
Low-Rank Tensor Approximation of Multivariate Functions for Quantum Chemistry Applications

Prashant Rai

Sandia National Laboratories, Livermore, CA

Joint work with :

M. Hermes, S. Hirata (UIUC)

K. Sargsyan, H. Najm (Sandia)

Dec 3, 2016

BASCD, Stanford



**Sandia
National
Laboratories**

High Dimensional Integrals are Difficult!

What we want?

$$I(u) = \int_{\Omega} u(\mathbf{x}) d\mathbf{x}$$

$$\mathbf{x} = (x_1, \dots, x_m); \quad m = 10, 100, 1000 \dots$$

Find $I(u)$ by sampling the $u(\mathbf{x})$ at random/well chosen points

Why is it difficult?

The **amount of information** (or samples) needed to integrate a high dimensional function increases **exponentially with dimension**

Is there a way? Monte Carlo

- Sample the function at a large number of (quasi) random points
- Compute average as $I(u)$
- Convergence rate independent of dimension ☺
- Dependence on variance of the function ☹

High Dimensional Integrals are Difficult!

What we want?

$$I(u) = \int_{\Omega} u(\mathbf{x}) d\mathbf{x}$$
$$\mathbf{x} = (x_1, \dots, x_m); \quad m = 10, 100, 1000 \dots$$

Find $I(u)$ by sampling the $u(\mathbf{x})$ at random/well chosen points

Why is it difficult?

The **amount of information** (or samples) needed to integrate a high dimensional function increases **exponentially with dimension**

Is there a way? Monte Carlo

- Sample the function at a large number of (quasi) random points
- Compute average as $I(u)$
- Convergence rate independent of dimension ☺
- Dependence on variance of the function ☹

High Dimensional Integrals are Difficult!

What we want?

$$I(u) = \int_{\Omega} u(\mathbf{x}) d\mathbf{x}$$
$$\mathbf{x} = (x_1, \dots, x_m); \quad m = 10, 100, 1000 \dots$$

Find $I(u)$ by sampling the $u(\mathbf{x})$ at random/well chosen points

Why is it difficult?

The **amount of information** (or samples) needed to integrate a high dimensional function increases **exponentially with dimension**

Is there a way? Monte Carlo

- Sample the function at a large number of (quasi) random points
- Compute average as $I(u)$
- Convergence rate independent of dimension 😊
- Dependence on variance of the function 😞

Key Idea: Separated Integration

Integration Problem

$$I(u) = \int_{\Omega} u(\mathbf{x})\rho(\mathbf{x})d\mathbf{x}$$

- $u(\mathbf{x}) = u(x_1, \dots, x_m)$
- $\rho(\mathbf{x}) = \rho(x_1) \cdots \rho(x_m)$

Low rank approximation of $u(\mathbf{x})$

$$u(\mathbf{x}) \approx v(x_1, \dots, x_m) = \sum_{\mu=1}^r v_{\mu}^{(1)}(x_1) \cdots v_{\mu}^{(m)}(x_m)$$

Separated Integration

$$I(u) \approx I(v) = \sum_{\mu=1}^r \left(\int_{\Omega_1} v_{\mu}^{(1)}(x_1)\rho(x_1)dx_1 \cdots \int_{\Omega_m} v_{\mu}^{(m)}(x_m)\rho(x_m)dx_m \right)$$

How to construct a low rank approximation of $u(\mathbf{x})$?

Functional Representation

Linear Approximation

$$u(x) \approx \sum_{j=1}^n u_j \phi_j(x)$$

- $u_j \in \mathbb{R}$ are coefficients
- $\phi_j(x)$ are basis functions

How should we choose basis set?

- Simplicity: polynomial, trigonometric functions
- Low Cardinality: small n

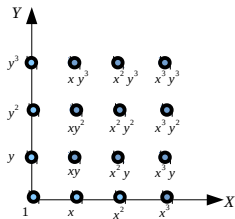
Problem

In high dimensions, both are competing objectives !

