

Parametric and Uncertainty Computations with Tensor Product Representations

Hermann G. Matthies, Alexander Litvinenko, Oliver Pajonk, Bojana V. Rosić,
and Elmar Zander

Institute of Scientific Computing,
Technische Universität, Braunschweig, Germany
`wire@tu-bs.de`

Abstract Computational uncertainty quantification in a probabilistic setting is a special case of a parametric problem. Parameter dependent state vectors lead via association to a linear operator to analogues of covariance, its spectral decomposition, and the associated Karhunen-Loève expansion. From this one obtains a generalised tensor representation. The parameter in question may be a tuple of numbers, a function, a stochastic process, or a random tensor field. The tensor factorisation may be cascaded, leading to tensors of higher degree. When carried on a discretised level, such factorisations in the form of low-rank approximations lead to very sparse representations of the high dimensional quantities involved. Updating of uncertainty for new information is an important part of uncertainty quantification. Formulated in terms of random variables instead of measures, the Bayesian update is a projection and allows the use of the tensor factorisations also in this case.

Keywords: uncertainty quantification, parametric problems, low-rank tensor approximation, Bayesian updating

1 Introduction

Situations where one is concerned with uncertainty quantification often come in the following guise: we are investigating some physical system which is modelled by an evolution equation for its state:

$$\frac{\partial}{\partial t}u(t) = A(p; u(t)) + f(p; t), \quad (1)$$

where $u(t) \in \mathcal{V}$ describes the state of the system at time $t \in [0, T]$ lying in a Hilbert space \mathcal{V} (for the sake of simplicity), A is an operator modelling the physics of the system, and f is some external influence (action / excitation / loading). The model depends on some parameter $p \in \mathcal{P}$; in the context of uncertainty quantification the actual value of p is uncertain. Often this uncertainty is modelled by giving the set \mathcal{P} a probability measure. Evaluation and quantification of the uncertainty will often involve functionals of the state $\Psi(u(p; t))$, and the functional dependence of u on p becomes important. Similar situations arise in

design, where p may be a design parameter still to be chosen, and one may seek a design such that a functional $\Psi(u(p; t))$ is, for example, maximised.

The situation just sketched involves a number of objects which are functions of the parameter values. While evaluating $A(p)$ or $f(p)$ for a certain p may be straightforward, one may easily envisage situations where evaluating $u(p)$ or $\Psi(u(p))$ may be very costly as it may involve some very time consuming simulation or computation, like for example running a climate model.

As will be shown in the following Section 2, any parametric object such as $u(p)$, $A(p)$, or $f(p)$ may be seen as an element of a tensor product space. This can be used to find very sparse approximations to those objects, and in turn much cheaper ways to evaluate the objects for other parameter values. In particular this may be used in the uncertainty quantification to large advantage, like computing means, covariances, exceedance probabilities, etc. For this the dependence of $A(p)$ and $f(p)$ on p has to be propagated to the solution or state vector $u(p)$. This is called the *forward* problem, the resolution of which will be sketched in Section 3, e.g. see [14,15] and the references therein.

The situation we would like to address finally is actually a bit more complicated: In a situation as just described, we observe a function of the state $Y(u(p), p)$, and from this observation we would like to identify the corresponding p . This is called the *inverse* problem, and as the mapping $p \mapsto Y$ is usually not invertible, it is usually *ill-posed*. We embed this in a larger class by modelling our knowledge about p with the help of probability theory, and in a Bayesian manner our task becomes to estimate conditional expectations, e.g. see [21] and the references therein. The problem is now *well-posed*, but at the price of ‘only’ obtaining probability distributions on the possible values of p . The resolution of the inverse or identification problem will be addressed in Section 4.

2 Parametric problems

Let $r : \mathcal{P} \rightarrow \mathcal{V}$ be a parametric description of one of the objects alluded to in the introduction, where \mathcal{P} is some set, and \mathcal{V} for the sake of simplicity is assumed as a separable Hilbert space with inner product $\langle \cdot | \cdot \rangle_{\mathcal{U}}$ (the meaning of the index \mathcal{U} will soon become clear). What we desire is a simple representation / approximation of that function, which avoids solving (1) every time one wants to know $r(p)$ for a new $p \in \mathcal{P}$, i.e. a *response surface* or surrogate model, sometimes also called an *emulator*, whereas the solver for (1) is termed a *simulator*.

