

Expensive methods, cusps, integrals & other stuff

Pierre-François Loos,^{1,2}

Anthony Scemama,¹ Yann Garniron,¹ Michel Caffarel,¹

Giuseppe Barca,² and Peter Gill.²

¹Laboratoire de Chimie et Physique Quantiques, UMR5626, Université Paul Sabatier, Toulouse, France

²Research School of Chemistry, Australian National University, Canberra, Australia

University of Basel

28th Sept 2017

Collaborators and Funding

■ Collaborators:



Anthony
Scemama



Yann
Garniron



Michel
Caffarel



Giuseppe
Barca



Peter
Gill

■ Job & Money: ANU, ARC and CNRS



Australian Government
Australian Research Council



Accurate energies with quantum Monte Carlo (QMC)

Trial wave function for QMC

$$\Psi_T(\mathbf{R}) = e^{J(\mathbf{R})} \sum_I c_I D_i^\uparrow(\mathbf{R}^\uparrow) D_i^\downarrow(\mathbf{R}^\downarrow)$$

- The **multideterminant part** is obtained via the (selected FCI) **CIPSI algorithm**
Giner et al. JCP 142 (2015) 044115
- We use a “minimal” (nodeless) **Jastrow $J(\mathbf{R})$**

A concrete example

The water molecule [Caffarel et al. JCP 144 (2016) 151103]

TABLE I. Number of determinants and corresponding variational energies for CIPSI expansions used in DMC for each cc-pCVnZ ($n = 2$ to 5) basis set. Last column: Deviations of the variational energy to the best FCI estimates of Almora-Diaz.²⁷ Energies in atomic units.

Basis set	FCI size	# dets used in DMC	E_0^{var}	FCI, Almora-Diaz ²⁷	Deviation
cc-pCVDZ	$\sim 10^{10}$	172 256	-76.282 136	-76.282 865	0.0007
cc-pCVTZ	$\sim 2 \cdot 10^{14}$	640 426	-76.388 287	-76.390 158	0.0018
cc-pCVQZ	$\sim 2 \cdot 10^{17}$	666 927	-76.419 324	-76.421 148	0.0018
cc-pCV5Z	$\sim 7 \cdot 10^{19}$	1 423 377	-76.428 550	-76.431 105	0.0025

TABLE II. All-electron DMC energies (in a.u.) obtained with CIPSI nodes for each basis set. Second column: Increase of CPU time due to the use of the large multideterminant expansion.

Basis set[Ndets]	$T_{CPU}(Ndets)/T_{CPU}(1det)$	E_0^{DMC}
cc-pCVDZ[172 256]	~ 101	-76.415 71(20)
cc-pCVTZ[640 426]	~ 185	-76.431 82(19)
cc-pCVQZ[666 927]	~ 128	-76.436 22(14)
cc-pCV5Z[1 423 377]	~ 235	-76.437 44(18)

TABLE III. Comparison of nonrelativistic ground-state total energies of water obtained with the most accurate theoretical methods. Energies in a.u.

Clark et al. ²⁰ DMC (upper bound)	-76.436 8(4)
This work, DMC (upper bound)	-76.437 44(18)
Almora-Diaz, ²⁷ CISDTQnSx (upper bound)	-76.434 3
Helgaker et al., ²⁹ R12-CCSD(T)	-76.439(2)
Muller and Kutzelnigg, ³⁰ R12-CCSD(T)	-76.437 3
Almora-Diaz, ²⁷ FCI + CBS	-76.438 6(9)
Halkier et al., ³¹ CCSD(T) + CBS	-76.438 6
Bytautas and Ruedenberg, ³² FCI + CBS	-76.439 0(4)
This work, DMC + CBS	-76.438 94(12)
Experimentally derived estimate ²⁵	-76.438 9

Section 2

CIPSI

The CIPSI method

CIPSI = CI using a Perturbative Selection made Iteratively

- Based on old idea by Bender and Davidson (1969)
- Further developments in Toulouse many years ago (Malrieu, Evangelisti, Daudey, Spiegelman, etc)
- CIPSI is a good candidate for **massively parallel wave function calculations** (PhD E. Giner and Y. Garniron)
- Open-source code: **QUANTUM PACKAGE** (A. Scemama)
https://github.com/LCPQ/quantum_package

Color code

Internal vs External

- **Red:** Variational/Internal wave function
- **Green:** External wave function (perturbation)

CIPSI algorithm

- 1 Define a *reference* wave function:

$$|\Psi\rangle = \sum_{I \in \mathcal{D}} c_I |I\rangle \quad E_{\text{var}} = \frac{\langle \Psi | \mathcal{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle}$$

- 2 Generate *external determinants*:

$$\mathcal{A} = \left\{ (\forall I \in \mathcal{D}) (\forall \hat{T} \in \mathcal{T}_1 \cup \mathcal{T}_2) : |\alpha\rangle = \hat{T}|I\rangle \right\}$$

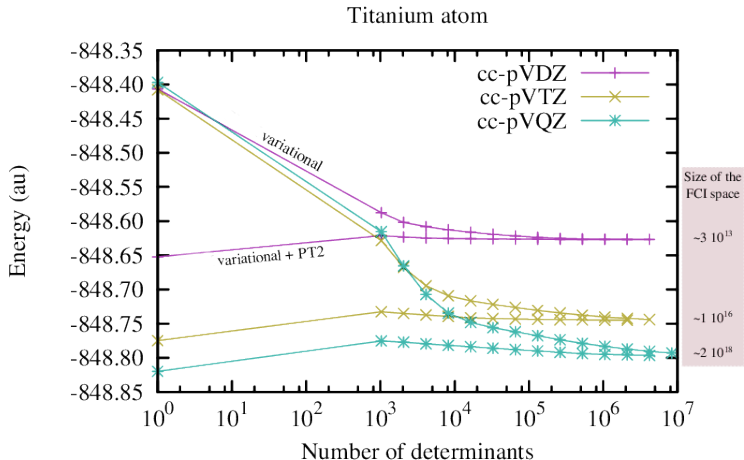
- 3 Second-order perturbative contribution of each $|\alpha\rangle$:

$$\delta E(\alpha) = \frac{|\langle \Psi | \mathcal{H} | \alpha \rangle|^2}{E_{\text{var}} - \langle \alpha | \mathcal{H} | \alpha \rangle}$$

- 4 Select $|\alpha\rangle$ with largest $\delta E(\alpha)$ and add them to \mathcal{D}
- 5 Diagonalize \mathcal{H} in $\mathcal{D} \Rightarrow$ update $|\Psi\rangle$ and E_{var}
- 6 Iterate

Giner, Scemama & Caffarel, JCP 142 (2015) 044115

CIPSI on the Titanium atom



Few remarks...

- When all $|I\rangle$ are selected, we obtain the **FCI energy**
- CIPSI is more an **algorithm** than a **method**
- CIPSI generates various wave function methods:
CID, CISD, CISDT, CAS, CASSD, MRCI, etc.
- Most of wave function methods can be performed à la CIPSI

PT2 correction

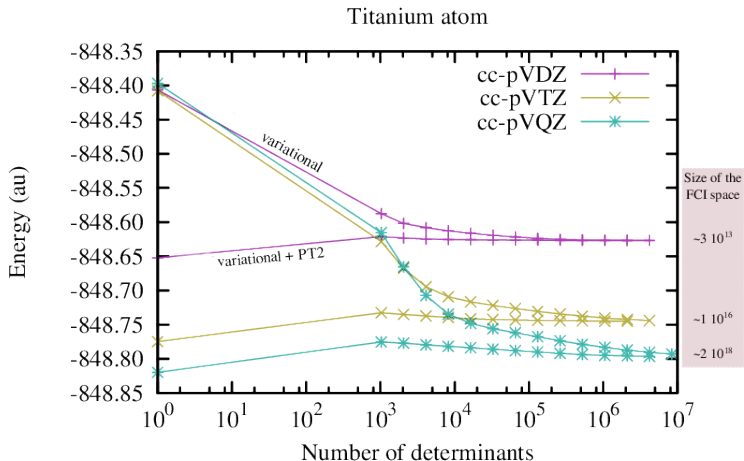
How do we know how far we are from the “true” FCI?

