Universal Hashing via Integer Arithmetic Without Primes, Revisited

Martin Dietzfelbinger(✉)

Department of Computer Science and Automation,
Technische Universität Ilmenau, Ilmenau, Germany
martin.dietzfelbinger@tu-ilmenau.de

Abstract. Let $u, m \geq 1$ be arbitrary integers and let $r \geq (u-1)m$ be any multiple of m. Consider the multiset ("*class*") $\mathcal{H}^{\text{lin}}_{u,m,r} = \{h_{a,b} \mid 0 \leq a, b < r\}$ of *hash functions* from $U = \{0, \ldots, u-1\}$ to $M = \{0, \ldots, m-1\}$, where $h_{a,b}(x) = \lfloor ((ax+b) \bmod r)/(r/m) \rfloor$, for $x \in U$. In a STACS paper of 1996 it was shown that $\mathcal{H}^{\text{lin}}_{u,m,r}$ is $\frac{5}{4}$-approximately 2-wise independent and (error-free) 2-wise independent if in addition r is a power of a prime number. Here, we revisit this result. We prove slightly stronger bounds (in part also shown by Woelfel (1999) with different methods) and we discuss applications that have appeared in the meantime, e.g., complexity lower bounds for integer multiplication, fine-grained complexity inside the class P, and the usefulness and limitations of these simple hash classes in combination with other applications of hashing.

Dedicated to Juraj Hromkovič on the occasion of his 60th birthday.

1 Introduction

Let $u, m \geq 1$ be arbitrary integers and let $r \geq (u-1)m$ be any multiple of m. Consider the "*class*" $\mathcal{H}^{\text{lin}}_{u,m,r} = \{h_{a,b} \mid 0 \leq a, b < r\}$ of *hash functions* from $U = [u] = \{0, \ldots, u-1\}$ to $M = [m] = \{0, \ldots, m-1\}$, where

$$h_{a,b}(x) = \lfloor ((ax+b) \bmod r)/(r/m) \rfloor, \text{ for } x \in U.$$

In a STACS paper [10] of 1996 it was shown that $\mathcal{H}^{\text{lin}}_{u,m,r}$ is $\frac{5}{4}$-approximately 2-wise independent and (error-free) 2-wise independent if in addition r is a power of a prime number. In the present paper, we revisit the contribution of 1996, describe improvements of the results and discuss ramifications and applications that have appeared in the meantime.

1.1 Background: Two-Wise Independence and Universality

Two-wise independent random variables and hash functions have a multitude of applications. We mention just two: universal hashing for use in data structures

H.-J. Böckenhauer et al. (Eds.): Hromkovič Festschrift, LNCS 11011, pp. 257–279, 2018.
https://doi.org/10.1007/978-3-319-98355-4_15

[7,13,18,26,29,40] and amplifying randomness and derandomization [2,8,11,23,28]. For a classical survey of applications, see [24].

Given sets U (the *universe*, usually finite) and M (the *range*, finite), we call a function $h\colon U \to M$ a *hash function*. Elements of U are called *keys*. We consider "*classes*" (multisets) $\mathcal{H}$ of hash functions. Such a class is called 1-*universal* if for h chosen uniformly at random from $\mathcal{H}$ the collision probability is bounded by $1/|M|$, i.e., if we have

$$\mathbf{Pr}(h(x_1) = h(x_2)) \leq \frac{1}{|M|}, \text{ for } x_1, x_2 \in U \text{ distinct.} \tag{1}$$

Class $\mathcal{H}$ is called 2-*wise independent* if for h chosen at random from $\mathcal{H}$ we have that for arbitrary distinct $x_1, x_2 \in U$ the random variables $h(x_1)$ and $h(x_2)$ are uniformly and independently distributed in M, more precisely:

$$\mathbf{Pr}(h(x_1) = i_1 \wedge h(x_2) = i_2) = \frac{1}{|M|^2}, \text{ for } x_1, x_2 \in U \text{ distinct, } i_1, i_2 \in M. \tag{2}$$

For applications, it is usually sufficient if (1) and (2) hold up to some relative error, i.e., if we require that $\mathbf{Pr}(h(x_1) = h(x_2)) \leq c/|M|$ in (1) for some $c \geq 1$ (such a class is called c-*universal*, which corresponds to the original definition of the term in [7]) or $(2-c)/|M|^2 \leq \mathbf{Pr}(h(x_1) = i_1 \wedge h(x_2) = i_2) \leq c/|M|^2$ (2) for some $c \in [1,2)$ (then we say the class is c-*approximately* 2-*wise independent*).

By 1996, simple methods for constructing such classes had been known for quite some time. We list examples of such constructions.[1]

(a) $U = [p]$, $M = [m]$ for a prime p and some $m \leq p$, $h \in \mathcal{H}$ is given by $h(x) = ((ax+b) \bmod p) \bmod m$, for $a, b \in [p]$. This class is c-approximately 2-wise independent for $c = (\lceil p/m \rceil\, m/p)^2 \leq 1 + 3m/p$. For $a, b \in [p]$, $a \neq 0$, it is 1-universal [7,18,40].

(b) $U = \mathbb{F}$ is a finite field of prime characteristic p, $M = \mathbb{Z}_p^\mu$ for some μ, $\xi\colon \mathbb{F} \to M$ is some $\mathbb{Z}_p$-linear projection, $h \in \mathcal{H}$ is given by $h(x) = \xi(ax+b)$, for $a, b \in \mathbb{F}$. This class is 2-wise independent.

(c) $U = \mathbb{F}^\mu$ for a finite field $\mathbb{F}$, $M = \mathbb{F}^\nu$, $h \in \mathcal{H}$ given by $h(x) = A \cdot x + b$, for $A \in \mathbb{F}^{\nu \times \mu}$, $b \in M = \mathbb{F}^\nu$. This class is 2-wise independent.

(d) $U = \mathbb{F}^\mu$ for a finite field $\mathbb{F}$, $M = \mathbb{F}^\nu$, $h \in \mathcal{H}$ given by $h(x) = a \circ x + b$, for $a \in \mathbb{F}^{\mu+\nu-1}$, $b \in M = \mathbb{F}^\nu$, where $\circ$ denotes convolution. This class is 2-wise independent [25].

We note that for implementing such classes we either need prime numbers or representations of the arithmetic in finite fields that have size $|U|$ or at least size $|M|$, or, for (c) or (d), we have to carry out vector-matrix multiplication or a convolution over some finite field $\mathbb{F}$, the most natural case being $\mathbb{F}_2 = \{0,1\}$. The main purpose of [10] was to provide a construction using only plain integer

[1] For a natural number n, we denote the set $\{0,1,\ldots,n-1\}$ by $[n]$; for a prime number p, we denote by $\mathbb{Z}_p$ the field of size p, with ground set $[p]$.

arithmetic without the need for prime numbers to obtain two-wise independent hash families.

A simple first step in this direction had been made in [12]. There, for $k \geq \mu \geq 1$ and sets $U = [2^k]$ and $M = [2^\mu]$ the class $\mathcal{H}^{\text{mult}}_{2^k,2^\mu}$ of *multiplicative functions*

$$h_a \colon U \ni x \mapsto \left\lfloor (ax \bmod 2^k)/2^{k-\mu} \right\rfloor \in M$$

for $0 < a < 2^k$ odd was studied. This class is *not* 2-wise independent; hash values $h(x)$ for a key x are not even uniformly distributed. However, it is 2-universal. Its big advantage is that functions from the class can be evaluated very efficiently on a computer—one only needs one multiplication and one or two shift operations. The construction in [10] is a variant of this class.

1.2 Definitions and Properties of the Class

Notation. The universe is $U = [u]$, the range is $M = [m]$, for positive integers u and m. We will calculate in the ring $\mathbb{Z}_r$ of integers with operations modulo r, where $r \geq 2$ is a suitable integer. In order to keep the notation simple, we will identify the ground set of $\mathbb{Z}_r$ with $[r]$. Arithmetic operations modulo r are denoted by $\cdot_r$, $+_r$, $-_r$. If a positive integer x divides an integer y, we write $x \mid y$, otherwise $x \nmid y$.

Definition 1. *Let u, m, and $r \geq m$ be given, where $m \mid r$. Let $k = r/m$.*

(a) *For $a, b \in \mathbb{Z}_r$ define*

$$\begin{aligned} g_{a,b}(x) &= a \cdot_r x +_r b, \textit{ for } x \in U, \textit{ and} \\ h_{a,b}(x) &= \lfloor g_{a,b}(x)/k \rfloor, \textit{ for } x \in U. \end{aligned}$$

(b) *The class of* linear hash functions *from U to M modulo r is the multiset*

$$\mathcal{H}^{\text{lin}}_{u,m,r} = \{h_{a,b} \mid a, b \in \mathbb{Z}_r\}.$$

This class has size r^2; representing (choosing) a function from the class requires $2\lceil \log r \rceil$ (random) bits. Following [10], Woelfel [41,42] gave several related constructions of substantially smaller subclasses that exhibit behaviour similar to that of $\mathcal{H}^{\text{lin}}_{u,m,r}$.

The basic result of [10] and of this paper is that $\mathcal{H}^{\text{lin}}_{u,m,r}$ is approximately 2-wise independent if r is large enough. (This will be proved in Sect. 3.)

