Pulse Optimisation

Anurag Saha Roy
CQT, National University of Singapore

Contents

- Algorithms for Global Blackbox Noisy Optimisation
- Synthetic Functions for testing Algorithms
- Initial Benchmarking Results
- Preliminary Suggestions for Algorithm
- Preview of Algorithm Benchmarking Suite

Contents

List of Algorithms

- CMA
- Random Search
- Noisy Bandit
- Powell’s
- SPSA
- Differential Evolution
- PSA
- (1+1)
- Particle Swarm
- Nelder Mead
- Bayesian Optimisation
- Estimation of Distribution

Covariance Matrix Adaptation –Evolution Strategy

- Evolutionary Algorithm
- Update candidate solution from a multivariate normal distribution
- Update mean and covariance of the distribution during every iteration to sample better solutions
- Adaptation of Covariance Matrix allows learning of a second order model of the objective function
- Suitable for ill-conditioned functions

arXiv:1604.00772

List of Algorithms

Random Search

- Sample solutions from the entire search space
- Use a uniform probability distribution for sampling
- Each candidate solution is independent of any prior solutions
- Gives quick reasonable solutions for small parameter spaces

Pseudo-Code

List of Algorithms

Noisy Bandit

- Designed to make best use of a finite number of noisy evaluations
- Budget more evaluations of best-performing solutions (v/s weaker solutions) in order to get a better estimate (try to beat noise)
- Choose an action (candidate solution), calculate feedback, reward and regret.
- Update action to minimise regret

Bandit Theory
Noisy Bandit

Powell’s Method

- Conjugate Direction Method, ideally for local minimisation
- Performs bi-directional search along each vector from a set of search vectors
- The candidate is updated by a linear combination of the minima found along each search vector
- This linear combination is made a new search vector in the set and the most successful one from previous run removed

Powell Method Proof

List of Algorithms

Simultaneous Perturbation of Stochastic Approximation-SPSA

- Gradient Approximation (GA) Method
- Requires only 2 objective function measurements for GA irrespective of size of parameter space
- Simultaneously vary all the variables (as opposed to individually varying in Finite-Difference(FD) methods)
- Stochastic Optimisation using FD and SP achieve same accuracy but SP requires lower function evaluations

SPSA

List of Algorithms

Differential Evolution

- Evolutionary Method
- Maintain a set of candidate solutions
- Create new solution by combining existing solutions
- Accept/Reject new solution based on fit/score

DE for Global Optimisation

List of Algorithms

Porcellio-Scaber Algorithm

- Bio-inspired algorithm (based on survival rules of the PS worm)
- Suitable for Constrained Optimisation
- Applicable to mixed discrete-continuous cases
- Mimics an Exploration/Exploitation trade-off
- Generate position of the PS worm and update its environment
- Select position for most suitable environment

arXiv:1709.09840

List of Algorithms

(1+1) Evolutionary Algorithm

- Also known as random hill-climbing or metropolis algorithm at zero temperature.
- Key aspects – crossover, recombination, mutation & replacement (common steps for all EA)
- For a single parameter function with binary bitstring input, set probability p = 1/n (n is size of bitstring)
- Choose a random initial bit string (x_0)
- Independently flip bits in x_0 with probability p to get x1
- Replace if fitness of x_0 isn't better than fitness of x_1
- The replacement strategy is called 1+1
- Suitable for global, noisy optimisation

List of Algorithms

Analysis of 1+1

Particle Swarm Optimisation

- Based on Swarm Intelligence.
- Have a set of candidate solutions and update based on their position and velocity
- The particles try to move towards both the local(particle) optimum and the global(swarm) optimum
- The velocity and the position are updated after each iteration

Kennedy-Eberhart 1995
Application – Poli 2007
Bonyadi-Michalewicz 2017

List of Algorithms

Nelder Mead Method

- Direct Search Method
- Also known as downhill simplex method
- Maintain an (n+1) dimension simplex for n-dimension problem
- Check at each vertex of simplex and discard the worst vertex
- New point is obtained from the remaining points by either stretching or shrinking
- Number of iterations for a sufficiently good optimum can be very high

Details
Usage

List of Algorithms

Bayesian Optimisation

- Use a Gaussian Process (GP) to produce a Statistical Model of the objective function
- Obtain a Bayesian posterior probability distribution to predict potential values at a candidate point
- Update probability distribution every time the function is evaluated at a candidate point
- Acquisition function is used to generate new candidates based on the distribution & candidate evaluation
- Also possible to use other non-GP surrogate models
- Unsuitable for high-dimensions

