

Weight optimization in HDMR with perturbation expansion method

Burcu Tunga¹ · Metin Demiralp²

Received: 19 April 2015 / Accepted: 5 August 2015 / Published online: 11 August 2015
© Springer International Publishing Switzerland 2015

Abstract High dimensional model representation method forms an effective divide-and-conquer method used for the truncated representation of a multivariate function, having N independent variables, in terms of certain number ($<2^N$) of less variate functions. The main aim of this method is not to use all these functions in the representation and to obtain an approximation to the given problem. This results in a need of having a good convergence performance just as it is expected in the other numerical methods. This work aims to increase the convergence rate of HDMR approximation by optimizing the weight function, which appears in the method as Prof. Rabitz suggested first, with the help of the perturbation expansion and fluctuationlessness theory. This work also includes a computational procedure with the help of a testing function to better understand the steps of the proposed method.

Keywords Optimization · Approximation · Multivariate functions · High dimensional model representation · Perturbation expansion

1 Introduction

High dimensional model representation method (HDMR), which is used for approximating a given multivariate function through sum of less variate functions, was proposed by Sobol [1] and has played an important role in multivariate modelling

✉ Burcu Tunga
tungab@itu.edu.tr

¹ Mathematics Engineering Department, Faculty of Arts and Sciences, Istanbul Technical University, 34469 Maslak, Istanbul, Turkey

² Informatics Institute, Computational Science and Engineering Program, Istanbul Technical University, 34469 Maslak, Istanbul, Turkey

problems for the last twenty years. HDMR is used for different purposes in different scientific areas such as engineering [2–4], physics [5], chemistry [6]. One of the purposes of using HDMR is basically decomposition of a given multivariate function and then its utilization at truncations for approximating those functions. Although HDMR method has the ability to exactly represent a given multivariate function, in general, to avoid the computational complexity, the scientists prefer to use HDMR as an approximation method in the literature by taking only a few components of its expansion into consideration. Thus, we need to obtain an acceptable representation efficient as much as possible through the HDMR method.

The main task of the HDMR method is to uniquely determine the structures of the components appearing in the HDMR expansion. To do this, a number of multiple integrations under a multivariate weight function should be evaluated over a hyperprismatic grid. The only user dependent selection is the weight selection in the standard HDMR method developed by Sobol and then extended by Rabitz [7–9] and Demiralp [10–12]. Various other scientists have been also working on HDMR [13–16]. Hence, the performance of HDMR in approximately representing a given multivariate function is extremely affected by the weight function selection process. A weight optimization process should be defined for this purpose. This process consists of nonlinear equations and, to solve these equations, we prefer to use perturbation expansion method [17, 18] in this work to achieve better results than the optimization obtained through the fluctuationlessness approximation theory [19–21].

It has been noticed that optimizing the weight selection process in HDMR increases the power of the HDMR based methods in representing a multivariate function. Hence, the main purpose of this work is to develop an alternative method to optimize the weight of HDMR. This new method includes the perturbation expansion method to be used in solving the optimization equations of the weight selection process.

Since HDMR is used for approximation, the performance level of this approximation should be measured by using some tools. One possible tool to observe the convergence of the obtained approximation through HDMR is to use additivity measurers which were defined and utilized first by Demiralp. When we take only the constant HDMR component into consideration and construct constant HDMR approximation, then the additivity measurer that measures the quality of such representation is the constancy measurer. The performance of the univariate HDMR approximation is investigated by looking at the first order additivity measurer. It has the same philosophy for higher variate approximations. In addition, we know from the literature that these measurers construct a practically well-ordered sequence taking the values between 0 and 1. If the value of an additivity measure becomes very close to 1, it means that the obtained HDMR approximation is a successful representation for the given problem. In this manner, the philosophy of weight optimization in this work depends on making the value of the constancy measurer very close to 1. A value very close to 1 for the constancy measurer will let the values of higher level additivity measurers much more closer to 1. This results in a well designed optimization process for the weight of HDMR.

The general structure of the weight function to be optimized is selected from a sub-space of the considered Hilbert space and is defined as the linear combination of functions that are the members of an orthogonal basis set spanning that sub-space.

The optimization process includes nonlinear structures to be solved at the end of the process and this problem corresponds to an eigenvalue problem. In this work, to solve these equations perturbation expansion method is used. This method can be used in various forms in the solution process depending on the needs in details.

One way to construct a perturbation scheme is to use an intermediate entity defined over a Rayleigh quotient appearing in the equations obtained after constancy maximization. An appropriately chosen value of this entity is assumed to be most dominant part of the equation(s) and all other terms are suggested perturbation. The scaling of these terms by a perturbation parameter leads us to a Neumann type perturbation scheme. Unfortunately, we have observed that this way does not work well since the convergence is almost out of control. This approach seemed to be incompatible to the nonlinearities of the equations. Only very restricted subsets of dominance values could be used and enforced us to take so many perturbation expansion terms. This situation was quite contradictory to the nature of perturbation approximation where only first few terms are desired to be taken into consideration for practicality reasons. We abolished this scheme.

The second way we have attempted to progress in was not using an artificial perturbation parameter. The deviation in the value of a parameter defined as a Rayleigh quotient has been considered as perturbation. This needs the determination of intersection points of certain mathematical expressions. It takes us to quite comprehensive iterative procedures and almost blocks out to get analyticity. We have also aborted this approach.

The third way is based on mathematical fluctuation theory combined with a perturbation scheme. This seems to be working well. This has been kept at the focus of this work.

The paper is organized as follows. Section 2 includes brief details of the HDMR method while the Sect. 3 is about the weight optimization process in which the related relations about the optimization equations are given. The solution of the optimization equations formulated in Sect. 4 through the perturbation expansion on the structure of a matrix spectrum is described in the section chapter. Section 5 consists of the computational procedure with a numerical implementation to show the details of the proposed method step-by-step. The numerical examples constructed to investigate the performance of our new method are given in Sect. 6. Finally, the last section includes the concluding remarks of this work.

