Quantum-inspired identification of complex cellular automata

[Link to publication record in Manchester Research Explorer](#)

Citing this paper
Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

OPEN ACCESS

Quantum-inspired identification of complex cellular automata

Matthew Ho,[1, 2, *] Andri Pradana,[1, †] Thomas J. Elliott,[3, 2, 1, ‡]

Lock Yue Chew,[1, 2, §] and Mile Gu[1, 2, 4, ¶]

[1] *School of Physical and Mathematical Sciences,*

Nanyang Technological University, Singapore 637371, Singapore

[2] *Complexity Institute, Nanyang Technological University, Singapore 637335, Singapore*

[3] *Department of Mathematics, Imperial College London, London SW7 2AZ, United Kingdom*

[4] *Centre for Quantum Technologies. National University of Singapore,*

3 Science Drive 2, Singapore 117543, Singapore

(Dated: March 14, 2023)

Elementary cellular automata (ECA) present iconic examples of complex systems. Though described only by one-dimensional strings of binary cells evolving according to nearest-neighbour update rules, certain ECA rules manifest complex dynamics capable of universal computation. Yet, the classification of precisely which rules exhibit complex behaviour remains somewhat an open debate. Here we approach this question using tools from quantum stochastic modelling, where quantum statistical memory – the memory required to model a stochastic process using a class of quantum machines – can be used to quantify the structure of a stochastic process. By viewing ECA rules as transformations of stochastic patterns, we ask: *Does an ECA generate structure as quantified by the quantum statistical memory, and can this be used to identify complex cellular automata?* We illustrate how the growth of this measure over time correctly distinguishes simple ECA from complex counterparts. Moreover, it provides a spectrum on which we can rank the complexity of ECA, by the rate at which they generate structure.

I. INTRODUCTION

We all have some intuition of complexity. When presented with a highly ordered periodic sequence of numbers, we can often spot a simple pattern. Meanwhile, highly disordered processes, such as a particle undergoing Brownian motion, can be described using tractable mathematical models [1, 2]. Other processes – such as stock markets, living systems, and universal computers – lack such simple descriptions and are thus considered complex; their dynamics often lie somewhere

* hosh0021@e.ntu.edu.sg

† andr0071@e.ntu.edu.sg

‡ physics@tjelliott.net

§ lockyue@ntu.edu.sg
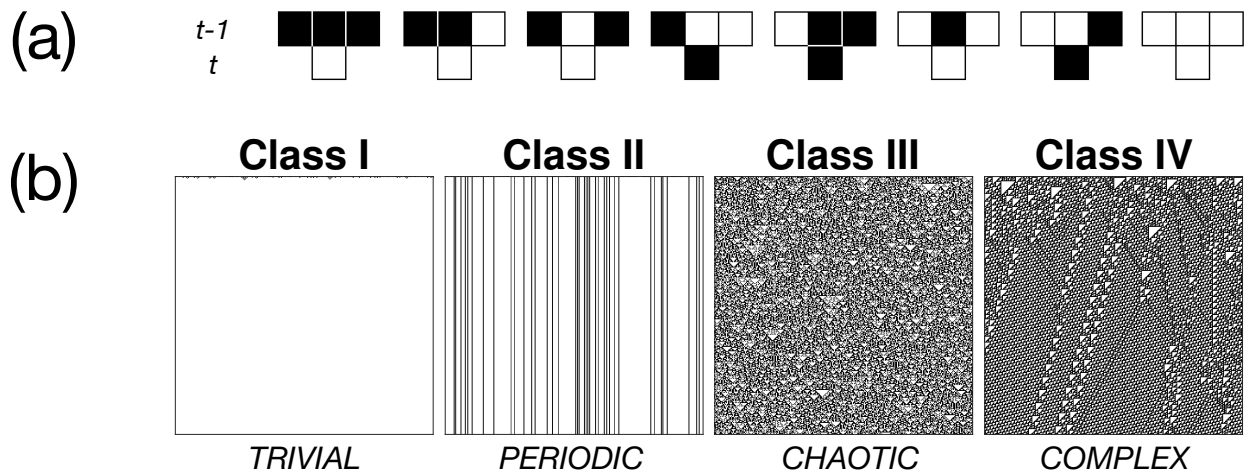
¶ mgu@quantumcomplexity.org

