

# Karhunen-Loève Expansion for the Discrete Transient Heat Transfer Equation

Tarik Fahlaoui\*

*Sorbonne Universités, UTC, EA 2222, LMAC, F-60205 Compiègne, France.*

Received September 29, 2017; Accepted January 8, 2018

---

**Abstract.** In this paper we perform and analyze a Karhunen-Loève expansion on the solution of a discrete heat equation. Unlike the continuous case, several choices can be made from the numerical scheme to the numerical time integration. We analyze some of these choices and compare them. In literature, it is shown that the KL-expansion's error depends on the singular values of the matrix (or the operator) which we attempt to compress, but there is few results on the decay of these singular values. The core of this article is to prove the exponential decay of the singular values. To achieve this result, the analysis is conducted in a classical way by considering the spatial correlation of the temperature. And then, we analyze the problem in a more general view by using the Krylov matrices with Hermitian argument, which are a generalization of the well-known Vandermonde matrices. Some computations are made using MATLAB to ensure the performance of the Karhunen-Loève expansion. This article presents an application of this work to an identification problem where the data are disturbed by a Gaussian white noise.

**AMS subject classifications:** 35Q79, 65M60, 15A18, 15B05

**Key words:** Karhunen-Loève expansion, discrete temperature field.

---

## 1 Introduction

In many applications, we compute the same partial differential equation with different parameters. For instance, let us assume that we want to find the more adequate material to our target application. In order to do so, we run a simulation of the PDE in which the material is represented by one or several parameters, changing for each computation these parameters. However, computing the solution each time is costly. We could construct a basis with a much smaller dimension than with a regular method, by choosing solutions with special parameters. The reduction of the dimension is theoretically ensured by the exponential decay of the Kolmogorov  $N$ -width, for linear PDEs and some

---

\*Corresponding author. *Email address:* [tarik.fahlaoui@utc.fr](mailto:tarik.fahlaoui@utc.fr) (T. Fahlaoui)

non-linear PDEs (see [29]). This result allows us to find the solution to our PDE in a space of dimension  $N$  such that the error between the solution of this equation and the approximate solution computed with the reduced basis decreases fastly with  $N$  (see [6], [16] for more details). Two main methods are usually used : the *greedy method* which consist in constructing a hierarchical basis provided by an estimate error (see [15]), and the *Proper Orthogonal Decomposition (POD)* which compresses a set of snapshot (solutions to the PDE for fixed parameters) to withhold the main information (see [16], [17]). For parametric PDEs, we can show that the performance of the *greedy method* is related to the Kolmogorov  $N$ -width (see [16]).

In this article, we focus on Time-dependent equations, for which a POD basis can be constructed. The idea remains the same : provided some snapshots (which can be solutions at either different time steps or different space points), we perform a *Singular Value Decomposition (SVD)* on the correlation matrix. This compression method appears in many fields under various names such as *Principal Components Analysis (PCA)* in data analysis (see [14]), or *Karhunen-Loève expansion* in statistics to compress a stochastic process ([3]). The error estimation of the POD in the literature shows that the performance of this method is related to the decay of the eigenvalues of the correlation matrix ([25] or [26]). For experimental temperature fields, a POD was performed and analyzed in [27]. However, there is no proof on the decay rate of the eigenvalues. In [28], they explicit this decay by using the smoothness of the solution. Nevertheless, in [1] for bi-variate functions, and in [5] for the heat transfer equation, by using some algebraic properties satisfied by the spatial correlation matrix, they obtain an exponential decay.

For anisotropic parameters or in dimension greater than 1, the analytic solution of the PDE is not accessible. Based on the analysis made in the continuous case, we construct a Karhunen-Loève expansion for a discrete solution and show the efficiency of the method by analyzing the decay of the eigenvalues.

We first construct the Karhunen-Loève expansion for a discrete solution. We then focus in Section 2 on the transient heat transfer equation discretized in time by two well-known numerical schemes. By expanding the discrete temperature field in the basis form by the eigenvectors of the Laplace operator, we give an estimate on the error between the temperature field and the truncated KL-expansion. We adopt a new point a view in Section 3, by using the algebraic properties satisfied by the Krylov matrices with Hermitian arguments. Section 4 is dedicated to numerical results : some computations of the eigenvalues to support the theoretical analysis made in Section 2, and an application of this work to the identification of parameters when the data are disturbed by a Gaussian white noise.

### 1.1 The Karhunen-Loève expansion for discrete solution

Let us consider a discrete temperature field in time  $T^n$ , represented by a vector of size  $N_t$  such that each component belongs to  $L^2(\Omega)$ . For each step time, the only assumption on the function  $T^n(\mathbf{x})$  is to belong to  $L^2$  for the space variable. Yet, when we solve a PDE or

when we work with experimental data, the solution is discrete in space. However by using an interpolation function, we are able to define the solution as a  $L^2$  function. Our goal is to compress the data in order to keep only the essential information. In other words, we want to find a basis of small dimension which minimize the  $L^2$ -error between the discrete temperature field and its projection on this basis. The construction of this basis is well-known under the name of Karhunen-Loève expansion, and it has been handle in the continuous case for bi-variate functions ([1]). In discrete case, some issues arise and the goal of this paragraph is to construct a Karhunen-Loève expansion for discrete function and analyze the performance of this compression method. When we work with discrete time, the first thing is to handle integration in time. Several quadrature rules allow this. In order to remain general, we don't fix the quadrature weight and just set two numbers  $\zeta_0$  and  $\zeta_1$  such that

$$\int_0^b \phi(t) dt \approx \sum_{n=0}^{N_t-1} \zeta_0 \phi^n + \zeta_1 \phi^{n+1} = \zeta_0 \phi^0 + (\zeta_0 + \zeta_1) \sum_{n=1}^{N_t-1} \phi^n + \zeta_1 \phi^{N_t}.$$

Providing this quadrature rule, we can now define an inner product on  $\mathbb{R}^{N_t}$  which can be viewed as the analogous of the  $L^2$  inner product in the continuous case. It should be noted that the two weights  $\omega_0$  and  $\omega_1$  have to be non zero. An example of quadrature rule is given by the trapezoidal rule for which  $\zeta_0 = \zeta_1 = \frac{\Delta t}{2}$ . We have now all the tools to construct our Karhunen-Loève expansion. Before beginning the construction, let us say some words on the link between the Karhunen-Loève expansion presented here and the Singular Value Decomposition (SVD) applied on the matrix storing the degree of freedom, in the case where the temperature field is computed by the finite element method. Assuming a matrix  $\mathcal{T}$  whose raw components are the degree of freedom (or the value at the grid point) and the line components correspond to the time step. Thanks to the regularity of the solution, we could hope to obtain almost as much information by storing fewer data. A way to keep only the main information is to compute the right and left singular vectors and the singular value of  $\mathcal{T}$ . This allows us to obtain  $\mathcal{T} = V \Sigma \Phi^T$ , where  $V$  is the matrix of the left singular value (corresponding to space) and satisfies  $V^T V = \mathbf{I}_d$ ,  $\Phi$  is the matrix of the right singular value (corresponding to time) and satisfies  $\Phi^T \Phi = \mathbf{I}_d$ , and  $\Sigma$  is a diagonal matrix containing the singular value. An other formulation can be :  $T^j(x_i) = \sum_k \sigma_k v_k^i \phi_k^j$ . And since the temperature field is known to be regular, we can hope to have a fast decay of the singular value. An inconvenient of the SVD is that it give us singular vector that are orthonormal for the identity matrix while if the solution is obtain by the finite element method, we would expect the left singular vector to be orthonormal for the mass matrix. And equivalently the right singular vector to be orthonormal for the inner product introduced previously. For this reason we introduce the Karhunen-Loève expansion. We first construct the KL-expansion, by introducing some integral operator. And then we show how the performance of this expansion is related to the decay of the singular value.

As mentioned before, we equip the space  $\mathbb{R}^{N_t+1}$  with the inner product  $(\cdot, \cdot)_\omega$  defined

by

$$(\phi, \psi)_\omega = \zeta_0 \phi^0 \psi^0 + (\zeta_0 + \zeta_1) \sum_{n=1}^{N_t-1} \phi^n \psi^n + \zeta_1 \phi^{N_t} \psi^{N_t} \quad \forall \phi, \psi \in \mathbb{R}^{N_t+1},$$

in the following, we consider a vector  $\omega \in \mathbb{R}^{N_t+1}$  such that  $\omega_0 = \zeta_0$ ,  $\omega_n = \zeta_0 + \zeta_1$  for  $1 \leq n \leq N_t - 1$  and  $\omega_{N_t} = \zeta_1$ . Thus we can rewrite the inner product under the condensed form

$$(\phi, \psi)_\omega = \sum_{n=0}^{N_t} \omega_n \phi^n \psi^n.$$

We first define the operator  $B_h$  by

$$\left\{ \begin{array}{l} B_h: L^2(\Omega) \rightarrow \mathbb{R}^{N_t+1}, \\ v \mapsto \left( \int_{\Omega} T^n(\mathbf{x}) v(\mathbf{x}) d\mathbf{x} \right)_{0 \leq n \leq N_t} \end{array} \right. .$$

Roughly speaking, this is the matrix-vector product in continuous case, for a matrix which line components are the step time and the row components stand for the space dependence. As evoked previously, we can work with the singular values and the singular vectors of the operator  $B_h$ . However we prefer to work with an auto-adjoint operator in order to use the Spectral theorem. By definition, the adjoint operator  $B_h^*$  is

$$\left\{ \begin{array}{l} B_h^*: \mathbb{R}^{N_t+1} \rightarrow L^2(\Omega), \\ \psi \mapsto \sum_{n=0}^{N_t} \omega_n T^n(\mathbf{x}) \psi^n. \end{array} \right.$$

We then define the auto-adjoint  $A_h := B_h^* B_h$ ,

$$\left\{ \begin{array}{l} A_h: V_h \rightarrow V_h, \\ v \mapsto \int_{\Omega} \left( \sum_{n=0}^{N_t} \omega_n T^n(\mathbf{x}) T^n(\xi) \right) v(\xi) d\xi. \end{array} \right.$$

Write under this form, the kernel of  $A_h$  is  $K_h(\mathbf{x}, \xi) := \Delta t \sum_{n=0}^{N_t} \omega_n T^n(\mathbf{x}) T^n(\xi)$ , and thus  $A_h$  is an integral operator ([18]) and so a compact operator. By applying the Hilbert-Schmidt theorem,  $A_h$  admits eigenvectors  $v_k$  which form an Hilbert basis of  $L^2(\Omega)$  and are related to non-negative eigenvalues  $\lambda_k$ . Finally, Mercer's theorem ([12]) allows the following representation of  $K_h$

$$K_h(\mathbf{x}, \xi) = \sum_{k \geq 0} \lambda_k v_k(\mathbf{x}) v_k(\xi).$$

The final result of introducing this operator is that we can write  $T^n$  under the following form

$$T^n(\mathbf{x}) = \sum_{k \geq 0} \sigma_k \phi_k^n v_k(\mathbf{x}), \tag{1.1}$$

where  $\sigma_k$  stands for the singular value of the operator  $B_h$  (or equivalently the square root of the  $k$ -th eigenvalue of  $A_h$ ), and  $\phi_k^n$  is the projection of  $T^n$  on the eigenvector  $v_k$ . And the two basis  $(v_k)_k$  and  $(\phi_k^n)_k$  are respectively orthonormal for the  $L^2$ -product scalar and the discrete product scalar  $\langle \cdot, \cdot \rangle_\omega$ .

The aim of the KL-expansion is to compress the information. Here we deal with some snapshots of temperature at different time steps. The natural issue is to give an estimation of the truncation error. Let us define  $T_M^n$  as the truncated function of expression (1.1). We can then show the following classical result in literature.