2 The HDMR method

A given multivariate function can be rewritten in terms of less variate functions by using the following HDMR expansion.

$$f(x_1, \dots, x_N) = f_0 + \sum_{i_1=1}^N f_{i_1}(x_{i_1}) + \sum_{\substack{i_1, i_2=1 \\ i_1 < i_2}}^N f_{i_1 i_2}(x_{i_1}, x_{i_2}) + \dots + f_{1\dots N}(x_1, \dots, x_N) \quad (1)$$

Here, f_0 , $f_{i_1}(x_{i_1})$ and $f_{i_1 i_2}(x_{i_1}, x_{i_2})$ are called constant term, univariate terms and bivariate terms respectively. The higher variate terms are named in the similar manner. If we want to obtain the exact representation for the given multivariate function, we have to use all the right hand side terms of the equation given in (1). However, this way is not preferred because of the computational complexity of such evaluations especially when N grows. Hence, in the literature, the first few terms of HDMR expansion are used for the approximate representation. Generally, this way results in convergence problems. Although there exist a few approaches to overcome this problem, we prefer to develop a weight optimization in this work.

To be able to determine the right hand side terms we can use vanishing conditions which are first defined by Sobol on the unit interval $[0, 1]$ and under the unit constant weight function. These conditions can be explicitly given as follows

$$\int_{a_{i_k}}^{b_{i_k}} dx_{i_k} W_{i_k}(x_{i_k}) f_{i_1 \dots i_s}(x_{i_1}, \dots, x_{i_s}) = 0, \quad i_1 \leq i_k \leq i_s \quad (2)$$

The weight function, $W(x_1, \dots, x_N)$, appearing in the vanishing conditions satisfy the following relation.

$$W(x_1, \dots, x_N) \equiv \prod_{i_1=1}^N W_{i_1}(x_{i_1}), \quad \int_{a_{i_1}}^{b_{i_1}} dx_{i_1} W_{i_1}(x_{i_1}) = 1, \quad x_{i_1} \in [a_{i_1}, b_{i_1}], \quad 1 \leq i_1 \leq N \quad (3)$$

To uniquely determine the general structures of the right hand side components is the main task of HDMR based methods. To this end, the first step is to obtain the general structure of the constant component. Applying an N -tuple integration to the both sides of the HDMR expansion under the vanishing conditions given in (2) gives us the following relation for the constant HDMR component

$$f_0 = \int_{a_1}^{b_1} dx_1 \cdots \int_{a_N}^{b_N} dx_N W(x_1, \dots, x_N) f(x_1, \dots, x_N) \quad (4)$$

The structure of univariate components is determined through $N - 1$ -tuple integration in which the integration on the related independent variable (the argument of the function) is discarded.

$$f_i(x_i) = \int_{a_1}^{b_1} dx_1 W_1(x_1) \cdots \int_{a_{i-1}}^{b_{i-1}} dx_{i-1} W_{i-1}(x_{i-1}) \int_{a_{i+1}}^{b_{i+1}} dx_{i+1} W_{i+1}(x_{i+1}) \\ \times \cdots \int_{a_N}^{b_N} dx_N W_N(x_N) f(x_1, \dots, x_N) - f_0, \quad 1 \leq i \leq N \quad (5)$$

Although we do not want to give more details about the above formula, we emphasize on that there are some approximants which are defined by making truncation at some level in HDMR method. These approximants are explicitly given as follows.

$$\begin{aligned}
 s_0(x_1, \dots, x_N) &= f_0 \\
 s_1(x_1, \dots, x_N) &= s_0(x_1, \dots, x_N) + \sum_{i_1=1}^N f_{i_1}(x_{i_1}) \\
 &\vdots \\
 s_k(x_1, \dots, x_N) &= s_{k-1}(x_1, \dots, x_N) + \sum_{\substack{i_1 \dots i_k=1 \\ i_1 < \dots < i_k}}^N f_{i_1 \dots i_k}(x_{i_1}, \dots, x_{i_k}), \\
 &1 \leq k \leq N
 \end{aligned} \tag{6}$$

To investigate the quality of these approximations, the following ‘‘Additivity Measurers’’ are defined

$$\begin{aligned}
 \sigma_0 &\equiv \frac{1}{\|f\|^2} \|f_0\|^2 \\
 \sigma_1 &\equiv \frac{1}{\|f\|^2} \sum_{i=1}^N \|f_i\|^2 + \sigma_0 \\
 &\vdots \\
 \sigma_N &\equiv \frac{1}{\|f\|^2} \|f_{12\dots N}\|^2 + \sigma_{N-1}
 \end{aligned} \tag{7}$$

These additivity measurers have the following practically well-ordered structure. In true mathematical sense equalities should not appear to get well-ordered structure in this formula. However, for practicality, we may use this expression for our purposes here.

$$0 \leq \sigma_0 \leq \dots \leq \sigma_N = 1 \tag{8}$$

3 The weight optimization process

High Dimensional Model Representation method uses a finite expansion that expresses the given multivariate function in terms of less variate functions. Hence, this method provides us an easy calculation in scientific problems having multivariate functions. However, the important point in this method is that the number of calculations should not be increased by using many terms from the expansion. That is, to have low mathematical and computational complexity, the HDMR expansion should be truncated at some level and it is better to have a number (which is as less as possible) of HDMR components in the truncation which corresponds to an approximate representation of the given multivariate function. To increase the quality of this approximation, the weight function optimization process can be performed. Although there may be some other ways of obtaining better approximation, in this study the constancy measurer is used in the optimization process and the equations obtained through this process are solved by using perturbation expansion method.