FIG. 1. (a) The state of each cell in an ECA takes on a binary value, and is deterministically updated at each timestep according to the current states of it and its nearest neighbours. A rule number codifies this update by converting a binary representation of the updated states to decimal. Shown here is Rule 26 $(00011010)_2$. (b) The evolution of an ECA can be visualised via a two-dimensional grid of cells, where each row corresponds to the full state of the ECA at a particular time, and columns the evolution over time. Depicted here are examples of ECA for each of Wolfram's four classes.

between order and disorder [3]. Yet, a quantitative criterion for identifying precisely what is complex remains challenging even in scenarios that appear deceptively simple.

Elementary cellular automata (ECA) provide an apt example of systems that may conceal vast complexities. Despite featuring dynamics involving only nearest-neighbour update rules on a one-dimensional chain of binary cells (see Fig. 1), they can exhibit remarkably rich behaviour. Wolfram's initial classification of their behaviour revealed both simple ECAs that were either highly ordered (e.g., periodic dynamics) or random (e.g., chaotic dynamics), as well as others complex enough to encode universal computers [4–6]. Yet, an identification and classification of which ECA rules are complex is not as clear cut as one may expect; Wolfram's initial classification was soon met with myriad of alternative approaches [7–18]. While they agree in extremal cases, the classification of many borderline ECA lacks a common consensus.

In parallel to these developments, there has also been much interest in quantifying and understanding the structure within stochastic processes. In this context, the statistical complexity has emerged as a popular candidate [19, 20]. It asks *"How much information does a model need to store about the past of a process for statistically-faithful future prediction?"*. By this measure, a completely-ordered process that generates a homogeneous sequence of 0s has no complexity, as there is only one past and thus nothing to record. A completely random sequence also has no

complexity – all past observations lead to the same future statistics, and so a model gains nothing by tracking any past information. However, modelling general processes between these extremes – such as those that are highly non-Markovian – can require the tracking of immense amounts of data, and thus indicate high levels of complexity. These favourable properties, together with the clear operational interpretation, has motivated the wide use of statistical complexity as a quantifier of structure in diverse settings [21–25].

Further developments have shown that even when modelling classical data, the most efficient models are quantum mechanical [26–32]. This has led to a quantum analogue of the statistical complexity – the quantum statistical memory – with distinct qualitative and quantitative behaviour [28, 30, 33–40] that demonstrates even better alignment with our intuitive notions of what is complex [33, 41], and a greater degree of robustness [42].

Here we ask: *Can these developments offer a new way to identify and classify complex ECAs?* To approach this, at each timestep t of the ECA's evolution we interpret the ECA state as a stochastic string, for which we can assign a quantum statistical memory cost $C_q^{(t)}$. We can then chart the evolution of this over time, and measure how the amount of information that must be tracked to statistically replicate the ECA state grows over time. Asserting that complex ECA dynamics should yield ever more complex strings, and thus enable continual growth in structure, we propose to then interpret the growth of $C_q^{(t)}$ as a measure of an ECA's complexity. Our results indicate that this interpretation has significant merit: ECA that were unanimously considered to be complex in prior studies exhibit continual growth in $C_q^{(t)}$, while those unanimously considered simple do not. Meanwhile, its application to more ambiguous cases offers a new perspective of their relative complexities, and indicates that certain seemingly-chaotic ECA may possess some structure.

II. BACKGROUND

A. Stochastic Processes, Models, and Structure

To formalise our approach, we first need to introduce some background on stochastic processes and measures of structure. A bi-infinite, discrete-time stochastic process is described by an infinite series of random variables Y_i, where index i denotes the timestep. A consecutive sequence is denoted by $Y_{l:m} := Y_l Y_{l+1} \ldots Y_{m-1}$, such that if we take 0 to be the present timestep, we can delineate a past $\overleftarrow{Y} := \lim_{L\to\infty} Y_{-L:0}$ and future $\overrightarrow{Y} := \lim_{L\to\infty} Y_{0:L}$. Associated with this is a set

of corresponding variates y_i, drawn from a distribution $P(\overleftarrow{Y}, \overrightarrow{Y})$. A stationary process is one that is translationally invariant, such that $P(Y_{0:\tau}) = P(Y_{k:\tau+k}) \forall \tau, k \in \mathcal{Z}$.

