Use of Tensor Formats in Elliptic Eigenvalue Problems

Wolfgang Hackbusch, Boris N. Khoromskij*
Max-Planck-Institut für Mathematik in den Naturwissenschaften,
Inselstr. 22-26, D-04103 Leipzig, Germany; \{wh, bokh\}@mis.mpg.de

Stefan Sauter
Universität Zürich, Switzerland; stas@math.uzh.ch

Eugene E. Tyrtyshnikov†
Institute of Numerical Mathematics, Russian Academy of Sciences,
Gubkina 8, 119333 Moscow, Russia; tee@inm.ras.ru

October 16, 2008

Abstract
We investigate approximations by finite sums of products of functions with separated variables to eigenfunctions of multivariate elliptic operators, and especially conditions providing an exponential decrease of the error with respect to the number of terms. The results of the consistent use of tensor formats can be regarded as a base for a new class of iterative eigensolvers with almost linear complexity in the univariate problem size.

The results of numerical experiments clearly indicate the linear-logarithmic scaling of low-rank tensor method in the univariate problem size. The algorithms work equally well for the computation of both, minimal and maximal eigenvalues of the discrete elliptic operators.

AMS Subject Classification: 65F30, 65F50, 65N35, 65F10
Key words: elliptic operators, spectra, eigenfunctions, separation of variables, separable approximations, matrix approximations, low-rank matrices, Kronecker products, multi-dimensional matrices, tensors, numerical methods, iterative algorithms

1 Introduction
Recent advances in tensor approximation methods applied to the functions and operators in $\mathbb{R}^d$ lead to the natural idea of solving multi-dimensional boundary and eigenvalue problems

---

*Work of BNK was partially supported by FIM, ETH Zuerich during his visit of the ETH/University of Zuerich in May - July, 2008.

†This work was partially supported by DFG, Russian Fund of Basic Research (grant 08-01-00115) and a Priority Research Grant of the Department of Mathematical Sciences of the Russian Academy of Sciences.
in tensor-product formats. This idea has been first time formulated in [2] in the very general setting. The recent results on efficient methods of tensor approximation and their application in electronic structure calculations can be found in [11, 3, 8, 15, 18] and others.

In this paper we investigate approximations by finite sums of products of functions with separated variables to eigenfunctions of elliptic operators with smooth coefficients, and especially conditions providing an exponential decrease of error in the number of terms.

We consider the model eigenvalue problem: Find a pair \((\lambda, u) \in \mathbb{R} \times H^1_0(\Omega) \setminus \{0\}\) such that

\[
\Lambda u = \lambda u \quad \text{in } \Omega, \\
u = 0 \quad \text{on } \partial \Omega
\]

with the elliptic differential operator \(\Lambda\) of the form

\[
\Lambda := -\text{div}(A \text{grad } u) + \langle b, \nabla u \rangle + cu,
\]

where \(\Omega \in \mathbb{R}^d\) is some bounded or unbounded tensor-product domain, and the operator coefficients \(A, b, c\) in (1.2) are supposed to be smooth (analytic) in \(\Omega\).

We will prove that the eigenfunctions of problem (1.1) allow separable approximation that converges exponentially in the number of terms.

Problem (1.1) is discretised by the Galerkin FEM with tensor-product basis functions, so that the arising stiffness and mass matrices of size \(n^{\otimes d}\) (with \(d\)-fold product \(n^{\otimes d} = n \times \ldots \times n\)) are represented in the low Kronecker rank format with the storage requirements and computational complexity of order \(O(dn)\).

In this paper, for the ease of presentation, we use simple iterative solvers such as the power method or the Lanczos iteration, though algorithms of better choice can be easily adapted to our concept. Due to the above mentioned approximation results for the continuous solutions, and relying on the rank-structured representation of all matrices involved, we propose to solve the corresponding high-dimensional algebraic eigenvalue problem of the size \(n^{\otimes d}\) in the low tensor-rank format. To this end, we introduce the so-called “truncated iterations”, where most of the intermediate vectors have to be approximated in some fixed rank-structured tensor product form. The corresponding rank truncation performed at each iteration makes use of recently developed methods (cf. [18, 22, 8]). For the class of rank structured matrices, our algorithm can be shown to have storage and complexity bounds of order \(O(r^d n + r n d)\), or even \(O(dRrn)\), where \(r, R\) are the small (often fixed) rank parameters with the theoretical bounds \(r = O(\log n)\) and \(R = O(\log n |\log \varepsilon|)\).

We notice that in the case \(d = 2\) the rank truncation operator is realised by the “truncated SVD” method applied to the rank-\(R\) matrix. It is, in practice, a finite algorithm (with complexity at most \(O(nR^2 + R^3)\)) providing the best rank-\(r\) approximation to the current iterand (see numerics in Sections 2 and 6.2). Hence, numerical results for \(d = 2\) can be viewed as the reference cases, demonstrating nearly optimal performance of the proposed techniques.

The results of the consistent use of tensor formats can be regarded as a base for a new class of iterative eigensolvers with almost linear complexity in the univariate problem size.

