Branch Mispredictions

[KS06] Kanela Kaligosi, Peter Sanders. ***How Branch Mispredictions Affect Quicksort***.
Proc. 14th Annual European Symposium on Algorithms (ESA),
Lecture Notes in Computer Science, Volume 4168, 780-791, Springer, 2006.

[BM05] Gerth Stølting Brodal, Gabriel Moruz.
Tradeoffs Between Branch Mispredictions and Comparisons for Sorting Algorithms.
Proc. 9th International Workshop on Algorithms and Data Structures (WADS),
Lecture Notes in Computer Science, Volume 3608, 385-395, Springer, 2005.

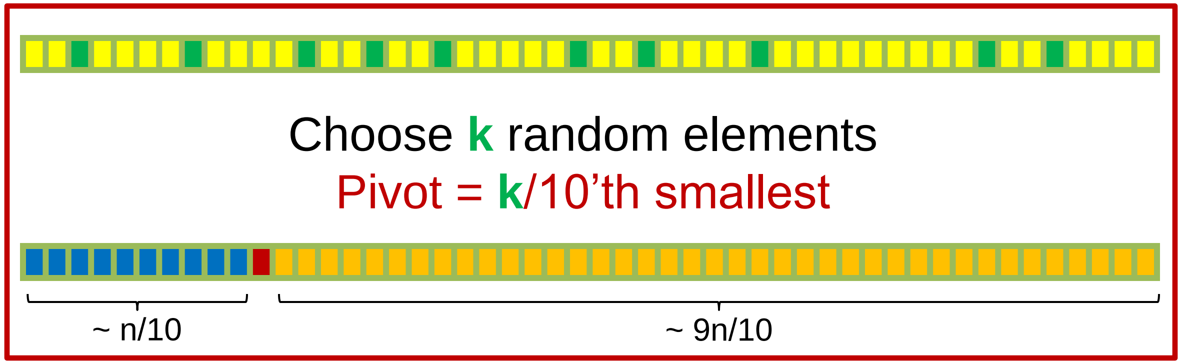
[BFM05] Gerth Stølting Brodal, Rolf Fagerberg, Gabriel Moruz. ***On the Adaptiveness of Quicksort***.
In Proc. 7th Workshop on Algorithm Engineering and Experiments (ALENEX), pages 130-140, 2005.

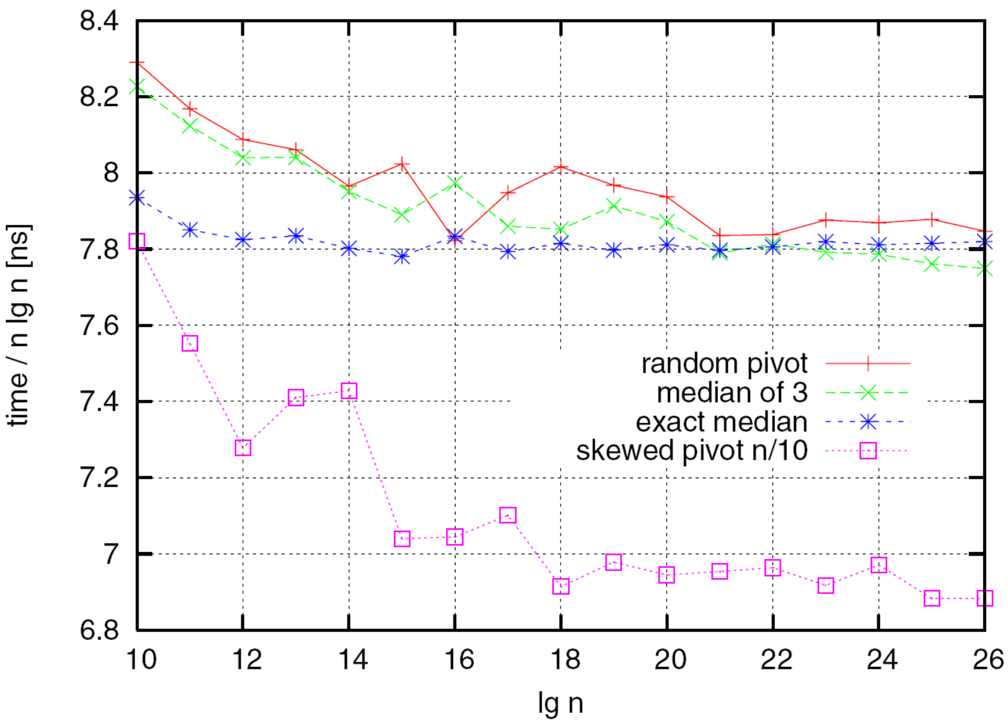
[BM06] Gerth Stølting Brodal, Gabriel Moruz. ***Skewed Binary Search Trees***.
In Proc. 14th Annual European Symposium on Algorithms (ESA),
Lecture Notes in Computer Science, Volume 4168, 708-719, Springer, 2006.

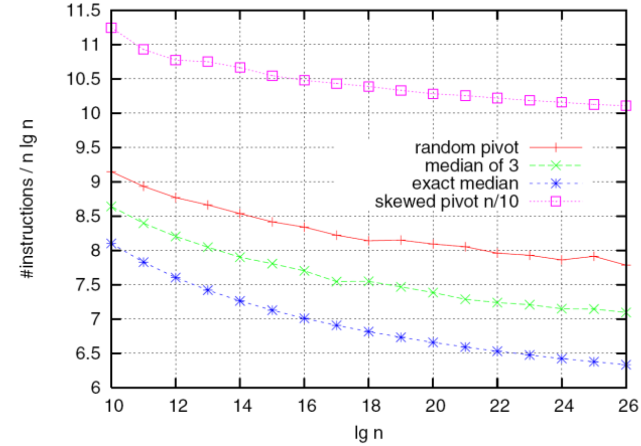
“How Branch Mispredictions Affect Quicksort”

