

# $\varepsilon$ -Cost Sharding: Scaling Hypergraph-Based Static Functions and Filters to Trillions of Keys

Sebastiano Vigna  
Dipartimento di Informatica  
Università degli Studi di Milano  
Milan, Italy

April 6, 2025

## Abstract

We describe a simple and yet very scalable implementation of static functions (VFUNC) and of static filters (VFILTER) based on hypergraphs. We introduce the idea of  $\varepsilon$ -cost sharding, which allows us to build structures that can manage trillions of keys, at the same time increasing memory locality in hypergraph-based constructions. Contrarily to the commonly used HEM sharding method [6],  $\varepsilon$ -cost sharding does not require to store of additional information, and does not introduce dependencies in the computation chain; its only cost is that of few arithmetical instructions, and of a relative increase  $\varepsilon$  in space usage. We apply  $\varepsilon$ -cost sharding to the classical MWHC construction [20], but we obtain the best result by combining Dietzfelbinger and Walzer’s *fuse graphs* [13] for large shards with *lazy Gaussian elimination* [15] for small shards. We obtain large structures with an overhead of 10.5% with respect to the information-theoretical lower bound and with a query time that is a few nanoseconds away from the query time of the non-sharded version, which is the fastest currently available within the same space bounds. Besides comparing our structures with a non-sharded version, we contrast its tradeoffs with *bumped ribbon* constructions [14], a space-saving alternative to hypergraph-based static functions and filters, which provide optimum space consumption but slow construction and query time (though construction can be parallelized very efficiently). We build offline a trillion-key filter using commodity hardware in just 60 ns/key.

## 1 Introduction

*Static functions* are data structures designed to store arbitrary functions from finite sets to integers; that is, given a set of  $n$  keys  $S$ , and for each  $x \in S$  a value  $v_x$ , with  $0 \leq v_x < 2^b$ , a static function will retrieve  $v_x$  given  $x \in S$  in constant time. Closely related are *minimal perfect hash functions* (MPHFs), where only the set  $S$  is given, and the data structure yields an injective numbering of  $S$  into the first  $n$  natural numbers.

While these tasks can be easily implemented using hash tables, static functions and MPHFs are allowed to return *any* value if the queried key is not in the original set  $S$ ; this relaxation enables to break the information-theoretical lower bound of storing the set  $S$ . Indeed, constructions for static functions achieve just  $O(nb)$  bits of space and MPHFs  $O(n)$  bits of space, regardless of the size of the keys. This property makes static functions and MPHFs powerful techniques when handling, for instance, large sets of strings, and they are important building blocks of space-efficient data structures.

Historically, the first construction of static functions was has been proposed by Majewski, Havas, Wormald, and Czech [20]; it is based on the idea of solving a random linear system greedily: the solution is guaranteed to exist for sufficiently sparse systems because of *peelability* properties

of an associated random hypergraph [21]. Their technique was later refined (and sometimes rediscovered [10]) to reduce the space usage [15, 13].

The main problem with hypergraph peeling is that it does not scale well due to its poor cache behavior. Visiting a hypergraph is a linear operation, but the memory access pattern is very irregular, and the data structure is not amenable to parallelization.

A different approach is given by variants of the *ribbon* construction [14], which again uses a random linear system, but of a specific form that is not peelable as a hypergraph, but it is easy to solve by Gaussian elimination, which makes it possible to build static functions and filter with a space that is within 1% from the optimum. The construction is also highly parallelizable. The downside, as we will see, is that for large  $b$  and large key sets construction and queries are slow.

As noted by Dietzfelbinger and Pagh [12], every construction for a static function provides a construction for a static *filter*, which is a folklore name for an approximate dictionary with false positives. It is sufficient to store a static function mapping every key to a  $b$ -bit hash, and at query time check that the output of the function is equal to the hash of the key. Thus, a construction for static functions is a construction for static filters (but not vice versa).

In this paper, we propose a new way to build hypergraph-based (in particular, fuse-graphs-based [13]) static functions for very large key sets. The static functions and filters we obtain use about 10.5% more space than the information-theoretical lower bound for large key sets, so they are not competitive with ribbon-based constructions in terms of space, but they provide much faster query and construction times, which for most realistic large-scale applications are more relevant than a 9% size gain. Moreover, for large key sets their construction can be parallelized, and becomes very fast. Our structures can be easily built offline, and we report timings for the offline construction of large filters (up to a trillion keys).

All implementations discussed in this paper are distributed as free software as part of the Rust crate `sux`.<sup>1</sup>

## 2 Notation and tools

We use von Neumann’s definition and notation for natural numbers, identifying  $n$  with  $\{0, 1, \dots, n-1\}$ , so  $2 = \{0, 1\}$  and  $2^b$  is the set of  $b$ -bit numbers. We use  $\ln$  for the natural logarithm,  $\lg$  for the binary logarithm, and  $\log$  in Big-Oh notation.

A  $k$ -*hypergraph* (or  $k$ -*uniform hypergraph*) on a vertex set  $V$  is a subset  $E$  of  $\binom{V}{k}$ , the set of subsets of  $V$  of cardinality  $k$ . An element of  $E$  is called a *hyperedge*; since we will only be discussing hypergraphs, we will call it simply an *edge*. The *degree* of a vertex is the number of edges that contain it. An *induced subgraph* is given by a subset of vertices and all edges incident on those vertices (and no other vertices). The  $t$ -*core* of a hypergraph is its maximal induced subgraph having degree at least  $t$ .

A hypergraph is *peelable* if it is possible to sort its edges in a list so that for each edge there is a vertex that does not appear in the following elements of the list (such sorting can be easily found in linear time using a visit). If we fix an Abelian group  $\mathbf{A}$  and assign to each edge  $e$  of a  $k$ -hypergraph a value  $v_e \in \mathbf{A}$ , we can associate with the hypergraph a system of equations in which variables are vertices and edges represent equations: more precisely, an edge  $e \subseteq V$  represents the equation  $\sum_{v \in e} v = v_e$ . This correspondence induces a bijection between systems with  $|E|$  equations, each containing  $k$  variables, and  $k$ -hypergraphs with values assigned to edges. We will switch frequently between the two views.

---

<sup>1</sup><https://crates.io/crates/sux>

### 3 The Problems, and Previous Work

Let us first state our problems more formally.

**Problem 1 (Static functions)** *We are given a set of keys  $S$ ,  $|S| = n$ , out of a universe  $U$ , and a function  $f : S \rightarrow 2^b$  associating a  $b$ -bit value with each key. Our problem is to store information so that given  $x \in U$  we can compute in constant time a value that will be  $f(x)$  if  $x \in S$ .*

The key point here is that we give no condition when  $x \notin S$  (except that the computation must still be performed in constant time), so the information-theoretical lower bound for storing  $S$  as a subset of  $U$  does not apply. However, storing the set of values requires in general  $nb$  bits.

In some literature, the alternative term *retrieval problem* is used to denote the same problem [12].

**Problem 2 (Static approximate dictionaries, AKA static filters)** *We are given a set of keys  $S$ ,  $|S| = n$ , out of a universe  $U$ , and a  $\rho \leq 1$ . Our problem is to store information so that given  $x \in U$  we can compute in constant time whether  $x \in S$  with a false-positive error rate of  $\rho$ .*

In general, storing a function with a  $b$ -bit output requires at least  $nb$  bits, whereas storing a filter requires at least  $n \lg(1/\rho)$  bits [9]. In both cases we will use the term *overhead* to denote the increase in space with respect to the optimum.

In some literature, the alternative term *approximate membership problem* is used to denote the same problem [12].

#### 3.1 MWHC

In their seminal paper [20], Majewski, Wormald, Havas, and Czech (MWHC hereinafter) introduced the first compact construction for static functions using the connection between systems of equations and hypergraphs. To store a function  $f : S \rightarrow 2^b$  with  $b$ -bit values, they generate a random system with  $n = |S|$  equations in  $\lceil cn \rceil$  variables, for some *expansion factor*  $c \geq 1$ , of the form

$$w_{h_0(x)} \oplus w_{h_1(x)} \oplus \cdots \oplus w_{h_{k-1}(x)} = f(x) \quad x \in S, \quad (1)$$

Here  $h_i : U \rightarrow [cn]$  are  $k$  fully random hash functions, and the  $w_i$ 's are variables in some Abelian group with operation  $\oplus$ . Due to bounds on the peelability of random hypergraphs, when  $k \geq 3$  if the expansion factor  $c$  (i.e., essentially the ratio between the number of variables and the number of equations) is above a certain threshold  $\gamma_k$  the system can be almost always triangulated in linear time by peeling the associated hypergraph, and thus solved; in other words, we have both a probabilistic guarantee that the system is solvable and that the solution can be found in linear time.

