Periodic structure optimization via local heat pulse - quench cycles employing the GULP code

A. Möbius[1] and J.C. Schön[2]

[1]*Leibniz Institute for Solid State and Materials Research Dresden*
[2]*MPI für Festkörperforschung Stuttgart*

1. Introduction: Experience from combinatorial optimization
2. Optimization by local heat pulses for a continuous problem
3. Conclusions

CompPhys11, Leipzig, 25.11.2011

1. Introduction: Experience from combinatorial optimization

Demands on good heuristic algorithm:

- Robust
- Fast
- Low programming effort
- Small number of adjustable parameters
- Universally applicable

$\rightarrow$ In short: **KISS**

Consider simulated annealing: Highly advantageous concerning last three points, but speed, and therefore reliability in finding very deep local minima, not convincing.

Our thermal cycling = repeated disturbance by incomplete heating, mostly only for a part of the degrees of freedom, followed by quench

PHYSICAL REVIEW
LETTERS

VOLUME 79 | 1 DECEMBER 1997 | NUMBER 22

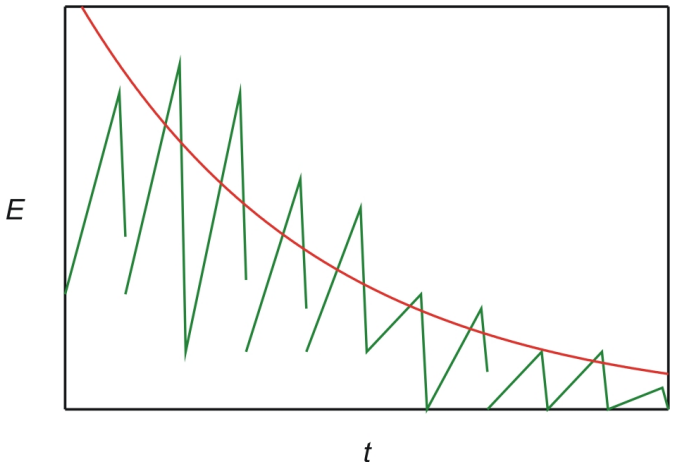
Optimization by Thermal Cycling

A. Möbius,[1] A. Neklioudov,[1] A. Díaz-Sánchez,[1,2] K. H. Hoffmann,[3] A. Fachat,[3] and M. Schreiber[3]
[1]*Institut für Festkörper- und Werkstofforschung, D-01171 Dresden, Germany*
[2]*Departamento de Física, Universidad de Murcia, E-30071 Murcia, Spain*
[3]*Institut für Physik, Technische Universität Chemnitz, D-09107 Chemnitz, Germany*
(Received 16 June 1997)

An optimization algorithm is presented which consists of cyclically heating and quenching by Metropolis and local search procedures, respectively. It works particularly well when it is applied to an archive of samples instead of to a single one. We demonstrate for the traveling salesman problem that this algorithm is far more efficient than the usual simulated annealing; our implementation can compete concerning speed with recent, very fast genetic local search algorithms. [S0031-9007(97)04649-8]

Aim: modify simulated annealing **making use of current altitude** above ground level.

Realization: **cyclically heating and quenching** with decreasing amplitude instead of slowly cooling down, **starting from lowest state** found.



Basic feature: **Incomplete heating** to retain gains of previous cycles.

Additional features: **Complex moves** can be incorporated in quenching.

Parallel search treating set of local minima, in order
(i) to reduce risk of getting trapped in "high" local minimum, and
(ii) to focus on "sensitive" degrees of freedom.

Relatives: Iterated Lin-Kernighan concept (Martin et al., 1992), genetic local search (Freisleben and Merz, 1996).

Comparison to basin hopping

Free to choose:

A) Exit old local minimum, in generalisation of move class in simulated annealing:

- kind of disturbance of current local minimum
- magnitude of disturbance

B) Determine new local minimum, in analogy to cooling schedule in simulated annealing:

- acceptance rate of new local minima
- time schedule of acceptance rate

In our approach: Focussing on A), and simplest approach for B), that is quench.

Travelling salesman problem (TSP)

Search for smallest round trip through given set of cities = NP complete problem

Simple approach: reversing of subtours => "long-range interaction" to be considered, but immediate surroundings of city particularly important. => Sorting helpful

TSP and e.g. Coulomb glass share many features such as huge number and hierarchy of local minima

=> Algorithms which were developed for one problem can be easily translated to other tasks and tested there

Padberg-Rinaldi problem of 532 US cities

Remark: Multi-start complex local search can be superior to simulated annealing, which can only incorporate moves of lower complexity

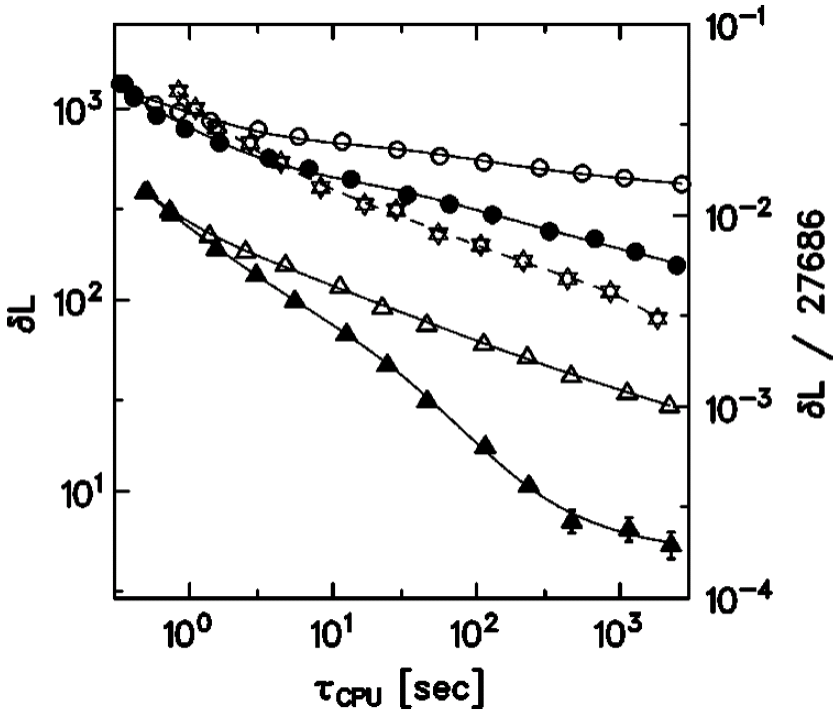
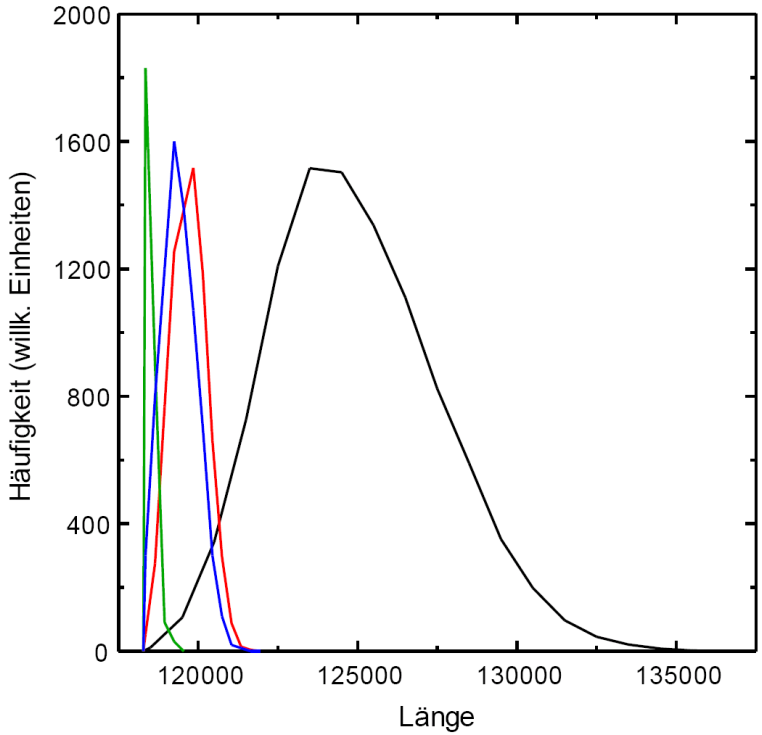


