

2007 Summer School on Computational Materials Science

Quantum Monte Carlo: From Minerals and Materials to Molecules

July 9 –19, 2007 • University of Illinois at Urbana–Champaign

<http://www.mcc.uiuc.edu/summerschool/2007/qmc/>



Order(N) methods in QMC

Dario Alfè

[d.alfè@ucl.ac.uk]

Monte Carlo methods

- Evaluation of multidimensional integrals:
 - Quadrature methods, error $\sim M^{-4/d}$
 - Monte Carlo methods, error $\sim M^{-0.5}$

$$I = \int d\mathbf{R} g(\mathbf{R})$$

$$\wp(\mathbf{R}) \geq 0; \quad \int d\mathbf{R} \wp(\mathbf{R}) = 1$$

$$I = \int d\mathbf{R} f(\mathbf{R}) \wp(\mathbf{R}); \quad f(\mathbf{R}) = g(\mathbf{R}) / \wp(\mathbf{R})$$

$$I = \lim_{M \rightarrow \infty} \left\{ \frac{1}{M} \sum_{m=1}^M f(\mathbf{R}_m) \right\} \approx \frac{1}{M} \sum_{m=1}^M f(\mathbf{R}_m)$$

Variational Monte Carlo

$$E_V = \frac{\int \Psi_T^*(\mathbf{R}) \hat{H} \Psi_T(\mathbf{R}) d\mathbf{R}}{\int \Psi_T^*(\mathbf{R}) \Psi_T(\mathbf{R}) d\mathbf{R}} \geq E_0$$

$$E_V = \frac{\int |\Psi_T(\mathbf{R})|^2 \left[\Psi_T(\mathbf{R})^{-1} \hat{H} \Psi_T(\mathbf{R}) \right] d\mathbf{R}}{\int |\Psi_T(\mathbf{R})|^2 d\mathbf{R}}$$

$$\wp(\mathbf{R}) = |\Psi_T(\mathbf{R})|^2 / \int |\Psi_T(\mathbf{R})|^2 d\mathbf{R}$$

$$E_V \approx \frac{1}{M} \sum_{m=1}^M E_L(\mathbf{R}_m); \quad E_L(\mathbf{R}_m) = \Psi_T(\mathbf{R}_m)^{-1} \hat{H} \Psi_T(\mathbf{R}_m)$$

Diffusion Monte Carlo

$$-\frac{\partial \Psi(\mathbf{R}, t)}{i\partial t} = (\hat{H} - E_T) \Psi(\mathbf{R}, t)$$

Extracting the ground state: substitute $\tau = it$

$$\tau \rightarrow \infty, \quad \Psi(\mathbf{R}, \tau) \rightarrow \Phi_0(\mathbf{R})$$

Fixed nodes:

$$\tau \rightarrow \infty, \quad \Psi(\mathbf{R}, \tau) \rightarrow \Phi_0^{FN}(\mathbf{R})$$

$$\Psi_T(\mathbf{R}) = \exp[J(\mathbf{R})] \sum_n c_n D_n^\uparrow \{\phi_i(\mathbf{r}_j)\} D_n^\downarrow \{\phi_i(\mathbf{r}_j)\}$$

Cost of evaluating Ψ_T

$$\Psi_T(\mathbf{R}) = \exp[J(\mathbf{R})] \sum_n c_n D_n^\uparrow \{\phi_i(\mathbf{r}_j)\} D_n^\downarrow \{\phi_i(\mathbf{r}_j)\}$$

- N electrons
- N single particle orbitals ϕ_i
- \sim N basis functions for each ϕ_i if

$$\phi_i(\mathbf{r}_j) = \sum_{\mathbf{G}} c_{\mathbf{G}} \exp\{-i\mathbf{G} \cdot \mathbf{r}_j\}$$

Cost proportional to \mathbf{N}^3

Cost of evaluating Energy E

Variance of E is proportional to $N \rightarrow$

Cost of total energy proportional to N^4

HOWEVER

Variance of E/atom is proportional to $1/N \rightarrow$

Cost of energy/atom proportional to N^2

(relevant to free energies in phase transitions, surface energies,)

Cost of evaluating Ψ_T

$$\Psi_T(\mathbf{R}) = \exp[J(\mathbf{R})] \sum_n c_n D_n^\uparrow \{\phi_i(\mathbf{r}_j)\} D_n^\downarrow \{\phi_i(\mathbf{r}_j)\}$$

- N electrons
- N single particle orbitals ϕ_i
- \sim N basis functions for each ϕ_i if

$$\phi_i(\mathbf{r}_j) = \sum_{\mathbf{G}} c_{\mathbf{G}} \exp\{-i\mathbf{G} \cdot \mathbf{r}_j\}$$

Cost proportional to \mathbf{N}^3

Cost of evaluating Ψ_T

$$\Psi_T(\mathbf{R}) = \exp[J(\mathbf{R})] \sum_n c_n D_n^\uparrow \{\phi_i(\mathbf{r}_j)\} D_n^\downarrow \{\phi_i(\mathbf{r}_j)\}$$

- N electrons
- N single particle orbitals ϕ_i
- $\sim o(1)$ basis functions for each ϕ_i if

$$\phi_i(\mathbf{r}_j) = \sum_l c_l f_l(\mathbf{r}_j) \quad \text{if } f_l(\mathbf{r}) \text{ is localised}$$