The data structure is then a solution of the system (i.e., the values of the variables  $w_i$ ): storing the solution makes it clearly possible to compute  $f(x)$  in constant time. The space usage is approximately  $\gamma_k b$  bits per key. The constant  $\gamma_k$  attains its minimum at  $k = 3$  ( $\gamma_3 \approx 1.2217$ ), providing static functions with 23% space overhead [21].

The original paper [20] used modular arithmetic for  $\oplus$ . All modern constructions use instead  $b$ -bit vectors and XOR, as in practical applications there is no advantage in using values of  $m$  that are not powers of 2, and modulo operations are very slow.

As we noted in the introduction, Dietzfelbinger and Pagh [12] argue that every construction for a static function provides a construction for a static filter by mapping each key to a  $b$ -bit hash, providing a false-positive error rate of  $2^{-b}$ . This observation led to implementations of static filters based on the MWHC construction, such as those in the Sux4J library of succinct data structures [1]. The same structures were a few years later named by Graf and Lemire “Xor filter” [16]; their paper, however, does not make any comparison with previous work. Moreover, “Xor filter”, albeit catchy,

is a bit of a misnomer, as basically all constructions of this kind are based on the XOR operation, and, in fact, a “filter” is expected to be mutable (like a Bloom filter). We will thus use the MWHC acronym to refer to the construction and keep around the “static” label to avoid confusion.

### 3.2 HEM

Botelho, Pagh, and Ziviani [6] introduced a practical external-memory algorithm called Heuristic External Memory (HEM) to construct static functions (and MPHFs) for sets that are too large to store their hypergraph in memory. The idea is that of mapping all keys to random-looking *signatures* of  $c \lg n$  bits that guarantee uniqueness with high probability (typically, 128-bit good-quality signatures are sufficient for all applications). At that point, signatures are sorted and divided into shards using the  $k$  higher bits, resulting in  $2^k$  shards; alternatively, one can generalize and map a larger number of high bits to a range that is not a power of two using inversion in fixed-point arithmetic [15], which makes it possible to size the shards more accurately. One then computes a separate structure for the keys in each shard and stores all such structures. To do so, one has to keep track of *offsets*, that is, the starting position of the representation of each structure. Moreover, for each shard one keeps track of a *local seed*, which is necessary to obviate a possible failure of the hypergraph construction in each shard. While devised for MPHFs based on the MWHC technique, HEM can be made to work with any construction technique based on hypergraphs, and it has been for quite some time the standard method to approach the construction of large-scale static functions and filters, also because sharding, besides making it possible to build structures larger than the available memory, increases the locality of the construction and enables parallel computations.

### 3.3 GOV

A first improvement over MWHC is due to Genuzio, Ottaviano, and Vigna [15], which propose to generate random systems in the same fashion as MWHC, but at the threshold of solvability of linear systems on  $\mathbf{F}_2$ , which is lower than that of peelability, and moreover decreases towards one as the number of variables increases [24]. Gaussian elimination, however, is cubic, so the authors shard heavily the keys using a variant of HEM and solve each shard using *lazy Gaussian elimination*, a heuristic technique that makes it possible to solve quickly systems with dozens of thousands of variables. The overhead in this case is about 10% with 3 variables per equation, or 4% using 4 variables. Queries are however slowed down by having to pass through the two steps of the HEM query processing.

### 3.4 Fuse Graphs

Dietzfelbinger and Walzer [13] revolutionized the field by introducing a class of random graphs, *fuse graphs*, that are almost always peelable if sufficiently sparse, as in the general case, but have peelability thresholds close to the solvability threshold of linear system on  $\mathbf{F}_2$  for sufficiently large key sets. This approach instantly made it possible to build static functions by peeling with overhead similar to GOV, but with a much faster construction time. Fuse graphs are currently the state of the art among hypergraph-based constructions, albeit they have some limitations that we will discuss later.

### 3.5 Bumped Ribbons

Finally, Dillinger, Hübschle-Schneider, Sanders, and Walzer introduced *bumped ribbons* [14], a class of systems of equations of particular form that can be solved in practical linear time, and in which a good locality makes it possible to use a high number of variables, thus bypassing the limits

of GOV and lowering the space overhead below 1%. This is a goal that it is not possible to reach by hypergraph peeling, but, as we will see, using a large number of variables has sometimes a high cost in terms of construction and query time. We will use the acronym proposed by the authors, *BuRR* (for *Bumped Ribbon Retrieval*), to refer to the construction.

### 3.6 Other Constructions

There are many other constructions that contend the space of static functions and filters, in some cases offering even mutability: for example, blocked Bloom filters [25], Morton filters [8], quotient filters [19], and so on. A very large number of these variants has been painstakingly tested in the *BuRR* paper [14], with a very detailed analysis of the tradeoffs within each construction, so we will not repeat the same analysis here. The conclusion is that it is difficult to beat the speed of hypergraph-based constructions (in particular, the MWHC construction, therein called “Xor filter”) due to the simplicity of their operation, and to the fact that the issue independent memory accesses.

Our goal is thus to make hypergraph-based structures scalable to arbitrary sizes on commodity hardware.

## 4 $\epsilon$ -Cost Sharding

We will now show how to build static functions and filters based on hypergraph peeling on a very large scale with a guaranteed space overhead, obtaining fast versions that scale with size, overcoming the problem of the large amount of memory required for their construction, and their poor cache behavior.

We already discussed the HEM approach to the construction of static functions for large sets of keys. HEM has a significant overhead on query time for two main reasons:

- shards will be approximately of the same size, but not all of the same size, and thus it is in general necessary to keep track of their sizes—more commonly, of the cumulative sum of their sizes;
- if the construction of the data structure is probabilistic, each shard may fail, which requires storing a *local seed* per shard that will be used to restart the construction in case it fails.

The consequence is that it is necessary to store these two pieces of data (typically, conflated in a single 64-bit word, as a few bits of seeding are sufficient), and at query time retrieving them might be costly in terms of cache misses; more importantly, they introduce a memory dependency in the chain of operations that are necessary to compute the output. For example, in the GOV construction [15] the necessity of creating small shards that can be solved yields a fairly large offset/seed array, which has a significant impact on query time.<sup>2</sup>

We thus suggest to perform  *$\epsilon$ -cost sharding*: the observation we make is that the process of sharding is essentially a *balls-and-bins problem* when we treat the shard assignment as random: using the tight result of Raab and Steger [26] on the size of the bin with the maximum number of balls in a bin, we will now show how, given an error  $\epsilon$ , it is possible to compute easily and accurately a maximum number of bins such that the ratio between the maximum and the average number of balls in a bin is bounded by  $1 + \epsilon$ : at that point, given that (as it usually happens, and as it happens in our use cases) it is possible to build a structure on a shard using slightly more space than it is exactly necessary, we can proceed by building structures using as space measure the size

---

<sup>2</sup>We note that *BuRR* does not have this problem because, as the authors note [14], the difference between shards can be conflated in the bumping process.

of the largest shard, knowing that we will not increase the space usage by more than  $1 + \varepsilon$ , and dispensing completely with the offset array and the associated memory dependency.<sup>3</sup>

For what matter the local seed, we can again play with shard size so that we can have success with all the shards with the same seed with good probability. This requirement is not in contrast with the previous one, as larger shards mean even more reduced space usage. A failure means simply that the whole construction has to be restarted.

Depending on the type of construction at hand, one might have analytical bounds on the probability of success. If it is not the case, one can simply resort to simulation to estimate the size of shards which implies a good probability of success. In any case, unless there are no other bounds on the shard size, one can limit the number of shards, in general, to a few thousand, as after that point the additional memory used at construction time will be negligible with respect to the size of the data structure.

