Supplementary Material for The Numerical Evaluation of Slater Integrals on GPU

Duy-Khoi Dang, Leighton W. Wilson, and Paul M. Zimmerman[*]

[*]paulzim@umich.edu

University of Michigan

930 N University Ave.

Ann Arbor, MI 48109

Contents

S1   One-electron integrals

The one electron integrals are of the form

$$O_{\mu\nu} = \langle \chi_\mu | \hat{O}_1 | \chi_\nu \rangle = \int \chi_\mu(\mathbf{r}) \hat{O}(\mathbf{r}) \chi_\nu(\mathbf{r}) \mathrm{d}\mathbf{r}, \tag{S1}$$

where operators of interest are the overlap ($\hat{O} = 1$), kinetic energy ($\hat{O} = \frac{1}{2}\nabla^2$), and nuclear attraction ($\hat{O} = \frac{Z_A}{R_{1A}}$) operators. For a set of grid points x and grid weights $w(x)$, the integral in Equation $S1$ can be evaluated by

$$O_{\mu\nu} = N_{\chi_\mu} N_{\chi_\nu} \sum_x \bar{\chi}_\mu(x) \hat{O}(\bar{\chi}_\nu(x)) w(x), \tag{S2}$$

where $\bar{\chi}_\lambda$ is the basis function with normalization constant N_{χ_λ} factored out.

S2   Radial component of the Slater Coulomb Potential

The radial component of the Slater Coulomb potential has the form[1]

$$I_{nl}(r) = r^{-l-1} \int_0^r (r')^{n+l+1} e^{-\zeta r'} \mathrm{d}r' + r^l \int_r^\infty (r')^{n-l} e^{-\zeta r'} \mathrm{d}r', \tag{S3}$$

which has analytic forms for each n, l of interest in quantum chemistry. The expressions for I_{nl} up to $n = 3$ are listed below. As n, l increase, the expression demonstrates numerical instability resulting in non-smooth STO integration, as discussed in the main text.

$$I_{1,0}(r) = \frac{\left(-\zeta r + 2e^{\zeta r} - 2\right) e^{-\zeta r}}{\zeta^3 r}$$

$$I_{2,0}(r) = \frac{\left(-\zeta^2 r^2 - 4\zeta r + 6e^{\zeta r} - 6\right) e^{-\zeta r}}{\zeta^4 r}$$

$$I_{2,1}(r) = \frac{3\left(-\zeta^3 r^3 - 4\zeta^2 r^2 - 8\zeta r + 8e^{\zeta r} - 8\right) e^{-\zeta r}}{\zeta^5 r^2}$$

$$I_{3,0}(r) = \frac{\left(-\zeta^3 r^3 - 6\zeta^2 r^2 - 18\zeta r + 24e^{\zeta r} - 24\right)e^{-\zeta r}}{\zeta^5 r}$$

$$I_{3,1}(r) = \frac{3\left(-\zeta^4 r^4 - 6\zeta^3 r^3 - 20\zeta^2 r^2 - 40\zeta r + 40e^{\zeta r} - 40\right) e^{-\zeta r}}{\zeta^6 r^2}$$

$$I_{3,2}(r) = \frac{5\left(-\zeta^5 r^5 - 6\zeta^4 r^4 - 24\zeta^3 r^3 - 72\zeta^2 r^2 - 144\zeta r + 144e^{\zeta r} - 144\right) e^{-\zeta r}}{\zeta^7 r^3}$$

S3 Basis set specifications

The STO subshell and exponent are listed for the main and auxiliary basis for each atom considered in the main paper. Basis sets are modified versions of those provided in the ADF package.[2]

DZP basis set specifications

H basis specification

Main basis		Auxiliary basis								
1s	0.76	1s	3.16	2s	1.50	2p	1.75	4f	3.00
1s	1.28	1s	2.09	2p	4.00	3d	4.00	5g	4.00
2p	1.25	1s	1.38	2p	2.65	3d	2.50		

B basis specification

Main basis		Auxiliary basis								
1s	6.50	1s	13.00	3s	2.56	3p	1.68	5g	4.50
1s	4.08	2s	14.79	3s	1.87	3d	6.08	5g	5.50
2s	1.00	2s	10.16	3s	1.36	3d	3.69		
2s	1.56	2s	6.98	2p	9.10	3d	2.24		
2p	1.70	2s	4.80	2p	5.17	3d	1.36		
2p	0.76	3s	4.83	3p	4.36	4f	5.00		
3d	1.50	3s	3.52	3p	2.71	4f	3.50		

C basis specification

Main basis		Auxiliary basis								
1s	7.68	1s	15.36	3s	3.08	3p	2.06	5g	4.50
1s	5.00	2s	17.53	3s	2.25	3d	7.20		
2s	1.98	2s	12.07	3s	1.64	3d	4.40		
2s	1.24	2s	8.31	2p	9.88	3d	2.69		
2p	2.20	2s	5.73	2p	5.80	3d	1.64		

```
2p   0.96        3s   5.78        3p   5.05        4f   5.40
3d   2.20        3s   4.22        3p   3.23        4f   3.55
```

O basis specification

```
Main basis          Auxiliary basis
1s   9.80        1s   19.60       3s   4.15        3p   2.84        5g   4.50
1s   6.80        2s   22.60       3s   3.05        3d   8.80
2s   1.70        2s   15.68       3s   2.24        3d   5.58
2s   2.82        2s   10.89       2p   12.86       3d   3.53
2p   3.06        2s   7.56        2p   7.68        3d   2.24
2p   1.30        3s   7.69        3p   6.78        4f   6.20
3d   2.00        3s   5.65        3p   4.39        4f   3.70
```

