Fluctuationlessness theorem to approximate univariate functions' matrix representations

METİN DEMİRALP
İTU Informatics Institute
Department of Computational Sci. and Eng.
İTU Ayazağa Campus, 34469 İstanbul
TURKEY (TÜRKİYE)
demiralp@be.itu.edu.tr

Abstract: Matrix representation of functions are required to convert an operator related problem to its algebraic counterpart over certain vectors and matrices. The problems involving operators which are purely or partially algebraic are most frequently encountered ones in applications. The algebraic operator here has a Hilbert space domain defined over square integrable univariate functions on a specified interval and its action on its argument is just multiplication by a function. We focus on univariate functions for simplicity in this very first step although the generalization to multivariance seems to be rather straightforward. The main purpose of this work is to introduce a conjecture to facilitate the numerical approximation of the matrix representation of the above algebraic operator and then to prove it to get an important theorem which seems to be capable of opening new very efficient horizons in numerical analysis and in its applications. Theorem states that the matrix representation of a univariate function is the image of the matrix representation of the independent variable under the same function for a finite Hilbert space. Illustrative numerical implementations are also given.

Key–Words: Matrix Representation, Fluctuation Operator, Hilbert Spaces, Projection Operators, Algebraic Multiplication Operators

1 Introduction

Matrix representations of the operators play important roles in the development of algebraic methods for the numerical solution of various problems. Especially, linear operator involving problems can be treated by using those operators' matrix representations on a linear vector space spanned by a finite number of basis functions to get approximate solutions and it is expected that the approximation's quality will be increased as the dimension of the linear vector space increases. Amongst these types of problems, quantum mechanical or non-equilibrium statistical mechanical ones are linear in nature unless an unknown external influence whose structure is to be determined together with the other unknowns of the system (like the external field as a controlling agent in quantum optimal control problems) is added to the problem [1–11]. In these areas, the algebraic operators multiplying their arguments by certain functions' values are frequently encountered. For example, in Schrödinger's equation potential term or external field's scalar component are in this category. The Liuoiville equation [12] also involves this kind of operators. The matrix representations of these operators have the elements defined through certain Hilbert spaces' inner products which are defined via integrals whose evaluations may not be easy in all encountered cases and necessitate approximations. Hence the approximation of the matrix representations of these operators are quite important not only for the theory but also for the practical applications.

Expected values of certain functions are also important to get insight for certain probabilistic events and they can be evaluated through quadratic forms or Rayleigh quotients constructed over these functions' matrix representations over a finite Hilbert space.

The matrix representation of a function, in its abstract definition, does not impose any limitation on the multivariance as long as appropriate Hilbert spaces with appropriate inner products are in use. Hence, what we are going to develop here does not conceptually differ from univariance to multivariance. However, for explicit representation reasons, it is better to confine ourselves on the univariate functions, although the multivariate functions, which may bring certain level complications in formulae, can be treated by using same tools as well.

Expected values of the operators are derived from their matrix representations over basis sets appropriately defined for the Hilbert space of the problem under consideration. The exact definition requires a complete basis set which has generally denumerably infinite number of elements. However, actual applications can handle only finite dimensional sets, and hence, the matrix representation is approximated by its counterpart defined on an appropriately chosen finite subspace of the Hilbert space under consideration. As the dimension of this finite subspace is increased the quality of the matrix representation is also increased and after certain specific dimension one can expect that the approximate representation becomes same as the exact representation within a prescribed numerical accuracy. The choice of the finite subspace is very important to get a rapid convergence and in fact it is a matter of expertise on the structuring of the basis set elements. Especially, the creation or imitation of the anticipated behavior of the unknowns in the Hilbert space under consideration, at certain specific values of the independent variables (like singular points) in the basis set elements, is desired. Beyond this fact, all basis sets give slow or fast convergence in the matrix representation sequences constructed through finite subspaces of the Hilbert space under consideration.

Now, to be more specific, we can consider the Hilbert space $\mathcal{H}$ of continuos (continuity is imposed here for simplicity, by paying sufficient care integrable discontinuities can be added into analysis) and therefore square integrable functions over the interval $[\,a,b\,]$ and denote its basis set, which is orthonormal with respect to the inner product defined through an integration weighed under the nonnegative function $w(x)$ over the interval $[\,a,b\,]$, by

$$\mathcal{U} \equiv \{u_i(x)\}_{i=1}^{\infty}, \qquad x \in [\,a,b\,] \tag{1}$$

The orthonormality means

$$(u_i, u_j) \equiv \int_a^b dx w(x) u_i(x) u_j(x) = \delta_{ij}, \qquad 1 \le i,j < \infty \tag{2}$$

where $w(x)$ may vanish at most a finite number of times in the interval and the x dependence of the basis functions is not explicitly shown in the notation of the inner product at the left hand side since x plays the role of the dummy integration (and therefore indexing) variable, and, δ_{ij} stands for the Kroenecker's delta symbol which produces 1 when $i = j$ otherwise 0.

Now, we can define the subspace $\mathcal{H}_n$ spanned by the following finite subset

$$\mathcal{U}_n \equiv \{u_i(x)\}_{i=1}^{n}, \quad x \in [\,a,b\,], \quad 1 \le n < \infty \tag{3}$$

which means that an arbitrary function $g(x)$ in this subspace can be uniquely defined as the following linear combination

$$g(x) \equiv \sum_{i=1}^{n} g_i u_i(x), \qquad x \in [\,a,b\,] \tag{4}$$

where g_is symbolize the constants to specify the function $g(x)$. The expected value of an arbitrary linear operator $\mathcal{L}$ which maps $\mathcal{H}_n$ onto itself, with respect to the function $g(x)$, is given through the following equality

$$E_g(\mathcal{L}) \equiv \frac{(g, \mathcal{L}g)}{(g,g)} = \frac{\mathbf{g}_n^T \mathbf{L}_n \mathbf{g}_n}{\mathbf{g}_n^T \mathbf{g}_n}, \qquad \mathcal{L} : \mathcal{H}_n \to \mathcal{H}_n \tag{5}$$

where

$$\mathbf{g}_n^T \equiv [\, g_1 \dots g_n \,] \tag{6}$$

and $\mathbf{L}_n$ stands for an $n \times n$ matrix whose general term is defined as

$$L_n^{(i,j)} \equiv \mathbf{e}_i^T \mathbf{L}_n \mathbf{e}_j \equiv (u_i, \mathcal{L}u_j), \quad 1 \le i,j \le n \tag{7}$$

and it is the matrix representation of the operator $\mathcal{L}$ on $\mathcal{H}_n$. In this formula $\mathbf{e}_i$ represents the i-th standard unit vector, in n dimensional cartesian space, whose only nonzero element is located at the i-th position and equals to 1.

