Simply modified GKL density classifiers that reach consensus faster

J. Ricardo G. Mendonça

Version of Record: https://www.sciencedirect.com/science/article/pii/S0375960119303512
Manuscript_976d0a5b4fe1c40fee1d19621dc49620

Simply modified GKL density classifiers that reach consensus faster

J. Ricardo G. Mendonça[a,b,*]

[a] *LPTMS, UMR 8626, CNRS, Université Paris-Sud, Université Paris Saclay, 91405 Orsay CEDEX, France*
[b] *Escola de Artes, Ciências e Humanidades, Universidade de São Paulo, Rua Arlindo Bettio 1000, 03828-000 São Paulo, SP, Brazil*

Abstract

The two-state Gacs-Kurdyumov-Levin (GKL) cellular automaton has been a staple model in the study of complex systems due to its ability to classify binary arrays of symbols according to their initial density. We show that a class of modified GKL models over extended neighborhoods, but still involving only three cells at a time, achieves comparable density classification performance but in some cases reach consensus more than twice as fast. Our results suggest the time to consensus (relative to the length of the CA) as a complementary measure of density classification performance.

Keywords: Cellular automata, density classification problem, spatially distributed computing, emergence

In 1978, Gacs, Kurdyumov, and Levin (GKL) introduced the density classification problem for cellular automata (CA) in the literature [1–3]. The problem consists in classifying arrays of symbols according to their initial density using local rules, and is completed successfully if all the cells of the CA converge to the initial majority state in linear time in the size of the input array. Density classification is a nontrivial task for CA composed of autonomous and memoryless cells because the cells have to achieve a global consensus cooperating locally; emergence of collective behavior is required. The GKL two-state model, or GKL-II for short, became a staple model in the theory of complex systems related with the concepts of communication, efficiency, and emergence [4–6]. It has been demonstrated that the density classification problem cannot be solved correctly 100% of the times by uniform two-state CA, although no upper bound on the maximum possible efficiency has been set [7, 8]. Solutions involving nonuniform CA and less strict criteria for what a solution to the problem means exist [9, 10]. Recent reviews on the density classification problem for CA are given in [11, 12].

The GKL-II CA is a finite one-dimensional array of $n \geq 4$ cells under periodic boundary conditions evolving by the action of a transition function $\Phi_{\rm II}\colon \{0,1\}^n \to \{0,1\}^n$ that given the state $\boldsymbol{x}^t = (x_1^t, \ldots, x_n^t)$ of the CA at instant t determines its state $\boldsymbol{x}^{t+1} = \Phi_{\rm II}(\boldsymbol{x}^t)$ at instant $t+1$ by the majority rule

$$x_i^{t+1} = \begin{cases} \mathrm{maj}(x_{i-3}^t, x_{i-1}^t, x_i^t), & \text{if } x_i^t = 0, \\ \mathrm{maj}(x_i^t, x_{i+1}^t, x_{i+3}^t), & \text{if } x_i^t = 1, \end{cases} \tag{1}$$

where $\mathrm{maj}(p,q,r) = \lfloor \frac{1}{2}(p+q+r) \rfloor$ for 0-1 variables p, q, r. The CA classifies density if $\boldsymbol{x}^t \to \mathbf{0} = (0,\ldots,0)$ or $\mathbf{1} = (1,\ldots,1)$ depending whether, respectively, the initial density $\rho^0 = n^{-1}\sum_i x_i^0 < 1/2$ or $\rho^0 > 1/2$. We do not require a definite behavior when $\rho^0 = 1/2$. The CA is supposed to reach consensus in $O(n)$ time steps. In [1, 3], the authors prove that the GKL-II CA on the infinite lattice $\mathbb{Z}$ displays the eroder property, washing out finite islands of the minority phase in finite time and eventually leading the CA to one of the two invariant states $\mathbf{0}$ or $\mathbf{1}$. In an array of $n = 149$ cells (odd length to avoid ties), GKL-II scores an average density classification performance of 81.5% over random initial conditions with each cell initialized in the state 0 or 1 equally at random, taking on average $86 \sim 0.576n$ time steps to reach consensus. Details on the GKL-II performance are given in [4–6, 11–14].

We now modify the neighborhood in the GKL-II. Instead of evaluating the majority vote of cell i with its nearest $i \pm 1$ and third $i \pm 3$ neighbours, we pick neighbors $i \pm j$ and $i \pm k$, with $k > j \geq 1$. The rules for the modified CA read

$$x_i^{t+1} = \begin{cases} \mathrm{maj}(x_{i-k}^t, x_{i-j}^t, x_i^t) & \text{if } x_i^t = 0, \\ \mathrm{maj}(x_i^t, x_{i+j}^t, x_{i+k}^t) & \text{if } x_i^t = 1. \end{cases} \tag{2}$$

We refer to this CA as GKL(j,k); GKL$(1,3)$ recovers

*Permanent address: Escola de Artes, Ciências e Humanidades, Universidade de São Paulo, SP, Brazil. Email: jricardo@usp.br.

