Préconditionement de la boucle auto-cohérente pour de grands systèmes inhomogènes

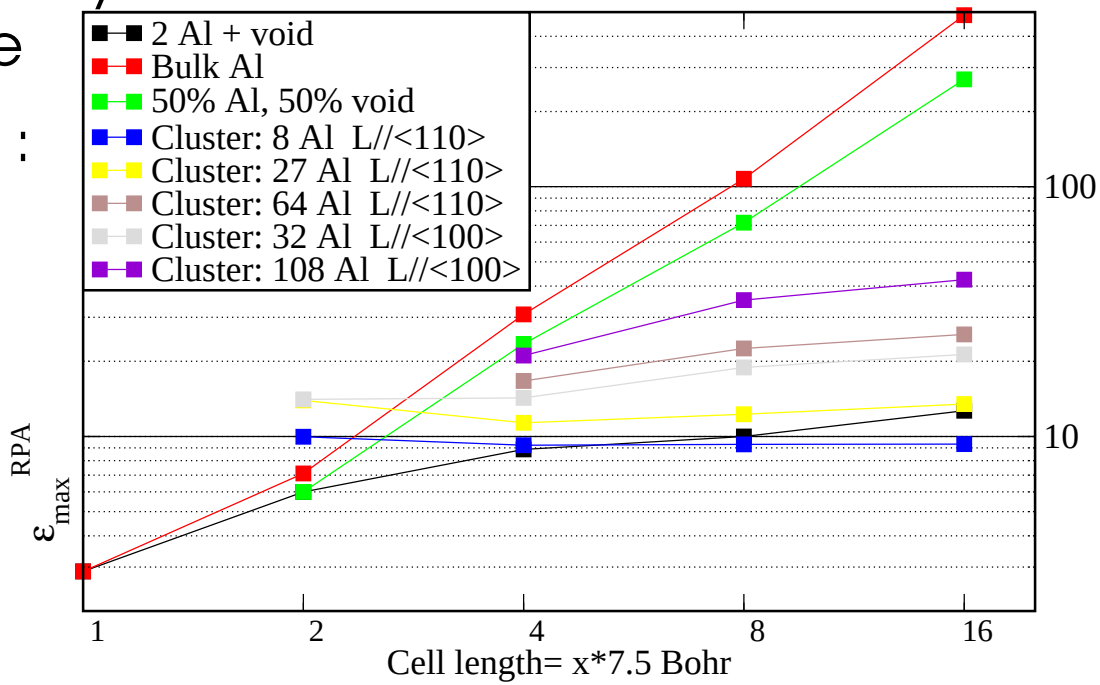
P.-M. Anglade X. Gonze

Plan

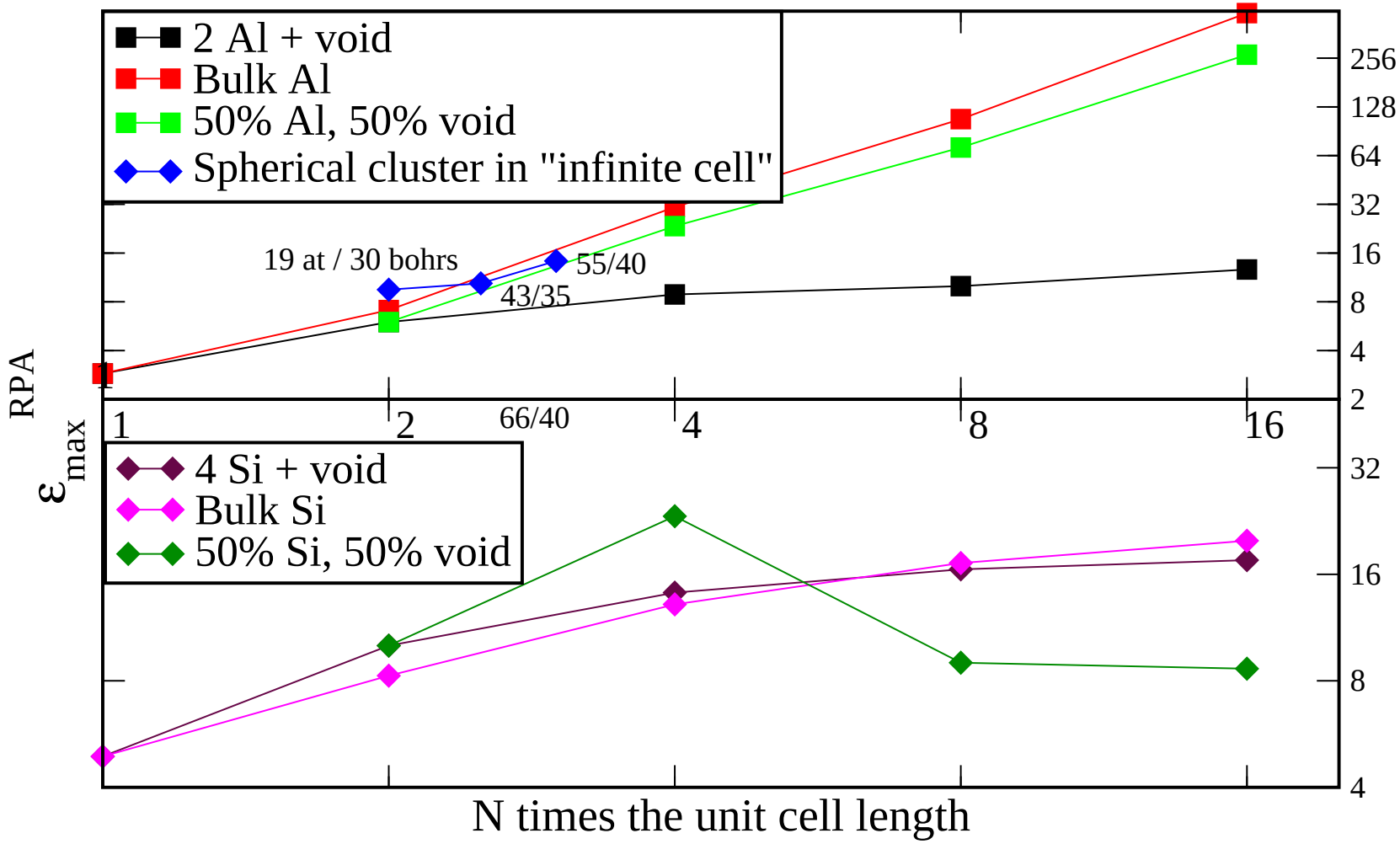
- Reminder: Why would anyone uses preconditioners?
  - Convergence problems
  - History coping algorithms and preconditioners
- From Ho's Newton-Raphson Method to *extrapolar*
  - Newton-Raphson on SCF cycles
  - The extrapolar approximation
- From Kerkers's preconditioner to *RSKerker*
  - Kerkers's preconditioner for metals
  - Abinit's extended Kerker's scheme
  - RSKerker
- Advanced mixing and preconditioners coupling
- Conclusion

Reminder: Convergence problems

- Hartree potential react to error in the charge ($\propto r^2 \propto \frac{1}{G^2}$) in metals.
- The matrix of coefficient discribing this reaction (the hessian) is the dielectric permittivity matrix.
- Reaction can amplify the error.
- Basically the largest eigenvalue of the dielectric permittivity matrix will show the difficulty for getting convergence
- Al as a metal prototype :