The matrix $\mathbf{L}_n$ depends on the nature of the operator $\mathcal{L}$ and this dependence prevents universality. Our purpose is to develop a representation of this matrix in terms of rather universal matrices like the matrix representation of the independent variable or certain derived entities through rather simple relations. Before attempting to do so we focus our attention on a specific class of operators, algebraic operators whose actions on their operands is just multiplication by a function which may be multivariate although we prefer to deal with the univariate ones for simplicity. The extension what we are going to obtain for the univariate case to the multivariate cases seems to be conceptually straightforward, however, there will be of course a certain level of mathematical complication in the formulae. This extension is out of the scope of this paper.

The paper is organized as follows. The second section focuses on a theorem for the approximation of the matrix representation of a product of algebraic operators whose actions on their arguments are just the multiplications by functions. This is what we need to prove the fluctuationlessness theorem on the approximation of matrix representation of the multiplying–by–a–function type algebraic operators. The definitions of the fluctuation operator and the first order fluctuation matrix of a function which are necessary for the proof of this theorem are also given in this section. The third section focuses on the matrix representations of the natural number powers of the independent variable. The definitions of universal fluctuation matrices are also given in this section. The purpose of this section together with the previous one is to prepare the necessary background to the fluctuationlessness theorem for the matrix representation of an algebraic operator which multiplies its operand by a function's value. The fourth section involves the explanation and the proof of the fluctuationlessness theorem. The fifth section covers certain illustrative applications. The sixth section including concluding remarks finalizes the paper.

2 Approximate factorization for the matrix representation to a binary product of function operators

Let us consider the operator $\widehat{x}$ whose domain is $\mathcal{H}$ (defined in the previous section) and its action on a function $g(x)$ from this Hilbert space is defined as

$$\widehat{x}g(x) = xg(x), \quad x \in [\,a, b,\,], \quad g(x) \in \mathcal{H} \tag{8}$$

If we consider the subspace $\mathcal{H}_n$ which was introduced in the previous section and define the projection operators

$$P_j g(x) \equiv u_j(x)\,(u_j, g)\,, \;\; 1 \le j \le n, \;\; g(x) \in \mathcal{H} \tag{9}$$

where P_j projects to the one dimensional subspace spanned by $u_j(x)$ in $\mathcal{H}_n$ and the multidimensional projection operator

$$P^{(n)} \equiv \sum_{j=1}^{n} P_j \tag{10}$$

which maps to $\mathcal{H}_n$ then we can write the following identity where $\mathcal{I}$ represents the unit operator on $\mathcal{H}$

$$\mathcal{I} \equiv P^{(n)} + \left[\mathcal{I} - P^{(n)}\right] \tag{11}$$

The interpretation of the operator $\left[\,\mathcal{I} - P^{(n)}\,\right]$ may require a little bit more details. To this end, we can write

$$g(x) \equiv P^{(n)}g(x) + \left[\mathcal{I} - P^{(n)}\right] g(x), \;\; g(x) \in \mathcal{H} \tag{12}$$

where $P^{(n)}g(x)$ belongs to $\mathcal{H}_n$ while $\left[\mathcal{I} - P^{(n)}\right] g(x)$ lies in its complementary space. Since $P^{(n)}$, as being a projection operator, is idempotent; one can easily show that

$$\begin{aligned} P^{(n)}\left[\mathcal{I} - P^{(n)}\right] &= P^{(n)} - {P^{(n)}}^2 \\ &= P^{(n)} - P^{(n)} = \widehat{0} \end{aligned} \tag{13}$$

is a valid equality where $\widehat{0}$ stands for the zero operator on $\mathcal{H}$. (13) guarantees that $P^{(n)}g(x)$ is orthogonal to $\left[\,\mathcal{I} - P^{(n)}\,\right] g(x)$ for any function in $\mathcal{H}$. Therefore (12) corresponds to an orthogonal decomposition. This is quite natural because the basis set of $\mathcal{H}_n$ and its complementary space are orthogonal. Now, one can write

$$\left[\mathcal{I} - P^{(n)}\right] g(x) = \sum_{i=n+1}^{\infty} g_i u_i(x), \;\; g(x) \in \mathcal{H} \tag{14}$$

which has infinite number of terms oscillating around zero because of orthogonalities of the basis functions. These infinite number of oscillations result in somehow arbitrary increases and decreases from zero and we may call them fluctuations to use the terminology of probability related sciences like quantum mechanics, statistical mechanics. By referring to this terminology we will call $\left[\,\mathcal{I} - P^{(n)}\,\right]$ "Fluctuation Operator".

(14) implies that the real numbers defined by $\left\|\left[\,\mathcal{I} - P^{(n)}\,\right]g\right\|$ $(n = 1, 2, ...)$ values form a nonincreasing sequence whose limit is zero. That is,

$$\lim_{n\to\infty} \left\|\left[\mathcal{I} - P^{(n)}\right] g\right\| = 0 \qquad g(x) \in \mathcal{H} \tag{15}$$

This means that the image of $g(x)$ under the operator $\left[\,\mathcal{I} - P^{(n)}\,\right]$ must diminish as n grows unboundedly.