- Second-order Epstein-Nesbet correction:

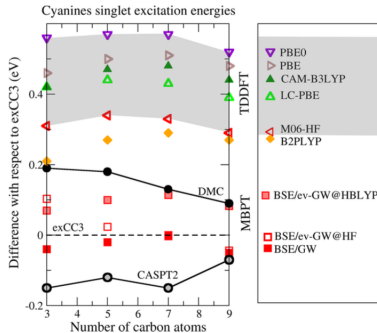
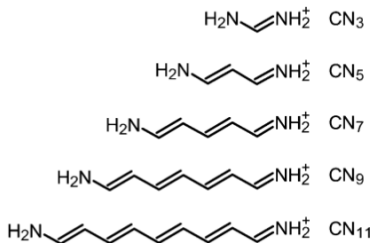
$$E_{PT2} = \sum_{\alpha} \delta E(\alpha)$$

- $|\alpha\rangle$'s with largest $\delta E(\alpha)$ have been added to Ψ previously
 ⇒ only small contributions remaining
- \mathcal{A} increases with \mathcal{D}
 ⇒ a very large number of very small contributions
- In practice, not all $|\alpha\rangle$ are generated
 ⇒ cheaper selection but E_{PT2} is biased.

CIPSI on the Titanium atom



Excited states of cyanines



Boulanger et al. JCTC 10 (2014) 1212

Excited states of cyanines

Computational details: CAS in π space (aug-cc-VDZ)

Cyanines	Size of CAS	N_{det}	Size of PT2	E_{PT2}
CN3	(2,31)	101 732	(18,111)	2×10^{-5}
CN5	(4,49)	1 435 094	(28,173)	2×10^{-4}
CN7	(6,67)	19 554 962	(38,235)	6×10^{-4}
CN9	(8,85)	10 892 791	(48,297)	4×10^{-3}
CN11	(10,103)	2 595 979	(58,359)	running

Excited states of cyanines

Vertical excitation energy of cyanines (in eV)

Method	Cyanines					Reference
	CN3	CN5	CN7	CN9	CN11	
TD-DFT calculations						
TD-B2PLYP	7.30	5.05	3.92	3.25	2.80	Send
TD-CAM-B3LYP	7.55	5.26	4.12	3.44	2.97	Send
TD-MO6-2X		5.23	4.09	3.41	2.95	Jacquemin
Reference calculations						
exCC3	7.16	4.84	3.65	2.96	2.53	Send
DMC	7.38(2)	5.03(2)	3.83(2)	3.09(2)	2.62(2)	Send
CASSCF	7.62	5.30	3.89	2.98	2.37	Send
CASPT2	7.19	4.69	3.53	2.81	2.46	Send
GW/BSE		4.80	3.63	2.96	2.48	Boulanger
Wave function methods						
FCI	7.27					
FCI(π)+PT2(σ)	7.65	4.99	3.95	3.35		
FCI(π)	7.62	5.27	3.93	3.12	2.68	
QMC methods						
DMC/FCI(π)	7.57(3)	5.03(11)				

Send et al. JCTC 7 (2011) 444; Jacquemin et al. JCTC 8 (2012) 1255; Boulanger et al. JCTC 10 (2014) 1212

Section 3

MR-PT2

Hybrid PT2 correction

THE JOURNAL OF CHEMICAL PHYSICS **147**, 034101 (2017)

Hybrid stochastic-deterministic calculation of the second-order perturbative contribution of multireference perturbation theory

Yann Garniron, Anthony Scemama,^{a)} Pierre-François Loos, and Michel Caffarel

Laboratoire de Chimie et Physique Quantiques, Université de Toulouse, CNRS, UPS, Toulouse, France

(Received 14 March 2017; accepted 26 June 2017; published online 17 July 2017)

A hybrid stochastic-deterministic approach for computing the second-order perturbative contribution $E^{(2)}$ within multireference perturbation theory (MRPT) is presented. The idea at the heart of our hybrid scheme—based on a reformulation of $E^{(2)}$ as a sum of elementary contributions associated with each determinant of the MR wave function—is to split $E^{(2)}$ into a stochastic and a deterministic part. During the simulation, the stochastic part is gradually reduced by dynamically increasing the deterministic part until one reaches the desired accuracy. In sharp contrast with a purely stochastic Monte Carlo scheme where the error decreases indefinitely as $t^{-1/2}$ (where t is the computational time), the statistical error in our hybrid algorithm displays a polynomial decay $\sim t^{-n}$ with $n = 3-4$ in the examples considered here. If desired, the calculation can be carried on until the stochastic part entirely vanishes. In that case, the exact result is obtained with no error bar and no noticeable computational overhead compared to the fully deterministic calculation. The method is illustrated on the F_2 and Cr_2 molecules. Even for the largest case corresponding to the Cr_2 molecule treated with the cc-pVQZ basis set, very accurate results are obtained for $E^{(2)}$ for an active space of (28e, 176o) and a MR wave function including up to 2×10^7 determinants. *Published by AIP Publishing.*
[\[http://dx.doi.org/10.1063/1.4992127\]](http://dx.doi.org/10.1063/1.4992127)

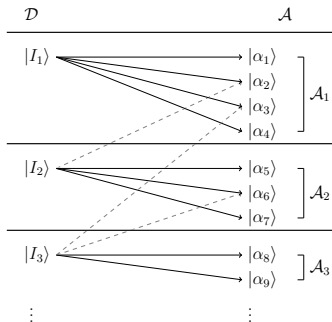
Divide & Conquer

Partition of \mathcal{A}

- Sort $|I\rangle$ according to their weights

$$w_I = \frac{c_I^2}{\langle \Psi | \Psi \rangle}$$

- Make disjoint groups \mathcal{A}_I of $|\alpha\rangle$ originating from the same generator $|I\rangle$
- Each \mathcal{A}_I has its own contribution ϵ_I to E_{PT2}



The crux

Partition of E_{PT2}

$$\begin{aligned}
 E_{PT2} &= \sum_{\alpha \in \mathcal{A}} \frac{|\langle \Psi | \mathcal{H} | \alpha \rangle|^2}{E_{var} - \langle \alpha | \mathcal{H} | \alpha \rangle} \\
 &= \sum_{I \in \mathcal{D}} \sum_{\alpha_I \in \mathcal{A}_I} \frac{|\langle \Psi | \mathcal{H} | \alpha_I \rangle|^2}{E_{var} - \langle \alpha_I | \mathcal{H} | \alpha_I \rangle} \\
 &= \sum_{I \in \mathcal{D}} \epsilon_I
 \end{aligned}$$

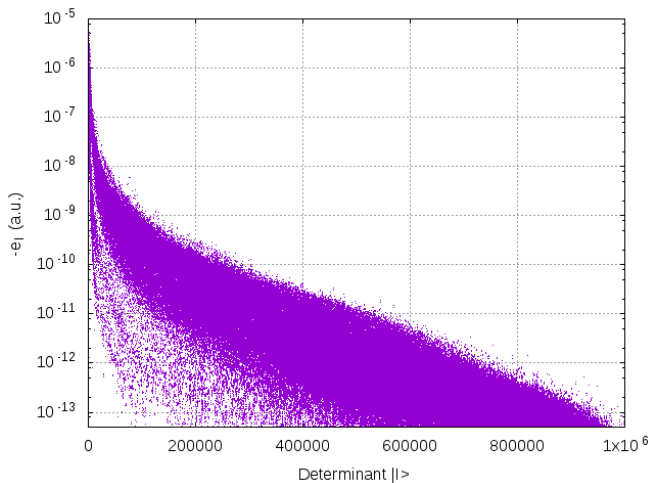
Consequences

Properties of ϵ_I

$$\epsilon_I = \sum_{\alpha_I \in \mathcal{A}_I} \frac{|\langle \Psi | \mathcal{H} | \alpha_I \rangle|^2}{E_{\text{var}} - \langle \alpha_I | \mathcal{H} | \alpha_I \rangle}$$

- $\langle \Psi | \mathcal{H} | \alpha_I \rangle = \sum_{J \geq I} c_J \langle J | \mathcal{H} | \alpha_I \rangle$
- $\langle \alpha_I | \mathcal{H} | \alpha_I \rangle$ is always high in energy, otherwise $|\alpha_I\rangle$ would be better in the **variational space** and **PT is questionable**
- $\forall I \in \mathcal{D} : \epsilon_I \leq 0$
- $|\epsilon_I|$ is expected to decay as c_I^2
- The computational cost decreases with I (less α 's in \mathcal{A}_I as I increases)

F_2 , cc-pVDZ, $N_{\text{det}} = 10^6$



Be lazy and you'll be rewarded...

