Trajectory Methods in Minimization

Georgiana Diana Ojoc

October 29, 2019

Abstract

This paper reviews the application of three trajectory methods to minimization problems, namely First Improvement Hill Climbing, Best Improvement Hill Climbing and Simulated Annealing. Four benchmark functions with variable number of arguments are analysed, namely Ackley Function, Michalewics Function, Rastrigin Function and Schwefel Function.

1 Introduction

Global optimization is the selection of a best element (with regard to some criterion) from some set of available alternatives. Optimization problems of sorts arise in all quantitative disciplines, from computer science and engineering, to geophysics and economics, and the development of solution methods has been of interest in mathematics for centuries.

Besides combinatorial algorithms and convergent iterative methods, there are **heuristics**, and these techniques will be addressed in this report, because a heuristic is any algorithm designed for solving a problem more quickly when classic methods are too slow, or for finding an approximate solution when classic methods fail to find any exact solution.

2 Motivation

Hill climbing finds the global optimum out of the search space for convex problems, but for other problems it will find only local optima, meaning solutions that cannot be improved upon by any neighboring configurations, which are not necessarily the best possible solution out of all possible solutions.

To attempt to avoid getting stuck in local optima, one could use **simulated annealing**, which is suitable for problems where finding an approximate global optimum is more important than finding a precise local optimum.

The relative simplicity of hill climbing makes it a popular first choice amongst optimizing algorithms. Although more advanced algorithms such as simulated annealing may give better results, in some situations hill climbing works just as well. Hill climbing can often produce a better result than other algorithms when the amount of time available to perform a search is limited, such as with real-time systems.

3 Method

Each argument is represented as a randomly generated bit vector, so that the neighbors are chosen by bitwise negation.

In first improvement hill climbing, the first closer value is chosen, whereas in best improvement hill climbing all neighbours are compared and the closest to the solution is chosen.

In simulated annealing, the temperature is set to 1000, and the cooling rate is equal to 0.1.

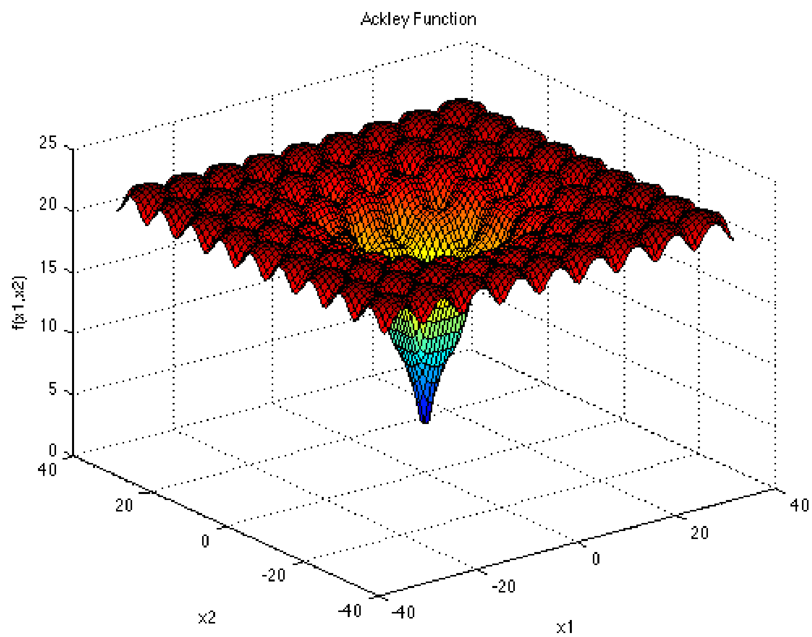
4 Results

All algorithms generated used 100 iterations and they have been run 32 times for 5, 10 and 30 dimensions. The elapsed time is measured in seconds.

Ackley Function

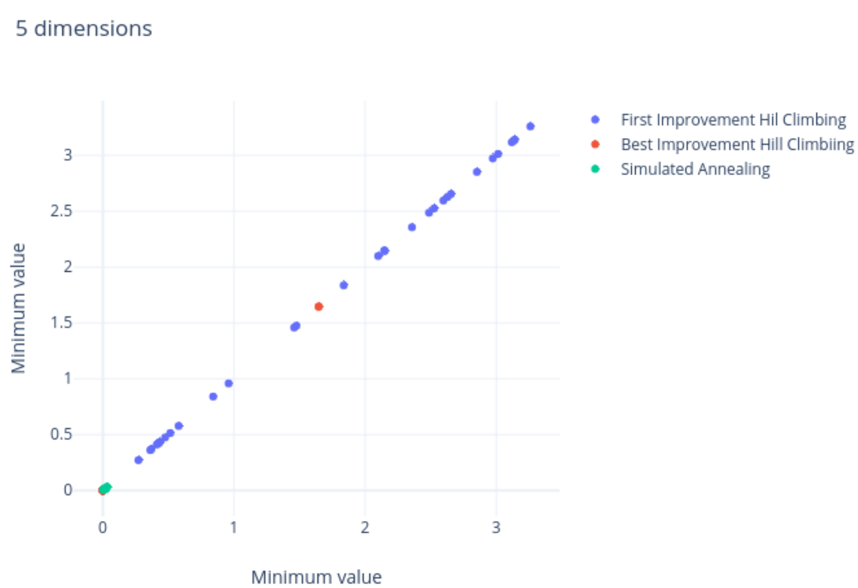
Ackley's function is a widely used multimodal test function.