Theorem 1. (i) *If $m, u \geq 2$ and $r \geq (u-1)m$ is a multiple of m, then $\mathcal{H}^{\text{lin}}_{u,m,r}$ is $\frac{9}{8}$-approximately 2-wise independent. More precisely, for $i \in M$, $x \in U$ we have*

$$\mathbf{Pr}(h(x) = i) = \frac{1}{m},$$

and for arbitrary $i_1, i_2 \in M$ and distinct $x_1, x_2 \in U$ we have

$$\frac{2-c}{m^2} \leq \frac{1}{cm^2} \leq \mathbf{Pr}(h(x_1) = i_1 \wedge h(x_2) = i_2) \leq \frac{c}{m^2},$$

where $c = c(u, m, r) = (4z^2 + 4z + 1)/(4z^2 + 4z)$*, for some* $z \geq \lfloor r/((u-1)m) \rfloor$*.*

(ii) *If* r *and* m *are powers of a prime* p*, and* $r \geq um/p$*, then* $\mathcal{H}^{\text{lin}}_{u,m,r}$ *is* 2*-wise independent. (The most natural value for* p *is* 2*.)*

Section 2 contains technical preparations involving "gap matrices", which form the basis for the proof of the main theorem in Sect. 3. In Subsect. 4.1 the main result is extended to keys that are represented as sequences of numbers, a standard variation of hash classes, which leads to a more efficient evaluation for very long keys. In Subsect. 4.3 we show that the approach is sufficient for constructing sequences in $M = [m]$ that are sufficiently close to two-independence for carrying out two-independent sampling in the sense of [8]. As an example, we show that for $r \geq u^{3/2} \cdot m$ the sequence

$$\lfloor (ax + b) \bmod r)/(r/m) \rfloor\,,\ 0 \leq x < u,$$

where $a, b \in [r]$ are chosen uniformly at random, is suitable. Subsection 4.4 deals with the problem of "collapsing the universe": Given a subset $S \subseteq U$ of n keys from $U = [u]$, transform the long keys $x \in U$ into ones (x') of length $O(\log\log u + \log n)$ such that this transformation is one-to-one on S. We give a construction that uses just a linear hash function from $\mathcal{H}^{\text{lin}}_{u,m,r}$ to achieve the length $\log r = O(\log\log u + \log n)$.

Finally, Sect. 5 discusses various applications of and observations about the hash class $\mathcal{H}^{\text{lin}}_{u,m,r}$ that have appeared in the literature since it was first described in 1996.

2 Gap Matrices

This section provides bounds on the number of 1's in 0–1 matrices with a certain shape.

Definition 2. *Let* $\gamma, k, \ell \geq 1$ *be integers. A* $k \times \ell$ *matrix* $A = (a_{ij})_{0 \leq i < k, 0 \leq j < \ell}$ *with entries from* $\{0, 1\}$ *is called a* gap-γ (diagonal) matrix *if the* 1*'s are arranged in diagonals of (horizontal/vertical) distance* γ*, i.e., if there is some* $t \in [\gamma]$ *such that*

$$a_{ij} = 1 \iff j - i \equiv t \pmod{\gamma},\ \textit{for}\ 0 \leq i < k, 0 \leq j < \ell.$$

Figure 1 shows examples of gap-γ square matrices. By N_A we denote the number of 1's in a 0–1 matrix A. In a $k \times \ell$ matrix we expect N_A to be about $k\ell/\gamma$, if $\gamma \leq k, \ell$. Even in $k \times k$ matrices there will be deviations if $\gamma \nmid k$, as demonstrated in Fig. 1. We provide bounds on the relative deviation $N_A/(k\ell/\gamma)$, for $k \times \ell$ matrices A.

Proposition 1. *Assume* A *is a gap-*γ *matrix of dimensions* $k \times \ell$ *with* $k, \ell \geq \gamma$*. Then we have the following:*

(a) *If* $\gamma \mid k$ *or* $\gamma \mid \ell$*, then* $N_A = k\ell/\gamma$*.*

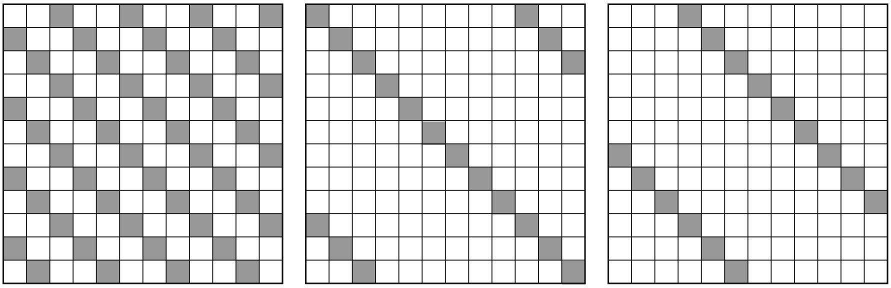

Fig. 1. Some $k \times k$ gap-γ matrices, with $k = 12$. 1's are represented as grey squares, 0's as white ones. *Left*: $\gamma = 3$, which divides k. There are exactly k^2/γ 1's. *Middle*: $\gamma = 9$, with $18 = \frac{9}{8}k^2/\gamma$ many 1's. *Right*: $\gamma = 9$, with $15 = \frac{15}{16}k^2/\gamma$ many 1's.

(b) *If* $z = \lfloor k/\gamma \rfloor = \lfloor \ell/\gamma \rfloor$, *then*

$$\frac{1}{c_z} \;\le\; \frac{N_A}{k^2/\gamma} \;\le\; c_z, \textit{ for } c_z = 1 + \frac{1}{4z(z+1)}.$$

(c) *If* $y = \lfloor k/\gamma \rfloor \le z = \lfloor \ell/\gamma \rfloor$, *then*

$$2 - c_{y,z} \;\le\; \frac{N_A}{k\ell/\gamma} \;\le\; c_{y,z}, \textit{ for } c_{y,z} = 1 + \frac{1}{4y(z+1)}.$$

Remark 1. (i) Note that $1/c_z = 4z(z+1)/(4z(z+1)+1) = 1 - 1/(2z+1)^2$ and $2 - c_{y,z} = 1 - 1/(4y(z+1))$. In (b), the upper bound is $\le \frac{9}{8}$, the lower bound $\ge \frac{8}{9}$. Asymptotically, the upper and lower bounds in (b) and (c) are $1 \pm O(\gamma^2/k^2)$ and $1 \pm O(\gamma^2/(k\ell))$, respectively.

(ii) A bound $|N_A/(k\ell/\gamma) - 1| \le \gamma^2/(4k\ell)$ for $k \times \ell$ gap-γ matrices was proved in [10, Lemma 8(c)]. It depends on the concrete values of k, ℓ, and γ if that bound or the one from Proposition 1(b) and (c) is stronger. For γ slightly smaller than $k = \ell$ (hence $y = z = 1$) the new bound is smaller by essentially a factor of 2.

(iii) The bounds in (b) are optimal. (See Remark 2 below.)

Proof (of Proposition 1). (a) If $\gamma = \ell$, each row contains exactly one 1, and hence $N_A = k$. (For an example see Fig. 2.) If $\gamma \mid \ell$, we may partition A into ℓ/γ many disjoint $k \times \gamma$ gap-γ matrices. So $N_A = (\ell/\gamma) \cdot k = k\ell/\gamma$. The cases $\gamma = k$ and $\gamma \mid k$ are similar.

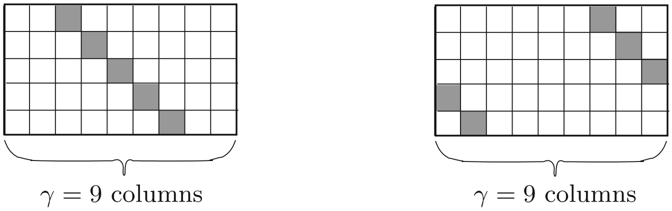

Fig. 2. A gap-γ matrix with $k = 5, \gamma = \ell = 9$ contains k many 1's.

(b) We can assume w.l.o.g. that $k \le \ell$. (Otherwise consider the transpose of A.) Because of (a), we may assume that $\gamma \nmid k$ and $\gamma \nmid \ell$. Let $k' = k \bmod \gamma$ and $\ell' = \ell \bmod \gamma$. With $\delta = k'/\gamma$ and $\varepsilon = \ell'/\gamma$ we have $0 < \delta \le \varepsilon < 1$ and $k = (z+\delta)\gamma$ and $\ell = (z+\varepsilon)\gamma$.

We partition A into four matrices, splitting after the $z\gamma$'th column and $z\gamma$'th row. We get a $(k-k') \times (\ell-\ell')$ submatrix A_1, a $(k-k') \times \ell'$ submatrix A_2, a $k' \times (\ell-\ell')$ submatrix A_3, and a $k' \times \ell'$ submatrix A_4. Since γ divides $k-k'$ and $\ell-\ell'$, part (a) gives that

$$N_{A_1} + N_{A_2} + N_{A_3} = \frac{(k-k')(\ell-\ell') + (k-k')\ell' + k'(\ell-\ell')}{\gamma} = \frac{k\ell - k'\ell'}{\gamma}. \quad (3)$$

So

$$\frac{N_A}{k\ell/\gamma} - 1 = \frac{N_{A_4}}{k\ell/\gamma} - \frac{k'\ell'}{k\ell} = \frac{N_{A_4} - k'\ell'/\gamma}{k\ell/\gamma}. \quad (4)$$

We need to bound this error term.

Upper Bound. Since no row in A_4 contains more than one 1, we have $N_{A_4} \le k'$. Hence

$$\frac{N_{A_4} - k'\ell'/\gamma}{k\ell/\gamma} \le \frac{k' - k'\ell'/\gamma}{k\ell/\gamma} = \frac{\delta\gamma - \delta\varepsilon\gamma}{(z+\delta)(z+\varepsilon)\gamma} = \frac{\delta(1-\varepsilon)}{(z+\delta)(z+\varepsilon)}.$$

Since $\delta \le \varepsilon$, this implies that $(N_{A_4} - k'\ell'/\gamma)/(k\ell/\gamma) \le \delta(1-\delta)/(z+\delta)^2$. Standard methods yield that the last fraction is maximal for $\delta = \delta_z = z/(1+2z)$; the corresponding maximal value is $\delta_z(1-\delta_z)/(z+\delta_z)^2 = 1/(4z(1+z))$. With (4) we get $N_A/(k\ell/\gamma) \le 1 + 1/(4z(1+z))$, which is the claimed upper bound.

Lower Bound. According to (4), we must show that $(k'\ell'/\gamma - N_{A_4})/(k\ell/\gamma) \le 1/(2z+1)^2$.

