Information Flow in Cellular Automata

Stan Palasek

Sonoran Science Academy
Tucson, Arizona

Long-time correlations between components of stochastic physical systems have been observed to be stronger than an exponential decay model would predict. Simple programs should likewise exhibit slow information dissipation at large t if they are to underlie statistical systems as Stephen Wolfram's *A New Kind of Science* suggests. Here it is shown that information about the initial conditions of subsystems of many nontrivial cellular automata leaves the subsystem initially at an exponential rate until a certain threshold is reached at which the information loss slows and qualitatively changes form. This phenomenon appears among not only the elementary cellular automata (ECAs) but also more complicated rules that share features with physical systems.

1. Introduction

In 1973, D. Levesque and L. Verlet showed in a computer simulation of argon at its triple point that the transport coefficients, and with them the autocorrelation functions, decay with a tail longer than a simple hard-sphere model predicts [1]. This same tail was later observed theoretically and experimentally in the physics literature on diffusion [2, 3], fluid mechanics [4–6], and atomic spins [7], as well as in disciplines as varied as queuing theory [8], network theory [9], and biology [10, 11]. Heavy distribution tails are notably prevalent in internet retail where low-demand niche products occupy an inordinately large segment of the market [12].

In his 1983 paper "Statistical Mechanics of Cellular Automata," Wolfram observed the decay of the two-point correlation function between the center initial cell and the same cell after t steps of elementary cellular automaton (ECA) rule 18 evolution. This was essentially the covariance function between the initial and final states, yielding a seemingly exponential time series that alternated between positive and negative correlations [13]. If real-life statistical physics and fluid dynamics are to have a basis in simple programs as Wolfram suggests throughout his *New Kind of Science* [14], the long-time correlations in a computation with a simple underlying rule should exhibit a correspondingly heavy long-time tail.

Here long-term behavior in information time series is classified as it is for probability density and mass functions: the exponential model serves as the null hypothesis when maximum entropy and a constant dissipation rate are assumed. A heavy-tailed distribution has been rigorously defined as one that cannot be bounded by an exponential, while long tails are a subset of heavy tails, both of which encompass subexponential tails [15]. The subtle differences between these terms are indiscernible with the numerical methods employed here, so they will be used interchangeably (these distinctions are often not made, e.g. [16]).

2. Mutual Information in Cellular Automata

A random linear configuration of $2t+1$ cells of k colors uniquely determines a sequence of l cells after t steps of cellular automaton evolution. The mapping of the center l to the final configuration, however, is dependent on the stochastic selection of the outer $2t$ initial cells. Thus, examining this evolution gives an indication as to how much of the information was preserved. Here mutual information will be chosen as the metric for correlation as it remains unaffected by biased distributions of colors produced by particular ECA rules. For two random variables X and Y, mutual information is defined as

$$I(X;\, Y) \equiv \left\langle \ln \frac{p(X,\, Y)}{p(X)\, p(Y)} \right\rangle = \sum_{x \in X,\, y \in Y} p(x,\, y) \ln \frac{p(x,\, y)}{p(x)\, p(y)}$$

measured in nats, where $\langle \cdots \rangle$ is the expectation operator, $p\,(X,\, Y)$ is the joint probability distribution of X and Y, and the sum is taken over the entire sample space [17].

3. Information Arrays

Sampling the space of k^{2t+1} initial conditions for an $l = 1$ evolution, the dispersion of information can be realized about the center initial cell. In Figure 1 the flow of information for five ECAs is shown with their evolutions from a single black cell and from random configurations. The gray level is the relative fourth root of the mutual information. The first two columns show 50 steps of evolution and the third shows 10 steps, sampling 10^7 initial configurations.

