Approximation Methods in Quantum Chemistry: Application to SCF and MP2

Frank Neese

Max Planck Institut für Chemische Energiekonversion Stiftstr. 34-36 D-45470 Mülheim an der Ruhr Frank.Neese@cec.mpg.de



Lecture 1: Algorithm Design

"Getting the exact same number faster"

Lecture 2: Approximation Methods

"Getting Approximately the right number even faster"

Approximations in Quantum Chemistry

Well justified approximation may lead to many orders of magnitude speedups, make impossible calculations possible and may not deteriorate the results

In fact, today there are next to NO quantum chemical calculations done that do not introduce approximations









Corollaries

First your approximation needs to be **ROBUST** and THEN it *can* be efficient!

Approximations should be **well controlled**, e.g. by a **threshold T** such that at T=0 the result is identical with the approximation free result

When you **document speedups**, the proper point of reference is the best alternative method of comparable accuracy and NOT a highly inefficient implementation!

When you document speedups, you *must* **keep the desired target accuracy** and not relax thresholds

You should NOT make general rules based on one calculation on one system. Your statements should be backed up by at least a minimal amount of **statistical analysis** on a range of systems

Target Precision

It is very important that YOU define a **target precision** that - when reached - renders the approximate result useful. You have to **demonstrate** that you reached that goal.

Examples:

- ✓ You approximate a coupled cluster energy and make an error of >100 kcal/mol
- Not useful! Your result is worse than a Hartree-Fock calculation.
- ✓ You approximate excitation energies with an error between 100 cm⁻¹ and 10000 cm⁻¹
- Probably not useful, in particular if the error is inconsistent between states
- ✓ You approximate vibrational frequencies and introduce an error of 10 cm⁻¹
- May or may not be useful. For most everyday applications that is good enough, for high precision studies (e.g. astrophysics) not good enough
- You approximate the total energy and introduce errors of 1-2 kcal/mol in energy differences.
- May or may not be useful. E.g. does it matter whether one isomer is 30.12 or 31.26 kcal/mol higher than another? But if one transition state is 1 kcal/mol higher or lower than another it changes the stereospecificity of the reaction!
- ✓ You calculate the electronic energy to a prevision of 0.03 kcal/mol but you have an error of 5 kcal/mol in the solvation energy or the entropy.
- Your could probably get away with a less accurate electronic energy

Numerical Precision and Thresholding

Machine Precision & Numerical Algorithms

The computer can only represent numbers with finite precision!

Floating point numbers with 8 byte size (double precision) are accurate to an absolute value of $\sim 10^{-16}$

- No quantum chemical result can be more accurate than this!
- In actual calculations there is a propagation of round-off errors which can become problematic!
- Numerical algorithms have to be chosen carefully
- ✓ BLAS operations are uncritical

√...

- ✓ **Eigenvalue** solvers (e.g. LAPACK) become inaccurate for large matrices
- Gram-Schmidt orthogonalization is problematic and singular value decomposition is preferred.
- ✓ Matrix inversion is numerically unstable and should be replaced by diagonalization, SVD, Cholesky decomposition, ... in most if not all cases the explicit construction of the inverse matrix is not necessary!

Example: Near Linearly Dependent Basis Sets



Benzene /aug'-cc-pVDZ

(additional diffuse s-functions down to exponent 0.0001)

- 246 basis functions
- Smallest eigenvalue of the overlap matrix 8.5x10⁻¹⁶

We can remove the linearly dependent eigenvalues and vectors according to a threshold $\boldsymbol{\tau}$

Energy does not really converge before the calculation becomes numerically unstable

Some basis sets can not be used to obtain converged results

 τ =10⁻⁸-10⁻¹⁰ gives about 7 significant digits

10-439-230.725 367 7410-616-230.725 957 2110-89-230.725 982 27
10 ⁻⁶ 16 -230.725 957 21 10 ⁻⁸ 9 -230.725 982 27
10 ⁻⁸ 9 -230.725 982 <u>27</u>
10 ⁻¹⁰ 5 -230.725 982 <u>44</u>
10 ⁻¹¹ 4 -230.725 982 4 <u>5</u>
<10 ⁻¹² 0 explodes

Self Consistent Field

$$\mathbf{F}(\mathbf{c})\mathbf{c}_i = \varepsilon_i \, \mathbf{S} \mathbf{c}_i$$

$$S_{\mu\nu} = \left\langle \mu \,|\, \nu \right\rangle$$





"Mean Field" Hartree-Fock

$$F_{\mu\nu} = h_{\mu\nu} + \sum_{\kappa\tau} P_{\kappa\tau}(\mu\nu \mid \kappa\tau) - \frac{c_{x}}{2} \sum_{\kappa\tau} P_{\kappa\tau}(\mu\kappa \mid \nu\tau) + \int \mu(\mathbf{r}) \frac{\delta E_{XC}[\rho]}{\underbrace{\delta\rho(\mathbf{r})}_{V_{XC}(\mathbf{r})}} \nu(\mathbf{r}) d\mathbf{r}$$

One Electron

Coulomb

HF Exchange

XC Potential

Typical Timings of a Traditional HF Calculation

Not converged



Reasonable conclusion: try to minimize integral calculation time

Almlöf's Revolutionary Proposals

For several decades progress in quantum chemistry was prevented by the large number of two-electron integrals that need to be calculated.



Even if the integrals can be stored for a 1000 basis function calculation, the I/O penalty is huge and the CPU remains largely idle while waiting for data to arrive from the hard drive

The integral bottleneck was finally overcome by Almlöfs revolutionary proposals

- 1. Do not store integrals. Recalculate when needed (direct SCF)
- 2. Split the calculation of the Coulomb and exchange terms and use the most efficient approximation for each rather than use the same integrals for both.

Negligible Integrals: Gaussian Product Theorem





In a large system there are only O(N), significant' Gaussian products

These significant products interact via an 1/r operator (which never becomes small). *Hence there are O(N²) non-negligible integrals*

Integral Estimates

In order to skip negligible contributions, we must be able to recognize small integrals **BEFORE** we calculate them

Almlöf originally used the prefactor of the Gaussian product theorem. However, Häser and Ahlrichs used the Schwartz inequality to show:

$$\left| (\mu_{A} \nu_{B} \mid \kappa_{C} \tau_{D}) \right| \leq \sqrt{(\mu_{A} \nu_{B} \mid \mu_{A} \nu_{B})} \sqrt{(\kappa_{C} \tau_{D} \mid \kappa_{C} \tau_{D})}$$



Cheap upper bound! Universally used.



Problem: does not take into account the multipolar structure of the charge distributions and hence may overestimate the integral strongly



Improved upper bounds can be formulated (Gill; Ochsenfeld and coworkers)

Kohn's Conjecture and the Density Matrix



The decay is exponential, but slow. 10⁻¹⁰ is only reached at 20-25 Angström!

Intrinsic Scaling of Coulomb and Exchange

Assuming exponential decay of the density, Almlöf realized that the intrinsic scaling of the Coulomb and exchange terms is different:

Coulomb:





Conclusion: Use the most efficient way to calculate or approximation each term separately!

