SAND2018-2510J

# Compressed sparse tensor based quadrature for vibrational quantum mechanics integrals

P. Rai, K.Sargsyan and H. Najm

*Sandia National Laboratories, Livermore CA, 94551*

**Abstract**

A new method for fast evaluation of high dimensional integrals arising in quantum mechanics is proposed. The method is based on sparse approximation of a high dimensional function followed by a low-rank compression. In the first step, we interpret the high dimensional integrand as a tensor in a suitable tensor product space and determine its entries by a compressed sensing based algorithm using only a few function evaluations. Secondly, we implement a rank reduction strategy to compress this tensor in a suitable low-rank tensor format using standard tensor compression tools. This allows representing a high dimensional integrand function as a *small* sum of products of low dimensional functions. Finally, a low dimensional Gauss-Hermite quadrature rule is used to integrate this low-rank representation, thus alleviating the curse of dimensionality. Numerical tests on synthetic functions, as well as on energy correction integrals for water and formaldehyde molecules demonstrate the efficiency of this method using very few function evaluations as compared to other integration strategies.

*Keywords:* Low rank, High dimensional, Tensor approximation, Quantum chemistry, Sparse approximation, Quadrature

## 1. Introduction

Electronic structure calculations have been developed as a powerful tool that is used in several fields including chemical sciences and biochemistry, as well as material and energy sciences. In *ab initio* electronic structure calculations, for instance, computation

[1]Corresponding author: pmrai@sandia.gov

(and often storage) of six-dimensional integrals involving two-body (i.e., Coulomb) interactions ("two-electron integrals") is necessary, creating a severe bottleneck of such calculations for larger molecules [1]. In quantum dynamics, accurate estimation of energy and vibrational frequencies of molecules requires integration of functions whose dimensionality increases linearly with $n$, the number of nuclei in the molecule. For example, the potential energy surfaces (PESs) are often a part of the integrand, and their dimensionality increases as $(3n-6)$. Thus efficient ways to approximate and integrate high dimensional functions that exploit their special structure, if it exists, are needed.

Numerical approximation or integration of these PESs, in practice, can be carried out with sampling techniques with the function as a black-box. For example, one probes the PES at different configuration of atoms in a molecule with standard quantum chemistry software packages such as NWChem [2]. Many methods to represent a PES using a set of *ab initio* data points exist. Some of the black-box fitting methods include splines [3, 4], modified-Shepard interpolation [5], interpolating moving least squares [6, 7], neural networks [8, 9, 10, 11, 12, 13], and reproducing kernel Hilbert space [14]. These methods, although efficient with PES approximation of smaller systems, do not necessarily scale well with system size and may suffer from severe computational constraints. Also, the functional form of the approximation may not render it in a way that is easy to integrate or employ further in a given computational pipeline.

Methods for accurate PES approximation of bigger molecules with relatively few samples are needed. Mathematically, for high dimensional functions, application of standard approximation approaches are often not sufficient due to the curse of dimensionality, i.e. when the underlying approximation space increases exponentially with dimension. One usually employs a class of methods that exploit specific structures of high dimensional functions, such as smoothness or sparsity (see chapter 1 of [15] for a brief survey). In this work, we first exploit the sparsity property [16, 17] of the PES. Sparse approximation methods rely on the fact that a good approximation of the function can be obtained by considering a linear combination of a finite subset of basis functions in a well-constructed dictionary [18, 19]. Here, the basis set consists of tensor products of orthogonal polynomials, and the dictionary is a subset with a constraint imposed on the total order of these multivariate polynomials. The approximation obtained can then

be interpreted as a *sparse tensor*. We then exploit the *low rank structure* [20, 21, 22] in a consequent step by applying a rank reduction strategy on the sparse tensor. Here, a high rank representation of a function is compressed as a *small* sum of products of low dimensional functions [23, 24]. In this paper, we are concerned with integrands that model physical phenomena (see section 2), which commonly involve a limited degree of nonlinear internal coupling, and thereby admit sparsity and/or low rank structures.

Finally, integrals of these low dimensional functions are evaluated using a Gauss-Hermite quadrature rule. The proposed approach thus presents a synthesis of three ideas: sparse approximation, low rank compression and quadrature for the resolution of an integration problem in quantum chemistry. This approach is efficient, and has a strong potential for scalability in order to estimate the energy and frequencies of large molecules. Note that methods that explicitly make use of low rank structure of the PES have recently been proposed [25, 26].

The outline of the paper is as follows. In Section 2, we recall formulations in quantum chemistry that lead to first and second order corrections to the zero point energy. In this section, we also outline the application of the proposed approach to obtain quantum chemistry integrals. In Section 3, we discuss the tensor interpretation of functions, and its approximation using least-squares with sparse regularization, followed by low rank compression of sparse tensors. In Section 4, we first illustrate the application of this approach on standard benchmark integration problems, followed by its application to the quantum chemistry integration problem motivated in Section 2. Finally, we conclude with perspectives in Section 5.

## 2. Motivation: Quantum Chemistry Integration Problem

In this section, we motivate the utility of the compressed sparse tensor based approximation of potential energy surfaces (PES). The PES is required in high-dimensional integrals that are evaluated in quantum chemistry for estimating anharmonic zero-point energies [27, 28]. We then outline our approach of separated integration for this quantum chemistry application.

### 2.1. First and second order energy corrections in quantum chemistry

Let us first list integrals in the XVH2 formalism of quantum chemistry. The reader is referred to the original papers [27, 28] for the derivations of this formalism.

The first-order perturbation correction, $I^{(1)}$, to the zero-point energy is given by

$$I^{(1)} = \int_{-\infty}^{+\infty} e(\boldsymbol{x}) P^{(1)}(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} \tag{1}$$

with

$$e(\boldsymbol{x}) = \prod_{i=1}^{m} e^{-\omega_i x_i^2}, \tag{2}$$

$$P^{(1)}(\boldsymbol{x}) = \Delta V(\boldsymbol{x}), \tag{3}$$

where $\boldsymbol{x} = \{x_1, \ldots, x_m\}$ is the $m$-dimensional set of normal coordinates, $\omega_i$ is the $i$th harmonic frequency. The fluctuation potential $\Delta V(\boldsymbol{x})$ is given by

$$\Delta V(\boldsymbol{x}) = V(\boldsymbol{x}) - V_{\mathrm{ref}} - \frac{1}{2}\sum_{i=1}^{m} \omega_i^2 x_i^2, \tag{4}$$

where $V(\boldsymbol{x})$ is the $m$-dimensional PES and $V_{\mathrm{ref}}$ is its value at the equilibrium geometry, which is the electronic energy at the equilibrium geometry of the molecule. The dimension of the integrand in (1) is $m = 3n-6$, i.e. it depends linearly on the number of atoms ($n$) in the molecule.

The second order correction to energy involves $2m$-dimensional integrals of the form,

$$I^{(2)} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e(\boldsymbol{x}, \boldsymbol{x}') P^{(2)}(\boldsymbol{x}, \boldsymbol{x}') \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{x}' \tag{5}$$

with

$$e(\boldsymbol{x}, \boldsymbol{x}') = \prod_{i=1}^{m} e^{-\omega_i (x_i^2 + x_i'^2)}, \tag{6}$$

$$P^{(2)}(\boldsymbol{x}, \boldsymbol{x}') = \Delta V(\boldsymbol{x}) \Delta V(\boldsymbol{x}') G(\boldsymbol{x}, \boldsymbol{x}'), \tag{7}$$

where $G(\boldsymbol{x}, \boldsymbol{x}')$ is a real-space Green's function given by

$$G(\boldsymbol{x}, \boldsymbol{x}') = \sum_{\substack{n_1=0 \\ (n_1,n_2,\ldots,n_m)\neq(0,0,\ldots,0)}}^{n_{\max}} \cdots \sum_{n_m=0}^{n_{\max}} \prod_{i=1}^{m} \frac{N_{n_i}^2 h_{n_i}(\omega_i^{1/2} x_i) h_{n_i}(\omega_i^{1/2} x_i')}{-\sum_{i=1}^{m} n_i \omega_i}. \tag{8}$$

Here, $N_{n_i}$ is a normalization coefficient, $h_{n_i}$ is the Hermite polynomial[2] of degree $n_i$, and $n_{max} \in \mathbb{N}_0$ is the highest quantum number for each quantum mode indicated here using $i$. Thus, the integrand factor $P^{(2)}$ in (7) is a polynomial in $\boldsymbol{x}$ and $\boldsymbol{x}'$. It is clear that the dimensions of integrands in (1) and (5) grow linearly with the number of atoms in a molecule. With this integration problem at hand, we detail our approach of separated integration for both $I^{(1)}$ and $I^{(2)}$, in the following subsection.

