Explicit electron correlation by a combined use of Gaussian-type orbitals and Gaussian-type geminals

T. Helgaker,
Centre for Theoretical and Computational Chemistry,
Department of Chemistry, University of Oslo, Norway

Pål Dahle, Department of Chemistry, University of Oslo, Norway

Dan Jonsson, Department of Physics, Stockholm University, Sweden

Peter Taylor, Department of Chemistry, University of Warwick, UK

Yngve Öhrn Symposium

International Conference of Computational Methods in Sciences and Engineering

(ICCMSE 2007)

Marbella Corfu Hotel, Corfu, Greece

September 25–30, 2007

CI expansions and explicit correlation

- Experimental ionization potential of helium: 24.59 eV
  - Unsöld 1927: 20.41 eV (first-order perturbation theory)
  - Hylleraas 1928: 24.47 eV (CI expansion: slow convergence!)
  - Hylleraas 1929: 24.58 eV (explicit correlation: fast convergence!)
  - the question of CI expansions vs. explicit correlation is still with us today

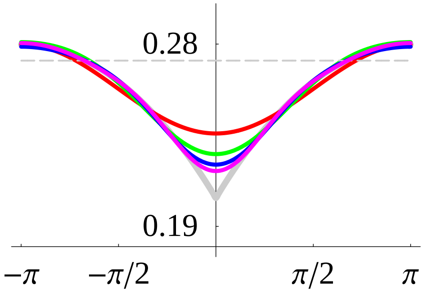
Overview

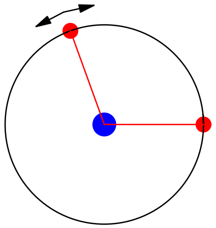
- We shall consider
  - Coulomb hole and energy convergence
  - extrapolation and explicit correlation
  - comparison of explicitly correlated methods (GTG, R12, GGn) and extrapolation
  - some conclusions
- Dahle, Helgaker, Jonsson and Taylor, PCCP **9**, 3112 (2007)

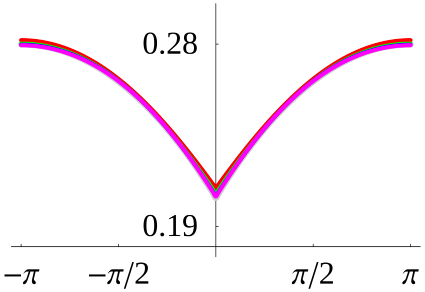
The Coulomb hole

- When electrons approach one another, the wave function behaves in special manner
  - large (infinite) repulsion is canceled by large (infinite) negative kinetic energy
  - for singlet pairs, a Coulomb hole is created, with a cusp at the point of coalescence
- Below we have plotted the Coulomb hole in helium for two classes of wave functions
  - CI wave functions with one shell included at a time: $1s$, $2s2p$, $3s3p3d$, $\ldots$
  - the same wave functions with a single term linear in r_{12} added (CI-R12)

$$\Psi_N^{\mathrm{CI}} \quad \rightarrow \quad \Psi_N^{\mathrm{CI}} + c_{12} r_{12} \Psi_{1s^2}$$



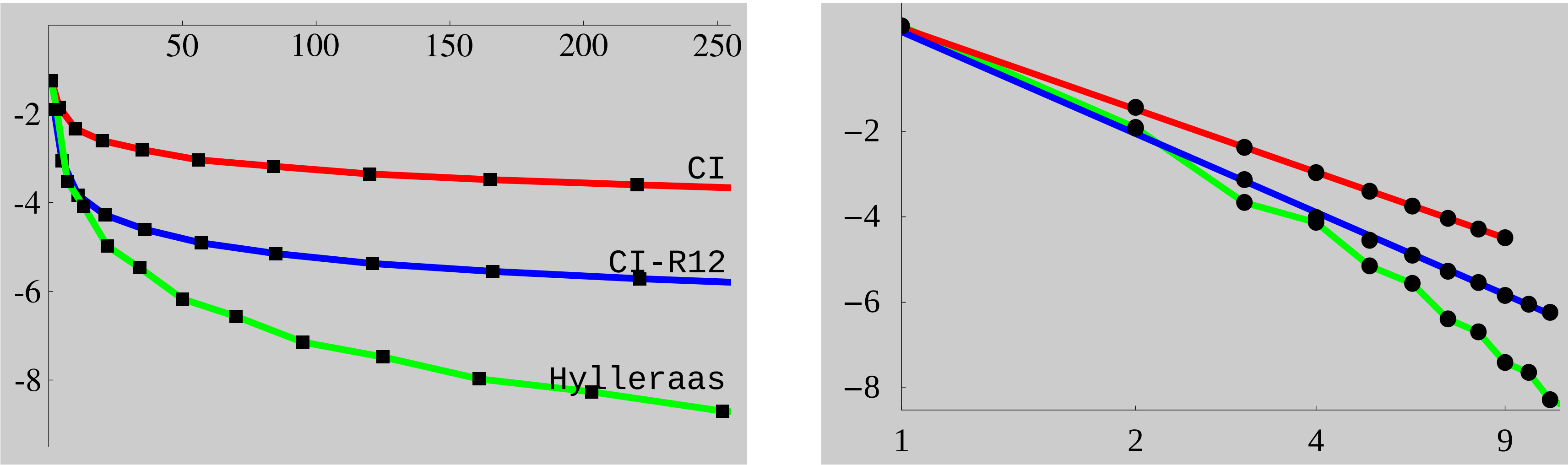




  - the CI wfs. can only indirectly describe the Coulomb hole and converge slowly
  - the CI-R12 wfs. contain a cusp by construction and model the Coulomb hole well
- For high accuracy in the energy, we need a good description of the Coulomb hole

Basis-set convergence

- Calculations on the helium atom using single-zeta Slater functions
  - standard CI expansion, CI-R12 expansion, the Hylleraas expansion



  - Left: log–lin plots of the error in the energy against the number of terms
  - Right: log–log plots of energy contributions against the principal quantum number n
- The standard CI expansion converges slowly
  - each new shell contributes n^{-4} energy for CI
  - convergence is very smooth
- The inclusion of a single R12 term reduces the energy error dramatically

The principal expansion and basis-set extrapolation

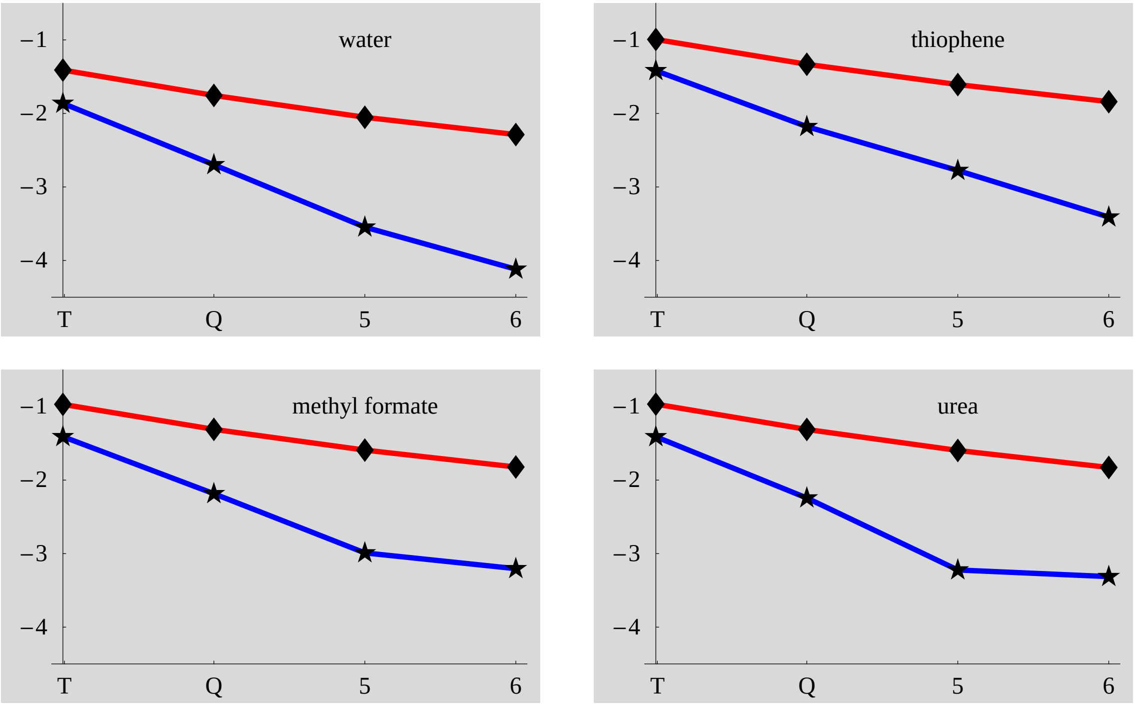
- The principal expansion: include orbitals in full shells of principal quantum number n
  - we can now easily estimate the omitted contributions and hence the basis-set limit

$$E_\infty = E_X + A \sum_{n=X+1}^{\infty} n^{-4} \approx E_X + AX^{-3}$$

  - the unknown constant A is eliminated by carrying out two separate calculations

$$E_\infty = \frac{X^3 E_X - Y^3 E_Y}{X^3 - Y^3}$$ ← two-point extrapolation formula

