



Factorized structure of the long-range two-electron integrals tensor and its application in quantum chemistry

Siwar Badreddine, Igor Chollet, Laura Grigori

► To cite this version:

Siwar Badreddine, Igor Chollet, Laura Grigori. Factorized structure of the long-range two-electron integrals tensor and its application in quantum chemistry. 2022. <hal-03826012>

HAL Id: hal-03826012

<https://hal.science/hal-03826012v1>

Preprint submitted on 23 Oct 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

FACTORIZED STRUCTURE OF THE LONG-RANGE TWO-ELECTRON INTEGRALS TENSOR AND ITS APPLICATION IN QUANTUM CHEMISTRY

Siwar Badreddine*

Igor Chollet†

Laura Grigori‡

October 23, 2022

ABSTRACT

We introduce two new approximation methods for the numerical evaluation of the long-range Coulomb potential and the approximation of the resulting high dimensional Two-Electron Integrals tensor (TEI) with long-range interactions arising in molecular simulations. The first method exploits the tensorized structure of the compressed two-electron integrals obtained through two-dimensional Chebyshev interpolation combined with Gaussian quadrature. The second method is based on the Fast Multipole Method (FMM). Numerical experiments for different medium size molecules on high quality basis sets outline the efficiency of the two methods. Detailed algorithmic is provided in this paper as well as numerical comparison of the introduced approaches.

Keywords Two-electron integrals · tensor compression · Numerical integration · Interpolation · Fast Multipole Method (FMM) · Chebyshev Polynomials · Quantum chemistry

1 Introduction

In this paper we are interested in the numerical evaluation of the long-range Coulomb interaction and the approximation of the resulting Two-Electron Integrals (TEI) tensor. The evaluation of the two-electron integrals is considered as a challenging problem in quantum chemistry. These integrals are essential to approximate the solution of the so-known Schrodinger equation for a general N-body system [1] arising in electronic and molecular structure calculations. This equation describes the state function of a quantum-mechanical system which is given in the time-independant form as follows [2],

$$\mathbf{H}\psi = E\psi, \quad (1.1)$$

where \mathbf{H} is the Hamiltonian operator that can be described by the sum of three terms: the kinetic energy, the Coulomb interaction between electrons and nuclei, and the electron-electron Coulomb repulsion [3, 4], ψ is the wave-function or the state-function, and E is the full energy of the system. Under the Born-Oppenheimer approximation, i.e the motion of atomic nuclei and electrons can be treated separately given that the nuclei are much heavier than the electrons [3, 4], finding an exact, analytic solution of the Schrodinger equation becomes intractable for systems with more than one electron [3]. Therefore, additional assumptions are considered such as the Hartree-fock strategy and the Galerkin approximation procedure [4]. These assumptions yield to include the evaluation of the two-electron integrals such that, given the finite basis set $\{g_\mu\}_{1 \leq \mu \leq N_b}$, $g_\mu \in H^1(\mathbb{R}^3)$, these integrals are defined by [5]

$$\mathcal{B}(\mu, \nu, \kappa, \lambda) = \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{g_\mu(\mathbf{x})g_\nu(\mathbf{x})g_\kappa(\mathbf{y})g_\lambda(\mathbf{y})}{\|\mathbf{x} - \mathbf{y}\|} d\mathbf{x}d\mathbf{y}, \text{ with } \mu, \nu, \kappa, \lambda \in \{1, \dots, N_b\}. \quad (1.2)$$

These six-dimensional integrals are the entries of a fourth-order tensor, referred to as \mathcal{B} , with $O(N_b^4)$ entries with N_b being the number of basis functions $\{g_\mu\}_{1 \leq \mu \leq N_b}$. Many works exist in the literature for the analytic evaluation of these integrals using certain types of basis functions, mainly Slater type functions and Gaussian-type functions [6, 7]. Considerable efforts have been devoted to minimize the cost of the integrals evaluation which is a challenging computational problem since it requires the evaluation N_b^4 six-dimensional integrals that are singular due to the presence of the Coulomb potential $\frac{1}{\|\mathbf{x} - \mathbf{y}\|}$ and where N_b increases drastically with the molecular system size. An alternative approach to tackle this problem is to develop methods dealing with smooth potential. We consider in our work an approach that relies on the range-seperation of the Coulomb potential

*Laboratoire Jacques-Louis Lions, Sorbonne Université, INRIA Alpines, (siwar.badreddine@inria.fr).

†Laboratoire Analyse, Géométrie et Applications, (igor.chollet@inria.fr).

‡Laboratoire Jacques-Louis Lions, Sorbonne Université, INRIA Alpines, (laura.grigori@inria.fr).

[6, 8, 9, 10, 11, 12, 13, 14, 15, 16] where the last is split into a smooth range-part and a complementary diverging part. The splitting is done through the function $\text{erf}(\omega \|\mathbf{x} - \mathbf{y}\|)$ with ω being the range separator parameter. This separation writes

$$\frac{1}{\|\mathbf{x} - \mathbf{y}\|} = \frac{\text{erfc}(\omega \|\mathbf{x} - \mathbf{y}\|)}{\|\mathbf{x} - \mathbf{y}\|} + \frac{\text{erf}(\omega \|\mathbf{x} - \mathbf{y}\|)}{\|\mathbf{x} - \mathbf{y}\|}, 0 \leq \omega < \infty, \quad (1.3)$$

with

$$\frac{\text{erf}(\omega \|\mathbf{x} - \mathbf{y}\|)}{\|\mathbf{x} - \mathbf{y}\|} = \frac{2}{\|\mathbf{x} - \mathbf{y}\| \sqrt{\pi}} \int_0^{\omega \|\mathbf{x} - \mathbf{y}\|} \exp(-t^2) dt, \text{ and } \text{erfc}(\omega \|\mathbf{x} - \mathbf{y}\|) = 1 - \text{erf}(\omega \|\mathbf{x} - \mathbf{y}\|), \quad (1.4)$$

where $\mathbf{x} = (x_1, x_2, x_3)$, $\mathbf{y} = (y_1, y_2, y_3) \in \mathbb{R}^3$, ω is a positive parameter that controls the separation range. The long-range contribution in equation (1.4) is a smooth function such that, for small ω , the singularity is eliminated at $\|\mathbf{x} - \mathbf{y}\| = 0$. When $\omega = 0$, the long-range part vanishes and when $\omega \rightarrow \infty$, it approaches the Coulomb potential $\frac{1}{\|\mathbf{x} - \mathbf{y}\|}$. The short range contribution (the complementary function in equation (1.3)) has singularity at $\|\mathbf{x} - \mathbf{y}\| = 0$. The long-range part, that we denote $K(\mathbf{x}, \mathbf{y}) = \frac{\text{erf}(\omega \|\mathbf{x} - \mathbf{y}\|)}{\|\mathbf{x} - \mathbf{y}\|}$, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$, is treated usually through employing numerical integration in Fourier space [17]. Following equation (1.3), the two-electron integrals tensor can be expressed as the sum of two terms

$$\mathcal{B}(\mu, \nu, \kappa, \lambda) = \underbrace{\int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\text{erf}(\omega \|\mathbf{x} - \mathbf{y}\|)}{\|\mathbf{x} - \mathbf{y}\|} g_\mu(\mathbf{x}) g_\nu(\mathbf{x}) g_\kappa(\mathbf{y}) g_\lambda(\mathbf{y}) d\mathbf{x} d\mathbf{y}}_{\mathcal{B}^{lr}(\mu, \nu, \kappa, \lambda)} \quad (1.5)$$

$$+ \underbrace{\int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\text{erfc}(\omega \|\mathbf{x} - \mathbf{y}\|)}{\|\mathbf{x} - \mathbf{y}\|} g_\mu(\mathbf{x}) g_\nu(\mathbf{x}) g_\kappa(\mathbf{y}) g_\lambda(\mathbf{y}) d\mathbf{x} d\mathbf{y}}_{\mathcal{B}^{sr}(\mu, \nu, \kappa, \lambda)}, \quad (1.6)$$

with $\mu, \nu, \kappa, \lambda \in \{1, \dots, N_b\}$, \mathcal{B}^{lr} is referring to the long-range two-electron integrals tensor and \mathcal{B}^{sr} is referring to the short-range two-electron integrals tensor. In this paper, we focus on the numerical evaluation of the long-range kernel $K(\mathbf{x}, \mathbf{y})$ and on the approximation of the long-range two-electron integrals given by

$$\mathcal{B}^{lr}(\mu, \nu, \kappa, \lambda) = \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} g_\mu(\mathbf{x}) g_\nu(\mathbf{x}) K(\mathbf{x}, \mathbf{y}) g_\kappa(\mathbf{y}) g_\lambda(\mathbf{y}) d\mathbf{x} d\mathbf{y}, \mu, \nu, \kappa, \lambda \in \{1, \dots, N_b\}. \quad (1.7)$$

We use finite linear combinations of primitive Gaussians as basis functions $\{g_\mu\}_{1 \leq \mu \leq N_b}$. Such basis functions are expressed as linear combinations of I_μ primitive Gaussians functions [18]

$$g_\mu(\mathbf{x}) = \sum_{j=1}^{I_\mu} c_j \prod_{l=1}^3 g_\mu^{(j)}(x_l), \mathbf{x}_l \in \mathbb{R}, I_\mu \in \mathbb{N}, \quad (1.8)$$

where primitive Gaussians are defined by

$$g_\mu^{(j)}(x_l) = (x_l - \mathbf{r}_l)^{p_{\mu_l}} \exp(-\mu_j (x_l - \mathbf{r}_l)^2), \mathbf{x}_l \in \mathbb{R}, \mu \in \{1 \dots N_b\}, \quad (1.9)$$

where N_b is the number of basis functions defined in (1.8), c_j refers to a normalization constant, μ_j is a parameter whose reference value is, for instance, given in [19], \mathbf{r}_l refers to the coordinates of atom nucleus that is known in practice, and the p_{μ_l} are exponents depending on the chosen basis function. These basis functions $\{g_\mu\}_{1 \leq \mu \leq N_b}$ correspond in chemistry to approximations of the atomic orbitals. In addition, we consider restrictions of these basis functions to sufficiently large compact support $[-b, b]^3 \subset \mathbb{R}^3$ such that we have

$$\mathcal{B}^{lr}(\mu, \nu, \kappa, \lambda) = \int_{[-b, b]^3} \int_{[-b, b]^3} g_\mu(\mathbf{x}) g_\nu(\mathbf{x}) K(\mathbf{x}, \mathbf{y}) g_\kappa(\mathbf{y}) g_\lambda(\mathbf{y}) d\mathbf{x} d\mathbf{y}, \mu, \nu, \kappa, \lambda \in \{1, \dots, N_b\}. \quad (1.10)$$

We note that the range-separation representation of the Coulomb potential is important in molecular simulations to describe non-local correlation effects and to allow an accurate evaluation of the long-range two-electron integrals while keeping the computational cost reasonably low [8].

In this work, we introduce two numerical approaches for the numerical evaluation of the smooth long-range interaction and the approximation of the long-range two electron integrals tensor. First, instead of performing a naive numerical computation of $K(\mathbf{x}, \mathbf{y})$ over $N \times N \times N$ 3D Cartesian grids, we consider two-dimensional Chebyshev interpolation method using only $N^{\frac{1}{3}} \times N^{\frac{1}{3}}$ isotropic Chebyshev grids combined with Gaussian-quadrature rule in order to approximate $K(\mathbf{x}, \mathbf{y})$. We refer to this approach as TA for Tensorized Approximation and we denote the approximation method for the evaluation of the long-range two-electron integrals by LTEI-TA. This numerical approximation yields to a tensorized expression of the six-dimensional integrals with erf -interaction leading to substantial time complexity reduction to evaluate one integral of the form $\mathcal{B}^{lr}(\mu, \nu, \kappa, \lambda)$.

However, these six-dimensional integrals are only one element from the fourth-order tensor \mathcal{B}^{lr} . This means that for a basis set consisting of N_b basis functions, there are $\mathcal{O}(N_b^4)$ integrals to evaluate. Therefore, we introduce, using LTEI-TA approach, a new alternative way to approximate these integrals by means of a factorized representation of the fourth-order tensor $\mathcal{B}^{lr} \in \mathbb{R}^{N_b \times N_b \times N_b \times N_b}$, leading to an efficient application of the matricization of \mathcal{B}^{lr} to a vector with a significant reduction in time complexity to $\mathcal{O}(\epsilon N^{4/3})$, $\epsilon \ll N$ instead of $\mathcal{O}(N^2)$ given a naive computation. These complexities may be further reduced due to properties of Gaussian-type functions. Hence, we propose to express the high dimensional fourth-order tensor \mathcal{B}^{lr} in a more compressed format by using screening techniques and low-rank approximation methods. Second, we consider Chebyshev interpolation combined with Fast Multipole Method (FMM) [20, 21] leading to linear time complexity when computing the FMM-accelerated matrix vector product involving the two-electron integrals tensor. This method is referred to as LTEI-FMM. We provide detailed comparison between the two approaches and discuss to what extent the relative performances of these methods make them attractive for different application cases. In order to test the performance of our algorithm, we use the data sets of molecular properties calculated from quantum chemistry for some moderate size molecules. These data sets are extracted from *quantum package* [22].

The paper is organized as follows. In Section 2 we introduce notations, problem definitions, and properties. In Section 3, we describe our new tensorized method to approximate $K(\mathbf{x}, \mathbf{y})$ and we present our LTEI-TA scheme for the element-wise evaluation of the two-electron integrals based on the underlying tensorized structure. We describe also using LTEI-TA a factorized expression of the two-electron integrals tensor and we derive error bounds and theoretical complexities for the approximation process we use. In Section 4 we demonstrate that our kernel $K(\mathbf{x}, \mathbf{y})$ is asymptotically smooth, so that we can benefit from fast hierarchical methods (especially Fast Multipole Methods) in order to efficiently evaluate the two-electron integrals decompositions. Hence, we reformulate these decompositions as N -body problems on non-uniform particle distributions. In Section 5, we propose an application case in electronic calculations by using the decompositions of the two-electron integrals tensor obtained through the new introduced approaches. In Section 6, further compression techniques are also presented, extending screening approaches and low rank approximation methods to our new decompositions. Finally, results of numerical tests of both methods are presented as well as a summary of our findings. We use Julia open-source language to test the new approximation method TA and the evaluation scheme LTEI-TA⁴ and the C++ library *defmm*⁵ for LTEI-FMM.

2 Preliminaries

This section introduces our notations as well as several definitions and properties that will be used in the paper. The matrix operations notations are defined in Section 2.2.

2.1 Notations

We use the following notations:

- $\mathcal{B} \in \mathbb{R}^{I_1 \times I_2 \times I_3 \times I_4}$ is a fourth-order tensor with modes I_1, I_2, I_3, I_4 .
- $\mathbf{B}_{(j)} \in \mathbb{R}^{I_j \times I_1 I_2 \dots I_{j-1} I_{j+1} \dots I_4}$ is the mode- j matricization of \mathcal{B} and $\mathbf{B}_{(1, \dots, j)} \in \mathbb{R}^{I_1 I_2 \dots I_j \times I_{j+1} \dots I_4}$ is the mode- $(1, \dots, j)$ matricization of \mathcal{B} , with $j \in \{1, 2, 3, 4\}$ (see (2.19) for more details).
- Scalars are either lowercase letters $x, y, z, \alpha, \beta, \gamma$ or uppercase Latin letters N, M, T . Vectors \mathbf{b} are denoted by lowercase boldface letters, matrices \mathbf{B} are denoted by uppercase boldface letters.
- b_i is the i -th element of the vector \mathbf{b} , $\mathbf{B}_{(i_1, i_2)}$ is the (i_1, i_2) th entry of the matrix \mathbf{B} , $\mathcal{B}_{(i_1, i_2, i_3, i_4)}$ is the (i_1, i_2, i_3, i_4) th entry of the tensor $\mathcal{B} \in \mathbb{R}^{I_1 \times I_2 \times I_3 \times I_4}$.
- $\mathbf{B}[:, j]$ (Julia/Matlab notations) denotes the subvector containing the column of \mathbf{B} indexed by j , $\mathbf{B}[j, :]$ (Julia notation) denotes the subvector containing the row of \mathbf{B} indexed by j , $\mathcal{B}[:, :, j]$ denotes the submatrix extracted from \mathcal{B} at index j , $\mathcal{B}[:, :, :, j]$ denotes the subtensor extracted from \mathcal{B} at index j .
- $\|\mathbf{x} - \mathbf{y}\| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2}$ is the euclidean distance between two points $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$ with coordinates $(x_1, x_2, x_3), (y_1, y_2, y_3)$ respectively.
- $|x|$ is the absolute value of x .
- \otimes is the kronecker product, \odot is the Hadamard product, \diamond is the row-wise Khatri-rao product, and $*$ is the column-wise Khatri-rao product.
- $\|f\|_{\infty, S} := \sup\{|f(s)| : s \in S\}$.

2.2 Definitions and properties

We give in the following several definitions and properties that we use in the subsequent sections. In the different approximations derived in this paper, the product of two Gaussian type functions is often used. Therefore, we recall the general product rule between two Gaussian functions.

⁴https://github.com/sbadred/LTEI_TA.jl.git

⁵<https://github.com/IChollet/defmm>

Proposition 1 ([18]). Let $g_1(\mathbf{x}) = \exp(-c_1 |\mathbf{x} - \mathbf{r}|^2)$, $g_2(\mathbf{x}) = \exp(-c_2 |\mathbf{x} - \tilde{\mathbf{r}}|^2)$ be Gaussian functions with $\mathbf{x}, \mathbf{r}, \tilde{\mathbf{r}} \in \mathbb{R}^3$, $c_1, c_2 \in \mathbb{R}$. The product of these functions is

$$g_{12}(\mathbf{x}) = g_1(\mathbf{x})g_2(\mathbf{x}) = \exp\left(\frac{-c_1 c_2}{c_1 + c_2} |\mathbf{x} - \mathbf{r}_{12}|^2\right) \exp\left(-(c_1 + c_2) |\mathbf{x} - \mathbf{r}_{12}|^2\right), \quad (2.1)$$

where $\mathbf{r}_{12} = \frac{c_1}{c_1 + c_2} \mathbf{r} + \frac{c_2}{c_1 + c_2} \tilde{\mathbf{r}}$.

We also have recourse to two-dimensional Chebyshev interpolation. Therefore, we give the expressions of the Chebyshev polynomials as well as the Chebyshev coefficients.

Definition 2.1 (Two dimensional Chebyshev interpolation [23, 24]). For a given continuous function $f(x, y)$ on $[a, b]^2$, $a, b \in \mathbb{R}$, the two-dimensional Chebyshev interpolation of this function is given by its interpolating polynomial that we denote

$$\tilde{f}(x, y) = \sum_{n, m=0}^N \alpha_{nm} T_n(x) T_m(y), \quad (2.2)$$

where N is the number of interpolation nodes, $T_n(x) = \cos(n \arccos(x))$, $x \in [a, b]$, $n \in \{1, \dots, N\}$ are the Chebyshev polynomials,

$$\alpha_{nm} = \frac{c_{nm}}{N^2} \sum_{k, k'=1}^N f(x_k, y'_k) T_n(x_k) T_m(y'_k), \quad c_{n, m} = \begin{cases} 1 & \text{if } m = n = 0 \\ 2 & \text{if } m \neq n = 0 \text{ or } n \neq m = 0 \\ 4 & \text{if } m \neq 0, n \neq 0 \end{cases} \quad (2.3)$$

are Chebyshev interpolation coefficients. The nodes x_k, y'_k form the Chebyshev two-dimensional grids such as Chebyshev-Gauss points (first kind)

$$x_k = \cos \theta_k, \quad \theta_k = \frac{(2k-1)\pi}{2N}, \quad k = 1, \dots, N, \quad (2.4)$$

or Chebyshev-Lobatto points (second kind)

$$x_k = \cos \phi_k, \quad \phi_k = \frac{(k-1)\pi}{N-1}, \quad k = 1, \dots, N. \quad (2.5)$$

The following proposition gives the interpolation error of the two-dimensional Chebyshev approximation.