Typical Observed Scaling



System Size



Conclusion: Coulomb dominates quickly. True asymptotic scaling reached relatively late w.r.t. system size

Fock-Matrix Construction and Pre-Screening

$$\begin{split} F_{\mu\nu} &= h_{\mu\nu} + G_{\mu\nu} \quad G_{\mu\nu} = \sum_{\kappa\tau} P_{\kappa\tau} \left[\left(\mu\nu \mid \kappa\tau \right) - \left(\mu\kappa \mid \nu\tau \right) \right] \\ & \text{G=0} \\ & \text{loop } \mu \\ & \text{loop } \kappa \\ & \text{loop } \tau \leq \kappa \left(\mu\nu \leq \kappa\tau \right) \\ & \text{test=sqrt} \left(\mu\nu \mid \mu\nu \right) \text{sqrt} \left(\kappa\tau \mid \kappa\tau \right) \max \left(\mathbf{P} \left(\kappa, \tau \right) , \mathbf{P} \left(\nu, \tau \right) \right) , \dots \\ & \text{if } \left(\text{test} < \underline{\text{Thresh}} \right) \quad \text{skip} \\ & \text{else} \\ & \text{Calculate } \left(\mu\nu \mid \kappa\tau \right) \\ & \text{add } G \left(\mu, \nu \right) + = \mathbf{P} \left(\kappa, \tau \right) \left(\mu\nu \mid \kappa\tau \right) \quad (\text{Coulomb}) \\ & \text{add } G \left(\mu, \kappa \right) - = \mathbf{P} \left(\nu, \tau \right) \left(\mu\nu \mid \kappa\tau \right) \quad (\text{Exchange}) \\ & \quad (\text{and permutations of indices}) \\ & \text{end } (\text{else}) \end{split}$$

end loops $\mu,\!\nu,\,\kappa,$



Only contributions >= Thresh go into the Fock matrix Better than testing for small integrals alone since P can be large

Recursive Fock-Matrix Construction

Subtract the Fock matrices of two iterations from each other:

$$F_{\mu\nu}^{(n+1)} - F_{\mu\nu}^{(n+1)} = \sum_{\kappa\tau} \underbrace{\left(P_{\kappa\tau}^{(n+1)} - P_{\kappa\tau}^{(n)}\right)}_{\Delta P_{\kappa\tau}} \left[(\mu\nu \mid \kappa\tau) - (\mu\kappa \mid \nu\tau) \right]$$

$$F_{\mu\nu}^{(n+1)} = F_{\mu\nu}^{(n+1)} + \sum_{\kappa\tau} \Delta P_{\kappa\tau} \left[(\mu\nu \mid \kappa\tau) - (\mu\kappa \mid \nu\tau) \right]$$

Giving



Prescreening becomes more efficient as the calculation is approaching convergence as ΔP is approaching zero



However: Recursive build leads to error accumulation and hence needs to be restarted every ~10 SCF cycles

Typical Timings

Total SCF time: 0 days 2 h	ours 13	min 19 sec			
Total time	• • • •	7999.924	sec		
Sum of individual times	••••	7998.252	sec	(1	.00.0%)
Fock matrix formation		7990.164	sec	(99.9 %)
Diagonalization	• • • •	0.984	sec	(0.0%)
Density matrix formation		0.320	sec	(0.0%)
Population analysis	• • • •	0.845	sec	(0.0%)
Initial guess	• • • •	2.892	sec	(0.0%)
Orbital Transformation	• • • •	0.000	sec	(0.0%)
Orbital Orthonormalization	• • • •	0.000	sec	(0.0%)
DIIS solution	• • • •	1.220	sec	(0.0%)
SOSCF solution	• • • •	1.015	sec	(0.0%)

Iteration	Time	for	Fock	: (s	sec)	1
1		522				_
2		436				
3		447				
9		278				
10		268				
11		253				
(*) TConv=10 ⁻	⁻¹⁰ , Tł	nresh	n 10-	13		
(**) ~600 bas	sis fu	ıncti	ons	up	to	f

- Overall timing strongly dominated by integral evaluation.
- Incremental build leads to speedups
 ~2 in later iterations

Direct SCF Precision

The first decision is when to consider a calculation to be converged and how to check on convergence! (total energy, density, error, orbital gradient, orbital rotation angles, ...).

Pragmatic: total energy change $< T_{conv}$ and one electron energy change $< 10^{3}T_{conv}$ *(TurboMole)*





Good: Error in Energy is always slightly lower than the requested tolerance!

Influence of the Fock Matrix Neglect Threshold



The neglect threshold should be smaller than convergence threshold! (e.g. in ORCA 0.01*T_{Conv} is enforced). Typical accuracy in SCF energies 10⁻⁷-10⁻¹² Eh

Balancing Cost and Accuracy



✓ How much does improved accuracy cost in terms of computer time?

T _{Conv} (Eh)	Thresh (Eh)	#(Iter)	Time (sec)	Energy (Eh)	Error (Eh)
10 -6	10 ⁻¹⁰	11	3391	x.991 303 <u>8</u> 90 741	5,4x10 ⁻⁸
10 ⁻⁷	1 0 ⁻¹¹	12	3678	x.991 303 94 <u>1</u> 877	3,1x10 ⁻⁹
10 -8	1 0 ⁻¹¹	13	4542	x.991 303 94 <mark>8</mark> 517	3,5x10 ⁻⁹
10 -9	10 ⁻¹²	16	5759	x.991 303 94 <mark>4</mark> 568	4,4x10 ⁻¹⁰
10 ⁻¹⁰	10 ⁻¹³	18	7813	x.991 303 94 <u>4</u> 971	4,2x10 ⁻¹¹
10 ⁻¹¹	10 ⁻¹⁴	18	9574	x.991 303 945 0 <u>0</u> 4	9,0x10 ⁻¹²
10 ⁻¹²	10 ⁻¹⁵	20	11060	x.991 303 945 013	-

SCF Convergence and Frequencies

- ✓ You may rightfully argue that for thermochemistry any accuracy in the SCF energies better than, say, 0.01 kcal/mol (e.g. 10⁻⁵ Eh) is good enough.
- ✓ However, SCF accuracy becomes a bigger issue for each additional derivative your taking!
- How accurate do you need your SCF to be for analytic frequencies to be accurate to 1 cm⁻¹?

	T _{conv} (Eh)	$\Delta\omega_{\max}$ (cm ⁻¹)	$\Delta\omega_{ m RMS}$ (cm ⁻¹)
	10-6	4.41	0.83
	10-7	3.92	0.41
	10 ⁻⁸	0.68	0.05
	10-9	0.67	0.05
	10 ⁻¹⁰	0.41	0.03

- Convergence to at least 10⁻⁸ Eh is necessary to have frequencies accurate to 1 cm⁻¹
- D. Bykov, T. Petrenko, R. Izsak, S. Kossmann, U. Becker, E. Valeev, FN Molecular Physics 113 (2015) 1961.

Quantum Chemical Approximations

- 1. Numerical Integration
- 2. Auxiliary Basis Set Expansion
- 3. Semi-Numerical Techniques
- 4. Multipole Techniques

Numerical Integration

Numerical Integration in Quantum Chemistry

Numerical integration in general:

$$\int f(x) \, dx \approx \sum_g w_g f(x_g)$$
 Weights ,Roots'

Very computer friendly, highly parallelizable, exposes locality clearly

For **Atoms**:

Make use of polar coordinates r, θ, ϕ :

$$x = r \sin \theta \sin \phi$$
$$y = r \sin \theta \cos \phi$$
$$z = r \cos \theta$$



Χ

Numerical integration divided into angular and radial integration:

$$\int f(\mathbf{r}) d\mathbf{r} \approx \sum_{g_r} \mathbf{w}_{g_r} \sum_{g_a} \mathbf{w}_{g_a} f(\mathbf{r}_{g_r}, \mathbf{\theta}_{g_a}, \mathbf{\phi}_{g_a})$$

Numerical Integration: Radial Integration

Radial integration

- Pretty much anything works.
- ✓ Need to map interval 0..infinity to something 0..1

✓ Gauss-Chebyshev roots and weights:

$$\begin{split} x_{g} &= \frac{n_{r} + 1 - 2g}{n_{r} + 1} + \frac{2}{\pi} \Biggl\{ 1 + \frac{2}{3} \sin \Biggl(\frac{g\pi}{n_{r} + 1} \Biggr)^{2} \Biggr\} \sin \Biggl(\frac{g\pi}{n_{r} + 1} \Biggr) \cos \Biggl(\frac{g\pi}{n_{r} + 1} \Biggr) \\ w_{g_{r}} &= \frac{16}{3 \Bigl(\frac{n_{r}}{n_{r} + 1} \Bigr)} \sin \Biggl(\frac{g\pi}{n_{r} + 1} \Biggr)^{4} \frac{\xi_{A}}{\ln \Bigl(2 \Bigr) \Bigl(1 - x_{g} \Bigr)} r_{g}^{2} \end{split}$$

 $n_r = 15\epsilon + 5r - 40$ r=row of periodic table ϵ =requested accuracy

,Atomic size'

(~60 points for 6 digit accuracy)

of radial points

 ξ_{A}

 n_r



Attention: Very steep primitive basis functions with an ,extent' smaller than radial grid point spacing cause deep trouble!

