiv.org/abs/1809.05567.
35. Matthies HG. Analysis of probabilistic and parametric reduced order models. arXiv: 1807.02219 [math.NA]. 2018. http://arxiv.org/1807.02219.
36. Matthies HG, Litvinenko A, Pajonk O, Rosić BV, Zander E. Parametric and uncertainty computations with tensor product representations, uncertainty quantification in scientific computing. In: Dienstfrey A, Boisvert R, editors. IFIP advances in information and communication technology, vol. 377. Boulder: Springer; 2012. p. 139–50. https://doi.org/10.1007/978-3-642-32677-6.
37. Matthies HG, Ohayon R. Analysis of parametric models for coupled systems. arXiv: 1806.07255 [math.NA]. 2018. http://arxiv.org/1806.07255.

38. Matthies HG, Ohayon R. Analysis of parametric models—linear methods and approximations. arXiv: 1806.01101 [math.NA]. 2018. http://arxiv.org/1806.01101.
39. Quarteroni A, Manzoni A, Negri F. Reduced basis methods for partial differential equations: an introduction. Berlin: Springer; 2015. https://doi.org/10.1007/978-3-319-15431-2.
40. Quarteroni A, Rozza G, editors. Reduced order methods for modeling and computational reduction, MS&A–modeling, simulation & applications, vol. 9. Berlin: Springer; 2014. https://doi.org/10.1007/978-3-319-02090-7_8.
41. Raissi M, Perdikaris P, Karniadakis GE. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. J Comput Phys. 2019;378:686–707. https://doi.org/10.1016/j.jcp.2018.10.045.
42. Regazzoni F, Dedè L, Quarteroni A. Machine learning for fast and reliable solution of time-dependent differential equations. J Comput Phys. 2019;397:108852. https://doi.org/10.1016/j.jcp.2019.07.050.
43. Schwab C, Zech J. Deep learning in high dimension: neural network expression rates for generalized polynomial chaos expansions in UQ. Anal Appl. 2019;17:19–55. https://doi.org/10.1142/S0219530518500203.
44. Soize C, Ghanem R. Physics-constrained non-Gaussian probabilistic learning on manifolds. Int J Numer Methods Eng. 2019;. https://doi.org/10.1002/nme.6202.
45. Venturi L, Ballarin F, Rozza G. A weighted POD method for elliptic PDEs with random inputs. J Sci Comput. 2019;81:136–53. https://doi.org/10.1007/s10915-018-0830-7.
46. Zancanaro M, Ballarin F, Perotto S, Rozza G. Hierarchical model reduction techniques for flow modeling in a parametrized setting. arXiv: 1909.01668 [math.NA], 2019. http://arxiv.org/abs/1909.01668.

Publisher's Note