Worst-Case Efficient Sorting with QuickMergesort

Stefan Edelkamp¹ and <u>Armin Weiß²</u>

¹King's College London, UK

²FMI, Universität Stuttgart, Germany

San Diego, January 7, 2019

Comparison-based sorting: Quicksort, Heapsort, Mergesort

Comparison-based sorting: Quicksort, Heapsort, Mergesort



Running times (divided by $n \log n$) for sorting integers. Left: random inputs.

Comparison-based sorting: Quicksort, Heapsort, Mergesort



Running times (divided by $n \log n$) for sorting integers. Left: random inputs. Right: random with large elements in the middle and end.

Quicksort, Heapsort, Mergesort

Algorithm	Fast on average	"in place"	$\mathcal{O}(n \log n)$ worst case
Quicksort	\checkmark	1	×
Heapsort	×	1	1
Mergesort	\checkmark	×	\checkmark

Algorithm	Fast on average	"in place"	$\mathcal{O}(n \log n)$ worst case
Quicksort	\checkmark	1	×
Heapsort	×	1	1
Mergesort	\checkmark	×	✓

Algorithm	Fast on average	"in place"	$\mathcal{O}(n \log n)$ worst case
Quicksort	\checkmark	1	×
Heapsort	×	1	1
Mergesort	1	×	1

• Make Quicksort worst-case efficient: Introsort, median-of-medians pivot selection

Algorithm	Fast on average	"in place"	$\mathcal{O}(n \log n)$ worst case
Quicksort	\checkmark	1	×
Heapsort	×	1	1
Mergesort	\checkmark	×	✓

- Make Quicksort worst-case efficient: Introsort, median-of-medians pivot selection
- Make Heapsort fast: Bottom-up Heapsort (not very fast)

Algorithm	Fast on average	"in place"	$\mathcal{O}(n \log n)$ worst case
Quicksort	1	1	×
Heapsort	×	1	1
Mergesort	1	×	1

- Make Quicksort worst-case efficient: Introsort, median-of-medians pivot selection
- Make Heapsort fast: Bottom-up Heapsort (not very fast)
- Make Mergesort in-place:
 - block based merging (stable implementations: Grailsort, Wikisort)

Algorithm	Fast on average	"in place"	$\mathcal{O}(n \log n)$ worst case
Quicksort	\checkmark	1	×
Heapsort	×	1	\checkmark
Mergesort	\checkmark	×	\checkmark

- Make Quicksort worst-case efficient: Introsort, median-of-medians pivot selection
- Make Heapsort fast: Bottom-up Heapsort (not very fast)
- Make Mergesort in-place:
 - block based merging (stable implementations: Grailsort, Wikisort)
 - rotation based merging (stable, but $\mathcal{O}(n \log^2 n)$)

Algorithm	Fast on average	"in place"	$\mathcal{O}(n \log n)$ worst case
Quicksort	\checkmark	1	×
Heapsort	×	1	1
Mergesort	\checkmark	×	1

- Make Quicksort worst-case efficient: Introsort, median-of-medians pivot selection
- Make Heapsort fast: Bottom-up Heapsort (not very fast)
- Make Mergesort in-place:
 - block based merging (stable implementations: Grailsort, Wikisort)
 - rotation based merging (stable, but $O(n \log^2 n)$)
 - use one half as buffer to sort the other half (In-situ Mergesort [Elmasry, Katajainen, Stenmark 2012], unstable)

Outline:

- QuickMergesort
- Our improvements and theoretical bounds
- Experiments

- 1: procedure QUICKSORT($A[\ell, ..., r]$)
- 2: **if** $r > \ell$ **then**
- 3: pivot \leftarrow choosePivot $(A[\ell, \ldots, r])$
- 4: $\operatorname{cut} \leftarrow \operatorname{partition}(A[\ell, \ldots, r], \operatorname{pivot})$
- 5: Quicksort($A[\ell, \ldots, \operatorname{cut} 1]$)
- 6: Quicksort(A[cut,...,r])
- 7: end if
- 8: end procedure

