Efficient calculation of integrals in mixed ramp-Gaussian basis sets

Laura K. McKemmish

Citation: The Journal of Chemical Physics **142**, 134104 (2015); doi: 10.1063/1.4916314
View online: http://dx.doi.org/10.1063/1.4916314
View Table of Contents: http://aip.scitation.org/toc/jcp/142/13
Published by the American Institute of Physics

Efficient calculation of integrals in mixed ramp-Gaussian basis sets

Laura K. McKemmish[a)]
Department of Physics and Astronomy, University College London, London, United Kingdom and Research School of Chemistry, Australian National University, Canberra, Australia

(Received 11 February 2015; accepted 16 March 2015; published online 2 April 2015)

Algorithms for the efficient calculation of two-electron integrals in the newly developed mixed ramp-Gaussian basis sets are presented, alongside a Fortran90 implementation of these algorithms, RampItUp. These new basis sets have significant potential to (1) give some speed-up (estimated at up to 20% for large molecules in fully optimised code) to general-purpose Hartree-Fock (HF) and density functional theory quantum chemistry calculations, replacing all-Gaussian basis sets, and (2) give very large speed-ups for calculations of core-dependent properties, such as electron density at the nucleus, NMR parameters, relativistic corrections, and total energies, replacing the current use of Slater basis functions or very large specialised all-Gaussian basis sets for these purposes. This initial implementation already demonstrates roughly 10% speed-ups in HF/R-31G calculations compared to HF/6-31G calculations for large linear molecules, demonstrating the promise of this methodology, particularly for the second application. As well as the reduction in the total primitive number in R-31G compared to 6-31G, this timing advantage can be attributed to the significant reduction in the number of mathematically complex intermediate integrals after modelling each ramp-Gaussian basis-function-pair as a sum of ramps on a single atomic centre. © *2015 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4916314]

I. INTRODUCTION

Over the last couple of decades, quantum chemistry has become a mainstream tool for simulating chemical systems in a wide variety of sciences including organic chemistry, biology, condensed matter physics, astronomy, and more. Key to this widespread utilisation is the development of very fast methods of performing quantum chemistry calculations (particularly density functional theory (DFT)) using Gaussian basis functions. Specifically, the evaluation of two-electron integrals[1] is the bottleneck for most Hartree-Fock (HF) and DFT calculations; modern quantum chemistry was largely made possible by key developments in integral evaluation algorithms,[2–6] particularly recurrence relations,[7–9] in the 1980s and 1990s combined with implementation of these algorithms in widely available, near "black-box" codes.

A pure Gaussian of exponent α and angular momentum quantum numbers ℓ, m is

$$\mathcal{G}_{\alpha\ell m}(\mathbf{r}) = N^{\mathcal{G}}_{\alpha\ell m} \exp(-\alpha r^2)\, r^{\ell} Y_{\ell m}(\theta,\phi), \tag{1}$$

where $Y_{\ell m}$ are real spherical harmonics and $N^{\mathcal{G}}_{\alpha\ell m} = \sqrt{\frac{(\ell+1)!}{(2\ell+2)!}} \frac{(8\alpha)^{\ell+3/2}}{\sqrt{\pi}}$. A s-Gaussian centered at A will be denoted by s^A_α and a generic Gaussian by $\mathcal{G}$.

Despite their widespread utilisation, Gaussian basis functions have some key shortcomings. In particular, it was recently shown[10] that the inability of Gaussian basis functions to model the nuclear-electron cusp[11] leads to their sub-exponential convergence behaviour,[12–16] i.e., the error in the energy of a n-term all-Gaussian approximation to a hydrogenic $1s$ wavefunction scales as $\exp(\pi\sqrt{3n})$. This can be clearly illustrated by considering the common Pople basis set, 6-31G. Even though chemistry occurs in the valence region (and is described by the "3" and "1" basis functions), 6 Gaussian primitives are required to describe the core adequately.

This author and collaborators recently proposed[17] the use of a novel type of basis function, the ramp, to describe electron distribution in the core region of atoms. A ramp with degree n and radius 1 is given by

$$\mathcal{R}_{n\ell m}(\mathbf{r}) = N^{\mathcal{R}}_{n\ell m} \begin{cases} (1-r)^n\, r^{\ell} Y_{\ell m}(\theta,\phi) & : r \le 1 \\ 0 & : r > 1 \end{cases}, \tag{2}$$

where, for convenience, we define $N^{\mathcal{R}}_{n\ell m} = \sqrt{\frac{(2n+2\ell+3)!}{(2n)!(2\ell+2)!}}$. A ramp function has the normalization $\langle \mathcal{R}_{n\ell m} | \mathcal{R}_{n\ell m} \rangle = 1$. A S-ramp centered at A will be denoted by S^A_n and a generic ramp by $\mathcal{R}$. S_n ramps have a cusp (i.e., a discontinuous first derivative at $r = 0$), which allows them to capture the behaviour of molecular orbitals close to nuclei[11] better than Gaussian functions.

Specifically, we introduced the R-31G basis set as an alternative to the commonly used 6-31G basis set; the former is obtained by replacing the 6 Gaussian primitives by 1 ramp and 1 Gaussian primitive. The value of n in Eq. (2) was chosen to maximise the overlap between the "6" and "R" basis functions; in practice, this meant that $n = Z + 1$ for the first-row atoms. As the valence basis functions were unchanged, the chemistry (relative energies) produced by UHF/R-31G and UHF/6-31G calculations was identical for small molecules to within 1 kcal/mol. We provided reasons to suggest that calculations with the R-31G basis set might be faster than with the 6-31G basis set, but deferred detailed timing considerations to this manuscript.

a)Author to whom correspondence should be addressed. Electronic mail: laura.mckemmish@gmail.com

Ramps can efficiently describe the core region with very few basis functions. Mixed ramp-Gaussian basis sets are therefore suitable for calculations that have traditionally been very difficult to do with all-Gaussian basis sets, in particular,

- nuclear magnetic resonance properties, such as chemical shift and J-J coupling,
- relativistic corrections,
- all-electron basis sets for heavy atoms,
- convergence to the complete Hartree-Fock basis set limit or spectroscopic accuracy for various properties, such as total energies.

There are at least two major, and very different, types of applications in which mixed ramp-Gaussian basis sets can find utility:

1. Fast, everyday calculations where they will speed up Hartree-Fock and DFT calculations for large molecules, using small to moderate size basis sets.
2. Specialised or high accuracy quantum chemistry calculations as discussed above.

The first application requires the speed of two-electron integral calculations in a mixed ramp-Gaussian basis set to be faster than in comparable quality all-Gaussian basis sets. This is a strong, but (as we shall see) feasible demand on this new basis set class. The second has more modest speed requirements; mixed ramp-Gaussian basis sets will be useful as long as the calculation times are reasonable (say within a factor of 5 of a comparable sized all-Gaussian basis set calculation) because they deliver far superior results. The usefulness of non-Gaussian basis sets with improved cusp properties is illustrated most starkly by considering the current use[18] of Slater basis sets[19–21] for specific purposes despite the very long integral evaluation times,[22,23] as well as more generally in the Amsterdam Density Functional (ADF) program.[24] Thus, despite more than 80 yr of investigation,[25–28] research is still undertaken[29–41] to improve integral evaluation for Slater-type orbitals to make these calculations competitive with all-Gaussian calculations. Given this, mixed ramp-Gaussian basis sets arguably encapsulate the best of both worlds: characteristics similar to all-Slater basis sets with the potential to match or better all-Gaussian calculation speeds.

Some might argue that mixed ramp-Gaussian basis sets are not suitable for the first application because the time spent on integrals involving the core is negligible compared to the total cost of the calculation. In Sec. II, we address this issue by comparing the time for a 6-31G, 2-31G, and 1-31G calculation in a series of moderate-sized (18-55 heavy atoms) molecules. We see that the R-31G basis set could provide savings of up to 40% (if the speed of ramp-containing integrals is negligible), though savings of up to 20%-30% are a more reasonable estimate of the potential savings available in highly optimised code using this basis set. The impact of this improvement should be placed in context of the widespread use of density functional theory worldwide.

Similar mixed ramp-Gaussian basis sets were initially investigated by Bishop[42,43] and Steiner[44–51] in the 1960s-1980s. Unfortunately, the molecular calculations performed by Steiner (up to H_2S) are too small to indicate whether or not his methodology was feasible for systems with large numbers of atoms, and he published no further results in this area after 1987.

This paper describes methods for efficient two-electron integral evaluation in mixed ramp-Gaussian basis sets, improving significantly on the original methodologies employed by Steiner,[48,50] inspired partly by recent developments in the efficient evaluation of all-Gaussian integrals, such as density fitting.[52–58]

Section III details the necessary pre-processing at the shell-pair level for efficient two-electron integral calculation. We redefine the shell-pair cutoff criteria for all-Gaussian shell-pairs to allow an analogous criterion to be constructed for ramp-Gaussian shell-pairs. Of particular note is the decision to model all ramp-Gaussian basis-function-pairs (BFPs) as a sum of ramps in a density-fitting type approach. This leads to the replacement of the concept of a shell pair (used to group all-Gaussian BFPs, $\mathfrak{G}$) by the concept of a nuclear centered group (used to group ramp-containing BFPs, $\mathfrak{R}$). A nuclear centered group contains all $\mathfrak{R}$ with a ramp on a single atomic center. This is a much larger grouping that in shell-pairs and allows reuse of intermediate quantities to a far greater extent.

Section IV details the methodology used to evaluate intermediates for two-electron integrals. Shell pairs and ramp-containing nuclear-centered groups with themselves and each other. We distinguish between integrals where the component shell pairs have significant overlap (**short-range integrals**) and when they have negligible overlap (**long-range integrals**). All complex mathematical operations are done at this stage of the two-electron integral evaluation process.

Section V details the way in which this intermediate quantities are combined to give full two-electron integrals, using simple multiply-adds and memory look-ups. In particular, we show the differences in loop structures between different classes of integrals. We demonstrate clearly that the intermediates for $\langle\mathfrak{G}|r_{12}^{-1}|\mathfrak{G}\rangle$ have to be calculated once for every pair of Gaussian shell pairs, whereas the intermediates for $\langle\mathfrak{R}|r_{12}^{-1}|\mathfrak{G}\rangle$ only have to be calculated for every set of atom/Gaussian shell-pair and the intermediates for $\langle\mathfrak{R}|r_{12}^{-1}|\mathfrak{R}\rangle$ are only calculated once for each pair of atoms. This significantly improves the speed of the calculation of two-electron integrals involving ramp-containing BFPs and therefore increases the competitiveness of mixed ramp-Gaussian basis sets compared to all-Gaussian basis sets.

These methods have been implemented in the standalone Fortran90 program, RampItUp, to produce a fully functional Hartree-Fock package to evaluate HF energies for molecules containing first-row atoms, with S-ramps, s-Gaussians, and p-Gaussians. It can also be used to produce one- and two-electron integrals for other quantum chemistry routines, e.g., Q-CHEM's DFT and MP2 packages. This program is freely available online on the author's website, or by contacting the author via email.

In Sec. VI, we perform timing comparisons between the unrestricted Hartree-Fock matrix evaluation time for UHF/R-31G and UHF/6-31G in moderate sized molecules, and compare the timings for calculation of different integral types in UHF/R-31G calculations.