Key observation

N_{det} contributions $\epsilon_I \Rightarrow$ can be stored in memory

Lazy Evaluation (Wikipedia)

In programming language theory, lazy evaluation, or call-by-need is an evaluation strategy which delays the evaluation of an expression until its value is needed (non-strict evaluation) and which also avoids repeated evaluations (sharing).

Monte Carlo with Lazy Evaluation

Monte Carlo sampling

$$E_{\text{PT2}} = \sum_{I \in \mathcal{D}} \epsilon_I = \sum_{I \in \mathcal{D}} p_I \frac{\epsilon_I}{p_I} = \left\langle \frac{\epsilon_I}{p_I} \right\rangle_{p_I}$$

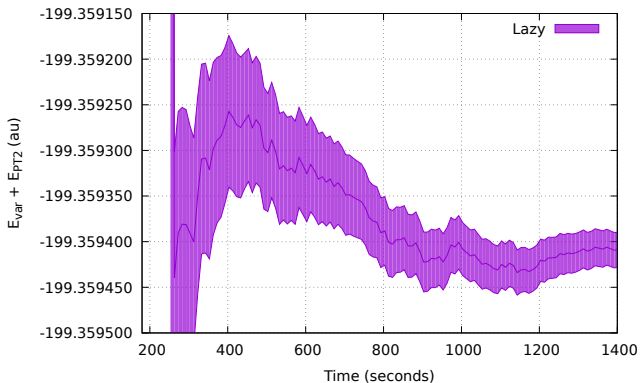
- Draw a generator determinant $|I\rangle$ with probability

$$p_I = \frac{c_I^2}{\sum_{J \in \mathcal{D}} p_J}$$

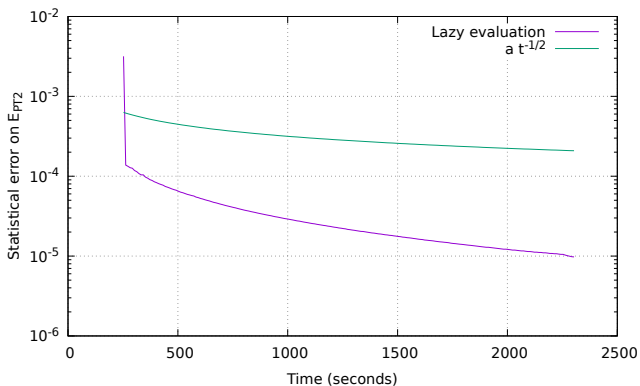
- Increment n_I , the number of evaluations of ϵ_I
- Lazy evaluation : if ϵ_I is not already computed, compute and store it

$$E_{\text{PT2}} \sim \sum_{I \in \mathcal{D}} \frac{n_I}{N_{\text{samples}}} \frac{\epsilon_I}{p_I} \quad \text{decay rate: } \mathcal{O}\left(N_{\text{samples}}^{-1/2}\right)$$

F₂, cc-pVQZ, $N_{\text{det}} = 5 \times 10^6$

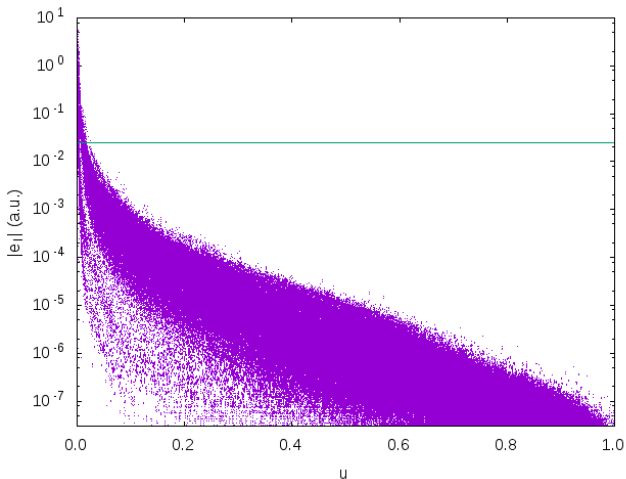


F_2 , cc-pVQZ, $N_{\text{det}} = 5 \times 10^6$



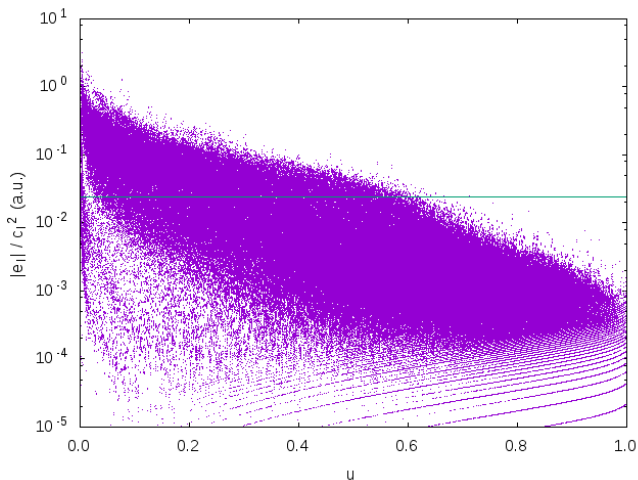
F_2 , cc-pVDZ, $N_{\text{det}} = 10^6$

Uniform sampling : $p_I = 1/N_{\text{det}}$, $E_{\text{PT2}} = \langle N_{\text{det}} \times \epsilon_I \rangle_{p_I}$



F_2 , cc-pVDZ, $N_{\text{det}} = 10^6$

Improved sampling : $p_I = c_I^2$, $E_{\text{PT2}} = \langle \epsilon_I / p_I \rangle_{p_I}$



Variance reduction

Statistical comb

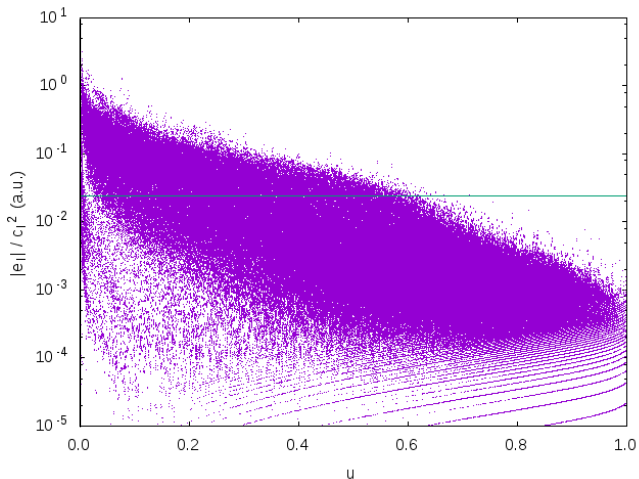
- **Statistical** comb: split \mathcal{D} into M equiprobable sets

$$E_{\text{PT2}} = \sum_{I \in \mathcal{D}} \epsilon_I = \sum_{k=1}^M \sum_{I_k \in \mathcal{D}_k} \epsilon_{I_k}$$

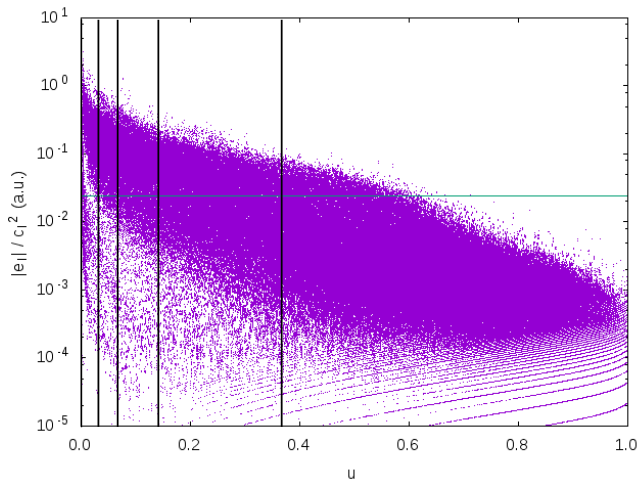
- New Monte Carlo estimator:

$$E_{\text{PT2}} = \left\langle \frac{1}{M} \sum_{k=1}^M \frac{\epsilon_{I_k}}{p_{I_k}} \right\rangle_{(p_{I_1}, \dots, p_{I_M})}$$