- 1: procedure QUICKSORT($A[\ell, ..., r]$)
- 2: **if** $r > \ell$ **then**
- 3: pivot \leftarrow choosePivot $(A[\ell, \ldots, r])$
- 4: $\operatorname{cut} \leftarrow \operatorname{partition}(A[\ell, \ldots, r], \operatorname{pivot})$
- 5: Quicksort($A[\ell, \ldots, \operatorname{cut} 1]$)
- 6: Quicksort(A[cut,...,r])
- 7: end if
- 8: end procedure

- 1: procedure QUICKSORT($A[\ell, ..., r]$)
- 2: **if** $r > \ell$ **then**
- 3: pivot \leftarrow choosePivot $(A[\ell, \ldots, r])$
- 4: $\operatorname{cut} \leftarrow \operatorname{partition}(A[\ell, \ldots, r], \operatorname{pivot})$
- 5: Quicksort($A[\ell, \ldots, \operatorname{cut} 1]$)
- 6: Quicksort(A[cut,...,r])
- 7: end if
- 8: end procedure
 - After line 4:



- 1: procedure QUICKSORT($A[\ell, ..., r]$)
- 2: **if** $r > \ell$ **then**
- 3: pivot \leftarrow choosePivot $(A[\ell, \ldots, r])$
- 4: $\operatorname{cut} \leftarrow \operatorname{partition}(A[\ell, \ldots, r], \operatorname{pivot})$
- 5: Quicksort($A[\ell, \ldots, \operatorname{cut} 1]$)
- 6: Quicksort $(A[\operatorname{cut},\ldots,r])$
- 7: end if
- 8: end procedure
 - After line 4:



• After line 5:



- 1: procedure QUICKSORT($A[\ell, ..., r]$)
- 2: **if** $r > \ell$ **then**
- 3: pivot \leftarrow choosePivot $(A[\ell, \ldots, r])$
- 4: $\operatorname{cut} \leftarrow \operatorname{partition}(A[\ell, \ldots, r], \operatorname{pivot})$
- 5: Quicksort($A[\ell, \ldots, \operatorname{cut} 1]$)
- 6: Quicksort(A[cut,...,r])
- 7: end if
- 8: end procedure
 - After line 4:



• After line 5:



• After line 6: both parts sorted recursively with Quicksort



- 1: procedure QUICKSORT($A[\ell, ..., r]$)
- 2: **if** $r > \ell$ **then**
- 3: pivot \leftarrow choosePivot $(A[\ell, \ldots, r])$
- 4: $\operatorname{cut} \leftarrow \operatorname{partition}(A[\ell, \ldots, r], \operatorname{pivot})$
- 5: Quicksort($A[\ell, \ldots, \operatorname{cut} 1]$)
- 6: Quicksort(A[cut,...,r])
- 7: end if
- 8: end procedure

- procedure QUICKSORT($A[\ell, ..., r]$) 1:
- 2: if $r > \ell$ then
- pivot \leftarrow choosePivot $(A[\ell, \ldots, r])$ 3:
- $\underbrace{\mathsf{Mergesort}}_{\mathsf{Quicksort}} (A[\ell, \dots, r], \mathsf{pivot})$ 4:
- 5:
- Quicksort(A[cut, ..., r])6:
- 7: end if
- end procedure 8:

- 1: procedure QUICKMERGESORT($A[\ell, ..., r]$)
- 2: **if** $r > \ell$ **then**
- 3: pivot \leftarrow choosePivot $(A[\ell, \ldots, r])$
- 4: $\operatorname{cut} \leftarrow \operatorname{partition}(A[\ell, \ldots, r], \operatorname{pivot})$
- 5: Mergesort($A[\ell, \ldots, \operatorname{cut} 1]$)
- 6: QuickMergesort(A[cut, ..., r])
- 7: end if
- 8: end procedure