FIG. 1. Effect of embedding IPTLS in multistart local search: relation between computing time τ_{CPU} (in seconds) and average deviation, $\delta L = L_{mean} - 27686$, of the obtained approximate solution from the optimum tour length for the Padberg-Rinaldi 532 cities problem, att532. ○ (●) and △ (▲), multistart local search based on move classes (a) and (d), respectively, without (with) IPTLS; ☆, SA. In all cases, averages were taken from 100 runs; fluctuations (1σ region) are indicated by error bars if they exceed the symbol size. The lines, full for multistart local search, and dashed for SA, are guides to the eye only.

*A. M., B. Freisleben, P. Merz, M. Schreiber, PRE **59** (1999) 4667*

Remark: Simplest disturbance of old approximation = partial randomisation



Result by Franz Besold, obtained in two week school student internship (grade 9) by C++ code written from scratch

127 Augsburg beer garden problem.
Quench by shifting individual cites as well as turning the direction of subchains, without distance sorting.

Comparison of length distribution for

unmodified quench results,

best of 100 quenches,

threshold accepting, and

repeated randomization of subchains of (roughly) 30 cities, followed by quench, and choosing better of starting and final tour of cycle for continuation.

Advantages: very simple, possibly in some cases more efficient than thermal cycling

Efficiency of thermal cycling for 532 US city problem

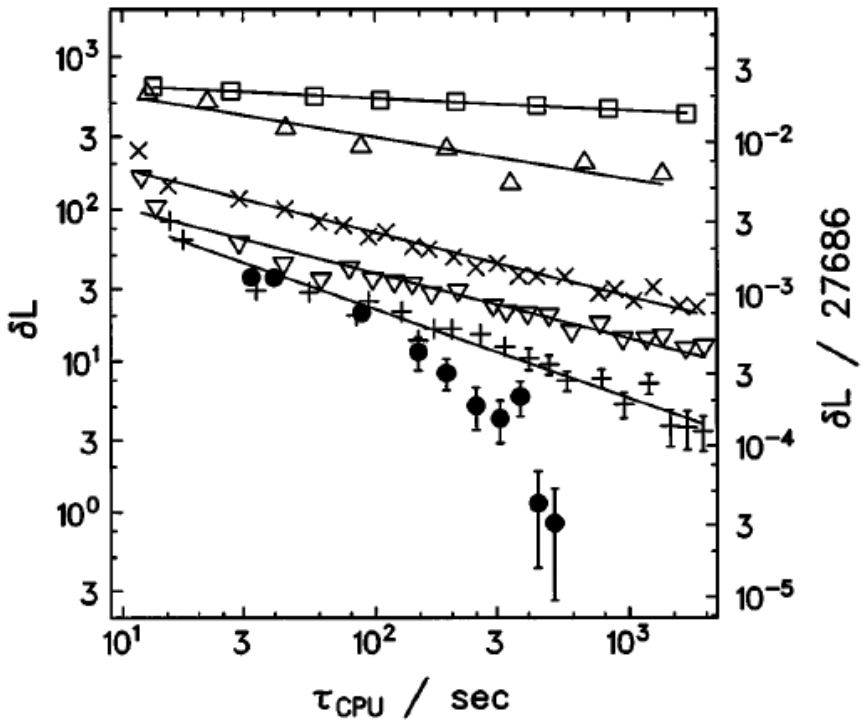
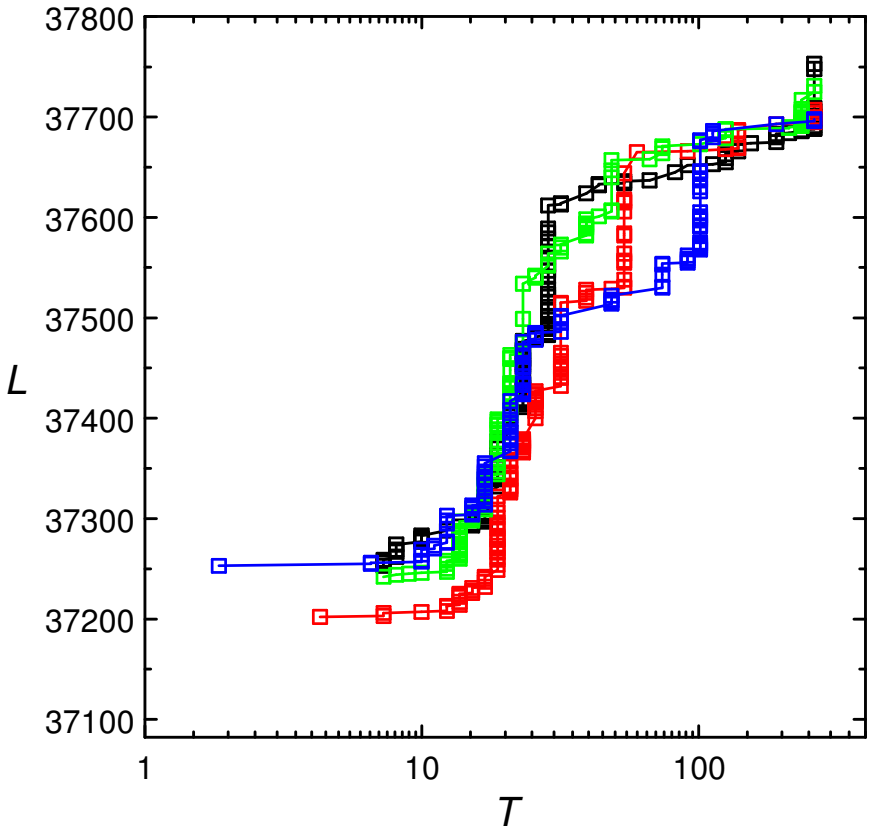