We recall that Raab and Steger [26] prove that when throwing  $n$  balls at random in  $s$  bins the maximum number  $M$  of balls in a bin satisfies

$$\text{Prob} \left[ M > \frac{n}{s} + \alpha \sqrt{2 \frac{n}{s} \ln s} \right] = o(1)$$

for  $\alpha > 1$ ,  $n = \omega(s \log s)$ ,  $n = O(s \text{ polylog}(s))$  (this bounds model well our case, and if  $n$  is asymptotically larger, the bounds are even more in our favor). We can then ask when the ratio between the maximum and the mean value  $n/s$  is at most  $1 + \varepsilon$ , and we obtain

$$\frac{\frac{n}{s} + \alpha \sqrt{2 \frac{n}{s} \ln s}}{\frac{n}{s}} \leq 1 + \varepsilon,$$

which turns into

$$\alpha \sqrt{2 \frac{s}{n} \ln s} \leq \varepsilon,$$

so

$$s \ln s \leq \frac{n \varepsilon^2}{2 \alpha^2},$$

and finally

$$\ln s \leq W \left( \frac{n \varepsilon^2}{2 \alpha^2} \right), \quad (2)$$

where  $W$  is Lambert's function. This is the bound on the number of shards we should use to keep the ratio between the maximum and the mean number of balls in a bin below  $1 + \varepsilon$ .

If we are just interested in power-of-two sharding, we can use the approximation  $W(x) \approx \ln x - \ln \ln x$  to obtain that we have to use at most

$$\left\lceil \lg \frac{n \varepsilon^2}{2 \alpha^2} - \lg \ln \frac{n \varepsilon^2}{2 \alpha^2} \right\rceil \quad (3)$$

higher bits when defining shards. In practice, choosing  $\alpha = 1$  works well.

Note that using the much rougher approximation  $W(x) \approx \ln x$  one can interpret (2) as saying that to get a space cost of  $\varepsilon$  shards should be at most  $n \varepsilon^2 / 2$ , that is, shards should be at least of size  $\approx 2 / \varepsilon^2$ , given that  $n \gg n \varepsilon^2 / 2$ , that is,  $\varepsilon^2 \ll 2$ ; despite being really rough, these approximations are useful to have an intuition about the range of validity and the value of the bound.

We are now going to apply this idea to two hypergraph-based constructions.

<sup>3</sup>We remark that a similar observation has been made by Ragnar Groot Koerkamp in [18], Section 3.2, when discussing partitioning techniques for the construction of minimal perfect hash functions, albeit without obtaining (2).

## 4.1 MWHC

We have applied  $\varepsilon$ -cost sharding to the MWHC construction [20], using 3-hypergraphs. The asymptotic behavior of random 3-hypergraphs kicks in very quickly, and graphs are peelable with overwhelming probability even for relatively small sizes.

While the MWHC construction has been mostly obsoleted by fuse graphs [13], due to the significantly lower space overhead, it is an interesting term of comparison because, as we will see, fuse graphs offer some obstacles to  $\varepsilon$ -cost sharding that random 3-hypergraphs do not. In particular, we set  $\varepsilon = 0.01$  in this case, leading, as we will see, to very fine sharding and the fastest parallel construction.

3-hypergraphs almost never exhibit a nontrivial 2-core for even moderate  $n$ . When we cannot peel a hypergraph, the reason is always the same: a duplicate edge. This phenomenon reflects the difference between the model in which the peelability thresholds are proven (all hypergraphs without duplicate edges with a certain ratio of edges to vertices are equiprobable) and model obtained by the random process generating the hypergraphs in the data structures we use (edges are generated at random with replacement, possibly yielding hypergraphs with duplicate edges).

We can however compute analytically the maximum number of shards for which the probability of a duplicate edge is below a certain threshold, as such probability are well known. If we consider a sequence of random uniform values  $x_0, x_1, x_2, \dots$  out of  $n$ , and we look for the first repeated value, that is, for the least  $t$  such that  $x_t = x_i$  for some  $i < t$ , the probability mass function of  $t$  for a uniform random source with  $n$  possible outputs is

$$\text{Prob}_n(t) = \frac{t}{n} \prod_{i=0}^{t-1} \left(1 - \frac{i}{n}\right).$$

The cumulative distribution function is even simpler:

$$F_n(t) = \sum_{i=0}^t \text{Prob}_n(i) = 1 - \prod_{i=0}^{t-1} \left(1 - \frac{i}{n}\right).$$

Both expressions can be bounded easily as follows:

$$\prod_{i=0}^{t-1} \left(1 - \frac{i}{n}\right) = e^{\sum_{i=0}^{t-1} \ln(1-i/n)} \leq e^{\sum_{i=0}^{t-1} -i/n} = e^{-t(t-1)/2n}. \quad (4)$$

In fact, when  $t \ll n$  the bound is a good approximation, so  $e^{-t^2/2n}$  approximates the probability that there is no duplicate value after generating uniformly  $t$  values out of  $n$ .

In our case, if  $c$  is the expansion factor,  $k$  the arity of the hypergraph, and  $m$  the number of keys of a shard, we have  $(cm/k)^k$  possible edges.<sup>4</sup> The probability that there are no duplicates for  $k = 3$  is thus

$$\exp\left(-\frac{m^2}{2(cm/3)^3}\right) = \exp\left(-\frac{1}{2(c/3)^3 m}\right).$$

When sharding, the  $n$  keys will be divided in  $S$  shards, each of which will have a size very close to  $n/S$ . Since we need to bound the probability at the same time for all shards, we want to impose that for some  $\eta < 1$

$$\exp\left(\frac{-1}{2(c/3)^3(n/S)}\right)^S \geq 1 - \eta.$$

---

<sup>4</sup>We use the  $k$ -partite construction [7].

Taking logarithms,

$$S \cdot \frac{-1}{2(c/3)^3(n/S)} \geq \ln(1 - \eta),$$

that is,

$$S \leq \sqrt{-2n(c/3)^3 \ln(1 - \eta)}, \quad (5)$$

which if we assume  $S = 2^h$  becomes

$$h \leq \frac{1}{2} (\lg n + 1 + 3 \lg(c/3) + \lg(-\ln(1 - \eta))) \approx \frac{1}{2} (\lg n + 1 + 3 \lg(c/3) + \lg \eta)$$

For example, we can divide  $10^{12}$  key into 8192 shards with  $c = 1.23$  and  $\eta = 10^{-3}$ . We will combine this bound with (3) when sharding.

## 4.2 Fuse Graphs

Let us first quickly recall the definition of a fuse graph of arity  $k$  [13]: given a set of  $n$  keys, we select an expansion factor  $c$  and a number of segments  $\ell$ . Then, we divide  $\lceil cn \rceil$  vertices in  $\ell + k - 1$  segments of the same size (this might require some rounding); finally, we choose a random edge by choosing a random segment  $s \in [0.. \ell)$  and  $k$  random vertices in the segments  $s, s + 1, \dots, s + k - 1$ . We will focus on the case  $k = 3$ .

For large  $n$  and suitable  $c$  and  $\ell$ , fuse graphs are almost always peelable. The first problem is that the asymptotic behavior kicks in fairly slowly—you need a few million vertices. Second, while empirically large graphs with, say,  $k = 3$  are peelable with  $c \approx 1.105$ , the exact values for  $\ell$  are not known analytically. The second problem is that for small  $n$  the expansion factor becomes significantly larger, reaching that of general random uniform hypergraphs.

Graf and Lemire [17] propose a very optimized implementation of fuse graphs for filters that we used as the base of our own implementation. In particular, they provide empirical estimates for  $c$  and  $\ell$  when  $k = 3$  or  $k = 4$  that we use partially. We depart from their estimates in two cases:

- For small sizes (below 800 000 keys) we resort to  $\varepsilon$ -cost sharding and lazy Gaussian elimination to keep  $c$  at 1.12; this entails a construction time that is, per key, an order of magnitude slower, but it is still practical given the small size.
- When sharding large key sets, we use  $c = 1.105$  and compute  $\ell$  using the estimate

$$0.41 \cdot \ln n \ln \ln n - 3 \quad (6)$$

that is  $O(\log n \log \log n)$ , rather than the estimate from [17],

$$\log_{3.33} n + 2.25, \quad (7)$$

which is  $O(\log n)$ . Having larger segments reduces the locality of the construction, thus slowing down the sequential construction time, but, as we will discuss now, makes it possible much finer sharding.

Even if we do not have precise probability estimations, within our assumption  $c = 1.105$  fuse graph almost never exhibits a nontrivial 2-core for large  $n$  (we never create shards smaller than 10 million keys). When we cannot complete a peeling, once again the reason is a duplicate edge. Duplicate edges are much more likely than in the previous case because the choice of vertices is limited to the ones in certain segments, and the problem is only exacerbated by  $\varepsilon$ -cost sharding, as we are building more structures with the same seed, using smaller segments.