- 1: procedure QUICKMERGESORT($A[\ell, ..., r]$)
- 2: **if** $r > \ell$ **then**
- 3: pivot \leftarrow choosePivot $(A[\ell, \ldots, r])$
- 4: $\operatorname{cut} \leftarrow \operatorname{partition}(A[\ell, \ldots, r], \operatorname{pivot})$
- 5: Mergesort($A[\ell, \ldots, \text{cut} 1]$)
- 6: QuickMergesort(A[cut,...,r])
- 7: end if
- 8: end procedure
 - After line 4:



• After line 5:



• After line 6: both parts sorted recursively with QuickMergesort



- 1. Partition according to some pivot element.
- 2. Sort one part with Mergesort.
- 3. Sort the the other part recursively with QuickMergesort.

- 1. Partition according to some pivot element.
- 2. Sort one part with Mergesort.
- 3. Sort the the other part recursively with QuickMergesort.

- 1. Partition according to some pivot element.
- 2. Sort one part with Mergesort.
- 3. Sort the the other part recursively with QuickMergesort.

- 1. Partition according to some pivot element.
- 2. Sort one part with Mergesort.
- 3. Sort the the other part recursively with QuickMergesort.

- 1. Partition according to some pivot element.
- 2. Sort one part with Mergesort.
- 3. Sort the the other part recursively with QuickMergesort.



- 1. Partition according to some pivot element.
- 2. Sort one part with Mergesort.
- 3. Sort the the other part recursively with QuickMergesort.



- 1. Partition according to some pivot element.
- 2. Sort one part with Mergesort.
- 3. Sort the the other part recursively with QuickMergesort.



- 1. Partition according to some pivot element.
- 2. Sort one part with Mergesort.
- 3. Sort the the other part recursively with QuickMergesort.



- 1. Partition according to some pivot element.
- 2. Sort one part with Mergesort.
- 3. Sort the the other part recursively with QuickMergesort.



- 1. Partition according to some pivot element.
- 2. Sort one part with Mergesort.
- 3. Sort the the other part recursively with QuickMergesort.



- 1. Partition according to some pivot element.
- 2. Sort one part with Mergesort.
- 3. Sort the the other part recursively with QuickMergesort.



- 1. Partition according to some pivot element.
- 2. Sort one part with Mergesort.
- 3. Sort the the other part recursively with QuickMergesort.



- 1. Partition according to some pivot element.
- 2. Sort one part with Mergesort.
- 3. Sort the the other part recursively with QuickMergesort.



- 1. Partition according to some pivot element.
- 2. Sort one part with Mergesort.
- 3. Sort the the other part recursively with QuickMergesort.



- 1. Partition according to some pivot element.
- 2. Sort one part with Mergesort.
- 3. Sort the the other part recursively with QuickMergesort.


- 1. Partition according to some pivot element.
- 2. Sort one part with Mergesort.
- 3. Sort the the other part recursively with QuickMergesort.



- 1. Partition according to some pivot element.
- 2. Sort one part with Mergesort.
- 3. Sort the the other part recursively with QuickMergesort.



QuickMergesort needs $\leq n \log n - 0.83n$ (resp. $n \log n - 1.24n$) +o(n) comparisons on average with median-of-three (resp. median of \sqrt{n}).

QuickMergesort needs $\leq n \log n - 0.83n$ (resp. $n \log n - 1.24n$) +o(n) comparisons on average with median-of-three (resp. median of \sqrt{n}).

Still quadratic worst-case (with median-of-three)! 😀

QuickMergesort needs $\leq n \log n - 0.83n$ (resp. $n \log n - 1.24n$) +o(n) comparisons on average with median-of-three (resp. median of \sqrt{n}).