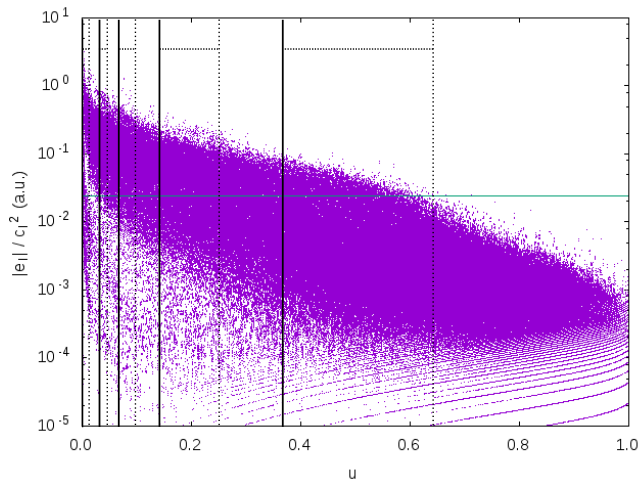
F_2 , cc-pVDZ, $N_{\text{det}} = 10^6$



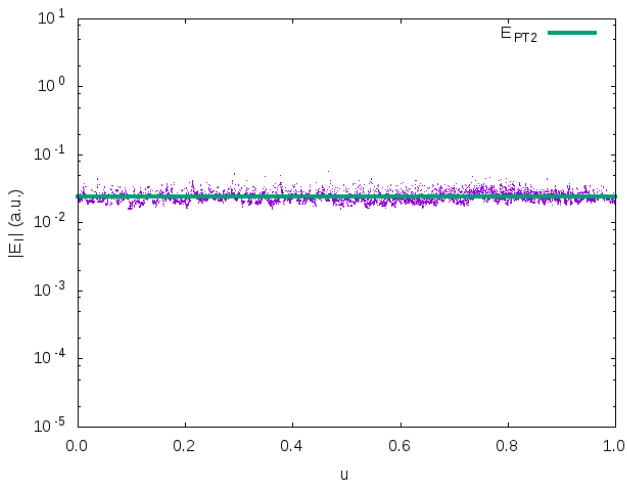
F_2 , cc-pVDZ, $N_{\text{det}} = 10^6$



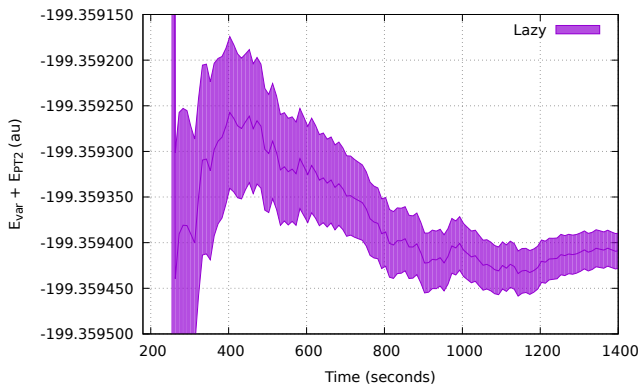
F_2 , cc-pVDZ, $N_{\text{det}} = 10^6$



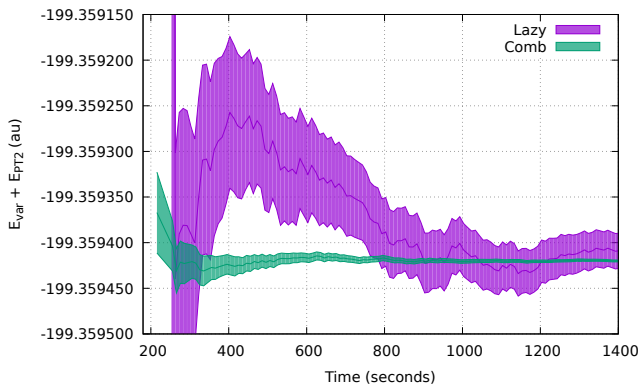
F_2 , cc-pVDZ, $N_{\text{det}} = 10^6$



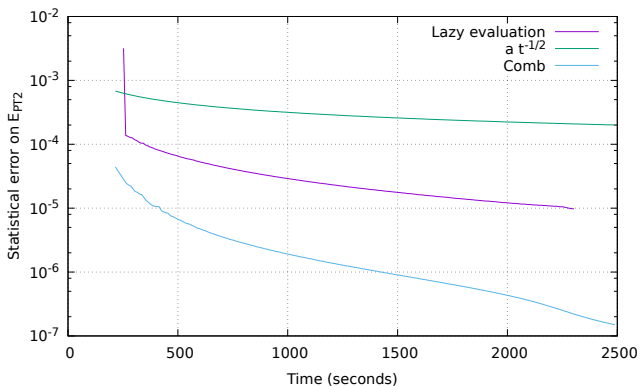
F_2 , cc-pVDZ, $N_{\text{det}} = 10^6$



F_2 , cc-pVDZ, $N_{\text{det}} = 10^6$



F_2 , cc-pVDZ, $N_{\text{det}} = 10^6$



Hybrid deterministic/stochastic scheme

Hybrid strategy

- When all $|I\rangle$'s have been drawn, the **exact** (i.e. deterministic) E_{PT2} value is obtained
- The exact result (i.e. with no statistical error) can be reached in a finite time
- Usually, we recover **90% of the wave function norm** with few determinants
- Compute the few first contributions ϵ_I , and perform MC for the rest

$$E_{PT2} = \sum_{I \in \mathcal{D}_D} \epsilon_I + \left\langle \frac{1}{M} \sum_{k=1}^M \frac{\epsilon_{I_k}}{p_{I_k}} \right\rangle_{(p_{I \in \mathcal{D}_S})}$$

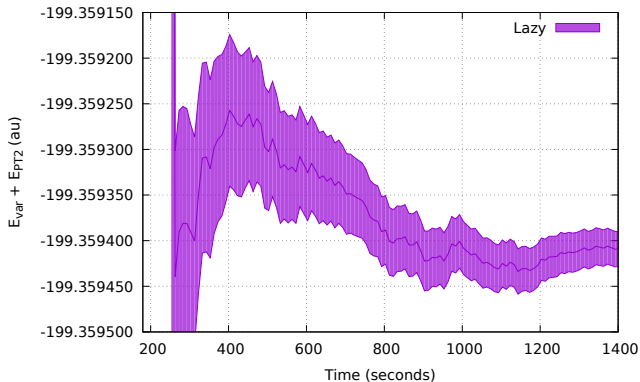
Hybrid deterministic/stochastic scheme

At each MC step,

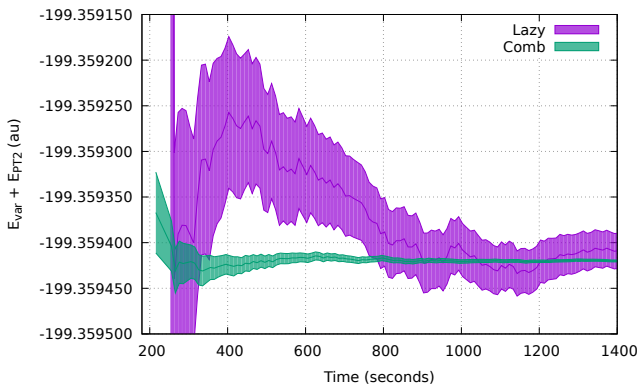
- Draw a random number
- Find $|I\rangle$'s selected by comb and increment n_I 's
- Compute ϵ_I (if required)
- **Daemon**: compute the first non-computed ϵ_I in the determinant list
- If a box is “full” \Rightarrow move it to deterministic space
- At each time t , we have

$$E_{\text{PT2}}(t) = \sum_{I \in \mathcal{D}_D(t)} \epsilon_I + \sum_{I \in \mathcal{D}_S(t)} \frac{1}{M(t)} \frac{n_I(t)}{N_{\text{samples}}(t)} \frac{\epsilon_I}{p_I}$$

