RAIRO-Theor. Inf. Appl. 58 (2024) 10
https://doi.org/10.1051/ita/2024008

RAIRO - Theoretical Informatics and Applications
www.rairo-ita.org

FINDING AUTOMATIC SEQUENCES WITH FEW CORRELATIONS

Vincent Jugé[1,*] and Irène Marcovici[2]

Abstract. Although automatic sequences are algorithmically very simple, some of them have pseudo-random properties. In particular, some automatic sequences such as the Golay–Shapiro sequence are known to be 2-uncorrelated, meaning that they have the same correlations of order 2 as a uniform random sequence. However, the existence of ℓ-uncorrelated automatic sequences (for $\ell \geqslant 3$) was left as an open question in a recent paper of Marcovici, Stoll and Tahay. We exhibit binary block-additive sequences that are 3-uncorrelated and, with the help of analytical results supplemented by an exhaustive search, we present a complete picture of the correlation properties of binary block-additive sequences of rank $r \leqslant 5$, and ternary sequences of rank $r \leqslant 3$.

Mathematics Subject Classification. 68R15, 11B85, 11K36.

Received December 14, 2022. Accepted February 23, 2024.

1. Introduction

A *k-automatic sequence* is a sequence that can be computed by a finite automaton in the following way: the nth term of the sequence is a function of the state reached by the automaton after reading the representation of the integer n in base k. Alternatively, a k-automatic sequence can also be described with the help of an infinite fixed point of a k-uniform morphism. We refer to the book of Allouche and Shallit [1] for a complete survey on automatic sequences. It is known that some automatic sequences present pseudo-random properties. In particular, a succession of works [2–4] has shown that different generalisations of the Golay–Shapiro sequence (also known as the Rudin–Shapiro sequence, see Ex. 3.4) have the same correlations of order 2 as a sequence of symbols chosen uniformly and independently at random. On the other hand, as the subword complexity of an automatic sequence is at most linear (see [1], Sect. 10.3), it is clear that automatic sequences cannot look "too much" like random sequences. In this work, we continue to address the question of "how random" an automatic sequence can look.

As in the references cited above, we focus on *block-additive* automatic sequences [3, 5, 6], also known as *digital sequences* [7]. They are obtained by sliding the representation of the integer n in base k with a window of length r, and summing the *weights* of the subwords read, for a given weight function; see Section 2 for a formal definition.

Keywords and phrases: Combinatorics on words, automata sequence, well-distributed sequence.

[1] Univ Gustave Eiffel, CNRS, LIGM, F-77454 Marne-la-Vallée, France.
[2] Univ Rouen Normandie, CNRS, Normandie Univ, LMRS UMR 6085, F-76000 Rouen, France.

[*] Corresponding author: vincent.juge@univ-eiffel.fr

In [3], the existence of a block-additive sequence being *ℓ-uncorrelated* for an integer $\ell \geqslant 3$ (*i.e.*, having the same correlations of order ℓ as a uniform random sequence) was left as an open question. In Section 3, we prove that, when ℓ is even, a binary block-additive sequence is ℓ-uncorrelated if and only if it is $(\ell+1)$-uncorrelated (Thm. 3.3). As a consequence, all the binary sequences that are known to be 2-uncorrelated are also 3-uncorrelated.

Then, in Section 4, we present a semi-decision algorithm providing a criterion for being *ℓ-correlated* (Alg. 1 and Thm. 4.1). Conversely, Section 5 provides new explicit criteria to ensure that a function is 2-uncorrelated. These criteria are based on a broadening of the notion of *fibre* introduced in [3], and extend the *difference condition* that was defining the previous generalisations of the Golay–Shapiro sequence. Combining these results, an exhaustive search allows us to obtain a complete description of the correlation properties of binary block-additive sequences of rank $r \leqslant 5$, and ternary sequences of rank $r \leqslant 3$ (Thm. 5.13). Finally, we complete the panorama of uncorrelated sequences by giving in Section 6 some relations between uncorrelated sequences, and we conclude by some discussion and open questions in Section 7.

2. Definitions

Below, let $\mathbb{N}$ denote the set of non-negative integers. For all integers $k \geqslant 0$, let Σ_k denote the set $\{0, 1, \ldots, k-1\}$ and let $\mathbb{Z}_k$ denote the set $\mathbb{Z}/k\mathbb{Z}$. For all finite sets $\mathscr{S}$, let $|\mathscr{S}|$ denote the cardinality of $\mathscr{S}$, *i.e.*, the number of elements of $\mathscr{S}$.

In general, let $\mathbf{0}$ (resp., $\mathbf{1}$) denote a tuple whose coordinates are all equal to 0 (resp., to 1), and let $\mathbf{1}_i$ denote a tuple whose ith coordinate is 1 and the other coordinates are 0; the dimension of the tuple is left implicit. Moreover, given a sequence $(u_n)_{n \geqslant 0}$ and a tuple $\boldsymbol{\delta} = (\delta_1, \delta_2, \ldots, \delta_\ell)$ of non-negative integers, let $u_{n+\boldsymbol{\delta}}$ denote the tuple $(u_{n+\delta_1}, u_{n+\delta_2}, \ldots, u_{n+\delta_\ell})$. We also say that $\boldsymbol{\delta}$ is *increasing* if $\delta_1 < \delta_2 < \cdots < \delta_\ell$, and that $\boldsymbol{\delta}$ is *initial* if it is increasing and $\delta_1 = 0$.

Finally, for all integers $n \geqslant 0$ and $k \geqslant 1$, let $(n \bmod k)$ denote the unique element $x \in \Sigma_k$ such that k divides $n - x$. If $k \geqslant 2$, let $\langle n \rangle_k$ denote the *little-endian* representation of n in base k, *i.e.*, the unique sequence $(x_i)_{i \geqslant 0}$ with values in Σ_k such that

$$n = \sum_{i \geqslant 0} x_i k^i.$$

Definition 2.1. Let $k \geqslant 2$ and $r \geqslant 1$ be integers, and let $f\colon \Sigma_k^r \to \mathbb{Z}_k$ be a function such that $f(\mathbf{0}) = 0$. For all $n \geqslant 0$, let u_n be the element of $\mathbb{Z}_k$ defined by

$$u_n = \sum_{i \geqslant 0} f(x_i, x_{i+1}, \ldots, x_{i+r-1}),$$

where $(x_i)_{i \geqslant 0} = \langle n \rangle_k$. The sequence $(u_n)_{n \geqslant 0}$ is said to be *block-additive* in base k with rank r, and we say that this sequence is *associated* with the function f.

The block-additive sequence $(u_n)_{n \geqslant 0}$ associated with a function f also has the following characterisation.

Remark 2.2. Let $\mathscr{A}$ be the automaton over the alphabet Σ_k with state set $Q = \mathbb{Z}_k \times \Sigma_k^{r-1}$, initial state $q_0 = (0, \mathbf{0})$ and transition function $\Delta\colon Q \times \Sigma_k \to Q$ defined by

$$\Delta\colon\ ((v, (x_1, x_2, \ldots, x_{r-1})), i) \to (v + f(i, x_1, x_2, \ldots, x_{r-1}), (i, x_1, x_2, \ldots, x_{r-2})).$$

The value of u_n is obtained by letting the automaton $\mathscr{A}$ read the infinite word $(x_i)_{i\geqslant 0} = \langle n\rangle_k$ *from right to left* (because we us little-endian representations):[1] $\mathscr{A}$ goes through each tuple $(x_i, x_{i+1}, \ldots, x_{i+r-1})$, stores its $r-1$ coordinates $(x_i, x_{i+1}, \ldots, x_{i+r-2})$ and accumulates, in the first component of each state, the sum of the values of $f(x_i, x_{i+1}, \ldots, x_{i+r-1})$ it has already encountered.

Let $\phi\colon Q^* \to Q^*$ be the morphism of monoids that sends a state $s \in Q$ to the word $\phi(s) = \Delta(s,0)\Delta(s,1)\ldots\Delta(s,k-1) \in Q^k$. Projecting on their first component the letters of the infinite fixed-point $\phi^\omega(q_0) \in Q^{\mathbb{N}}$ provides us with the infinite word $u_0u_1u_2\ldots$, which we identify with the sequence $(u_n)_{n\geqslant 0}$ itself. The automaton $\mathscr{A}$ is said to *generate* the sequence $(u_n)_{n\geqslant 0}$.

Definition 2.3. Let $\ell \geqslant 1$ be an integer and let $\mathscr{S}$ be a finite set. Given a sequence $\mathbf{u} = (u_n)_{n\geqslant 0}$ with values in $\mathscr{S}$, an increasing tuple $\boldsymbol{\delta} \in \mathbb{N}^\ell$ and a subset $\mathbf{S}$ of $\mathscr{S}^\ell$ called a *pattern set*, we introduce the set

$$\mathscr{D}^{\mathbf{u}}_{\boldsymbol{\delta}}(\mathbf{S}) = \{n \in \mathbb{N}\colon u_{n+\boldsymbol{\delta}} \in \mathbf{S}\}.$$

We define the *frequency* of the pattern set $\mathbf{S}$ in the sequence $(u_{n+\boldsymbol{\delta}})_{n\geqslant 0}$ as the real number

$$\mathsf{freq}^{\mathbf{u}}_{\boldsymbol{\delta}}(\mathbf{S}) = \lim_{N\to+\infty} \frac{|\mathscr{D}^{\mathbf{u}}_{\boldsymbol{\delta}}(\mathbf{S}) \cap \Sigma_N|}{N},$$

when this limit exists.

Note that, if the automaton $\mathscr{A}$ that generates $\mathbf{u}$ is strongly connected, the sequence $\mathbf{u}$ is a morphic primitive sequence, which ensures that all the densities are well-defined.

We say that $\mathbf{u}$ is *ℓ-uncorrelated* if $\mathsf{freq}^{\mathbf{u}}_{\boldsymbol{\delta}}(\mathbf{S}) = |\mathbf{S}|/|\mathscr{S}|^\ell$ for all tuples $\boldsymbol{\delta}$ and all sets $\mathbf{S} \subseteq \mathscr{S}^\ell$. Otherwise, we say that $\mathbf{u}$ is *ℓ-correlated*.

A sequence whose terms are chosen independently and uniformly at random in $\mathbb{Z}_n$ is almost-surely ℓ-uncorrelated for every integer $\ell \geqslant 0$. However, no automatic sequence, and in particular no block-additive sequence, can be ℓ-uncorrelated for every $\ell \geqslant 0$, since this would in particular require the sequence to be normal, while the subword complexity of an automatic sequence is at most linear.