F basis specification

```
Main basis          Auxiliary basis
1s   10.88       1s   21.76       3s   3.95        3p   3.42        5g   4.50
1s   7.70        2s   24.39       3s   2.85        3d   9.70
2s   3.22        2s   16.56       3s   2.05        3d   6.16
2s   1.92        2s   11.24       2p   1.48        3d   3.91
2p   3.52        2s   7.63        2p   14.40       3d   2.48
2p   1.48        3s   7.60        3p   7.53        4f   6.50
3d   2.00        3s   5.48        3p   5.90        4f   3.75
Additional Functions in extended Auxiliary Basis
2P   4.76        3D   0.62
2P   2.64        3D   18.71
2P   1.26        4F   2.97
2P   0.65        4F   1.44
3D   18.71       5G   3.12
3D   1.52
```

Cr basis specification

```
Main basis          Auxiliary basis
1S   27.25       1S   54.50       7S   2.68        6D   3.62
1S   21.70       2S   60.57       7S   2.13        7D   2.60
2S   9.20        2S   40.83       7S   1.70        4F   15.60
2S   6.05        3S   40.40       2P   40.45       5F   8.92
2P   13.20       3S   28.95       3P   27.34       5F   4.30
2P   8.25        4S   27.44       4P   18.68       6F   2.54
3S   5.25        4S   20.46       5P   12.96       5G   9.75
3S   3.30        4S   15.25       5P   7.53        5G   5.17
3P   4.65        5S   14.16       6P   5.30        5G   2.74
3P   2.80        5S   10.86       6P   3.22
3D   5.70        5S   8.33        7P   2.30
3D   2.70        6S   7.65        3D   29.95
```

```
3D   1.24        6S   6.00        4D   20.59
4S   1.75        6S   4.70        5D   14.37
4S   1.00        6S   3.68        5D    8.39
4P   1.30        7S   3.36        6D    5.94
```

TZ2P basis set specifications

Li basis specification

```
Main basis                      Auxiliary basis
1s   4.24        2p   0.60       1s   8.48       3p   0.92       4f   3.50
1s   2.26        2p   1.20       1s   5.43       3d   4.72       5g   3.50
2s   2.36        3d   1.20       2p   6.60       3d   2.08
2s   0.68        4f   1.80       2p   2.98       3d   0.92
2s   0.46                        2p   1.34       4f   5.00
```

F basis specification

```
Main basis                      Auxiliary basis
1s  10.88        2p   4.54       1s  21.76       3s   2.85       3d   9.70
1s   7.70        2p   2.30       2s  24.39       3s   2.05       3d   6.15
2s   3.24        2p   1.24       2s  16.56       3s   1.48       3d   3.91
2s   1.94        3d   2.00       2s  11.24       2p  14.40       3d   2.48
2s   0.74        4f   3.00       2s   7.63       2p   7.53       4f   6.50
                                 3s   7.60       3p   5.90       4f   3.75
                                 3s   5.48       3p   3.42       5g   4.50
                                 3s   3.95       3p   1.98
```

TZP basis set specifications For the following TZP basis sets, the same auxiliary basis was used as specified in the DZP basis.

C basis specification

```
1S   7.68
1S   5.00
2S   1.28
2S   2.10
2S   4.60
2P   0.82
2P   1.48
2P   2.94
3D   2.20
```

H basis specification

```
1S   0.69
1S   0.92
1S   1.58
2P   1.25
```

S4 Geometries

CH$_4$

```
C  -0.00000  0.00000  0.00000
H   0.35665 -1.00881  0.00000
H   0.35667  0.50440  0.87365
H   0.35667  0.50440 -0.87365
H  -1.07000  0.00001  0.00000
```

C$_2$H$_6$

```
C  -0.25667 -0.36298 -0.62870
H   0.09998 -1.37179 -0.62870
H   0.10000  0.14142 -1.50235
H  -1.32667 -0.36296 -0.62870
C   0.25667  0.36298  0.62870
H  -0.09839  1.37235  0.62772
H  -0.10160 -0.14029  1.50235
H   1.32667  0.36127  0.62968
```

C$_3$H$_8$

```
C  -0.16408 -0.97376 -0.82820
H   0.19257 -1.98257 -0.82820
H   0.19259 -0.46936 -1.70185
H  -1.23408 -0.97375 -0.82820
C   0.34926 -0.24780  0.42920
H  -0.00901 -0.75107  1.30285
H   1.41926 -0.24951  0.43018
C  -0.16175  1.20494  0.42779
H  -1.23175  1.20665  0.42644
H   0.19461  1.70923  1.30163
H   0.19682  1.70831 -0.44567
```

BH$_3$: B-H bond length 1.18000 Å
BF$_3$: B-F bond length 1.46000 Å
CF$_4$: C-F bond length 1.35000 Å

C_4H_{10}

```
C  -0.25772 -1.45229 -1.25677
H   0.09893 -2.46110 -1.25677
H   0.09895 -0.94789 -2.13042
H  -1.32772 -1.45227 -1.25677
C   0.25562 -0.72633  0.00063
H  -0.10265 -1.22960  0.87428
H   1.32562 -0.72804  0.00161
C  -0.25539  0.72641 -0.00077
H  -1.32539  0.72812 -0.00213
H   0.10318  1.22979 -0.87423
C   0.25751  1.45221  1.25690
H   1.32750  1.45043  1.25830
H  -0.09748  2.46161  1.25588
H  -0.10114  0.94888  2.13036
```