  - practical realization: cc-pVXZ
  - logarithmic errors in plain and extrapolated energies relative to R12
  - from: Klopper, Manby, Ten-no and Valeev, *Int. Rev. Phys. Chem.* **25**, 427 (2006)



Explicit correlation

- Extrapolation improves our description significantly
  - it is very simple and does not require new coding
- A more fundamental solution is the use of explicitly correlated methods
  - insert $f(r_{12})$ explicitly into the wave function
  - accurate total energies obtainable without extrapolation
- We shall now consider several such explicitly correlated methods
  - the GTG method of Szalewicz, Jeziorski, Monkhorst and Zabolitzky (1982)
  - the R12 method of Kutzelnigg and Klopper (1985, 1986)
  - the GGn model (a mixed GTO–GTG model) explored by us
- We shall consider small systems, asking the questions
  - what error reduction can be expected from these methods?
  - how does it compare with basis-set extrapolation?
- It is sufficient to consider second-order Møller–Plesset (MP2) theory
  - the doubles contributions converge more slowly than higher excitations

Møller–Plesset theory

- The zero-order system is represented by the Fock operator

$$F_1\phi_i(1) = \varepsilon_i\phi_i(1) \qquad \leftarrow \text{spin orbitals and energies}$$

  - the zero-order wave function is a determinant

- To first order in perturbation theory, the electrons are correlated pairwise:

$$\phi_{ij}(1,2) = \det|\phi_i(1)\phi_j(2)| \quad \rightarrow \quad \phi_{ij}(1,2) + Q_{12}u_{ij}(1,2)$$

  - the first-order pair function u_{ij} may or may not depend explicitly on r_{12}
  - the strong-orthogonality (SO) operator Q_{12} ensures orthogonality to occupied pairs

$$Q_{12} = [1 - P_{\mathrm{occ}}(1)]\,[1 - P_{\mathrm{occ}}(2)]\,, \quad P_{\mathrm{occ}} = \textstyle\sum_i |\phi_i\rangle\langle\phi_i|$$

- The first-order corrections are obtained by minimizing the Hylleraas functional

$$J[u_{ij}] = 2\underbrace{\left\langle u_{ij}\left|Q_{12}r_{12}^{-1}\right|\phi_{ij}\right\rangle}_{\text{3-electron integrals}} + \underbrace{\langle u_{ij}\,|Q_{12}\,(F_1 + F_2 - \varepsilon_i - \varepsilon_j)\,Q_{12}|\,u_{ij}\rangle}_{\text{5(4)-electron integrals}}$$

  - this is Sinanoglu's SO functional

- The MP2 correlation energy may be written as the sum of pair energies

$$E_{\mathrm{corr}} = \textstyle\sum_{ij}\varepsilon_{ij}, \quad \varepsilon_{ij} = \left\langle \tilde{u}_{ij}\left|Q_{12}r_{12}^{-1}\right|\phi_{ij}\right\rangle$$

Møller–Plesset theory—choice of pair function

- In standard orbital-based MP2 theory, we use a CI-type expansion of each pair function:

$$u_{ij} = \sum_{ab} C_{ij}^{ab} \phi_{ab}, \quad Q_{12} u_{ij} = u_{ij}$$

  - strong orthogonality is ensured and only two-electron integrals arise
  - slow convergence

- In explicitly-correlated MP2 theory, the pair functions depend explicitly on r_{12}:

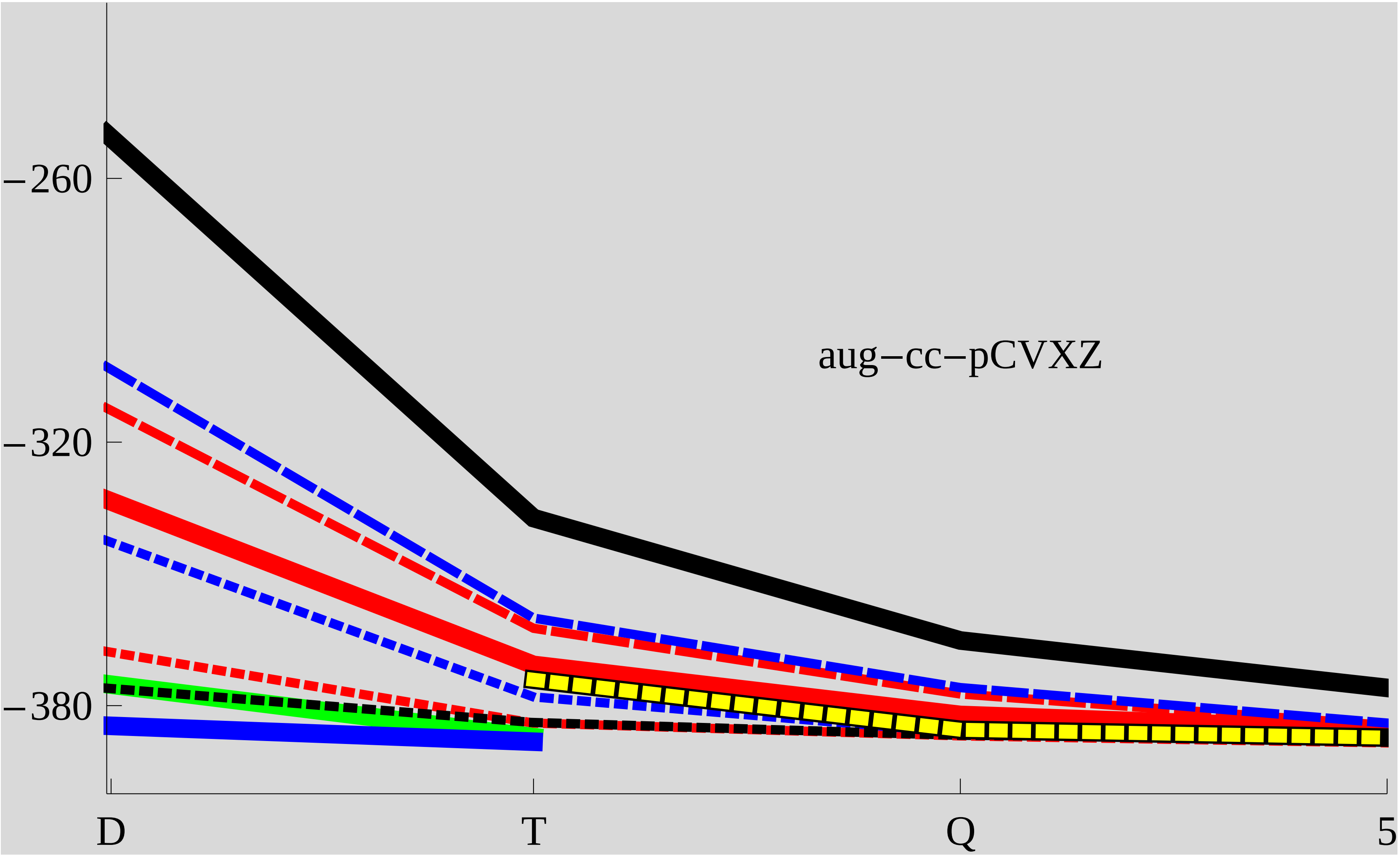
$$u_{ij} = u_{ij}(r_{12}), \quad Q_{12} u_{ij} \neq u_{ij}$$