Table 1: Best density classification performances of GKL(j,k) in the range $1 \le j \le 5$, $j < k \le 15$ in an array of $n = 299$ cells averaged over 10^6 random initial configurations near the critical density $\rho^0 = 1/2$. The uncertainty in the performance $\langle f \rangle$ is ± 0.0004. GKL-II figures are displayed in bold for comparison.

(j,k)	$(4,12)$	$(3,9)$	$(2,6)$	$(5,15)$	$(\mathbf{1},\mathbf{3})$	$(1,9)$	$(1,11)$	$(2,14)$	$(2,10)$	$(3,15)$	$(1,7)$	$(1,5)$
$\langle f \rangle$	0.7926	0.7922	0.7921	0.7920	**0.7917**	0.7893	0.7875	0.7874	0.7873	0.7869	0.7868	0.7865
$\langle t^* \rangle / n$	0.5843	0.5848	0.5844	0.5849	**0.5848**	0.2633	0.2289	0.3270	0.4123	0.4123	0.3269	0.4122

the original GKL-II model. To the best of our knowledge these models have never been considered in the literature before. We measured the average density classification performance $\langle f \rangle$ of GKL(j,k) over 10^6 random initial states close to the critical density ($x_i^0 = 0$ or 1 equally at random) in an array of $n = 299$ cells to minimize finite-size effects that show up in the rules with larger (j,k). Our results appear in Table 1. We see that the GKL(j,k) with $k = 3j$, i. e., the GKL$(j,3j)$ models, all display virtually the same density classification and relative time to consensus ($\langle t^* \rangle / n$) performances. Otherwise, the GKL$(1,9)$ and GKL$(1,11)$ models display almost the same density classification performance as GKL-II but achieve consensus in about half the time. Explicitly, GKL$(1,11)$ is just about 0.53% less efficient than GKL-II but is ~ 2.5 times faster. From Table 1 we conclude that if quality is critical, then GKL$(4,12)$ is the best CA in its class, while if one needs speed, then GKL$(1,9)$ or GKL$(1,11)$ becomes the CA of choice.

Figure 1 displays the average classification performance of the GKL$(1,k)$ CA as a function of the imbalance $\delta = \frac{1}{2}(n_1 - n_0)$ between the number of cells in states 1 and 0 in the initial configuration. Here the initial density $\rho^0 = 1/2 + \delta/n$ is fixed but the configurations are random. By symmetry, the performance of the CA depends only on the magnitude of δ, not on its sign. The data show that the density classification performance of all these CA are close over a range of initial densities, differing significantly, however, on the time to consensus. Space-time diagrams of some GKL(j,k) CA are displayed in Figure 2.

We do not currently have a sound explanation for the efficient combinations of j, k found. The efficiency of the GKL$(j,3j)$ can be related with that of GKL$(1,3)$ in one or more sublattices, although the fast convergence of GKL$(1,9)$ and GKL$(1,11)$ cannot be immediately related with any sublattice dynamics. Intuitively, in the GKL(j,k) CA information about the dynamics of the interfaces between islands of 0s and 1s can jump over longer distances (i. e., move faster) with increased $k - j$. Data from Table 1 for the time to consensus for GKL$(1,k)$ with $k = 3$, 5, 7, 9, and 11 corroborates this idea. Note that the metric $\langle t^* \rangle / n$ is

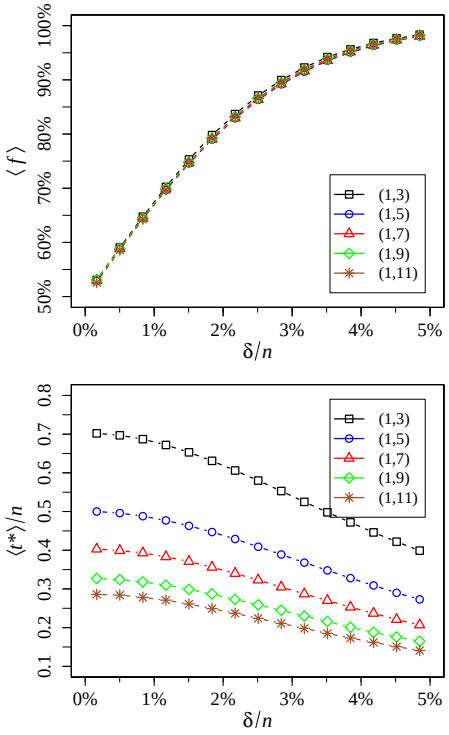

Figure 1: Density classification performance $\langle f \rangle$ and time to consensus $\langle t^* \rangle / n$ averaged over 10^6 random initial configurations of some GKL$(1,k)$ CA of length $n = 299$ as a function of the relative imbalance $\delta/n = \frac{1}{2}(n_1 - n_0)/n$ in the initial configurations. Error bars ($\pm 0.5\%$ or smaller) are of the order of the sizes of the symbols shown.

