



OPEN

## On complexity of colloid cellular automata

Andrew Adamatzky<sup>1,4✉</sup>, Nic Roberts<sup>2</sup>, Raphael Fortulan<sup>1</sup>, Noushin Raeisi Kheirabadi<sup>1</sup>, Panagiotis Mougkogiannis<sup>1</sup>, Michail-Antisthenis Tsompanas<sup>1</sup>, Genaro J. Martínez<sup>3</sup>, Georgios Ch. Sirakoulis<sup>1,4</sup> & Alessandro Chiolerio<sup>1,5</sup>

The colloid cellular automata do not imitate the physical structure of colloids but are governed by logical functions derived from them. We analyze the space-time complexity of Boolean circuits derived from the electrical responses of colloids—specifically ZnO (zinc oxide, an inorganic compound also known as calamine or zinc white, which naturally occurs as the mineral zincite), proteinoids (microspheres and crystals of thermal abiotic proteins), and their combinations in response to electrical stimulation. To extract Boolean circuits from colloids, we send all possible configurations of two-, four-, and eight-bit binary strings, encoded as electrical potential values, to the colloids, record their responses, and infer the Boolean functions they implement. We map the discovered functions onto the cell-state transition rules of cellular automata—arrays of binary state machines that update their states synchronously according to the same rule—creating the colloid cellular automata. We then analyze the phenomenology of the space-time configurations of the automata and evaluate their complexity using measures such as compressibility, Shannon entropy, Simpson diversity, and expressivity. A hierarchy of phenomenological and measurable space-time complexity is constructed.

**Keywords** Cellular automata, Unconventional computing, Colloids, Liquid computers

A liquid computer is a device that uses incompressible fluid to process information via mechanical, electrical, optical, or chemical means. The implementation of computation in liquid media has a history spanning over 120 years, from hydraulic algebraic machines developed in the 1900s to fluid maze solvers and droplet logics in the late 2000s. For an overview, please see<sup>1</sup>. Advantages of liquid computing include reconfigurability and flexibility, scalability, potential for reduced energy consumption, bio-compatibility and integration with biological systems, intrinsic parallelism, innovative data storage and retrieval, and novel computation paradigms. While liquid-based computers are still largely experimental and face several technical challenges, they offer intriguing advantages that could revolutionise various fields of computing and technology<sup>2–5</sup>.

Recently a new sub-domain of liquid computing emerged—computing with colloids (mixtures where microscopically dispersed insoluble particles liquids). The rise of colloid computers started from the liquid cybernetic systems, conceptualised as colloidal autonomous soft holonomic processors have demonstrated intriguing features, including autolographic capabilities<sup>2,3</sup>. Our previous experiments with ZnO colloids under controlled laboratory conditions demonstrated their potential as electrical analog neurons, successfully implementing synaptic-like learning and Pavlovian reflexes<sup>2,5</sup>. Additionally, the computational capabilities of Fe<sub>3</sub>O<sub>4</sub> ferrofluid for digit recognition further exemplify the versatility of liquid-based systems<sup>4</sup>.

One of the key developments in colloid computing became mining of Boolean circuits in colloids<sup>6,7</sup>. The technique is based on selecting a pair of input sites, applying all possible combinations of inputs, where logical values are represented by electrical characteristics of input signals, to the sites and recording outputs, represented by electrical responses of the substrate, on a set of the selected output sites. The approach belongs to the family of reservoir computing<sup>8–12</sup> and *in materia* computing<sup>13–17</sup> techniques of analysing computational properties of physical and biological substrates.

In our experimental laboratory studies<sup>6,7</sup> we discovered a range of 4-, 6- and 8-ary Boolean functions. In present paper, we evaluate dynamics and complexity of the functions using one-dimensional cellular automata (CA). CA, despite their simple rules and structure, can exhibit complex behaviour. This makes them an excellent tool for evaluating the inherent complexity of *n*-ary Boolean functions by mapping the functions onto the

<sup>1</sup>Unconventional Computing Laboratory, UWE, Bristol, UK. <sup>2</sup>Department of Engineering and Technology, University of Huddersfield, Huddersfield, UK. <sup>3</sup>Escuela Superior de Cómputo, Instituto Politécnico Nacional, Mexico City, Mexico. <sup>4</sup>Democritus University of Thrace, DUTH University Campus, 67100 Xanthi, Greece. <sup>5</sup>Bioinspired Soft Robotics, Istituto Italiano di Tecnologia, Via Morego 30, 16163 Genoa, Italy. ✉email: Andrew.Adamatzky@uwe.ac.uk

CA rules and observing the resulting dynamics. CA can generate a variety of patterns based on initial states and transition rules. By encoding  $n$ -Boolean functions into CA rules, we study the patterns that emerge, providing a visual and dynamic representation of the function's complexity. This is particularly useful for understanding how simple functions can lead to complex behaviours and vice versa.