To quantify structure within such stochastic processes, we make use of computational mechanics [19, 20] – a branch of complexity science. Consider a model of a stochastic process that uses information from the past to produce statistically-faithful future outputs. That is, for any given past $\overleftarrow{y}$, the model must produce a future $\overrightarrow{y}$, one step at a time, according to the statistics of $P(\overrightarrow{Y}|\overleftarrow{y})$. Since storing the entire past is untenable, operationally, this requires a systematic means of encoding each $\overleftarrow{y}$ into a corresponding memory state $S_{\overleftarrow{y}}$, such that the model can use its memory to produce outputs according to $P(\overrightarrow{Y}|S_{\overleftarrow{y}}) = P(\overrightarrow{Y}|\overleftarrow{y})$. Computational mechanics then ascribes the complexity of the process to be the memory cost of the simplest model, i.e., the smallest amount of information a model must store about the past to produce statistically-faithful future outputs. This memory cost is named the *statistical complexity* C_μ.

In the space of classical models, this minimal memory cost is achieved by the ε-machine of the process. They are determined by means of an equivalence relation $\overleftarrow{y} \sim \overleftarrow{y}' \iff P(\overrightarrow{Y}|\overleftarrow{y}) = P(\overrightarrow{Y}|\overleftarrow{y}')$, equating different pasts iff they have coinciding future statistics. This partitions the set of all pasts into a collection of equivalence classes $\mathcal{S}$, called causal states. An ε-machine then operates with memory states in one-to-one correspondence with these equivalence classes, with an encoding function ε that maps each past to a corresponding causal state $S_j = \varepsilon(\overleftarrow{y})$. The resulting memory cost is given by the Shannon entropy over the stationary distribution of causal states:

$$C_\mu = H[P(S_j)] = -\sum_j P(S_j) \log_2 P(S_j), \tag{1}$$

where $P(S_j) = \sum_{\overleftarrow{y}|\varepsilon(\overleftarrow{y})=S_j} P(\overleftarrow{y})$.

When the model is quantum mechanical, further reduction of the memory cost is possible, by mapping causal states to non-orthogonal quantum memory states $S_j \rightarrow |\sigma_j\rangle$. These quantum memory states are defined implicitly through an evolution operator, such that the action of the model at each timestep is to produce output y with probability $P(y|S_j)$, and update the memory state [29, 31]. Specifically, this evolution is given by

$$U|\sigma_j\rangle|0\rangle = \sum_y \sqrt{P(y|S_j)}|\sigma_{\lambda(y,j)}\rangle|y\rangle \tag{2}$$

where U is a unitary operator also implicitly defined by this expression. Here, $\lambda(y, j)$ is a deterministic update function that updates the memory state to that corresponding to the causal state of the updated past. Sequential application of U then replicates the desired statistics (see Fig. 2).
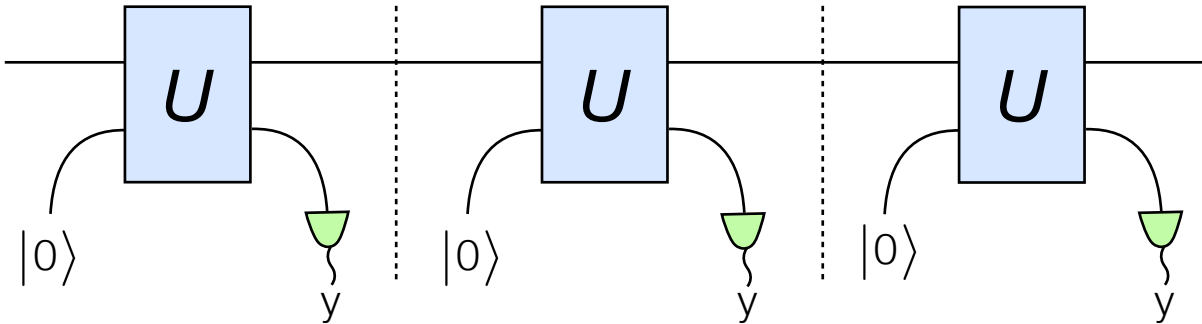
FIG. 2. A quantum model consists of a unitary operator U acting on a memory state $|\sigma_j\rangle$ and blank ancilla $|0\rangle$. Measurement of the ancilla produces the output symbol, with the statistics of the modelled process realised through the measurement statistics.