**Proposition 1.1.**

$$\sum_{n=0}^{N_t} \omega_n \|T^n - T_M^n\|_{L^2(\Omega)}^2 = \left( \sum_{k>M} \sigma_k^2 \right)^{\frac{1}{2}}.$$

*Proof.* The proof is obvious by using the orthonormality of the eigenvectors  $(v_k)_k$  and  $(\phi_k)_k$ .  $\square$

Thus, the quality of the KL-expansion is related to the fast decay of the eigenvalues of the operator  $A_h$ . In other words, we have to study the spectrum of  $A_h$ . We can analyze directly the eigenvalue problem since we have

$$\int_{\Omega} K_h(\mathbf{x}, \boldsymbol{\xi}) v(\boldsymbol{\xi}) d\boldsymbol{\xi} = \lambda v(\mathbf{x}). \quad (1.2)$$

Unfortunately, we are not able to solve this system set in this form. The first way to rewrite the problem in a more suitable form is to expand each function ( $K_h$  and  $v$ ) in an orthonormal basis of  $L^2(\Omega)$ . Thus, the integral over  $\Omega$  disappears and we can work with an infinite matrix, whose eigenvectors are the coefficients of the eigenvectors of  $A_h$  in the orthonormal basis. The natural basis to expand the temperature field is the basis composed of the eigenvectors of the Spatial operator over which  $T^n$  is computed. In the next section, we consider the transient heat transfer equation and give an expression of the matrix on which we expect to show the fast decay of the eigenvalues.

## 2 The governing equation

Let us consider classical assumption such that  $\Omega$  is an open bounded domain in  $\mathbb{R}^d$  with smooth boundary  $\partial\Omega$ ,  $b > 0$  corresponding to the final time. The equation is defined by

$$\begin{cases} \partial_t T - \nabla \cdot (\gamma \nabla T) + \beta T = S, & \text{in } \Omega \times (0, b), \\ \gamma \nabla T \cdot \mathbf{n} = 0, & \text{on } \partial\Omega \times (0, b), \\ T(\cdot, 0) = a(\cdot) & \text{in } \Omega, \end{cases} \quad (2.1)$$

where the unknown  $T$  is the temperature field, which belongs to  $L^2(0, b; H^1(\Omega))$ . The diffusion term  $\gamma$  is a symmetric and positive tensor, and  $\beta$  is a scalar and represents

the heat transfer coefficient. Finally, we assume the right hand side  $S$  to be a separated function on time and space, i.e.

$$S(\mathbf{x}, t) = f(\mathbf{x})\theta(t).$$

However, in practice the solution to equation (2.1) is just known point-wise in time. We then begin to present an analysis in a discrete context. A first choice when we are interested in computing equation (2.1) is the numerical time scheme. In order to understand the importance of this choice, we consider two well-known schemes which are unconditionally stable.

### 2.1 Backward Euler scheme

The first scheme we work with is the Backward Euler scheme, which is

$$\begin{cases} \frac{T^{n+1}-T^n}{\Delta t} - \nabla \cdot (\gamma \nabla T^{n+1}) + \beta T^{n+1} = S^{n+1}, & \text{in } \Omega \quad \forall 0 \leq n \leq N-1, \\ \gamma \nabla T^{n+1} \cdot \mathbf{n} = 0 & \text{on } \partial\Omega, \quad \forall 0 \leq n \leq N-1, \\ T^0(\cdot) = a(\cdot) & \text{in } \Omega \end{cases} \quad (2.2)$$

with  $N\Delta t = b$ .

According to the Spectral Theorem ([13, Chapter 2], Theorem 5.1), we can rewrite  $T^n$  in the orthonormal basis formed by the eigenvectors  $(e_k)_{k \geq 1}$  of the Laplace operator  $\mathcal{K} := -\nabla \cdot \gamma \nabla + \beta$ . Thus, expanding the discrete temperature field in this basis gives us

$$T^n(\mathbf{x}) = \sum_{k \geq 1} T_k^n e_k(\mathbf{x}) \quad \text{for all } 0 \leq n \leq N-1. \quad (2.3)$$

We expand the spatial part of  $S$  and  $a$  in a similar way.

$$S^n = \theta^n \sum_{k \geq 1} f_k e_k(\mathbf{x}), \quad a = \sum_{k \geq 1} a_k e_k(\mathbf{x}). \quad (2.4)$$

An advantage to use the basis  $(e_k)_k$  is that it satisfies the steady heat transfer equation and thus we can give an explicit expression for  $T_k^n$  depending only on  $a_k$ ,  $\theta^n$ ,  $f_k$  and the eigenvalues of  $\mathcal{K}$  which we denote by  $r_k$ . Straightforward computations over the equation (2.2) yield the following expression for  $T^n$ ,

$$T^n = \left(\mathbf{I}_d + \Delta t \mathcal{K}\right)^{-n} a + \Delta t \sum_{i=1}^n \left(\mathbf{I}_d + \Delta t \mathcal{K}\right)^{i-n-1} S^i,$$

where  $\left(\mathbf{I}_d + \Delta t \mathcal{K}\right)^{-1}$  is the inverse of  $\left(\mathbf{I}_d + \Delta t \mathcal{K}\right)$ . Using the definition of an inverse and the eigenvalue problem satisfied by the couple  $(e_k, r_k)$ , we have

$$\sum_{k \geq 1} T_k^n e_k(\mathbf{x}) = \sum_{k \geq 1} a_k \left(\frac{1}{1 + \Delta t r_k}\right)^n e_k(\mathbf{x}) + \Delta t \sum_{i=1}^n \theta^i \sum_{k \geq 1} f_k \left(\frac{1}{1 + \Delta t r_k}\right)^{n-i+1} e_k(\mathbf{x})$$

and finally multiplying this equation by  $e_k$  and integrating over  $\Omega$ , we obtain the following expression of  $T_k^n$

$$T_k^n = a_k c_k^n + \Delta t \sum_{i=1}^n \theta^i f_k c_k^{n-i+1}, \tag{2.5}$$

where  $c_k = \frac{1}{1+\Delta t \lambda_k}$ .

### 2.2 Crank-Nicholson scheme

Let us now examine the expression of  $T_k^n$  when the equation (2.1) is discretized by a Crank-Nicholson scheme, given below

$$\begin{cases} \frac{T^{n+1}-T^n}{\Delta t} - \frac{1}{2} \left( \nabla \cdot (\gamma \nabla T^{n+1} + \gamma \nabla T^n) \right) + \frac{1}{2} \beta (T^{n+1} + T^n) = \frac{1}{2} (S^{n+1} + S^n), \\ \gamma \nabla T^{n+1} \cdot \mathbf{n} = 0 \\ T^0(\cdot) = a(\cdot) \end{cases} \begin{matrix} \text{in } \Omega, & \forall 0 \leq n \leq N-1, \\ \text{on } \partial\Omega, & \forall 0 \leq n \leq N-1, \\ \text{in } \Omega. \end{matrix} \tag{2.6}$$

We can write  $T^n$  as

$$T^n = \left( \mathbf{I}_d + \frac{\Delta t}{2} \mathcal{K} \right)^{-n} \left( \mathbf{I}_d - \frac{\Delta t}{2} \mathcal{K} \right)^n a + \frac{\Delta t}{2} \sum_{i=1}^n \left( \mathbf{I}_d + \frac{\Delta t}{2} \mathcal{K} \right)^{i-n-1} \left( \mathbf{I}_d - \frac{\Delta t}{2} \mathcal{K} \right)^{n-i} (S^i + S^{i-1}),$$

and using (2.3) and (2.4), it yields

$$T_k^n = a_k c_k^n + \frac{\Delta t}{2} \sum_{i=1}^n (\theta^i + \theta^{i-1}) f_k c_k^{n-i+1} \left( \frac{1}{1 - \frac{\Delta t}{2} \lambda_k} \right), \tag{2.7}$$

where  $c_k = \frac{1 - \frac{\Delta t}{2} \lambda_k}{1 + \frac{\Delta t}{2} \lambda_k}$ .

We first give the main result of this section, which is the performance of the Karhunen-Loève expansion.

**Theorem 2.1.** *Let  $T^n$  be a discrete temperature field computed by the Backward Euler scheme or the Crank-Nicholson scheme.  $T_M^n$  is the  $M$ -truncated temperature obtained by the Karhunen-Loève expansion, i.e.*

$$T_M^n(\mathbf{x}) = \sum_{k \leq M} \sigma_k \phi_k^n v_k(\mathbf{x}).$$

Then we have

$$\sum_{n=0}^{N_t} \omega_n \|T_M^n - T^n\|_{L^2(\Omega)}^2 \leq C_1 \epsilon + C_2 \left( \sum_{k \geq M} \exp\left(-\frac{\pi^2 k}{v \log\left(\frac{4b}{a}\right)}\right) \right)^{\frac{1}{2}},$$

where  $C_1$  just depends on  $N_t$  and  $\omega_n$ , and  $v$  is not greater than 18 and depends on the initial condition and the source term.  $C_2$  depends on  $v, r_k, N_t, \Delta t$ , the weights  $\xi_0$  and  $\xi_1$  and the initial condition and the source term.

**Remark 2.1.** This result shows the exponential performance of the Kahrunen-Loève expansion. The first term of the RHS is the price of acting with finite matrices. However, it does not reduce the result since we can find  $\epsilon$  such that this term is negligible with respect to the second term.

Before proving Theorem 2.1, we rewrite the eigenvalue problem (1.2) under a more suitable form thanks to the decomposition (2.5) and (2.7) of  $T^n$ . Recalling that  $A_h$  is define by

$$(A_h v)(\mathbf{x}) = \int_{\Omega} K_h(\mathbf{x}, \boldsymbol{\xi}) v(\boldsymbol{\xi}) d\boldsymbol{\xi},$$

we can expand  $K_h$  and  $v$  in  $(e_k)_k$  in order to eliminate the integral over  $\Omega$ . Considering  $\mathcal{G}$  the matrix which contains the coefficient of the kernel  $K_h$  in  $(e_k)_k$ , the operator  $A_h$  can be written

$$(A_h v)(\mathbf{x}) = \sum_{k,l} \mathcal{G}_{k,l} v_l e_k(\mathbf{x}).$$

The next step is to use this expression in the eigenvalue problem, multiplying by  $e_k$  and integrating over  $\Omega$  and we recover a matricial eigenvalue problem. Here we just define  $\mathcal{G}$  as the matrix containing the coefficients of the kernel, by use of (2.5) and (2.7), we can give an explicit form of this matrix. We recall that the kernel  $K_h$  is defined as

$$K_h(\mathbf{x}, \boldsymbol{\xi}) = \sum_{n=0}^{N_t} \omega_n T^n(\mathbf{x}) T^n(\boldsymbol{\xi}).$$

Thus, it is easy to see that the matrix  $\mathcal{G}$  is defined by

$$\mathcal{G}_{k,l} = \sum_{n=0}^{N_t} \omega_n T_k^n T_l^n. \tag{2.8}$$

In order to work with temperature field provided from a Backward Euler scheme or a Crank-Nicholson scheme, we set

$$\tilde{\theta}^i = \begin{cases} \Delta t \theta^i, & \text{for the Backward Euler scheme,} \\ \frac{\Delta t}{2} (\theta^i + \theta^{i-1}), & \text{for the Crank-Nicholson scheme} \end{cases}$$

and we set

$$\alpha_k = \begin{cases} 1, & \text{for the Backward Euler scheme,} \\ \left(1 - \frac{\Delta t}{2} r_k\right)^{-1}, & \text{for the Crank-Nicholson scheme.} \end{cases}$$

Thus for the two schemes,  $T_k^n$  has the same form

$$T_k^n = a_k c_k^n + \alpha_k \sum_{i=1}^n \tilde{\theta}^i f_k c_k^{n-i+1}. \tag{2.9}$$



The next step is to introduce the expression of  $T_k^n$  in the expression (2.10) of the matrix  $\mathcal{G}$  and remove the sum over  $n$ . Using the properties of geometric series and switching the sum over  $i$  and  $n$ , we have the following expression for  $\mathcal{G}$ .