Case 1: $\delta + \varepsilon \le 1$, or equivalently $k' + \ell' \le \gamma$. — In this case we have

$$\frac{k'\ell'/\gamma - N_{A_4}}{k\ell/\gamma} \le \frac{k'\ell'}{k\ell} = \frac{\delta\varepsilon}{(z+\delta)(z+\varepsilon)}. \quad (5)$$

We observe that the last expression cannot be maximal if $\delta < \varepsilon$. (Assume $\delta < \varepsilon$. Define a mapping

$$D\colon [0, \varepsilon-\delta] \ni \zeta \mapsto \ln\left(\frac{(\delta+\zeta)(\varepsilon-\zeta)}{(z+(\delta+\zeta))(z+(\varepsilon-\zeta))}\right)$$
$$= \ln(\delta+\zeta) + \ln(\varepsilon-\zeta) - \ln(z+\delta+\zeta) - \ln(z+\varepsilon-\zeta).$$

Since

$$\frac{d}{d\zeta} D(\zeta)\Big|_{\zeta=0} = \frac{1}{\delta} - \frac{1}{\varepsilon} - \frac{1}{z+\delta} + \frac{1}{z+\varepsilon} = \frac{z}{\delta(z+\delta)} - \frac{z}{\varepsilon(z+\varepsilon)} > 0,$$

the value $\delta\varepsilon/((z+\delta)(z+\varepsilon))$ cannot be maximal.) Thus we only have to bound $\delta^2/(z+\delta)^2$ over all $\delta \le \frac{1}{2}$. Clearly, this maximum is $\left(\frac{1}{2}\right)^2/\left(z+\frac{1}{2}\right)^2 = 1/(2z+1)^2$.

Case 2: $\delta + \varepsilon > 1$, or equivalently $k' + \ell' > \gamma$. — In this case A_4 contains at least $k' + \ell' - \gamma = (\delta + \varepsilon - 1)\gamma$ many 1's. Hence

$$\frac{k'\ell'/\gamma - N_{A_4}}{k\ell/\gamma} \le \frac{\delta\varepsilon\gamma - (\delta + \varepsilon - 1)\gamma}{(z+\delta)(z+\varepsilon)\gamma} = \frac{(1-\delta)(1-\varepsilon)}{(z+\delta)(z+\varepsilon)}. \tag{6}$$

As above, we observe that the last expression cannot be maximal for $\delta < \varepsilon$. (Define $D\colon [0, \varepsilon - \delta] \ni \zeta \mapsto \ln(1 - (\delta + \zeta)) + \ln(1 - (\varepsilon - \zeta)) - \ln(z + (\delta + \zeta)) - \ln(z + (\varepsilon - \zeta))$. Then

$$\begin{aligned}\frac{d}{d\zeta} D(\zeta)\Big|_{\zeta=0} &= -\frac{1}{1-\delta} + \frac{1}{1-\varepsilon} - \frac{1}{z+\delta} + \frac{1}{z+\varepsilon} \\ &= \frac{z+1}{(1-\varepsilon)(z+\varepsilon)} - \frac{z+1}{(1-\delta)(z+\delta)}.\end{aligned}$$

This is positive, since the function $\tau \mapsto (1-\tau)(z+\tau)$ is decreasing for $\tau \in (0,1)$.) Thus we only have to bound $(1-\delta)^2/(z+\delta)^2$ over all $\delta > \frac{1}{2}$. This expression is bounded by $\left(\frac{1}{2}\right)^2/\left(z+\frac{1}{2}\right)^2 = 1/(2z+1)^2$.

(c) We now turn to matrices that are not necessarily almost square: Let A be a $k \times \ell$ gap-γ matrix with $y = \lfloor k/\gamma \rfloor \le z = \lfloor \ell/\gamma \rfloor$. Since the other cases have been covered already, we may assume that $y < z$ and that $\gamma \nmid k$ and $\gamma \nmid \ell$. Let $\delta = k/\gamma - y$ and $\varepsilon = \ell/\gamma - z$. Then $0 < \delta, \varepsilon < 1$. We extend A to an almost square gap-γ matrix $\bar{A}$ by adding $(z-y)\gamma$ rows at the bottom. Let $\bar{k} = (z+\delta)\gamma$ be the number of rows in $\bar{A}$. Applying part (b) yields

$$\frac{4z(z+1)}{(2z+1)^2} \;\le\; \frac{N_{\bar{A}}}{\bar{k}\ell/\gamma} \;\le\; 1 + \frac{1}{4z(z+1)}. \tag{7}$$

Further, by Proposition 1(a), we have $N_{\bar{A}} - N_A = (z-y)\ell$.

Upper Bound. We have

$$\begin{aligned}\frac{N_A}{k\ell/\gamma} &\le \left(1 + \frac{1}{4z(z+1)}\right)\frac{\bar{k}}{k} - \frac{(z-y)\ell}{k\ell/\gamma} \\ &= 1 + \frac{1}{4z(z+1)} + \left(1 + \frac{1}{4z(z+1)}\right)\frac{(z-y)\gamma}{(y+\delta)\gamma} - \frac{z-y}{y+\delta} \\ &= 1 + \frac{1}{4z(z+1)} + \frac{z-y}{4z(z+1)(y+\delta)} \\ &\le 1 + \frac{1}{4z(z+1)}\left(1 + \frac{z-y}{y}\right) = 1 + \frac{1}{4y(z+1)}.\end{aligned}$$

Lower Bound. By (7) we have

$$\begin{aligned}\frac{N_A}{k\ell/\gamma} &\ge \frac{4z(z+1)}{(2z+1)^2} \cdot \frac{\bar{k}}{k} - \frac{(z-y)\ell}{k\ell/\gamma} \\ &= 1 - \frac{1}{(2z+1)^2} + \left(1 - \frac{1}{(2z+1)^2}\right)\frac{z-y}{y+\delta} - \frac{z-y}{y+\delta} \\ &= 1 - \frac{1}{(2z+1)^2}\left(1 + \frac{z-y}{y+\delta}\right) \ge 1 - \frac{1}{4z(z+1)} \cdot \frac{z}{y} = 1 - \frac{1}{4y(z+1)}.\end{aligned}$$

□

Remark 2. There are arbitrarily large $k \times k$ matrices for which the bounds in Proposition 1(b) are optimal. For the *upper bound* we first consider $z = 1$. In the middle $k \times k$ matrix of Fig. 1 we have $\gamma = 9$ and $k = 12$, and $N_A = 18 = \frac{9}{8}k^2/\gamma$ many 1's. In the same way one gets $12h \times 12h$ gap-$9h$ matrices for arbitrary $h \geq 1$ that realize excess ratio $c_1 = \frac{9}{8}$. For arbitrary z, we define a $k \times k$ matrix for $k = 2z(z+1)$, as follows. We choose $\gamma = 2z+1$ and obtain $k' = k - \gamma z = z$. If we place 1's on the main diagonal, the $k' \times k'$ submatrix in the lower right corner has k' many 1's; hence the total number of 1's is $(k^2 - k'^2)/\gamma + k'$. One easily checks that $((k^2 - k'^2)/\gamma + k')/(k^2/\gamma)$ evaluates to $1 + 1/(4z(1+z)) = c_z$. For the *lower bound* we first consider $k = 9$ and $\gamma = 6$, hence $z = 1$. We obtain a matrix A with the minimum number of 1's by placing them in the diagonals running from $(0,3)$ to $(5,8)$ and from $(3,0)$ to $(8,5)$, which gives $N_A = 12$, hence $N_A/(k^2/\gamma) = 12/(9^2/6) = \frac{8}{9} = 1/c_1$. Again, it is easy to find larger matrices with ratio $1/c_1$ and examples for arbitrary z with ratio $1/c_z$.

3 Proof of Universality Properties

In this section we show that $\mathcal{H}^{\text{lin}}_{u,m,r}$ is approximately two-wise independent in general and two-wise independent if r is a prime power.

We will assume throughout this section that $U = [u] = \{0, \ldots, u-1\}$ for some $u \geq 2$ and $M = [m] = \{0, \ldots, m-1\}$ for some $m \geq 2$, and that $r = km$ is a multiple of m.

As preparation for the proof, we provide some observations and lemmas.

Fact 1. *Let $z \in \mathbb{Z}_r$ and let $\gamma = \gcd(z, r)$. Then for arbitrary $t \in \mathbb{Z}_r$ the following holds:*

$$|\{x \in \mathbb{Z}_r \mid x \cdot_r z = t\}| = \begin{cases} \gamma & \text{if } \gamma \mid t; \\ 0 & \text{otherwise.} \end{cases}$$

Proof. Using the fact ("Bézout's Lemma") that we may write $\gamma = \alpha r + \beta z$ for certain $\alpha, \beta \in \mathbb{Z}$, one sees that the range $m_z(\mathbb{Z}_r)$ of the mapping $m_z \colon \mathbb{Z}_r \ni x \mapsto x \cdot_r z \in \mathbb{Z}_r$ is (γ), the set of all multiples of γ in $\mathbb{Z}_r$. The fundamental homomorphism theorem from group theory applied to the group homomorphism m_z (with respect to the additive group structure $(\mathbb{Z}_r, +_r, 0)$) yields that $\left|m_z^{-1}(t)\right|$ is the same for all $t \in m_z(\mathbb{Z}_r) = (\gamma)$. Since $|(\gamma)| = r/\gamma$, the claim follows. □

Recall that $h(x) = \lfloor g(x)/k \rfloor$, where $g(x) = g_{a,b}(x) = a \cdot_r x +_r b$, for $a, b \in \mathbb{Z}_r$ chosen uniformly at random. We make some simple observations on the distributions of $g(x)$ and $h(x)$.

Lemma 1. *For each $x \in U$ the following holds:*

(a) *$g(x)$ is independent of a and uniformly distributed in $\mathbb{Z}_r$;*
(b) *$h(x)$ is independent of a and uniformly distributed in M.*

Proof. (a) For $s \in \mathbb{Z}_r$ and $\alpha \in \mathbb{Z}_r$ fixed we calculate:

$$\mathbf{Pr}(g(x) = s \mid a = \alpha) = \mathbf{Pr}(b = s - \alpha x \mid a = \alpha) = \frac{1}{r}.$$

Thus $g(x)$ and a are independent. Since the events $\{a = \alpha\}$, $\alpha \in \mathbb{Z}_r$, partition the probability space, we also get $\mathbf{Pr}(g(x) = s) = 1/r$.
(b) This follows immediately from (a), since k divides r, and hence the operation $s \mapsto \lfloor s/k \rfloor$ maps exactly k elements of $\mathbb{Z}_r$ to each element i of M. □

From here on, assume that $x_1, x_2 \in U$ with $x_2 < x_1$ are fixed. Let

$$z = x_1 - x_2 \ (= x_1 -_r x_2) \quad \text{and} \quad \gamma = \gcd(z, r).$$

Then $0 < z < u$ and $1 \le \gamma < u$. Now we can describe the joint distribution of $g(x_1)$ and $g(x_2)$ (see Fig. 3 below).