## Methods

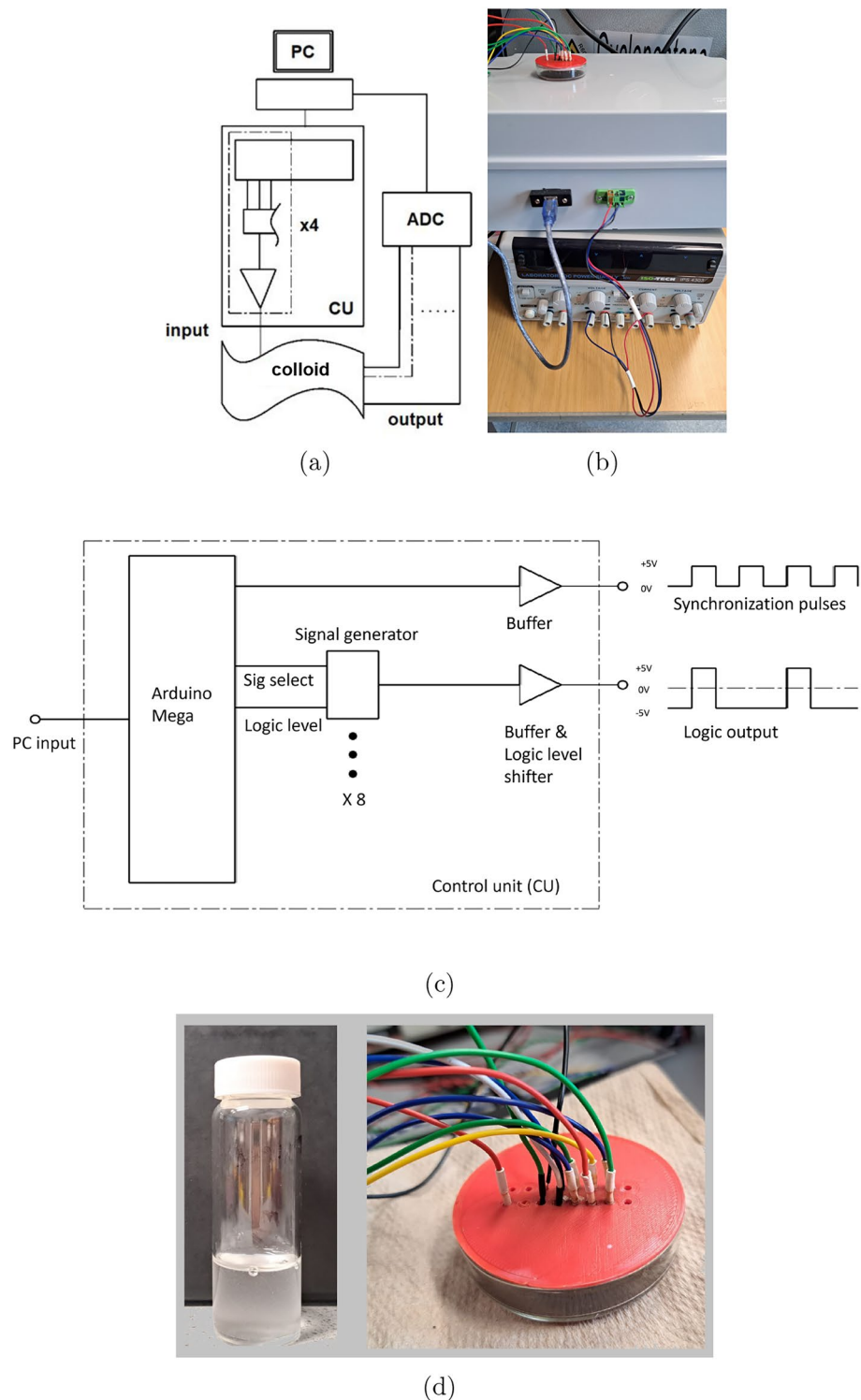
Experimental techniques on mining Boolean functions are described in full details in<sup>6,7</sup>. Here we briefly outline an overall approach, based on the example of<sup>7</sup>. Colloids of ZnO and proteinoids have been prepared as detailed in<sup>6,7</sup>. The hardware was built around an Arduino Mega 2560 (Elegoo, China) and a series of AD9833 programmable signal generators (Analog, USA). This setup can send sequences of 2, 4, and 8-bit strings to the colloid sample, with the strings encoded as step voltage inputs:  $-5$  V representing a logical '0' and  $5$  V representing a logical '1'. In Fig. 1a, a PC programs a Control Unit (CU) and receives readings from an analog-to-digital converter (ADC). The CU, shown as a grey box connected to a standard laboratory power supply in Fig. 1b, contains the Arduino Mega and multiple amplifiers. To generate the 2, 4, and 8-bit strings without redesigning or rewiring the CU, multiple programmable signal generators were incorporated. This is abstracted in Fig. 1c, where only one generator and its output are depicted for simplicity. Activation of these generators is controlled by the Arduino Mega, which is programmed through the PC and also depicted within the CU entity in Fig. 1c. To search for 2-, 4-, and 8-input Boolean circuits, we used respective electrodes. These were  $10\ \mu\text{m}$  platinum rods inserted  $5$  mm apart into the colloid container. Data acquisition (DAQ) probes, separated by  $5$  mm, fed 2 differential outputs to a Pico 24 ADC. Its 3rd channel received a pulse on each input state change. Refer to Fig. 1 for the apparatus schematic. The strings counted from binary 00 to 11, 0000 to 1111, or 00000000 to 11111111, changing state every  $15$  s. All possible electrode states were tested. For two bits, states sequentially altered every  $15$  s between 00, 01, 10, and 11. Similarly, all states of the four- and eight-bit strings were sequentially applied. Samples from 2 channels were taken at  $1$  Hz throughout the experiment. Peaks for each channel were located for 10 thresholds, from  $100$  to  $600$  mV in  $50$  mV steps, for each input state from 0000 to 1111. Most commonly found Boolean functions extracted from ZnO nanoparticle are listed in Table 1. Boolean functions derived in<sup>7</sup> are presented in Table 1(abc), and the functions derived in<sup>6</sup> in Table 1(def). Most frequent Boolean functions discovered in proteinoid colloids are shown in Table 2(abc) and mixture of ZnO and proteinoids in Table 2(def).

We evaluate complexity of the functions discovered via complexity of the space-time configurations of one-dimensional cellular automata (CA) governed by these functions. We call these CA 'colloid cellular automata' because their space-time evolution is governed by Boolean functions implemented by colloids and their mixtures in laboratory experiments. We consider an array  $Z$  of finite state automata, called cells, where every cell takes states '0' or '1' and updates its state depending on the states of its two, four or eight immediate neighbours. All cells update their states by the same rule and in discrete time. For example, a cell with index  $i$ ,  $x_i \in Z$ , updates its state at time  $t$  as a function of states of its two neighbours  $x^{t+1} = f(x_{i-1}^t, x_{i+1}^t)$  (representing variables  $A$  and  $B$  in Tables 1ad and 2ad), four neighbours:  $x^{t+1} = f(x_{i-2}^t, x_{i-1}^t, x_{i+1}^t, x_{i+2}^t)$  (representing variables  $A \dots C$  in Tables 1be and 2be), or eight neighbours  $x^{t+1} = f(x_{i-4}^t, x_{i-3}^t, x_{i-2}^t, x_{i-1}^t, x_{i+1}^t, x_{i+2}^t, x_{i+3}^t, x_{i+4}^t)$  (representing variables  $A \dots H$  in Tables 1cf and 2cf). For example the function  $f_{25} = (A \cdot B \cdot D \cdot \overline{C}) + (C \cdot D \cdot \overline{A} \cdot \overline{B})$  is represented in a cell-state transition rule as  $x^{t+1} = (x_{i-2}^t \cdot x_{i-1}^t \cdot x_{i+2}^t \cdot x_{i+1}^t) + (x_{i+1}^t \cdot x_{i+2}^t \cdot x_{i-2}^t \cdot x_{i-1}^t)$ . We evolved automata of 500 cells in 500 iterations of simultaneous cell-state transition.

To evaluate complexity of the cellular automata we used Shannon entropy<sup>19–21</sup>, Simpson's diversity (commonly used in ecological studies to evaluate biodiversity of populations<sup>22–24</sup>), Lempel-Ziv complexity<sup>25</sup>, space filling and expressiveness<sup>26,27</sup>. Let matrix  $L$  represent time-space configuration of a 1D CA governed by state-transition rules derived from colloids. Let  $W = \{0, 1\}^9$  be a set of all possible configurations of a 9-node neighbourhood  $B_x$  including the central node  $x$ . Let  $B$  be a configuration of matrix  $L$ , we calculate a number of non-quietest neighbourhood configurations as  $\eta = \sum_{x \in L} \epsilon(x)$ , where  $\epsilon(x) = 0$  if for every resting  $x$  all its neighbours are in the state '0', and  $\epsilon(x) = 1$  otherwise. The Shannon entropy  $H$  is calculated as  $H = - \sum_{w \in W} (v(w)/\eta \cdot \ln(v(w)/\eta))$ , where  $v(w)$  is a number of times the neighbourhood configuration  $w$  is found in configuration  $B$ . Simpson's diversity  $S$  is calculated as  $S = \sum_{w \in W} (v(w)/\eta)^2$ . Simpson diversity linearly correlates with Shannon entropy for  $H < 3$ ; relationships becomes logarithmic for higher values of  $H$  as we previously demonstrated in<sup>28</sup>. The assessment of Lempel-Ziv complexity (compressibility), denoted as  $LZ$ , is based on the size of space-time configurations saved as PNG files representing configurations. This approach suffices because the 'deflation' algorithm utilised in PNG lossless compression, as outlined in<sup>29–31</sup>, is a derivative of the classical Lempel–Ziv 1977 algorithm, as described in<sup>25</sup>. Space filling  $D$  is a ratio of non-zero entries in  $B$  to the total number of cells/nodes. This is used to estimate expressiveness. Expressiveness  $E$  is calculated as the Shannon entropy  $H$  divided by space-filling ratio  $D$ , the expressiveness reflects the 'economy of diversity'.