Proposition 2 (Interpolation error [25]). Let $\tilde{f}(x, y)$ be an interpolating polynomial of $f(x, y)$ on $[a, b]^2$ at Chebyshev N interpolation nodes and suppose that the partial derivatives $\partial^{N+1} f(x, y) / \partial x^{N+1}$ and $\partial^{N+1} f(x, y) / \partial y^{N+1}$ exist and are continuous for all $(x, y) \in [a, b]^2$. We have

$$|f(x, y) - \tilde{f}(x, y)| \leq \frac{\left(\frac{b-a}{2}\right)^{N+1}}{2^N (N+1)!} c_1 + \frac{\delta \left(\frac{b-a}{2}\right)^{N+1}}{2^N (N+1)!} c_2, \quad (2.6)$$

where

$$c_1 = \max_{\xi \in [a, b]} \left| \frac{\partial^{N+1} f(\xi, y)}{\partial \xi^{N+1}} \right|, \quad c_2 = \max_{(\xi, \eta) \in [a, b]^2} \left| \frac{\partial^{N+1} f(\xi, \eta)}{\partial \eta^{N+1}} \right|, \quad \text{and } \delta = \max_{s \in [a, b]} \sum_{i=0}^N |L_{i, N}(s)|. \quad (2.7)$$

The so-called Lebesgue constant δ grows only logarithmically if Chebyshev interpolation nodes are used, $L_{i, N}(s)$ are Lagrange polynomials of degree N .

The following proposition recalls the upper bound of Gaussian-quadrature rule error.

Proposition 3 (Quadrature error, Section 5.2 [26]). Let $[a, b]$ be a real closed interval of length $|b - a| > 0$ and let $f \in \mathbb{C}^{2N_q}([a, b])$, $N_q \geq 1$, the integration of f over $[a, b]$ can be given as follows, using Gaussian quadrature rule

$$\int_{[a, b]} f(x) dx = \int_{-1}^1 f\left(\frac{b-a}{2}z + \frac{a+b}{2}\right) \frac{dx}{dz} dz = \frac{b-a}{2} \sum_{i=1}^{N_q} w_i f\left(\frac{b-a}{2}z_i + \frac{a+b}{2}\right) + R_{N_q}, \quad (2.8)$$

where w_i and x_i are the weights and nodes of the quadrature rule, N_q is the number of quadrature points and R_{N_q} refers to the Gaussian quadrature error. This last quantity verifies

$$|R_{N_q}| \leq \frac{|b-a|^{2N_q+1} (N_q!)^4}{(2N_q+1) [(2N_q)!]^3} \left\| \frac{d^{2N_q}}{ds^{2N_q}} f(s) \right\|_{\infty, [a, b]}. \quad (2.9)$$

We recall now several matrix products that are used in this paper. The *Hadamard product* between matrices \mathbf{A} and $\mathbf{B} \in \mathbb{R}^{I \times J}$ is $\mathbf{A} \odot \mathbf{B} \in \mathbb{R}^{I \times J}$ defined as

$$\mathbf{A} \odot \mathbf{B} = \begin{bmatrix} a_{11}b_{11} & a_{12}b_{12} & \cdots & a_{1J}b_{1J} \\ a_{21}b_{21} & a_{22}b_{22} & \cdots & a_{2J}b_{2J} \\ \vdots & \vdots & \ddots & \vdots \\ a_{I1}b_{I1} & a_{I2}b_{I2} & \cdots & a_{IJ}b_{IJ} \end{bmatrix}. \quad (2.10)$$

The *Kronecker product* of matrices $\mathbf{A} \in \mathbb{R}^{I_1 \times J_1}$ and $\mathbf{B} \in \mathbb{R}^{I_2 \times J_2}$ is $\mathbf{A} \otimes \mathbf{B} \in \mathbb{R}^{I_1 I_2 \times J_1 J_2}$ defined as

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \cdots & a_{1J_1}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \cdots & a_{2J_1}\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{I_11}\mathbf{B} & a_{I_12}\mathbf{B} & \cdots & a_{I_1J_1}\mathbf{B} \end{bmatrix} \quad (2.11)$$

We also use the compact product notation $\otimes_{k=1}^d$. Given d matrices $\mathbf{A}_k \in \mathbb{R}^{I_k \times J_k}$, $k \in \{1, \dots, d\}$, we have

$$\otimes_{k=1}^d \mathbf{A}_k = \mathbf{A}_1 \otimes \mathbf{A}_2 \otimes \dots \otimes \mathbf{A}_d \in \mathbb{R}^{\prod_{k=1}^d I_k \times \prod_{k=1}^d J_k}. \quad (2.12)$$

Consider two matrices $\mathbf{A} = [\mathbf{A}[1, :]^\top \quad \mathbf{A}[2, :]^\top \quad \dots \quad \mathbf{A}[I_1, :]^\top]^\top \in \mathbb{R}^{I_1 \times J_1}$ and $\mathbf{B} = [\mathbf{B}[1, :]^\top \quad \mathbf{B}[2, :]^\top \quad \dots \quad \mathbf{B}[I_2, :]^\top]^\top \in \mathbb{R}^{I_2 \times J_2}$, where $\mathbf{A}[k, :] \in \mathbb{R}^{1 \times J_1}$ and $\mathbf{B}[k, :] \in \mathbb{R}^{1 \times J_2}$ for $k \in \{1, \dots, I_1\}$. The row-wise Khatri-Rao product $\mathbf{A} \diamond \mathbf{B}$ is a matrix of dimension $I_1 \times (J_1 J_2)$ defined as

$$\mathbf{A} \diamond \mathbf{B} = [\mathbf{A}[1, :]^\top \otimes \mathbf{B}[1, :]^\top \quad \mathbf{A}[2, :]^\top \otimes \mathbf{B}[2, :]^\top \quad \dots \quad \mathbf{A}[I_1, :]^\top \otimes \mathbf{B}[I_1, :]^\top]^\top. \quad (2.13)$$

Given d matrices $\mathbf{A}_k \in \mathbb{R}^{I_k \times J_k}$, $k \in \{1, \dots, d\}$, we use the notation

$$\diamond_{k=1}^d \mathbf{A}_k = \mathbf{A}_1 \diamond \mathbf{A}_2 \diamond \dots \diamond \mathbf{A}_d \in \mathbb{R}^{I_1 \times \prod_{k=1}^d J_k}. \quad (2.14)$$

Consider two matrices $\mathbf{A} = [\mathbf{A}[:, 1] \quad \mathbf{A}[:, 2] \quad \dots \quad \mathbf{A}[:, J_1]] \in \mathbb{R}^{I_1 \times J_1}$ and $\mathbf{B} = [\mathbf{B}[:, 1] \quad \mathbf{B}[:, 2] \quad \dots \quad \mathbf{B}[:, J_2]] \in \mathbb{R}^{I_2 \times J_2}$, where $\mathbf{A}[:, k] \in \mathbb{R}^{I_1 \times 1}$ and $\mathbf{B}[:, k] \in \mathbb{R}^{I_2 \times 1}$ for $k \in \{1, \dots, J_1\}$. The column-wise Khatri-Rao product $\mathbf{A} * \mathbf{B}$ is a matrix of dimension $(I_1 I_2) \times J_1$ defined as

$$\mathbf{A} * \mathbf{B} = [\mathbf{A}[:, 1] \otimes \mathbf{B}[:, 1] \quad \mathbf{A}[:, 2] \otimes \mathbf{B}[:, 2] \quad \dots \quad \mathbf{A}[:, J_1] \otimes \mathbf{B}[:, J_1]], \quad (2.15)$$

where $\mathbf{A}[:, k] \otimes \mathbf{B}[:, k]$ for $k \in \{1, \dots, J_1\}$ defines the Kronecker product between vectors $\mathbf{A}[:, k]$ and $\mathbf{B}[:, k]$. That is, each column of $\mathbf{A} * \mathbf{B}$ is the Kronecker product between the respective columns of the two input matrices \mathbf{A} and \mathbf{B} . The relation between column-wise and row-wise Khatri-Rao product is the following

$$(\mathbf{A} * \mathbf{B})^\top = \mathbf{A}^\top \diamond \mathbf{B}^\top. \quad (2.16)$$

We give several useful relations among these matrix products that we use in our derivations.

Proposition 4 ([27]). Consider matrices $\mathbf{A} \in \mathbb{R}^{I_1 \times J_1}$, $\mathbf{B} \in \mathbb{R}^{I_1 \times J_2}$, $\mathbf{C} \in \mathbb{R}^{J_1 \times J_3}$, and $\mathbf{D} \in \mathbb{R}^{J_2 \times J_4}$, then

$$(\mathbf{A} \diamond \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = (\mathbf{A}\mathbf{C}) \diamond (\mathbf{B}\mathbf{D}). \quad (2.17)$$

Consider matrices $\mathbf{A} \in \mathbb{R}^{I_1 \times J_1}$ and $\mathbf{B} \in \mathbb{R}^{I_1 \times J_2}$ and $\mathbf{C} \in \mathbb{R}^{J_1 \times J_3}$, and $\mathbf{D} \in \mathbb{R}^{J_2 \times J_4}$, then

$$(\mathbf{A} \diamond \mathbf{B})(\mathbf{C} * \mathbf{D}) = (\mathbf{A}\mathbf{C}) \odot (\mathbf{B}\mathbf{D}). \quad (2.18)$$

In this paper, we use the concept of matricization, also called tensor unfolding [28]. The mode- j matricization of a tensor $\mathcal{B} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_d}$, $d \in \mathbb{N}$, referred to as $\mathbf{B}_{(j)} \in \mathbb{R}^{I_j \times I_1 I_2 \dots I_{j-1} I_{j+1} \dots I_d}$, $j \in \{1, \dots, d\}$, can be defined by the following mapping

$$\mathcal{B}(i_1, i_2, \dots, i_d) = \mathbf{B}_{(j)}(i_j, i_1 i_2 \dots i_{j-1} i_{j+1} \dots i_d) = \mathbf{B}_{(j)}(i_j, \bar{i}), \quad (2.19)$$

with $\bar{i} = 1 + \sum_{\substack{k=1 \\ k \neq j}}^d \left((i_k - 1) \prod_{\substack{m=1 \\ m \neq j}}^{k-1} I_m \right)$. For example, if $d = 4$, the mode-1 matricization of $\mathcal{B} \in \mathbb{R}^{I_1 \times I_2 \times I_3 \times I_4}$ which is denoted by $\mathbf{B}_{(1)} \in \mathbb{R}^{I_1 \times I_2 I_3 I_4}$ can be defined by the following mapping

$$\mathcal{B}(i_1, i_2, i_3, i_4) = \mathbf{B}_{(1)}(i_1, i_2 i_3 i_4) = \mathbf{B}_{(1)}(i_1, \bar{i}), \quad (2.20)$$

with $\bar{i} = 1 + (i_2 - 1)I_1 + (i_3 - 1)I_2 + (i_4 - 1)I_2 I_3$. The mode-(1, 2) matricization of $\mathcal{B} \in \mathbb{R}^{I_1 \times I_2 \times I_3 \times I_4}$ which is denoted by $\mathbf{B}_{(1,2)} \in \mathbb{R}^{I_1 I_2 \times I_3 I_4}$ can be denoted entry-wise as follows

$$\mathcal{B}(i_1, i_2, i_3, i_4) = \mathbf{B}_{(1,2)}(i_1 i_2, i_3 i_4). \quad (2.21)$$

3 Long-range TEI tensor factorization through Tensorized Approximation (LTEI-TA)

In this section we introduce a new numerical method that allows to evaluate efficiently the two-electron integrals through the factorization of the long-range Coulomb potential. This method, that we refer to as TA, factorizes the fourth order long-range two-electron integrals tensor \mathcal{B}^{lr} through the approximation of the long-range kernel $K(\mathbf{x}, \mathbf{y})$ with two-dimensional Chebyshev interpolation and Gaussian quadrature. Error bounds for the numerical approximation of the long-range two-electron integrals are also provided. approximated six-dimensional integral.

3.1 The element-wise evaluation of the TEI tensor

We first describe the efficient evaluation of the six-dimensional integrals $\mathcal{B}^{lr}(\mu, \nu, \kappa, \lambda)$ defined in (1.7). We start by presenting our approach for computing the long-range $K(\mathbf{x}, \mathbf{y})$ defined as

$$K(\mathbf{x}, \mathbf{y}) = \frac{\exp(\omega \|\mathbf{x} - \mathbf{y}\|)}{\|\mathbf{x} - \mathbf{y}\|} = \frac{2}{\sqrt{\pi}} \frac{\int_{[0, \omega \|\mathbf{x} - \mathbf{y}\|]} \exp(-t^2) dt}{\|\mathbf{x} - \mathbf{y}\|}, \mathbf{x}, \mathbf{y} \in \mathbb{R}^3. \quad (3.1)$$

Let $t = s \|\mathbf{x} - \mathbf{y}\|$. With this change of variable, we obtain

$$K(\mathbf{x}, \mathbf{y}) = \frac{2}{\sqrt{\pi}} \int_{[0, \omega]} \exp(-s^2 \|\mathbf{x} - \mathbf{y}\|^2) ds, \mathbf{x}, \mathbf{y} \in \mathbb{R}^3. \quad (3.2)$$

Using the Gaussian quadrature rule (see Proposition 3), we can evaluate numerically the integral in (3.2) as

$$\int_{[0, \omega]} \exp(-s^2 \|\mathbf{x} - \mathbf{y}\|^2) ds = \frac{\omega}{2} \int_{[-1, 1]} \exp\left(-\left(\frac{\omega}{2} + \frac{\omega}{2}z\right)^2 \|\mathbf{x} - \mathbf{y}\|^2\right) dz \approx \frac{\omega}{2} \sum_{i=1}^{N_{q_1}} w_i \exp\left(-\left(\frac{\omega}{2} + \frac{\omega}{2}z_i\right)^2 \|\mathbf{x} - \mathbf{y}\|^2\right), \quad (3.3)$$

where w_i are the Gaussian quadrature weights, z_i are the Gaussian quadrature nodes, and N_{q_1} is the number of quadrature points. The coordinates of \mathbf{x} and \mathbf{y} are denoted by $(x_1, x_2, x_3), (y_1, y_2, y_3)$ respectively. The exponential term in (3.3) can be written as

$$\exp\left(-\left(\frac{\omega}{2} + \frac{\omega}{2}z_i\right)^2 \|\mathbf{x} - \mathbf{y}\|^2\right) = \prod_{l=1}^3 \exp\left(-\left(\frac{\omega}{2} + \frac{\omega}{2}z_i\right)^2 (x_l - y_l)^2\right), l \in \{1, 2, 3\}. \quad (3.4)$$

Given the truncated computational box $[-b, b]^3, b \in \mathbb{R}$, each function of the form $\exp\left(-\left(\frac{\omega}{2} + \frac{\omega}{2}z_i\right)^2 (x_l - y_l)^2\right), i \in \{1, \dots, N_{q_1}\}, l \in \{1, 2, 3\}$ is smooth, differentiable (hence continuous) on $[-b, b]^2$, so that it is an excellent candidate for two-dimensional Chebyshev interpolation. According to Definition 2.1, the interpolated function can be written as

$$\exp\left(-\left(\frac{\omega}{2} + \frac{\omega}{2}z_i\right)^2 (x_l - y_l)^2\right) \approx \sum_{n_l, m_l=1}^{N_i} \alpha_{n_l m_l}^{(i)} T_{n_l}^{(i)}(x_l) T_{m_l}^{(i)}(y_l), \quad (3.5)$$

where N_i is the number of interpolation nodes for $i \in \{1, \dots, N_{q_1}\}$, $x_l, y_l \in [-b, b]$, and $l \in \{1, 2, 3\}$. We recall that among the advantages of using two-dimensional Chebyshev interpolation method is that forming two-dimensional Chebyshev grids $N_i \times N_i$ for each function (3.5) takes $\mathcal{O}(N_i^2)$ storage complexity, where N_i is the number of interpolation points needed. Furthermore, Chebyshev-Lobatto nodes can be obtained in linearithmic time using Fast Fourier Transform (FFT) [29]. This is one of the reasons for which we use Chebyshev basis. Our implementation that we discuss in more details in Section 7 uses FFTW [30] routine in Julia and the *chebfun2* library [24] to find the number of interpolation points N_i of the functions in (3.5). By replacing (3.5) and (3.3) in (3.2), the numerical approximation of the kernel $K(\mathbf{x}, \mathbf{y})$ becomes

$$K(\mathbf{x}, \mathbf{y}) \approx \frac{\omega}{\sqrt{\pi}} \sum_{i=1}^{N_{q_1}} w_i \left(\sum_{n_1, m_1, \dots, n_3, m_3=1}^{N_i} \prod_{l=1}^3 \alpha_{n_l m_l}^{(i)} T_{n_l}^{(i)}(x_l) T_{m_l}^{(i)}(y_l) \right), \quad (3.6)$$

where $\omega \geq 0$ is the parameter that regulates the separation range of the long-range/short-range interactions, $\alpha_{n_l m_l}^{(i)}$ are the N_i Chebyshev nodes, $T_{n_l}^{(i)}(x_l), T_{m_l}^{(i)}(y_l)$ are the Chebyshev polynomials (see Definition 2.1) and w_i are the Gaussian quadrature weights with $i \in \{1, \dots, N_{q_1}\}$. All along this paper, we denote N the maximum number of interpolation points in the tensorized Chebyshev grid in all directions such that $N = (\max(N_i)_{i \in \{1, \dots, N_{q_1}\}})^3$. The precomputation cost here to approximate the kernel (3.6) is $\mathcal{O}(N_{q_1} N^{\frac{1}{3}} (\log(N^{\frac{1}{3}}) + N^{\frac{1}{3}})) : \mathcal{O}(N_{q_1} N^{\frac{1}{3}} \log(N^{\frac{1}{3}}))$ FLOPS for the evaluation of the Chebyshev coefficient matrices using FFT algorithm, linearithmic in the number of interpolation points in a single direction $N^{\frac{1}{3}}$ and linear in the number of quadrature points, and $\mathcal{O}(N_{q_1} N^{\frac{2}{3}})$ FLOPS for forming the Chebyshev two-dimensional grids.

