An Empirical Analysis of Entropy Search in Batch Bayesian Optimisation

Author: Petre-Alexandru Hautelman
(p.a.hautelman-1@student.tudelft.nl)

Responsible Professor: Matthijs Spaan
Supervisor: Joery de Vries

Introduction 1

Batch Bayesian Optimisation with Classical Methods:
- Probability of Improvement, Expected Improvement, Upper Confidence Bound
- can be applied to batch using Markov Chain Monte Carlo methods
- selection set is populated sequentially

$$\alpha_{PI-MCMC}(x|\{x_{q'}\}_{q'=1}^{q}) = \int_{\mathcal{X}^q} [\alpha_{PI}(x|D_n \cup \{x_{q'}, y_{q'}\}_{q'=1}^{q}) (\{y_{q'}\}_{q'=1}^{q}|D_n, \{x_{q'}\}_{q'=1}^{q}) dy_1 .. dy_q]$$

- double greedy, myopic selection

Entropy Search methods:
- non-greedy by design
- considers information-theoretic concepts: information gain, information entropy

$$\alpha_{ES}(x) = H[p(x^*|D_n)] - E_{p(y|D_n,x)}[H(p(x^*|D_n \cup \{x,y\}))] \quad \textbf{(1)}$$

- easily applied to batch

$$\alpha(\{x_q\}_{q=1}^{Q}) = \sum_{q=1}^{Q} \alpha(x_q|\{x'_q\}_{q'=1}^{q})$$

Predictive Entropy Search (PES) [1]:
- derived from equation **(1)**

$$\alpha_{PES}(x) = H[p(y|D_n,x)] - \mathbb{E}_{p(x^*|D_n)}[H(p(y|D_n,x,x^*))]$$

- compute the change in the entropy of the predictive distribution at the optimum's position
- analytically tractable

Max-Value Entropy Search (MES) [2]:
- considers entropy over maximum function value

$$\alpha_{MES}(x) = H[p(y|D_n,x)] - \mathbb{E}_{p(y^*|D_n)}[H(p(y|x,D_n,y^*))]$$

- marginalized over sampled maximum values
- robust and efficient implementation

Joint Entropy Search (JES) [3]:
- combines ideas from PES and JES

$$\alpha_{JES}(x) = H[p(y|D_n,x)] - \mathbb{E}_{p(x^*,y^*|D_n)}[H(p(y|D_n \cup (x^*,y^*), x, y^*))]$$

Literature gap:
- systematic study on the performance of PES, MES, JES in parallel optimisation across various environment settings
- batch sizes, function dimensions, noise levels, types of objective functions

Research question 2

How do Entropy Search algorithms perform under various environment specifications, and what are the factors influencing their performance?

3 Methodology

Batch sizes (q): {2, 5, 10, 25}

Noise levels: {0%, 5%, 10%, 20%, 40%}
- percentage of range of values of objective function

Input dimensions (D): {2, 5, 10, 25, 50}

Function shapes:
- unimodal: Easom, Zakharov, Sum of Different Powers
- multimodal: Griewank, Ackley, Schwefel

Fractional Factorial Design:
- used to assemble the testing configurations
- exposes information about the most important features of the studied problem
- 26 testing environments

Metrics:
- simple regret
- cumulative regret

Algorithms:
- **qPES, qMES, qJES** as implemented in BOTorch
- random and qEI
- results averaged over 5 runs