Review
Tutorial (1807.02811)
Visualisation

List of Algorithms

Estimation of Distribution

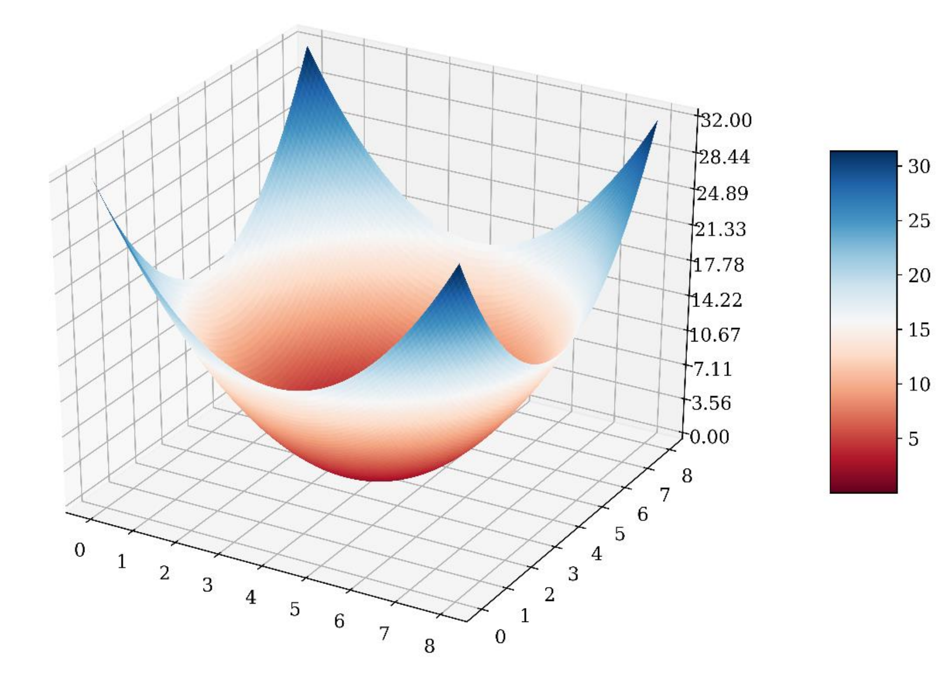
- Population based search for global optimisation
- Mix of Population-based & Probability-based algorithms
- Incorporate probabilistic modelling of promising candidate solutions
- Simulation of the induced models is used to generate new populations as opposed to crossover or mutation
- Suitable for noisy optimisation