$$\begin{aligned} \mathcal{G}_{k,l} = & \frac{1}{1-c_k c_l} \left[ \left( \zeta_0 + \zeta_1 c_k c_l + \zeta_1 c_k^{N_t} c_l^{N_t} - (2\zeta_1 + \zeta_0) c_k^{N_t+1} c_l^{N_t+1} \right) a_k a_l \right. \\ & + \left( \zeta_0 c_l \sum_{i=1}^{N_t} \tilde{\theta}^i c_k^i + \zeta_1 c_l^2 \sum_{i=1}^{N_t} \tilde{\theta}^i c_k^{i+1} - \zeta_0 c_k^{N_t} \sum_{i=1}^{N_t} \tilde{\theta}^i c_l^{N_t+1-i} - \zeta_1 c_k^{N_t+1} \sum_{i=1}^{N_t} \tilde{\theta}^i c_l^{N_t+2-i} \right) a_k f_l \alpha_l \\ & + \left( \zeta_0 c_k \sum_{i=1}^{N_t} \tilde{\theta}^i c_l^i + \zeta_1 c_k^2 \sum_{i=1}^{N_t} \tilde{\theta}^i c_l^{i+1} - \zeta_0 c_l^{N_t} \sum_{i=1}^{N_t} \tilde{\theta}^i c_k^{N_t+1-i} - \zeta_1 c_l^{N_t+1} \sum_{i=1}^{N_t} \tilde{\theta}^i c_k^{N_t+2-i} \right) a_l f_k \alpha_k \\ & + \sum_{i,j=1}^{N_t} \tilde{\theta}^i \tilde{\theta}^j \left( \zeta_0 c_k^{1-i+i \wedge j} c_l^{1-j+i \wedge j} + \zeta_1 c_k^{2-i+i \wedge j} c_l^{2-j+i \wedge j} - \zeta_0 c_k^{N_t+1-i} c_l^{N_t+1-j} \right. \\ & \left. - \zeta_1 c_k^{N_t+2-i} c_l^{N_t+2-j} \right) f_k f_l \alpha_k \alpha_l \left. \right]. \end{aligned} \tag{2.10}$$

This form has the advantage of separating the terms of subscripts  $k$  and  $l$ , except for the factor  $\frac{1}{1-c_k c_l}$  which can be useful as we show later. In the last line corresponding to the product of the RHS terms, the term indexed by  $k$  and  $l$  are not separated. To avoid this problem, we have by symmetry

$$\sum_{i,j=1}^{N_t} \tilde{\theta}^i \tilde{\theta}^j c_k^{1-i+i \wedge j} c_l^{1-j+i \wedge j} = \frac{1}{2} \left( \sum_{i,j=1}^{N_t} \tilde{\theta}^i \tilde{\theta}^j c_k c_l^{1+|i-j|} + \sum_{i,j=1}^{N_t} \tilde{\theta}^i \tilde{\theta}^j c_k^{1+|i-j|} c_l \right),$$

and we can finally write  $\mathcal{G}$  over a form that allows us to analyze it.

The proof of this theorem will be acted in four steps. We first prove that working with finite matrix does not change the result, then we state that the truncated matrix of  $\mathcal{G}$  satisfies a Lyapunov equation with a displacement of low rank. We give an upper bound on the Zolotarev number of the spectrum of the Lyapunov operator, and we finally conclude with an upper bound on the first eigenvalue of the truncated matrix of  $\mathcal{G}$ .

All the previous analysis was dealt with an infinite matrix  $\mathcal{G}$ . Unfortunately, there are only few results on eigenvalues for infinite matrices. However we can cope with this problem. By assumption, the temperature field  $T^n$  belongs to the space  $L^2(\Omega)$  and using the orthonormality of the basis  $(e_k)$  (the eigenvectors of the Laplace operator), we can conclude that

$$\forall \epsilon > 0, \quad \exists N(\epsilon) \in \mathbb{N}^* \text{ such that } \sum_{k > N(\epsilon)} |T_k^n|^2 < \epsilon, \quad \forall 0 \leq n \leq N_t. \tag{2.11}$$

If this result does not allow us to cut the matrix  $\mathcal{G}$ , it gives us an estimation of the error provided by the truncation of the matrix  $\mathcal{G}$ . Indeed, let the truncated sum  $T^{n,N(\epsilon)} = \sum_{k \leq N(\epsilon)} T_k^n e_k(\mathbf{x})$  and the truncated sum  $T_M^n$  introduced in proposition (1.1), then we have

$$\begin{aligned} & \sum_{n=0}^{N_t} \omega_n \|T^n - T_M^n\|_{L^2(\Omega)}^2 \\ & \leq \sum_{n=0}^{N_t} \omega_n \|T^n - T^{n,N(\epsilon)}\|_{L^2(\Omega)}^2 + \sum_{n=0}^{N_t} \omega_n \|T^{n,N(\epsilon)} - T_M^n\|_{L^2(\Omega)}^2. \end{aligned} \tag{2.12}$$

For the first term in RHS, we choose  $\epsilon$  in such a way that it is negligible and the second term will be handled in proposition (1.1) and the analysis of the cut matrix of  $\mathcal{G}$ . A more elegant way to show that working with finite matrices does not change the result, is to study the difference between the eigenvalues of the operator  $A_h$  and the eigenvalues of the cut matrix of  $\mathcal{G}$ . Let us introduce the operator  $A_h^\epsilon: L^2(\Omega) \mapsto L^2(\Omega)$  such that

$$\left(A_h^\epsilon v\right)(\mathbf{x}) = \int_{\Omega} K_h^\epsilon(\mathbf{x}, \boldsymbol{\xi}) v(\boldsymbol{\xi}) d\boldsymbol{\xi}, \tag{2.13}$$

where we define the kernel  $K_h^\epsilon$  as

$$K_h^\epsilon(\mathbf{x}, \boldsymbol{\xi}) = \sum_{n=0}^{N_t} \omega_n T^{n,\epsilon}(\mathbf{x}) T^{n,\epsilon}(\boldsymbol{\xi}) \tag{2.14}$$

and  $T^{n,\epsilon}(\mathbf{x}) = \sum_{k \leq N(\epsilon)} T_k^n e_k(\mathbf{x})$ . Under this notation, the operator  $A_h^\epsilon$  corresponds to the finite eigenvalue problem, i.e. the matrix resulting from the decomposition of  $A_h^\epsilon$  in the eigenvectors of the Laplace operator is nothing else than the truncated matrix of  $\mathcal{G}$ . The operator  $A_h^\epsilon$  is also a compact auto-adjoint operator. Moreover, since  $T^n(\mathbf{x})$  belongs to  $L^2(\Omega)$ , the operator  $A_h$  and  $A_h^\epsilon$  belongs to the trace class operator. In order to measure the difference between the eigenvalues of  $A_h$  and  $A_h^\epsilon$  under the assumption (2.11), we use the perturbation theory for compact operator. Thus we define the operator  $E = A_h - A_h^\epsilon$ . It is also an auto-adjoint operator and we have

$$Tr(E) = \left(\sum_{n=0}^{N_t} \omega_n\right) \sum_{k > N(\epsilon)} (T_k^n)^2 < C\epsilon.$$

We have then the following inequality ([24] and [23])

$$|\lambda_k(A_h) - \lambda_k(A_h^\epsilon)| \leq \|E\|,$$

and since  $\|E\| \leq Tr(E)$ , we have the following estimation

$$|\lambda_k(A_h) - \lambda_k(A_h^\epsilon)| < C\epsilon$$

with  $C = \sum_{n=0}^{N_t} \omega_n$ . This statement tells us that if we cut the matrix  $\mathcal{G}$  far enough for the assumption (2.11) to become true, then the error between the  $N(\epsilon)$  first eigenvalues of the operator  $A_h$  and the eigenvalues of the truncated matrix is of the order  $\epsilon$ .

**Remark 2.2.** The decomposition of the temperature field in the basis formed by the eigenvectors of the Laplace operator, is a theoretical process. Thus we can take  $N(\epsilon)$  as large as needed to recover the assumption (2.11). For the computation, we use an other basis to expand the temperature field. If  $T^n$  is computed by the finite element method, we choose the shape function. And thus we get the following eigenvalue problem :

$$M\mathcal{G}Mv = \lambda Mv,$$

where  $M$  denotes the mass matrix and  $\mathcal{G}$  the coefficient of  $K_h$  in the shape function. Using the Cholesky factorization for positive semi-definite matrix  $M = L_M D L_M^T$ , and setting  $w = D^{\frac{1}{2}} L_M^T v$ , the eigenvalue can be written

$$D^{\frac{1}{2}} L_M^T \mathcal{G} L_M D^{\frac{1}{2}} w = \lambda w,$$

which is a  $N_{dof} \times N_{dof}$  eigenvalue problem, where  $N_{dof}$  is the number of degree of freedom.

To make the notation less cluttered, we assume that the matrix  $\mathcal{G}$  is a finite matrix, i.e. we fix an integer  $K \in \mathbb{N}^*$  and assume that  $\mathcal{G} \in \mathbb{R}^{K \times K}$ . We have a first statement on  $\mathcal{G}$  :

**Proposition 2.1.** The matrix  $\mathcal{G}$  can be written as a Hadamard product between a matrix, diagonally congruous to a matrix  $\mathcal{P}$ , and a sum of tensor product of vector, i.e.

$$\mathcal{G} = \Delta \mathcal{P} \Delta \circ \sum_{i=1}^v \alpha_{i,1} \otimes \alpha_{i,2},$$

where the matrix  $\mathcal{P}$  is defined by

$$\mathcal{P}_{i,j} = \frac{d_i + d_j}{x_i + x_j}, \quad \forall 1 \leq i \leq K, x_i \in [a, b], \quad \text{with } 0 < a < b < \infty.$$

*Proof.* Setting  $d_i = \frac{1}{2}$ ,  $x_i = \frac{1-c_i}{1+c_i}$  and  $\Delta = \text{diag}(\frac{\sqrt{2}}{1+c_i})$  yields

$$(\Delta \mathcal{P} \Delta)_{k,l} = \frac{1}{1 - c_k c_l},$$

and since we have written  $\mathcal{G}$  in a way that all terms are separated, the result follows. For the Backward Euler scheme,  $c_k = \frac{1}{1 + \Delta t r_k}$  then  $x_k = \frac{\Delta t r_k}{2 + \Delta t r_k}$ . The eigenvalues  $(r_k)_k$  grow to infinity and are positive, thus we have :  $r_1 \leq \dots \leq r_K$ . And thus

$$a = \frac{\Delta t r_1}{2 + \Delta t r_1}, \quad b = \frac{\Delta t r_K}{2 + \Delta t r_K}.$$

For the Crank-Nicholson scheme, we have  $a = \frac{\Delta t}{2} r_1$  and  $b = \frac{\Delta t}{2} r_K$  □

**Remark 2.3.** Here  $\nu$  depends on the initial state and the RHS. If neither of them is identically zero, then  $\nu = 18$ . If the RHS is zero,  $\nu = 4$  and if  $a = 0$ ,  $\nu = 6$ . As we will see later,  $\nu$  corresponds to the rank of the matrix  $\mathcal{G}$ . Roughly speaking, the smoother the solution  $T^n$  is, the smaller  $\nu$  is. It is then natural that this number depends on the operator  $\mathcal{K}$ , the initial state, and the RHS.

This expression for  $\mathcal{G}$  will be useful in order to show the fast decay of its eigenvalues. Indeed, the matrix  $\mathcal{P}$  is a Pick matrix ([4]), which appears in interpolation problems, except that the coefficient  $x_i$  can be equal. Thus, we can write a Lyapunov equation on which  $\mathcal{P}$  is the solution. And we will use this equation to show the exponential decrease of  $\mathcal{G}$ 's eigenvalue. We first begin by showing that  $\mathcal{G}$  is solution to a Lyapunov equation :

**Proposition 2.2.** The matrix  $\mathcal{G}$  satisfies a Lyapunov equation with a displacement of rank  $\nu$ , i.e.

$$D_x \mathcal{G} - \mathcal{G}(-D_x) = MN^*, \quad M \in \mathbb{R}^{K \times \nu}, \quad N \in \mathbb{R}^{K \times \nu},$$

where  $D_x = \text{diag}(\frac{1-c_i}{1+c_i})$ .