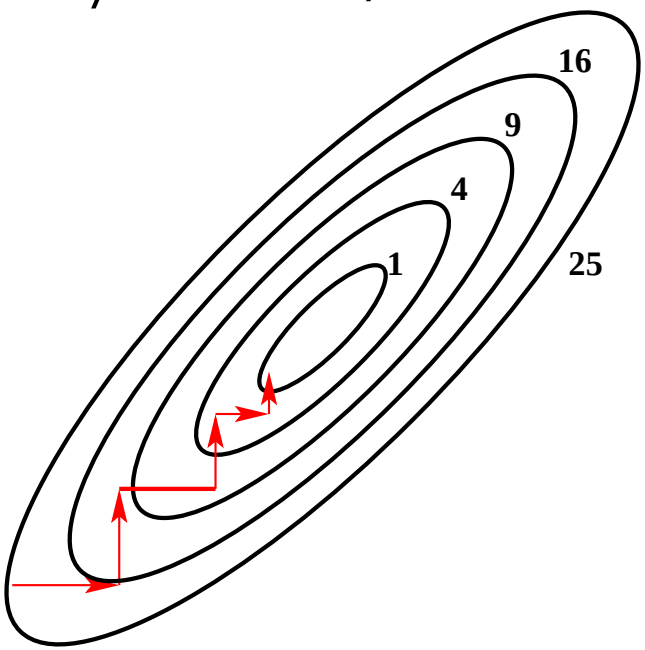


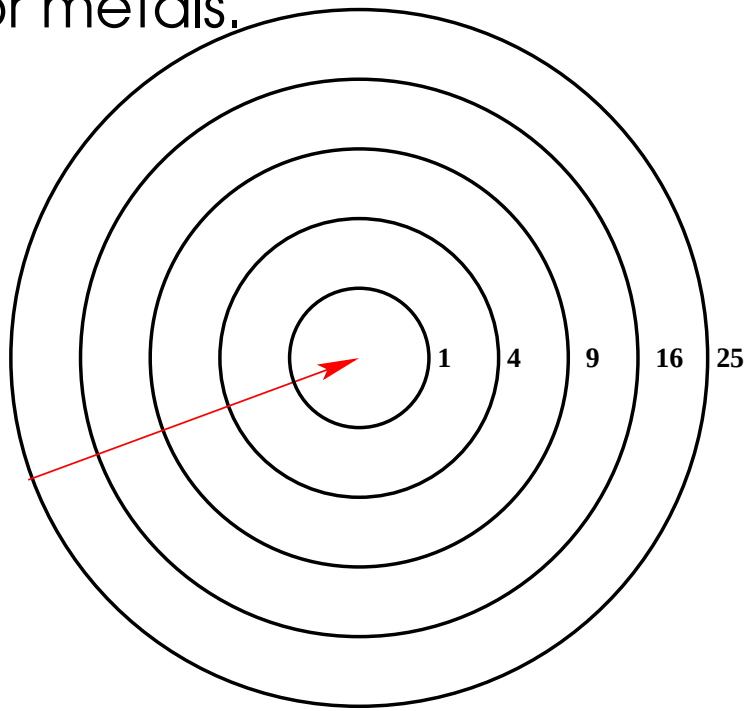
Reminder: Convergence problems



Reminder: Why would anyone uses preconditioners? 1/2

- Hessians of large systems have large maximum eigenvalues.
- Simple mixing is not enough because the problem complexity is a matter of condition number (CN)$= \frac{\lambda_{max}}{\lambda_{min}}$.
  - The smallest eigenvalues (λ_{min}) depends on the XC. They are more or less local properties. Usually $\lambda_{min} \simeq 1$
  - The largest eigenvalues (λ_{max}) basically grows with system size, without limits for metals.





Reminder: Why would anyone uses preconditioners? 2/2

- Advanced mixing schemes (Broyden, CG, Pulay) use a single mixing coefficient per step which is not enough to cope with difficult problem.
  - Best mixing schemes promise convergence in N step for a N eigenvalue quadratic problem;
  - First order preconditioners offer convergence in 1 step for the same problem.
- To evaluate the efficiency of a preconditioner we must stick with simple mixing:
  - DFT problems are not quadratic;
  - Potentialy difficult to combine preconditioners and advanced mixing schemes (more on that later).
- To sum up:
  - Preconditioners are mandatory for large scale problems;
  - They must be examined alone (simple mixing).

Exact Newton-Raphson's preconditioning 1/2

- K.-M. Ho *et. al.*, PRB 25, no. 6, p. 4260 (1982)
- $$\chi(G,G') = -2\sum_{n,n',k}(w_{n,k}-w_{n,k'})\frac{\langle nk|e^{-iG.r}|n'k\rangle\langle nk|e^{iG.r}|n'k\rangle}{E_{n',k}-E_{n,k}}$$

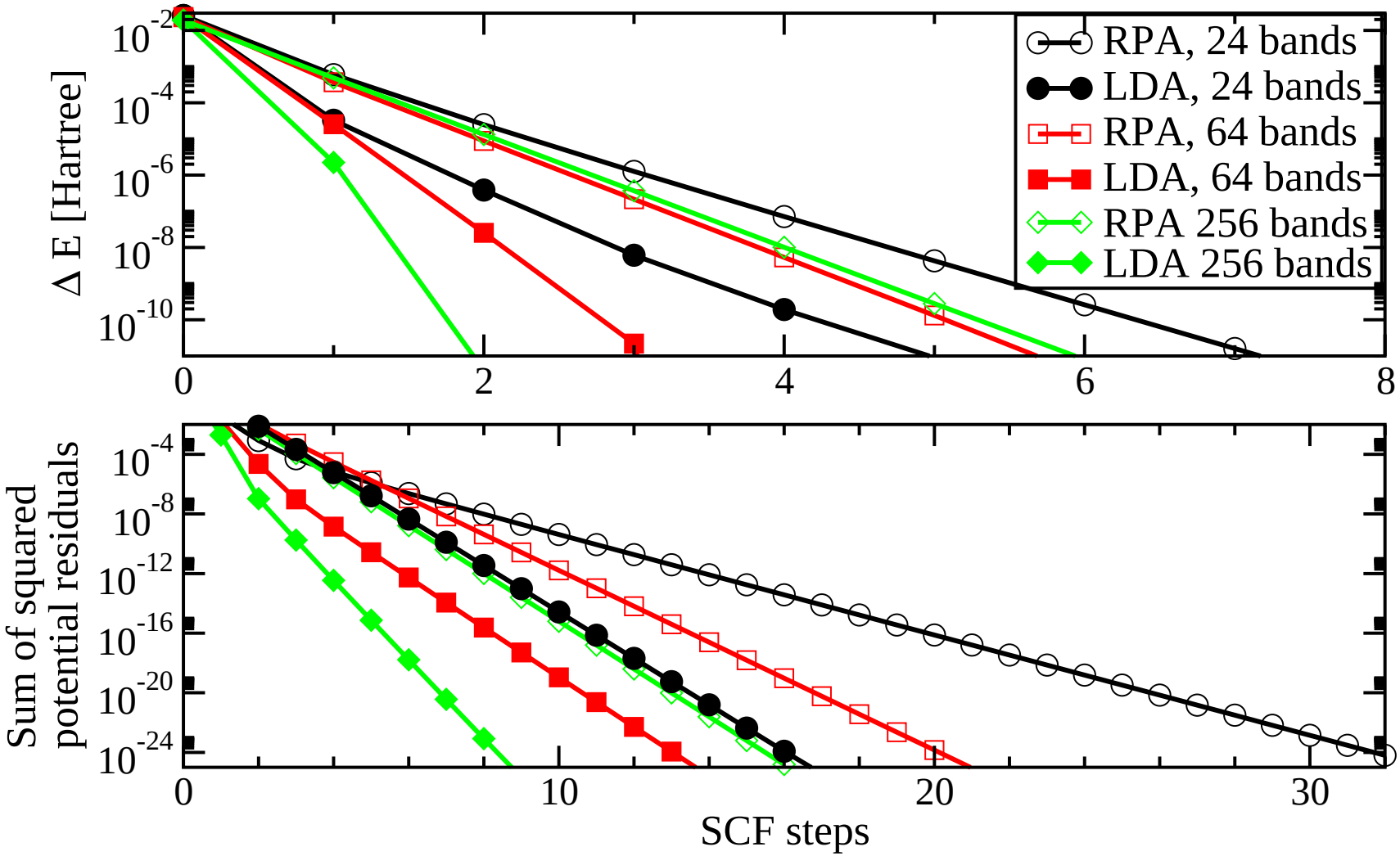
$$\varepsilon_{LDA}(G,G') = \delta_{G,G'} - (K_H + K_{xc})\chi(G,G')$$

$$\varepsilon_{RPA}(G,G') = \delta_{G,G'} - K_H\chi(G,G')$$

$$\Delta V_{opt} = \varepsilon^{-1}(V_{out} - V_{in})$$

Exact Newton-Raphson's preconditioning 2/2

- Linear cell of bulk Si (8 atoms) with no longitudinal symmetry; simple mixing ($\alpha = 1$).