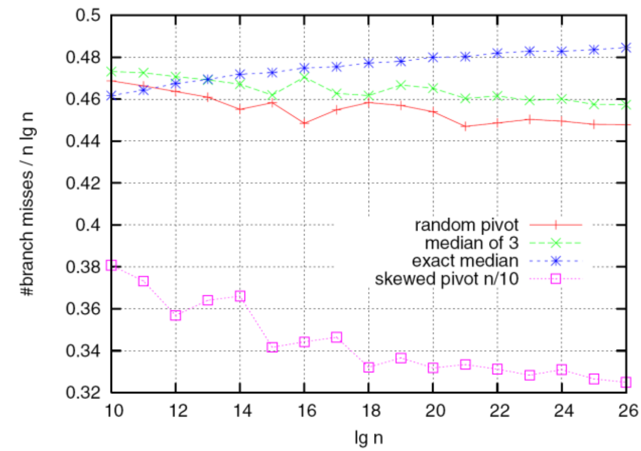
Kanela Kaligosi and Peter Sanders

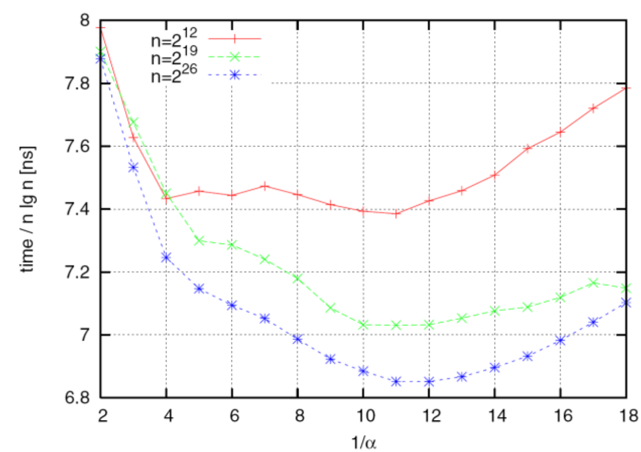
14th Annual European Symposium on Algorithms (ESA 2006)











3GHz Pentium 4 Prescott, GCC 3.3, PAPI, avg. 100 runs with random permutations of [1..n]

Algorithm 1. Sort array part $a[\ell..r]$

```
Procedure quicksort(ℓ, r : integer);
  if r > ℓ then
    i = ℓ; j = r; x = pivot();
    repeat
      while a[i] < x do i++ ; endwhile {Loop I}
      while a[j] > x do j-- ; endwhile {Loop J}
      if i ≤ j then swap(a[i], a[j]);
    until j ≤ i
    quicksort(ℓ, i − 1);
    quicksort(i + 1, r);
  end if
```

Table 1. Number of branch mispredictions

	random pivot	α-skewed pivot
static predictor	$\frac{\ln 2}{2} n \lg n + \mathcal{O}(n)$, $\frac{\ln 2}{2} \approx 0.3466$	$\frac{\alpha}{H(\alpha)} n \lg n + \mathcal{O}(n)$, $\alpha < 1/2$ $\frac{1-\alpha}{H(\alpha)} n \lg n + \mathcal{O}(n)$, $\alpha \geq 1/2$
1-bit predictor	$\frac{2\ln 2}{3} n \lg n + \mathcal{O}(n)$, $\frac{2\ln 2}{3} \approx 0.4621$	$\frac{2\alpha(1-\alpha)}{H(\alpha)} n \lg n + \mathcal{O}(n)$
2-bit predictor	$\frac{28\ln 2}{45} n \lg n + \mathcal{O}(n)$, $\frac{28\ln 2}{45} \approx 0.4313$	$\frac{2\alpha^4-4\alpha^3+\alpha^2+\alpha}{(1-\alpha(1-\alpha))H(\alpha)} n \lg n + \mathcal{O}(n)$

H(α)=-(α·lg(α)+(1- α) ·lg(1- α)

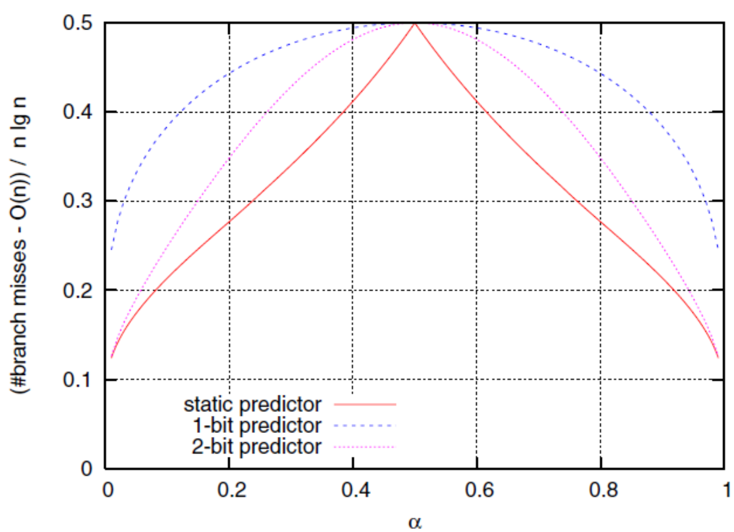


Fig. 2. The α-dependent coefficients of $n \lg n$ for varying α

Gerth Stølting Brodal, Gabriel Moruz.
Tradeoffs Between Branch Mispredictions and Comparisons for Sorting Algorithms.
Proc. 9th International Workshop on Algorithms and Data Structures (WADS),
Lecture Notes in Computer Science, Volume 3608, 385-395, Springer, 2005.

Thm O($d \cdot n \cdot \log n$) comparisons $\Rightarrow$ $\Omega(n \cdot \log_d n)$ mispredictions

Assumptions:

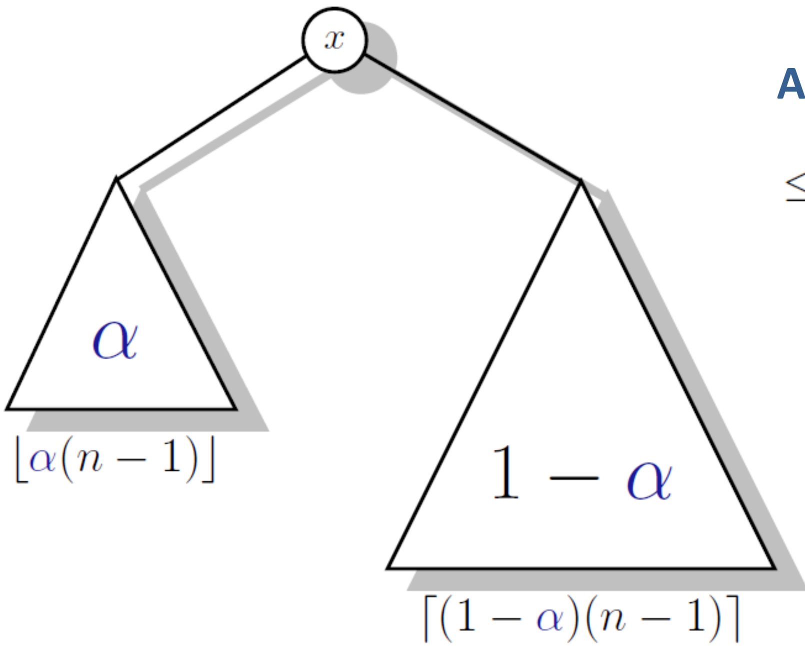
- Deterministic comparison based sorting algorithm
- Each comparison followed by a conditional branch