In the supplementary material,[59] we detail preliminary investigations into the gradient and second derivative integrals demonstrating that no singularities arise and that similar integral evaluation techniques as considered in the main paper are applicable for the derivative integrals.

II. MAXIMUM AND BENCHMARK POTENTIAL SAVINGS OF R-31G OVER 6-31G

In this section, we investigate timings in all-Gaussian basis sets with different numbers of core primitives to get an idea of the savings possible with a rampified basis set.

A key component of the speed of modern quantum chemistry codes is the screening of two-electron integrals, usually using the Schwarz inequality as an upper bound to the integral value. As a preliminary, proof-of-principle code, for simplicity, RampItUp does not implement this screening. We, thus, further investigate whether screening significantly influences this potential time saving of R-31G compared to 6-31G.

R-31G calculations cannot be faster than 1-31G calculations; therefore, the maximum possible time saving of R-31G over 6-31G is given by

$$\text{Maximum \% Saving} = \frac{\text{6-31G time} - \text{1-31G time}}{\text{6-31G time}}. \quad (3)$$

Table I shows that both with and without screening, the maximum possible saving is between 19% and 41%; screening therefore is expected to not have a significant influence on the time savings with a R-31G basis set over a 6-31G basis.

It is useful to define the 2-31G basis set as the R-31G basis with the ramp primitive replaced by a Gaussian primitive with exponent A that maximises $\langle S_n | s_A \rangle$. For simplicity, the coefficients of the two Gaussian primitives used in the basis set definition of 2-31G are inherited from the coefficients of the ramp and Gaussian in R-31G, with the automated renormalisation of basis functions in Q-CHEM used to ensure normalisation of the new "2" function. For carbon, the value $A = 21.751$ is used (where $n = 7$ for the R-31G basis set for carbon).

A good indication of the actual savings that could be achieved in an optimal code is to compare a 2-31G to 6-31G calculation time: the 2-31G calculation time will be equal to the R-31G calculation time if the integrals in a mixed ramp-Gaussian basis set are assumed to be exactly as fast as in an all-Gaussian basis set. Thus, the benchmark savings of R-31G compared to 6-31G are given by

$$\text{Benchmark \% Saving} = \frac{\text{6-31G time} - \text{2-31G time}}{\text{6-31G time}}. \quad (4)$$

Table I gives a benchmark of 19%-34% savings for optimal R-31G code over 6-31G code based solely on the reduction of the number of primitives. This seems slightly smaller when screening is applied than without screening, but this effect is less than the influence of different molecules. Actual savings will differ from this value depending on algorithmic and implementation details of the integral evaluation.

It is worth commenting on how these potential savings change for different basis sets. In larger basis sets (especially those with many polarisations or diffuse basis sets), less calculation time is spent on integrals that contain the core basis function; therefore, possible savings are reduced. However, for heavier elements, there are more core basis functions to rampify and savings should thus be greater.

Rampification of a different class of basis set (e.g., Pople,[60] Dunning,[61] Jensen,[62] and ANO[63]) is expected to yield very similar results to those described for 6-31G vs R-31G, that is, the chemical (relative) energetics will be very similar, and there is the potential for modest potential time savings. Note that all basis sets with general contraction should be converted to their segmented contracted version (e.g., following the Jensen procedure[64]) before rampification.

Current investigations of mixed ramp-Gaussian basis sets have focused on replacing all-Gaussian basis sets with new basis sets that aim to produce very similar chemistry. This has been done so that the effect of rampification can be isolated from other basis set design considerations, and permit more careful comparisons. However, future development of entirely re-optimised ramp-Gaussian basis sets is desirable, and could increase time savings.

TABLE I. Maximum possible and benchmark percentage time savings of R-31G over 6-31G with and without two-electron screening for a set of molecules. All timings in this table were performed with 1 cpu on a Fujitsu Primergy cluster, using production version Q-CHEM4.2 with the two minor modifications outlined in Sec. VI. Details of molecules are in Table III.

	% saving without screening		*% saving with screening*	
	Max	Benchmark	Max	Benchmark
Linear molecules				
Alkane-30	38	29	32	25
Alkane-55	38	24	19	19
Stearic acid	33	26	37	25
Montanic acid	34	22	36	26
Octatriacontanoic acid	33	27	44	31
Non-linear molecules				
Fullerene-20	37	34	39	33
Fullerene-28	41	27	30	24
Cholesterol	34	25	29	21

III. SHELL-PAIR PROCESSING

Our first paper[17] describes methods of calculating required one-electron integrals. As this is not a time critical part of the overall code, we will not discuss efficient calculation of these integrals; optimisation is usually not essential beyond ensuring memory requirements are kept sufficiently low.

At the shell-pair level, there are two tasks that are critical to ensuring fast two-electron integral evaluation:

- The decision to keep or discard each possible shell-pair must be made.
- Often, simplification and/or modelling of each significant shell-pair is performed.

We will look at both of these in turn in this section.

A. Density ramps

A density ramp with degree n and radius 1 is given by

$$\mathbb{R}_{n\ell m}(\mathbf{r}) = N^{\mathbb{R}}_{n\ell m} \begin{cases} (1-r)^n\, r^\ell Y_{\ell m}(\theta,\phi) & : r \le 1 \\ 0 & : r > 1 \end{cases},$$

where for convenience we define

$$N^{\mathbb{R}}_{n\ell m} = \frac{(n+2\ell+3)!}{n!(2\ell+2)!}. \tag{5}$$

The density ramp is of the same form as the orbital ramp as defined in Eq. (2) except that it has the "unit multipole" normalization, *i.e.*, $\langle \mathbb{R}_{n\ell m} | r^\ell Y_{\ell m}\rangle = 1$. An S-type density ramp centered at A will be denoted by $\mathbb{S}^A_n$ and a generic density ramp by $\mathbb{R}$.

B. Negligible and significant shell-pairs

The first task in shell-pair processing in preparation for two-electron integral evaluation is to reduce the number of basis-function-pairs (BFPs), by removing all that contribute negligibly to one- and two-electron integrals.

$\mathcal{R}\mathcal{R}$ BFPs are very easy to consider. Two non-concentric ramps have strictly zero overlap if the interatomic distance is greater than 2 a_0 (as is assumed in this paper for bonds not including H (which doesn't have ramps)). Note that even if the bond distance is less than 2 a_0, significant errors will not immediately arise by neglecting this kind of shell pair because the two basis functions are very tight and thus the overlap is low. For example, if we have a C-C bond ($n = 7$), the overlap between two normalised ramps only exceeds 10^{-10} at bond lengths shorter than 1.62 a_0.

In Q-CHEM, traditionally $\mathcal{G}\mathcal{G}$ shell-pairs are neglected based on the prefactor after the Gaussian product rule is applied, where the Gaussian product rule is given by[65] $s^{\mathbf{A}}_\alpha s^{\mathbf{B}}_\beta = G_{AB} s^{\mathbf{P}}_\zeta$, where $\zeta = \alpha+\beta$, $G_{AB} = \exp[-\alpha\beta\zeta^{-1}|\mathbf{A}-\mathbf{B}|^2]$ and $\mathbf{P} = (\alpha\mathbf{A}+\beta\mathbf{B})/\zeta$. The Q-CHEM shell-pair cut-off criteria takes two primitive shell-pairs $\mathcal{G}^{\mathbf{A}}_{\alpha_i}$ and $\mathcal{G}^{\mathbf{B}}_{\alpha_j}$ with contraction coefficients D_i and D_j and eliminates this $\mathcal{G}\mathcal{G}$ shell-pair if $\left| D_i\, D_j\, N^{\mathcal{G}}_{\alpha_i 00}\, N^{\mathcal{G}}_{\beta_j 00}\, e^{-\frac{\alpha_i\beta_j}{\zeta_{ij}}|\mathbf{A}-\mathbf{B}|^2} \right| \le 10^{-\text{thresh}}$, where thresh is set in the program input depending on the desired accuracy; 8-10 is usual for energy calculations. (All calculations in this paper use thresh = 10.)

However, it is not obvious how to translate this criteria to $\mathcal{R}\mathcal{G}$ shell-pairs which don't combine in the same way. In the interests of fair timing comparison, we have altered the criteria for shell-pair cut-offs in Q-CHEM to directly involve the overlap integral of the radial component of the $\mathcal{G}\mathcal{G}$ shell pairs, *i.e.,* the shell-pairs are eliminated if

$$\begin{aligned} &\left| D_i\, D_j \langle s^{\mathbf{A}}_{\alpha_i} | s^{\mathbf{B}}_{\beta_j}\rangle \right| \\ &= \left| D_i\, D_j\, N^{\mathcal{G}}_{\alpha_i 00}\, N^{\mathcal{G}}_{\beta_j 00} \left(\frac{\pi}{\zeta_{ij}}\right)^{3/2} e^{-\frac{\alpha_i\beta_j}{\zeta_{ij}}|\mathbf{A}-\mathbf{B}|^2} \right| \le 10^{-\text{thresh}}. \end{aligned} \tag{6}$$

With this new criteria for neglect of $\mathcal{G}\mathcal{G}$ integrals, we can form a completely analogous criteria for $\mathcal{R}\mathcal{G}$ BFPs where shell-pairs are eliminated if $\left| D_i\, D_j \langle S^{\mathbf{A}}_n | s^{\mathbf{B}}_{\beta_j}\rangle \right| \le 10^{-\text{thresh}}$.

C. Simplifying and modelling basis-function-pairs

One of the most critical tasks at the basis-function-pair level in terms of its influence of the overall time of the quantum chemistry calculations is to simplify the representation of each BFP.

Simplifying a $\mathcal{R}\mathcal{R}$ BFP is straightforward using the ramp product rule whereby two concentric ramps collapse into a single ramp; for S-ramps, $S_{n_1} S_{n_2} = (2\pi^{1/2})^{-1} S_{n_1+n_2}$.

The radial components of a ss BFP can be combined by using the Gaussian product rule.

Both the ramp product rule and the Gaussian product rule convert BFPs from a function with two centers and two parameters (ramp degrees and exponents) to a function with a single center and single parameter (with a trivial normalisation factor). The inability of two exponentials to simplify in a similar fashion has been the key reason for the difficulty of calculating Slater two-electron integrals.[18]

There is no similar rule for simplifying ramp-Gaussian BFPs. Instead, we choose to model ramp-Gaussian basis-function-pairs as a sum of density ramps on the ramp atomic center, i.e.,

$$\mathcal{R}^{\mathbf{A}}\mathcal{G}^{\mathbf{B}} \approx \sum_{k=1}^{K} c_{n_k\ell_k m_k} \mathbb{R}^{\mathbf{A}}_{n_k\ell_k m_k}, \tag{7}$$

where the coefficients, $c_{n_k\ell_k m_k}$, are found by least-squares fitting and the set of all $\mathbb{R}^{\mathbf{A}}_{n_k\ell_k m_k}$ is the model basis set.