Using the same framework as in the previous section, if  $c$  is the expansion factor,  $k$  the arity of the hypergraph,  $m$  the number of keys of a shard, and  $s_c$  the function (parameterized by the expansion factor) returning the size of a segment given the number of keys, we have  $(cm/s_c(m) - k + 1)s_c(m)^k \approx cm \cdot s_c(m)^{k-1}$  possible edges. In our case the probability that there are no duplicates is thus approximately

$$\exp\left(-\frac{m^2}{2cm \cdot s_c(m)^2}\right) = \exp\left(-\frac{m}{2c \cdot s_c(m)^2}\right)$$

As in the previous section, when sharding, the  $n$  keys will be divided in  $S$  shards, each of which will have a size very close to  $n/S$ . Since we need to bound the probability at the same time for all shards, we want to ask that

$$\exp\left(\frac{-n/S}{2c \cdot s_c(n/S)^2}\right)^S \geq 1 - \eta$$

Taking logarithms,

$$S \cdot \frac{-n/S}{2c \cdot s_c(n/S)^2} \geq \ln(1 - \eta),$$

that is,

$$s_c(n/S) \geq \sqrt{-\frac{n}{2c \ln(1 - \eta)}}.$$

We take a base-2 logarithm on both sides and replace  $\lg(s_c(n/S))$  with (6), obtaining

$$\alpha \ln\left(\frac{n}{S}\right) \ln \ln\left(\frac{n}{S}\right) + \beta \geq \frac{1}{2} \lg \frac{-n}{2c \ln(1 - \eta)},$$

and then

$$\ln \ln\left(\frac{n}{S}\right) \geq W\left(\frac{1}{2\alpha} \left(\lg \frac{-n}{2c \ln(1 - \eta)} - 2\beta\right)\right)$$

where  $W$  is Lambert's function. Solving for  $S$ , and recalling that  $e^{W(x)} = x/W(x)$ , we finally obtain

$$S \leq n \cdot \exp\left(-\frac{\frac{1}{2\alpha} \left(\lg \frac{-n}{2c \ln(1 - \eta)} - 2\beta\right)}{W\left(\frac{1}{2\alpha} \left(\lg \frac{-n}{2c \ln(1 - \eta)} - 2\beta\right)\right)}\right), \quad (8)$$

which if we assume  $S = 2^h$  becomes

$$h \leq \lg n - \frac{\frac{1}{2\alpha} \left(\lg \frac{-n}{2c \ln(1 - \eta)} - 2\beta\right)}{\ln 2 \cdot W\left(\frac{1}{2\alpha} \left(\lg \frac{-n}{2c \ln(1 - \eta)} - 2\beta\right)\right)}. \quad (9)$$

For example, we can divide  $10^{12}$  key into 2048 shards with  $\eta = 10^{-3}$ .

The right side is of the form  $\lg n - O(\lg n / \lg \lg n)$ , whereas (3) is of the form  $\lg n - O(\lg \lg n)$ , so, in the long run, the dominating bound will be (9). Note that in this case it is important that Lambert's function is computed accurately, and not using approximations such as  $W(x) \approx \ln x - \ln \ln x$ .

We remark that using the estimate (7) instead of (6) would lead to little sharding, even for trillion of keys.

### 4.3 Peel by Signature, Peel by Index

An important issue to manage when peeling hypergraphs is memory usage, albeit this issue is often overlooked in the literature. Since one of the purposes of sharding is to parallelize peelings, memory usage becomes even more important.

Let us briefly remember how the peeling visit works: one first chooses a supporting data structure  $S$ , which can be a queue or a stack; we will also need an edge stack  $E$  where we accumulate peeled edges.

Then, we (almost) follow the standard visit algorithm: we iterate over the vertices, and when we find a vertex of degree one we enqueue it in  $S$ . We now dequeue vertices from  $S$  until it is empty, and for each dequeued vertex of degree one we remove the only incident edge, pushing it on  $E$ , updating the incidence lists of the other vertices in the edge (if the vertex has degree zero, we discard it). If any of such vertices has now become of degree one, they are enqueued to  $S$ ; this maintains the invariant that each vertex on the stack has at most degree one. Once  $S$  is empty, we continue to iterate over the vertices. At the end of the visit, the peeling is successful if  $E$  contains all edges. At that point, edges can be popped from  $E$ , and for each edge we can assign a value to the vertex from which the edge was peeled; this information must also be stored in  $E$ ; rather than storing the vertex, it is sufficient to store the index of the vertex in the edge, which we call the *side* of the vertex.<sup>5</sup>

There are many variants that can be more performant on some architectures: for example, one can pre-load  $S$  with all vertices of degree one. Note that, as in the standard case, the stack-based visit is not a depth-first visit: a depth-first visit would require to push on the stack also the state of the iteration on the other vertices. In any case, however, each vertex is enqueued in  $S$  at most once.

Whereas the random hypergraphs of the MWHC construction have a very bad locality (edges are, indeed, almost entirely random), fuse graphs have edges whose vertices are relatively near, as they lie in consecutive segments: in fact, the smaller the segments, the higher the locality.

The choice of visit for the peeling process can be crucial: stack-based visits are known, in general, to have a higher locality than queue-based visits, and this difference is extremely pronounced in the case of fuse graphs, as noted by Graph and Lemire [17]: not only the peeling visit is much faster using a stack, but also the subsequent assignment of values becomes faster, because in  $E$  nearby edges often belong to the same segment.

All current constructions for hypergraph-based structures start by hashing the keys to unique signatures (usually, to 64 bits for small sizes, 128 for large sizes) and then work on the signatures. This approach avoids costly recalculations each time we have to turn a key into an edge, because we can simply generate the edge from the signature. It also couples well with  $\epsilon$ -cost sharding, because we can re-use the sharding signatures, taking care of not using the sharding bits, as they will be the same for all signatures in the same shard.

If we need to represent a static function, beside each signature we also have to keep track of the associated value; in the case of filters, we can compute hashes of the desired size directly from the signature [1]. One has first to pass through all the signatures to generate the edges and store them in incidence lists, one for each vertex, and again in this phase fuse graphs help, as count sorting the signatures by the segment of their associated edge increases the locality of the construction [17].

A significant reduction of the space used for incidence lists can be obtained using *Djamal's XOR trick* [4, 5]: rather than storing an edge list per vertex, it is sufficient to XOR together the edges in the list, because we follow an edge during the visit only if the vertex we discover is of degree one, and thus the XOR contains just the edge we are interested in.

There are at this point two main strategies:

---

<sup>5</sup>Technically, edges are sets, but in practice they are represented as a sequence of vertices, so sides are well defined.

- Peel-by-signature: incidence lists are stored as XOR of signatures (and possibly associated values). This is, for example, the strategy used in [17].
- Peel-by-index: incidence lists are stored as XOR of indices into the list of signatures (and possibly values).

In the peel-by-signature case, if the goal is that of peeling completely the graph it is possible to drop the signature/value pairs and just keep the graph representation. Thus the memory usage is reduced to a signature/value pair per vertex. In the peel-by-index case, one has to keep around the signature/value pairs to retrieve an edge given its index, so the memory usage is a signature/value pair per key plus an index per vertex, but after a failed peeling it is possible to generate and solve a system of equations taking care of the 2-core, that is, the unpeeled edges. There is some small additional space to store degrees and sides, but it is common to both representations.

The other important difference is locality. In the peel-by-signature case during the visit we have direct access to signatures, from which we can generate edges associated with keys, whereas in the peel-by-index case we have a costly memory indirection from an edge index to a signature.

The other choice to make is how to represent the stack of peeled edges. One possibility is that of, once again, store signature/value pairs, which provide high locality in the assignment phase [17]. The memory cost of the visit is then an index per vertex for the visit stack, plus a signature/value per for each edge. We call this strategy the *high-memory visit*, as it more than doubles the memory usage with respect to the graph representation.

As noted in [15], however, if one takes care of leaving in the incidence list of a vertex the edge that was peeled from that vertex (it is sufficient to zero virtually the degree for the peeling algorithm to work),  $E$  can accumulate vertex indices, rather than edges, as the edges (in whichever form) can be retrieved from incidence lists. At that point, for queue-based visits  $E$  and  $S$  can be stored one after another in the same vector, because for each vertex dequeued from  $S$  there is at most a vertex pushed on  $E$ .