Cost proportional to \mathbf{N}^2

Cost of evaluating Energy E

Variance of E is proportional to $N \rightarrow$

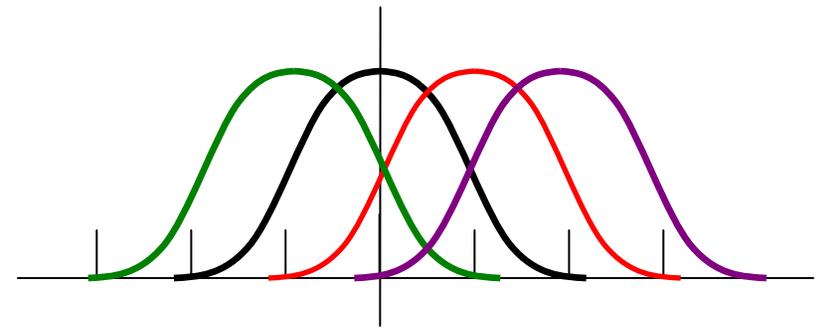
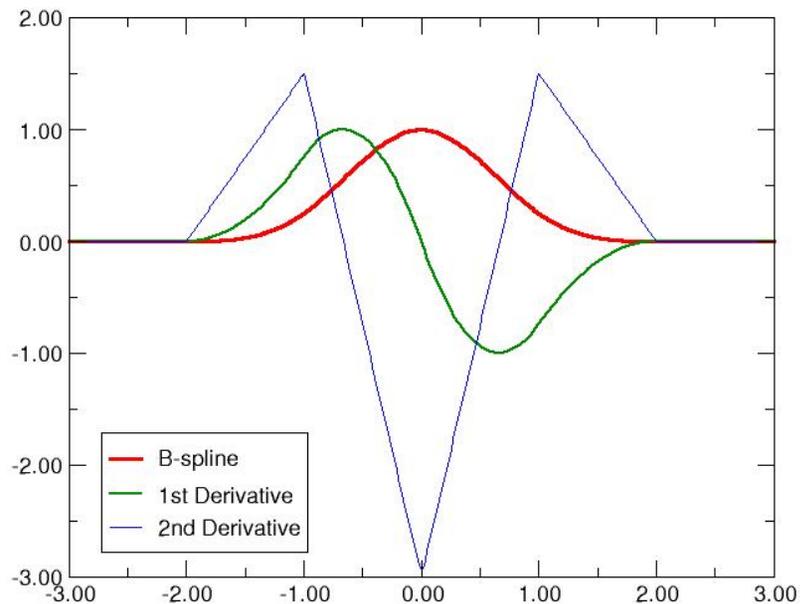
Cost of total energy proportional to N^3

HOWEVER

Variance of E/atom is proportional to $1/N \rightarrow$

Cost of energy/atom proportional to N
(relevant to free energies in phase transitions, surface energies,)

B-splines: Localised functions sitting at the points of a uniform grid



$$f(x) = 1 - \frac{3}{2}x^2 + \frac{3}{4}|x|^3 \quad 0 \leq |x| \leq 1$$

$$= \frac{1}{4}(2 - |x|)^3 \quad 1 \leq |x| \leq 2$$

E. Hernández, M. J. Gillan and C. M. Goringe, Phys. Rev. B, **20** (1997)

D. Alfè and M. J. Gillan, Phys. Rev. B, Rapid Comm., **70**, 161101, (2004)

Cost of evaluating Ψ_T

$$\Psi_T(\mathbf{R}) = \exp[J(\mathbf{R})] \sum_n c_n D_n^\uparrow \{\phi_i(\mathbf{r}_j)\} D_n^\downarrow \{\phi_i(\mathbf{r}_j)\}$$

- N electrons
- N single particle orbitals ϕ_i
- $\sim o(1)$ basis functions for each ϕ_i if

$$\phi_i(\mathbf{r}_j) = \sum_l c_l f_l(\mathbf{r}_j)$$

Cost proportional to \mathbf{N}^2

Cost of evaluating Ψ_T

$$\Psi_T(\mathbf{R}) = \exp[J(\mathbf{R})] \sum_n c_n D_n^\uparrow \{\phi_i(\mathbf{r}_j)\} D_n^\downarrow \{\phi_i(\mathbf{r}_j)\}$$

- N electrons
- $\sim o(1)$ single particle localised orbitals ϕ_i , if ϕ_i is localised
- $\sim o(1)$ basis functions for each ϕ_i if

$$\phi_i(\mathbf{r}_j) = \sum_l c_l f_l(\mathbf{r}_j)$$

Cost proportional to **N** (Linear scaling)

Cost of evaluating Energy E

Variance of E is proportional to $N \rightarrow$

Cost of total energy proportional to N^2

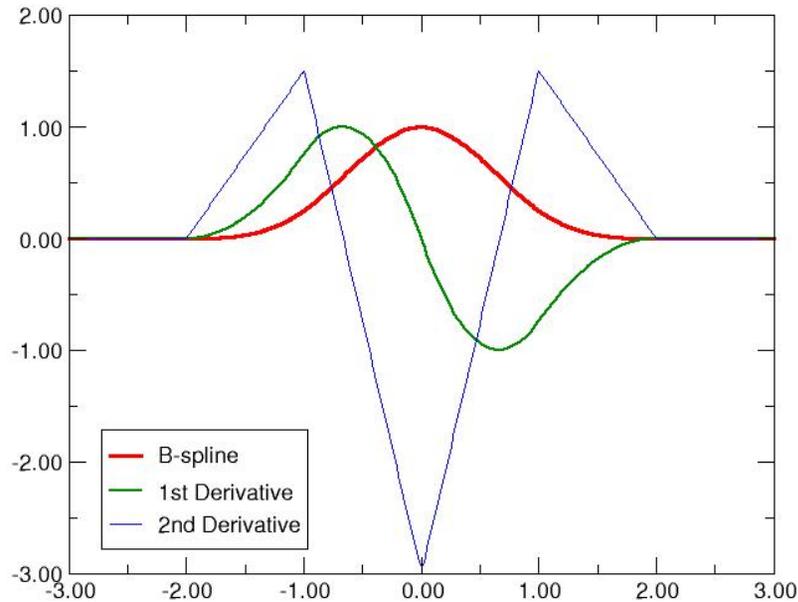
HOWEVER

Variance of E/atom is proportional to $1/N \rightarrow$

Cost of energy/atom independent on $N!$

(relevant to free energies in phase transitions, surface energies,)

B-splines (blips)



- Defined on a uniform grid
- Localised: $f(x)=0$ for $|x|>2$
- Continuous with first and second derivative continuous