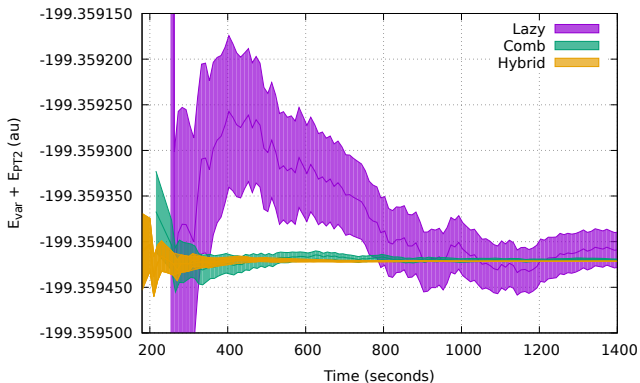
F_2 , cc-pVDZ, $N_{\text{det}} = 10^6$



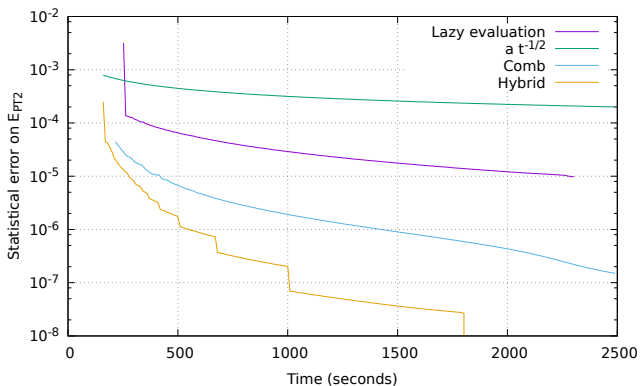
F_2 , cc-pVDZ, $N_{\text{det}} = 10^6$



F_2 , cc-pVDZ, $N_{\text{det}} = 10^6$



F_2 , cc-pVDZ, $N_{\text{det}} = 10^6$



F₂ molecule at equilibrium geometry

System specification

- 16 correlated electrons with RHF MOs
- cc-pVDZ: 28 MOs
- cc-pVTZ: 60 MOs
- cc-pVQZ: 110 MOs

Results

	N_{det}	E_{var}	$E_{\text{var}} + E_{\text{PT2}}$	E_{PT2}
cc-pVDZ	2×10^6	-199.098 015	-199.099 412 -199.099 41(9)	-0.001 397
cc-pVTZ	2×10^6	-199.286 288	-199.298 119(1) -199.297 7(1)	-0.011 831(1)
cc-pVQZ	1×10^7	-199.349 290	-199.361 355(1) -199.359 8(2)	-0.012 065(1)

In gray: i-FCI-QMC results of Cleland *et al*, JCTC 2012

Chromium dimer Cr_2 at equilibrium geometry

System specification

- 38 correlated electrons (10 frozen MOs)
- cc-pVDZ: 86 MOs
- cc-pVTZ: 136 MOs
- cc-pVQZ: 186 MOs
- CASSCF (12,12) MOs
- Selected FCI: $N_{\text{det}} = 2 \times 10^7$

Chromium dimer Cr_2 , $N_{\text{det}} = 2 \times 10^7$, 800 CPU cores

Timing

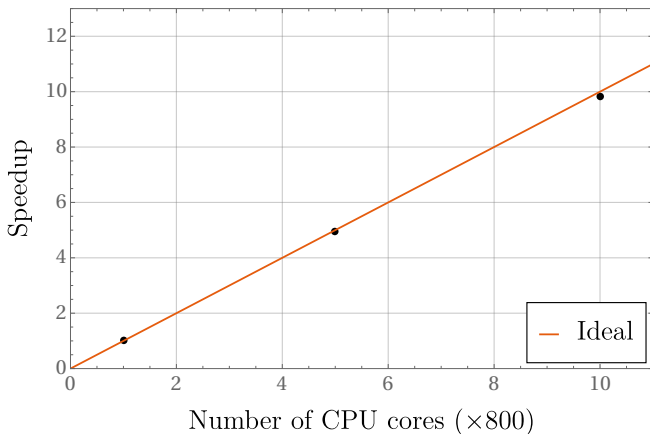
Basis	E_{PT2}	Walltime
cc-pVDZ	-0.068 3(1)	14 min
	-0.068 36(1)	55 min
	-0.068 361(1)	2.4 hr
	-0.068 360 604	3 hr
cc-pVTZ	-0.124 4(5)	19 min
	-0.124 7(1)	58 min
	-0.124 63(1)	3.5 hr
	-0.124 642(1)	8.7 hr
	—	~ 15 hr (estimated)
cc-pVQZ	-0.155 8(5)	56 min
	-0.155 9(1)	2.5 hr
	-0.155 95(1)	9.0 hr
	-0.155 952(1)	18.5 hr
	—	~ 29 hr (estimated)

Chromium dimer Cr_2 , $N_{\text{det}} = 2 \times 10^7$

CIPSI with CAS, MR-CISD and FCI

	N_{det}	E_{var}	$E_{\text{var}} + E_{\text{PT2}}$	E_{PT2}
cc-pVDZ				
CAS(12,12)	3×10^5	-2086.650 896	-2087.287 47(5)	-0.636 57
MR-CISD	2×10^6	-2087.156 659	-2087.273 230(4)	-0.116 571
FCI	6×10^7	-2087.244 972	-2087.298 530	-0.053 558
cc-pVTZ				
CAS(12,12)	5×10^5	-2086.655 594	-2087.588 64(5)	-0.933 05(5)
FCI	2×10^7	-2087.449 781	-2087.574 423(1)	-0.124 676(1)
cc-pVQZ				
CAS(12,12)	6×10^5	-2086.643 210	-2087.690 55(4)	-1.047 34(4)
FCI	2×10^7	-2087.513 373	-2087.669 330(1)	-0.155 957(1)

Parallel efficiency: $Cr_2/cc\text{-pVQZ}$, $N_{\text{det}} = 2 \times 10^7$



Perspectives

- 1 Large systems: use JM-MRPT2 (Giner *et al*) instead of Epstein-Nesbet
 - Needs a partition of the MO space (CAS)
 - Perturbbers are Slater determinants (decontracted formalism)
 - Size-consistent
 - Less sophisticated than NEVPT2, but of comparable quality
- 2 CIPSI
 - Speed-up selection
 - Stochastic Shifted- B_k method
- 3 Stochastic multi-reference coupled cluster
 - Use same algorithm for triples and quadruples
- 4 Parallelism
 - Target: 10^6 CPU cores
 - Hybrid CPU/GPU

Section 4

FCI-F12

Back to slide number 1...

Trial wave function for QMC

$$\Psi_T(\mathbf{R}) = e^{J(\mathbf{R})} \sum_I c_I D_i^\uparrow(\mathbf{R}^\uparrow) D_i^\downarrow(\mathbf{R}^\downarrow)$$

- The **multideterminant part** is obtained via the (selected FCI) **CIPSI algorithm**
Giner et al. JCP 142 (2015) 044115
- We use a “minimal” (nodeless) **Jastrow $J(\mathbf{R})$**

What we don't want...

- ☹️ redundant work between **Jastrow** and **multideterminant part**
- ☹️ chemical intuition and/or user input
- ☹️☹️ stochastic optimization

What we want...

