

Search-Tree Scaling in Parallel Monte Carlo Tree Search

Scott Cheng¹, Meng-Yu Tsai², Ding-Yong Hong³, Mahmut Taylan Kandemir¹

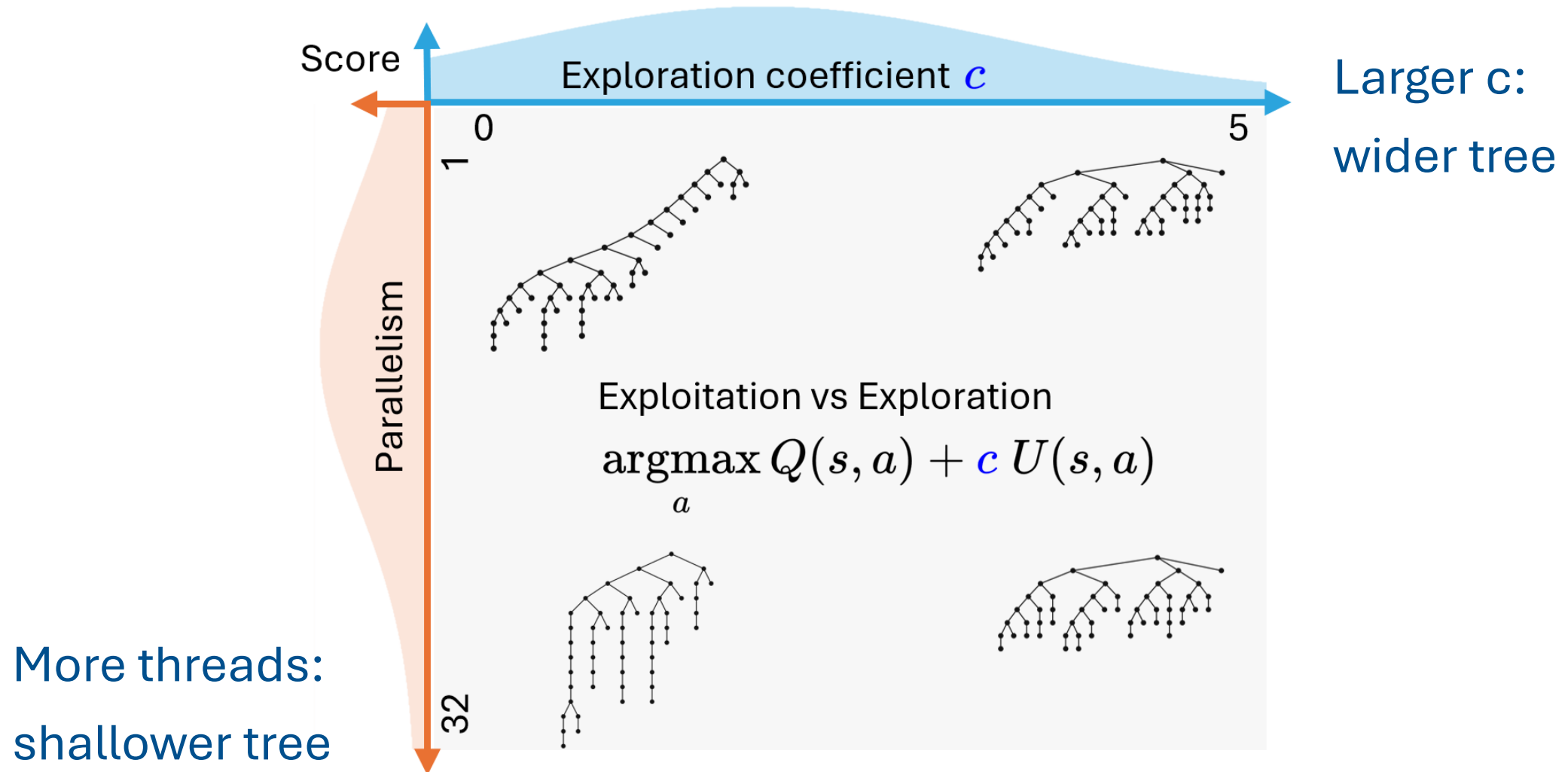
¹ The Pennsylvania State University, USA