The memory cost of such quantum models is referred to as the *quantum statistical memory* [43]. Paralleling its classical counterpart, this is given by the von Neumann entropy of the steady-state of the quantum model's memory $\rho = \sum_j P(S_j)|\sigma_j\rangle\langle\sigma_j|$:

$$C_q = -\text{Tr}(\rho \log_2 \rho). \tag{3}$$

The non-orthogonality of the quantum memory states ensures that in general $C_q < C_\mu$ [26], signifying that the minimal past information needed for generating future statistics – if all methods of information processing are allowed – is generally lower than C_μ. This motivates C_q as an alternative means of quantifying structure [26, 33, 34, 44], where it has shown stronger agreement with intuitions of complexity [33, 41]. In addition to this conceptual relevance, the continuity of the von Neumann entropy makes C_q much more well-behaved compared to C_μ, such that a small perturbation in the underlying stochastic process leads to only a small perturbation in C_q, which becomes particularly relevant when inferring complexity from a finite sample of a process [42].

Since a quantum model can be systematically constructed from the ε-machine of a process, a quantum model and associated C_q can be inferred from data by first inferring the ε-machine. However, the quantum model will then inherit errors associated with the classical inference method, such as erroneous pairing or separation of pasts into causal states. For this reason, a quantum-specific inference protocol was recently developed [42] that bypasses the need to first construct an ε-machine, thus circumventing some of these errors. It functions by scanning through the stochastic process in moving windows of size $L + 1$, in order to estimate the probabilities $P(Y_{0:L+1})$, from which the marginal and conditional distributions $P(Y_{0:L})$ and $P(Y_0|Y_{-L:0})$ can be determined.

From these, we construct a set of inferred quantum memory states $\{|\varsigma_{y_{-L:0}}\rangle\}$, satisfying

$$U|\varsigma_{y_{-L:0}}\rangle|0\rangle = \sum_{y_0} \sqrt{P(y_0|y_{-L:0})}|\varsigma_{y_{-L+1:1}}\rangle|y_0\rangle. \tag{4}$$

for some suitable unitary operator U. When L is greater than or equal to the Markov order of the process, and the probabilities used are exact, this recovers the same quantum memory states as the exact quantum model Eq. (2), where the quantum memory states associated to two different pasts are identical iff the pasts belong to the same causal state. Otherwise, if L is sufficiently long to provide a 'good enough' proxy for the Markov order, and the data stream is long enough for accurate estimation of the $L+1$-length sequence probabilities, then the quantum model will still be a strong approximation with a similar memory cost. From the steady-state of these inferred quantum memory states, the quantum statistical memory C_q can be inferred [42].

However, the explicit quantum model need not be constructed as part of the inference of the quantum statistical memory. The spectrum of the quantum model steady-state is identical to that of its Gram matrix [45]. For the inferred quantum model, this Gram matrix is given by

$$G_{y_{-L:0}y'_{-L:0}} = \sqrt{P(y_{-L:0})P(y'_{-L:0})} \sum_{y_{0:L}} \sqrt{P(y_{0:L}|y_{-L:0})P(y_{0:L}|y'_{-L:0})}. \tag{5}$$