Still quadratic worst-case (with median-of-three)! 😀

Solution: choose the median as pivot:

QuickMergesort needs $\leq n \log n - 0.83n$ (resp. $n \log n - 1.24n$) +o(n) comparisons on average with median-of-three (resp. median of \sqrt{n}).

Still quadratic worst-case (with median-of-three)! 😅

Solution: choose the median as pivot:

• Quickselect resp. Introselect.

QuickMergesort needs $\leq n \log n - 0.83n$ (resp. $n \log n - 1.24n$) +o(n) comparisons on average with median-of-three (resp. median of \sqrt{n}).

Still quadratic worst-case (with median-of-three)! 😅

Solution: choose the median as pivot:

- Quickselect resp. Introselect.
 - → In-situ Mergesort [Elmasry, Katajainen, Stenmark 2012]

QuickMergesort needs $\leq n \log n - 0.83n$ (resp. $n \log n - 1.24n$) +o(n) comparisons on average with median-of-three (resp. median of \sqrt{n}).

Still quadratic worst-case (with median-of-three)! 😅

Solution: choose the median as pivot:

Quickselect resp. Introselect. Problem: worst case
 → In-situ Mergesort [Elmasry, Katajainen, Stenmark 2012]

QuickMergesort needs $\leq n \log n - 0.83n$ (resp. $n \log n - 1.24n$) +o(n) comparisons on average with median-of-three (resp. median of \sqrt{n}).

Still quadratic worst-case (with median-of-three)! 😅

Solution: choose the median as pivot:

- Quickselect resp. Introselect. Problem: worst case
 → In-situ Mergesort [Elmasry, Katajainen, Stenmark 2012]
- Median-of-medians algorithm [Blum, Floyd, Pratt, Rivest, Tarjan 1973]

QuickMergesort needs $\leq n \log n - 0.83n$ (resp. $n \log n - 1.24n$) +o(n) comparisons on average with median-of-three (resp. median of \sqrt{n}).

Still quadratic worst-case (with median-of-three)! 😅

Solution: choose the median as pivot:

- Quickselect resp. Introselect. Problem: worst case
 ~> In-situ Mergesort [Elmasry, Katajainen, Stenmark 2012]
- Median-of-medians algorithm [Blum, Floyd, Pratt, Rivest, Tarjan 1973]

Theorem (Folklore)

The median-of-medians algorithms needs at most 20n + o(n) comparisons to find the median of n elements.

QuickMergesort needs $\leq n \log n - 0.83n$ (resp. $n \log n - 1.24n$) +o(n) comparisons on average with median-of-three (resp. median of \sqrt{n}).

Still quadratic worst-case (with median-of-three)! 😅

Solution: choose the median as pivot:

- Quickselect resp. Introselect. Problem: worst case
 ~> In-situ Mergesort [Elmasry, Katajainen, Stenmark 2012]
- Median-of-medians algorithm [Blum, Floyd, Pratt, Rivest, Tarjan 1973]

Theorem (Folklore)

The median-of-medians algorithms needs at most 20n + o(n) comparisons to find the median of n elements.

Need to find the median of *n*, $\frac{n}{2}$, $\frac{n}{4}$,... elements $\rightarrow 40n$ comparisons

Key observation: we do not need the exact median.

Key observation: we do not need the exact median.

Sufficient if one third is smaller/greater than the pivot:

Key observation: we do not need the exact median.

Sufficient if one third is smaller/greater than the pivot:

- form groups of 3 elements
- compute the median of each group
- compute the median of the n/3 medians



Key observation: we do not need the exact median.

Sufficient if one third is smaller/greater than the pivot:

- form groups of 3 elements
- compute the median of each group
- compute the median of the n/3 medians



Key observation: we do not need the exact median.

Sufficient if one third is smaller/greater than the pivot:

- form groups of 3 elements
- compute the median of each group
- compute the median of the n/3 medians



Theorem

Basic MoMQuickMergesort needs at most $n \log n + 13.8n + o(n)$ comparisons.



