Virtual-Memory Powersort

Finn Moltmann
Philipps-Universität Marburg, Germany

Tamio-Vesa Nakajima
Philipps-Universität Marburg, Germany

Sebastian Wild
Philipps-Universität Marburg, Germany
University of Liverpool, UK

Abstract

We give a more space-efficient implementation of adaptive mergesort: Virtual-Memory Powersort. Using internal buffering techniques, we significantly reduce the memory consumption of the algorithm; specifically, for sorting n objects the required buffer area is reduced from space for $n/2$ objects to $O(\sqrt{n \log n})$ objects. While this space-efficiency can be achieved (indeed reduced to $O(1)$) conceptually very easily with known inplace merging algorithms, using these as a drop-in replacement for the standard merge algorithm incurs a substantial slow-down. Virtual-Memory Powersort, by contrast, uses the *same* number of moves and comparisons as previous Powersort implementations up to an additive $O(n)$ term. We report on an empirical running-time study comparing our implementation against other Powersort variants and state-of-the-art stable sorting methods, demonstrating that *almost* in-place stable sorting can be achieved with negligible overhead in many scenarios.

2012 ACM Subject Classification Theory of computation → Data structures design and analysis

Keywords and phrases adaptive sorting, inplace sorting, inplace merging, library sort, virtual memory, internal buffering, Powersort, Timsort

Digital Object Identifier 10.4230/LIPIcs.ESA.2026.14

Related Version *Previous Version*: `https://arxiv.org/abs/2605.27147`

Supplementary Material *Software (Source Code)*: `https://git.sr.ht/~tamionv/vm-powersort` [24] archived at `swh:1:snp:8e65d4e51fc7e84ac78dde26a19bea76773d8ce8`

Funding *Sebastian Wild*: Supported by EPSRC grant EP/X039447/2.

Acknowledgements Note: Where we were aware of both a conference and a full paper, we cite both to give credit to both the definitive version and the venue of first dissemination.

1 Introduction

Sorting is an elementary building block that most standard libraries offer an implementation of. For general purpose methods, the specification often prescribes a *stable* sorting method, i.e., one where elements that compare equal w.r.t. the sorting criteria remain in the same relative order as in the input. Apart from stability, the desiderata for a library sort are to be fast and to use little (or no) extra memory. Unfortunately, these three properties seem to be at odds: while Quicksort is usually the fastest generic option when implemented *in place*,[1] it is not stable. Mergesort is easily implemented as a stable method and can be made almost as

[1] Strictly speaking, "in place" should mean no extra memory at all; usually $O(\log n)$ (words of) space are accepted as "in place". In practice, a stack of $O(\log n)$ pointers – as recursion stack in Quicksort and top-down Mergesort, or as run stack in Timsort/Powersort – has negligible memory footprint and is indeed is often implemented as a static fixed-size array.

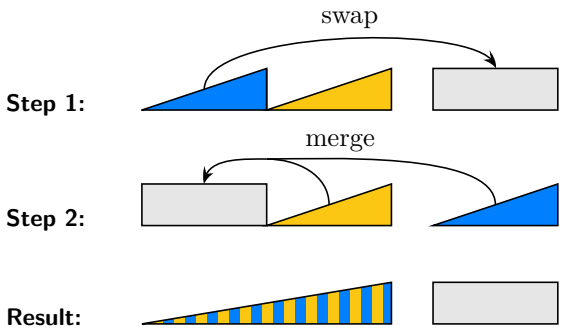


Figure 1 Standard "copy"-merging procedure where one of the two runs fits into the buffer as used in CPython's Powersort implementation. (Buffer and input need not be adjacent in memory.)

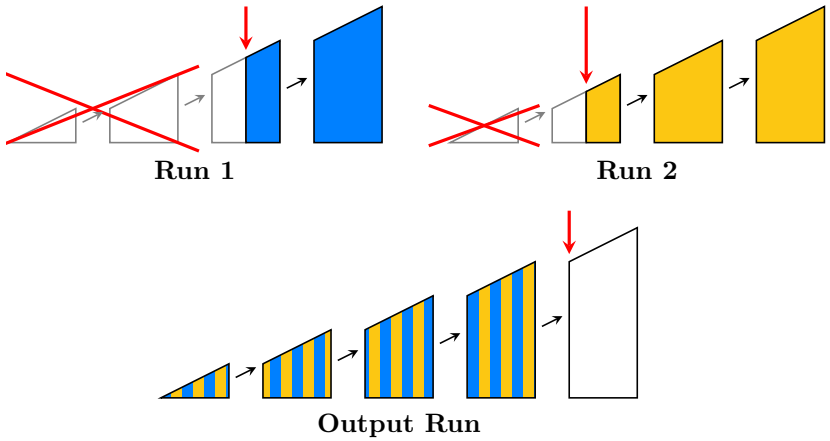


Figure 2 Merging two runs in virtual memory. Both input runs and the output run are contiguous in virtual address space, but physically broken into *pages* which need not be contiguous in physical memory. Whenever a page is entirely consumed form an input run (grayed/crossed-out in the figure), it is released to be reused as a new output page for further merging.

fast as Quicksort, but typically at the expense of a linear-size buffer of extra memory. In-place variants of Mergesort are known, but they either sacrifice stability [18, 15, 30, 4, 32, 5] or are substantially more complicated and slower in practice [20, 8].

Orthogonally to the triad of stability, time, and space, standard libraries have increasingly embraced the idea of *adaptive sorting*, where the sorting process becomes faster when the input is already partially ordered. Asymptotically optimal algorithms have been devised for many thinkable measures of presortedness [7, 28, 29], but not always is the overhead worth the potential savings [6]. One exception has been exploiting natural *runs*, i.e., contiguous sorted segments in the input. This approach has found wide-spread use via Timsort [27] and is used for stable sorting in most popular software frameworks (CPython, PyPy, OpenJDK, Android Runtime, Swift, ...). Several of these frameworks have replaced Timsort's original merge policy by *Powersort* [25, 1]. All these Timsort/Powersort implementations still use extra space for $n/2$ objects to facility efficient stable merging; see Figure 1. This limits the use of these successful adaptive stable methods to use cases where ample extra space is available.

We show that a seasoned technique, *virtual memory*,[2] can be efficiently combined with Powersort to yield $O(\sqrt{n \log n})$-extra-space variants with almost no running-time overhead over the copy-based version as used in CPython. The key idea of virtual memory is to break the physical memory (our input array) into logical *pages* (each holding P objects) and to represent contiguous runs by (potentially non-contiguous) sequences of pages. It is a folklore result that merging and Mergesort can be implemented (e.g., in external memory) by reading from and writing to only one block of data from the input resp. output runs. We apply the same method at page granularity. Moreover, whenever a page of an input run has been exhausted, we declare that memory area a *free page*, to be reused as a future output page (see Figure 2). The novelty of our results lies in making this efficient for *internal sorting*.

Merging page by page, yields the output as a *virtual run*: sorted in virtual order, but with pages potentially in arbitrary locations. A permutation step, swapping contents of entire pages, can produce the run as a contiguous block in physical memory and thus turn the virtual-memory merge into a conventional one; however, this permuting step is rather costly. For the final merge this extra cost is inevitable, but for *intermediate* merges we can leave pages permuted: If each page stores the physical location of its successor page, we can sequentially traverse a virtual run almost as efficiently as a run contiguous in physical memory.