The rest of the paper is organised as follows. In Section 2 we give motivating numerical illustrations for the Laplace operator in \(d = 2\), which clearly indicate a spectacular gain by the truncated Lanczos iteration compared with the standard full-format implementation. In Section 3, we briefly describe the tensor product formats for representing multivariate
functions of the continuous and discrete arguments. Section 4 proves the existence of separable approximation for the eigenfunctions of an elliptic operator posed in \( \mathbb{R}^d \). This result is of principal significance for understanding the rigorous mathematical basis for applying tensor methods in multidimensional setting. Section 5 describes the favourable properties of tensor formats in discrete elliptic eigenvalue problems, while Section 6 presents numerical illustrations for certain spectral problems in dimensions \( d = 2, 3 \).

2 Motivating Numerics

In this research we are motivated by the nice solution structure for the 2D Laplace operator

\[
\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}
\]

and the eigenproblem

\[-\Delta u(x, y) = \lambda u(x, y), \quad (x, y) \in [0, \pi]^2.\]

If \( u \) is zero on the boundary, then the eigenvalues and eigenfunctions are

\[
\lambda_{kl} = k^2 + l^2, \quad u_{kl}(x, y) = \sin kx \sin ly, \quad k, l = 1, 2, \ldots.
\]

A matrix counterpart of the Laplace operator can be taken in the form

\[
Mu = \lambda u, \quad M = A \otimes I + I \otimes A,
\]

where \( \otimes \) denotes the Kronecker (tensor) product, \( I \) is the identity and

\[
A = \begin{bmatrix}
2 & -1 & -1 & \cdots & -1 & 2 \\
-1 & 2 & \cdots & \cdots & \cdots & -1 \\
\vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\
-1 & 2 & \cdots & \cdots & 2 & -1 \\
\end{bmatrix}.
\]

Let \( A \) and \( I \) be of order \( n \). Then the matrix eigenvalue problem \( Mx = \lambda x \) can be solved explicitly. The eigenvalues are

\[
\lambda_{kl} = 4 \sin^2 \frac{\pi k}{2(n + 1)} + 4 \sin^2 \frac{\pi l}{2(n + 1)}, \quad 1 \leq k, l \leq n,
\]

and the corresponding eigenvectors are exactly represented as tensor products:

\[
x_{kl} = u^k \otimes v^l,
\]

where \( u^k \) and \( v^l \) are \( n \)-dimensional vectors with the entries

\[
u^k_s = \sin \frac{\pi ks}{n + 1}, \quad 1 \leq s \leq n; \quad v^l_t = \sin \frac{\pi lt}{n + 1}, \quad 1 \leq t \leq n.
\]

In general, the cost of traditional iterative eigenvalue algorithms is higher than linear, possibly \( O(n^2) \). The tensor structure of eigenvectors allows to modify the eigensolvers so that the cost of one iteration reduces to \( O(n) \).

As an example of an iterative eigensolver we discuss the Lanczos method.
Algorithm 2.1 Let $M$ be a real symmetric matrix of order $N$. The following steps produce the Ritz values approximating the eigenvalues of $M$:

- Choose an initial vector $p_1$ with $||p_1|| = 1$ and set $p_0 = 0$, $b_0 = 0$.
- For $k = 1, 2, ...$ compute
  
  \[
  z_k = Mp_k, \\
  a_k = (z_k, p_k), \\
  q_k = z_k - a_k p_k - b_{k-1} p_{k-1}, \\
  b_k = ||q_k||, \\
  p_{k+1} = q_k / b_k. 
  \]

- Compute the Ritz values as the eigenvalues of the projected $k \times k$ matrix
  
  \[ M_k = P_k^T M P_k, \quad P_k = [p_1, ..., p_k], \]

  which is the symmetric tridiagonal matrix consisting of the values $a_k, b_k$.

Now we consider the same computations with the following vector truncation. If $x \in \mathbb{R}^{n^2}$, $T_\varepsilon(x)$ is defined as a vector $y$ of the form

\[
 y = \sum_{t=1}^{r} u_t \otimes v_t, \quad u_t, v_t \in \mathbb{R}^n, 
\]

with minimal $r$ such that

\[ ||x - y|| \leq \varepsilon \]

for a fixed $\varepsilon$. Note that if $x$ is in the above tensor format and a matrix $M$ is of the tensor form (cf. Section 3.5)

\[ M = \sum_{s=1}^{R} A_s \otimes B_s, \quad R \ll n, \]

then the matrix-vector multiplication $z = M y$ has a tensor format with $R r$ summands. A natural idea to reduce this number is to replace $z$ with $T_\varepsilon(z)$. Then, the Lanczos algorithm transforms to the following tensor-format scheme.