To accomplish weight optimization process, we choose the weight function as the square of a linear combination of the functions which form an orthonormal basis set spanning a subspace of Hilbert space under consideration. The reason why we use the square of the whole relation here is to guarantee the positiveness of the weight function. Because we now have a positive function and this function has the property that vanishes over the interval of its independent variables at only certain finite number of points, this function can be said to be a true weight function in mathematical sense.

$$W(x) = \left(\sum_{j=1}^n \alpha_j w_j(x) \right)^2 \quad (9)$$

Here α_j s are arbitrary parameters and they will be used as the basic unknowns of the optimization process. In this study, this optimization process will be achieved through only constancy measurer. Definitely, it is possible to use the other measurers in the optimization process. However, we know that when we push the value of the constancy measurer to 1, the others also naturally get much closer to 1. To this end, the relation for the constancy measurer can be rewritten by using the weight structure given in (9) as follows.

$$\sigma_0 = \frac{\left(\sum_{j=1}^n \sum_{k=1}^n \alpha_j \alpha_k \int_a^b dx w_j(x) w_k(x) f(x) \right)^2}{\sum_{j=1}^n \sum_{k=1}^n \alpha_j \alpha_k \int_a^b dx w_j(x) w_k(x) f(x)^2} \quad (10)$$

The weight factors appearing in the HDMR method supply the integral normalization condition given as relation (3). That is,

$$\int_a^b dx \sum_{j=1}^m \sum_{k=1}^m \alpha_j \alpha_k w_j(x) w_k(x) = 1 \quad (11)$$

Here, it is emphasized on that $w_j(x)$ functions are the elements of an orthonormal basis set. Hence, the following equation can be written

$$\int_a^b dx w_j(x) w_k(x) = \delta_{jk} \quad 1 \leq j, k \leq m \quad (12)$$

If the equation given in (11) is rewritten by using above relation, the following equation is obtained for the α_j parameters.

$$\sum_{j=1}^m \sum_{k=1}^m \alpha_j \alpha_k \delta_{jk} = \sum_{j=1}^m \alpha_j^2 = 1 \quad (13)$$

To begin the optimization process, a cost functional has to be defined as follows

$$J(\alpha_1, \dots, \alpha_m, \lambda) = \frac{\left(\sum_{j=1}^m \sum_{k=1}^m \alpha_j \alpha_k a_{jk}^{(1)}\right)^2}{\sum_{j=1}^m \sum_{k=1}^m \alpha_j \alpha_k a_{jk}^{(2)}} + \lambda \left(\sum_{j=1}^m \alpha_j^2 - 1\right) \tag{14}$$

where

$$a_{jk}^{(1)} = \int_a^b dx w_j(x) w_k(x) f(x), \quad a_{jk}^{(2)} = \int_a^b dx w_j(x) w_k(x) f(x)^2 \tag{15}$$

The second step for the optimization process is to differentiate the cost functional with respect to the independent variables, λ and α_i s, separately. When this step is conducted and we set the result of each differentiation to 0, below equations are obtained

$$\sum_{j=1}^m \alpha_j^2 = 1 \tag{16}$$

$$-2\mu \sum_{k=1}^m \alpha_k a_{ik}^{(1)} + \mu^2 \sum_{k=1}^m \alpha_k a_{ik}^{(2)} = \lambda \alpha_i \tag{17}$$

where

$$\mu = \frac{\sum_{j=1}^m \sum_{k=1}^m \alpha_j \alpha_k a_{jk}^{(1)}}{\sum_{j=1}^m \sum_{k=1}^m \alpha_j \alpha_k a_{jk}^{(2)}} \tag{18}$$

The equation given in (17) is an eigenvalue problem. The α parameters appearing in the equation of this eigenvalue problem are the unknowns that are needed to specify the optimized weight of HDMR. To determine the values of the α parameters, we need to obtain the value of μ . The relation (18) which is given for μ also includes the unknown α parameters and this case corresponds to the solution of a nonlinearly structured equation. As very well known as solving a nonlinear problem is not an easy task and to bypass the disadvantages coming from nonlinearity we will use the fluctuation matrices defined in the fluctuationlessness approximation theory [19–21]. The fluctuation matrices are the matrices whose elements are defined through the following inner products

$$\mathbf{M}_n(\widehat{f})_{jk} = (w_j, f w_k), \quad \mathbf{M}_n(\widehat{f^2})_{jk} = (w_j, f^2 w_k), \quad 1 \leq j, k \leq n \tag{19}$$

where the symbols \widehat{f} and $\widehat{f^2}$ correspond to the algebraic multiplication operators which multiply their operands by the functions $f(x)$ and $f(x)^2$ respectively.

To this end, Eqs. (17) and (18) are rewritten by using fluctuation matrices. The final forms of these equations are obtained as follows

$$-2\mu \mathbf{M}_n(\widehat{f})\boldsymbol{\alpha} + \mu^2 \mathbf{M}_n(\widehat{f^2})\boldsymbol{\alpha} = \lambda \boldsymbol{\alpha} \quad (20)$$

$$\mu = \frac{\boldsymbol{\alpha}^T \mathbf{M}_n(\widehat{f})\boldsymbol{\alpha}}{\boldsymbol{\alpha}^T \mathbf{M}_n(\widehat{f^2})\boldsymbol{\alpha}} \quad (21)$$

The final results of this problem will be obtained by applying the perturbation expansion method to the matrix form of these equations. For this purpose, the following relation can be taken into consideration

$$\mathbf{M}_n(\widehat{f}[\mathbf{I} - \mathbf{P}^{(n)}]\widehat{f}) = \mathbf{M}_n(\widehat{f^2}) - \mathbf{M}_n(\widehat{f})^2 \quad (22)$$

then the eigenvalue equation given in (20) and the accompanying relation given in (21) become the following form by using n dimensional fluctuation matrix multiplied by an ε parameter.

$$-2\mu \mathbf{M}_n(\widehat{f})\boldsymbol{\alpha} + \mu^2 \mathbf{M}_n(\widehat{f})^2 \boldsymbol{\alpha} + \varepsilon \mu^2 \mathbf{M}_n(\widehat{f}[\mathbf{I} - \mathbf{P}^{(n)}]\widehat{f})\boldsymbol{\alpha} = \lambda \boldsymbol{\alpha} \quad (23)$$

$$\mu(\varepsilon)\boldsymbol{\alpha}(\varepsilon)^T \mathbf{M}_n(\widehat{f})^2 \boldsymbol{\alpha}(\varepsilon) + \varepsilon \boldsymbol{\alpha}(\varepsilon)^T \mathbf{M}_n(\widehat{f}[\mathbf{I} - \mathbf{P}^{(n)}]\widehat{f})\boldsymbol{\alpha}(\varepsilon) = \boldsymbol{\alpha}(\varepsilon)^T \mathbf{M}_n(\widehat{f})\boldsymbol{\alpha}(\varepsilon) \quad (24)$$