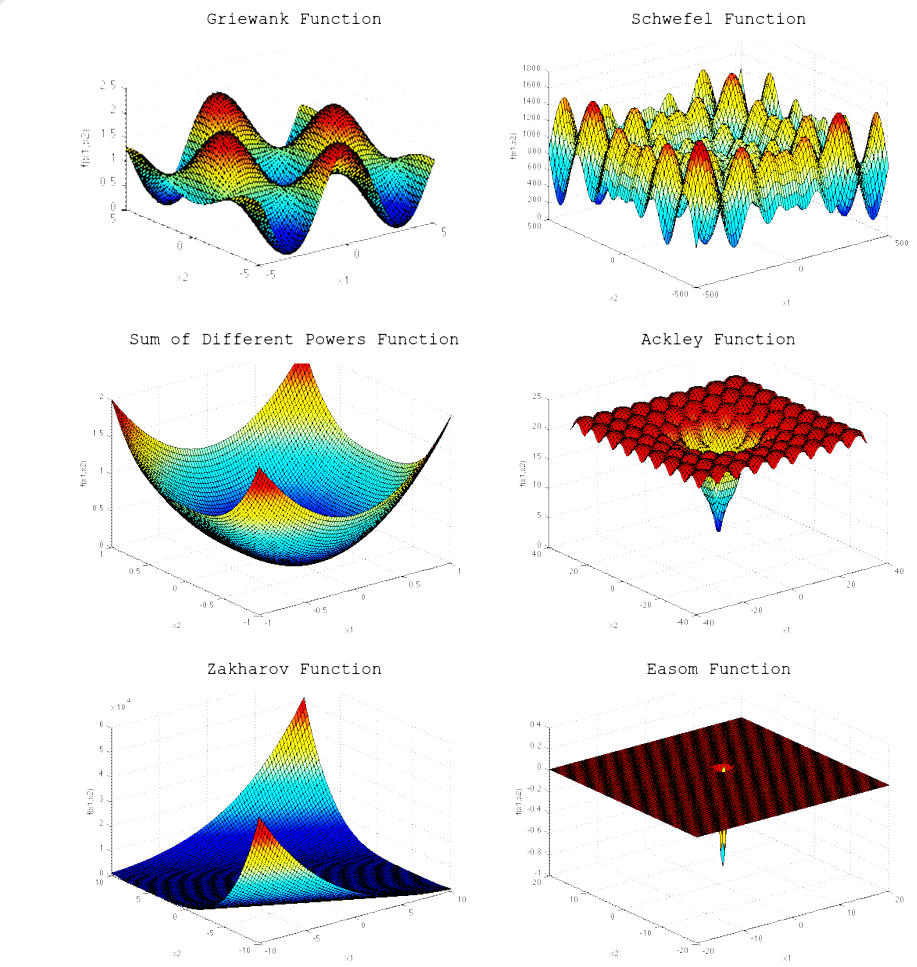

Figure 1: 2-dimensional representation of the functions used for evaluation

4 Results

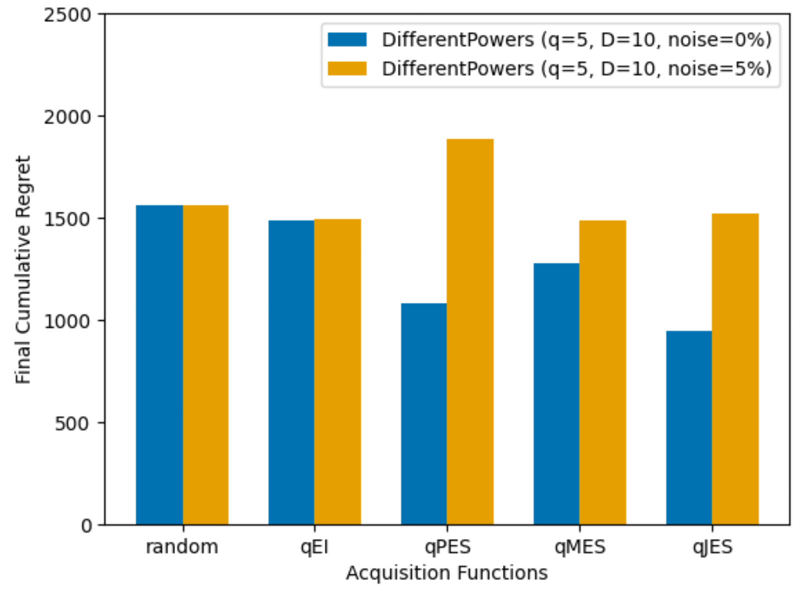

Figure 2: Performance of acquisition functions in high-dimensional unimodal function after 50 iterations

ES has better performance for batch BO

Noise is the most influential parameter
- especially in high-dimensional spaces

Increased efficiency for unimodal functions with discernable slope
- optimisation failed for Easom, which has a small optimum area surrounded by a flat outer region

qJES is most robust to greater input dimensions
- qPES also looks promising pending further testing

Function	Average runtime (s)
qEI	1.14
qPES	13.62
qMES	3.13
qJES	24.95

Table 1: Average runtime of studied algorithms. Note that qPES was run with GPU acceleration