We note that this consideration can be extended to stack-based visits, by using a single vector to store both  $S$  and  $E$ : one stack grows upwards, the other downwards. The visit and assignment phases can then be performed at the price of an index per vertex, at the cost of a slightly slower assignment phase: we call this strategy the *low-memory visit*. In the case of static functions using 128-bit signatures, the low-memory visit uses an order of magnitude less memory than the high-memory visit.

Depending on the hardware and on the size of the key set, sometimes the low-memory visit is much faster than the high-memory visit, in spite of the lesser locality. While we prefer the high-memory visit for low parallelism, when significant parallelism is possible we switch to the low-memory visit, which avoids out-of-memory errors due to excessive allocation, and in general reduces by a significant factor the memory that is necessary for the construction, improving the overall performance. When we plan to use lazy Gaussian elimination we peel by index, and then use the low-memory visit.

#### 4.4 Subsigning

We propose to use a technique that we call *subsigning* to speed up construction, and, marginally, query time. The idea is to use just a part of the bits of the signature, the *local signature*, to generate edges (and hashes for filters) for a shard, as the reduced number of keys in a shard makes the probability of a collision much smaller. Most of the operations discussed in the previous section use signatures, so being able to reduce their size means reducing memory usage and operations.

One has however to be careful to avoid that subsigning leads to unwanted collisions of local signatures. For example, if we use 128-bit signatures and we subsign with 64 bits we can build

structures with MWHC up to roughly a hundred billion keys, and slightly less for fuse graphs. In practice, one treats duplicate local signatures similarly to duplicate signatures, with the difference that the user can circumvent the problem by using more bits for subsigning. In the case of filters, however, we are now going to see that this is not necessary.

## 4.5 Duplicate Signatures and Filters

A final useful observation is that in the case of filters if the hashes used by the filter to check for false positives are derived from the sharding signature, and not directly computed from the keys, as pioneered by the Sux4J implementations [1], duplicate signatures can be simply removed. Keys with the same signature are mapped to the same hash, so there will be no difference in the behavior of the filter. In fact, we can even deduplicate local signatures if they are used for generating both edges and hashes.

Sorting the signatures to find duplicates slows down slightly the construction, but we can use one of the many recent excellent parallel radix sort implementations (see, e.g., [22, 2]). If we also take care of using the first bits after the sharding bits (if any) to compute the segment for fuse graphs using a monotone function, like inversion in fixed-point arithmetic, then sorting for duplicates can play the same role of the count sort by segment. In fact, if the following further bits are used to select the first vertex, the radix sort provides even more locality to the generation phase, getting back some of the overhead of radix sorting with respect to count sorting. The strategy used by Graf and Lemire [17] to generate an edge uses indeed inversion in fixed-point arithmetic, thus satisfying this property, albeit they do not exploit it in their implementation.

## 5 Experiments

We ran experiments on medium-sized to very large integer (a trillion keys) sets; our keys are always 64-bit integers, as they have the shortest hash time. We have in mind use cases with billions of keys and a sizable output, such as index functions<sup>6</sup> on several billion keys, but as in [14] we mostly perform experiments on filters as they subsume the case of functions, and software for benchmarking is available from the authors.

We used a `c7i.metal-24x1` Amazon AWS instance with 48 Intel® Xeon® Platinum 8488C cores and 192 GiB of RAM, endowed with 16TB high-performance SSD storage. Parallel construction code was provided 32 threads to work with. We used `g++ 14.2.0` and `rustc 1.85`.

We have the following setups:

- For power-of-two hash sizes  $b$  (8, 16, and 32 bits), we compare BuRR against fuse graphs with 64- and 128-bit signatures, and again VFILTER, both in the version based on the MWHC construction, and on the version based on fuse graphs and lazy Gaussian elimination (which however at this size is not used). Fuse graphs with 128 bit signatures are not useful in practice, but their benchmarks highlight the cost of different components of the construction. All such structure store their values in a slice of the correct unsigned integer type for the size of the output. BuRR cannot be built at the largest size ( $2^{33}$ ) because of insufficient the core memory. The results are displayed in Figure 1.
- We perform the same tests for non-power-of-two hash sizes (9, 17, and 33 bits). The results are displayed in Figure 2. We store the values of the other constructions in a bit array.

---

<sup>6</sup>An *index function* (AKA *order-preserving minimal perfect hash function*) maps elements of a list to their position in the list.

- We perform again the same tests on queries, but starting 96 instances of the test in parallel, saturating the cores and using a significant part of the memory. The BuRR paper claims that in this kind of test, which should simulate a high parallel workload. BuRR becomes significantly faster than hypergraph-based constructions.
- Finally, we compare static functions mapping integers to themselves: in this case we compare a very fast minimal perfect hash, PTHash-HEM [23], followed by a bit array, with VFUNC in the two settings above. The interest for this experiment stems from a consideration of Demaine, Mayer, Pagh, and Pătraşcu [11] that a MPHf is stronger than a static function, as one can obtain a static function from a MPHf by indexing an array with the result of the MPHf, but not the other way around; this consideration is in contrast with an observation by Dietzfelbinger and Pagh [12] that many kinds of retrieval data structures issue independent memory accesses, which are parallelized efficiently by modern processors. The results are displayed in Figure 4.
- In Figure 5 we compare the sequential construction time for fuse graphs in the sharded (VFILTER) and unsharded version.

Note that in this setting BuRR has an almost negligible space overhead ( $\approx 0.1\%$ ) whereas VFUNC/VFILTER have a space overhead of 10.5% ( $\epsilon = 0.001$ ), and the MWHC construction has a space overhead of 24% ( $\epsilon = 0.01$ ). For the PTHash + array case the overhead is very small, but variable, as it depends on the number of keys.

We try key sets with size  $2^{20}, 2^{21}, \dots, 2^{33}$ , stopping at  $2^{31}$  for constructions using 64-bit hashes, and at  $2^{32}$  for Burr and 128-bit fuse graphs as core memory is not sufficient.

We used the parallel BuRR implementation made available by the authors [3], suitably modified to run tests similar to the Rust counterpart.<sup>7</sup> Since BuRR has several parameters whose influence of the behavior of the filter is not always predictable, in particular in combination, we used the settings provided by the authors named a “fast filter configuration“, which should provide a “good all-around preset for filters”.

For all other constructions, we use the Rust implementation we developed.<sup>8</sup> The implementation uses Rust traits to abstract over the size of the signatures (64 or 128 bits) and the sharding/edge generation logic (MWHC, fuse, etc.), making it possible to obtain monomorphized, optimized versions of all constructions (including bare fuse graphs) using the same code base.

We report the construction time after the input has been read. All constructions depend on some hashing applied to the key, and different implementations have different ways to feed the keys (e.g., the BuRR code assumes all keys are in RAM, whereas we do not).

The results are quite clear-cut:

- For what matters query time, constructions based on hypergraphs are the fastest by a wide margin, in particular for non-power-of-two hash sizes. VFILTER (i.e.,  $\epsilon$ -cost sharding plus fuse graphs) is slightly slower than fuse graphs because of the sharding logic, but the difference is of a few nanoseconds.
- The MWHC construction has ups and downs because as the key set grows, bad cache behavior makes the construction slower, but when (2) increases, we can increase the sharding bits and use twice as many threads.
- Serial construction of BuRR is quite slow compared to all other approaches, but BuRR construction parallelizes in an almost uncanny way, providing almost constant per-key

<sup>7</sup><https://github.com/vigna/BuRR/commit/cf7cd8310b98321f64f3fb9a531b0a9e469bdc15>

<sup>8</sup><https://github.com/vigna/sux-rs/commit/5222dfe073d6e3d5daa4d6b0c4834f81dba5bf1f>

Keys	Shards	Hashing	Construction	Query
$10^{10}$	$2^7$	10.31	10.95	69.97
$10^{11}$	$2^9$	14.22	42.65	75.47
$10^{12}$	$2^{11}$	15.13	45.60	111.32

Table 1: Offline construction and query time for very large 1-bit VFILTER instances. All timings are in ns/key.

construction time, although passing the billion of keys VFILTER becomes enough sharded to have shorter construction times.