Grid spacing 1:

$$f(x) = 1 - \frac{3}{2}x^2 + \frac{3}{4}|x|^3 \quad 0 \leq |x| \leq 1$$

$$= \frac{1}{4}(2 - |x|)^3 \quad 1 \leq |x| \leq 2$$

Grid spacing a :

$$X_i = ia$$

$$\Theta(x - X_i) = f((x - X_i) / a)$$

Blips in three dimensions:

$$\Theta(\mathbf{r} - \mathbf{R}_i) = \Theta(x - X_i)\Theta(y - Y_i)\Theta(z - Z_i)$$

Single-particle orbital representation:

$$\psi_n(\mathbf{r}) = \sum_i a_{ni} \Theta(\mathbf{r} - \mathbf{R}_i)$$

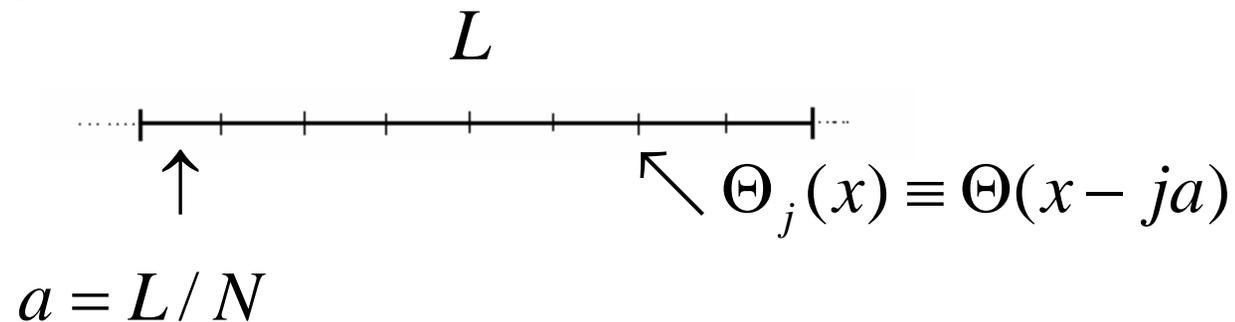
For each position r there are only 64 blip functions that are non zero. By contrast, in a plane wave representation:

$$\psi_n(\mathbf{r}) = \sum_{\mathbf{G}} c_{n\mathbf{G}} e^{i\mathbf{G}\cdot\mathbf{r}}$$

The number of plane waves in MgO is ~ 3000 per atom!
(HF pseudopotentials, 100 Ha PW cutoff)

Approximate equivalence between blips and plane waves

One dimension:



$$\phi_n(x) = e^{2\pi i n x / L} = e^{i k_n x} \quad k_n = 2\pi n / L; \quad k_{\max} = \pi / a$$

$$\chi_n(x) = \sum_{j=0}^{N-1} \Theta_j(x) e^{2\pi i j n / N} \quad -\frac{1}{2} N + 1 \leq n \leq \frac{1}{2} N$$

$$\langle \phi_n | \chi_m \rangle = \int_0^L dx \sum_{j=0}^{N-1} \theta_j(x) e^{2\pi i j m / N} e^{-i k_n x} = \quad k_n = \frac{2\pi n}{aN}$$

$$= \sum_{j=0}^{N-1} e^{2\pi i (m-n) j / N} \int_0^L \theta_j(x) e^{2\pi i n (j-x/a) / N} dx = (\text{use PBC}) x - ja \rightarrow x, \Theta_j \rightarrow \Theta_0$$

$$= \sum_{j=0}^{N-1} e^{2\pi i (m-n) j / N} \int_0^L \theta_0(x) e^{-2\pi i n x / aN} dx = N \delta_{nm} \int_0^L f(x/a) e^{-2\pi i n x / aN} dx =$$

$$= N \delta_{nm} a \int_0^{L/a} f(y) e^{-2\pi i n y / N} dy = L \delta_{nm} \mathcal{F}(2\pi n / N)$$

$$\mathcal{F}(q) = \int_{-\infty}^{+\infty} f(x) e^{-iqx} dx = \int_{-2}^{+2} f(x) e^{-iqx} dx = \frac{3}{q^4} [3 - 4 \cos(q) + \cos(2q)]$$

Now consider:

$$\alpha_n = \frac{\langle \phi_n | \chi_n \rangle}{\sqrt{\langle \phi_n | \phi_n \rangle^2 \langle \chi_n | \chi_n \rangle^2}}$$

If χ_n and ϕ_n were proportional then $\alpha_n = 1$, and blip waves would be identical to plane waves. Let's see what value of α_n has for a few values on n .

n	k_n	α_n
0	0	1.00
1	$\pi/6a$	0.9986
3	$\pi/2a$	0.9858
6	π/a	0.6532

Take $N=12$, for example

$$k_n = 2\pi n / Na$$

$$k_{\max} = \pi / a$$

Therefore if we replace ϕ_n with χ_n we make a small error at low k but maybe a significant error for large values of k . Let's do it anyway, then we will come back to the size of the error and how to reduce it.

Representing single particle orbitals in 3 d

$$\psi_n(\mathbf{r}) = \sum_{\mathbf{G}} c_{n\mathbf{G}} e^{i\mathbf{G}\cdot\mathbf{r}}$$

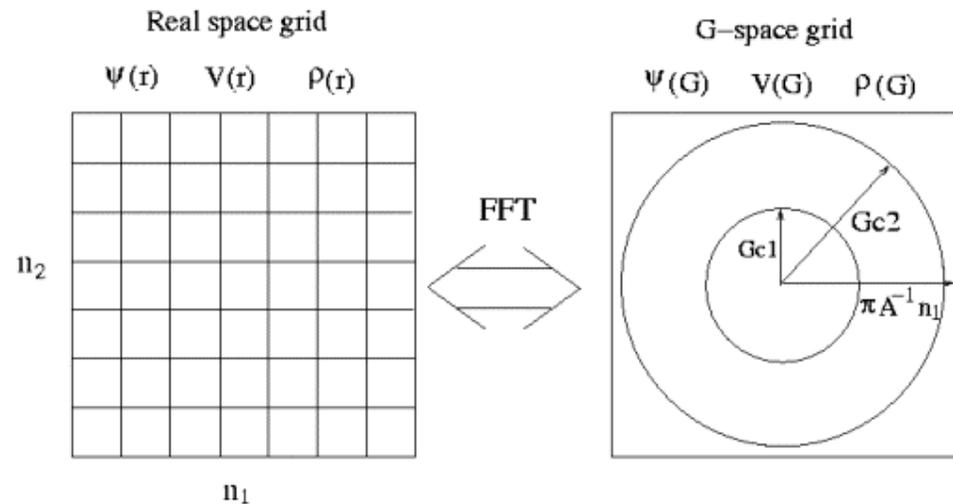
$$\psi_n(\mathbf{r}) = \sum_j a_{nj} \Theta_j(\mathbf{r})$$

