An efficient implementation of the GOSTSHYP pressure model by applying shell-bounding gaussian 1-electron-3-center integral screening

Felix Zeller,[†] Eric Berquist,[‡] Evgeny Epifanovsky,[‡] and Tim Neudecker[*,†,¶,§]

†*University of Bremen, Institute for Physical and Theoretical Chemistry, Leobener Str. NW2, D-28359 Bremen, Germany*

‡*Q-Chem Inc., 6601 Owens Drive, Suite 105, Pleasanton, CA 94588, United States of America*

¶*Bremen Center for Computational Materials Science, University of Bremen, Am Fallturm 1, D-28359 Bremen, Germany*

§*MAPEX Center for Materials and Processes, University of Bremen, Bibliothekstr. 1, D-28359 Bremen, Germany*

E-mail: neudecker@uni-bremen.de

Abstract

We implemented a screening algorithm for one-electron-three-center (1e3c) overlap integrals over contracted gaussian-type orbitals (CGTOs) into the Q-Chem program package. The respective bounds were derived using shell-bounding gaussians (SBGs) and the *Obara-Saika* recurrence relations. Using integral screening, we reduced the computational scaling of the Gaussians On Surface Tesserae Simulate HYdrostatic Pressure (GOSTSHYP) model in terms of calculation time and memory usage to a linear relationship with the tesserae used to discretize the surface area. Further code

improvements allowed for additional performance boosts. To demonstrate the better performance, we calculated the compressebility of fullerenes up to C_{180}, where we were originally limited to C_{40} due to the high RAM usage of GOSTSHYP.

Introduction

Over the past decades high pressure chemistry has made significant progress in fields such as organic synthesis, biochemistry, materials science and spectroscopy.[1–5] While enormous progress has been made in experimental setups, especially using diamond-anvil cells,[1–4] another driving force in high pressure chemistry is the synergy between experiment and theory.[6–15] Pressure is most commonly modeled by manipulating box parameters[16–22] in calculations with periodic boundary conditions, which, e.g., allowed the description of pressure-induced changes in crystal structures.[6,23–29]

Since pressure is a macroscopic phenomenon, its description in electronic structure theory for single atoms or molecules is not straightforward. The simplest method to describe the effect of pressure on a single molecule is by adding external forces to its nuclear gradient.[15,30,31] The oldest method using this approach is the Generalized Force-Modified Potential Energy Surface (G-FMPES) method,[31] in which forces pull each atom towards the molecular centroid. With the eXtended Hydrostastatic Compression Force Field (X-HCFF) method[30] we recently proposed a new mechanochemical model, which obtains the required forces by the classical definition of hydrostatic pressure acting on the van der Waals (VDW) surface of a molecule. The great advantage of the X-HCFF method is that it allows to apply a user-defined pressure, rather than a guess. An established alternative to mechanochemical approaches is the eXtreme Pressure Polarizable Continuum Model (XP-PCM),[32–34] which simulates the *Pauli* repulsion of a molecule's electron gas with a surrounding medium. Pressure is then applied by shrinking the molecule-shaped cavity, while the pressure acting on a system is quantified as the negative partial derivative of the electronic energy with respect to the volume.

Recently, we proposed the GOSTSHYP method,[35] which, similar to X-HCFF, uses an ansatz based on a *Lebedev* grid, to generate the discretized VDW surface and ultimately the cavity. However, GOSTSHYP is more sophisticated than X-HCFF in that a pressure poten-

tial term is directly added to the molecular Hamiltonian. This potential term is calculated as a large number of one-electron-three-center (1e3c) overlap integrals between gaussian potentials, located on the discretized VDW surface, and the electron gas. The number of 1e3c overlap integrals formally scales as $\mathcal{O}(n_{bsf}^2 \cdot n_{tess})$, where n_{bsf} is the number of basis functions and n_{tess} is the number of tesserae used to discretize the VDW sphere. Unfortunately, even for small molecules, hundreds of tesserae are needed, which severely increases the computational cost, both with regard to memory and calculation time.[35]

The classical approach in quantum chemistry to reduce calculation times for large amounts of integrals is to employ a pre-screening algorithm based on integral bounds. This is a standard approach in all quantum chemistry packages, where Self-Consistent Field (SCF) calculations are heavily bottlenecked by the large amount of electron repulsion integrals (ERIs). Traditionally, ERI screening algorithms are based on the *Cauchy-Schwarz* inequality,[36,37] however, in recent years more efficient bounds were developed, e.g. by using multipole expansions.[38–41] While most development of integral bounds focuses on ERIs,[38–46] Thompson and Ochsenfeld recently proposed a rather general bound formalism, applicable to various one- and multielectron integrals, based on space partitioning.[42]

Another type of integrals leading to steep computational scaling are effective core potential (ECP) overlap integrals, due to them containing spherical projectors. Over the last years a few screening procedures were developed for ECPs.[47–49] Most notably, McKenzie et al. implemented a powerful screening algorithm for 1e3c ECP integrals based on shell-bounding gaussians (SBGs) in the Q-Chem software package, formally reducing the scaling from $\mathcal{O}(n_{bsf}^2 N)$ to $\mathcal{O}(N)$, where N is the number of ECPs.[47,50] The bounds for unprojected ECP integrals, as derived by McKenzie et al., are equivalent to bounds for 1e3c integrals with a mid center s-type gaussian function, as used to calculate the GOSTSHYP pressure potential. Since 1e3c integrals with up to f-type mid center gaussian-type orbitals (GTOs) are needed to calculate the GOSTSHYP gradient, we found it the natural choice to use the

bounds derived by McKenzie et al. as a basis to screen 1e3c overlap integrals, by generalizing these bounds for mid-center GTOs with arbitrary angular momentum using the Obara-Saika recurrence relations (RRs).[51] This allows us to reduce the scaling of GOSTSHYP to $\mathcal{O}(n_{tess})$.

The rest of the paper is structured as follows: In section 2 we describe the theory behind GOSTSHYP and SBG-based bounds and afterwards show our approach to using the bounds by McKenzie et al. for screening 1e3c overlap integrals over GTOs . In section 3 we shortly describe specifics of our implementation into the Q-Chem program package,[50] followed by the discussion of a small benchmark, showing the improvements in calculation time, in section 4. To demonstrate the improved applicability of GOSTSHYP to large molecules, in the last section we present compressibilities for fullerenes up to C_{180}, where we were previoulsly restricted to C_{40} due to memory limitations.[35]

Theory

The purpose of this section is a brief recapitulation of the GOSTSHYP model, the introduction of the notation and the derivation of the bounds used to screen the 1e3c overlap integrals used in this work.

