EFFICIENTLY TRAINING NEURAL NETWORKS FOR IMPERFECT INFORMATION GAMES BY SAMPLING INFORMATION SETS

Timo Bertram, Johannes Fürnkranz, Martin Müller
Johannes Kepler University Linz, University of Alberta

KI 2024

RECONNAISSANCE
BLIND CHESS

IMPERFECT INFORMATION GAMES


11.560
0.9%
1.286
0.12%
2.286
1.51%
0.11%
0.156
1.0287
2.344
0.1204
0.1902

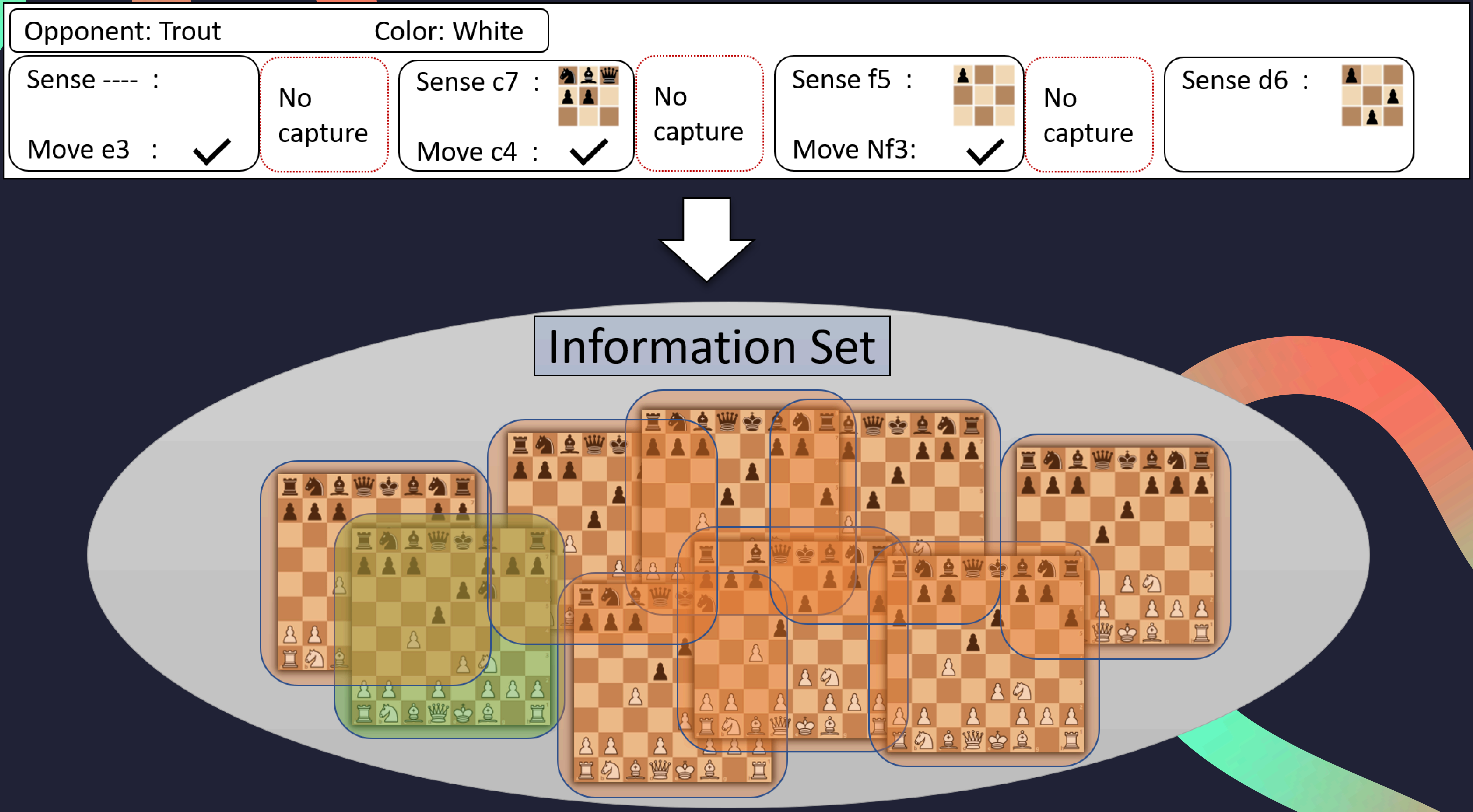