Algorithm 2.2 The Lanczos steps with tensor approximation of computed vectors reads as follows:

- Choose an initial vector $p_1$ so that $p_1 = T_\varepsilon(p_1)$ with $||p_1|| = 1$ and set $p_0 = 0$, $b_0 = 0$.
- For $k = 1, 2, ...$ compute
  
  \[
  z_k = Mp_k, \\
  a_k = (z_k, p_k), \\
  q_k = T_\varepsilon(z_k - a_k p_k - b_{k-1} p_{k-1}), \\
  b_k = ||q_k||, \\
  p_{k+1} = q_k / b_k. 
  \]
Compute the Ritz values from the generalized eigenvalue problem

\[ M_k u = \lambda N_k u, \]

where \( M_k = P_k^T M P_k, \ N_k = P_k^T P_k. \)

We can compare the performance of these two algorithms. The truncation parameter was chosen so that every vector in the Krylov subspace basis is approximated in the tensor format so that the tensor rank (the number of tensor-product terms in the sum) does not exceed 7 or the relative approximation accuracy is less than \( 10^{-2} \). In the table below we can see the timings obtained on a 2.7 GHz workstation:

<table>
<thead>
<tr>
<th>( n )</th>
<th>1000</th>
<th>2000</th>
<th>4000</th>
<th>6000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lanczos time (sec)</td>
<td>2.8</td>
<td>12.1</td>
<td>76.7</td>
<td>224.9</td>
</tr>
<tr>
<td>Tensor Lanczos time (sec)</td>
<td>0.4</td>
<td>0.7</td>
<td>1.5</td>
<td>2.2</td>
</tr>
</tbody>
</table>

The accuracy in the maximal eigenvalue was about \( 10^{-3} \) in both methods, even a bit better with truncations. The table shows that the time of the tensor-truncated Lanczos grows linearly in \( n \) while the original Lanczos time increases actually faster as \( n^2 \). For \( n = 6000 \) we observe an acceleration by a factor 100.

Thus, we are interested to find if we might have the same effect when using tensor formats in matrix eigensolvers for more general elliptic problems,

\[ \Lambda u = \lambda u, \quad u : [0, \pi]^d \rightarrow \mathbb{R}, \]

where \( \Lambda \) is the elliptic operator with smooth enough (or separable) coefficients. For this purpose, we next introduce the tensor representations.

### 3 Description of tensor formats

#### 3.1 Tensor spaces and tensor representations

Several continuous and discrete spaces considered in this paper are tensor spaces of order \( d \), where in our application \( d \) equals the spatial dimension of the eigenvalue problem. Let

\[ \mathbb{W} = W_1 \otimes W_2 \otimes \ldots \otimes W_d \quad (3.1) \]

be the notation for the underlying tensor space. By definition, each \( w \in \mathbb{W} \) can be written as a sum

\[ w = \sum_k w_k^{(1)} \otimes w_k^{(2)} \otimes \ldots \otimes w_k^{(d)} \quad (w_k^{(j)} \in W_j). \quad (3.2) \]

An interesting subset are those elements which require only \( R \) terms. They form the set

\[ \mathcal{M}_R = \left\{ w \in \mathbb{W} : w = \sum_{k=1}^R w_k^{(1)} \otimes w_k^{(2)} \otimes \ldots \otimes w_k^{(d)}, w_k^{(j)} \in W_j \right\}. \]
We say that elements \( w \in \mathcal{M}_R \) with \( w \notin \mathcal{M}_{R-1} \) have the tensor rank \( R \). Obviously, tensors \( w \in \mathcal{M}_R \) can be represented by the description of \( Rd \) elements \( w_k^{(j)} \in W_j \). Hence, the cost is linear in \( d \).

As in the Galerkin method, the replacement of the spaces \( W_j \) by subspaces \( V_j \subset W_j \) (1 \( \leq j \leq d \)) is of interest. They lead to the tensor subspace
\[
\mathbb{V} = V_1 \otimes V_2 \otimes \ldots \otimes V_d \subset \mathbb{W}.
\]

Setting \( r_j := \text{dim} V_j \) and choosing \( a \) (without loss of generality orthonormal) basis \( \{ \phi_k^{(j)} : 1 \leq k \leq r_j \} \) of \( V_j \), we can represent each \( v \in \mathbb{V} \) by
\[
v = \sum_k b_k \phi_k^{(1)} \otimes \phi_k^{(2)} \otimes \ldots \otimes \phi_k^{(d)},
\]
where the multi-index \( k = (k_1, \ldots, k_d) \) runs over all \( 1 \leq k_j \leq r_j, 1 \leq j \leq d \).

Let \( r = (r_1, \ldots, r_d) \in \mathbb{N}^d \) be a tuple of dimensions. We define
\[
\mathcal{M}_r = \left\{ v \in \mathbb{W} \text{ as in } (3.4) \text{ with } b_k \in \mathbb{R} , \{ \phi_k^{(j)} : 1 \leq k \leq r_j \} \subset W_j \text{ orthonormal system} \right\}.
\]

A representation of \( w \in \mathcal{M}_r \) in the form (3.4) is also called a Tucker representation of Tucker rank \( r \). A representation of \( w \in \mathcal{M}_r \) in the Tucker format (3.4) requires \( \prod_{j=1}^d n_j \) reals and the representation of \( \sum_{j=1}^d n_j \) vectors \( \phi_k^{(j)} \).