$$\chi_{\mathbf{G}}(\mathbf{r}) = \sum_j \Theta_j(\mathbf{r}) e^{i\mathbf{G}\cdot\mathbf{R}_j}$$

$$e^{i\mathbf{G}\cdot\mathbf{r}} ; \gamma_{\mathbf{G}} \chi_{\mathbf{G}}(\mathbf{r})$$

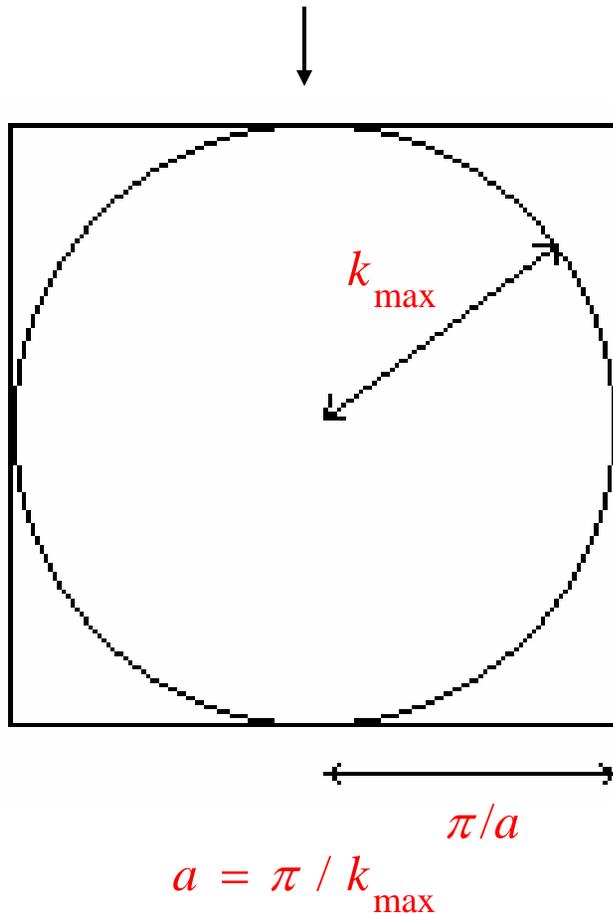
$$\psi_n(\mathbf{r}) ; \sum_{\mathbf{G}} \gamma_{\mathbf{G}} c_{n\mathbf{G}} \chi_{\mathbf{G}}(\mathbf{r}) = \sum_{\mathbf{G}} c_{n\mathbf{G}} \gamma_{\mathbf{G}} \sum_j \Theta_j(\mathbf{r}) e^{i\mathbf{G}\cdot\mathbf{R}_j} = \sum_j a_{nj} \Theta_j(\mathbf{r})$$

$$a_{nj} = \sum_{\mathbf{G}} \gamma_{\mathbf{G}} c_{n\mathbf{G}} e^{i\mathbf{G}\cdot\mathbf{R}_j} \quad (\text{note that in DA \& MJG, PRB } \mathbf{70}, 16101 (2004) \gamma_{\mathbf{G}} \text{ should be } 1/\gamma_{\mathbf{G}})$$



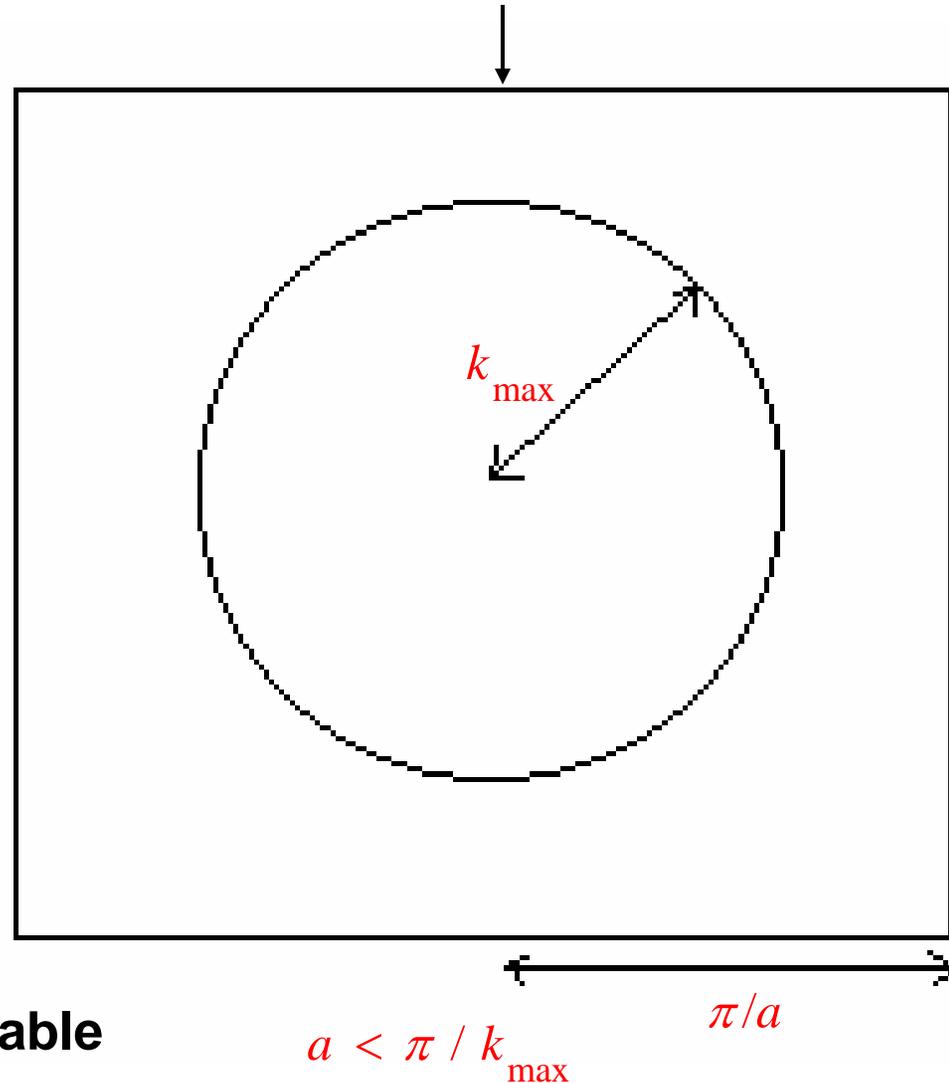
Improving the quality of the B-spline representation