Note that by itself, the virtual-memory technique does not yield a fully inplace algorithm, since we need a small number of extra pages to bootstrap the process. Moreover, we need extra space for one pointer per page to store successor pages. In theory, at most 3 extra page are sufficient to handle merges, if we allow runs to start and end mid-page; that however complicates the merging logic substantially. Our second insight is that for Powersort specifically, we can ensure that each written run starts at a page boundary (and ends in a potentially partially used page) if we have space for $O(\log n)$ additional pages: at any point in time, we only have $O(\log n)$ written runs on the run stack of Powersort. These are the only runs that use partial pages; the not-yet-discovered runs in the suffix of the input do not yet start at a page boundary.[3]

Using loop unrolling and a careful implementation of memory accesses, our VM Powersort implementation executes, for most of its operation, exactly the same code as standard Powersort. Indeed, since we typically do *fewer* data movements in VM Powersort, when sorting large objects (that are not too expensive to compare), VM Powersort can be *faster* than a CPython-style Powersort.

Our closer study of Powersort's merging choreography is of interest for another reason: When allowed a size-n buffer, we can implement all but the last merges as an *out-of-place merge*, reading data from one area and writing the output to a different area, avoiding any extraneous copies. (The CPython implementation does such extraneous copies for every merge). For non-adaptive mergesort, this is a well-known folklore trick sometimes called "Ping-Pong Merge" [2], where alternating levels of Mergesort's recursion tree switch the roles of two arrays (read from A, write to B, resp. vice versa). For run-adaptive Mergesort, the "level" of a merge may not be known a priori and hence this trick becomes harder to execute. The specific merge policy of Powersort, however, allows for a *Pingpong Powersort* implementation that avoids all extraneous copies of elements up to *one* move per element.

[2] The algorithmic idea is described, e.g., as "block table" by [33, §5.3] in the context of external sorting. We borrow the name from the "virtual runs" of Graefe [12, §3.4] from sorting in database systems.

[3] We note that the virtual-memory technique is otherwise not specific to Powersort in any way, and could likely be applied to other stack-based adaptive mergesort variants.

Related Work

Making sorting use little extra space has long been of interest, especially in combination with stable sorting. A comprehensive survey is thus not possible here and we restrict ourselves to key results for mergesort. For a pedagogical overview of various inplace merging methods, as well as other inplace sorting algorithms, see Wegner's lecture notes [31].

Interestingly, one can obtain an inplace sorting method without making merging itself inplace; this works via a technique of internal buffer, using part of the input as buffer for merges. This can be exploited via unbalanced merging [18, 15] or via quicksort-style partitioning [30, 4, 32, 5]; neither of these approaches is *stable*, though. Already these methods are classified as "not efficient in practice" by Katajainen et al. [15].

A line of work starting with Kronrod [18] designed merge methods without extra space. The first resp. more practical ones were not stable [19, 14], but various *stable* inplace merging methods have been devised, as well [26, 13]. While these methods remained of mostly theoretical interest, the simple divide-and-conquer based merge method of Dudzinski and Dydek [3] is the basis of the inplace merge method implemented as part of the GNU Compiler Collection's STL implementation (more details in 2); theoretically speaking its running time of $\Theta(n \log n)$ for merging two runs of size n is suboptimal, though. Later refinements of Dudzinski and Dydek's method [16, 17, 9] improve upon that but are, again, more complicated.

In theory, one can combine stable and inplace merging even with a linear number of moves [8], but the complicated resulting method is likely unpractical and has not been implemented to our knowledge. The most competitive available practical implementation of inplace, stable sorting may be Wikisort [20], which is based on Kim and Kutzner's refinement [17].

2 Algorithms

Both Powersort and Timsort are stable, run-adaptive variants of Mergesort. They operate by one outer iteration over the array from left to right while maintaining a stack of runs. In each step, they find the next run of already sorted data to be put on top of the stack. A certain set of rules specific to the algorithm, the *merge policy*, determines whether we execute (potentially several) merges of runs at the top of stack before adding the newly detected run. Timsort's original merge policy uses a combination of rules comparing certain lengths among the 4 topmost runs. Powersort's merge policy [25] enforces increasing *"(run-boundary) powers"* (defined below) on the stack, simulating Mehlhorn's nearly optimal binary search tree algorithm [22, 23]. This more robust merge policy achieves optimal adaptivity up to an additive linear term with very low running-time overhead.

Throughout, we let T denote the size, in words, of the data type we are sorting. We always count memory in words. By *moves* we mean moving one element of the data type we are sorting to a different location in memory.

In this section, we detail the various algorithms we will use. Let us begin with a very abstract version of Powersort – this formulation will strip away details that are not relevant to the current work, as well as be general enough to capture all the variants of Powersort we will discuss. This algorithm is given in 1. Throughout this algorithm, we will assume some mechanism to maintain a small number *buffers* of bounded total length. These buffers should act as though they were double ended queues (e.g. with the typical operations on such queues i.e. PushFront, PushBack, PopFront, PopBack). We will not directly use the operations PushFront, PushBack, PopFront, PopBack, only the abstract operations NewBuffer, ExtractRun, MergeBuffers, Copy and Power. The meanings of these operations will now be described:

Algorithm 1 Abstract Powersort.

```
PowerSort(input[0, n))
    // Adaptively sort input[0, n).
1   work := NewBuffer(input[0, n)).
2   stack := a stack of pairs of buffers and powers.
3   curr := ExtractRun(work).
4   while ¬Empty(work)
5       next := ExtractRun(work)
6       p = Power(n, curr, next)
7       while Top(stack) == (top, p′) and p′ ≥ p
8           curr := MergeBuffers(top, curr)
9           Pop(stack)
10      Push(stack, (curr, p))
11      curr := next
12  while Top(stack) = (top, p)
13      curr := MergeBuffers(top, curr)
14  Copy(curr, input[0, n)).
```

NewBuffer(s). Instantiate a buffer using the memory from array s; this operation lets our buffering scheme take ownership of s.

ExtractRun(b). Extracts a maximal non-decreasing run from b; in principle this can also use e.g. insertion sort to insure a minimum run length.

MergeBuffers(b, b'). Merge buffers b and b', assumed to be in non-decreasing order, into a result buffer. This may destroy b and b'.

Copy(b, s). Copies the buffer b into the array s. This assumes that s was the input to a NewBuffer earlier; this relinquishes ownership of s.

***Power*(n, b, b').** A function which assigns an integer to a pair of buffers b, b'. This will be described below; for now, assume it it chosen so as to make 1 efficient.

Our efficient implementations will reuse memory to actually maintain these buffers – however for 1, the buffers simply behave like (double-ended) queues. Furthermore, every buffer will always contain elements from a contiguous subsegment of the original array – for ease of presentation, we implicitly assume that each buffer remembers this subsegment.[4]

The main important fact about powersort is that the *merge cost* is bounded by the *run-length entropy*. First let us formally define merge cost.

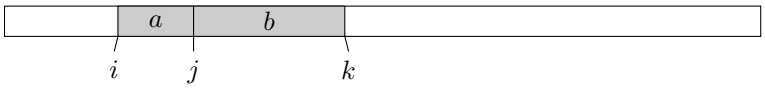
▶ **Definition 2.1.** *For a merge-based sorting algorithm, the* merge cost *is the sum of the lengths of all the inputs to all the merges used within the algorithm.*

Throughout, we let $\mathcal{H}$ denote the run-length entropy for our input sequence – formally, for an input consisting of runs of lengths $\ell_1, \dots, \ell_r$, we have $\mathcal{H} = \sum_{i=1}^{r} \frac{\ell_i}{n} \lg \frac{n}{\ell_i}$ (which coincides with the entropy of the distribution of selecting a random run proportional to its length).