Now we can define the following operators to proceed

$$\begin{aligned} \mathcal{L}_{f_1}g(x) &\equiv f_1\left(\widehat{x}\right)g(x) \equiv f_1\left(x\right)g(x), \\ &\quad g(x) \in \mathcal{H} \end{aligned} \tag{16}$$

$$\begin{aligned} \mathcal{L}_{f_2}g(x) &\equiv f_2\left(\widehat{x}\right)g(x) \equiv f_2\left(x\right)g(x), \\ &\quad g(x) \in \mathcal{H} \end{aligned} \tag{17}$$

from where the following product operator is defined

$$\mathcal{L}_{f_1 f_2}g(x) \;\equiv\; \mathcal{L}_{f_1}\mathcal{L}_{f_2}g(x), \quad g(x) \in \mathcal{H} \tag{18}$$

where the functions $f_1(x)$ and $f_2(x)$ are assumed to be analytic on the interval $[\,a, b\,]$ although we do not need the actual analiticity in the proof of the theorem of this section. We will need it in the proof of the fluctuationlessness theorem.

The general terms of the matrix representations of the operators in (16) and (17) over $\mathcal{H}_n$ can be written through the following equalities

$$\mathbf{e}_i^T \mathbf{L}_{f_1}^{(n)} \mathbf{e}_j \equiv \left(u_i, f_1\left(\widehat{x}\right) u_j\right), \qquad 1 \le i,j \le n \quad (19)$$

$$\mathbf{e}_i^T \mathbf{L}_{f_2}^{(n)} \mathbf{e}_j \equiv \left(u_i, f_2\left(\widehat{x}\right) u_j\right), \qquad 1 \le i,j \le n \quad (20)$$

and the following one for the product operator in (18)

$$\mathbf{e}_i^T \mathbf{L}_{f_1 f_2}^{(n)} \mathbf{e}_j \equiv \left(u_i, f_1\left(\widehat{x}\right) f_2\left(\widehat{x}\right) u_j\right), \quad 1 \le i,j \le n \quad (21)$$

Now by using the identity given in (11) here we can get the following equality

$$\left(u_i, f_1\left(\widehat{x}\right) f_2\left(\widehat{x}\right) u_j\right) \equiv \left(u_i, f_1\left(\widehat{x}\right) P^{(n)} f_2\left(\widehat{x}\right) u_j\right) + \left(u_i, f_1\left(\widehat{x}\right)\left[\mathcal{I} - P^{(n)}\right] f_2\left(\widehat{x}\right) u_j\right), \quad 1 \le i,j \le n \quad (22)$$

Since the fluctuation operator realizes a projection it is identical to its square. On the other hand, both the fluctuation operator and the algebraic operator $f_1\left(\widehat{x}\right)$ are hermitian. These permit us to write

$$\left(u_i, f_1\left(\widehat{x}\right)\left[\mathcal{I} - P^{(n)}\right] f_2\left(\widehat{x}\right) u_j\right) = \left(\left[\mathcal{I} - P^{(n)}\right] f_1\left(\widehat{x}\right) u_i, \left[\mathcal{I} - P^{(n)}\right] f_2\left(\widehat{x}\right) u_j\right), \quad 1 \le i,j \le n \quad (23)$$

which means

$$\left|\left(u_i, f_1\left(\widehat{x}\right)\left[\mathcal{I} - P^{(n)}\right] f_2\left(\widehat{x}\right) u_j\right)\right| \le \left\|\left[\mathcal{I} - P^{(n)}\right] f_1\left(\widehat{x}\right) u_i\right\| \left\|\left[\mathcal{I} - P^{(n)}\right] f_2\left(\widehat{x}\right) u_j\right\|, \quad 1 \le i,j \le n \quad (24)$$

because of the Cauchy–Schwarz inequalities. If we define the following operators $\mathcal{F}_1$ and $\mathcal{F}_2$ which are in fact the reduced–by–idempotency forms of the operators at the rightmost norm terms of above equation, and, will be called "First Order Fluctuation Operators"of the functions $f_1(x)$ and $f_2(x)$ because the fluctuation operator appears only once in the structures of these operators

$$\mathcal{F}_1 \equiv f_1\left(\widehat{x}\right)\left[\mathcal{I} - P^{(n)}\right] f_1\left(\widehat{x}\right), \quad (25)$$

$$\mathcal{F}_2 \equiv f_2\left(\widehat{x}\right)\left[\mathcal{I} - P^{(n)}\right] f_2\left(\widehat{x}\right) \quad (26)$$

then we can symbolize their matrix representations on $\mathcal{H}_n$ by $\mathbf{F}_1^{(n)}$ and $\mathbf{F}_2^{(n)}$ respectively. This enables us to rewrite (24) as follows

$$\left|\left(u_i, f_1\left(\widehat{x}\right)\left[\mathcal{I} - P^{(n)}\right] f_2\left(\widehat{x}\right) u_j\right)\right| \le \left(\mathbf{e}_i^T \mathbf{F}_1^{(n)} \mathbf{e}_i\right)^{\frac{1}{2}} \left(\mathbf{e}_j^T \mathbf{F}_2^{(n)} \mathbf{e}_j\right)^{\frac{1}{2}}, \quad 1 \le i,j \le n \quad (27)$$

Function based fluctuation operators $\mathcal{F}_1$ and $\mathcal{F}_2$ somehow measure the smoothness of the functions $f_1(x)$ and $f_2(x)$ and more fluctuation means less smoothness and vice versa.