### 3.2 Tensor approximation, tensor truncation

For any element \( w \in \mathbb{W} \) there are numbers \( R \) and tuples \( r \) such that \( w \in \mathcal{M}_{R^*} \) and \( w \in \mathcal{M}_{r^*} \), but the ranks \( R^* \) and \( r^* \) may be rather large. A representation by one of the tensor formats (3.2) or (3.4) is only of interest if the respective ranks are small enough. Therefore, given \( w \in \mathbb{W} \) we search for approximations \( v \in \mathcal{M}_R \) or \( v \in \mathcal{M}_r \) with suitably small ranks \( R \) or \( r \).

When we fix the set \( \mathcal{S} := \mathcal{M}_R \) or \( \mathcal{S} := \mathcal{M}_r \), the smallest error is described by
\[
\sigma(w, \mathcal{S}) := \inf_{v \in \mathcal{S}} \| w - v \|.
\]

In the following we will give examples where \( \sigma(w, \mathcal{S}) \) decays exponentially with the rank \( R \) or \( \min \{ r_j : 1 \leq j \leq d \} \), respectively. In general, the infimum in the definition of \( \sigma(w, \mathcal{S}) \) cannot be replaced by a minimum, since for \( \mathcal{S} = \mathcal{M}_R \) a minimiser is not necessarily existing.

Although, the minimiser exists for \( \mathcal{S} = \mathcal{M}_r \), its computation can be performed only approximately. Therefore, in practice, one has to determine a \( v \in \mathcal{S} \) such that \( \| w - v \| \) comes close to \( \sigma(w, \mathcal{S}) \). The replacement of \( w \) by such a \( v \in \mathcal{S} \) is called the tensor truncation to \( \mathcal{S} \) and denoted by
\[
w \mapsto v = \begin{cases} T_Rw & \text{if } \mathcal{S} = \mathcal{M}_R, \\ T_rw & \text{if } \mathcal{S} = \mathcal{M}_r. \end{cases}
\]

Heuristic methods for computing the rank structured approximations in different problem settings are discussed in \([5, 19, 27, 18, 8, 7]\).

In the particular case of \( d = 2 \), the difficulties mentioned above do not appear. The minimiser of \( \inf_{v \in \mathcal{S}} \| w - v \| = \min_{v \in \mathcal{S}} \| w - v \| \) is the result of the truncated singular value decomposition. Furthermore, the representations by \( \mathcal{M}_R \) and \( \mathcal{M}_r \) with \( r = (R, R) \) coincide.
3.3 Application to function spaces

Let \( I = I_1 \times I_2 \times \ldots \times I_d \) be the product of (possibly infinite) intervals \( I_j \subset \mathbb{R} \). Then \( L^2(I) \) is the tensor space \( L^2(I_1) \otimes L^2(I_2) \otimes \ldots \otimes L^2(I_d) \). The tensor product \( w = \bigotimes_{j=1}^d w^{(j)} \) of \( w^{(j)} \in L^2(I_j) \) corresponds to the pointwise product \( w(x) = \prod_{j=1}^d w^{(j)}(x_j) \).

If \( w \) is an analytical function in all variables \( x_j \), approximations by polynomials may lead to small errors. In the case of a uniform degree \( r - 1 \), the subspaces \( V_j \subset W_j \) in (3.3) are \( \mathbb{P}_{r-1} \) and any \( v \in V = \bigotimes_{j=1}^d V_j \) has the Tucker rank \( r = (r, \ldots, r) \). The error, which is an upper bound of \( \sigma(w, \mathcal{M}_r) \), depends on the decay of the higher derivatives. The analysis in Section ?? will show exponential decay of \( \sigma(w, \mathcal{M}_r) \).

Multivariate functions depending on the Euclidean norm as, e.g., the classical potentials \( 1/|x| \), \( e^{-\lambda|x|}/|x| \), \( e^{-\lambda|x|} \) can be rather well approximated in \( \mathcal{M}_{R} \) leading to exponential decay of \( \sigma(w, \mathcal{M}_{R}) \) with respect to \( R \). For its computation and analysis see [12, 11, 23, 14, 17, 15, 4].

3.4 Application to grid functions

Galerkin discretisations with tensor product basis functions or finite difference schemes lead to grid points \( x_i = \left(x^{(1)}_i, \ldots, x^{(d)}_i\right) \), where \( i \in I = I_1 \times \ldots \times I_d \). Hence, the grid values \( u(x_i) = u_i \) belong to \( \mathbb{R}^I \) which is the tensor space
\[
\mathbb{R}^I = \mathbb{R}^{I_1} \otimes \mathbb{R}^{I_2} \otimes \ldots \otimes \mathbb{R}^{I_d},
\]
i.e., \( W_j = \mathbb{R}^{I_j} \) from (3.1) have the dimension \( n_j := \# I_j \). For simplicity we assume \( n_j = n \) for all \( 1 \leq j \leq d \).