We consider now the finite six-dimensional integral $\mathcal{B}^{lr}(\mu, \nu, \kappa, \lambda)$ defined in (1.10) on the same truncated computational box $[-b, b]^3 \times [-b, b]^3, b \in \mathbb{R}$ with $\mu, \nu, \kappa, \lambda \in \{1, \dots, N_b\}$, where N_b is the number of basis functions that we defined in (1.8) and

b is the size of the computational box that is chosen according to the most slowly decaying basis functions. We discuss this aspect in more details in Section 6.2. By replacing $K(\mathbf{x}, \mathbf{y})$ with its approximation from (3.6), the numerical approximation of $\mathcal{B}^{lr}(\mu, \nu, \kappa, \lambda)$, denoted by $\mathcal{B}_{LTEI-TA}^{lr}(\mu, \nu, \kappa, \lambda)$, writes

$$\mathcal{B}_{LTEI-TA}^{lr}(\mu, \nu, \kappa, \lambda) = \frac{\omega}{\sqrt{\pi}} \sum_{i=1}^{N_{q1}} w_i \left(\int_{[-b,b]^3} \int_{[-b,b]^3} g_\mu(\mathbf{x}) g_\nu(\mathbf{x}) g_\kappa(\mathbf{y}) g_\lambda(\mathbf{y}) \left(\sum_{n_1, m_1, \dots, n_3, m_3=1}^{N_i} \prod_{l=1}^3 \alpha_{n_l m_l}^{(i)} T_{n_l}^{(i)}(\mathbf{x}_l) T_{m_l}^{(i)}(\mathbf{y}_l) \right) d\mathbf{x} d\mathbf{y} \right). \quad (3.7)$$

To obtain an efficient factorized representation of $\mathcal{B}_{LTEI-TA}^{lr}(\mu, \nu, \kappa, \lambda)$, we further consider the separability of the Gaussian primitives. Let $g_{\mu\nu}(\mathbf{x}) = g_\mu(\mathbf{x})g_\nu(\mathbf{x})$, $g_{\kappa\lambda}(\mathbf{y}) = g_\kappa(\mathbf{y})g_\lambda(\mathbf{y})$ such that according to (1.8) we have (showing only $g_{\mu\nu}$ expression)

$$g_{\mu\nu}(\mathbf{x}) = g_\mu(\mathbf{x})g_\nu(\mathbf{x}) = \sum_{j_1=1}^{I_\mu} \sum_{j_2=1}^{I_\nu} c_{j_1} c_{j_2} \prod_{l=1}^3 g_\mu^{(j_1)}(\mathbf{x}_l) g_\nu^{(j_2)}(\mathbf{x}_l) = \sum_{j=1}^{I_{\mu\nu}} c_j \prod_{l=1}^3 g_{\mu\nu}^{(j)}(\mathbf{x}_l), \quad (3.8)$$

where $I_{\mu\nu} = I_\mu I_\nu$, $c_j = c_{j_1} c_{j_2}$, $g_{\mu\nu}^{(j)}(\mathbf{x}_l) = g_\mu^{(j_1)}(\mathbf{x}_l) g_\nu^{(j_2)}(\mathbf{x}_l)$. Expressing the three dimensional function $g_{\mu\nu}(\mathbf{x})$ as a sum of separable functions is important to reduce the evaluation cost of $\mathcal{B}_{LTEI-TA}^{lr}(\mu, \nu, \kappa, \lambda)$ such that after replacing the Gaussian basis functions in (3.7) by their separable expression (3.8) we obtain

$$\mathcal{B}_{LTEI-TA}^{lr}(\mu, \nu, \kappa, \lambda) = \frac{\omega}{\sqrt{\pi}} \sum_{i=1}^{N_{q1}} w_i \left(\int_{[-b,b]^3} \int_{[-b,b]^3} g_{\mu\nu}(\mathbf{x}) g_{\kappa\lambda}(\mathbf{y}) \left(\sum_{n_1, m_1, \dots, n_3, m_3=1}^{N_i} \prod_{l=1}^3 \alpha_{n_l m_l}^{(i)} T_{n_l}^{(i)}(\mathbf{x}_l) T_{m_l}^{(i)}(\mathbf{y}_l) \right) d\mathbf{x} d\mathbf{y} \right) \quad (3.9)$$

$$= \frac{\omega}{\sqrt{\pi}} \sum_{i=1}^{N_{q1}} w_i \left(\int_{[-b,b]^3} \int_{[-b,b]^3} \left(\sum_{j=1}^{I_{\mu\nu}} \sum_{j'=1}^{I_{\kappa\lambda}} c_j c_{j'} \prod_{l=1}^3 g_{\mu\nu}^{(j)}(\mathbf{x}_l) g_{\kappa\lambda}^{(j')}(\mathbf{y}_l) \right) \left(\sum_{n_1, m_1, \dots, n_3, m_3=1}^{N_i} \prod_{l=1}^3 \alpha_{n_l m_l}^{(i)} T_{n_l}^{(i)}(\mathbf{x}_l) T_{m_l}^{(i)}(\mathbf{y}_l) \right) d\mathbf{x} d\mathbf{y} \right) \quad (3.10)$$

$$= \frac{\omega}{\sqrt{\pi}} \sum_{i=1}^{N_{q1}} w_i \sum_{j=1}^{I_{\mu\nu}} \sum_{j'=1}^{I_{\kappa\lambda}} c_j c_{j'} \underbrace{\sum_{\substack{n_1, n_2, n_3 \\ m_1, m_2, m_3=1}}^{N_i} \prod_{l=1}^3 \left(\alpha_{n_l m_l}^{(i)} \int_{[-b,b]} g_{\mu\nu}^{(j)}(\mathbf{x}_l) T_{n_l}^{(i)}(\mathbf{x}_l) d\mathbf{x}_l \int_{[-b,b]} g_{\kappa\lambda}^{(j')}(\mathbf{y}_l) T_{m_l}^{(i)}(\mathbf{y}_l) d\mathbf{y}_l \right)}_{\approx \mathbf{F}^{(i)}(j, j')}. \quad (3.11)$$

We note that the expression of $\mathcal{B}_{LTEI-TA}^{lr}(\mu, \nu, \kappa, \lambda)$ in (3.11) involves the numerical evaluation of one dimensional integrals. We associate each such integral with the element of a matrix and obtain two matrices $\mathbf{W}_{\mu\nu}^{(i,l)} \in \mathbb{R}^{I_{\mu\nu} \times N_i}$ and $\mathbf{W}_{\kappa\lambda}^{(i,l)} \in \mathbb{R}^{I_{\kappa\lambda} \times N_i}$ defined entry-wise as

$$\mathbf{W}_{\mu\nu}^{(i,l)}(j, n_l) = \int_{[-b,b]} g_{\mu\nu}^{(j)}(\mathbf{x}_l) T_{n_l}^{(i)}(\mathbf{x}_l) d\mathbf{x}_l \text{ and } \mathbf{W}_{\kappa\lambda}^{(i,l)}(j', m_l) = \int_{[-b,b]} g_{\kappa\lambda}^{(j')}(\mathbf{y}_l) T_{m_l}^{(i)}(\mathbf{y}_l) d\mathbf{y}_l. \quad (3.12)$$

We use one-dimensional Gaussian quadrature rule for the evaluation of (3.12). Their approximation is denoted by $\tilde{\mathbf{W}}_{\mu\nu}^{(i,l)}(j, n_l)$ (resp. $\tilde{\mathbf{W}}_{\kappa\lambda}^{(i,l)}(j', m_l)$). We further define matrices $\mathbf{F}^{(i)}$, $i \in \{1, \dots, N_{q1}\}$, as displayed in (3.11). By replacing the expressions of $\tilde{\mathbf{W}}_{\mu\nu}^{(i,l)}$ and $\tilde{\mathbf{W}}_{\kappa\lambda}^{(i,l)}$, we obtain

$$\mathbf{F}^{(i)}(j, j') = \sum_{\substack{n_1, n_2, n_3 \\ m_1, m_2, m_3=1}}^{N_i} \prod_{l=1}^3 \left(\alpha_{n_l m_l}^{(i)} \tilde{\mathbf{W}}_{\mu\nu}^{(i,l)}(j, n_l) \tilde{\mathbf{W}}_{\kappa\lambda}^{(i,l)}(j', m_l) \right). \quad (3.13)$$

By changing the order of summation in (3.13) and exploiting Khatri-Rao as well as Kronecker structures (see their definitions in Section 2.2), we obtain the factorized representation of $\mathcal{B}_{LTEI-TA}^{lr}$ as given in the following theorem.

Theorem 1. *The long-range two-electrons integrals has a factorized representation that writes*

$$\mathcal{B}_{LTEI-TA}^{lr}(\mu, \nu, \kappa, \lambda) = \frac{\omega}{\sqrt{\pi}} \sum_{i=1}^{N_{q1}} w_i \sum_{j=1}^{I_{\mu\nu}} \sum_{j'=1}^{I_{\kappa\lambda}} c_j c_{j'} \mathbf{F}^{(i)}(j, j'), \quad (3.14)$$

where $\mathbf{F}^{(i)} \in \mathbb{R}^{I_{\mu\nu} \times I_{\kappa\lambda}}$

$$\mathbf{F}^{(i)} = (\odot_{l=1}^3 \tilde{\mathbf{W}}_{\mu\nu}^{(i,l)}) (\otimes_{l=1}^3 \mathbf{A}^{(i)}) (\odot_{l=1}^3 \tilde{\mathbf{W}}_{\kappa\lambda}^{(i,l)})^\top = \odot_{l=1}^3 \tilde{\mathbf{W}}_{\mu\nu}^{(i,l)} \mathbf{A}^{(i)} \tilde{\mathbf{W}}_{\kappa\lambda}^{(i,l)\top}, \quad (3.15)$$

where $\mathbf{A}^{(i)} \in \mathbb{R}^{N_i \times N_i}$ are the Chebyshev coefficients matrices such that $\mathbf{A}^{(i)}(n_l, m_l) = \alpha_{n_l m_l}^{(i)}$ for $n_l, m_l \in [1, \dots, N_i]$, $l \in \{1, 2, 3\}$ with $\alpha_{n_l m_l}^{(i)}$ defined in (3.5), $\tilde{\mathbf{W}}_{\mu\nu}^{(i,l)} \in \mathbb{R}^{I_{\mu\nu} \times N_i}$ and $\tilde{\mathbf{W}}_{\kappa\lambda}^{(i,l)} \in \mathbb{R}^{I_{\kappa\lambda} \times N_i}$ are the numerical approximation of the one-dimensional integrals defined in (3.12).

Algorithm 1 computes the approximated entries $\mathcal{B}_{LTEI-TA}^{lr}(\mu, \nu, \kappa, \lambda)$ (3.14) given the coefficient matrix obtained from the two-dimensional Chebyshev interpolation $\mathbf{A}^{(i)} \in \mathbb{R}^{N_i \times N_i}$, for $i \in \{1, \dots, N_{q_1}\}$ and for any given pairs of $\mu, \nu, \kappa, \lambda$. This approach allows to reduce the storage complexity (resp. arithmetic complexity) to $\mathcal{O}\left(\sum_{i=1}^{N_{q_1}} N_i(N_i + I_{\mu\nu} + I_{\kappa\lambda})\right) \sim \mathcal{O}\left(N_{q_1} N^{\frac{1}{3}}(N^{\frac{1}{3}} + I_{\mu\nu} + I_{\kappa\lambda})\right)$ (resp. $\mathcal{O}\left(\sum_{i=1}^{N_{q_1}} N_i I_{\kappa\lambda}(N_i + I_{\mu\nu})\right) \sim \mathcal{O}\left(N_{q_1} N^{\frac{1}{3}} I_{\kappa\lambda}(N^{\frac{1}{3}} + I_{\mu\nu})\right)$), with $N^{\frac{1}{3}} = (\max(N_i)_{i \in \{1, \dots, N_{q_1}\}})$, instead of $\mathcal{O}(N(N + I_{\mu\nu} + I_{\kappa\lambda}))$ (resp. $\mathcal{O}(N I_{\kappa\lambda}(N + I_{\mu\nu} + I_{\kappa\lambda}))$), using naïve tensorized three dimensional quadrature on the computational box $[-b, b]^3$. Numerical results for this element-wise factorization are summarized in Section 7.

Algorithm 1 Compute $\mathcal{B}_{LTEI-TA}^{lr}(\mu, \nu, \kappa, \lambda)$

Input: Chebyshev coefficient matrices $\mathbf{A}^{(i)}$, $\mu, \nu, \kappa, \lambda$, w_i for $i \in \{1, \dots, N_{q_1}\}$.

Output: $\mathcal{B}_{LTEI-TA}^{lr}(\mu, \nu, \kappa, \lambda)$

Compute $\tilde{\mathbf{W}}_{\mu\nu}^{(i,1)}, \tilde{\mathbf{W}}_{\mu\nu}^{(i,2)}, \tilde{\mathbf{W}}_{\mu\nu}^{(i,3)} I_{\mu\nu} \times N_i$ matrices (according to (3.12)).

Compute $\tilde{\mathbf{W}}_{\kappa\lambda}^{(i,1)}, \tilde{\mathbf{W}}_{\kappa\lambda}^{(i,2)}, \tilde{\mathbf{W}}_{\kappa\lambda}^{(i,3)} I_{\kappa\lambda} \times N_i$ matrices (according to (3.12)).

Set $s = 0$.

for $i=1$ to N_{q_1} **do**

$$\mathbf{F}^{(i)} = \odot_{l=1}^3 \tilde{\mathbf{W}}_{\mu\nu}^{(i,l)} \mathbf{A}^{(i)} \tilde{\mathbf{W}}_{\kappa\lambda}^{(i,l)\top}.$$

$$s = s + w_i \sum_{j=1}^{I_{\mu\nu}} \sum_{j'=1}^{I_{\kappa\lambda}} c_j c_{j'} \mathbf{F}^{(i)}(j, j').$$

end for

$$\mathcal{B}_{LTEI-TA}^{lr}(\mu, \nu, \kappa, \lambda) = \frac{\omega}{\sqrt{\pi}} s.$$

3.2 Error bound of the two-electron integrals numerical approximation

In what follows, we give a theoretical error bound associated with the element-wise numerical approximation of $\mathcal{B}^{lr}(\mu, \nu, \kappa, \lambda)$ introduced in (3.14).

Proposition 5. *The element-wise error ϵ between the long-range two-electron integrals $\mathcal{B}^{lr}(\mu, \nu, \kappa, \lambda)$, given a finite box $[-b, b]^3$, and it's approximation $\mathcal{B}_{LTEI-TA}^{lr}$ can be bounded as follows*

$$|\epsilon| := |\mathcal{B}^{lr}(\mu, \nu, \kappa, \lambda) - \mathcal{B}_{LTEI-TA}^{lr}(\mu, \nu, \kappa, \lambda)| \leq c_1 \sup_{\mathbf{x}, \mathbf{y} \in [-b, b]^3} \left(\left\| \frac{d^{2N_{q_1}}}{ds^{2N_{q_1}}} f(s, \mathbf{x}, \mathbf{y}) \right\|_{\infty, [0, \omega]} \right) + \frac{\omega}{\sqrt{\pi}} c_2, \quad (3.16)$$

where we define the multivariate function

$$f(s, \mathbf{x}, \mathbf{y}) = \exp(-s^2 \|\mathbf{x} - \mathbf{y}\|^2), s \in [0, \omega], \mathbf{x}, \mathbf{y} \in [-b, b]^3. \quad (3.17)$$

ϵ is the approximation error, N_{q_1} is the number of quadrature points, c_1 and c_2 are defined in the following proof.

Proof. We start by introducing the following function

$$h(z_i) = \int_{[-b, b]^3} \int_{[-b, b]^3} g_{\mu\nu}(\mathbf{x}) g_{\kappa\lambda}(\mathbf{y}) \exp\left(-\left(\frac{\omega}{2} + \frac{\omega}{2} z_i\right)^2 \|\mathbf{x} - \mathbf{y}\|^2\right) d\mathbf{x} d\mathbf{y} \quad (3.18)$$

$$= \sum_{j=1}^{I_{\mu\nu}} \sum_{j'=1}^{I_{\kappa\lambda}} c_j c_{j'} \left(\prod_{l=1}^3 \int_{[-b, b]^2} g_{\mu\nu}^{(j)}(x_l) g_{\kappa\lambda}^{(j')}(y_l) \exp\left(-\left(\frac{\omega}{2} + \frac{\omega}{2} z_i\right)^2 (x_l - y_l)^2\right) dx_l dy_l \right), \quad (3.19)$$

with $z_i, i \in \{1, \dots, N_{q_1}\}$ being the Gaussian quadrature nodes. The upper bound of ϵ can be found as follows

$$|\epsilon| = \left| \mathcal{B}^{lr}(\mu, \nu, \kappa, \lambda) - \mathcal{B}_{LTEI-TA}^{lr}(\mu, \nu, \kappa, \lambda) \right| \leq \underbrace{\left| \mathcal{B}^{lr}(\mu, \nu, \kappa, \lambda) - \frac{\omega}{\sqrt{\pi}} \sum_{i=1}^{N_{q_1}} w_i h(z_i) \right|}_{\epsilon_1} + \underbrace{\left| \frac{\omega}{\sqrt{\pi}} \sum_{i=1}^{N_{q_1}} w_i h(z_i) - \mathcal{B}_{LTEI-TA}^{lr}(\mu, \nu, \kappa, \lambda) \right|}_{\epsilon_2}, \quad (3.20)$$

Using Proposition 3, triangle inequality, and Stirling formula given by $n! \approx \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$, ϵ_1 is bounded as follows

$$\epsilon_1 \leq c_1 \sup_{\mathbf{x}, \mathbf{y} \in [-b, b]^3} \left(\left\| \frac{d^{2N_{q_1}}}{ds^{2N_{q_1}}} f(s, \mathbf{x}, \mathbf{y}) \right\|_{\infty, [0, \omega]} \right), c_1 = \frac{2e N_{q_1}}{\sqrt{\pi}} b^6 \|g_{\mu\nu}\|_{\infty, [-b, b]^3} \|g_{\kappa\lambda}\|_{\infty, [-b, b]^3}, \quad (3.21)$$

with $e_{N_{q_1}} = \frac{\omega^{2N_{q_1}+1} e^{2N_{q_1}} (N_{q_1} \pi)^{\frac{1}{2}}}{2^{6N_{q_1}+1} N_{q_1}^{2N_{q_1}} (2N_{q_1}+1)}$. The error bound of ϵ_2 needs a more detailed explanation. We replace $\mathcal{B}_{LTI-TA}^{lr}(\mu, \nu, \kappa, \lambda)$ by its expression defined in (3.14) such that

$$|\epsilon_2| = \left| \frac{\omega}{\sqrt{\pi}} \sum_{i=1}^{N_{q_1}} w_i \left(h(z_i) - \sum_{j=1}^{I_{\mu\nu}} \sum_{j'=1}^{I_{\kappa\lambda}} c_j c_{j'} \mathbf{F}^{(i)}(j, j') \right) \right| \leq \frac{\omega}{\sqrt{\pi}} \sum_{i=1}^{N_{q_1}} |w_i| \left| h(z_i) - \sum_{j=1}^{I_{\mu\nu}} \sum_{j'=1}^{I_{\kappa\lambda}} c_j c_{j'} \mathbf{F}^{(i)}(j, j') \right|, \quad (3.22)$$

with $\mathbf{F}^{(i)}(j, j')$ being defined in (3.13). Using the triangle inequality, the expression of $\left| h(z_i) - \sum_{j=1}^{I_{\mu\nu}} \sum_{j'=1}^{I_{\kappa\lambda}} c_j c_{j'} \mathbf{F}^{(i)}(j, j') \right|$, for $i \in \{1, \dots, N_{q_1}\}$, can be bounded as follows

$$\left| h(z_i) - \sum_{j=1}^{I_{\mu\nu}} \sum_{j'=1}^{I_{\kappa\lambda}} c_j c_{j'} \mathbf{F}^{(i)}(j, j') \right| \leq \sum_{j=1}^{I_{\mu\nu}} \sum_{j'=1}^{I_{\kappa\lambda}} c_j c_{j'} \left| \prod_{l=1}^3 \int_{[-b, b]^2} g_{\mu\nu}^{(j)}(x_l) g_{\kappa\lambda}^{(j')}(y_l) e^{-\left(\frac{\omega}{2} + \frac{\omega}{2} z_i\right)^2 (x_l - y_l)^2} dx_l dy_l - \mathbf{F}^{(i)}(j, j') \right| \quad (3.23)$$

In order to evaluate the bound of (3.23), one needs to evaluate the error bound of the following expression using Proposition 2 and Proposition 3

$$\left| \int_{[-b, b]^2} g_{\mu\nu}^{(j)}(x_l) g_{\kappa\lambda}^{(j')}(y_l) e^{-\left(\frac{\omega}{2} + \frac{\omega}{2} z_i\right)^2 (x_l - y_l)^2} - \sum_{n_l, m_l} \alpha_{n_l m_l}^{(i)} \tilde{\mathbf{W}}_{\mu\nu}^{(i, l)}(j, n_l) \tilde{\mathbf{W}}_{\kappa\lambda}^{(i, l)}(j', m_l) \right| \leq \beta_i, l \in \{1, 2, 3\}, \quad (3.24)$$

where for $i \in \{1, \dots, N_{q_1}\}$, $j \in \{1, \dots, I_{\mu\nu}\}$, and $j' \in \{1, \dots, I_{\kappa\lambda}\}$, β_i is defined as follows

$$\begin{aligned} \beta_i = & (2b)^2 \left\| g_{\mu\nu}^{(j)} \right\|_{\infty, [-b, b]} \left\| g_{\kappa\lambda}^{(j')} \right\|_{\infty, [-b, b]} e_{N_i} \\ & + e_{N_{q_2}} \left(\left\| g_{\kappa\lambda}^{(j')} \right\|_{\infty, [-b, b]} \left\| \frac{d^{2N_{q_2}} \left(g_{\mu\nu}^{(j)} T_{n_1}^{(i)} \right) (x)}{dx^{2N_{q_2}}} \right\|_{\infty, [-b, b]} + N_{q_2} \max_{1 \leq i \leq N_{q_2}} (w_i) \left\| g_{\mu\nu}^{(j)} \right\|_{\infty, [-b, b]} \left\| \frac{d^{2N_{q_2}} \left(g_{\kappa\lambda}^{(j')} T_{m_1}^{(i)} \right) (y)}{dy^{2N_{q_2}}} \right\|_{\infty, [-b, b]} \right), \end{aligned} \quad (3.25)$$

with $e_{N_{q_2}} = \frac{(2b)^{2N_{q_2}+1} e^{2N_{q_2}} (N_{q_2} \pi)^{\frac{1}{2}}}{2^{6N_{q_2}+1} N_{q_2}^{2N_{q_2}} (2N_{q_2}+1)}$. The term e_{N_i} is defined as follows

$$e_{N_i} = \frac{b^{N_i+1}}{2^{1+N_i}} \frac{e^{-N_i+1}}{\sqrt{2\pi} (N_i+1) \left(1 + N_i^{\frac{1}{3}} (N_i+1)\right)} \left[\max_{-b \leq \xi \leq b} \left| \frac{\partial^{N_i+1} \mathcal{F}(z_i, \xi, y)}{\partial \xi^{N_i+1}} \right| + \delta \max_{-b \leq \xi, \eta \leq b} \left| \frac{\partial^{N_i+1} \mathcal{F}(z_i, \xi, \eta)}{\partial \eta^{N_i+1}} \right| \right], \quad (3.27)$$

where $\mathcal{F}(z_i, x, y) = e^{-\left(\frac{\omega}{2} + \frac{\omega}{2} z_i\right)^2 (x-y)^2}$ with z_i being the Gaussian quadrature points and δ is defined in (2). Now, by factorizing (3.23) and using (3.24), one arrives at the desired error bound of ϵ_2

$$\epsilon_2 \leq \frac{\omega}{\sqrt{\pi}} c_2, \quad (3.28)$$

with

$$c_2 = N_{q_1} \sup_{1 \leq i \leq N_{q_1}} \left(|w_i| \sum_j \sum_{j'} c_j c_{j'} (2b)^4 \left\| g_{\mu\nu}^{(j)} \right\|_{\infty, [-b, b]}^2 \left\| g_{\kappa\lambda}^{(j')} \right\|_{\infty, [-b, b]}^2 \sup_{1 \leq n_1, m_1 \leq N_i} \left((1 + N_i \alpha_{n_1, m_1}^{(i)} + (N_i \alpha_{n_1, m_1}^{(i)})^2) \beta_i \right) \right). \quad (3.29)$$

□

All along this work, we consider a fixed number of quadrature points N_{q_2} for the evaluation of (3.12) and the study is not being done on the parameter N_{q_2} since the computations using these one-dimensional Gaussian quadrature to evaluate (3.12) are involved in the precomputation steps. As we notice here, the approximation error depends on the value of ω , the number of quadrature points N_{q_1} , the regularity of the function f , the Gaussian-type functions and on the dimension of the hypercube.