Since the norms in the right hand side of (24) tend to vanish when n grows up to infinity, as long as $\mathcal{U}$ is complete; the fluctuation operator containing term at the right hand side of (22) becomes negligible beside the fluctuation operator free term. Hence we can write

$$\left(u_i, f_1\left(\widehat{x}\right) f_2\left(\widehat{x}\right) u_j\right) \approx \left(u_i, f_1\left(\widehat{x}\right) P^{(n)} f_2\left(\widehat{x}\right) u_j\right) = \sum_{k=0}^{n} \left(u_i, f_1\left(\widehat{x}\right) u_k\right)\left(u_k, f_2\left(\widehat{x}\right) u_j\right), \quad 1 \le i,j \le n \quad (28)$$

We call this approximate equality "Fluctuationlessness Approximation". In matrix representation terminology, this approximation can be written by using the matrices given in (19), (20), and (21) as follows

$$\mathbf{L}_{f_1 f_2}^{(n)} \approx \mathbf{L}_{f_1}^{(n)} \mathbf{L}_{f_2}^{(n)}, \quad 1 \le n < \infty \quad (29)$$

This is the mathematical statement of the following theorem:

Theorem 1 *The matrix representation of a binary product of function type operators is the product of the individual matrix representations of those operators with the same ordering of their operator counterparts, at the fluctuationless limit.*

which will be called "Fluctuationlessness Theorem for a Binary Product".

3 Matrix representations for natural number powers of the independent variable

We can use Theorem 1 to approximate the matrix representations of the natural number powers of the operator $\widehat{x}$. If we denote the matrix representation of $\widehat{x}^k$ on the subspace $\mathcal{H}_n$ by $\mathbf{X}_k^{(n)}$ then we can write

$$\mathbf{e}_i^T \mathbf{X}_k^{(n)} \mathbf{e}_j \equiv \left(u_i, \widehat{x}^k u_j\right), \quad 1 \le i,j \le n, \quad k = 0,1,2,... \quad (30)$$

We can write the following approximation formula for $\mathbf{X}_2^{(n)}$ by considering the operator $\widehat{x}^2$ as the product of $\widehat{x}$ with itself.

$$\mathbf{X}_2^{(n)} \approx \mathbf{X}^2, \qquad \mathbf{X} \equiv \mathbf{X}_1^{(n)} \quad (31)$$

Now we can proceed one step ahead and consider the operator $\widehat{x}^3$ as the product $\widehat{x}^2\widehat{x}$. This allows us to write the following approximations in accordance with Theorem 1 and (31).

$$\mathbf{X}_3^{(n)} \approx \mathbf{X}_2^{(n)}\mathbf{X}_1^{(n)} \approx \mathbf{X}^3 \tag{32}$$

Being inspired by (31) and (32) we can propose the following general approximation formula

$$\mathbf{X}_k^{(n)} \approx \mathbf{X}^k, \quad k = 0, 1, 2, ... \tag{33}$$

which matches (31) and (32) for $k = 2$ and $k = 3$ respectively. It becomes exact when $k = 0$ and $k = 1$ as can be easily proven. Therefore all conditions for a mathematical induction are provided. By assuming that (33) holds for certain k value we can write

$$\mathbf{X}_{k+1}^{(n)} \approx \mathbf{X}_k^{(n)}\mathbf{X} \approx \mathbf{X}^{k+1}, \quad k = 0, 1, 2, ... \tag{34}$$

This completes the proof of the following theorem

Theorem 2 *The matrix representation of a natural number power of the operator $\widehat{x}$ is the same natural number power of the matrix representation of $\widehat{x}$, at the fluctuationless limit.*

which will be called “Fluctuationlessness Approximation for Natural Number Power of the Independent Variable”.

The smoothness of the function x^k decreases as the natural number k increases. This implies that the error coming from the ignorance of the fluctuations grows parallel to the increase in k.

4 Fluctuationlessness Theorem for algebraic operators which multiply their operands by a function

Consider the following algebraic operator whose action on its operand from $\mathcal{H}$ is the multiplication by a given function $f(x)$ which is assumed to be analytic on the interval $[\,a, b\,]$

$$\mathcal{L}_f g(x) \equiv f\left(\widehat{x}\right) g(x) \equiv f(x)g(x), \qquad g(x) \in \mathcal{H}_n \tag{35}$$

The analiticity of $f(x)$ on the interval $[\,a, b\,]$ guarantees the convergence of the following power series expanded at $x = (a + b)/2$ for entire interval $[\,a, b\,]$

$$f(x) = \sum_{k=0}^{\infty} f_k \left(x - \frac{a+b}{2}\right)^k \tag{36}$$

where f_k is equivalent to the quotient of the function’s k–th order derivative’s value at $x = \frac{a+b}{2}$ by $k!$. This power series expansion implies

$$f\left(\widehat{x}\right) = \sum_{k=0}^{\infty} f_k \left(\widehat{x} - \frac{a+b}{2}\widehat{I}\right)^k \tag{37}$$

Since it is a linear procedure to take the matrix representation, the matrix representation of the operator will be an infinite sum over the matrix representation of the summand of (37). This means that we can write

$$\mathbf{F}^{(n)} = \sum_{k=0}^{\infty} f_k \sum_{\ell=0}^{k} \binom{k}{\ell} \left(-\frac{a+b}{2}\right)^{\ell} \mathbf{X}_{k-\ell}^{(n)} \tag{38}$$

where we can use fluctuationless theorem on the natural number powers of the independent variable, that is,

$$\mathbf{X}_{k-\ell}^{(n)} \approx \mathbf{X}^{k-\ell}, \;\; k = 0, 1, 2..., \;\; \ell = 0, 1, ..., k \tag{39}$$

which enables us to get an analytical approximate structure for the finite sum in (38)

$$\sum_{\ell=0}^{k} \binom{k}{\ell} \left(-\frac{a+b}{2}\right)^{\ell} \mathbf{X}_{k-\ell}^{(n)} \approx \left(\mathbf{X} - \frac{a+b}{2}\mathbf{I}^{(n)}\right)^k, \qquad k = 0, 1, 2, ... \tag{40}$$

where we have used the fact that the zeroth power of $\mathbf{X}$ is $n \times n$ unit matrix $\mathbf{I}^{(n)}$ which is also equivalent to the rightmost matrix in the above equation when k meets 0.