The associated conditional probabilities $P(Y_{0:L}|Y_{-L:0})$ can either be estimated from compiling the $P(Y_0|Y_{-L:0})$ using L as a proxy for the Markov order, or directly by frequency counting of strings of length of $2L$ in the data stream. Then, the quantum inference protocol yields an estimated quantum statistical memory $\tilde{C}_q$:

$$\tilde{C}_q = -\mathrm{Tr}(G \log_2 G). \tag{6}$$

B. Classifying Cellular Automata

The state of an ECA can be represented as an infinite one-dimensional chain of binary cells that evolve dynamically in time. At timestep t, the states of the cells are given by $x_i^t \in \mathcal{A} = \{0,1\}$, where i is the spatial index. Between each timestep, the states of each cell evolve synchronously according to a local update rule $x_i^t = \mathcal{F}(x_{i-1}^{t-1}, x_i^{t-1}, x_{i+1}^{t-1})$. There are $2^{2^3} = 256$ different possible such rules, each defining a different ECA [4]. A rule number is specified by converting the string of possible outcomes for each past configuration from binary into decimal as illustrated in Fig. 1(a). After accounting for symmetries (relabelling 0 and 1, and mirror images), 88 independent rules remain. Each of these rules can yield very different behaviour, motivating their grouping into

classes. One such popular classification is that of Wolfram [4, 5], describing four distinct classes (see Fig. 1(b)):

Class I: *Trivial.* The ECA evolves to a uniform state.

Class II: *Periodic.* The ECA evolves into a stable or periodic state.

Class III: *Chaotic.* The ECA evolves into a seemingly random state with no discernable structure.

Class IV: *Complex.* The ECA forms structures in its state that can interact with each other.

Classes I, II and III are all considered simple, in the sense their future behaviour is statistically easy to describe. However, Class IV ECAs are deemed complex, as they enable highly non-trivial forms of information processing, including some capable of universal computation [6].

Consider for example, two extremal cases that typify simplicity and complexity respectively:

Rule 30: Used for generating pseudo-random numbers [46, 47], it is an iconic example of a chaotic Class III ECA.

Rule 110: Proven capable of universal computation by encoding information into 'gliders' [6], it is the iconic example of a complex Class IV ECA.

Yet the boundaries between classes, especially that of Classes III and IV, still lacks universal consensus. This has motivated several diverse methods to better classify ECA, including the analysis of power spectra [8, 16, 17], Langton parameters [7, 10, 12–14], filtering of space-time dynamics [15], hierarchical classification [11], computing the topology of the Cantor space of ECA [48–51], and mean field theory [9]. These schemes feature discrete classes but do not yield identical conclusions, highlighting the inherent difficulties in discerning complexity and randomness.

Illustrative examples of ECA with ambiguous classification include:

Rule 54: This rule exhibits interacting glider systems like Rule 110 [21, 52, 53], but its capacity for universal computation is unknown.

Rule 18: Assigned by Wolfram to Class III, though subsequent studies indicate it contains anomalies known as kinks that can propagate and annihilate other kinks [54, 55]. This indicates some level of structure.

Ambiguous rules such as these complicate attempts that seek to determine a border between simple and complex ECA. Rather, we may instead envision trying to place such ECA on a spectrum, with one end having ECA that are clearly in Class III, and the other ECA clearly in Class IV,

with the more ambigious cases on a gradated scale in between. To our knowledge, ours is the first approach that refines ECA classification into such a hybrid mix of discrete classes with a continuous spectrum.

III. A STOCHASTIC PERSPECTIVE

A. Overview and Analytical Statements

We will use the tools of computational mechanics to classify ECA, placing them on such a spectrum as described above. To do so, we first need to describe ECA in the language of stochastic processes. Observe that if we initialise an ECA at random, then the state of the ECA at $t = 0$ is described by a spatially-stationary stochastic process $\overleftrightarrow{Y}^{(0)}$ – specifically, a completely random process which has $C_q^{(0)} = 0$. The evolution of the ECA transforms this into a new stochastic process; let $\overleftrightarrow{Y}^{(t)}$ describe the stochastic process associated with the ECA state at timestep t, where the spatial index of the state takes the place of the temporal index in the process. As the update rules for ECA are translationally invariant, the $\overleftrightarrow{Y}^{(t)}$ are also spatially-stationary, and thus possess a well-defined complexity measure $C_q^{(t)}$. We emphasise that the evolution of the ECA remains deterministic; the stochasticity in the processes extracted from the ECA state emerges from a scanning of successive cells in the ECA state at a given point in time without knowledge of any prior state of the ECA (including its initial, random, state).