  - strong orthogonality is not ensured and many-electron integrals arise
  - faster convergence

- A variety of explicitly correlated methods have been developed, depending on
  - the choice of correlation function $u_{ij}(r_{12})$
    - linear correlation function: $r_{12}\phi_{ab}$
    - Gaussian Gaussian function: $\exp\left(-\gamma_v r_{12}^2\right) \phi_{ab}$
  - the treatment of strong orthogonality Q_{12}
    - operators that project out more than Q_{12}
    - operators that project out less than Q_{12} (compensated for by penalty functions)

- The models have been dictated not only by physics but also by difficulties in integration

Explicit correlation—a plethora of methods (neon atom)



Gaussian-type geminal (GTG) theory

- In GTG theory, the pair functions are expanded in Gaussian-type geminals

$$u_{ij} = \mathcal{A}_{\text{anti}} \sum_v c_v \exp\left[-\alpha_v (\mathbf{r}_1 - \mathbf{P}_v)^2 - \beta_v (\mathbf{r}_2 - \mathbf{Q}_v)^2 - \gamma_v r_{12}^2\right] \sigma_1 \sigma_2$$

  - all exponents α_v, β_v, γ_v and centers $\mathbf{P}_v$, $\mathbf{Q}_v$ variationally optimized
  - a difficult nonlinear optimization of pair energies

- To avoid four-electron integrals for such u_{ij}, Szalewicz *et al.* [CPL **91**, 169 (1982)] modified Sinanoglu's SO functional

$$W[u_{ij}] = 2\left\langle u_{ij} \middle| Q_{12} r_{12}^{-1} \middle| \phi_{ij} \right\rangle + \left\langle u_{ij} \middle| \cancel{Q_{12}} \left(\tilde{F}_1 + \tilde{F}_2 - \varepsilon_i - \varepsilon_j\right) \cancel{Q_{12}} \middle| u_{ij} \right\rangle$$

where the shifted Fock operators introduce a penalty and are given by

$$\tilde{F} = F + \eta_{ij} P_{\text{occ}}, \quad \eta_{ij} = \tfrac{1}{2}\left(\varepsilon_i + \varepsilon_j\right) - \varepsilon_1 + \eta, \quad \eta > 0$$

  - only two- and three-electron integrals now remain

- This weak-orthogonality (WO) functional is an upper bound to the SO functional:

$$W[u_{ij}] \geq J[u_{ij}] \geq \varepsilon_{ij}$$

  - equality for the exact first-order pair function only
  - orthogonality controlled by a penalty function—requires a flexible pair function

R12 theory

- In Kutzelnigg's and Klopper's R12 theory, the pair function is taken to have the form

$$u_{ij} = \sum_{ab} C_{ij}^{ab} \phi_{ab} + Q_{12} \sum_{kl} c_{ij}^{kl} r_{12} \phi_{kl}$$

  - a combined conventional and explicitly correlated expansion
  - only parameters (expansion coefficients) are optimized (conventional GTOs are used)
- To avoid three- and four-electron integrals, resolution of identity (RI) is invoked
  - this dramatically improves performance—applicable to large molecules
- Recently, more general correlation factors have been explored (F12 theory)
  - $f(r_{12}) = 1 - \exp(-\gamma r_{12})$ (Ten-no, 1994)
  - $f(r_{12}) = 1 - \sum_v c_v \exp\left(-\gamma_v r_{12}^2\right)$ (May and Manby, 1994)
- Various flavors (Ansätze) of R12/F12 theories exist, depending on:
  - Q_{12} projection against all MOs or only the occupied MOs (Ansätze 1 and 2)
  - the omission or inclusion of certain exchange commutators (Ansätze A and B)
  - the (non)assumption of the extended Brillouin theorem (EBT): $F\phi_a = \varepsilon_a \phi_a$

The GGn model: a mixed GTO–GTG pair-function expansion

- We have investigated a method that combines elements of GTG and R12 theories

$$u_{ij} = \underbrace{\sum_{ab} C_{ij}^{ab} \phi_{ab}}_{\text{GTO part}} + \underbrace{\sum_{pq} \sum_{v} c_{ij}^{pq,v} \exp(-\gamma_v r_{12}^2) \phi_{pq}}_{\text{GTG part}}$$

- In common with GTG theory:
  - we use the WO functional, avoiding 4-electron integrals but not 3-electron integrals
  - we use Gaussian geminals for explicit correlation
- In common with R12 theory:
  - we combine explicit correlation with a traditional GTO expansion
  - we avoid nonlinear optimization (fixed exponents and positions)
- There are three levels of GGn theory, depending on what geminals are included:

GG0: include only “ground-state” geminals $\exp(-\gamma_v r_{12}^2)\phi_{ij}$

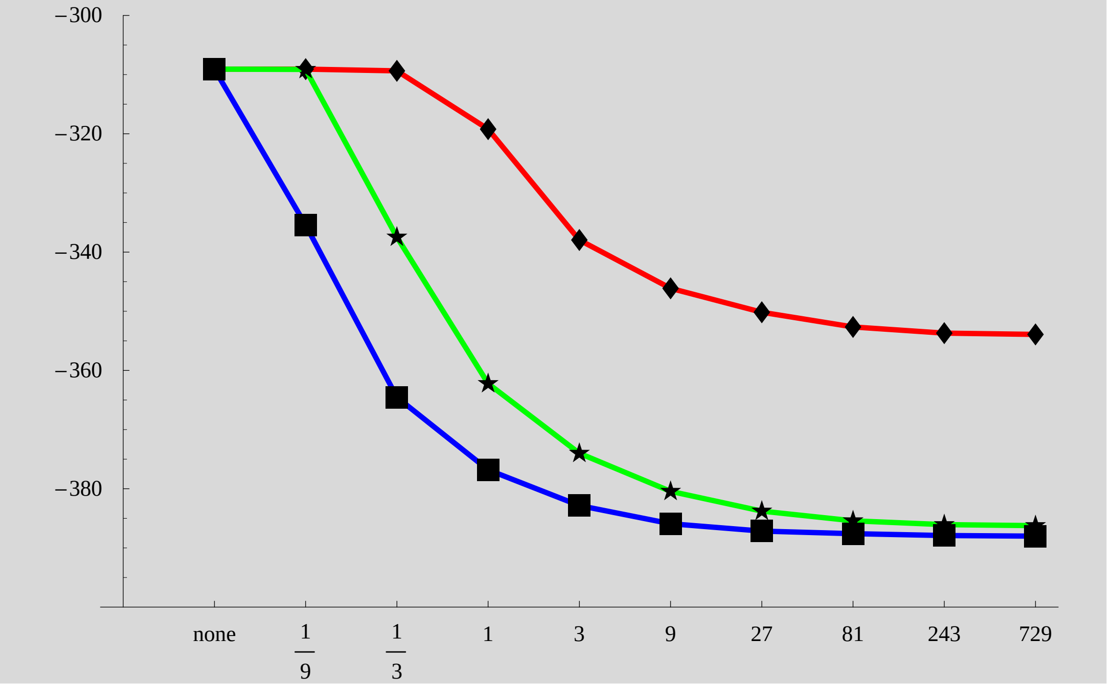
GG1: include also “singly-excited” geminals $\exp(-\gamma_v r_{12}^2)\phi_{ai}$

GG2: include also “doubly-excited” geminals $\exp(-\gamma_v r_{12}^2)\phi_{ab}$

- GREMLIN code written by Pål Dahle (DALTON module)