- Number of variables: n variables
- Definition: $f(\mathbf{x}) = f(x_1, ..., x_n) = 20 + e - 20e^{-0.2\sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}} - e^{\frac{1}{n}\sum_{i=1}^{n} cos(2\pi x_i)}$
- Search domain: $-15 \leqslant x_i \leqslant 30,\ i = 1, 2, ..., n$
- Global minimum: $x^* = (0, 0, ..., 0),\ f(x^*) = 0$
- Function graph: for $n = 2$



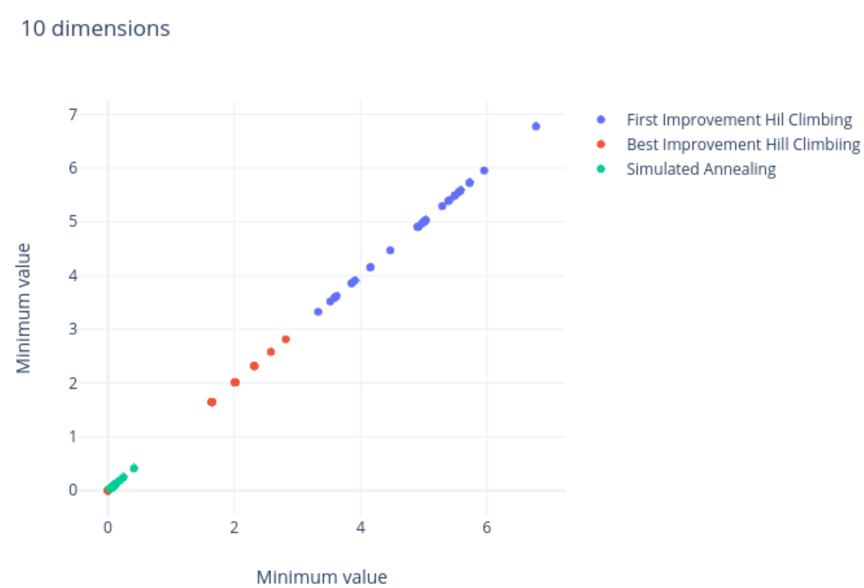
5 dimensions *Minimum value*	**Next Ascent Hill Climbing**	**Steepest Ascent Hill Climbing**	**Simulated Annealing**
Lowest value	0.271644	0.000000	0.006228
Highest value	3.259373	1.646224	0.031405
Mean	1.753228	0.102889	0.016966
Median	2.122435	0.000000	0.016935
Standard deviation	1.076154	0.404864	0.006094

5 dimensions *Elapsed seconds*	**Next Ascent Hill Climbing**	**Steepest Ascent Hill Climbing**	**Simulated Annealing**
Lowest value	0.380186	0.338660	1.100486
Highest value	0.449096	1.642838	1.145398
Mean	0.404842	0.713785	1.121963
Median	0.402426	0.642902	1.118713
Standard deviation	0.016443	0.310509	0.011663



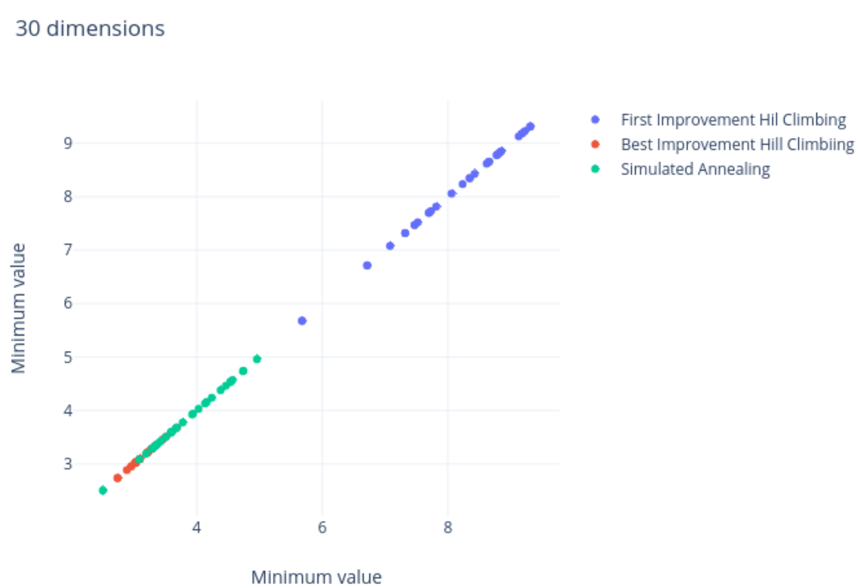
10 dimensions Minimum value	**Next Ascent Hill Climbing**	**Steepest Ascent Hill Climbing**	**Simulated Annealing**
Lowest value	3.326332	0.000000	0.028927
Highest value	6.775652	2.814350	0.409925
Mean	4.814584	1.630285	0.097009
Median	4.980963	1.646224	0.082380
Standard deviation	0.841747	0.775751	0.070873