C_5H_{12}

```
C  -0.20131 -2.03735 -1.50082
H   0.15534 -3.04616 -1.50082
H   0.15536 -1.53295 -2.37447
H  -1.27131 -2.03734 -1.50082
C   0.31203 -1.31139 -0.24341
H  -0.04624 -1.81466  0.63024
H   1.38203 -1.31310 -0.24243
C  -0.19899  0.14135 -0.24482
H  -1.26898  0.14305 -0.24617
H   0.15959  0.64472 -1.11828
C   0.31391  0.86715  1.01285
H   1.38391  0.86537  1.01425
H  -0.04473  0.36382  1.88632
C  -0.19700  2.31993  1.01139
H  -1.26700  2.32170  1.00998
H   0.15935  2.82421  1.88523
H   0.16165  2.82326  0.13793
```

$Cr(CO)_6$

```
Cr -0.44964 -0.05396  0.00000
C  -0.44964  1.84604  0.00000
C   1.45036 -0.05396  0.00000
C  -0.44964 -0.05396  1.90000
C  -0.44964 -0.05396 -1.90000
C  -2.34964 -0.05396  0.00000
O   2.56576 -0.05396  0.00000
O  -0.44964  2.96144  0.00000
O  -0.44964 -0.05396 -3.01540
O  -0.44964 -0.05396  3.01540
O  -3.46504 -0.05396  0.00000
C  -0.44964 -1.95396  0.00000
O  -0.44964 -3.06936  0.00000
```

$F^- + CH_3F$

```
C   0.33338 -0.23574  0.00000
F   1.48429 -1.04955  0.00000
H   0.34198  0.37985  0.87916
H   0.34198  0.37985 -0.87916
H  -0.53718 -0.86347  0.00000
F  -1.70265  1.20396  0.00000
```

$[FCH_3F]^-$

```
C  -0.00151 -0.00151 -0.00151
F   1.12836  1.12836  1.12836
H  -0.44307 -0.44307  0.87143
H  -0.44307  0.87143 -0.44307
H   0.87143 -0.44307 -0.44307
F  -1.12662 -1.12662 -1.12662
```

Cyclobutadiene D_{2h}

```
C  2.02818 -0.10367 -0.56968
C  3.38504 -0.11014 -0.53143
C  3.38648  1.46992 -0.30910
C  2.02955  1.47649 -0.34724
H  1.25399 -0.85949 -0.69767
H  4.15799 -0.87319 -0.61720
H  4.16066  2.22574 -0.18090
H  1.25658  2.23961 -0.26231
```

Cyclobutadiene D_{4h}

```
C  1.94150   0.34778 -0.37424
C  3.41620   1.78907 -0.14334
C  3.40850   0.33689 -0.35264
C  1.94920   1.79996 -0.16495
H  1.16468 -0.41143 -0.49587
H  4.17721 -0.43380 -0.45150
H  4.19302  2.54828 -0.02171
H  1.18049  2.57065 -0.06608
```

S5 Performance and Kernel Analysis

The performance behaviors of the 2080-Ti and GV100 can be explained by comparing the underlying hardware. Specifically, the GV100 has $\sim 40\%$ higher theoretical memory bandwidth than the 2080-Ti and $\sim 17\%$ more streaming multiprocessors (SM) than the 2080-Ti, however, the 2080-Ti benefited from 27% higher clock speed throughout the timing tests. The greater performance of the 2080-Ti in mixed precision integral evaluation therefore suggests that the GPU kernels are compute bound. The performance inversion observed in full double precision integration is primarily a result of the substantially fewer double precision execution units in the 2080-Ti compared to the GV100. To further verify whether that the SlaterGPU kernels are compute bound, the integrals for C_5H_{12} were computed using a Nvidia A4000 GPU at various locked memory and SM clock speeds. Figure S1 shows the relative performance loss by lowering the memory and SM clock speeds separately. Full timing details are provided in Table S1. In mixed and double precision, lowering the SM clock speed by 14% (1560 MHz to 1335 MHz) and 30% (1560 MHz to 1095 MHz) leads to ERI performance loss of $\sim 12\%$ and $\sim 25\%$, respectively. On the other hand, lowering the memory clock by 24% only leads to $\sim 5\%$ and $< 1\%$ performance loss in mixed and double precision, respectively. Thus, the current implementation of the ERI GPU kernels are indeed compute bound.
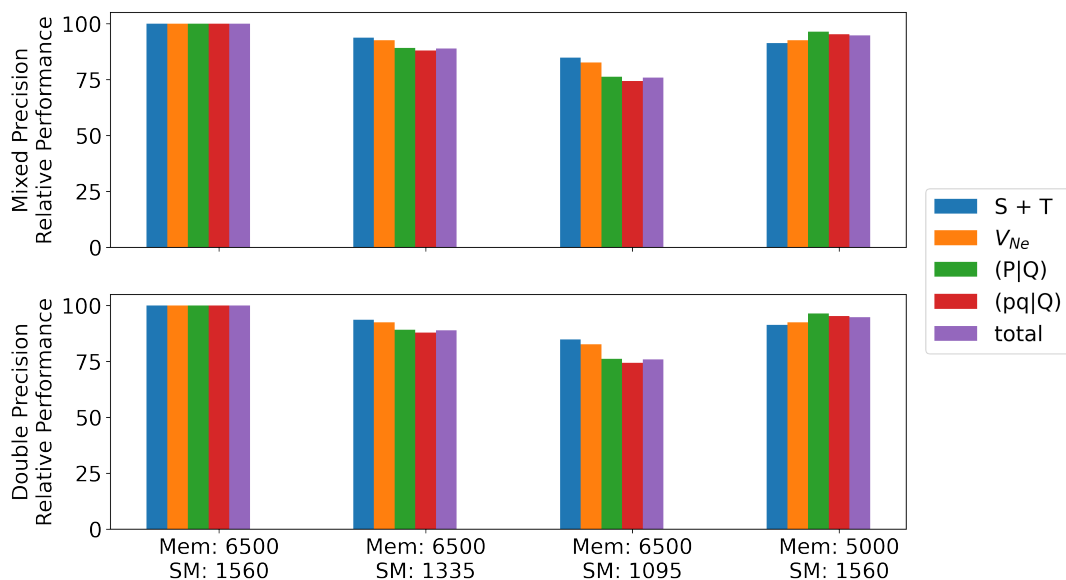