▶ **Theorem 2.2** ([1, Thm 3.3, Rem 3.4])**.** *Consider an integer n and two runs a, b, such that*

$$1 \le \text{Left}(a) \le \text{Right}(a) = \text{Left}(b) \le \text{Right}(b) \le n.$$

[4] All segments are indexed closed on the left and open on the right.



Suppose $i = \textsc{Left}(a)$, $j = \textsc{Right}(a) = \textsc{Left}(b)$, $k = \textsc{right}(b)$. *If we set* $\textsc{Power}(n, a, b)$ *to be the smallest integer* p *such that there is a* $c \in \mathbb{Z}$ *with* $\frac{i+j}{2n} < c \cdot 2^{-p} \leq \frac{j+k}{2n}$ *then, within algorithm 1 the following hold:*

1. *The stack has height bounded by* $\lceil \log_2 n \rceil + 1$.
2. *The merge cost is bounded by* $n(\mathcal{H} + 2)$.

We will focus on four different implementations of 1, which we detail below. For all that follows, let $\mathcal{M}$ denote the merge cost of our algorithm. Note that many of our implementations will use $\mathcal{M} + (n-1)$ comparsions – this is because run detection takes an extra $n-1$ comparisons; for brevity, we drop the -1.

CPython-style Powersort. This is our baseline. In this implementation, buffers are merely subsegments of the original memory, in the natural order. NewBuffer, ExtractRun and Copy are purely metadata operations. Only Merge remains, which (i) copies the smaller buffer to auxiliary storage, then (ii) merges the two buffers back into the original space they used. This approach uses, in the worst case, $\mathcal{M} + n$ comparisons and $1.5\mathcal{M}$ moves, as well as $\frac{n}{2} \cdot T$ extra memory.

Note that our code uses the same buffer-copy strategy as the CPython implementation of Powersort, but in other aspects (e.g., non-galloping merge) chooses what is efficient for a `C++` generic sort with moderately expensive comparisons.

Pingpong Powersort (new). In this implementation, buffers on the stack are maintained in auxiliary memory, while *curr*, *next*, *work* are left in the original storage. Altogether, in the worst case we use $\mathcal{M} + n$ comparisons, $\mathcal{M} + n$ moves, and nT extra memory. We will explain this approach in more detail in 3.

Virtual-Memory (VM) Powersort (new). In this implementation, we use a variant of internal buffering to eliminate most of the extra space usage of our algorithm. We will explain this approach in detail in 4; for now, we simply state that this approach will use $\mathcal{M} + n$ comparisons, $\mathcal{M} + 3n$ moves, and $O(\sqrt{nT \log n})$ extra memory.

Inplace-Merge Powersort. In this implementation, we store all buffers within the original array, but achieve a space usage of $O(\log n)$ extra words of space via the inplace merge method `std::__merge_without_buffer`,[5] the (nonstandard) inplace merge implementation of the GCC STL implementation. It is based on the divide-and-conquer method of [3] and has running time $O(n \log n)$ for merging two runs of size $\Theta(n)$. Within the STL, it is used in `std::stable_sort` in case no extra memory is available.

We close with an informal description of our new algorithms. Pingpong Powersort uses the initial array $A[0..n)$ and an auxiliary array $B[0..n)$ as buffers: each run will be stored in A or B, and we merge runs from A and/or B using the corresponding positions of A and/or B as working space.

Virtual-memory Powersort, by contrast, treats the input array as the (first and largest) part of a collection of $n/P + \log_2(n) + O(1)$ *pages* of P elements each. (For the purposes of this paragraph, assume we sort objects 1 word long.) Each run will occupy a (not necessarily contiguous) list of full pages and maybe one partially filled page (and needs to also store pointers to these pages). As of yet undiscovered runs lie in their original position in the

[5] Source at: `https://github.com/gcc-mirror/gcc/blob/master/libstdc%2B%2B-v3/include/bits/stl_algo.h#L2436C5-L2436C27`

input array, so they do not waste space; only the $\log_2(n) + O(1)$ runs on the run stack each require one partially filled page, and we need $O(1)$ extra pages while merging. The partial pages and pointers cost $O(n/P + P \log n)$ extra words of memory. Setting $P = \Theta(\sqrt{n/\log n})$ minimizes the extra memory requirement, yielding $O(\sqrt{n \log n})$.

Galloping Merges

CPython's Timsort/Powersort implementation uses a *galloping merge* [27, 21], i.e., exponential searches for the minimum of the smaller array in the larger array. Using such a merge method makes the *comparison* cost of many run-adaptive mergesort variants approach the *dual*-runlength entropy [11, 10]; however, the data movement remains mostly unchanged.

Since galloping merges increase the variability of running times and is mostly effective if comparisons are very expensive [25], we do not include galloping in our merge methods. It is, however, an interesting question for future work to evaluate the overhead of virtual memory for galloping, since on one hand a check for page boundaries needs to be added, on the other hand, there is an enticing novel opportunity for speedups through virtual memory: in case an entire page can be skipped over, we can make this page part of the output using only metadata changes, *entirely avoiding* data movement for this page.

3 Pingpong Powersort (Few moves)

Let us first give a very simple implementation of 1 which sorts using $\mathcal{M} + n$ moves and $O(nT)$ extra memory. This implementation allocates a new buffer of n elements, each of size T, call it *stackBuf*. Consider some buffer *buf*, which contains elements from the range $input[i, j)$. (Recall that each buffer only holds a contiguous subrange of the original array.) If *buf* is one of *work*, *curr*, *next*, then we keep it in the memory range $input[i, j)$. Otherwise (i.e. if *buf* is kept in *stack*), we keep it in the memory range $stackBuf[i, j)$.

Let us describe how realize the operations from 1 with this implementation.

NewBuffer(s). This operation does nothing in this implementation: *work* ought already to be stored in *input*.

ExtractRun(b). This operation simply looks for the first pair of adjacent elements in the wrong order in *work*, then returns that. The result doesn't ever need to be moved (since it is stored either in *curr* or *next*).

MergeBuffers(b, b'). We observe that we only ever merge a buffer from *stack* with *curr*. For some values of $i \le j \le k$, the first occupies the memory range $stackBuf[i, j)$ and the second the memory range $input[j, k)$. Our goal will be to merge these two ranges into $input[i, k)$ – this can be done simply by merging left-to-right and outputting from i towards k in *input* (as in Step 2 of Figure 1). This takes at most $k - i$ moves i.e. as many moves as elements in the output.

Copy(b, s). This operation does nothing: *curr* is already in *input*.

A careful reader will have noticed that we have not yet described one detail: given the above description of where we store our buffers, we need to copy buffers from *input* to *stackBuf* whenever pushing it to the stack. However, this happens at most once for every element in the input array. Hence, as we do these n extra moves and otherwise do exactly one move for each comparison, we see that we used at most $\mathcal{M} + n$ moves as required.

In fact, we can sometimes do slightly better. Suppose the stack is such that, in line 7 of 1, we need to pop (and merge) at least one element. As described before, our algorithm extracts runs from the top of the stack, merging them with *curr*, storing them in memory

range within *input*. However, at the very last pop, we can instead directly merge *curr* into *stackBuf*: this time we need merely to merge right-to-left. Any such run eliminates its unnecessary moves. (This seems to happen very often in our experiments, cf. Figure 6.)