The representation of \( w \in \mathcal{M}_R \) needs a storage of \( Rdn \), while \( w \in \mathcal{M}_r \) with \( r = (r, \ldots, r) \) requires \( rdn + rd \) data.

3.5 Application to matrices

The index sets \( I_1, \ldots, I_d \) and \( J_1, \ldots, J_d \) give rise to the two tensor spaces \( X := \mathbb{R}^{I_1} \otimes \ldots \otimes \mathbb{R}^{I_d} \) and \( Y := \mathbb{R}^{J_1} \otimes \ldots \otimes \mathbb{R}^{J_d} \). Given matrices \( A^{(j)} \in \mathbb{R}^{I_j \times J_j} \) (\( 1 \leq j \leq d \)), its Kronecker product \( \mathcal{A} := A^{(1)} \otimes \ldots \otimes A^{(d)} \) is defined as the mapping
\[
\mathcal{A} : X \to Y, \quad x = x^{(1)} \otimes \ldots \otimes x^{(d)} \mapsto \mathcal{A}x = A^{(1)} x^{(1)} \otimes \ldots \otimes A^{(d)} x^{(d)} \in Y.
\]

4 Regularity for elliptic eigenvalue problems

4.1 Polynomial approximation of analytic functions

To understand the separability property of eigenfunctions we analyse their regularity with respect to some classes of functions which allow the holomorphic extension to the complex plane.

The error estimates will be derived for the function set \( \mathcal{A}_{M, \rho}(I) \) and for its multidimensional counterpart \( \mathcal{A}_{M, \rho}(I^d) \), \( d \geq 2 \). The definition of the space \( \mathcal{A}_{M, \rho}(I^d) \) requires several
steps. For the interval $I := (-1, 1)$ and $\rho > 1$, Bernstein's regularity ellipse is given by (cf. [1])

$$E_\rho := \{ z \in \mathbb{C} : |z - 1| + |z + 1| \leq \rho + \rho^{-1} \}. $$

Its semi-axes are $a = \frac{\rho + \rho^{-1}}{2}$ and $b = \frac{\rho - \rho^{-1}}{2}$, implying $a + b = \rho$.

**Definition 4.1** Let $I = (-1, 1)$ and $M > 0$, $\rho > 1$ be given constants. $A_{M, \rho}(I)$ is the class of functions $f \in C^\infty(I)$ having a holomorphic extension to $E_\rho(I)$ such that

$$|f(z)| \leq M \quad \forall z \in E_\rho(I).$$

Next, we introduce the multidimensional analogue of $A_{M, \rho}(I)$ on the tensor domain $I^d := (-1, 1)^d$. Let $E_\rho^{(j)} := I \times \ldots \times I \times E_\rho \times I \times \ldots \times I$ with $E_\rho$ to be inserted at the $j$th position.

**Definition 4.2** For given constants $M > 0$, $\rho > 1$, the set $A_{M, \rho}(I^d)$ consists of all functions $f \in C^\infty(I^d)$ having holomorphic extensions to $E_\rho^{(j)}$, for all $1 \leq j \leq d$, and satisfying

$$\max_{1 \leq j \leq d} \{ \sup_{x \in E_\rho^{(j)}} |f(x)| \} \leq M.$$

The following remark recalls the well-known fact that controlling all higher derivatives of a function implies that it belongs to $A_{M, \rho}(I)$ (see e.g., [20] for the proof).

**Remark 4.3** Assume that a function $u : I \to \mathbb{R}$ satisfies for some $C_u, \gamma_u \geq 0$

$$\left\| \frac{\partial^n u}{\partial x^n} \right\|_{L^\infty(I)} \leq C_u \gamma_u^n n! \quad \text{for all } n \in \mathbb{N}_0. \quad (4.1)$$

Then $u \in A_{M, \rho}(I)$ holds with $\rho = 1 + \gamma_u^{-1} > 1$, $M = C \cdot C_u$.

For the continuous multivariate functions $f = f(x_1, \ldots, x_d) : \mathbb{R}^d \to \mathbb{R}$, we use the tensor product interpolant

$$I_N f = I_N^{x_1} \cdots I_N^{x_d} f \in P_N[I^d],$$

where $I_N^{x_i}$ ($1 \leq i \leq d$) denotes the interpolation polynomial of degree $N$ with respect to the variables $x_i \in I := [-1, 1]$ interpolating $f$ with respect to the variables $x_i \in I := [-1, 1]$ at the Chebyshev nodes.

**Proposition 4.4** Let $M > 0$ and $\rho > 1$ be given. For all $f \in A_{M, \rho}(I^d)$ and $N > 1$ the estimate

$$\| f - I_N f \|_{C^0(I^d)} \leq c M (\log N)^d \rho^{-N} \quad (4.2)$$

holds.