We can now focus on the spectrum of $n \times n$ matrix $\mathbf{X}$. Its spectrum composes of real values because of its symmetry. On the other hand, a quadratic form constructed over $\mathbf{X}$ is equivalent to the integral of the product of a function, which is positive since it is the square of a linear combination of the basis functions, by the independent variable x over the interval $[\,a, b,]$. Hence a Rayleigh quotient constructed over $\mathbf{X}$ is bounded by b and a respectively from above and below. This means that that the spectrum of $\mathbf{X}$ is located in the interval $[\,a, b\,]$ as we expect. This fact guarantees that the power series over the rightmost power term in (40) with linear combination coefficient f_k converge. This, however, implies that

$$\mathbf{F}^{(n)} \approx \sum_{k=0}^{\infty} f_k \left(\mathbf{X} - \frac{a+b}{2}\mathbf{I}^{(n)}\right)^k \equiv f\left(\mathbf{X}\right) \tag{41}$$

because of the analyticity of $f(x)$ over the entire interval $[\,a, b\,]$. This result can be stated in the following theorem.

Theorem 3 *The matrix representation of an algebraic multiplication–by–a–univariate–function type operator, whose function is analytic on the interval* $[a, b]$, *over* $\mathcal{H}_n$ *is the image of the matrix representation of the independent variable over* $\mathcal{H}_n$ *under the function of the considered algebraic operator, at the fluctuationlessness limit.*

We call this "Fluctuationlessness Theorem for Functions". This is a very important issue because it brings the universality to the evaluation of the matrix representations at least approximately. This is because of (41) where the matrix representation needs the evaluation of $\mathbf{X}$ first and this procedure does not depend on the structure of $f(x)$, hence it is universal. There are many ways to evaluate the matrix $f(\mathbf{X})$. However, the best one seems to be using the spectral method where the spectral decomposition of the matrix $\mathbf{X}$ is constructed first and then the eigenvalues in the formula changed by the images of the eigenvalues under f. This enables us to deal with the solution of the eigenvalue problem of the matrix $\mathbf{X}$ first. It is of course universal, that is, does not depend on f. The dependence on f appears only when we construct the images of the eigenvalues under f, and obviously, this is rather simple operation since it does not require the solution of any set of equations.

Last discussions imply that $f(\mathbf{X})$ can be expressed as a linear combination of the f values evaluated at the eigenvalues of $\mathbf{X}$ with matrix coefficients each of which is the projection matrix projecting to the eigenspace spanned by the eigenvector corresponding to $\mathbf{X}$'s eigenvalue used as the f's evaluation point. This means that fluctuation free matrix representation stated in the fluctuationless theorem above has a quadrature [13–24] like structure.

5 Illustrative numerical implementations

In this section we are going to apply fluctuationlessness theorem (Theorem 3) on three different functions chosen from three essential categories to discuss the efficiency of the theorem on different functional characters. Our purpose is to compare the actual matrix representation of a function $f(x)$, $\mathbf{F}^{(n)}$, with its fluctuationless approximation, $f(\mathbf{X})$. We could do this elementwise but it would leave us with presentation of many values, plots, the number of which would increase as n grows, that is, it would be too much comprehensive for the presentation in a scientific paper. Instead, we could compare one or two, but just a few, global features like the norms of these two entities, to get just a few real valued properties for each n value. We are going to use two basic functionals: (1) Increase in Diagonal Square Dominancy and (2) Relative Error. The first one of these functionals is based on the increase in the following entity we call "Diagonal Square Dominancy Measurer"

$$\mu_{\text{dsd}}\left(\mathbf{F}^{(n)}\right) \equiv \frac{\text{Tr}\left(\text{Diag}\left(\mathbf{F}^{(n)}\right)^2\right)}{\text{Tr}\left({\mathbf{F}^{(n)}}^2\right)} \tag{42}$$

where we have used the symmetry of the matrix $\mathbf{F}^{(n)}$. The numerator and the denominator of (42)'s right hand side are in fact the Frobenius norm squares of the diagonal matrix of $\mathbf{F}^{(n)}$ and itself. The reason why we do not use square roots of the numerator and denominator as is done in the square norm definition lies in the fact that square rooting is avoided (if possible) in especially symbolic computer programming or scripting although it does not seem to be problematic at the first glance. To complete the definition of the first functional mentioned above we need to use the spectral components of the matrix $\mathbf{X}$. If we denote the i–th eigenvalue and the corresponding normalized eigenvector of $\mathbf{X}$ by ξ_i and $\mathbf{x}_i$ respectively then we can define an orthornomal matrix, $\mathbf{Q}$, whose columns from left to right are the vectors $\mathbf{x}_i$s in ascending subindex. The linear mapping on the cartesian counterpart of $\mathcal{H}_n$ under this matrix results in an n dimensional rotation in axes such that the matrix $\mathbf{X}$ becomes diagonal under this mapping so does the matrix $f(\mathbf{X})$. A similar action is expected for the matrix $\mathbf{F}^{(n)}$ which is approximated by $f(\mathbf{X})$, in accordance with the fluctuationlessness theorem. Although we may not expect an exact diagonalization in $\mathbf{F}^{(n)}$ after the rotation, we may expect an increase in the diagonal square dominancy of the same matrix. If we define

$$\mathbf{F}_{\text{rot}}^{(n)} \equiv \mathbf{Q}^T \mathbf{F}^{(n)} \mathbf{Q} \tag{43}$$

and therefore

$$\mu_{\text{dsd}}\left(\mathbf{F}_{\text{rot}}^{(n)}\right) \equiv \frac{\text{Tr}\left(\text{Diag}\left(\mathbf{F}_{\text{rot}}^{(n)}\right)^2\right)}{\text{Tr}\left({\mathbf{F}_{\text{rot}}^{(n)}}^2\right)} \tag{44}$$

then we can arrive at the following functional which measures the increase (we call it increase since we expect so, but, depending on the structure of the function f, it may decrease seldomly) in the matrix $\mathbf{F}^{(n)}$ after the rotation

$$\mathcal{I}_{\text{dsd}}\left(\mathbf{F}^{(n)}\right) \equiv \mu_{\text{dsd}}\left(\mathbf{F}_{\text{rot}}^{(n)}\right) - \mu_{\text{dsd}}\left(\mathbf{F}^{(n)}\right) \tag{45}$$

This is the first one of what we want to use in quality investigations of this section. We call $\mathcal{I}_{\rm dsd}\left(\mathbf{F}^{(n)}\right)$ "Diagonal Square Dominancy Increase".