One relatively well-known way, particularly in statistical estimation [9], turns the problem into one of approximation of a linear mapping: let $\mathcal{U} = \overline{\text{span}} r(\mathcal{P}) = \overline{\text{span}} \text{im } r \subseteq \mathcal{V}$ be the smallest closed subspace of \mathcal{V} which is spanned by all the vectors $\{r(p) | p \in \mathcal{P}\}$. Then to each such function $r : \mathcal{P} \rightarrow \mathcal{U}$ one may associate a linear map

$$R : \mathcal{U} \ni u \mapsto \langle r(\cdot) | u \rangle_{\mathcal{U}} \in \mathbb{R}^{\mathcal{P}} \quad (2)$$

where $\mathbb{R}^{\mathcal{P}}$ denotes the space of real-valued functions on \mathcal{P} . By construction, R is injective. This may be used to define an inner product on $\text{im } R$ as

$$\forall \phi, \psi \in \text{im } R : \quad \langle \phi | \psi \rangle_{\mathcal{R}} := \langle R^{-1} \phi | R^{-1} \psi \rangle_{\mathcal{U}}, \quad (3)$$

and let \mathcal{R} be the completion of $\text{im } R$ with that inner product. It is obvious that R is a unitary map between the Hilbert spaces \mathcal{U} and \mathcal{R} .

Up to now, no structure on the set \mathcal{P} has been assumed, whereas on \mathcal{U} the inner product is assumed to measure what is important for the state $r(p) \in \mathcal{U}$. This is carried via the map R in (2) onto the space of scalar functions \mathcal{R} on the set \mathcal{P} , and the inner product there measures essentially the same thing as the one on \mathcal{U} .

2.1 Reproducing kernel Hilbert space

This is a first representation, and \mathcal{R} is called a reproducing kernel Hilbert space (RKHS) [7] with *reproducing kernel* $\varkappa \in \mathbb{R}^{\mathcal{P}} \times \mathbb{R}^{\mathcal{P}}$

$$\varkappa(p_1, p_2) := \langle r(p_1) | r(p_2) \rangle_{\mathcal{U}}. \quad (4)$$

It is straightforward to verify that it defines an obviously continuous (on \mathcal{R}) point-evaluation functional $\mathcal{R} \ni \phi \mapsto \phi(p) = \langle \varkappa(p, \cdot) | \phi \rangle_{\mathcal{R}} \in \mathbb{R}$, hence the name.

In other settings like classification or machine learning, e.g. with support vector machines, where $p \in \mathcal{P}$ has to be classified as belonging to certain subsets of \mathcal{P} , the space \mathcal{V} and the map $r : \mathcal{P} \rightarrow \mathcal{V}$ may often be freely chosen. This is then referred to as the “kernel trick”, and classification may be achieved by mapping these subsets with r into \mathcal{U} and separating them with hyperplanes—a linear classifier.

In terms of representation, one may now choose a basis $\{\varphi_m\}_{m \in \mathbb{N}}$ in \mathcal{R} , which may be assumed to be a complete orthonormal system (CONS). With the CONS $\{y_m \mid y_m = R^{-1}\varphi_m\}_{m \in \mathbb{N}}$ in \mathcal{U} , the operator R , its inverse R^{-1} , and the parametric element $r(p)$ become

$$R = \sum_m \varphi_m \otimes y_m; \quad R^{-1} = \sum_m y_m \otimes \varphi_m; \quad r(p) = \sum_m y_m \varphi_m(p), \quad (5)$$

exhibiting the tensorial nature of the representation mapping. With such a basis one may define a unitary map from ℓ_2 to \mathcal{R} and via R^{-1} further to \mathcal{U} :

$$\ell_2 \ni \mathbf{a} = (a_1, a_2, \dots) \mapsto \sum_m a_m \varphi_m \mapsto \sum_m a_m y_m \in \mathcal{U}. \quad (6)$$

Note that this representation is linear in the new ‘parameters’ $(a_1, a_2, \dots) \in \ell_2$. Model reductions may be achieved by choosing only subspaces of \mathcal{R} or ℓ_2 , or by approximating the map R^{-1} . This pattern of (5) or (6) repeats itself for all representations to follow.

2.2 Spectral decomposition

As a way of measuring of what is important on the set \mathcal{P} , assume that there is another inner product $\langle \cdot | \cdot \rangle_{\mathcal{W}}$ for scalar functions $\phi \in \mathbb{R}^{\mathcal{P}}$, and denote the Hilbert

space of functions with that inner product by \mathcal{W} . With this, one may define [9] a densely defined map C in \mathcal{U} through the bilinear form

$$\forall u, v \in \mathcal{U} : \quad \langle Cu|v \rangle_{\mathcal{U}} := \langle Ru|Rv \rangle_{\mathcal{W}}. \quad (7)$$

The map $C = R^*R$ (the adjoint is taken w.r.t. the \mathcal{W} -inner product, by abuse of notation we shall still call the map R) may be called the ‘correlation’ operator. By construction it is injective, positive, and self-adjoint.