Before applying the perturbation expansion method, the expression of λ must be determined in matrix representation form. To this end, some handy manipulations should be done. If the Eq. (20) is divided by μ and we use $\bar{\lambda}$ instead of $\frac{\lambda}{\mu}$, the following relation is obtained

$$-2\mathbf{M}_n(\widehat{f})\boldsymbol{\alpha} + \mu \mathbf{M}_n(\widehat{f^2})\boldsymbol{\alpha} = \bar{\lambda} \boldsymbol{\alpha} \quad (25)$$

Afterwards, if the explicit form of μ given in Eq. (21) is replaced into this relation and we multiply both sides of this new form from the left by $\boldsymbol{\alpha}^T$, the resulting structure is obtained as

$$-\boldsymbol{\alpha}^T \mathbf{M}_n(\widehat{f})\boldsymbol{\alpha} = \bar{\lambda} \boldsymbol{\alpha}^T \boldsymbol{\alpha} \quad (26)$$

Since $\boldsymbol{\alpha}^T \boldsymbol{\alpha} = 1$, the value of λ is found in the following final form

$$\lambda = -\frac{(\boldsymbol{\alpha}^T \mathbf{M}_n(\widehat{f})\boldsymbol{\alpha})^2}{\boldsymbol{\alpha}^T \mathbf{M}_n(\widehat{f^2})\boldsymbol{\alpha}} = -\sigma_0 \quad (27)$$

4 Perturbation expansion in the structure of matrix spectrum

The previous sections cover the determination process of the optimization equations and the construction of inner structures needed to successfully complete the solution process of these equations. This section includes the steps designed to solve the above-mentioned optimization equations. To do this, the equations given in (23) and (24) are reconstructed as follows through the perturbation expansion method

$$2\mu(\varepsilon)\mathbf{M}_n(\widehat{f})\boldsymbol{\alpha}(\varepsilon) - \mu(\varepsilon)^2\mathbf{M}_n(\widehat{f})^2\boldsymbol{\alpha}(\varepsilon) - \varepsilon\mu(\varepsilon)^2\mathbf{M}_n(\widehat{f}[\mathbf{I} - \mathbf{P}^{(n)}]\widehat{f})\boldsymbol{\alpha}(\varepsilon) = \sigma_0(\varepsilon)\boldsymbol{\alpha}(\varepsilon) \tag{28}$$

$$\mu(\varepsilon) = \frac{\boldsymbol{\alpha}(\varepsilon)^T \mathbf{M}_n(\widehat{f})\boldsymbol{\alpha}(\varepsilon)}{\boldsymbol{\alpha}(\varepsilon)^T \mathbf{M}_n(\widehat{f})^2\boldsymbol{\alpha}(\varepsilon) + \varepsilon\boldsymbol{\alpha}(\varepsilon)^T \mathbf{M}_n(\widehat{f}[\mathbf{I} - \mathbf{P}^{(n)}]\widehat{f})\boldsymbol{\alpha}(\varepsilon)} \tag{29}$$

The next step is to define the following relations for $\boldsymbol{\alpha}$, μ and σ_0 by using the perturbation parameter, ε

$$\boldsymbol{\alpha}(\varepsilon) = \sum_{i=0}^{\infty} \boldsymbol{\alpha}_i \varepsilon^i, \quad \mu(\varepsilon) = \sum_{i=0}^{\infty} \mu_i \varepsilon^i, \quad \sigma_0(\varepsilon) = \sum_{i=0}^{\infty} \sigma_0^{(i)} \varepsilon^i \tag{30}$$

Here, we need to make these relations approximately finite to be able to use in our solution process. For this purpose, the first m components of each relation will be taken into consideration.

$$\boldsymbol{\alpha}^{(m)}(\varepsilon) = \sum_{i=0}^m \boldsymbol{\alpha}_i \varepsilon^i, \quad \mu^{(m)}(\varepsilon) = \sum_{i=0}^m \mu_i \varepsilon^i, \quad \sigma_{0,m}(\varepsilon) = \sum_{i=0}^m \sigma_0^{(i)} \varepsilon^i \tag{31}$$

That is, we have m th order perturbation expansion. It means that $m + 1$ number of equations to be solved are obtained. Here, $\boldsymbol{\alpha}^{(m)}$ is a vector, $\mu^{(m)}$ and $\sigma_{0,m}$ are scalars.

When the relations given in (31) are inserted into the relations given in (28) and (29), $m + 1$ number equations appear for each relation. Since the aim in perturbation expansion method is not to select the m value greater than 3 or 4 for practicality issues, the mentioned number of equations does not bring additional computational complexity to the proposed method of this work. The obtained equations for relation (28) are as follows when we take $m = 2$

$$2\mu_0\mathbf{M}_n(\widehat{f})\boldsymbol{\alpha}_0 - \mu_0^2\mathbf{M}_n(\widehat{f})^2\boldsymbol{\alpha}_0 = \sigma_0^{(0)}\boldsymbol{\alpha}_0 \tag{32}$$

$$2\mu_0\mathbf{M}_n(\widehat{f})\boldsymbol{\alpha}_1 + \mu_1\mathbf{M}_n(\widehat{f})\boldsymbol{\alpha}_0 - \mu_0^2\mathbf{M}_n(\widehat{f})^2\boldsymbol{\alpha}_1 - 2\mu_0\mu_1\mathbf{M}_n(\widehat{f})^2\boldsymbol{\alpha}_0 - \mu_0^2\mathbf{M}_n(\widehat{F})\boldsymbol{\alpha}_0 = \sigma_0^{(0)}\boldsymbol{\alpha}_1 + \sigma_0^{(1)}\boldsymbol{\alpha}_0 \tag{33}$$

