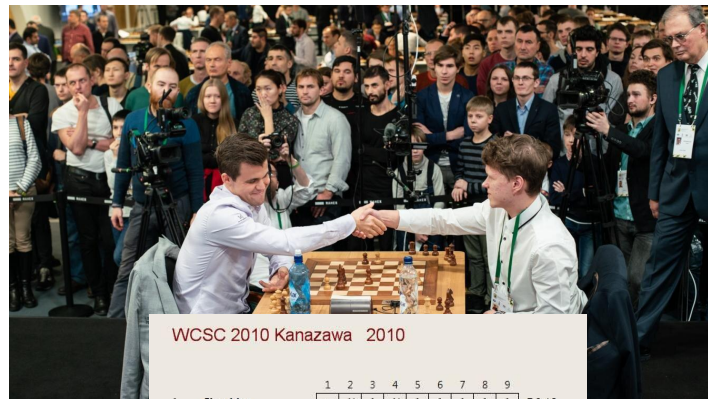


Active Sampling for Pairwise Comparisons via Approximate Message Passing and Information Gain Maximization

Aliaksei Mikhailiuk, Clifford Wilmot, Maria Perez-Ortiz, Dingcheng Yue, Rafal K. Mantiuk

Problem overview - Inferring Scale



WCSC 2010 Kanazawa 2010

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | |
|-----------------------|---|---|---|---|---|---|---|---|---|---------|
| 1 Shredder | * | ½ | 1 | ½ | 1 | 1 | 1 | 1 | 1 | 7.0 / 8 |
| 2 Rondo | ½ | * | ½ | 1 | ½ | 1 | 1 | 1 | 1 | 6.5 / 8 |
| 3 Thinker | 0 | ½ | * | ½ | 1 | ½ | 1 | 1 | 1 | 5.5 / 8 |
| 4 Pandix Breakthrough | ½ | 0 | ½ | * | ½ | | 1 | 1 | 1 | 4.5 / 7 |
| 5 Deep Junior | 0 | ½ | 0 | ½ | * | | 1 | 1 | 1 | 4.0 / 7 |
| 6 Jonny | 0 | 0 | ½ | | | * | | 1 | 1 | 2.5 / 5 |
| 7 Darmanios | 0 | 0 | 0 | 0 | 0 | | * | 1 | 1 | 2.0 / 7 |
| 8 Fridolin | 0 | 0 | 0 | 0 | 0 | 0 | | * | 1 | 1.0 / 8 |
| 9 Hector For Chess | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | * | 0.0 / 8 |

(33 Games)

Problem overview



Now
University of Cam... Change

- ☒ Delivery
- ☐ Pickup

Sort

Hygiene rating

- ☐ Hygiene rating: 5 (69)
- ☐ Hygiene rating: 4 (23)

Up to 25% off View all (29)



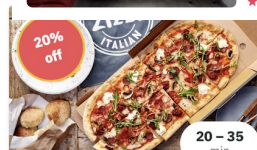
The Athenian

★ 4.4 Very Good (50+) · Greek · Mediterranean · Wraps · Dessert
2.4 miles away · £3.99 delivery
💖 Spend £10, get 25% off



The Athenian Plant Based

★ 4.5 Excellent (50+) · Greek · Wraps · Healthy · Mediterranean
2.4 miles away · £3.99 delivery
💖 Spend £10, get 25% off



Zizzi

★ 4.4 Very Good (50+) · Italian · Pizza
1.2 miles away · £3.49 delivery
💖 20% off entire menu



Chinese Canteen

★ 4.6 Excellent (50+) · Chinese · Noodles · Asian · Dumplings
1.1 miles away · £2.99 delivery
💖 Service Guarantee



The Olive Gro

★ 4.6 Excellent (50+) · Pizza · Medite
1.7 miles away
💖 Spend £15



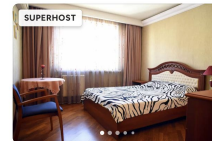
Minsk 7 Dec - 13 Dec

143 stays · 7 Dec - 13 Dec · 1 guest

Stays in Minsk

Cancellation flexibility Type of place Price Instant Book More filters

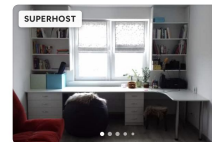
Review COVID-19 travel restrictions before you book. [Learn more](#)



Private room in Minsk
(Л) Комната, комфортная, центр города
2 guests · 1 bedroom · 1 bed · 2 shared bathrooms
Wifi · Kitchen · Washing machine

★ 4.96 (53)

£10 / night
£58 total



Private room in Minsk
Уютная комната недалеко от м Уручье с видом н...
2 guests · 1 bedroom · 1 bed · 1 shared bathroom
Wifi · Kitchen · Washing machine · Self check-in

★ 4.95 (92)

£9 / night
£33 total

Motivation

- Pairwise comparisons are often used in crowdsourcing.

Motivation

- Pairwise comparisons are often used in crowdsourcing.
- Drawback - for n conditions there are $n(n-1)/2$ pairs.

Motivation

- Pairwise comparisons are often used in crowdsourcing.
- Drawback - for n conditions there are $n(n-1)/2$ pairs.
- How can we minimize the number of comparisons?

Motivation

- Pairwise comparisons are often used in crowdsourcing.
- Drawback - for n conditions there are $n(n-1)/2$ pairs.
- How can we minimize the number of comparisons?
- Two types of methods:

Motivation

- Pairwise comparisons are often used in crowdsourcing.
- Drawback - for n conditions there are $n(n-1)/2$ pairs.
- How can we minimize the number of comparisons?
- Two types of methods:
 - Sorting based (Swiss chess tournament system, Adaptive Rectangular Design (ARD))

Motivation

- Pairwise comparisons are often used in crowdsourcing.
- Drawback - for n conditions there are $n(n-1)/2$ pairs.
- How can we minimize the number of comparisons?
- Two types of methods:
 - Sorting based (Swiss chess tournament system, Adaptive Rectangular Design (ARD))
 - Information gain based - assume data model (Hybrid-MST, Crowd-BT, HR-Active)

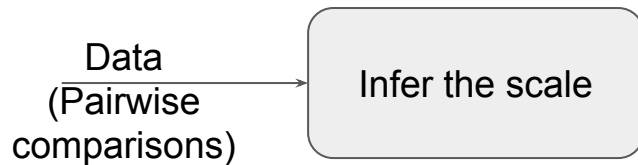
Motivation

- Pairwise comparisons are often used in crowdsourcing.
- Drawback - for n conditions there are $n(n-1)/2$ pairs.
- How can we minimize the number of comparisons?
- Two types of methods:
 - Sorting based (Swiss chess tournament system, Adaptive Rectangular Design (ARD))
 - Information gain based - assume data model (Hybrid-MST, Crowd-BT, HR-Active, **ASAP**)

Algorithm

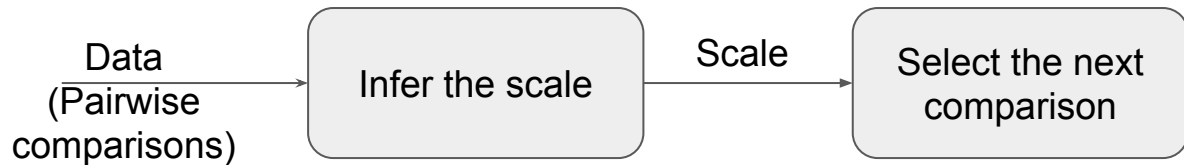
Data
(Pairwise
comparisons)