### 2.2. Separated integration

Consider an integration problem of a high dimensional function $u(\boldsymbol{x}), \boldsymbol{x} = (x_1, \ldots, x_m)$. That is, we would like to estimate the quantity

$$I[u] = \int_{-\infty}^{+\infty} u(\boldsymbol{x})\rho(\boldsymbol{x})\mathrm{d}\boldsymbol{x}, \tag{9}$$

where $\rho(\boldsymbol{x})$ is a non-negative weight function that is integrable and multiplicatively separable, i.e.,

$$\rho(\boldsymbol{x}) \geq 0, \tag{10}$$

$$\int_{-\infty}^{+\infty} \rho(\boldsymbol{x})\mathrm{d}\boldsymbol{x} < \infty, \tag{11}$$

and

$$\rho(\boldsymbol{x}) = \prod_{i=1}^{m} \rho^{(i)}(x_i). \tag{12}$$

In order to estimate $I[u]$, we first search for an accurate approximation of the form

$$u(\boldsymbol{x}) \approx \tilde{u}(\boldsymbol{x}) = \sum_{k=1}^{r} \prod_{i=1}^{m} w_k^{(i)}(x_i), \tag{13}$$

where $r$ is the so-called separation rank, which should be as small as possible, $w_k^{(i)}(x_i)$ is the $k$th univariate function in the $i$th dimension, to be determined (see the next section). Under the assumption that $u(\boldsymbol{x})$ admits such a representation with sufficiently small $r$,

[2] To be precise, we refer to the physicists' Hermite polynomials, where $h_n(x) = (-1)^n e^{x^2} \frac{\mathrm{d}^n}{\mathrm{d}x^n} e^{-x^2}$.

we have

$$I[u] \approx \sum_{k=1}^{r} \prod_{i=1}^{m} \int_{-\infty}^{+\infty} w_k^{(i)}(x_i)\rho_i(x_i)\,\mathrm{d}x_i, \tag{14}$$

which can be evaluated by quadrature at a computational cost of $O(rmp)$, where $p$ is the degree of univariate polynomial functions $w_k^{(i)}(x_i)$, increasing only linearly with dimension. We now apply this strategy for $I^{(1)}$ and $I^{(2)}$ from Eqs. (1) and (5), respectively.

In $I^{(1)}$, the weight function $e(\boldsymbol{x})$ is a separable Gaussian, and $P^{(1)}(\boldsymbol{x})$ has the factor of $\Delta V(\boldsymbol{x})$ which can be expressed in a low-rank format

$$\Delta V(\boldsymbol{x}) \approx \sum_{k=1}^{r_1} \prod_{i=1}^{m} \Delta V_k^{(i)}(x_i) \tag{15}$$

with a separation rank $r_1$. Thus $I^{(1)}$ can be evaluated as a sum-of-products of one-dimensional integrals,

$$I^{(1)} \approx \sum_{k=1}^{r_1} \prod_{i=1}^{m} \int_{-\infty}^{+\infty} e^{-\omega_i x_i^2} \Delta V_k^{(i)}(x_i)\,\mathrm{d}x_i, \tag{16}$$

using Gauss-Hermite quadrature at a cost $O(r_1 mp)$ that scales linearly with the dimension $m$.

In $I^{(2)}$, $G(\boldsymbol{x}, \boldsymbol{x}')$ appears as an integrand factor which can also be represented with a low rank approximation

$$G(\boldsymbol{x}, \boldsymbol{x}') \approx \sum_{k=1}^{r_2} \prod_{i=1}^{m} G_k^{(i)}(x_i, x_i'), \tag{17}$$

with a separation rank $r_2$. Substituting (15) and (17) in (5), $I^{(2)}$ can be evaluated as a sum-of-products of two-dimensional integrals,

$$I^{(2)} \approx \sum_{k_1=1}^{r_1} \sum_{k_2=1}^{r_1} \sum_{k_3=1}^{r_2} \prod_{i=1}^{m} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-\omega_i (x_i^2 + x_i'^2)} \Delta V_{k_1}^{(i)}(x_i) \Delta V_{k_2}^{(i)}(x_i') G_{k_3}^{(i)}(x_i, x_i')\,\mathrm{d}x_i\,\mathrm{d}x_i', \tag{18}$$

at a cost $O(r_1^2 r_2 mp)$ that again scales linearly with dimension, albeit with a larger prefactor compared to the cost of $I^{(1)}$.

For accurate, efficient and scalable computation of $I^{(1)}$ and $I^{(2)}$ using separated integration with (16) and (18), we require two conditions to be satisfied. Firstly, in a non

intrusive setting where we can only evaluate $\Delta V(\boldsymbol{x})$ at a limited number of geometrical configurations of the molecule (e.g. using NWChem), low rank approximations in (15) should be sufficiently accurate using as few evaluations of $\Delta V(\boldsymbol{x})$ as possible. Secondly, the separation rank $r_1$ in (15) and $r_2$ in (17) must be small for sufficiently accurate approximation in order to reduce computation time of quadrature integration of the associated one- or two-dimensional integrals. In the following section, we detail our approach that strives to satisfy both of the above conditions.

## 3. Integration Using Low Rank Compression of Sparse Tensor

Methods proposed in the literature that exploit sparsity on a multivariate tensor product basis set for approximation of functions in a black box setting have been very successful [16, 17, 18, 19]. However, in high dimension, the method still suffers from the curse of dimensionality due to exponential growth of basis functions. More recently, methods that exploit low rank structure have been proposed [29, 30, 31] for approximation of high dimensional functions which, in certain cases, alleviate the curse of dimensionality. In [32, 15], authors exploit sparsity within the low rank structure. Low rank methods, however, use heuristic alternating least squares algorithm for optimizing a non-linear objective function thus requiring more samples for an accurate solution. The novelty of this work is that we first exploit sparsity in the function by resolution of a problem formulated as optimization of a linear objective function. Having thus obtained a sparse tensor with known entries, we apply low rank tensor compression resulting in a separated representation with a small rank which is faster to integrate. The proposed 2-step procedure ultimately yields a low rank approximation of the integrand similar to the method proposed in [25] where a low rank approximation is obtained directly from function evaluations. However, the proposed approach is more effective because it first uses a sparse least squares procedure to accurately estimate coefficients of the linear representation of the integrand on a tensor product basis set. Having obtained an accurate tensor representation, it can then be adequately compressed. We thus take advantage of both sparsity and low rank structures more effectively.

In this section, we detail our approach in a general setting. In Section 3.1, we discuss

*a priori* selection of approximation space. Sparse approximation using least-squares in this space is then discussed in Section 3.2 followed by low rank compression of the sparse solution in Section 3.3.

### *3.1. Selection of approximation space*

A multivariate function $u(\boldsymbol{x})$ can be expanded on a multidimensional tensor product basis as

$$u(\boldsymbol{x}) \approx \tilde{u}(\boldsymbol{x}) = \sum_{j_1=0}^{p_1} \cdots \sum_{j_m=0}^{p_m} v_{j_1,\ldots,j_m} \phi_{j_1}^{(1)}(x_1) \cdots \phi_{j_m}^{(m)}(x_m), \tag{19}$$

where $\phi_{j_i}^{(i)}(x_i)$ is the $j_i$th basis function in the $i$th coordinate, $x_i$. Here, the total number of coefficients of the multidimensional basis needed to characterize $\tilde{u}(\boldsymbol{x})$ is given by $P = \prod_{k=1}^{m} p_k$. Let us, for the sake of simplicity, choose the same number of basis, $p$, in each dimension. For a high dimensional function, $m$ is large, and the exponential increase in the number of basis functions $P = p^m$ is referred to as the *curse of dimensionality*. Thus, a least squares based approximation of $u(\boldsymbol{x})$ will require at least $P$ evaluations of $u(\boldsymbol{x})$ which, in general, may not be feasible. Here, we wish to take advantage of the possible sparsity structure of the function on the chosen approximation bases in two steps. First, we reduce the approximation space *a priori* by considering only a subset of multidimensional bases. Then, we obtain a sufficiently accurate approximation, for the purpose of integration, on the reduced approximation space with only a few function evaluations making use of sparse regularization techniques.