3.3 A new decomposition of TEI tensor \mathcal{B}^{lr} through TA approach

As already discussed in the introduction, one of the main steps in many methods in quantum chemistry involves the application of the two-electron integrals tensor $\mathcal{B}^{lr} \in \mathbb{R}^{N_b \times N_b \times N_b \times N_b}$ to a vector with N_b^2 elements or a set of such vectors. To perform efficiently this contraction operation, we introduce in this section a factorized representation of the fourth-order two-electron integrals tensor \mathcal{B}^{lr} that expands the factorized representation of its elements summarized in Theorem 1. We show also that the obtained tensorized structure is beneficial to accelerate contraction operations.

3.3.1 Factorized expression of \mathcal{B}^{lr}

In what follows we derive the factorized representation of \mathbf{B}^{lr} (mode-(1,2) matricization of $\mathcal{B}^{lr} \in \mathbb{R}^{N_b \times N_b \times N_b \times N_b}$). We slightly modify the expression of the approximation of the two-electron integrals (see Theorem 1) by changing the order of summation to obtain

$$\mathcal{B}_{LTEI-TA}^{lr}(\mu, \nu, \kappa, \lambda) = \frac{\omega}{\sqrt{\pi}} \sum_{i=1}^{N_{q1}} w_i \left[\sum_{\substack{n_1, n_2, n_3 \\ m_1, m_2, m_3=1}}^{N_i} \left(\sum_{j=1}^{I_{\mu\nu}} c_j \prod_{l=1}^3 \tilde{\mathbf{W}}_{\mu\nu}^{(i,l)}(j, n_l) \right) \prod_{l=1}^3 \alpha_{n_l m_l}^{(i)} \left(\sum_{j'=1}^{I_{\kappa\lambda}} c_{j'} \prod_{l=1}^3 \tilde{\mathbf{W}}_{\kappa\lambda}^{(i,l)}(j', m_l) \right) \right]. \quad (3.30)$$

We introduce the matrices $\mathbf{M}_{TA}^{(i)} \in \mathbb{R}^{N_b^2 \times N_i^3}$ with single entries $\sum_{j=1}^{I_{\mu\nu}} c_j \prod_{l=1}^3 \tilde{\mathbf{W}}_{\mu\nu}^{(i,l)}(j, n_l), \mu, \nu \in \{1, \dots, N_b\}, n_l \in \{1, \dots, N_i\}, l \in \{1, 2, 3\}$ such that the approximation of mode-(1,2) matricization of $\mathcal{B}_{LTEI-TA}^{lr}$, referred to as $\mathbf{B}_{LTEI-TA}^{lr}$, writes

$$\mathbf{B}_{LTEI-TA}^{lr} = \frac{\omega}{\sqrt{\pi}} \sum_{i=1}^{N_{q1}} w_i \mathbf{M}_{TA}^{(i)} (\otimes_{l=1}^3 \mathbf{A}^{(i)}) \mathbf{M}_{TA}^{(i)\top} \in \mathbb{R}^{N_b^2 \times N_b^2} \text{ and } \mathbf{M}_{TA}^{(i)}(\mu\nu, n_1 n_2 n_3) = \sum_{j=1}^{I_{\mu\nu}} c_j \prod_{l=1}^3 \tilde{\mathbf{W}}_{\mu\nu}^{(i,l)}(j, n_l), i \in \{1, \dots, N_{q1}\}, \quad (3.31)$$

where $\mathbf{A}^{(i)} = \left(\alpha_{n_l m_l}^{(i)} \right)_{n_l, m_l \in [1, \dots, N_i]}, l \in \{1, 2, 3\}$ are the coefficient matrices obtained from the two-dimensional Chebychev interpolation (see Definition 3.5).

3.3.2 Fast evaluation of tensor products

In practice, we only need to compute the matrix $\mathbf{M}_{TA}^{(i)}$ with the maximum number of interpolation points N . We denote this matrix by $\mathbf{M}_{TA, max} \in \mathbb{R}^{N_b^2 \times N}$. In fact, the other matrices $\mathbf{M}_{TA}^{(i)} \in \mathbb{R}^{N_b^2 \times N_i^3}, N_i^3 \leq N, i \in \{1, \dots, N_{q1}\}$ have common entries with $\mathbf{M}_{TA, max}$. For example, given the two following matrices $\mathbf{M}_{TA}^{(i)} \in \mathbb{R}^{N_b^2 \times N_i}$ and $\mathbf{M}_{TA}^{(j)} \in \mathbb{R}^{N_b^2 \times N_j}$ with $N_i < N_j$ and $i, j \in \{1, \dots, N_{q1}\}$, we have

$$\mathbf{M}_{TA}^{(i)}(\mu\nu, n_1 n_2 n_3) = \mathbf{M}_{TA}^{(j)}(\mu\nu, n_1 n_2 n_3), n_1, n_2, n_3 \in \{1, \dots, N_i\}. \quad (3.32)$$

This can also be illustrated in Figure 1. Therefore, the storage complexity for storing $\mathbf{M}_{TA, max}$ is $\mathcal{O}(NN_b^2)$.

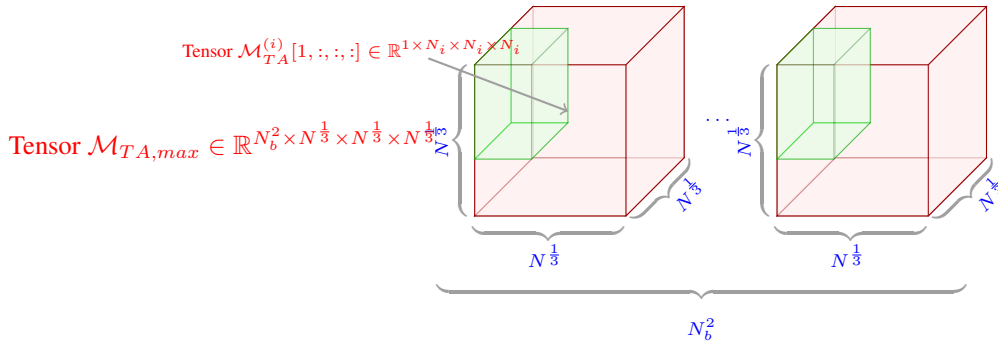


Figure 1: $\mathcal{M}_{TA, max} \in \mathbb{R}^{N_b^2 \times N^{\frac{1}{3}} \times N^{\frac{1}{3}} \times N^{\frac{1}{3}}}$ is the tensorization of $\mathbf{M}_{TA, max} \in \mathbb{R}^{N_b^2 \times N}$.

By doing so, we can extract $\mathcal{M}_{TA}^{(i)} \in \mathbb{R}^{N_b^2 \times N_i \times N_i \times N_i}$ tensors that we unfold back to matrices $\mathbf{M}_{TA}^{(i)} \in \mathbb{R}^{N_b^2 \times N_i^3}$ by mode-1 matricization defined in (2.19). We can exploit the tensorized structure of the factorized long-range two-electron integral tensor in equation (3.31) to reduce the application cost of the product between the tensorized form $\otimes_{l=1}^3 \mathbf{A}^{(i)} \in \mathbb{R}^{N_i^3 \times N_i^3}$ and $\mathbf{M}_{TA}^{(i)} \in \mathbb{R}^{N_b^2 \times N_i^3}$ from $\mathcal{O}(N_i^6 N_b^2)$ to $\mathcal{O}(N_i^4 N_b^2)$. Given the Definition 2.19, the product $\left(\otimes_{l=1}^3 \mathbf{A}^{(i)} \right) \mathbf{M}_{TA}^{(i)\top}$ can be defined entry-wise by

$$\begin{aligned} \left(\left(\otimes_{l=1}^3 \mathbf{A}^{(i)} \right) \mathbf{M}_{TA}^{(i)\top} \right)(n, j) &= \sum_{m_1, m_2, m_3=1}^{N_i} \left(\prod_{l=1}^3 \mathbf{A}^{(i)}(n_l, m_l) \right) \mathbf{M}_{TA}^{(i)\top}(m_1 m_2 m_3, j) \\ &= \sum_{m_1, m_2, m_3=1}^{N_i} \left(\prod_{l=1}^2 \mathbf{A}^{(i)}(n_l, m_l) \right) \left(\mathbf{A}^{(i)} \mathbf{M}_{TA, (4)}^{(i)\top} \right)(n_3, m_1 m_2 j), \end{aligned} \quad (3.33)$$

where $\mathbf{M}_{TA,(4)}^{(i)\top}$ is the mode-4 matricization (see Definition 2.19) of the fourth order tensor $\mathcal{M}_{TA}^{(i)} \in \mathbb{R}^{N_b^2 \times N_i \times N_i \times N_i}$. From (3.33), we notice that we need to perform three times the matrix-matrix products of sizes $N_i \times N_i$ and $N_i \times N_i^2 N_b^2$, leading to an overall time complexity of $\mathcal{O}(3N_i^4 N_b^2) \sim \mathcal{O}(N_i^4 N_b^2)$. If we want to compute the whole tensor, we need to sum over $i \in \{1, \dots, N_{q1}\}$ which yields to a complexity of $\mathcal{O}\left(N^{\frac{4}{3}} N_b^2\right)$ with $N = \left(\max(N_i)_{i \in \{1, \dots, N_{q1}\}}\right)^3$.

Indeed, in practical applications the whole two-electron integrals tensor does not need to be evaluated but it is rather kept in its tensorized structure to benefit from fast matrix operations when applying it to vectors or matrices. We will discuss in more details an application case in Section 5. An important point when implementing these tensor product evaluations is that the presented method can benefit from BLAS operations [31]. Indeed, (3.33) can be interpreted as the application of a sequence of products of permutation matrices and block-diagonal matrices (with the same blocks $\mathbf{A}^{(i)}$ along the diagonal) to $\mathbf{M}_{TA,(k)}^{(i)}$, $k \in \{2, 3, 4\}$. Matrix-vector products with block-diagonal matrices of this form can be numerically reformulated as matrix-matrix products between one of these diagonal blocks and a matrix composed of the concatenation of subvectors of the original one [32]. Since matrix-matrix products can be performed more efficiently than matrix-vector products using BLAS routines (namely BLAS-3 instead of BLAS-2), this optimization results in efficient implementations. In our case, we have even larger concatenation of subvectors because we apply these tensor products to matrices (not simply vectors), resulting in even better exploitation of BLAS-3 routines.

4 Long-range TEI tensor factorization through Fast Multipoles Methods (LTEI-FMM)

In what follows, we recall briefly Fast Multipole Methods FMM and its application in our problem after demonstrating that our kernel is *asymptotically smooth*. As many methods taking advantage of tree space decomposition [20, 21, 33], FMM rely on an important property of usual kernels. We discuss also the similarities and differences between LTEI-TA and LTEI-FMM approaches to approximate \mathcal{B}^{lr} .

Definition 4.1 (Definition 5.1 in [34]). *A kernel $K(\cdot, \cdot) : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}$ is said to be asymptotically smooth if there exist two constants c_1, c_2 and a singularity degree $\sigma \in \mathbb{N}_0$ such that $\forall z \in \{\mathbf{x}_l, \mathbf{y}_l\} \in \mathbb{R}, \forall n \in \mathbb{N}_0, \forall \mathbf{x} \neq \mathbf{y}$,*

$$\left| \frac{\partial^n}{\partial z^n} K(\mathbf{x}, \mathbf{y}) \right| \leq n! c_1 (c_2 \|\mathbf{x} - \mathbf{y}\|)^{-n-\sigma}.$$

Based on this property, efficient hierarchical schemes can be derived for the evaluation of N -body problems involving asymptotically smooth kernels.

4.1 Fast Multipole Methods

Considering two point clouds with $N_{\mathbf{x}}, N_{\mathbf{y}}$ points, where we denote these clouds by $\{\mathbf{x}_n\}_{n=1}^{N_{\mathbf{x}}}, \{\mathbf{y}_n\}_{n=1}^{N_{\mathbf{y}}} \subset \mathbb{R}^3$ (whose elements are referred to as 3D points or *particles*), $q : \{\mathbf{y}_n\}_{n=1}^{N_{\mathbf{y}}} \rightarrow \mathbb{C}$ and an asymptotically smooth function $K : (\mathbb{R}^3 \times \mathbb{R}^3) \setminus \{\mathbf{0}\} \rightarrow \mathbb{C}$, one may express the associated N -body problem as the computation of $p : \{\mathbf{x}_n\}_{n=1}^{N_{\mathbf{x}}} \rightarrow \mathbb{C}$ such that

$$p(\mathbf{x}) := \sum_{\mathbf{y} \in \{\mathbf{y}_n\}_{n=1}^{N_{\mathbf{y}}}} K(\mathbf{x}, \mathbf{y}) q(\mathbf{y}). \quad (4.1)$$

Computing p naively requires $\mathcal{O}(N^2)$ floating point operations, with $N = \max(N_{\mathbf{x}}, N_{\mathbf{y}})$. Thanks to hierarchical methods, such as *hierarchical matrices* or *Fast Multipole Methods* (FMM), this complexity can be reduced to $\mathcal{O}(N \log N)$ or even $\mathcal{O}(N)$ (but at the cost of an error we can control). These methods rely on decompositions of $\{\mathbf{x}_n\}_{n=1}^{N_{\mathbf{x}}}$ and $\{\mathbf{y}_n\}_{n=1}^{N_{\mathbf{y}}}$ into groups of particles whose interaction can be efficiently performed through low-rank matrix approximations if their distance is sufficiently large compared to their radius. For non-oscillatory kernels K , FMMs are able to reach the $\mathcal{O}(N)$ complexity, so that they are attractive algorithm for efficiently solving N -body problems.

Among the different formulation of FMMs, we seek for particular features needed for our application case. Indeed, the method has to:

- perform efficiently (actually in a linear time with respect to the number of points) on highly non-uniform point distributions, such as the three-dimensional Chebyshev grids,
- handle the kernel K (which is non-standard kernel in the FMM community),
- be able to reach the precision required in realistic chemistry applications.

4.2 Application to two-electron integrals (TEI)

First, in order to exploit FMM on the two-electron integrals, one has to check that the underlying kernel is asymptotically smooth (see Definition 4.1). In our case, we want the FMM to act on the long-range kernel $K(\mathbf{x}, \mathbf{y})$, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$ (see (3.2)), which leads us to demonstrate the result of Proposition 6.

Proposition 6. $K(\mathbf{x}, \mathbf{y}) = \frac{\text{erf}(\omega \|\mathbf{x} - \mathbf{y}\|)}{\|\mathbf{x} - \mathbf{y}\|}$, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3, 0 \leq \omega < \infty$ is asymptotically smooth.

Proof. Given the function $K(\mathbf{x}, \mathbf{y}) = \frac{\text{erf}(\omega \|\mathbf{x} - \mathbf{y}\|)}{\|\mathbf{x} - \mathbf{y}\|}$, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$, $0 \leq \omega < \infty$, we want to evaluate the function's partial derivative upper bound with respect to $x_1 \in \mathbb{R}$ such that $\forall n \in \mathbb{N}_0, \forall \mathbf{x} \neq \mathbf{y}$, the n th derivative of $K(\mathbf{x}, \mathbf{y})$ with respect to x_1 writes

$$\frac{\partial^n}{\partial x_1^n} K(\mathbf{x}, \mathbf{y}) = \frac{\partial^n}{\partial x_1^n} \left(\frac{2}{\sqrt{\pi}} \int_0^\omega \exp(-s^2 \|\mathbf{x} - \mathbf{y}\|^2) ds \right) \quad (4.2)$$

$$= \frac{2^{n+1}}{\sqrt{\pi}} n! \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \frac{(-1)^{n-2k} 2^{-2k} (x_1 - y_1)^{n-2k}}{k!(n-2k)!} s^{2n-2k} \exp(-s^2 \|\mathbf{x} - \mathbf{y}\|^2) ds. \quad (4.3)$$

If n is even, the term under the integral in (4.2) is positive. Otherwise, it is either negative or positive. Therefore, (4.2) can be bounded by the absolute value of the n th derivative of the Coulomb potential that writes

$$\frac{\partial^n}{\partial x_1^n} \frac{1}{\|\mathbf{x} - \mathbf{y}\|} = \frac{2^{n+1}}{\sqrt{\pi}} n! \int_0^\infty \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \frac{(-1)^{n-2k} 2^{-2k} (x_1 - y_1)^{n-2k}}{k!(n-2k)!} s^{2n-2k} \exp(-s^2 \|\mathbf{x} - \mathbf{y}\|^2) ds, \quad (4.4)$$

and

$$\frac{\partial^n}{\partial x_1^n} K(\mathbf{x}, \mathbf{y}) \leq \left| \frac{\partial^n}{\partial x_1^n} \frac{1}{\|\mathbf{x} - \mathbf{y}\|} \right|. \quad (4.5)$$

Since $\frac{1}{\|\mathbf{x} - \mathbf{y}\|}$ is asymptotically smooth [35, 36], this shows that $K(\mathbf{x}, \mathbf{y})$ is also asymptotically smooth.