*Proof.* Using the expression (2.1), we can rewrite  $\mathcal{G}$  by

$$\mathcal{G} = \sum_{i=1}^{\nu} D_{\alpha_{i,1}} \Delta \mathcal{P} \Delta D_{\alpha_{i,2}} \tag{2.15}$$

with the notation  $D_{\alpha_{i,1}} = \text{diag}(\alpha_{i,1})$ . The next step is to show that  $\mathcal{P}$  satisfies a Lyapunov equation by using the Kronecker product and the *vec*-operator. Let us set  $D_x = \text{diag}(x_i)$ , with the coefficient  $x_i$  defined above. We have

$$\left[ (\mathbf{I} \otimes D_x) + (D_x \otimes \mathbf{I}) \right] = \begin{pmatrix} D_x + x_1 \mathbf{I} & & & \\ & \ddots & & \\ & & \ddots & \\ & & & D_x + x_n \mathbf{I} \end{pmatrix}.$$

Multiplying by *vec*( $\mathcal{P}$ ), it yields

$$\begin{aligned} \left[ (\mathbf{I} \otimes D_x) + (D_x \otimes \mathbf{I}) \right] \text{vec}(\mathcal{P}) &= \sum_{j=1}^n (D_x + x_j \mathbf{I}) \mathcal{P}_{\cdot,j} \\ &= \sum_{i,j=1}^n (x_i + x_j) \frac{d_i + d_j}{x_i + x_j}, \end{aligned}$$

and finally  $\mathcal{P}$  satisfies the following Lyapunov equation ([19])

$$D_x \mathcal{P} + \mathcal{P} D_x = de^* + ed^* \text{ with } e^* = (1, \dots, 1).$$

We set  $d_i = \frac{1}{2}$  in the previous proposition, so that

$$D_x \mathcal{P} + \mathcal{P} D_x = ee^*.$$

By commutative properties of the multiplication of two diagonal matrices and summing over  $i$  we find the result. □

**Proposition 2.3.** The eigenvalues of the matrix  $\mathcal{G}$  have an exponential decay, i.e.

$$\lambda_{j+\nu k}(\mathcal{G}) \leq 4 \left[ \exp\left(\frac{\pi^2}{2\log(4\frac{b}{a})}\right) \right]^{-2k} \lambda_j(\mathcal{G}).$$

*Proof.* In Proposition 2.2, we have shown that  $\mathcal{G}$  satisfies the Lyapunov equation

$$D_x \mathcal{G} - \mathcal{G}(-D_x) = MN^*, \quad M \in \mathbb{R}^{K \times \nu}, \quad N \in \mathbb{R}^{K \times \nu},$$

where  $D_x = \text{diag}(x_i)$  and  $\forall 1 \leq i \leq K$ ,  $x_i \in [a, b]$  with  $0 < a < b < \infty$ . If we denote by  $\sigma(D_x)$  the spectrum of  $D_x$ , it is clear that  $\sigma(D_x) \subset [a, b]$ ,  $\sigma(-D_x) \subset [-b, -a]$ ,  $[a, b] \cap [-b, -a] = \emptyset$  and  $D_x$  is a symmetric matrix. By applying [2, Theorem 2.1.], we have

$$\lambda_{j+\nu k}(\mathcal{G}) \leq Z_k([-b, -a], [a, b]) \lambda_j(\mathcal{G}),$$

where  $Z_k$  is the Zolotarev number. And finally using the upper bound [2, Corollary 3.2.], we have

$$Z_k([-b, -a], [a, b]) \leq 4 \left[ \exp\left(\frac{\pi^2}{2\log(4\frac{b}{a})}\right) \right]^{-2k}, \quad 0 < a < b < \infty,$$

which concludes the proof. □

We can express the exponential decay of the eigenvalues of  $\mathcal{G}$  as

$$\lambda_{k+1}(\mathcal{G}) \leq 4 \left[ \exp\left(\frac{-\pi k^2}{\nu \log(4\frac{b}{a})}\right) \right] \lambda_1(\mathcal{G}),$$

which gives us the upper bound in Theorem 2.1, where we begin at  $M$  the sum over  $k$ .

**Remark 2.4.** The exponential decay is related to the value of  $\frac{b}{a}$  which is nothing else than  $\kappa(D_x)$ , the condition number of the matrix  $D_x$ . If  $\kappa(D_x)$  tends to infinity, we lose the exponential convergence. In the case where the RHS is equal to zero, the Lyapunov equation satisfied by  $\mathcal{G}$  becomes

$$D_x \mathcal{G} + \mathcal{G} D_x = BB^T$$

with  $B \in \mathbb{R}^{K \times 4}$ . Since  $D_x$  has eigenvalues contained in a strip around the real axis and is diagonal, we have the following decay for  $\mathcal{G}$  when  $\kappa(D_x)$  tends to infinity ([20])

$$\lambda_k(\mathcal{G}) \lesssim \exp\left(\frac{-\pi}{\sqrt{8}}\right)^{\sqrt{k}}. \quad (2.16)$$

When the RHS is non zero, we don't have a similar result, however the decay of the eigenvalues  $(\lambda_k)_k$  remains exponential for large  $K$  (for more details refer to [1, Remark A.1.]).

When we explicit the quotient  $\frac{b}{a}$  for each time scheme, we see that it is related to the quotient between the maximal and minimal eigenvalues of the Laplace operator. Indeed, we have  $\frac{b}{a} = \left(\frac{r_k}{r_1}\right) \left(\frac{2+\Delta tr_1}{2+\Delta tr_k}\right)$  for the Backward Euler scheme and  $\frac{b}{a} = \frac{r_k}{r_1}$  for the Crank-Nicholson scheme.

When we construct our matrix  $\mathcal{G}$ , we fix an integer  $K$  in order to work with finite matrix. And since the sequence  $(r_k)_k$  grows to infinity, the quotient  $\frac{b}{a}$  grows with  $K$ . Thus, the upper bound (2.16) ensures the exponential decay of the eigenvalue, even if the decay is slow. This issue will be supported by numerical experiment in Section 5.

We have then show that working with finite matrices yields an error which can be control to be negligible, and we show the exponential decay of the eigenvalues  $(\lambda_k(\mathcal{G}))_k$ . We finally give an upper bound of the first eigenvalue  $\lambda_1(\mathcal{G})$ .

**Proposition 2.4.** The first eigenvalue of  $\mathcal{G}$  can be bounded as

$$\lambda_1(\mathcal{G}) \leq C \sum_{i=1}^v \|\alpha_{i,1}\|_2 \|\alpha_{i,2}\|_2.$$

*Proof.* Since  $\lambda_1(\mathcal{G}) \leq \|\mathcal{G}\|_F$ , we have to give an upper bound of the Frobenius norm of  $\mathcal{G}$ . From (2.15), we have the following form for  $\mathcal{G}$

$$\mathcal{G} = \sum_{i=1}^v D_{\alpha_{i,1}} \Delta \mathcal{P} \Delta D_{\alpha_{i,2}}.$$

Thus, we have

$$\begin{aligned} \|\mathcal{G}\|_F^2 &= \sum_{k,l} \left| \sum_{i=1}^v (\alpha_{1,i})_k (\Delta \mathcal{P} \Delta)_{k,l} (\alpha_{2,i})_l \right|^2 \leq \sum_{k,l} \sum_{i=1}^v \left| (\alpha_{1,i})_k \frac{\sqrt{2}}{1+c_k} \frac{1}{x_k+x_l} \frac{\sqrt{2}}{1+c_l} (\alpha_{2,i})_l \right|^2 \\ &\leq \sum_{k,l} \sum_{i=1}^v \left| (\alpha_{1,i})_k^2 \frac{2}{(1+c_k)^2} \frac{1}{2x_k} \right| \left| (\alpha_{2,i})_l^2 \frac{2}{(1+c_l)^2} \frac{1}{2x_l} \right| \\ &= \sum_{i=1}^v \sum_{k,l} \left| (\alpha_{1,i})_k^2 \frac{1}{(1+c_k)(1-c_k)} \right| \left| (\alpha_{2,i})_l^2 \frac{1}{(1+c_l)(1-c_l)} \right| \\ &= \sum_{i=1}^v \left( \sum_k \left| (\alpha_{1,i})_k^2 \frac{1+2\Delta tr_k + \Delta t^2 r_k^2}{2\Delta tr_k + \Delta t^2 r_k^2} \right| \right) \left( \sum_l \left| (\alpha_{2,i})_l^2 \frac{1+2\Delta tr_l + \Delta t^2 r_l^2}{2\Delta tr_l + \Delta t^2 r_l^2} \right| \right) \\ &\leq \left( \frac{1+2\Delta tr_1 + \Delta t^2 r_1^2}{2\Delta tr_1 + \Delta t^2 r_1^2} \right)^2 \sum_{i=1}^v \left( \sum_k \left| (\alpha_{1,i})_k \right|^2 \right) \left( \sum_l \left| (\alpha_{2,i})_l \right|^2 \right). \end{aligned}$$

Using the fact that the eigenvalues of the Laplace operator grow to infinity. And finally, setting  $C = \frac{1+2\Delta t r_1 + \Delta t^2 r_1^2}{2\Delta t r_1 + \Delta t^2 r_1^2}$  gives us the expected result.  $\square$

The Theorem 2.1 shows the performance of the Karhunen-Loève expansion applied to the discrete heat transient equation. However the KL-expansion was constructed by using the operator  $A_h: L^2(\Omega) \mapsto L^2(\Omega)$ . Assuming that the data are provided by a finite element method, we have a matrix  $T$  which stores the degree of freedom at each time step. Computing the eigenvalue problem  $A_h v_k = \lambda_k v_k$  is equivalent to computing an  $N_{dof} \times N_{dof}$  eigenvalue problem. But if we define  $A_h$  as  $A_h = B_h B_h^*$  the eigenvalue problem's size becomes  $N_t + 1 \times N_t + 1$  which can be expected to be much smaller. This arises the following issue : does the KL-expansion remain good? And as stated in this section, it is equivalent to determine if the eigenvalues have an exponential decay. An other issue is the case of noisy data. Intuitively, there is no reason to expect an exponential decay since adding a white noise to the data amounts to add information. To tackle such questions, we will work directly with the singular value problem :  $B_h v_k = \sigma_k \phi_k$ .

### 3 A more general approach: Krylov matrices with Hermitian argument

As made in the previous section, we begin by constructing the matrix on which we made the singular value decomposition. We recall the definition of the operator from which we constructed the KL-expansion.

The operator  $B_h: L^2(\Omega) \rightarrow \mathbb{R}^{N_t+1}$  is defined by

$$(B_h v)_n = \int_{\Omega} T^n(\mathbf{x}) v(\mathbf{x}) d\mathbf{x}. \quad (3.1)$$

His adjoint operator  $B_h^*$  is defined from  $\mathbb{R}^{N_t+1} \rightarrow L^2(\Omega)$  by

$$(B_h^* \phi)(\mathbf{x}) = \sum_{0 \leq n \leq N_t} \omega_n T^n(\mathbf{x}) \phi_n. \quad (3.2)$$

We are interested in the singular value problem

$$B_h v_k = \sigma_k \phi_k \quad \text{and} \quad B_h^* \phi_k = \sigma_k v_k.$$

This singular value problem can be viewed as the continuous case of the singular value decomposition applied to the matrix  $T$  which stores the data. As we will later see, the singular vectors computed are orthonormal for a matrix which can be different to the identity.

To write the matrix under a suitable form, we will assume that there exists an orthonormal basis  $e_k$  of  $L^2(\Omega)$  such that the coefficients of the discrete temperature field in this basis have the following form

$$T^n(\mathbf{x}) = \sum_k \left( (\mathcal{K}_1)_{k,n} + (\mathcal{K}_2 \Theta)_{k,n} \right) e_k(\mathbf{x}). \quad (3.3)$$

Where  $\mathcal{K}_1$  and  $\mathcal{K}_2$  are Krylov matrices with Hermitian argument, i.e. of the form  $\mathcal{K}_{k,n} = (A^{n-1}\mathbf{x})_k$  where  $A$  is a square Hermitian matrix and  $\mathbf{x}$  a vector. We assume also that  $\Theta$  is invertible.