A natural choice of basis functions for functional representation are polynomials. The tensor product space of multidimensional polynomials with degree per-dimension $p$ over the domain $\mathcal{X}$ is denoted as

$$\mathbb{Q}_p(\mathcal{X}) = \text{span}\left\{ \prod_{i=1}^{m} \phi_{j_i}^{(i)}(x_i) : \mathbf{j} \in \mathbb{N}_0^m, |\mathbf{j}|_\infty := \max_{i \in \{1\ldots m\}} j_i \leq p \right\}, \tag{20}$$

where $\phi_{j_i}^{(i)}(x_i)$ is a polynomial of degree $j_i$ and $\mathbf{j} = (j_1, \ldots, j_m) \in \mathcal{J} = \times_{i=1}^{m} \{0, \ldots, p\}$ with $\dim(\mathbb{Q}_p(\mathcal{X}))=(p+1)^m$. The notation $\mathbb{N}_0^m$ refers to the set of vectors with non-negative integer elements. Another alternative is the tensor product space of multidimensional

polynomials of total degree $p$ is defined by:

$$\mathbb{P}_p(\mathcal{X}) = \text{span}\left\{\prod_{i=1}^{m} \phi_{j_i}^{(i)}(x_i) : \mathbf{j} \in \mathbb{N}_0^m, |\mathbf{j}|_1 := \sum_{i=1}^{m} j_i \leq p\right\} \tag{21}$$

with $\dim(\mathbb{P}_p(\mathcal{X})) = \tilde{P} = \frac{(m+p)!}{m!p!}$. If $\mathcal{X}$ is a product domain i.e. $\mathcal{X} = \mathcal{X}_1 \times \cdots \times \mathcal{X}_m$, $\mathbb{Q}_p(\mathcal{X})$ is a full tensorization of unidimensional polynomial spaces of degree $p$:

$$\mathbb{Q}_p(\mathcal{X}) = \mathbb{Q}_p(\mathcal{X}_1) \otimes \cdots \otimes \mathbb{Q}_p(\mathcal{X}_m). \tag{22}$$

The space $\mathbb{P}_p(\mathcal{X})$ is a subset of polynomial space $\mathbb{Q}_p(\mathcal{X})$, albeit still grows fast with dimensionality $m$. In this work, the first step in obtaining a sparsef approximation is to select *a priori* the approximation space as $\mathbb{P}_p(\mathcal{X})$ with fewer basis functions as compared to $\mathbb{Q}_p(\mathcal{X})$ for the same degree $p$. Thus we are explicitly choosing the coefficients corresponding to those basis function in $\mathbb{Q}_p(\mathcal{X})$ as zero that are not present in $\mathbb{P}_p(\mathcal{X})$. We can thus write the approximation of $u(\boldsymbol{x})$ as

$$u(\boldsymbol{x}) \approx \tilde{u}(\boldsymbol{x}) = \sum_{\mathbf{j} \in \tilde{\mathcal{J}}} v_{\mathbf{j}} \phi_{\mathbf{j}}(\boldsymbol{x}), \tag{23}$$

where $\mathbf{j} \in \tilde{\mathcal{J}}$ such that $|\mathbf{j}|_1 \leq p$ and $v_{\mathbf{j}}$ is real coefficient on the basis $\phi_{\mathbf{j}} = \prod_{i=1}^{m} \phi_{j_i}^{(i)}$.

### *3.2. Sparse approximation using least-squares*

Having chosen a reduced approximation space, we now take advantage of sparsity, if it exists. Here, we use sparse approximation based on the fact that a sufficiently accurate representation of $u(\boldsymbol{x})$ can be obtained by only considering a finite subset of basis functions $\mathcal{D} = \{\phi_{\mathbf{j}}; \mathbf{j} \in \tilde{\mathcal{J}}\}$:

$$u(\boldsymbol{x}) \approx \tilde{u}_n(\boldsymbol{x}) = \sum_{\mathbf{j} \in \tilde{\mathcal{J}}_n} v_{\mathbf{j}} \phi_{\mathbf{j}}(\boldsymbol{x}),\ \tilde{\mathcal{J}}_n \subset \tilde{\mathcal{J}},\ \#\tilde{\mathcal{J}}_n = n. \tag{24}$$

The notation # is used to represent cardinality of a set. The objective here is to obtain a sparse representation of $u(\boldsymbol{x})$ on a basis set $\{\phi_{\mathbf{j}} : \mathbf{j} \in \tilde{\mathcal{J}}\}$ with few evaluations of $u(\boldsymbol{x})$ on $\mathbf{x}^s, s = 1, \ldots, S$ realizations of $\boldsymbol{x}$ such that we obtain only $n$ non-zero coefficients. In this work, to obtain $\tilde{u}_n(\boldsymbol{x})$, we use least-squares with sparse regularization. The matrix representation of (23) for sample realizations $\{\mathbf{x}^s\}_{s=1}^{S}$ is given by

$$\boldsymbol{\Phi}\mathbf{v} \approx \mathbf{u}, \tag{25}$$

where $\mathbf{u} = (u(\mathbf{x}^s))_{s=1}^S \in \mathbb{R}^S$, $\mathbf{\Phi}(s,:) = (\phi_{\mathbf{j}}(\mathbf{x}^s))$ and $\mathbf{v} = (v_{\mathbf{j}})_{\mathbf{j}\in\tilde{\mathcal{J}}} \in \mathbb{R}^{\tilde{P}}$. The best n-term approximation can be written as

$$\min_{\substack{\mathbf{v}\in\mathbb{R}^{\tilde{P}} \\ \|\mathbf{v}\|_0 = n}} \|\mathbf{\Phi}\mathbf{v} - \mathbf{u}\|_2^2, \tag{26}$$

The above problem is non-convex and we instead solve

$$\min_{\mathbf{v}\in\mathbb{R}^{\tilde{P}}} \|\mathbf{\Phi}\mathbf{v} - \mathbf{u}\|_2^2 + \lambda\|\mathbf{v}\|_1, \tag{27}$$

where $\|\cdot\|_1$ is the $\ell_1$ norm. Problem (27) is a convex optimization problem for which several methods are available [33]. This is the basis of compressed sensing approach that made a breakthrough in signal processing nearly a decade ago [16, 17, 34]. If $u(\boldsymbol{x})$ admits an accurate sparse approximation, then, under some additional conditions, solving (27) gives a sparse solution $\mathbf{v}$. In this work, we use Lasso modified least angle regression (LARS) (see LARS presented in [35]) of SPAMS [36] software for solving (27). For the selection of regularization coefficient, we use fast leave-one-out cross validation proposed in [37].

### *3.3. Low rank compression of sparse solution*

The final step in the developed integration method is the application of quadrature rules for integrating separated functions. The approximation $u_n(\boldsymbol{x})$ is already in a separated form with separation rank $n$. However, even if $n$ is only moderately large, the number of quadrature integrations required can increase rapidly. For example, in the evaluation of $I^{(2)}$ in (18), the number of separated terms grows as $r_1^2 r_2$ with rank. In order to improve computation efficiency, we can reduce the separation rank considerably for a small loss of accuracy. We thus strive to find approximations of the form

$$\tilde{u}_n(\boldsymbol{x}) \approx \tilde{u}_r(\boldsymbol{x}) = \sum_{k=1}^{r\ll n} \alpha_k w_k(\boldsymbol{x}),\ \alpha_k \in \mathbb{R}, \tag{28}$$

where

$$w_k(\boldsymbol{x}) = \prod_{i=1}^{m} w_k^{(i)}(x_i),\ \text{and}\ w_k^{(i)}(x_i) = \sum_{j=1}^{p} v_{k,j}^{(i)} \phi_j^{(i)}(x_i). \tag{29}$$

The univariate functions $w_k^{(i)}(x_i)$, characterized by coefficient vectors $\mathbf{v}_k^{(i)} = (v_{k,1}^{(i)}, \ldots, v_{k,p}^{(i)})$, are represented as expansions on bases $\boldsymbol{\phi}^{(i)} = (\phi_1^{(i)}(x_i), \ldots, \phi_p^{(i)}(x_i))$. Finding low rank

approximations of the form (28) thus reduces to finding coefficient vectors $\mathbf{v}_k^{(i)},\ 1 \le i \le m, 1 \le k \le r$. The scalar coefficient $\alpha_k$ is obtained by normalizing $\mathbf{v}_k^{(i)}, 1 \le i \le m$ such that $\|\mathbf{v}_k^{(i)}\| = 1$. Thus, with this rank reduction procedure, we can greatly reduce the number of terms in the separated representation of the integrand thus reducing the computation time for quadrature based evaluation of the integrals.