An overview of the GOSTSHYP model

In the GOSTSHYP model, pressure is applied to a molecular system by adding a distortion potential term E_p to the Hamiltonian.[35] E_p is calculated as the sum of contractions of the density matrix $\boldsymbol{D}$ and 1e3c overlap integrals between the system's basis functions ϕ_i and gaussian pressure potentials $G_j(\boldsymbol{r})$

$$E_p = \sum_j \sum_{ab} D_{ab} \langle \phi_a | G_j(\boldsymbol{r}) | \phi_b \rangle. \tag{1}$$

The potential $G_j(\boldsymbol{r})$ is located at the center $\boldsymbol{r}_j$ of tessera j determined by a *Lebedev Grid*,[52] discretizing the system's VDW surface. $G_j(\boldsymbol{r})$ takes the form

$$G_j(\boldsymbol{r}) = p_j \exp\left(-\omega_j (\boldsymbol{r} - \boldsymbol{r}_j)^2\right) = p_j \tilde{G}_j. \tag{2}$$

The exponents ω_j are chosen such that the pressure potentials form a continuous field around the system under consideration, with $\omega_j = \frac{\pi \ln 2}{A_j}$, where A_j is the area of the tessera j. The amplitudes p_j are obtained by considering a force equilibrium between an outer force, $F_{outer} = A_j \cdot P$, scaling with the pressure P and acting perpendicular to the surface A_j, and the restoring force of the electron gas, $F_{inner} = \frac{\partial E_p}{\partial \boldsymbol{r}_j} \cdot \boldsymbol{n}_j$, in direction of the surface normal vector $\boldsymbol{n}_j$

$$A_j P = \sum_{ab} D_{ab}\ p_j \left(n_x \frac{\partial}{\partial x_j} \langle \phi_a | \tilde{G}_j | \phi_b \rangle + n_y \frac{\partial}{\partial y_j} \langle \phi_a | \tilde{G}_j | \phi_b \rangle + n_z \frac{\partial}{\partial z_j} \langle \phi_a | \tilde{G}_j | \phi_b \rangle \right), \tag{3}$$

leading to

$$p_j = \frac{A_j P}{\sum_{ab} D_{ab} F_{j,ab}}, \tag{4}$$

with the force term $F_{j,ab}$ being determined by overlap integrals with p-type pressure gaussians,

$$F_{j,ab} = 2\omega_j \left(n_x \langle \phi_a | (x - x_j) \tilde{G}_j | \phi_b \rangle + n_y \langle \phi_a | (y - y_j) \tilde{G}_j | \phi_b \rangle + n_z \langle \phi_a | (z - z_j) \tilde{G}_j | \phi_b \rangle \right). \tag{5}$$

The calculation of the overlap matrices $\boldsymbol{G}_j$ and $\boldsymbol{F}_j$ with elements $G_{j,ab} = \langle \phi_a | \tilde{G}_j | \phi_b \rangle$ and $F_{j,ab}$ causes the formal scaling of $\mathcal{O}(n_{bsf}^2 \cdot n_{tess})$ for GOSTSHYP. The *Fock* contribution can be obtained as the derivative $\frac{\partial E_p}{\partial D_{ab}}$ and the analytical nuclear gradient as the derivative of E_p with respect to the nuclear coordinates, leading to overlap integrals with up to f-type pressure gaussians. The interested reader is kindly referred to ref. 35 for the analytical expressions. Since up to f-type orbitals are needed to calculate the GOSTSHYP gradient, we had to generalize the expressions of McKenzie et al. for mid center GTOs with arbitrary angular momentum to use them in GOSTSHYP.

Notation

An unnormalized primitive gaussian-type orbital (GTO) is given as

$$|\boldsymbol{a}] = (x - A_x)^{a_x} (y - A_y)^{a_y} (z - A_z)^{a_z} \exp\left(-\alpha |\boldsymbol{r} - \boldsymbol{A}|^2\right), \tag{6}$$

where $\boldsymbol{A}^T = (A_x, A_y, A_z)$ is its center, α its exponent, $a = a_x + a_y + a_z$ its total angular momentum and $\boldsymbol{a}^T = (a_x, a_y, a_z)$ its angular momentum vector .[47] GTOs are usually grouped in a shell $|a]$, denoted by nonbold writing, which contain all GTOs with the same center, total angular momentum and exponent, only differing in the composition of their angular momentum vector. A contracted gaussian-type orbital (CGTO) is built as a linear

combination from primitives

$$|\boldsymbol{a}\rangle = \sum_{i}^{n_a} c_i |\boldsymbol{a}]_i, \tag{7}$$

where c_i are the contraction coefficients and n_a is the number of contracted primitives. Thus, a two-center integral of CGTOs is a sum of integrals between primitives:

$$\langle \boldsymbol{a}|\boldsymbol{b}\rangle = \sum_{i,j}^{n_a,n_b} c_i c_j [\boldsymbol{a}_i|\boldsymbol{b}_j]. \tag{8}$$

A shell-bounding gaussian (SBG) $|\mathring{a}]$ is an s-type GTO that bounds a primitive shell $|a]$,[47] so that

$$|\mathring{a}] \geq ||\boldsymbol{a}]| \qquad \forall\ |\boldsymbol{a}] \in |a]. \tag{9}$$

The SBG is always positive and is described as

$$|\mathring{a}] = N_a \exp(-\tilde{\alpha}(\boldsymbol{r} - \boldsymbol{A})^2), \tag{10}$$

with the effective exponent $\tilde{\alpha} = \alpha(1 - \sigma_a)$ and

$$N_a = \left(\frac{a}{2e\alpha\sigma_a}\right)^{\frac{a}{2}}. \tag{11}$$

σ_a is an arbitrary parameter that is used to minimize the bound.

Bound equations using shell-bounding gaussian

The following recapitulates the bound derivations of McKenzie et al..[47] While originally derived for ECP integrals,[47] the bound equations are also valid for 1e3c overlap integrals with a mid center s-type GTO.

Using eq. 9, a 1e3c integral of primitive GTOs $[\boldsymbol{a}|s|\boldsymbol{b}]$ with an s-type mid center GTO $|s]$, which also describes a shell composed of a single GTO and thus is written nonbold, can be bound as:

$$|[\boldsymbol{a}|s|\boldsymbol{b}]| \leq [\mathring{a}|s|\mathring{b}] \qquad \forall\ |\boldsymbol{a}], |\boldsymbol{b}] \in |a], |b]. \tag{12}$$

This allows to screen all integrals between the two shells $|a]$ and $|b]$ over $|s]$ by calculating one bound. Using the Gaussian Product Relation and Hölder's Inequality, the bound can be separated into

$$[\mathring{a}|s|\mathring{b}] \leq [\mathring{a}|s]_\infty [\mathring{b}], \tag{13}$$

where $[\mathring{a}|s]_\infty$ is the essential supremum of the product of $|\mathring{a}]$ and $|s]$ and $\mathring{b}$ the spatial integral of $\mathring{b}$. Eq. 13 can be made independent of the particular shell $|b]$ by minimizing with respect to σ_b, setting b to the highest available angular momentum $\hat{b}$ and β to the smallest exponent $\check{\beta}$ in the basis set. The result is

$$[\mathring{b}]_{max} = \left(\frac{(\hat{b}+3)^{\hat{b}+3}(\pi/3)^3}{(2e)^{\hat{b}}\check{\beta}^{\hat{b}+3}} \right). \tag{14}$$

The bound $[B_2]$ to eq. 13 is then obtained by minimizing $[\mathring{a}|s]_\infty$ with respect to σ_a

$$[B_2] = N_a \exp\left(-\frac{\tilde{\alpha}\gamma(\boldsymbol{A}-\boldsymbol{C})^2}{\tilde{\alpha}+\gamma} \right), \tag{15}$$

with

$$\sigma_a = \frac{a(\alpha+\gamma)^2}{2\alpha(\ \gamma^2(\boldsymbol{A}-\boldsymbol{C})^2 + a(\alpha+\gamma))}. \tag{16}$$

The 1e3c integrals thus can be screened by comparing $[B_2]$ with an arbitrary threshold τ:

$$|[\boldsymbol{a}|s|\boldsymbol{b}]| \leq [B_2][\mathring{b}]_{max} \leq \tau \qquad \forall\ |\boldsymbol{a}], |\boldsymbol{b}] \in |a], |b]. \tag{17}$$