Figure S1: Relative performance of Slater integration is plotted as memory at various memory and SM clock speeds. Integrals were computed for the C_5H_{12} molecule. The x-axis labels list the memory clock followed by the SM clock in MHz.

8

Table S1: Timings (in seconds) for Slater integration at various memory and SM clock speeds. Integrals were computed for the C_5H_{12} molecule.

Precision	Memory clock (MHz)	SM Clock (MHz)	$S+T$	V_{ne}	$(P\|Q)$	$(\mu\nu\|Q)$	Total
mixed	6500	1560	0.467	16.08	2.358	63.97	82.88
mixed	6500	1335	0.4986	17.38	2.646	72.76	93.28
mixed	6500	1095	0.5507	19.46	3.095	86.06	109.2
mixed	5000	1560	0.5113	17.38	2.446	67.15	87.49
double	6500	1560	0.2852	8.92	6.474	167.5	183.2
double	6500	1335	0.3016	9.541	7.433	190.7	208.0
double	6500	1095	0.326	10.5	8.889	225.3	245.0
double	5000	1560	0.3119	9.542	6.622	167.9	184.3

Hartree-Fock (HF) timings are also provided in Table S2. Compared to the GPU timings for the integrals, the HF timings take less than 10% of the overall compute time across all system sizes considered in mixed and double precision. These timings do not represent an optimized HF implementation.

Table S2: Hartree-Fock timings (in seconds) for various alkanes. Each atom contributes 46,200 grid points. Timings are recorded on a 2080-Ti GPU. Total time in HF does not include integral computation.

Mixed Precision					
Operation/Quantity	CH_4	C_2H_6	C_3H_8	C_4H_{10}	C_5H_{12}
Main basis size	35	60	85	110	135
Aux basis size	224	370	516	662	808
Time for S and T integrals (s)	4.45E−02	6.40E−02	9.75E−02	1.40E−01	1.95E−01
Time for V_{Ne} integrals (s)	2.03E−01	7.37E−01	1.94E+00	4.06E+00	7.32E+00
Time for 3-center integrals (s)	4.22E−01	2.05E+00	5.79E+00	1.25E+01	2.34E+01
Avg time for G Tensor formation (s)	1.51E−03	5.94E−03	1.62E−02	4.30E−02	8.01E−02
Avg time for AO Fock formation (s)	1.30E−05	1.48E−05	1.95E−05	3.95E−05	3.80E−05
Avg time for DIIS (s)	4.69E−04	6.22E−04	9.34E−04	1.10E−03	1.38E−03
Avg time for Fock diagonalization (s)	1.77E−04	2.57E−04	3.78E−04	4.74E−04	6.34E−04
Number of HF iterations	11	12	15	16	17
Total integral time (s)	7.67E-01	3.05E+00	8.20E+00	1.73E+01	3.18E+01
Total time in HF (s)	1.01E−01	2.72E−01	6.61E−01	1.43E+00	2.58E+00
Double Precision					
Operation/Quantity	CH_4	C_2H_6	C_3H_8	C_4H_{10}	C_5H_{12}
Main basis size	35	60	85	110	135
Aux basis size	224	370	516	662	808
Time for S and T integrals (s)	4.98E−02	8.05E−02	1.12E−01	1.63E−01	2.27E−01
Time for V_{Ne} integrals (s)	2.35E−01	8.82E−01	2.33E+00	4.86E+00	8.70E+00
Time for 3-center integrals (s)	1.98E+00	9.56E+00	2.69E+01	5.79E+01	1.07E+02
Avg time for G Tensor formation (s)	1.50E−03	5.93E−03	1.57E−02	3.39E−02	7.91E−02
Avg time for AO Fock formation (s)	1.32E−05	1.49E−05	1.95E−05	2.25E−05	3.60E−05
Avg time for DIIS (s)	4.71E−04	6.03E−04	9.03E−04	1.07E−03	1.35E−03
Avg time for Fock diagonalization (s)	1.76E−04	2.54E−04	3.78E−04	4.79E−04	6.35E−04
Number of HF iterations	11	12	15	16	17
Total integral time (s)	2.66E+00	1.15E+01	3.11E+01	6.58E+01	1.20E+02
Total time in HF (s)	9.88E−02	2.70E−01	6.32E−01	1.24E+00	2.67E+00

S6   Mixed-Precision Error

Additional plots due to mixed-precision errors are provided in Figures S3 and S4. Figures S5 and S6 plot the values of the 2-electron integrals as the internuclear distance between centers is scanned along various coordinate directions. Figures S3 and S5 use basis functions where $m = 0$, while in Figures S4 and S6, $m = 0$ for the left basis and $m = n-1$ for the right basis. In each of these figures, ζ is set to 1 for the left and right basis functions.