Let us denote a tensor $\tilde{\mathsf{U}}_n \in \mathbb{R}^{p\times\cdots\times p}$ whose entries are coefficients $u_{\mathbf{j}}, \mathbf{j} \in \tilde{\mathcal{J}}_n$ in the sparse approximation given by (24). With choice of bases $\phi_j^{(i)}, 1 \le i \le m, 1 \le j \le p$, function $\tilde{u}_n(\boldsymbol{x})$ can be identified with tensor $\tilde{\mathsf{U}}_n$. Finding the low rank approximation $\tilde{u}_r(\boldsymbol{x})$ in (28) then reduces to finding a low rank decomposition of the tensor $\tilde{\mathsf{U}}_n \approx \tilde{\mathsf{U}}_r$ such that

$$\tilde{\mathsf{U}}_r = \sum_{k=1}^{r} \alpha_k \left( \otimes_{i=1}^{m} \mathbf{v}_k^{(i)} \right). \tag{30}$$

The decomposition of the tensor in (30) is known as the canonical polyadic (CP) decomposition for which several tensor decomposition tools are available. In this work, we use the Tensor Toolbox developed by Sandia National Laboratories [38].

We now wish to integrate the low rank function $u_r(\boldsymbol{x})$ with respect to a separable measure $\rho(\boldsymbol{x})$ via quadrature as

$$I[u] \approx I[u_r] = \sum_{k=1}^{r} \alpha_k \left( \prod_{i=1}^{m} \int_{-\infty}^{+\infty} w_k^{(i)}(x_i) \rho^{(i)}(x_i) dx_i \right), \tag{31}$$

Let us denote $\gamma_{q_i}^{(i)}$ and $\mathrm{x}_i^{q_i}, 1 \le q_i \le Q_i$ as one dimensional quadrature weights and quadrature points along dimension $i$ for the measure $\rho^{(i)}(x_i)$. We can evaluate (31) as

$$I[u_r] = \sum_{k=1}^{r} \alpha_k \left( \prod_{i=1}^{m} \sum_{q_i=1}^{Q_i} \gamma_{q_i}^{(i)} w_k^{(i)}(\mathrm{x}_i^{q_i}) \right). \tag{32}$$

Thus we integrate our function using $rm$ one dimensional integrals each of which is evaluated using quadrature.

## 4. Applications

In Subsection 4.1, we first illustrate the proposed method on two Genz functions, benchmark functions for which analytical integrals are known [39]. Next, in Section 4.2, we apply our approach on the quantum chemistry problem motivated in Section 2.

We estimate two distinct error components at different stages of approximation. We consider a separate test sample set with 10000 evaluations of the function we wish to integrate to determine the accuracy of the sparse approximation. The relative approximation error of the sparse solution is defined as

$$\epsilon_s = \frac{\|\mathbf{u} - \tilde{\mathbf{u}}_n\|_2}{\|\mathbf{u}\|_2} = \frac{\sqrt{\sum_{s=1}^{10000}(u(\mathbf{x}^s) - \tilde{u}_n(\mathbf{x}^s))^2}}{\sqrt{\sum_{s=1}^{10000}(u(\mathbf{x}^s))^2}}, \tag{33}$$

where $\mathbf{u}$ is a vector containing evaluations of $u(\boldsymbol{x})$ and $\tilde{\mathbf{u}}_n$ is a vector of evaluations of $\tilde{u}_n(\boldsymbol{x})$ at sample points $\mathbf{x}^s$ in the test set.

Further, we define the error due to low rank compression as

$$\epsilon_c = \frac{\|\tilde{\mathbf{u}}_n - \tilde{\mathbf{u}}_r\|_2}{\|\tilde{\mathbf{u}}_n\|_2} = \frac{\sqrt{\sum_{s=1}^{10000}(\tilde{u}_n(\mathbf{x}^s) - \tilde{u}_r(\mathbf{x}^s))^2}}{\sqrt{\sum_{s=1}^{10000}(\tilde{u}_n(\mathbf{x}^s))^2}}, \tag{34}$$

where $\tilde{\mathbf{u}}_r$ is a vector containing evaluations of the low rank function $\tilde{u}_r(\boldsymbol{x})$ (defined in Eq. (28)) at test sample points. We also compare our approach with a direct approximation in canonical low rank tensor format (a greedy and non greedy algorithm with $\ell_1$ and $\ell_2$ regularization in alternating minimization scheme was proposed in [15] and [40] respectively). A comparison with this method is pertinent because of two reasons. Firstly, this approach gives a low rank approximation which is an essential ingredient for its application in the quantum chemistry problem discussed in Section 2 (this also motivates our choice of canonical tensor format over tree based format which have multilinear ranks). Secondly, it can potentially be applied for approximation of very high dimensional functions which is needed indeed for approximating PES of bigger molecules.

For the sake of comparison, we consider $\epsilon_s$ of direct low rank approximation evaluated on the same test sample set. Note that $\epsilon_c$ in direct approximation in canonical tensor format cannot be evaluated as the approximant is already in a compressed form.

Table 1: Genz functions and formulae for evaluating analytical integration. In the integral of the Guassian Genz function, $erf(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$.

| Genz Function | Formula: $u_{\boldsymbol{\beta},\mathbf{a}}(\boldsymbol{x})$ | $\int_0^1 u_{\boldsymbol{\beta},\mathbf{a}}(\boldsymbol{x}) d\boldsymbol{x}$ |
|---|---|---|
| Oscillatory | $\cos\left(2\pi a_1 + \sum_{i=1}^d \beta_i x_i\right)$ | $\cos\left(2\pi a_1 + \sum_{i=1}^d \beta_i\right) \prod_{i=1}^d \frac{\sin(\beta_i/2)}{\beta_i}$ |
| Gaussian | $\exp\left(-\sum_{i=1}^d \beta_i^2 (x_i - a_i)^2\right)$ | $\prod_{i=1}^d \frac{\sqrt{\pi}}{2\beta_i} (erf(\beta_i(1-a_i)) + erf(\beta_i a_i))$ |

To take into account the effect of $\epsilon_c$, let us consider the following identity

$$\begin{aligned} ||\mathbf{u} - \tilde{\mathbf{u}}_r||_2 &\leq ||\mathbf{u} - \tilde{\mathbf{u}}_n||_2 + ||\tilde{\mathbf{u}}_n - \tilde{\mathbf{u}}_r||_2 && (35) \\ \frac{||\mathbf{u} - \tilde{\mathbf{u}}_r||_2}{||\mathbf{u}||_2} &\leq \frac{||\mathbf{u} - \tilde{\mathbf{u}}_n||_2}{||\mathbf{u}||_2} + \frac{||\tilde{\mathbf{u}}_n - \tilde{\mathbf{u}}_r||_2}{||\tilde{\mathbf{u}}_n||_2} \frac{||\tilde{\mathbf{u}}_n||_2}{||\mathbf{u}||_2} && (36) \\ &\leq \epsilon_s + \epsilon_c \frac{||\tilde{\mathbf{u}}_n||_2}{||\mathbf{u}||_2} && (37) \end{aligned}$$

In (37), if $\epsilon_c$ is small as compared to $\epsilon_s$ and $\frac{||\tilde{\mathbf{u}}_n||_2}{||\mathbf{u}||_2} \approx 1$, then one can compare approximation error in low rank canonical tensor format with $\epsilon_s$. We also note that, in many applications, a cross validation based error estimator is required if the computational budget is limited, instead of a separate set of test samples.

In order to take variation due to sampling effects into account, for each sample size considered in the plotted results for obtaining a sparse low rank approximation, we illustrate $\epsilon_s$ and $\epsilon_c$ by connecting the median error with each error bar indicating the 25/75 quantiles obtained from 51 independent sample sets. Points in each set are sampled according to the distribution of input parameters indicated in the examples below.