This will lead to a list of shell triplets containing nonzero overlap integrals, which then can

be screened again using the more expensive three-center bound equation, which is obtained by directly evaluating eq. 12, using previously calculated effective exponents:

$$[B_3] = [\mathring{a}|s|\mathring{b}] \tag{18}$$

While bounds for CGTO integrals could be obtained straightforwardly using eq. 7, the following equation strongly reduces the computational cost:[47]

$$\langle B_3\rangle \leq \langle\mathring{a}\rangle\langle\mathring{b}\rangle[\mathring{\check{a}}|\check{s}|\mathring{\check{b}}], \tag{19}$$

with

$$\langle\mathring{a}\rangle = \sum_i^{n_a} |c_i^a| N_{ia} \tag{20}$$

and

$$[\mathring{\check{a}}|\check{s}|\mathring{\check{b}}] = \left(\frac{\pi}{\check{\check{\zeta}}}\right)^{(3/2)} \exp\left(-\check{\check{\alpha}}(\boldsymbol{A}-\boldsymbol{S})^2 - \check{\check{\beta}}(\boldsymbol{B}-\boldsymbol{S})^2 + \check{\check{\zeta}}\check{\check{\boldsymbol{R}}}^2\right), \tag{21}$$

where $\check{\check{\zeta}} = \check{\check{\alpha}} + \check{\check{\beta}} + \check{\gamma}$ and $\check{\check{\boldsymbol{R}}} = (\check{\check{\alpha}}(\boldsymbol{A}-\boldsymbol{S}) + \check{\check{\beta}}(\boldsymbol{B}-\boldsymbol{S}))/\check{\check{\zeta}}$. $\check{\check{\alpha}}$ denotes the usage of the smallest effective exponents within the CGTO and γ is the exponent of $|s]$.
Using this approach allowed McKenzie et al. to reduce the scaling of ECP integrals from from $\mathcal{O}(n_{nbsf}^2 N)$ to $\mathcal{O}(N)$, where N is the number of ECPs.[47]

Recurrence relations

Equations that allow to express integrals of an angular momentum l in terms of a sum over integrals with lower angular momentum are called RRs. By using RRs, integrals with arbitrary angular momenta can be efficiently computed as linear combinations of integrals over s-type GTOs. The general RR for a 1-electron-3-center integral was derived by Obara and Saika as[51]

$$[\boldsymbol{a}|\boldsymbol{c}+\mathbf{1}_i|\boldsymbol{b}] = (Q_i - C_i)[\boldsymbol{a}|\boldsymbol{c}|\boldsymbol{b}] + \frac{1}{2(\alpha+\beta+\gamma)} N_i(\boldsymbol{a})[\boldsymbol{a}-\mathbf{1}_i|\boldsymbol{c}|\boldsymbol{b}]$$
$$+\frac{1}{2(\alpha+\beta+\gamma)} N_i(\boldsymbol{b})[\boldsymbol{a}|\boldsymbol{c}|\boldsymbol{b}-\mathbf{1}_i] + \frac{1}{2(\alpha+\beta+\gamma)} N_i(\boldsymbol{c})[\boldsymbol{a}|\boldsymbol{c}-\mathbf{1}_i|\boldsymbol{b}]. \tag{22}$$

where $Q = \frac{\alpha\boldsymbol{A}+\beta\boldsymbol{B}+\gamma\boldsymbol{C}}{\alpha+\beta+\gamma}$. The notation $|\boldsymbol{a}-\mathbf{1}_i]$ denotes a GTO with an angular momentum vector reduced by one in the spatial direction of i compared to $|\boldsymbol{a}]$. $N_i(\boldsymbol{a})$, $N_i(\boldsymbol{b})$ and $N_i(\boldsymbol{c})$ correspond to sums of Kronecker Deltas and are evaluated for an exemplary d-function as $N_i(\mathbf{1}_j+\mathbf{1}_k) = N_i(\mathbf{1}_j) + N_i(\mathbf{1}_k) = \delta_{ij} + \delta_{ik}$. After 'shifting' the coordinate origin to the mid center shell to be consistent with the bound equations of McKenzie et al.,[47] a 1e3c integral with *bra* and *ket* SBGs $|\mathring{a}]$ and $|\mathring{b}]$ becomes

$$[\mathring{a}|\boldsymbol{c}+\mathbf{1}_i|\mathring{b}] = \tilde{R}_i[\mathring{a}|\boldsymbol{c}|\mathring{b}] + \frac{1}{2\tilde{\zeta}} N_i(\boldsymbol{c})[\mathring{a}|\boldsymbol{c}-\mathbf{1}_i|\mathring{b}], \tag{23}$$

where the two middle terms in eq. 22 vanish since $N_i(\mathbf{0}) = 0$. The corresponding two-center RR relation is obtained by replacing $\boldsymbol{b}$ with the identity:

$$[\mathring{a}|\boldsymbol{c}+\mathbf{1}_i] = \tilde{P}_i[\mathring{a}|\boldsymbol{c}] + \frac{1}{2\tilde{\kappa}} N_i(\boldsymbol{c})[\mathring{a}|\boldsymbol{c}-\mathbf{1}_i], \tag{24}$$

where $\tilde{\boldsymbol{P}} = \frac{\tilde{\alpha}(\boldsymbol{A}-\boldsymbol{C})}{\tilde{\alpha}+\gamma}$ and $\tilde{\kappa} = \tilde{\alpha}+\gamma$.

Bounds for 1e3c overlap integrals

We generalized the previously shown bounds for an arbitrary mid center CGTO using the Obara-Saika RRs,[51] using the ansatz:

$$[\boldsymbol{a}|\boldsymbol{c}|\boldsymbol{b}] \leq |[\mathring{a}|\boldsymbol{c}_{max}|\mathring{b}]|, \tag{25}$$

where $|\boldsymbol{c}_{max}]$ is the GTO in the shell $|c]$ maximizing the integral. Thus, as in the algorithm of McKenzie et al., only one bound has to be calculated to screen all 1e3c integrals of the

shells $|a]$, $|b]$ and $|c]$. In a shell, all exponents are equal and accordingly the Obara-Saika equations (eq 22-24) will only differ in the Cartesian part. Thus the angular momentum vector $\boldsymbol{c}_{max}$ of $|\boldsymbol{c}_{max}]$ is purely pointing towards $\tilde{P}_i$, the maximum component of $|\tilde{\boldsymbol{P}}|$, and can be written in the Obara-Saika notation as a sum of c $\mathbf{1}_i$ vectors

$$\boldsymbol{c}_{max} = \sum_{n=1}^{c} \mathbf{1}_{n,i} = \mathbf{1}_{1,i} + \mathbf{1}_{2,i} + \cdots + \mathbf{1}_{c,i} = c \cdot \mathbf{1}_i \qquad i \in x, y, z, \tag{26}$$

where i is the angular direction of the maximum component. To apply eq. 13, we now have to evaluate the integral $[\mathring{a}|\boldsymbol{c}_{max}]$. Applying eq. 24 once leads to

$$[\mathring{a}|\boldsymbol{c}_{max}] = \ |\tilde{P}_i|[\mathring{a}|(c-1)\mathbf{1}_i] + \frac{N_i((c-1)\mathbf{1}_i))}{2\tilde{\kappa}}[\mathring{a}|(c-2)\mathbf{1}_i], \tag{27}$$

which, using the relation $N_i(\mathbf{1}_i + \mathbf{1}_{\boldsymbol{j}}) = N_i(\mathbf{1}_i) + N_i(\mathbf{1}_{\boldsymbol{j}})$,[51] can be simplified to

$$|[\mathring{a}|\boldsymbol{c}_{max}]| = \ |\tilde{P}_i|[\mathring{a}|(c-1)\mathbf{1}_i] + \frac{c-1}{2\tilde{\kappa}}[\mathring{a}|(c-2)\mathbf{1}_i]. \tag{28}$$