## Results

CA presented by majority of functions from Tabs. 1 and 2 evolve to all-0 or all-1 state, an example of evolution to all-0 states is shown in Fig. 2a. These are 'trivial' functions. Let us consider the positions of the functions within Wolfram's classification of CA behaviour<sup>32</sup>. Most functions discovered belong to Class I, which is characterised by automata exhibiting simple dynamics and evolving to a stable state where all cells are in the same state. Functions  $f_1$ ,  $f_{10}$  (Fig. 2b),  $f_{39}$ ,  $f_{28}$ ,  $f_{35}$  (Fig. 2g),  $f_{40}$  (Fig. 2h),  $f_{12}$  (Fig. 3a),  $f_{14}$  (Fig. 3b),  $f_{19}$  (Fig. 3c),  $f_{20}$  (Fig. 3d),  $f_{37}$  (Fig. 4c),  $f_3$ ,  $f_6$ ,  $f_{38}$  (Fig. 2c),  $f_4$ ,  $f_9$ ,  $f_{13}$ ,  $f_{15}$ ,  $f_{17}$ ,  $f_{18}$  (Fig. 2d) and the function  $f_{37}$  (Fig. 4c) belong to the class II: the automata fall into global cells do not update their state or update them cyclically from '0' to '1'. Space-time dynamics of class III CA is by quasi-dynamics behaviour and difficult predictability of the successions of the global states. The following functions can be related to the class III CA:  $f_7$  (Fig. 2e),  $f_8$  (Fig. 2f),  $f_{21}$  (Fig. 4a),  $f_{36}$  and  $f_{45}$



**Fig. 1.** (a) A scheme of the experiments. PC—laptop for generating sequences; CU—control unit, the dashed section is a breakdown of a single channel; ADC—analogue to digital converter<sup>18</sup>. (b) experimental setup. (c) A schematic of the inside of the unit control box. (d) A close-up photo of the colloid dish and the electrodes interfacing it. From<sup>7</sup>.

(Fig. 4b). No functions from those discovered in laboratory experiments seem to belong to class IV, where the space-time dynamics of automata show gliders (compact patterns translating in space) with non-trivial

(a)	
<i>f</i>	$\phi$
$f_1 = \bar{A} + \bar{B}$	73
$f_2 = A + B$	45
$f_3 = \bar{A} + B$	37
$f_4 = A + \bar{B}$	33
$f_5 = A \cdot B$	8
$f_6 = B \cdot \bar{A}$	6
$f_7 = (A \cdot \bar{B}) + (B \cdot \bar{A})$	4
$f_8 = (A \cdot B) + (\bar{A} \cdot \bar{B})$	3
$f_9 = A \cdot \bar{B}$	3
$f_{10} = \bar{A} \cdot \bar{B}$	2
(b)	
<i>f</i>	$\phi$
$f_{11} = (A \cdot \bar{B}) + (B \cdot \bar{A} \cdot \bar{C}) + (B \cdot \bar{C} \cdot \bar{D})$	7
$f_{12} = (C \cdot D \cdot \bar{B}) + (A \cdot \bar{B} \cdot \bar{D}) + (B \cdot \bar{A} \cdot \bar{D}) + (D \cdot \bar{A} \cdot \bar{C})$	6
$f_{13} = (A \cdot \bar{B} \cdot \bar{D}) + (B \cdot \bar{A} \cdot \bar{C} \cdot \bar{D})$	6
$f_{14} = (\bar{A} \cdot \bar{D}) + (A \cdot B \cdot C \cdot D) + (B \cdot \bar{A} \cdot \bar{C}) + (C \cdot \bar{A} \cdot \bar{B})$	5
$f_{15} = (A \cdot \bar{B} \cdot \bar{D}) + (B \cdot \bar{A} \cdot \bar{C}) + (B \cdot \bar{C} \cdot \bar{D})$	5
$f_{16} = A \cdot D \cdot \bar{B} \cdot \bar{C}$	5
$f_{17} = A \cdot \bar{B} \cdot \bar{C} \cdot \bar{D}$	5
$f_{18} = (B \cdot C \cdot D) + (B \cdot C \cdot \bar{A}) + (C \cdot D \cdot \bar{A}) + (A \cdot \bar{B} \cdot \bar{C} \cdot \bar{D})$	5
$f_{19} = (A \cdot D \cdot \bar{B}) + (B \cdot D \cdot \bar{A}) + (A \cdot \bar{B} \cdot \bar{C}) + (B \cdot \bar{A} \cdot \bar{C}) + (D \cdot \bar{A} \cdot \bar{C})$	5
$f_{20} = (D \cdot \bar{A}) + (D \cdot \bar{B}) + (B \cdot \bar{A} \cdot \bar{C})$	5
(c)	
$f_{21} = (A \cdot B \cdot F \cdot \bar{C} \cdot \bar{E}) + (A \cdot D \cdot F \cdot \bar{C} \cdot \bar{E}) + (A \cdot G \cdot H \cdot \bar{B} \cdot \bar{C}) + (B \cdot D \cdot E \cdot \bar{A} \cdot \bar{F}) + (B \cdot E \cdot H \cdot \bar{A} \cdot \bar{C}) + (C \cdot D \cdot E \cdot \bar{B} \cdot \bar{F}) + (D \cdot E \cdot H \cdot \bar{B} \cdot \bar{C}) + (D \cdot F \cdot H \cdot \bar{A} \cdot \bar{B}) + (B \cdot C \cdot D \cdot F \cdot G \cdot \bar{H}) + (B \cdot C \cdot D \cdot G \cdot H \cdot \bar{F}) + (B \cdot E \cdot \bar{A} \cdot \bar{C} \cdot \bar{H}) + (C \cdot F \cdot \bar{B} \cdot \bar{C} \cdot \bar{H}) + (E \cdot H \cdot \bar{A} \cdot \bar{C} \cdot \bar{F}) + (F \cdot H \cdot \bar{B} \cdot \bar{D} \cdot \bar{E}) + (A \cdot C \cdot E \cdot G \cdot \bar{B} \cdot \bar{D}) + (A \cdot E \cdot F \cdot G \cdot \bar{C} \cdot \bar{D}) + (B \cdot D \cdot E \cdot G \cdot \bar{C} \cdot \bar{F}) + (B \cdot E \cdot F \cdot G \cdot \bar{A} \cdot \bar{D}) + (C \cdot F \cdot G \cdot H \cdot \bar{D} \cdot \bar{E}) + (A \cdot C \cdot D \cdot E \cdot F \cdot H \cdot \bar{G}) + (B \cdot \bar{C} \cdot \bar{E} \cdot \bar{G} \cdot \bar{H}) + (C \cdot \bar{B} \cdot \bar{D} \cdot \bar{F} \cdot \bar{G}) + (D \cdot \bar{C} \cdot \bar{E} \cdot \bar{F} \cdot \bar{H}) + (E \cdot \bar{A} \cdot \bar{B} \cdot \bar{D} \cdot \bar{H}) + (E \cdot \bar{B} \cdot \bar{D} \cdot \bar{G} \cdot \bar{H}) + (B \cdot C \cdot D \cdot \bar{A} \cdot \bar{E} \cdot \bar{G}) + (B \cdot D \cdot F \cdot \bar{C} \cdot \bar{G} \cdot \bar{H}) + (B \cdot D \cdot G \cdot \bar{A} \cdot \bar{C} \cdot \bar{E}) + (B \cdot E \cdot H \cdot \bar{C} \cdot \bar{F} \cdot \bar{G}) + (C \cdot D \cdot G \cdot \bar{B} \cdot \bar{E} \cdot \bar{H}) + (C \cdot E \cdot F \cdot \bar{D} \cdot \bar{C} \cdot \bar{H}) + (C \cdot G \cdot H \cdot \bar{A} \cdot \bar{E} \cdot \bar{F}) + (D \cdot E \cdot F \cdot \bar{A} \cdot \bar{B} \cdot \bar{C}) + (B \cdot D \cdot \bar{E} \cdot \bar{F} \cdot \bar{G} \cdot \bar{H}) + (B \cdot F \cdot \bar{A} \cdot \bar{D} \cdot \bar{E} \cdot \bar{G}) + (C \cdot D \cdot \bar{A} \cdot \bar{B} \cdot \bar{E} \cdot \bar{H}) + (C \cdot H \cdot \bar{D} \cdot \bar{E} \cdot \bar{F} \cdot \bar{G}) + (A \cdot C \cdot E \cdot G \cdot \bar{D} \cdot \bar{F} \cdot \bar{H}) + (A \cdot \bar{B} \cdot \bar{C} \cdot F \cdot \bar{G} \cdot \bar{H}) + (D \cdot \bar{A} \cdot \bar{B} \cdot \bar{C} \cdot \bar{E} \cdot \bar{G}) + (G \cdot \bar{A} \cdot \bar{C} \cdot \bar{D} \cdot \bar{E} \cdot \bar{H})$	
(d)	
Function	$\phi$
$f_{22} = A + B$	35
$f_{23} = A \cdot B$	3
(e)	
<i>f</i>	$\phi$
$f_{24} = A + B + C + D$	16
$f_{25} = (A \cdot B \cdot D \cdot \bar{C}) + (C \cdot D \cdot \bar{A} \cdot \bar{B})$	6
$f_{26} = (C \cdot D) + (A \cdot B \cdot D)$	4
$f_{27} = A + B + D$	2
(f)	
<i>f</i>	$\phi$
$f_{28} = \bar{A} + \bar{B} + \bar{C} + \bar{D} + \bar{E} + \bar{F} + \bar{G}$	4
$f_{29} = A \cdot C \cdot D \cdot E \cdot F \cdot G \cdot H \cdot \bar{B}$	2
$f_{30} = A + C + D + E + F + H + (B \cdot \bar{G}) + (G \cdot \bar{B})$	1
$f_{31} = C + D + E + F + H + (A \cdot B) + (A \cdot \bar{G}) + (B \cdot \bar{G}) + (G \cdot \bar{A} \cdot \bar{B})$	1