- Parallel workloads reduce the difference we notice between queries in the sequential case, which is expected, since memory access becomes slower, but BuRR remains by far the slowest structure (except in a couple of data points). We thus cannot replicate the claims of [3]. We believe that the two main reasons are a more modern memory architecture (DDR5 vs. DDR4) and more engineered implementation of the hypergraph-based constructions.
- Hypergraph-based constructions become increasingly slower as the key set becomes larger due to bad locality, but as soon as we enter the sharding phase construction time per key drops very quickly.
- One can see clearly the early start of sharding for MWHC compared to the fuse graphs used by VFUNC. Nonetheless, in the end the better locality of fuse graphs wins.
- Using an intermediate state-of-the-arc MPHf followed by access to an array of values is much slower than using a static function like VFUNC. The dependency chain in memory accesses kills performance.
- Sharding not only enables parallelism, but also improves construction time even in the sequential case for large shards due to improved locality.

We conclude that the 9% space gain of BuRR comes at a significant performance cost on query and sequential construction times, in particular in the case of large key sets or large output sizes, and that  $\epsilon$ -cost sharding has a minimal impact on query time, while keeping space occupancy under control.

We note that the comparisons in Figure 2 is slightly unfair to VFILTER, which in these benchmarks has a dynamically chosen output size, whereas the code for BuRR provided by the authors has the output bit width fixed at compile time, which might be considered acceptable for filter application, but it is entirely unrealistic for static functions (e.g., index functions have an output bit size that depends on the size of the key set).

## 6 Large-Scale Offline Construction

VFUNC and VFILTER can be indifferently built using in-memory or on-disk sharding. For all previous tests, all construction for all structures was in-memory, but we report in Table 1 the offline construction and query times for very large 1-bit VFILTER instances, showing that our approach scales to trillions of keys. Note that since the data structure is built sequentially, one shard at a

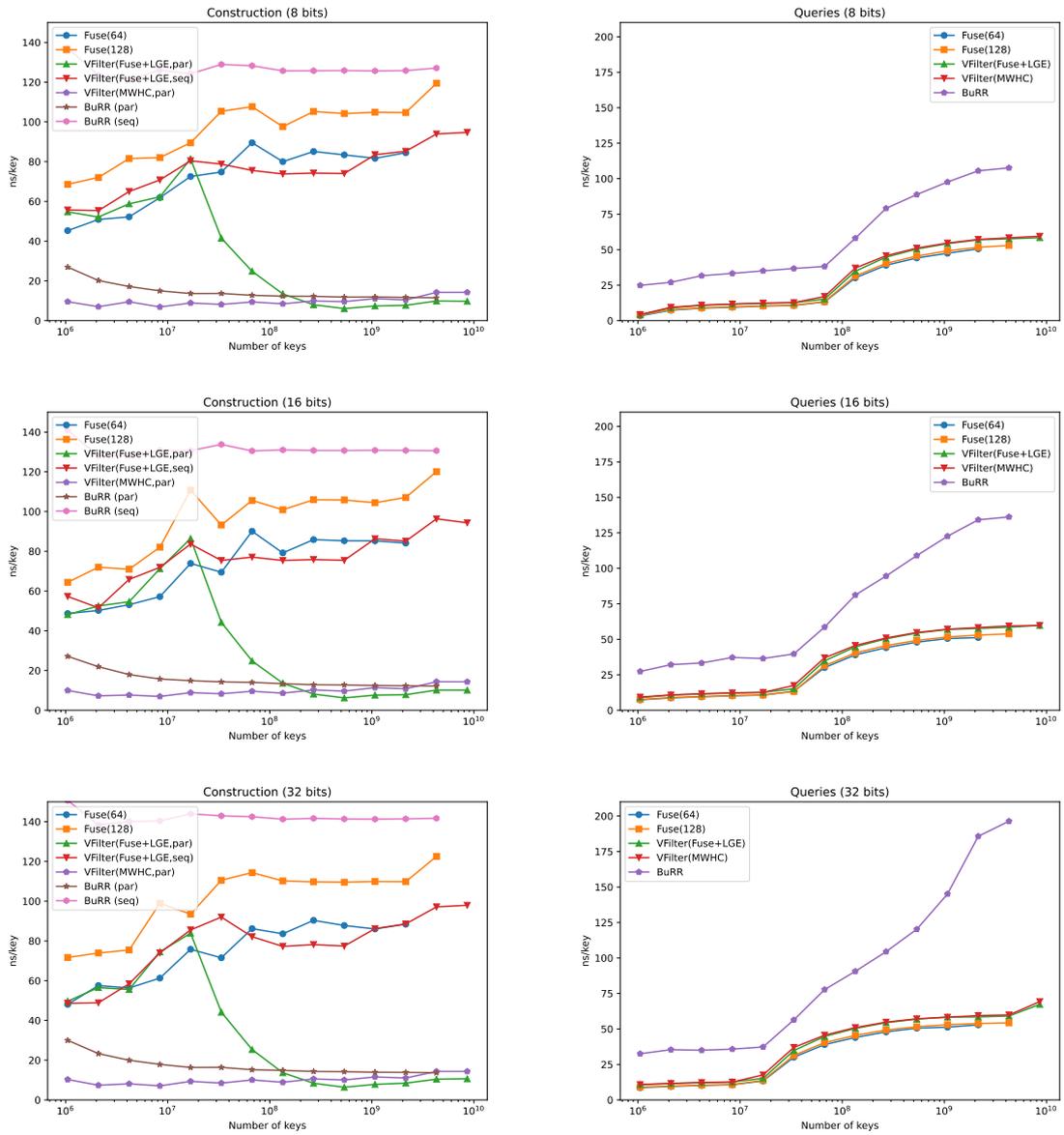


Figure 1: Construction and query times for power-of-two bit sizes. BuRR uses the sparse-coefficient variant.