This proof applies for all the other directions. \square

Hence, thanks to the asymptotically smooth behavior of K , FMM can be applied to this kernel and the far field contribution of the N -body problem can be efficiently approximated, especially by exploiting polynomial interpolation. Similar to the previous sections, we consider the finite six-dimensional integral $\mathcal{B}^{lr}(\mu, \nu, \kappa, \lambda)$ defined in (1.10) on a truncated computational box $[-b, b]^3 \times [-b, b]^3$, $b \in \mathbb{R}$ as follows

$$\mathcal{B}^{lr}(\mu, \nu, \kappa, \lambda) = \int_{[-b, b]^3} \int_{[-b, b]^3} g_{\mu\nu}(\mathbf{x}) K(\mathbf{x}, \mathbf{y}) g_{\kappa\lambda}(\mathbf{y}) d\mathbf{x} d\mathbf{y}. \quad (4.6)$$

Instead of applying Gaussian quadrature rule on the kernel $K(\mathbf{x}, \mathbf{y})$ as we did in the previous Section 3, we use Chebyshev polynomials evaluated in a six-dimensional Chebyshev grid, the low-rank approximation of $K(\mathbf{x}, \mathbf{y})$ can be written, as explained in [20], as follows

$$K(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^N L(\mathbf{x}_i, \mathbf{x}) \underbrace{\sum_{j=1}^N K(\mathbf{x}_i, \mathbf{y}_j) L(\mathbf{y}_j, \mathbf{y})}_{N\text{-body problem as in Eq. 4.1}}, \quad (4.7)$$

where N is the total number of Chebyshev interpolation points (we use the same N as the one introduced in Section 3), $\mathbf{x}_i = (\mathbf{x}_{i_1}, \mathbf{x}_{i_2}, \mathbf{x}_{i_3})$ and $\mathbf{y}_i = (\mathbf{y}_{i_1}, \mathbf{y}_{i_2}, \mathbf{y}_{i_3})$, for $i \in \{1, 2, \dots, N\}$, are 3-vectors of Chebyshev points with $i_l, j_l \in \{1, \dots, N\}$, $l \in \{1, 2, 3\}$. We also have

$$L(\mathbf{x}_i, \mathbf{x}) = L^{(1)}(\mathbf{x}_{i_1}, \mathbf{x}_1) L^{(2)}(\mathbf{x}_{i_2}, \mathbf{x}_2) L^{(3)}(\mathbf{x}_{i_3}, \mathbf{x}_3). \quad (4.8)$$

$$L^{(l)}(\mathbf{x}_{i_l}, \mathbf{x}_l) = \frac{1}{N^{\frac{1}{3}}} + \frac{2}{N^{\frac{1}{3}}} \sum_{k=2}^{N^{\frac{1}{3}}} T_k(\mathbf{x}_{i_l}) T_k(\mathbf{x}_l), \quad l \in \{1, 2, 3\}. \quad (4.9)$$

One may notice that the equation (4.9) appears as a simple reformulation of the interpolation presented in Definition (2.1), combining the equation (2.2) and the equation (2.3). The important point here is that we want the kernel to explicitly appear (evaluated on Chebyshev interpolation nodes) in the expression, so that a FMM algorithm can be derived, following [20, 32]. Chebyshev polynomials are used here as interpolation basis and were already defined in Definition 2.1. The long-range two-electron integrals in (1.10) can be written as follows

$$\begin{aligned} \mathcal{B}_{LTI-FMM}^{lr}(\mu, \nu, \kappa, \lambda) &= \sum_{i=1}^N \underbrace{\int_{[-b, b]^3} g_{\mu\nu}(\mathbf{x}) L(\mathbf{x}_i, \mathbf{x}) d\mathbf{x}}_{\mathbf{Z}_{\mu\nu}(1, \mathbf{x}_i)} \left(\sum_{j=1}^N K(\mathbf{x}_i, \mathbf{y}_j) \underbrace{\int_{[-b, b]^3} g_{\kappa\lambda}(\mathbf{y}) L(\mathbf{y}_j, \mathbf{y}) d\mathbf{y}}_{\mathbf{Z}_{\kappa\lambda}(\mathbf{y}_j, 1)} \right) \\ &= \sum_{i=1}^N \mathbf{Z}_{\mu\nu}(1, \mathbf{x}_i) \left(\sum_{j=1}^N K(\mathbf{x}_i, \mathbf{y}_j) \mathbf{Z}_{\kappa\lambda}(\mathbf{y}_j, 1) \right). \end{aligned} \quad (4.10)$$

Equation (4.10) can be written in matrix formulation as follows for fixed $\mu, \nu, \kappa, \lambda \in \{1, \dots, N_b\}$

$$\mathcal{B}_{LTEI-FMM}^{lr}(\mu, \nu, \kappa, \lambda) = \mathbf{Z}_{\mu\nu} \mathbf{K} \mathbf{Z}_{\kappa\lambda}^\top, \mathbf{Z}_{\mu\nu}, \mathbf{Z}_{\kappa\lambda} \in \mathbb{R}^{1 \times N}, \mathbf{K} \in \mathbb{R}^{N \times N} \quad (4.11)$$

with $\mathbf{K}(\mathbf{x}_i, \mathbf{y}_j) = \frac{erf(\omega \|\mathbf{x}_i - \mathbf{y}_j\|)}{\|\mathbf{x}_i - \mathbf{y}_j\|}, i, j \in \{1, \dots, N\}$. The last term into parenthesis in (4.10) corresponds to an N -body problem as in Equation (4.1), whose evaluation can be performed in $\mathcal{O}(N)$ FLOPS using FMM. One may notice that the FMM accuracy can be chosen accordingly to the interpolation error in equation (4.10). For all $\mu, \nu, \kappa, \lambda \in \{1, \dots, N_b\}$, the factorized representation of the mode-(1,2) matricization of the fourth-order tensor $\mathcal{B}_{LTEI-FMM}^{lr}$ (4.10) is then given by

$$\mathbf{B}_{LTEI-FMM}^{lr} = \mathbf{M}_{FMM} \mathbf{K} \mathbf{M}_{FMM}^\top \in \mathbb{R}^{N_b^2 \times N_b^2}, \mathbf{M}_{FMM} \in \mathbb{R}^{N_b^2 \times N}, \quad (4.12)$$

with $\mathbf{M}_{FMM}[\mu\nu, :] = \mathbf{Z}_{\mu\nu} \in \mathbb{R}^{1 \times N}$, for $\mu, \nu \in \{1, \dots, N_b\}$. Hence, the entire computation of (4.12) requires the application of the FMM method to each column of \mathbf{M}_{FMM} , the overall evaluation complexity of FMM becomes $\mathcal{O}(N \times N_b^2)$ to compute $\mathbf{K} \mathbf{M}_{FMM}^\top$.

Remark 1. The FMM formulation we opted for relies on precomputations (at a linear cost with respect to the number of particles) for the construction of low-rank approximations (see Section 4.1) that depends only on the particle distribution. Because the interpolation points are the same for each $\mathbf{Z}_{\kappa\lambda}(\mathbf{y}_j, 1)$, our particle distributions do not change, so these precomputations can be performed only once and reused for each FMM application.

4.3 Similarities and differences between LTEI-TA and LTEI-FMM approaches

In table 1 we summarize the approximated expressions of (1.10) obtained through LTEI-TA and LTEI-FMM approaches.

Table 1: Factorization of TEI

Approaches	LTEI-TA	LTEI-FMM
Distribution	N Chebyshev points	N Chebyshev points
Entry-wise evaluation:	$\mathcal{B}_{LTEI-TA}^{lr}(\mu, \nu, \kappa, \lambda) := \frac{\omega}{\sqrt{\pi}} \sum_{i=1}^{N_{q1}} \left(w_i \sum_{j=1}^{I_{\mu\nu}} \sum_{j'=1}^{I_{\kappa\lambda}} c_j c_{j'} \mathbf{F}^{(i)}(j, j') \right)$	$\mathcal{B}_{LTEI-FMM}^{lr}(\mu, \nu, \kappa, \lambda) := \mathbf{Z}_{\mu\nu} \mathbf{K} \mathbf{Z}_{\kappa\lambda}^\top$
Factorized representation:	$\mathbf{B}_{LTEI-TA}^{lr} := \frac{\omega}{\sqrt{\pi}} \sum_{i=1}^{N_{q1}} w_i \mathbf{M}_{TA}^{(i)} (\otimes_{l=1}^3 \mathbf{A}^{(i)}) \mathbf{M}_{TA}^{(i)\top} \in \mathbb{R}^{N_b^2 \times N_b^2}$	$\mathbf{B}_{LTEI-FMM}^{lr} := \mathbf{M}_{FMM} \mathbf{K} \mathbf{M}_{FMM}^\top \in \mathbb{R}^{N_b^2 \times N_b^2}$

We discuss here the differences and similarities between both approaches. On one hand, for TA approach, we start by applying a change of variable to the long-range kernel $K(\mathbf{x}, \mathbf{y})$ (3.2) in order to remove the term $\frac{1}{\|\mathbf{x} - \mathbf{y}\|}$, then we apply one-dimensional Gaussian quadrature (see (3.3)) with N_{q1} quadrature points. In addition to that, we apply two-dimensional Chebyshev interpolation which yields to obtaining a tensorized form obtained in (3.14), (3.31). Thus, we need to evaluate $\mathbf{M}_{TA, max} \in \mathbb{R}^{N_b^2 \times N}$ which involves the evaluation of one-dimensional integrals over $[-b, b]$. On the other hand, when applying interpolation directly on the original kernel K , one ends up with a three dimensional N -body problem that can be efficiently handled using FMM approach. Thus, we need to compute $\mathbf{M}_{FMM} \in \mathbb{R}^{N_b^2 \times N}$ which involves also the evaluation of one-dimensional integrals over $[-b, b]$. The similarities between both approaches consist in employing Chebyshev interpolation with the same total number of interpolation points N .

Remark 2. One may mention that for low level optimisations (such as explicit formula for the polynomials or fast FFT-based assembling of the interpolation coefficients), we opted for slightly different interpolation nodes in the two methods. Indeed, Gauss-Chebyshev-Lobatto nodes are used for LTEI-TA method while Chebyshev nodes are used for LTEI-FMM. These last points are defined as (showing only \mathbf{x}_{i_l} expression)

$$\mathbf{x}_{i_l} = \cos\left(\frac{2k-1}{2N^{\frac{1}{3}}}\pi\right), k \in \left\{1, \dots, N^{\frac{1}{3}}\right\}, l \in \{1, 2, 3\}. \quad (4.13)$$

However, for both cases, the same number of interpolation nodes is considered for a given targeted precision, $N^{\frac{1}{3}}$ per direction, so that this detail does not impact the complexity estimates and the comparison between them.

5 Application to electronic structure calculations

We describe in what follows an application case for the two-electron integrals tensor using LTEI-TA as well as LTEI-FMM. In quantum chemistry, one of the main steps in many methods is the construction of the Coulomb matrix [15, 37, 38]. We define in the following the long-range Coulomb matrix in the molecular orbital basis ϕ_i that are represented (approximately) as [3]

$$\phi_i = \sum_{\mu=1}^{N_b} q_{i\mu} g_\mu, i \in \{1, \dots, N_{orb}\}, \quad (5.1)$$

with $q_{i\mu}$ being the coefficients of the linear combinations over the basis functions $\{g_\mu\}_{1 \leq \mu \leq N_b}$. In this molecular orbital basis, the Coulomb long-range integral reads

$$\mathbf{J}^{lr}(i, j) = \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} K(\mathbf{x}, \mathbf{y}) \sum_{i=1}^{N_{orb}} |\phi_i(\mathbf{x})|^2 \sum_{j=1}^{N_{orb}} |\phi_j(\mathbf{y})|^2 d\mathbf{x} d\mathbf{y} \quad (5.2)$$

$$= \sum_{\mu, \nu, \kappa, \lambda=1}^{N_b} \sum_{i, j=1}^{N_{orb}} q_{i\mu} q_{i\nu} \left(\int_{\mathbb{R}^3} \int_{\mathbb{R}^3} K(\mathbf{x}, \mathbf{y}) g_\mu(\mathbf{x}) g_\nu(\mathbf{x}) g_\kappa(\mathbf{y}) g_\lambda(\mathbf{y}) q_{j\kappa} q_{j\lambda} d\mathbf{x} d\mathbf{y} \right). \quad (5.3)$$

Let us define the rectangular matrix $\mathbf{Q} \in \mathbb{R}^{N_{orb} \times N_b^2}$ with entries $Q(i, \mu\nu) = q_{i\mu} q_{i\nu}$ such that \mathbf{J}^{lr} writes in matrix notation as

$$\mathbf{J}^{lr} = \mathbf{Q} \mathbf{B}^{lr} \mathbf{Q}^\top \in \mathbb{R}^{N_{orb} \times N_{orb}}, \quad (5.4)$$

where $\mathbf{B}^{lr} \in \mathbb{R}^{N_b^2 \times N_b^2}$ is the mode-(1,2) matricization of \mathcal{B}^{lr} . A naive approach to evaluate (5.4), given ω , the matrix $\mathbf{Q} \in \mathbb{R}^{N_{orb} \times N_b^2}$, and the long-range two-electron integrals \mathcal{B}^{lr} , is to first compute the matrix product $\mathbf{B}^{lr} \mathbf{Q}^\top$ and then perform $\mathbf{Q} (\mathbf{B}^{lr} \mathbf{Q}^\top)$. The last has an arithmetic cost of $\mathcal{O}(N_b^4 N_{orb})$. Given a truncated computational box $[-b, b]^3$, one can use the factorized structure $\mathcal{B}_{LTEI-TA}^{lr}$ defined in (3.31) or $\mathcal{B}_{LTEI-FMM}^{lr}$ defined in (4.12) to evaluate (5.4) efficiently. Given the two approximation approaches (LTEI-TA and LTEI-FMM), we arrive at the following matrix representations

$$\mathbf{J}_{LTEI-TA}^{lr} = \frac{\omega}{\sqrt{\pi}} \sum_{i=1}^{N_{q1}} w_i \left(\mathbf{Q} \mathbf{M}_{TA}^{(i)} \right) \left(\otimes_{l=1}^3 \mathbf{A}^{(i)} \right) \left(\mathbf{Q} \mathbf{M}_{TA}^{(i)} \right)^\top \text{ and } \mathbf{J}_{LTEI-FMM}^{lr} = \left(\mathbf{Q} \mathbf{M}_{FMM} \right) \mathbf{K} \left(\mathbf{Q} \mathbf{M}_{FMM} \right)^\top. \quad (5.5)$$

We present in Table 2 an overview of the storage complexities obtained through LTEI-TA method as well as LTEI-FMM method to evaluate entries of the long-range two-electron integrals tensor and its application to evaluate the long-range Coulomb matrix defined in (5.4).

Table 2: Storage complexity comparison

	LTEI-TA	LTEI-FMM
Element-wise TEI	$\mathcal{O}(N^{\frac{1}{3}} N_{q1} (N^{\frac{1}{3}} + I_{\mu\nu} + I_{\kappa\lambda}))$	$\mathcal{O}(N)$
Application (5.4)	$\mathcal{O}(N^{\frac{2}{3}} (N_{orb} N^{\frac{1}{3}} + N_{q1}))$	$\mathcal{O}(N(1 + N_{orb}))$

The storage complexity of the element-wise evaluation for LTEI-FMM is a consequence of (4.11), i.e. linear with regard to the number of interpolation points N . The storage complexities for the evaluation of (5.4) are obtained as follows. For LTEI-TA approach

1. Instead of forming all matrices $\mathbf{Q} \mathbf{M}_{TA}^{(i)}$ for $i \in \{1..N_{q1}\}$, we form only (as explained in Section 3.3.2) $\mathbf{Q} \mathbf{M}_{TA, \max}$ that requires $\mathcal{O}(N_{orb} N)$ storage.
2. As discussed before, we keep $(\otimes_{l=1}^3 \mathbf{A}^{(i)})$ in tensorized form. Hence, forming all coefficient matrices $\mathbf{A}^{(i)}$ of size $N_i \times N_i$, for $i \in \{1..N_{q1}\}$ requires $\mathcal{O}(\sum_{i=1}^{N_{q1}} N_i^2) \sim \mathcal{O}(N_{q1} N^{\frac{2}{3}})$ storage.

So in total, the storage complexity is $\mathcal{O}(N^{\frac{2}{3}} (N_{orb} N^{\frac{1}{3}} + N_{q1}))$. For LTEI-FMM approach

1. Forming $\mathbf{Q} \mathbf{M}_{FMM}$ requires $\mathcal{O}(N_{orb} \times N)$ of storage.
2. Forming \mathbf{K} requires $\mathcal{O}(N)$ of storage.

So in total, the storage complexity is $\mathcal{O}(N(1 + N_{orb}))$. According to Table.2, the storage demand for this evaluation seems lower (in order) for LTEI-TA compared to LTEI-FMM. However, we cannot conclude on the best method in terms of storage complexity since N_{q1} and $N^{\frac{1}{3}}$ depend on the value of ω and the chosen computational box $[-b, b]^3$. This motivates numerical comparisons between the two approaches for different parameters (see Section 7).

6 Compression techniques for the factorized long-range TEI tensor

One of the main precomputation steps required to obtain the factorized representation of \mathcal{B}^{lr} is based on the evaluation of $\mathbf{M}_{TA, \max} \in \mathbb{R}^{N_b^2 \times N}$ (resp. $\mathbf{M}_{FMM} \in \mathbb{R}^{N_b^2 \times N}$) matrix. This step tends to be expensive in terms of both computational and memory requirements for molecules of moderate size, as we consider in our experiments. In this section we address this problem by discussing different approaches to compress $\mathbf{M}_{TA, \max}$, some of which can be applied to \mathbf{M}_{FMM} .

6.1 Compression by using low-rank methods

In many cases, the matrix $\mathbf{M}_{TA, \max}$ is numerically low-rank as we will discuss in the numerical experiments section (see Figure 10a). It is possible to reduce its dimensions by exploiting its low rank structure. We recall the *screening* technique

[39] which consists in simply discarding "negligible" pairs of Gaussian type basis functions as explained in 6.3. Low rank approximation methods such as truncated SVD [40] can be also applied directly on $\mathbf{M}_{TA,max}$ to further reduce its dimensions. We introduce in this section a different compression method that exploits the khatri-rao products and associated properties. Let $\tilde{\mathbf{W}}^{(i,l)} \in \mathbb{R}^{N_b^2 I_{\mu\nu,max} \times N_i}$ be defined by

$$\tilde{\mathbf{W}}^{(i,l)} = \left[\begin{array}{c} \tilde{\mathbf{W}}_{\mu\nu}^{(i,l)} \\ \mathbf{0} \end{array} \right] \left\{ \begin{array}{l} I_{\mu\nu} \times N_i \\ (I_{\mu\nu,max} - I_{\mu\nu}) \times N_i, I_{\mu\nu,max} = \max(I_{\mu\nu})_{1 \leq \mu, \nu \leq N_b} \end{array} \right. \quad (6.1)$$

Its low rank $R_{i,l}$ approximation can be written as:

$$\tilde{\mathbf{W}}^{(i,l)} \approx \mathbf{U}^{(i,l)} \mathbf{V}^{(i,l)\top}, \quad (6.2)$$

where $\mathbf{U}^{(i,l)} \in \mathbb{R}^{N_b^2 I_{\mu\nu,max} \times R_{i,l}}$ and $\mathbf{V}^{(i,l)\top} \in \mathbb{R}^{R_{i,l} \times N_i}$. Given the decomposition (6.2), Proposition 4 is used to obtain the following expression

$$\left(\diamond_{l=1}^3 (\tilde{\mathbf{W}}^{(i,l)}) \right) \left(\diamond_{l=1}^3 \mathbf{A}^{(i)} \right) \left(\ast_{l=1}^3 (\tilde{\mathbf{W}}^{(i,l)})^\top \right) = \left(\diamond_{l=1}^3 (\mathbf{U}^{(i,l)} \mathbf{V}^{(i,l)}) \right) \left(\mathbf{A}^{(i)} \otimes \mathbf{A}^{(i)} \otimes \mathbf{A}^{(i)} \right) \left(\ast_{l=1}^3 (\mathbf{U}^{(i,l)} \mathbf{V}^{(i,l)})^\top \right) \quad (6.3)$$

$$= \left(\diamond_{l=1}^3 \mathbf{U}^{(i,l)} \right) \left(\otimes_{l=1}^3 \mathbf{V}^{(i,l)} \right) \left(\mathbf{A}^{(i)} \otimes \mathbf{A}^{(i)} \otimes \mathbf{A}^{(i)} \right) \left(\otimes_{l=1}^3 \mathbf{V}^{(i,l)\top} \right) \left(\ast_{l=1}^3 (\mathbf{U}^{(i,l)})^\top \right) \quad (6.4)$$

$$= \left(\diamond_{l=1}^3 \mathbf{U}^{(i,l)} \right) \otimes_{l=1}^3 \left(\mathbf{V}^{(i,l)} \mathbf{A}^{(i)} \mathbf{V}^{(i,l)\top} \right) \left(\ast_{l=1}^3 (\mathbf{U}^{(i,l)})^\top \right). \quad (6.5)$$

By replacing the low rank approximation of the matrix $\tilde{\mathbf{W}}^{(i,l)}$ in the expression of $\mathbf{B}_{LTEI-TA}^{lr}$ in equation (3.31), we obtain

$$\mathbf{B}_{LTEI-TA}^{lr} \approx \frac{\omega}{\sqrt{\pi}} \sum_{i=1}^{N_{q1}} w_i \tilde{\mathbf{U}}^{(i)} \otimes_{l=1}^3 \left(\mathbf{V}^{(i,l)} \mathbf{A}^{(i)} \mathbf{V}^{(i,l)\top} \right) \tilde{\mathbf{U}}^{(i)\top}, \quad (6.6)$$

where $\tilde{\mathbf{U}}^{(i)} = \sum_{j=1}^{I_{\mu\nu,max}} c_j \mathcal{U}^{(i)}[j, :, :] \in \mathbb{R}^{N_b^2 \times \prod_{l=1}^3 R_{i,l}}$ with $\mathcal{U}^{(i)} \in \mathbb{R}^{I_{\mu\nu,max} \times N_b^2 \times \prod_{l=1}^3 R_{i,l}}$ the tensorization of $\left(\diamond_{l=1}^3 \mathbf{U}^{(i,l)} \right) \in \mathbb{R}^{I_{\mu\nu,max} N_b^2 \times \prod_{l=1}^3 R_{i,l}}$. In practice, we compute only the matrix $\tilde{\mathbf{U}}^{(i)} \in \mathbb{R}^{N_b^2 \times \max(\prod_{l=1}^3 R_{i,l})_{i \in \{1, \dots, N_{q1}\}}}$ with the maximum rank $\max(\prod_{l=1}^3 R_{i,l})_{i \in \{1, \dots, N_{q1}\}}$ as discussed in Section 3.3.2.