Applying Eq. 24 $k = c/2$ times will lead to

$$\begin{aligned}
[\mathring{a}|\boldsymbol{c}_{max}] =& \eta|\tilde{P}_i|^k[\mathring{a}|(c-k)\mathbf{1}_i] + \eta\frac{|\tilde{P}_i|^{k-1}}{2\tilde{\kappa}}[\mathring{a}|(c-k-1)\mathbf{1}_i] \\
&+ \cdots + \eta\frac{|\tilde{P}_i|}{(2\tilde{\kappa})^{k-1}}[\mathring{a}|(c-k-2)\mathbf{1}_i] + \eta\frac{1}{(2\tilde{\kappa})^k}[\mathring{a}|(c-2k)\mathbf{1}_i] \qquad (29) \\
=& \eta|\tilde{P}_i|^k[\mathring{a}|(c-k)\mathbf{1}_i] + \eta\frac{|\tilde{P}_i|^{k-1}}{2\tilde{\kappa}}[\mathring{a}|(c-k-1)\mathbf{1}_i] \\
&+ \cdots + \frac{|\tilde{P}_i|}{(2\tilde{\kappa})^{k-1}}[\mathring{a}|\mathbf{1}_i] + \eta\frac{1}{(2\tilde{\kappa})^k}[\mathring{a}|0 \cdot \mathbf{1}_i] \qquad (30)
\end{aligned}$$

where the prefactor η is individual for each term and arises from the combination of equal terms and the $(c-1)$ multiplicator. To reduce all integrals to s-type integrals, the RR has to be applied again, up to k times to the most left, $k-1$ times to the second most left term, and so on:

$$[\mathring{a}|\boldsymbol{c}_{max}] = \eta|\tilde{P}_i|^{2k}[\mathring{a}|s_c] + \eta\frac{|\tilde{P}_i|^{2(k-1)}}{2\tilde{\kappa}}[\mathring{a}|s_c] + \cdots + \eta\frac{|\tilde{P}_i|^2}{(2\tilde{\kappa})^{k-1}}[\mathring{a}|s_c] + \eta\frac{1}{(2\tilde{\kappa})^k}[\mathring{a}|s_c] \tag{31}$$

Here $|s_c]$ denotes an s-type primitive with the center and exponent of $|c]$. Since $2k = c$ the integral is expressed in the sum

$$[\mathring{a}|\boldsymbol{c}_{max}] = \sum_{0\leq k\leq \frac{c}{2}} \eta\frac{|\tilde{P}_i|^{c-2k}}{(2\tilde{\kappa})^k}[\mathring{a}|s_c] \tag{32}$$

To determine the prefactor η one has to think of the RR as a two case combinatory problem, where the first case will reduce the angular momentum by one and the second case will reduce the angular momentum by two and multiply the result with $m-1$, where m is the angular momentum of the specific integral the RR is applied to. Thus, the total number of paths which lead to a term specified by k in eq. 32 will be $c-k$, where k is the number of times the second case was chosen. Consequently the number of possibilities to create a term k is given by the binomial coefficient $\binom{c-k}{k}$.

The prefactor of a *specific* term will depend on the specific angular momenta for which the second case was applied, since they define the $m-1$ multiplicator. Accordingly, the determination of the exact prefactor η is not trivial. However one can easily define a maximum prefactor η_k^{max} for each term by assuming that the second case is always applied to the highest possible angular momentum integral:

$$\eta_k^{max} = \prod_{m=0}^{k-1}(c-1-2m)\binom{c-k}{k} > \eta_k. \tag{33}$$

Using η_k^{max}, the final bounds equation can be obtained by evaluating $[\mathring{a}|s_c]_\infty$ as

$$[B_2]_c = \rho_2(c)[B_2] = \left(\sum_{0\leq k\leq \frac{c}{2}}\prod_{m=0}^{k-1}(c-1-2m)\binom{c-k}{k}\frac{|\tilde{P}_i|^{c-2k}}{(2\tilde{\kappa})^k}\right)[B_2], \tag{34}$$

where $\rho_2(c)$ is the bound correction factor for the angular momentum c. The 3-center bound

equation can be derived analogously, using eq. 22 to evaluate eq. 25, yielding:

$$[B_3]_c = \rho_3(c)[B_3] = \left(\sum_{0 \leq k \leq \frac{c}{2}} \prod_{m=0}^{k-1} (c-1-2m) \binom{c-k}{k} \frac{|\tilde{R}_i|^{c-2k}}{(2\tilde{\zeta})^k} \right) [B_3]. \tag{35}$$

where $\rho_3(c)$ is the bound correction factor for the angular momentum c.

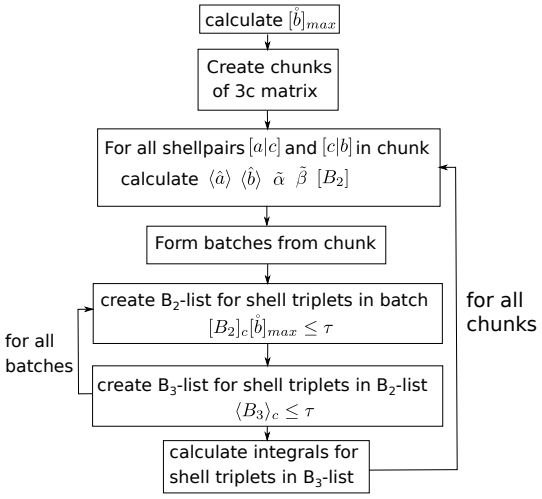
Implementation

As in the original GOSTSHYP implementation,[35] we implemented our screening algorithm together with additional code improvements in the Q-Chem[50] program package. The integral screening algorithm is located in *libqints*, Q-Chem's new integral library, and is generally applicable to 1e3c overlap integrals of gaussian functions. The screening algorithm is embedded into a 1e3c overlap integral engine, allowing to calculate specified chunks of the overlap matrix, which avoids the need to allocate memory for the complete overlap matrix. Additional code improvements were implemented directly into the GOSTSHYP model located in *libdistort*, a library concerned with modeling molecules under distortion potentials. Our code improvements were part of the Q-Chem 5.4.2 update.

The implementation of the screening algorithm follows the structure of the implementation of the ECP screener by McKenzie et al..[47] The general workflow of the overlap integral engine is shown in scheme 1. First, the basis constant $[\hat{b}_{max}]$ is calculated. Then a chunk of the overlap matrix is specified, for which the shell bounds $[B_2]$, effective components $\tilde{\alpha}$ and contractions $\langle \hat{a} \rangle$ are calculated. These shells are then processed in batches, where for each batch a list of non-negligible integrals is created using the $[B_2]$ bounds. Afterwards, listed integrals are compared to the tighter $[B_3]$ bounds and non-negligible integrals are evaluated using the Obara-Saika RRs.[51]

The general bounds, eqs. 34 and 35, contain a binomial coefficient, making them expensive to calculate. Therefore, we calculated and implemented the explicit η_k for integrals containing mid center p, d and f shells, which is the highest mid center angular momentum occurring in overlap integrals in GOSTSHYP.[35] The explicit bound correction factors $\rho(c)$ are shown in tab. 1.

The new integral engine allowed further code improvements. In the original implementation of GOSTSHYP, only the force matrix $\boldsymbol{F}$ (eq. 5) was kept in memory during a SCF calculation, which led to a significant demand of RAM. Our implementation allows to cal-



Scheme 1: Function of new 1e3c overlap integral engine with embedded integral screening.