In the next section we derive the regularity results for solutions of elliptic eigenvalue problems which will imply Remark 4.3.
4.2 Regularity for elliptic eigenvalue problems with smooth data

Let \( \Omega \subset \mathbb{R}^d \) be a Lipschitz domain. For \( \ell \in \mathbb{N} \), we define the norms

\[
\| \nabla^\ell u \|_{L^2(\Omega)}^2 := \sum_{|\alpha|=\ell} \frac{\ell!}{\alpha!} \| D^\alpha u \|_{L^2(\Omega)}^2 \quad \text{and} \quad \| \nabla^\ell u \|_{L^\infty(\Omega)} := \left\| \sum_{|\alpha|=\ell} \frac{\ell!}{\alpha!} |D^\alpha u|^2 \right\|_{L^\infty(\Omega)}.
\]

For \( u \) replaced by vectors or matrices, the absolute value \(|\cdot|\) is to be replaced by the Euclidean or spectral norm.

Consider the eigenvalue problem for the differential operator

\[
\Lambda u := -\nabla \cdot (A \nabla u) + \langle b, \nabla u \rangle + cu,
\]

where \( \langle a, b \rangle := \sum_{i=1}^d a_i b_i \) (without complex conjugation). We say that \( \Lambda \) is uniformly elliptic if \( A \in L^{\text{sym}}(\Omega, \mathbb{R}^d) \), \( b \in L^\infty(\Omega, \mathbb{R}^d) \), and \( c \in L^\infty(\Omega) \) satisfy

\[
0 < a_{\min} := \inf_{x \in \Omega} \inf_{v \in \mathbb{C} \setminus \{0\}} \frac{\langle A(x) v, v \rangle}{\langle v, v \rangle} \leq \sup_{x \in \Omega} \sup_{v \in \mathbb{R} \setminus \{0\}} \frac{\langle A(x) v, v \rangle}{\langle v, v \rangle} =: a_{\max} < \infty
\]

\[
0 \leq -\frac{1}{2} \text{div } b + c.
\]

We assume that the coefficients of the operator \( \Lambda \) are analytic; i.e., there exist positive constants \( C_A, C_b, C_c \) and \( \gamma_A, \gamma_b, \gamma_c \) such that

\[
\| \nabla^p A \|_{L^\infty(\Omega)} \leq C_A \gamma_A^p \| p! \| \quad \forall p \in \mathbb{N}_0,
\]

\[
\| \nabla^p b \|_{L^\infty(\Omega)} \leq C_b \gamma_b^p \| p! \| \quad \forall p \in \mathbb{N}_0,
\]

\[
\| \nabla^p c \|_{L^\infty(\Omega)} \leq C_c \gamma_c^p \| p! \| \quad \forall p \in \mathbb{N}_0.
\]

We consider two types of domain: Either \( \Omega = \mathbb{R}^d \) or \( \Omega \) is a bounded Lipschitz domain with analytic boundary, i.e., there is a finite family \( \mathcal{U} \) of open subset in \( \mathbb{R}^d \) along a family of bijective maps \( \{ \chi_U : \overline{B_1} \to \overline{U} \} \). Let us consider the eigenvalue problem: Find a pair \( (\lambda, u) \in \mathbb{R} \times H_0^1(\Omega) \setminus \{0\} \) such that

\[
\Lambda u = \lambda u \quad \text{in } \Omega,
\]

\[
u = 0 \quad \text{on } \partial \Omega
\]

with \( \Lambda \) as in (4.3). Let \( E(\lambda) \) denote the eigenspace for the eigenvalue \( \lambda \).

\footnote{\( B_1 \) denotes the unit ball in \( \mathbb{R}^d \) and \( B_1^0 := \{ x \in B_1 \mid x_d = 0 \} \). For \( \sigma \in \{ +, - \} \), we set \( B_1^\sigma := \{ x \in B_1 \mid \sigma x_d > 0 \} \).}
Theorem 4.5 Let \( \Omega \) be an analytic, bounded Lipschitz domain which satisfies (4.6). Assume that the coefficients \( A, b, c \) satisfy (4.5). Then, any eigenfunction \( u \in E(\lambda) \) of (4.7) (normalised to \( \|u\|_{L^2(\Omega)} = 1 \)) is analytic. There exist constants \( C, K > 0 \) depending only on the constants in (4.5), (4.6), and on \( a_{\min} \) and the spatial dimension \( d \) such that

\[
\| \nabla^{p+2} u \|_{L^2(\Omega)} \leq CK^{p+2} \max \left\{ p, \sqrt{\lambda} \right\}^{p+2} \quad \text{for all } p \geq 0. \tag{4.8}
\]

Proof.