$$2\mu_0\mathbf{M}_n(\widehat{f})\boldsymbol{\alpha}_2 + \mu_1\mathbf{M}_n(\widehat{f})\boldsymbol{\alpha}_1 + 2\mu_2\mathbf{M}_n(\widehat{f})\boldsymbol{\alpha}_0 - \mu_0^2\mathbf{M}_n(\widehat{f})^2\boldsymbol{\alpha}_2 - 2\mu_0\mu_1\mathbf{M}_n(\widehat{f})^2\boldsymbol{\alpha}_1 - \mu_1^2\mathbf{M}_n(\widehat{f})^2\boldsymbol{\alpha}_0 - 2\mu_0\mu_2\mathbf{M}_n(\widehat{f})^2\boldsymbol{\alpha}_0 - \mu_0^2\mathbf{M}_n(\widehat{F})\boldsymbol{\alpha}_1 - 2\mu_0\mu_1\mathbf{M}_n(\widehat{F})\boldsymbol{\alpha}_0 = \sigma_0^{(0)}\boldsymbol{\alpha}_2 + \sigma_0^{(1)}\boldsymbol{\alpha}_1 + \sigma_0^{(2)}\boldsymbol{\alpha}_0 \tag{34}$$

where $\mathbf{M}_n(\widehat{F}) \equiv \mathbf{M}_n(\widehat{f}[\mathbf{I} - \mathbf{P}^{(n)}]\widehat{f})$. Here, for simplicity, we select the m value as 2 to show the general characteristics of the obtained perturbation equations. Constancy measurer for each eigen-pairs by using first order perturbation expansion

The similar steps are applied to the relation (29) and the following equations are obtained

$$\mu_0 \alpha_0^T M_n(\widehat{f})^2 \alpha_0 = \alpha_0^T M_n(\widehat{f}) \alpha_0 \tag{35}$$

$$\begin{aligned} \mu_0 \alpha_0^T M_n(\widehat{f})^2 \alpha_1 + \mu_0 \alpha_1^T M_n(\widehat{f})^2 \alpha_0 + \mu_1 \alpha_0^T M_n(\widehat{f})^2 \alpha_0 + \alpha_0^T M_n(\widehat{F}) \alpha_0 \\ = \alpha_0^T M_n(\widehat{f}) \alpha_1 + \alpha_1^T M_n(\widehat{f}) \alpha_0 \end{aligned} \tag{36}$$

$$\begin{aligned} \mu_0 \alpha_0^T M_n(\widehat{f})^2 \alpha_2 + \mu_0 \alpha_1^T M_n(\widehat{f})^2 \alpha_1 + \mu_0 \alpha_2^T M_n(\widehat{f})^2 \alpha_0 + \mu_1 \alpha_0^T M_n(\widehat{f})^2 \alpha_1 \\ + \mu_1 \alpha_1^T M_n(\widehat{f})^2 \alpha_0 + \mu_2 \alpha_0^T M_n(\widehat{f})^2 \alpha_0 + \alpha_0^T M_n(\widehat{F}) \alpha_1 + \alpha_1^T M_n(\widehat{F}) \alpha_0 \\ = \alpha_0^T M_n(\widehat{f}) \alpha_2 + \alpha_1^T M_n(\widehat{f}) \alpha_1 + \alpha_2^T M_n(\widehat{f}) \alpha_0 \end{aligned} \tag{37}$$

The unknowns are $\alpha_0, \alpha_1, \alpha_2, \mu_0, \mu_1, \mu_2, \sigma_0^{(0)}, \sigma_0^{(1)}$ and $\sigma_0^{(2)}$ in these equations. To determine these unknowns, first, the equations given in (32) and (35) are taken into consideration. To solve these equations it is assumed that the fluctuation matrix, M_n , having $n \times n$ type satisfies the following eigenvalue problem

$$M_n(\widehat{f}) \mathbf{g}_i = \gamma_i \mathbf{g}_i \quad 1 \leq i \leq n \tag{38}$$

where γ_i s are the eigenvalues and \mathbf{g}_i s are the corresponding eigenvectors of M_n . Since the vector α_0 is equal to one of the eigenvectors, \mathbf{g}_i ($1 \leq i \leq n$), these eigenvectors can be used in Eq. (35) to determine the unknown μ_0

$$\mu_0 \mathbf{g}_i^T \gamma_i^2 \mathbf{g}_i = \mathbf{g}_i^T \gamma_i \mathbf{g}_i \tag{39}$$

where $\mathbf{g}_i^T \mathbf{g}_i = 1$.

$$\mu_0 \gamma_i^2 = \gamma_i \quad \rightarrow \quad \mu_0 = \frac{1}{\gamma_i}, \quad 1 \leq i \leq n \tag{40}$$

Hence μ_0 is found as it is equal to multiplicative inverse of one of the eigenvalues of matrix $M_n(\widehat{f})$. Afterwards, if μ_0 and α_0 are placed into Eq. (32), the value of $\sigma_0^{(0)}$ is obtained as 1. This results in as follows when we also take the relation (31) into consideration

$$2 \frac{1}{\gamma_i} \gamma_i \mathbf{g}_i - \frac{1}{\gamma_i^2} \gamma_i^2 \mathbf{g}_i = \sigma_0^{(0)} \mathbf{g}_i \quad \rightarrow \quad \sigma_0^{(0)} = 1 \quad \rightarrow \quad \sigma_{0,0} = 1 \tag{41}$$

Other unknowns can be found by following the same steps. The values of $\sigma_0^{(1)}$ and μ_1 are determined from the above Eqs. (33) and (36) respectively

$$\sigma_0^{(1)} = -\frac{1}{\gamma_i^2} \mathbf{g}_i^T M_n(\widehat{f}) \mathbf{g}_i \tag{42}$$

$$\mu_1 = -\frac{1}{\gamma_i^2} \mathbf{g}_i^T M_n(\widehat{f}) \mathbf{g}_i \tag{43}$$

To be able to determine the unknown vector α_1 , the Eq. (33) has to be rewritten as follows

$$\alpha_1 = \left(\frac{2}{\gamma_i} M_n(\hat{f}) - \frac{1}{\gamma_i^2} M_n(\hat{f})^2 - I \right)^{-1} \left(\frac{1}{\gamma_i^2} M_n(\hat{F}) g_i - \frac{1}{\gamma_i^2} g_i^T M_n(\hat{F}) g_i \right) \tag{44}$$

Here, we have to find the inverse of $\left(\frac{2}{\gamma_i} M_n(\hat{f}) - \frac{1}{\gamma_i^2} M_n(\hat{f})^2 - \sigma_0^{(0)} I \right)$ to evaluate α_1 . For this purpose, the following relation is taken into consideration

$$\left(\frac{2}{\gamma_i} M_n(\hat{f}) - \frac{1}{\gamma_i^2} M_n(\hat{f})^2 \right) g_k = \beta_k g_k, \quad I = \sum_{k=1}^n g_k g_k^T, \quad g_k^T g_k = 1 \tag{45}$$

where $k = 1, \dots, n$.