**Table 1.** Most commonly found Boolean functions,  $\phi$  is a frequency of the functions’ discovery, extracted from ZnO nanoparticle. Boolean functions derived in<sup>7</sup> are shown in (abc), and the functions derived in<sup>6</sup> in (def) (ad) Two-inputs, (be)Four-inputs. (cf) Eight inputs.

interactions between the gliders. CA governed by functions presented in Fig. 2c,d demonstrate travelling compact patterns however these patterns emerge due to asymmetry of the functions.

Complexity measures of the functions discussed are shown in Table 3. Complexity measures — Shannon entropy, Simpson index and expressiveness—are consistent with each other as seen in scatter plots for Shannon entropy *H* vs. expressiveness *E* (Fig. 5a), Shannon entropy *H* vs. Simpson index *S* (Fig. 5b). Person correlation coefficients  $r(H, E) = 0.57$ , coefficient of determination  $R^2 = 0.3234$ , shows moderate positive linear correlation

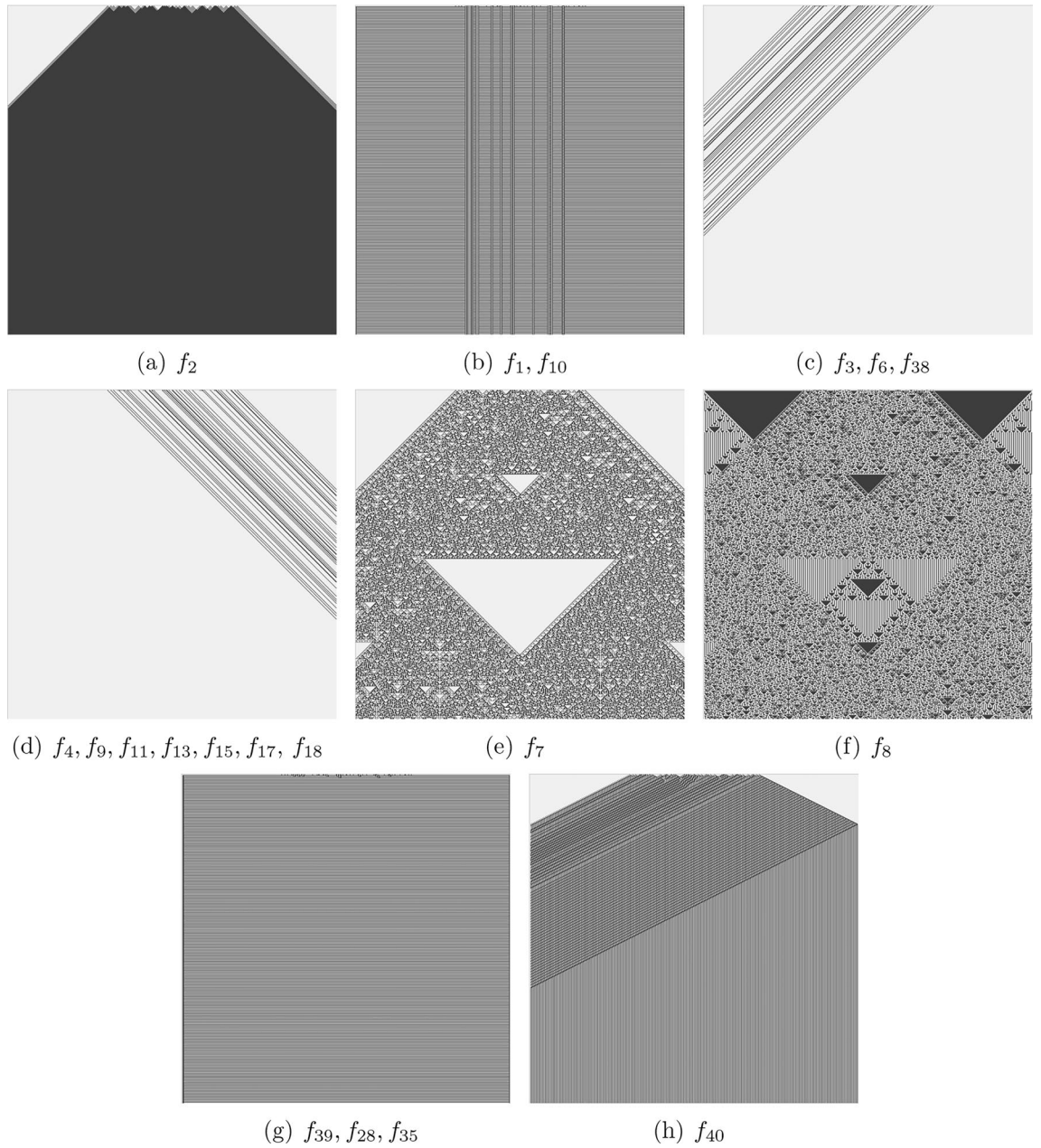
(a)	
$f$	$\phi$
0 (False)	19
$f_{22} = A + B$	16
$f_{23} = A \cdot B$	3
(b)	
$f$	$\phi$
$f_{24} = A + B + C + D$	23
$f_{32} = A \cdot B \cdot C \cdot D$	4
$f_{33} = A + B + (C \cdot D)$	3
$f_{34} = (A \cdot B) + (B \cdot D) + (C \cdot D) + (A \cdot \overline{C} \cdot \overline{D})$	3
(c)	
$f$	Count
$f_{35} = \overline{A} + \overline{B} + \overline{C} + \overline{D} + \overline{E} + \overline{F} + \overline{G} + \overline{H}$	19
$f_{36} = (A \cdot \overline{E}) + (B \cdot \overline{H}) + (C \cdot \overline{G}) + (D \cdot \overline{F}) + (E \cdot \overline{D}) + (F \cdot \overline{C}) + (G \cdot \overline{B}) + (H \cdot \overline{A})$	4
$f_{37} = A \cdot B \cdot C \cdot D \cdot E \cdot F \cdot H \cdot \overline{G}$	2
$(C \cdot \overline{B}) + (C \cdot \overline{D}) + (D \cdot \overline{E}) + (E \cdot \overline{G}) + (F \cdot \overline{H}) + (G \cdot \overline{F}) \vee$	
$f_{38} = (H \cdot \overline{E}) + (A \cdot B \cdot \overline{C}) + (A \cdot H \cdot \overline{B}) + (B \cdot H \cdot \overline{C})$	1
(d)	
$f$	$\phi$
$f_{23} = A \cdot B$	21
$f_{22} = A + B$	14
$B$	2
$A$	1
(e)	
$f$	$\phi$
$f_{39} = \overline{A} + \overline{B} + \overline{C}$	15
$f_{40} = (A \cdot \overline{B}) + (B \cdot \overline{C}) + (D \cdot \overline{A})$	7
$f_{41} = A \cdot B \cdot D \cdot \overline{C}$	5
$f_{42} = (A \cdot B \cdot \overline{C}) + (A \cdot D \cdot \overline{B}) + (B \cdot D \cdot \overline{C})$	4
(f)	
$f$	$\phi$
$f_{43} = A + B + C + D + E + F + G + H$	6
$f_{44} = A \cdot B \cdot C \cdot D \cdot E \cdot F \cdot G \cdot \overline{H}$	3
$f_{45} = (A \cdot \overline{D}) + (B \cdot \overline{G}) + (C \cdot \overline{F}) + (D \cdot \overline{E}) + (E \cdot \overline{C}) + (F \cdot \overline{B}) + (G \cdot \overline{A}) + (G \cdot \overline{H})$	2
$f_{46} = (A + C + D + E + (B \cdot F) + (B \cdot G) + (B \cdot H) + (F \cdot G) + (F \cdot H) + (G \cdot H)$	1