Charting the evolution of $C_q(t)$ with successive applications of the update rule can then help us quantify the amount of structure created or destroyed during the evolution of an ECA. We propose the following criterion:

An ECA is complex if $C_q^{(t)}$ grows with t without bound.

The rationale is that complex dynamics should be capable of generating new structure. Simple ECA are thus those that generate little structure, and correspondingly have $C_q^{(t)}$ stagnate. At the other extreme, ECA capable of universal computation can build up complex correlations between cells arbitrarily far apart as t grows, requiring us to track ever-more information; thus, $C_q^{(t)}$ should grow.

We can immediately make the following statements regarding all ECA in Classes I and II, and some in Class III – highlighting that our criterion also identifies simple ECA.

(1) $\lim_{t\to\infty} C_q^{(t)} = 0$ for all Class I ECA.

(2) $C_q^{(t)}$ must be asymptotically bounded for all Class II ECA. That is, there exists a sufficiently large T and constant K such that $C_q^{(t)} \leq K$ for all $t > T$.

(3) $C_q^{(t)} \approx 0$ for ECA suitable for use as near-perfect pseudo-random number generators.

Statement (1) follows from the definition that Class I ECA are those that evolve to homogeneous states; at sufficiently large t, $\overleftrightarrow{y}^{(t)}$ is a uniform sequence of either all 0s or all 1s, for which $C_q^{(t)} = 0$ (since there is only a single causal state). With Class II ECA defined as those that are eventually periodic with some periodicity τ (modulo spatial translation, which does not alter C_q), this implies that their $C_q^{(t)}$ similarly oscillates with at most the same periodicity τ, with some maximal value bounded by some constant K, thus statement (2) follows. Finally to see statement (3), we note that a perfect random number generator should have no correlations or structure in their outputs, and hence have $C_q^{(t)} = 0$; the continuity of C_q then ensures that near-perfect generators should have $C_q^{(t)} \approx 0$.

B. Numerical Methodology

For the ECA that lie between Class III and IV of interest to us here, we must deal with their potential computational irreducibility. There are thus generally no shortcuts to directly estimate the long-time behaviour of $C_q^{(t)}$ in such ECA. We therefore employ numerical simulation. We approximate each ECA of interest with a finite counterpart, where each time-step consists of $W = 64,000$ cells, such that the state of the ECA at time t is described by a finite data sequence $y_{0:W}^{(t)}$, up to a maximum time $t_{\max} = 10^3$ [56]. We can then infer the quantum statistical memory $C_q^{(t)}$ of this sequence, using the quantum inference protocol discussed above [42], with $L = 6$ for the history length. The workflow is illustrated in Fig. 3.