3. Correlations, block-additivity and base 2

It follows from Remark 2.2 that, if the automaton $\mathscr{A}$ that generates the block-additive sequence $\mathbf{u}$ is strongly connected, all the letters of Σ_k (*i.e.*, patterns of length 1) have a well-defined frequency in $\mathbf{u}$. Furthermore, this frequency is equal to $1/k$. Indeed, each state in Q is the target of k edges from $\mathscr{A}$, so that the uniform probability measure on Q is the unique stationary measure of the Markov chain associated with $\mathscr{A}$, where each edge has weight $1/k$. We extend this argument to prove that this property still holds for the subsequences of $\mathbf{u}$ of the form $(u_{a+k^bn})_{n\geqslant 0}$, as stated in the lemma below.

Lemma 3.1. *Let $(u_n)_{n\geqslant 0}$ be a block-additive sequence in base $k \geqslant 2$ whose generating automaton is strongly connected. For all integers $a \geqslant 0$, $b \geqslant 0$ and all $s \in \mathbb{Z}_k$,*

$$\lim_{N\to+\infty} \frac{|\{n \in \Sigma_N : u_{a+k^bn} = s\}|}{N} = \frac{1}{k}.$$

Proof. First, observe that if the result is true for some pair (a,b) with $a \geqslant k^b$, then it is also true for the pair $(a-k^b, b)$. So, we can assume without loss of generality that $a < k^b$. The sequence $(u_{a+k^bn})_{n\geqslant 0}$ is a result of a process very similar to the one described in Remark 2.2. Let us consider the automaton $\mathscr{A}$ that generates $\mathbf{u}$.

[1] This procedure is well-defined because $(x_i)_{i\geqslant 0}$ ends with 0 terms only, and the automaton does not leave its initial state when reading a 0. Alternatively, we might truncate the sequence $(x_i)_{i\geqslant 0}$ at least r positions after its rightmost non-zero term, and let the automaton read the resulting finite word from right to left.

Instead of transforming the word $\phi^\omega(q_0)$ into the sequence $(u_n)_{n\geqslant 0}$ by using directly the projection τ on the first component, we use the morphism $\tau_{a,b}\colon Q \to \mathbb{Z}_k$ defined by

$$\tau_{a,b} = \tau \circ \Delta(\,\cdot\,,\alpha_0) \circ \Delta(\,\cdot\,,\alpha_1) \circ \cdots \circ \Delta(\,\cdot\,,\alpha_{b-1}),$$

where $(\alpha_i)_{i\geqslant 0} = \langle a \rangle_p$ and each function $\Delta(\cdot, i)$ is identified with a function from Q to itself.

Since $\mathscr{A}$ has a self-loop around the state q_0, and identifying $\mathscr{A}$ with a Markov chain $\mathbf{M}$ in which each edge has probability $1/k$, this Markov chain is ergodic. Then, since each state in Q is the target of k edges from $\mathscr{A}$, the uniform probability measure on Q is the unique stationary measure of $\mathbf{M}$. Consequently, the letters of the infinite word $\phi^\omega(q_0)$ have the same density.

Moreover, the relations

$$\tau_{a,b}(u, x_1, x_2, \ldots, x_r) = u + \tau_{a,b}(0, x_1, x_2, \ldots, x_r),$$

hold for all states $(u, x_1, x_2, \ldots, x_r) \in Q$, which proves that all elements of $\mathbb{Z}_k$ have the same number of antecedents by $\tau_{a,b}$. Hence, the k possible letters of the word $\tau_{a,b} \circ \phi^\omega(q_0) = (u_{k^b n + a})_{n\geqslant 0}$ all have the same density. □

From this lemma, we derive the following characterisation of ℓ-uncorrelated block-automatic sequences.

Proposition 3.2. *Let $\mathbf{u} = (u_n)_{n\geqslant 0}$ be a block-additive sequence in base $k \geqslant 2$, and let $\ell \geqslant 2$ be an integer. The sequence $\mathbf{u}$ is ℓ-uncorrelated if and only if*

$$\mathsf{freq}^{\mathbf{u}}_{\boldsymbol{\delta}}(\mathbf{s} + \mathbb{Z}_k \mathbf{1}) = \frac{1}{k^{\ell-1}}$$

for all initial tuples $\boldsymbol{\delta} \in \mathbb{N}^\ell$ and all tuples $\mathbf{s} \in \mathbb{Z}_k^\ell$, where $\mathbf{s} + \mathbb{Z}_k\mathbf{1}$ denotes the set $\{\mathbf{s} + x\,\mathbf{1} \colon x \in \mathbb{Z}_k\}$.

Proof. The "only if" part is a direct consequence of Definition 2.3. Thus, we focus on the "if" part. Let $\mathbf{u}$ be a sequence that meets the requirements of Proposition 3.2. We wish to prove that $\mathbf{u}$ is ℓ-uncorrelated.

First, we prove that the automaton $\mathscr{A}$ that generates $\mathbf{u}$ is strongly connected. To do this, it is sufficient to show that, for all $x \in \mathbb{Z}_k$, the state $(x, \mathbf{0})$ is accessible from the initial state $q_0 = (0, \mathbf{0})$.

Let r be the rank of $\mathbf{u}$, and let $\boldsymbol{\delta}$ be an initial tuple such that $\delta_2 = k^{r-1}$. Since we assumed that $\mathsf{freq}^{\mathbf{u}}_{\boldsymbol{\delta}}(\mathbf{1}_2 + \mathbb{Z}_k\mathbf{1}) = k^{1-\ell} > 0$, there exists an integer n and an element t of $\mathbb{Z}_k$ such that $u_n = t$ and $u_{n+k^{r-1}} = t+1$. Noting $(x_i)_{i\geqslant 0}$ and $(x'_i)_{i\geqslant 0}$ the representations of n and $n + k^{r-1}$ in base k, we observe that $x_i = x'_i$ for all $i \leqslant r-2$. Thus, both states $q_1 = (t, (x_0, x_1, \ldots, x_{r-2}))$ and $q_2 = (t+1, (x_0, x_1, \ldots, x_{r-2}))$ are accessible in $\mathscr{A}$. Then, starting from q_1 and following $r-1$ times the transition with label 0, we arrive in a state $(y, \mathbf{0})$; starting from q_2, we would have arrived in the state $(y+1, \mathbf{0})$.

Consequently, the set $\mathscr{G} = \{x \in \mathbb{Z}_k \colon (x, \mathbf{0}) \text{ is accessible from } \mathbf{0} \text{ in } \mathscr{A}\}$ contains both y and $y+1$. Since $\mathscr{G}$ is a subgroup of $(\mathbb{Z}_k, +)$, it must coincide with $\mathbb{Z}_k$ itself, and $\mathscr{A}$ is strongly connected.

Now, let $\boldsymbol{\delta} = (\delta_1, \delta_2, \ldots, \delta_\ell) \in \mathbb{N}^\ell$ be an arbitrary initial tuple. Then, consider an element $\mathbf{s} = (s_1, s_2, \ldots, s_\ell)$ of $\mathbb{Z}_k^\ell$ and a real number $\varepsilon > 0$. Let $b \geqslant \log_k(\delta_\ell/\varepsilon)$ be an integer such that

$$|\mathscr{D}^{\mathbf{u}}_{\boldsymbol{\delta}}(\mathbf{s} + \mathbb{Z}_k\mathbf{1}) \cap \Sigma_{k^{b+r}}| \geqslant (k^{1-\ell} - \varepsilon)k^{b+r}.$$

Finally, let $\mathscr{S}$ be the set of integers $n \geqslant 0$ such that $(n \bmod k^b) \in \Sigma_{(1-\varepsilon)k^b}$.

For every integer $n \in \mathscr{S}$, and since $(1-\varepsilon)k^b \leqslant k^b - \delta_\ell$, the value of the tuple $u_{n+\boldsymbol{\delta}} - u_n\,\mathbf{1}$ depends only on $(n \bmod k^{b+r})$; indeed, for all $i \leqslant \ell$, only the b least significant bits of $n + \delta_i$ may differ from those of n. Consequently, for all integers $a \in \mathscr{S} \cap \mathscr{D}^{\mathbf{u}}_{\boldsymbol{\delta}}(\mathbf{s} + \mathbb{Z}_k\mathbf{1}) \cap \Sigma_{k^{b+r}}$ and all integers $n \geqslant 0$ such that $a = (n \bmod k^{b+r})$, we have $u_{n+\delta} = \mathbf{s}$ as soon as $u_n = s_1$. Furthermore, since $\mathscr{A}$ is strongly connected, Lemma 3.1 proves that there exists an integer N_0 such that

$$\frac{|\{n \in \Sigma_N : u_{a+k^{b+r}n} = s_1\}|}{N} \geqslant \frac{1-\varepsilon}{k}$$

for all integers $a < k^{b+r}$ and $N \geqslant N_0$.

Finally, consider an arbitrary integer $M \geqslant k^{b+r} N_0$, and let $N = \lfloor M/k^{b+r} \rfloor$. The set $\mathscr{S} \cap \mathscr{D}^{\mathbf{u}}_{\boldsymbol{\delta}}(\mathbf{s} + \mathbb{Z}_k \mathbf{1}) \cap \Sigma_{k^{b+r}}$ contains at least $(k^{1-\ell} - \varepsilon)k^{b+r} - k^r \varepsilon \geqslant (k^{-\ell} - \varepsilon)k^{b+r+1}$ elements a and, for each such a, the set $\mathscr{S} \cap \Sigma_M$ contains at least $(1-\varepsilon)N/k$ integers n for which $a = (n \bmod k^{b+r})$ and $u_n = s_1$. It follows that

$$\begin{aligned} |\mathscr{D}^{\mathbf{u}}_{\boldsymbol{\delta}}(\{\mathbf{s}\}) \cap \Sigma_M| &\geqslant |\mathscr{D}^{\mathbf{u}}_{\boldsymbol{\delta}}(\{\mathbf{s}\}) \cap \mathscr{S} \cap \Sigma_M| \\ &\geqslant (1-\varepsilon)(k^{-\ell} - \varepsilon)k^{b+r} N \\ &\geqslant (1-\varepsilon)(k^{-\ell} - \varepsilon)(M - k^{b+r}). \end{aligned}$$

We conclude that $\mathsf{freq}^{\mathbf{u}}_{\boldsymbol{\delta}}(\{\mathbf{s}\}) \geqslant k^{-\ell}$ by choosing ε arbitrarily small and M arbitrarily large. This inequality being valid for the k^ℓ elements $\mathbf{s}$ of $\mathbb{Z}_k^\ell$, it is in fact an equality. Finally, summing these equalities for all $\mathbf{s} \in \mathbf{S}$ proves that $\mathsf{freq}^{\mathbf{u}}_{\boldsymbol{\delta}}(\mathbf{S}) = |\mathbf{S}|/k^\ell$. □

Theorem 3.3. *Let ℓ be an even positive integer and let $\mathbf{u}$ be a block-additive sequence in base 2. This sequence is ℓ-uncorrelated if and only if it is $(\ell+1)$-uncorrelated.*

Proof. First, every $(\ell+1)$-uncorrelated sequence is clearly ℓ-uncorrelated. Conversely, let $\mathbf{u}$ be an ℓ-uncorrelated block-additive sequence in base 2, let $\boldsymbol{\delta} \in \mathbb{N}^{\ell+1}$ be an increasing tuple, and let $\mathbf{s}$ be an element of $\mathbb{Z}_2^{\ell+1}$. Let $|\mathbf{s}|_1$ denote the number of entries equal to 1 in the tuple $\mathbf{s}$, *i.e.*, $|\mathbf{s}|_1 = |\{i : s_i = 1\}|$.