² Independent

³ Academia Sinica, Taiwan



Exploration-Exploitation Tradeoff in MCTS



Characterization of the Asymptotic Tree Structure

Under regularity assumptions:

- Constant exploration coefficient c , optimal depth grows $H_{\text{opt}}(n) = \Theta(\sqrt{n})$
- Dynamic exploration coefficient $c(\cdot)$, $H_{\text{opt}}(n) = \Theta\left(\int_1^n \frac{dm}{c(m)\sqrt{m}}\right)$ for $c(m) = O(\sqrt{m})$

In particular, $c(m) = m^\alpha L(m)$ for some slowly-varying function L

$$H_{\text{opt}}(n) = \begin{cases} \Theta\left(\frac{\sqrt{n}}{c(n)}\right), & -\frac{1}{2} < \alpha < \frac{1}{2}, \\ \Theta\left(\log n + \int_1^n \frac{dm}{mL(m)}\right), & \alpha = \frac{1}{2}, \\ \Theta(\log n), & \alpha > \frac{1}{2}. \end{cases}$$

↑ smaller growth rate of exploration c
=> deeper tree, asymptotically

Characterization of the Asymptotic Tree Structure

Under regularity assumptions:

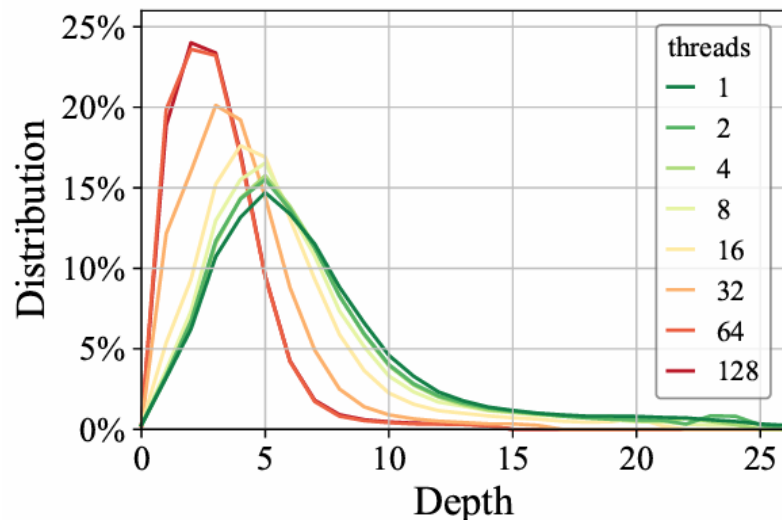
- Constant exploration coefficient c , optimal depth grows $H_{\text{opt}}(n) = \Theta(\sqrt{n})$
- Dynamic exploration coefficient $c(\cdot)$, $H_{\text{opt}}(n) = \Theta \left(\int_1^n \frac{dm}{c(m)\sqrt{m}} \right)$ for $c(m) = O(\sqrt{m})$
- Example 1: $c(m) = \sqrt{m}$
$$H_{\text{opt}}(n) = \Theta \left(\int_1^n \frac{1}{m} dm \right) = \Theta(\log n)$$
- Example 2: $c_{\text{puct}}(m) = \log \left(\frac{m + \text{pb_c_base} + 1}{\text{pb_c_base}} \right) + \text{pb_c_init}$,
(MuZero)
$$H_{\text{opt}}(n) = \Theta \left(\frac{\sqrt{n}}{c(n)} \right) = \Theta \left(\frac{\sqrt{n}}{\log n} \right)$$

smaller growth rate of exploration c
=> deeper tree, asymptotically

Characterization of the Asymptotic Tree Structure

Under the assumption on thread contention:

- Fixed thread count T , optimal depth grows $H_{\text{opt}} = \Theta(\sqrt{n})$
- General parallelism T ,
 - (1) If $T = o(\sqrt{n})$, $H_{\text{opt}}(n, T) = \Theta(\sqrt{n})$,
 - (2) If $T \gg \sqrt{n}$, $H_{\text{opt}}(n, T) = \Theta(n/T)$.