RECONNAISSANCE BLIND CHESS


1
2
3
4
5
6
7
8
h
g
f
e
d
c
b
a

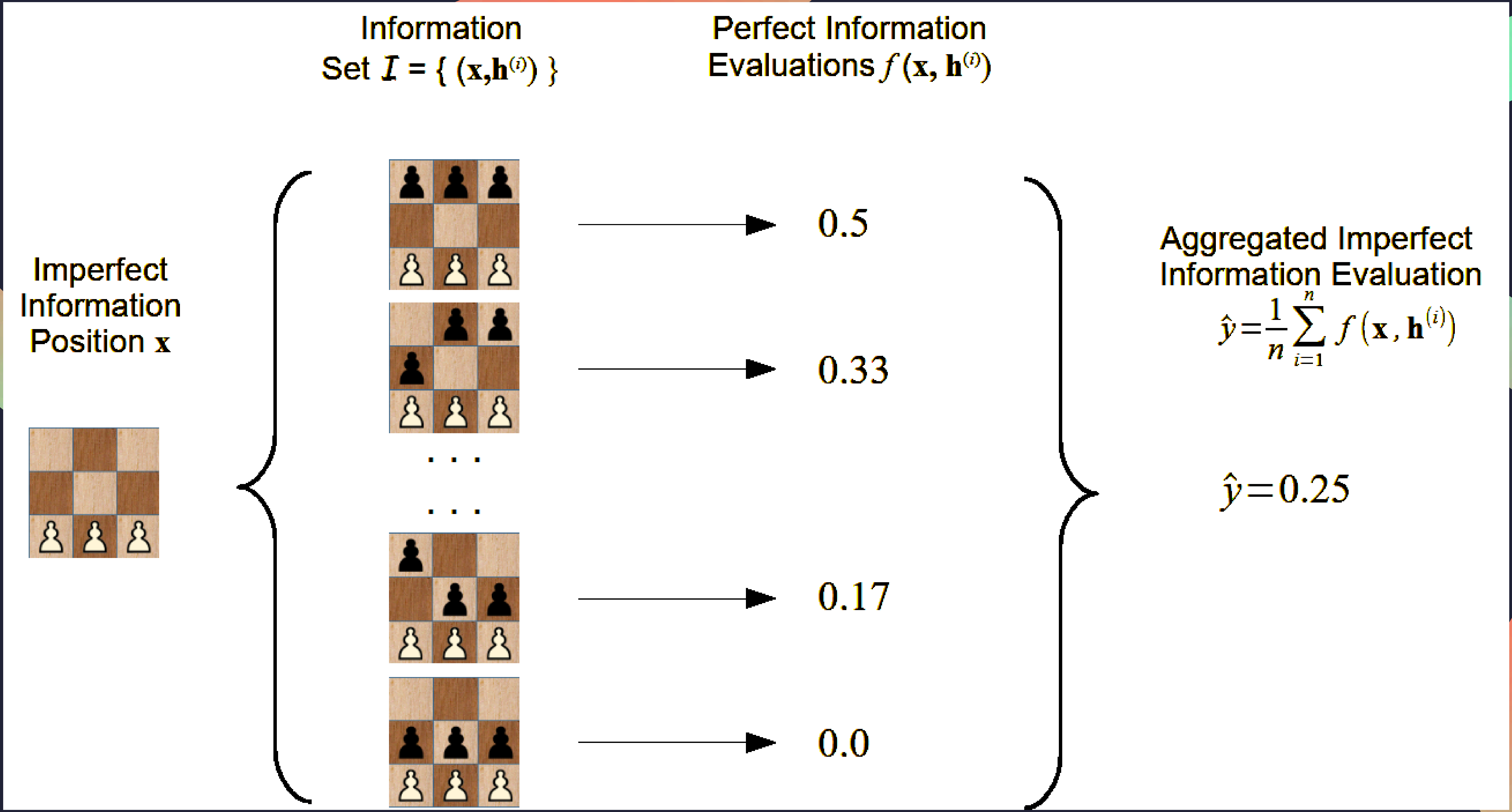

8
7
6
5
4
3
2
1
a b c d e f g h


8
7
6
5
4
3
2
1
a b c d e f g h

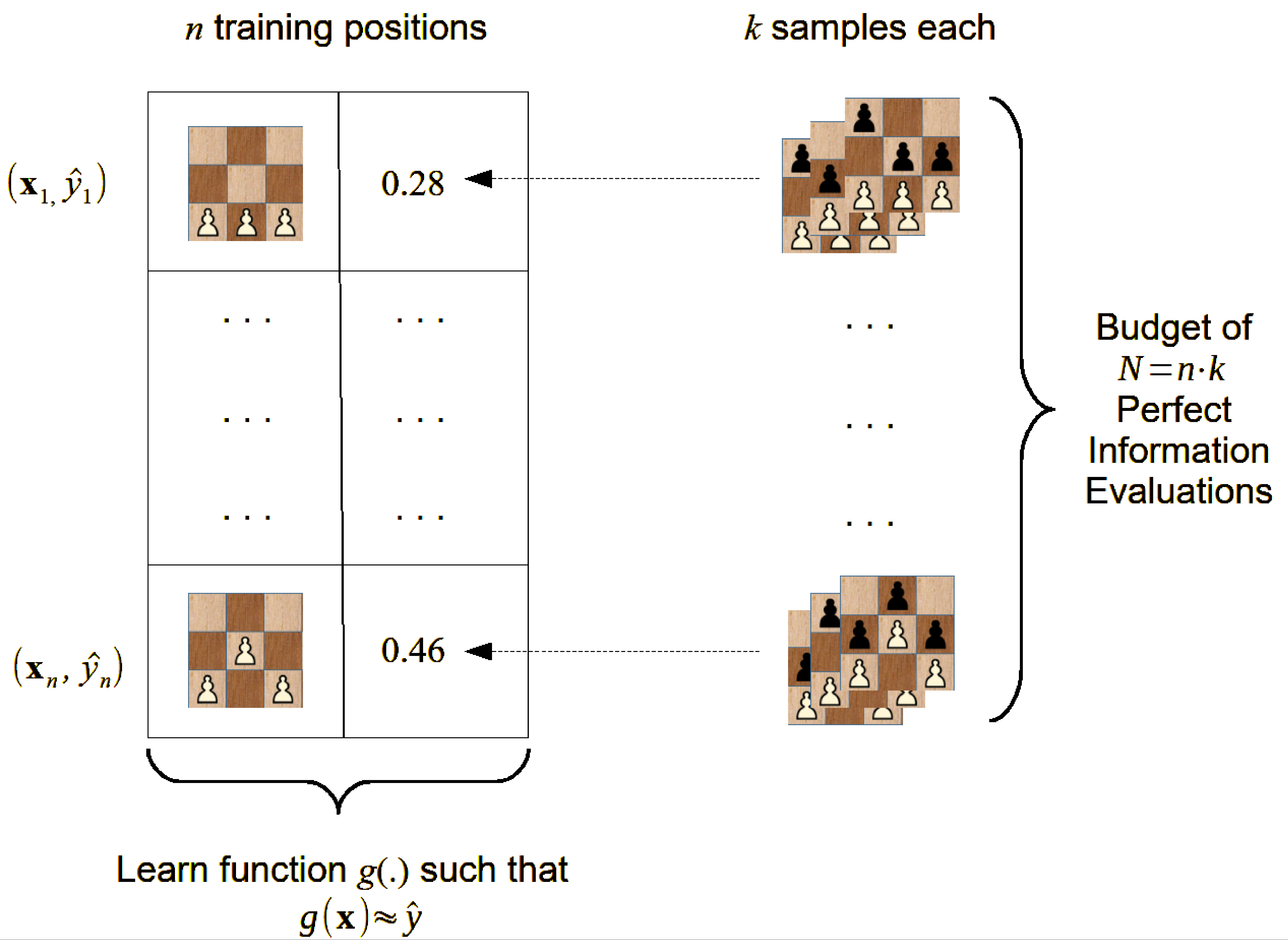

1
2
3
4
5
6
7
8
h g f e d c b a


Opponent: Trout
Color: White
Sense ---- :
Move e3 :
No capture
Sense c7 :
Move c4 :
No capture
Sense f5 :
Move Nf3:
No capture
Sense d6 :
Information Set