There are two key decisions involved in the modelling process: the metric used in the least-squares fitting and the choice of the model basis set (the equivalent of the auxiliary basis set[66] in density-fitting[55,67–69] of Gaussian-Gaussian products). Empirically, we find that modelling concentric $\mathcal{R}\mathcal{G}$ shell-pairs with high Gaussian exponents is most challenging, necessitating careful selection of both fitting metric and model basis set; we have discussed this case in detail.[17] Modelling of non-concentric shell-pairs is more forgiving of fitting metric and model basis set, though more careful selection of these will enable shorter model lengths whilst retaining accuracy. Improvements to this procedure will improve the short-range timings, but have no influence on the timings for long-range integrals.

Note that after simplification, all ramp-containing BFPs, $\mathfrak{R}$, contain no details about the Gaussian primitives, ramp degrees, or contraction coefficients of the two individual basis functions.

It is also useful to consider a long-range representation of $\mathfrak{R}$ that stores only the multipole moments of the BFPs. This representation can be written as

$$\mathfrak{R} = \sum_{\ell m} M_{\ell m} \mho_{\ell m}, \tag{8}$$

where $\mho_{\ell m}$ is the unit ℓm multipole vector and $M_{\ell m}$ is the magnitude of ℓm multipole.

We choose the multipole moment operator as given by

$$\hat{M}_{\ell m}(\rho(\mathbf{r})) = \int \rho(\mathbf{r}) r^\ell Y_{\ell m} d\mathbf{r}. \tag{9}$$

Note that this differs from the common definition used by, e.g., Stone[70] and Hättig[71] by excluding a factor of $\sqrt{\frac{4\pi}{2\ell+1}}$. This is to

simplify our mathematical presentation. Note this definition means that a "unit" $\mathfrak{V}_{00}$ multipole moment has a charge of $2\sqrt{\pi}$.

By definition of the density ramps, the

$$\hat{M}_{LM}(\mathbb{R}_{n\ell m}) = \begin{cases} 1 & : L = \ell \text{ and } M = m \\ 0 & : \text{otherwise} \end{cases}. \quad (10)$$

We can build on the short-range representation of $\mathfrak{R}$ to find the ℓm multipole moment of $\mathfrak{R}$, which is given by

$$M_{\ell m} = \hat{M}_{\ell m}(\mathfrak{R}) = \sum_n c_{n\ell m}. \quad (11)$$

It is also possible to find the multipole moments of Ss and Sp directly, rather than through their models.

Though clearly different basis-function-pairs require different ℓ, m multipole components in their multipole expansion, it is computationally more efficient to keep the size of the multipole expansion fixed. In RampItUp, we use multipoles up to $\ell = 4$ to retain accuracy in the integrals and final energies.

D. The nuclear-centered group: Replacing the concept of a shell-pair for $\mathcal{RG}$ and $\mathcal{RR}$ basis-function pairs

Figure 1 shows how Gaussian-Gaussian basis-function pairs are grouped into shell-pairs. Each shell-pair contains only a small number of basis-function-pairs, depending on angular momentum of each shell. For a fixed molecule size, the number of shell-pairs in a molecule grows with the number of basis-functions squared.

Shell-pairs are important for efficient calculation of all-Gaussian two-electron integrals because integrals involving each basis-function-pair within a shell-pair can usually be found from a single set of intermediate quantities. Calculating these intermediate quantities for the whole shell-pair together, and only differentiating between basis-function-pairs within a shell-pair as late as possible makes integral evaluation much more efficient.

In contrast, Figure 2 shows that all ramp-Gaussian BFPs can be divided into nuclear-centered groups based on the atomic center of the ramp. Each nuclear-centered-group contains all BFPs (both $\mathcal{RG}$ and $\mathcal{RR}$) involving the ramp on the atomic center, a much larger number than the number of BFPs in each shell-pair. The total number of nuclear-centered groups grows with the number of heavy atoms, much slower than the total number of shell-pairs. Very dense basis sets (e.g., a large number of functions on each atom) will have the same number of nuclear-centered groups as minimal basis sets.

All of the time consuming steps of integral evaluation, such as evaluation of erf and exp functions and numerical integration, are done once for each nuclear-centered group for a small subset of parameters associated with each density ramp or unit multipole moment and stored as a look-up table. This is analogous to the calculation and storage intermediate quantities for integrals involving a particular shell-pair. For a particular integral involving an individual ramp basis-function-pair, it becomes a simple digestion process; a set of multiple adds using values from the small look-up table and the model coefficients, $c_{n\ell m}$, or multipoles, $M_{\ell m}$, from the BFP information. This is analogous to using the intermediates for shell-pairs to produce all the related two-electron integrals. In this manner, the need for the concept of ramp shell-pair is superseded by this much larger grouping of ramp BFPs with common integral intermediates. Since there are fewer nuclear-centered groupings than shell-pairs, the reuse of intermediate data is greater for two-electron integrals involving ramp-containing BFPs $\mathfrak{R}$ than those involving Gaussian-Gaussian BFPs $\mathfrak{G}$.

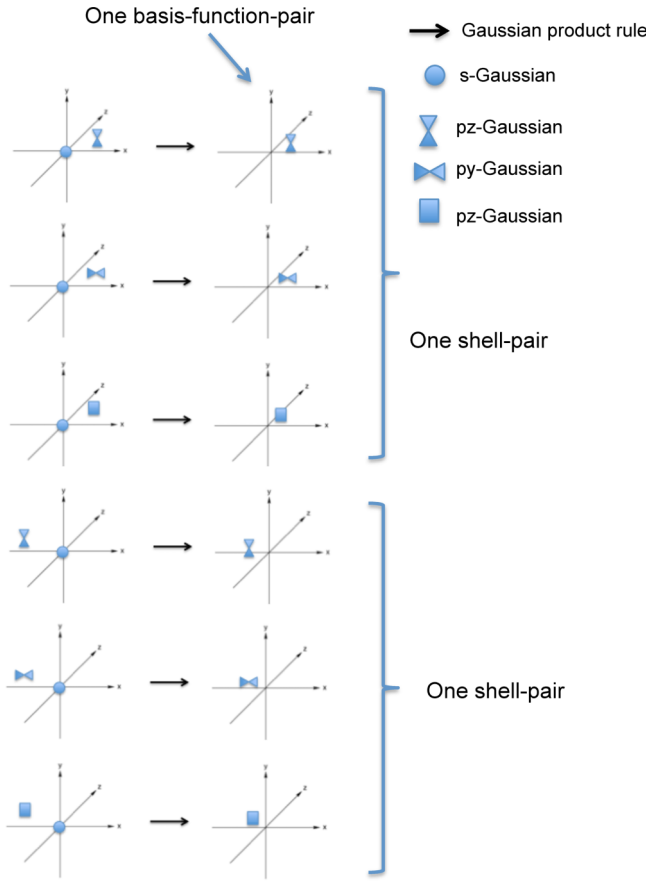

FIG. 1. Diagrammatic representation of a set of Gaussian-Gaussian basis-function-pairs into shell-pairs.

IV. INTERMEDIATE TWO-ELECTRON INTEGRAL QUANTITIES IN MIXED RAMP-GAUSSIAN BASIS SETS

Two-electron integrals are central to quantum chemistry calculations; their efficient evaluation is critical to fast Hartree-Fock, most DFT and many MP2 calculations. The form of this integral and the difficulty of evaluating it efficiently and accurately are determined by the basis set chosen for the calculation.

A mixed ramp-Gaussian basis set has Gaussian-Gaussian BFPs $\mathfrak{G}$ (that have been considered in depth over the last half century) and also introduces a new kind of basis-function-pair, $\mathfrak{R}$. There are two representations of this new ramp-containing basis-function-pair.

(a) Short-range representation—used when interacting with overlapping BFPs: $\mathfrak{R} = \sum_{n\ell m} c_{n\ell m} \mathbb{R}_{n\ell m}$.
(b) Long-range representation—used when interacting with non-overlapping (or negligibly overlapping) BFPs: $\mathfrak{R} = \sum_{\ell m} M_{\ell m} \mathfrak{V}_{\ell m}$.

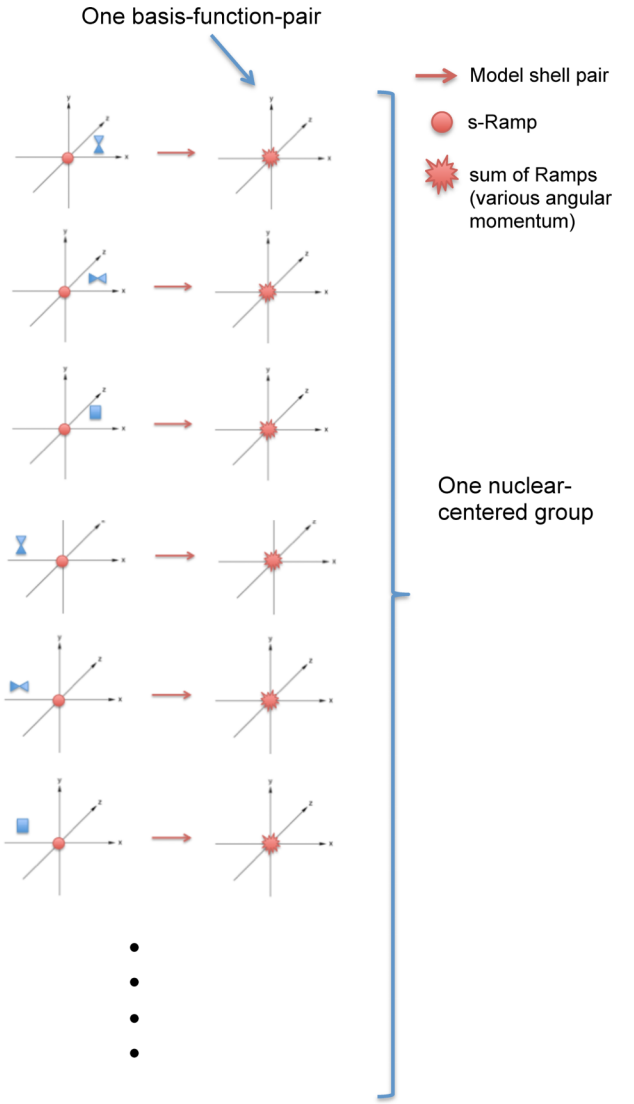

FIG. 2. Diagrammatic representation of a set of ramp-Gaussian basis-function-pairs into nuclear-centered groups.

For simplicity, in this section, we will consider the interactions of the individual components of the representations, i.e., $\mathbb{R}_{n\ell m}$ and $\mathfrak{U}_{\ell m}$, with each other and with Gaussian-Gaussian shell pairs. Section V details how to combine these intermediate quantities to get the full two-electron integrals. Note that the intermediate quantities contain all necessary complex mathematical operations, e.g., erf, exp, sqrt, divide, and trigonometric operators; combining these quantities requires only simple multiply-adds.

In this manuscript and the associated RampItUp code, we consider a single primitive component of Gaussian-Gaussian shell-pairs, i.e., $s^{\mathbf{Q}}$; future implementations should take advantage of contraction. However, it is not as important in the R-31G basis set as the 6-31G basis set because the average degree of contraction is much smaller.