10 dimensions Elapsed seconds	**Next Ascent Hill Climbing**	**Steepest Ascent Hill Climbing**	**Simulated Annealing**
Lowest value	2.133079	1.391911	2.581203
Highest value	2.489794	4.151548	2.894549
Mean	2.300421	2.350857	2.698487
Median	2.294451	2.205590	2.680843
Standard deviation	0.091338	0.704883	0.064962



30 dimensions *Minimum value*	**Next Ascent Hill Climbing**	**Steepest Ascent Hill Climbing**	**Simulated Annealing**
Lowest value	5.678198	2.738609	2.504724
Highest value	9.314037	3.682008	4.958828
Mean	8.045001	3.114515	3.831233
Median	8.290626	3.026937	3.671377
Standard deviation	0.975008	0.202957	0.577180

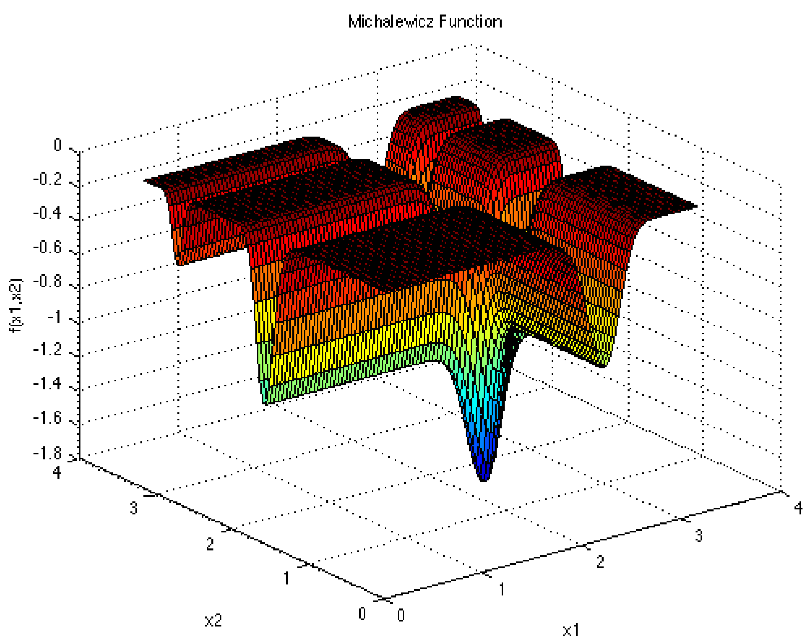
30 dimensions *Elapsed seconds*	**Next Ascent Hill Climbing**	**Steepest Ascent Hill Climbing**	**Simulated Annealing**
Lowest value	45.741335	31.796416	13.590985
Highest value	51.193883	587.651324	14.150504
Mean	47.753402	62.600942	13.747746
Median	47.867586	46.482138	13.702317
Standard deviation	0.942615	96.052137	0.138673



Michalewics Function

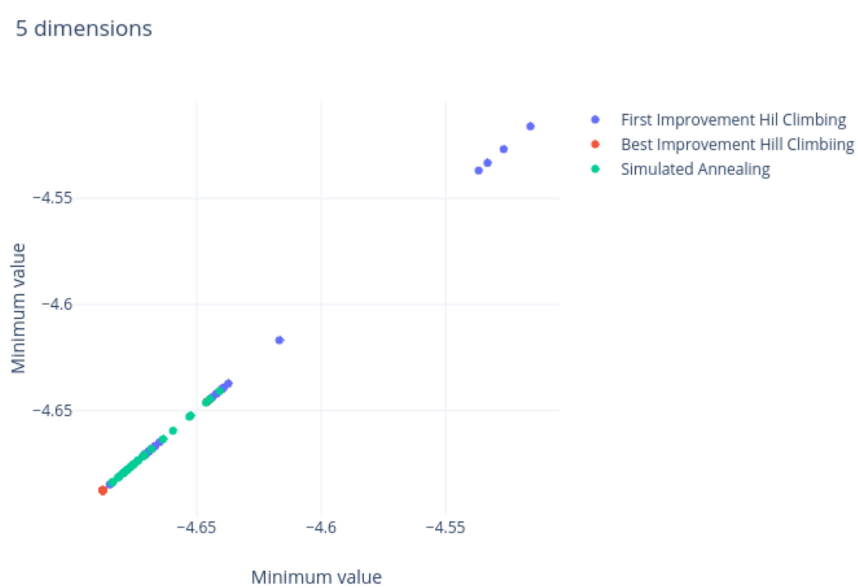
Michalewics' function is a multimodal test function. The exponent of the second sinus defines the "steepness" of the valleys or edges. Larger exponent leads to more difficult search. For very large exponent the function behaves like a needle in the haystack (the function values for points in the space outside the narrow peaks give very little information on the location of the global optimum).