4 Virtual-Memory Powersort (Almost Inplace)

In this section we will explain how to implement the buffers in 1 so as to use only $O(\sqrt{nT \log n})$ extra memory, as well as only $\mathcal{M} + 3n$ moves. We will first give a simple theoretical algorithm that achieves these constraints, and then optimise some implementation details in 4.1.

Our algorithm will manage memory in terms of *pages* of P words of memory each, where we choose $P \approx \sqrt{n/T \log_2 n}$.[6] Each buffer will be stored in a sequence of pages, where the contents of a page are in the correct order, but pages may be located at arbitrary locations. Each page by itself will also be stored contiguously in memory. Thus each buffer will own a list of pointers to its constituent pages. Our algorithm globally maintains an (initially empty) list of free pages, which we call the *free list*. Whenever a new page is requested we try to take it from the free list. If no free page is available, we allocate a new page.

Let us now walk through the implementation of all operations:

PushFront, PushBack. We try to insert into the first/last page if possible; if not, we get a new page for the buffer from the free list, and insert the new element there.

PopFront, PopBack. This operation is simple to implement; the only crucial detail is that if a full page becomes empty, we add it to the free list.

NewBuffer. This operation takes in memory not yet managed by our page system, and creates a buffer containing it. We handle this by segmenting the memory into pages.[7]

Copy. This operation is quite nontrivial in our case, since we are copying from a buffer to the original memory – while our buffer may still be using part of the original memory! To handle this, we do the following: we search in the page list of the buffer *curr* to see if the first page in *input* is used somewhere. If it is, we copy that page out to a new page. Then, we copy the correct contents into the first page of *input*. We then continue this approach for all further pages.[8,9]

ExtractRun, MergeRuns. These operations are implemented trivially in terms of the previous operations, i.e. one allocates a result run and builds it by using PushBack, while using PopFront to consume the input buffer(s).

Empty. This is a trivial metadata operation.

Let us now consider the total memory consumption of this approach. To do this, let us count how many pages we need at any point in the run of the algorithm, and how many pages we receive from segmenting the original memory. Furthermore, we need to count how much metadata memory we use.

Pages needed. Clearly there are at most $\lceil n/P \rceil$ *full* pages. Note that there are at most 2 partially full pages for every buffer. Furthermore, there are only $\lceil \log_2 n \rceil$ buffers on the stack, as well as $O(1)$ extra auxilliary buffers needed at any point during the run of the algorithm. Hence overall we need $n/P + O(\log_2(n))$ pages.

---

[6] For our practical implementation, we additionally chose P to be a power of 2, to make arithmetic faster.

[7] There may be a final partial page – this page will never be declared free and reused by our algorithm, so we may treat it simply as a normal page, even if we don't own all of it.

[8] One page in particular will hold what should be the contents of the partial page that may exist at the end of *work*. This page is even easier to handle, since we know that the memory at the end of *work* is not being used to hold anything, as we do not free a partial page.

[9] Note also that this theoretically uses quadratic time with respect to the number of pages, i.e. $O((N/P)^2)$ which for our choice is $O(nT \log_2 n)$. This can be optimised to theoretical $O(n)$ time easily by using e.g. a hash table, but in practice a linear scan is fast enough.

Metadata needed. For every buffer, we must keep a page list, containing pointers to each page within our buffer. Since this is stored as a linked list, each buffer requires $O(1)$ extra memory plus $O(1)$ extra memory for each full page. Hence, since there are only $\log_2 n + O(1)$ buffers overall, and $\lceil n/P \rceil$ full pages, this only requires $n/P + \log_2 n + O(1)$ words of space overall.

Pages received. From segmenting the original main memory, we gain $\lfloor n/P \rfloor = n/P - O(1)$ pages in the worst case.

Hence we need $O(n/P)$ extra words of metadata space, as well as $O(\log_2 n)$ extra pages in the worst case, each costing $P \cdot T$ words, for T the size of the sorted records in words. Setting $P \approx \sqrt{n/(T \log_2 n)}$ balances these two terms, for a final memory usage of $O(\sqrt{nT \log n})$.

4.1 Implementation considerations

Fewer partial pages. In order to simplify our implementation, we can observe that the only time we ever use PopFront is either on *work*, or when calling Merge. The practical relevance of this is that we can maintain as an invariant that all buffers (except *work* and during Merge) have all pages entirely full *except perhaps the last one*. This tightens the analysis on the number of extra pages needed by a factor of 2, significantly lowering memory consumption, as well as simplifying implementation logic.

Fast merging. The innermost loop of our algorithm is a merge – hence it makes sense to optimise the merging of two buffers as much as possible. If we were to naively merge by calling PopFront and PushBack, we would incur the cost of a second indirection at each step! Hence, we implement our algorithm *page by page*: while either input buffer has any contents left, we merge the first pages of the two inputs into the currently last page of the output.

Algorithm 2 Standard and page-based binary straight merging algorithm. The page-based method use one extra check (boxed) that the standard merge doesn't need.

PageMerge$(a[x_a, y_a), b[x_b, y_b), r[x_r, y_r))$

```
// Merge as much as possible of
// a[x_a, y_a) and b[x_b, y_b) into r[x_r, y_r).
while x_a < y_a ∧ x_b < y_b ∧ [x_r < y_r]
    if a[x_a] ≤ b[x_b]
        r[x_r] := a[x_a]
        x_a := x_a + 1
    else
        r[x_r] := b[x_b]
        x_b := x_b + 1
    x_r := x_r + 1.
```

StandardMerge$(a[x_a, y_a), b[x_b, y_b), r[x_r, -))$

```
// Merge a[x_a, y_a) and b[x_b, y_b) into r[x_r, -),
// assuming there is enough output space.
while x_a < y_a ∧ x_b < y_b
    if a[x_a] ≤ b[x_b]
        r[x_r] := a[x_a]
        x_a := x_a + 1
    else
        r[x_r] := b[x_b]
        x_b := x_b + 1
    x_r := x_r + 1.
```

Now, a naive page merging algorithm is as in 2. We imagine, within PageMerge, that the indices $x_a, y_a, x_b, y_b, x_r, y_r$ are passed by reference. This algorithm merges two (perhaps partial) pages of memory into a (perhaps partial) result page. We can merge two buffers by repeatedly calling this subprocedure. Note that the cost of the subprocedure dominates over the cost of bookkeeping around it, since we only call such a subprocedure when we finish processing pages – this is roughly once for every merge, plus once every roughly P steps.

Compare this with a normal merging algorithm, where the output range is assumed to be long enough to contain all the output we care to write (cf. 2). Note that StandardMerge does one fewer check than PageMerge per element; namely, it doesn't do the check $x_r < y_r$.

But, conveniently the value $y_r - x_r$ decreases by exactly 1 every loop iteration. Hence we can do loop unrolling to get rid of almost all of these checks. This modification makes the unrolled PageMerge and StandardMerge do the same operations most of the time – leading to our almost inplace algorithm doing almost the same innermost loop as the naive algorithm.

Pre-allocation of buffers. To reduce the cost of allocating and deallocating memory, as well as to improve memory locality, it is advantageous to allocate all the extra memory used by the algorithm once, at the beginning of the run. Simply allocate enough space for all the memory we could need in the worst case, and add the allocated pages to the free list to begin with.