Often the inner product $\langle \cdot | \cdot \rangle_{\mathcal{W}}$ comes from a measure ϖ on \mathcal{P} , so that \mathcal{W} may be taken as $L_2(\mathcal{P}, \varpi)$. One important class of problems is when ϖ is a probability measure on \mathcal{P} , i.e. $\varpi(\mathcal{P}) = 1$. Often the set has more structure, like being in a topological space, differentiable (Riemann) manifold, or a Lie group, which then may induce the choice of σ -algebra or measure. In all such cases one has $C = R^*R = \int_{\mathcal{P}} r(p) \otimes r(p) \varpi(dp)$. It is the factorisation of $C = R^*R$ which paves the way for further possibilities of representation. Most common is to use the spectral decomposition (e.g. [19,3]) of C :

$$Cu = \int_0^\infty \lambda \, dE_\lambda(u), \quad (8)$$

where E_λ is the corresponding projection valued spectral measure, with the spectrum $\sigma(C) \subseteq \mathbb{R}_+$. For the sake of simplicity assume that C has a pure point spectrum $\sigma_p(C) = \sigma(C)$ —the important case where C has also a continuous spectrum requires too many technical tools such as Gel’fand triplets (rigged Hilbert spaces) [3] and generalised eigenvectors to be treated in this short note—such that (8) may be written with the CONS of unit- \mathcal{U} -norm eigenvectors v_m :

$$Cu = \sum_m \lambda_m \langle v_m | u \rangle_{\mathcal{U}} v_m = \sum_m \lambda_m (v_m \otimes v_m) u. \quad (9)$$

From this follows the singular value decomposition of R , with $\lambda_m^{1/2} s_m := Rv_m$:

$$R = \sum_m \lambda_m^{1/2} (s_m \otimes v_m); \quad R^* = \sum_m \lambda_m^{1/2} (v_m \otimes s_m); \quad r(p) = \sum_m \lambda_m^{1/2} s_m(p) v_m, \quad (10)$$

where the last relation is the so-called Karhunen-Loève or proper orthogonal decomposition (POD). Observe that r —as well as R^* —is linear in the s_m . Similarly to (6), we have the—linear in \mathbf{a} —representation:

$$\ell_2 \ni \mathbf{a} = (a_1, a_2, \dots) \mapsto \sum_m a_m s_m \mapsto \sum_m \lambda_m^{1/2} a_m v_m \in \mathcal{U}. \quad (11)$$

An alternative formulation of the spectral decomposition (8) is [19] that C is unitarily equivalent to a multiplication operator:

$$C = VM_k V^* = (VM_k^{1/2})(VM_k^{1/2})^*, \quad (12)$$

where V is unitary between some $L_2(\mathcal{T})$ and \mathcal{U} , M_k is a multiplication operator on the measure space \mathcal{T} with a positive function $k(s) > 0$, and $M_k^{1/2} = M_{\sqrt{k}}$.

The essential range of k is the spectrum of C . This gives in the now familiar manner a representation on $L_2(\mathcal{T})$ through the choice of a CONS $\{\varsigma_m\}$. Setting $u_m := VM_{\sqrt{k}}\varsigma_m$, one obtains

$$(VM_k^{1/2}) = (VM_{\sqrt{k}}) = \sum_m u_m \otimes \varsigma_m \quad (13)$$

as tensorial representation.

2.3 Other factorisations of C

Other factorisations $C = B^*B$ —which are all unitarily equivalent—lead to analogous representations. Let $B : \mathcal{U} \rightarrow \mathcal{H}$ be an injective mapping into another Hilbert space \mathcal{H} . Pick a CONS $\{e_m\}$ in \mathcal{H} and set $f_m := B^*e_m$, then

$$B^* = \sum_m f_m \otimes e_m, \quad (14)$$

again a tensorial representation. All the representations considered so far are of this type. Similarly to (6), we have the—linear in \mathbf{a} —representation:

$$\ell_2 \ni \mathbf{a} = (a_1, a_2, \dots) \mapsto \sum_m a_m e_m \mapsto \sum_m a_m f_m \in \mathcal{U}. \quad (15)$$

For finite dimensional spaces, a favourite choice for such a decomposition of C is the Cholesky factorisation $C = LL^T$.

Another often used possibility to consider factorisations of the reproducing kernel, i.e. *kernel decompositions*. In an abstract way, one considers $\hat{C} := RR^*$ instead of $C = R^*R$, and $\hat{C} : \mathcal{W} \ni \phi \mapsto \langle \varkappa(p, \cdot), \phi \rangle_{\mathcal{W}} \in \mathcal{W}$. In case $\mathcal{W} = L_2(\mathcal{P}, \varpi)$ one has here the Fredholm equation $(\hat{C}\phi)(p_1) = \int_{\mathcal{P}} \varkappa(p_1, p_2)\phi(p_2) \varpi(dp)$. As this is similar to the decompositions considered for C it will be omitted for the sake of brevity.

2.4 Examples and interpretations

Some examples are now in order, so that one may see that the above description is in many cases an abstract statement of already very familiar constructions. For a general parameter space, the constructions provide a ‘response surface’, but in some cases this is known under a different name:

- If \mathcal{V} is a space of centred random variables (RVs), r is a random field or stochastic process indexed by \mathcal{P} , and the reproducing kernel is the covariance function.
- If the measure ϖ on \mathcal{P} is a probability measure ($\varpi(\mathcal{P}) = 1$), and r is a centred \mathcal{V} -valued RV, then C is the covariance operator.
- If $\mathcal{P} = \{1, 2, \dots, n\}$ and $\mathcal{R} = \mathbb{R}^n$, then \varkappa is the Gram matrix of the vectors v_1, \dots, v_n .