GTG exponents

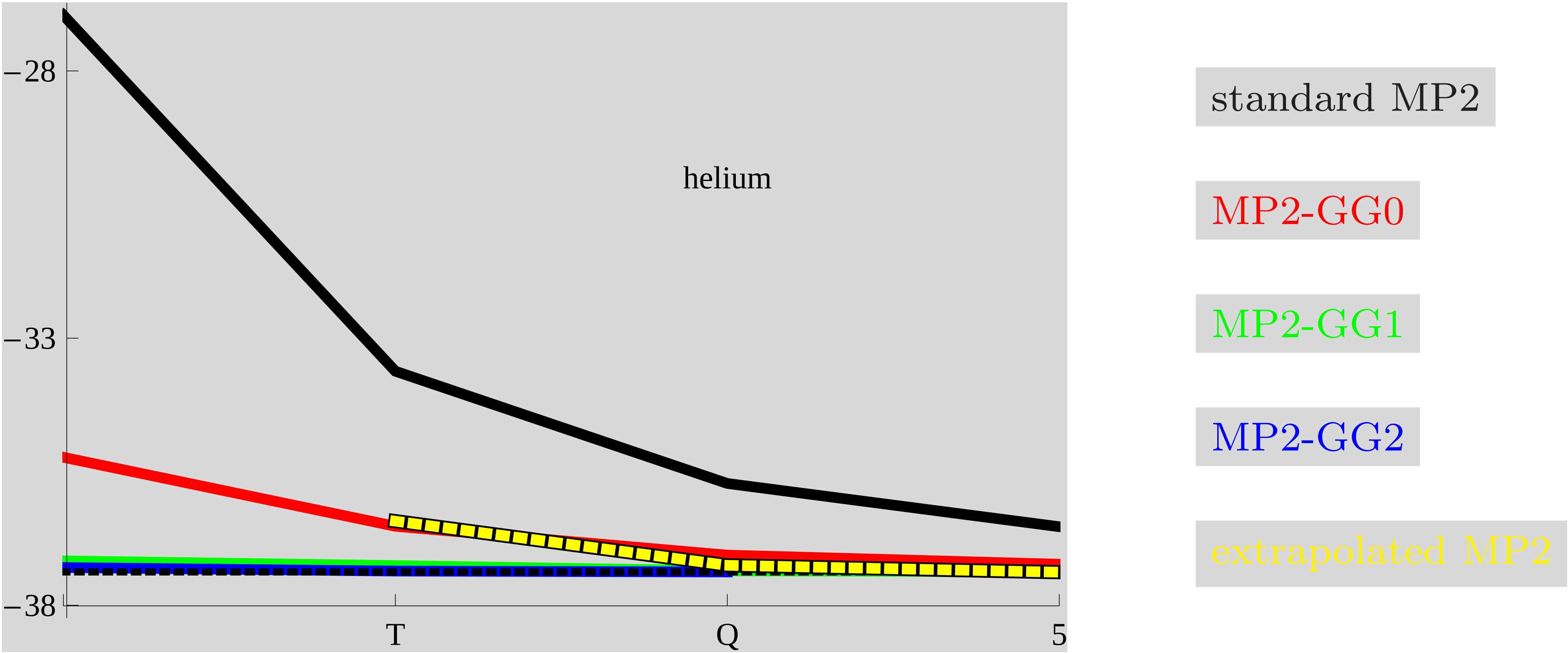
- We use nine even-tempered GTG exponents ($\gamma_v = 1/9,\ 1/3,\ 1,\ 3,\ 9,\ 27,\ 81,\ 243,\ 729$)
  - all-electron GG0, GG1, GG2 aug-cc-pCVTZ(spd) neon calculations



  - for GG0, only GTGs with $\gamma \geq 1$ are important
  - diffuse GTGs are important for excited GTGs (for GG1 and GG2)
  - the three steepest GTGs improve mainly the $1s^2$ energy
  - with all nine GTGs included, the energy is converged to within 0.1 mH

Helium

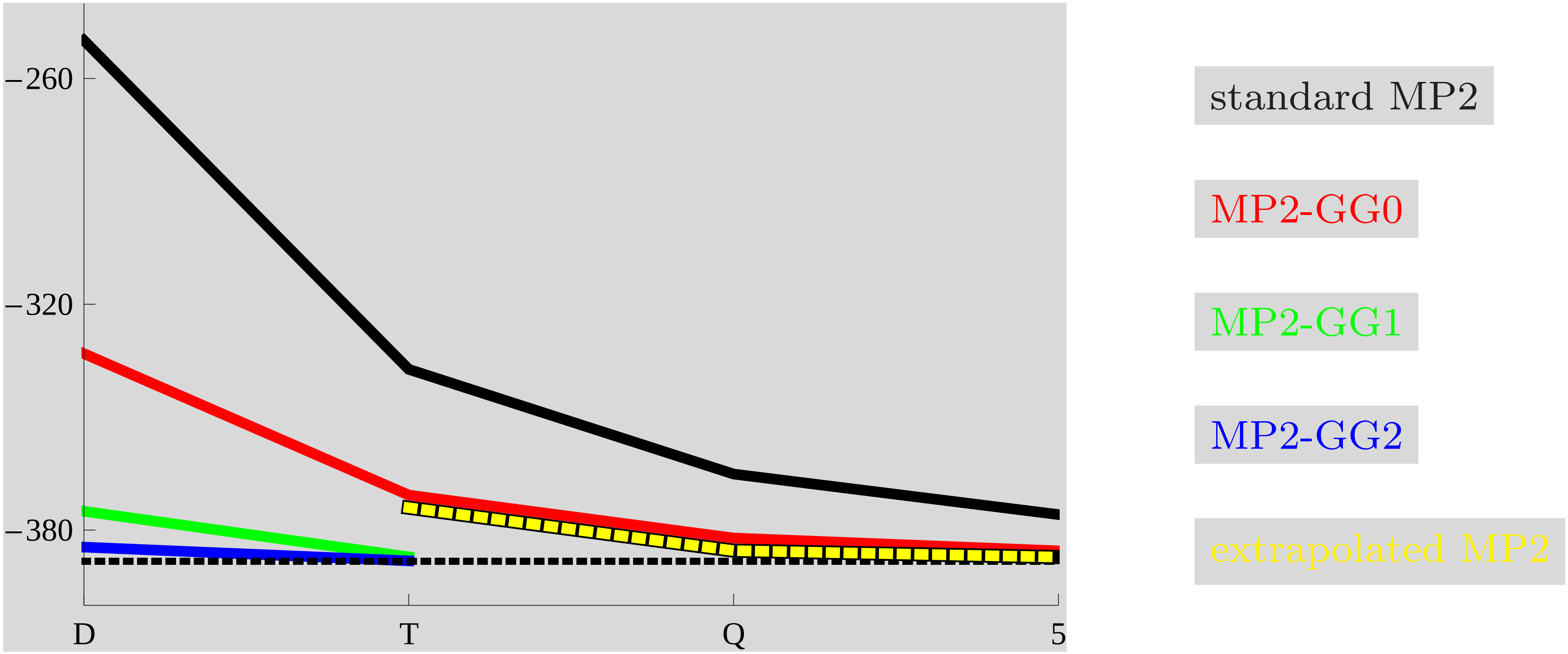
- A comparison of MP2-GGn/aug-cc-pVXZ with standard and extrapolated MP2 theory



- As expected, the use of geminals improves convergence considerably
  - errors reduced by an order of magnitude at the GG1 and GG2 levels of theory
  - note: the GG0 error reduction is no better than extrapolation!
- Our results are very close to the nonlinear GTG theory:
  - GG2: −37.3773 mH; Patkowski *et al.*: −37.3775 mH; extrapolation: −37.36 mH