- Number of variables: n variables
- Definition: $f(\mathbf{x}) = f(x_1, ..., x_n) = -\sum_{i=1}^{n} \sin(x_i) \sin^{20}(\frac{ix_i^2}{\pi})$
- Search domain: $0 \leqslant x_i \leqslant \pi$, $i = 1, 2, ..., n$
- Global minimum: $n = 5 : f(x^*) = -4.687658$, $n = 10 : f(x^*) = -9.660150$
- Function graph: for $n = 2$



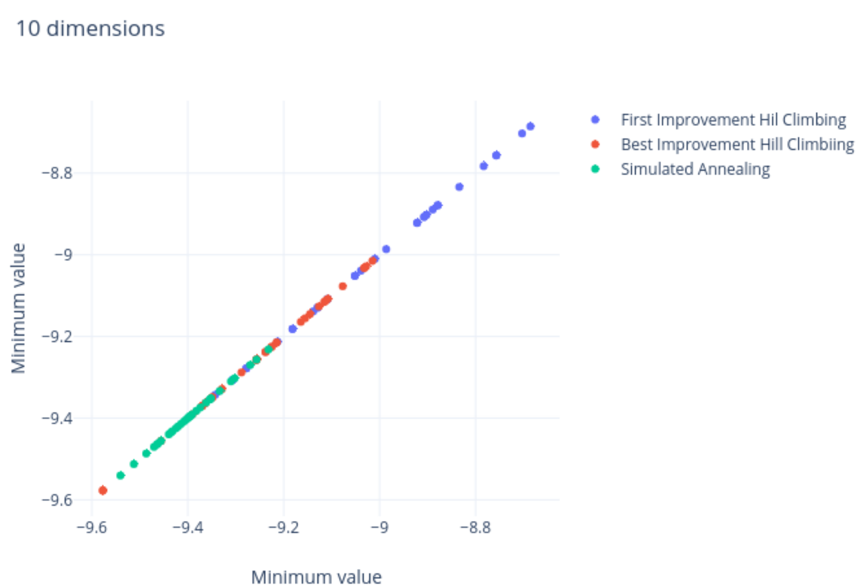
5 dimensions *Minimum value*	**Next Ascent Hill Climbing**	**Steepest Ascent Hill Climbing**	**Simulated Annealing**
Lowest value	-4.687506	-4.687658	-4.683960
Highest value	-4.516009	-4.645895	-4.640778
Mean	-4.644585	-4.682438	-4.670241
Median	-4.667276	-4.687658	-4.675162
Standard deviation	0.047689	0.014033	0.012620

5 dimensions *Elapsed seconds*	**Next Ascent Hill Climbing**	**Steepest Ascent Hill Climbing**	**Simulated Annealing**
Lowest value	0.454328	0.512152	1.799124
Highest value	0.490810	1.726793	1.963733
Mean	0.471344	0.909760	1.843238
Median	0.472403	0.857190	1.826547
Standard deviation	0.007361	0.339628	0.050879



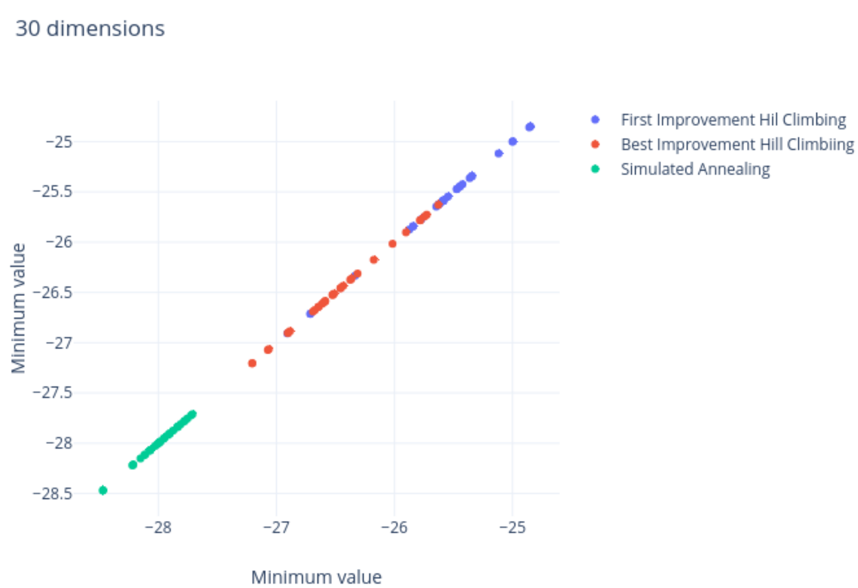
10 dimensions Minimum value	**Next Ascent Hill Climbing**	**Steepest Ascent Hill Climbing**	**Simulated Annealing**
Lowest value	-9.343045	-9.577276	-9.540572
Highest value	-8.685749	-9.015225	-9.231948
Mean	-9.017809	-9.239426	-9.394646
Median	-9.024617	-9.231762	-9.398126
Standard deviation	0.188814	0.147499	0.071194

10 dimensions Elapsed seconds	**Next Ascent Hill Climbing**	**Steepest Ascent Hill Climbing**	**Simulated Annealing**
Lowest value	2.938293	2.521623	5.226329
Highest value	3.312507	7.944412	5.825756
Mean	3.127907	4.235420	5.508513
Median	3.139609	3.584072	5.530187
Standard deviation	0.099865	1.403770	0.136778



30 dimensions *Minimum value*	**Next Ascent Hill Climbing**	**Steepest Ascent Hill Climbing**	**Simulated Annealing**
Lowest value	-26.907732	-27.205175	-28.470696
Highest value	-24.848757	-25.629206	-27.711190
Mean	-25.615436	-26.468908	-27.992412
Median	-25.585226	-26.524710	-27.984600
Standard deviation	0.485646	0.438445	0.190435