Algorithm



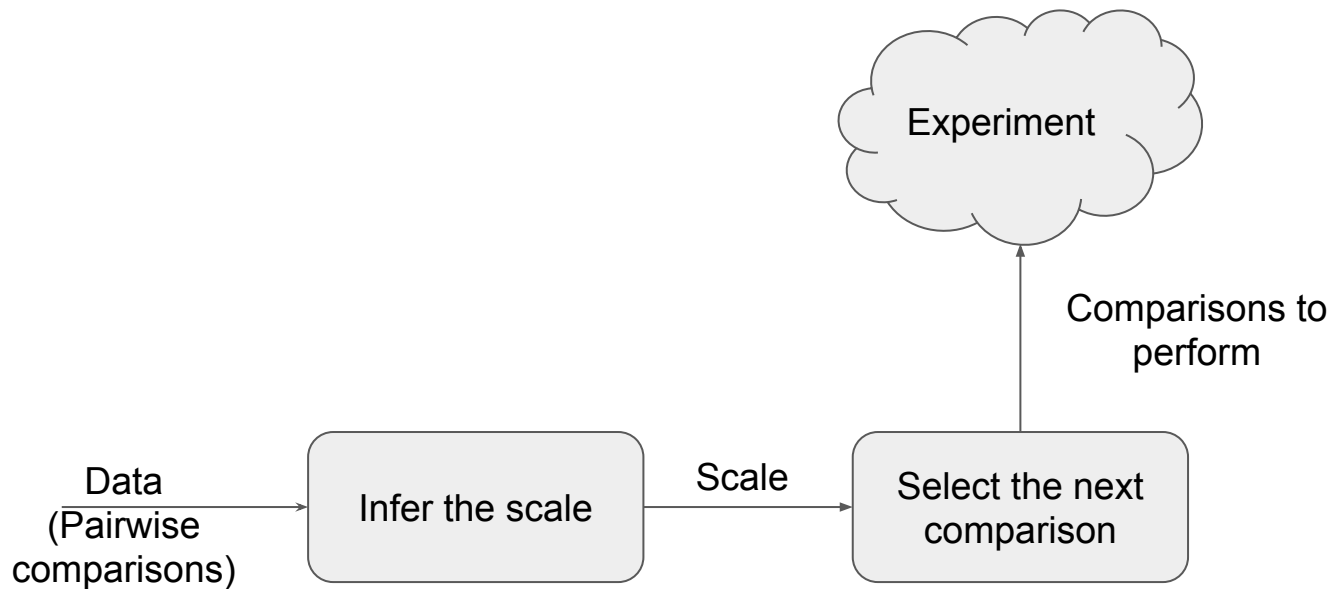
Most methods
approximate,
we do not

Algorithm

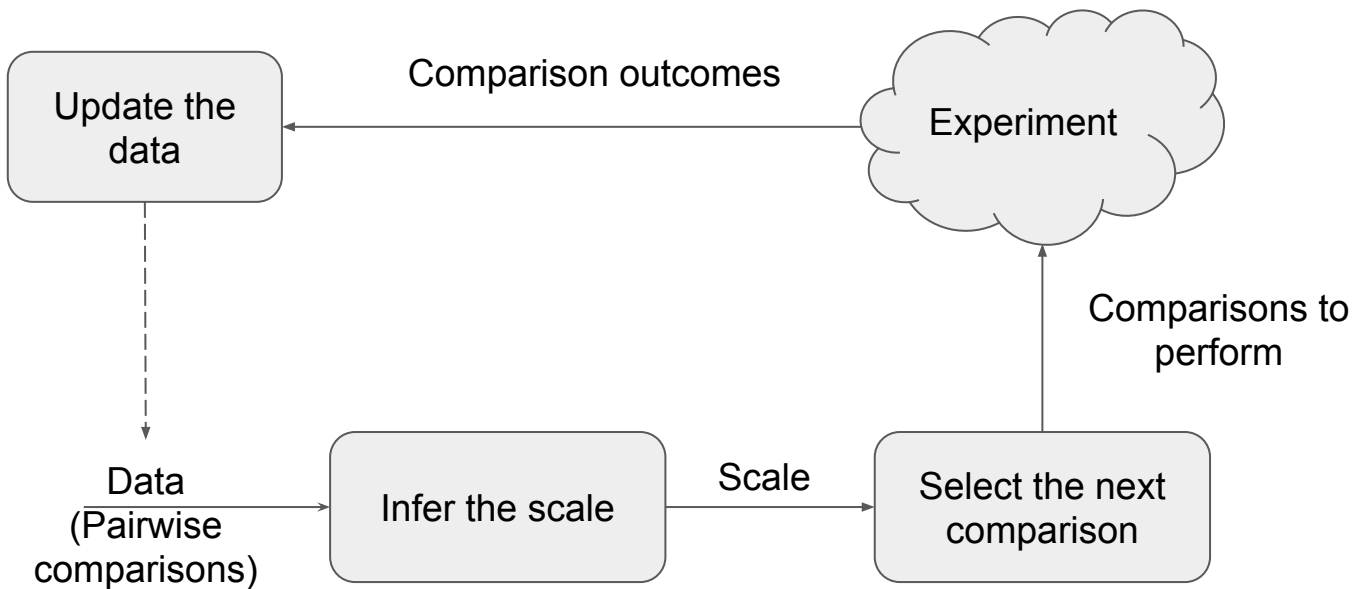


Most methods
do not improve
speed, we do

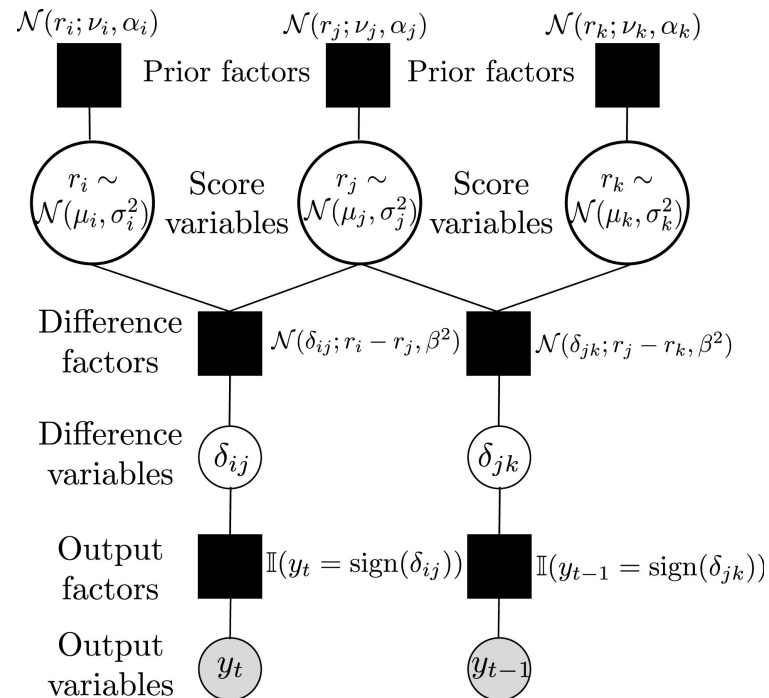
Algorithm



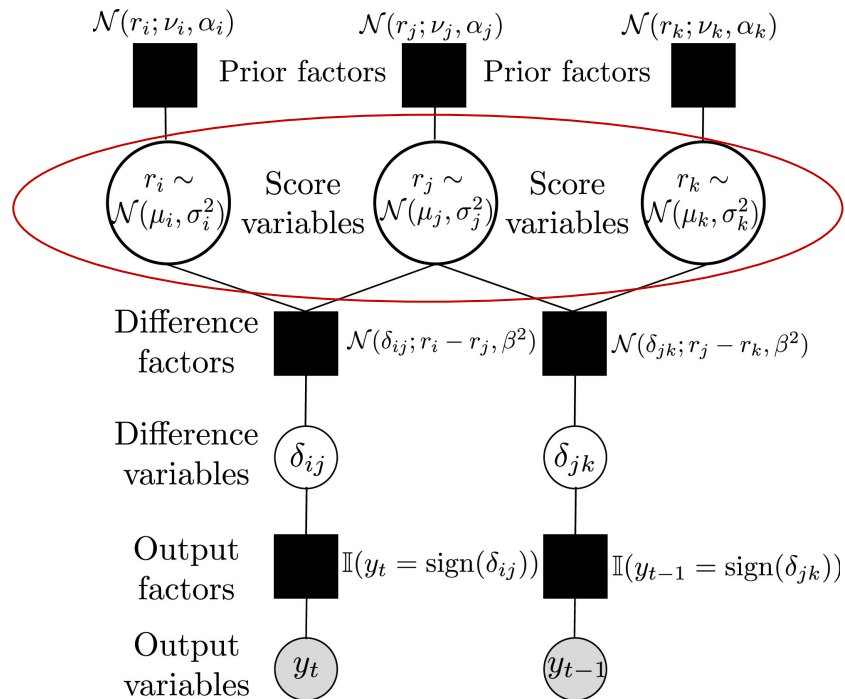
Algorithm



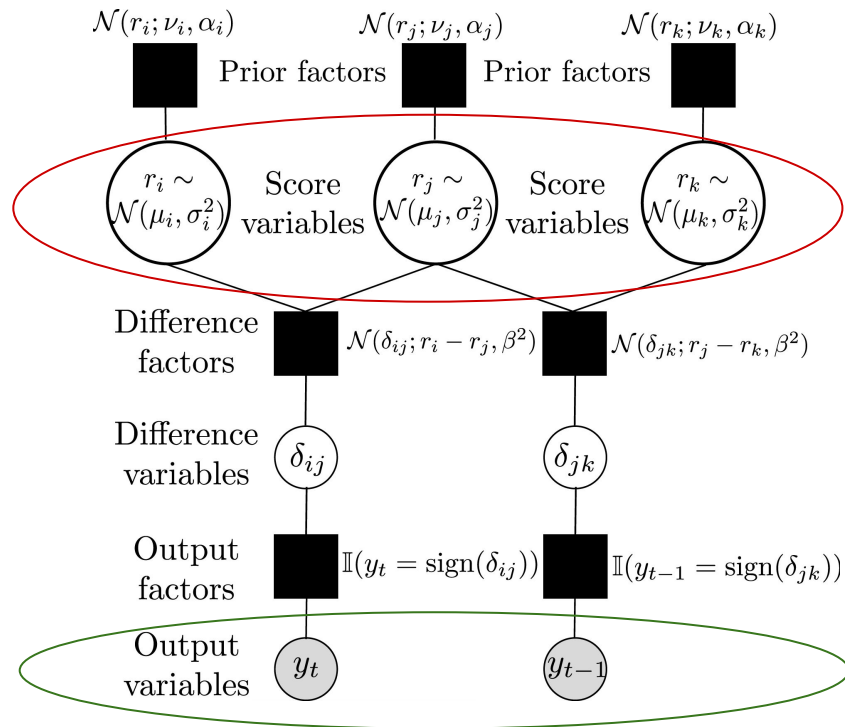
Bayesian Score Inference (TrueSkill based [1])



Bayesian Score Inference (TrueSkill based)

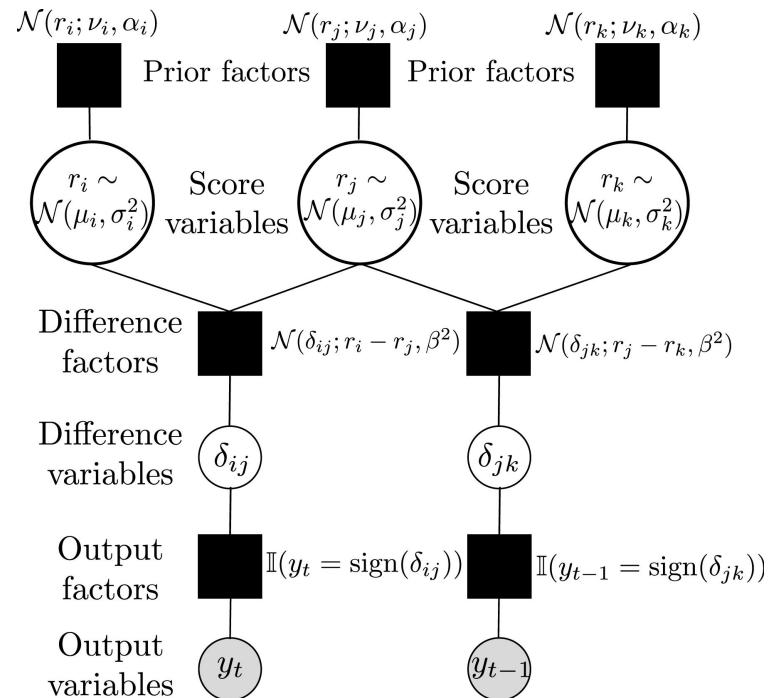


Bayesian Score Inference (TrueSkill based)



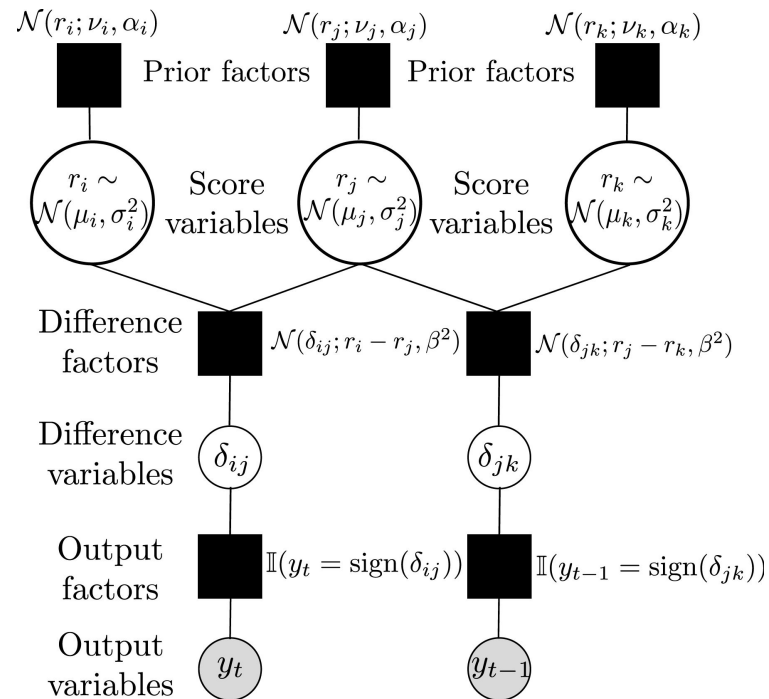
Bayesian Score Inference (TrueSkill based)

- Sum-product algorithm using expectation propagation via moment matching



Bayesian Score Inference (TrueSkill based)

- Sum-product algorithm using expectation propagation via moment matching
- $k \cdot O(n+t)$ complexity of inferring the scores (k - number of iterations)



Expected Information Gain (EIG)

Expected Information Gain (EIG)

Collected data so far

| | | | |
|---|---|---|---|
| 0 | 1 | 2 | 3 |
| 3 | 0 | 3 | 1 |
| 4 | 1 | 0 | 3 |
| 0 | 2 | 1 | 0 |

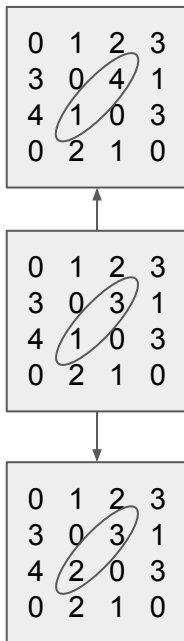
Expected Information Gain (EIG)