FIG. 1. Relation between CPU time, τ_{CPU} (in seconds for one PA8000 180 MHz processor of an HP K460), and deviation, $\delta L = L_{\mathrm{mean}} - 27\,686$, of the obtained mean approximate solution from the optimum tour length for the Padberg-Rinaldi 532 city problem. □: repeated quench to stability with respect to demand (a) defined in the text; △: simulated annealing; ×, ▽, +, •: thermal cycling with archives of various sizes, and local search concerning conditions (a), (b), (c), and (d), respectively. In all cases, averages were taken from 20 runs. Errors (1σ region) are presented if they exceed the symbol size. The straight lines are guides to the eye only.

Influence of heat-pulse *T* for xvb13584



Schedule: 3 trials at fixed temperature *T*, in case of success repetition, otherwise reduction of *T* by factor 0.9

=> Reduction of length difference to best known approximate solution by roughly "factor of 5" within 1 hour CPU time

=> There is / are optimum heat pulse temperature(s), where process works most effectively.

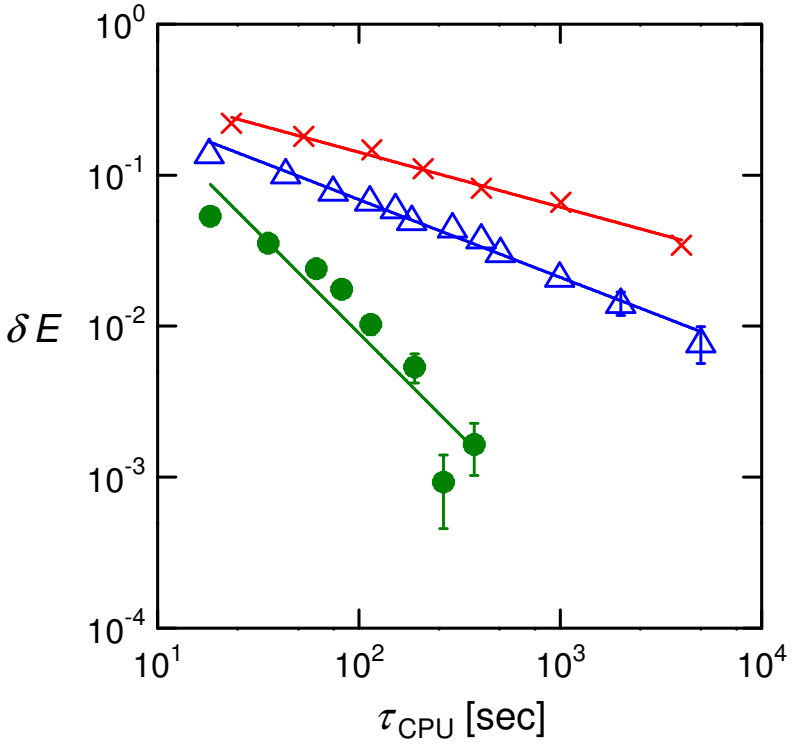
Application to Coulomb glass,

lattice model with 1000 sites and B = 1:

simulated annealing,

multi-start local search,

thermal cycling



*A. Díaz-Sánchez, A. M., M. Ortuno, A. Neklioudov, M. Schreiber, PRB **62** (2000) 8030*

Collective search

Strategies:

- Partition of computational effort into several search processes reduces failure risk, see Huberman et al., Science **275** (1997) 51. Optimum number depends on task, algorithm, and parameters.
- Compare set of approximate solutions to find sensible regions, e.g. in thermal cycling. This corresponds to elimination of backbones (only) in heating.
- Combination of best parts of two solutions. Can be used per se or as part of complex procedure as thermal cycling.
- Elimination of backbones in renormalisation procedure, see J. Houdayer and O. C. Martin, 1992 / 2008.

Unfortunately, for continuous problems, only first strategy can be simply implemented.

2. Optimization by local heat pulses for a continuous problem

Testproblem with 60 atoms: $Mg_{10}Al_4Ge_2Si_8O_{36}$ - lattice

Energy landscape defined by Buckingham and three-body potentials.

Parameters taken from:

A.R. Oganov, J.C. Schön, M. Jansen, S.M. Woodley, W.W. Tipton, R.G. Hennig, "Appendix: First blind test of inorganic crystal structure prediction methods" in "Modern methods of Crystal Structure Prediction", ed. A.R. Organov, Wiley, 2011

Results therein:

Random sampling: -1943.46 eV (best of 13029 minimizations)

Simulated annealing: -1949.10 eV (best of totally 685 minimizations along 9 SA runs)

Genetic algorithm: -1950.53 eV (best of totally 4610 minimizations along 2 GA runs)

Critical part: Local Search

We use:

General **U**tility **L**attice **P**rogram (GULP) by Julian D. Gale and co-workers, v. 3.4,
see J.D. Gale and A.L. Rohl, Molecular Simulation **29** (2003) 291.