**Remark 3.1.** From (2.9), the coefficients of  $T^n$  in the basis formed by the eigenvectors of the Laplace operator are

$$T_k^n = a_k c_k^n + \alpha_k \sum_{i=1}^n \tilde{\theta}^i f_k c_k^{n-i+1}.$$

If we set  $(\mathcal{K}_1)_{k,n} = (D_c^{n-1}a)_k$ , where  $D_c = \text{diag}(c_k)$  and  $a = (a_k)$ ,  $(\mathcal{K}_2)_{k,n} = (D_c^{n-1}\alpha \circ f)_k$ , and finally

$$\Theta = \Delta t \begin{pmatrix} 0 & \dots & \dots & \dots & 0 \\ 0 & \tilde{\theta}^1 & \dots & \tilde{\theta}^{N-1} & \tilde{\theta}^N \\ \vdots & 0 & \ddots & \vdots & \vdots \\ 0 & \dots & 0 & \tilde{\theta}^1 & \tilde{\theta}^2 \\ 0 & \dots & \dots & 0 & \tilde{\theta}^1 \end{pmatrix},$$

we recover the general form. Note here that  $\Theta$  is not invertible, however noting that  $\mathcal{K}_1 + \mathcal{K}_2\Theta = \tilde{\mathcal{K}}_1 + \mathcal{K}_2\tilde{\Theta}$  where  $(\tilde{\mathcal{K}}_1)_{k,n} = (D_c^{n-1}(a - \gamma\alpha \circ f))_k$  and  $\tilde{\Theta} = \gamma I_d + \Theta$ , such that  $\gamma + \tilde{\theta}^1 \neq 0$ , we can recover the form (3.3).

We can now construct the singular value problem by putting the expression (3.3) of  $T_k^n$  in (3.1) and (3.2). It yields

$$\begin{aligned} (Bv)_n &= \int_{\Omega} T^n(\mathbf{x})v(\mathbf{x})d\mathbf{x} \\ &= \sum_k \left( (\mathcal{K}_1)_{k,n} + (\mathcal{K}_2\Theta)_{k,n} \right) v_k \end{aligned}$$

and similarly

$$\begin{aligned} (B^*\phi)(\mathbf{x}) &= \sum_{0 \leq n \leq N_t} \omega_n T^n(\mathbf{x})\phi_n \\ &= \sum_{0 \leq n \leq N_t} \sum_k \omega_n \left( (\mathcal{K}_1)_{k,n} + (\mathcal{K}_2\Theta)_{k,n} \right) e_k(\mathbf{x}). \end{aligned}$$

Thus we get

$$Bv = \sigma\phi \iff \sum_k \left( (\mathcal{K}_1)_{k,n} + (\mathcal{K}_2\Theta)_{k,n} \right) v_k = \sigma\phi^n$$



and multiplying by  $e_l$  and integrating over  $\Omega$  the expression of  $B^*$ , it yields

$$B^* \phi = \sigma v \iff \sum_{0 \leq n \leq N_t} \omega_n \left( (\mathcal{K}_1)_{l,n} + (\mathcal{K}_2 \Theta)_{l,n} \right) = \sigma v_l.$$

To make the notation less cluttered, we define the matrix  $\mathcal{T}$  by  $\mathcal{T}_{k,n} = \left( (\mathcal{K}_1)_{k,n} + (\mathcal{K}_2 \Theta)_{k,n} \right)$ . Note that it is nothing else than the coefficient of  $T^n$  is the basis  $e_k$ . At the end, the matrices are written

$$B_{n,k} = \mathcal{T}_k^n \quad \text{and} \quad B_{k,n}^* = \mathcal{T}_k^n \omega_n.$$

Let  $L = \text{diag}(\omega_n)$ , we have

$$\mathcal{T}^T v_k = \sigma_k \phi_k \quad \text{and} \quad \mathcal{T} L \phi_k = \sigma_k v_k.$$

Setting  $\psi_k = L^{\frac{1}{2}} \phi_k$ , it yields

$$\mathcal{T}^T v_k = \sigma_k L^{-\frac{1}{2}} \psi_k \quad \text{and} \quad \mathcal{T} L^{\frac{1}{2}} \psi_k = \sigma_k v_k$$

and finally

$$L^{\frac{1}{2}} \mathcal{T}^T v_k = \sigma_k \psi_k \quad \text{and} \quad \mathcal{T} L^{\frac{1}{2}} \psi_k = \sigma_k v_k.$$

Thus we have just to focus on the singular values of the matrix  $\mathcal{T} L^{\frac{1}{2}}$ .

**Remark 3.2.** If  $(e_k)_k$  are the eigenvalues of the Laplace operator, we get :

$$\mathcal{G} = \left( \mathcal{T} L^{\frac{1}{2}} \right) \left( \mathcal{T} L^{\frac{1}{2}} \right)^T.$$

And similarly, the eigenvalue problem  $BB^* \phi = \lambda \phi$  expanded in the basis  $(e_k)_k$  will be

$$\left( \mathcal{T} L^{\frac{1}{2}} \right)^T \left( \mathcal{T} L^{\frac{1}{2}} \right) \psi = \lambda \psi,$$

where we recall  $\psi = L^{\frac{1}{2}} \phi$ .

### 3.1 Exponential decay

In order to recover the exponential decay of the singular values  $(\sigma_k)$ , we take advantage of the Krylov matrices, which singular values have an exponential decay. An example of Krylov matrices are the well-known Vandermonde matrices. However, as in the first section, we have to deal with finite matrices, and thus we assume  $\mathcal{T}$  belonging to  $\mathbb{R}^{K \times N_t + 1}$ . We finally recover an error estimate as the one given in Theorem 2.1, by the upper bound (2.12).

We begin this section by giving the main result :

**Proposition 3.1.** Assuming the number of time subdivisions  $N_t$  is odd, then the singular values of  $\mathcal{T}L^{\frac{1}{2}}$  decrease as such

$$\sigma_{k+1}(\mathcal{T}L^{\frac{1}{2}}) \leq \max \left\{ \max_{0 \leq n \leq N_t} \omega_n^{\frac{1}{2}}; \max_{0 \leq n \leq N_t} \omega_n^{\frac{1}{2}} \theta^n \right\} \left[ \exp \left( \frac{-k\pi^2}{16 \log \left( \frac{8 \lfloor (N_t+1)/2 \rfloor}{\pi} \right)} \right) \right] \bullet$$

$$\left( \sigma_1(\mathcal{K}_1) + \sigma_1(\mathcal{K}_2) \right)$$

where  $\theta^n$  are the diagonal coefficients of the matrix  $\Theta$ .

**Remark 3.3.** The assumption  $N_t$  odd is not essential. If  $N_t$  is even, the formula remains unchanged except for the numerator in the exponential, which becomes  $(-k+1)\pi^2$ .

Before beginning the proof, we can remark that the constant in the log term only depends on  $N_t$  which have no reason to tend to infinity.

*Proof.* The proof will be made in three parts. We first state the exponential decay for the singular value of a Krylov matrix, then we show the exponential decay of  $\mathcal{K}_1 L^{\frac{1}{2}}$  and  $\mathcal{K}_2 \Theta L^{\frac{1}{2}}$  by the Courant-Fischer theorem, and finally we conclude by the additive Weyl's theorem.

Let  $\mathcal{K}_{k,n} = \left( A^{n-1} \mathbf{x} \right)_k$  a  $N \times M$  Krylov matrix with Hermitian argument (i.e.  $A$  is an Hermitian matrix).

It yields ([2, Corollary 5.2])

$$\sigma_{j+2k}(\mathcal{K}) \leq 4 \left[ \exp \left( \frac{\pi^2}{4 \log \left( \frac{8 \lfloor \frac{N+1}{2} \rfloor}{\pi} \right)} \right) \right]^{-k+[N]_2} \sigma_j(\mathcal{K}),$$

where  $[N]_2$  is zero if  $N$  is even and 1 otherwise. By the Courant-Fischer Theorem ([21, p. 555]), we know that

$$\sigma_k(\mathcal{K}_1 L^{\frac{1}{2}}) = \min_{V \in G_{N_t-k+2}} \max_{\substack{\mathbf{x} \in V \\ \|\mathbf{x}\|_2=1}} \left\| \mathcal{K}_1 L^{\frac{1}{2}} \mathbf{x} \right\|_2,$$

where  $G_k$  denotes the set of all vectorial subspace of dimension equal to  $k$ . For the first singular value we have then

$$\sigma_1(\mathcal{K}_1 L^{\frac{1}{2}}) = \max_{\substack{\mathbf{x} \in \mathbb{R}^{N_t+1} \\ \|\mathbf{x}\|_2=1}} \left\| \mathcal{K}_1 L^{\frac{1}{2}} \mathbf{x} \right\|_2,$$

and then by setting  $\mathbf{y} = L^{\frac{1}{2}} \mathbf{x} / \|L^{\frac{1}{2}} \mathbf{x}\|_2^{-1}$ , it yields

$$\sigma_1(\mathcal{K}_1 L^{\frac{1}{2}}) = \|\mathcal{K}_1 \mathbf{y}\|_2 \|L^{\frac{1}{2}} \mathbf{x}\|_2 \leq \|L^{\frac{1}{2}}\|_2 \max_{\substack{\mathbf{x} \in \mathbb{R}^{N_t+1} \\ \|\mathbf{x}\|_2=1}} \|\mathcal{K}_1 \mathbf{x}\|_2 = \|L^{\frac{1}{2}}\|_2 \sigma_1(\mathcal{K}_1),$$

since  $\|\mathbf{y}\|_2 = 1$ . For the second singular value, we have

$$\sigma_2(\mathcal{K}_1 L^{\frac{1}{2}}) = \min_{V \in G_{N_t}} \max_{\substack{\mathbf{x} \in V \\ \|\mathbf{x}\|_2=1}} \|\mathcal{K}_1 L^{\frac{1}{2}} \mathbf{x}\|_2 \leq \max_{\substack{\mathbf{x} \in V' \\ \|\mathbf{x}\|_2=1}} \|\mathcal{K}_1 L^{\frac{1}{2}} \mathbf{x}\|_2, \quad \forall V' \in G_{N_t}$$

Setting again  $\mathbf{y} = L^{\frac{1}{2}} \mathbf{x} / \|L^{\frac{1}{2}} \mathbf{x}\|_2^{-1}$ ,  $\mathbf{y}$  belongs to  $V'$  and  $\|\mathbf{y}\|_2 = 1$ , thus we have

$$\max_{\substack{\mathbf{x} \in V' \\ \|\mathbf{x}\|_2=1}} \|\mathcal{K}_1 L^{\frac{1}{2}} \mathbf{x}\|_2 \leq \|L^{\frac{1}{2}}\|_2 \|\mathcal{K}_1 \mathbf{y}\|_2 \leq \|L^{\frac{1}{2}}\|_2 \max_{\substack{\mathbf{y} \in V' \\ \|\mathbf{y}\|_2=1}} \|\mathcal{K}_1 \mathbf{y}\|_2.$$

And since it is true for all  $V' \in G_{N_t}$  passing to the min, it yields

$$\sigma_2(\mathcal{K}_1 L^{\frac{1}{2}}) \leq \|L^{\frac{1}{2}}\|_2 \sigma_2(\mathcal{K}_1).$$

Doing the same for each  $k$  gives

$$\sigma_k(\mathcal{K}_1 L^{\frac{1}{2}}) \leq \|L^{\frac{1}{2}}\|_2 \sigma_k(\mathcal{K}_1).$$

And similarly we have

$$\sigma_k(\mathcal{K}_2 \Theta L^{\frac{1}{2}}) \leq \|\Theta L^{\frac{1}{2}}\|_2 \sigma_k(\mathcal{K}_2).$$

Finally, by the additive Weyl's inequality ([11]), we have

$$\sigma_{i+j-1}(\mathcal{T} L^{\frac{1}{2}}) \leq \sigma_i(\mathcal{K}_1 L^{\frac{1}{2}}) + \sigma_j(\mathcal{K}_2 \Theta L^{\frac{1}{2}}).$$

which concludes the proof. □

The Proposition 3.1 shows the performance of the KL-expansion in the case where we solved the eigenvalues  $BB^* \phi = \lambda \phi$  or  $B^* B v = \lambda v$ . When we have our eigenvectors  $(\phi_k)_k$  in the first case or  $(v_k)_k$  in the second case the next step is to project the temperature fields into the basis form by these eigenvectors, i.e. in the first case  $(v_k)_k$  will be defined by  $v_k(\mathbf{x}) = \langle T^n(\mathbf{x}), \phi_k^n \rangle_\omega$  and in the second case  $(\phi_k)_k$  will be defined by

$$\phi_k^n = \int_{\Omega} T^n(\mathbf{x}) v_k(\mathbf{x}) d\mathbf{x}.$$

An advantage to work with the singular value problem is to avoid this projection step, indeed performing a singular value decomposition on the matrix  $\mathcal{T} L^{\frac{1}{2}}$  gives us  $\mathcal{T} L^{\frac{1}{2}} \psi = \sigma v$ , and we have just to compute  $\phi = L^{-\frac{1}{2}} \psi$ , where we recall  $L^{\frac{1}{2}} = \text{diag}(w_n^{\frac{1}{2}})$ .