Collected data so far

| | | | |
|---|---|---|---|
| 0 | 1 | 2 | 3 |
| 3 | 0 | 3 | 1 |
| 4 | 1 | 0 | 3 |
| 0 | 2 | 1 | 0 |

Expected Information Gain (EIG)

Assume 2 selected over 3



| | | | |
|---|---|---|---|
| 0 | 1 | 2 | 3 |
| 3 | 0 | 4 | 1 |
| 4 | 1 | 0 | 3 |
| 0 | 2 | 1 | 0 |

Collected data so far

| | | | |
|---|---|---|---|
| 0 | 1 | 2 | 3 |
| 3 | 0 | 3 | 1 |
| 4 | 1 | 0 | 3 |
| 0 | 2 | 1 | 0 |

| | | | |
|---|---|---|---|
| 0 | 1 | 2 | 3 |
| 3 | 0 | 3 | 1 |
| 4 | 2 | 0 | 3 |
| 0 | 2 | 1 | 0 |

Assume 3 selected over 2

Assume all possible
outcomes

Expected Information Gain (EIG)

Assume 2 selected over 3

| | | | |
|---|---|---|---|
| 0 | 1 | 2 | 3 |
| 3 | 0 | 4 | 1 |
| 4 | 1 | 0 | 3 |
| 0 | 2 | 1 | 0 |

→ Infer scale

Collected data so far

| | | | |
|---|---|---|---|
| 0 | 1 | 2 | 3 |
| 3 | 0 | 3 | 1 |
| 4 | 1 | 0 | 3 |
| 0 | 2 | 1 | 0 |

→ Infer scale

| | | | |
|---|---|---|---|
| 0 | 1 | 2 | 3 |
| 3 | 0 | 3 | 1 |
| 4 | 2 | 0 | 3 |
| 0 | 2 | 1 | 0 |

→ Infer scale

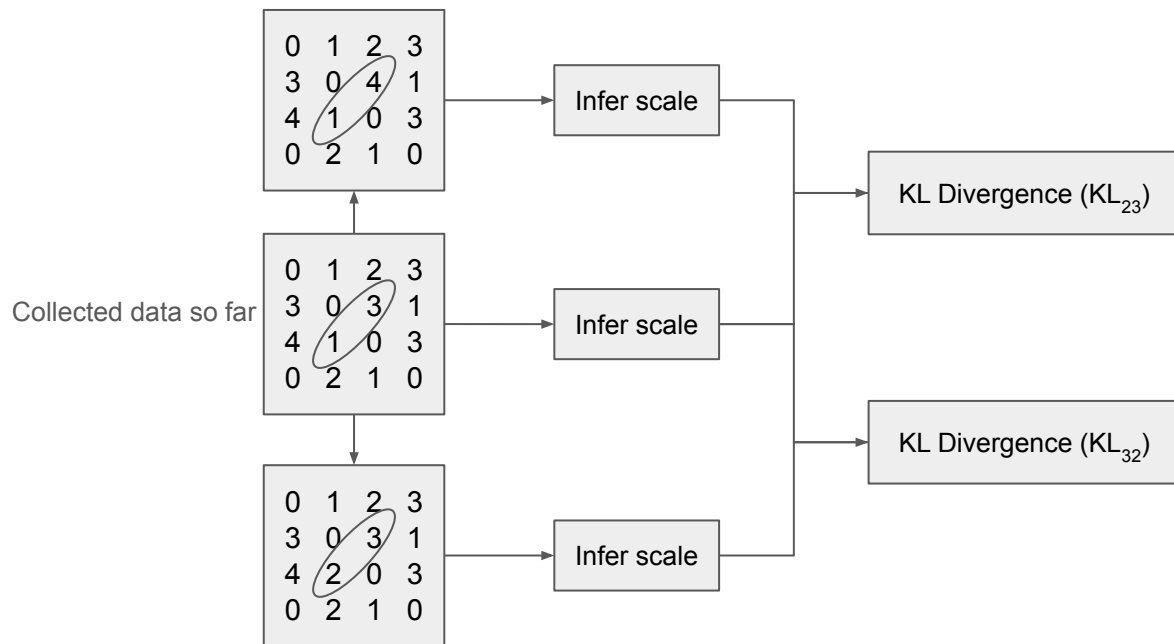
Assume 3 selected over 2

Assume all possible
outcomes

Infer
scores

Expected Information Gain (EIG)

Assume 2 selected over 3



Assume 3 selected over 2

Assume all possible
outcomes

Infer
scores

Calculate change
in the scale

Expected Information Gain (EIG)

Assume 2 selected over 3

| | | | |
|---|---|---|---|
| 0 | 1 | 2 | 3 |
| 3 | 0 | 4 | 1 |
| 4 | 1 | 0 | 3 |
| 0 | 2 | 1 | 0 |

| | | | |
|---|---|---|---|
| 0 | 1 | 2 | 3 |
| 3 | 0 | 3 | 1 |
| 4 | 1 | 0 | 3 |
| 0 | 2 | 1 | 0 |

| | | | |
|---|---|---|---|
| 0 | 1 | 2 | 3 |
| 3 | 0 | 3 | 1 |
| 4 | 2 | 0 | 3 |
| 0 | 2 | 1 | 0 |

Collected data so far

Infer scale

Infer scale

Infer scale

KL Divergence (KL_{23})

KL Divergence (KL_{32})

$$P(o_i \succ o_j | r_i, r_j) \triangleq \Phi \left(\frac{\mu_i - \mu_j}{\sqrt{2}\sigma_{ij}} \right)$$

$$p_{23}KL_{23} + p_{32}KL_{32}$$

EIG

Assume 3 selected over 2

Assume all possible
outcomes

Infer
scores

Calculate change
in the scale

Calculate expected
improvement

Efficiency considerations

- At every iteration t , there are $n(n-1)$ comparison outcomes to consider.

Efficiency considerations

- At every iteration t , there are $n(n-1)$ comparison outcomes to consider.
- We consider two strategies to improve the speed:

Efficiency considerations

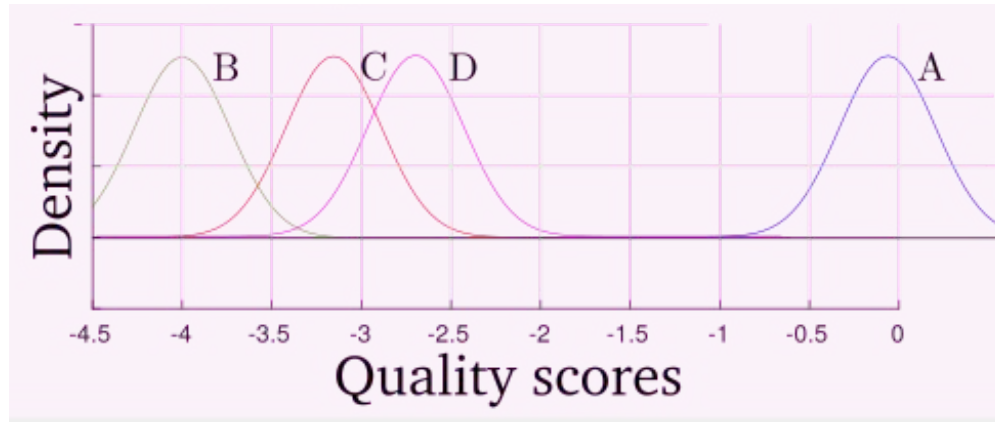
- At every iteration t , there are $n(n-1)$ comparison outcomes to consider.
- We consider two strategies to improve the speed:
 - Selective EIG evaluations