Big advantage of this code:

Efficient analytic calculation of gradients

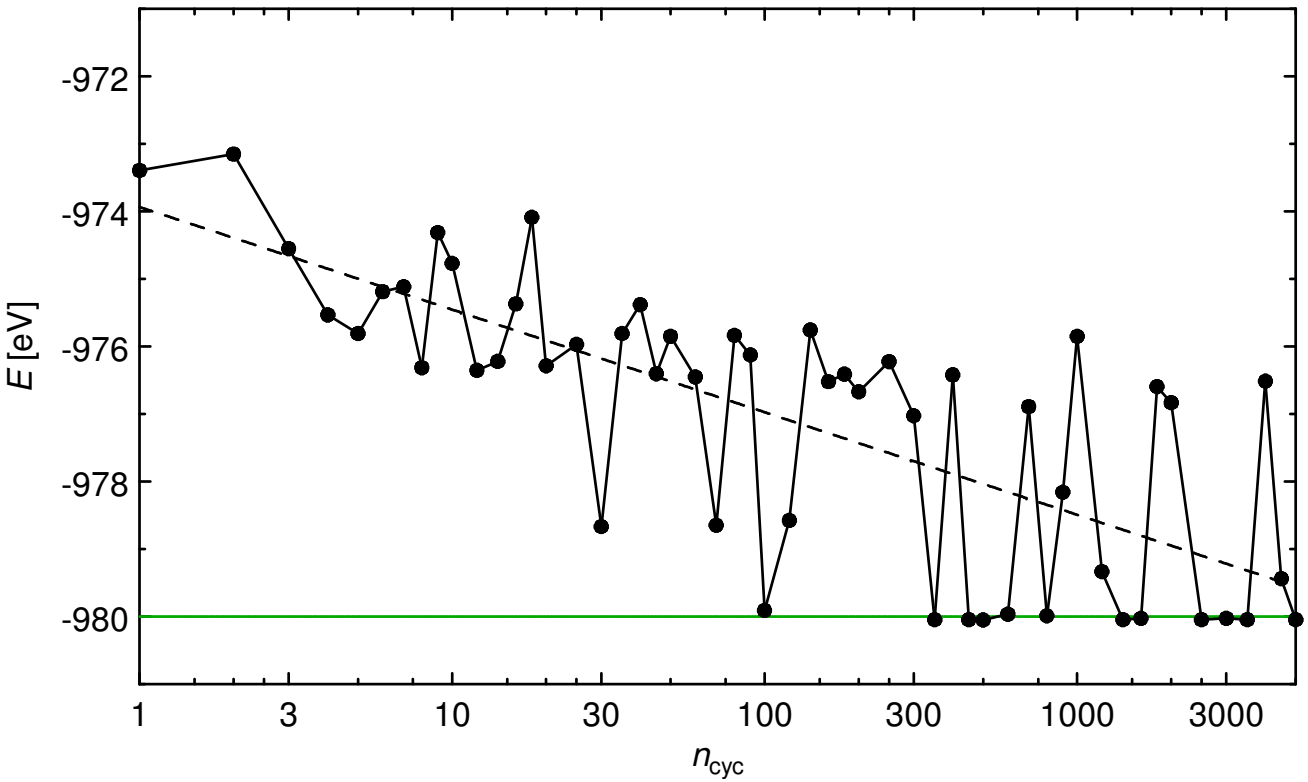
Our task: meta-program for preparing input files for GULP and analysing output

Our user experience:

Overall excellent code with a few problems in our application, partly arising from discontinuities of the considered energy landscape.

Numerical results

To gain first experience, let's simplify our life and consider half a cell: **$Mg_5Al_2GeSi_4O_{18}$**

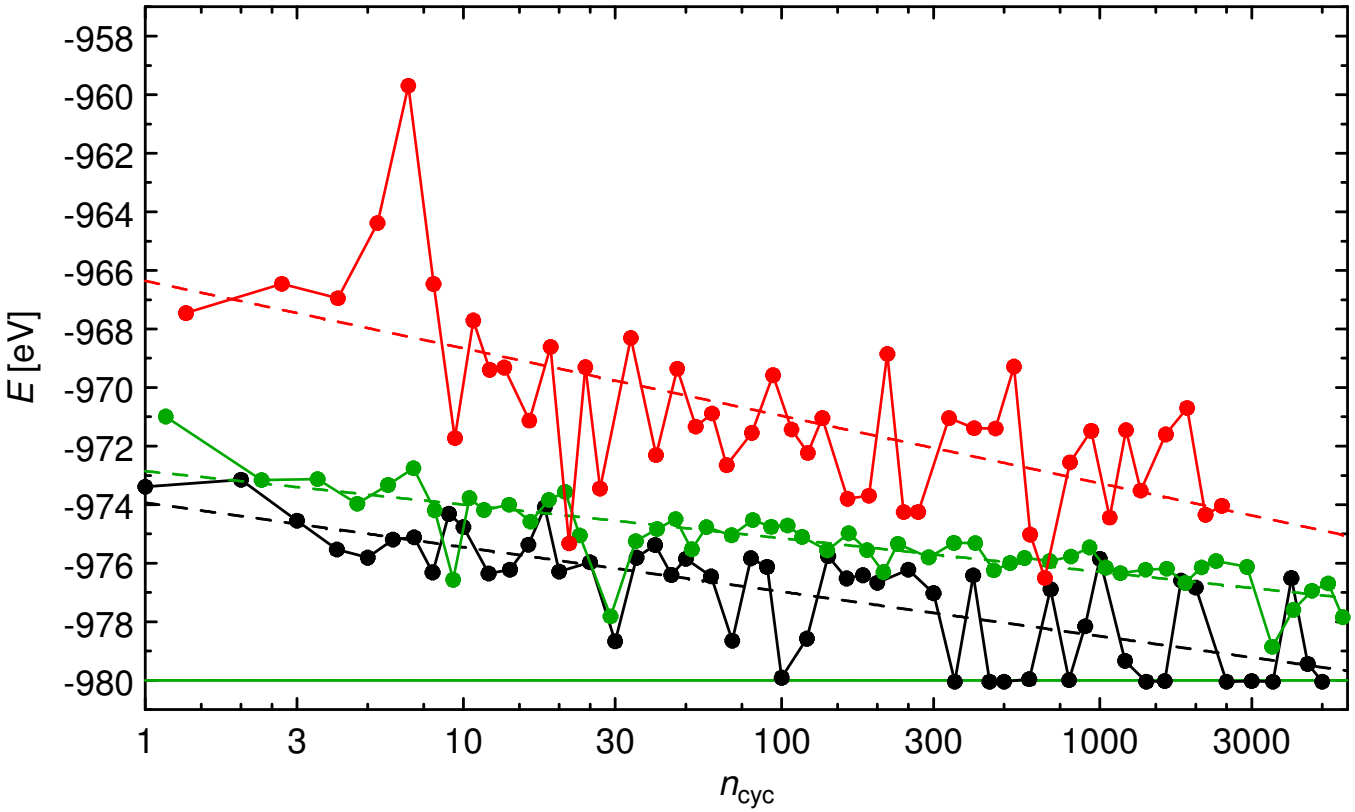


Amplitude starts at 1. It is stepwise reduced (factor 0.9). For each amplitude value, n_{cyc} cycles of heat-up (displacement of 5 atoms) and successive quench are performed.