Neon

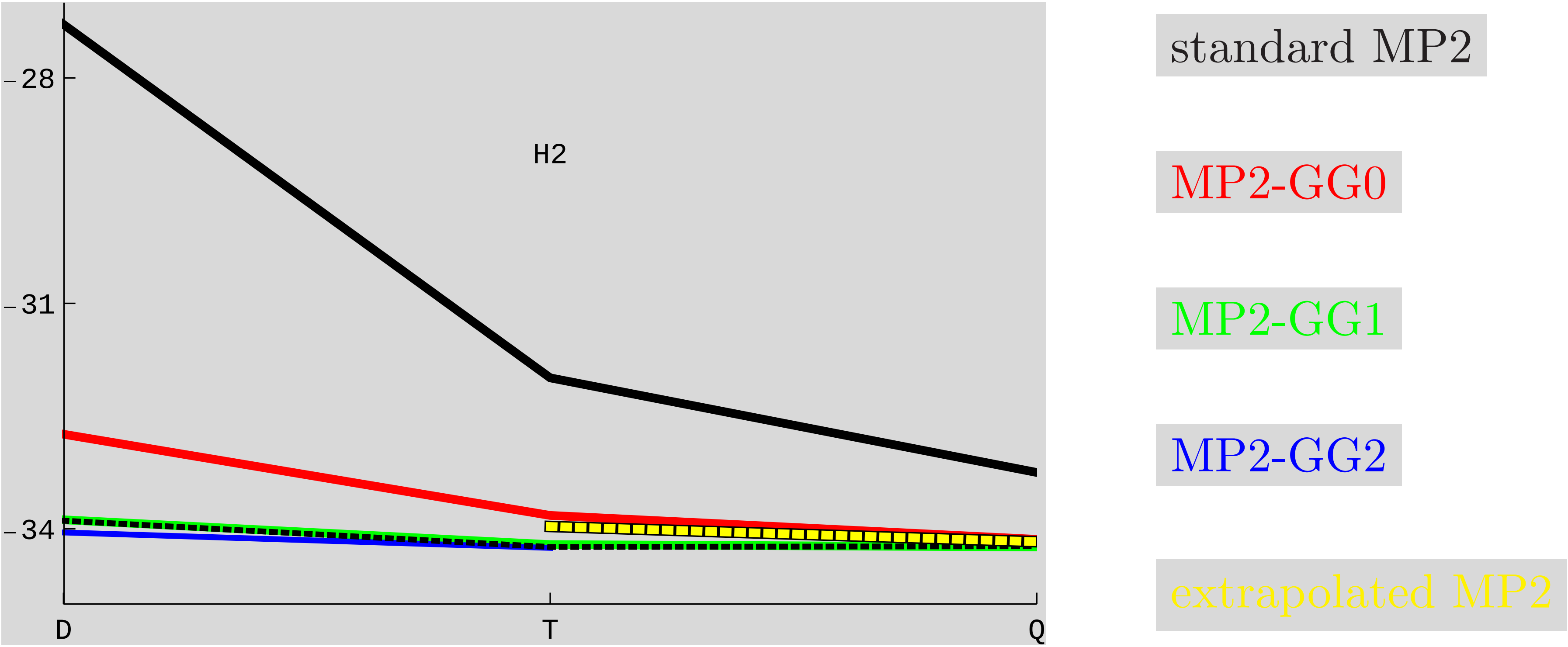
- A similar behaviour is observed for the neon atom in the aug-cc-pCVXZ basis:



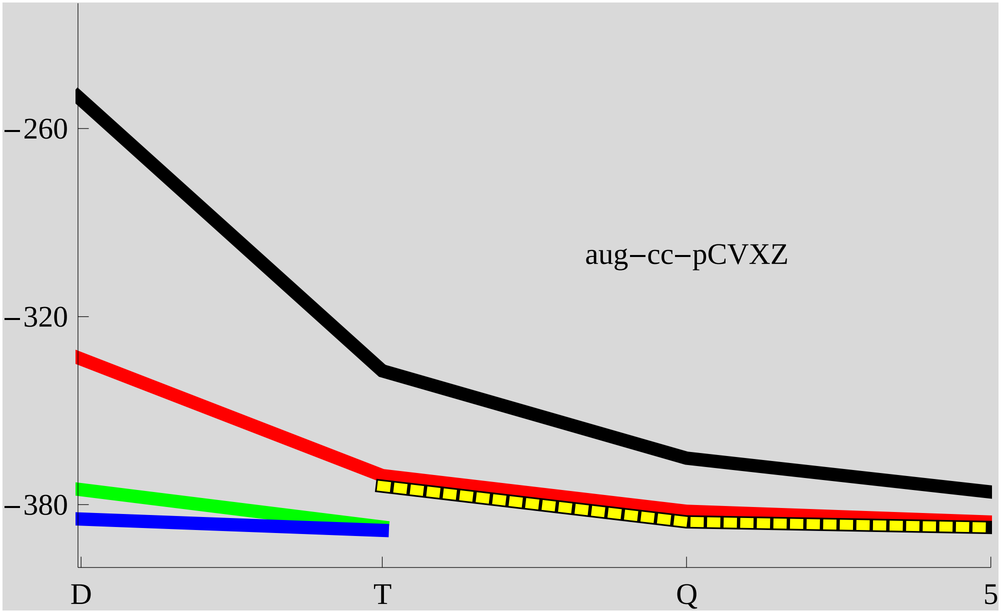
- GG0 theory reduces error by a factor of three or four relative to standard MP2 theory
  - but basis-set extrapolation gives similar results
- The GG2/aug-cc-pCVTZ value is −388.19 mH
  - this is the lowest variationally bounded literature value
  - 0.25% (1 mH) below the best extrapolated value

Hydrogen molecule

- The same pattern is observed for other systems
  - H_2 in aug-pVXZ basis

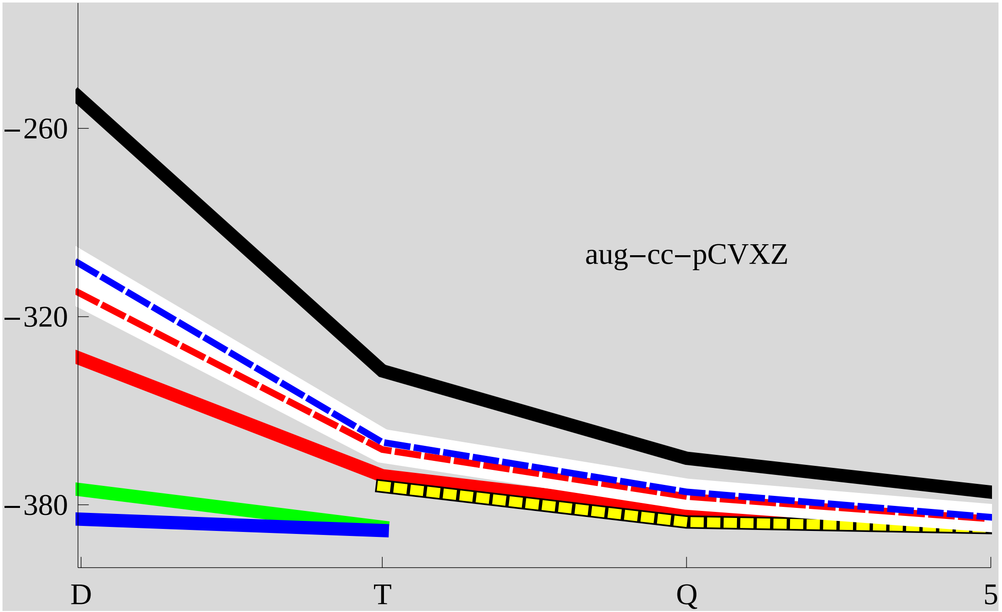


Neon: a comparison with R12/F12 methods



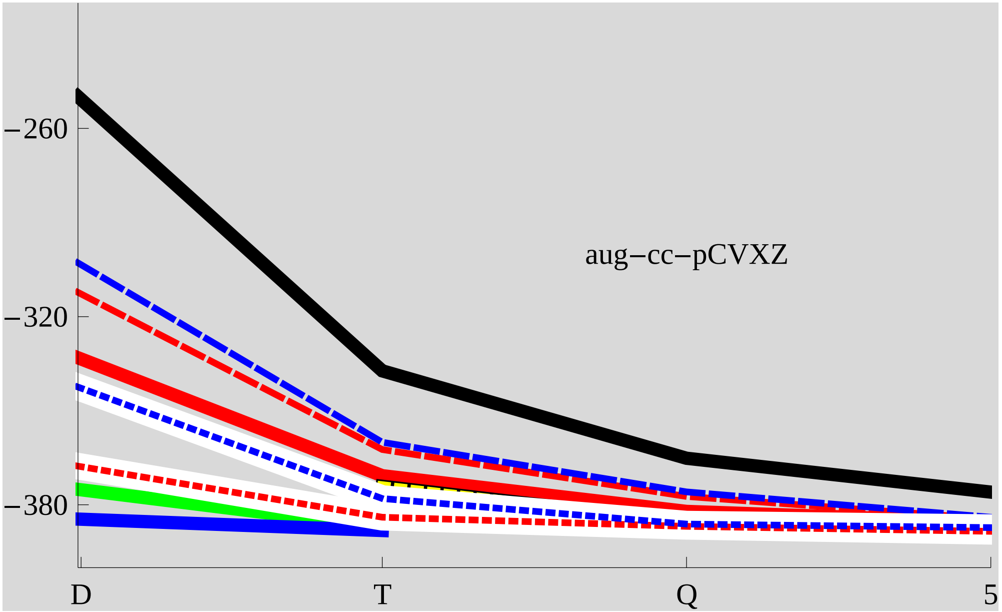
- R12 calculations: Klopper and Samson, JCP **116**, 6397 (2002)
  - Ansatz **1**: all MO pairs projected out; Ansatz **2**: occupied MO pairs projected out
  - Ansatz **A**: exchange commutator $[K, r_{12}]$ neglected; Ansatz **B**: $[K, r_{12}]$ included
- MP2-geminal calculations: Ten-no, JCP **121**, 117 (2004)
  - GTGs fitted to $\exp(-\zeta r_{12})$, numerical quadrature, $[K, r_{12}]$ neglected, EBC assumed
- The R12/2A', R12/2B and MP2-geminal energies are lower than the GG0 energy
  - MP2-geminal close to MP2-GG1! WO penalty? error cancellation?