### *4.1. Genz Functions*

Table 1 defines Oscillatory and Gaussian Genz functions considered in numerical tests in this section. For each sample set, weight ($\beta_i$) and shift ($a_i$) parameters are taken from standard uniform random distribution over $[-1, 1]$. The exact integration is analytically available and is used as a reference for comparison.

The input variables $\boldsymbol{x}$ of the Genz functions are uniformly sampled over the range $[0, 1]$. For both Genz functions, in order to illustrate the effect of dimensionality, we

consider two cases: $d = 5$ and $d = 10$. Also, we choose the sparse solution (see Section 3.2) with minimum $\epsilon_s$ by considering approximation in the space of multidimensional orthogonal Legendre polynomials $\mathbb{P}_6([-1,1]^d)$ (samples of input variables $\boldsymbol{x}$ are scaled to [-1,1] to take advantage of orthogonality of Legendre polynomials over [-1,1] with uniform measure).

Figure 1 shows evolution of $\epsilon_s$ with sample size. We find that this error decays monotonically with sample size. For both functions, if we keep an accuracy threshold of $\epsilon_s = 1.0 \times 10^{-2}$, we need $\approx 100$ samples for $d = 5$ and $\approx 1000$ for $d = 10$. Also, since the optimal sparse solution for the sample sets is obtained in $\mathbb{P}_6([-1,1]^d)$, we have $\dim(\mathbb{P}_6([-1,1]^{d=5})) = 462$ and $\dim(\mathbb{P}_6([-1,1]^{d=10})) = 8008$. This indicates that a sufficiently accurate sparse solution is obtained even when we under-sample the function by a factor of $\frac{462}{100} \approx 4.6$ for $d = 5$ and $\frac{8008}{1000} \approx 8$ for $d = 10$. Next, we observe that, for both Genz functions, sparse approximation in $\mathbb{P}_6([-1,1]^d)$ is more accurate than direct approximation in low rank canonical tensor format (here, the optimal rank$\leq 20$ is selected using cross validation) for the same sample size.

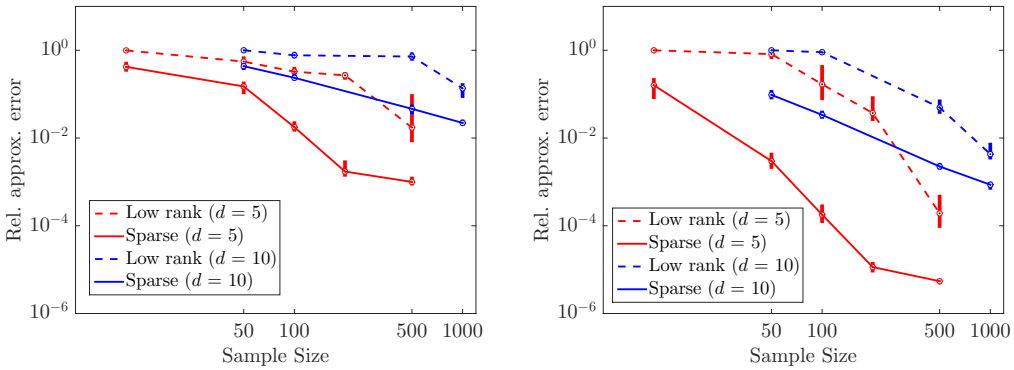


Figure 1: The relative approximation error $\epsilon_s$ in sparse reconstruction of Genz Oscillatory (left) and Genz Gaussian (right) function versus number of samples. The plot connects the medians, with each error bar indicating the 25/75 quantiles obtained from 51 independent ensembles.

Figure 2 plots the number of non-zero coefficients (i.e. the $\ell_0$ norm, $\|\cdot\|_0$, of the coefficient vector), which is also the separation rank before compression, of $\tilde{\mathbf{u}}_n$ with sample size. For the sake of comparison across dimension $d$, let us keep the same

accuracy threshold of $\epsilon_c = 1.0 \times 10^{-2}$. We find that, for both Genz functions, the number of non-zero coefficients is $\approx 100$ for $d = 5$ with a proportional increase to $\approx 200$ for $d = 10$.

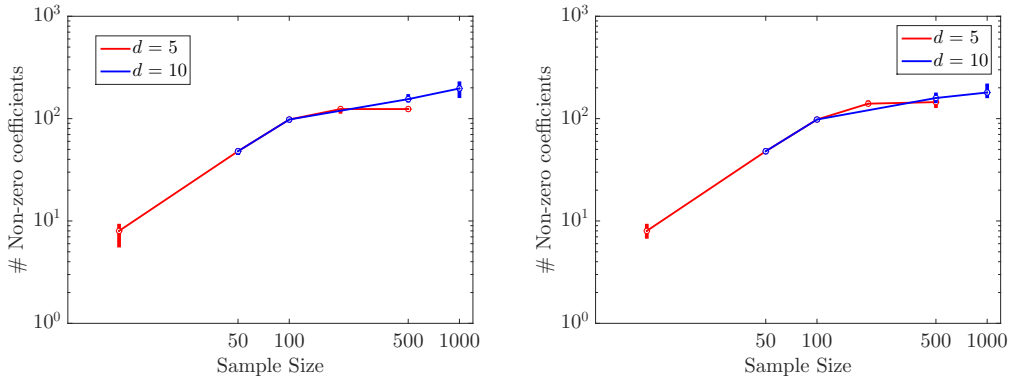


Figure 2: The number of non-zero coefficients in sparse reconstruction of Genz Oscillatory (left) and Genz Gaussian (right) functions as the number of samples ($S$) increases. The plot connects the medians, with each error bar indicating the 25/75 quantiles obtained from 51 independent ensembles.

The results for the next level of approximation, corresponding to the low rank compression of optimal sparse approximants of Genz functions, is shown in Figure 3. Here, we plot evolution of $\epsilon_c$ with rank of the compressed function $\tilde{\mathbf{u}}_r$. As before, we compress $\tilde{\mathbf{u}}_n$ obtained with 100 samples for $d = 5$ and 1000 samples with $d = 10$. In this case, we target a compression $O(10)$ i.e. $\text{rank}(\tilde{\mathbf{u}}_r) \approx 10$. We find that, for oscillatory Genz function, we obtain $\epsilon_c = 1.0 \times 10^{-4}$ for $d = 5$ and $\epsilon_c = 1.0 \times 10^{-2}$ for $d = 10$ thus admitting not only sparse, but also a low rank representation. Similarly, for Gaussian Genz function, for both $d = 5, 10$, we get a compression error of $\epsilon_c = 1.0 \times 10^{-3}$. Thus, in both Genz functions, $\epsilon_c$ is an order of magnitude smaller than $\epsilon_s$. For 10 dimensional Oscillatory Genz function, $\epsilon_s$ and $\epsilon_c$ are similar in magnitude for rank = 10. However, this can easily be adjusted by accepting a slightly smaller decrease in low rank compression by increasing the rank. Note that since $\epsilon_c$ is a relative error w.r.t the sparse solution $\tilde{\mathbf{u}}_n$ (and not w.r.t the true function), in both cases, the gain due to compression by a factor of $\approx 10$ for a small $\epsilon_c$ is considerable for improving computational efficiency in evaluating integrals using quadrature in the next step.

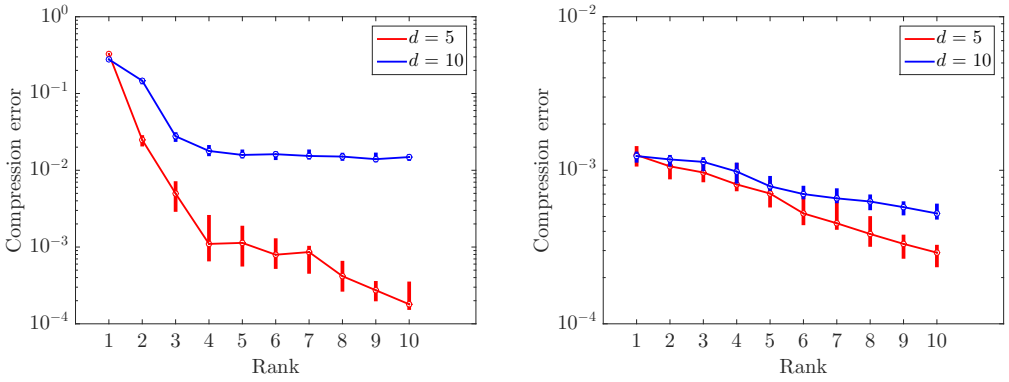


Figure 3: The compression error $\epsilon_c$ of sparse approximants of Genz Oscillatory (left) and Genz Gaussian (right) functions versus rank. The plot connects the medians, with each error bar indicating the 25/75 quantiles obtained from 51 independent ensembles.