A. Intermediate integrals for $\langle \mathfrak{G}|r_{12}^{-1}|\mathfrak{G}\rangle$

The evaluation of $\langle \mathfrak{G}|r_{12}^{-1}|\mathfrak{G}\rangle$ has been the subject of decades of research, for good reviews and detailed equations see Refs. 72 and 73. For our purposes, it is sufficient to note that a set of intermediate integrals have to be recalculated for every all-Gaussian shell-quartet. For large $\sqrt{(\zeta\eta(\zeta+\eta)^{-1})}R_{PQ}$ (i.e., negligible overlapping bra and ket), significant simplifications can be made, decreasing the calculation time per intermediate integral.

Later, in the loop structures discussed in Sec. V, these intermediate integrals are combined using vertical and horizontal recurrence relations in the PRISM algorithm[9] (or similar) to give the final two-electron all-gaussian integral.

B. Intermediate integrals for $\langle \mathfrak{R}|r_{12}^{-1}|\mathfrak{R}\rangle$

1. Short-range: Concentric bra and ket, $\langle \mathbb{R}^{\mathbf{A}}|r_{12}^{-1}|\mathbb{R}^{\mathbf{A}}\rangle$

For concentric ramps, only density ramps of the same angular momentum have a non-zero Coulomb interaction. In this case,

$$\langle \mathbb{R}^{\mathbf{A}}_{n_1\ell m}|\mathbb{R}^{\mathbf{A}}_{n_2\ell m}\rangle = N^{\mathbb{R}}_{n_1\ell m}N^{\mathbb{R}}_{n_2\ell m}\frac{4\pi}{2\ell+1}\frac{\Gamma[3+2\ell]\Gamma[3+n_t]}{\Gamma[6+2\ell+n_t]} \times\left(\frac{2l}{(n_1+1)(n_2+1)} + \frac{n_1^2+n_2^2+9n_t+3n_1n_2+16}{(n_1+1)(n_1+2)(n_2+1)(n_2+2)}\right), \quad (12)$$

where $n_t = n_1 + n_2$.

These intermediate values are computed once at the beginning of all RampItUp jobs and stored as a look-up table.

2. Long-range: Non-concentric bra and ket, $\langle \mathfrak{U}^{\mathbf{A}}|r_{12}^{-1}|\mathfrak{U}^{\mathbf{C}}\rangle$

It is most efficient to calculate the interaction of two non-concentric ramp-containing BFPs by interacting their real pure multipole moments.

With the definition, we choose for multipole moments $\mathfrak{U}$, the interaction of two unit $(0,0)$-multipoles at centers $\mathbf{A}$ and $\mathbf{C}$ is given by $\langle \mathfrak{U}_{00}(\mathbf{A})|r_{12}^{-1}|\mathfrak{U}_{00}(\mathbf{C})\rangle = \frac{4\pi}{R_{AC}}$.

We refer the reader to Hättig[71] for efficient recurrence relations for the calculation of $\langle \mathfrak{U}_{\ell m}(\mathbf{A})|r_{12}^{-1}|\mathfrak{U}_{\lambda\mu}(\mathbf{C})\rangle$, which are obtained by simple multiplication from the elements of the real spherical multipole interaction tensor discussed in Hättig, i.e., $\langle \mathfrak{U}_{\ell m}(\mathbf{A})|r_{12}^{-1}|\mathfrak{U}_{\lambda\mu}(\mathbf{C})\rangle = \frac{4\pi}{\sqrt{(2\ell+1)(2\lambda+1)}}T^{\lambda\mu}_{\ell m}$.

These intermediate values are computed once for each pair of atoms; this makes it a relatively rare operation.

C. Intermediate integrals for $\langle \mathfrak{R}|r_{12}^{-1}|\mathfrak{G}\rangle$

1. Short-range: Overlapping bra and ket, $\langle \mathbb{R}^{\mathbf{A}}|r_{12}^{-1}|s^{\mathbf{Q}}\rangle$

Integration over all angular degrees of freedom and one radial degree is relatively straightforward if the angular momentum of the ramp shell pair is specified. The remaining one-dimensional integral is smooth and over a compact domain of 0 to 1. For a single model component, the integral has the form

$$\langle \mathbb{R}^{\mathbf{A}}_{n\ell m}|r_{12}^{-1}|s^{\mathbf{Q}}_{\zeta}\rangle = \frac{8\pi^{3/2}}{\sqrt{1+2\ell}} N^{\mathcal{G}}_{\zeta 00} N^{\mathbb{R}}_{n\ell m} Y_{\ell m}[\Theta,\Phi] \times \int_0^1 (1-r)^n \, h_\ell[r;\zeta,R]dr, \quad (13)$$

where R,Θ,Φ are the spherical coordinates of $\mathbf{R}_{AQ}$ and $H_\ell[r;\zeta,R]$ contains terms dependent on the angular momentum ℓ, for example,

$$H_0[r;\zeta,R] = \frac{r}{8R\zeta^2}\left(\pi^{1/2}\zeta^{1/2}\left((r-R)\,\mathrm{erf}\,\left[(-r+R)\sqrt{\zeta}\right] + (r+R)\,\mathrm{erf}\,\left[(r+R)\sqrt{\zeta}\right]\right) - 2e^{-\left(r^2+R^2\right)\zeta}\sinh[2rR\zeta]\right). \quad (14)$$

We give expressions for $\ell = 1,2,3$ in the supplementary material.[59]

We evaluate these expressions using Gauss-Legendre quadrature, i.e.,

$$\langle \mathbb{R}^{\mathbf{A}}_{n\ell m}|r_{12}^{-1}|s^{\mathbf{Q}}_{\zeta}\rangle = \frac{8\pi^{3/2}}{\sqrt{1+2\ell}} N^{\mathcal{G}}_{\zeta 00} N^{\mathbb{R}}_{n\ell m} Y_{\ell m}[\Theta,\Phi] \times \sum_i \left\{ w_i \left(-\frac{(1-r)^{n+1}(1+r+nr)}{(1+n)(2+n)} \right) H_\ell[r_i;\zeta,R] \right\}, \quad (15)$$

where w_i and r_i are the weights and abscissas of Gauss-Legendre quadrature. This quadrature is most difficult when R is small and ζ is large; however, for realistic molecules, $R > 1.5$, and for a R-31G basis set, $\zeta < 40$. Future ramp-Gaussian basis sets should retain similar bounds; removing the very large exponent Gaussian primitives is one of the primary reasons for ramifying basis sets.

Some may view numerical quadrature as undesirable. However, in this case, we have a small, finite domain on a reasonably smooth 1D integral; therefore, the number of quadrature points to ensure high accuracy is quite small (we conservatively use 41 in all cases). Furthermore, this is an intermediate integral done outside the inner-most loop of the two-electron integral evaluation; it only has to be done once for every atom paired with every Gaussian shell-pair. Last, it is an intermediate component of a short-range integral; for a sufficiently large molecule, long-range integrals will dominate the timings.

Note that we must take care with the limits of these formulae when $R = 0$ due to the presence of terms like $\mathrm{erf}\,[\zeta(R+r)]/R$.

2. Long-range: Negligibly overlapping bra and ket, $\langle \mathfrak{U}^{\mathbf{A}}|r_{12}^{-1}|s^{\mathbf{Q}}\rangle$

The simplest case of the interaction of a unit multipole $\mathfrak{U}_{00}$ at $\mathbf{A}$ with a normalised s Gaussian at $\mathbf{Q}$ is given by

$$\langle \mathfrak{U}^{\mathbf{A}}_{00}|r_{12}^{-1}|s^{\mathbf{Q}}_{\zeta}\rangle = 2\sqrt{\pi}\left(\frac{2\pi}{\zeta}\right)^{3/4}\frac{1}{R}. \quad (16)$$

The interaction of a unit multipole $\mathfrak{U}_{\ell m}$ at $\mathbf{A}$ with a normalized s Gaussian at $\mathbf{Q}$ is given by

$$\langle \mathfrak{U}_{\ell m}(\mathbf{A})|r_{12}^{-1}|s^{\mathbf{Q}}_{\zeta}\rangle = \langle \mathfrak{U}_{\ell m}(\mathbf{A})|r_{12}^{-1}|\mathfrak{U}_{00}(\mathbf{Q})\rangle \hat{M}(s_\zeta), \quad (17)$$

$$= \left(\frac{\pi}{2}\right)^{1/4}\zeta^{-3/4} \times \langle \mathfrak{U}_{\ell m}(\mathbf{A})|r_{12}^{-1}|\mathfrak{U}_{00}(\mathbf{Q})\rangle. \quad (18)$$

This has to be calculated for once for each atom paired with each Gaussian shell-pair.

3. Higher Gaussian angular momentum integrals

The integrals involving sp and pp kets were found by Boys differentiation,[74] e.g., $px^{\mathbf{A}}_\alpha = \frac{1}{2\alpha}\frac{d}{dA_x}s^{\mathbf{A}}_\alpha$. These formula utilised the $\langle \Re|r_{12}^{-1}|s\rangle$ integrals and derivative of these integrals with respect to R, similar to the form used in all-Gaussian integrals.

General recurrence relations for arbitrarily large Gaussian angular momentum integrals have not yet been found, but should be of similar form to those for all-Gaussian integrals.[1,7–9] Note that recurrence relations are not used to relate ramps of differing angular momentum because Boys differentiation cannot be used to transform ramps of one angular momentum to a different higher angular momentum.

D. Significant or negligible shell-pair overlap

For each class of integrals, we have specified two different methodologies for calculating the fundamental interaction integral and noted that the choice between the two methods depends on whether the bra and ket overlap significantly or negligibly; in the latter case, we can simply interact the multipole moments of the bra and ket to find the integral much quicker than in the former case.

Just like in the shell-pair situation, we have to define what we mean by "significant" and "negligible" overlap; a definition that is too stringent (i.e., give more overlapping shell-pairs) will cause the program to run much slower, whereas too loose a definition (too few overlapping shell-pairs) will give inaccurate results.

For $\langle \Re|r_{12}^{-1}|\Re\rangle$, this decision is easy because non-concentric bra and kets have zero overlap due to the ramp's compact support.

For $\langle \mathfrak{G}|r_{12}^{-1}|\mathfrak{G}\rangle$ integrals in Q-CHEM, the parameter, $T = \frac{\zeta\eta}{\zeta+\eta}R^2_{PQ}$ is used to determined the cutoff, with $T_{\mathrm{crit}} = (\mathrm{thresh} + 2)\ln(10)$, where again thresh is set in the program input; 8-10 is usual. $T > T_{\mathrm{crit}}$ indicates negligible overlap while $T \leq T_{\mathrm{crit}}$ indicates significant overlap. This is justified by the presence of terms like $\mathrm{erf}(\sqrt{T})$ and $\exp(-T)$ which reach limiting values of 1 and 0, respectively, for large T.

There are similar limits we can take in evaluating $\langle \Re|r_{12}^{-1}|\mathfrak{G}\rangle$, for example, $\mathcal{T} = \zeta R^2_{PQ}$ where $R^2_{PQ} \gg 1$. Like in the all-Gaussian integral case, we compare $\mathcal{T}$ with T_{crit} to determine whether to use short-range or long-range algorithm to evaluate $\langle \Re|r_{12}^{-1}|\mathfrak{G}\rangle$.

E. Summary

Now that we have a methodology for calculating each of these intermediate quantities, it is informative to compare and

TABLE II. Number of intermediate two-electron integrals and related quantities for generic molecules and for benzene, $M = 6, N_g = 366, K = 25, \kappa = 235, c \approx 1/3$.