Since $\mathbf{u}$ is ℓ-uncorrelated, we know that $\mathsf{freq}^{\mathbf{u}}_{\boldsymbol{\delta}}(\mathbf{s}) + \mathsf{freq}^{\mathbf{u}}_{\boldsymbol{\delta}}(\mathbf{s} + \mathbf{1}_i) = 2^{-\ell}$ for all tuples $\mathbf{s} \in \mathbb{Z}_2^{\ell+1}$ and all integers $i \leqslant \ell + 1$. Hence, and since $\ell+1$ is odd, an immediate induction on $|\mathbf{s}|_1$ proves that $\mathsf{freq}^{\mathbf{u}}_{\boldsymbol{\delta}}(\mathbf{s}) = \mathsf{freq}^{\mathbf{u}}_{\boldsymbol{\delta}}(\mathbf{0})$ if $|\mathbf{s}|_1$ is even, and that $\mathsf{freq}^{\mathbf{u}}_{\boldsymbol{\delta}}(\mathbf{s}) = \mathsf{freq}^{\mathbf{u}}_{\boldsymbol{\delta}}(\mathbf{1})$ otherwise. It follows that

$$\mathsf{freq}^{\mathbf{u}}_{\boldsymbol{\delta}}(\mathbf{s} + \mathbb{Z}_2 \mathbf{1}) = \mathsf{freq}^{\mathbf{u}}_{\boldsymbol{\delta}}(\mathbb{Z}_2 \mathbf{1}) = 2^{-\ell}$$

for all $\mathbf{s} \in \mathbb{Z}_2^{\ell+1}$, and Proposition 3.2 then proves that $\mathbf{u}$ is $(\ell+1)$-uncorrelated. □

We prove now, by giving two examples, that the conclusions of Theorem 3.3 are no longer ensured if ℓ is odd or if $(u_n)_{n \geqslant 0}$ is block-additive in a base $k \geqslant 3$. These examples also serve as toy cases for Theorem 4.1 below.

Example 3.4. Let $(u_n)_{n \geqslant 0} \in \Sigma_2^{\mathbb{N}}$ be the block-additive sequence associated with the function $f \colon \Sigma_2^2 \mapsto \mathbb{Z}_2$ defined by $f(x_1, x_2) = x_1 x_2$. This sequence is known as the *Golay–Shapiro* or *Rudin–Shapiro* sequence [8], Remark 1. It is both 3-uncorrelated and 4-correlated.

Proof. It is known (see [9] and the subsequent generalisations [2–4]) that $(u_n)_{n \geqslant 0}$ is 2-uncorrelated, and thus Theorem 3.3 proves it is also 3-uncorrelated. Consider the tuples $\boldsymbol{\delta} = (0, 1, 2, 3) \in \mathbb{N}^4$ and $\mathbf{s} = (0, 0, 0, 1) \in \mathbb{Z}_2^4$. Let $(x_i)_{i \geqslant 0} = \langle n \rangle_2$ be the representation of an integer n in base 2. If $n \in 8\mathbb{N}$, *i.e.*, if $x_0 = x_1 = x_2 = 0$, one checks that

$$u_n = u_{n+1} = u_{n+2} = u_{n+3} - 1 = \sum_{i \geqslant 3} x_i x_{i+1}.$$

This shows that $u_{n+\boldsymbol{\delta}} = \mathbf{s} + u_n \mathbf{1}$, *i.e.*, that $n \in \mathscr{D}^{\mathbf{u}}_{\boldsymbol{\delta}}(\mathbf{s} + \mathbb{Z}_2 \mathbf{1})$. One checks similarly that $n \in \mathscr{D}^{\mathbf{u}}_{\boldsymbol{\delta}}(\mathbf{s} + \mathbb{Z}_2 \mathbf{1})$ if $n \in 16\mathbb{N} + 11$. Hence, $\mathsf{freq}^{\mathbf{u}}_{\boldsymbol{\delta}}(\mathbf{s} + \mathbb{Z}_2 \mathbf{1}) \geqslant 3/16 > 2^{-3}$, so that $(u_n)_{n \geqslant 0}$ is 4-correlated. □

Example 3.5. Let $(v_n)_{n \geqslant 0} \in \Sigma_3^{\mathbb{N}}$ be the block-additive sequence associated with the function $f \colon \Sigma_3^2 \mapsto \mathbb{Z}_3$ defined by $f(x_1, x_2) = x_1 x_2$. The sequence $(v_n)_{n \geqslant 0}$ is both 2-uncorrelated and 3-correlated.

Proof. It is known (same references as above) that $(v_n)_{n\geqslant 0}$ is 2-uncorrelated. Consider the tuple $\boldsymbol{\delta} = (0,1,2) \in \mathbb{N}^3$, and let $(x_i)_{i\geqslant 0} = \langle n\rangle_3$ be the representation of an integer n in base 3. If $n \in 9\mathbb{N}$, *i.e.*, if $x_0 = x_1 = 0$, one checks that

$$v_n = v_{n+1} = v_{n+2} = \sum_{i\geqslant 2} x_i x_{i+1};$$

this proves that $n \in \mathscr{D}^{\mathbf{v}}_{\boldsymbol{\delta}}(\mathbb{Z}_3\mathbf{1})$. Similarly, $n \in \mathscr{D}^{\mathbf{v}}_{\boldsymbol{\delta}}(\mathbb{Z}_3\mathbf{1})$ whenever $n \in 27\mathbb{N}+1$. It follows that $\mathsf{freq}^{\mathbf{v}}_{\boldsymbol{\delta}}(\mathbb{Z}_3\mathbf{1}) \geqslant 4/27 > 3^{-2}$, so that $(v_n)_{n\geqslant 0}$ is 3-correlated. □

4. Detecting correlations

In this section, we focus on the following problem. Given an integer $\ell \geqslant 1$ and a block-additive sequence $\mathbf{u} = (u_n)_{n\geqslant 0}$ in base $k \geqslant 2$, is $\mathbf{u}$ ℓ-correlated? We provide two partial results. First, we propose an algorithm for detecting ℓ-correlations, when they exist. This algorithm extends the method used for the sequences of Examples 3.4 and 3.5. Second, we propose a criterion that is sufficient for being 2- or 3-uncorrelated when $k = 2$.

Theorem 4.1. *Algorithm 1 is a semi-decision algorithm for deciding if a given block-additive sequence is ℓ-correlated. More precisely, when given integers $k \geqslant 2$, $r \geqslant 1$, $\ell \geqslant 1$ and a function $f\colon \Sigma_k^r \to \mathbb{Z}_k$ such that $f(\mathbf{0}) = 0$ as input, Algorithm 1 eventually returns **true** if the block-additive sequence associated with f is ℓ-correlated; otherwise, it runs forever.*

Algorithm 1: Detecting ℓ-correlations in block-additive sequences

Input : Integers $k \geqslant 2$, $r \geqslant 1$ and $\ell \geqslant 1$.
Function $f\colon \Sigma_k^r \to \mathbb{Z}_k$ such that $f(\mathbf{0}) = 0$.
Result : **true** if the block-additive sequence $(u_n)_{n\geqslant 0}$ associated with f is ℓ-correlated.

```
for m = 1, 2, 3, ...:
  for all initial tuples δ = (δ_1, ..., δ_ℓ) ∈ N^ℓ such that δ_ℓ ⩽ m:
    for all ℓ-tuples s = (s_1, ..., s_ℓ) ∈ Z_k^ℓ such that s_1 = 0:
      c(s) ← 0
    for n = 0, 1, ..., k^(m+r) − 1:
      if (n mod k^m) ∈ Σ_{k^m − m}:
        s ← u_{n+δ} − u_n 1
        c(s) ← c(s) + 1
        if c(s) > k^(m+r+1−ℓ):
          return true
```

Proof. Let $\mathbf{u} = (u_n)_{n\geqslant 0}$ be the block-additive sequence associated with the function f. For all $m \geqslant 1$, let

$$\mathscr{S}_m = \{n \in \mathbb{N}\colon (n \bmod k^m) \in \Sigma_{k^m - m}\}.$$

Provided that $\delta_\ell \leqslant m$, then, for all integers $n \in \mathscr{S}_m$, the value of $u_{n+\boldsymbol{\delta}} - u_n\,\mathbf{1}$, *i.e.*, the tuple $\mathbf{s} \in \mathbb{Z}_k^\ell$ such that $s_1 = 0$ and $n \in \mathscr{D}^{\mathbf{u}}_{\boldsymbol{\delta}}(\mathbf{s} + \mathbb{Z}_k\mathbf{1})$, depends only on $(n \bmod k^{m+r})$.

If $\mathbf{u}$ is ℓ-correlated, Proposition 3.2 proves that there exists a tuple $\mathbf{s} \in \mathbb{Z}_k^\ell$ and an initial tuple $\boldsymbol{\delta} \in \mathbb{N}^\ell$ such that $\mathsf{freq}^{\mathbf{u}}_{\boldsymbol{\delta}}(\mathbf{s} + \mathbb{Z}_k\mathbf{1}) \neq k^{1-\ell}$. Since the average of these densities is $k^{1-\ell}$, the largest one is larger than $k^{1-\ell}$, *i.e.*,

$$\mathsf{freq}^{\mathbf{u}}_{\boldsymbol{\delta}}(\mathbf{s} + \mathbb{Z}_k\mathbf{1}) \geqslant k^{1-\ell} + 2\varepsilon$$

for some tuple $\mathbf{s} \in \mathbb{Z}_k^\ell$ and some real number $\varepsilon \in (0,1)$. Without loss of generality, we even assume that $s_1 = 0$.