Curse of dimensionality in approximation

Approximation of a bivariate function $u(x, y)$

$$u(x, y) \approx \sum_{j_1=1}^n \sum_{j_2=1}^n u_{j_1 j_2} \phi_{j_1}^{(1)}(x) \phi_{j_2}^{(2)}(y)$$

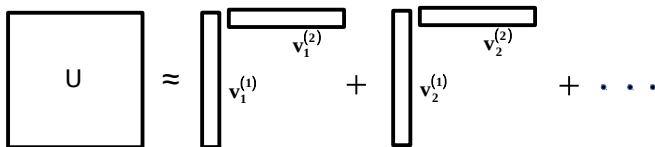


Approximation of a m -variate function $u(x_1, \dots, x_m)$

$$u(x_1, \dots, x_m) \approx \sum_{j_1=1}^n \cdots \sum_{j_m=1}^n u_{j_1 \dots j_m} \phi_{j_1}^{(1)}(x_1) \cdots \phi_{j_m}^{(m)}(x_m)$$

In Search of Low Rank Structures

Low rank structure from Singular Value Decomposition

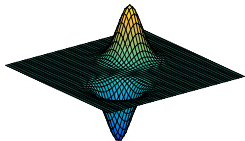

$$U \approx \begin{matrix} \text{column} \\ v_1^{(1)} \end{matrix} \begin{matrix} \text{row} \\ v_1^{(2)} \end{matrix} + \begin{matrix} \text{column} \\ v_2^{(1)} \end{matrix} \begin{matrix} \text{row} \\ v_2^{(2)} \end{matrix} + \dots$$

Separated representation of a function ($m = 2$)

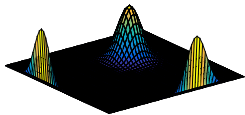
$$u(x, y) \approx v_1^{(1)}(x)v_1^{(2)}(y) + v_2^{(1)}(x)v_2^{(2)}(y) + \dots$$

Example

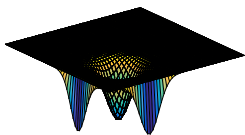
$u(x, y)$



$v_1^{(1)}(x)v_1^{(2)}(y)$

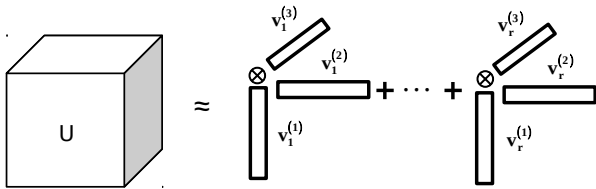


$v_2^{(1)}(x)v_2^{(2)}(y)$



Canonical Tensor Format

Generalization of SVD for $m = 3$



Representation of a function in canonical format

$$u(\mathbf{x}) \approx v(x_1, \dots, x_m) = \sum_{\mu=1}^r v_{\mu}^{(1)}(x_1) \cdots v_{\mu}^{(m)}(x_m); \quad v_{\mu}^{(i)}(x_i) = \sum_{j=1}^n v_{\mu,j}^{(i)} \phi_j^{(i)}(x_i)$$

Least-Squares Approximation

Functional representation

$$u(x) = \sum_{i=1}^n v_i \phi_i(x)$$
$$\mathbf{u} = \Phi \mathbf{v}$$

- $\Phi \in \mathbb{R}^{S \times n}$, $\Phi_{si} = \phi_i(x^s)$: Measurement matrix
- $\mathbf{v} \in \mathbb{R}^n$: Coefficient vector
- $\mathbf{u} \in \mathbb{R}^S$, $u_s = u(x^s)$: Vector of function evaluations