May not be good enough



Blips are systematically improvable

Reduce a and achieve better quality



Example: Silicon in the β -Sn structure, 16 atoms, 1st orbital

$$\alpha_n = \frac{\langle \psi_n^{B-splines} | \psi_n^{PW} \rangle}{\sqrt{\langle \psi_n^{B-splines} | \psi_n^{B-splines} \rangle \langle \psi_n^{PW} | \psi_n^{PW} \rangle}}$$

	ψ	$\Delta\psi$	$\nabla_x \psi$	$\nabla_y \psi$	$\nabla_z \psi$
$\alpha_1(a = \pi / k_{\max})$.999977	.892597	.991164	.991164	.996006
$\alpha_1(a = \pi / 2k_{\max})$.99999998	.995315	.999988	.999988	.999993

Energies in QMC (eV/atom)

(Silicon in β -Sn structure, 16 atoms, 15 Ry PW cutoff)

	PW	Blips ($a=\pi/k_{max}$)	Blips ($a=\pi/2k_{max}$)
E_k (DFT)	43.863		
E_{loc} (DFT)	15.057		
E_{nl} (DFT)	1.543		
E_k (VMC)	43.864(3)	43.924(3)	43.862(3)
E_{loc} (VMC)	15.057(3)	15.063(3)	15.058(3)
E_{nl} (VMC)	1.533(3)	1.525(3)	1.535(3)
E_{tot} (VMC)	-101.335(3)	-101.277(3)	-101.341(3)
σ (VMC)	4.50	4.74	4.55
T(s/step)	1.83	0.32	0.34
E_{tot} (DMC)	-105.714(4)	-105.713(5)	-105.716(5)
σ (VMC)	2.29	2.95	2.38
T(s/step)	2.28	0.21	0.25

Energies in QMC (eV/atom)

(MgO in NaCl structure, 8 atoms, 200 Ry PW cutoff)

	PW	Blips ($a=\pi/k_{max}$)	Blips ($a=\pi/2k_{max}$)
E_k (VMC)	178.349(49)	178.360(22)	178.369(22)
E_{loc} (VMC)	-225.191(50)	-225.128(24)	-225.177(23)
E_{nl} (VMC)	-17.955(25)	-17.974(11)	-17.976(11)
E_{tot} (VMC)	-227.677(8)	-227.648(4)	-227.669(4)
σ (VMC)	14	15	14.5
T(s/step)	7.8	5.6×10^{-2}	7.1×10^{-2}

Blips and plane waves give identical results, blips are 2 orders of magnitude faster on 8 atoms, 3 and 4 order of magnitudes faster on 80 and 800 atoms respectively (on 800 atoms: Blips 1 day, PW 38 years)

Achieving linear scaling

Maximally localised Wannier functions

(Marzari Vanderbilt)

[Williamson et al., PRL, **87**, 246406 (2001)]

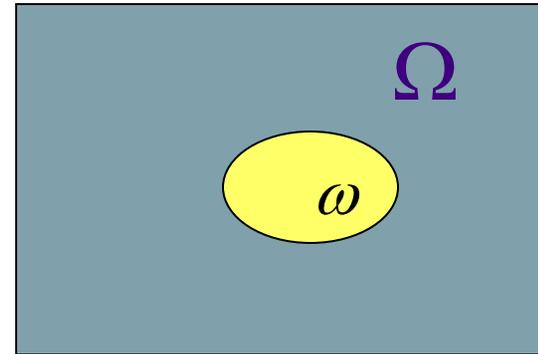
Unitary transformation:

$$\phi_i(\mathbf{r}) = \sum_{m=1}^N c_{mi} \psi_m(\mathbf{r})$$

And:

$$D\{\phi_i(\mathbf{r})\} = D\{\psi_m(\mathbf{r})\}$$

$$\begin{aligned} \Psi_T(\mathbf{R}) &= \exp[J(\mathbf{R})] \sum_n c_n D_n^\uparrow \{\psi_i(\mathbf{r}_j)\} D_n^\downarrow \{\psi_i(\mathbf{r}_j)\} \\ &= \exp[J(\mathbf{R})] \sum_n c_n D_n^\uparrow \{\phi_i(\mathbf{r}_j)\} D_n^\downarrow \{\phi_i(\mathbf{r}_j)\} \end{aligned}$$



Linear scaling obtained by truncating orbitals to zero outside their localisation region.

$$E_L(\mathbf{R}) = \Psi_T(\mathbf{R})^{-1} \hat{H} \Psi_T(\mathbf{R})$$

The VMC and DMC total energies are exactly invariant with respect to **arbitrary (non singular)** linear combinations of single-electron orbitals! **Transformation does not have to be unitary.**

Arbitrary linear combination:

$$\phi_i(\mathbf{r}) = \sum_{m=1}^N c_{mi} \psi_m(\mathbf{r})$$

And:

$$D\{\phi_m(\mathbf{r})\} = D\{c_{ij}\} D\{\psi_i(\mathbf{r})\} \quad \text{Provided} \quad D\{c_{ij}\} \neq 0$$

We have set of orbitals $\psi_n(\mathbf{r})$ extending over region Ω

Want to make orbital $\phi(\mathbf{r}) = \sum_n c_n \psi_n(\mathbf{r})$
maximally localised in sub-region $\omega \in \Omega$.