Table 1: Derived explicit bound corrections for mid center p, d and f shells, for use in eqs. 34 and 35.

c	$\rho_2(c)$	$\rho_3(c)$
p	$\tilde{P}$	$\tilde{R}$
d	$\frac{3}{2\kappa} + \tilde{P}^2$	$\frac{3}{2\tilde{\zeta}} + \tilde{R}^2$
f	$\tilde{P}^3 + \frac{9}{2\kappa}\tilde{P}$	$\tilde{R}^3 + \frac{9}{2\zeta}\tilde{R}$

culate the full overlap matrices $\boldsymbol{G}$ and $\boldsymbol{F}$ in slices $\boldsymbol{G}_j$ and $\boldsymbol{F}_j$ and store them compressed as sparse matrices containing only non-negligible integrals. This allowed us to speed up calculations by additionally storing the 1e3c overlap matrix in memory, while still significantly reducing the RAM usage of GOSTSHYP. Furthermore, using sparse matrices speeds up the calculation of the energy and Fock contribution as well as the amplitudes by reducing the number of operations needed to perform the respective matrix algebra. Since the calculation of energy contributions (eq. 1) and amplitudes (eq. 4) can be performed separately for each tessera j, we also parallelized those calculations, which yielded an additional performance boost.

As a drawback, the integral screening introduces numerical instabilities, caused by the definition of the pressure amplitudes p_j (eq. 4). Since the amplitudes scale with the inverse of the contraction of the density with $\boldsymbol{F}_j$, for small overlaps, the error introduced by the integral screening will become large and in the extreme case the screening will lead to a division

by zero. To prevent this, we embedded the screening of integrals involved on building $\boldsymbol{F}_j$ into an algorithm which ensures that a precision of at least 8 significant digits is kept after integral screening. For this purpose, the integrals are screened with a threshold of 10^{-18} and afterwards the largest absolute in $\boldsymbol{F}_j$ has to be at least eight orders of magnitude larger than the smallest absolute. If this condition is not fulfilled, the integrals are recomputed iteratively until the condition is fulfilled, whereby the thresh is lowered by three orders of magnitude in each iteration.

A similar problem is found in the calculation of the energy contribution, where small overlaps in the matrix elements $\boldsymbol{G}_j$ are multiplied with a large amplitude p_j. In this case, the introduced error by integral screening might be significantly larger than expected from a chosen threshold. This is prevented by including amplitudes larger 10^5 $a.u.$ into the integrals and recalculating $\boldsymbol{G}_j$ for this tessera.

These procedures make the influence of screening thresholds on calculation times hardly predictable and might actually increase the calculation time for larger thresholds. Thus, we decided to set the threshold for screening integrals building $\boldsymbol{G}_j$ to 10^{-14} and make the screening threshold not changeable for the user. The thresholds were chosen such that the differences in the Fock contribution, amplitudes and gradient contribution were below 10^{-12} $a.u.$ compared to the original implementation for our test cases of H_2, H_2O and LiH at a pressure of 50 GPa with a scaling factor of the atomic VDW radii of 1.2 and 302 tesserae per atom. To asses the impact of our integral screening and the chosen thresholds, we measured the timing for the calculation of the GOSTSHYP matrices of n-dodecane, where the calculation time was reduced from 60.3 s to 25.3 s, when running on one thread.

To check the deviations caused by the screening algorithm and to show the overall reduction in computational cost, we performed a benchmark, which will be discussed in the following section.

Benchmark

To evaluate the performance of the updated GOSTSHYP code, we conducted a benchmark containing geometry optimizations for 31 small to medium-sized molecules, shown in table 2, at pressures of 10, 35, 60, 85 and 110 GPa. All optimizations were performed at the HF/cc-pVDZ[53–58] level of theory using the Direct Inversion in Iterative Subspace (DIIS) algorithm.[59] Each calculation was performed once with the original implementation in Q-Chem 5.4.1 and once with the updated implementation (development version of Q-Chem 5.4.2), totalling 310 calculations. Calculations were performed on a workstation equipped with an Intel Core i7-8700K with 6 cores and 12 threads running at a clock speed of 3.7 GHz, using all available threads. The scripts to perform and evaluate the benchmark have been uploaded to GitHub. [1]

Table 2: Molecules contained in the benchmark set, with the corresponding number of basis functions n_{bsf} using the cc-pVDZ basis set.

molecule	n_{bsf}	molecule	n_{bsf}	molecule	n_{bsf}	molecule	n_{bsf}
annulene	342	cysteine	137	hydrogen	10	nitrobenzene	151
ammonia	29	d-glucose	228	isohexane	154	n-tetradecane	346
asparagine	166	d-mannose	228	leucine	191	octane	202
benzene	114	d-ribose	190	methane	34	oxygen	28
benzothiophene	160	dithioformic acid	60	methanol	48	p-xylol	162
carbon dioxide	42	formic acid	52	methionine	185	t-butyl-ethane	154
carbon disulfide	50	hexane	154	n-dodecane	298	water	24
CHBrClF	78	histidine	199	n-hexadecane	394		

We first conducted all calculations with a VDW sphere scaling factor of 1.0. However, we found severe convergence problems. In particular, 156 optimizations in total stopped due to reaching the limit of SCF cycles. We found the problematic SCF convergence to be linked to the appearance of negative amplitudes p_j. Those are nonphysical artifacts and can appear on edges of the Lebvedev grid, where two VDW spheres overlap.[35] The contribution of negative amplitudes to the pressure potential term E_p, eq. 1, and consequently to the Fock

[1] Code available at https://github.com/zellerf/gostshyp-542-benchmark

matrix is omitted in our code. Thus, if the number of negative amplitudes changes during the SCF procedure, the Fock matrix will change uncontinuously. This will introduce errors in the new Fock guess matrices and thus cause problematic convergence. Since negative amplitudes are far more likely to appear for small VDW-scaling factors, we recalculated the benchmark set with a scaling factor of 1.5, which reduced the number of calculations reaching the limit of SCF cycles to 12. In the case of the appearance of many negative amplitudes in a GOSTSHYP calculation we consequently suggest the increase of the scaling factor or the reduction of the DIIS subspace. In the following only calculations with a scaling factor of 1.5 are discussed.

Figs. 1 a-d show the average computation time per SCF cycle and gradient calculation, using our new screened GOSTSHYP implementation versus the original implementation. Since unscreened GOSTSHYP formally scales as $\mathcal{O}(n_{bsf}^2 \cdot n_{tess})$, regression lines for plots of calculation times and RAM usage vs. t n_{bsf} (Figs. 1a, c, e) were fitted using a quadratic model, while in the corresponding plots vs. n_{tess} (Figs. 1b, d, f) a linear model was used. Figs. 1b, 1d and 1f show that calculation times and RAM usage are rather well described by a linear relationship to n_{tess} for the new implementation. Since n_{bsf}, as n_{tess}, is correlated with the system size, calculation times and RAM usage should also show a linear relationship in the new implementation, when plotted vs. n_{nbsf}. It should be noted that the fits vs. n_{bsf} for the new implementation still contain a significant quadratic component for SCF and gradient times (figs. 1a and 1c). However, while GOSTSHYP is the most time consuming step, SCF and gradient calculations also contain the calculation of ERIs, which scale asymptotically to $\mathcal{O}(n_{bsf}^2)$,[60] leading to a polynomial relation. We are therefore confident to propose a reduction in scaling from $\mathcal{O}(n_{bsf}^2 \cdot n_{tess})$ to $\mathcal{O}(n_{tess})$ for GOSTSHYP. It should be noted that calculation times and RAM usage are still heavily dependent on the chosen basis set which will determine the slope of the linear scaling.