<table>
<thead>
<tr>
<th>( p )</th>
<th>Time/it.</th>
<th>( \delta_{\lambda_1} )</th>
<th>( \delta_{u_1} )</th>
<th>( \delta_{\lambda_2} )</th>
<th>( \delta_{u_2} )</th>
<th>( \delta_{\lambda_3} )</th>
<th>( \delta_{u_3} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>0.02</td>
<td>2.3 \cdot 10^{-6}</td>
<td>2.5 \cdot 10^{-4}</td>
<td>2.7 \cdot 10^{-5}</td>
<td>1.9 \cdot 10^{-3}</td>
<td>4.2 \cdot 10^{-6}</td>
<td>3.3 \cdot 10^{-3}</td>
</tr>
<tr>
<td>8</td>
<td>0.03</td>
<td>2.1 \cdot 10^{-7}</td>
<td>8.9 \cdot 10^{-5}</td>
<td>1.3 \cdot 10^{-6}</td>
<td>3.9 \cdot 10^{-4}</td>
<td>5.3 \cdot 10^{-7}</td>
<td>1.1 \cdot 10^{-3}</td>
</tr>
<tr>
<td>10</td>
<td>0.1</td>
<td>1.7 \cdot 10^{-8}</td>
<td>2.2 \cdot 10^{-6}</td>
<td>6.9 \cdot 10^{-8}</td>
<td>9.4 \cdot 10^{-5}</td>
<td>1.3 \cdot 10^{-8}</td>
<td>1.8 \cdot 10^{-4}</td>
</tr>
<tr>
<td>12</td>
<td>0.39</td>
<td>9.9 \cdot 10^{-10}</td>
<td>5.2 \cdot 10^{-8}</td>
<td>4.5 \cdot 10^{-9}</td>
<td>2.3 \cdot 10^{-5}</td>
<td>1.9 \cdot 10^{-9}</td>
<td>7.0 \cdot 10^{-5}</td>
</tr>
<tr>
<td>14</td>
<td>1.6</td>
<td>6.5 \cdot 10^{-11}</td>
<td>1.3 \cdot 10^{-9}</td>
<td>2.8 \cdot 10^{-10}</td>
<td>6.0 \cdot 10^{-7}</td>
<td>1.5 \cdot 10^{-11}</td>
<td>1.7 \cdot 10^{-8}</td>
</tr>
<tr>
<td>16</td>
<td>6.9</td>
<td>3.8 \cdot 10^{-12}</td>
<td>3.2 \cdot 10^{-9}</td>
<td>1.8 \cdot 10^{-11}</td>
<td>1.0 \cdot 10^{-7}</td>
<td>4.3 \cdot 10^{-11}</td>
<td>4.5 \cdot 10^{-6}</td>
</tr>
</tbody>
</table>

Table 4.1: Several smallest eigenvalues for 2D Laplacian.

Again, these results indicate the linear scaling of the tensor method in \( n \). The cost of one power iteration step increases like \( O(n \log n) \) as expected. Furthermore, we expect the asymptotic behaviour

\[
\delta_{\lambda_1} = O(\lambda h^2) \quad \text{and} \quad \delta_u = O(\sqrt{\lambda} h), \quad \text{as} \quad h \to 0,
\]

which are in agreement with the above presented calculations. Notice that in our case the true scaling factor between the refined grids is 16 for the eigenvalues and 4 for the eigenfunctions.

Example 3. We apply the tensor method to compute the minimal eigenvalue of the 3D Laplacian on large \( n \times n \times n \) grids with \( n = 2^p - 1 \). In this case, the action of the truncation operator \( T_1 \) is equivalent to the rank-1 Tucker approximation of the \( n \times n \times n \) rank-\( R \) tensors arising at each iterative step.

<table>
<thead>
<tr>
<th>( p )</th>
<th>Time/it.</th>
<th>( \delta_{\lambda} )</th>
<th>( \delta_u )</th>
<th>it.</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>0.03</td>
<td>2.0 \cdot 10^{-4}</td>
<td>1.5 \cdot 10^{-3}</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>0.05</td>
<td>1.6 \cdot 10^{-5}</td>
<td>9.4 \cdot 10^{-4}</td>
<td>4</td>
</tr>
<tr>
<td>10</td>
<td>0.12</td>
<td>7.8 \cdot 10^{-7}</td>
<td>1.2 \cdot 10^{-4}</td>
<td>5</td>
</tr>
<tr>
<td>12</td>
<td>0.51</td>
<td>4.9 \cdot 10^{-8}</td>
<td>3.4 \cdot 10^{-5}</td>
<td>5</td>
</tr>
<tr>
<td>14</td>
<td>2.2</td>
<td>3.1 \cdot 10^{-9}</td>
<td>9.3 \cdot 10^{-6}</td>
<td>5</td>
</tr>
<tr>
<td>16</td>
<td>10.6</td>
<td>1.9 \cdot 10^{-10}</td>
<td>2.8 \cdot 10^{-6}</td>
<td>5</td>
</tr>
<tr>
<td>17</td>
<td>22.3</td>
<td>4.8 \cdot 10^{-11}</td>
<td>1.6 \cdot 10^{-6}</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 4.2: Minimal eigenvalue for 3D Laplacian on large spatial grids.
We observe the asymptotic complexity $O(dn \log n)$ to achieve the theoretical error bounds for both the eigenvalues and eigenfunctions. Notice that the problem size on the finest grid (with $n = 2^{17}$, $d = 3$) exceeds $N = n^d = 2^{17d} \approx 10^{15}$, which is far beyond the facilities of modern super-computers. Hence tensor methods are mandatory for solving large scale multi-dimensional spectral problems.