**Table 2.** The four most common extracted sum-of-products Boolean expressions with varying thresholds for the dispersed proteinoids (abc) and mixture of ZnO and proteinoids (def), (ad) 2-bit input string, (be) 4-bit input string, (cf) 8-bit input string.

and  $r(H, S) = 0.9518, R^2 = 0.9059$ , shows strong positive linear correlation. While Shannon entropy  $H$  vs space filling  $D$  (Fig. 5c) show very weak negative correlations,  $r(H, D) = -0.1722, R^2 = 0.0297$ .

Based on measures calculated we can construct the following hierarchies of complexity:

- CA representing 2-ary functions.
  - LZ:  $\{f_8, f_7\} \gg \{f_1, f_3, f_4\} > f_{11} > f_2$
  - H:  $\{f_8, f_7\} \gg f_1 > f_2 \gg \{f_{11}, f_3, f_4\}$
  - S:  $\{f_8, f_7\} > f_1 \gg \{f_3, f_4\} > \{f_2, f_{11}\}$
  - E:  $f_7 > f_8 \gg \{f_1, f_4, f_3\} > \{f_2, f_{11}\}$
- CA representing 4-ary functions.
  - LZ:  $f_{20} > f_{19} > \{f_{12}, f_{14}\} > f_{40}$
  - H:  $f_{12} > \{f_{14}, f_{20}\} > \{f_{19}, f_{40}\}$
  - S:  $\{f_{12}, f_{20}\} > f_{14} > \{f_{19}, f_{40}\}$
  - E:  $f_{20} > \{f_{12}, f_{14}\} > \{f_{19}, f_{40}\}$

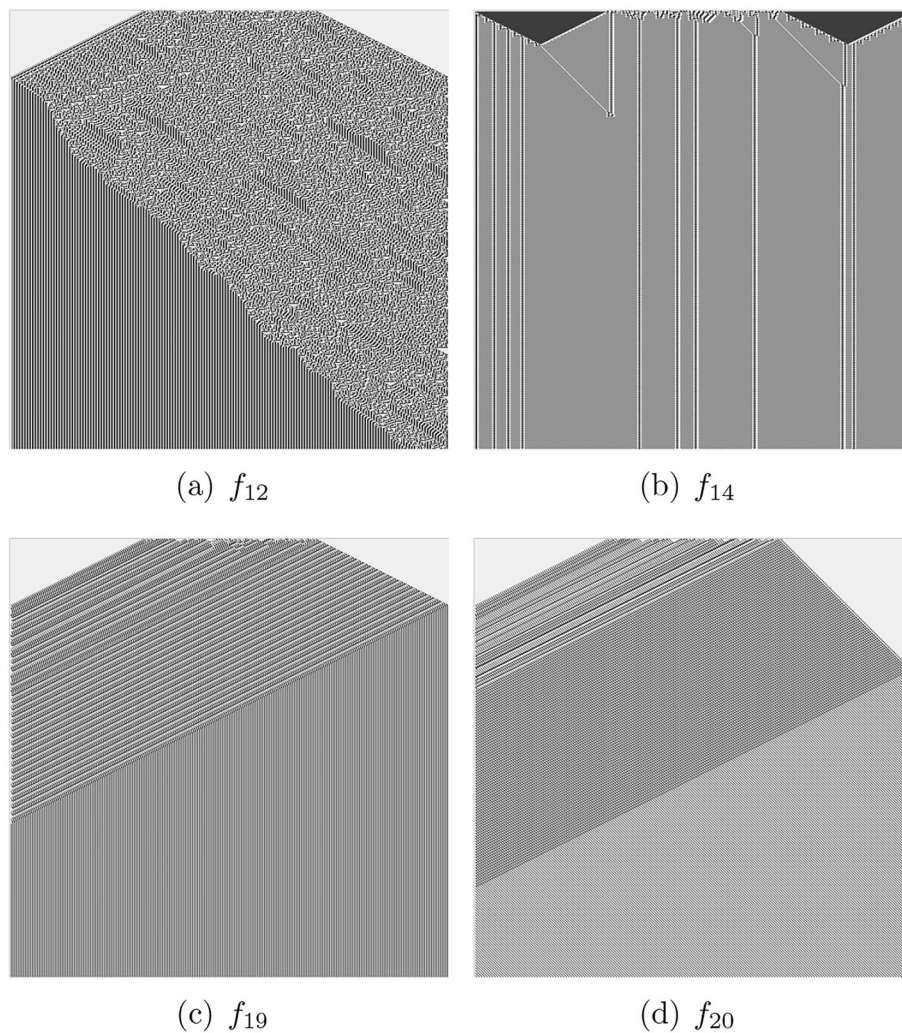


**Fig. 2.** Functions with two-arguments and those functions with four or eight arguments which produce alike patterns. Space (1D CA array) states are horizontal, and time (progressing from top to bottom) is vertical:  $x_1^0 x_2^0 \dots x_{500}^0 x_1^1 x_2^1 \dots x_{500}^1 \dots x_1^{500} x_2^{500} \dots x_{500}^{500}$ .

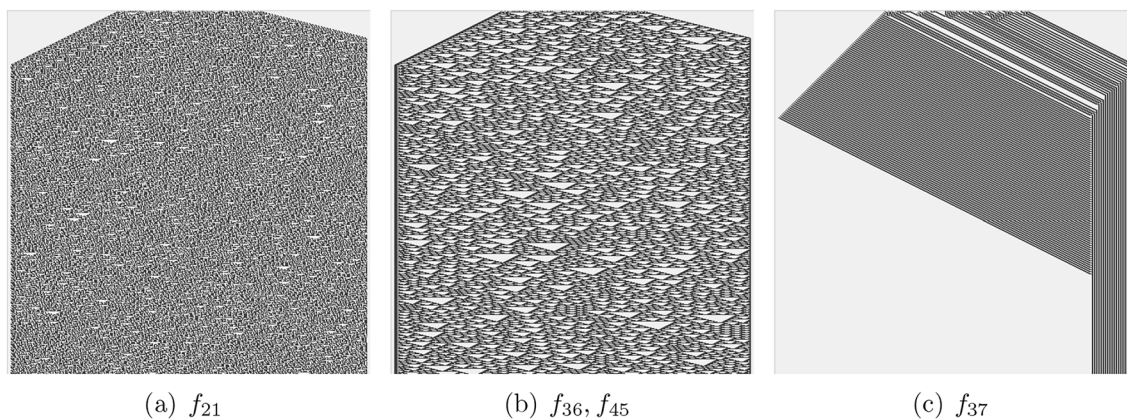
- CA representing 8-ary functions.