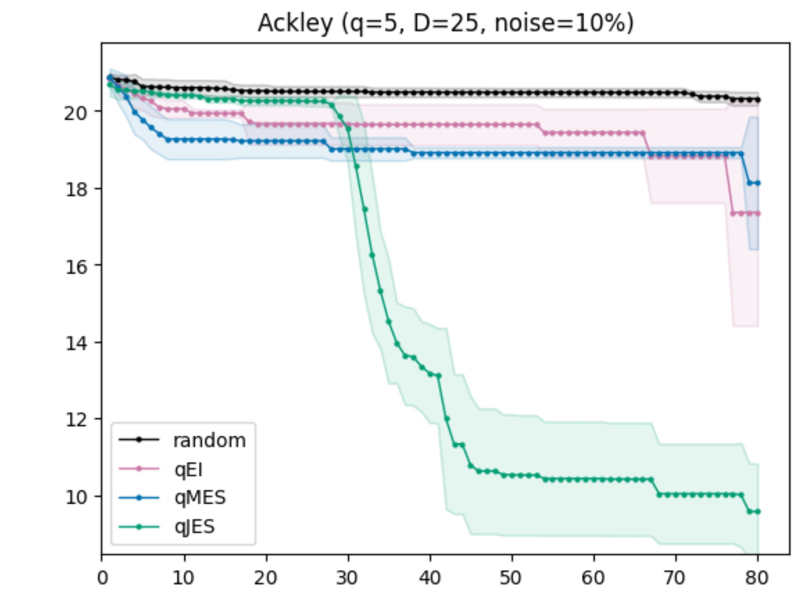

Figure 3: Ackley run

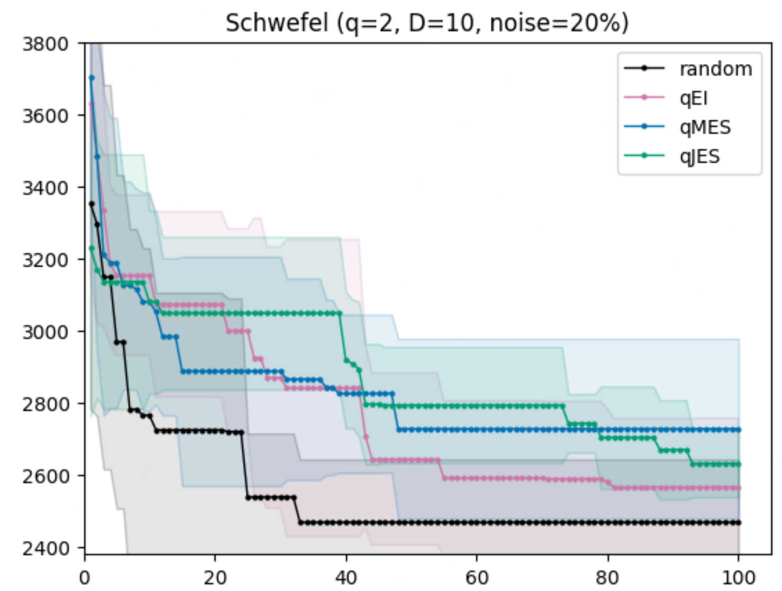

Figure 3: Schwefel run

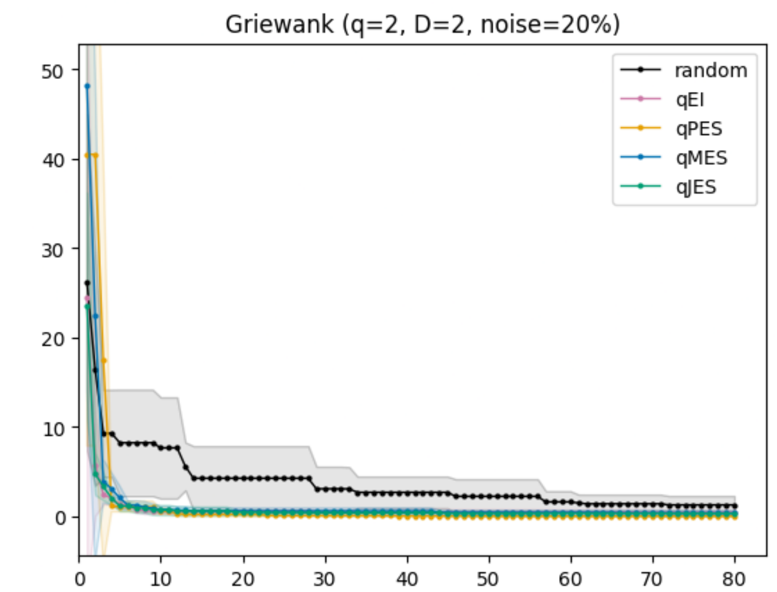

Figure 4: Griewank run

5 Future work

Explore more complex scenarios:
- input-dependant noise
- misspecified environments
- multi-objective optimisation

Different metrics:
- inference regret

Real-World Benchmarking

References:

[1] Jose Herńandez, Matthew Hoffman, and Zoubin Ghahramani. Predictive entropy search for efficient global optimization of black-box functions.
[2] Zi Wang and Stefanie Jegelka. Max-value entropy search for efficient bayesian optimization
[3] Carl Hvarfner, Frank Hutter, and Luigi Nardi. Joint entropy search for maximally-informed bayesian optimization.