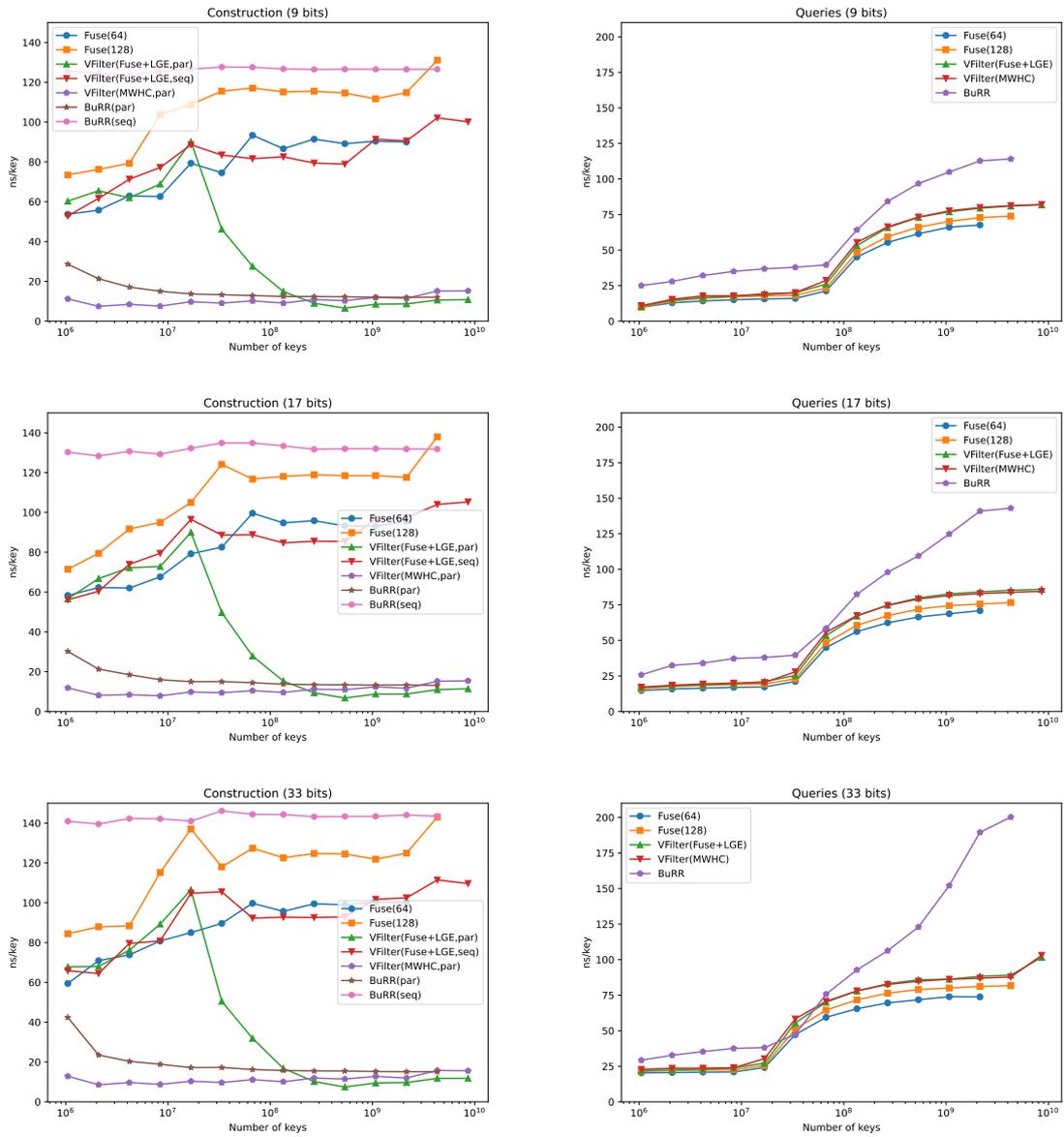


Figure 2: Construction and query times for power-of-two-plus-one bit sizes. BuRR uses the interleaved-coefficient variant. Note that the graphs on the right have a very different scale on the vertical axis.

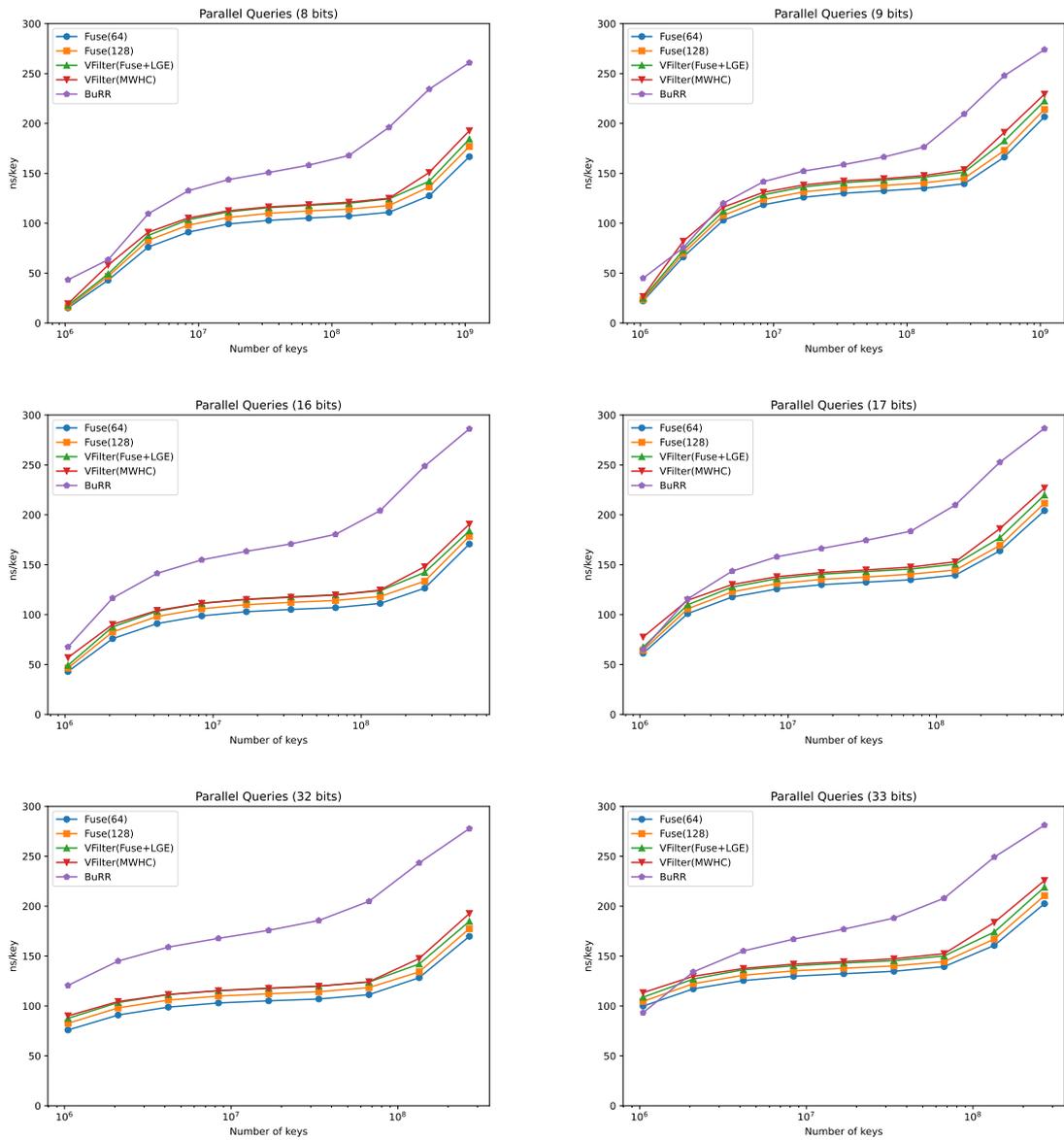


Figure 3: Query times under CPU load and memory stress.

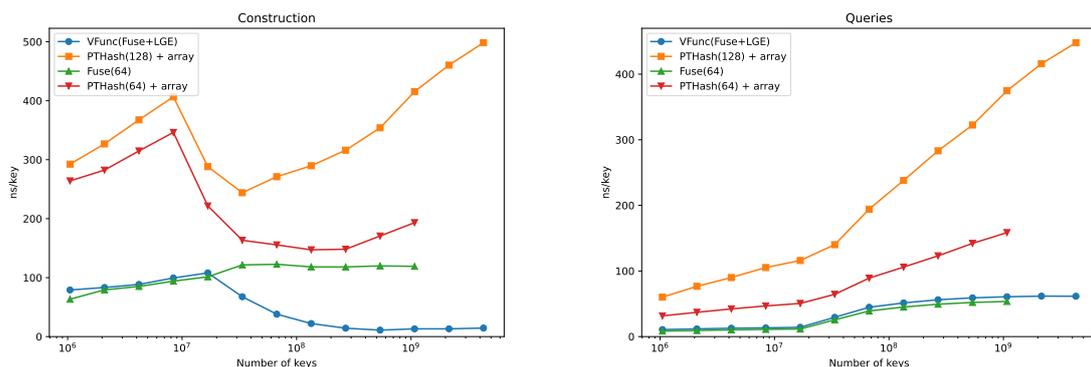


Figure 4: Construction and query times for index functions.

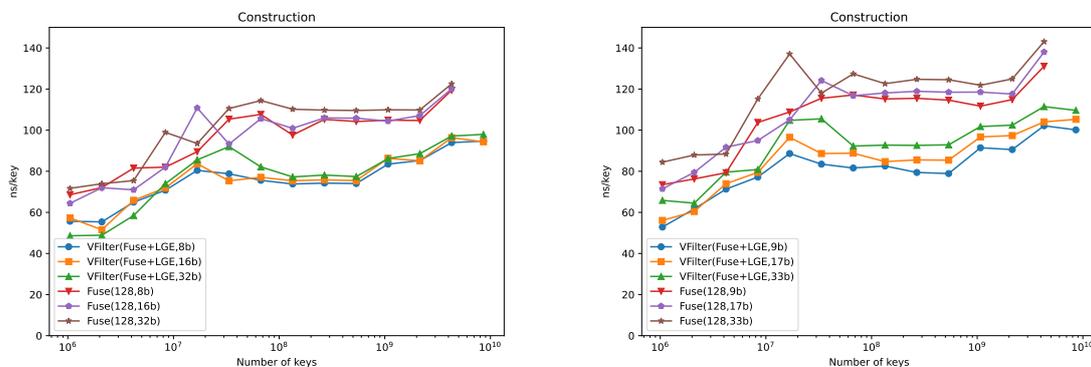


Figure 5: Construction time in the sequential case for sharded (VFILTER) and non-sharded fuse graphs. The spike at the end is due to the necessity of forcing low-memory visits to be able to build a structure with  $2^{32}$  keys within the memory available in the unsharded case; building a structure with  $2^{33}$  keys failed for insufficient memory.

time, it can be serialized on disk without ever being materialized in memory, and possibly just memory-mapped.