Consider some integer $m \geqslant \delta_\ell$ such that $\varepsilon > m/k^m$ and

$$|\mathscr{D}^{\mathbf{u}}_{\boldsymbol{\delta}}(\mathbf{s} + \mathbb{Z}_k \mathbf{1}) \cap \Sigma_{k^{m+r}}| \geqslant (k^{1-\ell} + \varepsilon) k^{m+r}.$$

By construction, $\mathbf{s} = u_{n+\boldsymbol{\delta}} - u_n\,\mathbf{1}$ for each integer $n \in \mathscr{D}^{\mathbf{u}}_{\boldsymbol{\delta}}(\mathbf{s} + \mathbb{Z}_k \mathbf{1})$. Moreover, at most mk^r elements of $\mathscr{D}^{\mathbf{u}}_{\boldsymbol{\delta}}(\mathbf{s} + \mathbb{Z}_k \mathbf{1}) \cap \Sigma_{k^{m+r}}$ belong to $\mathbb{N} \setminus \mathscr{S}_m$. The

$$(k^{1-\ell} + \varepsilon) k^{m+r} - mk^r > k^{m+r+1-\ell}$$

(or more) remaining elements all contribute to incrementing the counter $\mathsf{c}(\mathbf{s})$, and thus Algorithm 1 returns **true**.

Conversely, if Algorithm 1 returns **true**, there exist an increasing tuple $\boldsymbol{\delta} \in \mathbb{N}^\ell$, a tuple $\mathbf{s} \in \mathbb{Z}_k^\ell$ and an integer m such that

$$|\mathscr{S}_m \cap \mathscr{D}^{\mathbf{u}}_{\boldsymbol{\delta}}(\mathbf{s} + \mathbb{Z}_k \mathbf{1}) \cap \Sigma_{k^{m+r}}| > k^{m+r+1-\ell}.$$

But then, every integer n such that $(n \bmod k^{m+r}) \in \mathscr{S}_m \cap \mathscr{D}^{\mathbf{u}}_{\boldsymbol{\delta}}(\mathbf{s} + \mathbb{Z}_k \mathbf{1})$ belongs to $\mathscr{D}^{\mathbf{u}}_{\boldsymbol{\delta}}(\mathbf{s} + \mathbb{Z}_k \mathbf{1})$, and therefore

$$\mathsf{freq}^{\mathbf{u}}_{\boldsymbol{\delta}}(\mathbf{s} + \mathbb{Z}_k \mathbf{1}) > k^{m+r+1-\ell} / k^{m+r} = k^{1-\ell}.$$ □

5. Constructing 2-uncorrelated functions

A *difference condition*, sufficient for ensuring that a block-additive sequence of rank $r = 2$ is 2-uncorrelated, was developed in [3]. In this section, we extend that condition, and we identify some cases where our extended condition is also necessary. This condition is based on the notion of *fibres*.

Definition 5.1. A partition of $\mathbb{N}$ into a collection of finite sets $(\mathscr{F}_x)_{x \in X}$ is called a *fibration* if it satisfies the criterion C_1 below; the sets $\mathscr{F}_x$ are then called *fibres*.

C_1: The proportion of elements of Σ_N that belong to a part $\mathscr{F}_x$ included in Σ_N is arbitrarily close to 1 when N grows arbitrarily; in other words,

$$\lim_{N \to +\infty} \frac{|\{n \in \Sigma_N : \exists x \in X \text{ s.t. } n \in \mathscr{F}_x \text{ and } \mathscr{F}_x \subseteq \Sigma_N\}|}{N} = 1.$$

Then, let $k \geqslant 2$ and $\delta \geqslant 1$ be integers, and let $\mathbf{u}$ be a block-additive sequence in base k. We say that a finite subset $\mathscr{F}$ of $\mathbb{N}$ is δ-*balanced* for $\mathbf{u}$ if it satisfies the criterion C_2 below.

C_2: The k subsets $\{n \in \mathscr{F} : u_{n+\delta} - u_n = d\}$ obtained when $d \in \mathbb{Z}_k$ have the same cardinality.

Finally, we say that a fibration $(\mathscr{F}_x)_{x \in X}$ is δ-balanced for $\mathbf{u}$ if each fibre $\mathscr{F}_x$ is δ-balanced for $\mathbf{u}$.

Proposition 5.2. *Let $k \geqslant 2$, and let $\mathbf{u}$ be a block-additive sequence in base k. The sequence $\mathbf{u}$ is 2-uncorrelated if and only if it admits a δ-balanced fibration for all $\delta \geqslant 1$.*

Proof. Let $\delta \geqslant 1$ be an integer, let $\boldsymbol{\delta} = (0, \delta)$, and let $\varepsilon > 0$ be an arbitrary positive real number. For all $d \in \mathbb{Z}_k$, let $\mathbf{S}_d$ denote the subset $(0, d) + \mathbb{Z}_k \mathbf{1}$ of $\mathbb{Z}_k^2$.

First, if $\mathbf{u}$ is 2-uncorrelated, we know that $\mathsf{freq}^{\mathbf{u}}_{\boldsymbol{\delta}}(\mathbf{S}_d) = 1/k$ for all $d \in \mathbb{Z}_k$. This means there exists an integer $N \geqslant 0$ such that $|\mathscr{D}^{\mathbf{u}}_{\boldsymbol{\delta}}(\mathbf{S}_d) \cap \Sigma_m| \geqslant (1 - \varepsilon)m/k$ for all integers $m \geqslant N$. In particular, the set $\mathscr{D}^{\mathbf{u}}_{\boldsymbol{\delta}}(\mathbf{S}_d)$ is infinite.

For each integer $x \geqslant 1$, let $\mathscr{F}_x$ be the set that contains the x^{th} smallest elements of each of the k sets $\mathscr{D}^{\mathbf{u}}_{\boldsymbol{\delta}}(\mathbf{S}_d)$. The collection $(\mathscr{F}_x)_{x \geqslant 1}$ forms a partition of $\mathbb{N}$ and each set $\mathscr{F}_x$ clearly satisfies C_2. Moreover, for all $m \geqslant N$, at least $(1 - \varepsilon)m$ elements of Σ_m belong to some set $\mathscr{F}_x$ that is included in Σ_m. Hence, C_1 is also satisfied, and $(\mathscr{F}_x)_{x \geqslant 1}$ is a δ-balanced fibration for $\mathbf{u}$.

Conversely, assume that $\mathbf{u}$ admits a δ-balanced fibration for all integers $\delta \geqslant 1$. Consider some tuple $\boldsymbol{\delta} = (\delta_1, \delta_2)$ such that $0 = \delta_1 < \delta_2$, and let $(\mathscr{F}_x)_{x \in X}$ be a δ_2-balanced fibration for $\mathbf{u}$. Then, let $\mathbf{s}$ be a tuple in $\mathbb{Z}_k^2$, and let $d = s_2 - s_1$. Finally, for all $N \geqslant 0$, let Ω_N be the set $\{x \in X : \mathscr{F}_x \subseteq \Sigma_N\}$. By criterion C_1, for every $\varepsilon > 0$, and if N is large enough, we know that $\sum_{x \in \Omega_N} |\mathscr{F}_x| \geqslant (1-\varepsilon)N$. It follows, for such integers N, that

$$\begin{aligned} |\mathscr{D}^{\mathbf{u}}_{\boldsymbol{\delta}}(\mathbf{s} + \mathbb{Z}_k \mathbf{1}) \cap \Sigma_N| &= |\{n \in \Sigma_N : u_{n+\delta} - u_n = d\}| \\ &\geqslant \sum_{x \in \Omega_N} |\{n \in \mathscr{F}_x : u_{n+\delta} - u_n = d\}| \\ &\geqslant \sum_{x \in \Omega_N} |\mathscr{F}_x|/k \\ &\geqslant (1-\varepsilon)N/k. \end{aligned}$$

This inequality being valid for all tuples $\mathbf{s} \in \mathbb{Z}_k^2$ and for arbitrarily small values of ε, it follows that $\mathsf{freq}^{\mathbf{u}}_{\boldsymbol{\delta}}(\mathbf{s} + \mathbb{Z}_k \mathbf{1}) = 1/k$. By Proposition 3.2, this implies that the sequence $\mathbf{u}$ is 2-uncorrelated. □

Aiming to simplify the problem of deciding whether $\mathbf{u}$ is 2-uncorrelated, we look for simple fibrations. The conditions we are looking for are threefold: (i) the fibres into which $\mathbb{N}$ is partitioned should be as simple as possible; (ii) there should be a finite algorithm for checking whether $\mathbf{u}$ admits such a δ-balanced fibration; (iii) all known cases of 2-uncorrelated sequences should admit such a δ-balanced fibration.

Definition 5.3. Let $k \geqslant 2$, $a \geqslant 0$, $b \geqslant 0$ and $\delta \geqslant 1$ be integers. Given an integer $n \geqslant 0$, let $(x_i)_{i \geqslant 0} = \langle n \rangle_k$ and $(y_i)_{i \geqslant 0} = \langle n + \delta \rangle_k$ be the representations of n and $n + \delta$ in base k, and let $\mathsf{c}_k(n, n+\delta) = \max\{i \in \mathbb{N} : x_i \neq y_i\}$ be the *carry distance* of n and $n + \delta$.

We call (k, a, b, δ)*-fibre* of n the set $\mathscr{F}_{k,a,b,\delta}(n)$ that consists of those integers $m \geqslant 0$ whose representation $(z_i)_{i \geqslant 0}$ in base k satisfies the equality $x_i = z_i$ whenever $i < \mathsf{c}_k(n, n+\delta) - a$, $i \geqslant \mathsf{c}_k(n, n+\delta) + b$ or $x_i \neq y_i$.