Review

List of Algorithms

Contents

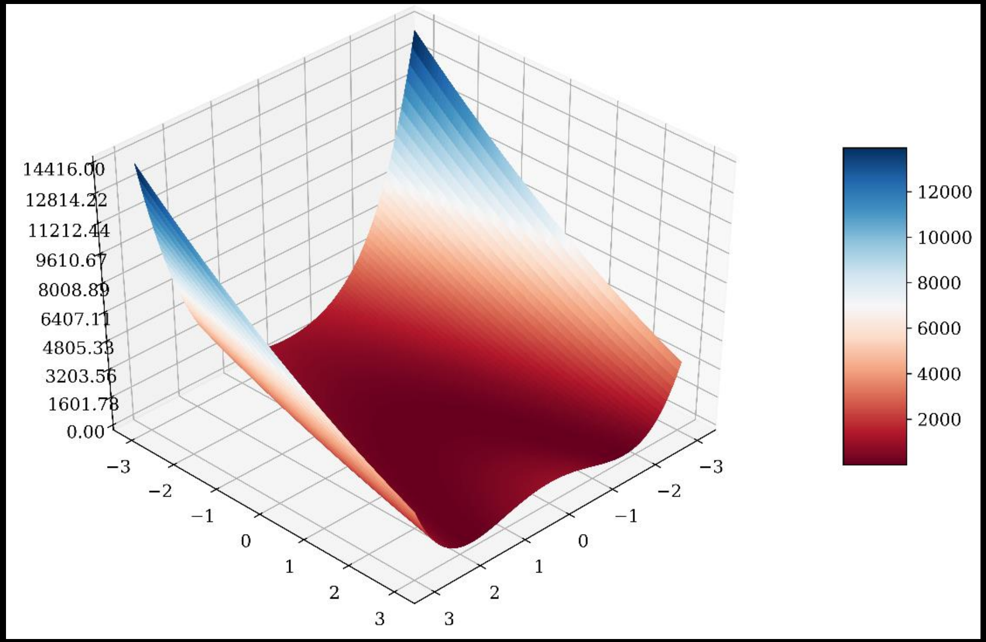
Translated Sphere



- Continuous
- Convex
- Parabolic

Details

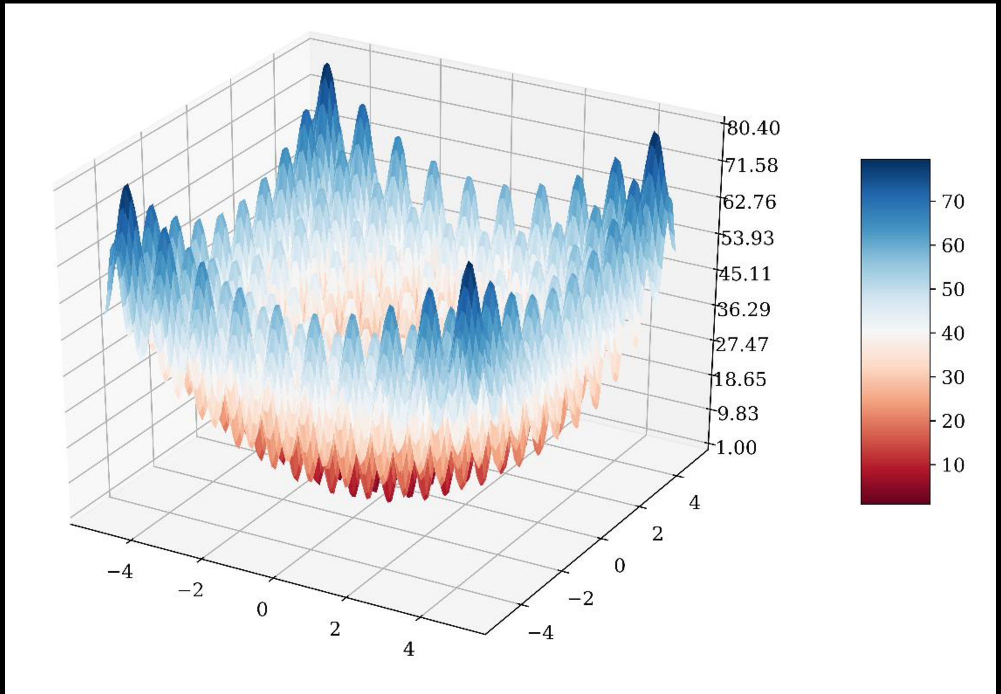
Rosenbrock



- Non-Convex
- Minima inside a narrow parabolic valley

Details

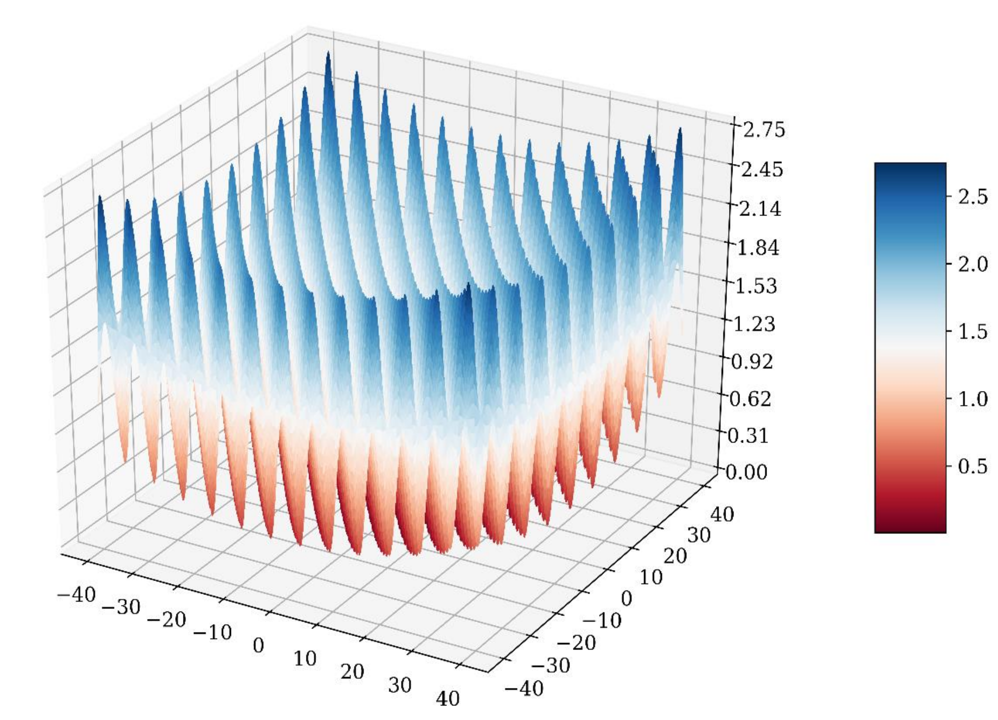
Ill-Conditioned



The condition number of a function measures how much the output value of the function can change for a small change in the input argument.

Nevergrad Ill-Conditioned Rastrigin

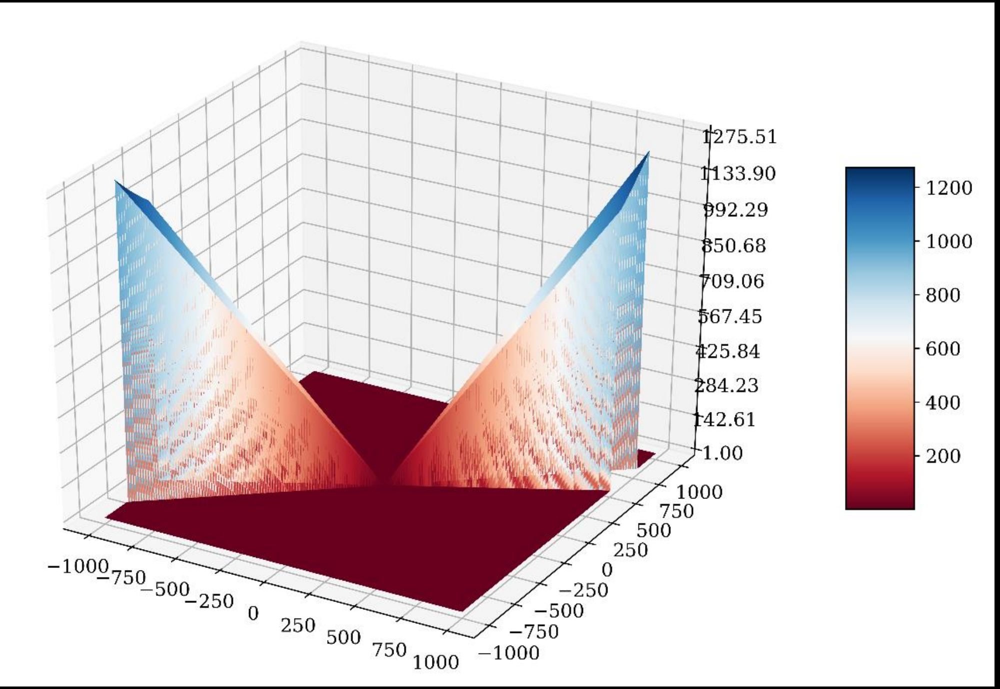
Multimodal



Infinitely many local optima as we get closer to the global optimum

[Griewank Multimodal]