- If $\mathcal{P} = [0, T]$ and $r(t), t \in [0, T]$, is the response of a dynamical system with state space \mathcal{V} , the R^* leads to the POD.
- If the two preceding items are combined, this gives the method of snapshots for the POD.
- If $\mathcal{P} = \{\omega_s \mid \omega_s \in \Omega\}$ are samples from some probability space Ω , then one gets the POD method for samples.

For the sake of simplicity we had restricted ourselves in the spectral decomposition (8) in Subsection 2.2 to the case of a pure point spectrum (9). If in the first item $\mathcal{P} = \mathbb{R}^d$, and the covariance function / reproducing kernel satisfies $\kappa(\mathbf{r}_1, \mathbf{r}_2) = c(\mathbf{r}_1 - \mathbf{r}_2)$ for $\mathbf{r}_1, \mathbf{r}_2 \in \mathbb{R}^d$, one calls the covariance *translation invariant*, and the random process for $d = 1$ *stationary* or the the random field *homogeneous*. In that case the eigenvalue equation with the operator $\hat{C} = RR^*$ is a convolution equation $(\hat{C}\phi)(\mathbf{r}_1) = \int_{\mathbb{R}^d} c(\mathbf{r}_1 - \mathbf{r}_2)\phi(\mathbf{r}_2) d\mathbf{r}_2$, which is well-known to be diagonalised by the (real) Fourier transform. This is an example of the spectral decomposition in (12), the function k for the multiplication operator M_k is the Fourier transform of c , and the point spectrum is typically empty [14].

3 Approximation and Propagation

When it comes to computing, two kinds of approximations will usually be employed: one is that the parametric dependence of the entities in (1) needs to be simplified to make it computationally accessible—often this is also termed a representation; the other approximation derives from the fact that even for a fixed parameter $p \in \mathcal{P}$ the system modelled by (1) can not be computationally treated without further approximation, i.e. because it is often an equation in an infinite dimensional space, e..g. a partial differential equation.

3.1 Representation, Approximation, and Model Reduction

In Section 2 were a number of examples on how to construct representations of the type

$$S : \mathcal{S} \rightarrow \mathcal{U} \tag{16}$$

with a Hilbert space \mathcal{S} which is used for the representation, such that $SS^* = C$, or equivalently $C = B^*B$ as in Subsection 2.3. At the core of all constructions was the mapping R in (2) and (10), which led to the operator $C = R^*R$ on \mathcal{U} , see (7). This mapping ‘linearises’ the problem, as one may choose new parameters on which the representation depends linearly, as was pointed out repeatedly, e.g. (11). Most representations are connected with the spectral decomposition in Subsection 2.2 or equivalently with the spectral kernel decomposition. All the representations shown could be written in a tensor product format. The possibilities alluded to for \mathcal{S} were the RKHS $\mathcal{R} \subset \mathbb{R}^{\mathcal{P}}$, see Subsection 2.1 and (3), with the reproducing kernel (4) and representation R^{-1} in (5), or the Hilbert space $\mathcal{Q} \subset \mathbb{R}^{\mathcal{P}}$ connected with the correlation operator C in (7), or Hilbert spaces

induced by the spectral decomposition like $L_2(\sigma(C))$ implicitly appearing in (8), or $L_2(\mathcal{T})$ in (12).

Other factorisations of C such as in Subsection 2.3, or of the reproducing kernel \varkappa with integral transforms, lead to representations which are not necessarily connected to the spectral decomposition and may be more convenient in certain circumstances.

On the other hand, through the magnitude of spectral values the spectral decomposition gives guidance on the relative importance of different subspaces of \mathcal{S} , and approximations of the representation map which may be computationally more advantageous, like low-rank approximations.

All these representations may be carried onto ℓ_2 in the manner of (14) in Subsection 2.3. Model reductions may be achieved through choice of a subspace of \mathcal{S} , and / or by approximating the representation map as alluded to above. Thus the quantity $r(p)$ is in all cases approximated by a tensor expression $r \approx \sum_j r_j \otimes \tau_j$, and the number of terms in the sum is termed the *rank* of the tensor, and this kind of versatile sparse approximation is also called a low-rank approximation [6].