When the data are provided by an experiment, there are disturbed. We frequently add a Gaussian white noise to the solution to numerically reproduce this phenomena. Thus we analyze the decay of the singular value when the temperature field are disturbed.

### 3.2 Disturbed data

Let consider  $T^n$  as in the previous section. We consider a Gaussian white noise  $\epsilon^n(\cdot)$  with variance denoted by  $\sigma^2$ . The disturbed temperature field  $\tilde{T}^n$  is defined by  $\tilde{T}^n = T^n + \epsilon^n$ . By assumption (3.3), we know that the coefficients of  $T^n$  in an orthonormal basis  $(e_k)_k$  of the space  $L^2(\Omega)$  are

$$T_k^n = \left( (\mathcal{K}_1)_{k,n} + (\mathcal{K}_2\Theta)_{k,n} \right).$$

Similarly, we can expand  $\epsilon^n(\cdot)$  in the basis  $(e_k)_k$ , i.e.

$$\epsilon^n(\mathbf{x}) = \sum_k \epsilon_k^n e_k(\mathbf{x}),$$

where  $\epsilon_k^n = \int_{\Omega} \epsilon^n(\mathbf{x}) e_k(\mathbf{x})$ . And finally the coefficients of  $\tilde{T}^n$  in  $(e_k)_k$  are given by

$$\tilde{T}_k^n = \left( (\mathcal{K}_1)_{k,n} + (\mathcal{K}_2\Theta)_{k,n} \right) + \int_{\Omega} \epsilon^n(\mathbf{x}) e_k(\mathbf{x}).$$

In the previous section, we have constructed our singular value problem which was  $\mathcal{T}L^{\frac{1}{2}}\psi = \sigma v$ , where the matrix  $\mathcal{T}$  was nothing else than the one containing the coefficients of  $T^n$  in the basis  $(e_k)_k$ . Here we have to take account of the noise, and so we denote by  $\tilde{\mathcal{T}}$  the new matrix. The new singular value problem becomes

$$\tilde{\mathcal{T}}L^{\frac{1}{2}}\psi = \sigma v \iff \left( \mathcal{K}_1 + \mathcal{K}_2\Theta + \mathcal{E} \right) L^{\frac{1}{2}}\psi = \sigma v,$$

where we denote  $\mathcal{E}$  by

$$\mathcal{E}_{k,n} = \int_{\Omega} \epsilon^n(\mathbf{x}) e_k(\mathbf{x}).$$

We have then the following result.

**Proposition 3.2.** The singular values of the matrix  $\tilde{\mathcal{T}}L^{\frac{1}{2}}$  are as such

$$\sigma_k(\tilde{\mathcal{T}}L^{\frac{1}{2}})^2 = \left( \sigma_k(\mathcal{T}L^{\frac{1}{2}}) + \tilde{\zeta}_k \right)^2 + \eta_k^2,$$

where we have

$$|\tilde{\zeta}_k| \leq \|P\mathcal{E}L^{\frac{1}{2}}\|_2, \quad \inf_{\|\mathbf{x}\|=1} \|(I-P)\mathcal{E}L^{\frac{1}{2}}\|_2 \leq \eta_i \leq \|(I-P)\mathcal{E}L^{\frac{1}{2}}\|_2$$

and  $P$  is the orthogonal projection onto the column space of  $\mathcal{T}L^{\frac{1}{2}}$ .

*Proof.* By construction, we have

$$\tilde{\mathcal{T}}L^{\frac{1}{2}} = \mathcal{T}L^{\frac{1}{2}} + \mathcal{E}L^{\frac{1}{2}}.$$

The result follows by [22]. □

Since the spectral norm of  $\mathcal{E}L^{\frac{1}{2}}$  is related to  $\sigma^2$  such that  $\|\mathcal{E}L^{\frac{1}{2}}\|_2 \lesssim \sigma^2$ , the Proposition 3.2 tells us that the singular values for the disturbed temperature field have a decay similar to the singular values for the non-noisy temperature field while the variance  $\sigma^2$  is inferior to the singular values. In other words, when the singular value corresponding to the non-disturbed temperature field is less than the variance, the decay stops to be exponential. This result will be supported by numerical computations in the next section.

## 4 Numerical computations

### 4.1 Eigenvalues of the operator $A = B^*B$

The performance of the KL-expansion is related to the fast decay of the eigenvalues ( or singular values) as stated by Proposition 1.1. In the first section, we prove that this decay is of the form

$$\lambda_{k+1}(\mathcal{G}) \leq 4 \left[ \exp\left(\frac{-\pi k^2}{\nu \log\left(\frac{4b}{a}\right)}\right) \right] \lambda_1(\mathcal{G}). \quad (4.1)$$

This result allows us to conclude on the fast decay of the eigenvalues, however a special interest should be taken in the constant  $\frac{b}{a}$ . Indeed, if  $\frac{b}{a}$  tends to infinity, the constant in square brackets tends to 1 and so we lose the fast decay. Fortunately, when this quotient tends to infinity we keep an exponential decay with a square root exponent for a zero RHS, i.e.

$$\lambda_k(\mathcal{G}) \lesssim \exp\left(\frac{-\pi}{\sqrt{8}}\right)^{\sqrt{k}} \quad (4.2)$$

and we can also show that we keep an exponential decay for a non zero RHS. To illustrate that, we consider the one dimensional case, where the eigenvalues of the Laplace operator are known. Let consider the following equation

$$\begin{cases} \partial_t T - \alpha \partial_{x^2} T + \beta T = S, & \text{in } (0, \pi) \times (0, b), \\ T(\cdot, 0) = a(\cdot), & \text{in } (0, \pi). \end{cases} \quad (4.3)$$

In this case the eigenvalues of the Laplace operator are  $r_k = \alpha k^2 + \beta$ . In order to work with finite matrices, we have chosen an integer  $K$  large enough so that the error committed by

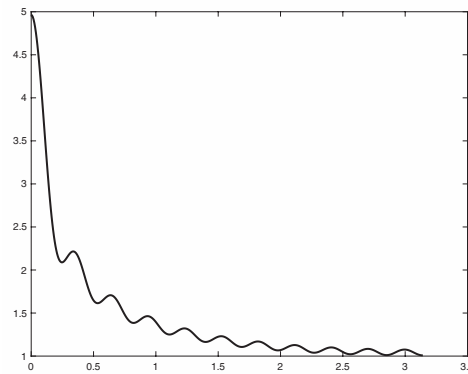


Figure 1: The initial state corresponding to the Fourier coefficients for a truncated number  $K=21$ .

using finite matrices does not affect the result. For a Crank-Nicholson time scheme, we have

$$\frac{b}{a} = \frac{r_K}{r_1} = \frac{\alpha K^2 + \beta}{\alpha + \beta}.$$

Similarly,

$$\frac{b}{a} = \frac{r_K(2 + \Delta tr_1)}{r_1(2 + \Delta tr_K)}$$

for the Backward Euler scheme. It is clear that  $b/a$  grows with  $K$ .

In the two cases, we expect that the eigenvalues decrease less quickly when  $K$  grows. We consider the initial state defined by

$$T_0(x) = \sum_{k=0}^{\infty} \frac{1}{\sqrt{m+1} \log(m+2)} \cos(kx). \tag{4.4}$$

Thus, the initial state belongs to  $L^2(\Omega)$  and have no more regularity. We cut the sum over  $k$  to  $K$  for different value of  $K$  and display the eigenvalues of the matrix  $\mathcal{G}$ , whose coefficients are given by (2.10). The function  $T_0$  truncated for  $K=21$ , is depicted in Fig. 1.

The final time is fixed to  $b=0.1$  and the time step is  $10^{-3}$ . We use the trapezoidal rule for the time integration. The thermal coefficients  $\alpha$  and  $\beta$  are fixed to 1. In Figure 2 we represent the twelve first eigenvalues of  $A$  for different  $K$ .

The eigenvalues seem to decrease to zero exponentially even if  $K$  becomes larger, which is consistent with the upper bound (4.2). We now consider a non zero RHS. For this purpose we define the separated function  $S$  by

$$S(x,t) = e^t \sum_{k=0}^{\infty} \frac{1}{(k+1) \log(k+2)} \cos(kx).$$

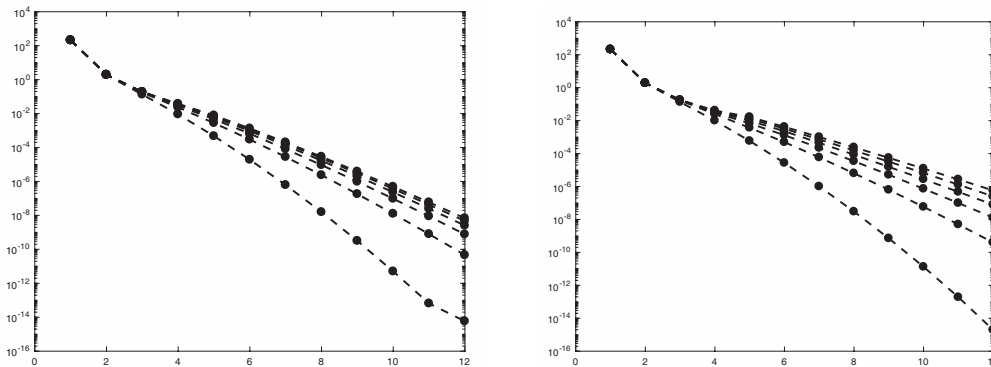


Figure 2: Eigenvalues of  $\mathcal{G}$  for a zero RHS,  $K$  varies from 15 to 100, for the Backward Euler scheme (left) and the Crank-Nicholson scheme (right).

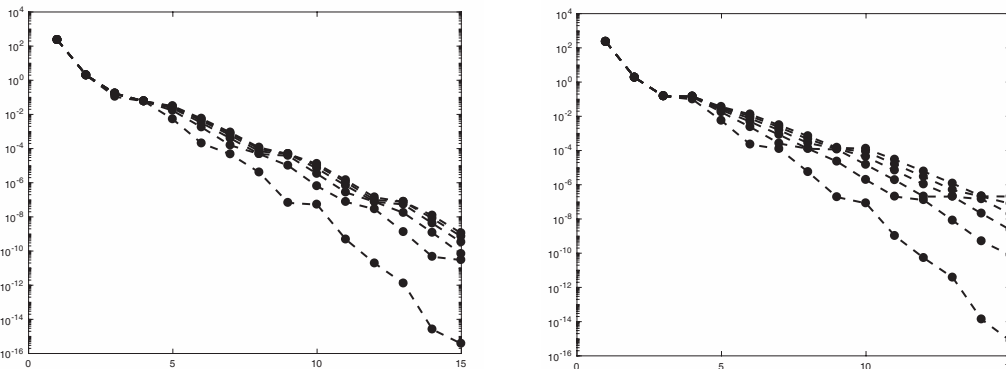


Figure 3: Eigenvalues of  $\mathcal{G}$  for a non zero RHS,  $K$  varies from 15 to 100, for the Backward Euler scheme (left) and the Crank-Nicholson scheme (right).

And we compute the eigenvalues of  $\mathcal{G}$  as shown above. The decay of the eigenvalues is depicted in Figure 3.

The eigenvalues decrease exponentially even if  $K$  grows.

### 4.2 Identification problem

We now present an application of the Karhunen-Loève expansion constructed and analyzed in this article. Let us suppose that we have a temperature field  $(T^n)_{0 \leq n \leq N}$  which is obtain numerically with a Backward Euler scheme or a Crank-Nicholson scheme for the time approximation and the finite element method with Lagrangian basis of degree 1 for the space approximation. In this case the temperature field is known as a matrix  $T$ , such that the columns are vectors containing the values of the temperature field at the mesh points, which is consistent with the data measured by a camera. The final goal is to recover the parameters  $(\alpha, \beta)$ . However, in practice, some noises appear due to the

camera's defects. We can then construct the following matrix with artificial noise

$$\tilde{T}_{i,j} = T_{i,j} + \epsilon_{i,j},$$

where  $\epsilon_{i,j}$  fits with the difference between the real temperature and the temperature measured by the camera. Thus,  $\epsilon_{i,j}$  will be small for a camera with a high precision. To recover the thermal parameters, we write the equation under the variational formulation and expand the solution by using the Karhunen-Loève expansion. Hence, we obtain more equations than unknowns. However, since the temperature field  $\tilde{T}$  does not solve the discrete heat equation, we have an ill-posed inverse problem. We must then choose few equations and use a least-squares method. The Karhunen-Loève expansion then allows us to keep the equations with the most information. This issue was handled in the continuous case in [5, 7–9].