Total number of	$\langle \mathbb{R}^{\mathbf{A}} \vert r_{12}^{-1} \vert \mathbb{R}^{\mathbf{A}} \rangle$	$\langle \mathcal{U}^{\mathbf{A}} \vert r_{12}^{-1} \vert \mathcal{U}^{\mathbf{C}} \rangle$	$\langle \mathbb{R}^{\mathbf{A}} \vert r_{12}^{-1} \vert s^{\mathbf{Q}} \rangle$	$\langle \mathcal{U}^{\mathbf{A}} \vert r_{12}^{-1} \vert s^{\mathbf{Q}} \rangle$	$\langle s^{\mathbf{P}} \vert r_{12}^{-1} \vert s^{\mathbf{Q}} \rangle$
Intermediates with 1 nuclear-centered group and 1 ss	$\approx K(\kappa/K)^2$	K^2	κ	K	1
Brakets with 1 nuclear-centered group and 1 ss	1[a]	M^2	cMN_g	$(1-c)MN_g$	N_g^2
All intermediates	$\approx K(\kappa/K)^2$	M^2K^2	$c\kappa MN_g$	$(1-c)KMN_g$	N_g^2
Intermediates in HF/R-31 + G for benzene	2025	22 500	172 020	36 600	133 956
Intermediates in HF/6-31 + G for benzene	0	0	0	0	54 819 216
Expensive operations	Γ, /	trig, $1/R$	erf, exp, trig, $1/R$	trig, $1/R$	erf, $1/R$

[a]Precomputed once at the beginning of the calculation and stored.

contrast these algorithms in order to answer questions such as how many of these intermediate quantities there are and how expensive they are to calculate.

To assist in this analysis, the following definitions are useful:

N_g = number of $\mathcal{GG}$ shell-pair primitives
M = number of non-hydrogen atoms
κ = the average total number of unique model components across all nuclear-centered groups of $\mathfrak{R}$
K = number multipole moment components calculated for each $\mathfrak{R}$ (25 in RampItUp)
c = fraction of overlapping shell-pairs vs. non-overlapping shell pairs in a large molecule.

In general, $\kappa \gg K$ and $N_g \gg M$. The number of nuclear-centered groups will also be M. c decreases for larger molecules.

Table II demonstrates that

- $\langle \mathbb{R}^{\mathbf{A}} \vert r_{12}^{-1} \vert \mathbb{R}^{\mathbf{A}} \rangle$ is both cheap and not computed often.
- $\langle \mathcal{U}^{\mathbf{A}} \vert r_{12}^{-1} \vert \mathcal{U}^{\mathbf{B}} \rangle$ has moderate cost but is not computed often.
- $\langle \mathbb{R}^{\mathbf{A}} \vert r_{12}^{-1} \vert s^{\mathbf{Q}} \rangle$ is relatively expensive and computed often. However, the prefactor c means that the number of these interactions will scale much more gently with the overall size of the system than the other kinds of intermediate quantities
- $\langle \mathcal{U}^{\mathbf{A}} \vert r_{12}^{-1} \vert s^{\mathbf{Q}} \rangle$ has moderate cost but is computed often, more so in large systems.
- $\langle s^{\mathbf{P}} \vert r_{12}^{-1} \vert s^{\mathbf{Q}} \rangle$ has moderate cost and is computed most often by far because the number of $\mathcal{GG}$ shell-pairs will generally be much greater than the number of atoms.
- There are far fewer intermediate quantities in a HF/R-31 + G basis set (less than 0.5×10^6 in benzene) compared to in a HF/6-31 + G basis set (almost 55×10^6 in benzene!)

V. EFFICIENT LOOP STRUCTURE FOR TWO-ELECTRON INTEGRALS IN MIXED RAMP-GAUSSIAN BASIS SETS

To note that the time consuming operations (involved in calculating the intermediate integrals) are done at different stages in the loop depending on the type of two-electron integral; at the pair of atoms level for $\langle \mathfrak{R} \vert r_{12}^{-1} \vert \mathfrak{R} \rangle$, at the atom/Gaussian-shell-pair level for $\langle \mathfrak{R} \vert r_{12}^{-1} \vert \mathfrak{G} \rangle$ and at the shell-quartet level for $\langle \mathfrak{G} \vert r_{12}^{-1} \vert \mathfrak{G} \rangle$. This has a significant influence on the cost of the two-electron integral computation. In the case of $\langle \mathfrak{R} \vert r_{12}^{-1} \vert \mathfrak{R} \rangle$ and $\langle \mathfrak{R} \vert r_{12}^{-1} \vert \mathfrak{G} \rangle$, some of the inner-most loops can be recast as matrix-matrix or matrix-vector operations that can be efficiently implemented as BLAS Level 3 and Level 2 subroutine, often significantly increasing the speed of the calculation.

A. Calculation of $\langle \mathfrak{G} \vert r_{12}^{-1} \vert \mathfrak{G} \rangle$

Algorithm I provides the loop structure for calculating all-Gaussian integrals. Note that the PRISM algorithm described in Ref. 1 is currently one of the most efficient and general ways of combining intermediate quantities to compute a full class of all-Gaussian shell-quartets. This is the algorithm used in Q-CHEM.

Note, in particular, that the integral $\langle \mathfrak{G} \vert r_{12}^{-1} \vert \mathfrak{G} \rangle$ depends on the particulars of each $\mathfrak{G}$ shell-pair and the time consuming elementary operations, such as erf, must be performed for each shell-quartet separately.

B. Calculation of $\langle \mathfrak{R} \vert r_{12}^{-1} \vert \mathfrak{R} \rangle$

1. Short-range: Concentric basis-function-pairs

The integral $\langle \mathfrak{R}^{\mathbf{A}} \vert r_{12}^{-1} \vert \mathfrak{R}^{\mathbf{A}} \rangle$ is given by a sum of the intermediate integrals

$$\langle \mathfrak{R}_1^{\mathbf{A}} \vert r_{12}^{-1} \vert \mathfrak{R}_2^{\mathbf{A}} \rangle = \sum_{\ell m} \left(\sum_{n_1} c_{n_1 \ell m} \sum_{n_2} c_{n_2 \ell m} \langle \mathbb{R}_{n_1 \ell m} \vert r_{12}^{-1} \vert \mathbb{R}_{n_2 \ell m} \rangle \right). \quad (19)$$

The intermediate integrals $\langle \mathbb{R}_{n_1 \ell m} \vert r_{12}^{-1} \vert \mathbb{R}_{n_2 \ell m} \rangle$ are stored in a look-up table.

The loop structure for evaluating $\langle \mathfrak{R}^{\mathbf{A}} \vert r_{12}^{-1} \vert \mathfrak{R}^{\mathbf{A}} \rangle$ is given in Algorithm II. Note that in the inner-most loops, only simple add-multiplies are required; all the more, expensive elementary

ALGORITHM I. Loop structure for $\langle \mathfrak{G} \vert r_{12}^{-1} \vert \mathfrak{G} \rangle$ integrals.

for each Gaussian shell pair **do**
  for each different Gaussian shell pair **do**
    Calculate $[0]^{\{m\}}$, $\langle \mathfrak{G} \vert r_{12}^{-1} \vert \mathfrak{G} \rangle$ using PRISM algorithm.
  end for
end for

ALGORITHM II. Loop structure for $\langle \mathfrak{R}^{\mathbf{A}} | r_{12}^{-1} | \mathfrak{R}^{\mathbf{A}} \rangle$.

Create look-up table of $\langle \mathbb{R}_{n_1\ell m} | r_{12}^{-1} | \mathbb{R}_{n_2\ell m} \rangle$ value for all values of n_1, n_2, ℓ in model basis set.
for atom **A** **do**
  for each $\mathfrak{R}_1^{\mathbf{A}}$ **do**
    for each different $\mathfrak{R}_2^{\mathbf{A}}$ **do**
      Evaluate $\langle \mathfrak{R}_1^{\mathbf{A}} | r_{12}^{-1} | \mathfrak{R}_2^{\mathbf{A}} \rangle$
    end for
  end for
end for

maths operations are in the evaluation of $\langle \mathbb{R}_{n_1\ell m} | r_{12}^{-1} | \mathbb{R}_{n_2\ell m} \rangle$ and not done within the loop structure.

This $\langle \mathfrak{R}^{\mathbf{A}} | r_{12}^{-1} | \mathfrak{R}^{\mathbf{A}} \rangle$ integral is rare; any reasonable implementation of Eq. (19) is adequate.

2. Long-range: Non-concentric ramps in the bra and ket

The integral $\langle \mathfrak{R}^{\mathbf{A}} | r_{12}^{-1} | \mathfrak{R}^{\mathbf{C}} \rangle$ is given by a sum of the intermediate integrals

$$\langle \mathfrak{R}^{\mathbf{A}} | r_{12}^{-1} | \mathfrak{R}^{\mathbf{C}} \rangle = \sum_{\ell_1 m_1} M^{\mathfrak{R}^{\mathbf{A}}}_{\ell_1 m_1} \left(\sum_{\ell_2 m_2} M^{\mathfrak{R}^{\mathbf{C}}}_{\ell_2 m_2} \langle \mathfrak{U}^{\mathbf{A}}_{\ell_1 m_1} | r_{12}^{-1} | \mathfrak{U}^{\mathbf{C}}_{\ell_2 m_2} \rangle \right), \quad (20)$$

$$= \sum_{\ell_1 m_1} M^{\mathfrak{R}^{\mathbf{A}}}_{\ell_1 m_1} \langle \mathfrak{U}^{\mathbf{A}}_{\ell_1 m_1} | r_{12}^{-1} | \mathfrak{R}^{\mathbf{C}} \rangle. \quad (21)$$

The loop structure used to evaluate this class of integrals is given by Algorithm III. Note that the expensive steps (evaluating $\langle \mathfrak{U}^{R}_{\ell_1 m_1} | r_{12}^{-1} | \mathfrak{U}^{R}_{\ell_2 m_2} \rangle$, including spherical harmonic evaluation, and divide) are done at the pair-of-atoms level; the inner most loops are simple multiply-adds.

If we retain a fixed number of ℓm multipoles for all $\mathfrak{R}$ in the same nuclear-centered group, the two inner-most loops of Algorithm III can be coded as two BLAS matrix-matrix multiplications.

C. Calculation of $\langle \mathfrak{R} | r_{12}^{-1} | \mathfrak{G} \rangle$

1. Short-range

The two-electron integral involving a contracted ramp-containing basis-function-pair with many different ramp

ALGORITHM III. Loop structure for $\langle \mathfrak{R}^{\mathbf{A}} | r_{12}^{-1} | \mathfrak{R}^{\mathbf{C}} \rangle$.