Lemma 2. (a) *For arbitrary $t \in \mathbb{Z}_r$ we have*

$$\mathbf{Pr}(g(x_1) -_r g(x_2) = t) = \begin{cases} \gamma/r & \textit{if } \gamma \mid t; \\ 0 & \textit{otherwise.} \end{cases}$$

(b) *For arbitrary $s_1, s_2 \in \mathbb{Z}_r$ we have*

$$\mathbf{Pr}(g(x_1) = s_1 \wedge g(x_2) = s_2) = \begin{cases} \gamma/r^2 & \textit{if } \gamma \mid (s_1 -_r s_2); \\ 0 & \textit{otherwise.} \end{cases}$$

Proof. (a) Clearly, for $g = g_{a,b}$, we have

$$g(x_1) -_r g(x_2) = a \cdot_r (x_1 -_r x_2) = a \cdot_r z.$$

Thus, the claim follows immediately from Fact 1.
(b) Since $g(x_1)$ and a are independent by Lemma 1(a) and $g(x_1) -_r g(x_2) = a \cdot_r z$ is a function of a, the random variables $g(x_1)$ and $g(x_1) -_r g(x_2)$ are independent. Thus, using part (a) we may calculate:

$$\begin{aligned} &\mathbf{Pr}(g(x_1) = s_1 \wedge g(x_2) = s_2) \\ &= \mathbf{Pr}(g(x_1) = s_1 \wedge g(x_1) -_r g(x_2) = s_1 -_r s_2) \\ &= \mathbf{Pr}(g(x_1) = s_1) \cdot \mathbf{Pr}(g(x_1) -_r g(x_2) = s_1 -_r s_2) \\ &\overset{(*)}{=} \begin{cases} (1/r) \cdot (\gamma/r) = \gamma/r^2 & \text{if } \gamma \text{ divides } s_1 -_r s_2; \\ (1/r) \cdot 0 = 0 & \text{otherwise.} \end{cases} \end{aligned}$$

(For $(*)$ we use both (a) and Lemma 1(a).) Thus (b) is proved. □

Finally we are ready to formulate and prove the central theorem about our hash functions. Let

$$\Gamma_{u,m,r} = \max(\{0\} \cup \{\gamma \in \{1, \ldots, u-1\} \mid \gamma \text{ divides } r \wedge \gamma \nmid k\}). \tag{8}$$

(The definition from [10] is changed in the clever way suggested in [41] to take the special case into account where all $\gamma \le u - 1$ that divide r also divide k.) Note that $\Gamma_{u,m,r} \le u - 1$.

Observation 1. *(a) If r is a power of a prime p and $r \geq um/p$, then $\varGamma_{u,m,r} = 0$.*
(b) If $r \geq (u-1)m$, then $\varGamma_{u,m,r} \leq k$.

Proof. (a) Since $r = mk$, the numbers m and k are also powers of p. Now if $\gamma < u$ divides r, it is a power of p itself, strictly smaller than $u \leq rp/m$. So γ divides $r/m = k$. This shows that $\varGamma_{u,m,r} = 0$. (b) This is trivial. □

Even if no assumption about r is made, we will always be able to make sure that $\varGamma_{u,m,r} \leq k$, so that the following is well-defined. With $\varGamma = \varGamma_{u,m,r}$, let

$$c_{u,m,r} = \begin{cases} 1 & \text{if } \varGamma = 0, \\ 1 + \dfrac{1}{4\lfloor k/\varGamma \rfloor (\lfloor k/\varGamma \rfloor + 1)} & \text{if } \varGamma > 0. \end{cases}$$

Note that $1 \leq c_{u,m,r} \leq \frac{9}{8}$ and that $c_{u,m,r} = 1 + O(u^2/k^2)$.

Theorem 2 (Main Theorem). *Let h be chosen at random from $\mathcal{H}^{\text{lin}}_{u,m,r}$, where $m \mid r$. Then the following holds, for arbitrary $i_1, i_2 \in M$ and distinct $x_1, x_2 \in U$:*
(a) *If r is a power of a prime p and $r \geq um/p$, then*

$$\mathbf{Pr}(h(x_1) = i_1 \wedge h(x_2) = i_2) = \frac{1}{m^2},$$

i.e., in this case $\mathcal{H}^{\text{lin}}_{u,m,r}$ is 2-wise independent.
(b) *If $r \geq (u-1)m$, then $\mathbf{Pr}(h(x_1) = i_1) = 1/m$ and*

$$\frac{2 - c_{u,m,r}}{m^2} \leq \frac{1}{c_{u,m,r}} \cdot \frac{1}{m^2} \leq \mathbf{Pr}(h(x_1) = i_1 \wedge h(x_2) = i_2) \leq \frac{c_{u,m,r}}{m^2},$$

i.e., in this case $\mathcal{H}^{\text{lin}}_{u,m,r}$ is $c_{u,m,r}$-approximately 2-wise independent.

Remark 3. A weaker version of the upper bound in (b), a similar lower bound, and (a) were shown in [10]. Woelfel [41] was the first to prove the upper bound in (b), with a different method.

Proof. Our aim is to find upper and lower bounds on

$$Q = Q(x_1, x_2, i_1, i_2) = \frac{\mathbf{Pr}(h(x_1) = i_1 \wedge h(x_2) = i_2)}{1/m^2}.$$

From the definition of $h_{a,b}$ it is immediate that with $I_i = \{ki, ki+1, \ldots, k(i+1)-1\}$, for $i \in M$, we have

$$\mathbf{Pr}(h(x_1) = i_1 \wedge h(x_2) = i_2) = \mathbf{Pr}(g(x_1) \in I_{i_1} \wedge g(x_2) \in I_{i_2}).$$

As usual, let $\gamma = \gcd(x_1 - x_2, r)$. Adding the equation given in Lemma 2(b) for all $s_1 \in I_{i_1}, s_2 \in I_{i_2}$, we obtain

$$\mathbf{Pr}(h(x_1) = i_1 \wedge h(x_2) = i_2) = |S_{i_1,i_2}| \cdot \gamma/r^2,$$

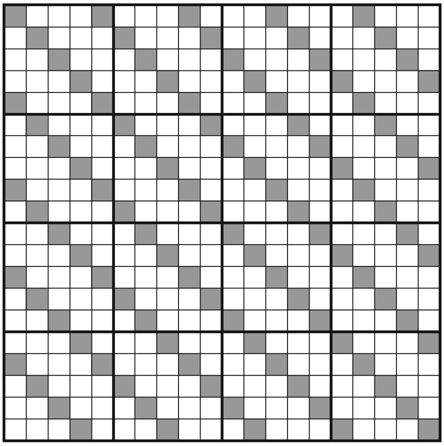

Fig. 3. The distribution of $(g(x_1), g(x_2))$ with $r = 20$, $\gamma = \gcd(x_1 - x_2, r) = 4$, $m = 4$, hence $k = 5$. We see $m^2 = 16$ many gap-4 matrices of dimension 5×5. Each grey position is attained with probability $\frac{4}{400} = \frac{1}{100}$. Some submatrices have probability $\frac{6}{100} < \frac{1}{16}$, some have probability $\frac{7}{100} > \frac{1}{16}$.

where $S_{i_1,i_2} = \{(s_1, s_2) \in I_{i_1} \times I_{i_2} \mid \gamma \text{ divides } s_1 -_r s_2\}$, hence

$$Q = \frac{|S_{i_1,i_2}| \cdot \gamma \cdot m^2}{r^2} = \frac{|S_{i_1,i_2}|}{k^2/\gamma}. \tag{9}$$

We can regard (the characteristic function of) S_{i_1,i_2} as being represented by a $k \times k$ matrix with row indices s_1 from I_{i_1} and column indices s_2 from I_{i_2}, with the entry for (s_1, s_2) being 1 if γ divides $s_1 -_r s_2$ and 0 otherwise. (Figure 3 shows all these matrices in one picture, for an example.) Clearly this is a gap-γ matrix as considered in Sect. 2, with shifted row and column numberings. So we can apply the estimates from there.

Case 1: $\gamma \mid k$. — In this case the quotient in (9) equals 1, by Proposition 1(a). Since the hypothesis of (a) implies that γ divides k (as seen in Observation 1), we have proved (a).

Case 2: $\gamma \nmid k$. — With Proposition 1(b) we may estimate:

$$\left(1 + \frac{1}{4\lfloor k/\gamma \rfloor (\lfloor k/\gamma \rfloor + 1)}\right)^{-1} \le Q \le 1 + \frac{1}{4\lfloor k/\gamma \rfloor (\lfloor k/\gamma \rfloor + 1)}. \tag{10}$$

We have $\gamma \le \mathit{\Gamma}$, hence $\lfloor k/\mathit{\Gamma} \rfloor (\lfloor k/\mathit{\Gamma} \rfloor + 1) \le \lfloor k/\gamma \rfloor (\lfloor k/\gamma \rfloor + 1)$, so (10) implies part (b). □

Remark 4. One can utilize the construction in Remark 2 of gap-γ matrices that realize the upper and lower bounds for the number of 1's to find u, m, and r for which the bounds in Theorem 2 are optimal. The details are left to the reader.

3.1 Generalization: r Not Divisible by m

We have proven the central universality properties of $\mathcal{H}^{\text{lin}}_{u,m,r}$. However, we always had to assume that m, the size of the range, is a divisor of r, which may not always be convenient. For example, we might wish to use as r some power of 2, because computer arithmetic in $\mathbb{Z}_r$ is particularly efficient in this case, but on the other hand wish to use some m that is not a power of 2. Another such case arises when we want to choose r randomly from some range and do not have control over the divisors of r. (See Subsect. 4.4 for such a situation.) For this reason we note here that the methods developed so far make it possible to construct approximately 2-wise independent classes also with a ring $\mathbb{Z}_r = \{0, \ldots, r-1\}$ where r is not a multiple of m.