One cycle needs roughly **80 s** on our cluster.

=> Optimization by repeated local heat pulses seems to work.

Comparison with **multi-start local search** and with simulated annealing.

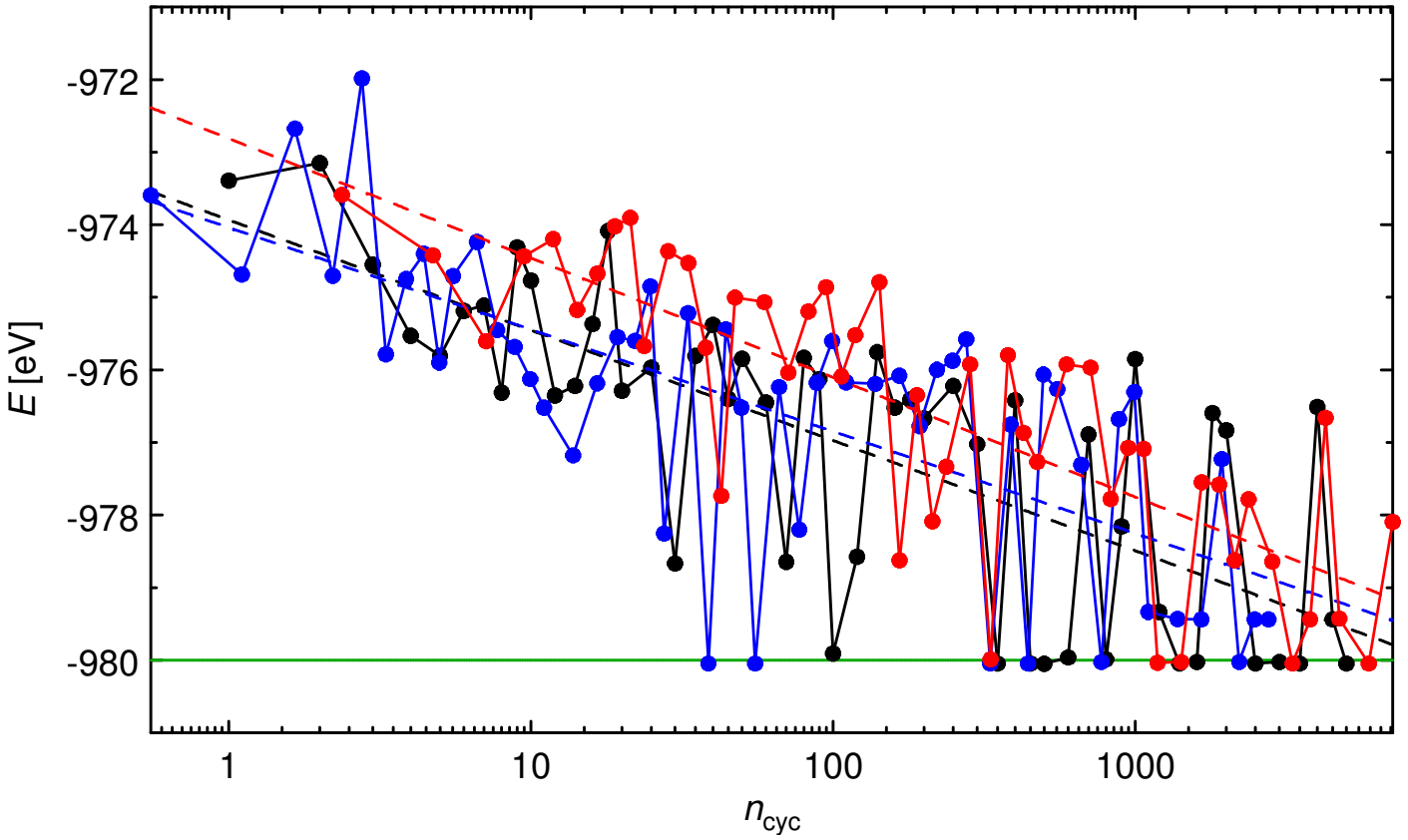


75000 random starts with successive quenches were performed in the longest multi-start local search run.

For a fair comparison, here, as well as in the following plots, cpu times (user) are appropriately rescaled. Thus the effort is related to n_{cyc} for heat-pulse optimization with standard parameters.

=> Only optimization by repeated local heat pulses reaches levels around -980 eV. Here, this method is clearly superior to multi-start local search as well as simulated annealing.

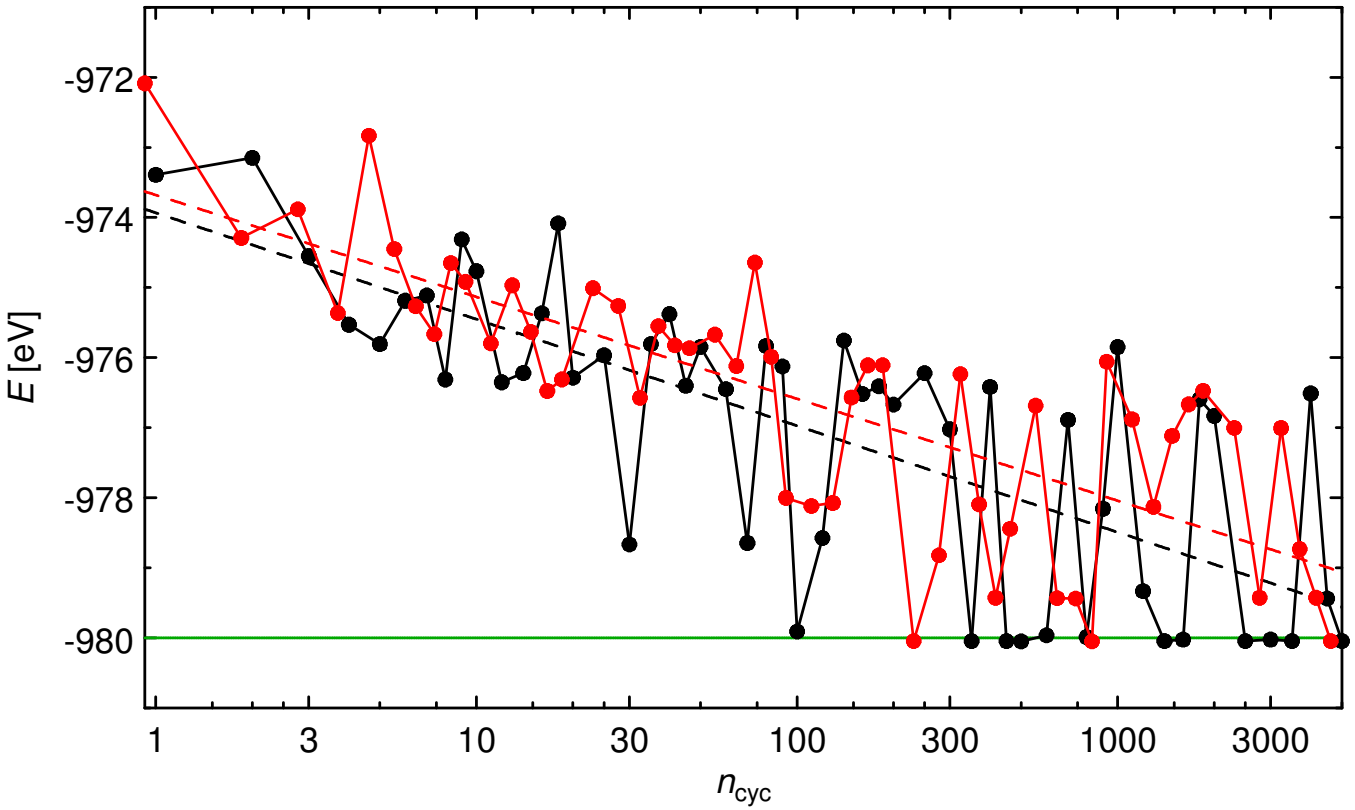
Comparison of optimization runs with 1, 5, and 30 atoms being displaced in the excitation.