S7   Lithium Fluoride Parallelity Test

Another test to check the mixed-precision error was performed with the HBCI method ($\varepsilon_1 = 0.5$ mHa, $\varepsilon_2 = 0.5$ μHa) on the lithium fluoride molecule. A triple-ζ doubly-polarized basis, denoted TZ2P, which contains a total of 48 basis functions and 128 auxiliary basis functions was used with no frozen core orbitals. The LiF bond was stretched from 1.4 Å to 3.0 Å. A plot of the LiF energies and non-parallelity error shown in Figure S2 demonstrates a smooth potential energy curve with negligible errors from using mixed-precision integrals on the order of μHa.
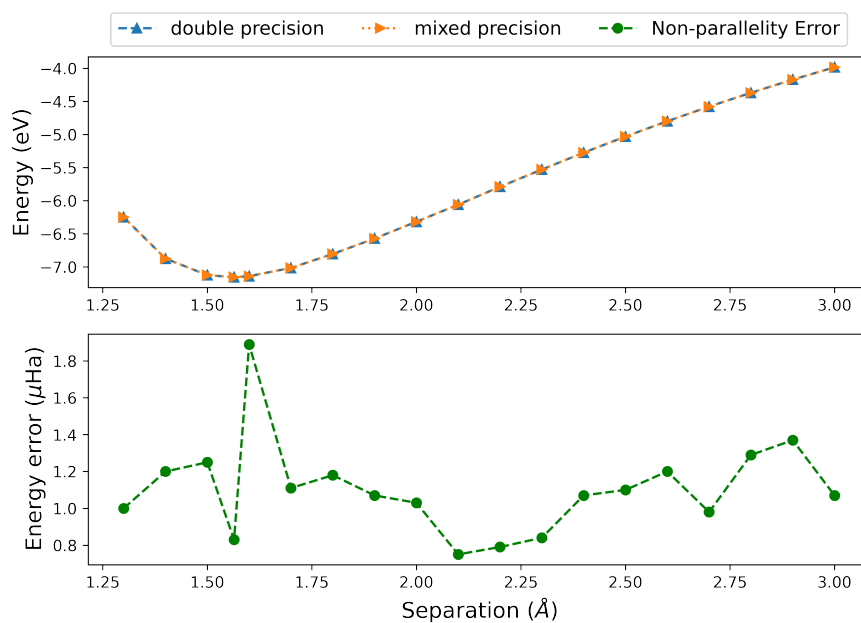
Figure S2: The LiF molecule is stretched from 1.4 Å to 3.0 Å. Energies (top) were computed using HBCI($\varepsilon_1 = 0.5$ mHa, $\varepsilon_2 = 0.1$ μHa) with the TZ2P basis. An additional point at the experimental bond distance[3] (1.564 Å) was included. Energies are relative to complete dissociation of the LiH bond. The non-parallelity error (bottom) from using mixed-precision integrals is also plotted.

S8   OpenACC Code Examples

Algorithm S1 Loop for computing the basis χ or potential V on a grid x with npts. The function f will refer to either χ, or V, and fx will refer to f evaluated on the grid x.

```
#pragma acc parallel loop present(fx[0:npts],x[0:npts])
for (int i = 0; i < npts; i++)
    fx[i] = f(x[i]);
```