Extrapolar 1/3

- (Best known to Abinit users as iprcel $\neq 0$).
- $$\chi(r,r') \propto \sum_k \sum_{n=1}^{n_{valence}} \sum_{n'=n_{valence}}^{\infty} \frac{\psi^*_{kn}(r)\psi_{kn'}(r)\psi^*_{kn'}(r)\psi_{kn}(r')}{E_{kn'} - E_{kn}}$$

$$\chi(r,r') \propto A + \sum_k \sum_{n=1}^{n_{valence}} \sum_{n'=n_{band}+1}^{\infty} \alpha \frac{\psi^*_{kn}(r)\psi_{kn'}(r)\psi^*_{kn'}(r)\psi_{kn}(r')}{\bar{E} - E_{kn}}$$

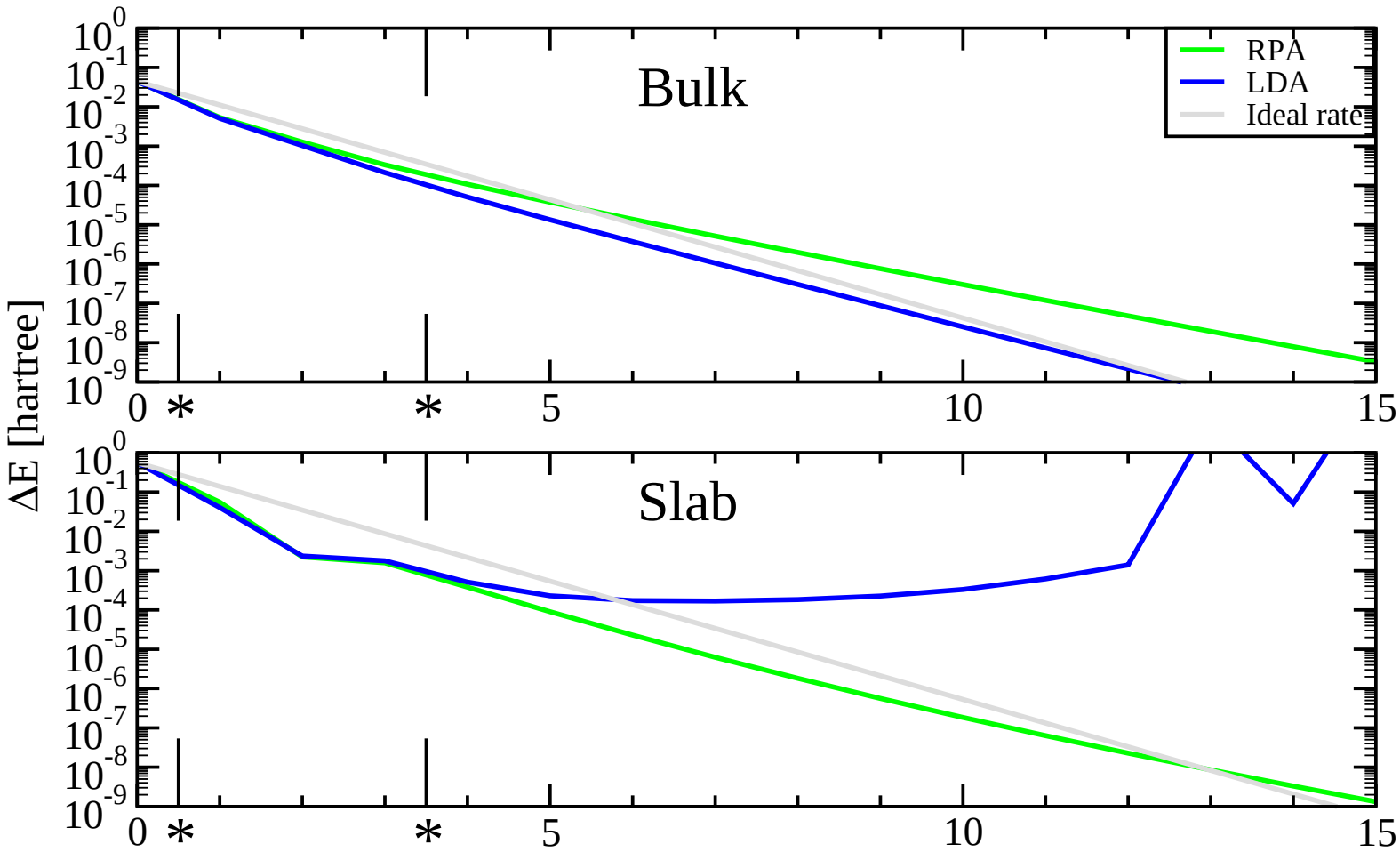
closure relation $\Downarrow$

$$\chi(r,r') \propto A + \sum_k \sum_{n=1}^{n_{valence}} \alpha \frac{\psi^*_{kn}(r)\psi_{kn}(r)}{\bar{E} - E_{kn}} - \sum_k \sum_{n=1}^{n_{valence}} \sum_{n'=1}^{n_{band}} \alpha \frac{\psi^*_{kn}(r)\psi_{kn'}(r)\psi^*_{kn'}(r)\psi_{kn}(r')}{\bar{E} - E_{kn}}$$

- Two (extra) parameters $\alpha = dielam$ et $\bar{E} = diegap$.

Extrapolar 2/3

- Divergence in vacuum due to $K_{xc} \propto \rho^{-\frac{2}{3}}$
- Linear cell of bulk (resp. slab) Si, 16 (resp. 8) atoms, with no longitudinal symmetry.



Extrapolar 3/3

- Advantages of extrapolar:
  - A (very) reduced plane wave basis can be used to compute the polarizability matrix (0.5 Ha may be enough, compared to typical 10-80 Ha cut-off !);
  - RPA formulation of extrapolar works for any system;
  - Parameters are almost perfectly transferable $\Rightarrow$ no tuning;
  - One to three empty band is OK $\Rightarrow$ Extrapolar is much faster to apply than Ho's method;
  - Extrapolar is perfectly suitable for system of about ~ 100 atoms.
- Drawbacks:
  - Must be used within RPA for inhomogeneous systems;
  - Do not converge as efficiently as Ho's method;
  - Still an $O(N^4)$ method (scaling figures later).

Kerker's preconditioner

- G.P. Kerker, PRB, vol. 23, nb. 6, p. 3082 (1981)
- Rewrite Poisson's equations:

$$(\nabla^2 - \lambda^2)V_{coul} = -4\pi\left(\rho + \frac{\lambda^2}{4\pi}V_{coul}\right)$$

- Apply this to an SCF cycle (in reciprocal space) :

$$V_{coul}^{(n+1)} = \frac{4\pi\rho^{(n)} + \lambda^2 V_{coul}^{(n)}}{4\pi^2 G^2 + \lambda^2}$$