3.2 Discretisation and Propagation

For brevity we follow [14], where more references may be found, cf. also the recent monograph [11]. For the sake of simplicity, let us concentrate on the time-independent or stationary version of (1), namely $A(p; u) = f(p)$. Usually this is some partial differential equation and has to be discretised, approximated, or somehow projected onto some finite dimensional subspace $\mathcal{V}_N \subset \mathcal{V}$, with $\dim \mathcal{V}_N = N$. The entities of (1) which are projected or induced on the corresponding \mathbb{R}^N will be denoted by boldface, such that the stationary, projected equation reads as

$$\mathbf{A}(p; \mathbf{u}) = \mathbf{f}(p). \quad (17)$$

To propagate the parametric dependence, choose a finite dimensional subspace of the Hilbert spaces mentioned in Subsection 3.1, say $\mathcal{S}_M \subset \mathcal{S}$ for the solution $\mathbf{u}(p)$ in (17). Via Galerkin projection or collocation, or other such techniques, the still parametric model (17) is thereby formulated on the tensor product $\mathcal{V}_N \otimes \mathcal{S}_M$, denoted as

$$\mathbf{A}(\mathbf{u}) = \mathbf{f}. \quad (18)$$

The solution of (18) is often computationally challenging, as $\dim \mathcal{V}_N \otimes \mathcal{S}_M = N \times M$ may be very large. One possibility for such high-dimensional problems are the low-rank approximations alluded to at the end of Subsection 3.1, by representing the entities in (17) such as \mathbf{A} , \mathbf{u} , and \mathbf{f} in a low-rank format. Several numerical techniques [17,4,10,16] have been developed recently to obtain an approximation to the solution $\mathbf{u} \approx \sum_j \mathbf{u}_j \otimes \mathbf{z}_j$ to (18) in this format by only ever operating on the data-sparse low-rank representation, thus allowing for an efficient resolution of the high-dimensional problem.

Once this has been computed, any other functional such as $\Psi(u(p))$ mentioned in Section 1 may be computed with relative ease. In case there is a probability measure on \mathcal{P} as given in the examples in Subsection 3.1, for example to quantify some uncertainty in the parameters, the functionals usually take the form of expectations, such that $\Psi(u) = \mathbb{E}(\psi(\mathbf{u}))$ becomes a mean, a variance, an exceedance probability, or other such quantity needed in an *uncertainty quantification*.

4 Identification and Inverse Problems

In the setting of (1) let us pose the following problem: Some components—let us denote these by q —of the parameters $p \in \mathcal{P}$ are not only uncertain, but we would like to infer what they are by making observations y_k at times $0 < t_1 < \dots < t_k \dots \in [0, T]$. But we can not observe the entity q directly—like in Plato’s cave allegory we can only see a ‘shadow’ of it, formally given by

$$Y : \mathcal{Q} \times \mathcal{V} \ni (q, u(t_k)) \mapsto z_k = Y(q; u(t_k)) \in \mathcal{Y}; \quad (19)$$

at least this is our model of what we are measuring. Usually the observation will deviate from what we expect to observe even if we knew the right q as (1) is only a *model*—so there is some model error ϵ , and the measurement will be polluted by some measurement error ε . Hence we observe $y_k = z_k + \epsilon + \varepsilon$. From this one would like to know what q and $u(t_k)$ are.

4.1 Identification

The mapping in (19) is usually not invertible and hence the problem is called ill-posed. One way to address this is via regularisation, but here we follow a different track. Modelling our lack-of-knowledge about q and $u(t_k)$ in a Bayesian way [21] by replacing them with a \mathcal{Q} - resp. \mathcal{V} -valued random variable (RV), the problem becomes well-posed [20]. But of course one is looking now at the problem of finding a probability distribution that best fits the data; and one also obtains a probability distribution, not just *one* pair q and $u(t_k)$.

The mathematical setup then is as follows: we assume that Ω is a measure space with σ -algebra \mathfrak{A} and with a probability measure \mathbb{P} , and that $q : \Omega \rightarrow \mathcal{Q}$ and $u : \Omega \rightarrow \mathcal{U}$ are random variables. For simplicity, we shall also require \mathcal{Q} to be a Hilbert space where each vector is a possible realisation. This is in order to allow to measure the distance between different q ’s as the norm of their difference, and to allow the operations of linear algebra to be performed.

In case the q ’s are not without constraints, or not in a vector space, then they should be mapped to such quantities. For example, if q is a diffusion tensor field, then it has to be symmetric and positive definite. The symmetric tensors are of course a subspace, but the manifold of positive definite ones is not, nor is it closed. But they can be given the structure of a Lie group and a Riemannian manifold [1], and then distance is measured as a the length of a path along

a geodesic. But the associated Lie algebra—the tangent space at the neutral element of group—is in one to one correspondence with the geodesics; hence one plays everything back to a vector space. A simple case of this are positive scalars; through the logarithm they are transformed into a vector space without constraints. The computations to be described should be performed in such a vector space.