Path Function



Function needs to follow a long, narrow path in order to reach optimum

Details

Criteria of Convergence for Noisy Functions

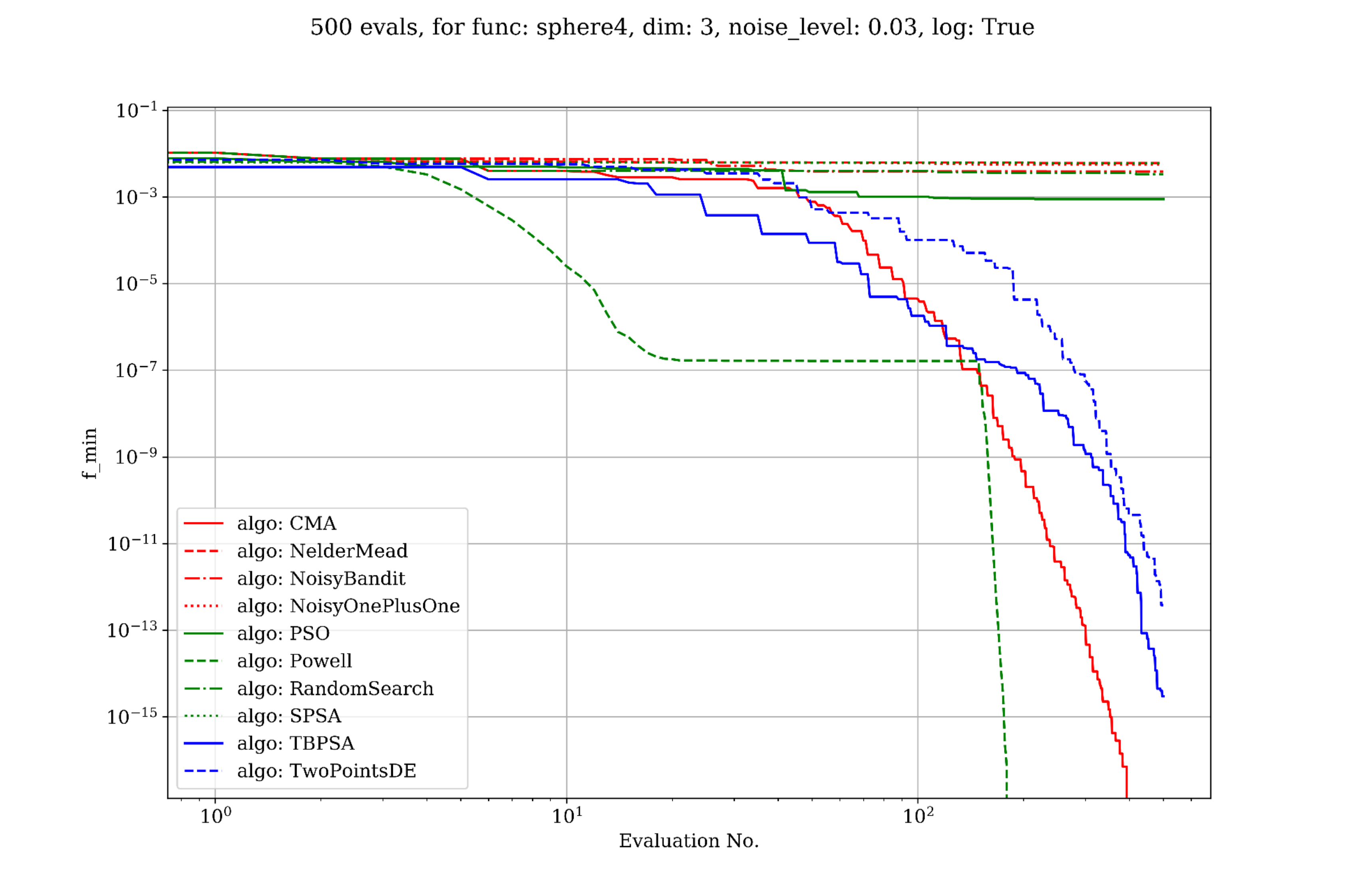
- Slope of Simple Regret
- Slope of Approximate Simple Regret
- Slope of Robust Simple Regret
- Slope of Average Simple Regret

Comparison
Detailed Review

Contents

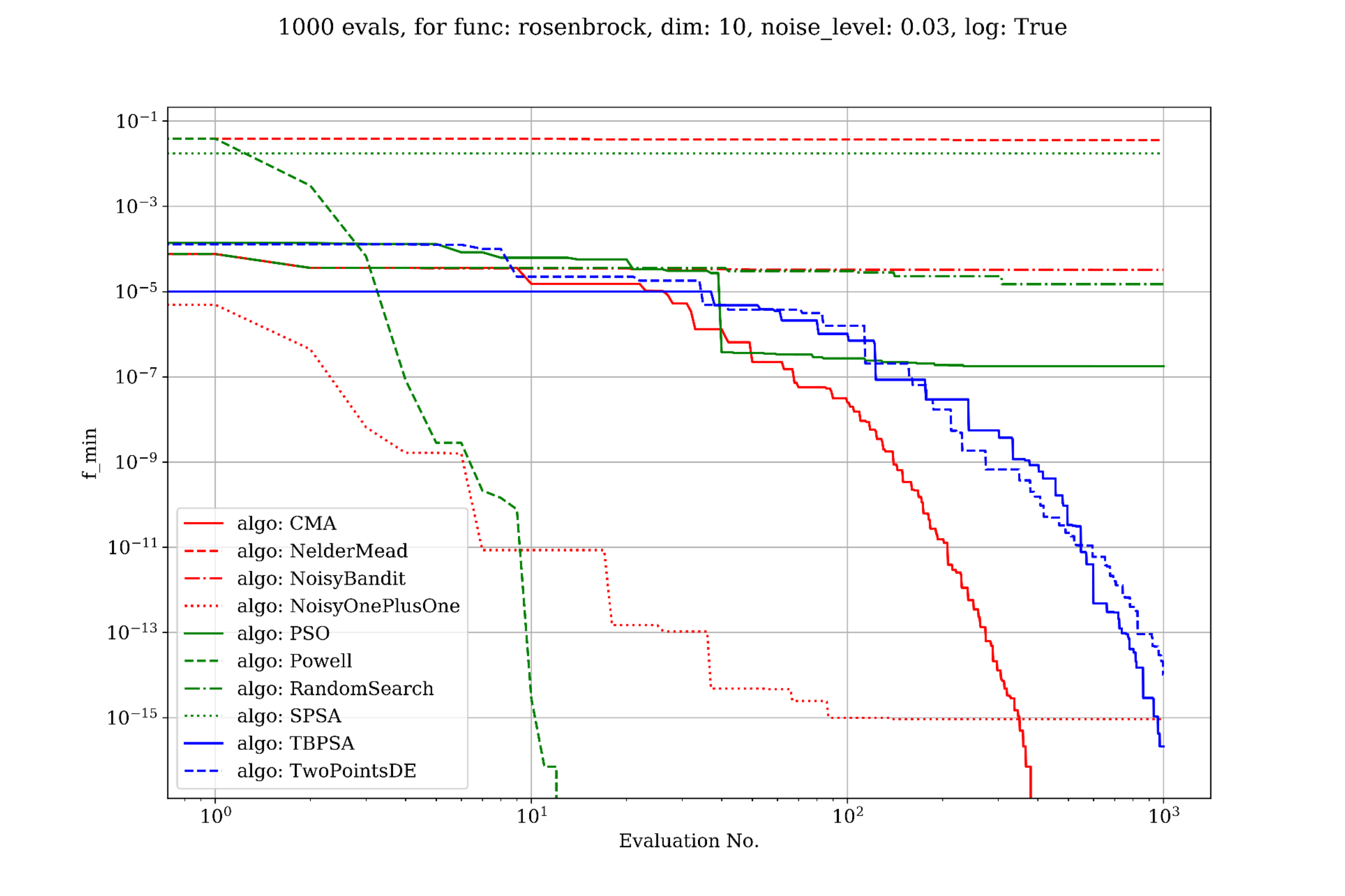
- Algorithms for Global Blackbox Noisy Optimisation
- Synthetic Functions for testing Algorithms
- **Initial Benchmarking Results**
- Preliminary Suggestions for Algorithm
- Preview of Algorithm Benchmarking Suite