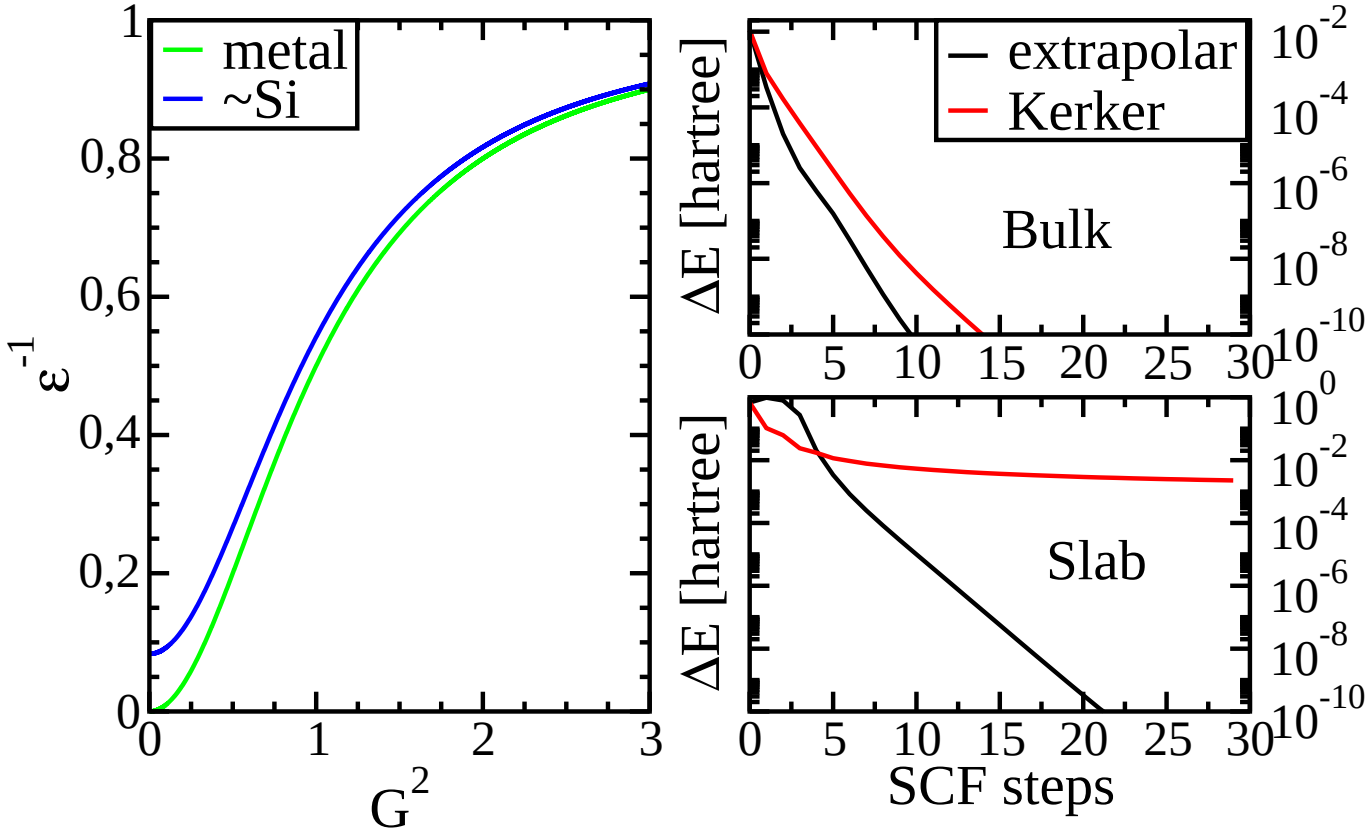
- We get (changing $\lambda \rightarrow \Lambda = \frac{2\pi}{\lambda} =$ dielng):

$$V_{coul,res}^{prc} = \frac{\Lambda^2 G^2}{\Lambda^2 G^2 + 1} V_{coul,res}$$

Abinit's extended Kerker's scheme

- Can be applied on the full trial potential.
- Can be tuned to account for finite charge sloshing (diemac= ε_r):

$$V_{res}^{prc} = \frac{\frac{1}{\varepsilon_r} + \Lambda^2 G^2}{\Lambda^2 G^2 + 1} V_{res}$$



RSKerker

- Kerker is equivalent in real space to:

$$\left(1-\frac{\Lambda^2}{4\pi^2}\nabla^2\right)V_{prc}(\boldsymbol{r})=\left(\frac{1}{\varepsilon_r}-\frac{\Lambda^2}{4\pi^2}\nabla^2\right)V_{res}(\boldsymbol{r})$$

- V_{prc} is minimizing:

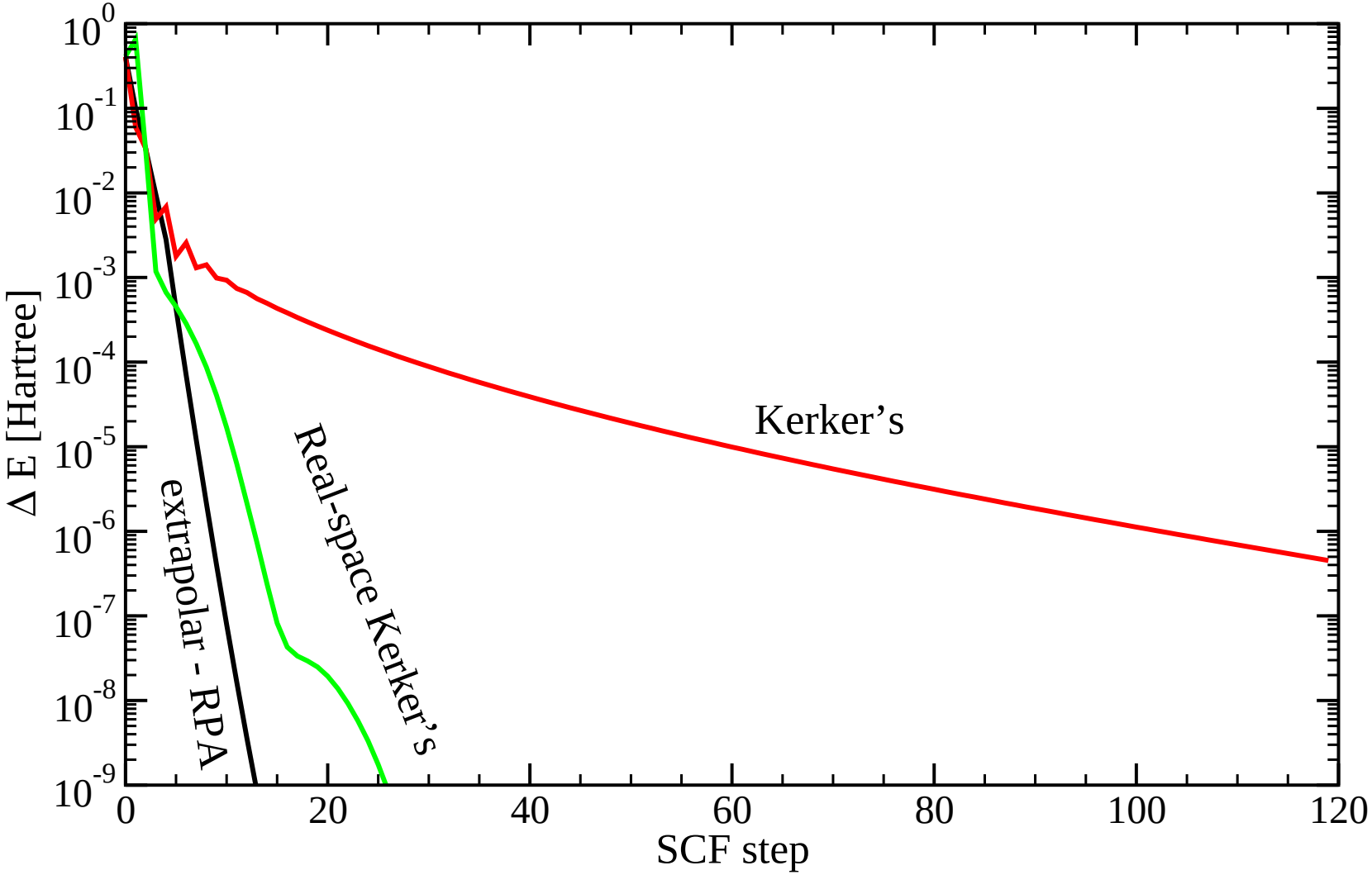
$$\text{pf}\,(V)=\frac{1}{2}V.\left(1-\frac{\Lambda^2}{4\pi^2}\nabla^2\right)V-V.\left(\frac{1}{\varepsilon_r}-\frac{\Lambda^2}{4\pi^2}\nabla^2\right)V_{res}$$