- 😊 a systematic, black-box procedure
- 😊 compact FCI expansions
- 😊 the exact (non-relativistic) energy, i.e. minimizing the **fixed-node error**
- 😊😊 a massively parallel implementation

⇒ **Explicitly-correlated FCI (selected if possible), i.e. FCI-F12 method**
Booth et al. JCP 137 (2012) 164112; Kersten et al. JCP 145 (2016) 054117

Wave function ansatz

$$|\Psi\rangle = \underbrace{|D\rangle}_{\text{conventional FCI}} + \underbrace{|F\rangle}_{\text{explicitly correlated}}$$

$$|D\rangle = \sum_I c_I |I\rangle$$

$$|F\rangle = \sum_I t_I \hat{Q} f |I\rangle$$

$$\text{projector: } \hat{Q} = \hat{1} - \sum_I |I\rangle\langle I|$$

$$\text{correlation factor: } f = \sum_{i<j} f_{ij}$$

Shiozaki et al. JCP 134 (2011) 034113

Correlation factor

Slater geminal fitted with Gaussian geminals

$$f_{12} = \frac{1 - \exp(-\lambda r_{12})}{\lambda} = r_{12} + \mathcal{O}(r_{12}^2)$$

$$\underbrace{\exp(-\lambda r_{12})}_{\text{Slater geminal}} \approx \sum_{\nu=1}^{N_{\text{GG}}} a_{\nu} \underbrace{\exp(-\lambda_{\nu} r_{12}^2)}_{\text{Gaussian geminal}}$$

Ten-no, CPL 398 (2004) 56; Tew & Klopper, JCP 123 (2005) 074101

Effective Hamiltonian theory: dressing the FCI matrix

$$\hat{H}|\Psi\rangle = E|\Psi\rangle \xrightarrow{\langle I|\times} c_I \left[H_{II} + c_I^{-1} \langle I|\hat{H}|F\rangle - E \right] + \sum_{J \neq I} c_J H_{IJ} = 0.$$

$$\hat{H}_{IJ} = \begin{cases} H_{II} + c_I^{-1} \langle I|\hat{H}|F\rangle, & \text{if } I = J, \\ H_{IJ}, & \text{otherwise.} \end{cases}$$

Dressing term

$$\langle I|\hat{H}|F\rangle = \sum_J t_J \left[\underbrace{\langle I|\hat{H}f|J\rangle - \sum_K H_{IK} f_{KJ}}_{\text{incomplete basis set correction}} \right],$$

⇒ The amplitudes t_I 's are obtained to satisfy the e-e cusp conditions
 Ten-no, JCP 121 (2004) 117

Three- and four-electron integrals over Gaussian geminals

THE JOURNAL OF CHEMICAL PHYSICS 147, 024103 (2017)

Three- and four-electron integrals involving Gaussian geminals: Fundamental integrals, upper bounds, and recurrence relations

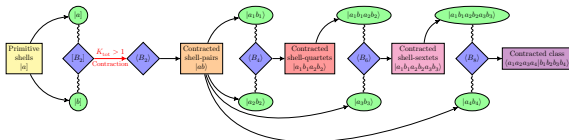
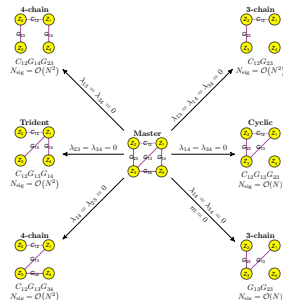
Giuseppe M. J. Barca¹ and Pierre-François Loos^{1,2,a}

¹Research School of Chemistry, Australian National University, Canberra, ACT 2601, Australia

²Laboratoire de Chimie et Physique Quantiques, Université de Toulouse, CNRS, UPS, Toulouse, France

(Received 26 April 2017; accepted 22 June 2017; published online 11 July 2017)

We report the three main ingredients to calculate three- and four-electron integrals over Gaussian basis functions involving Gaussian geminal operators: fundamental integrals, upper bounds, and recurrence relations. In particular, we consider the three- and four-electron integrals that may arise in explicitly correlated F12 methods. A straightforward method to obtain the fundamental integrals is given. We derive vertical, transfer, and horizontal recurrence relations to build up angular momentum over the centers. Strong, simple, and scaling-consistent upper bounds are also reported. This latest ingredient allows us to compute only the $\mathcal{O}(N^2)$ significant three- and four-electron integrals, avoiding the computation of the very large number of negligible integrals. *Published by AIP Publishing.* [<http://dx.doi.org/10.1063/1.4991733>]



Barca, Loos & Gill, in preparation

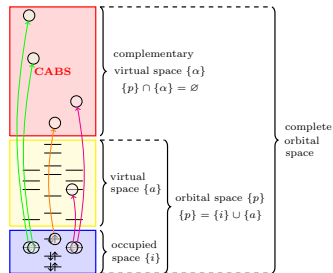
Resolution of the identity (RI) approximation

RI approximation

$$\hat{1} = \sum_{A \in \mathcal{A}} |A\rangle\langle A|$$

Dressing term

$$\begin{aligned} \langle I | \hat{H} | F \rangle &= \sum_J t_J \left[\sum_{A \in \mathcal{A}} H_{IA} f_{AJ} - \sum_{K \in \mathcal{D}} H_{IK} f_{KJ} \right] \\ &= \sum_J t_J \sum_{X \in \mathcal{C}} H_{IX} f_{XJ}, \end{aligned}$$



Valeev, CPL 395 (2004) 190

Self-consistent dressing of the FCI matrix

```

1: procedure DRESSED FCI-F12
2:
3:   Perform a (selected) FCI calculation to get  $|D\rangle = \sum_I c_I |I\rangle$ 
4:   Compute extra integrals, e.g.  $\langle I|\hat{H}f|J\rangle$ 
5:
6:   while  $|\Delta E| > \tau$  do
7:     Determine  $t_I$ 's to satisfy e-e cusp conditions
8:     Form the dressed Hamiltonian  $\hat{H}$ 
9:     Diagonalize  $\hat{H}$  to get a new set of  $c_I$ 's and energy  $E$ 
10:  end while
11:
12:  Return useful quantities for QMC, i.e.  $c_I$ 's and  $t_I$ 's
13:
14: end procedure

```

⇒ It can be “embedded” in the Davidson diagonalization

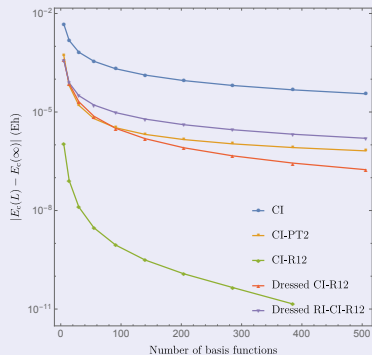
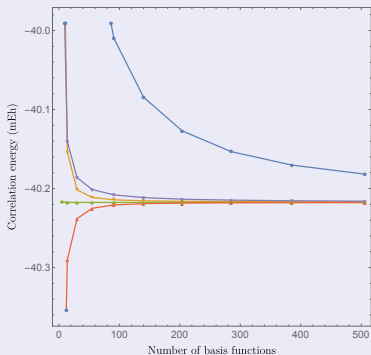
Illustrative calculations on a two-electron system

Paradigm: *"Two electrons on a glome"*

- Singlet pair of electrons on a surface of a 3-sphere of radius $R \Rightarrow$ **uniform density!**
Loos & Gill PRL 103 (2009) 123008; JCP 135 (2011) 214111
- Everything can be done analytically \Rightarrow **great to test approximation(s)!**
Loos & Gill PRA 79 (2009) 062517
- For $R = 1$, $E_c \approx -40.2 \text{ mE}_h \Rightarrow$ **similar to He!**
Loos & Gill PRL 105 (2010) 113001; CPL 500 (2010) 1

Illustrative calculations on a two-electron system

Convergence of the correlation energy: $R = 1$, $L_{RI} = 3L$ and $f_{12} = r_{12}$



Concluding remarks

Things to explore further...

- Design a selected version of the algorithm
- Three-electron integrals or RI tricks?
- More generally, **how do we get the nodes right?**

Future directions...

- Efficient and massively parallel implementation in **QUANTUM PACKAGE**
https://github.com/LCPQ/quantum_package
- We are currently working on a similar methodology to enforce **the electron-nucleus cusp** in QMC calculations
Loos, Scemama & Caffarel, in preparation.

Section 5

Many-electron integrals

Why do we want to compute these?

Potentially useful in many cases

- Gaussian-type geminals (GTGs) calculations
Szalewicz & Jeziorski, Mol. Phys. 108, 3191 (2010)
Komornicki & King, JCP 134 (2011) 244115
- GTGs and Gaussian-type orbitals (GTOs)
Persson & Taylor, JCP 105 (1996) 5915; TCA 97 (1997) 240
- Transcorrelated-type methods
Ten-no, CPL 330 (2000) 175
- Explicitly-correlated F12 methods: MP2-F12, CC-F12, etc.
(RI world)
Hattig et al. Chem Rev 112 (2012) 4
Ten-no & Noga, WIREs Comput Mol Sci 2 (2012) 114

What do we want to calculate?