Partition $\mathbb{Z}_r$ into m intervals $I_0, \ldots, I_{m-1}$, where $I_0, \ldots, I_{(r \bmod m)-1}$ have length $\overline{k} = \lceil r/m \rceil$ and $I_{r \bmod m}, \ldots, I_{m-1}$ have length $\underline{k} = \lfloor r/m \rfloor$. Let $h_{a,b}(x) = i$ if $g_{a,b}(x) \in I_i$. Arithmetically, this is expressed as follows: Precompute $\overline{k}$ and $\underline{k}$, as well as $B = r \bmod m$ and $T = B \cdot \overline{k}$. Then let

$$h_{a,b}(x) = \begin{cases} \lfloor g_{a,b}(x)/\overline{k} \rfloor & \text{if } g_{a,b}(x) < T; \\ B + \lfloor (g_{a,b}(x) - T)/\underline{k} \rfloor & \text{otherwise.} \end{cases}$$

The resulting class serves all hash values with almost the same probability in the case $r \gg m$, since

$$\mathbf{Pr}(h(x) = i) = \begin{cases} \lceil r/m \rceil / r \le (1 + 1/\lfloor r/m \rfloor)/m, & \text{for } 0 \le i < B; \\ \lfloor r/m \rfloor / r \ge (1 - 1/\lceil r/m \rceil)/m, & \text{for } B \le i < m. \end{cases}$$

The class of all $h_{a,b}$, $a, b \in [r]$, is again called $\mathcal{H}^{\text{lin}}_{u,m,r}$. This class satisfies approximate independence conditions, which can be proved in analogy to Theorem 2(b), using Proposition 1(b) for $\ell \in \{k-1, k, k+1\}$. (Assume $\gamma = \gcd(x_1 - x_2, r)$ as in the proof of Theorem 2. Following that proof, we will have $|I_{i_1}|, |I_{i_2}| \in \{\underline{k}, \overline{k}\}$. Since the error-free formula from Proposition 1(a) applies if at least one of the numbers $|I_{i_1}|$ and $|I_{i_2}|$ is divisible by γ, the quotients $y = \lfloor |I_{i_1}| / \gamma \rfloor$ and $z = \lfloor |I_{i_2}| / \gamma \rfloor$ can be assumed to be the same.) The reader is invited to work out the details of the proof of the following statement.

Proposition 2. *Assume $r \ge (u-1)m$, and consider the generalized version of class $\mathcal{H}^{\text{lin}}_{u,m,r}$ as just described. Let $x_1, x_2 \in U = [u]$ be distinct and $i_1, i_2 \in M = [m]$ be arbitrary. Then we have, with $c_{u,m,r}$ as defined for Theorem 2:*

$$2 - c_{u,m,r} \;\le\; \frac{\mathbf{Pr}(h(x_1) = i_1 \wedge h(x_2) = i_2)}{\mathbf{Pr}(h(x_1) = i_1)\,\mathbf{Pr}(h(x_2) = i_2)} \;\le\; c_{u,m,r}\,.$$

□

4 Variations of the Construction

4.1 Vectors as Keys and Range Elements

If keys are very long, it may be inconvenient in practice to treat them as integers in a universe $U = [u]$, since one has to carry out long integer multiplication. How can one deal with longer keys, e.g., given as vectors in a universe $U' = U^\ell$, for $U = [u]$, when the range is $M = [m]$? It is well known that 2-independence can be ensured by just using 2-independent hashing on the ℓ components separately and taking the sum of the results modulo m. In [10] it was sketched how to proceed just utilizing the ring $\mathbb{Z}_r$ with $r \geq (u-1)m$ and $k = r/m$, as before. An advantage is that fewer random bits are needed and that, as observed by Woelfel [41,42], a parsimonious extension to a range that is also a set of vectors is possible. A hash function is specified by a coefficient vector $\boldsymbol{a} = (a_0, \ldots, a_{\ell-1}) \in \mathbb{Z}_r^\ell$ and $b \in \mathbb{Z}_r$. The hash function $h_{\boldsymbol{a},b}$ is defined by

$$h_{\boldsymbol{a},b}((\xi_0, \ldots, \xi_{\ell-1})) = \left\lfloor \left(\sum_{0 \leq \lambda < \ell}^{(r)} a_\lambda \cdot_r \xi_\lambda +_r b \right) \Big/ k \right\rfloor,$$

for $\boldsymbol{x} = (\xi_0, \ldots, \xi_{\ell-1}) \in [u]^\ell$. (The superscript "$(r)$" indicates summation in $\mathbb{Z}_r$.) The resulting class exhibits the same universality properties as $\mathcal{H}^{\text{lin}}_{u,m,r}$, as stated in Theorem 2. The proof is very similar to that of Theorem 2. The idea is simple. Two different keys $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)} \in U^\ell$ must have one component in which they differ, $\xi_0^{(1)} \neq \xi_0^{(2)}$, say. One fixes $a_1, \ldots, a_{\ell-1}$ and studies the joint distribution of the pair

$$\mathbb{Z}_r^2 \ni (a_0, b) \mapsto \left\lfloor \Big(a_0 \cdot_r \xi_0^{(i)} +_r b +_r \underbrace{\sum_{1 \leq \lambda < \ell}^{(r)} a_\lambda \cdot_r \xi_\lambda^{(i)}}_{=C_i} \Big) / k \right\rfloor, \quad i = 1, 2,$$

of random variables. The values C_1 and C_2 are regarded as constant. The analysis is practically identical to that in the proof of Theorem 2.

Example: Assume we work on a computer with fast 64-bit arithmetic, the keys are bit strings of arbitrary, fixed length given as a sequence of words, and the range is $M = [2^\mu]$ for some $\mu \leq 32$. Then we may let $u = 2^{32}$ and $r = 2^{64}$. A key x is split into 32-bit pieces $\xi_0, \xi_1, \ldots, \xi_{\ell-1}$ for a suitable ℓ. A hash function is represented as a sequence $(a_0, \ldots, a_{\ell-1}, b)$ of 64-bit integers $a_0, \ldots, a_{\ell-1}, b$. The modulo r operation is for free, since the standard hardware carries out multiplication and addition modulo 2^{64}. The final division by $2^{64-\mu}$ is done by a shift.

This construction is helpful if the range M is not too big— the ring size r must be bigger than $(u-1)m$ or um/p in the case that r is a power of a prime number p. Woelfel [41,42] gave a more general, very elegant construction to remove this restriction. It resembles the convolution construction over finite fields studied in [25], but is made to work over the ring $\mathbb{Z}_r$. Assume the range is a set $M' = M^\varrho = [m]^\varrho$ of vectors. Of course, we could use ϱ many independent hash functions with range M to build one with range M^ϱ. However, then the number $(\ell+1)\varrho$ of required random coefficients from $\mathbb{Z}_r$ grows with the product of ℓ and ϱ.

To save random bits, and storage space, Woelfel proposed using convolution (or "polynomial multiplication") over $\mathbb{Z}_r$. A hash function is given by vectors

$\boldsymbol{a} = (a_0, \ldots, a_{\ell+\varrho-2})$ and $\boldsymbol{b} = (b_0, \ldots, b_{\varrho-1})$ with coefficients from $\mathbb{Z}_r$ For a key $\boldsymbol{x} = (\xi_0, \ldots, \xi_{\ell-1}) \in U^\ell$ and $0 \le \kappa < \varrho$ component κ of the hash value $h(\boldsymbol{x})$ is given by

$$(h_{\boldsymbol{a},\boldsymbol{b}}((\xi_0, \ldots, \xi_{\ell-1})))_\kappa = \left\lfloor \Bigl(\sum_{0\le\lambda<\ell}^{(r)} a_{\kappa+\lambda} \cdot_r \xi_\lambda +_r b_\kappa \Bigr) \Big/ k \right\rfloor.$$

The resulting class of hash functions from $U' = U^\ell$ to $M' = M^\varrho$ is called $\mathcal{H}^{\mathrm{conv},\ell,\varrho}_{u,m,r}$. In the "clean" case where r is a power of a prime number p and $r \ge um/p$ one can show that $\mathcal{H}^{\mathrm{conv},\ell,\varrho}_{u,m,r}$ is 2-wise independent. If we only have that m divides r and $r \ge (u-1)m$, then $\mathcal{H}^{\mathrm{conv},\ell,\varrho}_{u,m,r}$ is $(c_{u,m,r})^\varrho$-approximately 2-wise independent, for $c_{u,m,r}$ the approximation constant for class $\mathcal{H}^{\mathrm{lin}}_{u,m,r}$ as in Theorem 2. (For details, see [41,42]. There it is also discussed how the parameter spaces for the coefficients may be reduced.)

4.2 Higher Independence

In [10] it was proved that polynomials over $\mathbb{Z}_r$ can be used to obtain (approximately) d-wise independent classes for arbitrary fixed $d \ge 2$. Here, we give the relevant definitions and state the theorem, and refer the reader to [10] for the proof. Unfortunately, the construction seems to be mainly of theoretical interest. While it achieves higher independence without prime numbers or finite fields being involved (and without tabulation [37], which requires longer tables of random entries), the big disadvantage of the construction is that r has to be quite large, which makes for slow evaluation. —As before, we assume that a universe $U = [u]$ and a range $M = [m]$ are given.

Definition 3. *Let $d \ge 2$. A class $\mathcal{H}$ of hash functions from U to M is called* d-wise independent *if for arbitrary distinct keys $x_0, \ldots, x_{d-1} \in U$ and arbitrary $i_0, \ldots, i_{d-1} \in M$ we have*

$$\mathbf{Pr}(h(x_s) = i_s,\ \textit{for } 0 \le s < d) = \frac{1}{m^d},$$

for h chosen uniformly at random from $\mathcal{H}$. For $c \ge 1$, such a class is called c-approximately d-wise independent if for each key x the hash value $h(x)$ is uniformly distributed in M and if for distinct $x_0, \ldots, x_{d-1} \in U$ and arbitrary $i_0, \ldots, i_{d-1} \in M$ we have

$$\frac{2-c}{m^d} \le \mathbf{Pr}(h(x_s) = i_s,\ \textit{for } 0 \le s < d) \le \frac{c}{m^d}.$$

As before, we fix $r = km$ for some positive integer k, and consider the ring $\mathbb{Z}_r = [r]$ with arithmetic modulo r. Further, fix $d \ge 2$.