Optimization Problem

$$\hat{\mathbf{v}} = \min_{\mathbf{v} \in \mathbb{R}^n} \|\mathbf{u} - \Phi \mathbf{v}\|_2^2$$

$$(\Phi^{-1} \Phi) \hat{\mathbf{v}} = \Phi^{-1} \mathbf{u}$$

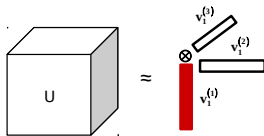
ALS in rank one format

- Consider a rank one approximation of a 3-d function

$$u(x) \approx v^{(1)}(x_1)v^{(2)}(x_2)v^{(3)}(x_3)$$

- ALS

- ALS in dimension 1: $u(x) \approx \left(\sum_{j=1}^n v_j^{(1)} \phi_j^{(1)}(x_1) \right) v^{(2)}(x_2)v^{(3)}(x_3)$
- ALS in dimension 2: $u(x) \approx v^{(1)}(x_1) \left(\sum_{j=1}^n v_j^{(2)} \phi_j^{(2)}(x_2) \right) v^{(3)}(x_3)$
- ALS in dimension 3: $u(x) \approx v^{(1)}(x_1)v^{(2)}(x_2) \left(\sum_{j=1}^n v_j^{(3)} \phi_j^{(3)}(x_3) \right)$



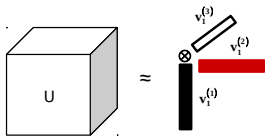
ALS in rank one format

- Consider a rank one approximation of a 3-d function

$$u(x) \approx v^{(1)}(x_1)v^{(2)}(x_2)v^{(3)}(x_3)$$

- ALS

- ALS in dimension 1: $u(x) \approx \left(\sum_{j=1}^n v_j^{(1)} \phi_j^{(1)}(x_1) \right) v^{(2)}(x_2)v^{(3)}(x_3)$
- ALS in dimension 2: $u(x) \approx v^{(1)}(x_1) \left(\sum_{j=1}^n v_j^{(2)} \phi_j^{(2)}(x_2) \right) v^{(3)}(x_3)$
- ALS in dimension 3: $u(x) \approx v^{(1)}(x_1)v^{(2)}(x_2) \left(\sum_{j=1}^n v_j^{(3)} \phi_j^{(3)}(x_3) \right)$



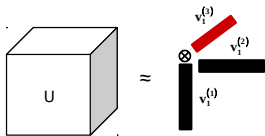
ALS in rank one format

- Consider a rank one approximation of a 3-d function

$$u(x) \approx v^{(1)}(x_1)v^{(2)}(x_2)v^{(3)}(x_3)$$

- ALS

- ALS in dimension 1: $u(x) \approx \left(\sum_{j=1}^n v_j^{(1)} \phi_j^{(1)}(x_1) \right) v^{(2)}(x_2)v^{(3)}(x_3)$
- ALS in dimension 2: $u(x) \approx v^{(1)}(x_1) \left(\sum_{j=1}^n v_j^{(2)} \phi_j^{(2)}(x_2) \right) v^{(3)}(x_3)$
- ALS in dimension 3: $u(x) \approx v^{(1)}(x_1)v^{(2)}(x_2) \left(\sum_{j=1}^n v_j^{(3)} \phi_j^{(3)}(x_3) \right)$



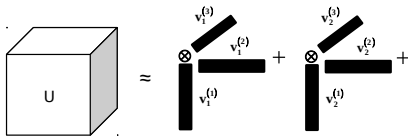
ALS in rank one format

- Consider a rank one approximation of a 3-d function

$$u(x) \approx v^{(1)}(x_1)v^{(2)}(x_2)v^{(3)}(x_3)$$

- ALS

- ALS in dimension 1: $u(x) \approx \left(\sum_{j=1}^n v_j^{(1)} \phi_j^{(1)}(x_1) \right) v^{(2)}(x_2)v^{(3)}(x_3)$
- ALS in dimension 2: $u(x) \approx v^{(1)}(x_1) \left(\sum_{j=1}^n v_j^{(2)} \phi_j^{(2)}(x_2) \right) v^{(3)}(x_3)$
- ALS in dimension 3: $u(x) \approx v^{(1)}(x_1)v^{(2)}(x_2) \left(\sum_{j=1}^n v_j^{(3)} \phi_j^{(3)}(x_3) \right)$