not unique—one could consider the alternative timings given by $\langle t^* \rangle / nk$, with k the radius of the CA, as well as $\langle t^* \rangle / nz$, with z the number of cells that enter the local rule ($z = 3$ for the GKL(j,k)). A characterization of the "computational mechanics" of the GKL(j,k) CA [4–6, 15] may help to understand their eroder mechanism and their efficiency better. It would also be of interest to assess the robustness of the GKL(j,k) against noise and whether the ensuing probabilistic CA display an ergodic-nonergodic transition, a long-standing unsettled issue for one-dimensional density classifiers [1–3, 12–18]. These and related questions (e. g., how $\langle f(n, \delta > 0) \rangle \to 1$ as $n \nearrow \infty$, see [8]) will be the subject of forthcoming publications.

Acknowledgments

The author thanks Nazim Fatès (LORIA, Nancy) for useful conversations and an anonymous reviewer for valuable suggestions improving the manuscript. The author also acknowledges the LPTMS for kind hospitality during a sabbatical leave in France and FAPESP (Brazil) for partial support through grant nr. 2017/22166-9.

References

References

[1] P. Gach, G. L. Kurdyumov, L. A. Levin, One-dimensional uniform arrays that wash out finite islands, Probl. Inf. Transm. 14 (3) (1978) 223–226.

[2] A. L. Toom, N. B. Vasilyev, O. N. Stavskaya, L. G. Mityushin, G. L. Kurdyumov, S. A. Pirogov, Discrete local Markov systems, in: R. L. Dobrushin, V. I. Kryukov, A. L. Toom (Eds.), Stochastic Cellular Systems: Ergodicity, Memory, Morphogenesis, Manchester University Press, Manchester, 1990, pp. 1–182.

[3] P. Gonzaga de Sá, C. Maes, The Gacs-Kurdyumov-Levin automaton revisited, J. Stat. Phys. 67 (3–4) (1992) 507–522.

[4] M. Mitchell, P. T. Hraber, J. P. Crutchfield, Revisiting the edge of chaos: Evolving cellular automata to perform computations, Complex Syst. 7 (2) (1993) 89–130.

[5] J. P. Crutchifeld, M. Mitchell, The evolution of emergent computation, Proc. Natl. Acad. Sci. USA 92 (23) (1995) 10742–10746

[6] J. P. Crutchfield, M. Mitchell, R. Das, The evolutionary design of collective computation in cellular automata, in: J. P. Crutchfield, P. Schuster (Eds.), Evolutionary Dynamics: Exploring the Interplay of Selection, Neutrality, Accident, and Function, Oxford University Press, New York, 2003, pp. 361–411.

[7] M. Land, R. K. Belew, No perfect two-state cellular automata for density classification exists, Phys. Rev. Lett. 74 (25) (1995) 5148–5150.

[8] A. Busić, N. Fatès, I. Marcovici, J. Mairesse, Density classification on infinite lattices and trees, Electron. J. Probab. 18 (2013) 51.

[9] H. Fukś, Solution of the density classification problem with two cellular automata rules, Phys. Rev. E 55 (3) (1997) R2081–R2084.

[10] M. Sipper, M. S. Capcarrere, E. Ronald, A simple cellular automaton that solves the density and ordering problems, Int. J. Mod. Phys. C 9 (7) (1998) 899–902.

[11] P. P. B. de Oliveira, On density determination with cellular automata: Results, constructions and directions, J. Cell. Autom. 9 (5–6) (2014) 357–385.

[12] N. Fatès, Stochastic cellular automata solutions to the density classification problem – When randomness helps computing, Theory Comput. Syst. 53 (2) (2013) 223–242.

[13] J. R. G. Mendonça, Sensitivity to noise and ergodicity of an assembly line of cellular automata that classifies density, Phys. Rev. E 83 (3) (2011) 031112.

[14] J. R. G. Mendonça, R. E. O. Simões, Density classification performance and ergodicity of the Gacs-Kurdyumov-Levin cellular automaton model IV, Phys. Rev. E 98 (1) (2018) 012135.

[15] P. Gács, I. Törmä, Stable multi-level monotonic eroders, arXiv:1809.09503 [math.PR].

[16] J. Mairesse, I. Marcovici, Around probabilistic cellular automata, Theor. Comput. Sci. 559 (2014) 42–72.