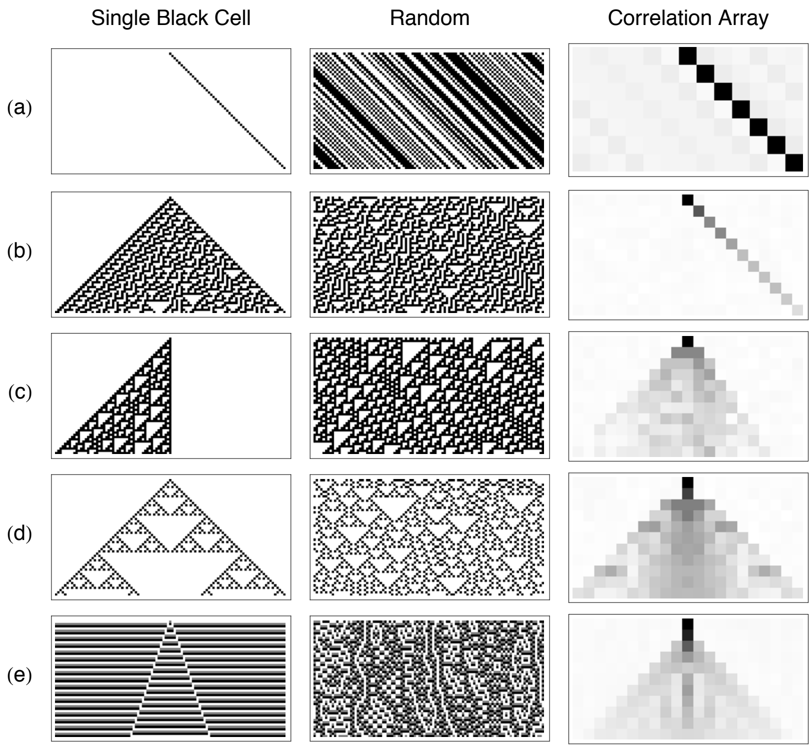

Figure 1. In order, the automata are (a) rule 240, which simply shifts the initial cells to the right; (b) rule 30, the standard for class 3 behavior from simple initial conditions; (c) rule 110, the standard for class 4 behavior and universality; (d) rule 18, the nested rule used in [13]; and (e) rule 1123289366095, a class 4 reversible rule referenced on p. 436 of [14].

4. Information Flow in Rule 30

Rule 30 interestingly transfers information strictly to the right, which gives its information array the appearance of rule 240 with dissipation. This is depicted in Figure 2 along with the exponential information loss along this diagonal.

This observation can be accounted for in the details of the rule used; under rule 30, a black left cell yields a white cell and a white left cell yields a black cell, both with a single exception. Therefore, rule 30 evolution can be computed with 75% (6/8) accuracy, which explains the exponential diagonal of information preservation. In the two exceptions to the sufficiency of the left cell, knowing additionally the center cell is still insufficient. Therefore, the observed dissipation is inevitable.

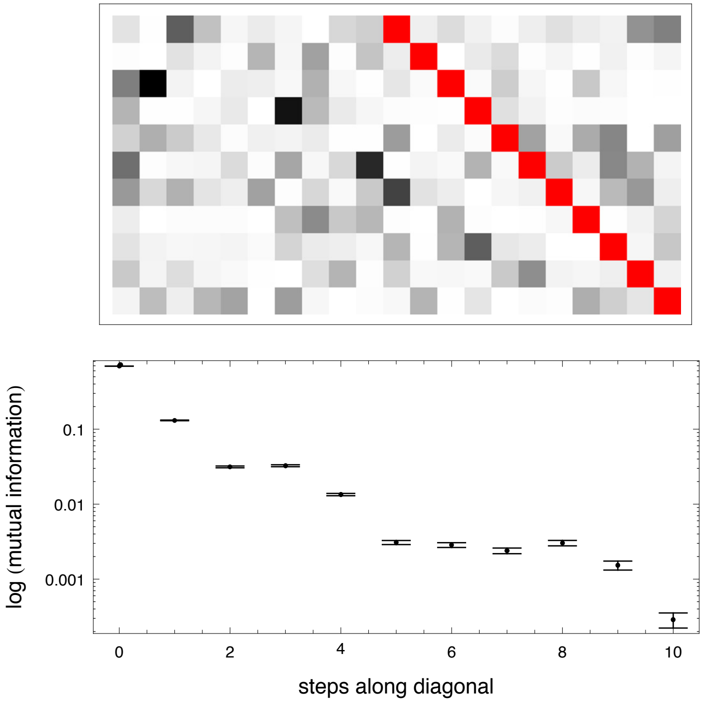

Figure 2. Information array for rule 30 with the strong diagonal signal replaced with red so the background noise can be seen (upper) and exponential information loss on that diagonal (lower). During 10 evolutions, 100 000 initial conditions were sampled. One standard deviation error bars are based on bootstrapped parameter estimates for which the data at each time is resampled 20 times.