Efficiency considerations

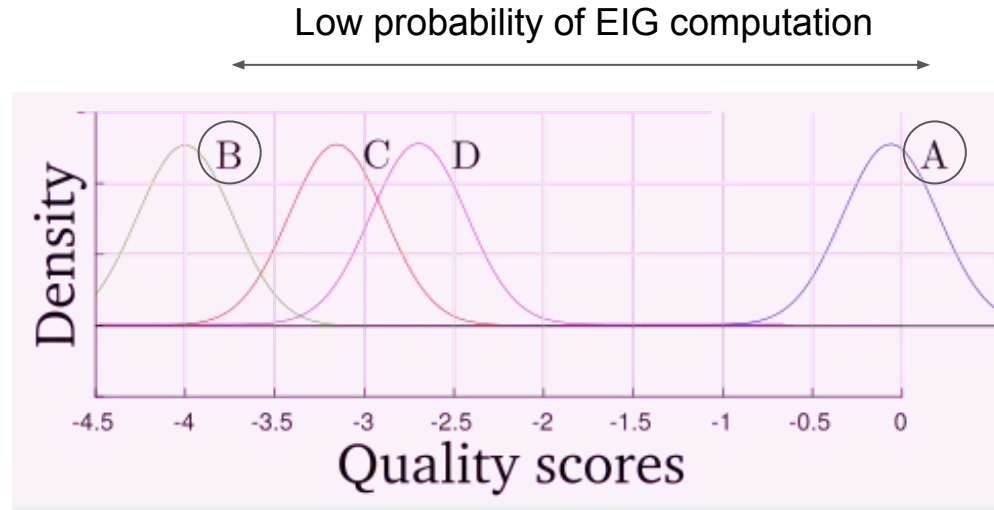
- At every iteration t , there are $n(n-1)$ comparison outcomes to consider.
- We consider two strategies to improve the speed:
 - Selective EIG evaluations
 - Batch mode with Minimum Spanning Tree (MST)

Selective EIG evaluations

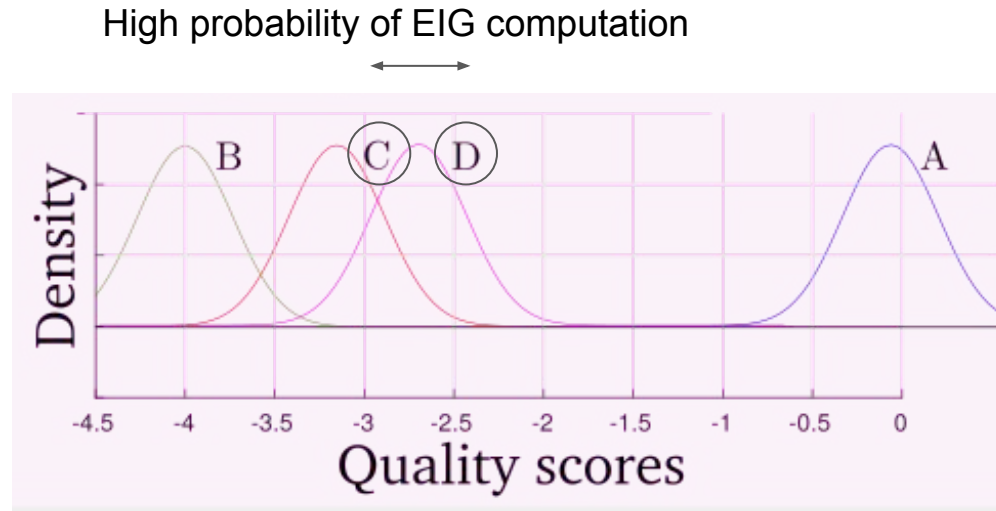
Selective EIG evaluations



Selective EIG evaluations



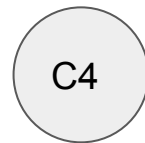
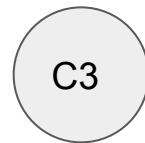
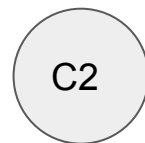
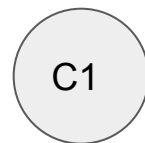
Selective EIG evaluations



Batch Mode

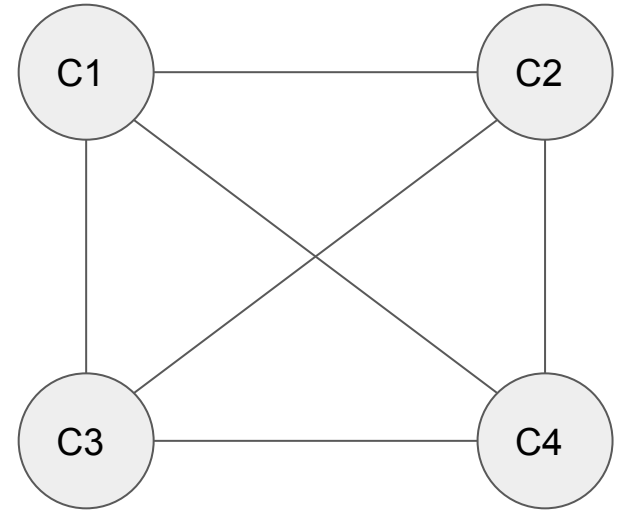
Batch Mode

- Vertices are conditions



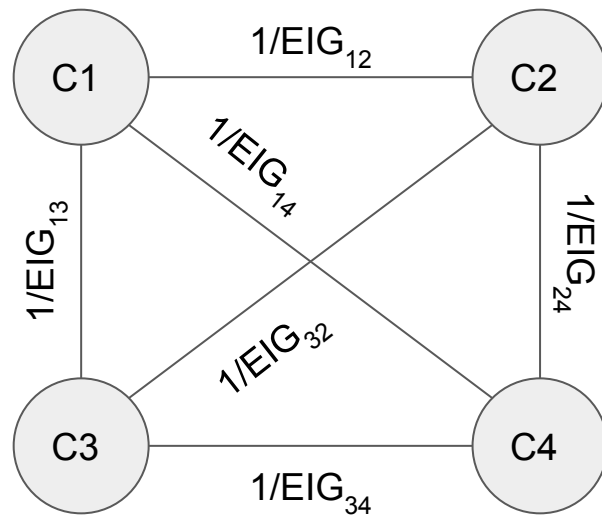
Batch Mode

- Vertices are conditions
- Edges are possible comparisons



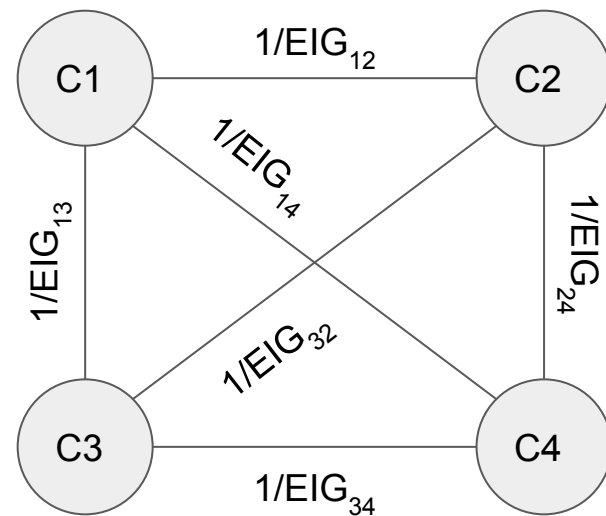
Batch Mode

- Vertices are conditions
- Edges are possible comparisons
- Weights are inverse of the expected information gain for comparison