Easy



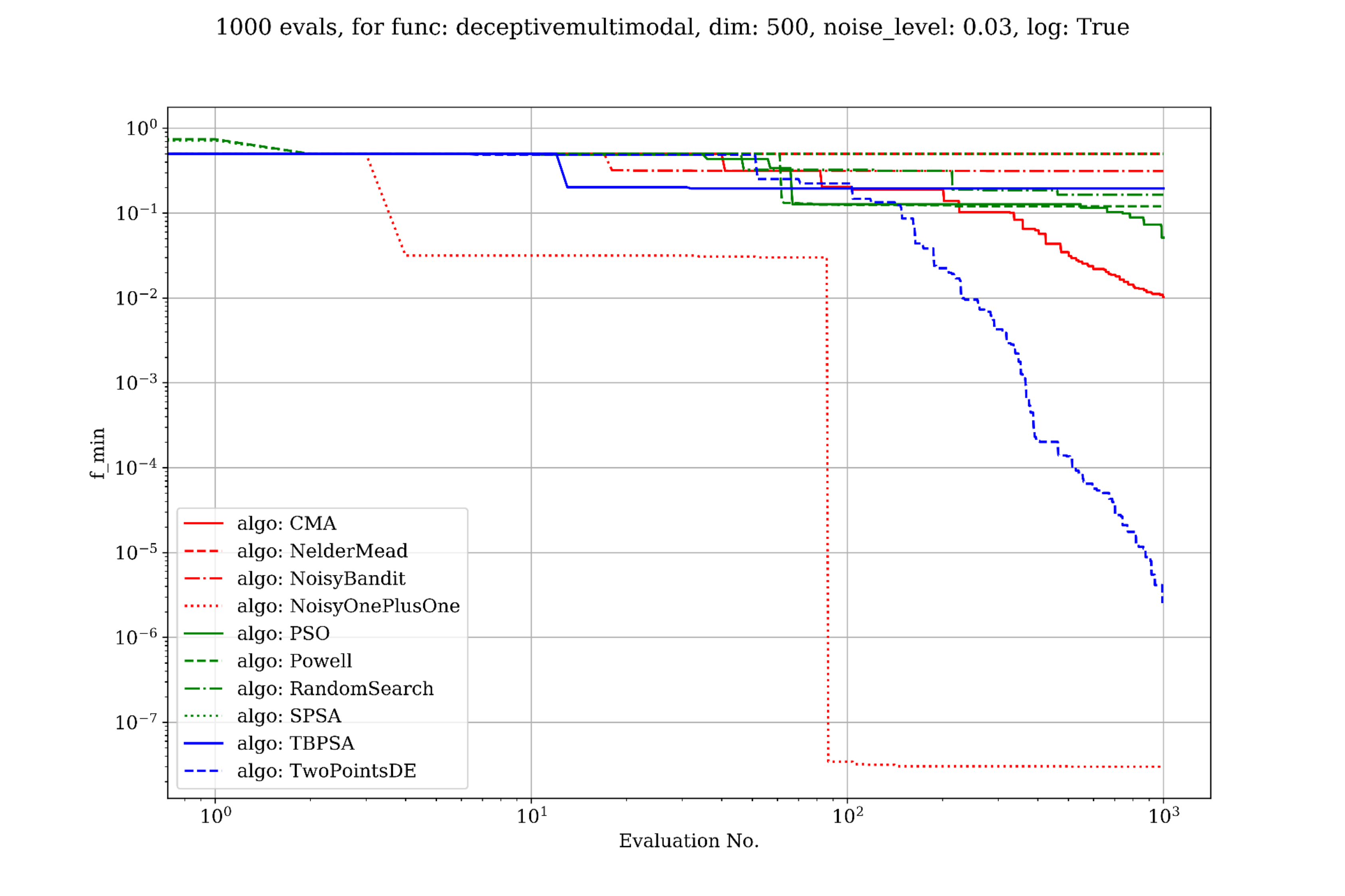
- Convex (translated sphere)
- noise 0.03
- dimension 3
- X: log-scale evaluations
- Y: log-scale goal