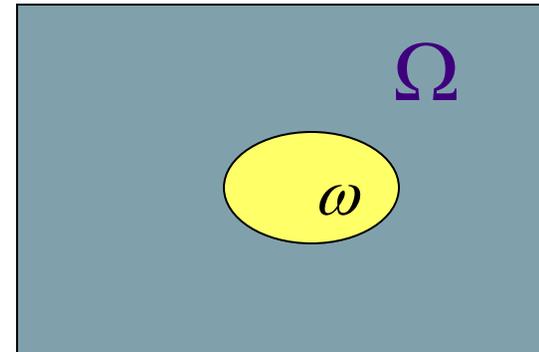
Maximise “localisation weight” P :

$$P = \int_{\omega} d\mathbf{r} |\phi(\mathbf{r})|^2 / \int_{\Omega} d\mathbf{r} |\phi(\mathbf{r})|^2$$

Express P as: $P = \sum_{m,n} c_m^* A_{mn}^{\omega} c_n / \sum_{m,n} c_m^* A_{mn}^{\Omega} c_n$

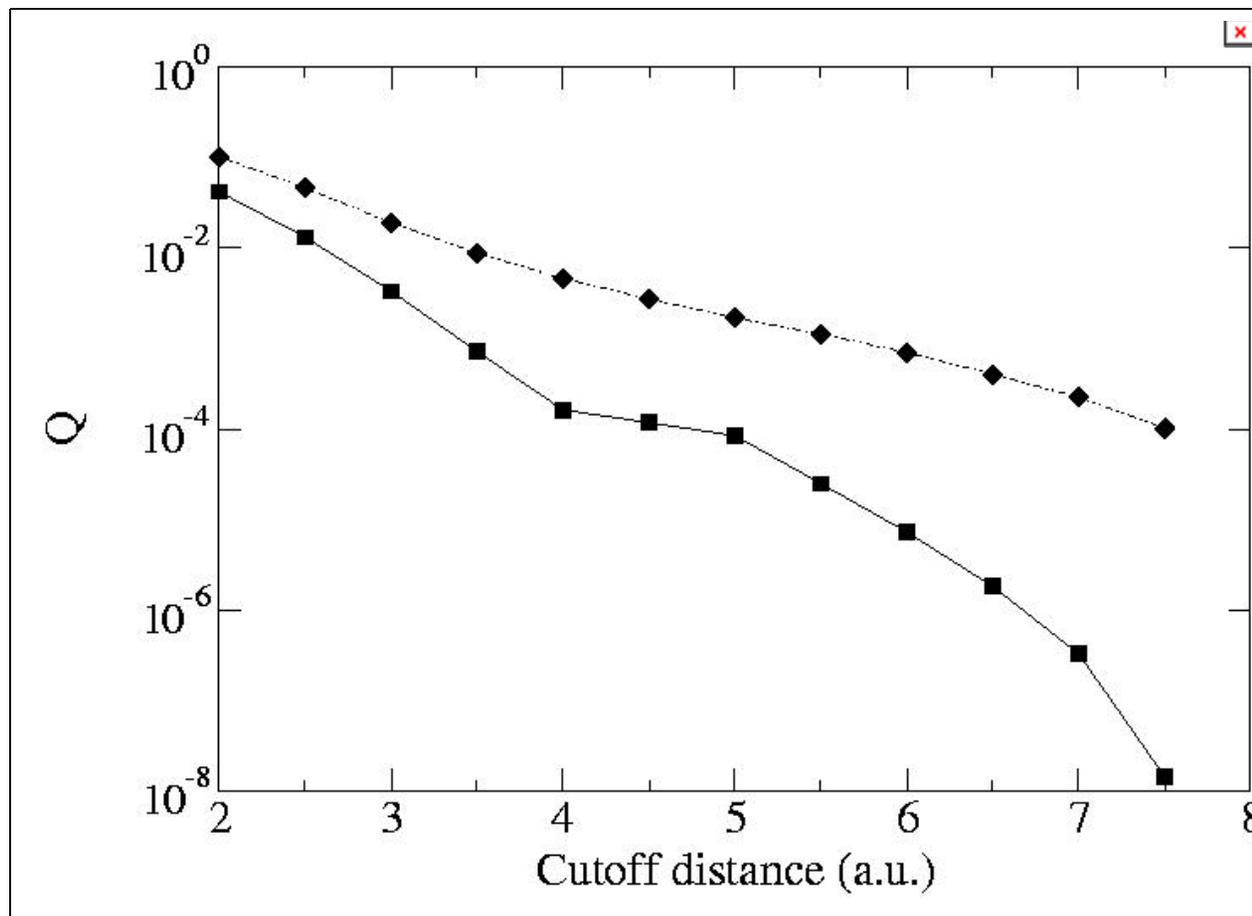
where: $A_{mn}^{\omega} = \int_{\omega} d\mathbf{r} \psi_m^* \psi_n$, $A_{mn}^{\Omega} = \int_{\Omega} d\mathbf{r} \psi_m^* \psi_n$

Weight P is maximal when c_n eigenvector of $\sum_n A_{mn}^{\omega} c_n = \lambda_{\alpha} \sum_n A_{mn}^{\Omega} c_n$
associated with maximum eigenvalue λ_{\max} , and $P = \lambda_{\max}$



Linear scaling obtained by truncating localised orbitals to zero outside their localisation region ω . Then for given \mathbf{r} number of non-zero orbitals $\phi_m(\mathbf{r})$ is $O(1)$.