Batch Mode

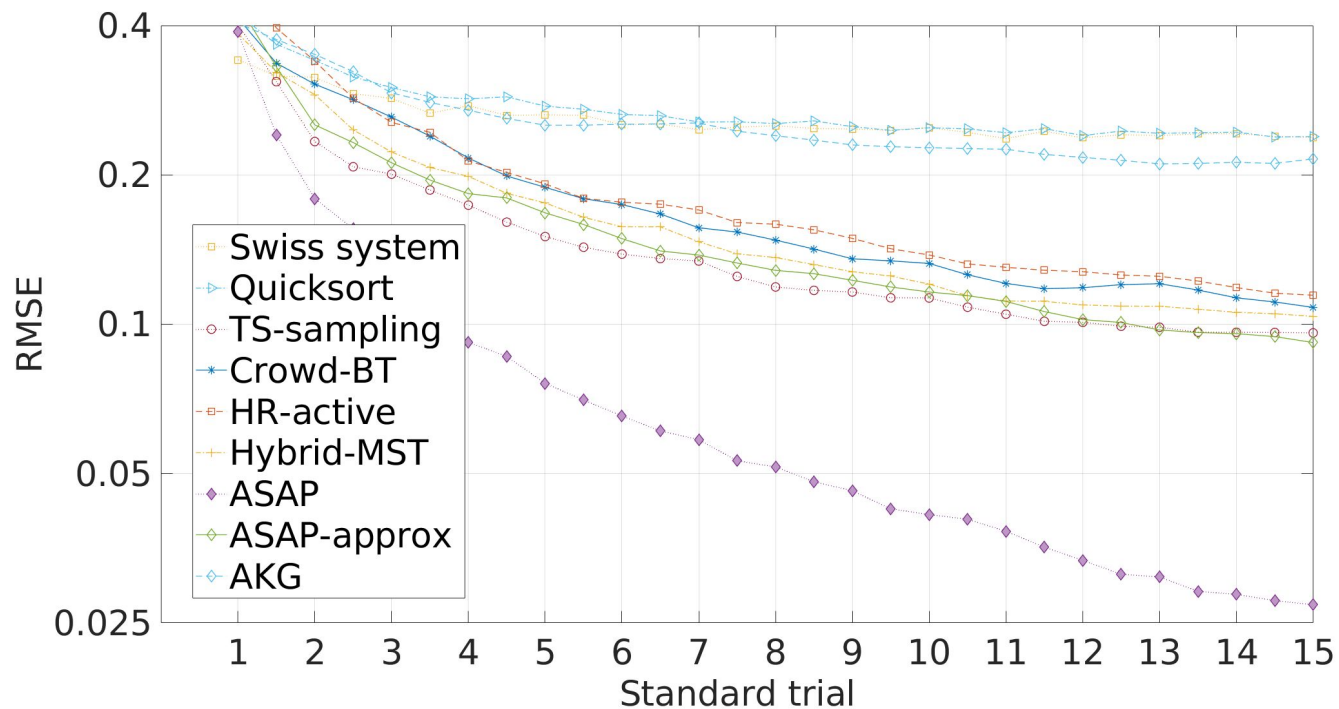
- Vertices are conditions
- Edges are possible comparisons
- Weights are inverse of the expected information gain for comparison
- Batch - comparisons forming edges in the minimum spanning tree (MST) [2]



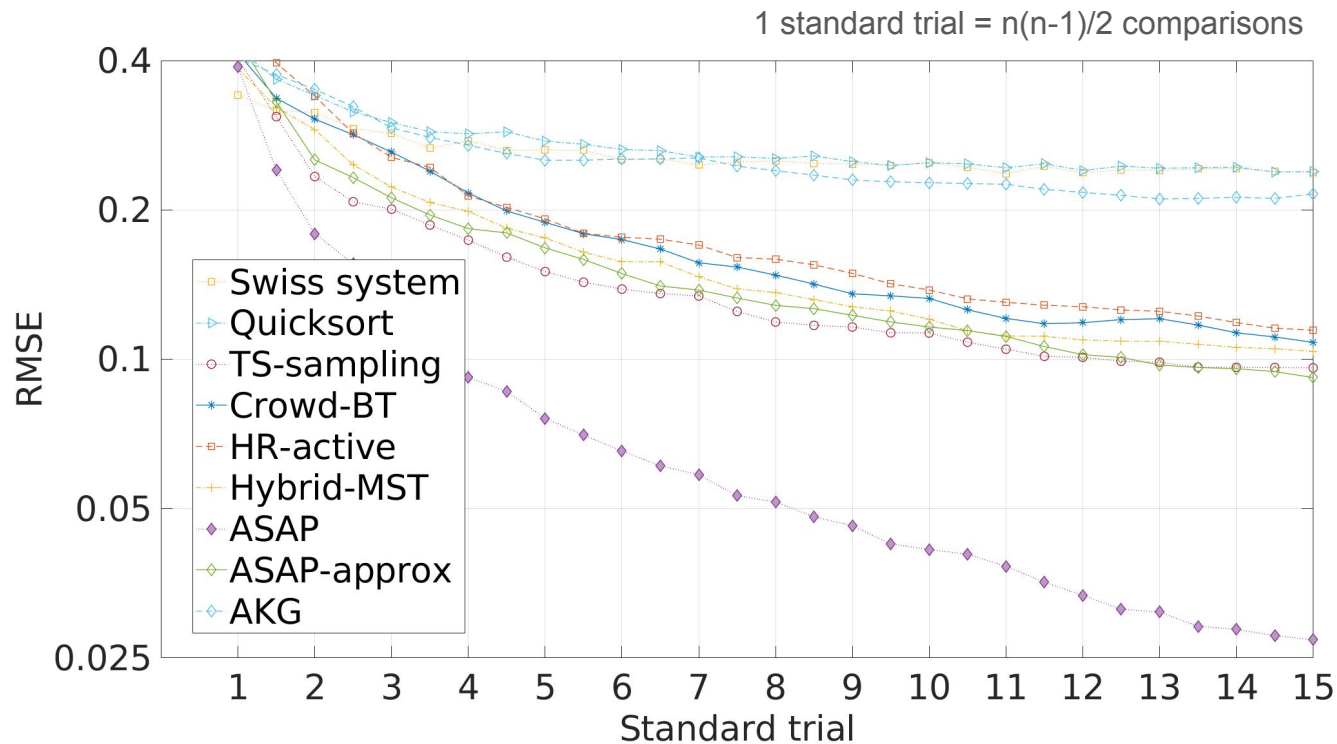
[2] J. Li, R. Mantiuk, J. Wang, S. Ling, and P. Le Callet, “Hybrid-MST: A hybrid active sampling strategy for pairwise preference aggregation,” NIPS, 31st Conference on Neural Information Processing Systems, 2018.