Quantum Chemistry Application

- Accurate prediction of vibrational spectra of molecules using perturbation theory require
 - Anharmonic approximation of potential energy surface
 - First and second order corrections to energy
 - First and second order corrections to frequencies per d.o.f
- Energy and frequency corrections are formulated as **non singular integrals** which can be **high order and multi-centered**

Objective

improve integration efficiency and scalability

First Order Corrections

$$I_1 = \int \Phi_0(\mathbf{x}) \Delta V(\mathbf{x}) \Phi_0(\mathbf{x}) \lambda_1(\mathbf{x}) d\mathbf{x}$$

$$\Delta V(\mathbf{x}) = V(\mathbf{x}) - V_{\text{ref}} - \frac{1}{2} \sum_{i=1}^m \omega_i^2 x_i^2$$

- $V(\mathbf{x})$: Potential energy (PE) containing up to n -th order force constants. As a default scenario, $n = 4$.
- V_{ref} : the minimum/reference PE, *i.e.* at the equilibrium geometry.
- ω_i : the i -th mode frequency of a reference mean-field theory. Found by solving Dyson equation.

$$\Phi_0(\mathbf{x}) = \prod_{i=1}^m \eta_{n_i}(x_i); \quad \eta_{n_i}(x_i) = C(n_i, \omega_i) e^{-\frac{\omega_i x_i^2}{2}} h_{n_i}(\sqrt{\omega_i} x_i)$$

- m : vibrational degrees of freedom. For H_2O (water), $m = 3$ and H_2CO (formaldehyde), $m=6$.
- $\eta_{n_i}(x_i)$: the harmonic-oscillator wave function with quantum number n_i along the i -th normal mode x_i .
- h_{n_i} : Hermite polynomial of degree n_i

First Order Corrections (contd..)

$$I_1 = \int \Phi_0(\mathbf{x}) \Delta V(\mathbf{x}) \Phi_0(\mathbf{x}) \lambda_1(\mathbf{x}) d\mathbf{x}$$

I_1	$\lambda_1(\mathbf{x}) = \prod_{i=1}^m \lambda_1^{(i)}(x_i)$
Energy (E_1)	$\lambda_1^{(i)}(x_i) = 1, 1 \leq i \leq m$
Frequencies ($\Sigma_1^{(i)}$)	$\lambda_1^{(i)}(x_i) = \frac{2^{1/2} \eta_2(x_i)}{\eta_0(x_i)}, \lambda_1^{(j)} = 1; j \neq i$

First order correction integrands can be reformulated as Gaussian times polynomials

$$I_1 = \int_{\mathbf{x}} e(\mathbf{x}) P_1(\mathbf{x}) d\mathbf{x}$$

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

$$P_1(\mathbf{x}) = \Delta V(\mathbf{x}) \prod_{i=1}^m \lambda_1^{(i)}(x_i)$$

Second Order Corrections

$$I_2 = \int_{\mathbf{x}} \int_{\mathbf{x}'} \Phi_0(\mathbf{x}) \Delta V(\mathbf{x}) G_2(\mathbf{x}, \mathbf{x}') \Delta V(\mathbf{x}') \Phi_0(\mathbf{x}') \lambda_2(\mathbf{x}, \mathbf{x}') d\mathbf{x} d\mathbf{x}'$$

using real-space Green's function

$$G_2(\mathbf{x}, \mathbf{x}') = \underbrace{e(\mathbf{x}, \mathbf{x}')}_{\text{Gaussian}} \underbrace{H(\mathbf{x}, \mathbf{x}')}_{\text{High rank polynomial on Hermite basis}}$$