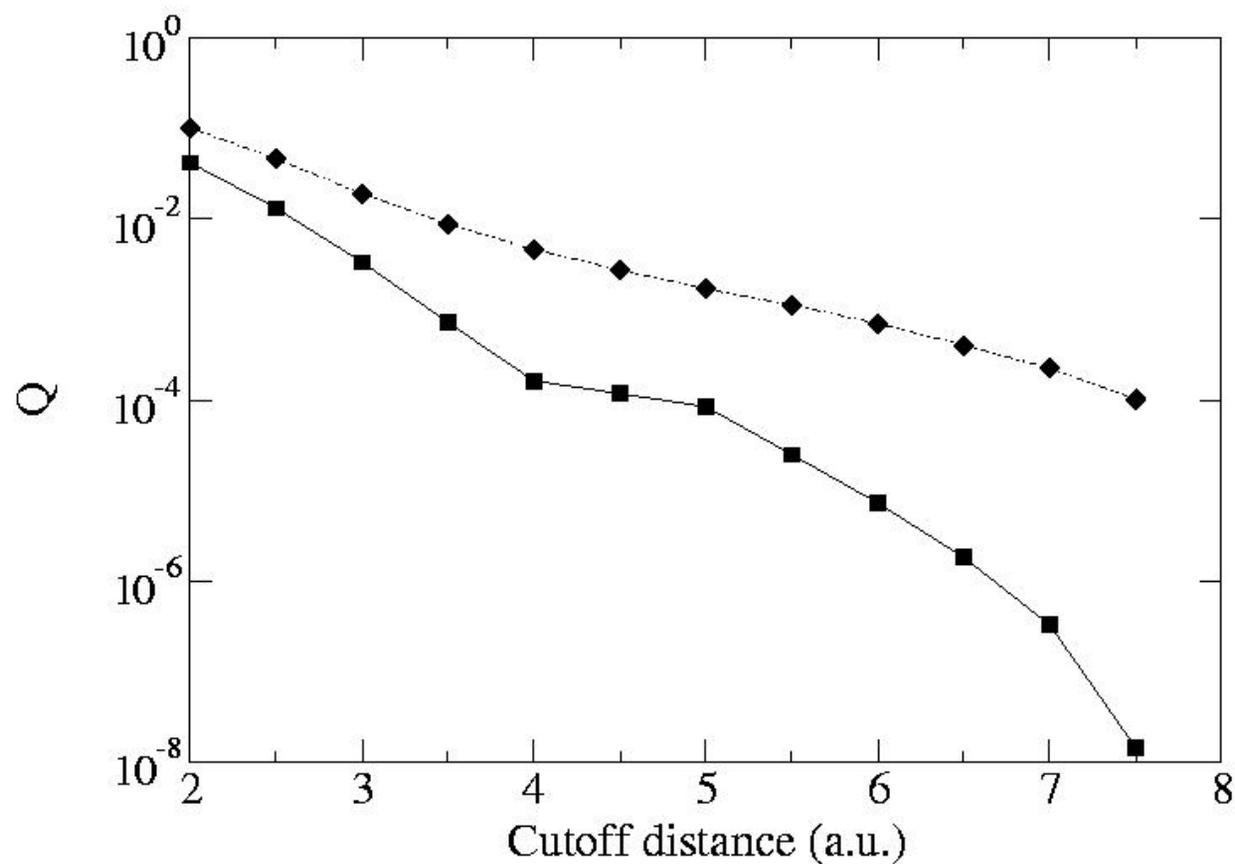
Localisation weight in MgO



Convergence of localisation weight $Q=1-P$ in MgO.
Squares: present method (D. Alfè and M. J. Gillan, *J. Phys. Cond. Matter* **16**, L305-L311 (2004))

Diamonds: MLWF (Williamson et al. *PRL*, **87**, 246406 (2001))