Numerical Integration: Angular Integration

✓ Lebedev grids: grids of octahedral symmetry that integrate spherical harmonics up to a given order L_{max} <u>exactly</u>.

N _{Points}	L _{max}
50	11
110	17
194	23
434	35
590	41
770	47



Since atom centered basis functions are spherical harmonics, these grids are optimal

Numerical Integration: Molecular Grids

For molecules, we can not just superimpose atomic grids. Rather we should make sure that the contributions of ,foreign' atoms are damped around a ,reference' atom

Voronoi cells

Integral:
$$I = \int F(\mathbf{r}) d\mathbf{r} \longrightarrow I = \sum_{A} I_{A} = \sum_{A} \int F_{A}(\mathbf{r}) d\mathbf{r}$$
 $F_{A}(\mathbf{r}) = p_{A}(\mathbf{r})F(\mathbf{r})$
Becke: $\mu_{AB}(\mathbf{r}) = (r_{A} - r_{B})R_{AB}^{-1}$ (=1 near A, =-1 near B) $\sum_{A} p_{A}(\mathbf{r}) = 1$
Switching function: $g(\mu_{AB}) = \frac{3}{2}\mu_{AB}(\mathbf{r}) - \frac{1}{2}\mu_{AB}(\mathbf{r})^{3}$ $f(\mu_{AB}) = g(g(g(\mu_{AB})))$
 $s(\mu_{AB}) = \frac{1}{2}[1 - f(\mu_{AB})]$ $\sum_{B \neq A} s(\mu_{AB}(\mathbf{r}))$ $p_{A}(\mathbf{r}) = \prod_{B \neq A} s(\mu_{AB}(\mathbf{r}))$ $p_{A}(\mathbf{r}) = \frac{P_{A}(\mathbf{r})}{\sum_{B} P_{B}(\mathbf{r})}$ $p_{A}(\mathbf{r}) = \frac{P_{A}(\mathbf{r})}{\sum_{B} P_{B}(\mathbf{r})}$

Numerical Integration: Grid Pruning

- Close to the nucleus, the integrand is typically close to spherical and one gets away with a smaller angular grid
- ✓ Far away from the nucleus the weights of the grid points will be strongly damped and one gets away with a smaller angular grid
- Divide the integration region in 5 sub-areas with a maximal angular integration grid in the bonding region at atomic size parameter R (depending on the row of the periodic table)

Gill's pruning algorithm:

Region 1: 0	- 0.25*R	Grid $L_{minimal}$
Region 2: 0.25*R	- 0.50*R	Grid $L_{Max}-2$
Region 3: 0.50*R	- 1.00*R	Grid $L_{Max}-1$
Region 4: 1.00*R	- 4.50*R	Grid L_{Max}
Region 5: >4.50*R		Grid $L_{Max}-1$

Numerical Integration: Linear Scaling Algorithm

- 1. Organize Grid points in spatially close **batches**
- 2. Find the **spatial extent** of the basis functions

3. Find the contributing basis functions for each grid point batch

$$F_{\mu\nu\in batch} \leftarrow \left\langle \mu \mid V[\rho] \mid \nu \right\rangle \qquad \rho(\mathbf{r}) = \sum_{\mu\nu\in batch} P_{\mu\nu}\mu(\mathbf{r})\nu(\mathbf{r})$$

Set to 0

Set to 0

loop grid point batches
$$O(N)$$
Evaluate basis function values $X_{\mu}(\mathbf{r}_{g})$ constEvaluate density $D(\mathbf{r}_{g}) = \Sigma_{\mu\nu} P_{\mu\nu}X_{\mu}(\mathbf{r}_{g}) X_{\nu}(\mathbf{r}_{g})$ constEvaluate potential $V(\mathbf{r}_{g})$ constLocal Fock matrix contributions $F_{\mu\nu} += V(\mathbf{r}_{g}) X_{\mu}(\mathbf{r}_{g}) X_{\nu}(\mathbf{r}_{g})$ constend loop grid point batches

Simple, convenient, efficient, general, ... much more than just XC integration!



Numerical Integration: Example

L _{Max}	IntAcc	#(Grid)	T (sec)	Error-Density	Error-Energy (Eh)
110	5.34	53046	95	-3,4x10 ⁻³	9,7x10 ⁻⁴
194	5.34	102466	158	+7,7x10 ⁻⁵	8,4x10 ⁻⁵
302	5.67	189941	268	+5,0x10 ⁻⁵	4,0x10 ⁻⁶
434	6.00	327912	432	-3,2x10 ⁻⁶	5,5x10 ⁻⁶
590	6.34	498480	657	-1,2x10 ⁻⁵	4,6x10 ⁻⁶
770	6.67	762322	942	+1,3x10 ⁻⁶	1,0x10 ⁻⁷
ref	15.00	9180185	12580	+4,5x10 ⁻⁷	_

- Error is *roughly* same order of magnitude than error in the density.
- Error is not bounded from above or below, can be positive or negative
- Error is essentially random and does not cancel in energy differences

Auxiliary basis set expansions (RI,DF)

Expansion of basis function products

Basis function products define charge distributions

$$\rho_{\mu_A\nu_B}(\mathbf{r}) = \mu_A(\mathbf{r})\nu_B(\mathbf{r})$$

✓ This is a two-center quantity. We aim to break it down into easier, one-center parts.

Charge distributions have multipole moments. However, rather than performing an actual multipole expansion, we expand the charge distribution in an atom-centered auxiliary basis set



Expansion of Basis Function Products

The expansion coefficients can be determined in many ways. It would be natural to simply minimize:
12

$$R_{\mu_A\nu_B}(\mathbf{r}) = \left[\rho_{\mu_A\nu_B}(\mathbf{r}) - \sum_{K_C} c_{\mu_A\nu_B}^{K_C} \eta_{K_C}(\mathbf{r})\right]$$

The first molecular DFT calculations with Slater basis sets by Baerends and co-workers used such an approach together with constraints that the fitted density integrates to the correct number of electrons

✓ In 1993 Almlöf showed that minimizing the **self-repulsion**:

$$\begin{split} R_{\mu_{A}\nu_{B}} &= \int \int \left\{ \rho_{\mu_{A}\nu_{B}}(\mathbf{r}_{1}) - \sum_{K_{C}} \boldsymbol{c}_{\mu_{A}\nu_{B}}^{K_{C}} \boldsymbol{\eta}_{K_{C}}(\mathbf{r}_{1}) \right\} \frac{1}{\left| \mathbf{r}_{1} - \mathbf{r}_{2} \right|} \left\{ \rho_{\mu_{A}\nu_{B}}(\mathbf{r}_{2}) - \sum_{K_{C}} \boldsymbol{c}_{\mu_{A}\nu_{B}}^{K_{C}} \boldsymbol{\eta}_{K_{C}}(\mathbf{r}_{2}) \right\} d\mathbf{r}_{1} d\mathbf{r}_{2} \\ &\equiv (\mu_{A}\nu_{B} \mid \mu_{A}\nu_{B}) + \sum_{K_{C}L_{D}} \boldsymbol{c}_{\mu_{A}\nu_{B}}^{K_{C}} \boldsymbol{c}_{\mu_{A}\nu_{B}}^{L_{D}} (K_{C} \mid L_{D}) - 2\sum_{K_{C}} \boldsymbol{c}_{\mu_{A}\nu_{B}}^{K_{C}} (\mu_{A}\nu_{B} \mid K_{C}) \end{split}$$

is an order of magnitude more accurate for a given auxiliary basis set.