To asses the error introduced by the integral screening we compared the energy difference

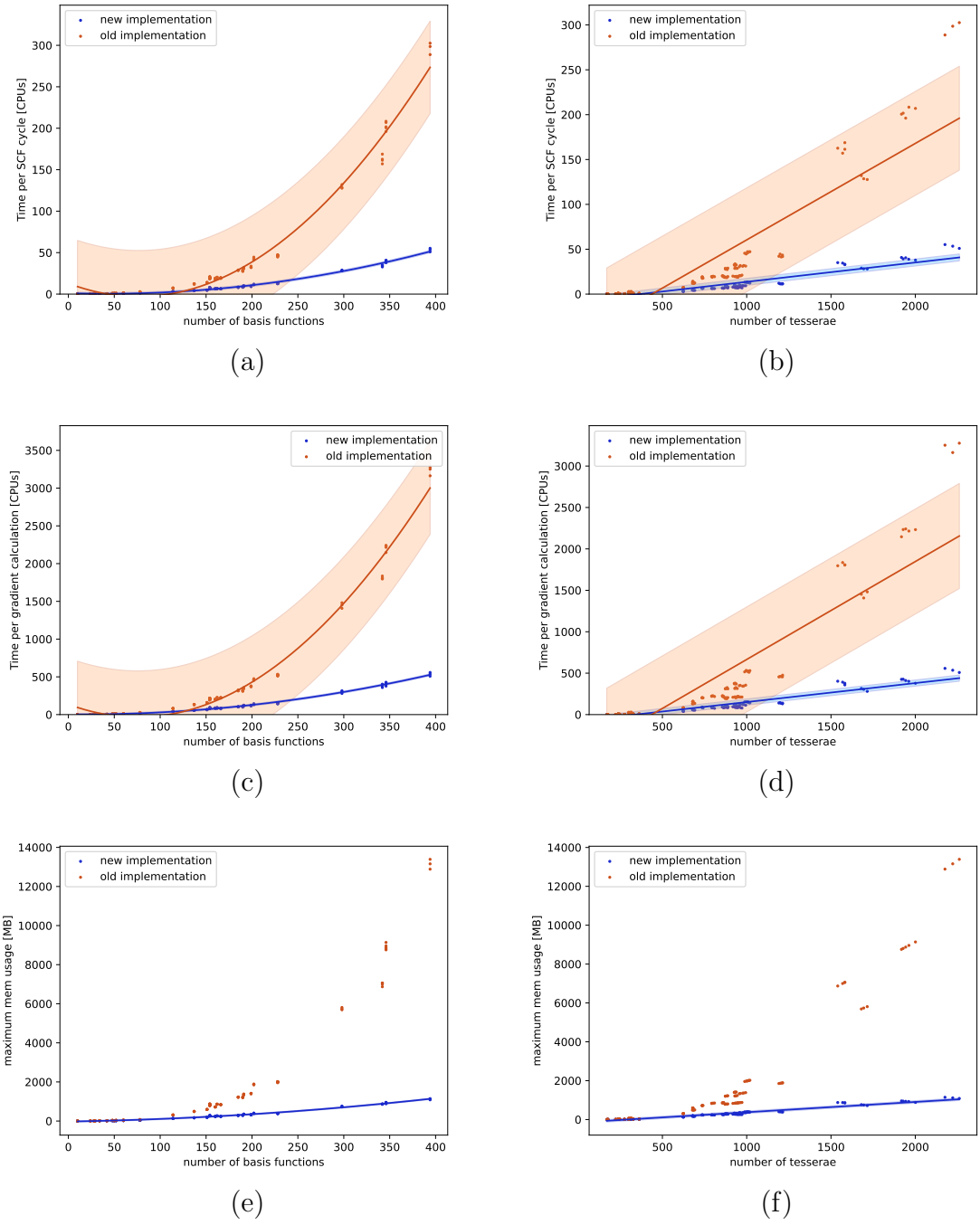


Figure 1: Calculation times using the screened and unscreened codes for SCF (a,b) and gradient computations (c,d) plotted against number of basis functions n_{bsf} (left) and number of tesserae n_{tess} (right). The graphs e and f show the maximum RAM usage. Since in the previous implementation no sparse matrices were used, the mem usage can be directly calculated from n_{tess} and n_{bsf} and thus no regression line was drawn. Shaded areas are within the root-mean-square deviation (RMSE).

between converged optimizations with the old and new implementations. As can be seen in fig. 2a, in all cases the energy difference is far below the set energy change convergence criterion of 10^{-6} *a.u.*. Accordingly, the error introduced by integral screening in GOSTSHYP can be assumed to be negligible.

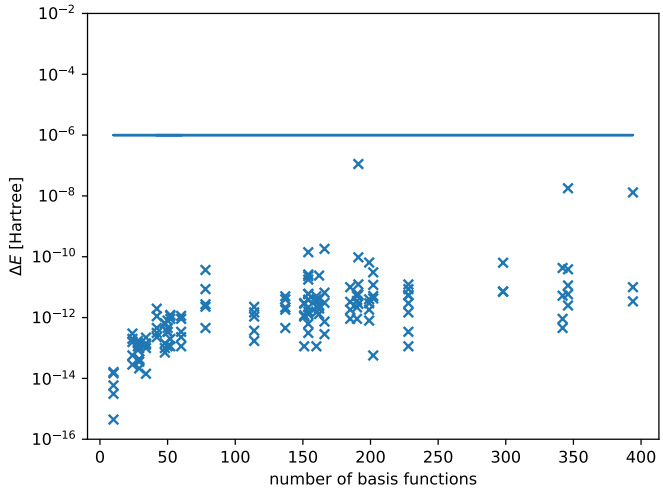


Figure 2: Deviations in energy for geometry optimization calculations with different maximum number of SCF cycles. The straight line marks the convergence criterion for the energy change in a geometry optimization. Unconverged calculations are masked.

Calculating the molecular volume of Fullerenes

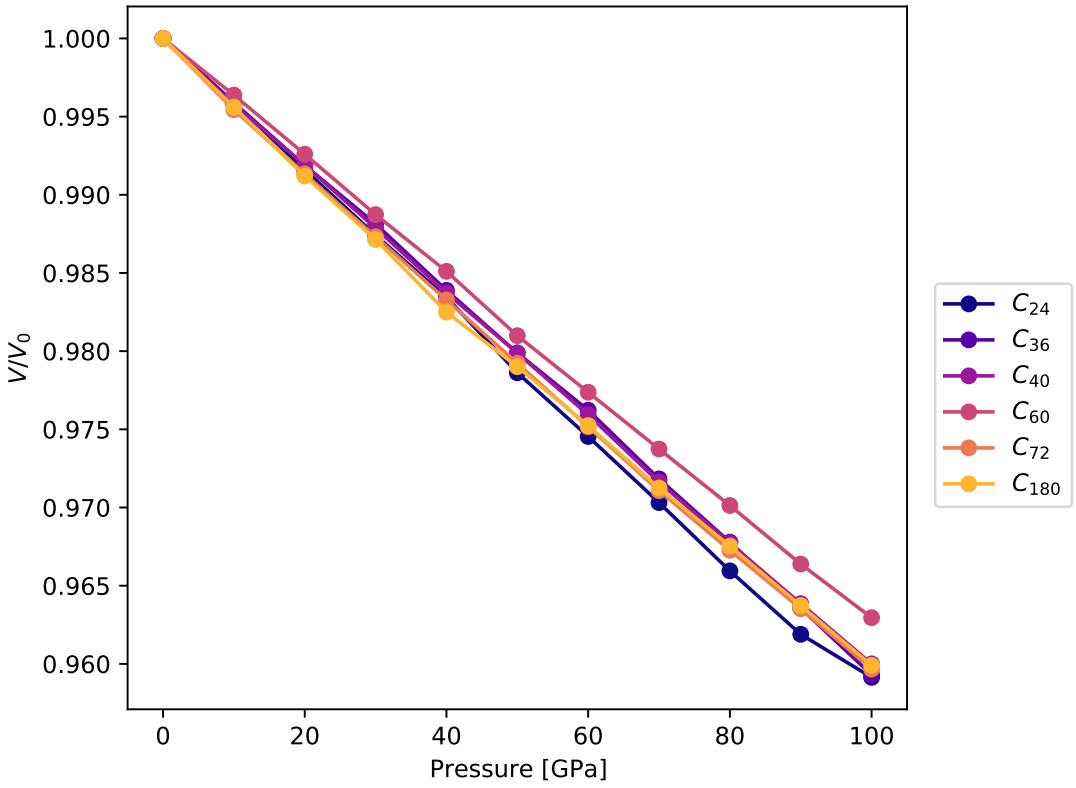


Figure 3: Compressibilities of fullerenes calculated by the relative reduction of the electron density from a GOSTSHYP calculation compared to the relaxed density.