The second functional can be explicitly defined either in the following form

$$\varepsilon_{\rm rel}\left(\mathbf{F}_{\rm rot}^{(n)}\right) \equiv \frac{\mathrm{Tr}\left(\left(\mathbf{F}^{(n)} - f\left(\mathbf{X}\right)\right)^2\right)}{\mathrm{Tr}\left(\mathbf{F}^{(n)2}\right)} \tag{46}$$

or in its equivalent form obtained after the rotation through $\mathbf{Q}$. We call it "Relative Error"

$$\begin{aligned}\varepsilon_{\rm rel}\left(\mathbf{F}_{\rm rot}^{(n)}\right) &\equiv \frac{\mathrm{Tr}\left(\left(\mathbf{F}_{\rm rot}^{(n)} - \mathbf{\Xi}\right)^2\right)}{\mathrm{Tr}\left(\mathbf{F}_{\rm rot}^{(n)2}\right)} \\ &\equiv \frac{\mathrm{Tr}\left(\left(\mathbf{F}_{\rm rot}^{(n)} - \mathbf{\Xi}\right)^2\right)}{\mathrm{Tr}\left(\mathbf{F}^{(n)2}\right)}\end{aligned} \tag{47}$$

where

$$\mathbf{\Xi} \equiv \mathbf{Q}^T f\left(\mathbf{X}\right)\mathbf{Q} \tag{48}$$

and we have used the fact that the trace operation is not affected by orthonormal transformations. We are going to use the $\mathcal{I}_{\rm dsd}\left(\mathbf{F}^{(n)}\right)$ and $\varepsilon_{\rm rel}\left(\mathbf{F}_{\rm rot}^{(n)}\right)$ functionals in the numerical implementations of this section as the quality observers.

For simplicity, we are going to use the interval $[\,0, 1\,]$ and unit weight, here, in the implementations. The independent variable of three functions we are going to investigate here is taken as scaled by a parameter α to work on a family of functions instead of a single function in each case. Thus, steepening or flattening of the functions can be controlled by this parameter and it becomes possible to see the role of the smoothness on the quality of the fluctuationless approximation.

MuPAD Computer Algebra System which has lost its free software nature quite recently has been used in numerical and symbolic calculations. Its necessary routines for calculations have been run in high precision when necessary and its plotting facilities were used to produce the figures.

Here we present totally six figures, three for each of $\mathcal{I}_{\rm dsd}\left(\mathbf{F}^{(n)}\right)$ and $\varepsilon_{\rm rel}\left(\mathbf{F}_{\rm rot}^{(n)}\right)$. Each figure have five curves painted in different colors and the color table are given at the right uppermost part of the figures. Colors will not be seen explicitly in black and white prints of the paper although its pdf files will contain them. The used colors are red, brick, green, blue, and black for $\mathcal{H}_1$, $\mathcal{H}_2$, $\mathcal{H}_3$, $\mathcal{H}_4$, and $\mathcal{H}_5$ respectively.

The first function is the exponential function's scaled argumented version

$$f_1(x) \equiv e^{\alpha x} \tag{49}$$

which has no singularity at any point in the finite regions of the complex domain of its real independent variable when it is extended to take complex values. That is, it is analytic over the interval $[\,0, 1\,]$ and also everywhere except infinity. It is also monotonously increasing as the argument moves from 0 to 1. The curvature of the plot increases when α grows over the positive real values.

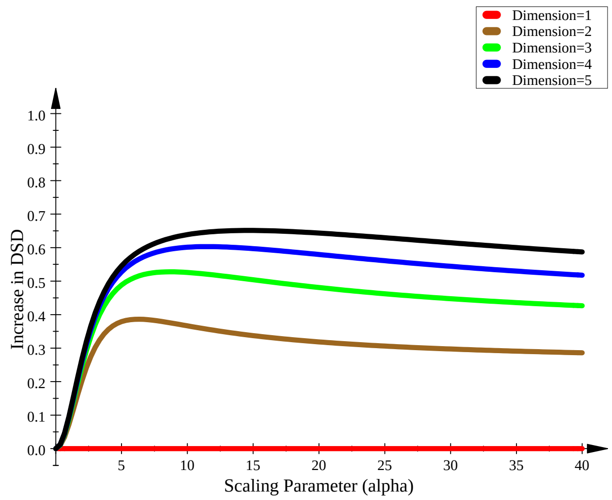

Figure 1: The increase in diagonal square dominancy for the function $f(x) \equiv e^{\alpha x}$ as the Hilbert subspace's dimension (n) increases in fluctuationlessness approximation.

The Figure 1 shows the increases in diagonal square dominancy after the rotations making independent variable's matrix representation diagonal, as the dimension of the related Hilbert space $\mathcal{H}_n$ increases. Whereas Figure 2 depicts the relative error's change with respect to α for first five dimensions. It can be easily noticed that the diagonal dominancy becomes better and the relative error tends to decrease starting from the small values of alpha as the dimension increases. The greater the alpha values the smaller the efficiency of fluctuationless theorem.