Create look-up table of intermediate integral value for all possible values of n_1, n_2, ℓ.
for atom **C** **do**
  for atom **A** **do**
    Evaluate $\langle \mathfrak{U}^{\mathbf{A}}_{\ell_1 m_1} | r_{12}^{-1} | \mathfrak{U}^{\mathbf{C}}_{\ell_2 m_2} \rangle$
    for each $\mathfrak{R}^{\mathbf{C}}$ **do**
      Evaluate $\langle \mathfrak{U}^{\mathbf{A}}_{\ell_1 m_1} | r_{12}^{-1} | \mathfrak{R}^{\mathbf{C}} \rangle$
      for each $\mathfrak{R}^{\mathbf{A}}$ **do**
        Evaluate $\langle \mathfrak{R}^{\mathbf{A}} | r_{12}^{-1} | \mathfrak{R}^{\mathbf{C}} \rangle$
      end for
    end for
  end for
end for

ALGORITHM IV. Loop structure for $\langle \mathfrak{R} | r_{12}^{-1} | s \rangle$ integrals with significant basis-function-pair overlap.

for each primitive ss shell-pair **do**
  for each atom **A** **do**
    Evaluate $\langle \mathbb{R}_{n\ell m} | r_{12}^{-1} | s^{\mathbf{Q}}_{\zeta} \rangle$
    for each $\mathfrak{R}$ **do**
      Evaluate $\langle \mathfrak{R}^{\mathbf{A}} | r_{12}^{-1} | s^{\mathbf{Q}}_{\zeta} \rangle$
    end for
  end for
end for

model components and a primitive Gaussian-Gaussian shell-pair is given by

$$\langle \mathfrak{R}^{\mathbf{A}} | r_{12}^{-1} | s^{\mathbf{Q}}_{\zeta} \rangle = \sum_{n\ell m} c_{n\ell m} \langle \mathbb{R}_{n\ell m} | r_{12}^{-1} | s^{\mathbf{Q}}_{\zeta} \rangle. \quad (22)$$

For effective evaluation, our program uses the loop structure shown in Algorithm IV. Again, integrals involving higher angular momentum Gaussians are evaluated using a similar loop structure. Note that the more expensive mathematical operations (evaluating $\langle \mathbb{R}_{\ell m} | r_{12}^{-1} | s^{\mathbf{Q}}_{\zeta} \rangle$ with erf, exp, quadrature, etc.) are done outside the inner-most loop, at the atom-Gaussian shell pair level. The innermost loop is simple multiply-adds.

It is also possible to code the inner-most loop of Algorithm IV as a BLAS matrix-vector multiplication. However, this may not be efficient if the total number of model basis functions used in a single nuclear-centered group is much larger than the number used in one particular short-range $\mathfrak{R}$ representation, i.e., one replaces a disjoint sum of $n\ell m$ by the full set of $n\ell m$ in the summation in Eq. (22).

2. Long-range

The two-electron integral involving a contracted ramp-containing basis-function-pair and a primitive Gaussian shell-pair is given by

$$\langle \mathfrak{R}^{\mathbf{A}} | r_{12}^{-1} | s^{\mathbf{Q}}_{\zeta} \rangle = \sum_{\ell m} M^{\mathfrak{R}^{\mathbf{A}}}_{\ell m} \langle \mathfrak{U}^{\mathbf{A}}_{\ell m} | r_{12}^{-1} | s^{\mathbf{Q}}_{\zeta} \rangle. \quad (23)$$

The loop structure to evaluate all two-electron integrals of this class is given in Algorithm V. Note that integrals involving higher angular momentum Gaussians are evaluated using a similar loop structure. Again, we see that the expensive operations are in evaluating $\langle \mathfrak{U}^{\mathbf{A}}_{\ell m} | r_{12}^{-1} | s^{\mathbf{Q}}_{\zeta} \rangle$ (which involves

ALGORITHM V. Loop structure for $\langle \mathfrak{R} | r_{12}^{-1} | s \rangle$ integrals with negligible basis-function-pair overlap.

for each primitive Gaussian shell pair **do**
  for each atom **A** **do**
    Evaluate $\langle \mathfrak{U}^{\mathbf{A}}_{\ell m} | r_{12}^{-1} | s^{\mathbf{Q}}_{\zeta} \rangle$
    for each $\mathfrak{R}^{\mathbf{A}}$ **do**
      Evaluate $\langle \mathfrak{R}^{\mathbf{A}} | r_{12}^{-1} | s^{\mathbf{Q}}_{\zeta} \rangle$
    end for
  end for
end for

spherical harmonics, divide, etc.) and the innermost loop is simple multiply-adds.

If we retain a fixed number of ℓm multipoles for all $\mathfrak{R}$ in the same nuclear-centered group, the inner-most loop of Algorithm V can be coded as a BLAS matrix-vector multiplication.

VI. TIMINGS

A. Method

We have written a Fortran90 program, RampItUp, to calculate all ramp-containing integrals using the algorithms discussed in this paper. To calculate all-Gaussian integrals, we use Q-CHEM.

All-Gaussian integrals' programs generally use direct integral evaluation to minimise memory requirements for large molecules; thus, we compare the time required for one Fock build using 6-31G vs R-31G basis sets for a range of medium sized molecules.

It is important that we do as much as possible to ensure a fair comparison; nevertheless, a completely fair comparison is impossible and so these numbers should be viewed only as indicative of potential time savings offered by mixed ramp-Gaussian basis sets over all-Slater basis sets and over conventional all-Gaussian basis sets.

1. Details of RampItUp

RampItUp calculates all necessary two-electron integrals for molecules with a basis set including S-ramps, s-Gaussians, and p-Gaussians. It builds the full Fock matrix for the mixed basis-set and then performs a UHF calculation. Basis functions that have mixed S-ramp and s-Gaussian content, such as the 1s basis function of the R-31G basis set, are correctly dealt with.

The program does screening of all BFPs, but no screening of two-electron integrals, unlike Q-CHEM. It is anticipated that screening of two-electron integrals in ramp-Gaussian integrals can be done via similar methods (such as the Schwarz inequality) as already used for all-Gaussian integrals.

RampItUp is a tool for preliminary exploration of the feasibility of fast two-electron ramp-containing integrals. Major effort was spent optimising the most common and time critical classes of integrals, particularly the $\langle \mathfrak{R}|r_{12}^{-1}|\mathfrak{G}\rangle$

TABLE III. Comparative timings (in s) for a set of molecules for HF/6-31G v.s. HF/R-31G, rounded to the nearest second or 2 significant figures. The percentage speedups are approximate only, especially for smaller molecules. These timings are done on a laptop with gfortran complied version of RampItUp and Q-CHEM. All integral timings are without screening. Where appropriate, approximate geometries are indicated by the letters: "L" for linear, "LB" for most linear with some branching, "P" for planar and "S" for spherical. Geometries for biomolecules and drugs are from ChemSpider[75] (with Taxol reoptimised with hydrogens added); ChemSpider IDs are listed in the table where appropriate. All other geometries were found with a simple force-field optimisation in IQmol.[76]

ID	Name	No C	No N,O,F	No H	Geometry	$\langle\mathfrak{R}\|\mathfrak{R}\rangle$	SR $\langle\mathfrak{R}\|\mathfrak{G}\rangle$	LR $\langle\mathfrak{R}\|\mathfrak{G}\rangle$	R-containing	1-31G	R-31G	6-31G	% saving
Linear alkane chains													
	alkane-30	30	0	62	L	0.32	2.7	4.3	7.4	23	30	33	10
	alkane-40	40	0	82	L	0.54	3.9	8.2	12	41	52	61	15
	alkane-45	45	0	92	L	0.72	4.3	11	15	53	68	77	13
	alkane-55	55	0	102	L	1.1	5.5	16	23	86	109	118	8
Saturated fatty acids													
5091	Stearic acid	18	2	36	L	0.13	1.5	1.5	3.1	7.3	10	11	9
10035	Arachidic acid	20	2	40	L	0.15	1.7	1.9	3.7	9.1	13	14	9
7923	Behenic acid	22	2	44	L	0.19	1.8	2.4	4.5	11	15	17	9
10724	Lignoceric acid	24	2	48	L	0.21	2.0	2.9	5.2	13	18	21	14
10037	Cerotic acid	26	2	52	L	0.26	2.2	3.6	5.9	16	22	24	8
10038	Montanic acid	28	2	56	L	0.29	2.6	4.1	6.8	18	25	28	11
10039	Melissic acid	30	2	60	L	0.33	2.7	4.8	7.6	21	29	32	9
18168	Lacceroic acid	32	2	64	L	0.38	3.0	5.6	8.7	24	33	37	11
85266	Geddic acid	34	2	68	L	0.43	3.0	6.4	9.8	27	37	42	12
4445723	Hexatriacontanoic acid	36	2	70	L	0.49	3.2	7.1	11	31	42	47	11
445725	Octatriacontanoic acid	38	2	74	L	0.68	3.5	9.1	13	35	48	52	8
Triglycerines													
62497	Triheptanoin	24	6	44	LB	0.48	4.3	5.0	9.6	20	30	29	−3
10393	Glycerol tricaprylate	27	6	50	LB	0.56	4.6	6.3	11	24	35	36	3
10394	Glycerin trilaurate	39	6	74	LB	0.93	6.1	12	19	47	66	68	3
10675	Trimyristin	45	6	86	LB	1.2	6.8	16	24	62	86	90	4
10674	Tripalmitin	51	7	98	LB	1.4	7.4	21	30	77	107	113	5
71374	Linolein	57	6	98	LB	1.7	7.9	24	33	83	116	126	10
4593733	Triolein	57	6	104	LB	7.3	27	67	35	94	129	139	8

classes. However, only moderate effort was spent optimising $\langle\mathfrak{R}|r_{12}^{-1}|\mathfrak{R}\rangle$ integrals and little attempt was made to optimise the calculation of one-electron integrals or the SCF diagonalisation and convergence procedure beyond what was essential for reasonable speed calculations. The calculation of all-Gaussian integrals in RampItUp is not heavily optimised; we thus consider the time for Q-CHEM to calculate these integrals in a 1-31G basis set calculation instead of the RampItUp time for these integrals.

2. Q-CHEM

Q-CHEM is a well-established quantum chemistry package which uses modern algorithms for efficient evaluation of all-Gaussian integrals.

There are some modifications to the code that needed to be performed, specifically

1. two-electron integral screening is turned off by directly modifying source code;
2. shell-pair cutoff has been modified to use ss overlap instead of the pair prefactor G_{AB}.

Suitable input parameters also ensured that the Q-CHEM calculation was unrestricted, always used direct scf (i.e., no storing of integrals), and always computed the full Fock matrix in every iteration.

B. Results for R-31G vs 6-31G

1. Breakdown of R-31G timings

The breakdown of the total time of the R-31G calculation into its component parts as shown in Tables III and IV is illuminating. In particular, it is confirmed that $\langle\mathfrak{R}|r_{12}^{-1}|\mathfrak{R}\rangle$ integrals are much quicker than the $\langle\mathfrak{R}|r_{12}^{-1}|\mathfrak{G}\rangle$ integrals.

The numbers also reveal that as the molecule gets larger, the calculation time for long-range integrals starts to dominate the short-range integrals. In linear molecules, this happens more quickly than for non-linear molecules. This occurs because there are comparatively fewer short-range integrals in linear molecules than non-linear molecules due to geometric considerations.

2. R-31G vs 6-31G timings with no screening

The initial data for timings with the preliminary proof-of-principle RampItUp code and Q-CHEM code, shown in Tables III and IV, demonstrate conclusively that R-31G calculations are competitive or faster than 6-31G calculations for large molecules. This is great news for this new basis set and encourages further research into optimisation of the implementation of mixed ramp-Gaussian basis sets.