Junlap analyzed the physical content of that approximation and pointed out that one fits the electric field of the charge distribution

EJ Baerends, DE Ellis P. Ros, Chem. Phys. 2 **1973,** 41; O Vahtras, J Almlöf MW Feyereisen, Chem. Phys. Lett. <u>213</u> **1993,** 514.; BI Dunlap, JWD Connolly JR Sabin, J. Chem. Phys. <u>71</u> **1979**, 3396.; BI Dunlap, J. Mol. Struct. Theochem <u>529</u> **2000,** 37

Expansion of Basis Function Products

The fit is straightforward:

$$\frac{\partial R_{\mu_A \nu_B}}{\partial c_{\mu_A \nu_B}^{K_C}} = 2 \sum_{L_D} c_{\mu_A \nu_B}^{L_D} (K_C \mid L_D) - 2(\mu_A \nu_B \mid K_C) \equiv 0$$

Thus:
$$\mathbf{c}_{\mu_A \nu_B} = \mathbf{V}^{-1} \mathbf{g}_{\mu_A \nu_B}$$

Linear equations

$$V_{_{K_{_{C}}L_{_{D}}}} = (K_{_{C}} \mid L_{_{D}}) \qquad \text{ 2-center repulsion}$$

$$g_{\mu_A \nu_B}^{K_C} = (\mu_A \nu_B \mid L_D)$$

3-center repulsion

EJ Baerends, DE Ellis P. Ros, Chem. Phys. 2 1973, 41; O Vahtras, J Almlöf MW Feyereisen, Chem. Phys. Lett. 213 1993, 514.; BI Dunlap, JWD Connolly JR Sabin, J. Chem. Phys. <u>71</u> 1979, 3396.; BI Dunlap, J. Mol. Struct. Theochem <u>529</u> 2000, 37

How to solve the equations?

If the auxiliary basis is wisely chosen, the inverse of V exists (V is positive definite).

However - matrix inversion is numerically almost never the best choice!

Possibility 1: Diagonalization

$$\begin{aligned} (\mu_{A}\nu_{B} \mid \kappa_{C}\tau_{D}) &\approx \sum_{K_{E}L_{F}} c_{\mu_{A}\nu_{B}}^{K_{E}} (K_{E} \mid L_{F}) c_{\kappa_{C}\tau_{D}}^{L_{F}} \\ &= \sum_{K_{E}K'_{E}L_{F}L'_{F}} (V_{K_{E}K'_{E}}^{-1} (\mu_{A}\nu_{B} \mid K'_{E}) V_{K_{E}L_{F}} V_{L_{F}L'_{F}}^{-1} (\kappa_{C}\tau_{D} \mid L'_{F}) \\ &= \sum_{K'_{E}L'_{F}} (\mu_{A}\nu_{B} \mid K'_{E}) (\mathbf{V}^{-1}\mathbf{V}\mathbf{V}^{-1})_{K'_{E}L'_{F}} (\kappa_{C}\tau_{D} \mid L'_{F}) \\ &= \sum_{K'_{E}L'_{F}} (\mu_{A}\nu_{B} \mid K'_{E}) (\mathbf{V}^{-1/2}\mathbf{V}^{-1/2})_{K'_{E}L'_{F}} (\kappa_{C}\tau_{D} \mid L'_{F}) \\ &= \sum_{K'_{E}} (\mu_{A}\nu_{B} \mid K'_{E}) (\kappa_{C}\tau_{D} \mid \overline{K}_{E}) \end{aligned}$$

$$(\mu_A \nu_B \mid \overline{K}_E) = \sum_{K'_E} (\mu_A \nu_B \mid K'_E) V_{K'_E \overline{K}_E}^{-1/2} \qquad \mathbf{V} \mathbf{U} = \mathbf{v} \mathbf{U} \to \mathbf{V}^{-1/2} = \mathbf{U} \mathbf{v}^{-1/2} \mathbf{U}^T$$

Solving the RI equations

Possibility 2: Cholesky decomposition

Since V is positive definite one can Cholesky decompose it

$\mathbf{V} = \mathbf{L}\mathbf{L}^T$

and use **L** in place of $V^{-1/2}$

One could also use SVD, but it is much slower. Cholesky is the numerically most stable and efficient numerical algorithm

The Cholesky vectors L can also be used to solve linear equations

$$\mathbf{V}\mathbf{c} = \mathbf{g}$$

extremely efficiently, if the the coefficients **c** and explicitly needed

How to construct auxiliary basis sets?

Insight comes from the single center problem. Basis functions are usually chosen as real spherical harmonics:

$$\begin{split} \mu_{A}(\mathbf{r})\nu_{A}(\mathbf{r}) &\propto \left\{ \exp(-\alpha_{\mu}r_{A}^{2})S_{l_{\mu}m_{\mu}}(\theta,\phi) \right\} \left\{ \exp(-\alpha_{\nu}r_{A}^{2})S_{l_{\nu}m_{\nu}}(\theta,\phi) \right\} \\ &= \exp(-(\alpha_{\mu}+\alpha_{\nu})r_{A}^{2})\sum_{\lambda=\left|l_{\mu}-l_{\nu}\right|}^{l_{\mu}+l_{\nu}}\sum_{\kappa=-\lambda}^{+\lambda}S_{\lambda\kappa}(\theta,\phi) \begin{pmatrix} l_{\mu} & l_{\nu} & | \lambda \\ m_{\mu} & m_{\nu} & | \kappa \end{pmatrix} \end{split}$$

- A product of two d functions creates s,p,d,f,g functions with exponents that are the sum of the individual exponents.
- A good aux basis contains functions with up to 2L_{max} (L_{max} is the highest angular momentum in the basis set.
- For each angular momentum I, there need to be gaussians with exponents between 2 α_{min} to $2\alpha_{max}$ where $\alpha_{min/max}$ are the smallest and largest exponents that ,create' functions with a given L.
- Best constructed as a **geometric series**: $\alpha_{i+1} = \alpha_i / \beta$ $\beta \approx 2-3$

(possibly further optimized by least square fitting to reference data)

Size of Auxiliary basis: 2-10 times the size of the orbital basis (MUCH better than N²!)

Standardized Fitting Basis Sets

Highly recommended fitting basis sets for various purposes:

- Coulomb fitting, def2/J bases
- MP2 def2/C bases
- Hartree-Fock exchange def2/JK bases
- 1. K. Eichkorn, O. Treutler, H. Ohm, M. Haser, and R. Ahlrichs, *Chem. Phys. Lett.* <u>242</u> (**1995**) 652.
- K. Eichkorn, F. Weigend, O. Treutler, and R. Ahlrichs, *Theor. Chem. Acc.* <u>97</u> (**1997**) 119.
- F. Weigend, M. Häser, H. Patzelt, and R. Ahlrichs, *Chem. Phys. Lett.* <u>294</u> (**1998**) 143.
- 4. F. Weigend, Phys. Chem. Chem. Phys. <u>4</u> (2002) 4285.
- 5. F. Weigend, Phys. Chem. Chem. Phys. <u>8</u> (2006) 1057.
- 6. F. Weigend, J. Comp. Chem. 29 (2008) 167.