Finally, Table 2 shows the relative error in the integration value obtained by applying Gauss Legendre quadrature rule on the compressed low rank form $\tilde{\mathbf{u}}_r$. The integration error shown is the mean value, ± standard deviation, obtained from 51 different sample sets. The number of samples indicated are the minimum sample size of Genz functions from an *a priori* set of sample sizes (indicated in Figure 1) for an approximation accuracy of $\epsilon_s = 1.0 \times 10^{-2}$. Similarly, for both functions, we consider the compressed function $\tilde{\mathbf{u}}_r$ with rank =10. For integration using quadrature (see (32)), the number of one dimensional Gauss Legendre quadrature points is $Q_i = (p+1)/2$, where $p$ is the degree of polynomial integrand. We find that, for both functions, an accurate estimate of the integral is obtained with the proposed method. Since we consider applications where the quantity of interest is the expectation of a function, we find that even with a relatively less accurate sparse approximation (error measured w.r.t point-wise accuracy) in a reduced approximation space with limited samples, we get a more accurate estimate of the integral. This is also the case, as will be shown in the next section, where we apply this approach on the quantum chemistry application proposed in Section 2.

### *4.2. Quantum chemistry problem*

In this section, we discuss the application of separated integration on the quantum chemistry problem described in Section 2. We remark that, unlike Genz functions

Table 2: Integration (relative) error of Genz Functions. The error values show mean ± standard deviation out of 51 replica simulations. The number of samples corresponds to a sparse approximation accuracy of $\epsilon_s = 1.0 \times 10^{-2}$.

| Dimension | # Samples | Rank of $\tilde{\mathbf{u}}_r$ | Oscillatory | Gaussian |
|---|---|---|---|---|
| 5 | 100 | 10 | $(2.68 \pm 1.6) \times 10^{-4}$ | $(1.12 \pm 0.9) \times 10^{-4}$ |
| 10 | 1000 | 10 | $(7.9 \pm 3) \times 10^{-4}$ | $(2.68 \pm 1.4) \times 10^{-5}$ |

which are defined on hypercubes, limits of quantum chemistry integrals $I^{(1)}$ and $I^{(2)}$ are $[-\infty, +\infty]$. The separated integration approach applies in both cases because of two reasons. Firstly, the measure of integrals $I^{(1)}$ and $I^{(2)}$ is Gaussian, which indicates an exponential decay in the probability of sampling points away from the mean. Secondly, for each molecule, we sample geometry coordinates $\boldsymbol{x}$ according to the standard normal distribution (zero mean and unit standard deviation) and scale them with the corresponding harmonic frequencies (obtained e.g. from a standard calculation in NWChem). This scaling, also employed in [27], captures samples in the domain of integrands (without privileging any particular coordinate) where approximation with better accuracy is desired for accurate estimation of integrals. We consider two molecules, water, $H_2O$ ($m = 3$), and formaldehyde, $H_2CO$ ($m = 6$). We choose total degree polynomial spaces of orthogonal Hermite basis functions with degree $p = 3$ and 4. Figure 4 plots the relative error of the sparse approximation of the PES, $\epsilon_s$, as a function of sample size. We also show error obtained using direct approximation in canonical low rank tensor format of rank $\leq 30$.

We first note that, with increase in sample size, $\epsilon_s$ is reduced dramatically in the approximation space with $p = 4$ for both water and formaldehyde. This is consistent with the established observation that a quartic force field (i.e. fourth order Taylor expansion) gives an accurate representation of the PES for these two molecules [41]. We also find that, for an accuracy of $1.0 \times 10^{-2}$, the proposed approach requires 40 samples for water molecule as compared to 150 samples needed for a direct approximation in canonical low rank format. Thus the sampling cost reduces by a factor of $\frac{150}{40} \approx 4.5$. In case of formaldehyde molecule, for the same accuracy of $1.0 \times 10^{-2}$, the cost reduction is by a

factor of $\frac{4000}{180} \approx 22.2$.

The number of basis functions in the reduced approximation space with $p = 4$ for water and formaldehyde are 35 and 210, respectively. We can consider the PES of these molecules to have a sparse representation in the reduced approximation space if a sufficiently accurate approximation (of the order $10^{-2}$) can be obtained with fewer samples as compared to the number of basis functions. In the case of water, we need approximately 35 PES evaluations to obtain an accurate representation. Here, the sparsity is achieved only with the *a priori* selection of the approximation space, and the PES does not admit a sparse solution on the reduced approximation space. However, in the case of formaldehyde, $\approx 180$ samples are sufficient for an accurate approximation thus taking advantage of sparsity structure on the reduced basis set.

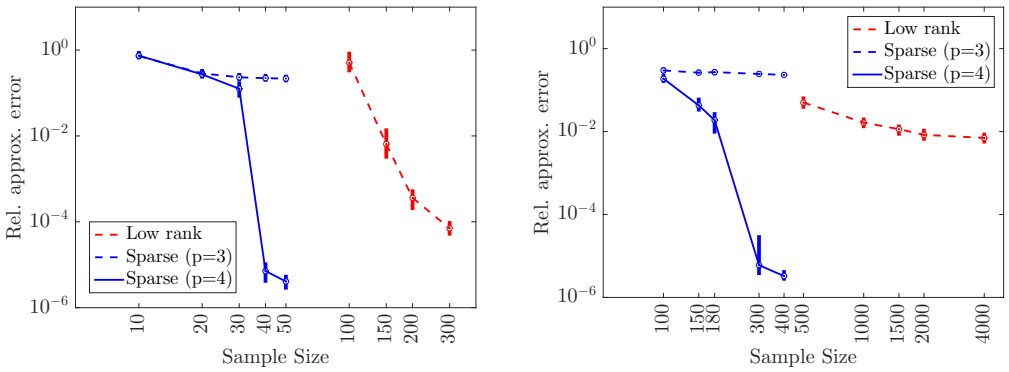


Figure 4: The relative approximation error in sparse PES of (a) water and (b) formaldehyde as a function of the Sample size. The plot connects the medians, with each error bar indicating the 25/75 quantiles obtained from 51 independent ensembles.

We now illustrate the application of low rank compression of the sparse PES in Figure 5, where we plot the compression error versus separation rank for water and formaldehyde. We find that $\epsilon_c$ decreases as separation rank $r$ is increased for both molecules. Considering that the number of non-zero coefficients in the sparse solution is equal to the number of basis functions in the reduced approximation space, low rank compression is able to achieve a reduction in separation rank of 35/7 and 210/30 for water and formaldehyde respectively for an error threshold of $\epsilon_c = 1.0 \times 10^{-2}$. Similarly,

the reduction in separation rank for an error threshold of $\epsilon_c = 1.0 \times 10^{-3}$ is 35/10 and 210/70 for a small loss of accuracy. Depending on the available computational budget and required accuracy, this procedure provides flexibility in truncating the separation rank of the PES, thus providing more flexibility in estimating energy corrections.

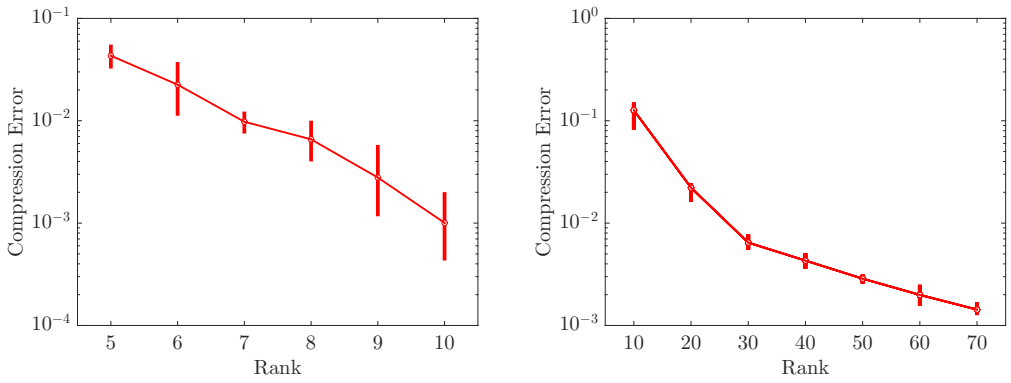


Figure 5: The compression error in low rank PES of (a) water and (b) formaldehyde as a function of separation rank. The plot connects the medians, with each error bar indicating the 25/75 quantiles obtained from 51 independent ensembles.