Neon: a comparison with R12/F12 methods



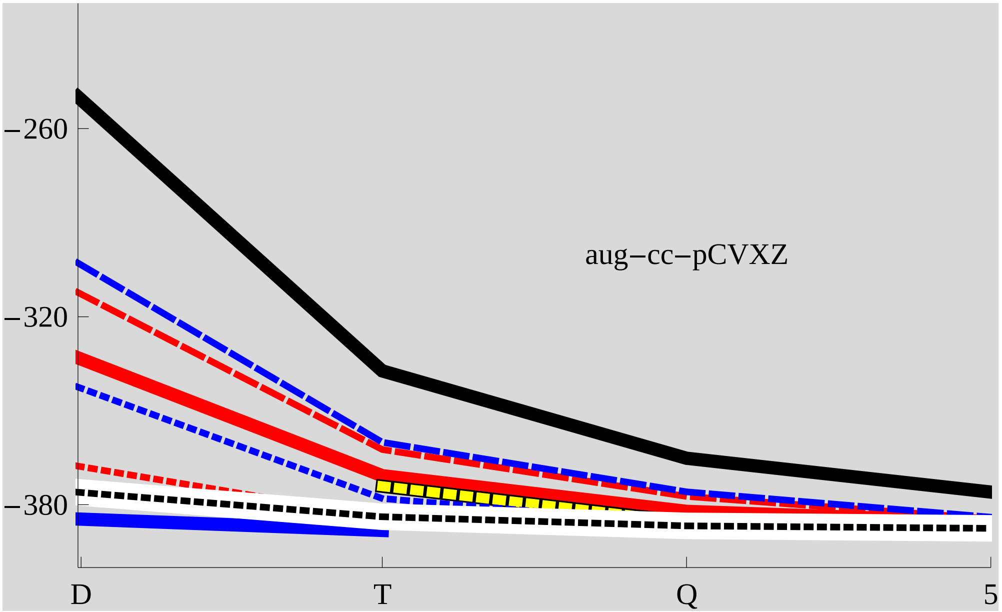
- R12 calculations: Klopper and Samson, JCP **116**, 6397 (2002)
  - Ansatz **1**: all MO pairs projected out ; Ansatz **2**: occupied MO pairs projected out
  - Ansatz **A**: exchange commutator $[K, r_{12}]$ neglected; Ansatz **B**: $[K, r_{12}]$ included
- MP2-geminal calculations: Ten-no, JCP **121**, 117 (2004)
  - GTGs fitted to $\exp(-\zeta r_{12})$, numerical quadrature, $[K, r_{12}]$ neglected, EBC assumed
- The R12/2A', R12/2B and MP2-geminal energies are lower than the GG0 energy
  - MP2-geminal close to MP2-GG1! WO penalty? error cancellation?

Neon: a comparison with R12/F12 methods



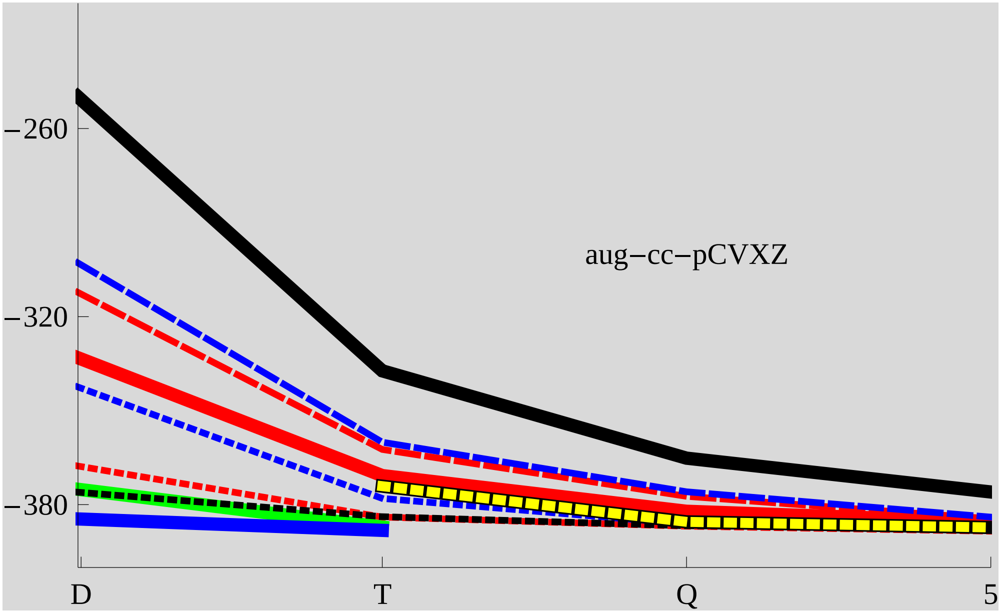
- R12 calculations: Klopper and Samson, JCP **116**, 6397 (2002)
  - Ansatz **1**: all MO pairs projected out; Ansatz **2**: occupied MO pairs projected out
  - Ansatz **A**: exchange commutator $[K, r_{12}]$ neglected; Ansatz **B**: $[K, r_{12}]$ included
- MP2-geminal calculations: Ten-no, JCP **121**, 117 (2004)
  - GTGs fitted to $\exp(-\zeta r_{12})$, numerical quadrature, $[K, r_{12}]$ neglected, EBC assumed
- The R12/2A', R12/2B and MP2-geminal energies are lower than the GG0 energy
  - MP2-geminal close to MP2-GG1! WO penalty? error cancellation?

Neon: a comparison with R12/F12 methods



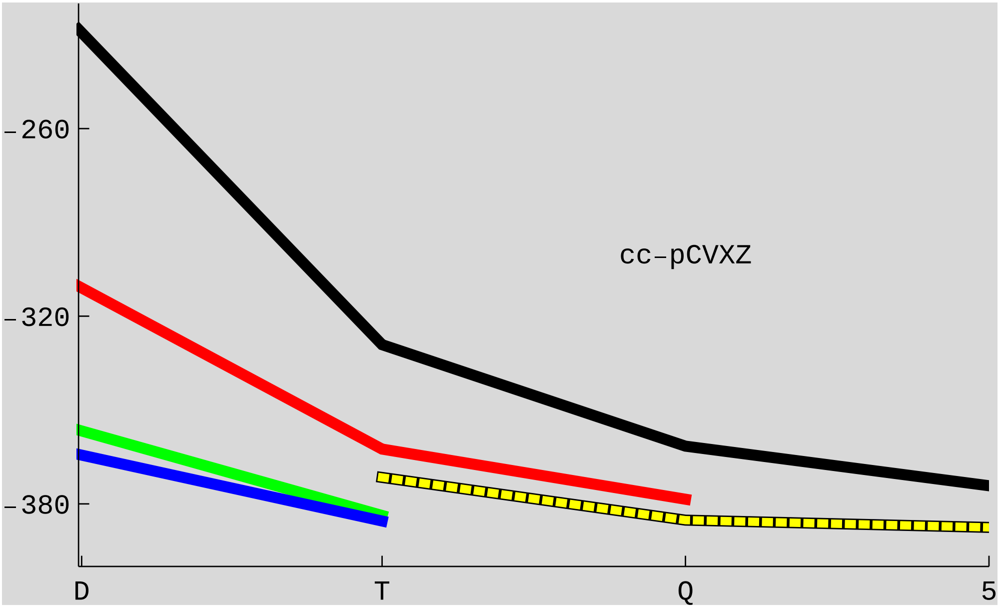
- R12 calculations: Klopper and Samson, JCP **116**, 6397 (2002)
  - Ansatz **1**: all MO pairs projected out; Ansatz **2**: occupied MO pairs projected out
  - Ansatz **A**: exchange commutator $[K, r_{12}]$ neglected; Ansatz **B**: $[K, r_{12}]$ included
- MP2-geminal calculations: Ten-no, JCP **121**, 117 (2004)
  - GTGs fitted to $\exp(-\zeta r_{12})$, numerical quadrature, $[K, r_{12}]$ neglected, EBC assumed
- The R12/2A', R12/2B and MP2-geminal energies are lower than the GG0 energy
  - MP2-geminal close to MP2-GG1! WO penalty? error cancellation?