We do not yet get the full savings that are possible in a R-31G basis set that were suggested by the BPS values in Sec. II. This indicates there may be room for improvement of R-31G timings with better algorithm design and/or implementation.

These initial results show a time saving for R-31G over 6-31G for some sufficient big molecules. The largest time savings are for the linear alkane chains of 40-55 carbons where there are savings of up to 15%.

There are many indications that R-31G has an advantage over 6-31G for long-range integrals but a disadvantage for short-range integrals. Specifically, R-31G gets more advantage over 6-31G with

- linear molecules, e.g., saturated fatty acids and alkane chains have more speed-up than the linear polycyclic aromatic hydrocarbons, which have more speed up than the other polycyclic aromatic hydrocarbons;
- more extended molecule, e.g., lipids give more advantage to R-31G than proteins.

This is reasonable; the short-range representation of $\mathfrak{R}$ is much longer than the long-range representation, whereas the two representations are the same length for $\mathfrak{G}$. This means the number of intermediates required for one short-range representation of $\mathfrak{R}$ is much larger than for one long-range representation of $\mathfrak{R}$. A reader might be concerned that three-dimensional structures, with timings shown in Table IV, actually take longer with R-31G than 6-31G. This occurs because, in 3D systems, the ratio of short to long-range integrals is greater than in 1D systems due to geometric proximity of atoms. As the size of the 3D system increases, however, the ratio of short to long-range integrals will decrease, and we expect that the savings already seen in the 1D systems will start to be observed in the 3D systems.

Based on these results, optimisation of the short-range $\langle\mathfrak{R}|r_{12}^{-1}|\mathfrak{G}\rangle$ integrals is a useful area of future development. We recommend two main avenues for improvements in this class of integrals.

- Reduction in the model length. For example, the current implementation of models for non-concentric $\mathscr{R}\mathscr{G}$ products in RampItUp actually uses the overlap metric rather than the anti-Coulomb metric for simplicity. Modification of this code to the anti-Coulomb metric as discussed in Ref. 17 should be a priority of future code development. Fortunately, the modelling subsection of the code is quite modular and improvements to the modelling process can be done without influencing other subsections of the code.
- Alternative implementation for the evaluation of the $\langle\mathbb{R}^{\mathbf{A}}_{n\ell m}|r_{12}^{-1}|s^{\mathbf{Q}}_{\zeta}\rangle$ integral and its higher Gaussian angular momentum counterparts, such as alternative more efficient quadrature approach, analytic formula, recurrence relations, and/or interpolation methods. For example, currently 41 quadrature points are used for all integrals to ensure accuracy; this could be made adaptive depending on ζ, R, and n.

Introduction of screening will reduce the number of long-range integrals without affecting the number of short-range integrals; this will probably mean that larger molecules are needed to show faster R-31G calculation than 6-31G calculations.

TABLE IV. Further timings for non-linear molecules. See caption for Table III.

ID	Name	No C	No N,O,F	No H	Geometry	$\langle\mathfrak{R}\|\mathfrak{R}\rangle$	SR $\langle\mathfrak{R}\|\mathfrak{G}\rangle$	LR $\langle\mathfrak{R}\|\mathfrak{G}\rangle$	R-containing	1-31G	R-31G	6-31G	% saving
Linear polycyclic aromatic hydrocarbons													
8347	Pentacene	22	0	14	P	0.25	2.7	2.3	5.2	6.7	12	11	−10
109666	Hexacene	26	0	16	P	0.40	3.5	3.9	7.3	9.8	17	16	−10
4574185	Heptacene	30	0	18	P	0.63	4.2	5.2	10	14	24	22	−11
4574182	Octacene	34	0	20	P	0.86	4.9	7.2	13	18	31	28	−11
5256923	Nonacene	38	0	22	P	0.91	5.7	9.3	16	23	39	36	−8
General polycyclic aromatic hydrocarbons													
8816	Triphenylene	18	0	12	P	0.18	2.4	1.5	4.1	4.9	9.0	7.7	−13
8761	Coronene	24	0	12	P	0.62	6.2	4.0	11	13	24	20	−20
60771	Ovalene	32	0	14	P	1.6	10	9.0	21	30	51	42	−21
Fullerenes													
	Fullerene-20	20	0	0	S	0.55	5.3	2.4	8.4	6.6	15	10	−44
	Fullerene-24	24	0	0	S	0.91	7.9	4.3	13	12	25	19	−35
	Fullerene-28	28	0	0	S	1.5	11	7.3	20	21	41	31	−32
	Fullerene-32	32	0	0	S	2.2	16	12	30	34	64	49	−31
	Fullerene-36	36	0	0	S	2.9	19	18	39	51	90	71	−28
	Fullerene-50	50	0	0	S	8.1	43	53	105	159	264	207	−28
	Fullerene-60	60	0	0	S	15	66	97	179	303	482	376	−28
Common drugs and biological molecules													
4447623	Hydrocodone	22	0	21		0.52	6.3	3.8	10	14	24	21	−14
5768	Sucrose	23	0	22		0.43	4.9	3.7	8.9	13	22	19	−16
39888	Paroxetime (Paxil)	24	0	20		0.51	4.2	3.5	8.2	11	19	17	−12
5775	Cholesterol	28	0	46		0.83	7.9	7.5	16	34	50	50	0
4514933	Lisinopril	29	0	31		0.61	5.7	5.7	12	21	33	31	−6
18508	Streptomycin	21	19	39		2.0	13	17	31	58	89	81	−10
54810	Atorvastatin	33	8	35		1.9	12	16	29	51	80	73	−10
4444685	Azadirachtin	35	16	44		5.6	28	42	75	133	209	184	−13
10368587	Taxol	47	15	51		7.2	30	57	94	176	270	245	−10

VII. CONCLUSIONS

In this manuscript, we have presented efficient methods of calculating the required two-electron integrals in a mixed ramp-Gaussian basis set. In particular, we introduce the concept of a nuclear-centered group to group all ramp-containing basis-function-pairs that contain the same intermediates, similar to, but much larger than, the shell-pair group of all-Gaussian basis-function-pairs. We present loop structures that highlight the major advantage of the nuclear-centered group; the computationally expensive mathematical operations to calculate the intermediate integrals have to be computed far less often for ramp-containing integrals than all-Gaussian integrals. Specifically,

- the intermediate integrals for $\langle\mathfrak{R}|r_{12}^{-1}|\mathfrak{R}\rangle$ have to be calculated once per pair of heavy atoms;
- the intermediate integrals for $\langle\mathfrak{R}|r_{12}^{-1}|\mathfrak{G}\rangle$ have to be calculated once per each set of atom/gaussian-shell pair;
- the intermediate integrals for $\langle\mathfrak{G}|r_{12}^{-1}|\mathfrak{G}\rangle$ have to be calculated once per gaussian shell-quartet.

Though we have only implemented integrals involving s and p-gaussians and S-ramps, the methodologies outlined in this paper will extend easily to higher angular momentum Gaussians and ramps. In particular, most of the difficulty involving new types of $\mathscr{R}\mathscr{G}$ shell-pairs can be reduced to modelling this new shell-pair as a sum of ramps; after this, all existing code infrastructure and algorithms for calculating two-electron integrals involving this shell-pair will work automatically.

We concluded this paper by comparing the time of UHF/R-31G and UHF/6-31G calculations of a variety of moderate-sized molecules, where considerable effort has been made to ensure fair timing comparisons. The comparative timings suggest that UHF/R-31G is a method that is faster than UHF/6-31G for long-range integrals but slower for short-range integrals; we see significant improvements of up to 15% for linear molecules (such as alkane chains and saturated fatty acids) of up to 50 carbon atoms, while calculations on more compact molecules (such as aromatic hydrocarbons, fullerenes, drugs and biological molecules) are slower. This behaviour suggests that for very large molecules (beyond

the reach of the current preliminary code primarily due to memory considerations), R-31G will always be faster than 6-31G, providing

1. a single package where calculation of both all-Gaussian integrals and ramp-containing integrals is optimised (currently, Q-CHEM and RampItUp separately fulfil these goals, but not in a single package with a single data structure, etc.);
2. standard speed-ups used for all-Gaussian integral calculation can also be implemented for ramp-containing integrals, most importantly screening of two-electron integrals via Schwarz or related inequalities but also other tasks such as taking advantage of symmetry and partial storage of integrals.

Note that for all systems, mixed ramp-Gaussian basis sets have timings of the same order of magnitude as all-Gaussian basis sets and will thus be much faster than a comparable sized all-Slater basis set. Thus, the algorithms and implementations of mixed ramp-Gaussian integrals will allow this type of basis set to replace Slater and specialised all-Gaussian basis set as a faster alternative to produce high accuracy calculations of core-dependent properties such as electron density at the nucleus, core correlation, NMR parameters, and total energies.

We want to end looking towards the future. In molecular quantum chemistry, a field that has been focused on Gaussian basis sets for decades, in a field where all alternatives basis sets have been found to be too slow despite considerable effort, the promise and potential of a new type of basis set is invigorating. And while this paper has gone a long way towards determining how to use this new basis set in fast calculations, already comparable to or faster than all-Gaussian basis sets, it is only one paper and it is the nature of science that more research makes better algorithms and faster calculations. Further improvements in speed of Gaussian basis set integral calculations are almost certainly incremental at this point; a new class of largely unexplored integrals' types surely offer much more opportunities for large factor speed-ups. We welcome this future research.

ACKNOWLEDGMENTS

The author thanks Andrew, Caleb, Emil, Jonathan, and Lorenzo for their helpful comments on various drafts of this manuscript. This research was undertaken with the assistance of resources provided at the NCI National Facility systems at the Australian National University through the National Computational Merit Allocation Scheme supported by the Australian Government.