APPROXIMATING INFORMATION SETS



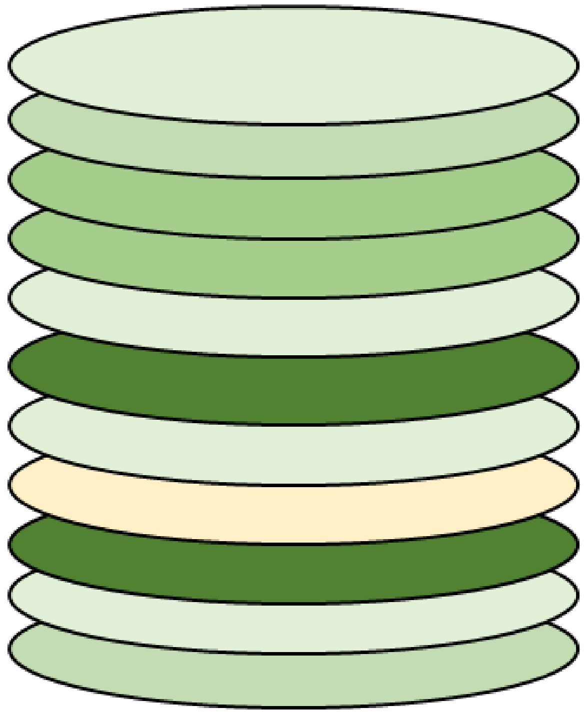
Information Set $I = \{ (\mathbf{x},\mathbf{h}^{(i)}) \}$
Perfect Information Evaluations $f(\mathbf{x}, \mathbf{h}^{(i)})$
Imperfect Information Position $\mathbf{x}$
0.5
0.33
0.17
0.0
Aggregated Imperfect Information Evaluation
$\hat{y} = \frac{1}{n}\sum_{i=1}^{n} f(\mathbf{x}, \mathbf{h}^{(i)})$
$\hat{y} = 0.25$