Figure 6 shows the compression error and the compression ratio (see below) of $G(\boldsymbol{x}, \boldsymbol{x}')$ in (8) as a function of the separation rank $r$ for water and formaldehyde. The compression ratio is the ratio of the number of parameters in the low-rank-decomposed (i.e., compressed) representation of $G(\boldsymbol{x}, \boldsymbol{x}')$ to the total number of parameters in the original representation:

$$\gamma = \frac{r n_{max} m}{n_{max}^m} \tag{38}$$

We see from Figure 6 that a low-rank compression with an accuracy of $1.0 \times 10^{-2}$ of $G(\boldsymbol{x}, \boldsymbol{x}')$ is obtained with $r = 10$ for water ($m = 3$) or with $r = 200$ for formaldehyde ($m = 6$). With $n_{max} = 8$ (as assumed in [27]), we find $\gamma = 0.37$ for water and $\gamma = 0.02$ for formaldehyde. Under the assumption that an accuracy of $1.0 \times 10^{-2}$ is sufficient for the sake of integration, this indicates that $G(\boldsymbol{x}, \boldsymbol{x}')$ is greatly compressed without any significant loss of information.

We compare the cost and accuracy of the proposed approach with those of our previous work reported in [25] that exploits only the low rank structure and Monte

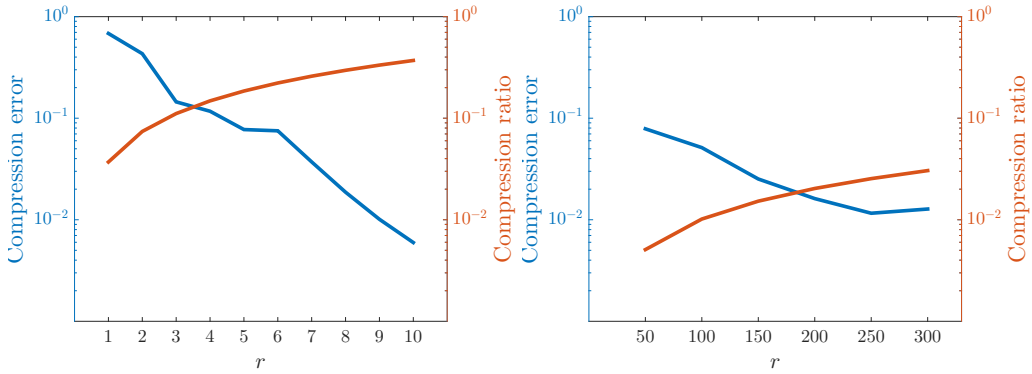


Figure 6: The compression error and compression ratio ($\gamma$) of low-rank-decomposed $G(\boldsymbol{x}, \boldsymbol{x}')$ as a function of the separation rank ($r$) for (a) water and (b) formaldehyde.

Carlo against reference calculations obtained from a high fidelity calculation. We show results for the first- and second-order corrections to the zero-point energies of water and formaldehyde.

For the water molecule (Table 3), we consider 40 PES evaluations for our approach as opposed to 150 for the low-rank-only method and $7.0 \times 10^5$ for Monte Carlo. With 40 samples, the zero-point energies come within tenths of $1\text{cm}^{-1}$ of the reference, which is sufficiently high accuracy for practical purposes. Moreover, the proposed method is three times faster than low rank and four orders of magnitude faster than Monte Carlo. For the formaldehyde molecule (Table 4), we consider 180 PES evaluations for the proposed method, as opposed to 4000 for the low rank method, to obtain energies with the same order of accuracy.

One of the most important questions from the quantum chemistry perspective is the scaling with respect to the size of the molecule. We find that, in the proposed method, the cost increases by a factor of $\frac{180}{40} = 4.5$ in going from water to formaldehyde, as compared to significantly higher increase of $\frac{4000}{150} \approx 27$ for the low rank only approach. Although the current dataset and range of problem size are both too small to draw any definitive general conclusions on scaling with molecule size, it is yet clear that, for these two molecules, our approach allows for significant reduction in cost scaling for zero-

Table 3: The first- and second-order anharmonic corrections $I^{(1)}$ and $I^{(2)}$ to the zero-point energy in $\mathrm{cm}^{-1}$ of the water molecule.

| | Sparse+Low Rank+Quadrature | Low rank+Quadrature [25] | Monte Carlo | Reference [28] |
|---|---|---|---|---|
| # Samples | 40 | 150 | $7.0\times10^5$ | |
| $I^{(1)}$ | $51.9\pm0.4$ | $51.5\pm0.3$ | $51.3\pm1.1$ | 51.6 |
| $I^{(2)}$ | $-120.5\pm0.1$ | $-120.5\pm0.2$ | $-119.1\pm0.7$ | $-120.6$ |

Table 4: The first- and second-order anharmonic corrections $I^{(1)}$ and $I^{(2)}$ to the zero-point energy in $\mathrm{cm}^{-1}$ of the formaldehyde molecule.

| | Sparse+Low Rank+Quadrature | Low rank+Quadrature [25] | Reference [28] |
|---|---|---|---|
| # Samples | 180 | 4000 | |
| $I^{(1)}$ | $-1.02\pm0.7$ | $-1.1\pm0.3$ | -1.0 |
| $I^{(2)}$ | $-77.7\pm0.10$ | $-77.7\pm0.3$ | -77.7 |

point energy calculations, while keeping the accuracy within $1\text{cm}^{-1}$. Finally, we note that, for both molecules, computation of compressive sensing and CP decomposition combined took less time than one sample evaluation of the potential energy surface using NWChem on the same machine. For a more quantitative complexity estimation, we refer to references [35] and [33] for LARS based sparse approximation and [42] for complexity of CP decomposition.

## 5. Conclusion

We presented a scalable general approach that takes advantage of sparsity and low rank structure to integrate high dimensional functions with few evaluations of the integrand. A sparse representation of the integrand is sought after in an approximation space chosen *a priori*. The sparse solution thus obtained is then compressed using low rank tensor decomposition to further reduce the number of terms in the separated representation. Finally, an appropriate quadrature rule is used to perform dimension-wise integration. We illustrated this method on benchmark integration problems, as well as a quantum chemistry application for calculating zero point energy corrections of molecules. The method achieves similar accuracy, with orders of magnitude fewer evaluations, as compared to existing methods in the literature.

In future work, we will enhance the approach to overcome the current limitations. Currently, we choose the approximation space based on the total degree of the multidimensional basis. We can, however, perform a more adaptive selection of the approximation basis. Indeed, selection of the basis using total degree approximation space excludes basis functions which could be significant for an accurate sparse representation. Application of this approach for larger molecules (e.g. $C_6H_6$) will lead to a better understanding of scaling behavior.

## 6. Acknowledgement

The authors thank So Hirata and Matthew Hermes at University of Illinois, Urbana Champaign and Tamara Kolda at Sandia National Laboratories, Livermore for fruitful discussions. Support for this work was provided through the Scientific Discovery

through Advanced Computing (SciDAC) program funded by the U.S. Department of Energy, Office of Science, Advanced Scientific Computing Research and Basic Energy Sciences under Award No. DE-FG02-12ER46875. Sandia National Laboratories is a multimission laboratory operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration. Sandia has major research and development responsibilities in nuclear deterrence, global security, defense, energy technologies and economic competitiveness, with main facilities in Albuquerque, New Mexico, and Livermore, California.

## References

[1] J. Almlöf, K. Faegri, K. Korsell, Principles for a direct SCF approach to LCAO-MO ab-initio calculations, J. Comput. Chem. 3 (3) (1982) 385–399.