- We can change $\varepsilon_r \longrightarrow \varepsilon_r(r)$
- First intuitive choice for $\varepsilon_r(r)$ is

$$\varepsilon_r(r)\propto\rho+Cste$$

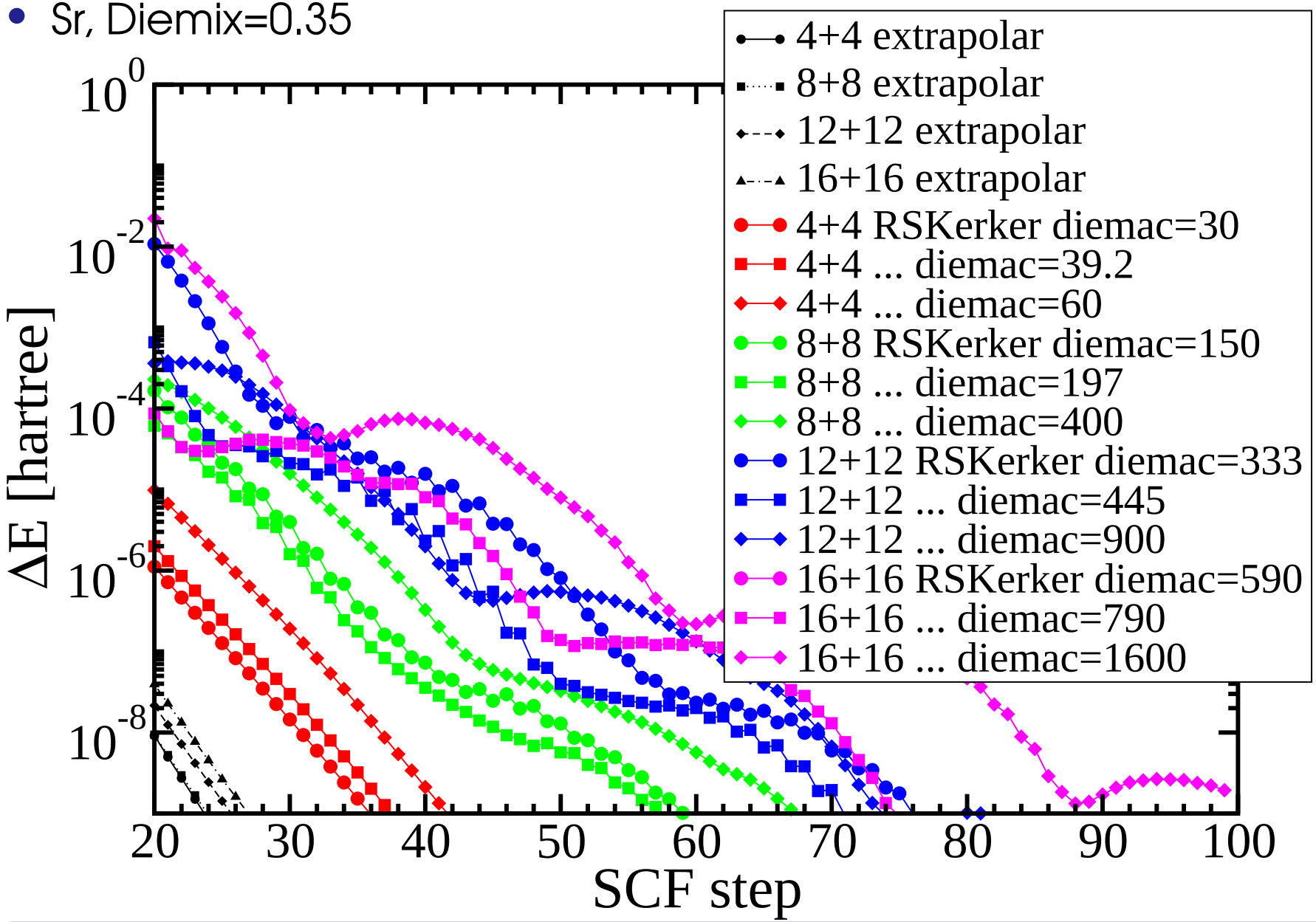
RSKerker

- Cell of 12 Al, 12 Vacuum.