#### 4.2.1 Karhunen-Loève expansion

We begin by just considering the perturbed matrix  $\tilde{T}$  and forget how it is obtained. If we construct our Karhunen-Loève expansion by expanding the temperature field in the shape functions, it yields

$$B = \tilde{T}^T M \quad \text{and} \quad B^* = \tilde{T} L,$$

where  $M$  denotes the mass matrix. As shown in the Section 4, we set  $\psi = L^{\frac{1}{2}} \phi$  and  $w = D^{\frac{1}{2}} L_M^T v$ , where we use the Cholesky factorization for semi definite symmetric matrices ( $M = L_M D L_M^T$ ). We then obtain

$$L^{\frac{1}{2}} \tilde{T}^T L_M D^{\frac{1}{2}} w = \sigma \psi \quad \text{and} \quad \left( L^{\frac{1}{2}} \tilde{T}^T L_M D^{\frac{1}{2}} \right)^T \psi = \sigma w.$$

Thus, we can apply the classical Singular Value Decomposition

$$L^{\frac{1}{2}} \tilde{T}^T L_M D^{\frac{1}{2}} = \Psi \Sigma W^T. \tag{4.5}$$

If we construct the matrix  $V = L_M^{-t} D^{-\frac{1}{2}} W$ , we get the coefficients of the eigenvectors  $v_k$  in the shape functions. Thus,  $v_k(x) = \sum_i V_{i,k} u_i(x)$ . Moreover,  $\phi_k^n = \Phi_{n,k}$  where  $\Phi = L^{-\frac{1}{2}} \Psi$ . Hence, we obtain the Karhunen-Loève expansion

$$\tilde{T}^n(x) = \sum_k \sigma_k \phi_k^n v_k(x). \tag{4.6}$$

This expansion allows us to represent the temperature field  $(\tilde{T}^n)_n$  by a sum of few elements with a small error since the singular values of  $B$  decrease fastly to zero. For numerical example, we just consider the case of a Backward Euler scheme. We thus compute



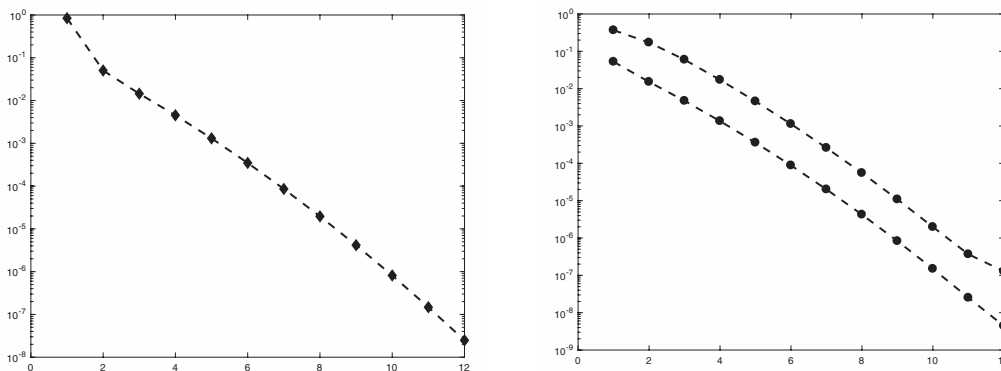


Figure 4: (left) Singular values of  $B$ . (right) Truncation error in norm  $L^2$  and in semi-norm  $H^1$ .

the singular values of  $B$  as explained before, and compare them to the error made by the truncated sum in the discrete anisotropic norms

$$\left( \sum_{n=0}^N \omega_n \|\tilde{T}^n - \tilde{T}_M^n\|_{L^2(\Omega)}^2 \right)^{\frac{1}{2}} \quad \text{and} \quad \left( \sum_{n=0}^N \omega_n \|\tilde{T}^n - \tilde{T}_M^n\|_{H^1(\Omega)}^2 \right)^{\frac{1}{2}}.$$

For the computation, we fix  $(\alpha, \beta) = (1, 1)$ ,  $S(x, t) = e^t |x - \frac{\pi}{2}|$  with the numbers of time points and space points equal to 101, and the initial time defined in (4.4). We first present this result when the variance is zero as shown in Figure 4.

This result show that the exponential decay of the eigenvalues leads to an exponential decay of the error. This is in accordance with the proposition 1.1 for the  $L^2$ -norm. For the semi  $H^1$ -norm we can deduce a similar result

$$\|\tilde{T}^n - \tilde{T}_M^n\|_{H^1(\Omega)}^2 = \sum_{k,l=M+1}^r \sigma_k \phi_k^n \sigma_l \phi_l^n \int_{\Omega} v'_k(x) v'_l(x),$$

where  $r$  is the minimum between the number of time point and space point. By summing over  $n$  it yields

$$\begin{aligned} \left( \sum_{n=0}^N \omega_n \|\tilde{T}^n - \tilde{T}_M^n\|_{H^1(\Omega)}^2 \right)^{\frac{1}{2}} &= \left( \sum_{k,l=M+1}^r \sigma_k \sigma_l \int_{\Omega} v'_k(x) v'_l(x) \sum_{n=0}^N \omega_n \phi_k^n \phi_l^n \right)^{\frac{1}{2}} \\ &= \left( \sum_{k=M+1}^r \sigma_k^2 \int_{\Omega} v'_k(x) v'_k(x) \right)^{\frac{1}{2}} \\ &\leq \max_{k \in \{M+1, r\}} |v_k|_{H^1(\Omega)} \left( \sum_{k=M+1}^r \sigma_k^2 \right)^{\frac{1}{2}} \end{aligned}$$

by the orthonormality of  $(\phi_k)_k$  for the weighted inner product  $(\phi, \psi)_\omega = \sum_{n=0}^N \omega_n \phi^n \psi^n$ . Moreover, the eigenvectors  $(v_k)_k$  belong to  $H^1(\Omega)$  ([5, Proposion 3.1]), thus the quantity  $\max_{k \in \{M+1, r\}} |v_k|_{H^1(\Omega)}$  is finite.

We now examine the case where the variance is non-zero. We compute the singular values for different variances, and display the incidence on the truncation error in Figure 5.

The reason why the singular values have no exponential decay comes from the Proposition 3.2. We have the following link between the singular values for a temperature field with and without noise

$$\sigma_k(\tilde{\mathcal{T}}L^{\frac{1}{2}})^2 = \left(\sigma_k(\mathcal{T}L^{\frac{1}{2}}) + \xi_k\right)^2 + \eta_k^2,$$

where  $\xi_k$  and  $\eta_k$  depend of the variance  $\sigma^2$ . Intuitively, the exponential decay of the singular values shows that the information contained in the matrix  $\tilde{\mathcal{T}}$  is poor. The exchange of heat is a smooth phenomenon, the variation of temperature near two time points or two space points is small. However, when we add some noise we change the regularity of the solution and we add a lot of information. It is then reasonable to lose the exponential decay of the singular value.

#### 4.2.2 Time-Space equation

Let us return to the identification problem. We identify the parameters thanks to the equation (4.3) discretized by a Backward Euler scheme, i.e

$$\begin{cases} \frac{T^{n+1}-T^n}{\Delta t} - \alpha \partial_{x^2} T + \beta T = S, & \text{in } (0, \pi) \times (0, b), \\ T(\cdot, 0) = a(\cdot), & \text{in } (0, \pi). \end{cases} \tag{4.7}$$

To reach this point, we proceed in two steps. The first one, assuming the temperature field to be non disturbed, is to write the equation satisfied by the parameters and by a least-square method to recover the parameters. Note that since the temperature  $T^n$  is the exact solution to our problem, we recover the parameters exactly. As a second step, we use the disturbed solution  $\tilde{T}^n$  and solve the equation by using the least-squares method as if the temperature field was non disturbed. In this case, an error is committed since  $\tilde{T}^n$  is not solution of (4.7). However, we are hoping to keep the most information by using the KL-expansion.

Assuming the temperature field to be computed by a FEM method and not disturbed by a Gaussian noise. Considering equation (4.7) in a weak form, we get

$$\int_{\Omega} \frac{T^{n+1}-T^n}{\Delta t} u + \alpha \int_{\Omega} (T^{n+1})' u' + \beta \int_{\Omega} T^{n+1} u = \theta^{n+1} \int_{\Omega} f u, \quad \text{for all } u \in H^1(\Omega).$$

The next step is to use the Karhunen-Love's expansion of  $(T^n)_n$  and set  $u = v_l$  ( $v_l$  belongs

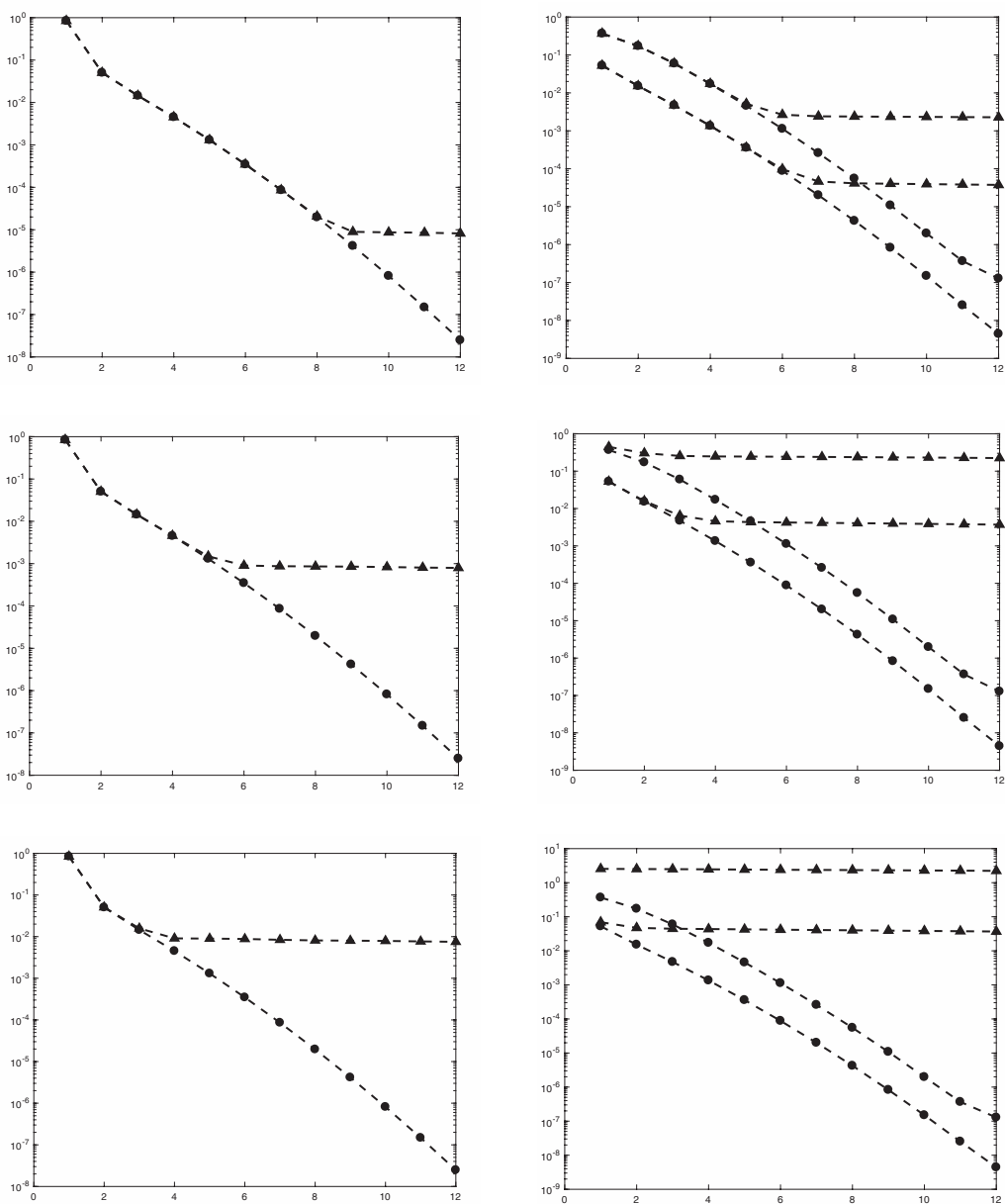


Figure 5: (left) Singular values of  $B$ , the circle corresponding to a zero variance, and the triangle corresponding to the case where the variance vary from  $10^{-4}$  to  $0.1$ . (right) The truncation error in norm  $L^2$  and  $H^1$  for non-zero variance (triangle), and zero variance (circle)

to  $H^1(\Omega)$  by [5, Proposion 3.1]). By the orthogonality of  $(v_k)_k$  we get

$$\sigma_l \frac{\phi_l^{n+1} - \phi_l^n}{\Delta t} + \alpha \sum_k \sigma_k \phi_k^{n+1} \int_{\Omega} v'_k v'_l + \beta \sigma_l \phi_l^{n+1} = \theta^{n+1} \int_{\Omega} f v_l.$$