Similar behaviors may be expected for the second function given by

$$f_2(x) \equiv \frac{41}{(1+\alpha x)\,(41-\alpha x)} \tag{50}$$

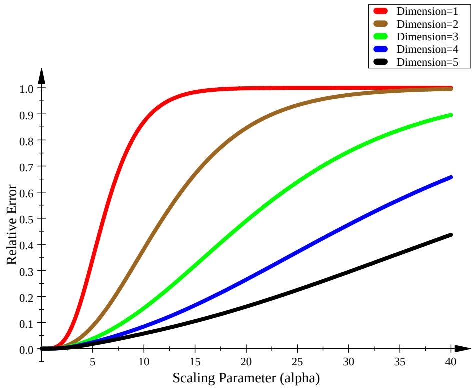

Figure 2: The relative error in norm square with respect to the matrix representation of the function $f(x) \equiv e^{\alpha x}$ as the Hilbert subspace's dimension (n) increases in fluctuationlessness approximation.

where there are two polar singularities at the right hand sides. They are located on the real axis outside the interval $[0, 1]$. The left pole is positioned at the point $-1/\alpha$ and moves from minus infinity to the origin as α grows unboundedly. However it never gets the value of 0. The second pole's location is $41/\alpha$ and it remains outside the interval $[0, 1]$ as long as α does not arrive at or exceeds 41. For $\alpha \geq 41$ the pole becomes an interior point of the interval $[0, 1]$ and destroys the proper integrability of the function f_2 over the interval $[0, 1]$. Although it may be possible to use improper integration which excludes interior poles this is not a natural way and it is out of the content of the fluctuationless theorem here. Hence we have depicted the curves for the real values of α between 0 and 40 here to avoid this singularity. The results are given in Figure 3 and 4.

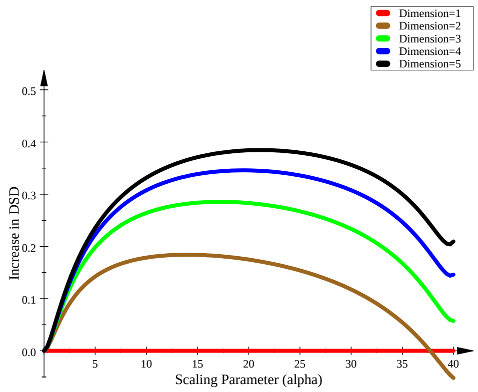

Figure 3: The increase in diagonal square dominancy for the function $f(x) \equiv 1/(1 + \alpha x)(41 - \alpha x)$ as the Hilbert subspace's dimension (n) increases in fluctuationlessness approximation.

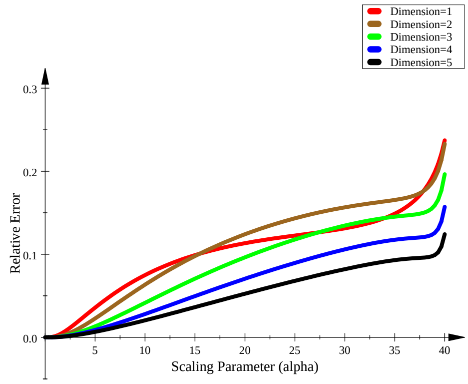

Figure 4: The relative error in norm square with respect to the matrix representation of the function $f(x) \equiv 1/(1 + \alpha x)(41 - \alpha x)$ as the Hilbert subspace's dimension (n) increases in fluctuationlessness approximation.

The sharp upward variation tendency around the singularity in the curves is apparent and evidently due to the pole at $\alpha = 41$.

The third and therefore last example is for the following oscilllatory function

$$f_3(x) \equiv \cos(\alpha x) \tag{51}$$

whose oscillation frequency increases as α tends to grow. The function is analytic everywhere except infinity. The oscillatory behavior is reflected to the plots which are given in Figure 5 and 6. However, the order of the fluctuationlessness aproximation, that is, n or the dimension of $\mathcal{H}_n$ suppresses all oscillatory behavior of the relative error starting from small values of α.

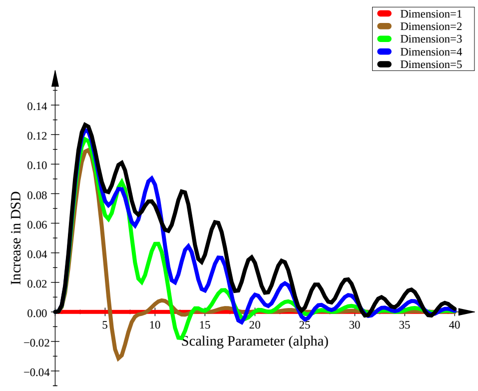

Figure 5: The increase in diagonal square dominancy for the function $f(x) \equiv \cos(\alpha x)$ as the Hilbert subspace's dimension (n) increases in fluctuationlessness approximation.

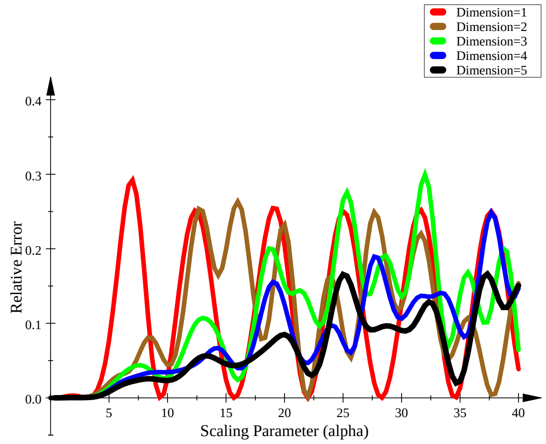

Figure 6: The relative error in norm square with respect to the matrix representation of the function $f(x) \equiv \cos(\alpha x)$ as the Hilbert subspace's dimension (n) increases in fluctuationlessness approximation.

In this example, there is no singularity issue, the negative effects on efficiency come from the oscillatory behaviour of the function and may require higher dimensional matrices to get the same efficiency as the monotonous functions.

One important point to be emphasized on here before closing the section is the removable singular structures especially in the first and third examples. These may not be easily handled by all kind of programming or scripting languages unless specific coding efforts are spent. This is perhaps one of the strongest reasons why symbolic scripting languages or computer algebraic systems are preferred to be used in the implementations of this paper.