5. Information Time Series

We compute the mutual information time series described in Section 2 for all ECAs for 15 steps of evolution, each time sampling 50 000 initial conditions. Blocks of length three are considered, as single cells do not maintain enough information for fruitful analysis. Of the 256 ECAs, about 120 exhibit exponential or near-exponential decay, while the rest quickly converge to $I = 0$ after some transient. Approximately 60% of the exponential curves show a sudden bend in their log plot (although there is some inherent subjectivity in enumerating classes as discussed by Wolfram in [14, p. 240]). Consider, for instance, rule 54, which is class 4 from random initial conditions, left-right symmetric, and supports several distinct background patterns. Figure 3 shows that mutual information indeed appears to decay exponentially for only the first five steps of evolution. On the sixth step, however, the sequence of R^2 coefficients falls off the plateau of strong relationships with $R^2 > 0.999$.

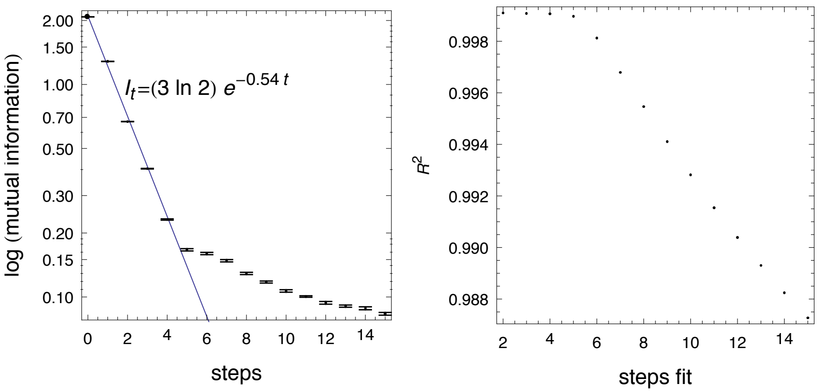

Figure 3. Log plot for mutual information for rule 54 sampling 100 000 initial conditions for 15 steps and the sequence of R^2 coefficients for an exponential fit of I_t from time 0 to t. The original strong exponential relationship evidently holds only for the first four to five steps.

This unquestionably exhibits a "long-time tail"; mutual information stays in this subsystem of length three longer than a simple exponential model predicts. Beginning with the transition from $t = 4$ to $t = 5$, not only does information dissipation slow, but the function generating the sequence I_t appears to qualitatively change form. This long tail is not unique to rule 54 and its equivalences, appearing frequently among the ECAs. Consider, for instance, the computationally universal and frequent subject of [14], rule 110. For large t, the information tail apparently takes the form of a power law (see Figure 4). Similar power laws as long-time tails have been observed in the literature in, for instance, [2, 7, 18].

The information dissipation behavior for each of the 88 nonequivalent ECAs is displayed in Table 1 to demonstrate the prevalence of the bent exponential form over pure exponential dissipation, or instances where all correlation is lost after a short transient. The "trivial" category includes automata that clearly either preserve all information, forget all information, or exhibit a predictable periodic pattern: typically class 1 in Wolfram's classification. The "rapid loss" automata are those that, after a transient where some local memory is present, lose any dependence upon the initial center cells. This is often due to a diagonal light cone like that of rule 30. The final two categories, "exponential" and "bent exponential," seem to be closely related and