4.2 Bayesian Updating

Bayes's theorem is commonly accepted as a consistent way to incorporate new knowledge into a probabilistic description [21]. The textbook statement of the theorem is about conditional probabilities

$$\mathbb{P}(I_q|M_y) = \frac{\mathbb{P}(M_y|I_q)}{\mathbb{P}(M_y)}\mathbb{P}(I_q), \quad (20)$$

where I_q is some subset of possible q 's, and M_y is the information provided by the measurement. As this becomes problematic when the set M_y has vanishing probability measure, Kolmogorov already defined conditional probabilities via conditional expectation [2]. But most computational approaches compute via the measures [13,20] Given the conditional expectation $\mathbb{E}(\cdot|M_y)$, the conditional probability is easily recovered as $\mathbb{P}(I_q|M_y) = \mathbb{E}(\chi_{I_q}|M_y)$, where χ_{I_q} is the characteristic function of the subset I_q .

The easiest point of departure for conditional expectation in our setting is to define it not just for one piece of measurement M_y , but for sub- σ -algebras of \mathfrak{A} . The connection with an event M_y is then that we take $\sigma(Y)$, the σ -algebra generated by Y . Observe that if $\mathfrak{S} \subseteq \mathfrak{A}$ is a sub- σ -algebra, then $L_2(\Omega, \mathfrak{S}, \mathbb{P})$ is a closed subspace of $L_2(\Omega, \mathfrak{A}, \mathbb{P})$.

For RVs with finite variance (elements of $L_2(\Omega, \mathfrak{A}, \mathbb{P})$) the conditional expectation $\mathbb{E}(\cdot|\mathfrak{S})$ is defined as the orthogonal projection onto $L_2(\Omega, \mathfrak{S}, \mathbb{P})$. It can then be extended as a contraction onto all $L_p(\Omega, \mathfrak{A}, \mathbb{P})$ with $p \geq 1$ [2]. In other words the Bayesian update may now be simply shown to be

$$\mathbb{E}(q|\sigma(Y)) = P_{\mathcal{Q}_n}(q) = \operatorname{argmin}_{\tilde{q} \in \mathcal{Q}_n} \|\tilde{q} - q\|_{L_2}^2, \quad (21)$$

where $\mathcal{Q}_n := L_2(\Omega, \sigma(Y), \mathbb{P})$ represents the new information, and $P_{\mathcal{Q}_n}$ is the orthogonal projector onto \mathcal{Q}_n . Already in [8] it was noted that the conditional expectation is the best estimate not only for the *loss function* ‘distance squared’, but for a much larger range of loss functions under certain distributional constraints. But for the above loss function this is independent of what distribution q might have.

Requiring the derivative of the loss function in (21) to vanish—equivalently remembering from elementary geometry that the line to the closest point is perpendicular to the approximating subspace—one arrives at the Galerkin orthogonality conditions

$$\forall \tilde{q} \in \mathcal{Q}_n : \langle q - \mathbb{E}(q|\sigma(Y)), \tilde{q} \rangle_{L_2} = 0. \quad (22)$$

To continue, note that the Doob-Dynkin lemma [2] assures us that if a RV like $\mathbb{E}(q|\sigma(Y))$ is measurable w.r.t. $\sigma(Y)$, then $\mathbb{E}(q|\sigma(Y)) = \psi(Y)$ for some measurable $\psi \in L_0(\mathcal{Y}; \mathcal{P})$. Hence $L_2(\Omega, \sigma(Y), \mathbb{P}) = L_2(\Omega, \mathfrak{A}, \mathbb{P}) \cap \overline{\text{span}}\{\phi(y) \mid y \in \mathcal{Y}, \phi \in L_0(\mathcal{Y}; \mathcal{Q})\}$, where $L_0(\mathcal{Y}; \mathcal{Q})$ is the vector space of measurable maps from \mathcal{Y} to \mathcal{Q} . In particular one sees that $\mathbb{E}(q|\sigma(Y)) \in L_0(\mathcal{Y}; \mathcal{Q})$. In the light of (22) the task of computing $\psi(y) := \mathbb{E}(q|\sigma(Y))$ may be phrased as: find $\psi \in L_0(\mathcal{Y}; \mathcal{Q})$ such that

$$\forall \phi \in L_0(\mathcal{Y}; \mathcal{Q}), y \in \mathcal{Y} : \quad \langle q - \psi(z), \phi(y) \rangle_{L_2} = 0. \quad (23)$$

The value $q_a := \psi(y) = \mathbb{E}(q|\sigma(Y))$ is called the *analysis, assimilated, or posterior* value, incorporating the new information. In case one has some *prior* approximation, also called a *forecast* q_f , this results in an affine shift of the subspaces involved, and hence $q_a = q_f + \psi(y - z_k)$ with $z_k = Y(q_f; u_f(t_k))$, e.g. see [8].