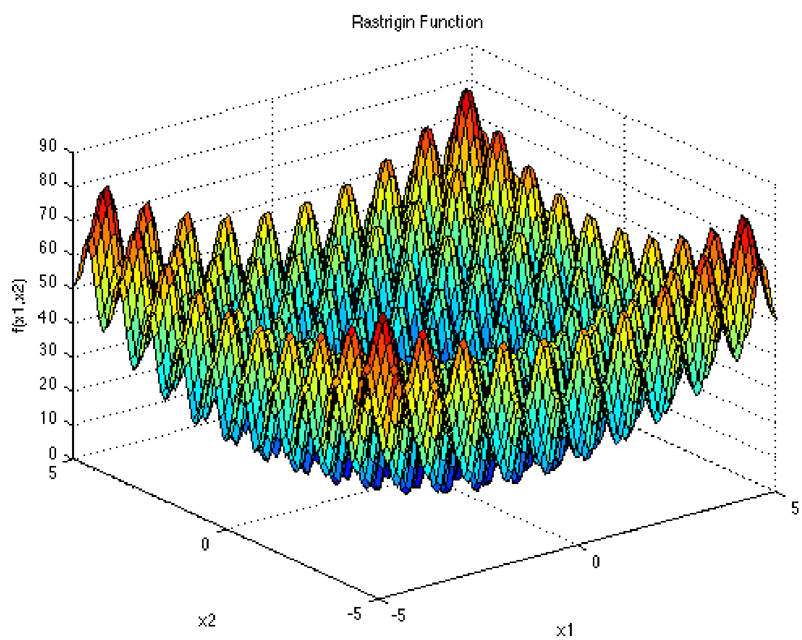
30 dimensions *Elapsed seconds*	**Next Ascent Hill Climbing**	**Steepest Ascent Hill Climbing**	**Simulated Annealing**
Lowest value	68.205660	62.247029	39.863832
Highest value	1107.560682	175.719554	43.304649
Mean	101.356755	104.936268	41.700896
Median	68.722056	103.348005	41.902054
Standard deviation	183.612021	28.154282	0.947249



Rastrigin Function

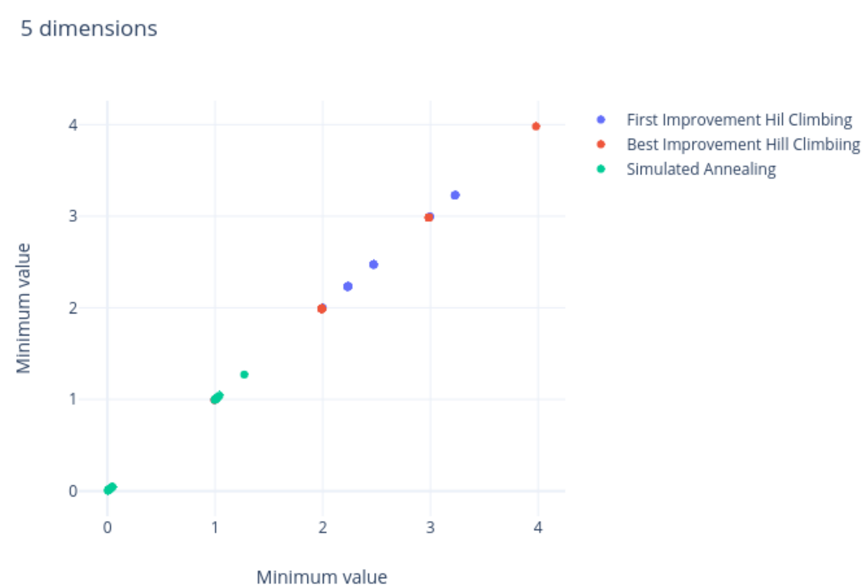
Rastrigin's function produces many local minima. Thus, the test function is highly multimodal. However, the location of the minima are regularly distributed.

- Number of variables: n variables
- Definition: $f(\mathbf{x}) = f(x_1, ..., x_n) = 10n + \sum_{i=1}^{n}(x_i^2 - 10cos(2\pi x_i))$
- Search domain: $-5.12 \leqslant x_i \leqslant 5.12,\ i = 1, 2, ..., n$
- Global minimum: $x^* = (0, 0, ..., 0),\ f(x^*) = 0$
- Function graph: for $n = 2$