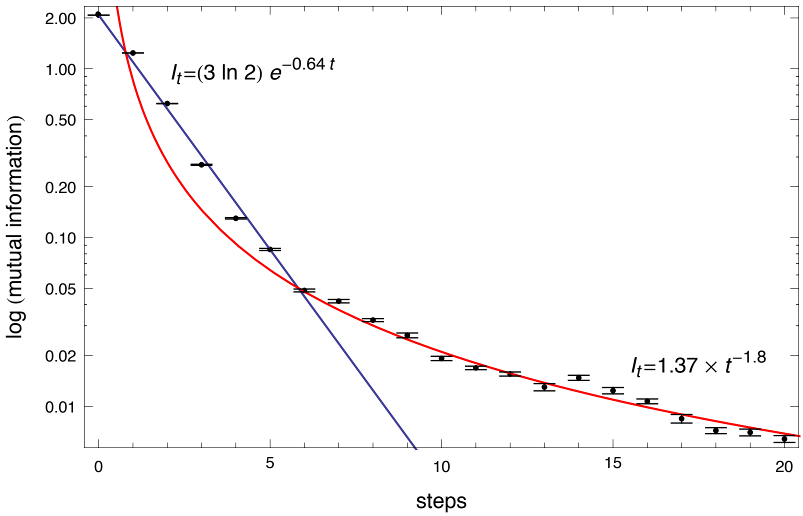

Figure 4. Log plot for mutual information for rule 110 sampling 1 000 000 initial conditions for 20 steps shown with exponential (blue) and power (red) fits.

may fall within a broader classification in which the abruptness of the corner and magnitude of the change in slope fall on a continuum. Both the initial and final slopes are given for the exponential rules designated as "bent."

6. Physical Implications

As mentioned in Section 1, Levesque et al. [1] found that at large times in a simulated system of argon gas, the shear viscosity defined by the Kubo correlation function exhibits a long-time tail. The observation of this tail made it possible to resolve inconsistencies in the evolution of the velocity autocorrelation function, making it a requisite feature of any model representing a statistical mechanical system. The simple programs examined here thus far, though in some cases capable of emulating any other computation, are not representative of any model that would be considered physical as they are not computationally reversible. None of the nontrivial ECAs have this property [14, p. 436]. There are, however, time-invertible three-color cellular automata that exhibit class 4 behavior from random initial conditions. One was already used as an example in the final row of Figure 1 and its time series is shown in Figure 5.

Rule	Dissipation Form	Decay Constant	Rule	Dissipation Form	Decay Constant
0	trivial	N/A	56	bent exponential	−0.601 → −0.061
1	trivial	N/A	57	bent exponential	−0.48 → −0.131
2	rapid loss	−0.76	58	bent exponential	−0.51 → −0.156
3	rapid loss	−0.381	60	rapid loss	−0.499
4	trivial	N/A	62	bent exponential	−0.436 → −0.13
5	trivial	N/A	72	trivial	N/A
6	exponential	−0.681	73	rapid loss	−0.695
7	rapid loss	−0.458	74	exponential	−0.764
8	trivial	N/A	76	trivial	N/A
9	exponential	−0.626	77	rapid loss	−0.471
10	rapid loss	−0.535	78	rapid loss	−0.415
11	exponential	−0.481	90	rapid loss	−1.255
12	trivial	N/A	94	rapid loss	−0.759
13	rapid loss	−0.473	104	rapid loss	−0.859
14	bent exponential	−0.444 → −0.127	105	rapid loss	−1.255
15	rapid loss	−0.499	106	rapid loss	−0.843
18	bent exponential	−0.754 → −0.072	108	rapid loss	−0.363
19	rapid loss	−0.344	110	bent exponential	−0.569 → −0.03
22	bent exponential	−0.931 → −0.07	122	bent exponential	−0.85 → −0.057
23	rapid loss	−0.47	126	bent exponential	−0.818 → −0.097
24	rapid loss	−0.822	128	exponential	−1.205
25	bent exponential	−0.471 → −0.082	130	rapid loss	−0.663
26	bent exponential	−0.845 → −0.094	132	rapid loss	−0.617
27	exponential	−0.358	134	exponential	−0.782
28	rapid loss	−0.427	136	exponential	−0.589
29	trivial	N/A	138	rapid loss	−0.499
30	rapid loss	−0.75	140	rapid loss	−0.224
32	exponential	−1.079	142	bent exponential	−0.451 → −0.135
33	rapid loss	−0.666	146	bent exponential	−0.815 → −0.047
34	rapid loss	−0.552	150	rapid loss	−1.255
35	bent exponential	−0.258 → −0.221	152	exponential	−0.522
36	trivial	N/A	154	rapid loss	−0.75
37	bent exponential	−0.816 → −0.005	156	rapid loss	−0.407
38	rapid loss	−0.538	160	exponential	−0.787
40	exponential	−0.772	162	rapid loss	−0.601
41	bent exponential	−0.866 → −0.394	164	rapid loss	−0.853
42	rapid loss	−0.486	168	bent exponential	−0.566 → −0.287
43	bent exponential	−0.45 → −0.07	170	rapid loss	−0.499
44	rapid loss	−0.529	172	bent exponential	−0.354 → 0
45	rapid loss	−0.84	178	bent exponential	−0.47 → 0
46	rapid loss	−0.475	184	bent exponential	−0.518 → −0.09
50	rapid loss	−0.356	200	rapid loss	−0.247
51	trivial	N/A	204	trivial	N/A
54	bent exponential	−0.535 → −0.039	232	rapid loss	−0.471