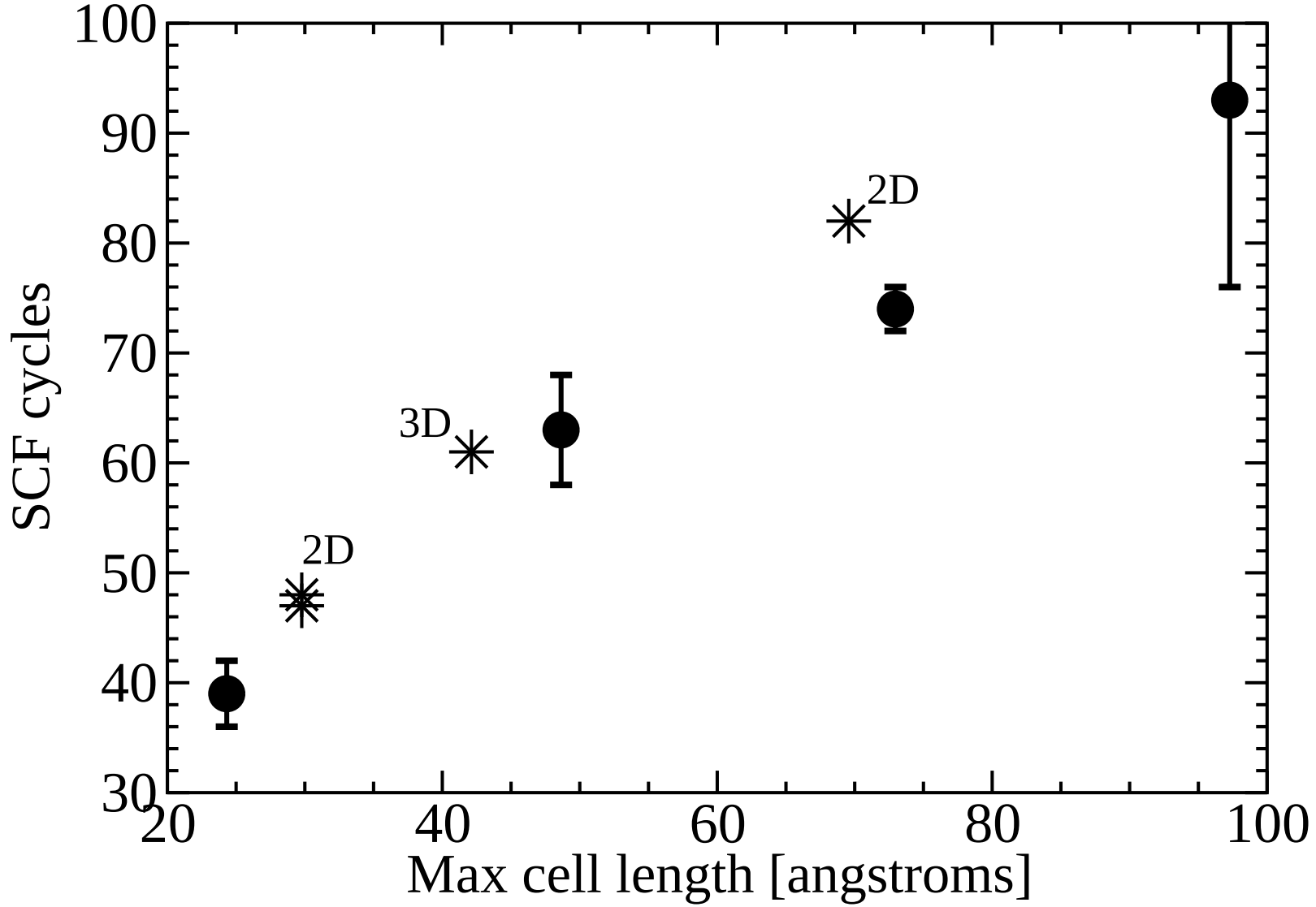
RSKerker

- Sr, Diemix=0.35



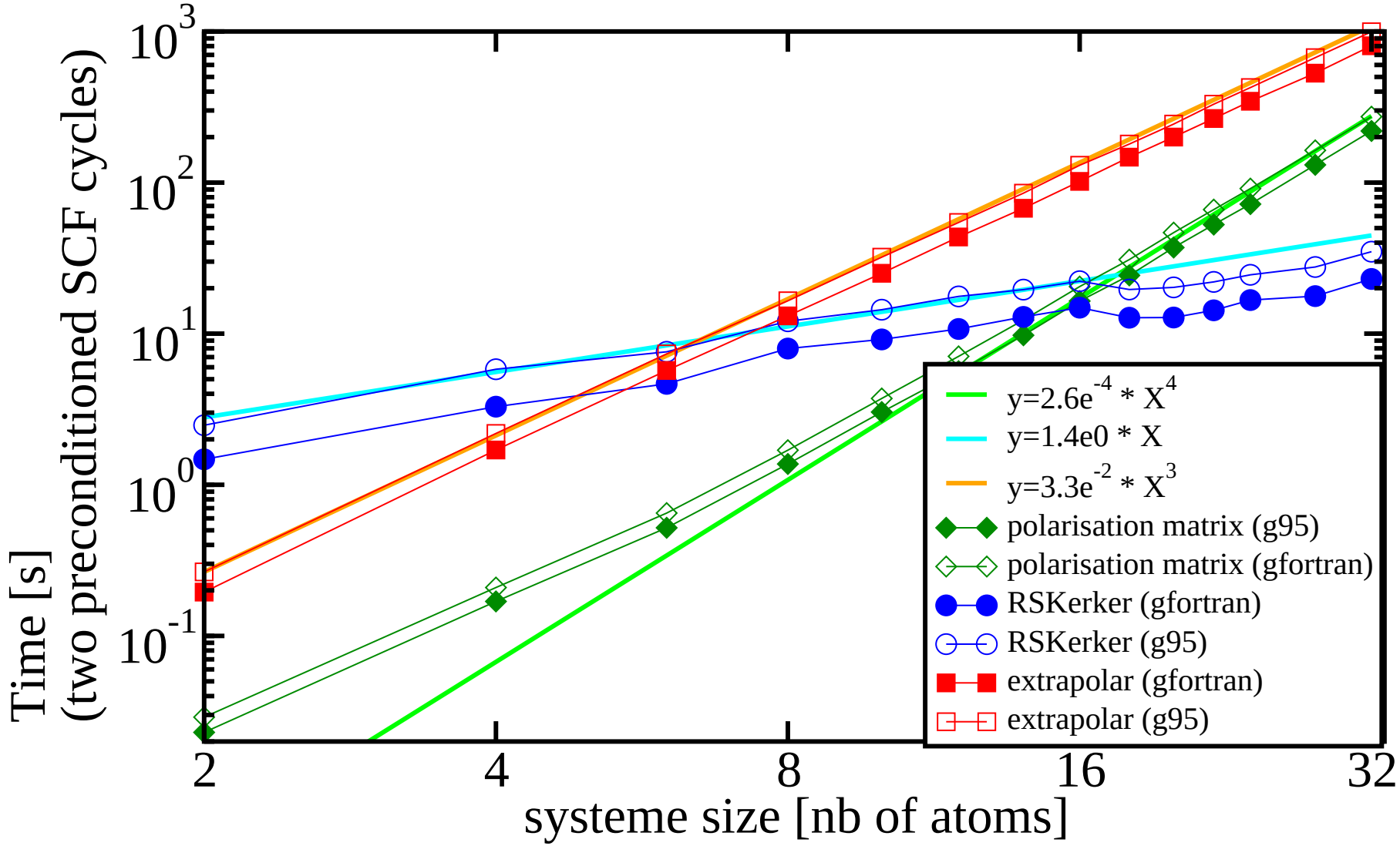
RSKerker

- Sr, Diemix=0.35