Example 5.4. If $k = 2$, $a = 3$, $b = 2$, $\delta = 25$ and $n = 332$, the binary expansions of n and $n + \delta$ are

$$\begin{aligned} (x_i)_{i \geqslant 0} &= 0\,0\,1\,1\,0\,\mathbf{0}\,1\,0\,1\,0\,0\ldots \\ \text{and} \quad (y_i)_{i \geqslant 0} &= 1\,0\,1\,0\,0\,\mathbf{1}\,1\,0\,1\,0\,0\ldots \end{aligned}$$

Their carry distance is $\mathsf{c}_2(332, 357) = 5$, as indicated in bold red: indeed, $x_5 < y_5$, whereas $x_i = y_i$ for all $i \geqslant 6$. Consequently, the fibre $\mathscr{F}_{2,3,1,25}(332)$ consists of those integers $m \geqslant 0$ whose representation $(z_i)_{i \geqslant 0}$ in base 2 is a sequence of the form $0\,0 \bullet 1 \bullet 0 \bullet 0\,1\,0\,0\ldots$, where each $\bullet$ may be a 0 or a 1. In other words,

$$\mathscr{F}_{2,3,2,25}(332) = \{264, 268, 280, 284, 328, 332, 344, 348\}.$$

The sets $\mathscr{F}_{k,a,b,\delta}(n)$, as their name suggests, form a fibration (with duplicates, *i.e.*, a given fibre may be equal to $\mathscr{F}_{k,a,b,\delta}(n)$ for more than one integer $n \geqslant 0$). This is the object of the following two results.

Lemma 5.5. *For all integers $k \geqslant 2$, $a \geqslant 0$, $b \geqslant 0$ and $\delta \geqslant 1$, the (k, a, b, δ)-fibres form a partition of $\mathbb{N}$.*

Proof. Let m be some element of a fibre $\mathscr{F}_{k,a,b,\delta}(n)$, and let $(u_i)_{i \geqslant 0} = \langle m \rangle_k$ and $(v_i)_{i \geqslant 0} = \langle m + \delta \rangle_k$ be the representations of m and $m + \delta$ in base k. An immediate induction on i proves that either $u_i = v_i = x_i = y_i$, or $u_i = v_i \neq x_i = y_i$, or $u_i = x_i \neq y_i = v_i$. Consequently, the carry distances $\mathsf{c}_k(n, n+\delta)$ and $\mathsf{c}_k(m, m+\delta)$ coincide with each other, and so do the fibres $\mathscr{F}_{k,a,b,\delta}(n)$ and $\mathscr{F}_{k,a,b,\delta}(m)$. The desired result follows. □

Lemma 5.6. *For all integers $N \geqslant 0$, there are no more than $k^{b+1} \delta \log_k(N)$ elements of Σ_N whose (k, a, b, δ)-fibre is not included in Σ_N.*

Proof. Consider some integer $n < N$ whose fibre $\mathscr{F}_{k,a,b,\delta}(n)$ is not included in Σ_N. Every element of this fibre is an integer m such that $\mathsf{c}_k(n,m) \leqslant \mathsf{c}_k(n,n+\delta)+b$. Moreover, every integer m such that $\mathsf{c}_k(n,m) < \mathsf{c}_k(n,N)$ belongs to Σ_N. Consequently, $\mathsf{c}_k(n,n+\delta)+b \geqslant \mathsf{c}_k(n,N)$.

Now, let $(z_i)_{i\geqslant 0} = \langle N\rangle_k$ be the expansion of N in base k. For all integers $d \geqslant b$, the integers $n < N$ such that $\mathsf{c}_k(n,N) = d$ are those integers whose expansion $(x_i)_{i\geqslant 0}$ in base k is such that $x_d < z_d$ and $x_i = z_i$ for all $i \geqslant d+1$. Therefore, there are exactly $z_d k^d$ such integers n.

Moreover, among any k^{d-b} consecutive integers, there are exactly $\min\{k^{d-b},\delta\}$ integers n such that $\mathsf{c}_k(n,n+\delta) \geqslant d-b$. Therefore, there are at most $\delta k^b z_d$ integers $n < N$ such that $\mathsf{c}_k(n,N) = d \leqslant \mathsf{c}_k(n,n+\delta)+b$. Summing this for all integers d, there are at most

$$\delta k^b(z_0+z_1+z_2+\cdots) \leqslant \delta k^{b+1}\log_k(N)$$

integers $n < N$ whose fibre is not included in Σ_N. □

Definition 5.7. Let $k \geqslant 2$, $a \geqslant 0$ and $b \geqslant 0$ be integers, and let $\mathbf{u}$ be a block-additive sequence in base k. We say that $\mathbf{u}$ is (a,b)*-strongly* 2*-uncorrelated* if, for each integer $\delta \geqslant 1$, the fibration $(\mathscr{F}_{k,a,b,\delta}(n))_{n\geqslant 0}$ is δ-balanced.

As a special case of Proposition 5.2, we obtain the following sufficient criterion for being 2-uncorrelated.

Proposition 5.8. *Every block-additive sequence that is* (a,b)*-strongly* 2*-uncorrelated for some integers* $a \geqslant 0$ *and* $b \geqslant 0$ *is* 2*-uncorrelated.*

It remains, however, to decide whether the block-additive sequence $\mathbf{u}$ associated with a given function $f\colon \Sigma_k^r \to \mathbb{Z}_k$ is (a,b)-strongly 2-uncorrelated. A first step in that direction consists in observing that we might simply choose $b = r$.

Lemma 5.9. *Let* $a \geqslant 0$*,* $b \geqslant 0$ *and* $r \geqslant 1$ *be integers, and let* $\mathbf{u}$ *be a block-additive sequence with rank* r*. If* $\mathbf{u}$ *is* (a,b)*-strongly* 2*-uncorrelated, it is also* (a,r)*-strongly* 2*-uncorrelated.*

Proof. Let $k \geqslant 2$ be the base in which $\mathbf{u}$ is block-additive. Since every $(k,a,b+1,\delta)$-fibre is a disjoint union of (k,a,b,δ)-fibres, the desired result is immediate if $b \leqslant r$.

Conversely, if $b \geqslant r$, consider some $(k,a,b+1,\delta)$-fibre $\mathscr{F}$, and let n be the smallest element of $\mathscr{F}$. Let $(x_i)_{i\geqslant 0} = \langle n\rangle_k$ be its decomposition in base k, and $\mathsf{c} = \mathsf{c}_k(n,n+\delta)$ be its carry distance with with $n+\delta$.

By minimality of n, we have $x_{\mathsf{c}+1} = x_{\mathsf{c}+2} = \cdots = x_{\mathsf{c}+b} = 0$. Consequently, $\mathscr{F}$ is the disjoint union of the sets $\mathscr{F}_{k,a,b,\delta}(n) + k^{\mathsf{c}+b}d$, where $0 \leqslant d < k$.

Now, consider some integer $m \in \mathscr{F}_{k,a,b,\delta}(n)$, some base-$k$ digit $d \in \Sigma_k$, and let $m' = m + k^{\mathsf{c}+b}d$. Let $(x_i)_{i\geqslant 0}$, $(y_i)_{i\geqslant 0}$, $(z_i)_{i\geqslant 0}$ and $(w_i)_{i\geqslant 0}$ be the decompositions of the integers m, $m+\delta$, m' and $m'+\delta$ in base k. By construction, we have $x_i = y_i = z_i = w_i$ for all $i \geqslant \mathsf{c}+1$, whereas $x_i = z_i$ and $y_i = w_i$ for all $i \leqslant \mathsf{c}+b-1$. Since $b \geqslant r$, it follows that

$$\begin{aligned} u_{m'+\delta} - u_{m'} &= \sum_{i\geqslant 0} f(w_i,w_{i+1},\ldots,w_{i+r-1}) - f(z_i,z_{i+1},\ldots,z_{i+r-1}) \\ &= \sum_{i=0}^{\mathsf{c}} f(w_i,w_{i+1},\ldots,w_{i+r-1}) - f(z_i,z_{i+1},\ldots,z_{i+r-1}) \\ &= \sum_{i=0}^{\mathsf{c}} f(y_i,y_{i+1},\ldots,y_{i+r-1}) - f(x_i,x_{i+1},\ldots,x_{i+r-1}) \\ &= u_{m+\delta} - u_m. \end{aligned}$$

In particular, for all $d \in \mathbb{Z}_k$, the set $\{m \in \mathscr{F}\colon u_{m+\delta} = u_m + d\}$ just consists in k copies of the set $\{m \in \mathscr{F}_{k,a,b,\delta}(n)\colon u_{m+\delta} = u_m + d\}$, shifted by $k^{\mathsf{c}+b}$. Consequently, if $\mathbf{u}$ is $(a,b+1)$-strongly 2-uncorrelated, it is also (a,b)-strongly 2-uncorrelated. □

It remains to verify whether $\mathbf{u}$ is (a,r)-strongly 2-uncorrelated. Below, we present this verification in detail when $(k,a,r)=(2,1,3)$. We will then describe this verification in general.

However, before we do so, let us slightly modify the representations of integers that we use. We note $\Sigma_{k\perp}$ the union $\Sigma_k \cup \{\perp\}$, and $\langle\langle n\rangle\rangle_k$ the unique sequence $(x_i)_{i\in\mathbb{Z}}$ with values in $\Sigma_{k\perp}$ such that

$$x_i = \perp \text{ for all } i \leqslant -1 \text{ and } n = \sum_{i\geqslant 0} x_i k^i.$$

In other words, the bi-infinite sequence $\langle\langle n\rangle\rangle_k$ consists in an infinity of terms $\perp$, followed by the usual representation $\langle n\rangle_k$ of n in base k. This bi-infinite sequence is called the *extended representation* of n in base k.

Then, we extend every function $f\colon \Sigma_k^r \to \mathbb{Z}_k$ to $\Sigma_{k\perp}^r$, by setting $f(\mathbf{x})=0$ whenever $\mathbf{x}$ has a coordinate equal to $\perp$. With these new notations, we simply have

$$u_n = \sum_{i\in\mathbb{Z}} f(x_i, x_{i+1}, \ldots, x_{i+r-1}).$$

Example 5.10. If $k=2$, $a=1$, $r=3$, and $f\colon \Sigma_2^3 \to \mathbb{Z}_2$ is a function such that $f(\mathbf{0})=0$, let us look at those constraints that f must satisfy to make its associated block-additive sequence $\mathbf{u}$ a $(1,3)$-strongly 2-uncorrelated sequence.