Neon: a comparison with R12/F12 methods



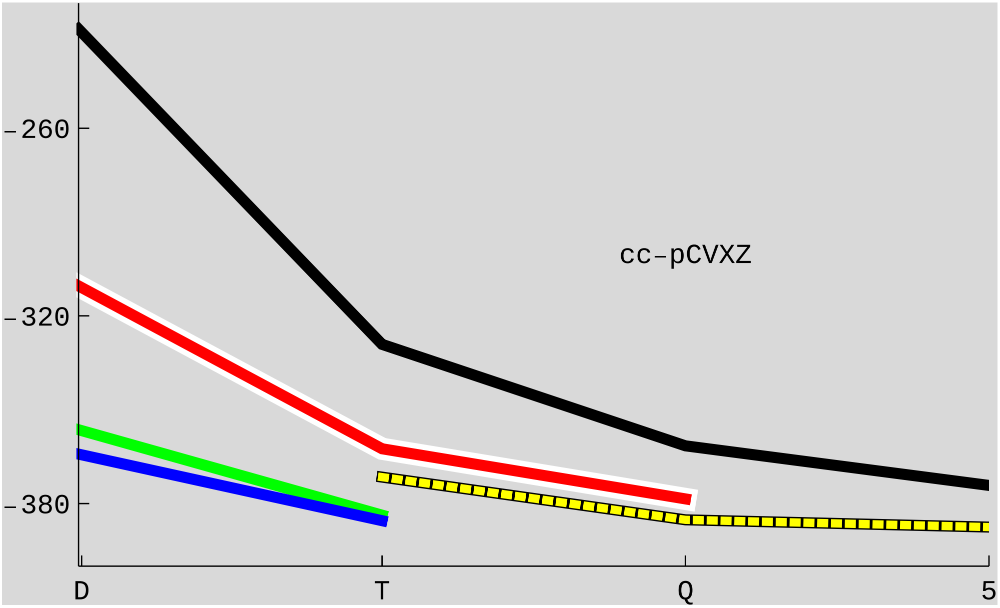
- R12 calculations: Klopper and Samson, JCP **116**, 6397 (2002)
  - Ansatz **1**: all MO pairs projected out; Ansatz **2**: occupied MO pairs projected out
  - Ansatz **A**: exchange commutator $[K, r_{12}]$ neglected; Ansatz **B**: $[K, r_{12}]$ included
- MP2-geminal calculations: Ten-no, JCP **121**, 117 (2004)
  - GTGs fitted to $\exp(-\zeta r_{12})$, numerical quadrature, $[K, r_{12}]$ neglected, EBC assumed
- The R12/2A', R12/2B and MP2-geminal energies are lower than the GG0 energy
  - MP2-geminal close to MP2-GG1! WO penalty? error cancellation?

Neon: sorting out the models



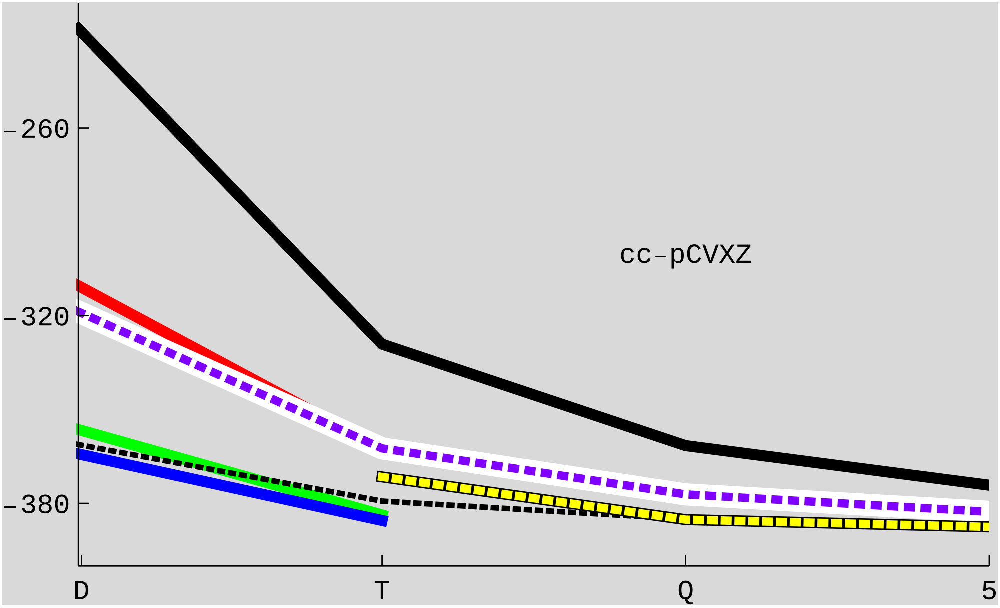
- GG0 converges slowly, slightly poorer than extrapolation
  - Gaussian correlation factor and WO functional
- R12-SO converges no faster than GG0
  - linear correlation factor and SO functional
- Ten-no's MP2-geminal model converges much faster—as fast as GG1
  - Gaussian correlation function and SO functional
- Conclusion: the WO functional and R12 factor are both poor in small basis sets

Neon: sorting out the models



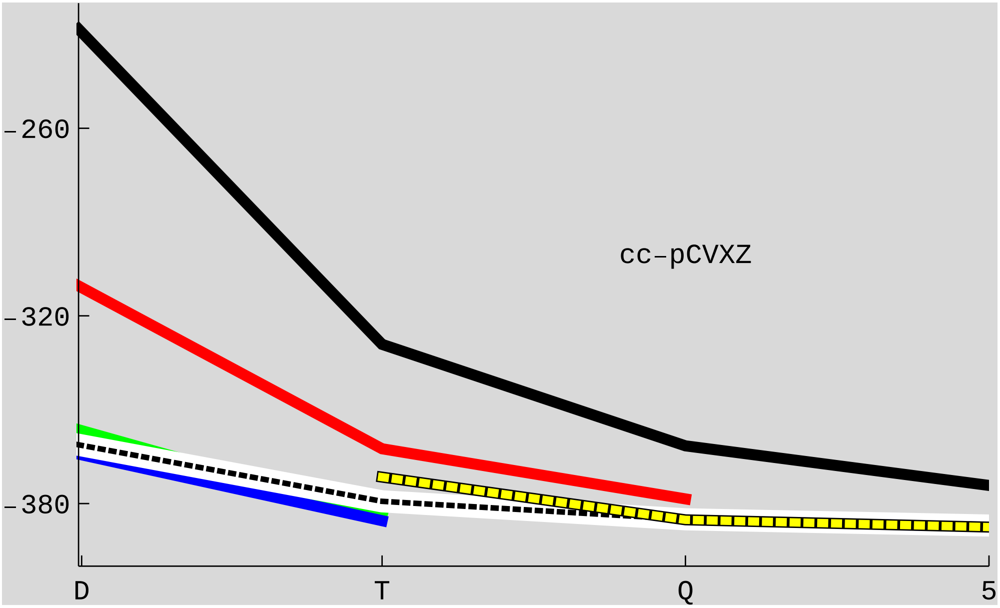
- GG0 converges slowly, slightly poorer than extrapolation
    - – Gaussian correlation function and WO functional
- R12-SO converges no faster than GG0
    - – linear correlation factor and SO functional
- Ten-no's MP2-geminal model converges much faster—as fast as GG1
    - – Gaussian correlation function and SO functional
- Conclusion: the WO functional and R12 factor are both poor in small basis sets