Proof (sketch)

- Consider decision tree
- Label edges with “0” or ”1”, “1” if branch misprediction
- $\leq k$ mispredictions $\Rightarrow$ all paths $\leq k$ edges labeled “1”
- Depth D decision tree $\Rightarrow$ #leaves $\leq \binom{D+k}{k}$
- Theorem follows from constraint $\binom{D+k}{k} \geq n!$

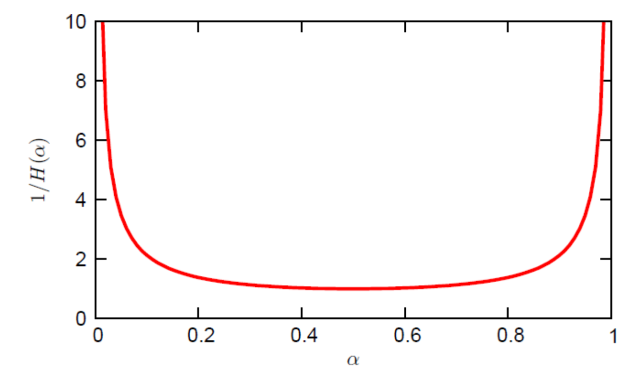
Upper bound: d-way MergeSort

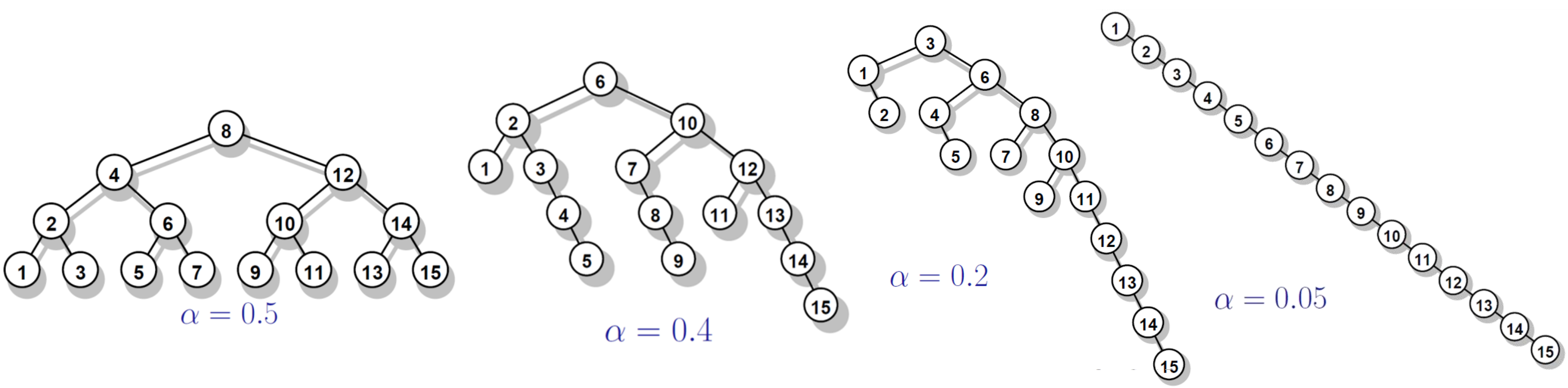
Gerth Stølting Brodal, Gabriel Moruz. ***Skewed Binary Search Trees***.
In Proc. 14th Annual European Symposium on Algorithms (ESA),
Lecture Notes in Computer Science, Volume 4168, 708-719, Springer, 2006.