Available from the EMSL basis set exchange

https://bse.pnl.gov/bse/portal

An AutoAux Algorithm

For each atom in the molecule

- \checkmark For each basis function calculate the radial expectation value <r>
- ✓ For each angular momentum. Calculate an effective single Gaussian with the same <r>

$$\alpha_{\rm eff} = \frac{2}{\pi} \left(\frac{2^{2l+1} ((l+1)!)^2}{(2l+2)!} \right)^2 \frac{1}{\langle r \rangle}$$

 \checkmark For each angular momentum find lowest and highest effective exponent

- ✓ For each basis function pair loop over I_{1+I₂} ... |I₁-I₂| and keep track of the lowest and highest exponent generated by the pair.
- ✓ Tweak the range of exponents for each aux basis angular momentum
- ✓ For each effective angular momentum in the aux basis generate a geometric progression of auxiliary functions:

$$\alpha_{_i} = \beta_{_l} \, \alpha_{_{i-1}}^{(l)}$$

 \checkmark Additional limited heuristics to limit the number of high L aux functions

Stoychev, G.; Auer, AA, FN 2016, submitted

An AutoAux Algorithm



Frozen-core MP2/def2-QZVPP

Similar accuracy to specific purpose optimized auxiliary basis sets while keeping performance with ~factor 2

Stoychev, G.; Auer, AA, FN 2016, submitted

Advantages of the RI/DF approximation

1. The error is small, smooth and bounded.

With reasonable auxiliary basis sets largely cancels upon taking energy differences (e.g. 0.02 kcal/mol in total energies). The RI-fitted Coulomb energy always undershoots the exact Coulomb energy

- 2. The **number of three center integrals** is much smaller than four center integrals Saves about one order of magnitude in storage, integral generation, ... Reduces the formal scaling of the algorithm
- 3. The three center integrals can be **calculated much faster** than four center integrals The calculation time per integral is up to a factor of 10 smaller owing to the much smaller FLOP/MOP count of 3-center vs 4-center integrals
- 4. The RI/DF approximation is an **approximate factorization**. Depending on the application, this may lead to reduced scaling algorithms through new and more compact, favorable intermediates

Semi-Numerical Techniques (COS, PS)

Locality of basis functions and Integration

Basis functions are inherently local and numerical integration is a natural way to exploit this locality

$$\mu_A(\mathbf{r})\nu_B(\mathbf{r}) =$$



=0 if the two spheres do not intersect

Hence, on a grid, non-zero basis functions for a given gridpoint and non-zero products are readily identified and can be skipped

However, two-electron integrals:

$$\int \int \mu_{A}(\mathbf{r}_{1}) \nu_{B}(\mathbf{r}_{1}) \frac{1}{\left|\mathbf{r}_{1}-\mathbf{r}_{2}\right|} \kappa_{C}(\mathbf{r}_{2}) \tau_{D}(\mathbf{r}_{2}) d\mathbf{r}_{1} d\mathbf{r}_{2}$$

Involve a nasty Coulomb singularity at $r_1=r_2$, which is unpleasant to integrate over numerically

Locality of Basis Functions and Integration

Rather consider the potential created by the second electron:

$$\begin{split} \int \int \mu_{A}(\mathbf{r}_{1})\nu_{B}(\mathbf{r}_{1}) \frac{1}{\left|\mathbf{r}_{1}-\mathbf{r}_{2}\right|}\kappa_{C}(\mathbf{r}_{2})\tau_{D}(\mathbf{r}_{2})d\mathbf{r}_{1} d\mathbf{r}_{2} \\ = \int \mu_{A}(\mathbf{r}_{1})\nu_{B}(\mathbf{r}_{1}) \frac{1}{\left|\mathbf{r}_{1}-\mathbf{r}_{2}\right|}\kappa_{C}(\mathbf{r}_{2})\tau_{D}(\mathbf{r}_{2})d\mathbf{r}_{2} d\mathbf{r}_{1} \\ \frac{V_{\kappa_{C}\tau_{D}}(\mathbf{r}_{1})}{V_{\kappa_{C}\tau_{D}}(\mathbf{r}_{1})} \end{split}$$

Which implies, that we may want to numerically integrate over \mathbf{r}_1 and analytically over \mathbf{r}_2

(Reasonably simple integrals, like electron-nuclear attraction)

Dealiasing and Overlap Fitting

The accuracy of the procedure can be improved by analyzing the error introduced by the numerical integration.

First replace:

$$\sum_{g} X_{\mu g} X_{\nu g} A^{g}_{\kappa\tau} \to \sum_{g} Q_{\mu g} X_{\nu g} A^{g}_{\kappa\tau}$$

Error analysis:

Action of an operator on a function:

Introduce a **complete basis** { φ } which has components inside ($\varphi_{\mu}^{\parallel}$) and outside (φ_{μ}^{\perp}) the actual orbital basis

$$\hat{O}\phi_{i} = \sum_{\mu} \left\langle \varphi_{\mu}^{\parallel} \,|\, \hat{O} \,|\, \phi_{i} \right\rangle \varphi_{\mu}^{\parallel} + \sum_{\mu} \left\langle \varphi_{\mu}^{\perp} \,|\, \hat{O} \,|\, \phi_{i} \right\rangle \varphi_{\mu}^{\perp}$$

 φ_{μ}^{\perp} = ,Dealiasing functions'

$$\text{Integrate:} \quad \left\langle \phi_{j} \mid \hat{O} \mid \phi_{i} \right\rangle = \sum_{\mu} \left\langle \varphi_{\mu}^{\parallel} \mid \hat{O} \mid \phi_{i} \right\rangle \left\langle \phi_{j} \mid \varphi_{\mu}^{\parallel} \right\rangle + \sum_{\mu} \left\langle \varphi_{\mu}^{\perp} \mid \hat{O} \mid \phi_{i} \right\rangle \left\langle \phi_{j} \mid \varphi_{\mu}^{\perp} \right\rangle$$
$$For analytic integration=0$$

R Izsák, FN, W Klopper, J Chem Phys <u>139</u>, **2013**, 094111 R Izsak, FN, J Chem Phys <u>135</u>, (**2011**) 144105.

Dealiasing and Overlap Fitting

For a numerically approximated integral

$$\begin{split} \left\langle \phi_{j} \mid \hat{O} \mid \phi_{i} \right\rangle - \left\langle \phi_{j} \mid \hat{O} \mid \phi_{i} \right\rangle^{N} &= \sum_{\mu} \left\langle \varphi_{\mu}^{\parallel} \mid \hat{O} \mid \phi_{i} \right\rangle \left[\left\langle \phi_{j} \mid \varphi_{\mu}^{\parallel} \right\rangle - \left\langle \phi_{j} \mid \varphi_{\mu}^{\parallel} \right\rangle^{N} \right] \\ &+ \sum_{\mu} \left\langle \varphi_{\mu}^{\perp} \mid \hat{O} \mid \phi_{i} \right\rangle \left[\left\langle \phi_{j} \mid \varphi_{\mu}^{\perp} \right\rangle - \left\langle \phi_{j} \mid \varphi_{\mu}^{\perp} \right\rangle^{N} \right] \end{split}$$

If the **numerically evaluated overlap** is *exact*, the first term vanishes and one is only left with the **dealiasing error**

$$\left\langle \phi_{j} \mid \hat{O} \mid \phi_{i} \right\rangle - \left\langle \phi_{j} \mid \hat{O} \mid \phi_{i} \right\rangle^{N} \xrightarrow{S_{j\mu}^{\parallel} - S_{j\mu}^{\parallel;N} = 0} \sum_{\mu} \left\langle \varphi_{\mu}^{\perp} \mid \hat{O} \mid \phi_{i} \right\rangle \left(S_{j\mu}^{\perp} - S_{j\mu}^{\perp;N} \right)$$

The exactly numerically integrated overlap can be ensured by choosing **Q** wisely:

$$\mathbf{Q} = \mathbf{S}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X} = \mathbf{S} \mathbf{S}_N^{-1} \mathbf{X}$$

Straightforward, computationally trivial and general for any orbital basis.