When we presented the GOSTSHYP model,[35] we calculated the compressibility of fullerenes up to C_{40}, by investigating the reduction in volume of the respective electron gas. Calculations for larger fullerenes were not possible, due to the originally steep scaling of RAM usage. After re-investigating those calculations we found an error in our submission script for C_{36} and C_{40}, where the electron densities in these calculations were not compressed by the correct pressure. Within this work we corrected those mistakes and expanded our investigation to larger fullerenes, most notably C_{180} demonstrating the reduction of computational cost for GOSTSHYP. The calculated compressibilities for pressures up to 100 GPa are shown in fig. 3. All fullerenes exhibit similar compressibilities, with the slight trend of larger fullerenes being less compressible. Unsurprisingly, C_{60} is the least compressible fullerene.

Conclusion and Outlook

In this paper we presented bounds for a screening algorithm for 1e3c overlap integrals over CGTOs. We used this screening algorithm, together with additional code improvements, to reduce the computational scaling of our recently proposed GOSTSHYP[35] method from $\mathcal{O}(n_{bsf}^2 \cdot n_{tess})$ to $\mathcal{O}(n_{tess})$. This allowed us to investigate significantly larger molecules, such as C_{180}, than with the original implementation. However, due to the sheer amount of additional integrals to be computed, GOSTSHYP still remains rather expensive in comparison to a pressure-free HF/DFT calculation. Therefore, future development will be devoted to modifications of the GOSTSHYP model to reduce the amount of integrals needed to calculate the Fock and gradient contributions. To allow users to balance computational time and memory usage, we also plan the implementation of a 'low memory' algorithm for GOSTSHYP.

Supporting Information Available

xyz files for fullerenes and molecules of the benchmark.

References

(1) Grochala, W.; Hoffmann, R.; Feng, J.; Ashcroft, N. W. The chemical imagination at work in very tight places. *Angewandte Chemie - International Edition* **2007**, *46*, 3620–3642.

(2) Schettino, V.; Bini, R. Constraining molecules at the closest approach: chemistry at high pressure. *Chemical Society Reviews* **2007**, *36*, 869–880.

(3) Jenner, G. Role of the Medium in High Pressure Organic Reactions. A Review. *Mini-Reviews in Organic Chemistry* **2004**, *1*, 9–26.

(4) Mao, H. K.; Chen, X. J.; Ding, Y.; Li, B.; Wang, L. Solids, liquids, and gases under high pressure. *Reviews of Modern Physics* **2018**, *90*, 15007.

(5) Mc Millan, P. F. Chemistry at high pressure. *Chemical Society Reviews* **2006**, *35*, 855–857.

(6) Rychkov, D. A. A short review of current computational concepts for high-pressure phase transition studies in molecular crystals. *Crystals* **2020**, *10*, 81.

(7) Zhang, T.; Shi, W.; Wang, D.; Zhuo, S.; Peng, Q.; Shuai, Z. Pressure-induced emission enhancement in hexaphenylsilole: A computational study. *Journal of Materials Chemistry C* **2019**, *7*, 1388–1398.

(8) Walker, M.; Morrison, C. A.; Allan, D. R. Nitric acid monohydrates at high pressure: An experimental and computational study. *Physical Review B - Condensed Matter and Materials Physics* **2005**, *72*, 224106.

(9) Ghosh, P. S.; Ali, K.; Arya, A. A computational study of high pressure polymorphic transformations in monazite-type LaPO4. *Physical Chemistry Chemical Physics* **2018**, *20*, 7621–7634.

(10) Colligan, M.; Forster, P. M.; Cheetham, A. K.; Lee, Y.; Vogt, T.; Hriljac, J. A. Synchrotron X-ray powder diffraction and computational investigation of purely siliceous zeolite Y under pressure. *Journal of the American Chemical Society* **2004**, *126*, 12015–12022.

(11) Wiebke, J.; Pahl, E.; Schwerdtfeger, P. Melting at high pressure: Can first-principles computational chemistry challenge diamond-anvil cell experiments? *Angewandte Chemie - International Edition* **2013**, *52*, 13202–13205.

(12) Hobday, C. L.; Woodall, C. H.; Lennox, M. J.; Frost, M.; Kamenev, K.; Düren, T.; Morrison, C. A.; Moggach, S. A. Understanding the adsorption process in ZIF-8 using high pressure crystallography and computational modelling. *Nature Communications* **2018**, *9*, 1429.

(13) Biedermann, N.; Speziale, S.; Winkler, B.; Reichmann, H. J.; Koch-Müller, M.; Heide, G. High-pressure phase behavior of SrCO3: an experimental and computational Raman scattering study. *Physics and Chemistry of Minerals* **2017**, *44*, 335–343.

(14) Gatta, G. D.; Tabacchi, G.; Fois, E.; Lee, Y. Behaviour at high pressure of Rb7NaGa8Si12O40·3H2O (a zeolite with EDI topology): a combined experimental–computational study. *Physics and Chemistry of Minerals* **2016**, *43*, 209–216.

(15) Stauch, T. Quantum chemical modeling of molecules under pressure. *International Journal of Quantum Chemistry* **2021**, *121*, e26208.

(16) Pašteka, L. F.; Helgaker, T.; Saue, T.; Sundholm, D.; Werner, H. J.; Hasanbulli, M.; Major, J.; Schwerdtfeger, P. Atoms and molecules in soft confinement potentials. *Molecular Physics* **2020**, *118*, e1730989.

(17) Novoa, T.; Contreras-García, J.; Fuentealba, P.; Cárdenas, C. The Pauli principle and the confinement of electron pairs in a double well: Aspects of electronic bonding under pressure. *Journal of Chemical Physics* **2019**, *150*, 204304.

(18) Gorecki, J.; Byers Brown, W. On the ground state of the hydrogen molecule-ion H2+ enclosed in hard and soft spherical boxes. *The Journal of Chemical Physics* **1988**, *89*, 2138–2148.

(19) LeSar, R.; Herschbach, D. R. Electronic and vibrational properties of molecules at high pressures. Hydrogen molecule in a rigid spheroidal box. *Journal of Physical Chemistry* **1981**, *85*, 2798–2804.

(20) Ley-Koo, E.; Rubinstein, S. The hydrogen atom within spherical boxes with penetrable walls. *The Journal of Chemical Physics* **1979**, *71*, 351–357.

(21) Ludeña, E. V. SCF calculations for hydrogen in a spherical box. *The Journal of Chemical Physics* **1977**, *66*, 468–470.

(22) Suryanarayana, D.; Weil, J. A. On the hyperfine splitting of the hydrogen atom in a spherical box. *The Journal of Chemical Physics* **1976**, *64*, 510–513.

(23) Miao, M. S. Caesium in high oxidation states and as a p-block element. *Nature Chemistry* **2013**, *5*, 846–852.

(24) Selli, D.; Baburin, I. A.; Martoňák, R.; Leoni, S. Novel metastable metallic and semiconducting germaniums. *Scientific Reports* **2013**, *3*, 1466.

(25) Huan, T. D.; Amsler, M.; Marques, M. A.; Botti, S.; Willand, A.; Goedecker, S. Low-energy polymeric phases of alanates. *Physical Review Letters* **2013**, *110*, 135502.