For displacement of 1 atom in heating, similar behaviour as for 5 atoms is found. Also for 3 and for 8 atoms (not shown), similar results are obtained.
But for 30 atoms, the convergence is clearly worse.

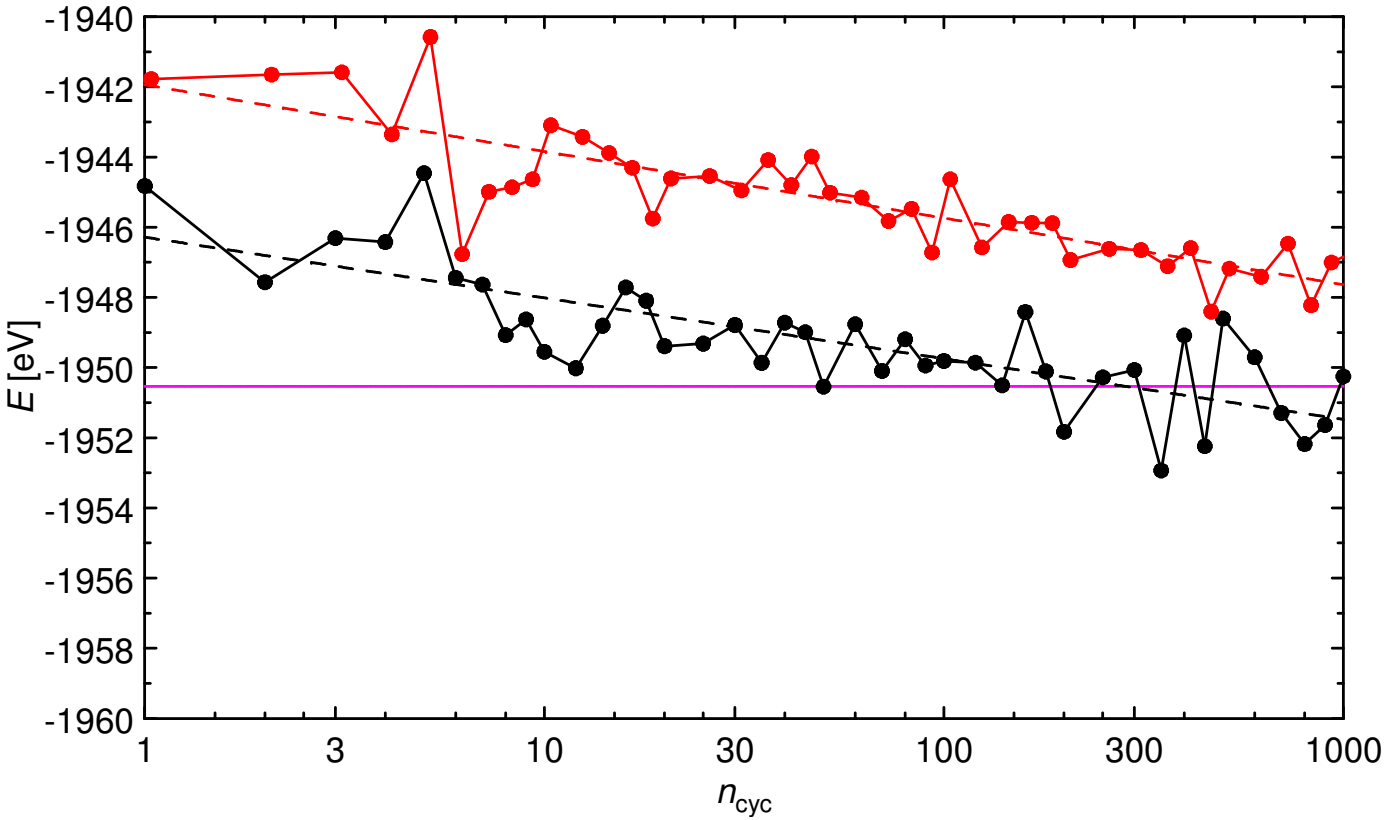
=> **Most efficient for small number of excited atoms, but value is not critical.**

Comparison of runs, in which 5 randomly chosen atoms and clusters of 5 atoms, respectively, are displaced.