R Izsák, FN, W Klopper, J Chem Phys <u>139</u>, **2013**, 094111 R Izsak, FN, J Chem Phys <u>135</u>, (**2011**) 144105.

In genuine *Pseudospectral* approaches, additional dealiasing functions must be introduced. They are specific for the underlying orbital basis and hence PS is not general R. A. Friesner, *Chem. Phys. Lett.* <u>116</u>, 39 (1985); R. A. Friesner, *J. Chem. Phys.* <u>85</u>, 1462 (1986).; R. A. Friesner, *J. Chem. Phys.* <u>86</u>, 3522 (1987).

Numerical Example



R Izsák, FN, W Klopper, J Chem Phys <u>139</u>, **2013**, 094111 R Izsak, FN, J Chem Phys <u>135</u>, (**2011**) 144105.

Approximations to SCF Equations

For a review see e.g. FN, in *Linear Scaling Methods in Quantum Chemistry*, edited by M. Papadopoulous et al. (**2011**), Springer, Heidelberg, p.227-262

Coulomb Term: Density Fitting and Split-RI-J

The RI-approximation works particularly well for Coulomb-type contractions:

$$J_{\mu\nu} = \sum_{\kappa\tau} P_{\kappa\tau}(\mu\nu \mid \kappa\tau)$$

$$\approx \sum_{\kappa\tau} \sum_{K,L} P_{\kappa\tau} c_{\mu\nu}^{K} c_{\kappa\tau}^{L} (K \mid L)$$

$$= \sum_{K} (\mu\nu \mid K) \sum_{L} V_{KL}^{-1} \sum_{\kappa\tau} P_{\kappa\tau}(\kappa\tau \mid L)$$

$$= \sum_{K} P_{K}(\mu\nu \mid K)$$

Requires generation (or read) of the three index integrals twice

Quadratic scaling with system size remains unaffected

Coulomb energy is variational

Implementation

g=0 Loop κ,τ Loop L	
Calculate (KT L)	def2-TZVP=667 BFs
g(L)=g(L)+P(KT)*(KT L)	def2-T7VP/J=1056 Aux BEs
end loops K,T,L	
Use Cholesky decomposition of V to solve $\mathbf{VP}=\mathbf{g}$	Traditional : 5800 sec
Loop μ,ν	RI(direct) : 172 sec
Loop K	
Calculate ($\mu\nu$ K)	RI(conv) : 44 sec
$J(\mu,\nu) = J(\mu,\nu) + P(K) * (\mu\nu K)$	Error : 0.5 mEh
end loops μ, ν, k	0.3 kcal/mc



Through the much smaller number of integrals and their cheapness, the savings are fairly spectacular (factor 10-100)



Enables a lot of computational chemistry! If exact exchange is not needed - Larger systems, faster turnaround, time, better basis sets, ...

Split-RI-J

Obviously, RI-J can profit in the same way from integrating the integral evaluation as Split-J (J-engine) can.

$$J_{\mu\nu}^{RI} = \sum_{tuv} E_{tuv}^{\mu\nu} \sum_{t'u'v'} R_{t+t',u+u',v+v'} \sum_{K} (-1)^{t'+u'+v'} E_{t'u'v'}^{K} X_{K}$$
$$X_{K} = \sum_{L} V_{KL}^{-1} \sum_{tuv} E_{tuv}^{\kappa\tau} \sum_{t'u'v'} R_{t+t',u+u',v+v'} \sum_{\kappa\tau} P_{\kappa\tau} (-1)^{t'+u'+v'} E_{t'u'v'}^{L}$$

Loop kt Calculate $T(q) = S_{kt} P(k, t) E(k, t, q) (-1)^{q}$ end Loop r Loop q Calculate R(p,q)V(p) += R(p,q) T(q)end Calculate E(r, p)q(r) += E(r, p)V(p)end Solve $c=V^{-1}q$ through Cholesky Loop r Calculate $X(p) = S_r c(r) E(r, p)$ end Loop m,n Loop p Calculate R(p,q)Y(q) = R(p, q) X(p)end Calculate E(m,n,q) $J(\mathbf{m},\mathbf{n}) += E(\mathbf{m},\mathbf{n},\mathbf{q}) Y(\mathbf{q})$ end



A Numerical Comparison



	$Def2-\underline{SV}(P)$	6-31G*	6-311G(<u>d,p</u>)	Def2-TZVP(-f)	Def2-TZVPP	Def2-QZVPP
N _{BAS}	350	350	522	654	976	1884
Exact-J	921	986	2772	9084	28881	271892
Split-J	513 (1.8)	632 (1.6)	1603 (1.7)	4206 (2.2)	22618 (1.3)	226576 (1.2)
RI-J(direct)	94 (9.8)	101 (9.8)	155 (17.9)	321 (28.3)	1346 (21.5)	39081 (7.0)
RI-J(semi-direct)	58 (15.9)	60 (16.4)	117 (23.7)	191 (47.6)	319 (90.5)	1795 (151.5)
RI-J(conv)	27 (34.1)	27 (36.5)	63 (44.0)	99 (91.8)	271 (106.6)	1243 (218,7)
Split-RI-J	62 (14.9)	67 (14.7)	110 (25.2)	164 (55.4)	374 (77.2)	38591 (7,0)

Single CPU of a MacPro 3.1; operating system OS X 10.5, 2 quad-core Intel XEON 3.0 GHZ CPUs that have 12 MB level 2 cache

If the RI integrals can be stored, the speedups are gigantic!

FN, in Linear Scaling Methods in Quantum Chemistry, M. Papadopoulous et al. (eds)(2011), Springer, Heidelberg, p.227-262

Coulomb Term: Cholesky Decomposition

Cholesky decomposition leads to equations that are very similar to RI-J

Bebe and Linderberg realized that the two-electron supermatrix

$$I_{\mu\nu,\kappa\tau} = (\mu\nu \,|\,\kappa\tau)$$

Is symmetric and positive (semi)definite. Hence one can Cholesky decompose it:

 $\mathbf{I} = \mathbf{L}\mathbf{L}^{T}$

Hence:

$$I_{\mu\nu,\kappa\tau} = \sum_{J} L_{\mu\nu,J} L_{\kappa\tau,J}$$

The ,Cholesky vectors' $L_{_{\mu
u,J}}$ take the role of the three-index integrals



There is a finite number $N_J << N_{Bas}^2$ of significant Cholesky vectors (e.g. 2-4N)

Large savings, if the Cholesky vectors can be made efficiently (Dalton, MOLCAS)

No auxiliary basis set necessary, precision can be finely controlled

NHF Beebe, J Linderberg, IJQC. 7, 683 1977. H Koch, A Sanchez de Meras, TB Pedersen, JCP, 2003, 118, 9481

Coulomb Term: Pseudo-Spectral Approximation

The pseudo-spectral approximation works well for the Coulomb term:

$$\begin{split} J_{\mu\nu} &= \sum_{\kappa\tau} P_{\kappa\tau} (\mu\nu \mid \kappa\tau) \\ &\approx \sum_{\kappa\tau} P_{\kappa\tau} \sum_{g} Q_{\mu g} X_{\nu g} A_{\kappa\tau}^{g} \\ &= \sum_{g} Q_{\mu g} X_{\nu g} \sum_{\kappa\tau} P_{\kappa\tau} A_{\kappa\tau}^{g} \\ &= \sum_{g} Q_{\mu g} X_{\nu g} V_{g} \end{split}$$

Basis for Friesner's pseudo spectral implementation in Jaguar

Coulomb Term: Poisson Equation

One step further, one can use fully numerical techniques for the Coulomb matrix. In order to avoid the numerical integration over the Coulomb singularity, one might solve the Poisson equation instead (Becke's NUMOL or Delly's DMOL)

Can be used with any type of basis function, even numerical orbitals!