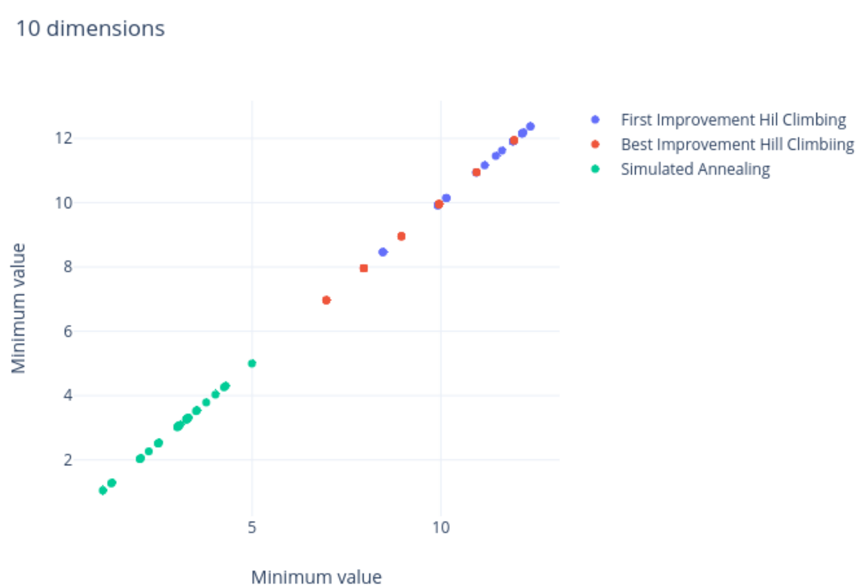
5 dimensions *Minimum value*	**Next Ascent Hill Climbing**	**Steepest Ascent Hill Climbing**	**Simulated Annealing**
Lowest value	0.994959	0.994959	0.005161
Highest value	3.230728	3.979836	1.271012
Mean	2.443212	2.300843	0.460844
Median	2.235768	1.989918	0.041462
Standard deviation	0.486063	0.641207	0.511944

5 dimensions *Elapsed seconds*	**Next Ascent Hill Climbing**	**Steepest Ascent Hill Climbing**	**Simulated Annealing**
Lowest value	0.228579	0.182609	0.995034
Highest value	0.239910	1.161846	1.028692
Mean	0.234747	0.494318	1.010188
Median	0.234624	0.324007	1.009101
Standard deviation	0.002887	0.328875	0.008828



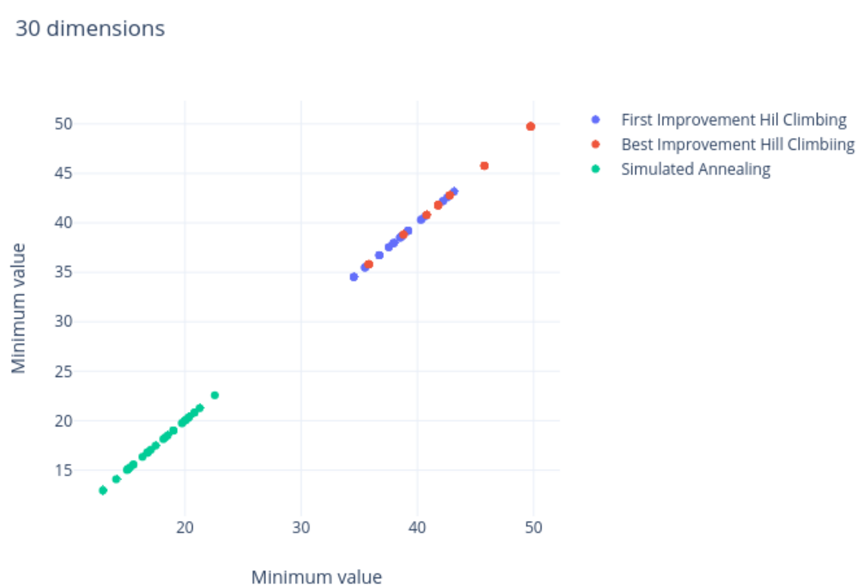
10 dimensions *Minimum value*	**Next Ascent Hill Climbing**	**Steepest Ascent Hill Climbing**	**Simulated Annealing**
Lowest value	8.461455	6.964708	1.047186
Highest value	12.369662	11.939504	4.995281
Mean	11.156503	9.141182	2.992019
Median	11.907784	8.954626	3.167510
Standard deviation	1.179549	1.483046	0.878797

10 dimensions *Elapsed seconds*	**Next Ascent Hill Climbing**	**Steepest Ascent Hill Climbing**	**Simulated Annealing**
Lowest value	1.093500	0.755659	2.161163
Highest value	1.266744	1.874628	2.543731
Mean	1.182207	1.314515	2.383484
Median	1.179192	1.348361	2.390092
Standard deviation	0.044147	0.315784	0.100580



30 dimensions Minimum value	**Next Ascent Hill Climbing**	**Steepest Ascent Hill Climbing**	**Simulated Annealing**
Lowest value	34.528340	35.818481	12.940115
Highest value	43.143851	49.747852	22.568770
Mean	38.320229	42.378966	17.798400
Median	38.496617	41.788210	17.809668
Standard deviation	2.328852	4.484311	2.402766