Medium



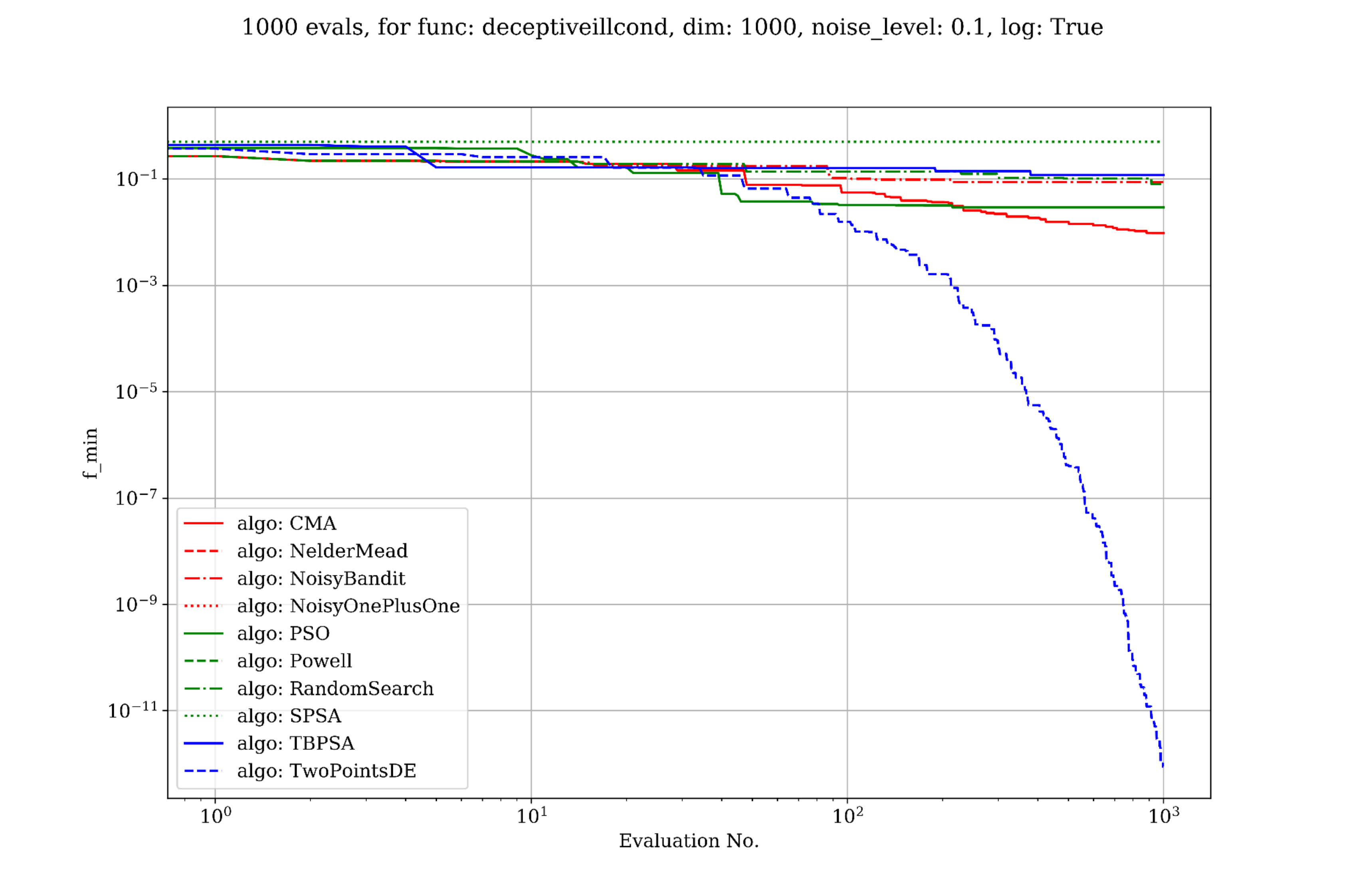
- Rosenbrock
- noise 0.03
- dimension 10
- X: log-scale evaluations
- Y: log-scale goal