Eliminating linked lists. Practical implementations using linked lists can suffer from (i) higher memory consumption, and (ii) reduced performance due to more metadata operations (e.g. dereferences, allocations, deallocations). Therefore, we optimise by storing page lists (as well as the free page list) as arrays. Here one can take advantage of the fact that (other than $O(1)$ page lists used for *curr*, *next* and for merging), all other runs are stored on a stack. Therefore, we can keep a stack of arrays of pointers to pages for the runs on the stack, where the page pointer arrays for the constituent buffers form subsegments of that one stack – we also pre-allocate all of the potential stack memory from the beginning of the run of our algorithm.

Short run merges. If we merge two runs who, together, can fit within a single page, there is no point in removing a new page from the free list for the answer and placing the pages of the runs we merge back to the free list. We can slightly optimise this case by just merging into one of the two already allocated pages.

5 Evaluation

We formulate the following hypotheses to test in our empirical evaluation. Note that the last two hypotheses can be confirmed by mathematical analysis and would hence not require an empirical validation; we use them as sanity checks in our evaluation.

- **(H1)** Almost inplace merging allows for a substantial reduction in memory consumption.
- **(H2)** Almost inplace merging has low running-time overhead.
- **(H3)** Fully inplace merging (as a blackbox, or weaved into an overall clever algorithm) is not competitive with non-inplace sorting.
- **(H4)** Avoiding extra copies resp. moves helps when sorting larger objects.
- **(H5)** The number of comparisons is unaffected by the run storage method.
- **(H6)** Data movement can be predicted accurately based in terms of the mergecost $\mathcal{M}$.

Experimental setup. We implemented the variants of Powersort in `C++` and compared them there with the GNU Compiler Collection (GCC) implementation of `std::stable_sort` and `Wikisort` across different scenarios. We tested three different metrics: the time needed by the algorithm, the number of comparisons used by the algorithm, and the number of assignments used by the algorithm. The full code of our experimental setup is included in the supplemental material mentioned below the abstract. (The code for memory calculation and for properly counting assignments is in a feature branch.)

CPU and compiler settings. We ran our experiments on a `Intel(R) Core(TM) i7-4770 CPU @ 3.40GHz`. We compiled with the `-O2` flag on `gcc version 10.2.1 20210110 (Debian 10.2.1-6)`. We realised late in testing that we were counting assignments incorrectly (i.e. not counting copy or move constructors, or move assignment); we retested using the code from branch `assignment-fix` from the repository.

Algorithm variants. We run 6 different algorithms:

1. Virtual-Memory Powersort,
2. CPython-style Powersort,
3. Pingpong Powersort,
4. Powersort with `C++` standard library inplace merge,
5. Wikisort [20],
6. `C++` standard library `std::stable_sort`.

The GCC implementation uses different algorithms depending on whether allocating a buffer of size n is successful, switching between methods in the recursion. For our running-time experiments, `std::stable_sort` was always provided with sufficient buffer space, so it executes a Bottom-up Pingpong Merge (with Insertion sort for initial run formation).

For all of our Powersort implementations, we forcibly pre-sort short runs using insertion sort, up to a certain minimum run length. This minimum run length is chosen according to Timsort's rule: to compute this minimum run length, start from $\ell = n$, then repeatedly set $\ell = \lceil \ell/2 \rceil$ until $\ell < 64$. Then ℓ is the required minimum run length. In this work, we consider only increasing runs; all Powersort variants could take advantage of decreasing runs, as well.

Input data types. We next describe *what* we sort. Observe that a data type is (primarily) characterised by two properties: the amount of memory it consumes (and thus requires to move), and the efficiency of pairwise comparisons. We therefore distinguish four distinct scenarios depending on whether comparison resp. move time is high resp. low.

Fast comparison, fast move. For this scenario, we use 32 bit integers. In our graphs, this data type is called `int`.

Slow comparison, fast move. For this scenario, we use pointers to arrays of length 30, containing 32-bit integers, ordered lexicographically. For this scenario, the first 29 integers are always set to 0, so on top of the cost of pointer dereferencing, each comparison compares all 30 pairs of integers, making comparisons rather costly. In our graphs this data type is called `ptr`.

Fast comparison, slow move. For this scenario, we use arrays of length 30 containing random 32-bit integers, ordered lexicographically. Here, all elements are potentially nonzero, so most comparisons will only compare one pairs of integers. In our graphs, this data type is called `randomBlob30`.

Slow comparison, slow move. For this scenario, we use arrays of length 30, containing 32-bit integers, ordered lexicographically, with only the last integer nonzero. In our graphs, this data type is called `zeroBlob30`.

For investigating running time and memory, we used all 4 data types. (For counting comparisons and assignments, the underlying data type is irrelevant; we use `int` for efficiency.)

Input permutations. All our randomness is generated using the `mt19937 C++` standard library pseudorandom generator, with a fixed seed (hence, the inputs we run our algorithms on are the same between all algorithms). Let $N = 10^7$. We sort input sequences whose length is distributed uniformly at random between $9N/10$ and N. After having fixed the run length,

we generate a random input sequence. Depending on the data type, we generate either single integers or arrays with either some values equal to 0, or all values random, as described above. Every integer we want to generate randomly is taken uniformly at random from the range $[100, 10^9]$. After generating a random input sequence, we introduce pre-sortedness into it in the following way; let $S \in \mathbb{N}$ be a parameter we fix. We repeatedly draw a geometrically distributed length and pre-sort a further run with that length. The geometric distribution has parameter $1/S$ – this means that the expected run length is $S - 1$, i.e. we expect to have S consecutive $<$ signs from the beginning of each run (ignoring the possibility that two consecutive runs happen to accidentally form a single run).

We tested our algorithms with $S \in \{2, 10^2, 10^3, 10^4, 10^5, 10^6\}$. We only considered one case with small S, since our algorithms use insertion sort to guarantee a minimum run length.

Number of iterations. For each combination of data type, algorithm, presortedness and metric, we ran 100 experiments and aggregated the results.

Memory consumption. Additionally to the experiments mentioned before, we computed the memory consumption (excluding stack space and constantly many local variables) for the CPython style implementation of Powersort, the few-moves implementation of Powersort, and the almost-inplace implementation of Powersort. Since this does not depend on the average run-length, we only computed it for $S = 2$.

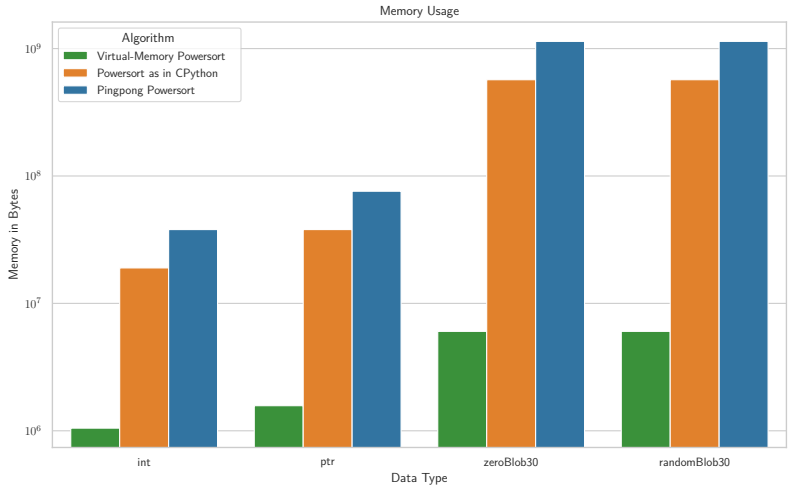
6 Results

We consider in turn the data relevant for our hypotheses from above. Figure 3 compares the extra memory allocated by the algorithms. Figure 4 and Figure 5 show the running-time results. Figure 6 and Figure 7 give the combinatorial cost measures (comparisons resp. assignments).