6.2 Adaptive approach for the choice of the integration domain $[-b, b]$

We discuss now an adaptive approach for the choice of the integration domain $[-b, b]$. For each pair of Gaussian functions, we identify its numerical support $[-b, b]$. We cluster together these numerical supports to obtain overall $N_{partitions}$ supports, $\{[-b_s, b_s]\}_{s \in \{1, \dots, N_{partitions}\}}$. For each pair of Gaussian functions (μ, ν) , $\mu, \nu \in \{1, \dots, N_b\}$, we proceed as follows: Given the general Gaussian product rule (Definition 1), the product of two primitive Gaussian type functions $g_{\mu\nu}^{(j)}(\mathbf{x}_l) = g_{\mu}^{(j_1)}(\mathbf{x}_l) g_{\nu}^{(j_2)}(\mathbf{x}_l)$ is

$$g_{\mu\nu}^{(j)}(\mathbf{x}_l) = g_{\mu}^{(j_1)}(\mathbf{x}_l) g_{\nu}^{(j_2)}(\mathbf{x}_l) = (\mathbf{x}_l - \mathbf{r}_l)^{p_{\mu l}} (\mathbf{x}_l - \mathbf{r}'_l)^{p_{\nu l}} \exp \left(-\frac{\mu_{j_1} \nu_{j_2}}{\mu_{j_1} + \nu_{j_2}} (\mathbf{r}_l - \mathbf{r}'_l)^2 \right) \sigma_{\mu_{j_1} \nu_{j_2}}(\mathbf{x}_l), \quad (6.7)$$

where

$$\sigma_{\mu_{j_1} \nu_{j_2}}(\mathbf{x}_l) = \exp \left(-(\mu_{j_1} + \nu_{j_2}) \left(\mathbf{x}_l - \frac{\mu_{j_1} \mathbf{r}_l + \nu_{j_2} \mathbf{r}'_l}{\mu_{j_1} + \nu_{j_2}} \right)^2 \right), j_1 \in \{1, \dots, I_{\mu}\}, j_2 \in \{1, \dots, I_{\nu}\}, \quad (6.8)$$

and (see (1.9))

$$g_{\mu}^{(j_1)}(\mathbf{x}_l) = (\mathbf{x}_l - \mathbf{r}_l)^{p_{\mu l}} \exp(-\mu_{j_1} (\mathbf{x}_l - \mathbf{r}_l)^2) \text{ and } g_{\nu}^{(j_2)}(\mathbf{x}_l) = (\mathbf{x}_l - \mathbf{r}'_l)^{p_{\nu l}} \exp(-\nu_{j_2} (\mathbf{x}_l - \mathbf{r}'_l)^2), \quad (6.9)$$

with $j = (j_1, j_2) \in \{1, \dots, I_{\mu\nu}\}$, $I_{\mu\nu} = I_{\mu} I_{\nu}$, $l \in \{1, 2, 3\}$, $\mu, \nu \in \{1, \dots, N_b\}$. The numerical support $[-b, b]$ is chosen according to a cutoff threshold $\tau_{adaptive} > 0$ such that

$$\sigma_{\mu_{j_1} \nu_{j_2}}(\mathbf{x}_l) \leq \tau_{adaptive}, l \in \{1, 2, 3\}. \quad (6.10)$$

To illustrate this adaptive approach, for a given pair (μ, ν) , we represent in Figure 2 (left) the exponential terms $\sigma_{\mu_{j_1} \nu_{j_2}}(\mathbf{x}_1)$ in the expression (6.7) for $j \in \{1, \dots, I_{\mu\nu}\}$ with respect to the first direction ($l=1$). The exponential decay of these functions enables us to limit the range of the numerical grid according to a chosen threshold $\tau_{adaptive}$. Through this adaptive technique, Figure 2 (right) illustrates the distribution of the numerical support (dimension b). Each bar represents the percentage of Gaussian function pairs (μ, ν) associated to the exponential terms $\sigma_{\mu_{j_1} \nu_{j_2}}(\mathbf{x}_l)$ lying in the range $[-b, b]$. It is showed that the distribution depends on the molecule choice as well as the number of basis functions N_b .

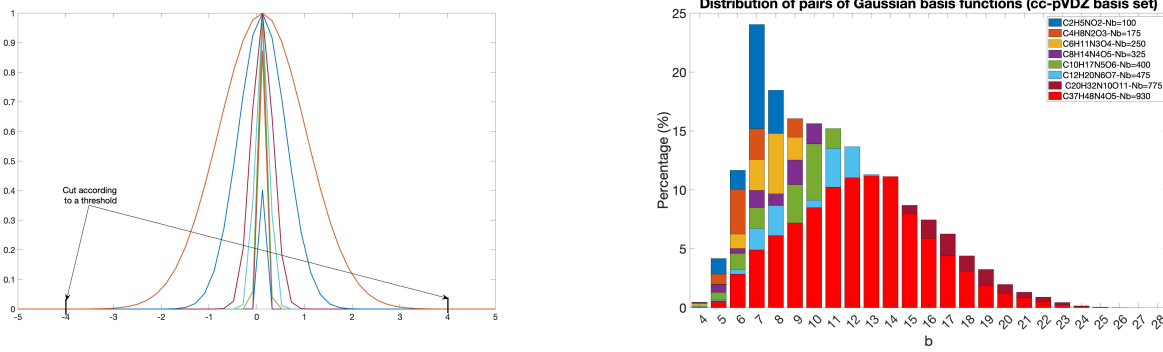


Figure 2: (Left) Identifying numerical supports of different pairs of Gaussian functions. Each color in the plot represents the exponential term $\sigma_{\mu_{j_1} \nu_{j_2}}(\mathbf{x}_1)$. Here the selected numerical support is $[-4, 4]$. (Right) Distribution of numerical supports $[-b, b]$ for a given threshold $\tau_{adaptive} = 10^{-20}$, the x-axis shows the dimension b of the box, the y-axis shows the percentage of the Gaussian function pairs.

The advantages of using this approach is that there is no need to fix in advance the size of the numerical box b since it depends on the Gaussian functions. Moreover, it is possible to reduce the storage demand since instead of storing the matrix $\mathbf{M}_{TA,max} \in \mathbb{R}^{N_b^2 \times N}$, smaller matrices of sizes $N_{b,s}^2 \times N_s$ are stored, where $N_{b,s}^2$ are the pairs of Gaussian functions associated to the integration domain $[-b_s, b_s]$ and N_s is the maximum number of Chebyshev interpolation points in the interval $[-b_s, b_s]$. We must point out that by using this adaptive method, multiple tensor contraction calculations need to be performed to compute (5.5) which will depend on the number of partitions \mathcal{P}_s . This can be costly if we consider a sequential algorithm. However, this adaptive approach offers a possibility to parallelize the evaluation of (5.5).

6.3 Compression by using Screening

It is possible to further reduce the dimensions of $\mathbf{M}_{TA,max}$ by exploiting the properties of the Gaussian type basis functions. In fact, given the product of two-primitive Gaussians introduced in (6.7), we notice that $g_{\mu\nu}^{(j)}(\mathbf{x}_l) = g_{\nu\mu}^{(j)}(\mathbf{x}_l)$, for $j \in \{1, \dots, I_{\mu\nu}\}$, $\mu, \nu \in \{1, \dots, N_b\}$ and $l \in \{1, 2, 3\}$. Therefore, there are only $\frac{N_b}{2}(N_b + 1)$ choices for N_b^2 combinations of μ and ν . We also apply the screening technique that is often used by chemists to reduce the computational cost of the evaluation of integrals [39]. From the Gaussian product rule (6.7), the higher the exponent of a primitive Gaussian, the faster the products with primitives from other centers decay with distance and the sooner they become negligible. Therefore, for large enough molecules, it is possible to discard a consistent number of pairs of primitive Gaussians which is illustrated in the numerical experiment section in Figure 10b. In practice, we discard the primitive pair that satisfies the following condition for a given threshold $\tau_{screening}$

$$\exp\left(-\frac{\mu_{j_1} \nu_{j_2}}{\mu_{j_1} + \nu_{j_2}} \sum_{l=1}^3 (\mathbf{r}_l - \mathbf{r}'_l)^2\right) \leq \tau_{screening}. \quad (6.11)$$

7 Numerical results

In this section, we evaluate numerically our novel method LTEI-TA⁶ by using a prototype implementation in Julia language version 1.5.3. We also compare it with LTEI-FMM method using *defmm* library [32]. The *defmm* library is a C++ code⁷ that is particularly well-suited for the two-electronic integrals context since it implements various important features with $\mathcal{O}(N)$ complexity on non-oscillatory kernels in both precomputation and application cost. More precisely, *defmm* is

- *kernel-independent*, meaning that the user has to provide only a routine evaluating $K(\mathbf{x}, \mathbf{y})$ to use the code and the handling of *erf* function can be added at minimal implementation effort,
- *adaptive*, meaning that the algorithm automatically adapts to the potential non-uniformity of the particle distribution. Similar performance was observed for *defmm* using non-oscillatory kernels applied on uniform and highly non-uniform distributions [32] (such as our tensorized Chebyshev grids),
- *convergent* for any asymptotically smooth kernel, including our kernel $K(\mathbf{x}, \mathbf{y})$ (see Proposition 6), as proven in [41].

An example of a call to *defmm* library is provided in A. *defmm* is compiled using the intel C++ compiler (version 19.1.2.254) and FFTW3 (since *defmm* relies on FFTs for the far field compression/evaluation). We remind that the evaluation algorithm in LTEI-TA, which is written in Julia, is based on matrix-matrix products, performed with optimized BLAS operations (see Section

⁶https://github.com/sbadred/LTEI_TA.jl

⁷<https://github.com/IChollet/defmm>

3.3.1) for the dense linear algebra computations. Hence, the effect of the programming language choice has a negligible impact for LTEI-TA. This justifies the comparison between c++ calls (*defmm*) and our implementation of LTEI-TA in Julia. We are also aware that results presented in the following correspond to prototypes in which we simply link *defmm* with outputs from our Julia code, regardless of further possible optimizations. All the calculations are carried out using Cleps cluster from Inria, Paris, France. This machine has 4 partitions. We use cpu-homogen partition which contains 20 nodes with hyper-threading such that we can allocate a maximum of 64 logical cores per node (Intel(R) Xeon(R) Silver 4214 CPU @ 2.20GHz) with a memory of 6GB per core. We start always by the data initialization step which consists in reading input files generated from *quantum package*. These files contain molecular properties: number of atoms, number of basis functions, coordinates of the nuclei, basis set parameters. For all molecules we use the “cc-pVDZ” Gaussian basis set [19].

7.1 Approximation error and computational cost

The following numerical results present the approximation errors with respect to different parameters ω , N , and N_{q_1} . We start by providing the approximation error for the element-wise evaluation of the long-range two-electron integrals tensor and then we provide the numerical error convergence obtained for the evaluation of the long-range Coulomb matrix as defined in (5.4) using both methods: LTEI-TA and LTEI-FMM.

7.1.1 Approximation error

First we provide convergence results of LTEI-TA method for the evaluation of the long-range two-electron integrals given in equation (3.14). For the following numerical tests, we consider small sized molecules : NH_3 and CO_2 , where we represent the mean relative error of 10^3 randomly chosen elements from the tensor \mathcal{B}^{lr} . On the left of the Figure 3, the maximum number of Chebyshev interpolation points N is fixed while on the right of the Figure 3 the number of quadrature points N_{q_1} is fixed.

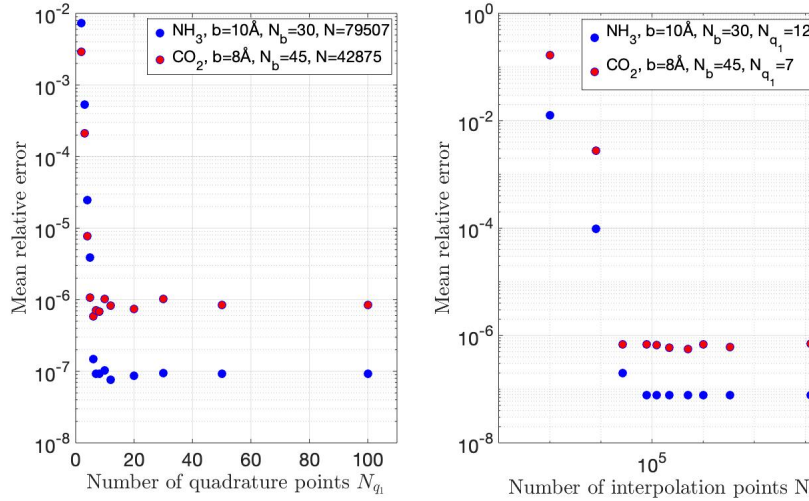


Figure 3: Approximation error of the long-range two-electron integrals using LTEI-TA, $\omega = 0.5$.

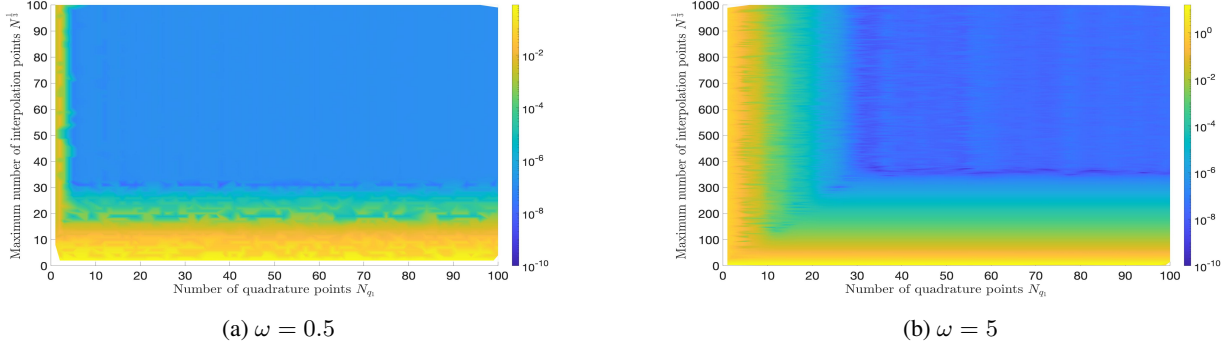


Figure 4: Approximation error of the element-wise evaluation of the two-electron integrals (3.14) with respect to $(\# \text{interpolation points per direction}, \# \text{quadrature points}) \equiv (N^{\frac{1}{3}}, N_{q1})$ for the optimal accuracy using NH3 molecule in the cc-pVDZ basis set for different values of ω . The colorbar shows the mean relative approximation error.

In Figure 3, with fixed $\omega = 0.5$, we notice the fast convergence of the relative error towards the value of $1e^{-7}$ for both subfigures such that the analytical results, generated from *quantum package*, and numerical results are in reasonably good agreement for both molecules. We note that the stagnation of the error is a consequence of the approximations used (Chebyshev interpolation and Gaussian quadrature rule), hence in order to optimize our method for a desired accuracy, we need to find a good compromise between the parameters $N^{\frac{1}{3}}$ and N_{q1} , as shown in Figure 4. Indeed, we note that in Figure 4, for each number of interpolation points, there is a number of quadrature points that allows to reach a small relative error (up to $1e^{-10}$). One may also notice that the minimal error is constrained by the choice of b , i.e. of the integration box, since the support of the primitive Gaussians are truncated.

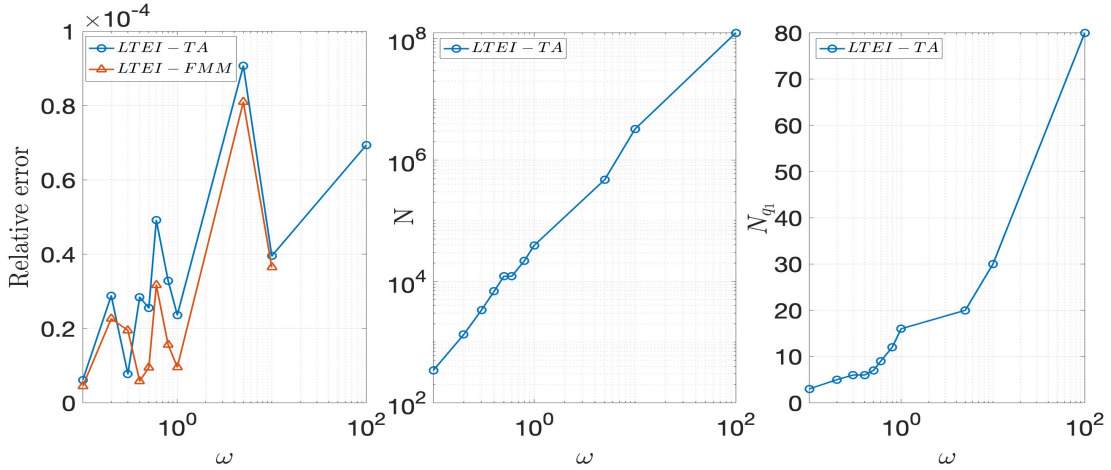


Figure 5: (Leftmost figure) The approximation error of the element-wise evaluation of the two-electron integrals with respect to ω for both approaches: LTEI-TA and LTEI-FMM. (Middle figure) The number of interpolation points N needed to reach the imposed accuracy (relative error smaller than $1e^{-4}$) with respect to ω . (Rightmost figure) The number of quadrature points N_{q1} needed to reach the imposed accuracy (relative error smaller than $1e^{-4}$) with respect to ω .

Figure 5 displays the number of interpolation points N (middle figure) and the number of quadrature points N_{q1} (rightmost figure) with respect to ω for computing a single entry of the long-range two-electron integrals tensor \mathcal{B}^{lr} through LTEI-TA and LTEI-FMM approaches. The entry $\mathcal{B}^{lr}(\mu, \nu, \kappa, \lambda)$ is chosen randomly and we impose that the relative error is smaller than 10^{-4} , where the relative error is defined as $\frac{|\mathcal{B}^{lr}(\mu, \nu, \kappa, \lambda) - \mathcal{B}_{LTEI-TA}^{lr}(\mu, \nu, \kappa, \lambda)|}{|\mathcal{B}^{lr}(\mu, \nu, \kappa, \lambda)|}$ for LTEI-TA, and as $\frac{|\mathcal{B}^{lr}(\mu, \nu, \kappa, \lambda) - \mathcal{B}_{LTEI-FMM}^{lr}(\mu, \nu, \kappa, \lambda)|}{|\mathcal{B}^{lr}(\mu, \nu, \kappa, \lambda)|}$ for LTEI-FMM, respectively. We observe that the number of interpolation points N and the number of quadrature points N_{q1} needed to reach the desired accuracy grow with ω , as it can be seen in the middle and rightmost figures. This is explained by the fact that when $\omega \rightarrow \infty$, LTEI-TA needs to approximate a nearly singular kernel, which increases its cost. The leftmost figure also shows that the accuracy of LTEI-TA and LTEI-FMM for the evaluation of an element of \mathcal{B}^{lr} is comparable for the same number

of interpolation points N . This is because both approaches are based on Chebyshev interpolation. We note that the quadrature in LTEI-TA is chosen to be at least as precise as the interpolation and the FMM error is controlled by a parameter [32] whose value is practically calibrated so that this error equals the numerical interpolation. Both methods thus lead to the expected accuracy.