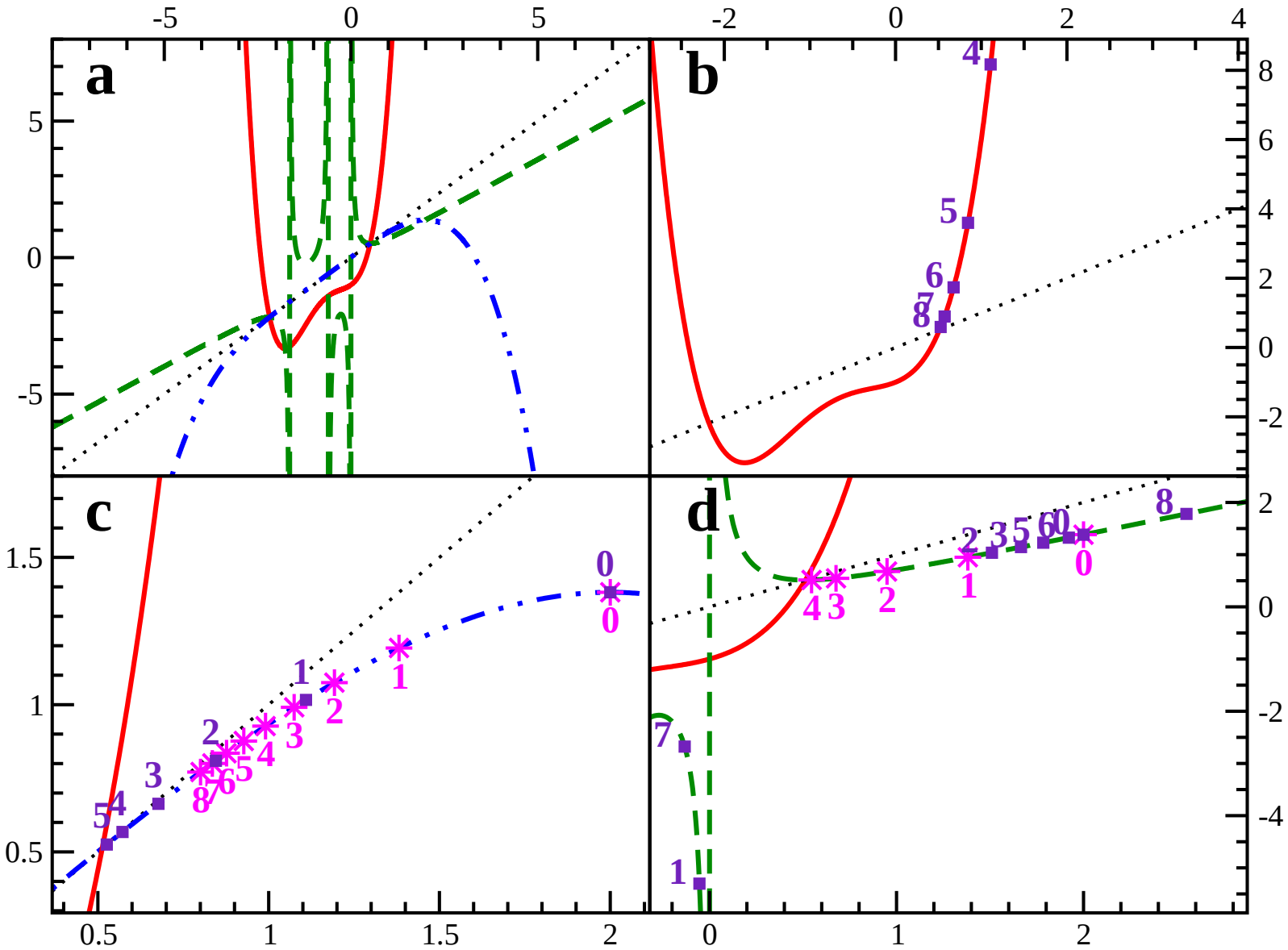


Scaling

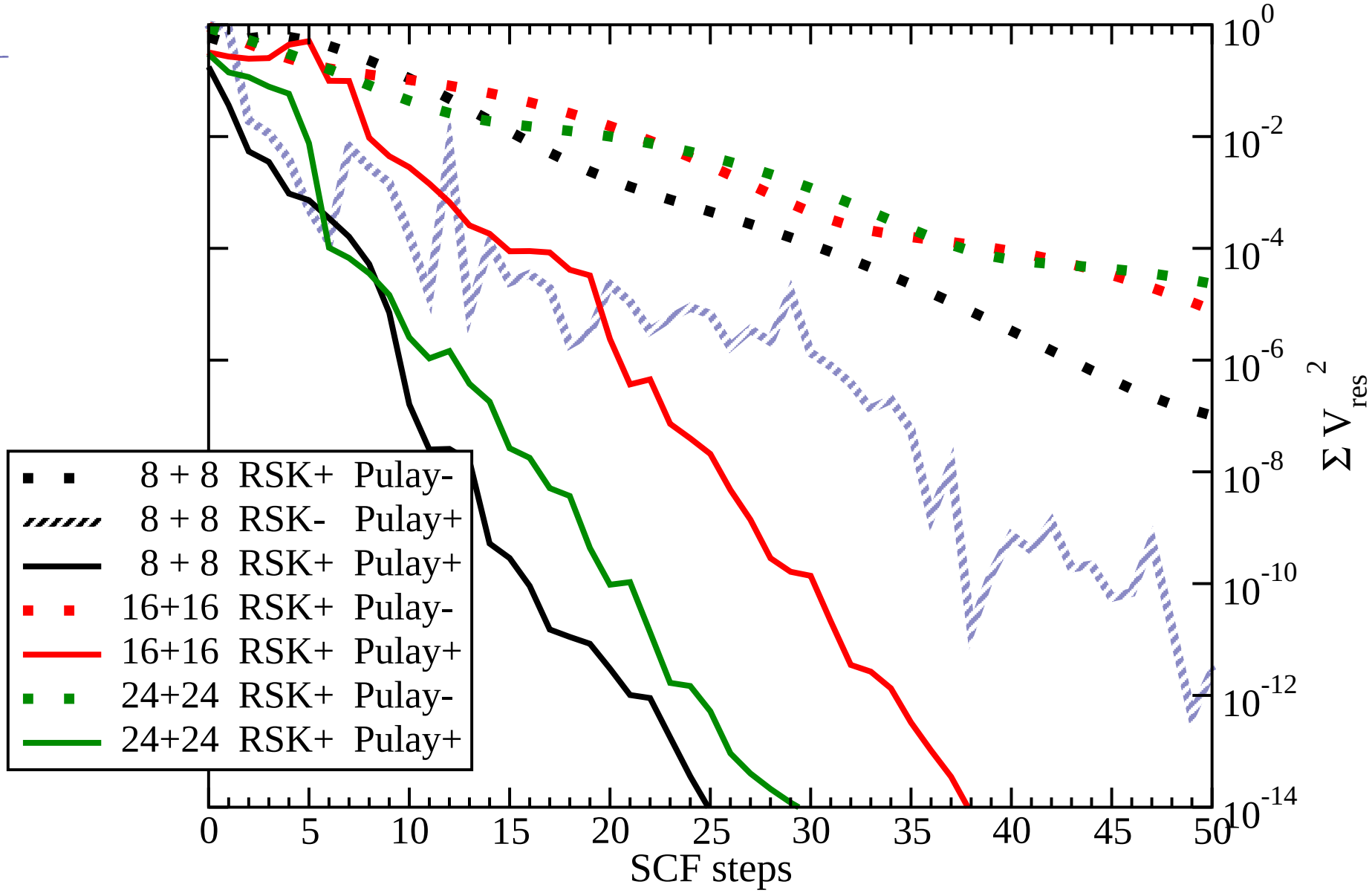
- Timings taken in Sr, half vacuum / half bulk linear cells (diecut=2 Ha, ecut=6 Ha)