[17] R. Fernández, P.-Y. Louis, F. R. Nardi, Overview: PCA models and issues, in: P.-Y. Louis, F. R. Nardi (Eds.), Probabilistic Cellular Automata, Springer, Cham, 2018, pp. 1–30.

[18] I. Marcovici, M. Sablik, S. Taati, Ergodicity of some classes of cellular automata subject to noise, arXiv:1712.05500 [math.PR].

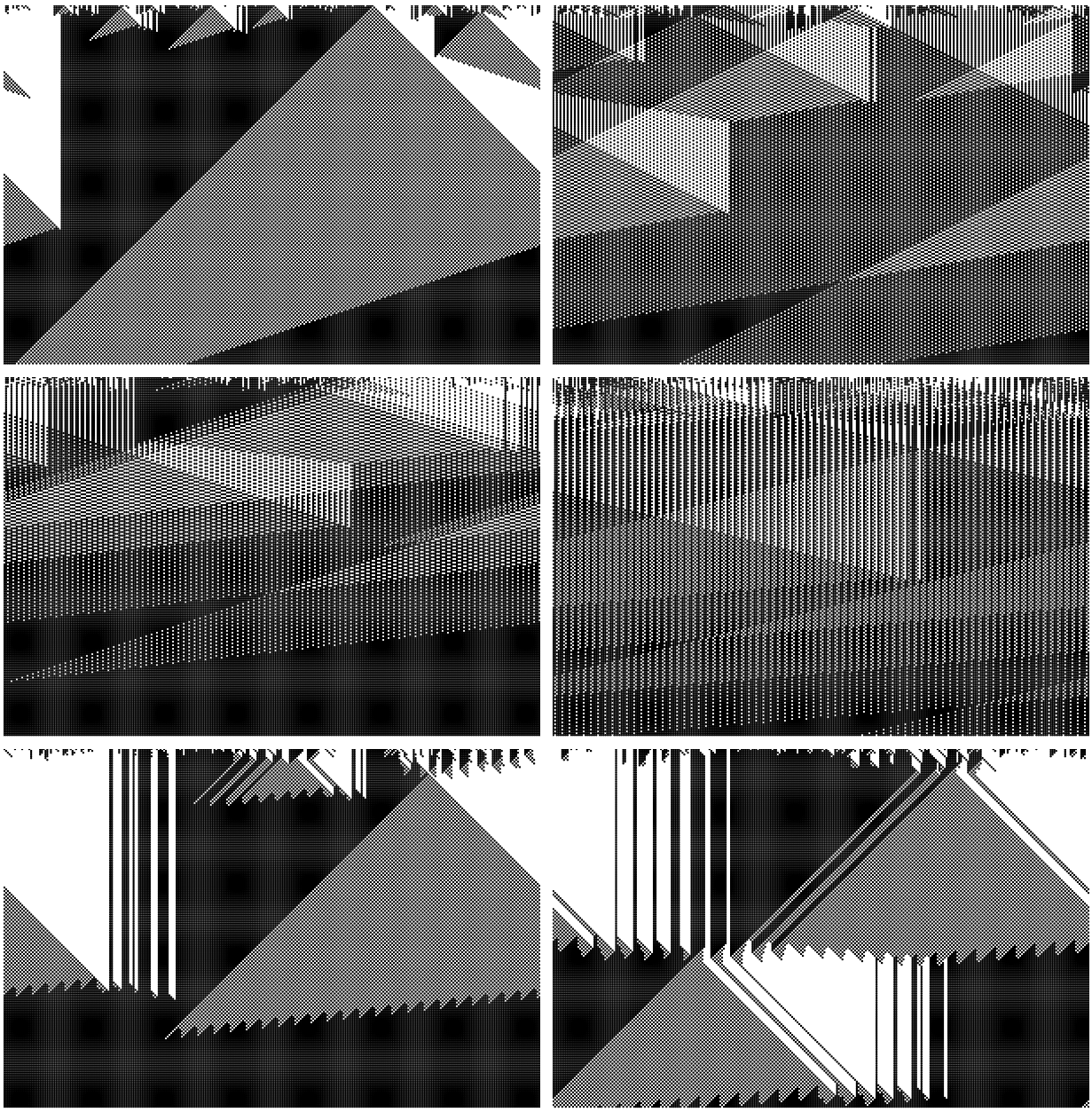

Figure 2: Space-time diagrams of GKL(j,k) with $n = 299$, $0 \leq t \leq 200$ (time flows downwards), and random initial conditions with $\rho^0 = 150/299$. From left to right, top to botton, $(j,k) = (1,3)$ (the usual GKL-II) and $(2,6)$, $(3,9)$ and $(4,12)$, $(1,9)$ and $(1,11)$. The diagrams displayed are those that reached or would eventually reach the majority state of all-1 cells (in black).