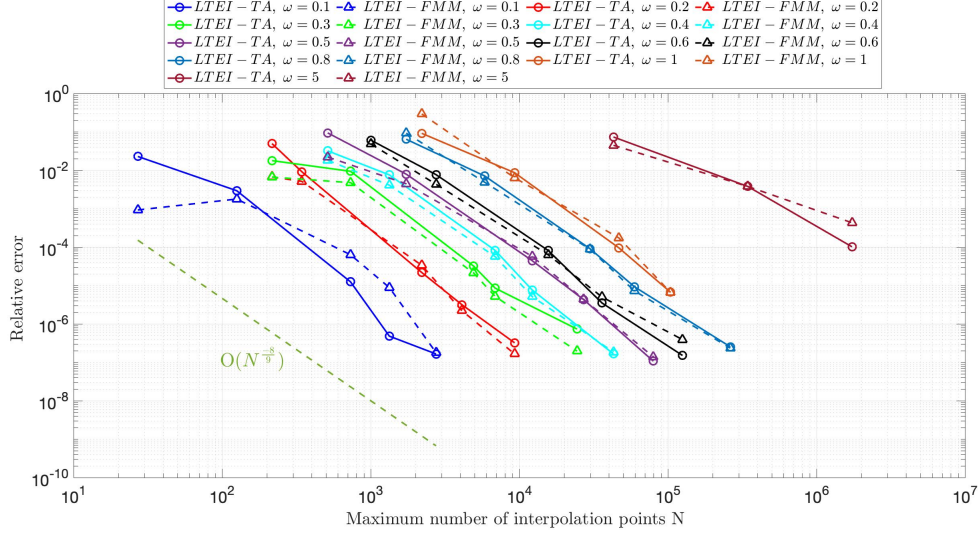


Figure 6: Approximation error of the evaluation of the long-range Coulomb matrix using LTEI-TA and LTEI-FMM with respect to the number of interpolation points N for various values of ω : convergence rate estimation. These calculations were carried for the glycine molecule with $N_b = 100$ in the cc-pVDZ basis set.

Figure 6 considers the evaluation of the long-range Coulomb matrix using LTEI-TA and LTEI-FMM approaches as described in (5.5). It displays the relative error with respect to the number of interpolation points N for different values of ω , where the relative error of LTEI-TA (resp. LTEI-FMM) is $\frac{\|\mathbf{J}^{lr} - \mathbf{J}_{LTEI-TA}^{lr}\|_2}{\|\mathbf{J}^{lr}\|_2}$ (resp. $\frac{\|\mathbf{J}^{lr} - \mathbf{J}_{LTEI-FMM}^{lr}\|_2}{\|\mathbf{J}^{lr}\|_2}$). We note that we were not able to evaluate theoretically the convergence rate of this evaluation with respect to the number of interpolation points N . We observe, however, that the numerical error seems to have an almost linear-scaling in the 3D tensorized interpolations grid size N for small values $\omega \in (0, 1)$. However, this scaling is lost for larger ω . Indeed, we expect our method to be far less efficient for very large ω since the underlying kernel tends to the (singular) Coulomb one when $\omega \rightarrow +\infty$.

7.1.2 Computational cost

We first discuss the execution time required for the evaluation of an element of the long-range tensor \mathcal{B}^{lr} , as displayed in Figure 7. The computational complexity of this evaluation is of order $\mathcal{O}(N_{q1} N^{\frac{1}{3}} I_{\kappa\lambda}(N^{\frac{1}{3}} + I_{\mu\nu}))$ as discussed in Section 3.1. For small values of N_{q1} and a few number of interpolations points $N^{\frac{1}{3}}$, we obtain linear scaling with respect to $N^{\frac{1}{3}}$ as shown in Figure 7. This is explained by the fact that the term $\mathcal{I}_{\kappa\lambda} \mathcal{I}_{\mu\nu}$ dominates the overall complexity for small ω . However, when ω increases, a quadratic complexity is observed with respect to $N^{\frac{1}{3}}$, which corresponds to $\mathcal{O}(N_{q1} N^{\frac{1}{3}} I_{\kappa\lambda}(N^{\frac{1}{3}} + I_{\mu\nu}))$. We also compare LTEI-TA with LTEI-FMM and with a naive numerical computation such that the two-electron integrals are computed with an integration over $N \times N \times N$ tensorized three dimensional Cartesian grids. We notice here that the LTEI-FMM approach has a linear scaling with regards to the number of interpolation points N as expected. We conclude that for the element-wise evaluation, LTEI-TA is the most efficient method.

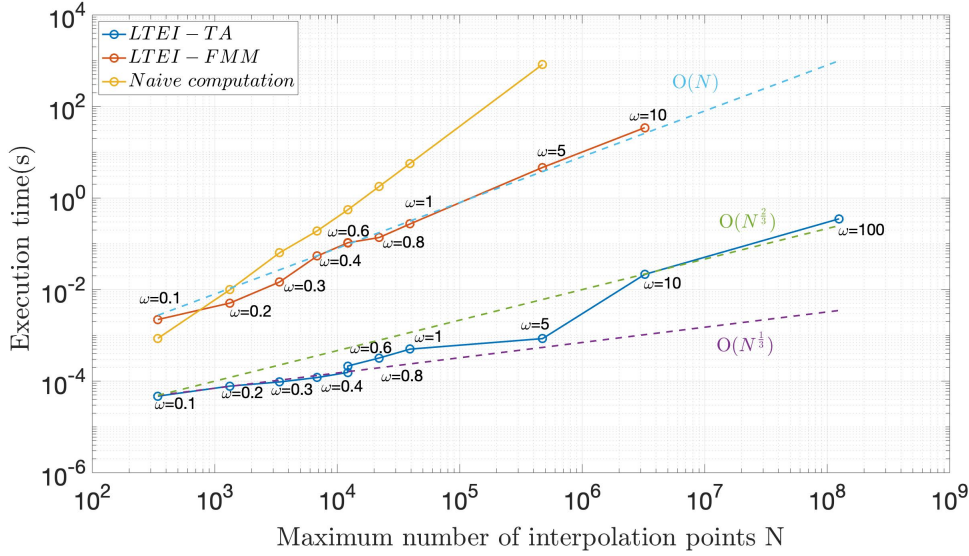


Figure 7: Computational time versus the maximum number of interpolation points N for different values of ω for the evaluation of the long-range two-electron integrals with relative error smaller than $\leq 1e^{-4}$.

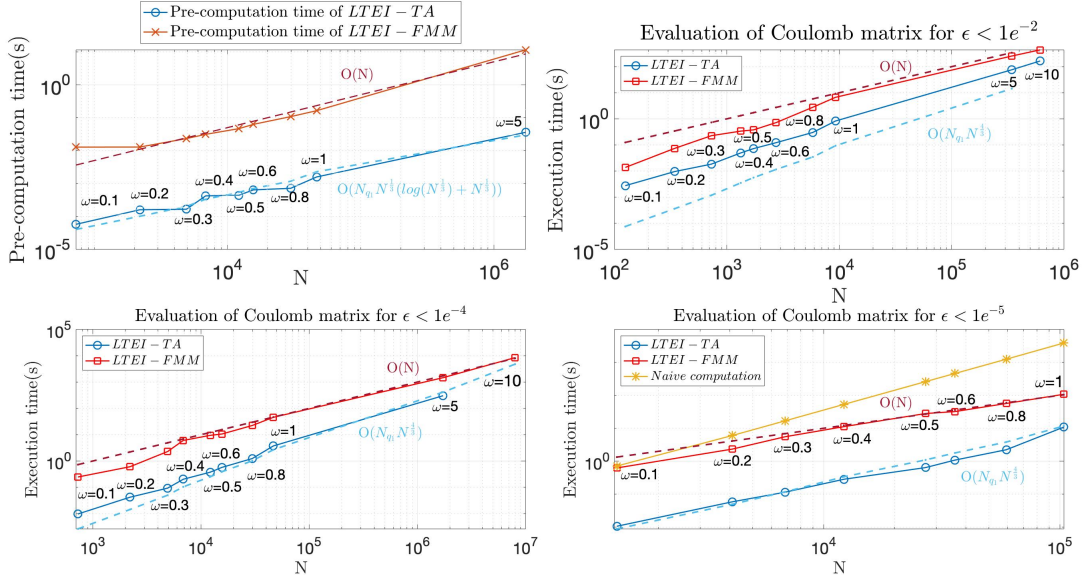


Figure 8: The leftmost plot represents the precomputation time for each approximation approach (LTEI-TA and LTEI-FMM) with respect to the maximum number of Chebyshev interpolation points N . We impose here that the relative error denoted by ϵ is smaller than $\leq 1e^{-4}$. We provide in the other plots a comparison in terms of the computational time required for the evaluation of (5.4) between both approaches by varying the error bound ϵ and ω . We use the Glycine molecule $C_2H_5NO_2$ with fixed $N_b = 100$ and $N_{orb} = 95$ in the cc-pVDZ basis set.

Second, we compare the precomputation cost required to approximate the long-range kernel $K(x, y)$, as given in (3.2), by using both approaches LTEI-TA and LTEI-FMM and by varying ω from 0.1 to 5. The results are displayed in the leftmost part of Figure 8. We observe that the runtime of LTEI-FMM depends linearly on the total number of interpolation points $\mathcal{O}(N)$, independently of the value of ω . LTEI-TA has also a precomputation time in accordance with the theory $\mathcal{O}(N_{q1} N^{\frac{1}{3}} (\log(N^{\frac{1}{3}}) + N^{\frac{1}{3}}))$ as explained in Section 3. We observe that LTEI-TA is two orders of magnitude faster than LTEI-FMM for all the considered values of ω (which is a consequence of its small precomputation complexity).

Third, we discuss the time required to evaluate the long-range Coulomb matrix, as given in equation (5.4), which involves the multiplication of the matricization of \mathbf{B}^{lr} with a matrix. Figure 8 illustrates the execution time with respect to the number of interpolation points N needed to achieve different relative errors for various values of ω for the evaluation of the Coulomb matrix. The relative error of LTEI-TA (resp. LTEI-FMM) is $\frac{\|\mathbf{J}^{lr} - \mathbf{J}_{LTEI-TA}^{lr}\|_2}{\|\mathbf{J}^{lr}\|_2}$ (resp. $\frac{\|\mathbf{J}^{lr} - \mathbf{J}_{LTEI-FMM}^{lr}\|_2}{\|\mathbf{J}^{lr}\|_2}$). We observe in Figure 8 that the evaluation of the long-range Coulomb matrix using LTEI-FMM approach scales linearly with the number of interpolation points $\mathcal{O}(N)$, but more than linearly for LTEI-TA. This reflects the complexity analysis of LTEI-TA method, $\mathcal{O}(N_{q_1} N^{\frac{4}{3}})$, provided in Section 3.3.2. However, LTEI-TA is still faster than LTEI-FMM for relatively small values of ω and for different relative errors. This numerical gain can be explained by the important prefactor of the LTEI-FMM approach: even if the complexity is linear, there is an important constant hidden in the big \mathcal{O} notations [42]. While for small values of ω , N_{q_1} is small and hence LTEI-TA is more efficient. However, LTEI-TA is not asymptotically competitive with respect to LTEI-FMM approach. Indeed, as ω controls the regularity of the *erf*-interaction function, when ω increases, LTEI-TA needs a larger number of interpolation points N as well as quadrature points N_{q_1} to achieve a given accuracy. As a consequence, LTEI-TA becomes more costly and less efficient than LTEI-FMM.

To summarize, these results demonstrate two major things: first, LTEI-TA is a numerically highly efficient method, able to outperform LTEI-FMM on tested cases. Second, we are able to reach the linear complexity (with regard to the total number of interpolation points) by exploiting LTEI-FMM, which allows to deal with more singular cases (with large values of ω). In the following, we want to study the efficiency of our numerical approaches for variable N_b .

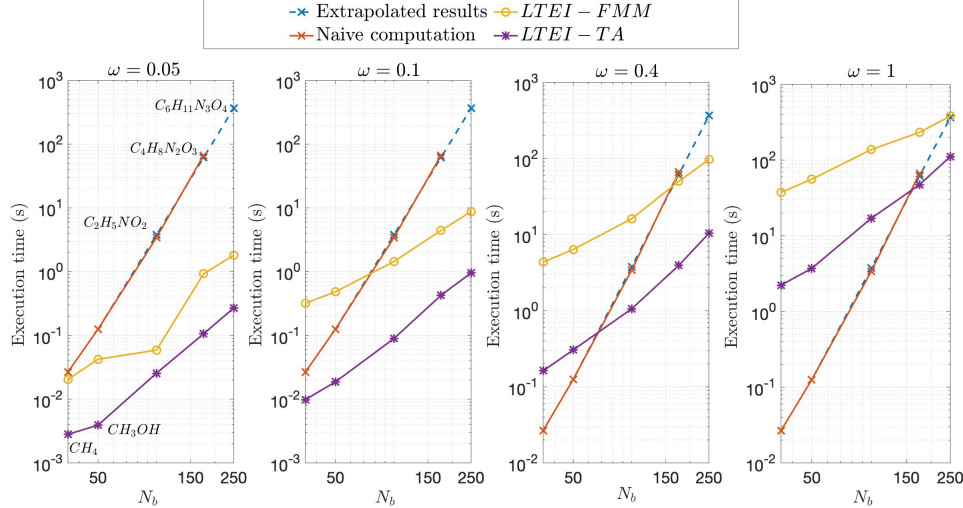


Figure 9: Execution time(s) required for the evaluation of (5.4) using the TEI tensor \mathbf{B}^{lr} for different values of N_b , for $\omega = 0.05, \omega = 0.1, \omega = 0.4$, and $\omega = 1$ with imposed relative error smaller than $1e^{-5}$.

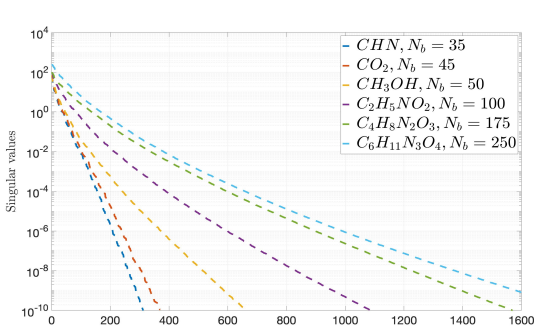
Figure 9 displays the execution times required to evaluate (5.4) with respect to the number of basis functions N_b for different values of ω and different molecules. We impose here that the relative errors of LTEI-TA and LTEI-FMM approaches for the evaluation of the long-range Coulomb matrix are smaller than $1e^{-5}$. We compare the running times between three approaches: the first approach is a direct computation of (5.4) given the matricization of \mathbf{B}^{lr} denoted by $\mathbf{B}^{lr} \in \mathbb{R}^{N_b^2 \times N_b^2}$ (times for $N_b > 175$ are obtained by extrapolation). The second (resp. third) approach exploits the factorized structure of \mathbf{B}^{lr} obtained through LTEI-TA (resp. LTEI-FMM) to compute (5.4). For small ω , we notice that a faster computation of (5.4) is obtained through LTEI-TA and LTEI-FMM methods: LTEI-TA is about one order of magnitude faster than LTEI-FMM. For important values of ω ($\omega=1$), the new introduced approaches, LTEI-TA and LTEI-FMM, are less efficient given the high number of interpolation points N needed as well as the number of quadrature points N_{q_1} for LTEI-TA method as we notice in Figure 8. However, when N_b increases, the tensor contractions using the direct method will be expensive and will have high memory demands (sometimes \mathbf{B}^{lr} is too large to store in memory). Therefore in some cases, it would be beneficial to use one of the new factorization methods to reduce the computational and storage cost. The numerical results are obtained for different molecules with different topologies. Therefore, in order to preserve the accuracy, in practice, we choose the size of the computational box $[-b, b]$ depending on the size of the molecule as well as the Gaussian functions decay as explained previously in Section 3.

7.2 Tensor compression techniques

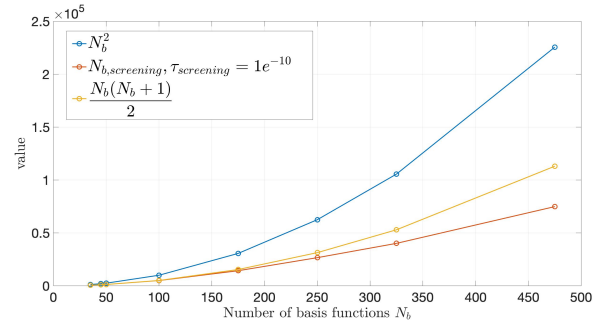
In this section we study numerically compression techniques to reduce the computation and storage cost of $\mathbf{M}_{TA,max} \in \mathbb{R}^{N_b^2 \times N}$ or $\mathbf{M}_{FMM} \in \mathbb{R}^{N_b^2 \times N}$ in order to speed up the evaluation of the Coulomb matrix (5.4). These techniques were discussed in Section 6. First, the number of basis functions N_b can be reduced by using screening techniques that exploit the symmetries of the pairs of basis functions as well as the properties of Gaussian type-functions. Indeed, Figure 10b shows that the number of pairs of Gaussian type basis functions N_b^2 can be reduced by using screening. Second, for small values of ω and different numbers of basis functions N_b , Figure 10a shows that the singular values of $\mathbf{M}_{TA,max}$ decay quickly, so $\mathbf{M}_{TA,max}$ can be approximated by a low-rank matrix. Therefore, we had recourse to three different approaches for the compression of $\mathbf{M}_{TA,max}$: the first approach, denoted by SVD, consists in approximating $\mathbf{M}_{TA,max}$ using ϵ -truncated SVD; the second approach, denoted by KR, exploits the Khatri-Rao product properties as discussed in Section 6.1; and the third approach, denoted by ADAP+KR, includes the partitioning of pairs of basis functions in terms of their numerical supports combined with KR approach as explained in Section 6.2.

Figure 11 (resp. Figure 12) displays the compression rate obtained between uncompressed $\mathbf{M}_{TA,max}$ matrix (resp. screened $\mathbf{M}_{TA,max}$ matrix) and its compressed representation, for different molecules with different number of basis functions N_b in the basis set cc-pVDZ. We notice that the best compression rate, i.e $(1 - \frac{\text{size of compressed version}}{\text{size of original}}) * 100$, is obtained through the ADAP+KR approach as observed in Figure 12 (86% for $N_b = 175$) compared to the other approaches SVD (75% for $N_b = 175$) and KR (83% for $N_b = 175$). We observe that for SVD, the larger N_b ($N_b \geq 50$), the better the compression. While screening techniques reduce the storage requirements of the matrix $\mathbf{M}_{TA,max}$ [39], better compression results are obtained when they are combined with additional techniques introduced here. Figure 13, shows the computational time required for the compression of $\mathbf{M}_{TA,max}$. The worst execution time is obtained for SVD method, in particular for large values of N_b ($N_b \geq 100$).

In summary, the adaptive approach leads to the best reduction in terms of storage while being the fastest among the tested methods. Moreover, the choice of the dimension of the computational box b does not have to be fixed in advance, since it depends on the pairs of Gaussian type-functions (1.8). We further investigate the accuracy of this method in Table 3. We display in this table the relative error obtained when approximating the Coulomb matrix (5.4) by using either $\mathbf{M}_{TA,max}$ compressed by the adaptive approach or a fixed computational box $[-b, b]$. The results show that the adaptive approach is more accurate than the ones obtained by fixing the computational box in advance. However, by using the adaptive method, the computation of the Coulomb matrix requires multiple matrix-matrix multiplications, and this can be more costly than fixing the computational box $[-b, b]$ in advance. However, since these multiplication can be performed in parallel, parallelization might be a key component to speed up the computation of the long-range Coulomb matrix (5.4).



(a)



(b)

Figure 10: (a) Singular values of $\mathbf{M}_{TA,max} \in \mathbb{R}^{N_b^2 \times N}$ for different molecules with $\omega = 0.1$ and $N_{q_1} = 3$. (b) Number of reduced pairs of basis functions obtained by exploiting symmetry (yellow curve), as well as symmetry+properties of Gaussian type functions with $\tau_{screening}=1e-10$ (red curve).

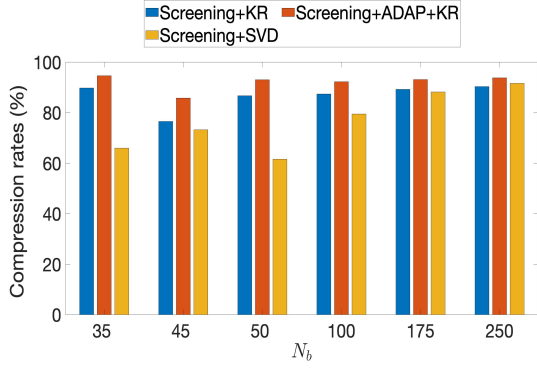


Figure 11: Compression rate between the original computed $\mathbf{M}_{TA,max}$ matrix and its compressed representation for $\omega = 0.3$ for different values of N_b , for the different molecules displayed in Figure 10a.

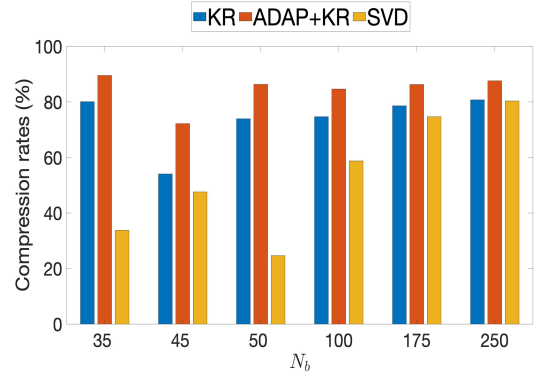


Figure 12: Compression rate between $\mathbf{M}_{TA,max}$ matrix (after screening) and its compressed representation for $\omega = 0.3$ for different values of N_b , for the different molecules displayed in Figure 10a.