The pseudo-inverse of the expression, $\left(\frac{2}{\gamma_i} M_n(\hat{f}) - \frac{1}{\gamma_i^2} M_n(\hat{f})^2 - \sigma_0^{(0)} I \right)$ is obtained as

$$\left(\frac{2}{\gamma_i} M_n(\hat{f}) - \frac{1}{\gamma_i^2} M_n(\hat{f})^2 - \sigma_0^{(0)} I \right)_d^{-1} = \sum_{\substack{j=1 \\ j \neq i}}^n \frac{1}{(\beta_j - 1)} \left[\frac{1}{\gamma_i^2} g_j^T M_n(\hat{F}) g_i \right] g_j g_j^T \tag{46}$$

This result is used in Eq. (44) the vector α_1 is obtained as follows.

$$\alpha_1 = \sum_{\substack{j=1 \\ j \neq i}}^N \frac{1}{(\beta_j - 1)\gamma_i^2} \left(g_j^T M_n(\hat{F}) g_i \right)^2 g_j \tag{47}$$

If we solve the equation given in (37), the value of $\sigma_0^{(2)}$ is obtained as

$$\sigma_0^{(2)} = 3\mu_1^2 \gamma_i - \frac{1}{\gamma_i^2} g_i^T M_n(\hat{F}) \alpha_1 - \sigma_0^{(1)} g_i^T \alpha_1 \tag{48}$$

Now, to determine the vector α_2 , the Eq. (37) is used.

$$\alpha_2 = \sum_{\substack{j=1 \\ j \neq i}}^N \frac{1}{(\beta_j - 1)\gamma_i^2} \left(g_j^T M_n(\hat{F}) g_i \right) g_j \times \left[\frac{1}{\gamma_i^2} g_j^T M_n(\hat{F}) \alpha_1 + \frac{2}{\gamma_i} \mu_1 g_j^T M_n(\hat{F}) g_i + \sigma_0^{(1)} g_j^T \alpha_1 \right] \tag{49}$$

If all obtained results are inserted into (31) and the ε parameter is set equal to 1, we produce results for a nonlinear problem by using second order perturbation expansion.

The next section covers a computational procedure through univariate functions for better understanding the algorithm.

5 The computational procedure

In the previous section, the theoretical details about how the weight optimization process is performed through constancy measurer with the help of perturbation expansion, to increase the quality of HDMR method, is given. In this section, to make the method more clear, the steps of this proposed method are given through a numerical example.

The inputs of the algorithm are the interval information of each independent variable and the analytical structure of a given function.

The output of the algorithm is an approximate analytical structure which represents the original function in terms of HDMR components. The proposed method of this work uses fluctuation free matrix representations and the perturbation expansion theory to determine the appropriate HDMR components for the given problem in which a better representation is obtained with respect to the method in which the plain HDMR method is used for the purpose.

The numerical example selected to describe the steps of our new method is as follows

$$f = \sqrt{x+1}, \quad -1 \leq x \leq 1 \quad (50)$$

This example is same as it is selected in the other related published paper of the authors to easily compare the performance of the proposed method of this work and the other method given in that work [19].

The steps of this new algorithm are given as follows:

1. Evaluate the fluctuation matrix of the given function.
 - Choose the dimension of the matrix and the basis set needed to construct the components of this matrix. Here, the dimension and the elements of the basis set are chosen as $n = 3$ and $\{1, x, x^2\}$ respectively. The elements of this basis set should be mutually orthonormal. This can be provided by using the Gram-Schmidt Orthonormalization Method [22]. The elements of this basis set has the following structures for $n = 3$ after the mentioned orthonormalization on the given basis set

$$w_1(x) = \sqrt{\frac{1}{2}}, \quad w_2(x) = \frac{\sqrt{6}x}{2}, \quad w_3(x) = \frac{3\sqrt{10}(x^2 - \frac{1}{3})}{4} \quad (51)$$

2. Determine the eigenvalues, γ_i , and the corresponding eigenvectors, \mathbf{g}_i , by using eigen-equation given in (38)

$$\gamma_1 = 1.327669153, \quad \gamma_2 = 0.9771259876, \quad \gamma_3 = 0.4207800884 \quad (52)$$

$$\mathbf{g}_1 = \begin{bmatrix} 0.539066018 \\ 0.7167974794 \\ 0.4422772905 \end{bmatrix}, \quad \mathbf{g}_2 = \begin{bmatrix} 0.6816365406 \\ -0.06282368898 \\ -0.7289888961 \end{bmatrix},$$

$$\mathbf{g}_3 = \begin{bmatrix} 0.4947519123 \\ -0.6944455037 \\ 0.52246147 \end{bmatrix} \quad (53)$$

where $\alpha_0[1] = g_1, \alpha_0[2] = g_2$ and $\alpha_0[3] = g_3$.

- Evaluate $\sigma_0^{(0)}$ and $\sigma_0^{(1)}$ by using (41) and (42) respectively. We know from (41) that the value of $\sigma_0^{(0)}$ is independent from the problem and is equal to 1. The values of $\sigma_0^{(1)}$ for each eigenvalue are obtained as follows

$$\begin{aligned} \sigma_0^{(1)}[1] &= -0.0061793943, & \sigma_0^{(1)}[2] &= -0.04511730997, \\ \sigma_0^{(1)}[3] &= -0.2908358637 \end{aligned} \tag{54}$$

where $\sigma_0^{(1)}[1], \sigma_0^{(1)}[2]$ and $\sigma_0^{(1)}[3]$ are obtained for the first, second and third eigenvalues respectively.