6 Concluding remarks

An important theorem which seems to be powerful to open new horizons in various problems of sciences and engineering is given with its proof in this paper. This very first form of it is limited to the matrix representations of algebraic operators whose actions on their operands are just multiplication with a specified univariate function which is assumed to be analytic on the interval of the inner product used in the definition of the matrix representation. However, there are many signals implying the possibility of extending this theorem to the multivariate functions and towards the integrable singularities, in our studies. Our current researches on many application of this theorem is also encouraging.

The fact that even lower dimensions like 3,4,5 give very promising results with high accuracy urges us to apply this theorem in diverse fields of problems. Although we have not given explicitly the relative error it is very small for alpha values around or less than 1. It is possible to get even ten decimal digits around these alpha values as long as the function is sufficiently smooth and the approximation becomes exact (as naturally expected) when the function under consideration is linear. On the other hand, there are certain implications that it will be possible to increase the accuracy via certain appropriate transformation for less efficient cases.

At this point, the function's smoothness and singularity free nature over the related interval are essential items to get higher accuracy. We believe that by taking these facts into consideration one can find many utilization area of this theorem.

Acknowledgements: Author is grateful to Turkish Academy of Sciences for its support.

References:

[1] M. Demiralp and H. Rabitz, Optimally Controlled Quantum Molecular Dynamics: A Perturbation Formulation and the Existence of Multiple Solutions, *Phys. Rev. A* 47, p. 809-816, (1993)

[2] M. Demiralp and H. Rabitz, Optimally Controlled Quantum Molecular Dynamics: The Effect of Nonlinearities on the Magnitude and Multiplicity of Control–Field Solutions, *Phys. Rev. A* 47, p. 831-837, (1993)

[3] M. Demiralp and H. Rabitz, Optimal Control in Classical Molecular Dynamics: A Perturbation Formulation and the Existence of Multiple Solutions, *J. Math. Chem.* 16, p. 185-209, (1994)

[4] M. Demiralp and H. Rabitz, Upper and Lower Bounds on the Control Field and the Quality of Achieved Optimally Controlled Quantum Molecular Motion, *J. Math. Chem.* 19, p. 337-357, (1995)

[5] M. Demiralp and H. Rabitz, Dispersion–free Wave Packets and Feedback Solitonic Motion in Controlled Quantum Dynamics, *Phys. Rev. A* 55, p. 673-677, (1997)

[6] M. Demiralp and H. Rabitz, Assessing Optimality and Robustness of Control Over Quantum Dynamics, *Phys. Rev. A* 57, p. 2420-2425, (1998)

[7] M. Demiralp and H. Rabitz, Assessing Optimality and Robustness for the Control of Dynamical Systems, *Phys. Rev. A* 61, p. 2569-2578, (2000)

[8] A. Kurşunlu, İ. Yaman, M. Demiralp, Optimal Control of One Dimensional Quantum Harmonic Oscillator Under an External Field With Quadratic Dipole Function and Penalty on Momentum: Construction of the Linearised Field Amplitude Integral Equation, *ANACM, Applied Numerical Analysis and Computational Mathematics, WILEY–VCH* 1, p. 270–279, (2003)

[9] G. Li, M. Artamonov, H. Rabitz, S.–W. Wang, P.G. Georgopoulos, and M. Demiralp, High Dimensional Model Representations Generated from Low Order Terms –lp–RS–HDMR, *J. Comput. Chem.* 24, p. 647–656, (2003)

[10] M. Demiralp and B. Tunga, Instantaneous Stability and Robustness Investigations in Quantum Optimal Control: Harmonic Oscillator Under Linear Dipole and Quadratic Control Agents, *Applied Numerical Analysis and Computational Mathematics* 2, p. 60–69, (2005)

[11] H. Rabitz, T.-S. Ho, M. Hsieh, R. Kosut, and M. Demiralp, Topology of optimally controlled quantum mechanical transition probability landscapes, *Physical Review A* 74, p. 012721-(1-9), (2006)

[12] R. Zwanzig, *Nonequilibrium Statistical Mechanics*, Oxford Univ. Press, 2001, p. 30

[13] D. Kincald, W. Cheney, Numerical Analysis (3rd edition) Brooks/Cole, ISBN 0-534-38905-8, (2002)

[14] M. Schatzman, Numerical Analysis: A Mathematical Introduction, Clarendon Press, Oxford, ISBN 0-19-850279-6, (2002)

[15] R.L. Burden, J.D. Faires, Numerical Analysis (7th edition), Brooks/Cole, ISBN 0-534-38216-9, (2000)

[16] M. Abramowitz and I.A. Stegun, eds, Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, Dover, New York, (1972)

[17] G.E. Forsythe, M.A. Malcolm, and C.B. Moler, Computer Methods for Mathematical Computations, Prentice–Hall, Englewood Cliffs, NJ, (1977)

[18] W.H. Press, B.P. Flannery, S.A. Teukolsky, W.T. Vetterling, Numerical Recipes in C, Cambridge University Press, Cambridge, UK, (1988)

[19] J. Stoer and R. Bulirsch, Introduction to Numerical Analysis, Springer–Verlag, New York, (1980)

[20] K.A. Atkinson, An Introduction to Numerical Analysis (2nd edition), John Wiley and Sons, ISBN 0-471-50023-2, (1989)

[21] A. Gil, J. Segura, N.M. Temme, Numerical Methods for Special Functions, SIAM, ISBN 978-0-898716-34-4, (2007)

[22] C.W. Clenshaw and A.R. Curtis, A method for numerical integration on an automatic computer, *Numerische Mathematik* 2, p. 197, (1960)

[23] J. Waldvogel, Fast construction of the Fejér and Clenshaw–Curtis quadrature rules, *BIT Numerical Mathematics* 43(1), p. 001-018, (2004)

[24] W.M. Gentleman, Implementing Clenshaw–Curtis quadrature I: Methodology and Experience, *Comm. of the ACM* 15(5), p. 337-342, (1972)