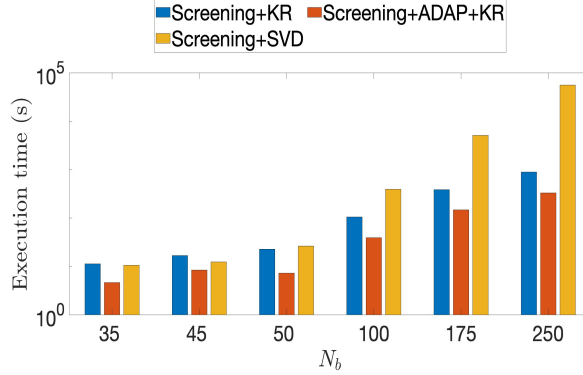


Figure 13: Execution time(s) of different compression methods defined in Section 6 for $\omega = 0.3$, for different values of N_b , for the different molecules displayed in Figure 10a.

Table 3: Adaptive method, $\omega = 0.5$

Molecule	$C_2H_5NO_2$	$C_4H_8N_2O_3$	$C_6H_{11}N_3O_4$
N_b	100	175	250
Adaptive approach	1.0354e-7	2.4882e-8	4.587e-7
$b = 15$	1.6058e-7	1.8332e-7	8.2245e-7
$b = 10$	3.7359e-07	0.001	0.02068

8 Conclusion

This paper introduces two new compression methods for the long-range kernel K and the approximation of the long-range six-dimensional two-electron integrals tensor. The first approach, referred to as LTEI-TA, relies on two-dimensional Chebyshev interpolation, Gaussian quadrature for numerical integration, and FFT for computing Chebyshev coefficients. The approximation of the long-range two-electron integrals tensor \mathcal{B}^{lr} by using this method allows to exploit a tensorized structure that leads to an efficient application of the matricization of \mathcal{B}^{lr} to evaluate the long-range Coulomb matrix for fixed N_b and N_{orb} , with $\mathcal{O}(N_{q1} N^{\frac{4}{3}})$ complexity, where N is the number of Chebyshev interpolation points and N_{q1} is the number of quadrature points. The second approach, referred to as LTEI-FMM, relies on kernel-independent Fast Multipole Methods, with $\mathcal{O}(N)$ complexity. It exploits the asymptotically smooth behaviour of the long-range kernel K . The storage and time complexity of the presented methods were analysed and compared numerically, exhibiting both the high efficiency of LTEI-TA and the linear complexity of LTEI-FMM. We further investigated the compression of \mathcal{B}^{lr} by using screening techniques, low-rank methods, and an adaptive

approach. LTEI-TA approach is particularly efficient for small values of ω , where ω is the separation parameter that controls the regularity of K . However, for large values of ω , in order to preserve accuracy, the number of interpolation points as well as the number of quadrature points becomes important for LTEI-TA and thus LTEI-FMM becomes more efficient.

As future work, we are planning to explore the potential of LTEI-TA for small values of ω in a range of quantum chemical contexts as post-HF models or hybrid approaches such as (long-range) DMRG–short-range DFT [43]. Concerning LTEI-FMM it would be interesting to consider more singular kernels than the one in this paper (such as $\frac{\text{erfc}(\omega|\mathbf{x}-\mathbf{y}|)}{|\mathbf{x}-\mathbf{y}|}$ or the Coulomb kernel directly), thus extending LTEI-FMM to the evaluation of the short-range two-electron integrals by studying appropriate singular quadrature. Such work might be also beneficial for Particle Mesh Ewald methods [16].

Acknowledgments

The authors are grateful to Julien Toulouse (Sorbonne university and CNRS), Emmanuel Giner (Sorbonne university) and Xavier Claeys (Sorbonne university) for valuable discussions. We are thankful to Emmanuel Giner for his assistance with the configuration of quantum package and the extraction of molecular data. The authors are also grateful to the CLEPS infrastructure from the Inria of Paris for providing resources and support. This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation program (grant agreement No 810367).

A The *defmm* library

The *defmm* library (<https://github.com/IChollet/defmm>) is a easy to use C++ implementation of the directional interpolation-based Fast Multipole Method exploiting equispaced interpolation combined with Fast Fourier Transforms. Mainly, *defmm* ensures a $\mathcal{O}(N)$ complexity independently of the particle distribution. Here, we provide an example of a short program calling *defmm*: only five lines are needed to construct and apply the FMM matrix to a vector.

```
#include "path to defmm/include/interface.hpp"
using namespace defmm;
int main(){

    const int DIM = 3 ; // Dimension
    const int ORDER = 4 ; // Interpolation order
    const int NCRIT = 32 ; // Number of particle per leaf cell
    const float KAPPA = 0. ; // Wavenumber (for oscillatory kernels)
    const int N = 41334; // Number of points

    // Get random charge vector
    Vec Q(N), P(N);
    for(int n = 0; n < N; n++){
        Q[n] = cplx(urand);}

    IBFMM_Mat<DIM> A; // FMM matrix
    A.addSourceParticlesINP("Y.inp",N); // Read source particles in Y.inp
    A.addTargetParticlesINP("X.inp",N); // Read target particles in X.inp
    A.prcmpt(ORDER,NCRIT,KAPPA); // Precompute
    gemv(A,Q,P); // Compute P = A Q

    return 0;
}
```

As a header-only library, *defmm* does not need to be compiled before calling. However, our library calls both BLAS and the FFTW3 library [30].

Input files for the listing of source and target particles (that can be the same) are given as a sequence of particle coordinates (one particle per line, coordinates separated by blanks).

B The Hartree-Fock exchange

In computational quantum chemistry, the efficient construction of the long-range exchange matrix in the Fock matrix is also interesting [5, 15, 15]. This matrix is calculated by using the long-range two-electron integrals tensor \mathcal{B}^{lr} . The long-range exchange matrix is given by

$$\mathbf{K}^{lr}(\mu, \nu) = 2 \sum_{j=1}^{N_{orb}} \sum_{\lambda, \kappa=1}^{N_b^2} q_{j\lambda} q_{j\kappa} \mathcal{B}^{lr}(\mu, \lambda, \kappa, \nu), \mu, \nu \in \{1, \dots, N_b\}, \quad (\text{B.1})$$

with $q_{j\lambda}$, $q_{j\kappa}$, and N_{orb} being defined in Section 5. Using the long-range two-electron integrals tensor \mathbf{B}^{lr} , The evaluation of $\mathbf{K}^{lr}(\mu, \nu)$ costs $\mathcal{O}(N_b^2 N_{orb})$. One can use the factorized structure $\mathbf{B}_{LTEI-TA}^{lr}$ defined in (3.31) to reduce the computational cost to $\mathcal{O}(NN_{orb}(N_b + N_{q_1} N_b^{\frac{1}{3}}))$ for LTEI-TA approach with N being the number of Chebyshev interpolation points and N_{q_1} being the number of quadrature points. We obtain the following representation

$$\mathbf{K}_{LTEI-TA}^{lr}(\mu, \nu) = 2 \sum_{j=1}^{N_{orb}} \left(\sum_{i=1}^{N_{q_1}} w_i \left(\mathbf{Q}_j \mathbf{M}_{TA,\mu}^{(i)} \right) \otimes_{l=1}^3 A^{(i)} \left(\mathbf{Q}_j \mathbf{M}_{TA,\nu}^{(i)} \right)^\top \right), \quad (\text{B.2})$$

where for a fixed $j \in \{1, \dots, N_{orb}\}$ and $\lambda \in \{1, \dots, N_b\}$, we have $\mathbf{Q}_j \in \mathbb{R}^{N_b}$ and $\mathbf{Q}_j(\lambda) = q_{j\lambda}$. The matrices $\mathbf{M}_{TA,\mu}^{(i)} \in \mathbb{R}^{N_b \times N_i^3}$, $i \in \{1, \dots, N_{q_1}\}$ are obtained by fixing the index $\mu \in \{1, \dots, N_b\}$ in the tensorized representation of $\mathbf{M}_{TA}^{(i)} \in \mathbb{R}^{N_b^2 \times N_i^3}$. These tensor representations are denoted by $\mathcal{M}_{TA}^{(i)} \in \mathbb{R}^{N_b \times N_b \times N_i^3}$ such that

$$\mathcal{M}_{TA}^{(i)}[\mu, :, :] = \mathbf{M}_{TA,\mu}^{(i)}. \quad (\text{B.3})$$

Figure 14 displays the execution times required to evaluate the long-range exchange matrix (B.1) with respect to the number of basis functions N_b , for small values of $\omega \in \{0.05, 0.1\}$. We impose that the relative error of LTEI-TA approach for this evaluation is smaller than $1e^{-5}$ and we compare the running times between a direct computation of (B.1) given $\mathbf{B}^{lr} \in \mathbb{R}^{N_b^2 \times N_b^2}$ and the factorized structure of \mathbf{B}^{lr} using $\mathbf{B}_{LTEI-TA}^{lr}$. It can be seen that in the case of small values of ω , we notice that a faster construction of (B.1) is obtained through LTEI-TA. Compression techniques introduced in Section 6, can be used here to get better running times.

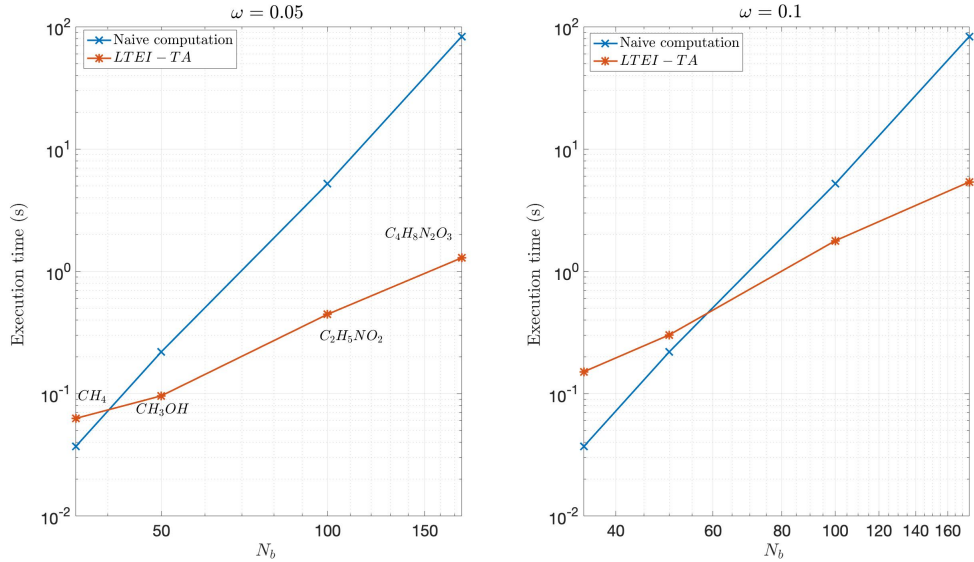


Figure 14: Execution time(s) required for the evaluation of (5.4) using the TEI tensor \mathbf{B}^{lr} for different values of N_b , for $\omega = 0.05$ and $\omega = 0.1$ with imposed relative error smaller than $1e^{-5}$.

References

- [1] V. Khoromskaia, B. N. Khoromskij, Tensor numerical methods in quantum chemistry: from Hartree–Fock to excitation energies, *Phys. Chem. Chem. Phys.* 17 (2015) 31491–31509.
- [2] E. Cancès, M. Defranceschi, W. Kutzelnigg, C. Le Bris, Y. Maday, Computational quantum chemistry: A primer, in: Special Volume, *Computational Chemistry*, volume 10 of *Handbook of Numerical Analysis*, Elsevier, 2003, pp. 3–270. URL: <https://www.sciencedirect.com/science/article/pii/S1570865903100038>. doi:doi:https://doi.org/10.1016/S1570-8659(03)10003-8.
- [3] S. Ashworth, Molecular quantum mechanics, 5th edn., by peter atkins and ronald friedman, *Contemporary Physics - CONTEMP PHYS* 53 (2012) 1–2.

- [4] Szabo, Attila, N. SÖstlund, Modern quantum chemistry : introduction to advanced electronic structure theory, in: Special Volume, Computational Chemistry, Mineola (N.Y.) : Dover publications, 1996, p. 481 / 481. URL: <http://lib.ugent.be/catalog/rug01:000906565>.
- [5] V. Khoromskaya, B. N. Khoromskij, R. Schneider, Tensor-structured factorized calculation of two-electron integrals in a general basis, *SIAM Journal on Scientific Computing* 35 (2013) A987–A1010.
- [6] J. Toulouse, Extension multidéterminantale de la méthode de Kohn-Sham en théorie de la fonctionnelle de la densité par décomposition de l'interaction électronique en contributions de longue portée et de courte portée, Theses, Université Pierre et Marie Curie - Paris VI, 2005. URL: <https://tel.archives-ouvertes.fr/tel-00550772>.
- [7] A. Savin, On degeneracy, near-degeneracy and density functional theory, volume 4, 1996, pp. 327–357. doi:doi:10.1016/S1380-7323(96)80091-4.
- [8] J. Toulouse, F. m. c. Colonna, A. Savin, Long-range–short-range separation of the electron-electron interaction in density-functional theory, *Phys. Rev. A* 70 (2004) 062505.
- [9] E. Giner, A new form of transcorrelated hamiltonian inspired by range-separated dft, *The Journal of Chemical Physics* 154 (2021) 084119.
- [10] A. Savin, Models and corrections: Range separation for electronic interaction—lessons from density functional theory, *The Journal of Chemical Physics* 153 (2020) 160901.
- [11] J. Toulouse, P. Gori-Giorgi, A. Savin, A short-range correlation energy density functional with multi-determinantal reference, *Theoretical Chemistry Accounts: Theory, Computation, and Modeling* 114 (2005) 305.
- [12] A. Lee, S. Taylor, J. Dombroski, P. Gill, Optimal partition of the coulomb operator, *Physical Review A - PHYS REV A* 55 (1997) 3233–3235.
- [13] A. Ferté, Théorie de la fonctionnelle de la densité avec une fonction d'onde multiréférence : Développement d'approximations pour la fonctionnelle de corrélation à courte portée utilisant la densité de paires à coalescence, 2018. Unpublished.
- [14] Lecours, Michael, Compact Sparse Coulomb Integrals using a Range-Separated Potential, Ph.D. thesis, University of Waterloo, 2021. URL: <http://hdl.handle.net/10012/17516>.
- [15] T. Limpanuparb, J. Milthorpe, A. Rendell, P. Gill, Resolutions of the coulomb operator: Vii. evaluation of long-range coulomb and exchange matrices, *Journal of Chemical Theory and Computation* 9 (2013) 863–867.
- [16] A. Simmonett, B. Brooks, T. Darden, Efficient and scalable electrostatics via spherical grids and treecode summation, 2022. doi:doi:10.26434/chemrxiv-2022-6xzql, unpublished.
- [17] O. Demel, M. J. Lecours, R. Habrovský, M. Noojien, Toward laplace mp2 method using range separated coulomb potential and orbital selective virtuals, *The Journal of chemical physics* 155 (2021) 154104.
- [18] Knowino, Gaussian type orbitals — knowino, an encyclopedia, 2010. URL: http://knowino.org/w/index.php?title=Gaussian_type_orbitals&oldid=3278.
- [19] B. P. Pritchard, D. Altarawy, B. Didier, T. D. Gibson, T. L. Windus, New basis set exchange: An open, up-to-date resource for the molecular sciences community, *Journal of Chemical Information and Modeling* 59 (2019) 4814–4820. PMID: 31600445.
- [20] W. Fong, E. Darve, The black-box fast multipole method, *Journal of Computational Physics* 228 (2009) 8712–8725.
- [21] L. Greengard, V. Rokhlin, A fast algorithm for particle simulations, *Journal of Computational Physics* 73 (1987) 325–348.
- [22] Y. Garniron, T. Applencourt, K. Gasperich, A. Benali, A. Ferté, J. Paquier, B. Pradines, R. Assaraf, P. Reinhardt, J. Toulouse, P. Barbareco, N. Renon, G. David, J.-P. Malrieu, M. Vêril, M. Caffarel, P.-F. Loos, E. Giner, A. Scemama, Quantum package 2.0: An open-source determinant-driven suite of programs, *Journal of Chemical Theory and Computation* 15 (2019) 3591–3609. PMID: 31082265.
- [23] E. Scheiber, On the chebyshev approximation of a function with two variables, 2015.
- [24] A. Townsend, L. N. Trefethen, An extension of chebfun to two dimensions, *SIAM Journal on Scientific Computing* 35 (2013) C495–C518.
- [25] Z. JafariBehbahani, M. Roodaki, Two-dimensional chebyshev hybrid functions and their applications to integral equations, *Beni-Suef University Journal of Basic and Applied Sciences* 4 (2015) 134–141.

- [26] M. Gupta, Numerical methods and software (david kahaner, cleve moler, and stephen nash), Siam Review - SIAM REV 33 (1991).
- [27] S. Liu, O. TRENKLER, Hadamard, khatri-rao, kronecker and other matrix products, International Journal of Information , Systems Sciences 4 (2008).
- [28] T. G. Kolda, B. W. Bader, Tensor decompositions and applications, SIAM Review 51 (2009) 455–500.
- [29] R. B. Platte, L. N. Trefethen, Chebfun: A New Kind of Numerical Computing, Springer Berlin Heidelberg, Berlin, Heidelberg, 2010, pp. 69–87. URL: https://doi.org/10.1007/978-3-642-12110-4_5. doi:doi:10.1007/978-3-642-12110-4_5.
- [30] M. Frigo, S. G. Johnson, The design and implementation of FFTW3, Proceedings of the IEEE 93 (2005) 216–231. Special issue on “Program Generation, Optimization, and Platform Adaptation”.
- [31] J. Dongarra, J. Croz, S. Hammarling, I. Duff, A set of level 3 basic linear algebra subprograms, ACM Transactions on Mathematical Software 16 (1990) 1–17.
- [32] I. Chollet, Symmetries and Fast Multipole Methods for Oscillatory Kernels, Theses, Sorbonne Université, 2021. URL: <https://tel.archives-ouvertes.fr/tel-03203231>.
- [33] J. Barnes, P. Hut, A hierarchical $O(N \log N)$ force-calculation algorithm, Nature 324 (1986) 446–449.
- [34] S. Chaillat, L. Desiderio, P. Ciarlet, Theory and implementation of \mathcal{H} -matrix based iterative and direct solvers for Helmholtz and elastodynamic oscillatory kernels, Journal of Computational Physics (2017).
- [35] M. Bebendorf, Hierarchical matrices, Lecture notes in computational science and engineering, v.63 (2008) 63 (2008).
- [36] W. Hackbusch, Hierarchical Matrices: Algorithms and Analysis, volume 49, 2015. doi:doi:10.1007/978-3-662-47324-5.
- [37] S. A. Losilla, M. A. Watson, A. Aspuru-Guzik, D. Sundholm, Construction of the fock matrix on a grid-based molecular orbital basis using gpgpus, Journal of Chemical Theory and Computation 11 (2015) 2053–2062. PMID: 26574409.
- [38] X. Xing, E. Chow, Fast coulomb matrix construction via compressing the interactions between continuous charge distributions, SIAM Journal on Scientific Computing 42 (2020) A162–A186.
- [39] J. A. Rosal Sandberg, New efficient integral algorithms for quantum chemistry, Ph.D. thesis, KTH, Theoretical Chemistry and Biology, 2014. QC 20140826.
- [40] P. C. Hansen, The truncatedsvd as a method for regularization, BIT Numerical Mathematics 27 (1987) 534–553.
- [41] I. Chollet, X. Claeys, P. Fortin, L. Grigori, A Directional Equispaced interpolation-based Fast Multipole Method for oscillatory kernels, 2022. URL: <https://hal.archives-ouvertes.fr/hal-03563005>, working paper or preprint.
- [42] M. J. Lu B, Cheng X, New-version-fast-multipole-method" accelerated electrostatic interactions in biomolecular systems, J Comput Phys. (2007).
- [43] E. D. Hedegård, S. Knecht, J. S. Kielberg, H. J. A. Jensen, M. Reiher, Density matrix renormalization group with efficient dynamical electron correlation through range separation, The Journal of Chemical Physics 142 (2015) 224108.