Definition 4. *For $\boldsymbol{a} = (a_0, \ldots, a_{d-1}) \in \mathbb{Z}_r^d$ define (arithmetic in $\mathbb{Z}_r$)*

$$g_{\boldsymbol{a}}(x) = \sum_{0\le\mu<d} a_\mu x^\mu,\ \textit{for } x \in U,$$

and

$$h_{\boldsymbol{a}}(x) = \lfloor g_{\boldsymbol{a}}(x)/(r/m) \rfloor\,,\ \textit{for } x \in U.$$

Further, let

$$\mathcal{H}^{\text{deg-}d}_{u,m,r} = \{h_{\boldsymbol{a}} \mid \boldsymbol{a} \in \mathbb{Z}_r^d\}.$$

Theorem 3. *If $r \geq m \cdot (u-1)^{\binom{d}{2}}$, then $\mathcal{H}^{\text{deg-}d}_{u,m,r}$ is c-approximately d-wise independent for $c = c(u,m,r,d) = (1 - m(u-1)^{\binom{d}{2}}/r)^{-d}$. More precisely, $h(x)$ is uniformly distributed in $[m]$, for each $x \in U$, and for arbitrary distinct elements $x_0, \ldots, x_{d-1} \in U$ and arbitrary values $i_0, \ldots, i_{d-1} \in M$ we have:*

(a) $\left(1 - \dfrac{m(u-1)^{\binom{d}{2}}}{r}\right)^d \leq \dfrac{\mathbf{Pr}(h(x_\lambda) = i_\lambda, 0 \leq \lambda < l)}{1/m^d} \leq \left(1 + \dfrac{m(u-1)^{\binom{d}{2}}}{r}\right)^d$*;*

(b) *if u, m, and r are powers of the same prime number $p \geq 2$, and $r \geq m(u/p)^{\binom{d}{2}}$, we even have d-wise independence, i.e.,*

$$\mathbf{Pr}(h(x_\lambda) = i_\lambda,\ \textit{for } 0 \leq \lambda < l) = \frac{1}{m^d}.$$

□

4.3 Two-Wise Independent Sampling

Here we describe how the function classes described before can be used for two-independent sampling in the sense of Chor and Goldreich [8]. There one has a finite set $M = [m]$ and a set $A \subseteq M$ of which one wants to find an arbitrary element. The only operation available regarding A is a membership test "is $x \in A$?" for $x \in M$. The most obvious search method is random sampling (keep choosing random elements of M until an element of A is found); however, this method has the disadvantage that in expectation $(1/\varrho)\log|M|$ random bits are needed, where $\varrho = |A|\,/\,|M|$ is the density of A in M (which sometimes is small). In order to save random bits, one employs a 2-wise independent sequence $X_0, \ldots, X_{u-1}$ of random variables, uniformly distributed in M. These elements are tested one after the other, until an element of A is found ("success"). In [8], Chor and Goldreich use finite fields based on prime numbers for constructing such sequences. They show that the probability that the first j trials all fail is bounded by $1/(j\varrho)$. If $\mathcal{H}$ is an arbitrary 2-wise independent class of hash functions from $U = [u]$ to M, one can use $X_i = h(i)$ for $0 \leq i < u$. The classes studied in the present paper can be used immediately if m and u are powers of the same prime number p and $r \geq um$. This requires $2\log r$ random bits for the coefficients a and b.

We wish to show with a slightly refined analysis that the classes $\mathcal{H}^{\text{lin}}_{u,m,r}$ can be used for arbitrary $M = [m]$ and $U = [u]$ if one chooses r sufficiently large. Let h be chosen uniformly at random from $\mathcal{H}^{\text{lin}}_{u,m,r}$. We let

$$X_i = h(i),\ \text{for } 0 \leq i < u.$$

For given $A \subseteq M$ we let

$$Y_i = \begin{cases} 1 & \text{if } X_i \in A, \\ 0 & \text{otherwise,} \end{cases}$$

for $0 \le i < u$. Since X_i is uniformly distributed, we have $\mathbf{E}(Y_i) = 1/\varrho$.

For $0 \le j \le u$, the number of successes one encounters in $X_0, \ldots, X_{j-1}$ is

$$Z_j = \sum_{0 \le i < j} Y_i .$$

Clearly, $\mathbf{E}(Z_j) = j\varrho$. The crux of Chor and Goldreich's method is to note that from the two-wise independence of the Y_i's it follows that $\mathbf{Var}(Z_j) \le \mathbf{E}(Z_j)$. The Chebychev-Cantelli inequality $\mathbf{Pr}(X \le \mathbf{E}(X) - t) \le \mathbf{Var}(X) / (\mathbf{Var}(X) + t^2)$ then implies that

$$\begin{aligned} &\mathbf{Pr}(X_i \notin A \text{ for all } i,\ 0 \le i < j) = \mathbf{Pr}(Z_j = 0) \le \mathbf{Pr}(Z_j \le \mathbf{E}(Z_j) - \mathbf{E}(Z_j)) \\ &\quad \le \frac{\mathbf{Var}(Z_j)}{\mathbf{Var}(Z_j) + \mathbf{E}(Z_j)^2} = \frac{1}{1 + \mathbf{E}(Z_j)^2 / \mathbf{Var}(Z_j)} \le \frac{1}{1 + \mathbf{E}(Z_j)} = \frac{1}{1 + j\varrho} . \end{aligned}$$

This is the desired bound on the failure probability of 2-wise independent sampling.

We show that for all sufficiently large r the required bound $\mathbf{Var}(Z_j) \le \mathbf{E}(Z_j)$ is true even though the random variables $X_0, \ldots, X_{u-1}$ are only approximately 2-wise independent.

Lemma 3. *If $r \ge u^{3/2} m$, then $\mathbf{Var}(Z_j) \le \mathbf{E}(Z_j)$, for $0 \le j < u$.*

Proof. We calculate:

$$\mathbf{Var}(Z_j) = \sum_{0 \le i < j} \mathbf{Var}(Y_i) + \sum_{\substack{0 \le i, i' < j \\ i \ne i'}} \mathbf{Cov}(Y_i, Y_{i'}) . \tag{11}$$

Clearly, $\mathbf{Var}(Y_i) = \varrho(1 - \varrho)$ for all i, $0 \le i < j$. We analyze a summand of the second sum in (11). Fix $i \ne i'$. For $s \in M$, $t \in M$, let

$$\chi_s(t) = \begin{cases} 1 & \text{if } t = s, \\ 0 & \text{otherwise.} \end{cases}$$

Then

$$\begin{aligned} \mathbf{Cov}(Y_i, Y_{i'}) &= \mathbf{E}((Y_i - \varrho)(Y_{i'} - \varrho)) \\ &= \mathbf{E}\left(\Big(\sum_{s \in A} (\chi_s(X_i) - 1/m) \Big) \Big(\sum_{t \in A} (\chi_t(X_{i'}) - 1/m) \Big) \right) \\ &= \sum_{s, t \in A} \mathbf{Cov}(\chi_s(X_i), \chi_t(X_{i'})) . \end{aligned} \tag{12}$$

For $s, t \in A$ arbitrary we have

$$\begin{aligned}\mathbf{Cov}(\chi_s(X_i), \chi_t(X_{i'})) &= \mathbf{E}(\chi_s(X_i) \cdot \chi_t(X_{i'})) - \mathbf{E}(\chi_s(X_i)) \cdot \mathbf{E}(\chi_t(X_{i'})) \\ &= \mathbf{Pr}(X_i = s \wedge X_{i'} = t) - 1/m^2 \leq (c_{u,m,r} - 1)/m^2,\end{aligned}$$

by Theorem 2(b). Since $r \geq (u-1)m$, we have $c_{u,m,r} - 1 \leq 1/(4\lfloor r/((u-1)m)\rfloor \cdot (\lfloor r/((u-1)m)\rfloor + 1)) \leq (um/r)^2$, and we get by summing up according to (11) and (12):

$$\begin{aligned}\mathbf{Var}(Z_j) &\leq j\varrho(1-\varrho) + j^2 \left|A\right|^2 (c_{u,m,r} - 1)/m^2 \\ &= j\varrho(1-\varrho) + j^2\varrho^2(c_{u,m,r} - 1) \\ &\leq j\varrho(1-\varrho) + j^2\varrho^2 \cdot (um/r)^2 \\ &\leq j\varrho(1-\varrho) + j\varrho^2(u^{3/2}m/r)^2.\end{aligned}$$

By the assumption $r \geq u^{3/2}m$ we get $\mathbf{Var}(Z_j) \leq j\varrho(1-\varrho) + j\varrho^2 = j\varrho = \mathbf{E}(Z_j)$, as desired. □

4.4 "Collapsing the Universe" Without Finite Fields

Assume $U = [u]$ is very large, meaning that if we represent keys as bit strings, they are very long. Very often, in applications like dictionaries, one wishes to replace a key x by a shorter key x' in a range $U' = [u']$, where a "collapse function" $h\colon U \to U'$ maps x to x'. For this to work, h must be one-to-one on the set $S \subseteq U$ of "relevant keys". With the results from Sect. 4 it is easy to find a good collapsing function with a description size of $O(\log u)$ bits. Collapse functions with smaller description sizes can be constructed deterministically [19, 33], or, usually much simpler, in a randomized way [11,35]. In the latter case the description size of h is $O(\log n + \log\log u)$ bits, and it requires choosing a random prime of this bitlength or knowing a finite field with elements of this bitlength. Here we demonstrate that class $\mathcal{H}^{\text{lin}}_{u,m,r}$ can be used to build collapse functions with such a small description size on the basis of modular arithmetic alone, without the need for prime numbers or finite field arithmetic.

Technically, assume $S = \{x_0, \ldots, x_{n-1}\} \subseteq U = [u]$ is given. We let $m = \lceil n^3 \log u \rceil$. The size r of the ring $\mathbb{Z}_r$ is chosen uniformly at random from $[m^2, 2m^2)$. Then $\lfloor r/m \rfloor \in [m, 2m)$. (Since usually r is not a multiple of m, we need to use the modified functions from Subsect. 3.1.) For the hash class $\mathcal{H}^{\text{lin}}_{u,m,r}$ to function in a 2-wise independent way on S, we only require, to make the proof of Theorem 2 work, that r is S-*good* in the following sense:

$$\varGamma_S = \max\{\gcd(x_j - x_i, r) \mid x_i, x_j \in S \text{ distinct}\} \leq r/m.$$

If this is the case, we will have $\mathbf{Pr}(h(x_i) = h(x_j)) = O(1/m) = O(1/n^3)$ for all distinct $x_i, x_j \in S$, and hence $\mathbf{Pr}(h \text{ is not one-to-one on } S) = O(1/n)$. Clearly, for r being S-good it is sufficient that

$$\gcd\Biggl(\prod_{0 \leq i < j < n} (x_j - x_i),\ r\Biggr) \leq r/m.$$

The following lemma implies that among the numbers in $[m^2, 2m^2)$ a constant fraction is S-good.