- Determine the value of constancy measurer, $\sigma_{0,1}$, for each eigenvalue by using the relation (31) and $\sigma_0^{(1)}[1], \sigma_0^{(1)}[2], \sigma_0^{(1)}[3]$ values.

$$\sigma_{0,1}[1] = 0.9938206057, \quad \sigma_{0,1}[2] = 0.95488269, \quad \sigma_{0,1}[3] = 0.7091641363 \tag{55}$$

- Evaluate the α_1 vectors. For this purpose, β_k values should be obtained by using the relation (45). Here, each γ_i produces $n \times n$ dimensional matrix, $\left(\frac{2}{\gamma_i} M_n(\hat{f}) - \frac{1}{\gamma_i^2} M_n(\hat{f})^2\right)$, so there are nxn number of β values. If we use these values and Eq. (47), we obtain n number of α_1 vectors.

$$\begin{aligned} \alpha_1[1] &= \begin{bmatrix} -0.002928938227 \\ 0.0007020365019 \\ 0.002432123682 \end{bmatrix}, & \alpha_1[2] &= \begin{bmatrix} -0.01378934274 \\ 0.005855699862 \\ 0.007316711143 \end{bmatrix}, \\ \alpha_1[3] &= \begin{bmatrix} -0.006411619398 \\ 0.002682877239 \\ 0.003466618179 \end{bmatrix} \end{aligned} \tag{56}$$

- Compute the $\sigma_0^{(2)}$ values by using the vectors, α_1 , and the relation (48).

$$\sigma_0^{(2)}[1] = 0.0001140984, \quad \sigma_0^{(2)}[2] = 0.0056209228, \quad \sigma_0^{(2)}[3] = 0.1049092968 \tag{57}$$

- Obtain the $\sigma_{0,2}$ value by using the relation given in (31). This value is the constancy measurer obtained through the second order perturbation expansion method. There will again be 3 values since the dimension is selected as 3.

$$\sigma_{0,2}[1] = 0.9939347041, \quad \sigma_{0,2}[2] = 0.9605036128, \quad \sigma_{0,2}[3] = 0.8140734331 \tag{58}$$

8. Evaluate the α_2 vectors by using the relation (49).

$$\alpha_2[1] = \begin{bmatrix} 0.000134137577 \\ -0.0001672080584 \\ 0.0001075011229 \end{bmatrix}, \alpha_2[2] = \begin{bmatrix} 0.0008996021449 \\ -0.001178910148 \\ 0.0008141835096 \end{bmatrix},$$

$$\alpha_2[3] = \begin{bmatrix} 0.002345385423 \\ -0.003032173772 \\ 0.002055581319 \end{bmatrix} \quad (59)$$

9. Identify the eigenvector that maximizes the constancy measurer evaluated in step 7. Depending on the results given in (58), this vector is the first eigenvector for the numerical example under consideration.
10. Evaluate the weight parameter, α by taking the result of the step 9 and the relation (31) into consideration.

$$\alpha = \alpha_0[1] + \alpha_1[1] + \alpha_2[1] \implies \alpha = \begin{bmatrix} 0.5362712174 \\ 0.7173323079 \\ 0.4448169153 \end{bmatrix} \quad (60)$$

11. Construct the optimized weight function by using the relation given in (9), the abovementioned vector, α and the elements of the basis set given in (51).

$$W_{opt} = (1.054975946x^2 + 0.878549065x + 0.02754236584)^2 \quad (61)$$

12. Evaluate the constant HDMR component of the given function through this optimized weight.

$$f_0 = 1.327683266 \quad (62)$$

6 Implementations

This section covers several implementations to investigate the performance of the proposed method of this work. The numerical results are obtained by using MuPAD [23], a computer algebra system, within 10-decimal-digits precision.

The first testing function is selected as exponential function

$$f_1(x) = e^x, \quad -1 \leq x \leq 1 \quad (63)$$

and the number of elements of basis set is chosen as 10. This results in using a 10×10 fluctuation matrix. This means that, we obtain 10 eigenvalues and corresponding eigenvectors. When we follow up the steps given in the previous section, we can evaluate constancy measurers obtained through the first and second order perturbation expansion respectively. Table 1 includes these constancy measurer values for each eigenpair of the fluctuation matrix. The value of $\sigma_{0,0}$ is 1 which comes from the theory of the method as mentioned before. Here, we need to select the eigenpair which gives the maximum value for the constancy measurer obtained through second order

Table 1 Constancy measurer values obtained for the eigenpairs of the first testing function

Eigenpairs #	1	2	3	4	5	6	7	8	9	10
$\sigma_{0,0}$	1	1	1	1	1	1	1	1	1	1
$\sigma_{0,1}$	0.9984	0.9919	0.9810	0.9675	0.9542	0.9455	0.9459	0.9578	0.9775	0.9950
$\sigma_{0,2}$	0.9985	0.9922	0.9827	0.9719	0.9610	0.9540	0.9496	0.9651	0.9746	0.9955

Table 2 σ_0 values for optimized and unoptimized weight functions ($n = 9$)

Testing functions	$1 - x^2$	$(1 + x)^2 e^x$	$\sin(x)$	$\sqrt{1 + x}$	$\ln(1 + \frac{x}{2})$
$\sigma_0 (W := \frac{1}{2})$	0.8333	0.4002	0	0.8888	0.0211
σ_0 (previous optimized weight)	0.9743	0.9950	0.9970	0.9996	0.9960
σ_0 (new optimized weight)	0.9978	0.9974	0.9977	0.9997	0.9968

perturbation expansion. The first eigenpair gives the highest constancy measurer value and it is highlighted as in Table 1. Next step is to use the first eigenvector to determine the optimized weight and this optimized weight function is obtained as follows

$$\begin{aligned}
 W_{opt} = & (35.14911413x^9 + 30.53700573x^8 - 53.38744191x^7 - 44.16006669x^6 \\
 & + 25.33028634x^5 + 19.18029593x^4 - 4.048023674x^3 - 2.533398576x^2 \\
 & + 0.1476361084x + 0.04821945896)^2 \tag{64}
 \end{aligned}$$

The constant HDMR component can be evaluated by using the abovementioned optimized weight function.

The following testing functions are defined to compare the performance of two methods, one is the proposed method of this study while the other is the HDMR method with optimized weight through the fluctuationlessness approximation method.

$$\begin{aligned}
 f_2(x) &= 1 - x^2 \\
 f_3(x) &= (1 + x)^2 e^x \\
 f_4(x) &= \sin(x) \\
 f_5(x) &= \sqrt{1 + x} \\
 f_6(x) &= \ln\left(1 + \frac{x}{2}\right) \tag{65}
 \end{aligned}$$

To make the comparison between the abovementioned two different methods, the value of n should be the same, that is, n value is taken as 9.