## 7 Conclusions

We presented two new static function/filter data structures, VFUNC and VFILTER, based on  $\epsilon$ -cost sharding applied to a combination of fuse graphs and lazy Gaussian elimination. While the space used by our data structures is 10.5% larger than the optimum (11 – 12% for small sizes), their query performance on very large key sets, where peeling a single fuse graphs is unfeasible, is the best currently available. Moreover, due to  $\epsilon$ -cost sharding they use less memory at construction time, and construction can be carried out in parallel or distributed fashion. For large-scale applications where query time is most relevant, they provide the fastest data structures available that can scale to trillions of keys.

There are several open problems and further research directions. It might be interesting to

experiment with 4-uniform fuse graphs, albeit having an additional memory access per key would significantly reduce performance, with a minor relative overhead decrease of  $\approx 5\%$ .

More interestingly, it would be useful to fill all the theoretical gaps in the construction of fuse graphs, which would lead to filling similar gaps in our data structures. Having precise probability bounds on the peelability of fuse graphs would make it possible to tune very finely the locality of the construction vs. the probability of duplicate edges. For example, slightly increasing the overhead leads experimentally to a wider range of peelable segment sizes, but without theoretical guarantees exploring the space of parameters is a daunting task. Indeed, the contrast between the estimates (6) and (7) suggests that the range of values of  $\ell$  for which fuse graphs are peelable increases significantly with  $n$ .

Another interesting bound to derive would be a precise estimation of the probability of duplicate local signatures when subsigning. The problem of the analysis lies in the fact that  $F_n(t)$  has very good estimates for  $t \ll n$ , but it is not easily tractable in the case that would be interesting for us.

## Acknowledgements

We thank Dario Moschetti for the Rust port of the lazy Gaussian elimination Java code from Sux4J, and Hans-Peter Lehmann for his help in using the BuRR code.

This work is supported by project SERICS (PE00000014) under the NRRP MUR program funded by the EU - NGEU. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the Italian MUR. Neither the European Union nor the Italian MUR can be held responsible for them.

## References

- [1] Sux4J 3.1. Succinct data structure for Java, 2014. <https://github.com/vigna/Sux4J/commit/a3325e107c18>.
- [2] Michael Axtmann, Sascha Witt, Daniel Ferizovic, and Peter Sanders. Engineering in-place (shared-memory) sorting algorithms. *ACM Trans. Parallel Comput.*, 9(1), January 2022.
- [3] Matthias Becht, Hans-Peter Lehmann, and Peter Sanders. Brief announcement: Parallel construction of bumped ribbon retrieval, 2024. arXiv:2411.12365.
- [4] Djamel Belazzougui, 2007. Personal electronic communication.
- [5] Djamel Belazzougui, Paolo Boldi, Giuseppe Ottaviano, Rossano Venturini, and Sebastiano Vigna. Cache-oblivious peeling of random hypergraphs. In *2014 Data Compression Conference (DCC 2014)*, pages 352–361. IEEE, 2014.
- [6] Fabiano C. Botelho, Rasmus Pagh, and Nivio Ziviani. Practical perfect hashing in nearly optimal space. *Information Systems*, 38(1):108–131, 2013.
- [7] Fabiano C. Botelho, Nicholas Wormald, and Nivio Ziviani. Cores of random  $r$ -partite hypergraphs. *Information Processing Letters*, 112(8):314–319, 2012.
- [8] Alex D. Breslow and Nuwan S. Jayasena. Morton filters: fast, compressed sparse cuckoo filters. *The VLDB Journal*, 29(2):731–754, 2020.
- [9] Larry Carter, Robert Floyd, John Gill, George Markowsky, and Mark Wegman. Exact and approximate membership testers. In *Proceedings of Symposium on Theory of Computation (STOC '78)*, pages 59–65. ACM Press, 1978.

- [10] Bernard Chazelle, Joe Kilian, Ronitt Rubinfeld, and Ayellet Tal. The Bloomier filter: an efficient data structure for static support lookup tables. In J. Ian Munro, editor, *Proceedings of the Fifteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2004*, pages 30–39. SIAM, 2004.
- [11] Erik D. Demaine, Friedhelm Meyer auf der Heide, Rasmus Pagh, and Mihai Pătraşcu. De dictionariis dynamicis paucis spatii utentibus (*lat.* on dynamic dictionaries using little space). In *LATIN*, volume 3887 of *Lecture Notes in Computer Science*, pages 349–361. Springer, 2006.
- [12] Martin Dietzfelbinger and Rasmus Pagh. Succinct data structures for retrieval and approximate membership (extended abstract). In Luca Aceto, Ivan Damgård, Leslie Ann Goldberg, Magnús M. Halldórsson, Anna Ingólfssdóttir, and Igor Walukiewicz, editors, *Automata, Languages and Programming, 35th International Colloquium, ICALP 2008, Proceedings, Part I: Track A: Algorithms, Automata, Complexity, and Games*, volume 5125 of *Lecture Notes in Computer Science*, pages 385–396. Springer, 2008.
- [13] Martin Dietzfelbinger and Stefan Walzer. Dense Peelable Random Uniform Hypergraphs. In Michael A. Bender, Ola Svensson, and Grzegorz Herman, editors, *27th Annual European Symposium on Algorithms (ESA 2019)*, volume 144 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 38:1–38:16, Dagstuhl, Germany, 2019. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.
- [14] Peter C. Dillinger, Lorenz Hübschle-Schneider, Peter Sanders, and Stefan Walzer. Fast Succinct Retrieval and Approximate Membership Using Ribbon. In Christian Schulz and Bora Uçar, editors, *20th International Symposium on Experimental Algorithms (SEA 2022)*, volume 233 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 4:1–4:20, Dagstuhl, Germany, 2022. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.
- [15] Marco Genuzio, Giuseppe Ottaviano, and Sebastiano Vigna. Fast scalable construction of ([compressed] static | minimal perfect hash) functions. *Information and Computation*, 273:104517, 2020.
- [16] Thomas Mueller Graf and Daniel Lemire. Xor filters: Faster and smaller than bloom and cuckoo filters. *Journal of Experimental Algorithmics (JEA)*, 25:1–16, 2020.
- [17] Thomas Mueller Graf and Daniel Lemire. Binary fuse filters: Fast and smaller than Xor filters. *ACM J. Exp. Algorithmics*, 27, 2022.
- [18] Ragnar Groot Koerkamp. PtrHash: Minimal perfect hashing at RAM throughput, 2025. arXiv:2502.15539.
- [19] Tobias Maier, Peter Sanders, and Robert Williger. Concurrent Expandable AMQs on the Basis of Quotient Filters. In Simone Faro and Domenico Cantone, editors, *18th International Symposium on Experimental Algorithms (SEA 2020)*, volume 160 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 15:1–15:13, Dagstuhl, Germany, 2020. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.
- [20] Bohdan S. Majewski, Nicholas C. Wormald, George Havas, and Zbigniew J. Czech. A family of perfect hashing methods. *Comput. J.*, 39(6):547–554, 1996.
- [21] Michael Molloy. Cores in random hypergraphs and boolean formulas. *Random Struct. Algorithms*, 27(1):124–135, 2005.

- [22] Omar Obeya, Endrias Kahssay, Edward Fan, and Julian Shun. Theoretically-efficient and practical parallel in-place radix sorting. In *The 31st ACM Symposium on Parallelism in Algorithms and Architectures, SPAA '19*, page 213–224, New York, NY, USA, 2019. Association for Computing Machinery.
- [23] Giulio Ermanno Pibiri and Roberto Trani. Parallel and external-memory construction of minimal perfect hash functions with PTHash. *IEEE Transactions on Knowledge and Data Engineering*, 36(3):1249–1259, 2023.
- [24] Boris Pittel and Gregory B. Sorkin. The satisfiability threshold for  $k$ -XORSAT. *Combinatorics, Probability and Computing*, (2):236–268, 2016.
- [25] Felix Putze, Peter Sanders, and Johannes Singler. Cache-, hash-, and space-efficient Bloom filters. *Journal of Experimental Algorithmics (JEA)*, 14, 2010.
- [26] Martin Raab and Angelika Steger. “Balls into bins” — a simple and tight analysis. In Michael Luby, José D. P. Rolim, and Maria Serna, editors, *Randomization and Approximation Techniques in Computer Science*, pages 159–170, Berlin, Heidelberg, 1998. Springer Berlin Heidelberg.