Let $\mathbf{x}=\langle\langle n\rangle\rangle_2$ and $\mathbf{y}=\langle\langle n+\delta\rangle\rangle_2$ be the extended representations of some integers n and $n+\delta$ in base 2, and let $\mathsf{c}=\mathsf{c}_2(n,n+\delta)$ be their carry distance. Then, let m be an element of the fibre $\mathscr{F}_{2,1,3,\delta}(n)$, and let $\mathbf{z}=\langle\langle m\rangle\rangle_2$ and $\mathbf{v}=\langle\langle m+\delta\rangle\rangle_2$. If we set

$$\begin{aligned}
\mathbf{X} &= \sum_{i\leqslant \mathsf{c}-4} f(x_i, x_{i+1}, x_{i+2})\\
\mathbf{Y} &= \sum_{i\leqslant \mathsf{c}-4} f(y_i, y_{i+1}, y_{i+2})\\
\mathbf{Z} &= \sum_{i\geqslant \mathsf{c}+1} f(z_i, z_{i+1}, z_{i+2})
\end{aligned}$$

and remembering that $x_\mathsf{c}=0$ and $y_\mathsf{c}=1$, we can observe that

$$\begin{aligned}
u_m = \mathbf{X}+\mathbf{Z} &+ f(x_{\mathsf{c}-3}, x_{\mathsf{c}-2}, z_{\mathsf{c}-1}) \\
&+ f(x_{\mathsf{c}-2}, z_{\mathsf{c}-1}, 0)\\
&+ f(z_{\mathsf{c}-1}, 0, z_{\mathsf{c}+1})\\
&+ f(0, z_{\mathsf{c}+1}, z_{\mathsf{c}+2})
\end{aligned}
\quad\text{and}\quad
\begin{aligned}
u_{m+\delta} = \mathbf{Y}+\mathbf{Z} &+ f(y_{\mathsf{c}-3}, y_{\mathsf{c}-2}, v_{\mathsf{c}-1})\\
&+ f(y_{\mathsf{c}-2}, v_{\mathsf{c}-1}, 1)\\
&+ f(v_{\mathsf{c}-1}, 1, z_{\mathsf{c}+1})\\
&+ f(1, z_{\mathsf{c}+1}, z_{\mathsf{c}+2}).
\end{aligned}$$

Consequently, we consider two cases:

- If $\mathsf{c}\geqslant 1$ and $x_{\mathsf{c}-1}=y_{\mathsf{c}-1}$, we have $z_{\mathsf{c}-1}=v_{\mathsf{c}-1}$. Thus, the difference

$$\begin{aligned}
u_{m+\delta}-u_m = \mathbf{Y}-\mathbf{X} &+ f(y_{\mathsf{c}-3}, y_{\mathsf{c}-2}, z_{\mathsf{c}-1}) - f(x_{\mathsf{c}-3}, x_{\mathsf{c}-2}, z_{\mathsf{c}-1})\\
&+ f(y_{\mathsf{c}-2}, z_{\mathsf{c}-1}, 1) \quad - f(x_{\mathsf{c}-2}, z_{\mathsf{c}-1}, 0)\\
&+ f(z_{\mathsf{c}-1}, 1, z_{\mathsf{c}+1}) \quad - f(z_{\mathsf{c}-1}, 0, z_{\mathsf{c}+1})\\
&+ f(1, z_{\mathsf{c}+1}, z_{\mathsf{c}+2}) \quad - f(0, z_{\mathsf{c}+1}, z_{\mathsf{c}+2})
\end{aligned}$$

  must take the value 0 four times and the value 1 four times when $z_{\mathsf{c}-1}$, $z_{\mathsf{c}+1}$ and $z_{\mathsf{c}+2}$ vary in $\mathbb{Z}_2$. Since the sums $\mathbf{X}$ and $\mathbf{Y}$ are independent of $z_{\mathsf{c}-1}$, $z_{\mathsf{c}+1}$ and $z_{\mathsf{c}+2}$, this amounts to demanding that the difference $(u_{m+\delta}-u_m)-(\mathbf{Y}-\mathbf{X})$ itself should take the values 0 and 1 equally often.

That family of differences depends only on the tuple $(x_{\mathsf{c}-3}, y_{\mathsf{c}-3}, x_{\mathsf{c}-2}, y_{\mathsf{c}-2})$, which can have 0, 2 or 4 coordinates $\perp$: it can be $(\perp,\perp,\perp,\perp)$, if $(n, n+\delta) = (0,2)$; or any tuple $(\perp,\perp,x,y)$ with $x, y \in \mathbb{Z}_2$, if $(n, n+\delta) = (x, 4+y)$; or any tuple (x,y,x',y') in $\mathbb{Z}_2^4$, if $(n, n+\delta) = (2^{\mathsf{c}-3}x + 2^{\mathsf{c}-2}x', 2^{\mathsf{c}-3}y + 2^{\mathsf{c}-2}y' + 2^{\mathsf{c}})$.

- If $\mathsf{c} = 0$ or $x_{\mathsf{c}-1} \neq y_{\mathsf{c}-1}$, we have $z_{\mathsf{c}-1} = x_{\mathsf{c}-1}$ and $v_{\mathsf{c}-1} = y_{\mathsf{c}-1}$. Thus, the difference

$$\begin{aligned} u_{m+\delta} - u_m = \mathbf{Y} - \mathbf{X} &+ f(y_{\mathsf{c}-3}, y_{\mathsf{c}-2}, y_{\mathsf{c}-1}) - f(x_{\mathsf{c}-3}, x_{\mathsf{c}-2}, x_{\mathsf{c}-1}) \\ &+ f(y_{\mathsf{c}-2}, y_{\mathsf{c}-1}, 1) \quad\;\; - f(x_{\mathsf{c}-2}, x_{\mathsf{c}-1}, 0) \\ &+ f(y_{\mathsf{c}-1}, 1, z_{\mathsf{c}+1}) \quad\;\; - f(x_{\mathsf{c}-1}, 0, z_{\mathsf{c}+1}) \\ &+ f(1, z_{\mathsf{c}+1}, z_{\mathsf{c}+2}) \quad\;\; - f(0, z_{\mathsf{c}+1}, z_{\mathsf{c}+2}) \end{aligned}$$

must take the value 0 twice and the value 1 twice when $z_{\mathsf{c}+1}$ and $z_{\mathsf{c}+2}$ vary in $\mathbb{Z}_2$; so must the difference $(u_{m+\delta} - u_m) - (\mathbf{Y} - \mathbf{X})$. The latter family of differences depends only on the tuple $(x_{\mathsf{c}-3}, y_{\mathsf{c}-3}, x_{\mathsf{c}-2}, y_{\mathsf{c}-2}, x_{\mathsf{c}-1}, y_{\mathsf{c}-1})$; as in the previous case, this tuple may be either $(\perp,\perp,\perp,\perp,\perp)$, or of the form $(\perp,\perp,\perp,\perp,x,y)$ with $x \neq y$, or $(\perp,\perp,x,y,x',y')$ with $x' \neq y'$, or any tuple (x,y,x',y',x'',y'') in $\mathbb{Z}_2^6$ with $x'' \neq y''$.

Therefore, overall, checking that $\mathbf{u}$ is $(1,3)$-strongly 2-uncorrelated amounts to computing finitely 21 families of 8 differences and 43 families of 4 differences, and checking that each family contains as many 0s as 1s.

Here is an intuition about how the previous example may be generalised. When considering the extended representations of two integers n and $n+\delta$ with carry distance c, we should just focus on a window of width $2r+a-1$ that consists of those digits in positions $\mathsf{c} - (a+r-1)$ to $\mathsf{c} + (r-1)$, and identify this window with a tuple in $\Sigma_{k\perp}^{2r+a-1}$. We shall then look at similar representations for the integers m and $m+\delta$ obtained when m varies in $\mathscr{F}_{k,a,r,\delta}(n)$; in each case, we will focus on the difference $(u_{m+\delta} - u_m) - (\mathbf{Y} - \mathbf{X})$, and demand that this difference takes every value in $\mathbb{Z}_k$ with the same frequency.

Formalising this intuition provides us with the following definition.

Definition 5.11. Let $k \geqslant 2$, $a \geqslant 0$ and $r \geqslant 1$ be integers, let $w = 2r+a-1$, and let $f\colon \Sigma_k^r \to \mathbb{Z}_k$ be a function. For each tuple $\mathbf{x} = (x_0, x_1, \ldots, x_{w-1})$ in $\Sigma_{k\perp}^w$, let $f(\mathbf{x})$ be the sum

$$\sum_{i=0}^{w-r} f(x_i, x_{i+1}, \ldots, x_{i+r-1}).$$

We say that a tuple $\mathbf{x} = (x_0, x_1, \ldots, x_{w-1})$ in $\Sigma_{k\perp}^w$ is *well-formed* if there exists an integer $t \leqslant w-r$ such that $x_i = \perp$ for all $i \leqslant t-1$ and $x_i \in \Sigma_k$ for all $i \geqslant t$; in that case, t is called the *weight* of $\mathbf{x}$. Then, we say that two well-formed tuples $\mathbf{x}$ and $\mathbf{y}$ in $\Sigma_{k\perp}^w$ are *related* if (i) they have the same weight; (ii) $x_{w-r} < y_{w-r}$; (iii) $x_i = y_i$ whenever $w-r < i < w$.

Given two related tuples $\mathbf{x}$ and $\mathbf{y}$ with weight t, we call *tuple fibre* of the pair $(\mathbf{x},\mathbf{y})$ the set $\mathscr{F}(\mathbf{x},\mathbf{y})$ that consists of those pairs $(\mathbf{z},\mathbf{v})$ of related tuples in Σ_k^w such that (i) $z_i = v_i = \perp$ for all $i \leqslant t-1$; (ii) $z_i = x_i$ and $v_i = y_i$ whenever $i \leqslant r-1$ or $x_i \neq y_i$; (iii) $z_i = v_i$ whenever $i \geqslant r$ and $x_i = y_i$.

Finally, the function f is called *a-strongly 2-uncorrelated* if, for every tuple fibre $\mathscr{F}$, the k sets given by $\{(\mathbf{z},\mathbf{v}) \in \mathscr{F}(\mathbf{x},\mathbf{y})\colon f(\mathbf{v}) = f(\mathbf{z}) + d\}$, where $d \in \mathbb{Z}_k$, have the same cardinalities.