LZ:  $f_{21} > f_{36} > f_{37}$   
 H:  $f_{21} > \{f_{37}, f_{36}\}$   
 S:  $f_{21} > f_{37} > f_{36}$   
 E:  $f_{37} \gg f_{21} > f_{36}$

For CA governed by 2-ary functions, LZ hierarchy shows that functions  $f_8$  and  $f_7$  have the highest complexity, significantly higher than the others,  $f_1, f_3$ , and  $f_4$  have moderate complexity,  $f_{11}$  lower, and  $f_2$  the lowest. In Shannon complexity hierarchy  $f_8$  and  $f_7$  again rank highest,  $f_1$  is slightly lower, followed by  $f_2$ ; functions  $f_{11}, f_3$ , and  $f_4$  rank lowest and are grouped together. Simpson index ordering indicates that  $f_8$  and  $f_7$  have the highest structural complexity,  $f_1$  follows, with  $f_3$  and  $f_4$  significantly lower, and  $f_2$  and  $f_{11}$  the lowest. Order of expressive complexity puts function  $f_7$  as the highest, slightly higher than  $f_8$ ; functions  $f_1, f_4$ , and  $f_3$  are moderate, while  $f_2$  and  $f_{11}$  rank the lowest.



**Fig. 3.** Functions with four-arguments and those functions with eight arguments which produce alike patterns. Space (1D CA array) states are horizontal, and time (progressing from top to bottom) is vertical.



**Fig. 4.** Functions with eight-arguments. Space (1D CA array) states are horizontal, and time (progressing from top to bottom) is vertical.

(a)						
<i>f</i>	<i>LZ</i>	<i>LZ/n</i>	<i>H</i>	<i>S</i>	<i>D</i>	<i>E</i>
$f_2$	4	2	0.1	0.02	0.98	0.1
$f_{11}$	9	4.5	0.07	0.02	0.98	0.1
$f_3$	14	7	0.05	0.03	0.1	0.5
$f_4$	14	7	0.05	0.03	0.1	0.5
$f_1$	16	8	0.5	0.2	0.9	0.6
$f_7$	57	28.5	1.9	0.8	0.4	4.8
$f_8$	61	30.5	1.9	0.8	0.5	3.8
(b)						
<i>f</i>	<i>LZ</i>	<i>LZ/n</i>	<i>H</i>	<i>S</i>	<i>D</i>	<i>E</i>
$f_{12}$	11	2.75	1.4	0.7	0.6	2.3
$f_{14}$	12	3	1.1	0.6	0.5	2.2
$f_{19}$	22	5.5	0.7	0.5	0.5	1.4
$f_{20}$	42	10.5	1.1	0.7	0.32	3.4
$f_{40}$	9	4.5	0.7	0.5	0.5	1.4
(c)						
<i>f</i>	<i>LZ</i>	<i>LZ/n</i>	<i>H</i>	<i>S</i>	<i>D</i>	<i>E</i>
$f_{37}$	11	1.375	1.2	0.7	0.1	12.0
$f_{36}$	36	4.5	1.1	0.5	0.5	2.2
$f_{21}$	65	8.125	1.9	0.8	0.5	3.8

**Table 3.** Complexity of space-time patterns generated by CA derived from non-trivial Boolean functions mined in ZnO and proteinoids' colloids: *LZ* is an LZ complexity measured via size of ZIP file of the space-time configurations, *LZ/n* is the complexity normalised by the input string size, *H* is Shannon entropy, *S* is Simpson diversity, *D* is a space filling, *E* is an expressiveness. (a) Two-arguments functions, (b) Four-arguments functions, (c) Eight-arguments functions.

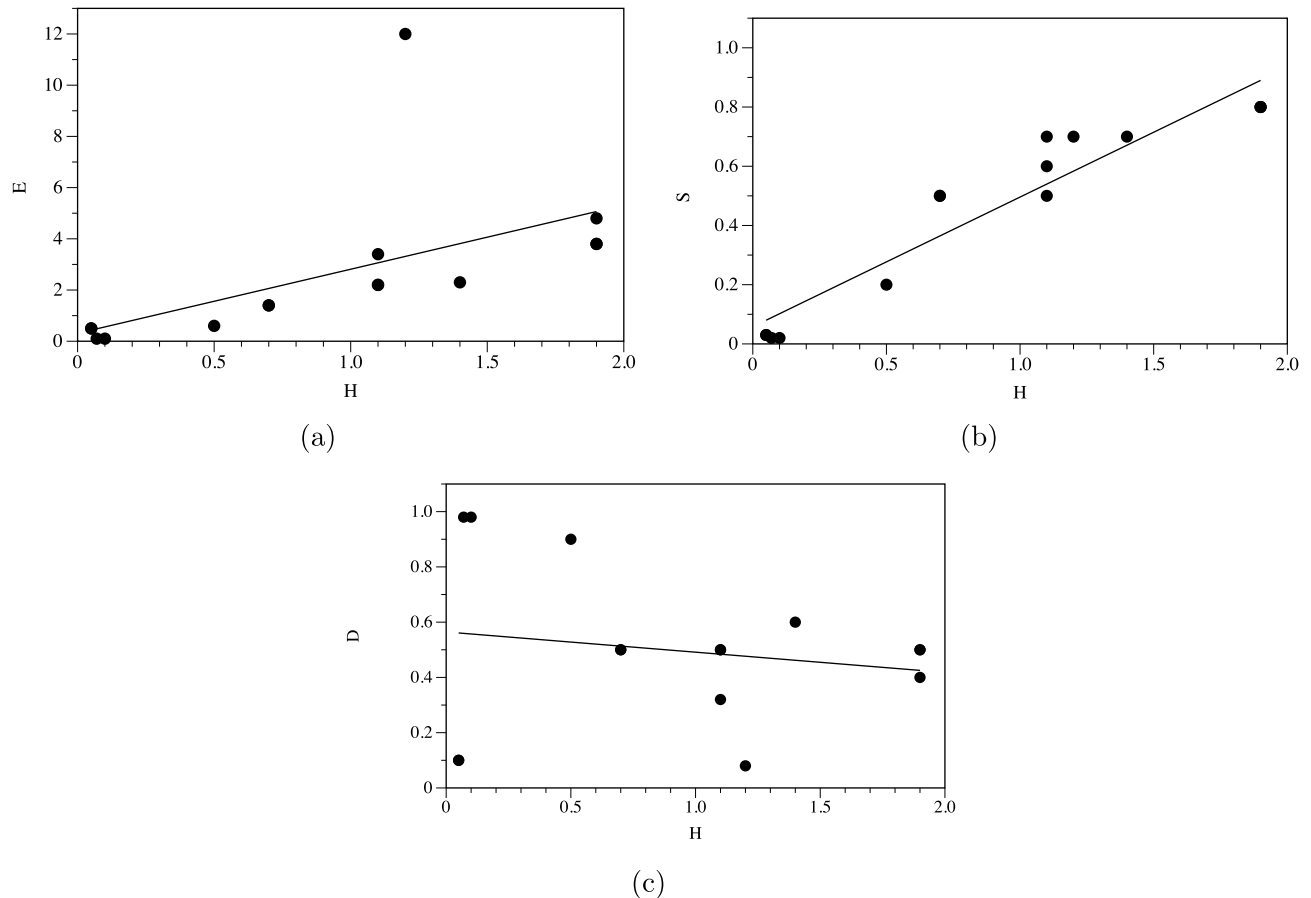
For CA governed by 4-ary functions, the order of compressibility demonstrates that function  $f_{20}$  has the highest complexity, followed by  $f_{19}$ , functions  $f_{12}$  and  $f_{14}$  have moderate complexity, and  $f_{40}$  the lowest. Shannon complexity demonstrates that function  $f_{12}$  ranks highest, with  $f_{14}$  and  $f_{20}$  following,  $f_{19}$  and  $f_{40}$  are the lowest. In Simpson hierarchy functions  $f_{12}$  and  $f_{20}$  are highest, followed by  $f_{14}$ , and  $f_{19}$  and  $f_{40}$  rank the lowest. In the expressiveness hierarchy function  $f_{20}$  is highest, with  $f_{12}$  and  $f_{14}$  in the middle, and  $f_{19}$  and  $f_{40}$  the lowest.