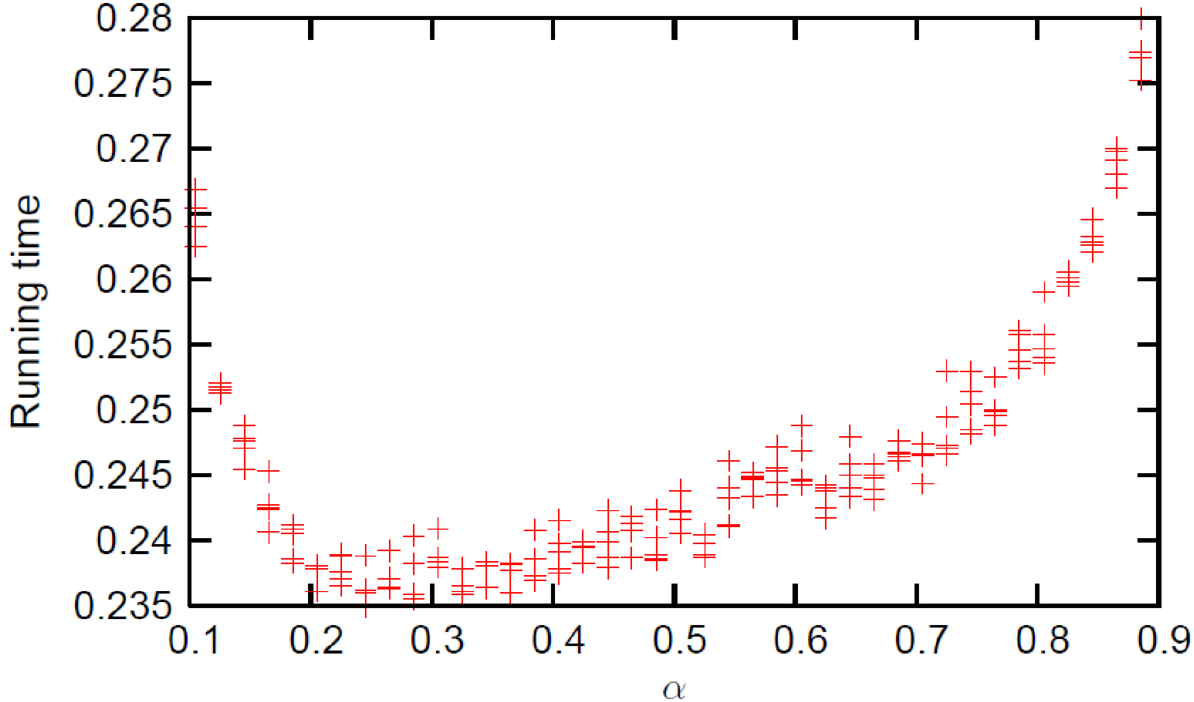


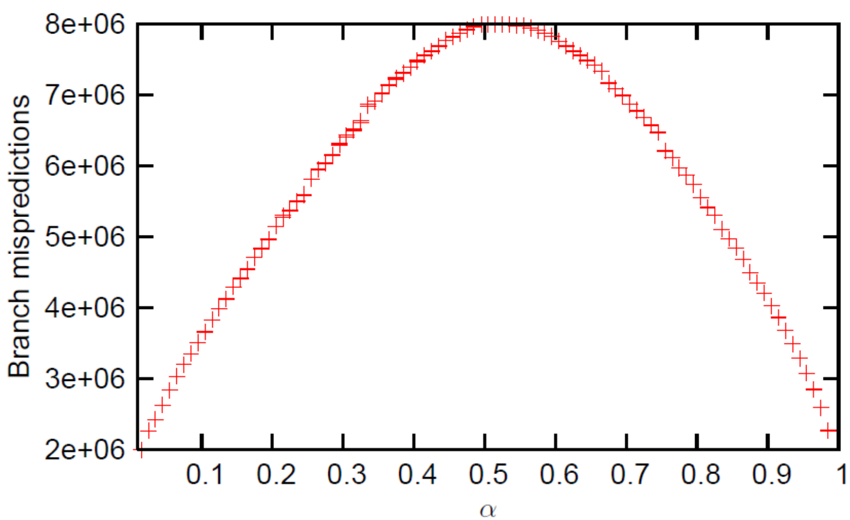
Average node depth

$$\leq \frac{1}{\underbrace{-\alpha \log_2 \alpha - (1-\alpha)\log_2(1-\alpha)}_{H(\alpha)}} \cdot \frac{n+1}{n} \cdot \log_2(n+1) - 2$$

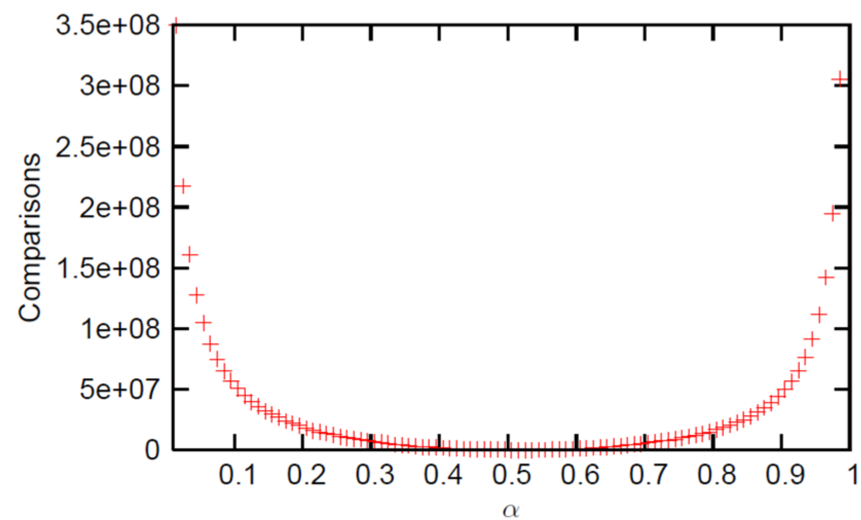






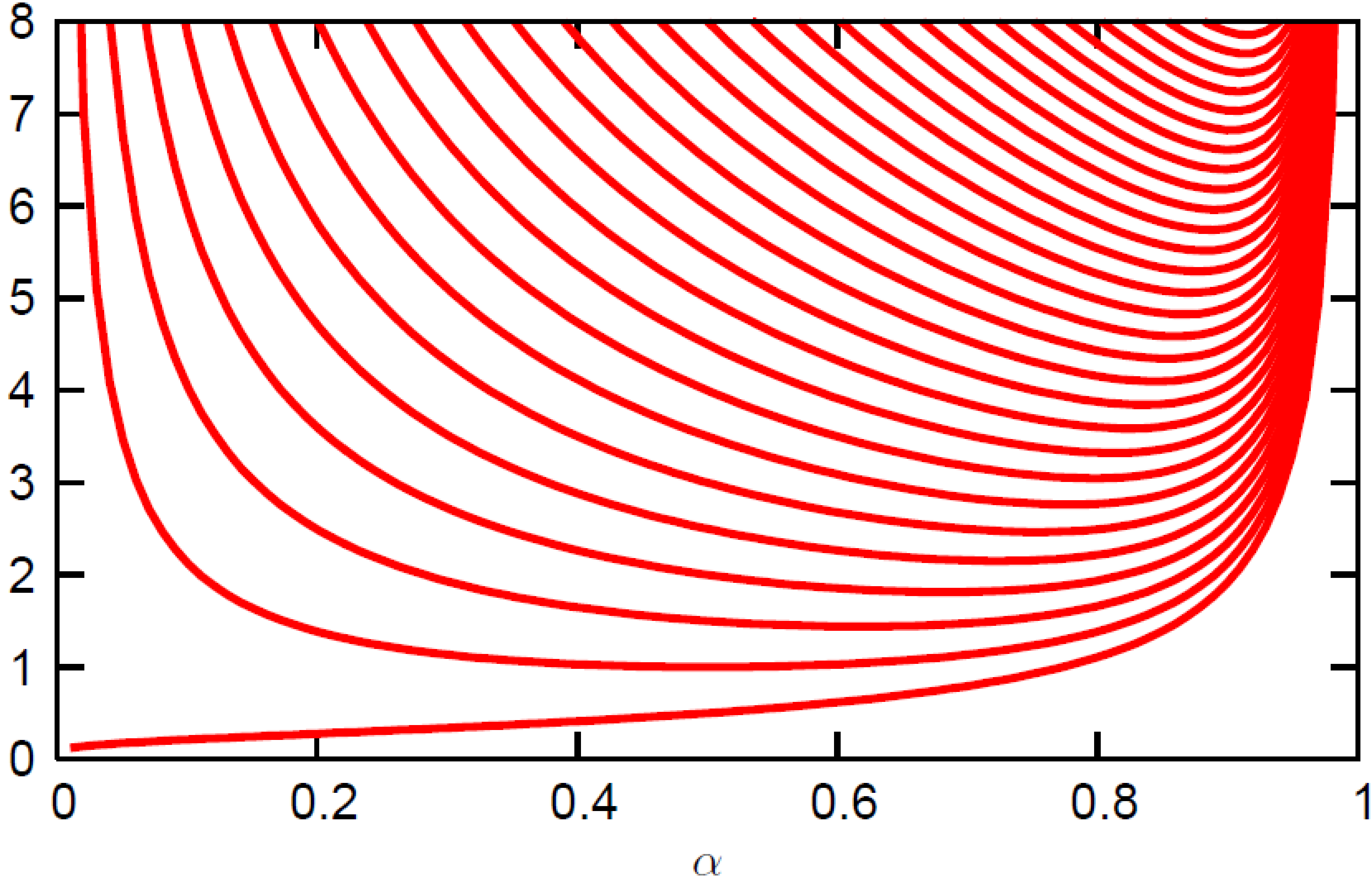


$n = 50.000$



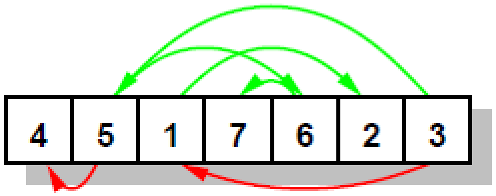
$n = 50.000$

cost$(\alpha) = (\alpha \cdot$ {left cost} $+ (1-\alpha) \cdot$ {right cost}$)/H(\alpha)$