Theorem 5.12. *Let $k \geqslant 2$, $a \geqslant 0$ and $r \geqslant 1$ be integers, let $f\colon \Sigma_k^r \to \mathbb{Z}_k$ be a function such that $f(\mathbf{0}) = 0$, and let $\mathbf{u}$ be the block-additive sequence associated with f. The sequence $\mathbf{u}$ is (a,r)-strongly 2-uncorrelated if and only if f is a-strongly 2-uncorrelated.*

Proof. Let $w = 2r+a-1$. Then, let $n \geqslant 0$ and $\delta \geqslant 1$ be integers, and let $(x_i)_{i\in\mathbb{Z}} = \langle\!\langle n \rangle\!\rangle_k$ and $(y_i)_{i\in\mathbb{Z}} = \langle\!\langle n+\delta \rangle\!\rangle_k$ be the extended representations of n and $n+\delta$ in base k. Also, let $\mathsf{c} = \mathsf{c}_k(n, n+\delta)$ be their carry distance, and let $\Omega = \{i \in \mathbb{Z}\colon \mathsf{c} - (a+r-1) \leqslant i \leqslant \mathsf{c} + (r-1)\}$. We map the pair $(n, n+\delta)$ to the pair of related tuples $(\mathbf{x},\mathbf{y})$ in $\Sigma_{k\perp}^w$ given by $\mathbf{x} = (x_i)_{i\in\Omega}$ and $\mathbf{y} = (y_i)_{i\in\Omega}$.

Let ϕ be this mapping, so that $\phi\colon (n, n+\delta) \to (\mathbf{x}, \mathbf{y})$. By construction, ϕ induces a bijection from the set $\{(m, m+\delta) \colon m \in \mathscr{F}_{k,a,r,\delta}(n)\}$ to $\mathscr{F}(\mathbf{x}, \mathbf{y})$. Moreover, if we set

$$\begin{aligned} \mathbf{X} &= \sum_{i \leqslant \mathsf{c}-(a+r)} f(x_i, x_{i+1}, \ldots, x_{i+r-1}), \\ \mathbf{Y} &= \sum_{i \leqslant \mathsf{c}-(a+r)} f(y_i, y_{i+1}, \ldots, y_{i+r-1}) \text{ and} \\ \mathbf{Z} &= \sum_{i \geqslant \mathsf{c}+1} f(x_i, x_{i+1}, \ldots, x_{i+r-1}), \end{aligned}$$

and if we note $(\mathbf{z}, \mathbf{v})$ the pair $\phi(m, m+\delta)$, we can observe, as in Example 5.10, that

$$(u_{m+\delta} - u_m) - (\mathbf{Y} - \mathbf{X}) = f(\mathbf{v}) - f(\mathbf{z}).$$

Consequently, the (k, a, r, δ)-fibre $\mathscr{F}_{k,a,r,\delta}(n)$ is δ-balanced for $\mathbf{u}$ if and only the k sets $\{(\mathbf{z}, \mathbf{v}) \in \mathscr{F}(\mathbf{x}, \mathbf{y}) \colon f(\mathbf{v}) = f(\mathbf{z}) + d\}$ where $d \in \mathbb{Z}_k$ have the same cardinalities.

It remains to prove that every tuple fibre $\mathscr{F}(\mathbf{x}, \mathbf{y})$ is the image of some set $\{(m, m+\delta) \colon m \in \mathscr{F}_{k,a,r,\delta}(n)\}$ by ϕ or, equivalently, that every pair of related tuples $(\mathbf{x}, \mathbf{y})$ in $\Sigma_{k\perp}^w$ is the image of some pair $(n, n+\delta)$ by ϕ. This last step is straightforward. Indeed, if t is the weight of both well-formed tuples $\mathbf{x} = (x_0, x_1, \ldots, x_{w-1})$ and $\mathbf{y} = (y_0, y_1, \ldots, y_{w-1})$, it suffices to choose

$$n = \sum_{i=t}^{w-1} k^{i-t} x_i \text{ and } n + \delta = \sum_{i=t}^{w-1} k^{i-t} y_i. \qquad \square$$

Finally, the usefulness of the notion of (a, r)-strong 2-correlation comes from the following partial classification result.

Theorem 5.13. *Let $k \geqslant 2$, $r \geqslant 1$ and $\ell \geqslant 2$ be integers, and let $\mathbf{u}$ be the block-additive sequence associated with a function $f\colon \Sigma_k^r \to \mathbb{Z}_k$ such that $f(\mathbf{0}) = 0$.*

1. *If $r = 1$, the sequence $\mathbf{u}$ is ℓ-correlated.*
2. *If $k = 2$, $2 \leqslant r \leqslant 5$ and $2 \leqslant \ell \leqslant 3$, the sequence $\mathbf{u}$ is ℓ-uncorrelated if and only if it is $(r-2, r)$-strongly 2-uncorrelated.*
3. *If $k = 2$, $2 \leqslant r \leqslant 5$ and $\ell \geqslant 4$, the sequence $\mathbf{u}$ is ℓ-correlated.*
4. *If $k = 3$ and $2 \leqslant r \leqslant 3$, the sequence $\mathbf{u}$ is 2-uncorrelated if and only if it is $(0, 2)$-strongly 2-uncorrelated.*
5. *If $k = 3$, $2 \leqslant r \leqslant 3$ and $\ell \geqslant 3$, the sequence $\mathbf{u}$ is ℓ-correlated.*

Proof. The proof of points 2 to 5 is computation-intensive, yet conceptually straightforward. It simply follows from an exhaustive search among the k^{k^r-1} possible functions f and associated sequences $\mathbf{u}$: in cases 2 and 4, Theorem 5.12 allows us to detect some $(r-2, r)$-strongly 2-uncorrelated sequences $\mathbf{u}$, which are thus 2- or even 3-uncorrelated; the other sequences treated in cases 2 and 4, and every sequence treated in cases 3 and 5, are proved to be ℓ-correlated by using Theorem 4.1 and Algorithm 1.

The code we used is available at [10].

https://github.com/VincentJuge1987/AutomaticSequences.

It uses Lemma 6.2 and similar acceleration techniques (*e.g.*, identifying as ℓ-correlated entire classes of functions whose differences are generated by the function g of Lemma 6.1, or by similar functions), thereby tackling cases 2 and 4 in 20 minutes, and cases 3 and 5 in 10 hours, on the authors' personal computers.

Thus, we focus on proving point 1. With that goal in mind, let us assume that $r = 1$. If $k = 2$ and $f(1) = 0$, we know that $u_n = 0$ for all $n \geqslant 0$, so that $\mathsf{freq}^{\mathbf{u}}_{(0,1)}(\{(0,0)\}) = 1$. If $k = 2$ and $f(1) = 1$, we know that $u_{n+1} - u_n = 1$ whenever $n \in 2\mathbb{N}$ or $n \in 8\mathbb{N}+3$, so that $\mathsf{freq}^{\mathbf{u}}_{(0,1)}((0,1)+\mathbb{Z}_2\mathbf{1}) \geqslant 5/8$. Thus, in both cases, $\mathbf{u}$ is 2-correlated.

Now, let us assume that $k \geqslant 3$. For all $s \leqslant k-2$, let $\Delta_f(s) = f(s+1) - f(s)$. If Δ_f coincides on two base-k digits s and t smaller than $k-1$, we can observe that $u_{n+1} - u_n \equiv \Delta_f(s) \pmod{k}$ whenever $n \equiv s$ or $t \pmod{k}$. Consequently, in that case, we have $\mathsf{freq}^{\mathbf{u}}_{(0,1)}((0,\Delta_f(s)) + \mathbb{Z}_k\mathbf{1}) \geqslant 2/k$, and $\mathbf{u}$ is 2-correlated.

If, however, Δ_f is injective, let d be the unique element of $\mathbb{Z}_k$ that does not belong to the range of Δ_f. There exists at most one base-k digit, say s_0, such that $s_0 < k-1$ and $\Delta_f(s_0) + f(0) - f(k-1) = d$. For every other base-$k$ digit $s < k-1$, and whenever $n \equiv sk + k - 1 \pmod{k^2}$, we still have $u_{n+1} - u_n \equiv \Delta_f(s) + f(0) - f(k-1) \not\equiv d \pmod{k}$. Thus, $\mathscr{D}^{\mathbf{u}}_{(0,1)}((0,d)+\mathbb{Z}_k\mathbf{1})$ is a subset of $k^2\mathbb{N} + \{s_0, k-1\}k + (k-1)$, and $\mathsf{freq}^{\mathbf{u}}_{(0,1)}((0,d)+\mathbb{Z}_k\mathbf{1}) \leqslant 2/k^2$, so that $\mathbf{u}$ is 2-correlated. ☐

6. Relations between uncorrelated sequences

Theorem 5.13 allows to compute the exhaustive list of binary functions of rank 3 for which the associated sequence is 3-uncorrelated. It turns out that there are 40 such functions, among the $2^7 = 128$ binary functions of rank 3 satisfying $f(0,0,0) = 0$. Thanks to the following general lemmas, we can, however, match the functions in groups of two, or even eight functions.

Lemma 6.1. *Let $k \geqslant 2$, $\ell \geqslant 2$ and $r \geqslant 2$ be integers. Let $g\colon \mathbb{Z}_k^r \to \mathbb{Z}_k$ be the function defined by*

$$g(a_1, \ldots, a_r) = a_1 - a_2,$$

and let $f\colon \mathbb{Z}_k^r \to \mathbb{Z}_k$ be a function such that $f(\mathbf{0}) = 0$. The sequence associated with f is ℓ-uncorrelated if and only if the sequence associated with $f+g$ is ℓ-uncorrelated.

Proof. Let $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ be the sequences associated with f, g and $f+g$. By definition of g, for every integer $n \geqslant 0$ with base-k expansion $(x_i)_{i\geqslant 0}$, we can observe that $v_n = \sum_{i\geqslant 0} x_i - x_{i+1} = x_0$. This means that $v_n \equiv n \pmod{k}$. Consequently, for every increasing tuple $\boldsymbol{\delta} \in \mathbb{N}^\ell$ and every integer $n \geqslant 0$,

$$w_{n+\boldsymbol{\delta}} = u_{n+\boldsymbol{\delta}} + v_{n+\boldsymbol{\delta}} = u_{n+\boldsymbol{\delta}} + n\,\mathbf{1} + \boldsymbol{\delta}.$$

In particular, for every tuple $\mathbf{s} \in \mathbb{Z}_k^\ell$, the term $w_{n+\boldsymbol{\delta}}$ belongs to $\mathbf{s} + \mathbb{Z}_k\mathbf{1}$ if and only if $u_{n+\boldsymbol{\delta}}$ belongs to $(\mathbf{s} - \boldsymbol{\delta}) + \mathbb{Z}_k\mathbf{1}$.