Table 1. Classification of the nonequivalent ECAs.

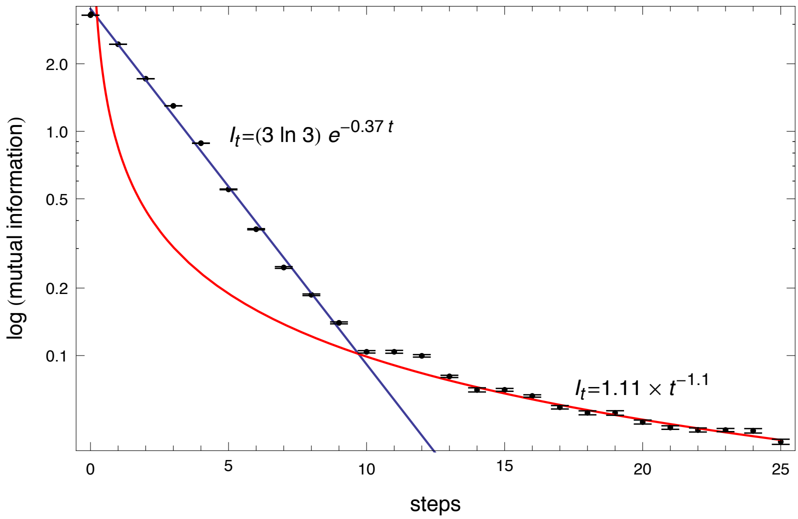

Figure 5. Log plot for mutual information for three-color class 4 reversible left-right symmetric rule 1123289366095 sampling 100 000 initial conditions for 25 steps shown with exponential (blue) and power (red) fits.

By definition, no information is lost when evolving based on a time-invertible rule like the one in Figure 5. Although there is dissipation from the $l = 3$ block examined in Figure 5, information again stays in the subsystem longer than the initial exponential decay would predict. It might seem that the dissipation is always exponential based on *ab initio* observations including, for instance, that trajectories in phase space diverge exponentially [19] and that there are k^{2t} times as many initial conditions as there are final ($k^{2t+l} \mapsto k^l$). Nonetheless, the abrupt threshold in this rule demonstrates dramatically the potential for simple discrete systems to emulate the information dynamics of processes in statistical physics. Unfortunately, the cited literature puts most of its focus on the nature of the information tails at large t rather than nonequilibrium effects like this power-exponential threshold. Therefore, it must now be determined which real or synthetic systems exhibit the same piecewise dissipation and whether it can be attributed to features of the relationships between their components.

Acknowledgments

I thank Taliesin Beynon, Todd Rowland, Stephen Wolfram, and the Wolfram Summer School for direction and useful discussions.

References

[1] D. Levesque, L. Verlet, and J. Kürkijarvi, "Computer 'Experiments' on Classical Fluids. IV. Transport Properties and Time-Correlation Functions of the Lennard–Jones Liquid Near Its Triple-Point," *Physical Review A*, 7(5), 1973 pp. 1690–1700. doi:10.1103/PhysRevA.7.1690.