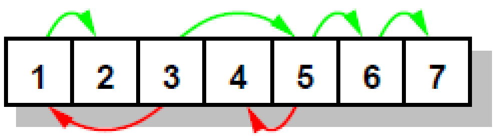


left cost = 1 and right cost = 0 .. 28

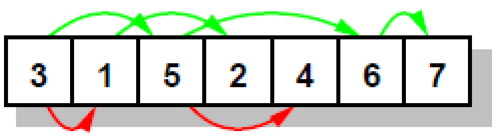
Memory Layouts



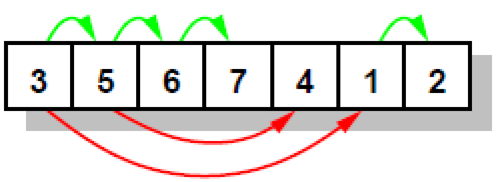
Random – $O(\frac{\log n}{H(\alpha)})$ I/Os



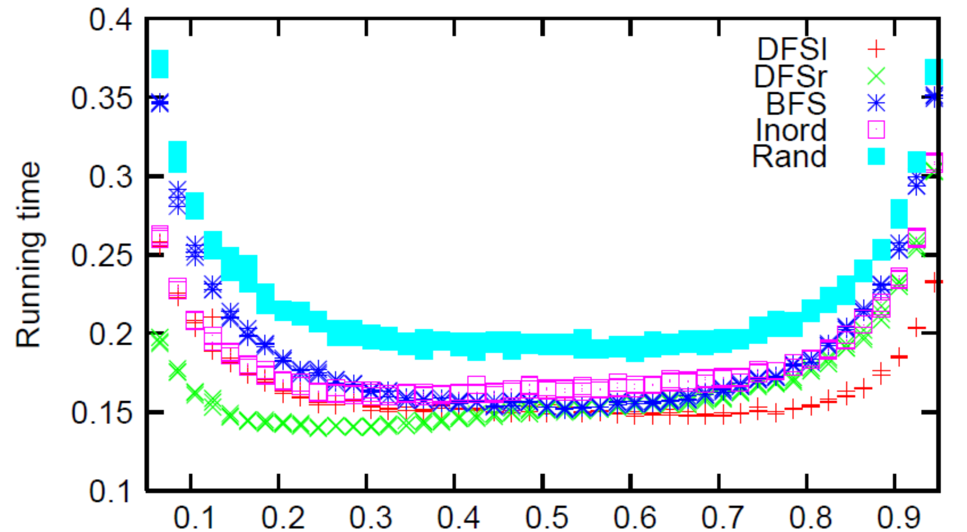
Inorder – $O(\frac{\log n}{H(\alpha)} - \log B)$ I/Os

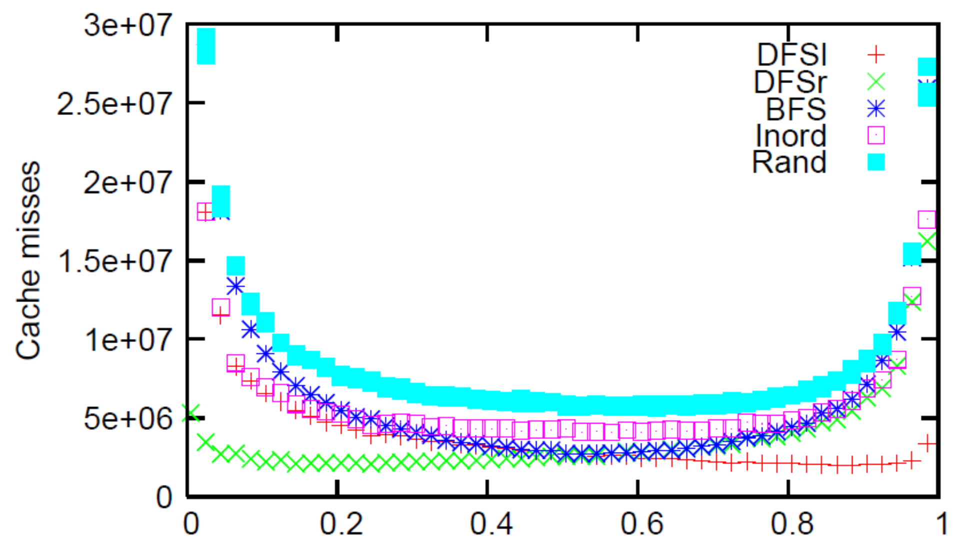


BFS – $O(\frac{\log n}{H(\alpha)} - \log B)$ I/Os



DFSr – $O(\frac{\alpha+(1-\alpha)/B}{H(\alpha)} \cdot \log n)$ I/Os.