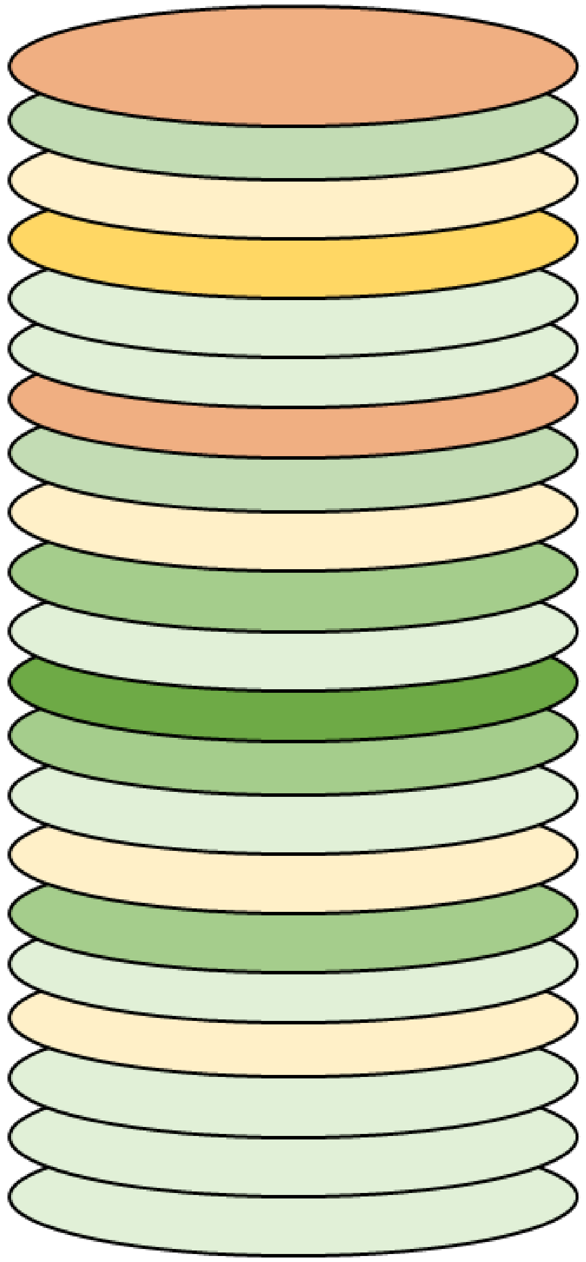


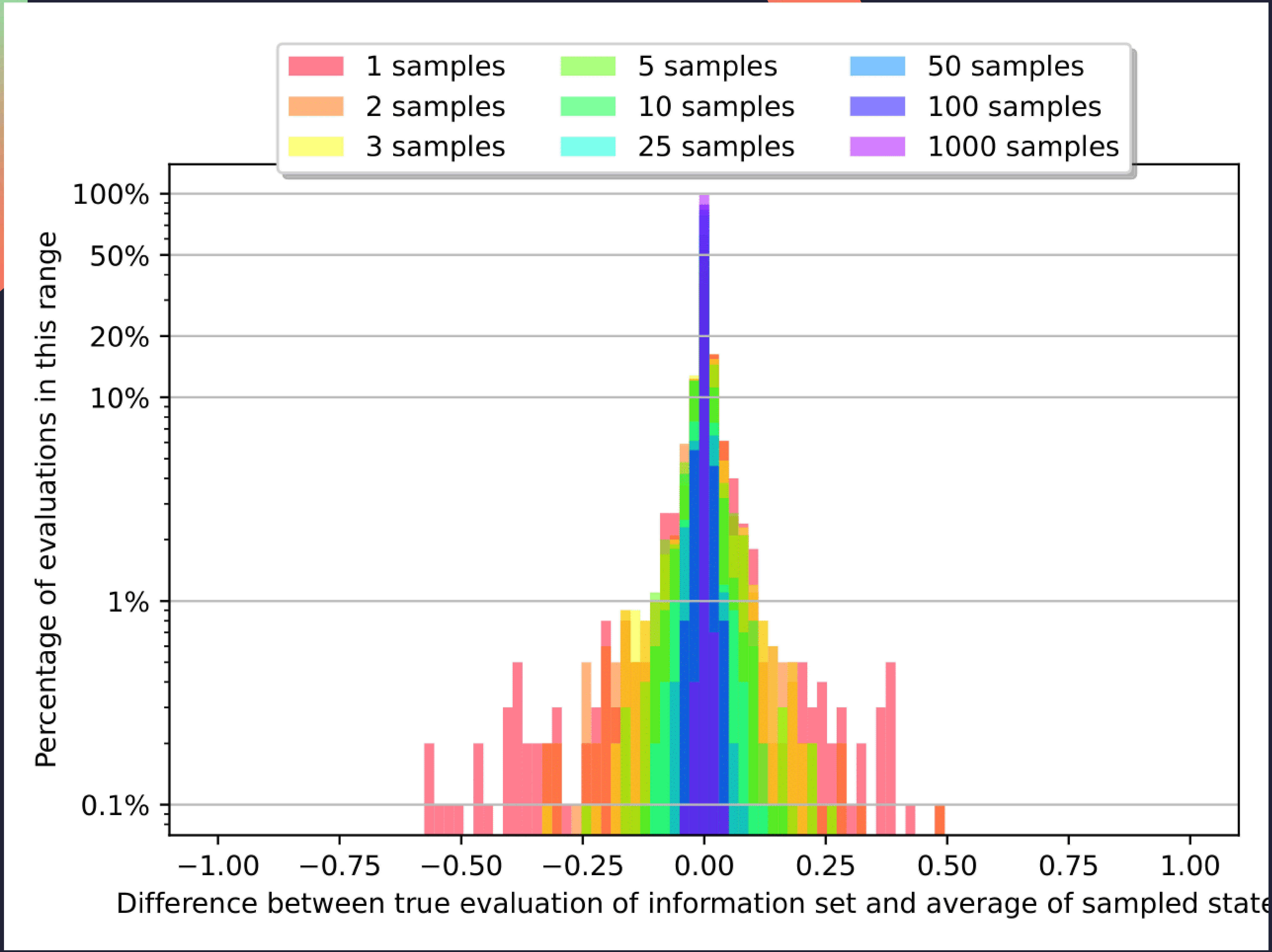
n training positions
k samples each
$(\mathbf{x}_1, \hat{y}_1)$
0.28
$(\mathbf{x}_n, \hat{y}_n)$
0.46
Budget of $N = n \cdot k$ Perfect Information Evaluations
Learn function $g(.)$ such that $g(\mathbf{x}) \approx \hat{y}$