Monte Carlo Simulation

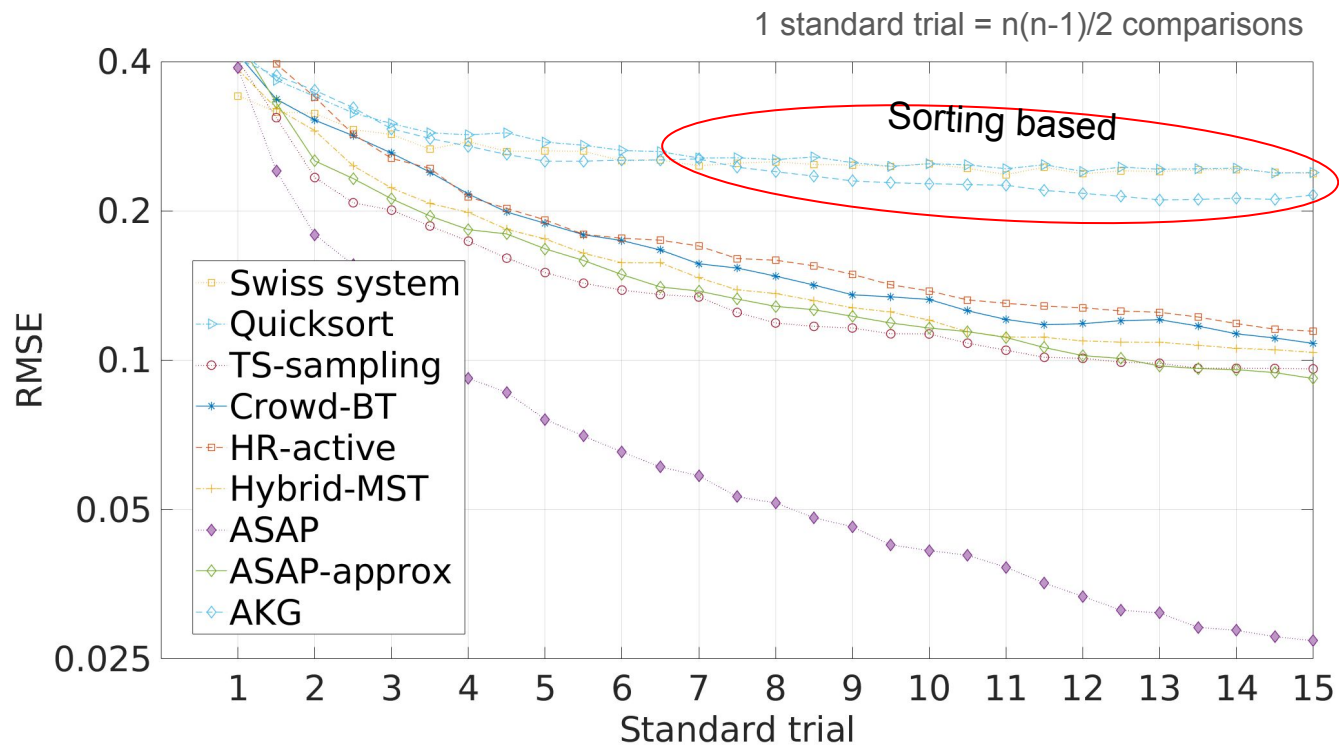
Monte Carlo Simulation



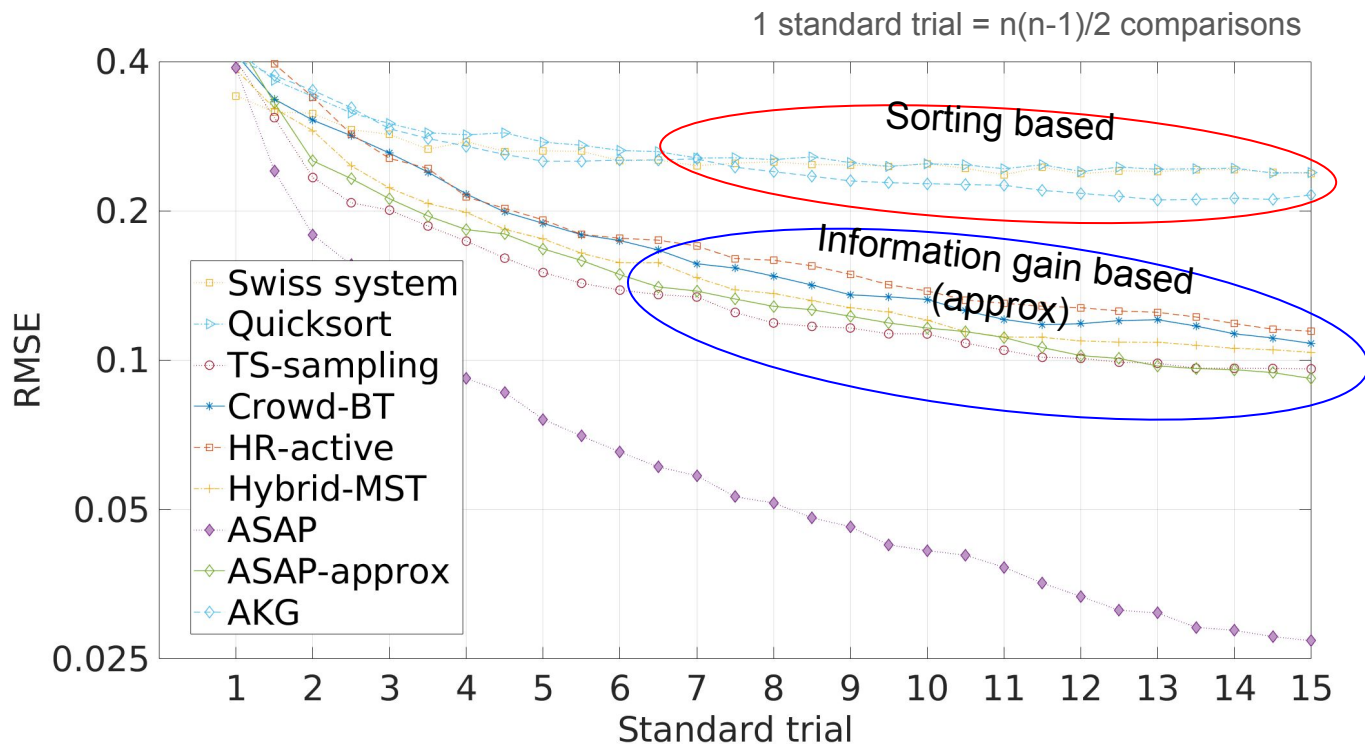
Monte Carlo Simulation



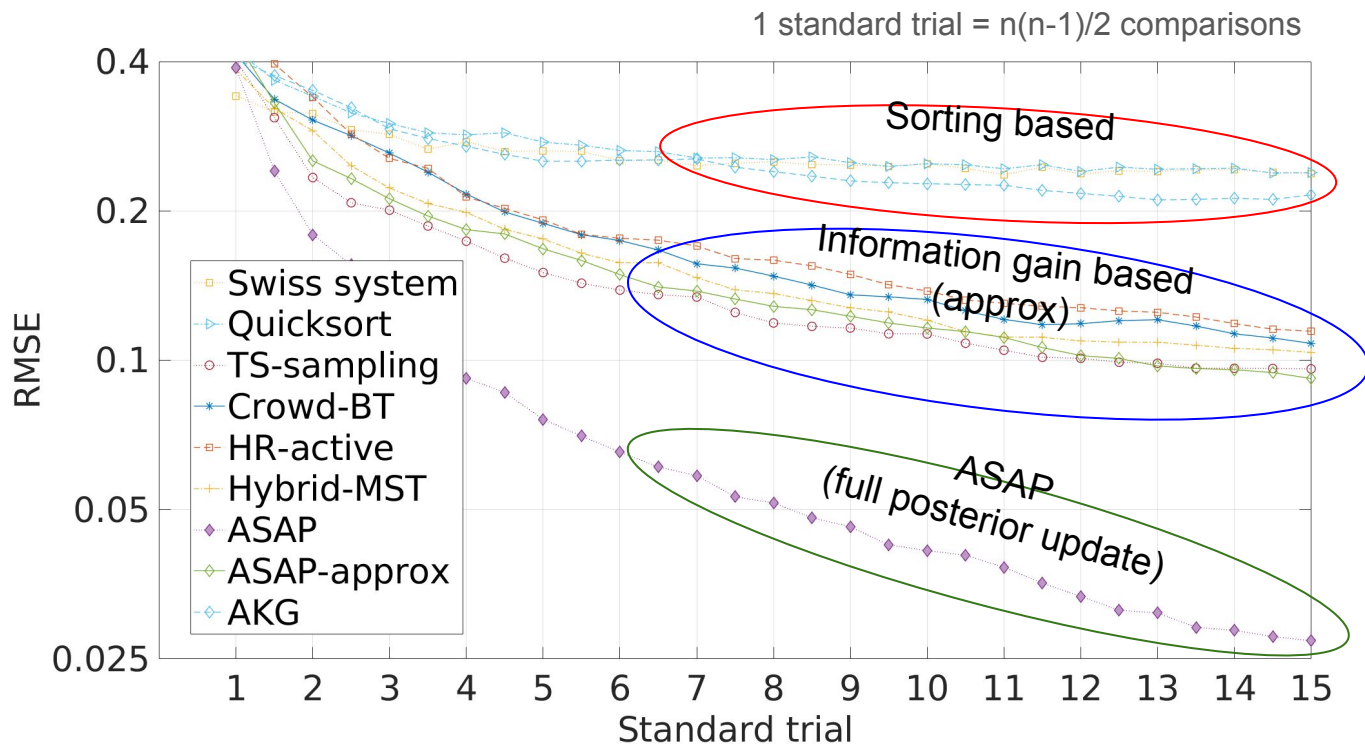
Monte Carlo Simulation



Monte Carlo Simulation

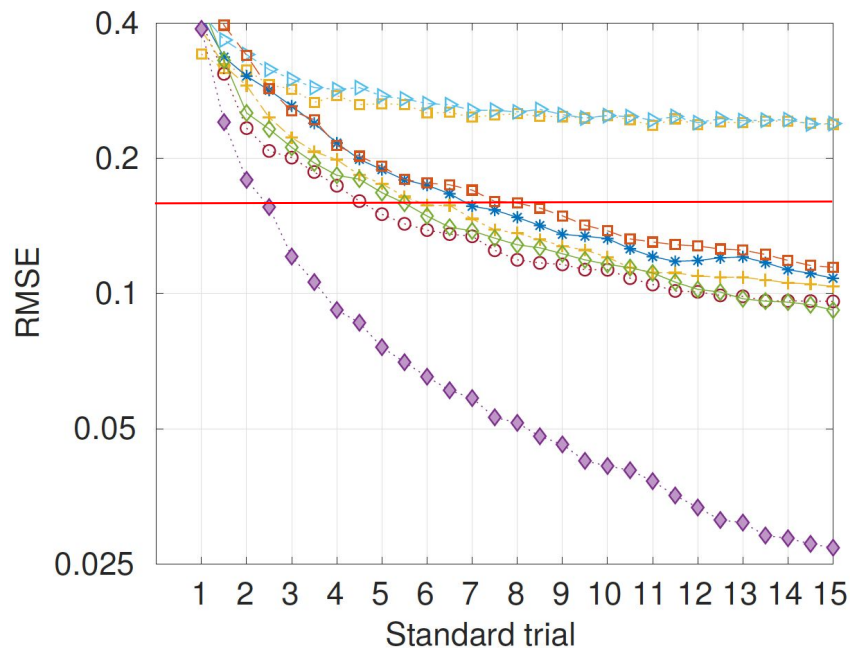


Monte Carlo Simulation



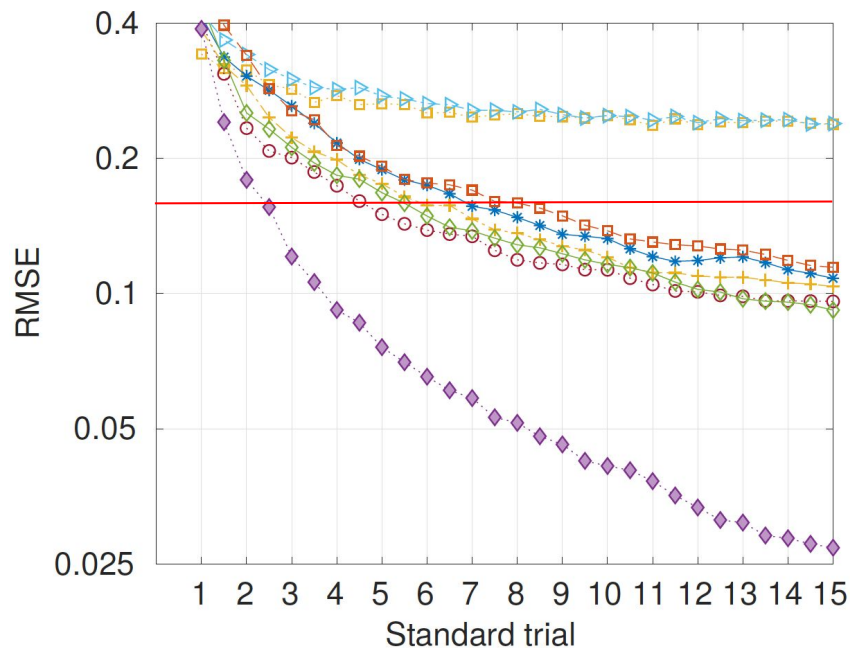
Experimental Effort

- Experimental Effort = time to achieve 0.15 RMSE

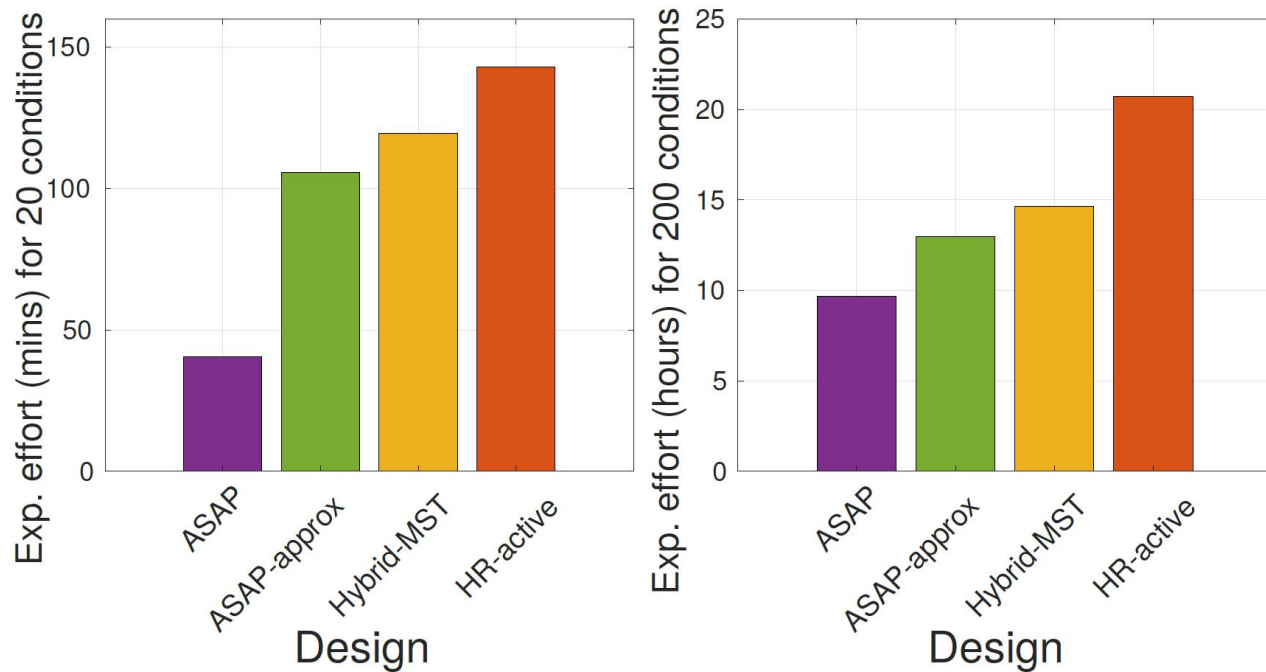


Experimental Effort

- Experimental Effort = time to achieve 0.15 RMSE
- Assume 5s decision time



Experimental Effort



Summary

- **Current methods** are sub-optimal, relying on **partial update** of the posterior distribution;
- **ASAP** computes the **full posterior update**, crucial to achieving **highest accuracy**;
- **ASAP** computes EIG for most informative pairs, **reducing the computational cost** by up to 80%;
- **ASAP** selects **batches** using a minimum spanning tree method.

Code: <https://github.com/gfxdisp/asap>

Contact: am2442@cam.ac.uk