(26) Wang, H.; Tse, J. S.; Tanaka, K.; Iitaka, T.; Ma, Y. Superconductive sodalite-like clathrate calcium hydride at high pressures. *Proceedings of the National Academy of Sciences of the United States of America* **2012**, *109*, 6463–6466.

(27) Amsler, M.; Flores-Livas, J. A.; Lehtovaara, L.; Balima, F.; Ghasemi, S. A.; MacHon, D.; Pailhès, S.; Willand, A.; Caliste, D.; Botti, S.; San Miguel, A.; Goedecker, S.; Marques, M. A. Crystal structure of cold compressed graphite. *Physical Review Letters* **2012**, *108*, 065501.

(28) Flores-Livas, J. A.; Amsler, M.; Lenosky, T. J.; Lehtovaara, L.; Botti, S.; Marques, M. A.; Goedecker, S. High-pressure structures of disilane and their superconducting properties. *Physical Review Letters* **2012**, *108*, 117004.

(29) Yao, Y.; Tse, J. S.; Klug, D. D. Structures of insulating phases of dense lithium. *Physical Review Letters* **2009**, *102*, 1155903.

(30) Stauch, T. A mechanochemical model for the simulation of molecules and molecular crystals under hydrostatic pressure. *Journal of Chemical Physics* **2020**, *153*, 134503.

(31) Subramanian, G.; Mathew, N.; Leiding, J. A generalized force-modified potential energy surface for mechanochemical simulations. *Journal of Chemical Physics* **2015**, *143*, 134109.

(32) Cammi, R.; Verdolino, V.; Mennucci, B.; Tomasi, J. Towards the elaboration of a QM method to describe molecular solutes under the effect of a very high pressure. *Chemical Physics* **2008**, *344*, 135–141.

(33) Cammi, R. A new extension of the polarizable continuum model: Toward a quantum chemical description of chemical reactions at extreme high pressure. *Journal of Computational Chemistry* **2015**, *36*, 2246–2259.

(34) Chen, B.; Hoffmann, R.; Cammi, R. The Effect of Pressure on Organic Reactions in Fluids—a New Theoretical Perspective. *Angewandte Chemie - International Edition* **2017**, *56*, 11126–11142.

(35) Scheurer, M.; Dreuw, A.; Epifanovsky, E.; Head-Gordon, M.; Stauch, T. Modeling Molecules under Pressure with Gaussian Potentials. *Journal of Chemical Theory and Computation* **2021**, *17*, 583–597.

(36) Häser, M.; Ahlrichs, R. Improvements on the direct SCF method. *Journal of Computational Chemistry* **1989**, *10*, 104–111.

(37) Whitten, J. L. Coulombic potential energy integrals and approximations. *The Journal of Chemical Physics* **1973**, *4496*, 4496–4501.

(38) Lambrecht, D. S.; Doser, B.; Ochsenfeld, C. Rigorous integral screening for electron correlation methods. *Journal of Chemical Physics* **2005**, *123*, 184102.

(39) Doser, B.; Lambrecht, D. S.; Kussmann, J.; Ochsenfeld, C. Linear-scaling atomic orbital-based second-order Møller-Plesset perturbation theory by rigorous integral screening criteria. *Journal of Chemical Physics* **2009**, *130*, 064107.

(40) Maurer, S. A.; Lambrecht, D. S.; Flaig, D.; Ochsenfeld, C. Distance-dependent Schwarz-based integral estimates for two-electron integrals: Reliable tightness vs. Rigorous upper bounds. *Journal of Chemical Physics* **2012**, *136*, 144107.

(41) Maurer, S. A.; Lambrecht, D. S.; Kussmann, J.; Ochsenfeld, C. Efficient distance-including integral screening in linear-scaling Møller-Plesset perturbation theory. *Journal of Chemical Physics* **2013**, *138*, 014101.

(42) Thompson, T. H.; Ochsenfeld, C. Integral partition bounds for fast and effective screening of general one-, two-, and many-electron integrals. *Journal of Chemical Physics* **2019**, *150*, 044101.

(43) Barca, G. M.; Loos, P. F. Three- and four-electron integrals involving Gaussian geminals: Fundamental integrals, upper bounds, and recurrence relations. *Journal of Chemical Physics* **2017**, *147*, 024103.

(44) Hollman, D. S.; Schaefer, H. F.; Valeev, E. F. A tight distance-dependent estimator for screening three-center Coulomb integrals over Gaussian basis functions. *Journal of Chemical Physics* **2015**, *142*, 154106.

(45) Irmler, A.; Pauly, F. Multipole-based distance-dependent screening of Coulomb integrals. *Journal of Chemical Physics* **2019**, *151*, 084111.

(46) Ye, H. Z.; Berkelbach, T. C. Tight distance-dependent estimators for screening two-center and three-center short-range Coulomb integrals over Gaussian basis functions. *Journal of Chemical Physics* **2021**, *155*, 124106.

(47) McKenzie, S. C.; Epifanovsky, E.; Barca, G. M.; Gilbert, A. T.; Gill, P. M. Efficient Method for Calculating Effective Core Potential Integrals. *Journal of Physical Chemistry A* **2018**, *122*, 3066–3075.

(48) Shaw, R. A.; Hill, J. G. Prescreening and efficiency in the evaluation of integrals over ab initio effective core potentials. *Journal of Chemical Physics* **2017**, *147*, 074108.

(49) Song, C.; Wang, L. P.; Sachse, T.; Preiß, J.; Presselt, M.; Martínez, T. J. Efficient implementation of effective core potential integrals and gradients on graphical processing units. *Journal of Chemical Physics* **2015**, *143*, 014114.

(50) Epifanovsky, E. et al. Software for the frontiers of quantum chemistry: An overview of developments in the Q-Chem 5 package. *Journal of Chemical Physics* **2021**, *155*, 084801.

(51) Obara, S.; Saika, A. Efficient recursive computation of molecular integrals over Cartesian Gaussian functions. *The Journal of Chemical Physics* **1985**, *84*, 3963–3974.

(52) Lange, A. W.; Herbert, J. M. Polarizable continuum reaction-field solvation models affording smooth potential energy surfaces. *Journal of Physical Chemistry Letters* **2010**, *1*, 556–561.

(53) Balabanov, N. B.; Peterson, K. A. Systematically convergent basis sets for transition metals. I. All-electron correlation consistent basis sets for the 3d elements Sc-Zn. *Journal of Chemical Physics* **2005**, *123*, 064107.

(54) Woon, D. E.; Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. III. The atoms aluminum through argon. *The Journal of Chemical Physics* **1993**, *98*, 1358–1371.

(55) Woon, D. E.; Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. V. Core-valence basis sets for boron through neon. *The Journal of Chemical Physics* **1995**, *103*, 4572–4585.

(56) Woon, D. E.; Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. IX. The atoms gallium through krypton. *The Journal of Chemical Physics* **1998**, 1358–1371.

(57) Hartree, D. R. The Wave Mechanics of an Atom with a Non-Coulomb Central Field Part I Theory and Methods. *Mathematical Proceedings of the Cambridge Philosophical Society* **1928**, *24*, 89–110.

(58) Hartree, D. R.; Hartree, W. Self-consistent field, with exchange, for beryllium. *Proceedings of the Royal Society of London. Series A - Mathematical and Physical Sciences* **1935**, *150*, 9–33.

(59) Pulay, P. Improved SCF convergence acceleration. *Journal of Computational Chemistry* **1982**, *3*, 556–560.

(60) Strout, D. L.; Scuseria, G. E. A quantitative study of the scaling properties of the Hartree–Fock method. *The Journal of Chemical Physics* **1995**, *102*, 8448–8452.