We would like to emphasise that it is the vector space setting of \mathcal{Q} and \mathcal{Y} which has made this well-known formulation possible, and it will also allow for easy numerical computation. To work with measures as in (20) is cumbersome, as probability measures are on the intersection of the unit sphere and the positive cone in the space of signed finite measures. A bit easier would be to work with RVs which are in a metric space, the conditional expectation then minimises the metric distance squared; but the Hilbert space setting is certainly the simplest instance of this. As we work in a vector space, we make another approximation to simplify the computations by replacing $L_0(\mathcal{Y}; \mathcal{Q})$ above by $\mathcal{L}(\mathcal{Y}, \mathcal{Q})$, the space of linear continuous maps. The Galerkin orthogonality condition (23) is then translated to: find $K \in \mathcal{L}(\mathcal{Y}, \mathcal{Q})$ such that [8,12]

$$\forall H \in \mathcal{L}(\mathcal{Y}, \mathcal{Q}), y \in \mathcal{Y} : \quad \langle q - Kz, Hy \rangle_{L_2} = 0, \quad (24)$$

and we set $\mathbb{E}(q|\sigma(Y))_\ell := K$, a linear approximation to $\mathbb{E}(q|\sigma(Y))$. As the projection is now onto the smaller space $\mathcal{Q}_\ell := L_2(\Omega, \mathfrak{A}, \mathbb{P}) \cap \overline{\text{span}}\{Hy \mid y \in \mathcal{Y}, H \in \mathcal{L}(\mathcal{Y}, \mathcal{Q})\} \subset L_2(\Omega, \sigma(Y), \mathbb{P})$, we are not using all the information available but the computation is simpler. In the case of prior information this is extended as before to [8,12]

$$q_a = q_f + K(y - z_k), \text{ with } K = C_{q,z}(C_z + C_\varepsilon)^{-1}. \quad (25)$$

This includes the errors $\epsilon + \varepsilon$ with covariance operator C_ε , and it is not difficult to show [8,12] that the optimal K is given by the well-known Kálmán gain in (25), where $C_z := \mathbb{E}(Y(q, u) \otimes Y(q, u))$ and $C_{q,z} = \mathbb{E}(q \otimes Y(q, u))$. In case $C_z + C_\varepsilon$ is not invertible or close to singularity, its inverse in (25) should be replaced by the Moore-Penrose pseudo-inverse. This update is in some ways very similar to the ‘Bayes linear’ approach [5].

4.3 Computing the Bayesian Update and an Example

As before in Subsection 3.2 for an actual computation the forward model is discretised like in (17) and (18). The space of possible measurements is also

discretised by $\mathcal{Y}_I \subset \mathcal{Y}$, as is the space of entities to be identified $\mathcal{Q}_J \subset \mathcal{Q}$, giving a discrete forward model and measurement operator

$$\frac{\partial}{\partial t} \mathbf{u}(\omega; t) + \mathbf{A}(\mathbf{q}(\omega); \mathbf{u}(\omega; t)) = \mathbf{f}(\omega; t); \quad \mathbf{z}_k = \mathbf{Y}(\mathbf{q}(\omega); \mathbf{u}(\omega; t_k)). \quad (26)$$

The update (25) is in this way also discretised to

$$\mathbf{q}_a(\omega) = \mathbf{q}_f(\omega) + \mathbf{K}(\mathbf{y}(\omega) - \mathbf{z}_k(\omega)), \quad \text{with } \mathbf{K} = \mathbf{C}_{q,z}(\mathbf{C}_z + \mathbf{C}_\varepsilon)^{-1}. \quad (27)$$

Completely analogous to how the dependence on $p \in \mathcal{P}$ was treated in Section 2 and Section 3, we represent the dependence on $\omega \in \Omega$ through a subspace $\mathcal{S} \subset L_0(\Omega, \mathbb{R}) \subset \mathbb{R}^\Omega$, i.e. a *random variable*, and in the discrete form by a finite dimensional subspace $\mathcal{S}_M \subset \mathcal{S}$. A popular choice for \mathcal{S}_M is Wiener’s *polynomial chaos*, orthogonal multi-variate Hermite polynomials in standard Gaussian RVs [7,14]. Looking at tensor products of these finite dimensional spaces, with the results of Section 2 and Section 3 the model and measurement equation (26) change to an analogue of (18), and (27) becomes $\mathbf{q}_a = \mathbf{q}_f + \mathbf{K}(\mathbf{y} - \mathbf{z}_k)$, where $\mathbf{K} = \mathbf{K} \otimes \mathbf{I}$ with \mathbf{K} from (27). Hence the update equation is naturally in a tensorised form, allowing to apply it directly to low-rank approximations as introduced in Section 3.