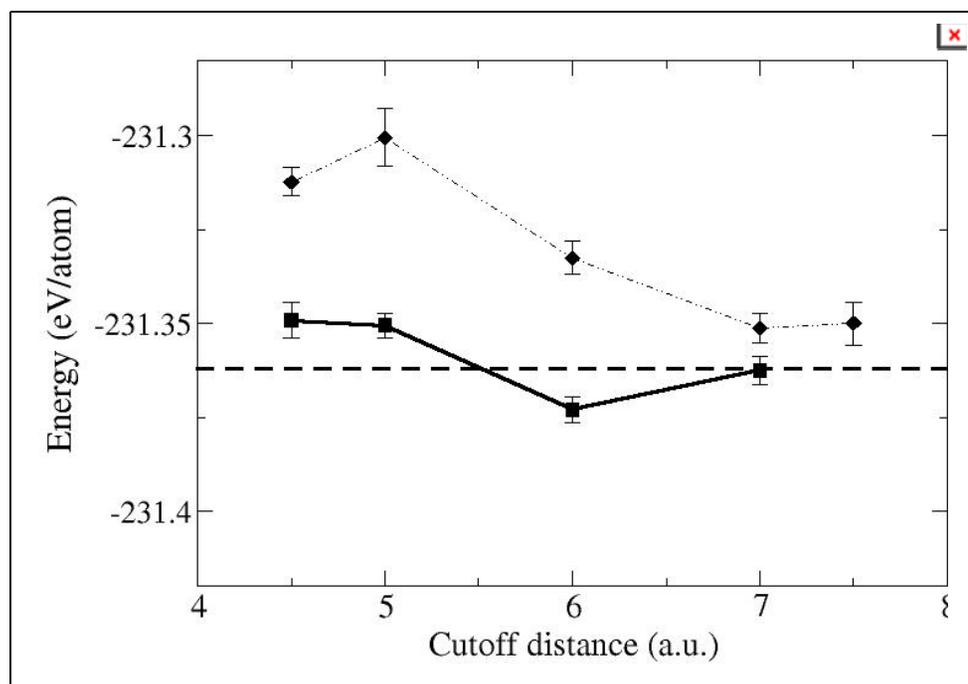
Localisation weight in MgO



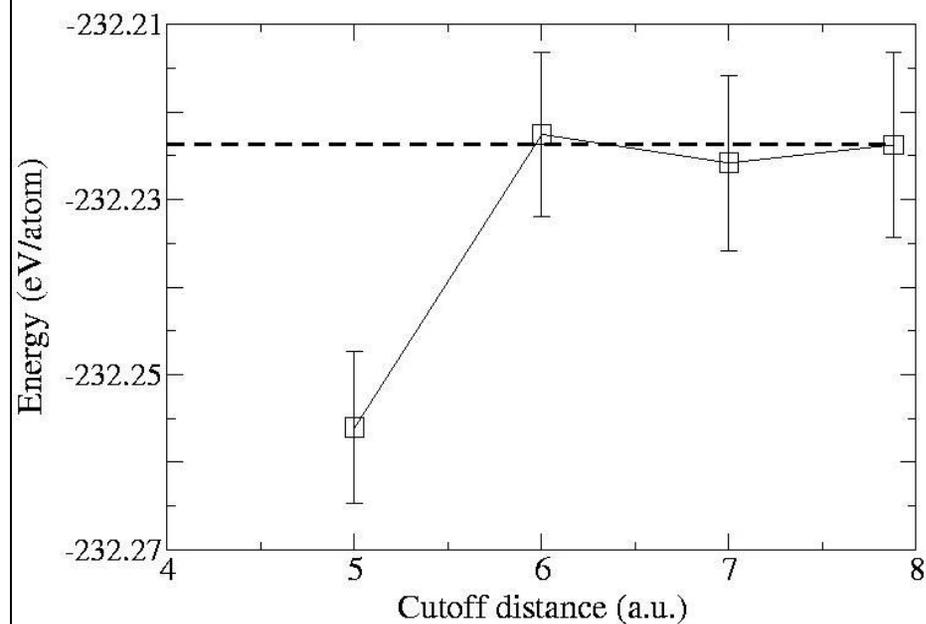
Convergence of localisation weight $Q=1-P$ in MgO.
 Squares: present method (D. Alfè and M. J. Gillan, *J. Phys. Cond. Matter* **16**, L305-L311 (2004))

Diamonds: MLWF (Williamson et al. *PRL*, **87**, 246406 (2001))

VMC and DMC energies in MgO

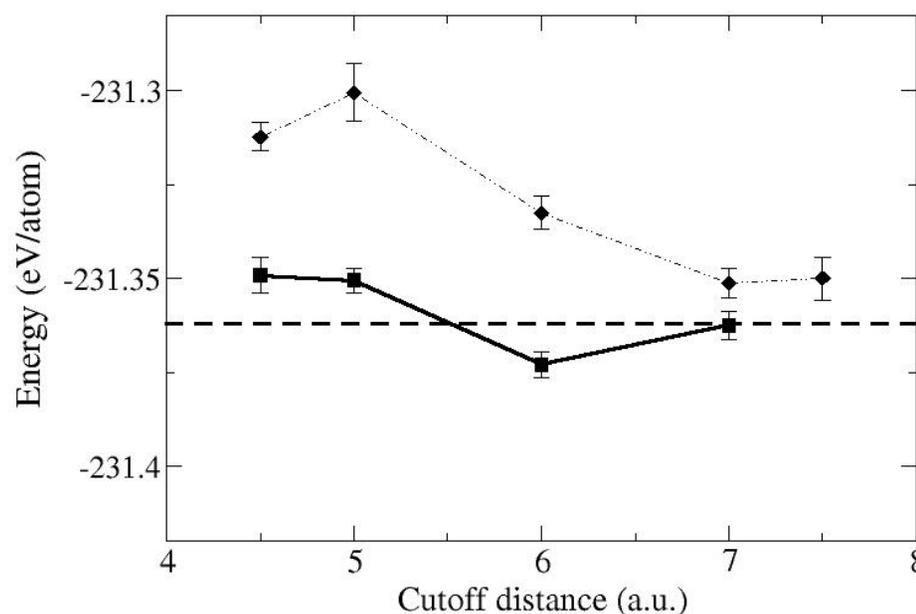


Convergence of VMC energy.
Squares: Non-orthogonal orbitals.
Diamonds: MLWF

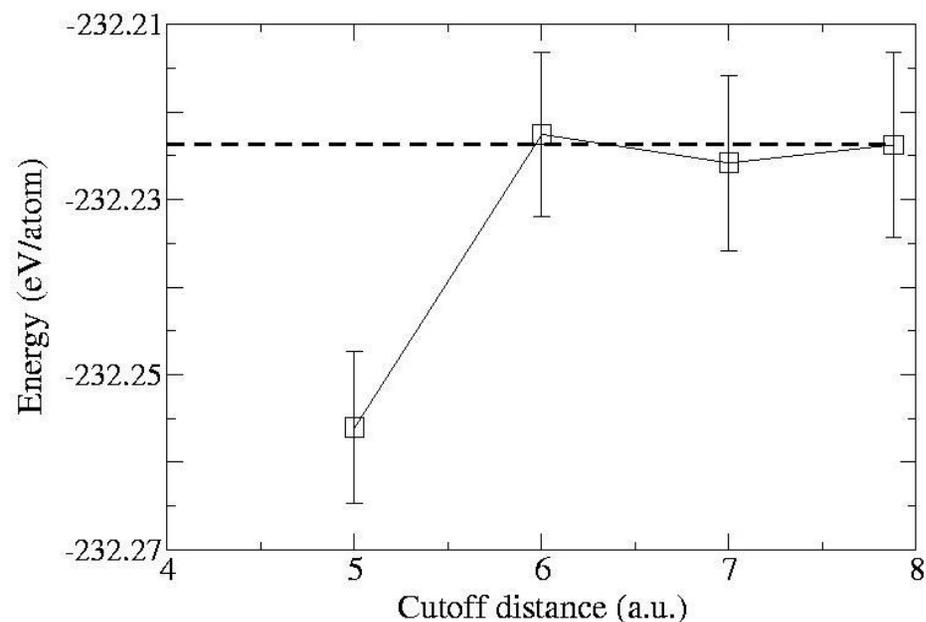


Convergence of DMC energy, non-orthogonal orbitals

VMC and DMC energies in MgO



Convergence of VMC energy.
Squares: Non-orthogonal orbitals.
Diamonds: MLWF



Convergence of DMC energy, non-orthogonal orbitals

Spherical cutoff at 6 a.u on a 7.8 a.u. cubic box (64 atoms):

- WF evaluation 4 times faster
- Memory occupancy 4 times smaller