[2] G. L. Paul and P. N. Pusey, "Observation of a Long-Time Tail in Brownian Motion," *Journal of Physics A*, **14**(12), 1981 p. 3301. doi:10.1088/0305-4470/14/12/025.

[3] J. W. Haus, K. W. Kehr, and K. Kitahara, "Long-Time Tail Effects on Particle Diffusion in a Disordered System," *Physical Review B*, **25**(7), 1982 pp. 4918–4921. doi:10.1103/PhysRevB.25.4918.

[4] D. Forster, D. R. Nelson, and M. J. Stephen, "Long-Time Tails and the Large-Eddy Behavior of a Randomly Stirred Fluid," *Physical Review Letters,* **36**(15), 1976 pp. 867–870. doi:10.1103/PhysRevLett.36.867.

[5] D. Forster, D. R. Nelson, and M. J. Stephen, "Large-Distance and Long-Time Properties of a Randomly Stirred Fluid," *Physical Review A*, **16**(2), 1977 pp. 732–749. doi:10.1103/PhysRevA.16.732.

[6] J. Stepišnik and P. T. Callaghan, "The Long Time Tail of Molecular Velocity Correlation in a Confined Fluid: Observation by Modulated Gradient Spin-Echo NMR," *Physica B: Condensed Matter*, **292**(3–4), 2000 pp. 296–301. doi:10.1016/S0921-4526(00)00469-5.

[7] S. Kambe, H. Yasuoka, A. Hayashi, and Y. Ueda, "Anomalous Long Time Tail Behavior of Spin-Spin Correlation in the Normal State of $Tl_2Ba_2CuO_y$," *Physical Review Letters*, **73**(1), 1994 pp. 197–200. doi:10.1103/PhysRevLett.73.197.

[8] J. Abate and W. Whitt, "Asymptotics for *M*/*G*/1 Low-Priority Waiting-Time Tail Probabilities," *Queueing Systems*, **25**(1–4), 1997 pp. 173–233. doi:10.1023/A:1019104402024.

[9] A-L. Barabási and R. Albert, "Emergence of Scaling in Random Networks." http://arxiv.org/abs/cond-mat/9910332.

[10] J. F. Nagle, "Long Tail Kinetics in Biophysics?" *Biophysical Journal*, **63**(2), 1992 pp. 366–370. doi:10.1016/S0006-3495(92)81602-8.

[11] P. Leiderman, M. Ben-Ziv, L. Genosar, D. Huppert, K. M. Solntsev, and L. M. Tolbert, "Study of the Long-Time Fluorescence Tail of the Green Fluorescent Protein," *Journal of Physical Chemistry B*, **108**(23), 2004 pp. 8043–8053. doi:10.1021/jp037786i.

[12] A. Elberse and F. Oberholzer-Gee, "Superstars and Underdogs: An Examination of the Long-Tail Phenomenon in Video Sales," *MSI Reports: Working Paper Series* 4, 2007, pp. 49–72.

[13] S. Wolfram, "Statistical Mechanics of Cellular Automata," *Reviews of Modern Physics*, **55**(3), 1983 pp. 601–644. doi:10.1103/RevModPhys.55.601.

[14] S. Wolfram, *A New Kind of Science*, Champaign, IL: Wolfram Media, Inc., 2002.

[15] S. Asmussen, *Applied Probability and Queues*, New York: Springer, 2003.

[16] A. Feldmann and W. Whitt, "Fitting Mixtures of Exponentials to Long-Tail Distributions to Analyze Network Performance Models," *Proceedings of the Sixteenth Annual IEEE Computer and Communications Societies (INFOCOM97)*, Vol. 3, Kobe, IEEE, 1997 pp. 1096–1104. doi:10.1109/INFCOM.1997.631130.

[17] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, Hoboken, NJ: Wiley-Interscience, 2006.

[18] B. U. Felderhof, "Effect of the Wall on the Velocity Autocorrelation Function and Long-Time Tail of Brownian Motion," *Journal of Physical Chemistry B*, **109**(45), 2005 pp. 21406–21412. doi:10.1021/jp051335b.

[19] W. Greiner, *Classical Mechanics: Systems of Particles and Hamiltonian Dynamics*, New York: Springer, 2003.