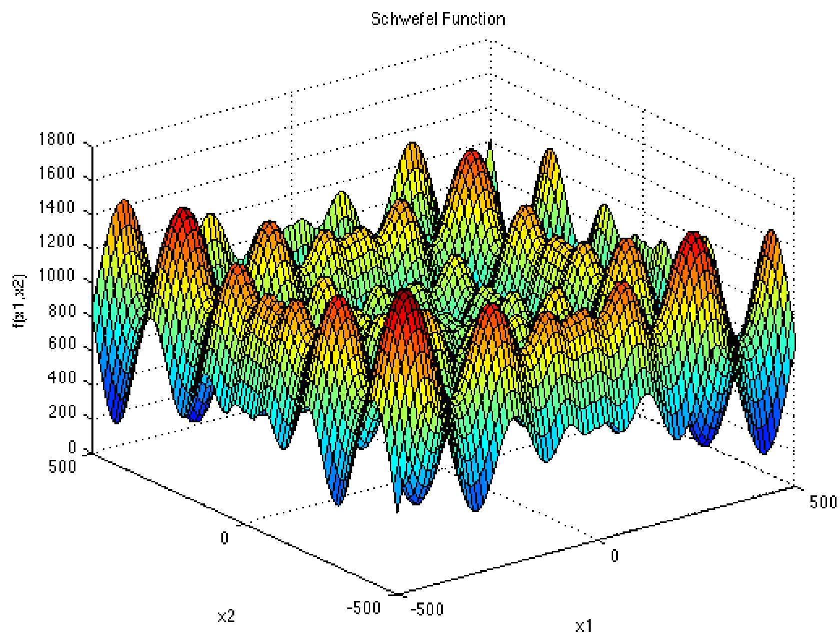
30 dimensions Elapsed seconds	**Next Ascent Hill Climbing**	**Steepest Ascent Hill Climbing**	**Simulated Annealing**
Lowest value	21.263931	25.544226	13.066887
Highest value	21.945812	42.342787	13.693978
Mean	21.512337	33.055792	13.251884
Median	21.445991	32.546085	13.207710
Standard deviation	0.178243	6.475660	0.149502



Schwefel Function

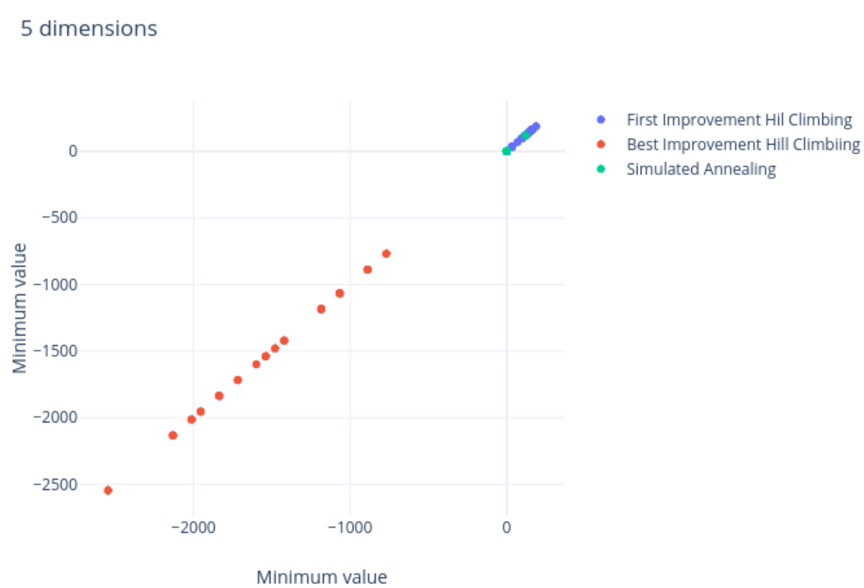
Schwefel's function is deceptive in that the global minimum is geometrically distant, over the parameter space, from the next best local minimum. Therefore, the search algorithms are potentially prone to convergence in the wrong direction.

- Number of variables: n variables
- Definition: $f(\mathbf{x}) = f(x_1, ..., x_n) = 418.9829n - \sum_{i=1}^{n} x_i \sin(\sqrt{|x_i|})$
- Search domain: $-500 \leqslant x_i \leqslant 500,\ i = 1, 2, ..., n$
- Global minimum: $x^* = (420.9687, 420.9687, ..., 420.9687),\ f(x^*) = 0$
- Function graph: for $n = 2$



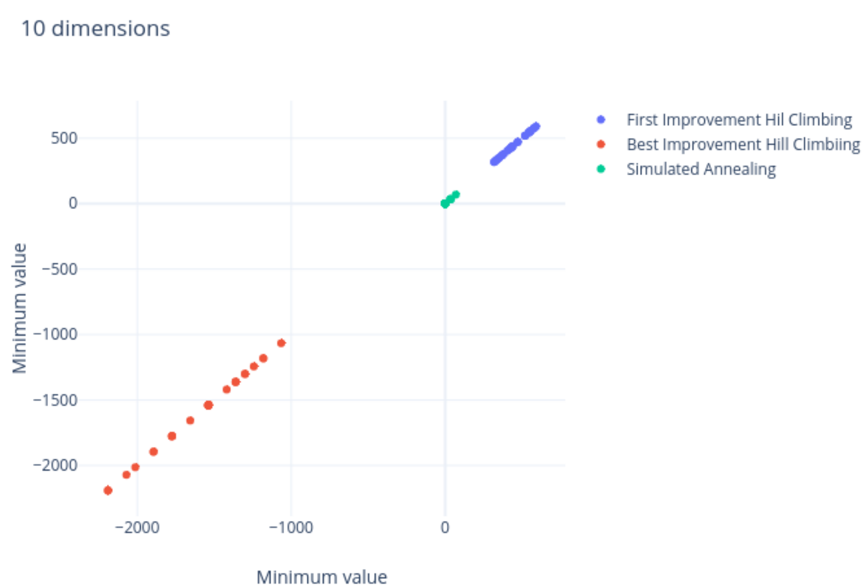
5 dimensions *Minimum value*	**Next Ascent Hill Climbing**	**Steepest Ascent Hill Climbing**	**Simulated Annealing**
Lowest value	0.104988	-2546.373939	0.005364
Highest value	187.019248	-769.835768	118.457059
Mean	113.511944	-1534.111869	3.803725
Median	124.466178	-1539.646112	0.119324
Standard deviation	50.981583	460.507656	20.922003