Lemma 4. *Let Y be a sufficiently large integer, and let $L \geq \ln Y$. Then (at least) a constant fraction of the numbers r in $[L^2, 2L^2)$ satisfy* $\gcd(r, Y) \leq r/L$.

Proof. Let

$$A = A_{L,Y} = \{r \in [L^2, 2L^2) \mid \gcd(r, Y) > r/L\}.$$

Then $A \subseteq B \cup C_L$, where

$$B = B_{L,Y} = \{r \in [L^2, 2L^2) \mid \gcd(r, Y) \,\text{has a prime factor}\, p > L\}$$

and

$$C_L = \{r \in [L^2, 2L^2) \mid p \leq L \,\text{for all prime factors}\, p \,\text{of}\, r\}.$$

Indeed, if $r \in [L^2, 2L^2) - (B \cup C_L)$, then r has a prime factor $p > L$ that does not divide Y, hence $\gcd(r, Y) \leq r/p < r/L$. We must estimate the sizes of B and C_L. First note that Y has at most $\ln Y / \ln L$ prime factors larger than L. Each such factor divides at most $L^2/L = L$ elements of $[L^2, 2L^2)$; thus, $|B| \leq L \cdot \ln Y / \ln L \leq L^2 / \ln L$.

In order to deal with C_L, we give a lower bound on the size of the complementary set $D_L = \{r \in [L^2, 2L^2) \mid r \,\text{has a prime factor larger than}\, L\}$. It is well known that $|D_L| = (\ln 2 - O(1/\log L)) \cdot L^2$ (for $L \to \infty$), the reason being the following: For each prime number $p \in (L, L^2]$ the set of multiples of p in $(L^2, 2L^2)$ has size at least $\lfloor L^2/p \rfloor \geq L^2/p - 1$. If we just add these figures, numbers $r \in [L^2, 2L^2)$ with two distinct prime factors $p_1, p_2 > L$ are counted twice. Note that in this situation we must have $p_1, p_2 < 2L$ and $r = p_1 p_2$; hence, by the prime number theorem, there are only $O((L/\log L)^2) = O(L/(\log L)^2)$ many such numbers r. The prime number theorem also entails that there are only $O(L^2/\log L)$ many primes in $(L, 2L^2]$. So we obtain:

$$|D_L| \geq \sum_{\substack{L<p\leq L^2 \\ p\,\text{prime}}} \left(\frac{L^2}{p} - 1\right) - O\left(\frac{L}{(\log L)^2}\right) = L^2 \cdot \sum_{\substack{L<p\leq L^2 \\ p\,\text{prime}}} \frac{1}{p} - O(L^2/\log L).$$

A well-known theorem from analytic number theory [20, p. 351, Theorem 427] says that $\sum_{p \leq x, p\,\text{prime}} 1/p = \ln \ln x + B_1 + E(x)$, where B_1 is a constant and $E(x) = O(1/\log x)$. It follows that

$$\sum_{\substack{L<p\leq L^2 \\ p\,\text{prime}}} \frac{1}{p} = \ln \ln L^2 - \ln \ln L - O(1/\log L) = \ln 2 - O(1/\log L)\,.$$

Summing up, we get that

$$|A| \leq |B| + |C_L| = |B| + (L^2 - |D_L|) = (1 - \ln 2) \cdot L^2 + O(L^2/\log L),$$

i.e., 69% of the r in $[L^2, 2L^2)$ satisfy $\gcd(r, Y) \leq r/L$, asymptotically. □

We apply Lemma 4 for $L = m$ and $Y = \prod_{0 \leq i < j < n}(x_j - x_i)$. Since $m = n^3 \ln u > \ln Y$, the assumptions are satisfied, and we conclude that asymptotically at least 69% of the r in $[m^2, 2m^2)$ are S-good.

Remark 5. In the context of algorithms that offer higher reliability for the running time, like the real-time dictionary from [14], the probability bounds provided by Lemma 4 are too weak. At present, all *reliable* collapse functions that only use $O(\log\log u + \log n)$ random bits involve the use of prime numbers.

5 Extensions and Applications

Smaller Pure Arithmetic Classes. Starting from the construction of Sect. 3, Woelfel [41,42] succeeded in giving smaller hash classes with similar universality properties, thus reducing the space, the number of random bits, and possibly the evaluation time. In particular, he observed that it is sufficient to choose b at random from $k \cdot [m] = \{ik \mid 0 \leq i < m\}$ to achieve the results in Sect. 3. Moreover, he showed that if r is a power of some prime p, one can have u as large as r, and one still gets a 1-universal class of hash functions from $[u]$ to $[m]$ with $m = r/k$, if one chooses a from $1 + p[r/p]$ and b from $p^{\lceil k/2 \rceil} \cdot [p^{\lfloor k/2 \rfloor}]$ uniformly at random. This class is both very efficient and small. Woelfel's work contains many more constructions for different universality concepts.

Lower Bounds on the Complexity of Multiplication. In [25], it is shown that 2-wise independent classes contain functions that are difficult to compute in several respects (time-space tradeoff $T \cdot S = \Omega(\log u \cdot \log m)$; quadratic $A \cdot T^2$ bounds for VLSI implementation; bounds for CREW PRAMs, boolean formulas, and constant-depth circuits). All these bounds transfer to integer multiplication in binary notation, by choosing r as a power of 2. Woelfel [43] and Bollig and Woelfel [6] used linear classes for proving lower bounds on the complexity of multiplication on several versions of branching programs, notably OBDDs. The central observation was that the universality properties imply that there are functions of the form $x \mapsto \lfloor((ax + b) \bmod r)/m\rfloor$ with moderately large r and $U = [u]$ only of size m^2 that are surjective. Bollig, Waack, and Woelfel [5] used a similar approach for lower bounds for multiplication in more general branching programs. Using similar hash classes, Sauerhoff and Woelfel [34] prove a time-space tradeoff for unrestricted, deterministic multi-way branching programs that compute the middle bit of integer multiplication.

Error-Correction Properties. For the purpose of constructing deterministic dictionaries, Miltersen [27] and Hagerup, Miltersen, and Pagh [19] utilized error-correction properties of 2-wise independent classes, in particular of the class from [12] and the classes studied in the present paper.

Limitations for Cuckoo Hashing and Other Applications. Pătraşcu and Thorup [31] proved that for min-wise hashing the linear classes fail badly. Dietzfelbinger and Schellbach [15,16] showed that using linear functions $h(x) =$

$((ax+b) \bmod p) \bmod m$ (for $p = u$ prime) and $h(x) = \lfloor((ax+b) \bmod r)/(r/m)\rfloor$ as in this paper in the naive way for cuckoo hashing will lead to failure of the data structure. On the other hand, Dietzfelbinger and Woelfel [17] showed how to run cuckoo hashing [32] with polynomials of constant degree in combination with lookup tables of random numbers as hash functions. With Aumüller these authors showed [3] how to modify the construction so that 1-universal classes, 2-wise independent classes, and tables are sufficient, so that now cuckoo hashing can be run with the linear classes from the present paper alone.

3SUM and Fine-Grained Complexity Inside P. 3SUM over an interval $[u]$ of integers, for sets of size n, is the following problem: Given three sets $A, B, C \subseteq [u]$, all of size n, decide whether there are $x \in A$, $y \in B$, $z \in C$ such that $x + y = z$. Depending on the model of computation, this problem is thought to have different complexities. The obvious deterministic algorithm (involving sorting and repeated merging) takes time $\Theta(n^2)$. If one specifies the model in more detail, one may consider the word RAM (random access machine with word length w). Here, randomization, bit-level parallelism, and other tricks make it possible to solve the problem faster, see, e.g., the seminal paper by Baran, Demaine, and Pătraşcu [4]. One tool in this faster (randomized) algorithm is universal hashing. The authors utilize a 1-universal hash class of functions from $[u]$ to $[m]$ that is “almost affine”, meaning that for every hash function h from the class there is a constant c_h such that for $x, y, z \in [u]$ with $x + y = z$ we have $h(x) +_m h(y) \in \{h(x+y) +_m c_h, h(x+y) +_m c_h +_m 1\}$. (The offset c_h is known from affine functions in vector spaces.) It is not hard to see (see Wang [39], who also considers k-SUM) that our class $\mathcal{H}^{\text{lin}}_{u,m,r}$ has this property. A second property that is needed is 1-universality, from which it follows that the number of keys mapped to overfull buckets is close to its expectation. The same kind of hash function was used in many works on low-level complexity, e.g., [1,22,30,38], in arguments that prove conditional lower bounds for dynamic data structures and string and graph problems.[2]

Upper Bound on Bucket Size. Knudsen [21] showed that the expected bucket sizes created by $\mathcal{H}^{\text{lin}}_{u,m,r}$ on a set S of size $|S| = m$ is $O(m^{1/3})$.

Efficiency. Thorup [36] and Dahlgaard et al. [9] experimentally explore the efficiency of different hash classes. Our class $\mathcal{H}^{\text{lin}}_{u,m,r}$, for u, m, r powers of 2, turns out to be very fast, in particular for values that combine well with the word length of the computer. One should beware, however, that the class must be used only for purposes where its suitability has been proved.

Acknowledgment. The author is grateful to Carl Pomerance for helping with the proof of Lemma 4 and to a reviewer for a careful reading of the manuscript and helpful suggestions.

[2] It should be noted that although the class of functions $x \mapsto (ax \bmod p) \bmod m$ for primes p is not suitable, the class of functions $x \mapsto (ax \bmod p)/\lfloor p/m\rfloor$ is also almost affine and $(1 + \frac{1}{m})$-universal, so it could also have been used for this application.