(H1): Significant reduction in memory consumption. As described, the required buffer area is reduced from space for $n/2$ objects to $O\left(\sqrt{n \log n}\right)$ objects, where n represents the number of objects to be sorted. The experimental results in Figure 3 show the reduction in memory: Almost-inplace Powersort achieves a memory reduction of several orders of magnitude compared to the CPython-style implementation and the few-moves implementation of Powersort.

(H2): Almost inplace merging has low overhead. We compare the run-time of almost-inplace Powersort with the other algorithms. We see in Figure 4 and Figure 5 that almost-inplace Powersort performs comparably with previous implementations of Powersort, and is uniformly faster than `Wikisort` or a black-box inplace implementation of inplace Powersort. Of particular note is that for `randomBlob30`, almost-inplace Powersort is even faster than the other algorithms. This might be due to touching an overall smaller range of memory addresses; here, the input already occupies 1GB of memory.

(H3): Fully inplace merging is not competitive. We see, in Figure 4, that an implementation of Powersort that uses a `C++` library inplace merge to achieve lower memory consumption is incredibly slow. This is probably due to the large number of assignments shown in Figure 6. Therefore, more clever techniques are needed. Wikisort, which also uses inplace merging as a subroutine, makes progress on this front, however it isn't fully run-adaptive (as seen by the higher comparison count, Figure 7), and still significantly slower.



Algorithm	`int`	`ptr`	`zeroBlob30`	`randomBlob30`
VM Powersort	1.0 MB	1.5 MB	6 MB	6 MB
CPython-style	18.1 MB	36.2 MB	543 MB	543 MB
Pingpong Powersort	36.2 MB	72.4 MB	1086 MB	1086 MB
Input size	36.2 MB	72.4 MB	1086 MB	1086 MB

Figure 3 Average memory usage in bytes of Virtual-Memory Powersort, Powersort as in CPython and Pingpong Powersort, for the different input types. This counts only main buffer allocations, not $O(\log(n))$-sized buffer stacks or $O(1)$ local variables. *The y-axis is logarithmic!*.

(H4): Avoiding extra copies resp. moves helps when sorting larger objects. Moving elements is very cheap for `int` and `ptr` data, but has substantial cost for `zeroBlob30` and `randomBlob30`. Comparing the relative ranking of the algorithms in Figure 5 across input types, we see that the move-heavy CPython-style Powersort and the move-optimized Pingpong Powersort perform almost the same for small types, whereas the former is substantially worse on large objects. Note that `std::stable_sort` on random inputs is slightly faster on `zeroBlob30`. This is likely due to its lower comparison cost: the adaptive algorithms' method for generating initial runs is not optimized for very expensive comparisons.

Hypotheses (H5) and (H6): Theoretical move and comparison counts. Our results for counting moves and comparisons closely match the theoretical estimation. Recall that VM Powersort should use $\mathcal{M} + 3n$ moves, Pingpong Powersort should use $\mathcal{M} + n$ moves, and the CPython-style Powersort should use $1.5\mathcal{M} + n$ moves; all of these should use $\mathcal{M} + n$ comparisons. We see this is well predicts the actual numbers, when ignoring $S = 2$ (where the initial run formation has a large influence). In particular, we see how closely the comparisons coincide in Figure 7.[10] In Figure 6, we see how (in our logarithmic plot) the trends for the three algorithms are clearly linear (which we expect, since for fixed n, $\mathcal{M}$ is logarithmic with respect to inverse run lengths), with slopes as expected (see the expected run-length

[10] Technically there is a very small difference, since some of the algorithms do some of their merges left-to-right or right-to-left respectively. This difference is very small.

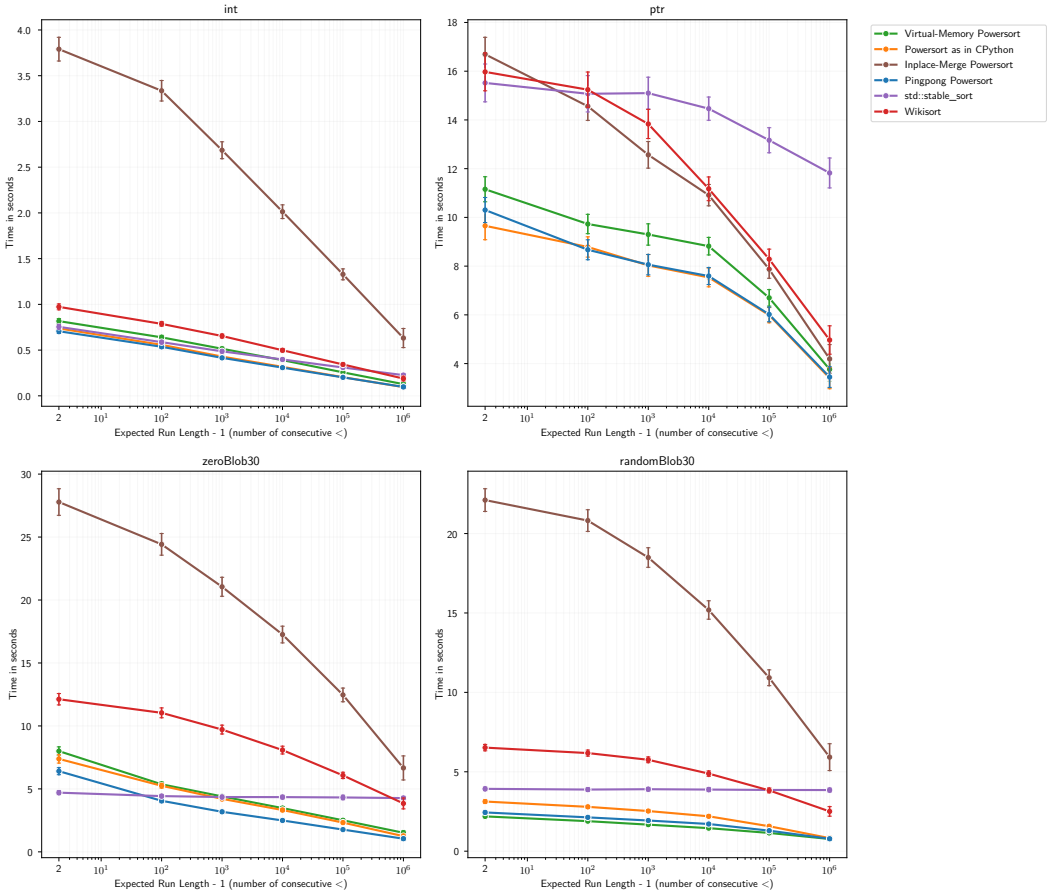


Figure 4 Average running time in seconds for the different input types. The input size is uniformly chosen between 9 million and 10 million, and the x-axis varies the presortedness.

entropy, in gray). The only unexpected thing is that Pingpong Powersort performs better than expected – likely since the optimisation mentioned in the last paragraph of 3 occurs often in our random inputs.