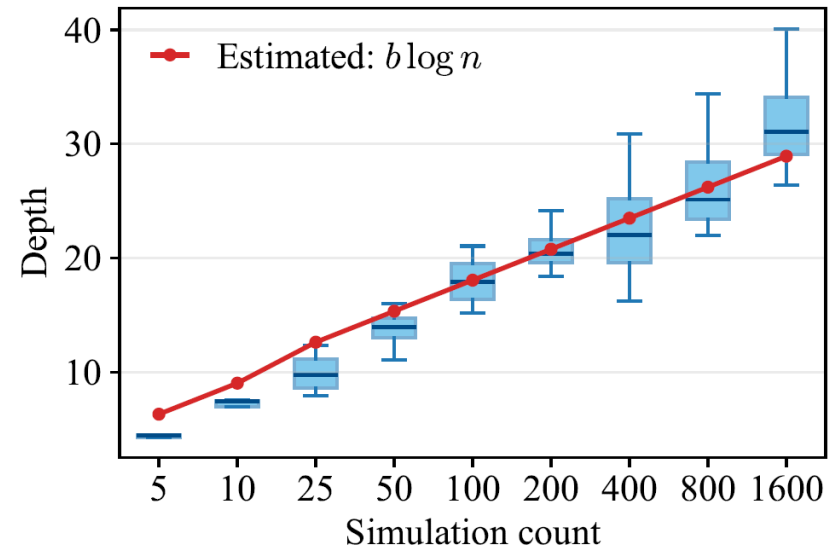
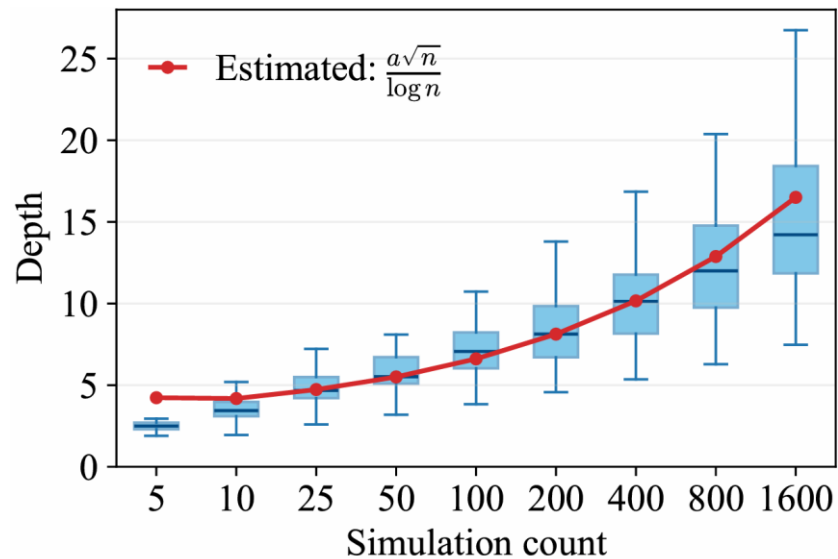


Empirical distribution of nodes over depth:

“moderate” level of parallelism if less than $\approx \sqrt{400} = 20$ threads

Evaluation: Empirical Scaling of Search Depth

Empirical depth closely matches the estimate with $R^2 = 0.86$



$$c_{\text{puct}}(m) = \log \left(\frac{m + \text{pb_c_base} + 1}{\text{pb_c_base}} \right) + \text{pb_c_init},$$

$$c_{\text{puct}}(m) = \sqrt{\left(\frac{m + \text{pb_c_base} + 1}{\text{pb_c_base}} \right)} + \text{pb_c_init}$$