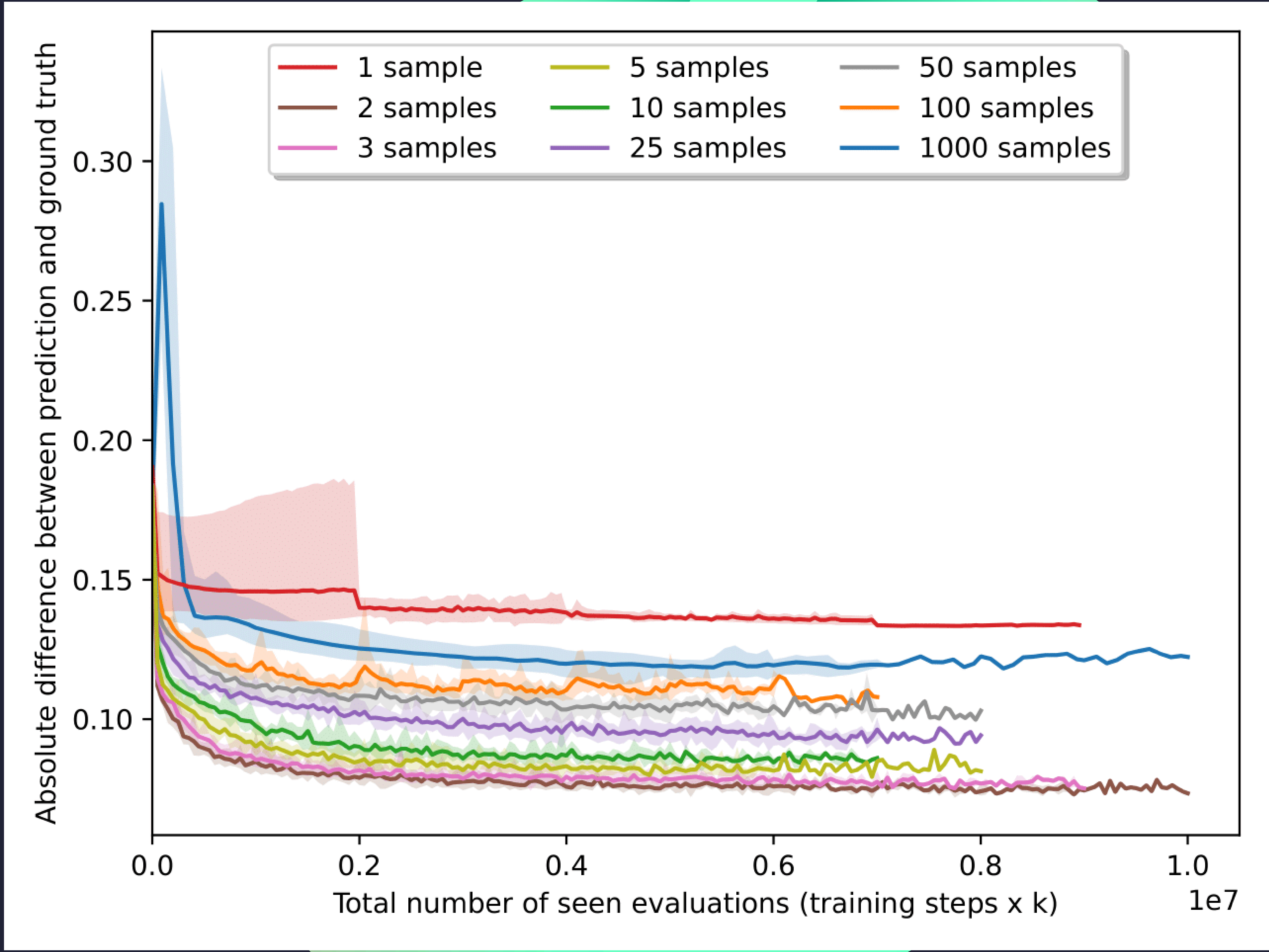
Larger k
Smaller n

Smaller k
Larger n

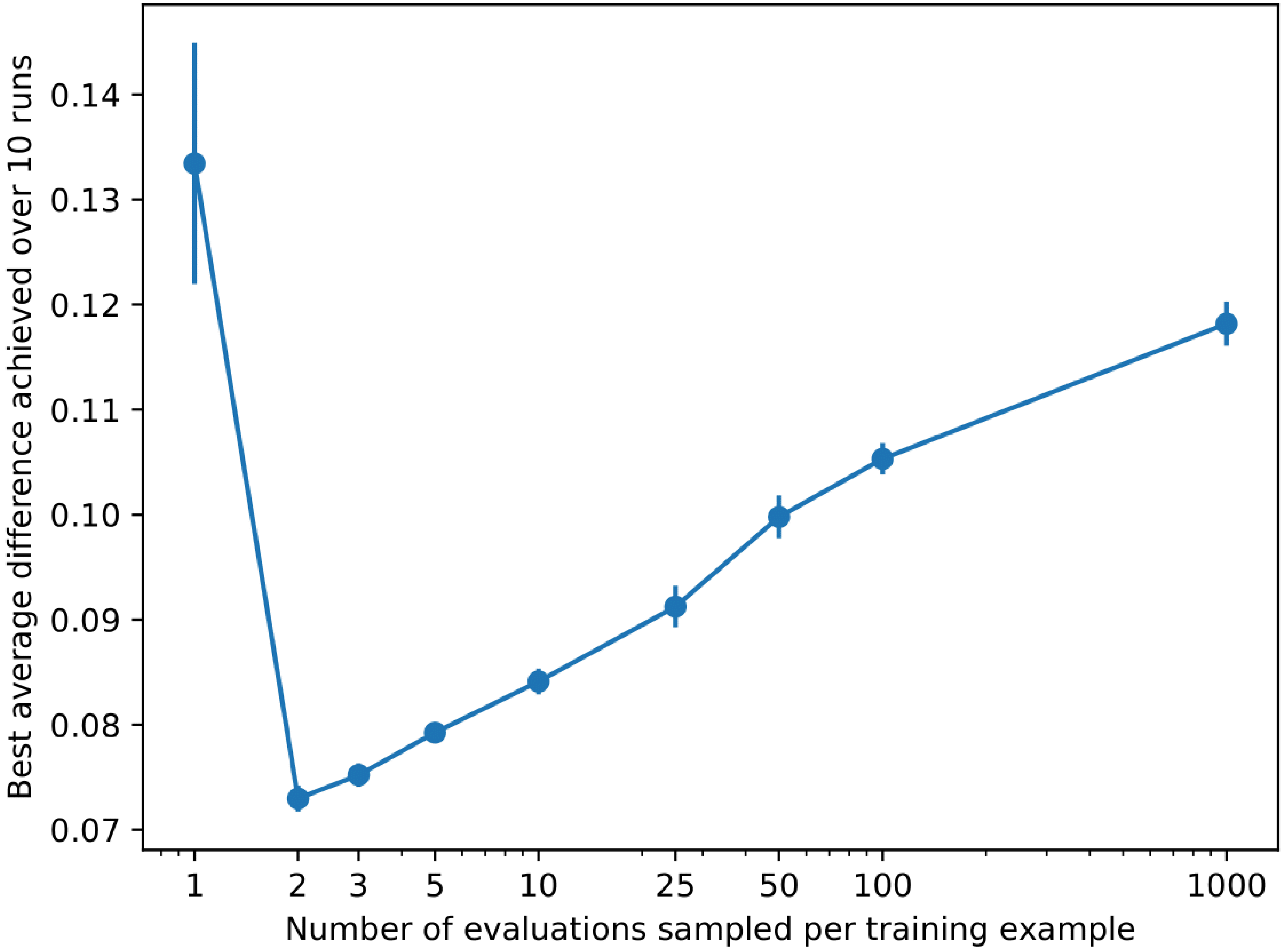





1 samples
2 samples
3 samples
5 samples
10 samples
25 samples
50 samples
100 samples
1000 samples
100%
50%
20%
10%
1%
0.1%
Percentage of evaluations in this range
−1.00
−0.75
−0.50
−0.25
0.00
0.25
0.50
0.75
1.00
Difference between true evaluation of information set and average of sampled state