7 Conclusion

We conclude from the support of our hypotheses that in `C++` generic sorting, Pingpong Powersort is a strong contender for a good `std::stable_sort` implementation when ample extra buffer space is available. Only one out of many scenarios favors the current Pingpong Bottom-up Mergesort: if comparisons are expensive, objects are big, *and* the input is a random permutation. In the context of adaptive sorting, we point out that Pingpong Bottom-up Mergesort can save some of the comparisons when natural runs exist, but its memory movement is not adaptive and hence it quickly falls behind when sorting larger objects.

Moreover, almost-inplace Powersort serves a wide middle ground where substantially less, but still some buffer space is acceptable. Its slowdown over the linear-buffer variants is very modest, but its quadratically smaller memory footprint makes fast, optimally-adaptive, stable sorting applicable to a much wider range of applications.

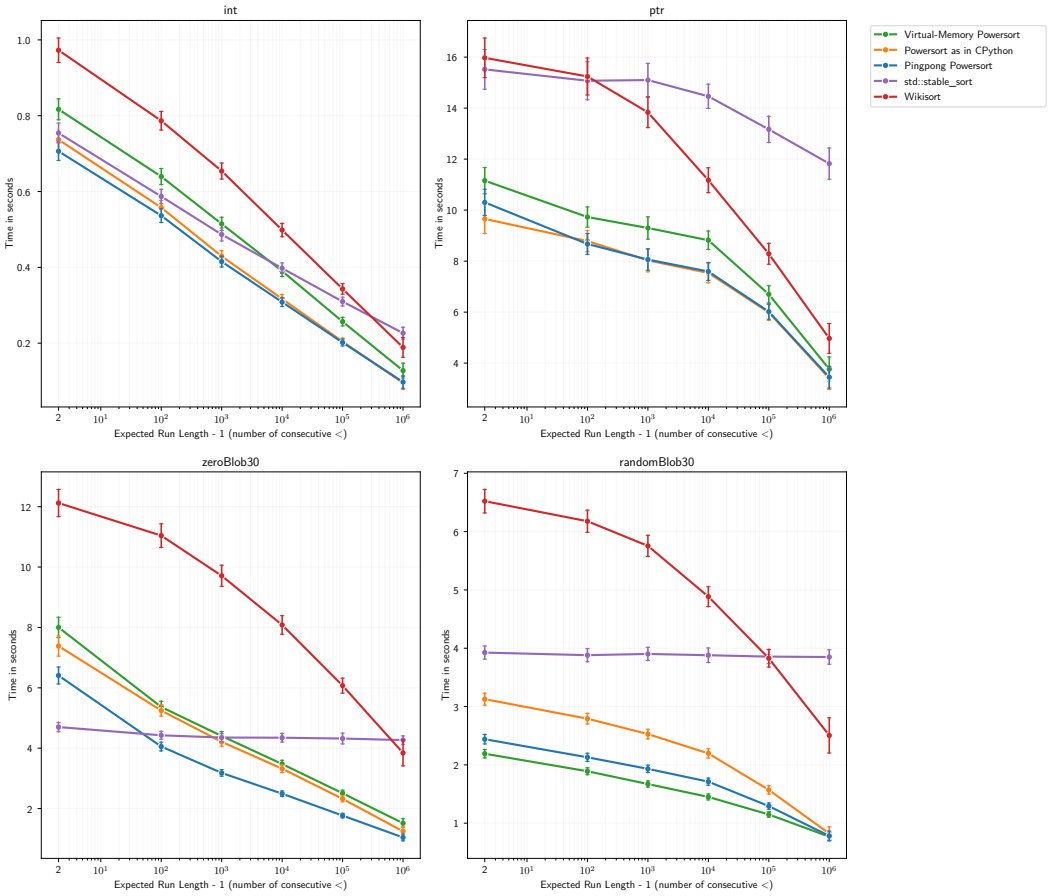


Figure 5 Same as Figure 4 but without the Inplace-Merge Powersort.

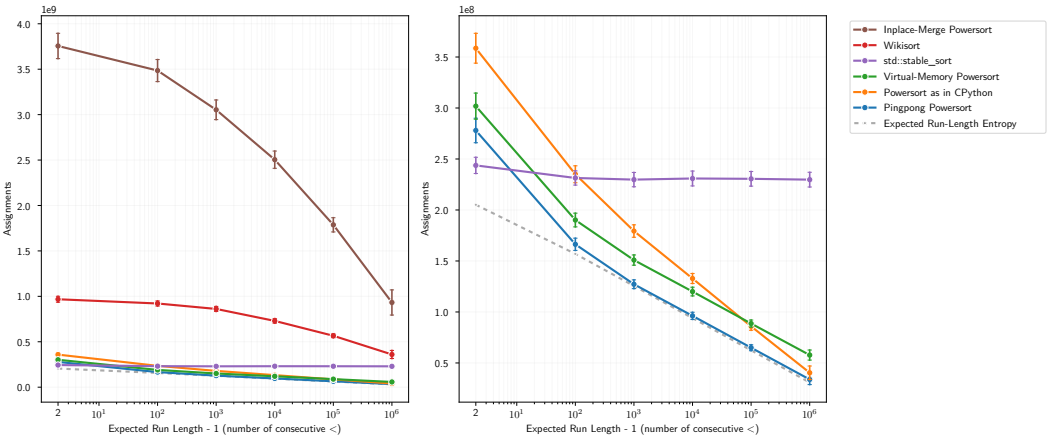


Figure 6 The average number of move/copy assignments and move/copy constructions for the same inputs as in Figure 4 with all algorithms (left) and without Wikisort and Inplace-Merge Powersort (right). The expected run-length entropy is calculated by the formula $EN\log_2(EN/(S+1))$, where $EN = 0.95 \times 10^7$ is the expected input length, and $S+1$ is the expected run length (ignoring consecutive runs that happen to become one run).

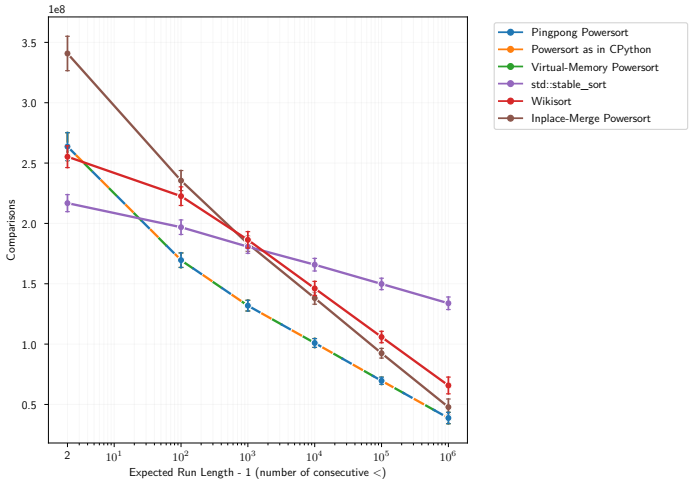


Figure 7 The average number of comparisons for the same inputs as in Figure 4. Pingpong, CPython-style and VM Powersort, execute (almost) the same comparisons.

References

1 William Cawley Gelling, Markus E. Nebel, Benjamin Smith, and Sebastian Wild. Multiway powersort. In *Symposium on Algorithm Engineering and Experiments (ALENEX)*, pages 190–200, 2023. doi:10.1137/1.9781611977561.ch16.