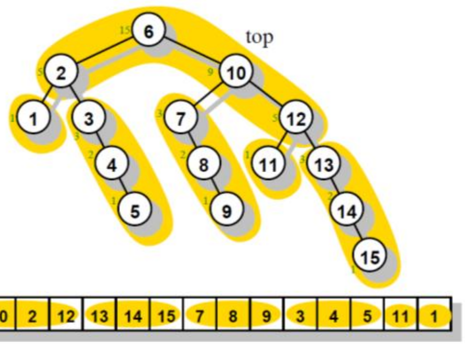




Blocked Memory Layouts

vEB

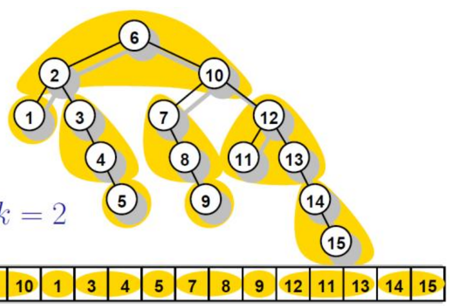
$O(\log_{B\alpha+1} n)$



***k*-level**

$O(\log_B n / H(\alpha))$

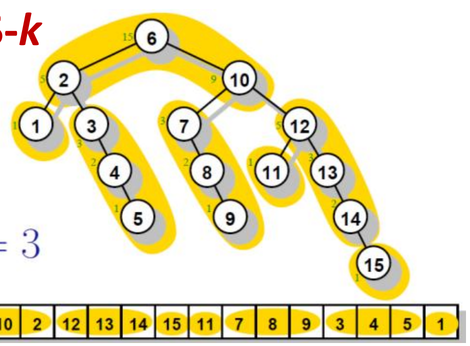
$k = 2$

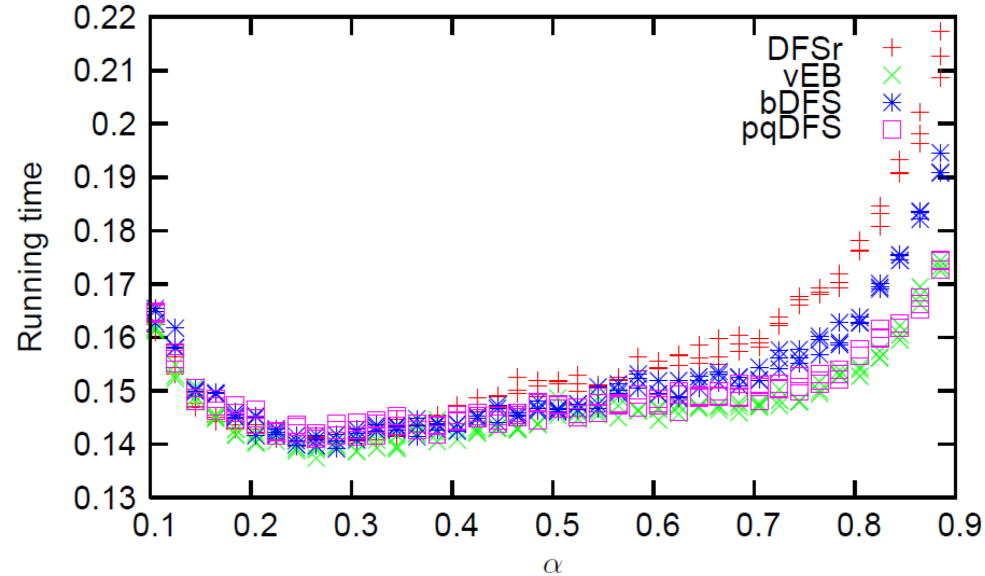


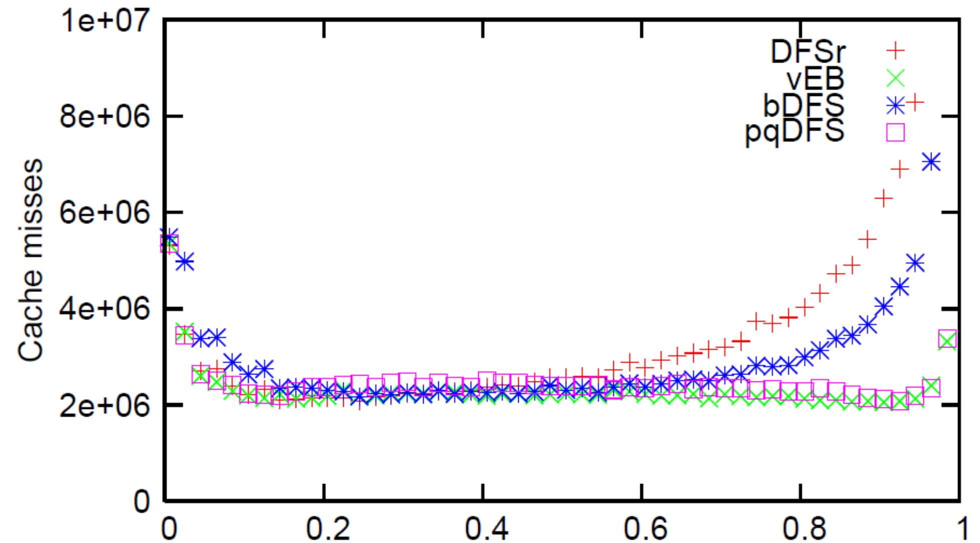
pqDFS-*k*

$O(\log_{B\alpha+1} n)$

$k = 3$

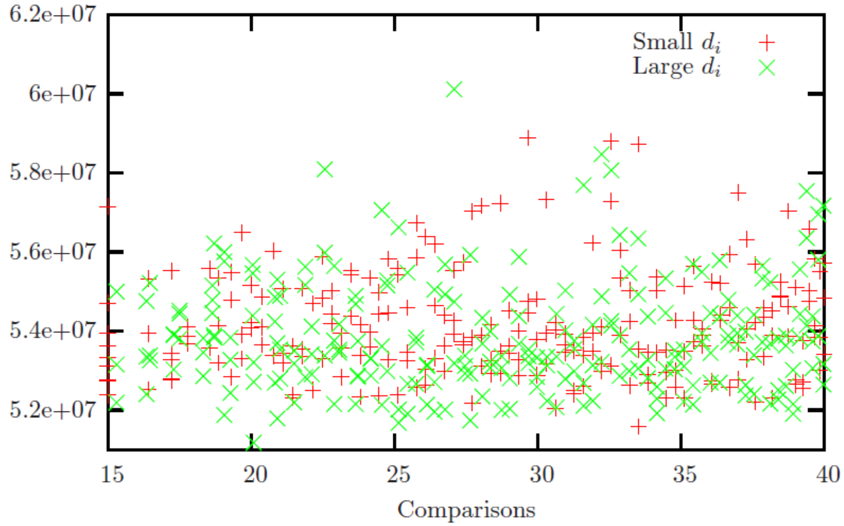


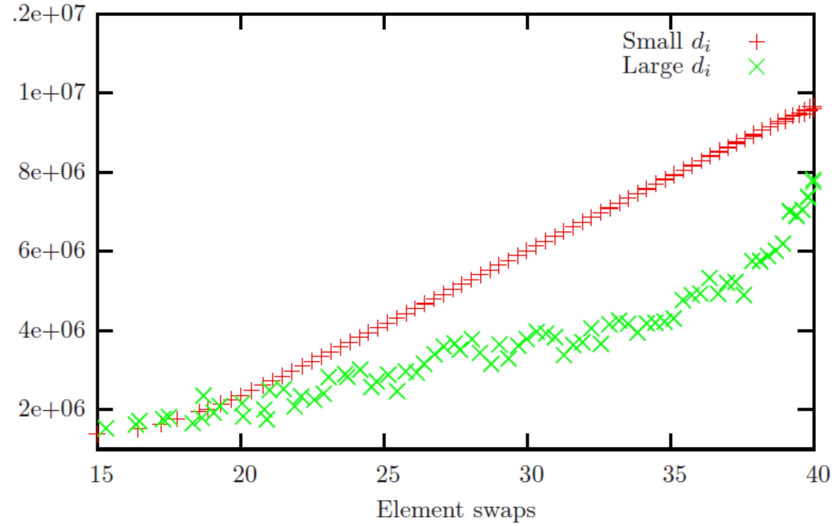