Neon: sorting out the models



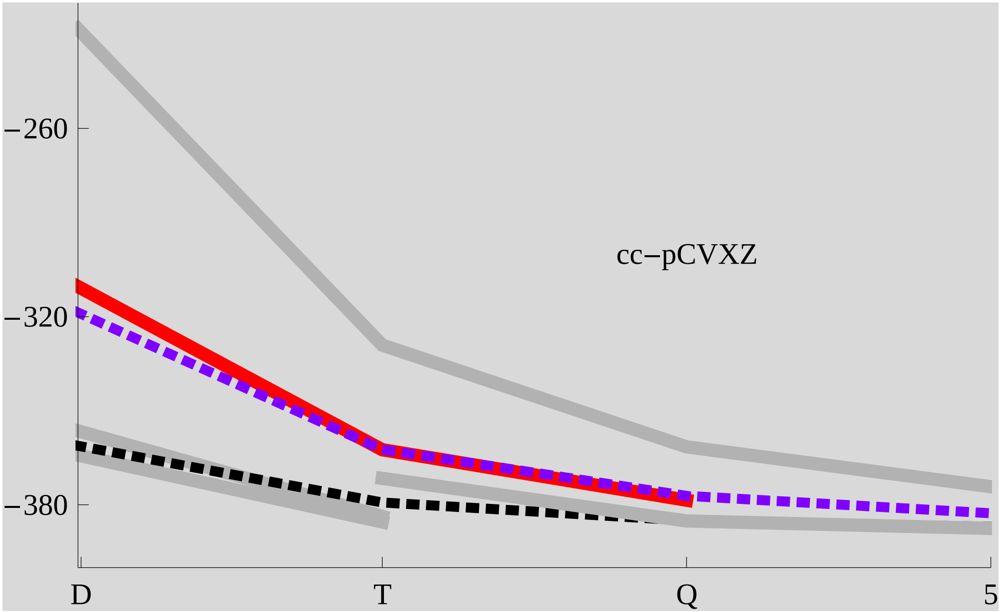
- GG0 converges slowly, slightly poorer than extrapolation
  - Gaussian correlation function and WO functional
- R12-SO converges no faster than GG0
  - linear correlation factor and SO functional
- Ten-no's MP2-geminal model converges much faster—as fast as GG1
  - Gaussian correlation function and SO functional
- Conclusion: the WO functional and R12 factor are both poor in small basis sets

Neon: sorting out the models



- GG0 converges slowly, slightly poorer than extrapolation
  - Gaussian correlation function and WO functional
- R12-SO converges no faster than GG0
  - linear correlation factor and SO functional
- Ten-no's MP2-geminal model converges much faster—as fast as GG1
  - Gaussian correlation function and SO functional
- Conclusion: the WO functional and R12 factor are both poor in small basis sets

Neon: sorting out the models



- GG0 converges slowly, slightly poorer than extrapolation
  - Gaussian correlation function and WO functional
- R12-SO converges no faster than GG0
  - linear correlation factor and SO functional
- Ten-no's MP2-geminal model converges much faster—as fast as GG1
  - Gaussian correlation function and SO functional
- Conclusion: the WO functional and R12 factor are both poor in small basis sets?

Some conclusions

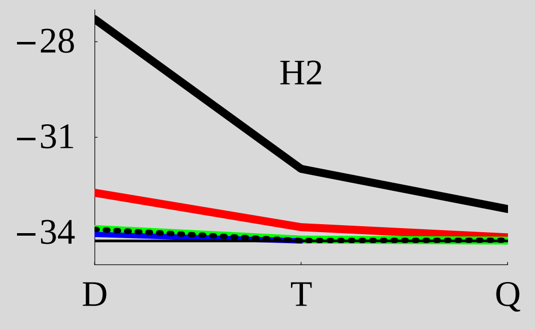
- Indeed, May *et al.* [PCCP **7**, 2710 (2005)] suggested a problem with the r_{12} factor
  - advocated the use of linear combinations of Gaussian geminals rather than r_{12}
- Very recent Tew, Klopper and Manby studied the WO penalty
  - large penalty for GG0 only—GG1, GG2 and GTG can model orthogonality
- One should probably avoid both linear r_{12} and WO with small basis sets
- All-electron correlation energies of Ne, HF, and H_2O in aug-cc-pCVXZ basis sets

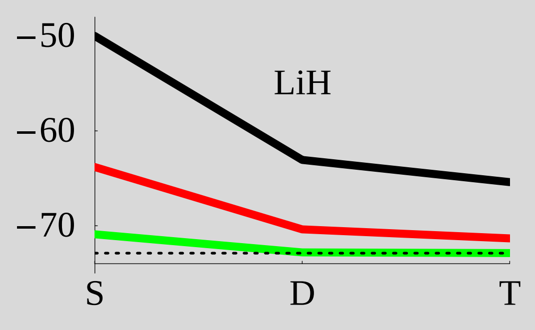
	std	ext	GG0	GG1	GG2	2A'	2B	Ten-no
DZ	68.2		87.8	97.2	99.2	96.2	91.1	96.9
TZ	88.3	96.8	96.3	99.8	100.0	99.1	98.0	99.1
QZ	94.8	99.6	98.5			99.8	99.4	99.7

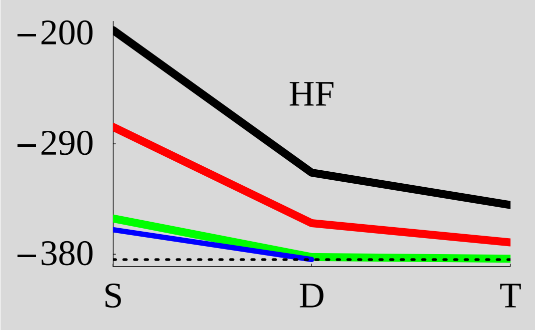
  - blue: more than 99% of the correlation energy has been recovered for all systems
- Extrapolation works well, recovering more than 99% at the aug-cc-pCV[TQ]Z level
  - the F12 methods outperform GG0—in particular, in a small basis
  - GG2 recovers 99% correlation energy in DZ basis and 100% in TZ basis
  - QZ basis needed to recover consistently more than 99% correlation energy

Small molecules

- The molecular correlation energies show the same performance







- We obtained the lowest ever energies for H_2 and HF
  - H_2: −34.252 mH with GG2/aug-cc-pVTZ
  - LiH:−72.877 mH with GG1/($14s9p4d3f/8s4p3d$) [72.890 mH by Bukowski *et al.*]
  - HF: −384.41 mH with GG2/aug-cc-pCVTZ(spd,spd)
- Basis sets
  - H_2: aug-cc-pVXZ
  - LiH: uncontracted ANO $(14s, 8s)$, $(14s9p, 8s4p)$, $(14s9p4d3f, 8s4p3d)$
  - HF: aug-cc-pCVTZ subspaces (sp, s), (sdp, sp) (sdf, spd)

MP2 correlation energies of small systems (mH)

system	this work[a]	current best	energy recovered
He	37.37729	37.37747[b]	99.9995%
Be	76.355	76.358[c]	99.996%
Ne	388.19	388.19	100%
H_2	34.252	34.252	100%
LiH	72.877	72.890[c]	99.98%
HF	384.41	384.41	100%

[a] GG2/TZ calculations except GG1/TZ for LiH

[b] Patkowski, Bukowski, Jeziorski and Szalewicz, personal communication

[c] Bukowski, Jeziorski and Rutkowski, JCP **110**, 4165 (1999)

Conclusions

- There are two solutions to the basis-set problem of orbital-based quantum chemistry
  - extrapolation techniques
  - explicitly correlated methods
  - both can deliver an error reduction by an order of magnitude or more
- The modern development of explicit correlation began with GTG methods
  - high accurate but applicable only to small systems (nonlinear optimization)
- With R12 theory, explicitly correlated methods became (almost) routine
  - many-electron integrals avoided by RI, applicable too large systems
- The GGn approach combines elements of both these approaches
  - WO functional (variationally bounded), explicit three-electron integration
  - the GG0 model recovers less correlation energy than does F12 theory
  - the GG1 and GG2 models recover more than 99% correlation energy in a TZ basis
  - in general QZ basis is needed for such an accuracy