Three-electron integral

$$\begin{aligned} \langle \mathbf{a}_1 \mathbf{a}_2 \mathbf{a}_3 | \mathbf{b}_1 \mathbf{b}_2 \mathbf{b}_3 \rangle &\equiv \langle \mathbf{a}_1 \mathbf{a}_2 \mathbf{a}_3 | f_{123} | \mathbf{b}_1 \mathbf{b}_2 \mathbf{b}_3 \rangle \\ &= \iiint \psi_{\mathbf{a}_1}^{\mathbf{A}_1}(\mathbf{r}_1) \psi_{\mathbf{a}_2}^{\mathbf{A}_2}(\mathbf{r}_2) \psi_{\mathbf{a}_3}^{\mathbf{A}_3}(\mathbf{r}_3) f_{123} \psi_{\mathbf{b}_1}^{\mathbf{B}_1}(\mathbf{r}_1) \psi_{\mathbf{b}_2}^{\mathbf{B}_2}(\mathbf{r}_2) \psi_{\mathbf{b}_3}^{\mathbf{B}_3}(\mathbf{r}_3) d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 \end{aligned}$$

Gaussian-type orbital (GTO)

$$\text{Primitive GTO} = |\mathbf{a}\rangle = (x - A_x)^{a_x} (y - A_y)^{a_y} (z - A_z)^{a_z} e^{-\alpha |\mathbf{r} - \mathbf{A}|^2}$$

$$\text{Contracted GTO} = |\mathbf{a}\rangle \equiv \psi_{\mathbf{a}}^{\mathbf{A}}(\mathbf{r}) = \sum_i^K D_i |\mathbf{a}\rangle_i$$

- Exponent α
- Center $\mathbf{A} = (A_x, A_y, A_z)$
- Angular momentum $\mathbf{a} = (a_x, a_y, a_z)$ and total angular momentum $a = a_x + a_y + a_z$

Three-electron operator

Three-electron operators

$$f_{123} = f_{12}g_{13}h_{23}$$

Two-electron operators

- Long-range Coulomb operator

$$C_{12} = r_{12}^{-1}$$

- Short-range Slater geminal

$$S_{12} = \exp(-\lambda r_{12})$$

- Short-range Gaussian geminal

$$G_{12} = \exp(-\lambda r_{12}^2)$$

- Short-range operator

$$E_{12} = r_{12} \operatorname{erfc}(\sqrt{\lambda} r_{12})$$

Asymptotic scaling

Operator		Scaling
Two-electron	Three-electron	
<i>S</i>	<i>SS, SSS, SSL</i>	$\mathcal{O}(N)$
<i>L</i>	<i>SL, SLL</i>	$\mathcal{O}(N^2)$
—	<i>LL, LLL</i>	$\mathcal{O}(N^3)$

S = short range

L = long range

Asymptotic scaling of two-electron integrals

Number of significant two-electron integrals for polyenes

$$N_{\text{sig}} = c N^\alpha$$

Molecule	N	C_{12}		G_{12}	
		N_{sig}	α	N_{sig}	α
propene	12	1 625	—	1 650	—
butadiene	16	5 020	3.9	5 020	3.9
hexatriene	24	24 034	3.9	23 670	3.8
octatetraene	32	63 818	3.4	52 808	2.8
decapentaene	40	119 948	2.8	81 404	1.9
dodecaexaene	48	192 059	2.6	109 965	1.6

Helgaker, Jorgensen & Olsen, Molecular Electronic-Structure Theory

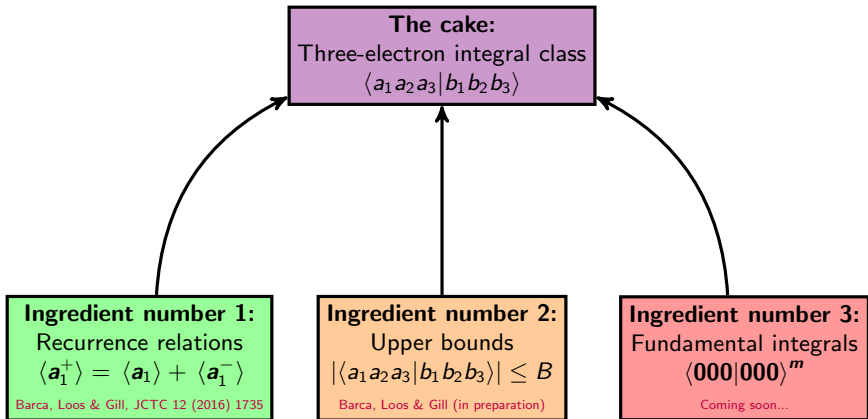
Asymptotic scaling of three-electron integrals

Number of significant three-electron integrals for polyenes

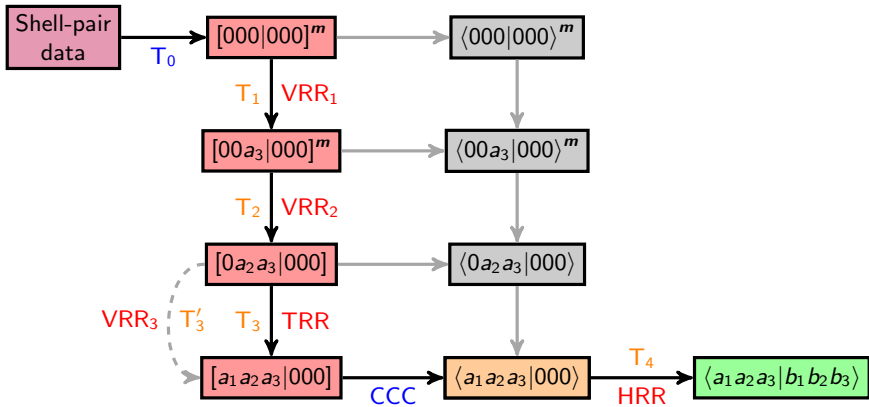
$$N_{\text{sig}} = c N^\alpha$$

1	cMolecule	N	$\mathcal{C}_{12}\mathcal{C}_{13}$		$\mathcal{G}_{12}\mathcal{C}_{13}$		$\mathcal{G}_{12}\mathcal{G}_{13}$	
			N_{sig}	α	N_{sig}	α	N_{sig}	α
	propene	12	123 480	—	243 071	—	123 480	—
	butadiene	16	650 034	5.8	1 288 614	5.8	649 796	5.8
	hexatriene	24	6 259 263	5.6	10 992 400	5.3	4 436 162	4.7
	octatetraene	32	22 875 778	4.5	31 511 030	3.7	9 273 218	2.6
	decapentaene	40	53 576 923	3.8	59 315 069	2.8	14 101 575	1.9
	dodecaexaene	48	101 224 185	3.5	94 176 325	2.5	18 927 362	1.6

Recipe for three-electron integrals



Late-contraction path algorithm (Head-Gordon-Pople & PRISM inspired)



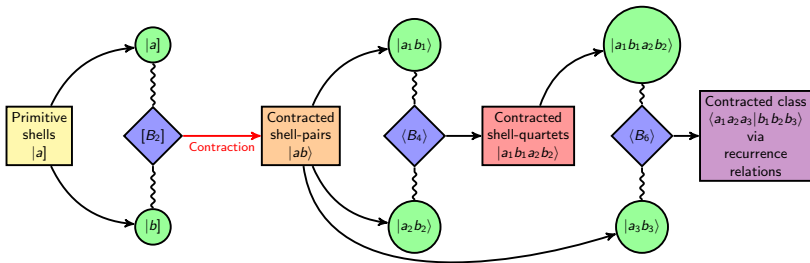
Recurrence relations for three-electron integrals

Number of terms in the recurrence relations for cyclic and chain operators

Step	RR type	Three-electron operator		Two-electron operator
		$f_{12} g_{13} h_{23}$ (cyclic)	$f_{12} g_{13}$ (chain)	f_{12}
T ₁	VRR ₁	8	6	4
T ₂	VRR ₂	10	7	6
T' ₃	VRR ₃	12	12	—
T ₃	TRR	6	6	4
T ₄	HRR	2	2	2

If one uses GTGs, the RRs become even simpler!
Barca & Loos, JCP 147 (2017) 024103

Screening algorithm for three-electron integrals



What is a “good” upper bound?