[1]P. M. W. Gill, Adv. Quantum Chem. **25**, 141–205 (1994).
[2]M. Dupuis, J. Rys, and H. F. King, J. Chem. Phys. **65**, 111–116 (1976).
[3]L. E. McMurchie and E. R. Davidson, J. Comput. Phys. **26**, 218–231 (1978).
[4]J. A. Pople and W. J. Hehre, J. Comput. Phys. **27**, 161–168 (1978).
[5]P. M. W. Gill, M. Head-Gordon, and J. A. Pople, J. Phys. Chem. **94**, 5564–5572 (1990).
[6]P. M. W. Gill, B. G. Johnson, and J. A. Pople, Int. J. Quantum Chem. **40**, 745–752 (1991).
[7]S. Obara and A. Saika, J. Chem. Phys. **84**, 3963–3974 (1986).
[8]M. Head-Gordon and J. A. Pople, J. Chem. Phys. **89**, 5777–5786 (1988).
[9]P. M. W. Gill and J. A. Pople, Int. J. Quantum Chem. **40**, 753–772 (1991).
[10]L. K. McKemmish and P. M. W. Gill, J. Chem. Theory Comput. **8**, 4891–4898 (2012).
[11]T. Kato, Commun. Pure Appl. Math. **10**, 151–177 (1957).
[12]W. Klopper and W. Kutzelnigg, J. Mol. Struct. (Theochem) **135**, 339–356 (1986).
[13]W. Kutzelnigg, Int. J. Quantum Chem. **51**, 447–463 (1994).
[14]W. Kutzelnigg, "Rate of convergence of basis expansions in Quantum Chemistry," in *Mathematical Methods in Quantum Chemistry* (2011), Vol. 8, pp. 1775–1785.
[15]W. Kutzelnigg, AIP Conf. Proc. **1504**, 15–30 (2012).
[16]W. Kutzelnigg, Int. J. Quantum Chem. **113**, 203–217 (2013).
[17]L. K. McKemmish, A. T. B. Gilbert, and P. M. W. Gill, J. Chem. Theory Comput. **10**, 4369–4376 (2014).
[18]P. Hoggan, M. Ruiz, and T. Ozdogan, *Quantum Frontiers of Atoms and Molecules* (Nova Publishing Inc., New York, 2011), pp. 61–89.
[19]I. Ema, J. G. De La Vega, G. Ramírez, R. López, J. Fernández Rico, H. Meissner, and J. Paldus, J. Comput. Chem. **24**, 859–868 (2003).
[20]E. Van Lenthe and E. J. Baerends, J. Comput. Chem. **24**, 1142–1156 (2003).
[21]D. P. Chong, E. Van Lenthe, S. Van Gisbergen, and E. J. Baerends, J. Comput. Chem. **25**, 1030–1036 (2004).
[22]J. Fernández Rico, R. López, I. Ema, and G. Ramírez, Int. J. Quantum Chem. **81**, 148–153 (2014).
[23]M. A. Watson, N. C. Handy, A. J. Cohen, and T. Helgaker, J. Chem. Phys. **120**, 7252–7261 (2004).
[24]G. te Velde, F. M. Bickelhaupt, E. J. Baerends, C. F. Guerra, S. J. A. van Gisbergen, J. G. Snijders, and T. Ziegler, J. Comput. Chem. **22**, 931–967 (2001).
[25]J. C. Slater, Phys. Rev. **36**, 57 (1930).
[26]J. Davenport, Phys. Rev. B **29**, 2896 (1984).
[27]E. t. Clementi and D.-L. Raimondi, J. Chem. Phys. **38**, 2686–2689 (1963).
[28]J. Fernández Rico, G. Ramírez, R. López, and J. I. Fernández-Alonso, Collect. Czech. Chem. Commun. **53**, 2250–2265 (1988).
[29]T. Özdoğan and M. Orbay, Int. J. Quantum Chem. **87**, 15–22 (2002).
[30]L. Berlu, H. Safouhi, and P. Hoggan, Int. J. Quantum Chem. **99**, 221–235 (2004).
[31]J. C. Cesco, J. E. Pérez, C. C. Denner, G. O. Giubergia, and A. E. Rosso, Appl. Numer. Math. **55**, 173–190 (2005).
[32]H. Safouhi and P. Hoggan, Int. J. Quantum Chem. **84**, 580–591 (2001).
[33]J. Fernández Rico, R. López, I. Ema, and G. Ramírez, J. Comput. Chem. **25**, 1987–1994 (2004).
[34]J. Fernández Rico, R. López, I. Ema, and G. Ramírez, J. Comput. Chem. **26**, 846–855 (2005).
[35]T. Özdoğan, Int. J. Quantum Chem. **100**, 69–79 (2004).
[36]H. Safouhi, J. Comput. Phys. **165**, 473–495 (2000).
[37]J. Fernandez Rico, R. Lopez, I. Ema, and G. Ramirez, Int. J. Quantum Chem. **108**, 1415–1421 (2008).
[38]J. E. Avery and J. S. Avery, Adv. Quantum Chem. **70**, 265–324 (2015).
[39]M. Lesiuk and R. Moszynski, Phys. Rev. E **90**, 063319 (2014).
[40]M. Lesiuk and R. Moszynski, Phys. Rev. E **90**, 063318 (2014).
[41]M. Lesiuk, M. Przybytek, M. Musiał, B. Jeziorski, and R. Moszynski, Phys. Rev. A **91**, 012510 (2015).
[42]D. M. Bishop, J. Chem. Phys. **40**, 1322–1325 (1964).
[43]D. M. Bishop, J. Chem. Phys. **48**, 291–300 (1968).
[44]E. Steiner and S. Sykes, Mol. Phys. **23**, 643–656 (1972).
[45]E. Steiner, Mol. Phys. **23**, 657–667 (1972).
[46]E. Steiner, Mol. Phys. **23**, 669–681 (1972).
[47]E. Steiner and B. C. Walsh, J. Chem. Soc., Faraday Trans. 2 **71**, 921–925 (1975).
[48]E. Steiner and B. C. Walsh, J. Chem. Soc., Faraday Trans. 2 **71**, 926–936 (1975).
[49]E. Steiner, J. Chem. Soc., Faraday Trans. 2 **76**, 391–404 (1980).
[50]E. Steiner, J. Chem. Soc., Faraday Trans. 2 **81**, 1101–1105 (1985).
[51]E. Steiner, J. Chem. Soc., Faraday Trans. 2 **83**, 783–790 (1987).
[52]E. J. Baerends, D. E. Ellis, and P. Ros, Chem. Phys. **2**, 41–51 (1973).
[53]J. L. Whitten, J. Chem. Phys. **58**, 4496–4501 (1973).
[54]B. I. Dunlap, J. W. D. Connolly, and J. R. Sabin, J. Chem. Phys. **71**, 3396–3402 (1979).
[55]R. A. Kendall and H. A. Früchtl, Theor. Chem. Acc. **97**, 158–163 (1997).
[56]F. R. Manby, J. Chem. Phys. **119**, 4607–4613 (2003).
[57]W. Klopper, F. R. Manby, S. Ten-No, and E. F. Valeev, Int. Rev. Phys. Chem. **25**, 427–468 (2006).

[58]H.-J. Werner and F. R. Manby, J. Chem. Phys. **124**, 054114 (2006).
[59]See supplementary material at http://dx.doi.org/10.1063/1.4916314 for preliminary investigations into gradients and second derivative integrals in mixed ramp-Gaussian basis sets.
[60]R. Ditchfield, W. J. Hehre, and J. A. Pople, J. Chem. Phys. **54**, 724–728 (1971).
[61]T. H. Dunning, J. Chem. Phys. **90**, 1007–1023 (1989).
[62]F. Jensen, J. Chem. Phys. **115**, 9113–9125 (2001).
[63]P.-O. Widmark, P.-Å. Malmqvist, and B. O. Roos, Theor. Chim. Acta **77**, 291–306 (1990).
[64]F. Jensen, J. Chem. Theory Comput. **10**, 1074–1085 (2014).
[65]A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry* (McGraw-Hill, New York, 1989), pp. 410–458.
[66]F. Weigend, M. Häser, H. Patzelt, and R. Ahlrichs, Chem. Phys. Lett. **294**, 143–152 (1998).
[67]F. Weigend, A. Köhn, and C. Hättig, J. Chem. Phys. **116**, 3175–3183 (2002).
[68]Y. Jung, A. Sodt, P. M. Gill, and M. Head-Gordon, Proc. Natl. Acad. Sci. U. S. A. **102**, 6692–6697 (2005).
[69]F. Aquilante, T. B. Pedersen, and R. Lindh, J. Chem. Phys. **126**, 194106 (2007).
[70]A. Stone, *The Theory of Intermolecular Forces* (Oxford University Press, 2013), pp. 43–56.
[71]C. Hättig, Chem. Phys. Lett. **260**, 341–351 (1996).
[72]P. M. Gill, B. G. Johnson, and J. A. Pople, Chem. Phys. Lett. **217**, 65–68 (1994).
[73]S. Reine, T. Helgaker, and R. Lindh, Wiley Interdiscip. Rev.: Comput. Mol. Sci. **2**, 290–303 (2012).
[74]S. F. Boys, Proc. Phys. Soc. London **A200**, 542–554 (1950).
[75]H. E. Pence and A. Williams, J. Chem. Educ. **87**, 1123–1124 (2010).
[76]Y. Shao, Z. Gan, E. Epifanovsky, A. T. Gilbert, M. Wormit, J. Kussmann, A. W. Lange, A. Behn, J. Deng, X. Feng, D. Ghosh, M. Goldey, P. R. Horn, L. D. Jacobson, I. Kaliman, R. Z. Khaliullin, T. Kuś, A. Landau, J. Liu, E. I. Proynov, Y. M. Rhee, R. M. Richard, M. A. Rohrdanz, R. P. Steele, E. J. Sundstrom, H. L. Woodcock, P. M. Zimmerman, D. Zuev, B. Albrecht, E. Alguire, B. Austin, G. J. O. Beran, Y. A. Bernard, E. Berquist, K. Brandhorst, K. B. Bravaya, S. T. Brown, D. Casanova, C.-M. Chang, Y. Chen, S. H. Chien, K. D. Closser, D. L. Crittenden, M. Diedenhofen, R. A. DiStasio, H. Do, A. D. Dutoi, R. G. Edgar, S. Fatehi, L. Fusti-Molnar, A. Ghysels, A. Golubeva-Zadorozhnaya, J. Gomes, M. W. Hanson-Heine, P. H. Harbach, A. W. Hauser, E. G. Hohenstein, Z. C. Holden, T.-C. Jagau, H. Ji, B. Kaduk, K. Khistyaev, J. Kim, J. Kim, R. A. King, P. Klunzinger, D. Kosenkov, T. Kowalczyk, C. M. Krauter, K. U. Lao, A. D. Laurent, K. V. Lawler, S. V. Levchenko, C. Y. Lin, F. Liu, E. Livshits, R. C. Lochan, A. Luenser, P. Manohar, S. F. Manzer, S.-P. Mao, N. Mardirossian, A. V. Marenich, S. A. Maurer, N. J. Mayhall, E. Neuscamman, C. M. Oana, R. Olivares-Amaya, D. P. O'Neill, J. A. Parkhill, T. M. Perrine, R. Peverati, A. Prociuk, D. R. Rehn, E. Rosta, N. J. Russ, S. M. Sharada, S. Sharma, D. W. Small, A. Sodt, T. Stein, D. Stück, Y.-C. Su, A. J. Thom, T. Tsuchimochi, V. Vanovschi, L. Vogt, O. Vydrov, T. Wang, M. A. Watson, J. Wenzel, A. White, C. F. Williams, J. Yang, S. Yeganeh, S. R. Yost, Z.-Q. You, I. Y. Zhang, X. Zhang, Y. Zhao, B. R. Brooks, G. K. Chan, D. M. Chipman, C. J. Cramer, W. A. Goddard, M. S. Gordon, W. J. Hehre, A. Klamt, H. F. Schaefer, M. W. Schmidt, C. D. Sherrill, D. G. Truhlar, A. Warshel, X. Xu, A. Aspuru-Guzik, R. Baer, A. T. Bell, N. A. Besley, J.-D. Chai, A. Dreuw, B. D. Dunietz, T. R. Furlani, S. R. Gwaltney, C.-P. Hsu, Y. Jung, J. Kong, D. S. Lambrecht, W. Liang, C. Ochsenfeld, V. A. Rassolov, L. V. Slipchenko, J. E. Subotnik, T. Van Voorhis, J. M. Herbert, A. I. Krylov, P. M. Gill, and M. Head-Gordon, Mol. Phys. **113**, 184–215 (2015).