Algorithm S2 Sample code for contracting V_P, μ, and ν into the 3-center integrals. Arrays corresponding to V_P, μ, and ν are assumed to have already been computed, and properly weighted.

```
int N2a = N*N*Naux;
#pragma acc parallel loop collapse(3) present(\
    V_p[0:nbas_atm1][0:npts],mu[0:nbas_atm2][0:npts],nu[0:nbas_atm3][0:npts],Vmunu[0:N2a])
for (int i = atom1_start; i < atom1_end; i++) { // loop over basis functions on atom 1
  for (int j = atom2_start; j < atom2_end; j++) { // loop over basis functions on atom 2
    for (int k = atom_3_start; k < atom3_end; k++) { // loop over basis functions on atom 3
      // shift i,j,k to properly index V_p, mu, nu from 0 to atom 1,2,3
      int i_shift = i - atom1_start;
      int j_shift = j - atom2_start;
      int k_shift = k - atom3_start;

      double sum = 0.;
      #pragma acc loop reduction(+:sum)
      for (int l = 0; l < npts; l++) {
        sum += v_p[i_shift][l] * mu[j_shift][l] * nu[k_shift][l];
      }
      Vmunu[i*N*N + j*N + k] = sum;
    }
  }
}
```
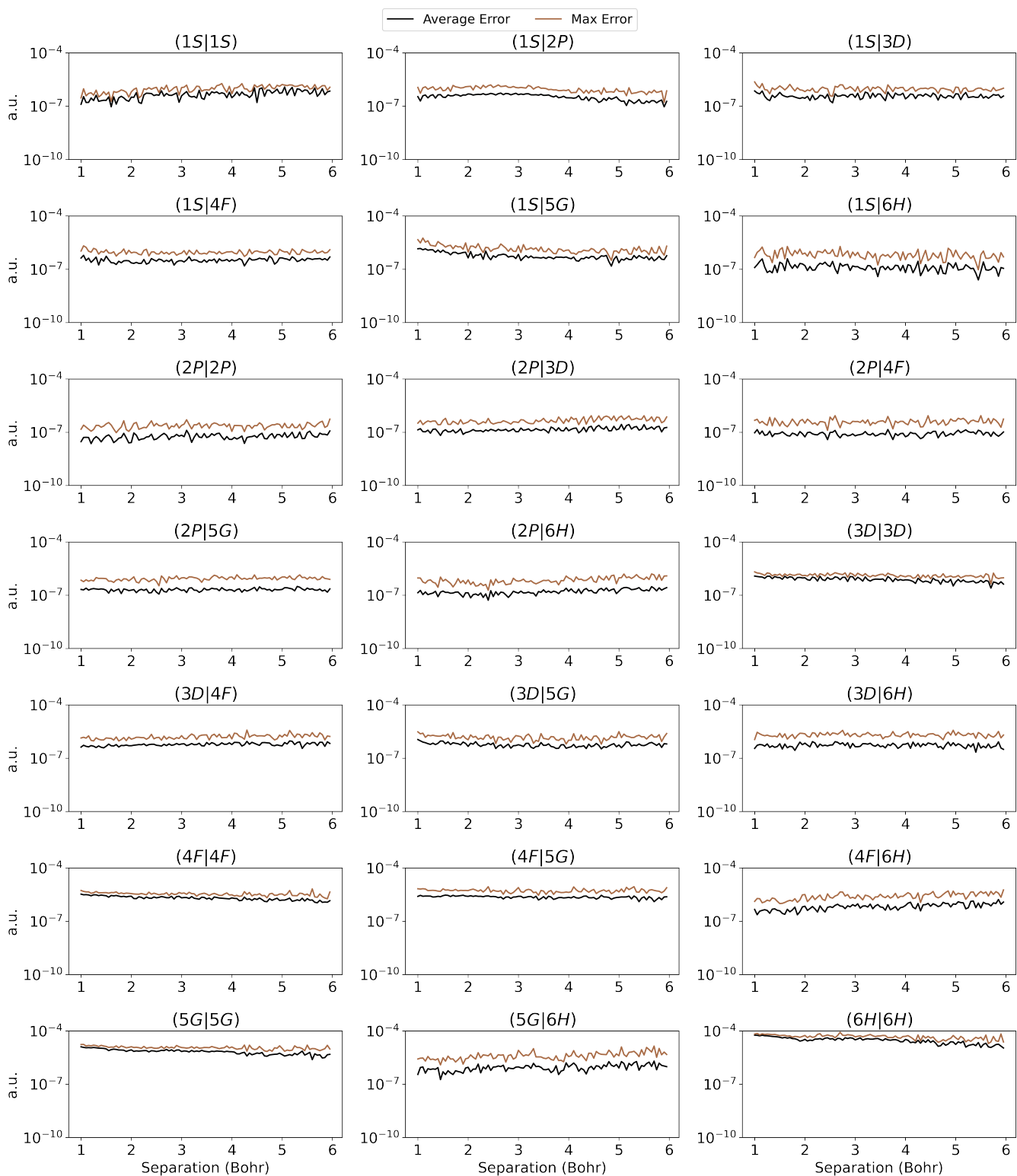
Figure S3: The max and average errors between mixed- and double-precision integral evaluation are plotted for various basis functions. All basis functions have $\zeta = 1$ and $m = 0$. The max and average errors are computed over internuclear distance scans based on the 16 all-positive directions of a Lebedev grid.
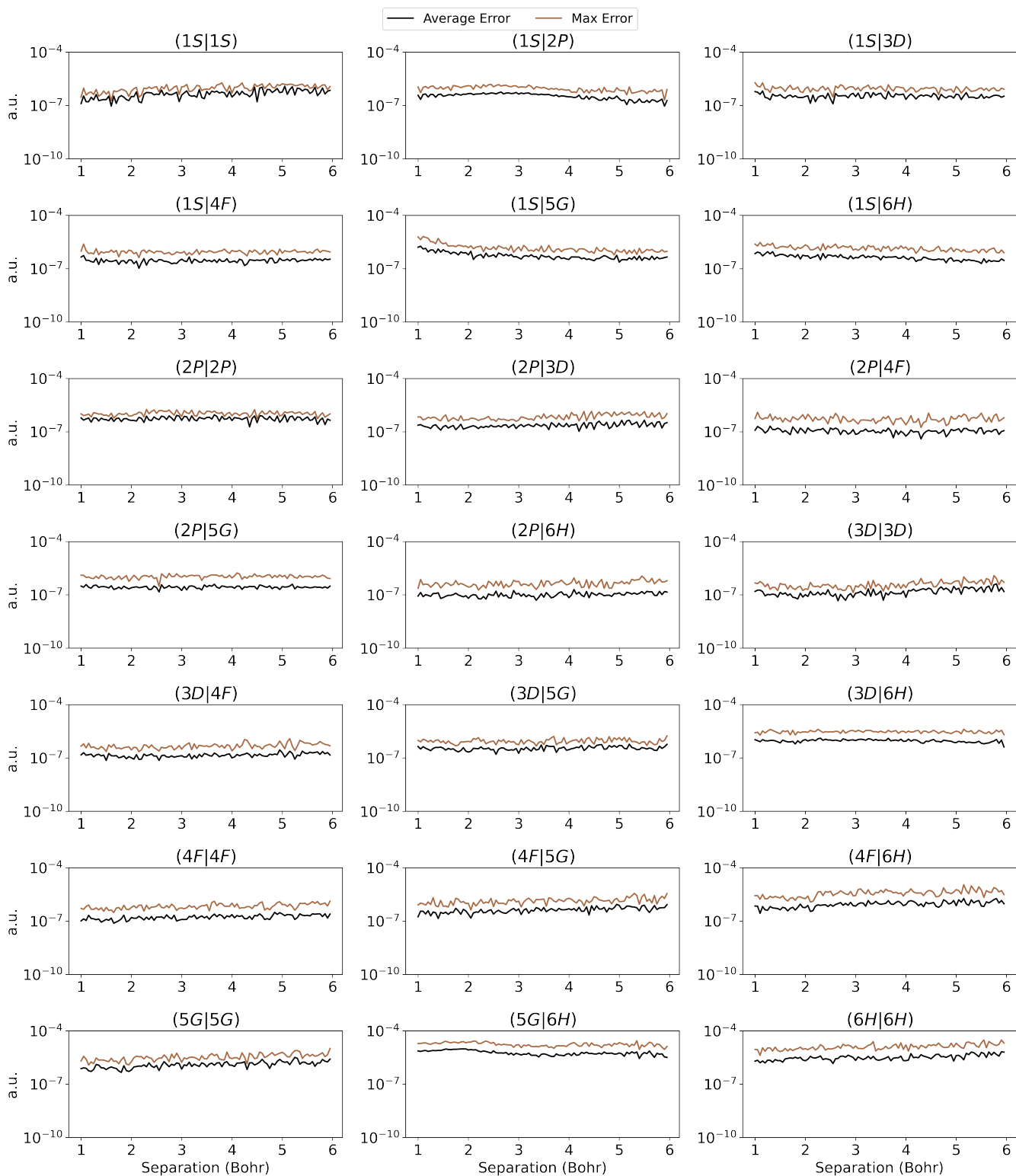
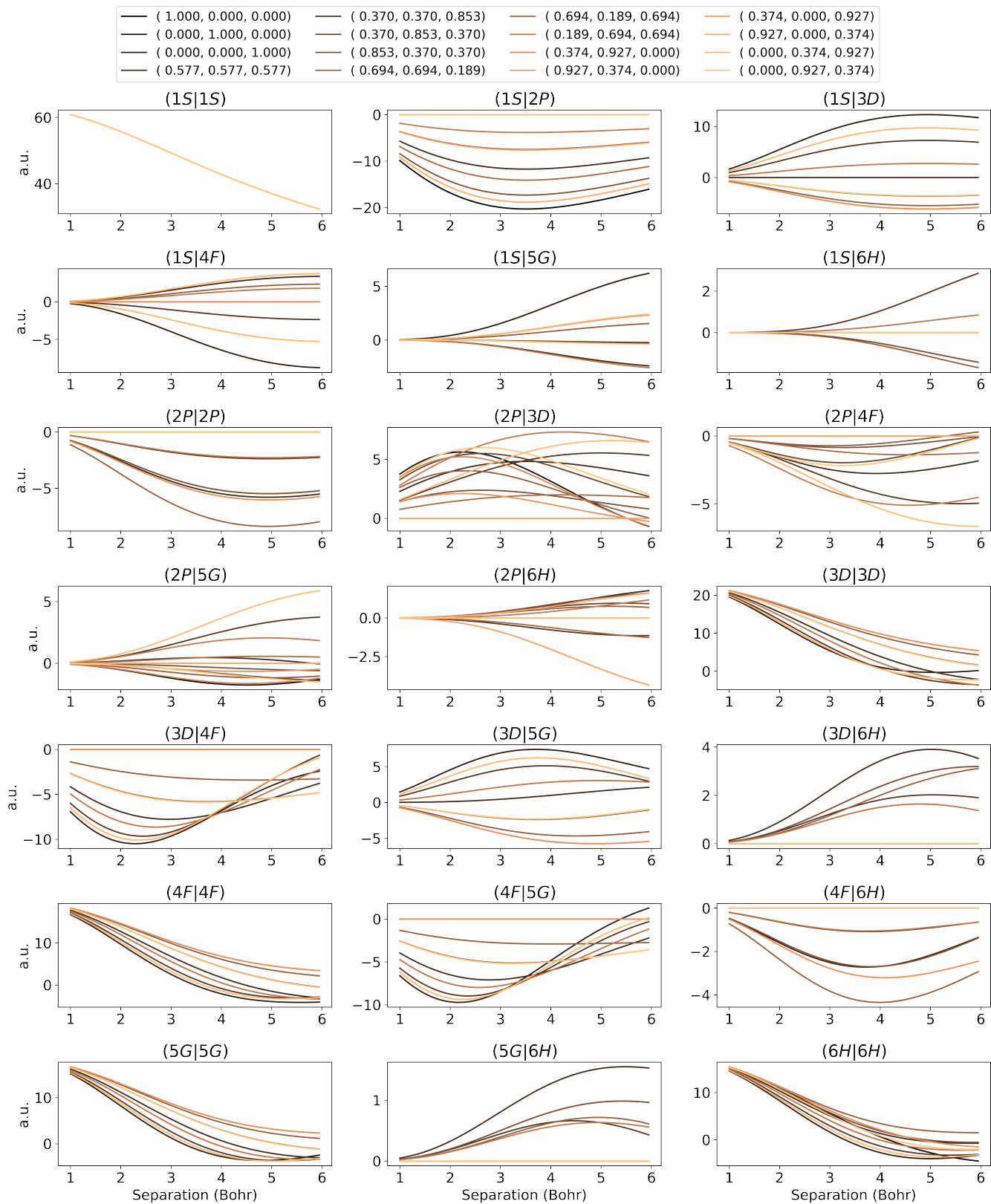Figure S4: The max and average errors between mixed- and double-precision integral evaluation are plotted for various basis functions. All basis functions have $\zeta = 1$. The left basis has $m = 0$ and the right basis has $m = n - 1$. The max and average errors are computed over internuclear distance scans based on the 16 all-positive directions of a Lebedev grid.