If $\mathbf{u}$ is ℓ-uncorrelated, Proposition 3.2 states that $\mathsf{freq}^{\mathbf{u}}_{\boldsymbol{\delta}}(\mathbf{s} + \mathbb{Z}_k\mathbf{1}) = k^{1-\ell}$. It follows that $\mathsf{freq}^{\mathbf{w}}_{\boldsymbol{\delta}}((\mathbf{s}+\boldsymbol{\delta}) + \mathbb{Z}_k\mathbf{1}) = k^{1-\ell}$. The tuple $\boldsymbol{\delta}$ being fixed, and by letting $\mathbf{s}$ vary in $\mathbb{Z}_k^\ell$, it follows that $\mathsf{freq}^{\mathbf{w}}_{\boldsymbol{\delta}}(\mathbf{t} + \mathbb{Z}_k\mathbf{1}) = k^{1-\ell}$ for all tuples $\mathbf{t} \in \mathbb{Z}_k^\ell$. The latter equality being valid for all initial tuples $\boldsymbol{\delta} \in \mathbb{N}^\ell$, Proposition 3.2 again proves that $\mathbf{w}$ is ℓ-uncorrelated.

Finally, if $\mathbf{w}$ is ℓ-uncorrelated, the same reasoning proves that the sequence $\mathbf{w}+\mathbf{v}$ is ℓ-uncorrelated. Repeating this argument, the sequences $\mathbf{w}+2\mathbf{v}, \mathbf{w}+3\mathbf{v}, \ldots$ are all ℓ-uncorrelated. Hence, the sequence $\mathbf{u} = \mathbf{w} + (k-1)\mathbf{v}$ is itself ℓ-uncorrelated. ☐

Lemma 6.2. *Let $k \geqslant 2$, $a \geqslant 0$ and $r \geqslant 2$ be integers. Consider some integer $i \leqslant r$ and some element q of $\mathbb{Z}_k$ and let $\mathsf{p}_{i\to q}\colon \mathbb{Z}_k^r \to \mathbb{Z}_k$ be the* selection-projection *function defined by*

$$\mathsf{p}_{i\to q}(a_1, \ldots, a_n) = \begin{cases} 1 & \text{if } a_i = q\ ; \\ 0 & \text{otherwise.} \end{cases}$$

Finally, let $f\colon \mathbb{Z}_k^n \to \mathbb{Z}_k$ be a function such that $f(\mathbf{0}) = 0$. If $q \neq 0$, the sequence associated with f is (a,r)-strongly 2-uncorrelated if and only if the sequence associated with $f + \mathsf{p}_{i\to q}$ is (a,r)-strongly 2-uncorrelated.

Proof. Let $\mathbf{u}$ and $\mathbf{v}$ be the sequences associated with f and $f + \mathsf{p}_{i\to q}$, and let $w = 2r + a - 1$. Let us assume that $\mathbf{u}$ is (a, r)-strongly 2-uncorrelated. Theorem 5.12 proves that f is a-strongly 2-uncorrelated, and we shall prove that so is $f + \mathsf{p}_{i\to q}$. To do that, let us extend $\mathsf{p}_{i\to q}$ to tuples in $\mathbb{Z}_{k\perp}^r$ by setting $\mathsf{p}_{i\to q}(\mathbf{x}) = 0$ whenever $\mathbf{x}$ has a coordinate equal to $\perp$.

For all pairs of related tuples $(\mathbf{x}, \mathbf{y}) = ((x_0, x_1, \ldots, x_{w-1}), (y_0, y_1, \ldots, y_{w-1}))$ and $(\mathbf{z}, \mathbf{v}) = ((z_0, z_1, \ldots, z_{w-1}), (v_0, v_1, \ldots, v_{w-1}))$ in $\Sigma_{k\perp}^w$ such that $(\mathbf{z}, \mathbf{v})$ belongs to the tuple fibre $\mathscr{F}(\mathbf{x}, \mathbf{y})$, note that $\mathsf{p}_{i\to q}(\mathbf{x})$ is equal to the number of coordinates x_j such that $i \leqslant j \leqslant w - r + i$ and $x_j = q$; the numbers $\mathsf{p}_{i\to q}(\mathbf{y})$, $\mathsf{p}_{i\to q}(\mathbf{z})$ and $\mathsf{p}_{i\to q}(\mathbf{v})$ can be expressed similarly. But then, for each such index j, either ($x_j = y_j$ and $z_j = w_j$) or ($x_j = z_j$ and $y_j = v_j$), so that the j^{th} coordinate does not contribute to the sum $(\mathsf{p}_{i\to q}(\mathbf{y}) - \mathsf{p}_{i\to q}(\mathbf{x})) - (\mathsf{p}_{i\to q}(\mathbf{v}) - \mathsf{p}_{i\to q}(\mathbf{z}))$.

Therefore, for each $d \in \mathbb{Z}_k$, the sets

$$\{(\mathbf{z}, \mathbf{v}) \in \mathscr{F}(\mathbf{x}, \mathbf{y}) \colon (f + \mathsf{p}_{i\to q})(\mathbf{v}) = (f + \mathsf{p}_{i\to q})(\mathbf{z}) + d\}$$

and

$$\{(\mathbf{z}, \mathbf{v}) \in \mathscr{F}(\mathbf{x}, \mathbf{y}) \colon f(\mathbf{v}) = f(\mathbf{z}) + d + \mathsf{p}_{i\to q}(\mathbf{x}) - \mathsf{p}_{i\to q}(\mathbf{t})\}$$

coincide with each other. Thus, the k sets obtained when $d \in \mathbb{Z}_k$ have the same cardinalities. This means that $f + \mathsf{p}_{i\to q}$ is a-strongly 2-uncorrelated, *i.e.*, that $\mathbf{v}$ is (a, r)-strongly 2-uncorrelated.

Conversely, if $\mathbf{v}$ is (a, r)-strongly 2-uncorrelated, so are the sequences associated with each function of the form $f + \lambda\mathsf{p}_{i\to q}$ when $\lambda \geqslant 1$; choosing $\lambda = k$ provides us with the function f itself, which completes the proof. ☐

Using the above observations, we can group the 3-uncorrelated binary functions of rank 3 in only 5 different classes of functions, each of cardinality 8. A representative from each class is given in the table below: every binary function of rank 3 associated with a 3-uncorrelated sequence can be written as

$$f + \epsilon_1\mathsf{p}_{1\to 1} + \epsilon_2\mathsf{p}_{2\to 1} + \epsilon_3\mathsf{p}_{3\to 1},$$

where f is one of the five functions from the table below, and $\epsilon_1, \epsilon_2, \epsilon_3 \in \mathbb{Z}_2$.

	Expression of $f(x_1, x_2, x_3)$	Patterns counted by f
1	x_1x_2	1 1 •
2	x_1x_3	1 • 1
3	$x_1(x_2 + x_3)$	1 0 1 and 1 1 0
4	$(x_1 + x_2)x_3$	0 1 1 and 1 0 1
5	$(x_1 + x_2)(x_2 + x_3)$	0 1 0 and 1 0 1

Note that the first function is in fact a function of rank 2, and the associated sequence is the Golay-Shapiro sequence, as mentioned in Example 3.4.

Similarly, we can group the 2-uncorrelated ternary functions of rank 2 in one single class of functions, of cardinality 2×3^4. These are the functions of the form

$$\epsilon_0 f + \epsilon_1\mathsf{p}_{1\to 1} + \epsilon_2\mathsf{p}_{1\to 2} + \epsilon_3\mathsf{p}_{2\to 1} + \epsilon_4\mathsf{p}_{2\to 2},$$

where $f(x_1, x_2) = x_1x_2$ is associated with the generalised Golay-Shapiro sequence, ϵ_0 is a non-zero element of $\mathbb{Z}_3$, and $\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4 \in \mathbb{Z}_3$.

7. Discussion and questions

Theorem 5.13 provides a complete picture of the correlation properties of binary block-additive sequences of rank $r \leqslant 5$, and ternary sequences of rank $r \leqslant 3$. As a consequence of Theorem 3.3, we have proven the existence of binary block-additive sequences that are 3-uncorrelated, but we have not found any 4-uncorrelated block-additive sequences, so that the existence of an ℓ-uncorrelated sequence for $\ell \geqslant 4$ remains an open question. An interesting approach would be to extend the notion of strongly ℓ-uncorrelated sequences to integers $\ell \geqslant 3$ and to see if the sufficient criterion thus obtained could provide some examples. Observe also that our exhaustive search has shown that the reciprocal of Proposition 5.8 is true for binary and ternary sequences of small rank, that is to say, all the 2-uncorrelated sequences of small rank ($r \leqslant 5$ for $k = 2$, $r \leqslant 3$ for $k = 3$) are (a, b)-strongly 2-uncorrelated for some integers $a \geqslant 0$ and $b \geqslant 0$. We believe that this property holds more generally but it remains to be proven.

Another possible extension of our work would be to study block-additive sequences of dimension greater than or equal to 2. In the same vein as the results on multi-dimensional sequences presented in [3], we can expect that all our results will adapt to the higher dimension without particular difficulty.

Data Availability Statement. The research data/code associated with this article are available in Zenodo, under the reference doi:10.5281/zenodo.10792181.

References

[1] J.-P. Allouche and J. Shallit, Automatic sequences. Theory, Applications, Generalizations. Cambridge University Press, Cambridge (2003).

[2] E. Grant, J. Shallit and T. Stoll, Bounds for the discrete correlation of infinite sequences on k symbols and generalized Rudin–Shapiro sequences. *Acta Arith.* **140** (2009) 345–368.

[3] I. Marcovici, T. Stoll and P.-A. Tahay, Discrete correlations of order 2 of generalized Golay–Shapiro sequences: a combinatorial approach. *Integers* **21** (2021) #A45.

[4] P.-A. Tahay, Discrete correlation of order 2 of generalized Rudin–Shapiro sequences on alphabets of arbitrary size. *Unif. Distrib. Theory* **15** (2020) 1–26.

[5] M. Drmota, P.J. Grabner and P. Liardet, Block additive functions on the Gaussian integers. *Acta Arith.* **135** (2008) 299–332.

[6] C. Müllner, The Rudin–Shapiro sequence and similar sequences are normal along squares. *Canad. J. Math.* **70** (2018) 1096–1129.

[7] E. Cateland, *Digital Sequences and k-regular Sequences.* Theses, Université Sciences et Technologies – Bordeaux I (1992).

[8] J.-P. Allouche, On a Golay–Shapiro-like sequence. *Unif. Distrib. Theory* **11** (2016) 205–210.

[9] M. Queffélec, Une nouvelle propriété des suites de Rudin–Shapiro. *Ann. Inst. Fourier (Grenoble)* **37** (1987) 115–138.

[10] V. Jugé, Detecting correlations in automatic sequences, (2024), 1, 10.5281/zenodo.10792181