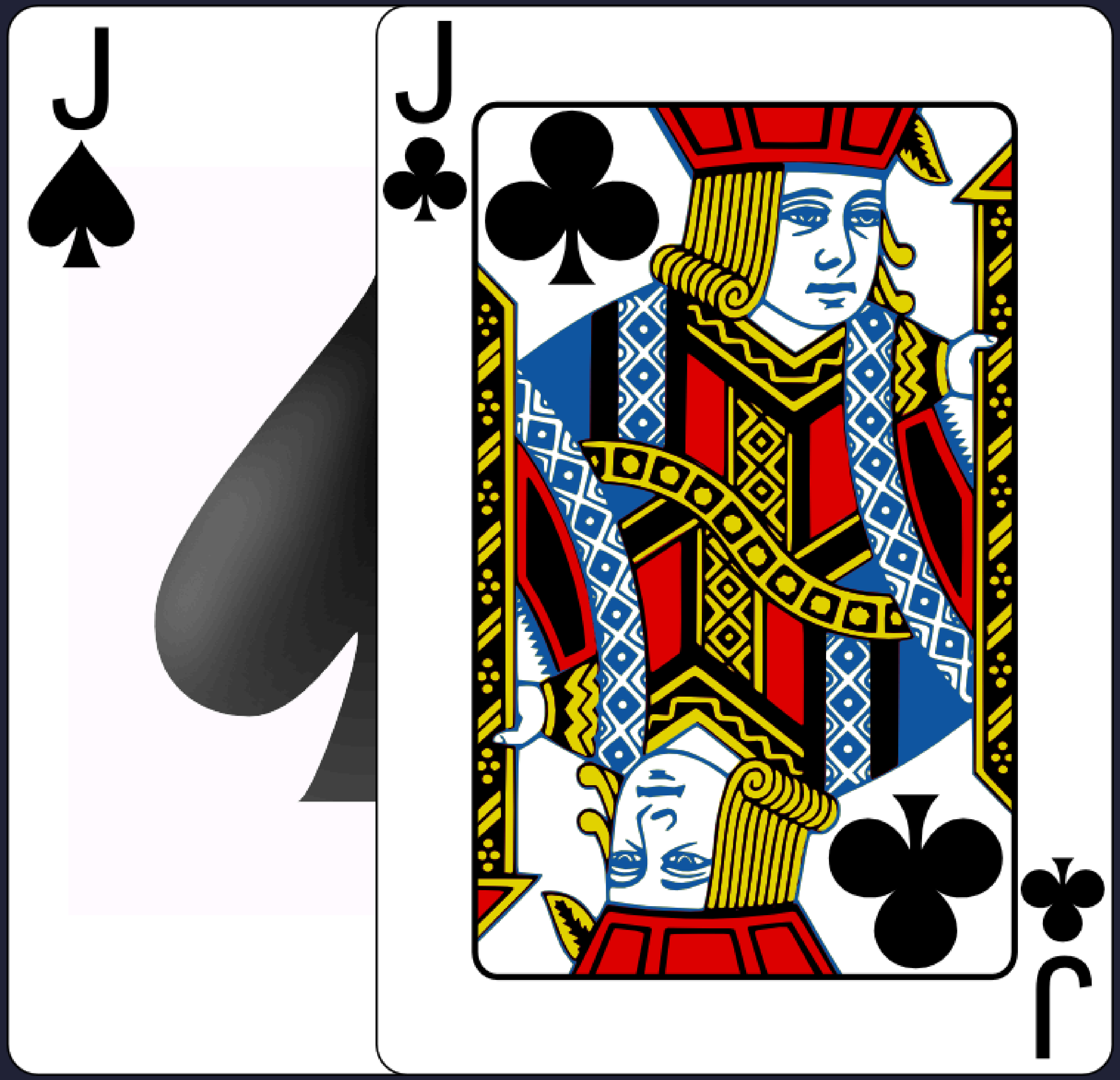


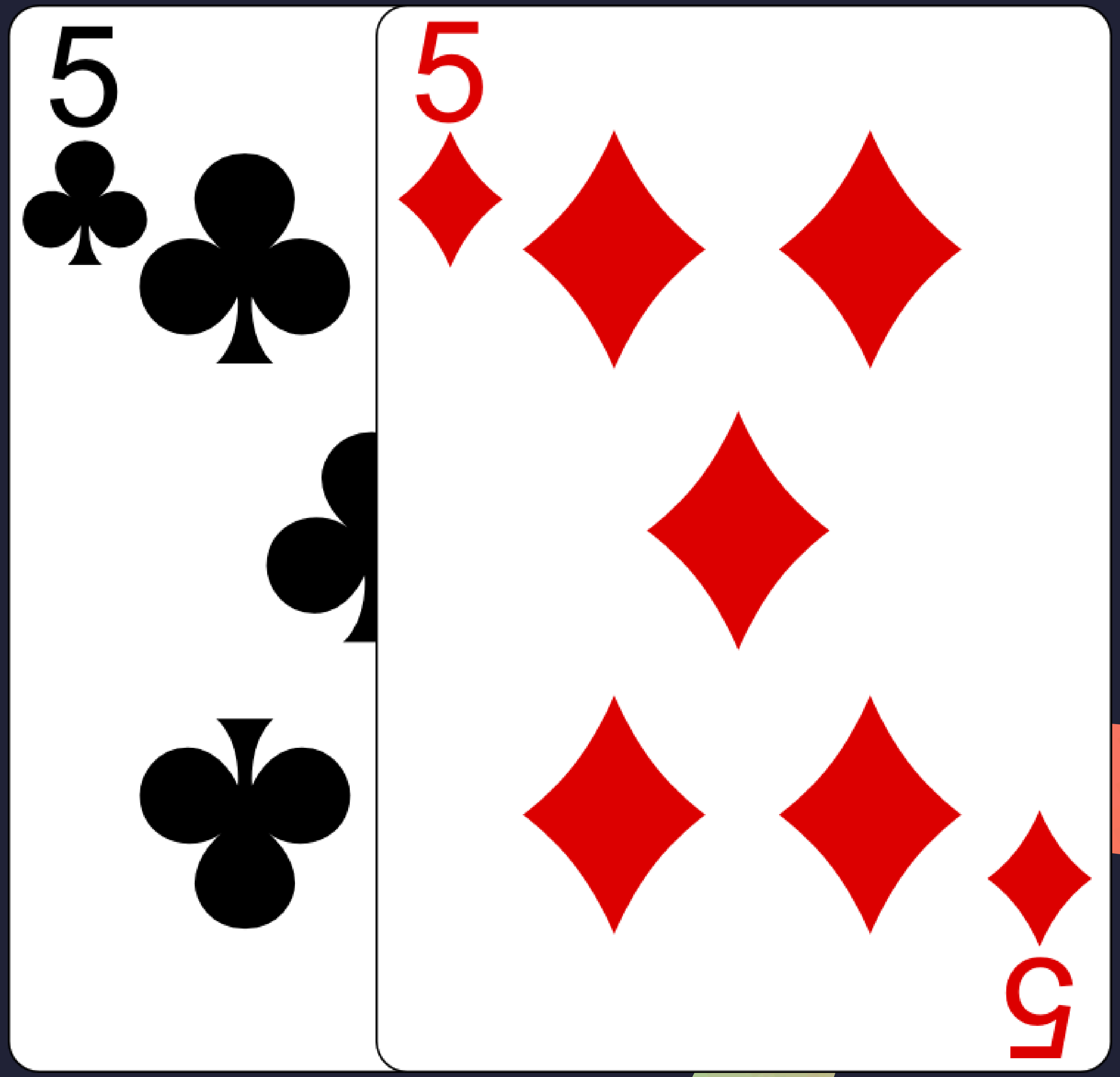
1 sample
2 samples
3 samples
5 samples
10 samples
25 samples
50 samples
100 samples
1000 samples
Absolute difference between prediction and ground truth
0.30
0.25
0.20
0.15
0.10
0.0
0.2
0.4
0.6
0.8
1.0
1e7
Total number of seen evaluations (training steps x k)