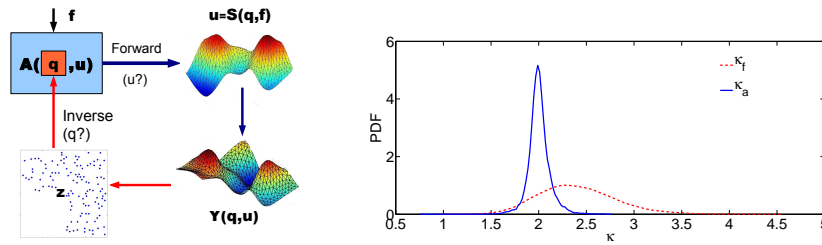


Figure 1. Updating schema and prior-posterior comparison

The following example of a diffusion equation on an L-shaped domain shows how the method works, for details and additional references please see [18]. A schematic view of the process is shown on the left picture in Fig. 1. The diffusion coefficient κ is the quantity to be identified; this is a positive quantity as alluded to before, so we consider its logarithm $q = \log(\kappa)$. The forward model in the top left produces a forecast on the system behaviour, shown on the top right. From this a forecast for the measurement is deduced, shown on the bottom right, and then compared with measurements—shown on the bottom left—to produce the Bayesian update. In the depiction on the right in Fig. 1, one may see how the prior or forecast distribution for $\kappa_f = \exp(q_f)$ is updated to the posterior or assimilated distribution $\kappa_a = \exp(q_a)$. For details please refer to [18].

References

1. Arsigny, V., Fillard, P., Pennec, X., Ayache, N.: Geometric means in a novel vector space structure on symmetric positive-definite matrices. *SIAM Journal on Matrix Analysis and Applications* 29(1), 328–347 (2006)
2. Bobrowski, A.: *Functional Analysis for Probability and Stochastic Processes*. Cambridge University Press, Cambridge (2005)
3. Dautray, R., Lions, J.L.: *Spectral Theory and Applications, Mathematical Analysis and Numerical Methods for Science and Technology*, vol. 3. Springer, Berlin (1990)
4. Doostan, A., Iaccarino, G.: A least-squares approximation of partial differential equations with high-dimensional random inputs. *Journal of Computational Physics* 228, 4332–4345 (2009)
5. Goldstein, M., Wooff, D.: *Bayes Linear Statistics—Theory and Methods*. Wiley Series in Probability and Statistics, John Wiley & Sons, Chichester (2007)
6. Hackbusch, W., Khoromskij, B.N.: Tensor-product approximation to operators and functions in high dimensions. *J. Complexity* 23(4-6), 697–714 (2007)
7. Janson, S.: *Gaussian Hilbert spaces*. Cambridge Tracts in Mathematics, 129, Cambridge University Press, Cambridge (1997)
8. Kálmán, R.E.: A new approach to linear filtering and prediction problems. *Transactions of the ASME—J. of Basic Engineering (Series D)* 82, 35–45 (1960)
9. Krée, P., Soize, C.: *Mathematics of Random Phenomena*. D. Reidel Publishing Co., Dordrecht (1986)
10. Krosche, M., Niekamp, R.: Low rank approximation in spectral stochastic finite element method with solution space adaption. *Informatikbericht 2010-03*, TU Braunschweig, Brunswick (2010), URL <http://www.digibib.tu-bs.de/?docid=00036351>
11. Le Maître, O.P., Knio, O.M.: *Spectral Methods for Uncertainty Quantification*. Scientific Computation, Springer, Berlin (2010)
12. Luenberger, D.G.: *Optimization by Vector Space Methods*. John Wiley & Sons, Chichester (1969)
13. Marzouk, Y.M., Najm, H.N., Rahn, L.A.: Stochastic spectral methods for efficient Bayesian solution of inverse problems. *Journal of Computational Physics* 224(2), 560–586 (2007)
14. Matthies, H.G.: Uncertainty quantification with stochastic finite elements. In: Stein, E., de Borst, R., Hughes, T.J.R. (eds.) *Encyclopaedia of Computational Mechanics*. John Wiley & Sons, Chichester (2007)
15. Matthies, H.G., Keese, A.: Galerkin methods for linear and nonlinear elliptic stochastic partial differential equations. *Computer Methods in Applied Mechanics and Engineering* 194(12-16), 1295–1331 (2005)
16. Matthies, H.G., Zander, E.: Sparse representations in stochastic dynamics. In: Papadrakakis, M., Stefanou, G., Papadopoulos, V. (eds.) *Computational Methods in Stochastic Dynamics*, pp. 247–265. Springer (2011)
17. Nouy, A., Le Maître, O.P.: Generalized spectral decomposition for stochastic nonlinear problems. *Journal of Computational Physics* 228(1), 202–235 (Jan 2009)
18. Rosić, B.V., Litvinenko, A., Pajonk, O., Matthies, H.G.: Direct Bayesian update of polynomial chaos representations. *Informatikbericht 2011-02*, TU Braunschweig, Brunswick (2011), URL <http://www.digibib.tu-bs.de/?docid=00039000>
19. Segal, I.E., Kunze, R.A.: *Integrals and Operators*. Springer-Verlag, Berlin (1978)
20. Stuart, A.M.: Inverse problems: A Bayesian perspective. *Acta Numerica* 19, 451–559 (2010)
21. Tarantola, A.: *Inverse Problem Theory and Methods for Model Parameter Estimation*. SIAM, Philadelphia (2004)