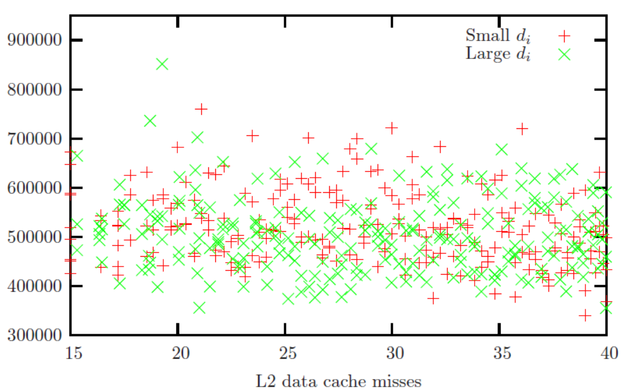


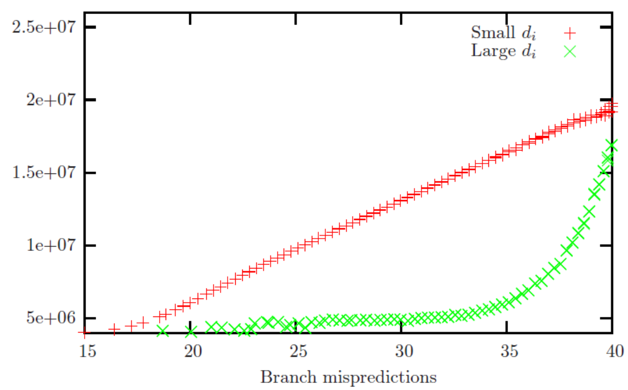
Gerth Stølting Brodal, Rolf Fagerberg, Gabriel Moruz. ***On the Adaptiveness of Quicksort***.
In Proc. 7th Workshop on Algorithm Engineering and Experiments (ALENEX), pages 130-140, 2005

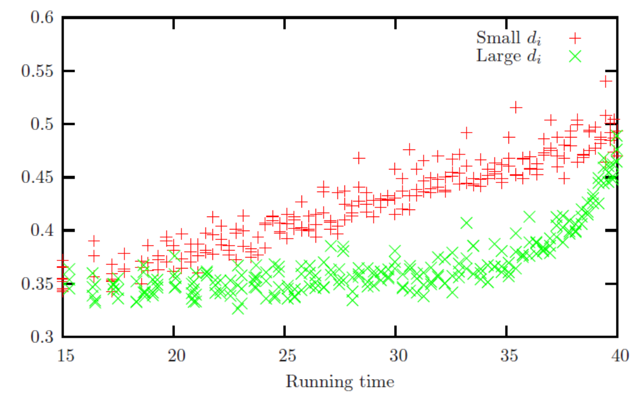
Thm Expected number of element swaps performed by randomized Quicksort is at most $n \cdot \ln(1 + 2 \cdot \text{INV}/n)$