Density:

$$ho(\mathbf{r}_{_{g}})=\sum_{A}
ho^{A}(\mathbf{r}_{_{g}})$$

Grid points are assigned to atoms and hence, the density on each grid point can be assigned to an atom as well

Multipole Expansion:

$$\rho^{A}(\mathbf{r}) = \sum_{lm} \rho_{lm}(r) S_{lm}(\theta, \phi)$$
$$\rho_{lm}(r) = \int S_{lm}(\theta, \phi) \rho^{A}(r, \theta, \phi) \sin \theta \, d\theta \, d\phi$$

Poisson Approach to the Coulomb term

Poisson equation:

$$\nabla^2 V_A(\mathbf{r}) = -4\pi\rho^A(\mathbf{r})$$

$$\begin{split} V_{A}(r,\theta,\phi) &= \sum_{lm} r^{-1} U_{lm}(r) S_{lm}(\theta,\phi) \\ \frac{d^{2} U_{lm}(r)}{dr^{2}} - \frac{l(l+1)}{r^{2}} U_{lm}(r) = -4\pi \rho_{lm}(r) \end{split}$$

Simple 1d differential equation; solved by standard methods

Total Coulomb potential:
$$V(\mathbf{r}) = \sum_{A} V_{A}(\mathbf{r})$$

Readily available on the grid

Coulomb matrix:

$$J_{\mu\nu} \approx \sum_{g} X_{g\mu} X_{g\nu} V(\mathbf{r}_{g})$$

No timings, but DMol is reported to be <u>extremely</u> efficient

Coulomb Term: Fast Multipole Method

Reaching linear or near-linear scaling in the Coulomb construction is possible by various variants of fast multipole methods. The basic idea is simple:

- 1. For a given pair of basis functions, divide the contributions to the Coulomb potential into a **near field** and a **far field** contribution
- 2. Treat the **near field contribution** with exact analytical integration
- 3. Treat the far field contribution using multipole expansions





Exchange: RI/DF Approximation

$$\begin{split} K_{\mu\nu} &= \sum_{\kappa\tau} P_{\kappa\tau} (\mu\kappa \mid \nu\tau) & \text{The exchange term is much harder to approximate! Let us try the RI method} \\ &\approx \sum_{\kappa\tau} \sum_{K,L} P_{\kappa\tau} c_{\mu\kappa}^{K} c_{\nu\tau}^{L} (K \mid L) & \text{approximate! Let us try the RI method} \\ &= \sum_{\kappa\tau} \sum_{K,L} P_{\kappa\tau} (\mu\kappa \mid K) (\nu\tau \mid L) V_{KL}^{-1} \sum_{\tau} c_{\tau i} (\nu\tau \mid L) & \text{BAD: No convenient v form an intermediate b summing two indices at the same time} \\ &= \sum_{i} \mathbf{X}^{i} \mathbf{Y}^{i+} & \text{Do Batches of i convenient } \mathbf{X}^{i} & \text{O}(N^{2}) - \mathbf{O}(N^{3}) \\ &= \sum_{i} \mathbf{X}^{i} \mathbf{Y}^{i+} & \text{Do Batches of i convenient } \mathbf{X}^{i} & \text{O}(N^{2}) - \mathbf{O}(N^{3}) \\ &= \sum_{i} \mathbf{X}^{i} \mathbf{Y}^{i+} & \text{Do Batches of i convenient } \mathbf{X}^{i} & \text{O}(N^{2}) - \mathbf{O}(N^{3}) \\ &= \sum_{i} \mathbf{X}^{i} \mathbf{Y}^{i+} & \text{Do Batches of i convenient } \mathbf{X}^{i} & \text{O}(N^{2}) \\ &= \sum_{i} \mathbf{X}^{i} \mathbf{Y}^{i+} & \text{Do Batches of i convenient } \mathbf{X}^{i} & \text{O}(N^{2}) \\ &= \sum_{i} \mathbf{X}^{i} \mathbf{Y}^{i+} & \text{Do Batches of i convenient } \mathbf{X}^{i} & \text{O}(N^{2}) \\ &= \sum_{i} \mathbf{X}^{i} \mathbf{X}^{i+} & \text{Do Batches of i convenient } \mathbf{X}^{i} & \text{O}(N^{2}) \\ &= \sum_{i} \mathbf{X}^{i} \mathbf{X}^{i+} & \text{Do Batches of i convenient } \mathbf{X}^{i} & \text{O}(N^{2}) \\ &= \sum_{i} \mathbf{X}^{i} \mathbf{X}^{i+} & \text{Do Batches of i convenient } \mathbf{X}^{i} & \text{O}(N^{2}) \\ &= \sum_{i} \mathbf{X}^{i} \mathbf{X}^{i+} & \text{Do Batches of i convenient } \mathbf{X}^{i} & \text{O}(N^{2}) \\ &= \sum_{i} \mathbf{X}^{i} \mathbf{X}^{i+} & \text{Do Batches of i convenient } \mathbf{X}^{i} & \text{O}(N^{2}) \\ &= \sum_{i} \mathbf{X}^{i} \mathbf{X}^{i+} & \text{Do Batches } \mathbf{X}^{i} & \text{O}(N^{2}) \\ &= \sum_{i} \mathbf{X}^{i} \mathbf{X}^{i+} & \text{Do Batches } \mathbf{X}^{i} & \text{O}(N^{2}) \\ &= \sum_{i} \mathbf{X}^{i} \mathbf{X}^{i+} & \text{Do Batches } \mathbf{X}^{i+} & \text{O}(N^{2}) \\ &= \sum_{i} \mathbf{X}^{i} \mathbf{X}^{i+} & \text{Do Batches } \mathbf{X}^{i+} & \text{O}(N^{2}) \\ &= \sum_{i} \mathbf{X}^{i} \mathbf{X}^{i+} & \text{D}^{i} \mathbf{X}^{i+} & \text{D}^{i} \mathbf{X}^{i+} & \text{D}^{i} \mathbf{X}^{i+} \\ &= \sum_{i} \mathbf{X}^{i} \mathbf{X}^{i+} & \text{D}^{i} \mathbf{X}^{i$$

D: No convenient way to *m* an intermediate by nming two indices away at same time

 $O(N^2) - O(N^3)$

 $O(N^2)$

 $O(N^4)$

is much harder to

BAD: exchange scales intrinsically linear, but we have $O(N^4)$ scaling!

Exchange: "Chain of Spheres" approach



F. Neese, F. Wennmohs, A. Hansen, and U. Becker, Chem. Phys. **356** (2009) 98.;R. Izsak, and F. Neese, J. Chem. Phys. **135**, 144105 (2011) 144105; R. Izsák, F. Neese, and W. Klopper, J. Chem. Phys. **139** (2013) 094111

Exchange: "Chain of Spheres" Implementation



Preselect interacting basis function pairs based on the overlapping spheres or significant density criteria

Loop ig over grid points	0 (N)
Calculate basis function values $X(\mu,g)$ for $\mu \subset \{g\}$	const
Construct secondary+tertiary shell-lists	const
Perform Matrix multiplication $\mathbf{F}(\kappa, g) = \mathbf{P}(\kappa, \tau) \mathbf{X}(\tau, g)$	const
Calculate analytic integrals $G(v,g) = \Sigma_{\kappa} F(\kappa,g) A(\kappa v,g)$	const
Perform matrix multiplication $K(\mu,\nu) = G(\nu,g) X(\mu,g)$	const
end	

Most expensive step: generation of analytic integrals

- → Hand optimized, highly efficient code
- → Loop unrolling, through angular momentum classes
- → Computation of common intermediates outside the inner loop over batches of grid points

FN, F. Wennmohs, A. Hansen, U. Becker, Chem. Phys. 356 (2009) 98. R. Izsak, and FN, JCP 135, 144105 (2011) 144105.