=> **For cluster excitations, efficiency may be worse.**

Let's rise to the challenge: complete cell with 60 atoms



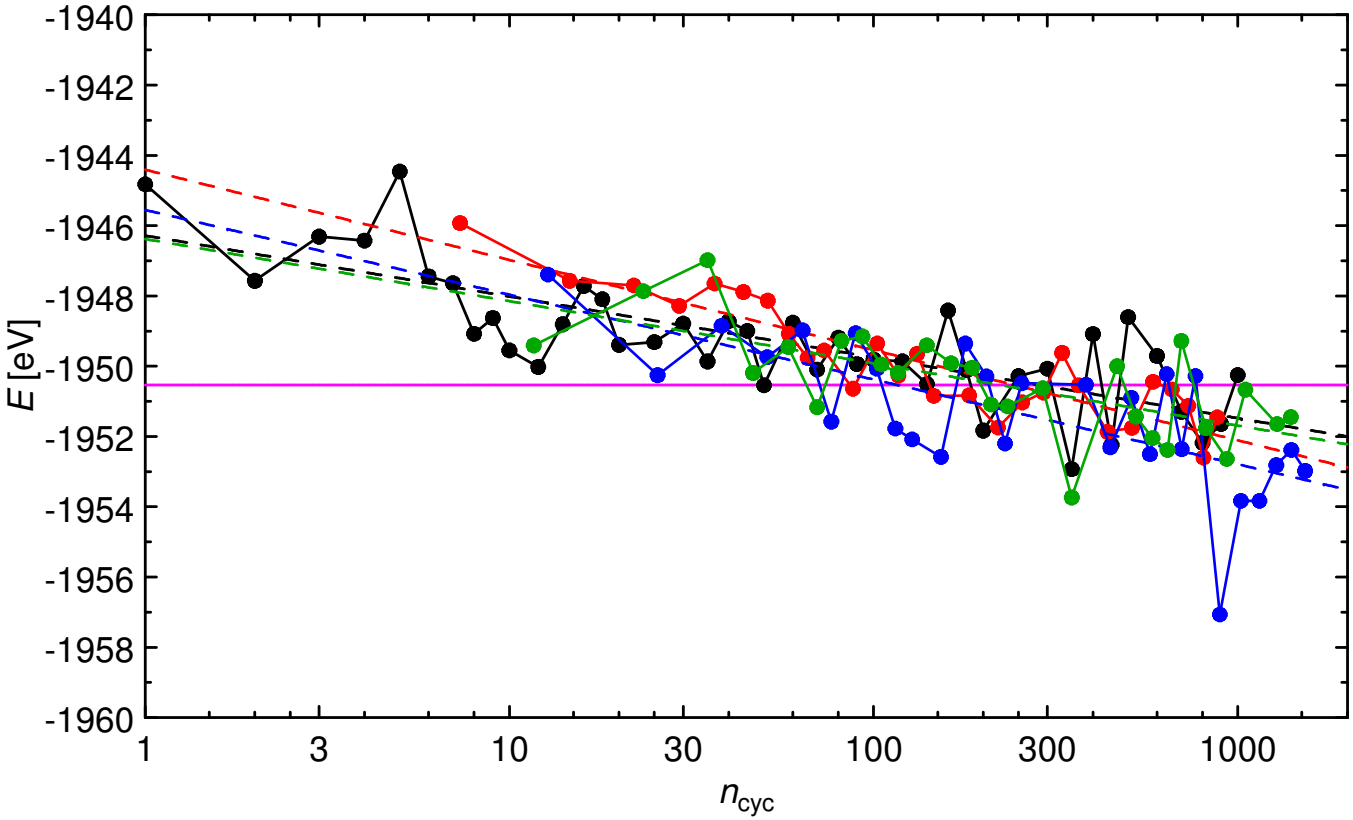
Comparison of optimization by local heat pulses with multi-start local search runs and with previously best known state.

Bottom = cell composed of two optimized half cells.

Cpu time range:
6 to 6000 min.

=> **Advantage of optimization by local heat pulses over multi-start local search considerably larger than for half cell.**

What about parallelization?

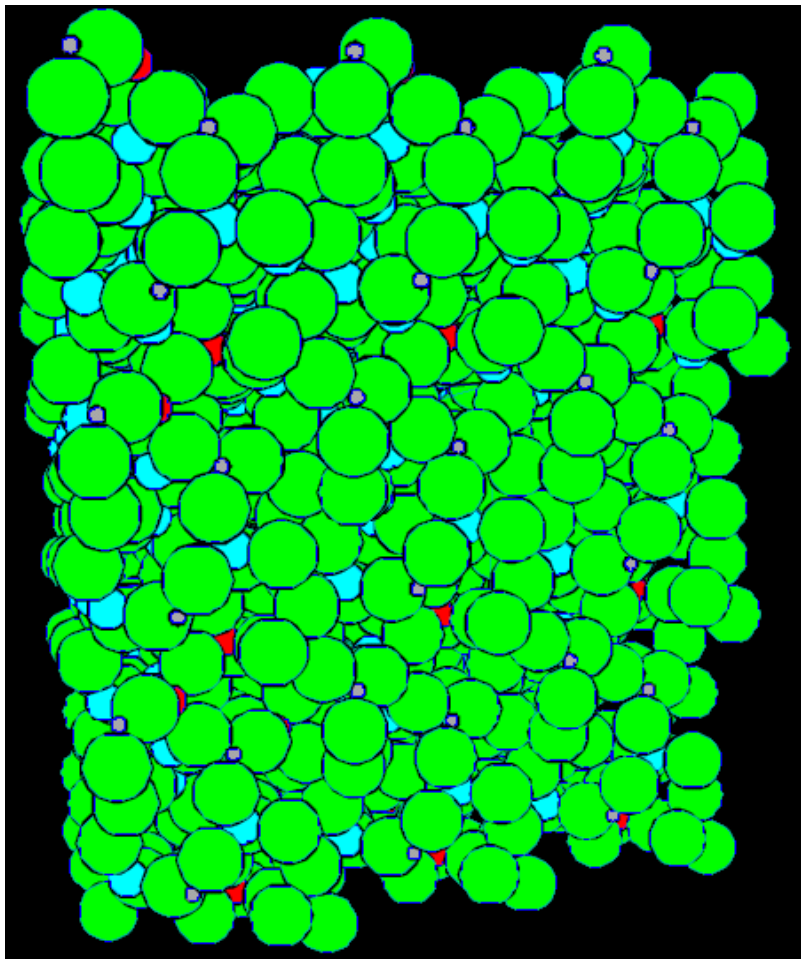


Comparison of standard procedure (one state considered) with results from 8 parallel runs without interaction, consideration of ensemble of 8 states, and basin hopping with fixed temperature.

Times relate to using only a single cpu.