$$e(\mathbf{x}, \mathbf{x}') = \prod_{i=1}^m e^{(i)}(x_i, x'_i); \quad e^{(i)}(x_i, x'_i) = e^{-\frac{\omega_i}{2}(x_i^2 + x_i'^2)}$$

$$H(\mathbf{x}, \mathbf{x}') = \sum_{n_1=1}^{n_{\max}} \cdots \sum_{n_m=1}^{n_{\max}} \frac{\prod_{i=1}^m C^2(n_i, \omega_i)}{-\sum_{i=1}^m n_i \omega_i} \prod_{i=1}^m h_{n_i}(\sqrt{\omega_i} x_i) h_{n_i}(\sqrt{\omega_i} x'_i)$$

Second Order Corrections (contd...)

$$I_2 = \int_{\mathbf{x}} \int_{\mathbf{x}'} \Phi_0(\mathbf{x}) \Delta V(\mathbf{x}) G_2(\mathbf{x}, \mathbf{x}') \Delta V(\mathbf{x}') \Phi_0(\mathbf{x}') \lambda_2(\mathbf{x}, \mathbf{x}') d\mathbf{x} d\mathbf{x}'$$

I_2	$\lambda_2(\mathbf{x}, \mathbf{x}') = \prod_{i=1}^m \lambda_2^{(i)}(x_i, x'_i)$
Energy (E_2)	$\lambda_2^{(i)}(x_i, x'_i) = 1$
Frequencies ($\Sigma_{2p}^{(i)}$)	$\lambda_2^{(i)}(x_i, x'_i) = \frac{(n_i+2)^{1/2}(n_i+1)^{1/2}\eta_{n_i+2}(x'_i)}{\eta_{n_i}(x'_i)}, \lambda_2^{(j \neq i)} = 1$
Frequencies ($\Sigma_{2b}^{(i)}$)	$\lambda_2^{(i)}(x_i, x'_i) = \frac{(n_i+1)\eta_{n_i+1}(x_i)\eta_{n_i+1}(x'_i)}{\eta_{n_i}(x'_i)}, \lambda_2^{(j \neq i)} = 1$

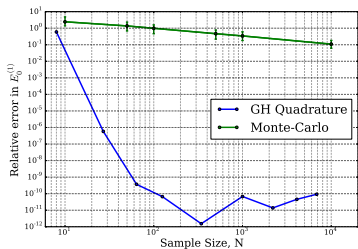
Second order correction integrands can also be formulated as Gaussian times polynomials

$$I_2 = \int_{\mathbf{x}} \int_{\mathbf{x}'} e(\mathbf{x}, \mathbf{x}') P_2(\mathbf{x}, \mathbf{x}') d\mathbf{x} d\mathbf{x}'$$

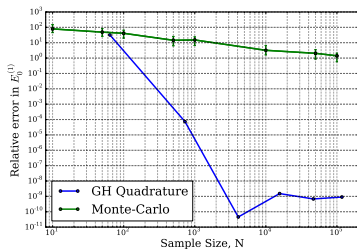
$$P_2(\mathbf{x}, \mathbf{x}') = \Delta V(\mathbf{x}) H(\mathbf{x}, \mathbf{x}') \Delta V(\mathbf{x}') \prod_{i=1}^m \lambda_2^{(i)}(x_i, x'_i)$$

E_1 Using Quadrature

Water



Formaldehyde



Quadrature based integration, in general, is not scalable

Second Order Correction

$$I_2 = \int_{\mathbf{x}} \int_{\mathbf{x}'} e(\mathbf{x}, \mathbf{x}') P_2(\mathbf{x}, \mathbf{x}') d\mathbf{x} d\mathbf{x}'$$

$$e(\mathbf{x}, \mathbf{x}') = \prod_{i=1}^m e^{(i)}(x_i, x'_i); \quad e^{(i)}(x_i, x'_i) = e^{-\omega_i(x_i^2 + x_i'^2)}$$

$$P_2(\mathbf{x}, \mathbf{x}') = \Delta V(\mathbf{x}) H(\mathbf{x}, \mathbf{x}') \Delta V(\mathbf{x}') \prod_{i=1}^m \lambda_2^{(i)}(x_i, x'_i)$$

$$\Delta V(\mathbf{x}) \approx \sum_{\mu_1=1}^{r_1} \prod_{i=1}^m \Delta V_{\mu_1}^{(i)}(x_i)$$