Medium-Hard



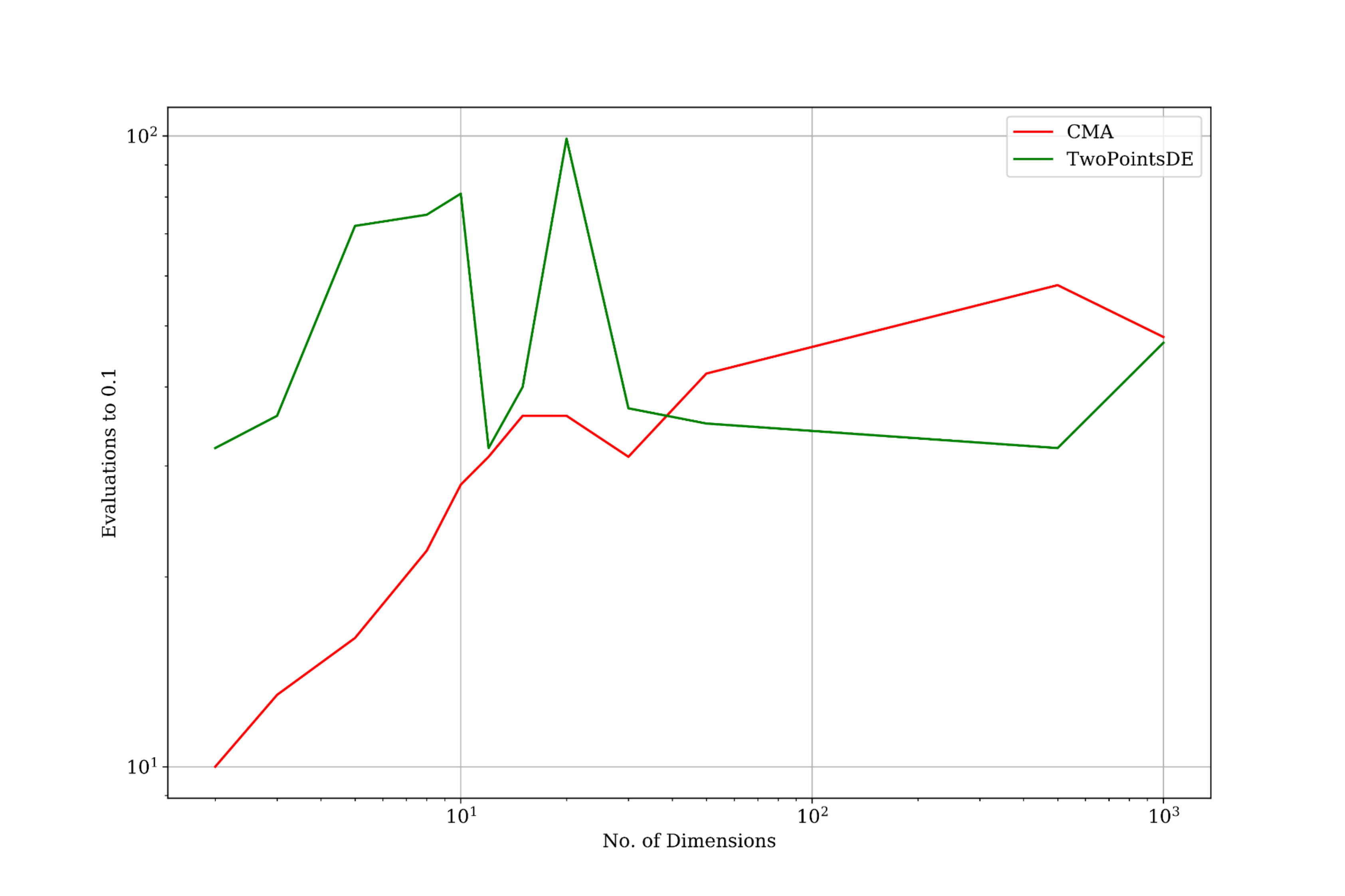
- Multimodal
- noise 0.03
- dimension 500
- X: log-scale evaluations
- Y: log-scale goal

Hard



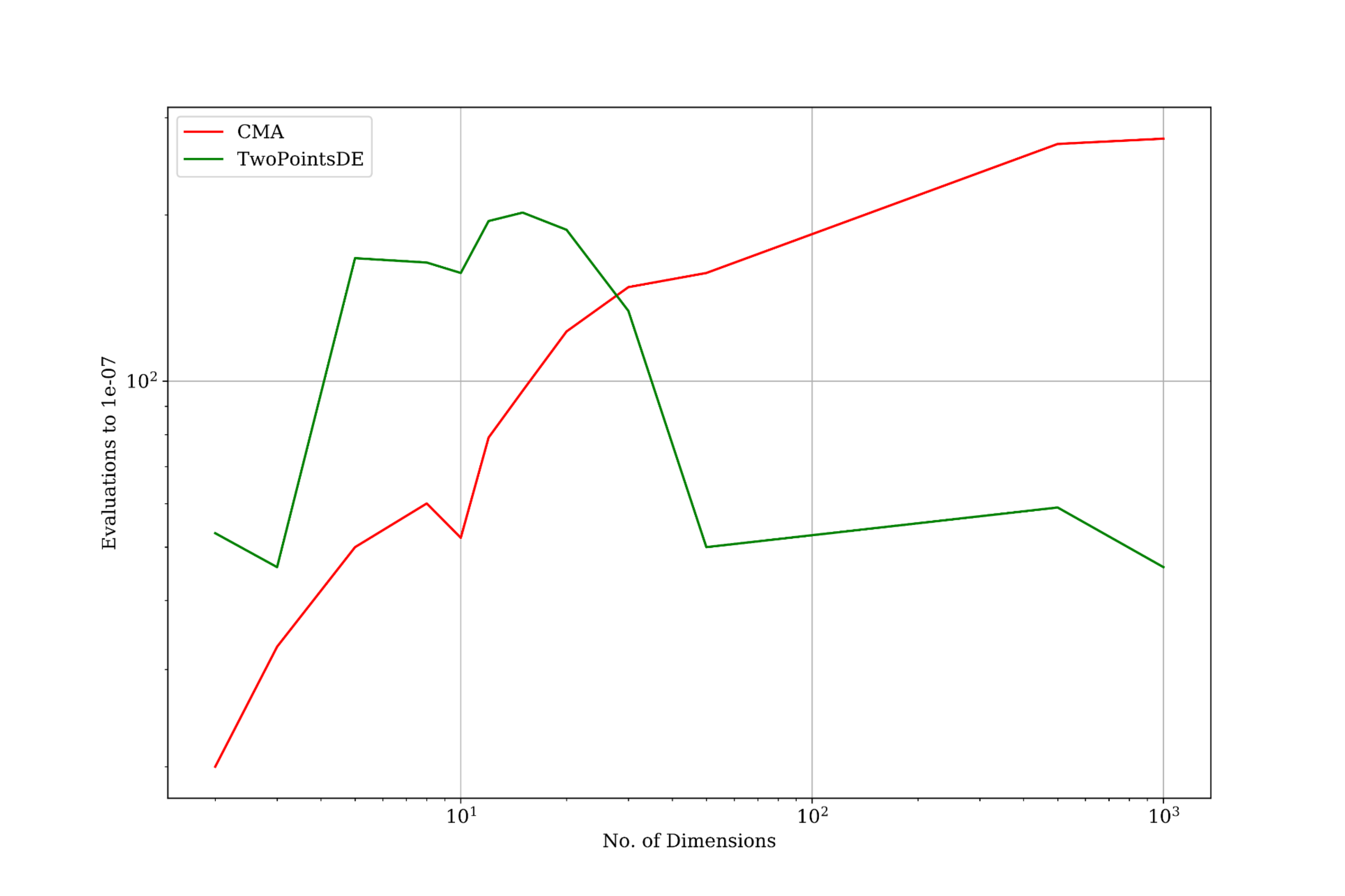
- Ill-Conditioned
- noise 0.1
- dimension 1000
- X: log-scale evaluations
- Y: log-scale goal