=> **Parallelization might lead to further improvement by avoiding getting stuck in high minima. In the case considered, all methods used yield similar results.**

Now the solid state research can begin:



Mg

Al

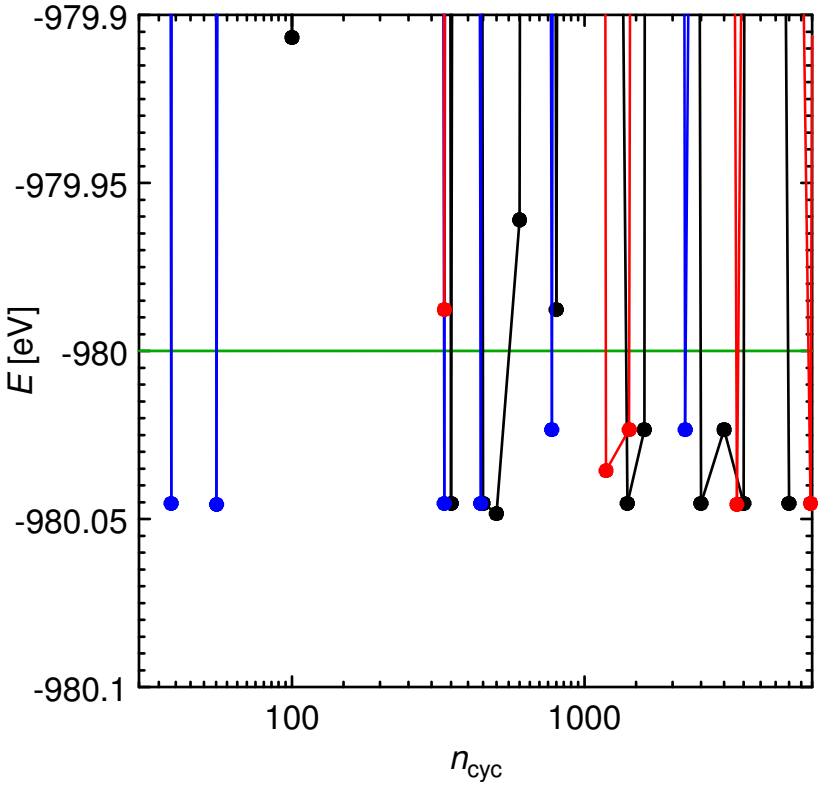
Si

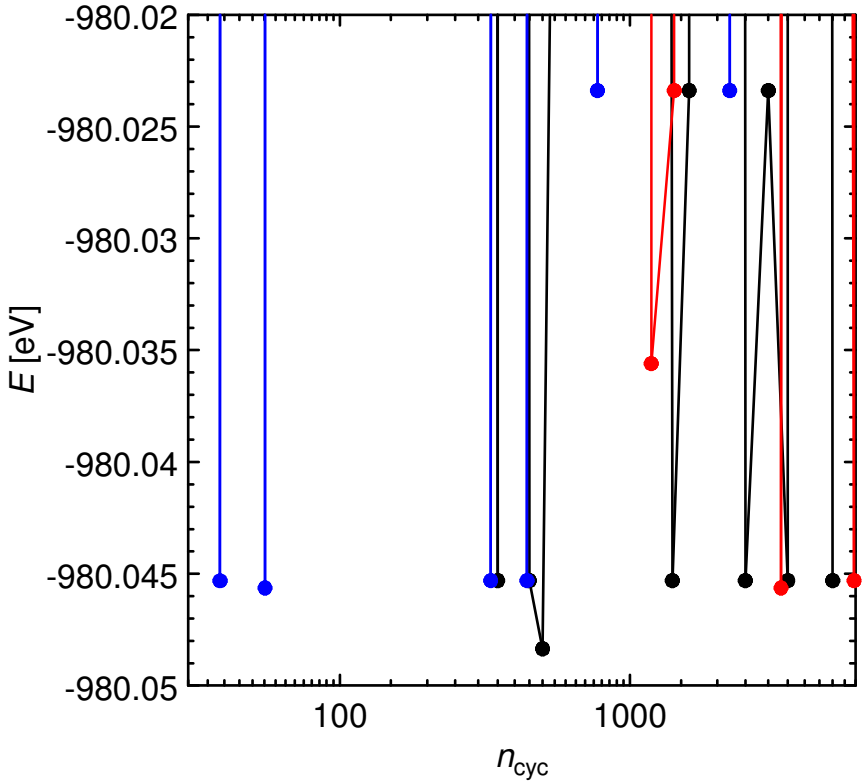
Ge

O

Fourfold coordinated Mg/(Al) cations may indicate an insufficient potential quality.

How can we understand the fine structure of the lowest half-cell states?





...

3. Conclusions

- Also for continuous variables, optimization by repeatedly “heating up” a small subset of the degrees of freedom and subsequent quenching has proved to be a robust and highly efficient algorithm.

- It can be easily parallelized.

- Talk reported on work under progress.

  => Remarks / questions are highly welcome!

- **Wanted: Colleagues, who try to apply this procedure to other tasks.**

  Support will be guaranteed.

Many thanks

to Karl Heinz Hoffmann for initiating this work,

to Ulrike Nitzsche for 7x24 IT support.

Problems in using GULP arising in our case:

Problem 1: According to documentation, GULP stops occasionally away from local minima.

Experience: If one starts far from equilibrium, convergence often slows down, or stops even though gradient is finite. Probable origin: Use of quasi-Newton algorithms inside GULP

Cure: Use rather low bound for number of iteration steps, restart GULP repeatedly.

```
START QUENCH

   ICYCLE   =         1
   ESTART =  -1681.809
   EFINAL   =  -1918.854

   ICYCLE   =         2
   ESTART =  -1918.854
   EFINAL   =  -1931.232

...

   ICYCLE   =         6
   ESTART =  -1942.293
   EFINAL   =  -1942.293

END QUENCH: EQF, ERROR =  -1942.293,   F
```

Problem 2: There is always a risk that the axes of the cell fold up.

Cure: After each minimization, try to reshape the cell by constructing **new basis vectors from sums / differences of old basis vectors**.

Principle: As long as possible, try to reduce the length of the basis vectors this way. Thus the angles of the basis tend towards orthogonality. Calculation of new fractional coordinates of atoms is simple: It consists of only adding / subtracting old components.

Moreover, if reshaping procedure detects too small a cell volume connected with the basis folding up, discard this minimization attempt.

Problem 3: The atoms are moved in finite steps. Occasionally the inter-atomic distances become too small so that deep unphysical energy minima are reached.

Cure: Check for such situations always before restarting GULP. If one of the inter-atomic distances is detected to be too small, **discard** this **minimization attempt**.

Problem 4: Repeatedly, the energy landscape was found to exhibit discontinuities. Currently the reason is unclear, handling of surface dipoles might cause them. => Further analysis needed.

Possible cure: Use of **simplex algorithm** for minimum search in these cases, not implemented yet.

Problem 5: Using GULP as canned software causes an intensive data transfer between RAM and disk while producing GULP input files and analysing its output.

Cure: Use **local disks**, in particular if calculations are performed on a cluster. Probably even better: use a **RAM disk**, not tried yet.

General question: Is quasi-Newton the most appropriate numerical approach far from equilibrium? Might simplex algorithm, line search approaches as Rosenbrock method, or Levenberg–Marquardt algorithm be more appropriate at the early stages of minimization?