### 4.2.1 Operator with variable coefficients

In this section we consider eigenvalue problems for the elliptic operators of the form

$$\Lambda u = \Delta u + c(x)u, \quad x \in \mathbb{R}^d$$

with smooth coefficient $c(x)$.

In the subsequent examples we consider discretisations represented by matrices of the form

$$A = M + qC, \quad q > 0,$$  \hspace{1cm} (4.9)

where $M > 0$ is the finite difference negative Laplacian in 2D as before, and $C$ is the low Kronecker rank matrix discretising the zero-order term $c(x) = \sum_{k=1}^{R} c_k^{(1)}(x_1) \cdot c_k^{(2)}(x_2) \geq 0$ as in (??). We suppose that

$$0 \leq \langle Cx, x \rangle \leq \langle x, x \rangle \quad \forall \ x \in V_n \setminus \{0\}.$$  

**Example 4.** In this example we apply the Lanczos algorithm for computing the largest eigenvalues as described in §2. Consider matrices of the form

$$A = M - \sum_{t=1}^{R} D_t \otimes D_t,$$

where $M$ is the negative discrete Laplacian and $D_t$ are diagonal matrices with positive entries. We approximate the maximal eigenvalue by the standard Lanczos and truncated tensor Lanczos methods for the following two examples:

(A) the entries of $D_t$ are grid values of the function $(1 + T_t(x))/10$, where $T_t$ is the Chebyshev polynomial of degree $t$;

(B) the entries of $D_t$ are random values uniformly distributed in $[0, 1]$.

We compare the results obtained after 50 iterations for both methods. The matrix size is $N = 300^2$, the truncation rank and the accuracy are set to 10 and $10^{-2}$, respectively.

We observe that the computed eigenvalues in both methods are close to the truncation accuracy even for the random case. This suggests that tensor tools may be applied to a much broader class of matrices than required by the theory described above.

**Example 5.** The minimal eigenvalue is calculated for the problem (4.9) with the matrix $C$ corresponding to the rank-1 potential $c(x) = \sin(\lambda x_1) \sin(\lambda x_2)$ in $(0, 1)^2$. We perform the truncated power iteration with the matrix $L = A^{-1}$, such that the respective matrix-vector
multiplication \( y = \mathcal{L}U \) is implemented by the “truncated” iterative procedure, \( y^p \to y \), as \( p \to \infty \), where
\[
y^{p+1} = y^p - \omega \mathcal{L}_R(Ay^p - U), \quad p = 0, 1, \ldots
\]
with certain rank-\( r \) initial guess \( y^0 \) and with a proper relaxation parameter \( \omega \in (0, 2) \). Here \( \mathcal{L}_R \) is the rank-\( R \) approximation to the inverse of the shifted Laplacian described in (??).

Table 6.6 presents the results on the sequence of grids \( n = 2^p, p = 10, 11, 12, 13 \), for the truncation rank \( r = 3 \). We give the total CPU time (sec.), the number of power iterations and the scaling factor between the neighbouring grids. We expect an \( O(n \log n) \) scaling provided that all nested iterations require the same number of loops and the same tensor rank for the preconditioner (of course, there are some fluctuations).

<table>
<thead>
<tr>
<th>( n )</th>
<th>Time</th>
<th>( \delta_{\text{Res}} )</th>
<th>Power iter.</th>
<th>Scaling</th>
</tr>
</thead>
<tbody>
<tr>
<td>1024</td>
<td>7.3</td>
<td>( 3.0 \cdot 10^{-4} )</td>
<td>10</td>
<td>-</td>
</tr>
<tr>
<td>2048</td>
<td>23.6</td>
<td>( 1.5 \cdot 10^{-4} )</td>
<td>14</td>
<td>2.3</td>
</tr>
<tr>
<td>4096</td>
<td>63.9</td>
<td>( 7.6 \cdot 10^{-5} )</td>
<td>14</td>
<td>2.7</td>
</tr>
<tr>
<td>8192</td>
<td>209.</td>
<td>( 3.8 \cdot 10^{-5} )</td>
<td>17</td>
<td>2.7</td>
</tr>
</tbody>
</table>

Table 4.5: Minimal eigenvalue for \(-\Delta + c(x)\) in 2D.

This table indicates the linear-logarithmic scaling in \( n \) as well as the robust convergence of the power iteration with the tensor modification.

### 4.3 Concluding remarks

The theoretical and numerical analysis of multi-dimensional eigenvalue problems presented in the paper clearly indicate that tensor structured methods for the approximation and solution of “smooth” spectral problems in \( \mathbb{R}^d \) yield a promising basis for efficient solution methods in the modern high dimensional applications. Moreover, it seem that these methods are not restricted to smooth problems.

**Acknowledgements.** The authors are thankful to Dipl. Ing. Cristóbal Bertoglio for assistance with the numerical computations of Tables 6.1, 6.2 and Example 5.
References