2 Badrish Chandramouli and Jonathan Goldstein. Patience is a virtue: revisiting merge and sort on modern processors. In *SIGMOD International Conference on Management of Data*, SIGMOD/PODS'14, pages 731–742. ACM, 2014. doi:10.1145/2588555.2593662.

3 Krzysztof Dudzinski and Andrzej Dydek. On a stable minimum storage merging algorithm. *Information Processing Letters*, 1981.

4 Stefan Edelkamp and Armin Weiß. QuickXsort: Efficient sorting with $n \log n - 1.399n + o(n)$ comparisons on average. In *Computer Science in Russia (CSR)*, pages 139–152, 2014. doi:10.1007/978-3-319-06686-8_11.

5 Stefan Edelkamp, Armin Weiß, and Sebastian Wild. QuickXsort – a fast sorting scheme in theory and practice. *Algorithmica*, 82(3):509–588, 2020. doi:10.1007/s00453-019-00634-0.

6 Amr Elmasry and Abdelrahman Hammad. Inversion-sensitive sorting algorithms in practice. *ACM Journal of Experimental Algorithmics*, 13, 2009. doi:10.1145/1412228.1455267.

7 Vladmir Estivill-Castro and Derick Wood. A survey of adaptive sorting algorithms. *ACM Computing Surveys*, 24(4):441–476, 1992. doi:10.1145/146370.146381.

8 Gianni Franceschini. Sorting Stably, in Place, with O(n log n) Comparisons and O(n) Moves. *Theory of Computing Systems*, 40(4):327–353, 2007. doi:10.1007/s00224-006-1311-1.

9 Viliam Geffert, Jyrki Katajainen, and Tomi Pasanen. Asymptotically efficient in-place merging. *Theoretical Computer Science*, 237(1-2):159–181, 2000. doi:10.1016/S0304-3975(98)00162-5.

10 Elahe Ghasemi, Vincent Jugé, and Ghazal Khalighinejad. Galloping in Fast-Growth Natural Merge Sorts. In *49th International Colloquium on Automata, Languages, and Programming (ICALP)*, volume 229 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 68:1–68:19, Dagstuhl, Germany, 2022. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.ICALP.2022.68.

11 Elahe Ghasemi, Vincent Jugé, Ghazal Khalighinejad, and Helia Yazdanyar. Galloping in fast-growth natural merge sorts. *Algorithmica*, 87(2):242–291, 2024. doi:10.1007/s00453-024-01285-6.

12 Goetz Graefe. Implementing sorting in database systems. *ACM Computing Surveys*, 38(3):10, 2006. doi:10.1145/1132960.1132964.

13 B-C. Huang and M. A. Langston. Fast stable merging and sorting in constant extra space. *The Computer Journal*, 35(6):643–650, 1992. doi:10.1093/comjnl/35.6.643.

14 Bing-Chao Huang and Michael A. Langston. Practical in-place merging. *Communications of the ACM*, 31(3):348–352, 1988. doi:10.1145/42392.42403.

15 Jyrki Katajainen, Tomi Pasanen, and Jukka Teuhola. Practical in-place mergesort. *Nordic Journal of Computing*, 3(1):27–40, 1996.

16 Pok-Son Kim and Arne Kutzner. Stable minimum storage merging by symmetric comparisons. In *European Symposium on Algorithms (ESA)*, volume 3221 of *LNCS*, pages 714–723, 2004. doi:10.1007/978-3-540-30140-0_63.

17 Pok-Son Kim and Arne Kutzner. Ratio based stable in-place merging. In *Theory and Applications of Models of Computation (TAMC)*, TAMC'08, pages 246–257, Berlin, Heidelberg, 2008. Springer-Verlag. doi:10.1007/978-3-540-79228-4_22.

18 M. A. Kronrod. An optimal ordering algorithm without a field of operation. *Soviet Math. Dokl.*, 10:744–746, 1969.

19 Heikki Mannila and Esko Ukkonen. A simple linear-time algorithm for in situ merging. *Information Processing Letters*, 18(4):203–208, 1984. doi:10.1016/0020-0190(84)90112-1.

20 Mike McFadden (BonzaiThePenguin). WikiSort. https://github.com/bonzaithepenguin/wikisort, 2021.

21 Peter McIlroy. Optimistic sorting and information theoretic complexity. In *SODA 1993*, pages 467–474. SIAM, 1993. URL: http://dl.acm.org/citation.cfm?id=313559.313859.

22 Kurt Mehlhorn. A best possible bound for the weighted path length of binary search trees. *SIAM Journal on Computing*, 6(2):235–239, 1977. doi:10.1137/0206017.

23 Kurt Mehlhorn. *Data Structures and Algorithms 1: Sorting and Searching*. Springer, 1984.

24 Finn Moltmann, Tamio-Vesa Nakajima, and Sebastian Wild. Virtual Memory Powersort. Software, version 1.0., Sebastian Wild: EPSRC grant EP/X039447/2, swhId: swh:1:snp:8e65d4e51fc7e84ac78dde26a19bea76773d8ce8 (visited on 2026-08-12). URL: https://git.sr.ht/~tamionv/vm-powersort, doi:10.4230/artifacts.27635.

25 J. Ian Munro and Sebastian Wild. Nearly-optimal mergesorts: Fast, practical sorting methods that optimally adapt to existing runs. In *European Symposium on Algorithms (ESA)*, volume 112 of *LIPIcs*, pages 63:1–63:16. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2018. doi:10.4230/LIPIcs.ESA.2018.63.

26 Luis Trabb Pardo. Stable Sorting and Merging with Optimal Space and Time Bounds. *SIAM Journal on Computing*, 6(2):351–372, 1977. doi:10.1137/0206025.

27 Tim Peters. Timsort (listsort.txt), 2002. URL: https://github.com/python/cpython/blob/3ddb856ed1fcfbfb750b00a60b9a5df76555751e/Objects/listsort.txt.

28 Ola Petersson and Alistair Moffat. A framework for adaptive sorting. In *Scandinavian Workshop on Algorithm Theory (SWAT)*, pages 422–433. Springer, 1992. doi:10.1007/3-540-55706-7_38.

29 Ola Petersson and Alistair Moffat. A framework for adaptive sorting. *Discrete Applied Mathematics*, 59(2):153–179, 1995. doi:10.1016/0166-218x(93)e0160-z.

30 Klaus Reinhardt. Sorting *In-Place* with a *Worst Case* complexity of $n \log n - 1.3n + o(\log n)$ comparisons and $\varepsilon n \log n + o(1)$ transports. In *International Symposium on Algorithms and Computation (ISAAC)*, pages 489–498, 1992. doi:10.1007/3-540-56279-6_101.

31 Lutz M. Wegner. Sorting – the Turku lectures, revised edition of the sorting algorithms presented in Turku, Finland May 11 – 15, 1987. Technical report, TUCS Turku Centre for Computer Science, 2014. URL: https://urn.fi/URN:NBN:fi-fe201402051375.

32 Sebastian Wild. Average cost of QuickXsort with pivot sampling. In *International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA)*, volume 110 of *LIPIcs*, pages 36:1–36:19. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2018. doi:10.4230/LIPIcs.AofA.2018.36.

33 Ian H. Witten, Alistair Moffat, and Timothy C. Bell. *Managing Gigabytes – Compressing and Indexing Documents and Images*. Morgan Kaufmann, 2nd edition, 1999.