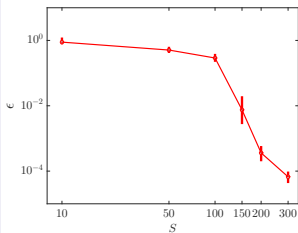
$$H(\mathbf{x}, \mathbf{x}') \approx \sum_{\mu_2=1}^{r_2} \prod_{i=1}^m H_{\mu_2}^{(i)}(x_i, x'_i)$$

$$I_2 \approx \sum_{\mu_1=1}^{r_1} \sum_{\mu_2=1}^{r_2} \sum_{\mu_3=1}^{r_1} \prod_{i=1}^m \int_{x_i} \int_{x'_i} e^{(i)}(x_i, x'_i) \underbrace{\Delta V_{\mu_1}^{(i)}(x_i) H_{\mu_2}^{(i)}(x_i, x'_i) \Delta V_{\mu_3}^{(i)}(x'_i) \lambda_2^{(i)}(x_i, x'_i)}_{\text{2 dimensional polynomial function}} d\mathbf{x} d\mathbf{x}'.$$

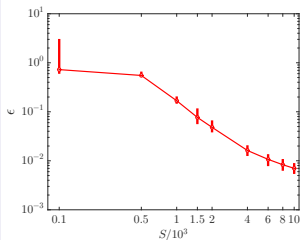
Approximation of ΔV

- Basis functions in each dimension: Monomials upto degree 4
- Optimal rank $r \in \{1, \dots, 30\}$
- Approximation error $\epsilon = \frac{\|u-v\|_{S'}}{\|u\|_{S'}}$, $S' = 1 \times 10^6$

Water ($m = 3$)



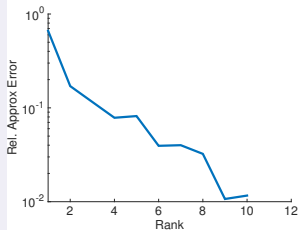
Formaldehyde ($m = 6$)



Compression of $H(\mathbf{x}, \mathbf{x}')$

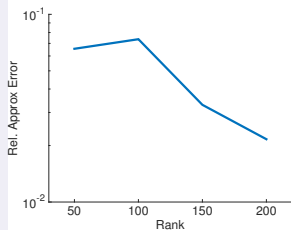
$$\text{Compression ratio } \gamma = \frac{n_{\max} \times r \times m}{n_{\max}^m}$$

Water



$$\gamma = 0.37$$

Formaldehyde



$$\gamma = 0.02$$

$H(\mathbf{x}, \mathbf{x}')$ is heavily over-parameterized in its original setting

Results

	Tensor	Monte Carlo	Reference
<i>#Samples</i>	150	7×10^5	
E_1	51.5 ± 0.3	51.3 ± 1.1	51.6
E_2	-120.5 ± 0.2	-119.1 ± 0.7	-120.6
ν_{min}	1566.7 ± 0.5	1566.3 ± 2.1	1566.9
ν_{max}	3767.3 ± 0.5	3768.5 ± 3.2	3767.4

Table: Water

	Tensor		Reference
<i>#Samples</i>	2000	4000	
E_1	-0.92 ± 0.3	-1.1 ± 0.3	-1.0
E_2	-77.3 ± 0.2	-77.7 ± 0.3	-77.7
ν_{min}	1167.4 ± 2.2	1166.4 ± 0.4	1166.4
ν_{max}	2871.9 ± 1.1	2870.4 ± 0.9	2870.8

Table: Formaldehyde

Acknowledgement

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.

Thank You