References

1. Abboud, A., Williams, V.V., Weimann, O.: Consequences of faster alignment of sequences. In: Esparza, J., Fraigniaud, P., Husfeldt, T., Koutsoupias, E. (eds.) ICALP 2014. LNCS, vol. 8572, pp. 39–51. Springer, Heidelberg (2014). https://doi.org/10.1007/978-3-662-43948-7_4
2. Alon, N., Goldreich, O., Håstad, J., Peralta, R.: Simple constructions of almost k-wise independent random variables. Random Struct. Algorithms **3**, 289–304 (1992). https://doi.org/10.1002/rsa.3240030308
3. Aumüller, M., Dietzfelbinger, M., Woelfel, P.: Explicit and efficient hash families suffice for cuckoo hashing with a stash. Algorithmica **70**(3), 428–456 (2014). https://doi.org/10.1007/s00453-013-9840-x
4. Baran, I., Demaine, E.D., Pǎtraşcu, M.: Subquadratic algorithms for 3SUM. Algorithmica **50**(4), 584–596 (2008). https://doi.org/10.1007/s00453-007-9036-3
5. Bollig, B., Waack, S., Woelfel, P.: Parity graph-driven read-once branching programs and an exponential lower bound for integer multiplication. In: Baeza-Yates, R., Montanari, U., Santoro, N. (eds.) Foundations of Information Technology in the Era of Network and Mobile Computing. ITIFIP, vol. 96, pp. 83–94. Springer, Boston, MA (2002). https://doi.org/10.1007/978-0-387-35608-2_8
6. Bollig, B., Woelfel, P.: A read-once branching program lower bound of $\Omega(2^{n/4})$ for integer multiplication using universal hashing. In: Proceedings of 33rd ACM STOC, pp. 419–424 (2001). https://doi.org/10.1145/380752.380835
7. Carter, J.L., Wegman, M.N.: Universal classes of hash functions. J. Comp. Syst. Sci. **18**, 143–154 (1979). https://doi.org/10.1016/0022-0000(79)90044-8
8. Chor, B., Goldreich, O.: On the power of two-point based sampling. J. Complex. **5**, 96–106 (1989). https://doi.org/10.1016/0885-064X(89)90015-0
9. Dahlgaard, S., Knudsen, M.B.T., Thorup, M.: Practical hash functions for similarity estimation and dimensionality reduction. In: Proceedings of NIPS 30, pp. 6618–6628 (2017). http://papers.nips.cc/paper/7239-practical-hash-functions-for-similarity-estimation-and-dimensionality-reduction.pdf
10. Dietzfelbinger, M.: Universal hashing and k-wise independent random variables via integer arithmetic without primes. In: Puech, C., Reischuk, R. (eds.) STACS 1996. LNCS, vol. 1046, pp. 567–580. Springer, Heidelberg (1996). https://doi.org/10.1007/3-540-60922-9_46
11. Dietzfelbinger, M., Gil, J., Matias, Y., Pippenger, N.: Polynomial hash functions are reliable. In: Kuich, W. (ed.) ICALP 1992. LNCS, vol. 623, pp. 235–246. Springer, Heidelberg (1992). https://doi.org/10.1007/3-540-55719-9_77
12. Dietzfelbinger, M., Hagerup, T., Katajainen, J., Penttonen, M.: A reliable randomized algorithm for the closest-pair problem. J. Algorithms **25**(1), 19–51 (1997). https://doi.org/10.1006/jagm.1997.0873
13. Dietzfelbinger, M., Karlin, A., Mehlhorn, K., Meyer auf der Heide, F., Rohnert, H., Tarjan, R.E.: Dynamic perfect hashing: upper and lower bounds. SIAM J. Comput. **23**(4), 738–761 (1994). https://doi.org/10.1137/S0097539791194094
14. Dietzfelbinger, M., Meyer auf der Heide, F.: Dynamic hashing in real time. In: Buchmann, J., Ganzinger, H., Paul, W.J. (eds.) Informatik: Festschrift zum 60. Geburtstag von Günter Hotz, Teubner-Texte zur Informatik, vol. 1, pp. 95–119. B. G. Teubner, Stuttgart-Leipzig (1992). https://doi.org/10.1007/978-3-322-95233-2_7
15. Dietzfelbinger, M., Schellbach, U.: On risks of using cuckoo hashing with simple universal hash classes. In: Proceedings of 20th ACM-SIAM SODA, pp. 795–804 (2009). https://doi.org/10.1137/1.9781611973068.87

16. Dietzfelbinger, M., Schellbach, U.: Weaknesses of cuckoo hashing with a simple universal hash class: the case of large universes. In: Nielsen, M., Kučera, A., Miltersen, P.B., Palamidessi, C., Tůma, P., Valencia, F. (eds.) SOFSEM 2009. LNCS, vol. 5404, pp. 217–228. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-540-95891-8_22
17. Dietzfelbinger, M., Woelfel, P.: Almost random graphs with simple hash functions. In: Proceedings of 35th ACM STOC, pp. 629–638 (2003). https://doi.org/10.1145/780542.780634
18. Fredman, M.L., Komlós, J., Szemerédi, E.: Storing a sparse table with $O(1)$ worst case access time. J. Assoc. Comput. Mach. **31**(3), 538–544 (1984). https://doi.org/10.1145/828.1884
19. Hagerup, T., Miltersen, P.B., Pagh, R.: Deterministic dictionaries. J. Algorithms **41**(1), 69–85 (2001). https://doi.org/10.1006/jagm.2001.1171
20. Hardy, G.H., Wright, E.M.: An Introduction to the Theory of Numbers. Oxford Science Publications, Clarendon Press, Oxford (1994)
21. Knudsen, M.B.T.: Linear hashing is awesome. CoRR arXiv:1706.02783 (2017). Technical report version of Knudsen, M.B.T.: Linear hashing is awesome. In: Proceedings of 57th IEEE FOCS, pp. 345–352 (2016). https://doi.org/10.1109/FOCS.2016.45
22. Kopelowitz, T., Pettie, S., Porat, E.: Higher lower bounds from the 3SUM conjecture. In: Proceedings of 27th ACM-SIAM SODA, pp. 1272–1287 (2016). https://doi.org/10.1137/1.9781611974331.ch89
23. Luby, M.: A simple parallel algorithm for the maximal independent set problem. SIAM J. Comput. **15**, 1036–1053 (1986). https://doi.org/10.1137/0215074
24. Luby, M., Wigderson, A.: Pairwise independence and derandomization. In: Foundations and Trends in Theoretical Computer Science, vol. 1, no. 4 (2005). https://doi.org/10.1561/0400000009
25. Mansour, Y., Nisan, N., Tiwari, P.: The computational complexity of universal hashing. Theor. Comput. Sci. **107**, 121–133 (1993). https://doi.org/10.1016/0304-3975(93)90257-T
26. Mehlhorn, K., Vishkin, U.: Randomized and deterministic simulations of PRAMs by parallel machines with restricted granularity of parallel memories. Acta Inform. **21**, 339–374 (1984). https://doi.org/10.1007/BF00264615
27. Miltersen, P.B.: Error correcting codes, perfect hashing circuits, and deterministic dynamic dictionaries. In: Proceedings of 9th ACM-SIAM SODA, pp. 556–563 (1998). https://dl.acm.org/citation.cfm?id=314613.314845
28. Nisan, N.: Pseudorandom generators for space-bounded computations. Combinatorica **12**(4), 449–461 (1992). https://doi.org/10.1007/BF01305237
29. Pagh, R.: Hash and displace: efficient evaluation of minimal perfect hash functions. In: Dehne, F., Sack, J.-R., Gupta, A., Tamassia, R. (eds.) WADS 1999. LNCS, vol. 1663, pp. 49–54. Springer, Heidelberg (1999). https://doi.org/10.1007/3-540-48447-7_5
30. Pătraşcu, M.: Towards polynomial lower bounds for dynamic problems. In: Proceedings of 42nd ACM STOC, pp. 603–610 (2010). https://doi.org/10.1145/1806689.1806772
31. Pătraşcu, M., Thorup, M.: On the k-independence required by linear probing and minwise independence. ACM Trans. Algorithms **12**(1), 8:1–8:27 (2016). https://doi.org/10.1145/2716317
32. Pagh, R., Rodler, F.F.: Cuckoo hashing. J. Algorithms **51**(2), 122–144 (2004). https://doi.org/10.1016/j.jalgor.2003.12.002

33. Ružić, M.: Making deterministic signatures quickly. ACM Trans. Algorithms **5**(3), 26:1–26:26 (2009). https://doi.org/10.1145/1541885.1541887
34. Sauerhoff, M., Woelfel, P.: Time-space tradeoff lower bounds for integer multiplication and graphs of arithmetic functions. In: Proceedings of 35th ACM STOC, pp. 186–195 (2003). https://doi.org/10.1145/780542.780571
35. Siegel, A.: On universal classes of extremely random constant-time hash functions. SIAM J. Comput. **33**(3), 505–543 (2004). https://doi.org/10.1137/S0097539701386216
36. Thorup, M.: Even strongly universal hashing is pretty fast. In: Proceedings of 11th ACM-SIAM SODA, pp. 496–497 (2000). https://dl.acm.org/citation.cfm?id=338219.338597
37. Thorup, M., Zhang, Y.: Tabulation-based 5-independent hashing with applications to linear probing and second moment estimation. SIAM J. Comput. **41**(2), 293–331 (2012). https://doi.org/10.1137/100800774
38. Williams, V.V., Williams, R.: Finding, minimizing, and counting weighted subgraphs. SIAM J. Comput. **42**(3), 831–854 (2013). https://doi.org/10.1137/09076619X
39. Wang, J.R.: Space-efficient randomized algorithms for K-SUM. In: Schulz, A.S., Wagner, D. (eds.) ESA 2014. LNCS, vol. 8737, pp. 810–829. Springer, Heidelberg (2014). https://doi.org/10.1007/978-3-662-44777-2_67
40. Wegman, M.N., Carter, J.L.: New classes and applications of hash functions. In: Proceedings of 20th IEEE FOCS, pp. 175–182 (1979). https://doi.org/10.1109/SFCS.1979.26
41. Woelfel, P.: Efficient strongly universal and optimally universal hashing. In: Kutyłowski, M., Pacholski, L., Wierzbicki, T. (eds.) MFCS 1999. LNCS, vol. 1672, pp. 262–272. Springer, Heidelberg (1999). https://doi.org/10.1007/3-540-48340-3_24
42. Woelfel, P.: Klassen universeller Hashfunktionen mit ganzzahliger Arithmetik. Diploma thesis, University of Dortmund (2000). (In German)
43. Woelfel, P.: New bounds on the OBDD-Size of integer multiplication via universal hashing. In: Ferreira, A., Reichel, H. (eds.) STACS 2001. LNCS, vol. 2010, pp. 563–574. Springer, Heidelberg (2001). https://doi.org/10.1007/3-540-44693-1_49