History and preconditioners

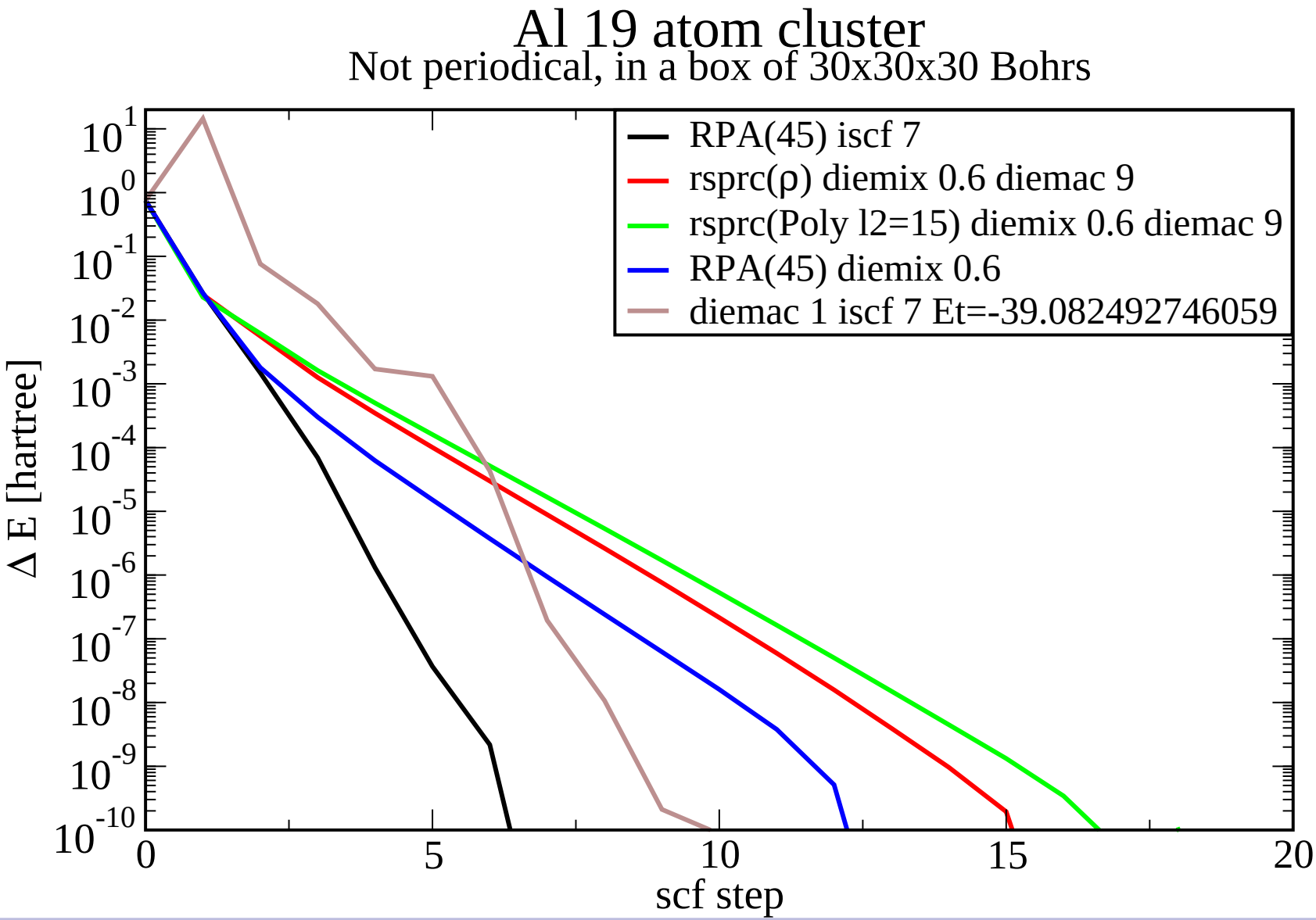


Advanced mixing & RSKerker

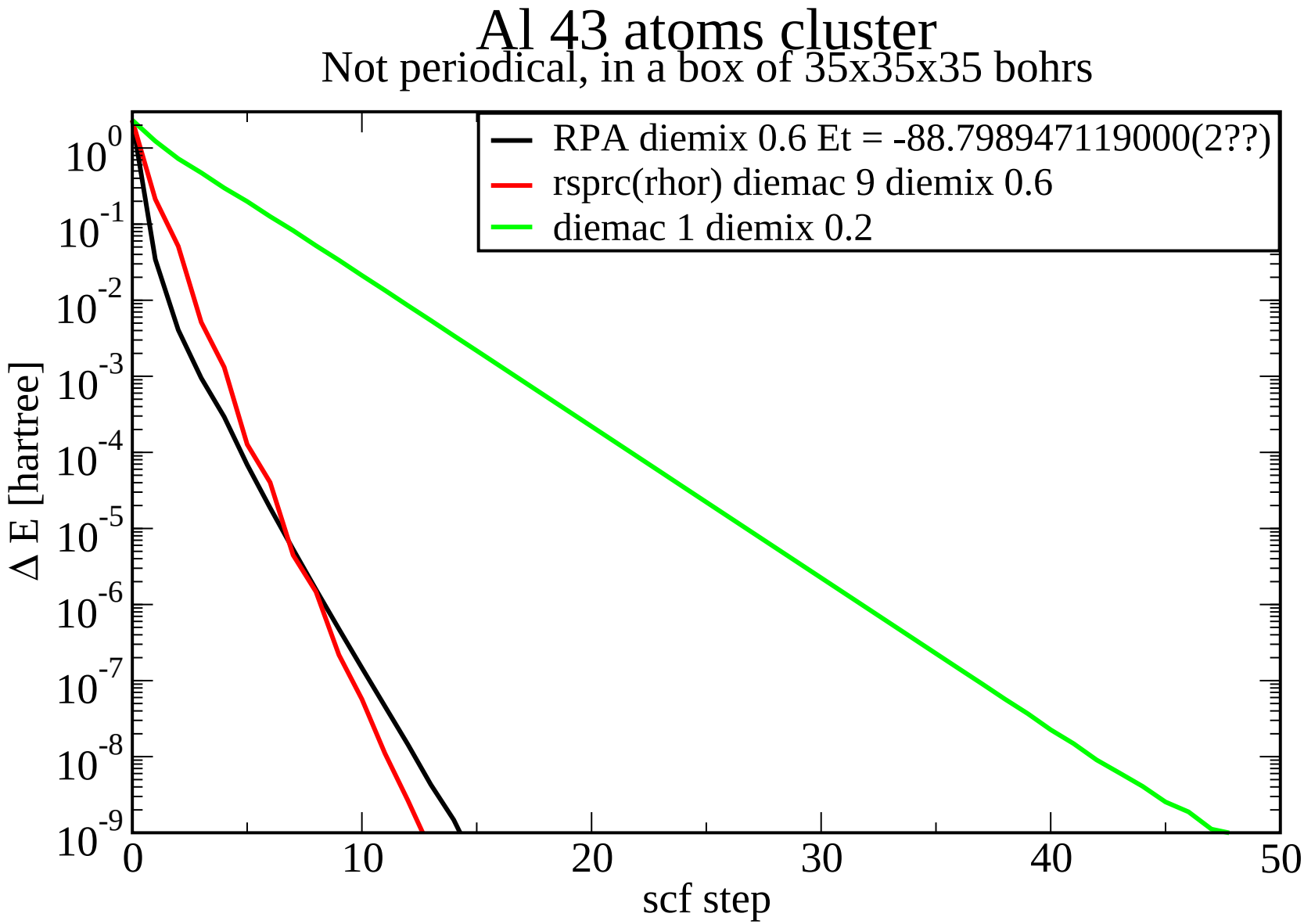


Conclusion

- Final words:
  - Extrapolar is perfect for all medium sized systems but the scaling will prevent the use for larger systems. However trying to keep reducing the prefactor might be fruitfull.
  - RSKerker is very promising for largest systems; yet it needs much better approximation for the real-space dielectric function. Uses of a $\Lambda(r)$ is also possible.
  - No clear scaling pattern when RSKerker is combined with Pulay mixing $\longrightarrow O(1)$?


Al 19 atom cluster
Not periodical, in a box of 30x30x30 Bohrs
RPA(45) iscf 7
rsprc(ρ) diemix 0.6 diemac 9
rsprc(Poly l2=15) diemix 0.6 diemac 9
RPA(45) diemix 0.6
diemac 1 iscf 7 Et=-39.082492746059
Δ E [hartree]
scf step
10^1
10^0
10^-1
10^-2
10^-3
10^-4
10^-5
10^-6
10^-7
10^-8
10^-9
10^-10
0
5
10
15
20

RSPRC


Al 43 atoms cluster
Not periodical, in a box of 35x35x35 bohrs
RPA diemix 0.6 Et = -88.798947119000(2??)
rsprc(rhor) diemac 9 diemix 0.6
diemac 1 diemix 0.2
Δ E [hartree]
scf step
10^0
10^-1
10^-2
10^-3
10^-4
10^-5
10^-6
10^-7
10^-8
10^-9
0
10
20
30
40
50

RSPRC

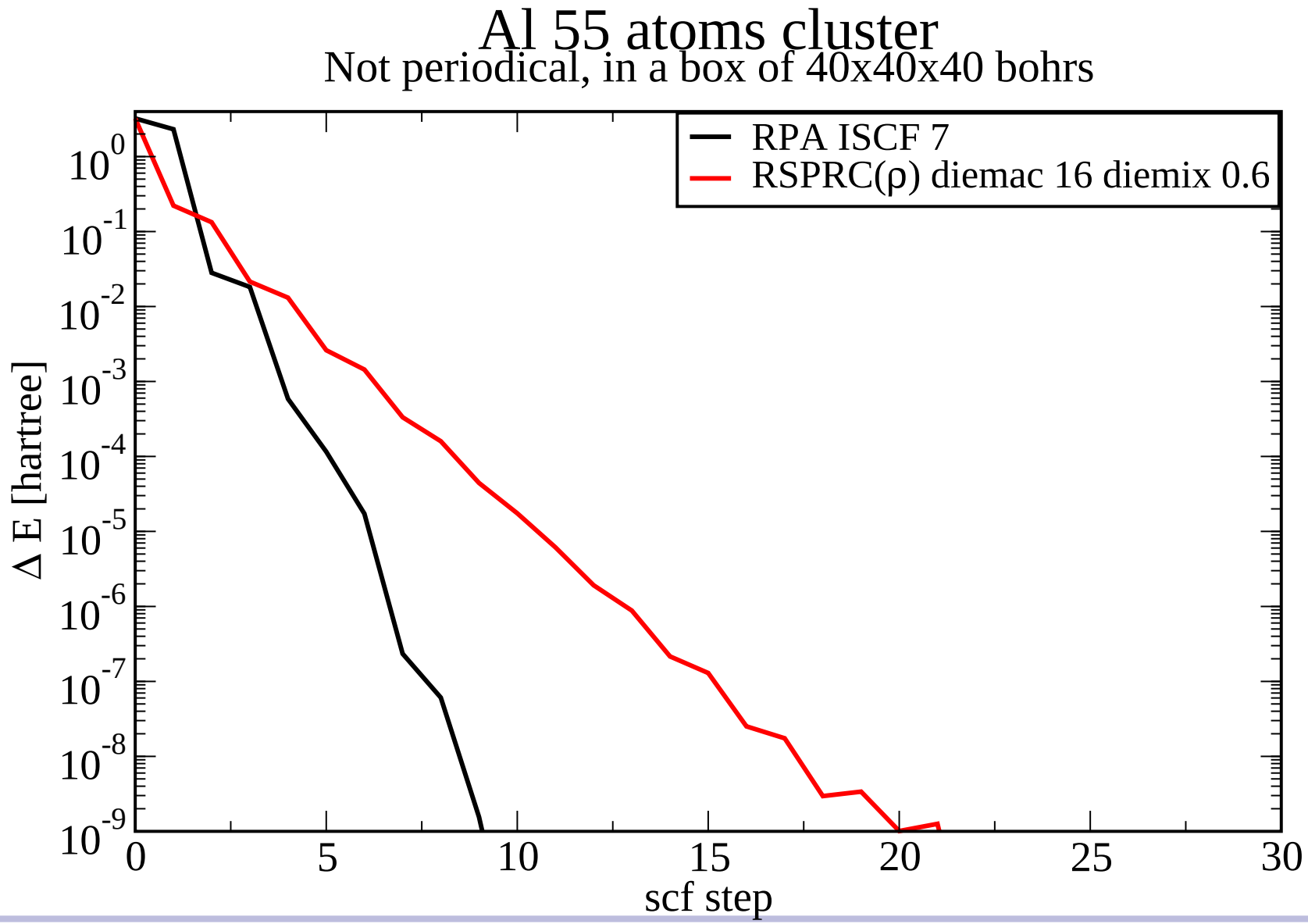

Al 55 atoms cluster
Not periodical, in a box of 40x40x40 bohrs
RPA ISCF 7
RSPRC(ρ) diemac 16 diemix 0.6
Δ E [hartree]
scf step
10^0
10^-1
10^-2
10^-3
10^-4
10^-5
10^-6
10^-7
10^-8
10^-9
0
5
10
15
20
25
30

Table:

	Si $\frac{\text{50\% vacuum}}{\text{50\% bulk}}$			Bulk Si		
type prc	λ_{max}	λ_{min}	CN	λ_{max}	λ_{min}	CN
RPA 24 bands	1.2	0.98	1.2	1.2	0.7	1.7
LDA 61 24 bands	NaN	NaN	NaN	1.2	0.82	1.5
RPA 64 bands	1.1	0.98	1.2	1.1	0.82	1.3
LDA 64 bands	2.3	0.99	2.3	1.2	0.92	1.3
Kerker Diemac 32	2.5	0.18	14	1.5	0.72	2.1
RSKerker Diemac 32	7.3	1.0	7.3	3.9	0.91	4.3
None	31	1.0	31	11	1.0	11