Non-gradient-based (Gradient Free) Function Optimization

Goal: Review non-gradient approaches.

Paper: Derivative-free optimization: A review of algorithms and comparison of software implementations, Book: Introduction to Derivative-Free Optimization

Maintain one or a small number of best candidates:

Hill climbing (including local search, local unimodal sampling, pattern search, random search, random optimization, simulated annealing)

Nelder Mead/Simplex/Amoeba method (Matlab fminsearch)

Iteratively fit deterministic surrogate functions to a number of samples and optimize the surrogate function

Response Surface Methodology (RSM)

Iteratively fit probability distributions to a number of samples and use the distributions to choose next samples

evolutionary algorithms, genetic algorithms, and ...

Paper: Derivative-free optimization: A review of algorithms and comparison of software implementations by Luis Miguel Rios and Nikolaos V. Sahinidis, Book: Introduction to Derivative-Free Optimization

Iteratively fit probability distributions to a number of samples and use the distributions to choose next samples

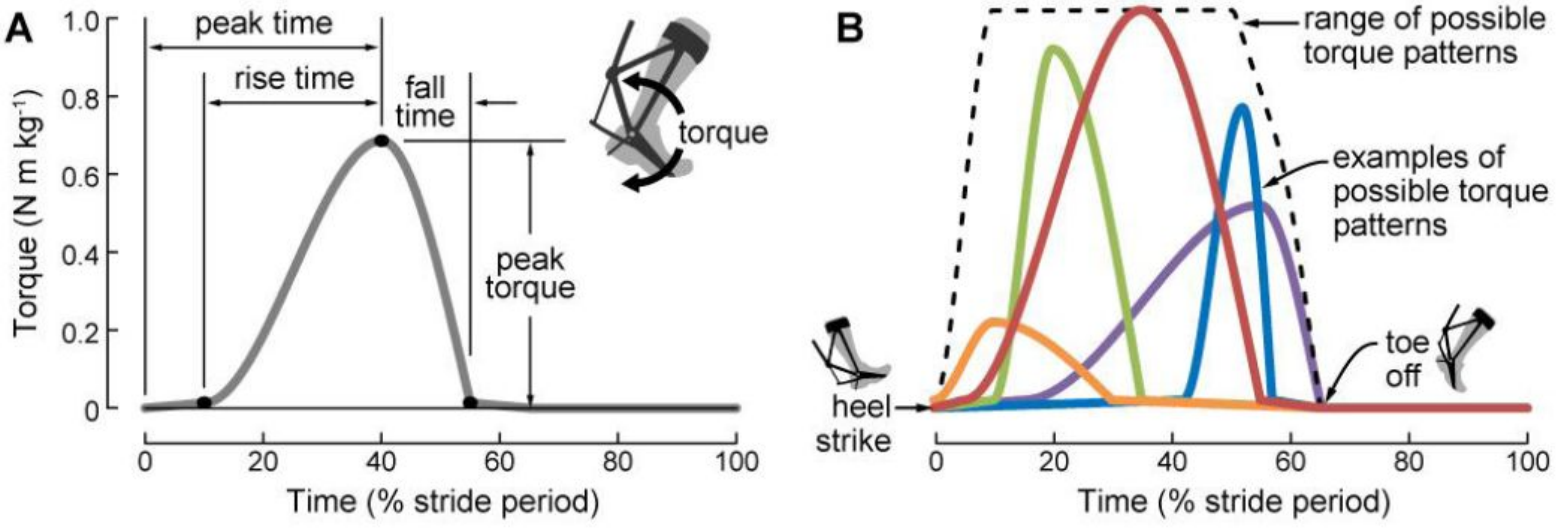
Covariance Matrix Adaptation Evolution Strategy.

Goal: Understand currently popular state of the art method.

See also Hansen web page. Example1, Ex2, Ex3, Ex4.