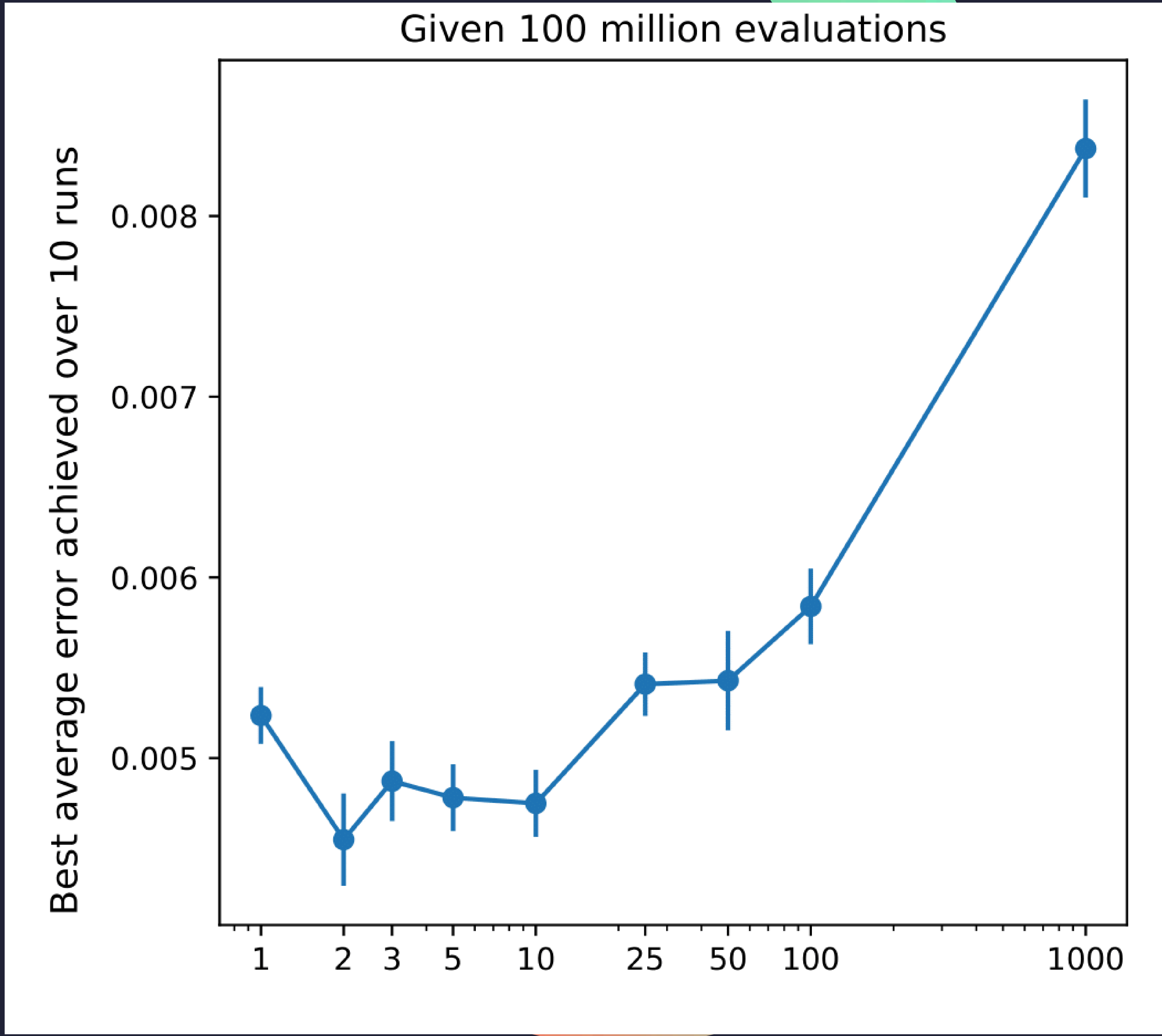


Best average difference achieved over 10 runs
0.14
0.13
0.12
0.11
0.10
0.09
0.08
0.07
1
2
3
5
10
25
50
100
1000
Number of evaluations sampled per training example

What is better?







Given 100 million evaluations
Best average error achieved over 10 runs
0.008
0.007
0.006
0.005
1
2
3
5
10
25
50
100
1000

Takeaways

We want to build training datasets based on subsets of information sets

There is a clear trade-off between approximating more sets and receiving better estimates

Using 2 samples per information set empirically worked best

THANK YOU!

TBERTRAM@FAW.JKU.AT @BERTRAMTIMO