Table 2 contains the constancy measurer values obtained through a constant weight as unoptimized weight and two different optimized weights one comes from the method given in Tunga and Demiralp [20], and the other comes from our new method.

The values of Table 2 show us that the optimized weight obtained through the proposed method of this work has the best performance.

7 Conclusion

Dealing with multivariate functions in scientific and engineering problems is extremely difficult. To overcome the difficulties, one way is to decompose the given multivariate problem into a number of less variate problems. This decomposition stands for divide-and-conquer philosophy and there exist various methods based on this philosophy. One of them is the High Dimensional Model Representation method. This method is given through a finite expansion in which a given multivariate function is rewritten by summation of a constant term, a number of univariate terms, another number of bivariate terms and so on. That is, a given multivariate function having N independent variables is represented by the summation of 2^N less variate functions which means that an exact representation of the given multivariate function is obtained when all components are used from the HDMR expansion.

The main disadvantage of the method is to have a huge number of components to be determined while the number of independent variables of the given problem increases to big values. To avoid an increase in computational and mathematical complexities we generally truncate the HDMR expansion at some level, say at the level of bivariate terms, and obtain an approximation to the given analytical structure. This results in a new problem that we should obtain acceptable approximations to represent the multivariate structure successfully. Investigating the HDMR algorithm carefully gives us the idea that optimizing the weight of the method provides to obtain better approximations. HDMR's weight optimization process was first proposed by the authors in another work [19] by using only the fluctuationlessness approximation theory. This work proposes a new method for weight optimization in HDMR to increase the performance of the approximations obtained through HDMR by using, this time, perturbation expansion method with the fluctuationlessness approximation theory. Numerical implementations show us that the weight optimization through the method offered in this manuscript gives the best approximations to the given problems when we compare our new method with an unoptimized version and the one that is optimized through the mentioned previous method which includes only the fluctuationlessness approximation theory. That is, the method proposed in this work takes the weight optimization philosophy a step forward and provides a more flexible algorithm such that when higher order perturbation expansions are used, better approximations can be obtained while at most the second order perturbation expansion is used here.

References

1. I.M. Sobol, Sensitivity estimates for nonlinear mathematical models. *Math. Model. Comput. Exp. (MMCE)* **1**, 407–414 (1993)
2. R. Chowdhury, B.N. Rao, Hybrid high dimensional model representation for reliability analysis. *Comput. Methods Appl. Mech. Eng.* **198**, 753–765 (2009)
3. M.C. Gomez, V. Tchijov, F. Leon, A. Aguilar, A tool to improve the execution time of air quality models. *Environ. Model. Softw.* **23**, 27–34 (2008)

4. I. Banerjee, M.G. Ierapetritou, Design optimization under parameter uncertainty for general black-box models. *Ind. Eng. Chem. Res.* **41**, 6687–6697 (2002)
5. B.N. Rao, R. Chowdhury, Probabilistic analysis using high dimensional model representation and fast fourier transform. *Int. J. Comput. Methods Eng. Sci. Mech.* **9**, 342–357 (2008)
6. T. Ziehn, A.S. Tomlin, A global sensitivity study of sulfur chemistry in a premixed methane flame model using HDMR. *Int. J. Chem. Kinet.* **40**, 742–753 (2008)
7. H. Rabitz, Ö. Alış, General foundations of high dimensional model representations. *J. Math. Chem.* **25**, 197–233 (1999)
8. Ö. Alış, H. Rabitz, Efficient implementation of high dimensional model representations. *J. Math. Chem.* **29**, 127–142 (2001)
9. G. Li, C. Rosenthal, H. Rabitz, High dimensional model representations. *J. Phys. Chem. A* **105**, 7765–7777 (2001)
10. M. Demiralp, High dimensional model representation and its application varieties. *Math. Res.* **9**, 146–159 (2003)
11. B. Tunga, M. Demiralp, Hybrid high dimensional model representation approximants and their utilization in applications. *Math. Res.* **9**, 438–446 (2003)
12. M.A. Tunga, M. Demiralp, Hybrid high dimensional model representation (HHDMR) on the partitioned data. *J. Comput. Appl. Math.* **185**, 107–132 (2006)
13. T. Ziehn, A.S. Tomlin, GUI-HDMR—a software tool for global sensitivity analysis of complex models. *Environ. Model. Softw.* **24**, 775–785 (2009)
14. J. Sridharan, T. Chen, Modeling multiple input switching of CMOS gates in DSM technology using HDMR. *Proc. Design Autom. Test Eur.* **1–3**, 624–629 (2006)
15. I. Banerjee, M.G. Ierapetritou, Parametric process synthesis for general nonlinear models. *Comput. Chem. Eng.* **27**, 1499–1512 (2003)
16. V.U. Unnikrishnan, A.M. Prasad, B.N. Rao, Development of fragility curves using high-dimensional model representation. *Earthq. Eng. Struct. Dyn.* **42**, 419–430 (2013)
17. A.H. Nayfeh, *Perturbation Methods* (Wiley, New York, 1973)
18. J. Kevorkian, J.D. Cole, *Perturbation Methods in Applied Mathematics* (Springer, New York, 1981)
19. B. Tunga, M. Demiralp, Constancy maximization based weight optimization in high dimensional model representation. *Numer. Algorithms* **52**, 435–459 (2009)
20. B. Tunga, M. Demiralp, Constancy maximization based weight optimization in high dimensional model representation for multivariate functions. *J. Math. Chem.* **49**, 1996–2012 (2011)
21. M. Demiralp, A new fluctuation expansion based method for the univariate numerical integration under gaussian weights. in *8-th International WSEAS Conference on Applied Mathematics*, 16–18 December 2005, Tenerife, Spain, pp. 68–73 (2005)
22. K. Hoffman, R. Kunze, *Linear Algebra* (Prentice Hall, New Jersey, 1971)
23. W. Oevel, F. Postel, S. Wehmeier, J. Gerhard, *The MuPAD Tutorial* (Springer, Berlin, 2000)