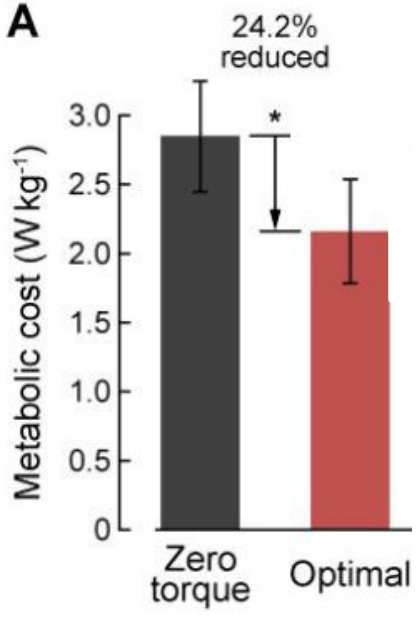
Example 1:
Assistive Exoskeleton
HILO: Human In the Loop Optimization (= robot learning)

Choosing parameters to optimize



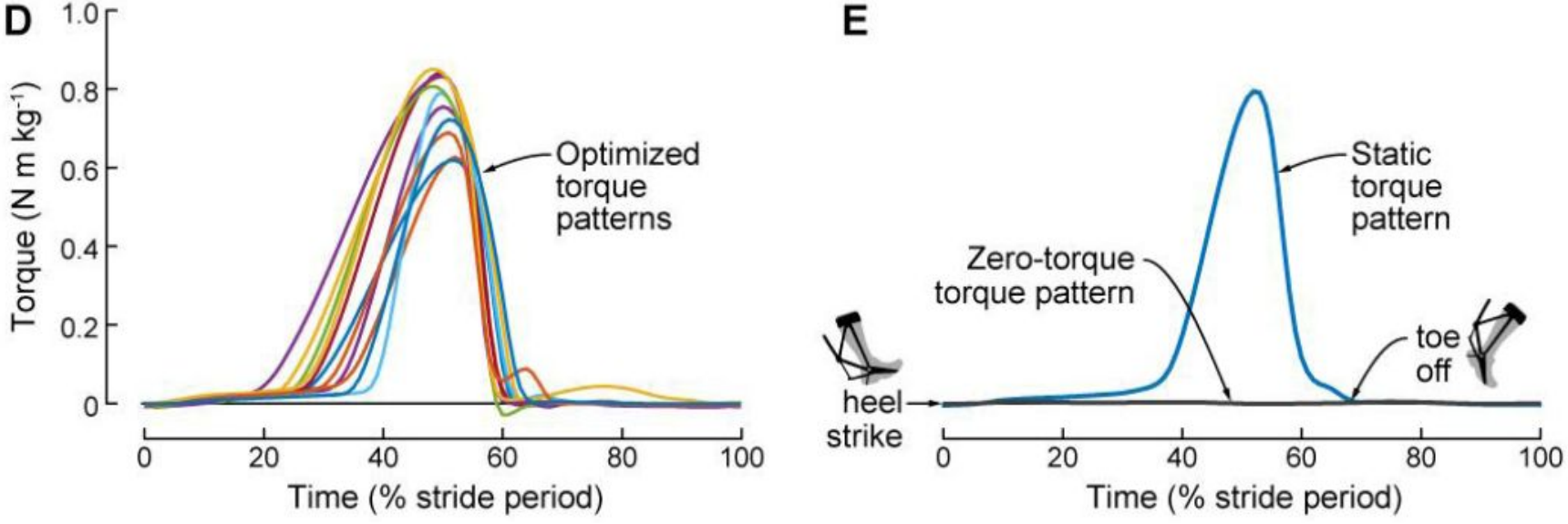
Zhang, Fiers, Witte, Jackson, Poggensee, Atkeson, and Collins. *Science* 2017

Online non-gradient-based optimization (CMA-ES) outperforms manual tuning (24% vs 7% improvement)



Zhang, Fiers, Witte, Jackson, Poggensee, Atkeson, and Collins. *Science* 2017

Optimized settings highly variable among subjects



Zhang, Fiers, Witte, Jackson, Poggensee, Atkeson, and Collins. *Science* 2017

Parallelizing Optimization

Limited scope to parallelize internals of gradient-based algorithms. Gradient-based algorithms are inherently serial.

Could optimize from many starting points in parallel (for any optimization approach).

Can evaluate optimization criterion for many points simultaneously so population-based optimization algorithms straightforwardly parallelize.