5 dimensions *Elapsed seconds*	**Next Ascent Hill Climbing**	**Steepest Ascent Hill Climbing**	**Simulated Annealing**
Lowest value	0.311774	0.345274	1.166670
Highest value	0.355430	1.066164	1.215562
Mean	0.322746	0.579244	1.193533
Median	0.320290	0.483143	1.193234
Standard deviation	0.009385	0.220177	0.010048



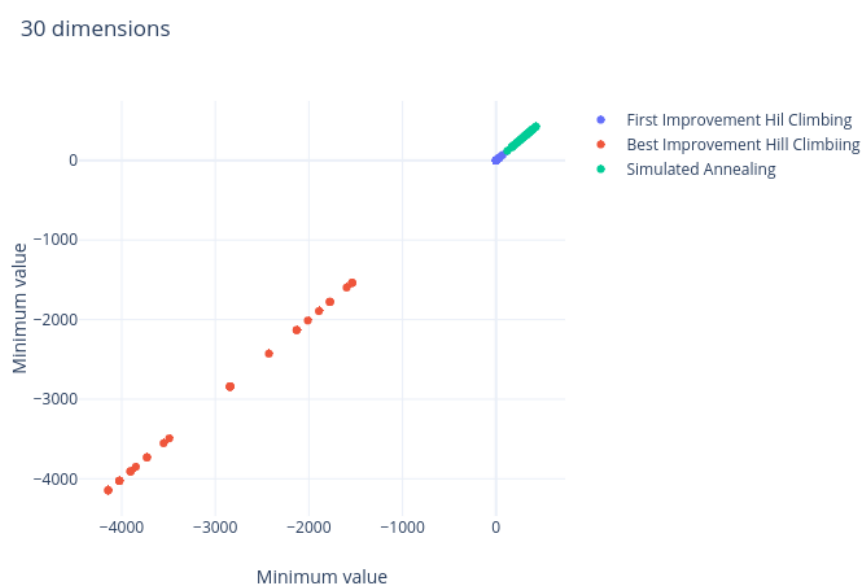
10 dimensions *Minimum value*	**Next Ascent Hill Climbing**	**Steepest Ascent Hill Climbing**	**Simulated Annealing**
Lowest value	318.245493	-2191.058871	0.139940
Highest value	590.017889	-1065.927094	68.828602
Mean	450.352956	-1584.881793	6.827592
Median	434.069567	-1539.662008	0.462836
Standard deviation	88.861298	288.008396	16.083109

10 dimensions *Elapsed seconds*	**Next Ascent Hill Climbing**	**Steepest Ascent Hill Climbing**	**Simulated Annealing**
Lowest value	1.752545	1.841768	3.174343
Highest value	2.080029	3.885900	3.696277
Mean	1.914168	2.879494	3.426429
Median	1.925839	2.833760	3.416261
Standard deviation	0.089681	0.685107	0.139101



30 dimensions *Minimum value*	**Next Ascent Hill Climbing**	**Steepest Ascent Hill Climbing**	**Simulated Annealing**
Lowest value	1419.395233	-4145.235784	119.700726
Highest value	2254.105152	-1538.136432	425.293153
Mean	1975.709800	-2926.939982	295.731647
Median	2002.961724	-2842.416569	311.421029
Standard deviation	204.189145	959.004882	80.957382

30 dimensions *Elapsed seconds*	**Next Ascent Hill Climbing**	**Steepest Ascent Hill Climbing**	**Simulated Annealing**
Lowest value	38.249965	34.826262	20.334840
Highest value	40.605872	50.380151	1341.703921
Mean	38.734819	42.271400	61.955161
Median	38.548092	41.354293	20.548915
Standard deviation	0.513578	5.012322	233.527809



5 Conclusions

For Ackley's function and Rastrigin's function, Simulated Annealing offers much better results than Next Ascent Hill Climbing and Steepest Ascent Hill Climbing. For 5 and 10 dimensions, Simulated Annealing is a little slower than the other two algorithms, but for 30 dimensions, it takes even a quarter of their time.

It is interesting that all methods give about the same results on Michalewics' function, because of the exponent of the second sinus is greater than 10.

Regarding Schwefel's function, none of these algorithms give a result close to the solution, because of the relatively high standard deviation of the minimum arguments.

References

- http://www-optima.amp.i.kyoto-u.ac.jp/member/student/hedar/Hedar_files/TestGO_files/Page364.htm
- https://www.sfu.ca/~ssurjano/ackley.html
- https://www.sfu.ca/~ssurjano/michal.html
- https://www.sfu.ca/~ssurjano/rastr.html
- https://www.sfu.ca/~ssurjano/schwef.html