In CA governed by 8-ary functions, compressibility hierarchy is the following. Function  $f_{21}$  has the highest complexity, followed by  $f_{36}$  and  $f_{37}$ . In the Shannon entropy and Simpson index hierarchies function  $f_{21}$  is highest, with  $f_{37}$  and  $f_{36}$  being equal and lower. Expressiveness hierarchy shows that function  $f_{37}$  is significantly higher, followed by  $f_{21}$ , and  $f_{36}$  the lowest.

Functions  $f_8$  and  $f_7$  are consistently ranked highest across multiple criteria for 2-ary functions, indicating their higher complexity or influence. Function  $f_{21}$  is ranked highest in the majority of criteria for 8-ary functions. Different criteria (*LZ*, *H*, *S*, *E*) can yield different hierarchies. For instance, in 4-ary functions,  $f_{12}$  is ranked highest by *H* and *S* but not by *LZ* or *E*. Expressiveness measure *E* seems to have distinct rankings compared to others, especially in the 8-ary functions.

Functions which produce CA patterns with absolute highest Liv-Zempel complexity, Shannon entropy and Simpson diversity are  $f_{21}$  (Table 1c and Fig. 4a),  $f_8$  (Table 1a and Fig. 2f) and  $f_7$  (Table 1a and Fig. 2e). A function with highest expressiveness is  $f_{37}$  (Table 2c and Fig. 4c). Whilst space-time configurations of CA governed by  $f_{37}$  shows complex local dynamics, the global dynamics is dull. This shows that the expressiveness might be not a reliable measure of global complexity. If we normalise values of *LZ*, *H* and *S* complexity measures by a number of terms or literals, we will find that functions  $f_7$  and  $f_8$  are most complex functions, relative to formula complexity and in terms of space-time dynamics, discovered in colloids.





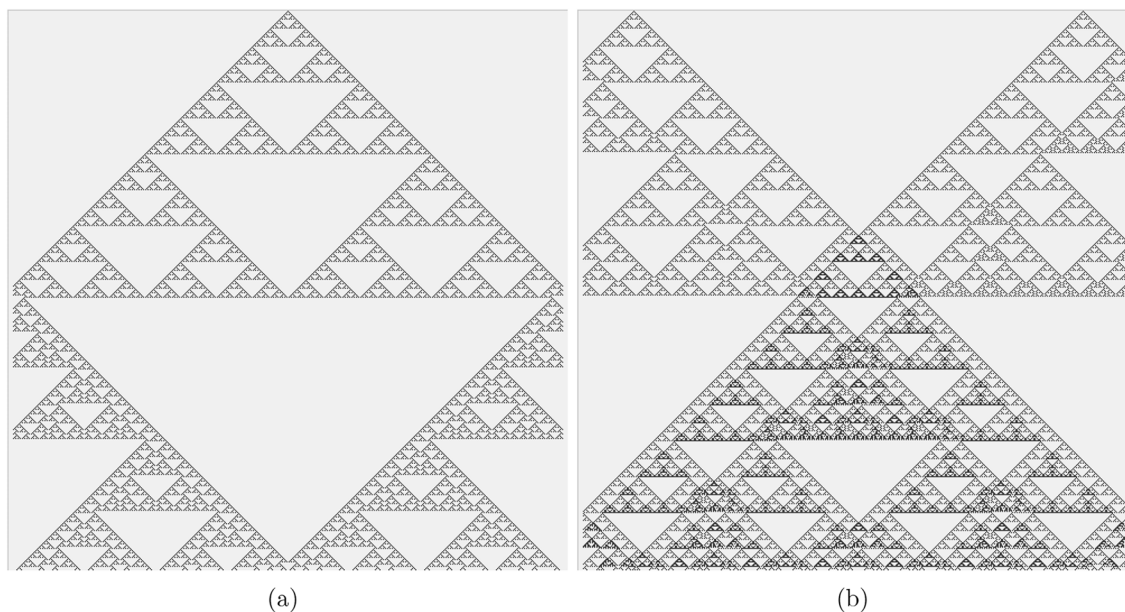
**Fig. 5.** (a) Shannon entropy  $H$  vs expressiveness  $E$ , linear approximation  $E = 0.30874 + 2.5032 * H$ . (b) Shannon entropy  $H$  vs Simpson index  $S$ , linear approximation  $S = 0.058092 + 0.43781 * H$ . (c) Shannon entropy  $H$  vs space filling  $D$ , linear approximation  $D = 0.56455 + (-0.073223) * H$ .

## Conclusion and discussion

The colloid automata—one-dimensional cellular automata (CA) governed by Boolean functions derived from ZnO, proteinoid, and their mixture, colloids—exhibit a rich spectrum of space-time evolution. Using complexity measures such as Lempel-Ziv complexity, Shannon entropy, Simpson diversity, and expressiveness, we can construct families of complexity hierarchies based on the space-time configurations of these colloid CA. These hierarchies reflect the inherent complexities of the Boolean functions and provide a means to compare and understand their behaviour across different dimensions.

The most complex, in terms of CA dynamics, functions discovered are XOR (function  $f_7$ ) and NOT XOR (function  $f_8$ ). The XOR gate is the most hard to find in natural non-linear systems, Boolean gate<sup>33–35</sup>. Moreover, already in 1960s it was demonstrated that a linear perceptron can no learn XOR function<sup>36</sup>. The use of XOR gates in modern circuit design offers several advantages, such as reduced representation size and improved testability, and optimal power consumption<sup>37</sup>. CA governed by XOR gate exhibit unpredictable dynamics, similar to that of that randomly generated patterns<sup>38</sup> and, when evolve from single non-zero state configurations produced fractal patterns—Sierpinski gasket<sup>39</sup>. An evolution of rule  $f_7$  CA started from a single cell in state ‘1’ is shown in (Fig. 6a), a reflection from absorbing boundaries is seen. The same evolution, but from two cells in state ‘1’ (Fig. 6b), shows a new fractal derived from a collision. The newly formed fractal pattern has a higher density of non-quietest cells than the parent fractal structures.

There are several limitations of the research which could be addressed in future studies. The research focuses solely on one-dimensional cellular automata. Extending this to two-dimensional or three-dimensional models could provide a more comprehensive understanding of the behaviour of colloid automata. The Boolean functions derived from ZnO, proteinoids, and their mixtures may not fully capture the complexities of information processing in real colloid systems. More sophisticated approaches incorporating physical and chemical interactions could yield more accurate results. The study is constrained by a finite set of states and rules, which might not encompass all possible behaviours of colloid systems. Exploring larger or infinite state spaces could reveal more complex dynamics. The reliance on specific complexity measures such as Lempel-Ziv complexity, Shannon entropy, Simpson diversity, and expressiveness might not capture all aspects of the system’s behaviour. Other measures or a combination of multiple metrics could provide a more holistic view. Measures like fractal dimension, Lyapunov exponents, or network-based metrics might offer new insights. Extending the research



**Fig. 6.** Space-time evolution of one-dimensional CA governed by  $f_7$ . (a) Initial configuration is a single cell (in the middle of the array) being in the state ‘1’, all others are ‘0’. (b) Two cells, based 200 cells to the left and to the right the array’s centre are at state ‘1’ at the beginning of evolution. Cells in state ‘1’ are black, ‘0’ are light-gray. Time goes top down.

to higher-dimensional cellular automata could provide deeper insights into the spatial-temporal patterns of information processing in colloid systems, potentially revealing new patterns and behaviours.