Figure S5: The mixed-precision 2-center ERIs are plotted for various basis functions. All basis functions have $\zeta = 1$ and $m = 0$. The ERIs are computed over internuclear distance scans based on the 16 all-positive directions of a Lebedev grid.
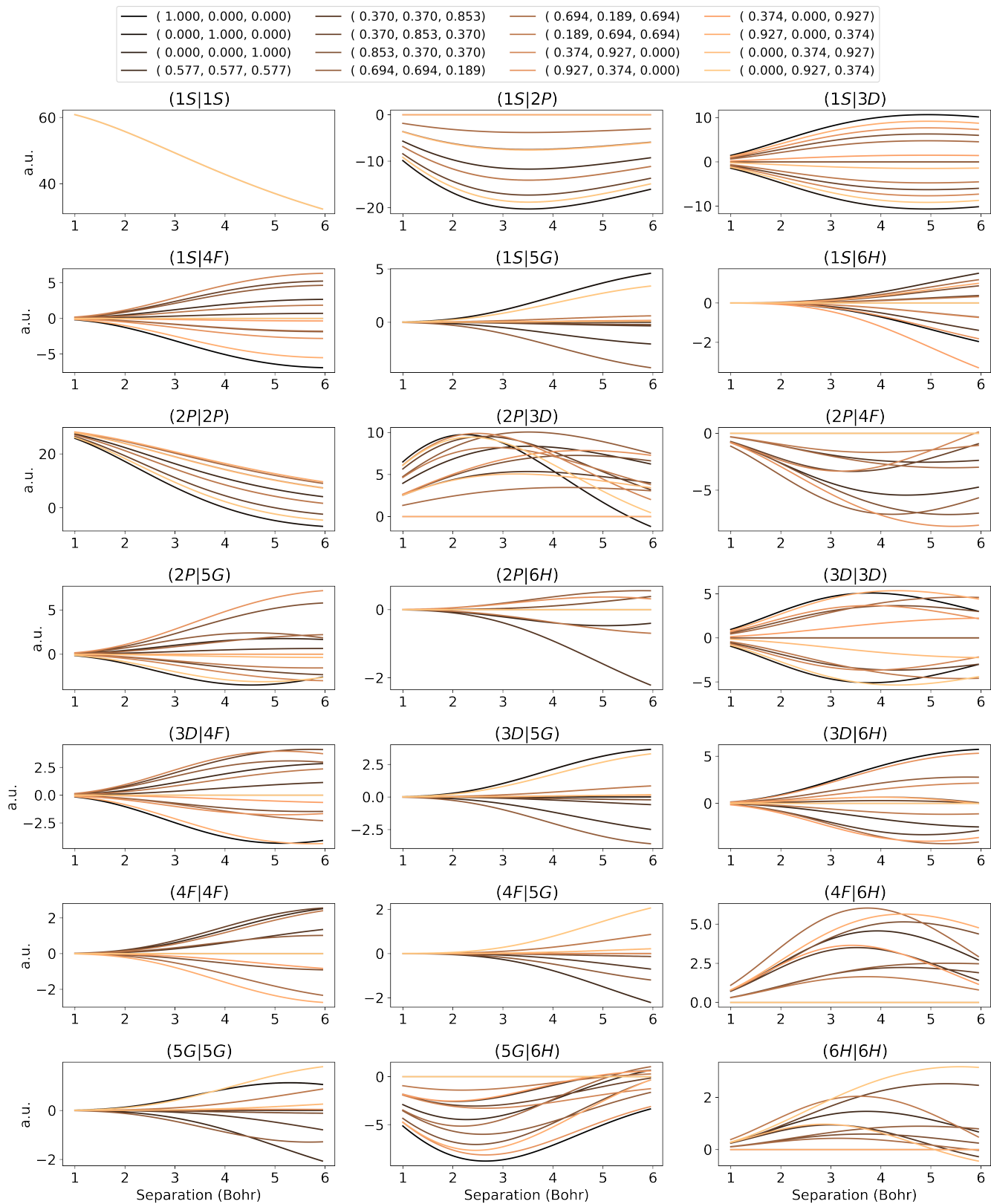
Figure S6: The mixed-precision 2-center ERIs are plotted for various basis functions. All basis functions have $\zeta = 1$. The left basis has $m = 0$ and the right basis has $m = n - 1$. The ERIs are computed over internuclear distance scans based on the 16 all-positive directions of a Lebedev grid.

References

(1) Cohen, A. J.; Handy, N. C. *The Journal of Chemical Physics* **2002**, *117*, 1470–1478.

(2) Van Lenthe, E.; Baerends, E. J. *Journal of Computational Chemistry* **2003**, *24*, 1142–1156.

(3) Lovas, F. J.; Tiemann, E. *Diatomic Spectral Database*; tech. rep. NIST Standard Reference Database 114; National Institute of Standards and Technology, 2005.