The RIJCOSX approximation



FN, F. Wennmohs, A. Hansen, U. Becker, Chem. Phys. 356 (2009) 98. R. Izsak, and FN, JCP 135, 144105 (2011) 144105.

Comparison: RI-JK and RIJCOSX

				SPEEDUP	ERROR	SPEEDUP	ERROR	
MOL	BASIS	N_{BAS}	$\mathrm{T}_{\mathrm{WALL}}$	RI-JK	RI-JK	RIJCOSX	RIJCOSX	
			(min)		(mEh)		(mEh)	
(gly) ₂	def2-SVP	166	1,8	2,4	0,4	1,5	-0,6	Small
	def2-TZVP(-df)	219	5,3	4,5	0,2	2,6	-0,4	system,big
	def2-TZVPP	391	42,9	7,9	0,3	4,9	-0,2	> RIK wins
	def2-QZVPP	753	464,3	9,1	0,3	4,5	0,5	
(gly) ₄	def2-SVP	308	10,2	1,8	0,8	2,2	-1,1	
	def2-TZVP(-df)	407	32,0	2,8	0,5	3,4	-0,5	
	def2-TZVPP	723	232,8	5,1	0,5	5,8	-0,1	
	def2-QZVPP	1389	2684,1	5,3	0,6	6,6	0,9	
(gly) ₈	def2-SVP	592	57,2	1,0	1,5	2,9	-1,7	
	def2-TZVP(-df)	783	208,4	2,3	0,9	5,6	-0,8	
	def2-TZVPP	1387	1379,6	5,0	1,0	9,7	-0,1	big system,
	def2-QZVPP	2661	12262,7	6,2	1,2	8,4	1,4	

Table 4: Comparison of RI-JK and RIJCOSX for RHF calculations on Glycine chains

S. Kossmann, and FN, Chem. Phys. Lett. 481 (2009) 240.

Application to MP2

MP2 Correlation Energy: Density Fitting

The RI approximation mixes particularly well with MP2

$$egin{aligned} &K^{ij}_{ab} = (ia \mid jb) \ &pprox \sum_{K} (ia \mid ar{K})(jb \mid ar{K}) & ext{Rate limiting O(N^5) step} \end{aligned}$$

Algorithm:

```
Integral Transformation
Calculate V<sup>-1/2</sup>
Loop K
Loop m<=n
Calculate I[K](m,n)=(mn|K)
end
Transform J[K](i,n)=(c<sub>occ</sub>+I[K])(i,n)
Transform I[K](i,a)=(J[K]c<sub>virt</sub>)(i,a)
end Loop K
Sort integrals to X[i](K,a)
Orthogonalize integrals with V<sup>-1/2</sup>
```

```
MP2
Loop i
Read X[i]
Loop j<=i
Read X[j]
Form K<sup>ij</sup>=X[i]<sup>+</sup>X[j]
Form Amplitudes T<sup>ij</sup>
Contribution to EMP2
End loops i,j
```

RI-MP2 Example



Diclophenac def2-TZVP (667 basis functions) 4 GB main memory used

BLAS optimized algorithm

RI algorithm

1732 sec

55 sec Factor >30

error ~0.01% = 0.53 mEh

Half transformation	:	1697.0	sec
AO-integral generation	:	1078.9	sec
Half transformation	:	354.0	sec
K-integral sorting	:	60.4	sec

Integral trafo	:	17.514	sec	(31.4%)
I/O of integrals	:	0.206	sec	(0.4%)
K(i,j) Operators	:	32.342	sec	(58.0%)
T(i,j) pair energies	:	4.056	sec	(7.3%)
V-calculation	:	0.040	sec	(0.1%)
V** (-1/2)	:	0.313	sec	(0.6%)



Superbly efficient for medium sized systems, but still O(N⁵) scaling

The Laplace-Transform MP2

Look at the MP2 expression:

$$E_{MP2} = \frac{1}{4} \sum_{ijab} \frac{\left| \left\langle ij \mid \mid ab \right\rangle \right|^2}{\varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j}$$

We know how to nicely factorize and approximate the nominator. But what about the denominator?

Almlöf realized that the **Laplace transform**:

$$\frac{1}{x} = \int_{0}^{\infty} \exp(-tx) dt$$

Can be used to get:

$$\frac{1}{\varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j} = \int_0^\infty \exp(-(\varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j)t)dt$$
$$= \int_0^\infty e^{-\varepsilon_a t} e^{-\varepsilon_b t} e^{\varepsilon_i t} e^{\varepsilon_j t} dt$$

J Almlöf Chem. Phys. Lett, 1991, 181, 4

Laplace Transform MP2

Leading to:

$$E_{MP2} = \frac{1}{4} \int \sum_{ijab} e^{-\varepsilon_a t} e^{-\varepsilon_b t} e^{\varepsilon_i t} e^{\varepsilon_j t} \left\langle ij \mid \mid ab \right\rangle \left\langle ij \mid \mid ab \right\rangle dt$$

Absorb the t-dependence in the orbitals:

$$\psi_{i}(t) = \psi_{i}(0)e^{\frac{1}{2}\varepsilon_{i}t}$$
$$\psi_{a}(t) = \psi_{a}(0)e^{-\frac{1}{2}\varepsilon_{a}t}$$

Giving:

$$E_{MP2} = \frac{1}{4} \sum_{ijab} \int \left\langle i(t)j(t) \mid \mid a(t)b(t) \right\rangle^2 dt$$

perform the t-integration numerically (6-10 points are enough)

Apply whatever approximation to the t-dependent integral or transform it to the AO basis

M. Häser, J. Almlöf, and G. E. Scuseria, Chem. Phys. Lett. <u>181</u>, 497 **1991**.

Laplace Tranform MP2

Linear scaling with full accuracy can be achieved

$$\begin{split} E_{MP2} &= \frac{1}{4} \sum_{ijab} \int \left\langle i(t)j(t) \mid \mid a(t)b(t) \right\rangle^2 dt \\ &= -\sum_{\alpha} \sum_{\mu\nu\kappa\tau} (\underline{\mu}\overline{\nu} \mid \underline{\kappa}\overline{\tau})^{(\alpha)} [2(\mu\nu \mid \kappa\tau) - (\mu\tau \mid \kappa\nu)] \qquad \text{a=quadrature points} \end{split}$$

$$(\underline{\mu}\overline{\nu} \mid \underline{\kappa}\overline{\tau})^{(\alpha)} = \sum_{\mu'\nu'\kappa'\tau'} (\mu'\nu' \mid \kappa'\tau') \overline{P}^{(\alpha)}_{\mu\mu'} \underline{P}^{(\alpha)}_{\nu\nu'} \overline{P}^{(\alpha)}_{\kappa\kappa'} \underline{P}^{(\alpha)}_{\tau\tau'}$$
$$\underline{P}^{(\alpha)}_{\mu\nu} = w^{1/4}_{\alpha} \sum_{i} c_{\mu i} c_{\nu i} e^{+\varepsilon_{i} t_{\alpha}}$$
$$\overline{P}^{(\alpha)}_{\mu\nu} = w^{1/4}_{\alpha} \sum_{a} c_{\mu a} c_{\nu a} e^{-\varepsilon_{a} t_{\alpha}}$$

2025 atoms and 20 371 basis functions (6-31G**)

SA. Maurer, DS Lambrecht, J Kussmann, C Ochsenfeld J Chem Phys, <u>138</u>, 014101 **2013**B Doser, J Zienau, L Clin, DS Lambrecht, C Ochsenfeld Z. Phys. Chem. Int. Ed., **2010**, <u>224</u>, 397