### Data availability

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

Received: 13 June 2024; Accepted: 3 September 2024

Published online: 17 September 2024

### References

1. Adamatzky, A. A brief history of liquid computers. *Philos. Trans. R. Soc. B* **374**(1774), 20180372 (2019).
2. Chiolerio, A. Liquid cybernetic systems: The fourth-order cybernetics. *Adv. Intell. Syst.* **2**(12), 2000120 (2020).
3. Chiolerio, A. & Quadrelli, M. B. Smart fluid systems: The advent of autonomous liquid robotics. *Adv. Sci.* **4**(7), 1700036 (2017).
4. Crepaldi, M. et al. Experimental demonstration of in-memory computing in a ferrofluid system. *Adv. Mater.* **35**(23), 2211406 (2023).
5. Kheirabadi, N. R., Chiolerio, A., Phillips, N. & Adamatzky, A. Learning in colloids: Synapse-like ZnO+ DMSO colloid. *Neurocomputing* **557**, 126710 (2023).
6. Fortulan, R., Kheirabadi, N. R., Mougkogiannis, P., Chiolerio, A. & Adamatzky, A. Reservoir computing with colloidal mixtures of ZnO and proteinoids. arXiv preprint [arXiv:2312.08130](https://arxiv.org/abs/2312.08130), (2023).
7. Roberts, N. et al. Logical circuits in colloids. arXiv preprint [arXiv:2307.02664](https://arxiv.org/abs/2307.02664), (2023).
8. Dale, M., Miller, J. F. & Stepney, S. Reservoir computing as a model for in-materio computing. In *Advances in Unconventional Computing*, 533–571 (Springer, 2017).
9. Dale, M., Miller, J. F., Stepney, S. & Trefzer, M. A. A substrate-independent framework to characterize reservoir computers. *Proc. R. Soc. A* **475**(2226), 20180723 (2019).
10. Konkoli, Z., Nichele, S., Dale, M. & Stepney, S. Reservoir computing with computational matter. In *Computational Matter*, 269–293. (Springer, 2018).
11. Lukoševičius, M. & Jaeger, H. Reservoir computing approaches to recurrent neural network training. *Comput. Sci. Rev.* **3**(3), 127–149 (2009).
12. Verstraeten, D., Schrauwen, B., d’Haene, M. & Strooband, D. An experimental unification of reservoir computing methods. *Neural Netw.* **20**(3), 391–403 (2007).
13. Miller, J. F. & Downing, K. Evolution in materio: Looking beyond the silicon box. In *Proceedings 2002 NASA/DoD Conference on Evolvable Hardware*, 167–176. (IEEE, 2002).
14. Miller, J. F., Harding, S. L. & Tufte, G. Evolution-in-materio: Evolving computation in materials. *Evol. Intell.* **7**(1), 49–67 (2014).
15. Miller, J. F., Hickinbotham, S. J. & Amos, M. In materio computation using carbon nanotubes. In *Computational Matter*, 33–43. (Springer, 2018).
16. Miller, J. F. The alchemy of computation: Designing with the unknown. *Nat. Comput.* **18**(3), 515–526 (2019).
17. Stepney, S. Co-designing the computational model and the computing substrate. In *International Conference on Unconventional Computation and Natural Computation*, 5–14. (Springer, 2019).
18. Roberts, N. & Adamatzky, A. Mining logical circuits in fungi. In *Fungal Machines: Sensing and Computing with Fungi*, 311–321 (Springer, 2023).
19. Eskov, V. M., Eskov, V. V., Vochmina, Y. V., Gorbunov, D. V. & Ilyashenko, L. K. Shannon entropy in the research on stationary regimes and the evolution of complexity. *Mosc. Univ. Phys. Bull.* **72**, 309–317 (2017).

20. Lin, J. Divergence measures based on the Shannon entropy. *IEEE Trans. Inf. Theory* **37**(1), 145–151 (1991).
21. Claude Elwood Shannon. A mathematical theory of communication. *Bell Syst. Tech. J.* **27**(3), 379–423 (1948).
22. Kim, B.-R. et al. Deciphering diversity indices for a better understanding of microbial communities. *J. Microbiol. Biotechnol.* **27**(12), 2089–2093 (2017).
23. Nagendra, H. Opposite trends in response for the Shannon and Simpson indices of landscape diversity. *Appl. Geogr.* **22**(2), 175–186 (2002).
24. Somerfield, P. J., Clarke, K. R. & Warwick, R. M. Simpson index. In *Encyclopedia of Ecology*, 3252–3255. (Elsevier, 2008).
25. Ziv, J. & Lempel, A. A universal algorithm for sequential data compression. *IEEE Trans. Inf. Theory* **23**(3), 337–343 (1977).
26. Adamatzky, A. & Chua, L. O. Phenomenology of retained refractoriness: On semi-memristive discrete media. *Int. J. Bifurcation Chaos* **22**(11), 1230036 (2012).
27. Redeker, M., Adamatzky, A. & Martinez, G. J. Expressiveness of elementary cellular automata. *Int. J. Mod. Phys. C* **24**(03), 1350010 (2013).
28. Adamatzky, A. Generative complexity of gray-scott model. *Commun. Nonlinear Sci. Numer. Simul.* **56**, 457–466 (2018).
29. Deutsch, P. & Gailly, J.-L. Zlib compressed data format specification version 3.3. Technical report (1996).
30. Howard, P. G. *The Design and Analysis of Efficient Lossless Data Compression Systems* (Brown University, 1993).
31. Roelofs, G. *PNG: The Definitive Guide* (O'Reilly & Associates, Inc., 1999).
32. Wolfram, S. Statistical mechanics of cellular automata. *Rev. Mod. Phys.* **55**(3), 601 (1983).
33. Adamatzky, A. & Bull, L. Are complex systems hard to evolve?. *Complexity* **14**(6), 15–20 (2009).
34. Boyar, J., Peralta, R. & Pochuev, D. On the multiplicative complexity of boolean functions over the basis  $(\wedge, \vee)$ . *Theoret. Comput. Sci.* **235**(1), 43–57 (2000).
35. Garcia-Morales, V. Universal map for cellular automata. *Phys. Lett. A* **376**(40–41), 2645–2657 (2012).
36. Minsky, M. L. & Papert, S. A. *Perceptrons: expanded edition*, (1988).
37. Ye, Y., Roy, K. & Drechsler, R. Power consumption in xor-based circuits. In *Proceedings of the ASP-DAC'99 Asia and South Pacific Design Automation Conference 1999 (Cat. No. 99EX198)*, 299–302. (IEEE, 1999).
38. Martinez, G. J., Adamatzky, A., Hoffmann, R., Désérable, D. & Zelinka, I. Patterns and dynamics of rule 22 cellular automaton. In *ACTIN COMPUTATION: Unlocking the Potential of Actin Filaments for Revolutionary Computing Systems*, 37–86. (World Scientific, 2024).
39. Sierpinski, W. Sur une courbe dont tout point est un point de ramification. *C.R. Acad. Sci.* **160**, 302–305 (1915).

## Acknowledgements

This project has received funding from the European Innovation Council And SMEs Executive Agency (EISMEA) under grant agreement No. 964388 “COGITOR: A new colloidal cybernetic system towards 2030”.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to A.A.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024