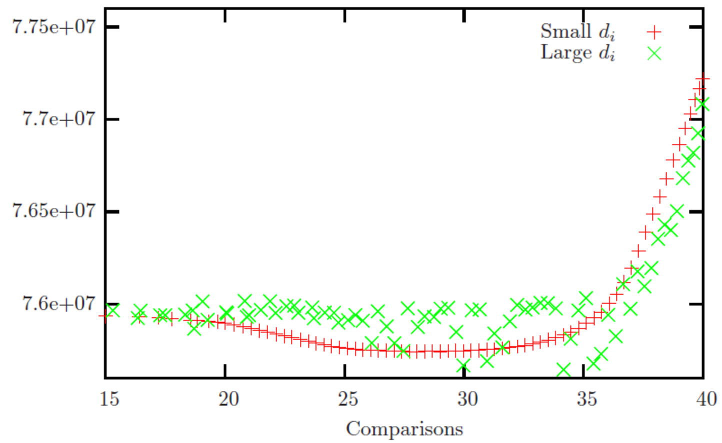


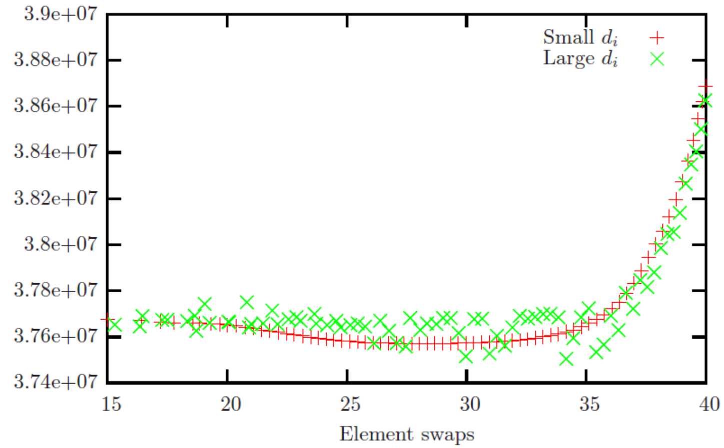


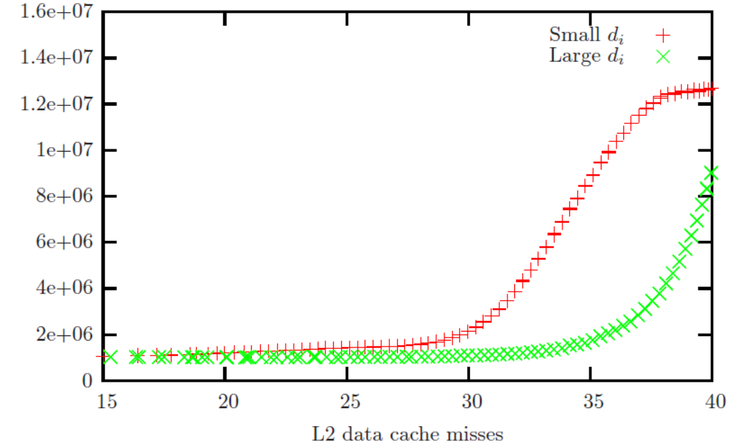


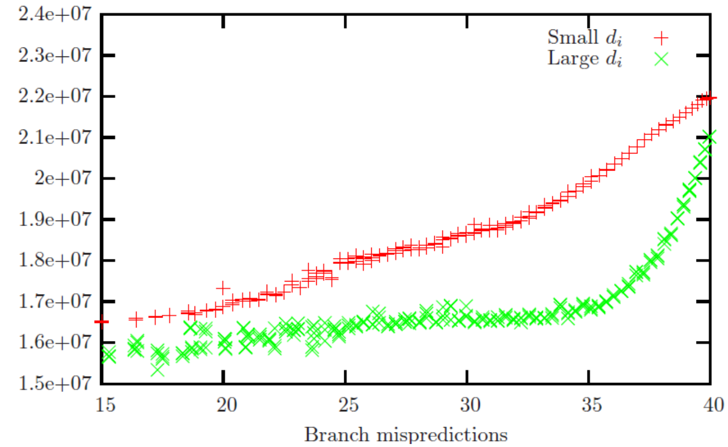


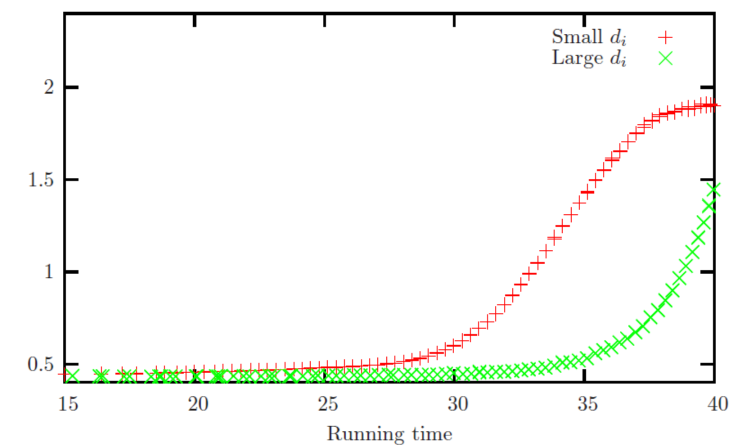
HeapSort


Small d_i
Large d_i
7.75e+07
7.7e+07
7.65e+07
7.6e+07
15
20
25
30
35
40
Comparisons

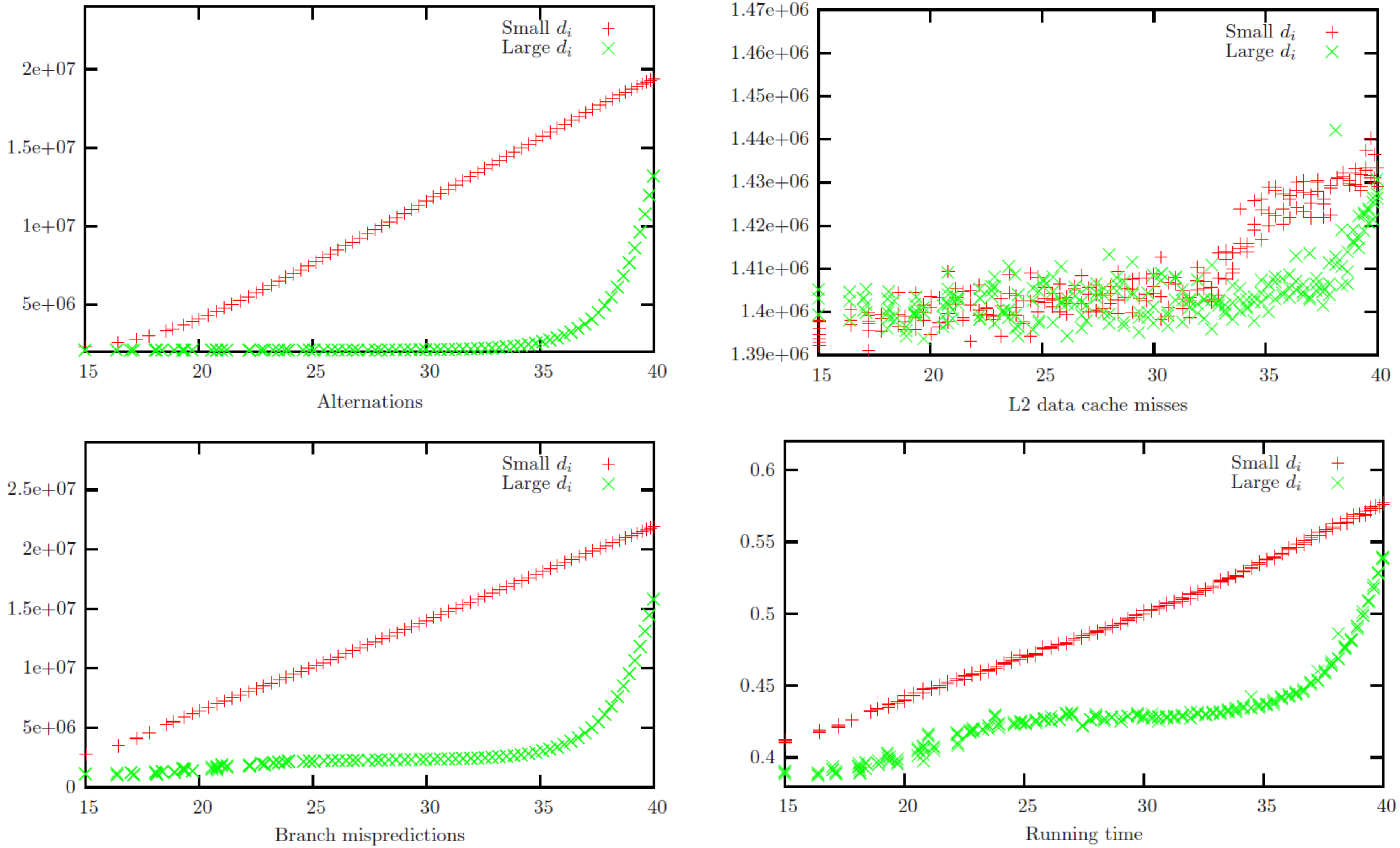

Small d_i
Large d_i
3.9e+07
3.88e+07
3.86e+07
3.84e+07
3.82e+07
3.8e+07
3.78e+07
3.76e+07
3.74e+07
15
20
25
30
35
40
Element swaps


Small d_i
Large d_i
1.6e+07
1.4e+07
1.2e+07
1e+07
8e+06
6e+06
4e+06
2e+06
0
15
20
25
30
35
40
L2 data cache misses


Small d_i
Large d_i
2.4e+07
2.3e+07
2.2e+07
2.1e+07
2e+07
1.9e+07
1.8e+07
1.7e+07
1.6e+07
1.5e+07
15
20
25
30
35
40
Branch mispredictions


Small d_i
Large d_i
2
1.5
1
0.5
15
20
25
30
35
40
Running time

MergeSort



Small d_i
Large d_i
Alternations
L2 data cache misses
Branch mispredictions
Running time