A good upper bound must be

- **simple**

⇔ significantly less computationally expensive than the true integral

- **strong**

⇔ as close as possible from the true integral in the threshold region
(10^{-14} – 10^{-8})

- **scaling-consistent**

⇔ $N_{\text{sig}} = \mathcal{O}(N_{\text{UB}})$ where N_{UB} is the number estimated by the UB

Shell and shell-pair bounding Gaussians

Shell bounding Gaussian (SBG)

$$\tilde{a}(\mathbf{r}) = \mathcal{N}_a e^{-(1-\sigma)\alpha|\mathbf{r}-\mathbf{A}|^2} \geq \|a\|$$

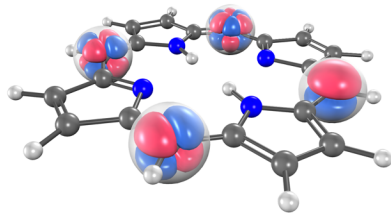
Shell-pair bounding Gaussian (SPBG)

$$\tilde{z}(\mathbf{r}) = \tilde{a}(\mathbf{r})\tilde{b}(\mathbf{r}) = h e^{-\xi|\mathbf{r}-\mathbf{Z}|^2} \geq \|ab\|$$

$$\zeta = \alpha + \beta$$

$$\mathbf{Z} = (\alpha\mathbf{A} + \beta\mathbf{B})/\zeta$$

$$\xi = (1 - \sigma)\zeta$$



© Andrew Gilbert aka Mr. IQMol

A tale of three strategies

■ Integral bound

We seek a single number that bounds a particular integral

$$|[a_1 a_2 | b_1 b_2]| \leq [a_1 a_1 | b_1 b_1]^{1/2} [a_2 a_2 | b_2 b_2]^{1/2}$$

With $\mathcal{O}(N^2)$ $[a_1 a_1 | b_1 b_1]$ bound factors, we can upper bound $\mathcal{O}(N^4)$ $[a_1 a_2 | b_1 b_2]$ integrals

■ Non-factorizable class bound

We seek a single number that bounds an entire class*

$$|[a_1 a_2 | b_1 b_2]| \leq [\check{a}_1 \check{a}_2 | \check{b}_1 \check{b}_2]$$

■ Factorizable class bound

We combine the two previous ideas

$$\begin{aligned} |[a_1 a_2 | b_1 b_2]| &\leq [\check{a}_1 \check{a}_2 | \check{b}_1 \check{b}_2] \\ &\leq [\check{a}_1 \check{a}_1 | \check{b}_1 \check{b}_1]^{1/2} [\check{a}_2 \check{a}_2 | \check{b}_2 \check{b}_2]^{1/2} \end{aligned}$$

*The simple $[ppp|ppp]$ class is made of 729 integrals!

Our home-made bounds

- We always use **class bounds via SPBGs**

$$|[a_1 a_2 a_3 | b_1 b_2 b_3]| \leq [\hat{a}_1 \hat{a}_2 \hat{a}_3 | \hat{b}_1 \hat{b}_2 \hat{b}_3]$$

⇒ σ is optimized with an **operator-specific metric**

- $[B_2]$, $[B_4]$ and $[B_6]$ are derived with **Hölder's inequality**

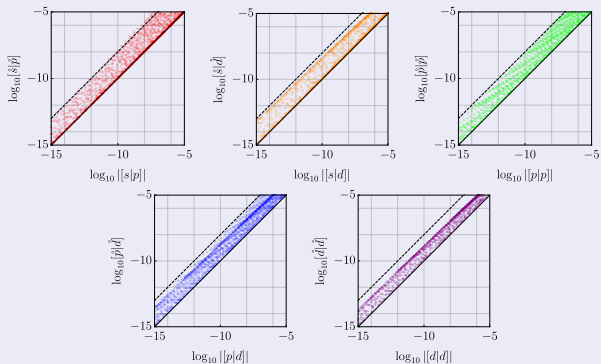
$$\left| \int \phi_1(\mathbf{r}) \phi_2(\mathbf{r}) d\mathbf{r} \right| \leq \left[\int |\phi_1(\mathbf{r})|^p d\mathbf{r} \right]^{1/p} \left[\int |\phi_2(\mathbf{r})|^q d\mathbf{r} \right]^{1/q},$$

for $p^{-1} + q^{-1} = 1$ and $p, q > 1$.

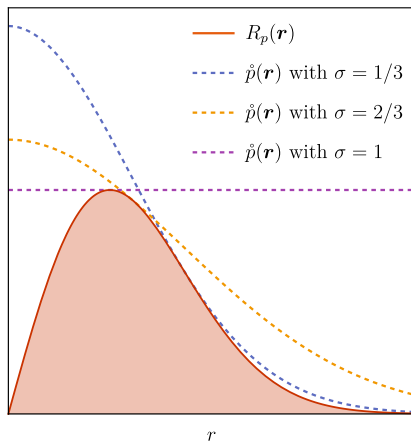
- We are **not** using Cauchy-Schwartz (i.e. $p = q = 2$)!
 - 1 they are usually weaker
 - 2 for three-electron integrals, they are not always simple

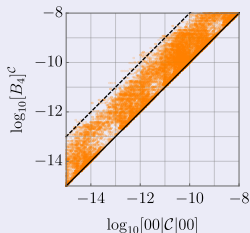
Performance of SPBGs

One-electron overlap integrals: $|[a|b]| \leq [a|b]$

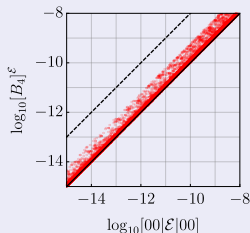


Performance of SPBGs



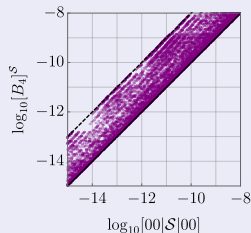
Performance of shell-quartet bounds $[B_4]$ for two-electron integralsLR operator \mathcal{C}_{12} 

(factorizable bound)

NB: tighter than
Cauchy-Schwartz!SR operator \mathcal{E}_{12} 

(non-factorizable bound)

$$\sqrt{\pi\lambda} \frac{\mathcal{E}_{12}}{\mathcal{G}_{12}} \leq 1$$

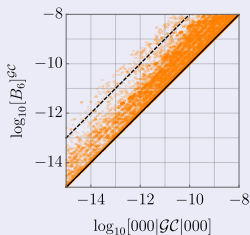
SR operator \mathcal{S}_{12} 

(non-factorizable bound)

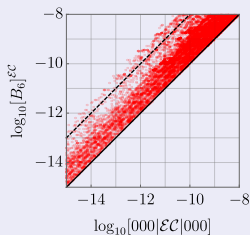
NB: we upper bound the
fundamental integral!

Performance of shell-sextet bounds $[B_6]$ for three-electron integrals

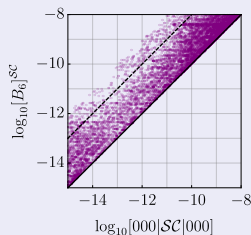
Operator $\mathcal{G}_{12}\mathcal{C}_{13}$



Operator $\mathcal{E}_{12}\mathcal{C}_{13}$



Operator $\mathcal{S}_{12}\mathcal{C}_{13}$



NB: All these integrals scale quadratically unlike $\mathcal{G}_{12}\mathcal{G}_{13}$, $\mathcal{E}_{12}\mathcal{E}_{13}$ or $\mathcal{S}_{12}\mathcal{S}_{13}$

Things to do...

Future work

- Efficient computation of **fundamental integrals**
(Slater-based operators in particular)
- Recurrence relations and upper bounds for **four-electron integrals**
(Barca & Loos, Adv. Quantum Chem. (in press))
- Specific recurrence relations for **GTGs** (three- and four-electron integrals)
(Barca & Loos, JCP 147 (2017) 024103)
- **Implementation** in quantum chemistry softwares
(under progress)