[2] M. Valiev, E. J. Bylaska, N. Govind, K. Kowalski, T. P. Straatsma, H. J. Van Dam, D. Wang, J. Nieplocha, E. Apra, T. L. Windus, et al., Nwchem: a comprehensive and scalable open-source solution for large scale molecular simulations, Computer Physics Communications 181 (9) (2010) 1477–1489.

[3] J. M. Bowman, J. S. Bittman, L. B. Harding, A binitio calculations of electronic and vibrational energies of hco and hoc, The Journal of chemical physics 85 (2) (1986) 911–921.

[4] S. Chapman, M. Dupuis, S. Green, Theoretical three-dimensional potential-energy surface for the reaction of be with hf, Chemical Physics 78 (1) (1983) 93–105.

[5] J. Ischtwan, M. A. Collins, Molecular potential energy surfaces by interpolation, The Journal of chemical physics 100 (11) (1994) 8080–8088.

[6] G. G. Maisuradze, D. L. Thompson, A. F. Wagner, M. Minkoff, Interpolating moving least-squares methods for fitting potential energy surfaces: Detailed analysis of one-dimensional applications, The Journal of chemical physics 119 (19) (2003) 10002–10014.

[7] Y. Guo, A. Kawano, D. L. Thompson, A. F. Wagner, M. Minkoff, Interpolating moving least-squares methods for fitting potential energy surfaces: Applications to classical dynamics calculations, The Journal of chemical physics 121 (11) (2004) 5091–5097.

[8] B. G. Sumpter, D. W. Noid, Potential energy surfaces for macromolecules. a neural network technique, Chemical physics letters 192 (5) (1992) 455–462.

[9] T. B. Blank, S. D. Brown, A. W. Calhoun, D. J. Doren, Neural network models of potential energy surfaces, The Journal of chemical physics 103 (10) (1995) 4129–4137.

[10] D. F. Brown, M. N. Gibbs, D. C. Clary, Combining ab initio computations, neural networks, and diffusion monte carlo: An efficient method to treat weakly bound molecules, The Journal of chemical physics 105 (17) (1996) 7597–7604.

[11] F. V. Prudente, J. S. Neto, The fitting of potential energy surfaces using neural networks. application to the study of the photodissociation processes, Chemical physics letters 287 (5) (1998) 585–589.

[12] H. Gassner, M. Probst, A. Lauenstein, K. Hermansson, Representation of intermolecular potential functions by neural networks, The Journal of Physical Chemistry A 102 (24) (1998) 4596–4605.

[13] S. Lorenz, A. Groß, M. Scheffler, Representing high-dimensional potential-energy surfaces for reactions at surfaces by neural networks, Chemical Physics Letters 395 (4) (2004) 210–215.

[14] T. Hollebeek, T.-S. Ho, H. Rabitz, Constructing multidimensional molecular potential energy surfaces from ab initio data, Annual review of physical chemistry 50 (1) (1999) 537–570.

[15] P. Rai, Sparse low rank approximation of multivariate functions–applications in uncertainty quantification, Ph.D. thesis, Ecole Centrale Nantes (2014).

[16] E. Candes, J. Romberg, T. Tao, Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information, IEEE Trans. Info. Theory 52(2) (2) (2006) 489–509.

[17] E. Candes, J. Romberg, T. Tao, Near optimal signal recovery from random projections: Universal encoding strategies?, IEEE Transactions on information theory 52(12) (2006) 5406–5425.

[18] G. Blatman, B. Sudret, Adaptive sparse polynomial chaos expansion based least angle regression, Journal of Computational Physics 230 (2011) 2345–2367.

[19] A. Doostan, H. Owhadi, A non-adapted sparse approximation of pdes with stochastic inputs, Journal of Computational Physics 230 (8) (2011) 3015–3034.

[20] W. Hackbusch, Tensor Spaces and Numerical Tensor Calculus, Vol. 42, Springer, 2012.

[21] B. N. Khoromskij, Tensors-structured numerical methods in scientific computing: Survey on recent advances, Chemom. Intell. Lab. Syst. 110 (1) (2012) 1–19.

[22] L. Grasedyck, D. Kressner, C. Tobler, A literature survey of low-rank tensor approximation techniques, GAMM-Mitteilungen 36 (1) (2013) 53–78.

[23] V. Khoromskaia, B. N. Khoromskij, Tensor numerical methods in quantum chemistry: from hartree–fock to excitation energies, Physical Chemistry Chemical Physics 17 (47) (2015) 31491–31509.

[24] U. Benedikt, A. A. Auer, M. Espig, W. Hackbusch, Tensor decomposition in post-hartree–fock methods. i. two-electron integrals and mp2, The journal of chemical physics 134 (5) (2011) 054118.

[25] P. Rai, K. Sargsyan, H. Najm, M. R. Hermes, S. Hirata, Low-rank canonical-tensor decomposition of potential energy surfaces: application to grid-based diagrammatic vibrational green's function theory, Molecular Physics 115 (17-18) (2017) 2120–2134.

[26] B. Ziegler, G. Rauhut, Efficient generation of sum-of-products representations of high-dimensional potential energy surfaces based on multimode expansions, J. Chem. Phys. 144 (11) (2016) 114114.

[27] M. R. Hermes, S. Hirata, Second-order many-body perturbation expansions of vibrational Dyson self-energies, J. Chem. Phys. 139 (3) (2013) 034111.

[28] M. R. Hermes, S. Hirata, Stochastic many-body perturbation theory for anharmonic molecular vibrations, J. Chem. Phys. 141 (8) (2014) 084105, *ibid.* **143**, 129903(E) (2015).

[29] A. A. Gorodetsky, S. Karaman, Y. M. Marzouk, Function-train: A continuous analogue of the tensor-train decomposition, arXiv preprint arXiv:1510.09088.

[30] P. Rai, M. Chevreuil, A. Nouy, J. S. Gupta, A regression based non-intrusive method using separated representation for uncertainty quantification, in: ASME 2012 11th Biennial Conference on Engineering Systems Design and Analysis, American Society of Mechanical Engineers, 2012, pp. 167–174.

[31] K. Konakli, B. Sudret, Polynomial meta-models with canonical low-rank approximations: numerical insights and comparison to sparse polynomial chaos expansions, Journal of Computational Physics 321 (2016) 1144–1169.

[32] M. Chevreuil, R. Lebrun, A. Nouy, P. Rai, A least-squares method for sparse low rank approximation of multivariate functions, SIAM/ASA J. Uncertain. Quantif. 3 (1) (2015) 897–921.

[33] F. Bach, R. Jenatton, J. Mairal, G. Obozinski, et al., Optimization with sparsity-inducing penalties, Foundations and Trends® in Machine Learning 4 (1) (2012) 1–106.

[34] D. L. Donoho, Compressed sensing, IEEE Transactions on information theory 52 (4) (2006) 1289–1306.

[35] B. Efron, T. Hastie, I. Johnstone, R. Tibshirani, et al., Least angle regression, The Annals of statistics 32 (2) (2004) 407–499.

[36] J. Mairal, F. Bach, J. Ponce, G. Sapiro, Online dictionary learning for sparse coding, in: Proceedings of the 26th annual international conference on machine learning, ACM, 2009, pp. 689–696.

[37] G. Blatman, B. Sudret, Adaptive sparse polynomial chaos expansion based on least angle regression, Journal of Computational Physics 230 (6) (2011) 2345–2367.

[38] MATLAB Tensor Toolbox Version 2.6, available online, B.W. Bader, T.G. Kolda, *et al.*, MATLAB Tensor Toolbox Version 2.6 (2015), available online at http://www.sandia.gov/~tgkolda/TensorToolbox (February 2015).
URL http://www.sandia.gov/~tgkolda/TensorToolbox/

[39] S. Surjanovic, D. Bingham, Virtual library of simulation experiments: Test functions and datasets, Retrieved May 9, 2017, from http://www.sfu.ca/~ssurjano.

[40] A. Doostan, A. Validi, G. Iaccarino, Non-intrusive low-rank separated approximation of high-dimensional stochastic models, Comput. Methods Appl. Mech. Eng. 263 (2013) 42–55.

[41] R. C. Fortenberry, X. Huang, A. Yachmenev, W. Thiel, T. J. Lee, On the use of quartic force fields in variational calculations, Chemical Physics Letters 574 (2013) 1–12.

[42] T. Kolda, B. Bader, Tensor decompositions and applications, SIAM Rev. 51 (3) (2009) 455–500. doi:10.1137/07070111X.