No of Evaluations to a Goal



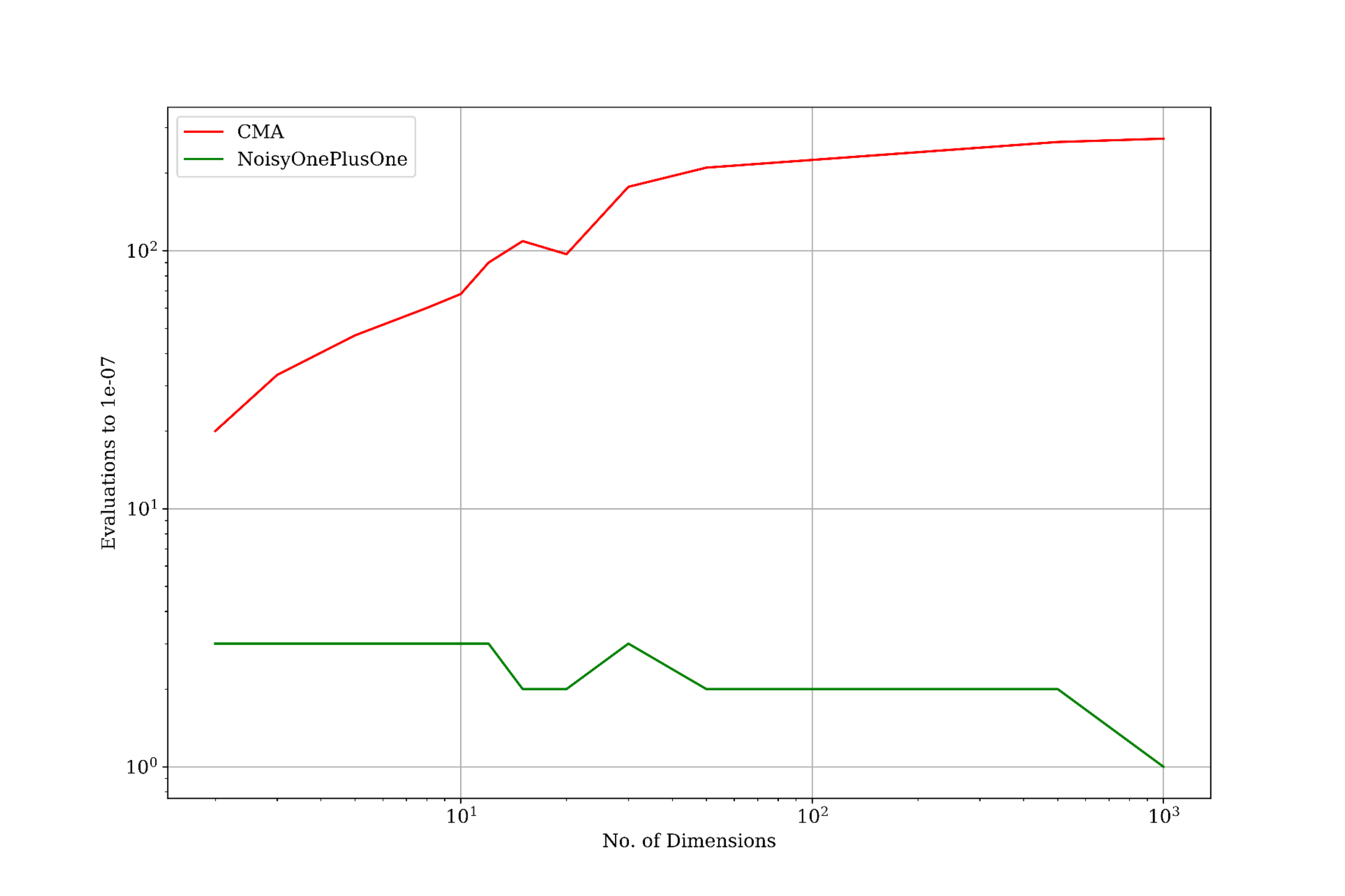
- Ill-Conditioned
- noise 0.03
- Goal 0.1

No of Evaluations to a Goal



- Rosenbrock
- noise 0.1
- Goal 1e-7

No of Evaluations to a Goal



- Rosenbrock
- noise 0.03
- Goal 1e-7

Contents

Preliminary Suggestions

- CMA-ES (Low & Medium Dimensions)
- Powell (Only Low Dimensions)
- 1+1 Evolutionary Algorithm (Decent for most cases)
- Particle Swarm/Differential Evolution (High-Dimensional & Ill-Conditioned)
- Bayesian Optimisation (High accuracy for constrained cases, not shown)
- Portfolio of Algorithms

Contents

- Algorithms for Global Blackbox Noisy Optimisation
- Synthetic Functions for testing Algorithms
- Initial Benchmarking Results
- Preliminary Suggestions for Algorithm
- **Preview of Algorithm Benchmarking Suite**