• Step 1 (merge from the left):

• Step 1 (merge from the left):



• Once the buffer is full, the final position for the largest element is "free".



• Expected result:

• Step 1 (merge from the left):



- Once the buffer is full, the final position for the largest element is "free".
- Step 2 (merge from the right):
- Result:



• Step 1 (merge from the left):



- Once the buffer is full, the final position for the largest element is "free".
- Step 2 (merge from the right):
- Result:



Need a guarantee that one fifth are smaller/greater than the pivot:

• Step 1 (merge from the left):



- Once the buffer is full, the final position for the largest element is "free".
- Step 2 (merge from the right):
- Result:



 M_1 M_2 M_3 M_4 M_5 M_6 \cdots M_ℓ M_1' M_2' \cdots M_ℓ' \sim median of pseudomedians of 15 elements

- Step 1 (merge from the left):
- Once the buffer is full, the final position for the largest element is "free".
- Step 2 (merge from the right):
- Result:

Need a guarantee that one fifth are smaller/greater than the pivot:

 \rightsquigarrow median of pseudomedians of 15 elements

Theorem

MoMQuickMergesort needs at most $n \log n + 4.57n + o(n)$ comparisons.







For merging sequences of different size a smaller buffer suffices:

• Step 1 (merge from the left):



For merging sequences of different size a smaller buffer suffices:

• Step 1 (merge from the left):



• Once the buffer is full, the final position for the largest element is "free".

For merging sequences of different size a smaller buffer suffices:

• Step 1 (merge from the left):



- Once the buffer is full, the final position for the largest element is "free".
- Step 2 (merge from the right):



For merging sequences of different size a smaller buffer suffices:

• Step 1 (merge from the left):



- Once the buffer is full, the final position for the largest element is "free".
- Step 2 (merge from the right):
- Result:





For merging sequences of different size a smaller buffer suffices:

• Step 1 (merge from the left):



- Once the buffer is full, the final position for the largest element is "free".
- Step 2 (merge from the right):
- Result:



~ trade-off: effort to find good pivots vs. increased merging costs

For merging sequences of different size a smaller buffer suffices:

- Step 1 (merge from the left):
- Once the buffer is full, the final position for the largest element is "free".
- Step 2 (merge from the right):
- Result:



~ trade-off: effort to find good pivots vs. increased merging costs

Undersampling: for $\theta \ge 1$ apply the median-of-pseudomedians-of-15 strategy to n/θ elements.



For merging sequences of different size a smaller buffer suffices:

- Step 1 (merge from the left):
- Once the buffer is full, the final position for the largest element is "free".
- Step 2 (merge from the right):
- Result:



 \rightsquigarrow trade-off: effort to find good pivots vs. increased merging costs

Undersampling: for $\theta \ge 1$ apply the median-of-pseudomedians-of-15 strategy to n/θ elements.

Theorem

MoMQuickMergesort with undersampling factor $\theta = 2.2$ needs at most $n \log n + 1.59n + o(n)$ comparisons.



• Experiments with random permutations of 32-bit integers (other data types in proceedings) in C++.

- Experiments with random permutations of 32-bit integers (other data types in proceedings) in C++.
- Not clear how to find worst-case instances.

- Experiments with random permutations of 32-bit integers (other data types in proceedings) in C++.
- Not clear how to find worst-case instances.
- Simulation of worst cases:

- Experiments with random permutations of 32-bit integers (other data types in proceedings) in C++.
- Not clear how to find worst-case instances.
- Simulation of worst cases:
 - Discard the computed pivots and compute worst-case pivots using Quickselect.

- Experiments with random permutations of 32-bit integers (other data types in proceedings) in C++.
- Not clear how to find worst-case instances.
- Simulation of worst cases:
 - Discard the computed pivots and compute worst-case pivots using Quickselect.
 - Minor details (e.g. random shuffle before Mergesort).