Setting  $(D\Phi)_{n+1,l} = \frac{\phi_l^{n+1} - \phi_l^n}{\Delta t}$ ,  $K$  the stiffness matrix, and  $\tilde{A}$  such that

$$(A)_{n,l} = \theta^n \int_{\Omega} f v_l,$$

we have the following equation

$$\left( D\Phi\Sigma + \alpha\Phi\Sigma V^T K V + \beta\Phi\Sigma - A \right)_{n+1,l} = 0, \tag{4.8}$$

where  $\Phi$  and  $V$  are the singular vectors defined in (4.6), and  $\Sigma$  the diagonal matrix of the singular values. We set

$$\mathcal{H}(\alpha, \beta) = D\Phi\Sigma + \alpha\Phi\Sigma V^T K V + \beta\Phi\Sigma - A.$$

By equation (4.8), the identification problem can be expressed as finding a couple  $(\alpha, \beta)$  such that

$$\|\mathcal{H}(\alpha, \beta)\|_F = 0.$$

We then search the couple which minimizes this quantity. And setting  $\frac{\partial}{\partial \alpha} \|\mathcal{H}(\alpha, \beta)\|_F = 0$  and  $\frac{\partial}{\partial \beta} \|\mathcal{H}(\alpha, \beta)\|_F = 0$  we get two equations for two parameters. However, when we consider the full matrix  $\mathcal{H}(\alpha, \beta)$ , the least-square procedure can fail because of the size of the matrix, even if we consider a non disturbed temperature field. In order to fix it, we can consider the truncated matrix  $\mathcal{H}_N(\alpha, \beta)$ , where we take the  $(N \times N)$ -first term. Truncating the matrix amounts to truncating the KL-expansion, and as viewed above it is a way of keeping the most information with the less amount of data. Thus, using the KL-expansion has a significance in this case.

For numerical computations, we consider two initial states, one singular and a second smoother. In order not to perturb too much the temperature field, we re-scale the initial state and the source term such that the temperature computed by the FEM is close to 100. Here, the Fourier coefficients are respectively given by

$$\frac{27.5}{\sqrt{m+1} \log(m+2)} \quad \text{and} \quad \frac{69}{(m+1)^{\frac{3}{2}} \log(m+2)}.$$

We depict in Figure 6 the two initial states. Intuitively, the singular initial state can be represented with fewer modes than the smooth one. The RHS will be  $S(x, t) = e^t |x - \frac{\pi}{2}|$  as the one used to show the decay of the singular value. The numbers of time points and space points are fixed to 500, and the domain on which we compute the solution is  $(0, \pi) \times (0, 0.1)$ . The variance will be varying between 0.01 and 1. Since we add a random matrix, the result can be different from one experience to another. In order to avoid this variation, the result will be an average of several computations. We take  $(\alpha, \beta) = (\exp(1), 3\pi) \approx (2.72, 9.42)$  as shown in [5]. The results are presented in Table 1. For our simulation, we add a noise on the temperature field and perform the Karhunen-Loève expansion on the disturbed field. In [5], the noise is introduced in the Gram matrix. The result depicted in Table 1 seems to be less sensitive to the noise than in [5].

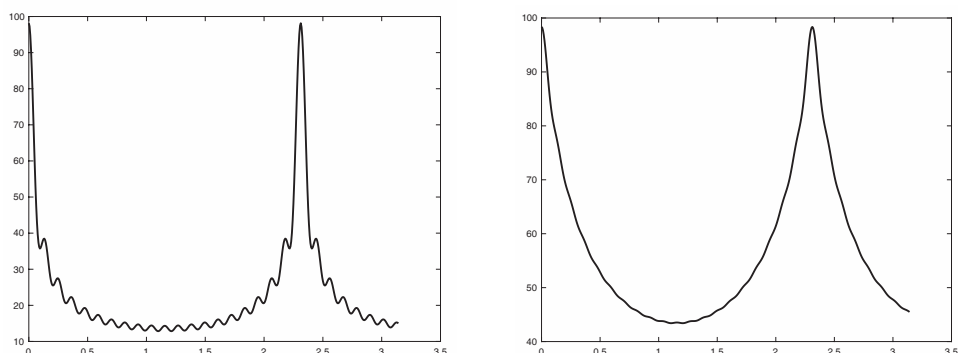


Figure 6: (left) Singular initial state. (right) Smooth initial state.

Table 1:

	Singular initial state	Smooth initial state
$var = 0.01$	(2.72,9.42)	(2.72,9.42)
$var = 0.5$	(2.03,9.70)	(1.88,9.52)
$var = 1$	(1.10,9.65)	(0.95,9.79)

## 5 Conclusion

This article aims to adapt the Karhunen-Loève expansion to the discrete transient heat transfer equation. An error estimate is given based on the exponential decay of the eigenvalues of the spatial correlation matrix. The key point in this paper is the decomposition of the temperature field in the Laplace operator's eigenvalues, which allows us to conclude on the performance of the Karhunen-Loève expansion in the case we consider the spatial correlation matrix. Since the field is assumed to be discrete in time, an other analysis is performed on the singular values. This allows us to show the performance of considering the time correlation matrix which give us a less costly eigenvalue problem to solve since the number of time points is generally smaller than the number of space points. Finally, we show the lack of decay in the case where the temperature field is disturbed by a Gaussian white noise. This is consistent with the fact that the data is enriched. All this results are successfully supported by some numerical experiments. Then, we present a practical way to compute the Karhunen-Loève expansion, in the case where the temperature field is computed by the finite element method. Finally the usefulness of the KL-expansion is enhanced by an direct application to an identification problem with noisy data.

## References

- [1] M. Azaiez and F. Ben Belgacem, Karhunen Loèves truncation error for bivariate functions, Computer Methods in Applied Mechanics and Engineering, 290 (Supplement C)(2015), 57 -

72.

- [2] B. Beckermann and A. Townsend, On the singular values of matrices with displacement structure, *SIAM Journal on Matrix Analysis and Applications*, 38(2017), 1227-1248.
- [3] A. Alexanderian, A brief note on the Karhunen-Loeve expansion.
- [4] D. Fasino and V. Olshevsky, How bad are symmetric Pick matrices?, *Advanced Signal Processing Algorithms, Architectures, and Implementations X*, 4116(2000), 147-156.
- [5] E. Ahusborde, M. Azaiez, F. Ben Belgacem and E. Palomo Del Barrio, Mercer's spectral decomposition for the characterization of thermal parameters, *J. Comput. Physics*, 29(4)(2015), 1-19.
- [6] A. Quarteroni, A. Manzoni and F. Negri, *Reduced Basis Methods for Partial Differential Equations: An Introduction*, Springer International Publishing, 2015.
- [7] A. Godin, Estimation sur des bases orthogonales des propriéts thermiques de matériaux htrogenes propriéts constantes par morceaux, Thse de doctorat, Bordeaux 1, 2013.
- [8] E. Palomo Del Barrio, J.-L. Dauvergne and C. Pradere, Thermal characterization of materials using Karhunen-Loeve decomposition techniques C Part I. Orthotropic materials, *Inverse Problems in Science and Engineering*, 20(8)(2012), 1115-1143.
- [9] E. Palomo Del Barrio, J.-L. Dauvergne and C. Pradere, Thermal characterization of materials using Karhunen-Loeve decomposition techniques C Part II. Heterogeneous materials, *Inverse Problems in Science and Engineering*, 20(8)(2012), 1145-1174.
- [10] M. Stewart, Perturbation of the SVD in the presence of small singular values, *Linear Algebra and its Applications*, 419(1)(2006), 53 - 77.
- [11] H. Weyl, Das asymptotische Verteilungsgesetz der Eigenwerte linearer partieller Differentialgleichungen (mit einer Anwendung auf die Theorie der Hohlraumstrahlung), *Mathematische Annalen*, 71(1912), 441-479.
- [12] J. Mercer, Functions of Positive and Negative Type, and their Connection with the Theory of Integral Equations, *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 209(1909), 415-446.
- [13] J. B. Conway, *A Course in Functional Analysis*, Springer New York, 1994.
- [14] I. T. Jolliffe, *Principal Component Analysis*, Springer Verlag, 1986.
- [15] A. Buffa, Y. Maday, A. T. Patera, C. Prud'homme and G. Turinici, A priori convergence of the Greedy algorithm for the parametrized reduced basis method, *ESAIM: Mathematical Modelling and Numerical Analysis*, 46(3)(2012), 595-603.
- [16] J. S. Hesthaven, G. Rozza and B. Stamm, *Certified Reduced Basis Methods for Parametrized Partial Differential Equations*, Springer International Publishing, 2015.
- [17] S. Volkwein, Proper orthogonal decomposition: Theory and reduced-order modelling, *Lecture Notes*, University of Konstanz, 4(4)(2013).
- [18] H. Brézis, P.G. Ciarlet and J.L. Lions, *Analyse fonctionnelle: théorie et applications*, Dunod, 1999.
- [19] K. Schacke, *On the Kronecker Product*, 2004.
- [20] L. Grubisic and D. Kressner, On the eigenvalue decay of solutions to operator Lyapunov equations, *Systems and Control Letters*, 73(Supplements C)(2014), 42 - 47.
- [21] C.D. Meyer, *Matrix Analysis and Applied Linear Algebra*, Society for Industrial and Applied Mathematics, 2000.
- [22] G.W. Stewart, A note on the perturbation of singular values, *Linear Algebra and its Applications*, 28(Supplement C)(1979), 213 - 216.
- [23] K. Fan, Maximum properties and inequalities for the eigenvalues of completely continuous operators, *Proceedings of the National Academy of Sciences*, 37(11)(1951), 760-766.

- [24] A. Guven, Quantitative Perturbation Theory for Compact Operators on a Hilbert Space. Ph.D Dissertation, Queen Mary University of London, 2016.
- [25] Z. Luo, J. Zhu, R. Wang and I.M. Navon, Proper orthogonal decomposition approach and error estimation of mixed finite element methods for the tropical Pacific Ocean reduced gravity model, *Computer Methods in Applied Mechanics and Engineering*, 196(41)(2007), 4184 - 4195.
- [26] M. Hinze and S. Volkwein, Proper orthogonal decomposition surrogate models for nonlinear dynamical systems: Error estimates and suboptimal control, *Dimension reduction of large-scale systems*, Springer, 2005, 261-306.
- [27] R. Ghosh and Y. Joshi, Error estimation in POD-based dynamic reduced-order thermal modeling of data centers, *International Journal of Heat and Mass Transfer*, 57(2)(2013), 698 - 707.
- [28] M. Griebel and H. Harbrecht, Approximation of bi-variate functions: singular value decomposition versus sparse grids, *IMA journal of numerical analysis*, Oxford University Press, 34(1)(2013), 28-54.
- [29] A. Cohen and R. Devore, Kolmogorov widths under holomorphic mappings, *IMA Journal of Numerical Analysis*, Oxford University Press, 36(1)(2016), 1-12.