Evaluation: Joint Tuning

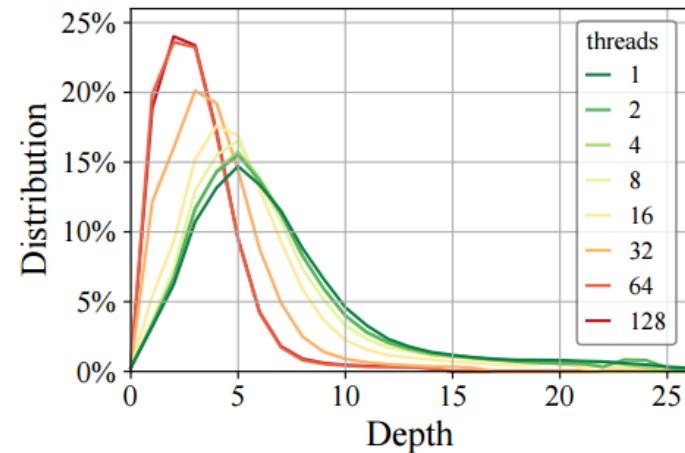
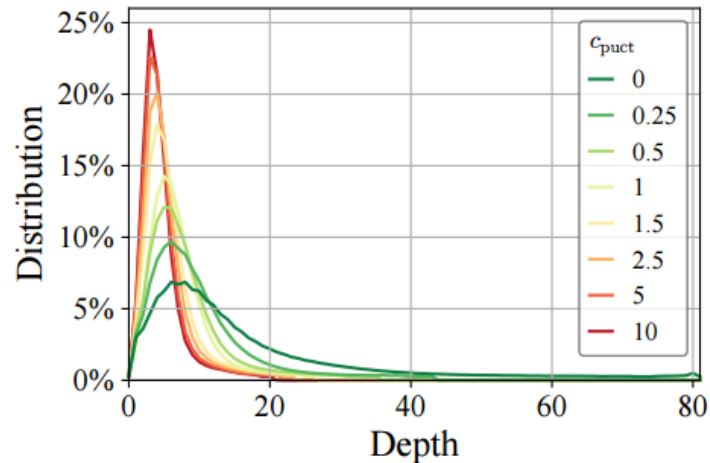
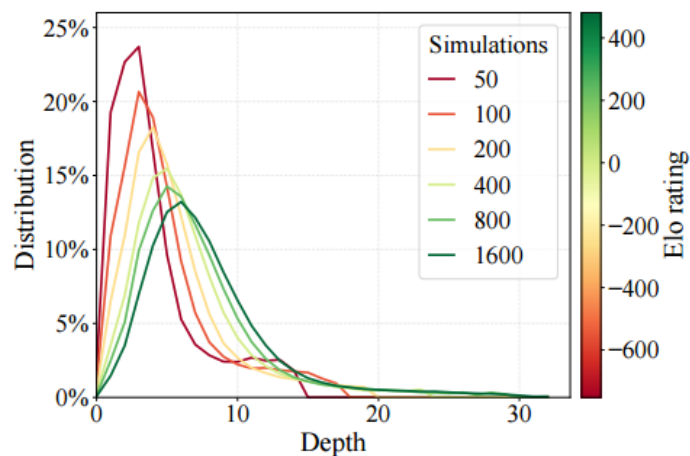
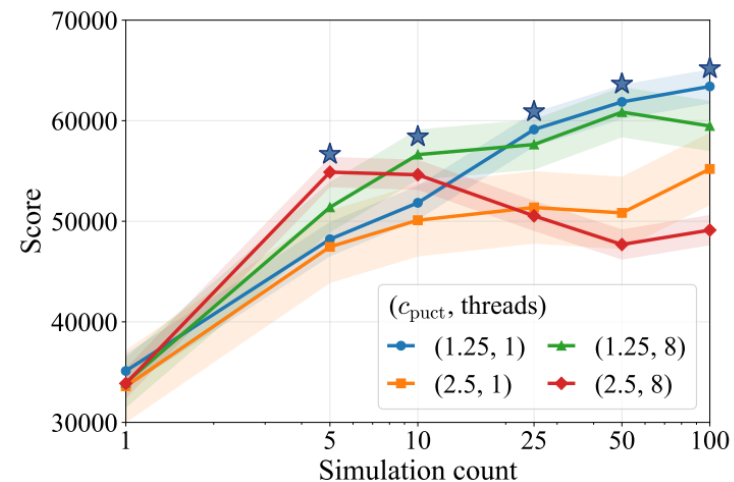
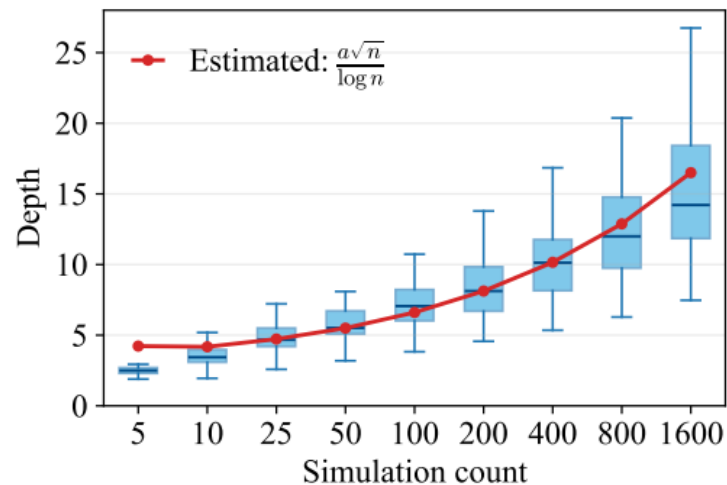
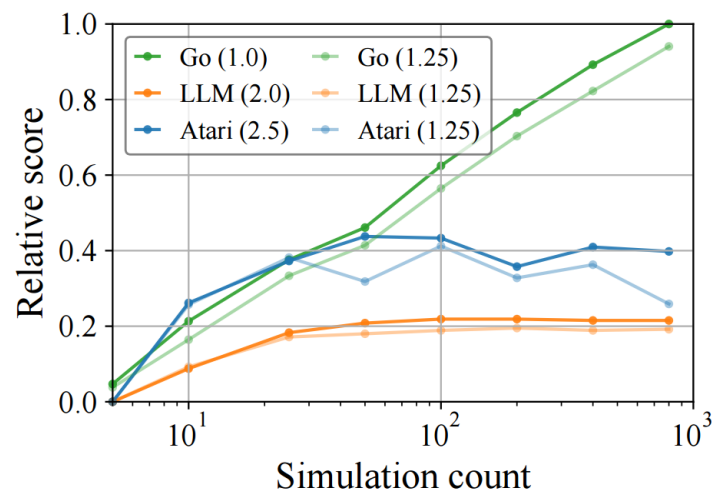
In Go experiments, the least-squares fit ($R^2 = 0.97$) gives

$$c_{\text{puct}}(T) = 1.498 - 0.192 \log_2 T.$$

i.e., each doubling of threads lowers the best c by ≈ 0.19

threads	c_{puct}				
	0.8	1.0	1.1	1.3	1.5
1	-18.1	11.8	22.3	27.9	30.0
2	-9.0	-0.7	1.4	21.6	-6.9
4	-24.4	-18.1	6.3	-12.5	-22.3
8	-38.4	-11.8	-19.5	-12.5	-34.2
10	-46.1	-53.2	-57.5	-46.2	-51.1

For more details... Please check our paper!





Thank you

Acknowledgements to:

- DOE grant and NSF grants*
- National Science and Technology Council of Taiwan
- Taiwan National Center for High-performance Computing (NCHC)
- Anonymous reviewers for their valuable comments

* The opinions expressed in this paper are those of the authors and not necessarily the views of the NSF or the DOE.