Counting comparisons

Algorithm	average case		worst case	
	exp.	theo.	exp.	theo.
bMQMS	2.772 ± 0.02			
MQMS	2.084 ± 0.001			
$MQMS_{11/5}$	0.246 ± 0.01			



Number of comparisons (linear term) of MoMQuickMergesort variants

Counting comparisons

Algorithm	average case		worst case	
	exp.	theo.	exp.	theo.
bMQMS	2.772 ± 0.02	-		
MQMS	2.084 ± 0.001	2.094		
$MQMS_{11/5}$	0.246 ± 0.01	0.275		



Number of comparisons (linear term) of MoMQuickMergesort variants
Counting comparisons

Algorithm	average case		worst case	
	exp.	theo.	exp.	theo.
bMQMS	2.772 ± 0.02	-	$13.05\ \pm 0.17$	13.8
MQMS	2.084 ± 0.001	2.094	4.220 ± 0.007	4.57
$\mathrm{MQMS}_{11/5}$	$\textbf{0.246} \pm 0.01$	0.275	1.218 ± 0.011	1.59



Number of comparisons (linear term) of MoMQuickMergesort variants and simulated worst cases.



Running times of different MoMQuickMergesort variants and their simulated worst cases for random permutations of 32-bit integers.

Running times



Running times for random permutations of 32-bit integers.

Running times



Running times for sorting integers.

Left: random inputs.

Right: Random with large elements in the middle and end.

Running times



Running times for sorting integers.

Left: random inputs.

Right: Random with large elements in the middle and end.

Conclusion

Algorithm	Fast on average	"in place"	$\mathcal{O}(n \log n)$ worst case
Quicksort	 Image: A set of the set of the	1	×
Heapsort	×	1	1
Mergesort	1	×	1
MoMQuickMergesort	 Image: A set of the set of the	1	\checkmark

¹Code available at https://github.com/weissan/QuickXsort

Algorithm	Fast on average	"in place"	$\mathcal{O}(n \log n)$ worst case
Quicksort	 Image: A set of the set of the	1	×
Heapsort	×	1	1
Mergesort	1	×	1
MoMQuickMergesort	1	1	1

 $\bullet \ \mathrm{MQMS}_{11/5}$ is an unstable sorting algorithm with

- $n \log n + 1.59n + o(n)$ comparisons in the worst case
- $n \log n + 0.275n + o(n)$ comparisons in the average case
- Implementation with stl-style interface¹.

¹Code available at https://github.com/weissan/QuickXsort

Algorithm	Fast on average	"in place"	$\mathcal{O}(n \log n)$ worst case
Quicksort	 Image: A set of the set of the	1	×
Heapsort	×	1	1
Mergesort	1	×	1
MoMQuickMergesort	1	1	1

 $\bullet \ \mathrm{MQMS}_{11/5}$ is an unstable sorting algorithm with

- $n \log n + 1.59n + o(n)$ comparisons in the worst case
- $n \log n + 0.275n + o(n)$ comparisons in the average case
- Implementation with stl-style interface¹.

Thank you!

¹Code available at https://github.com/weissan/QuickXsort

Experiments with larger records



Running times of MoMQuickMergesort (average and simulated worst case), hybrid QMS and other algorithms for random permutations 44-byte records with 4-byte keys.

Experiments sorting pointers



Running times of MoMQuickMergesort (average and simulated worst case), hybrid QMS and other algorithms for random permutations of pointers to records.

- Experiments with random permutations of 32bit integers (other data types in proceedings).
- running time and comparison count
- ullet \geq 100 measurements for each data point
- Test environment:
 - Intel Core i5-2500K CPU (3.30GHz) with 16GB RAM
 - Ubuntu Linux 64bit version 14.04.4
 - g++ (4.8.4) compiler with flags -O3 -march=native