For each ECA rule we generate an initial state for $t = 1$ where each cell is randomly assigned 0 or 1 with equal probability, and then evolve for $t_{\max}$ steps. We then apply the inference methods to the states at $t = 1, 2, 3, ..., 9, 10, 20, ..., 90, 100, 200, ..., t_{\max}$; evaluating at every timestep shows little qualitative difference beyond highlighting the short periodicity of some Class II rules. We repeat five times for each rule, and determine the mean and standard deviation of $C_q^{(t)}$. To avoid boundary effects from the edges of the ECA, to obtain an ECA state of width W for up to $t_{\max}$ timesteps we generate an extended ECA of width $W' = W + 2t_{\max}$ with periodic boundary conditions and keep only the centremost W cells; this is equivalent to generating a width W ECA with open boundaries
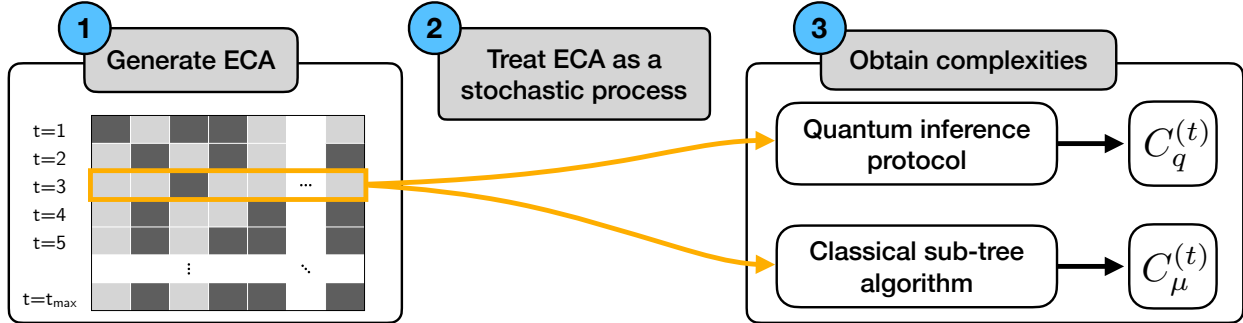
FIG. 3. An ECA is evolved from random initial conditions. Treating the ECA state at each timestep as a stochastic process, we then infer the quantum statistical memory $C_q^{(t)}$ and classical statistical complexity $C_\mu^{(t)}$. By observing how these measures change over time, we are able to deduce the complexity of the ECA rule.
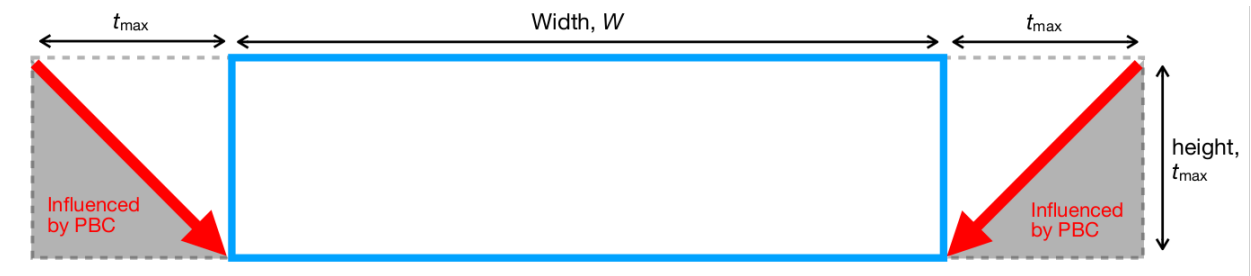

FIG. 4. Generation of finite-width ECA evolution with open boundaries via extended ECA with periodic boundaries.

(see Fig. 4). Note however that the choice of boundary condition showed little quantitative effect upon our results.

In Fig. 5) we present plots showing little difference in the qualitative features of interest with extensions up to $L = 8$ and $t_{max} = 10^5$, supporting our choice for these values. The exception to this is Rule 110, which appears to plateau at longer times. We believe this to be attributable to the finite width of the ECA studied – as there are a finite number of gliders generated by the initial configuration, over time as the gliders annihilate there will be fewer of them to interact and propagate further correlations.

For completeness, we also perform analogous inference for the classical statistical complexity $C_\mu^{(t)}$ using the sub-tree reconstruction method. See Appendix A for details.
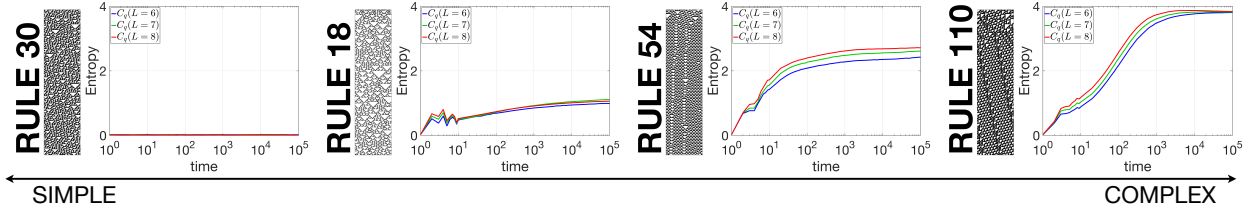
FIG. 5. $C_q^{(t)}$ plots for a selection of rules with longer L and larger $t_{\max}$. Plots shown for $(W = 64,000, L = 6)$, $(W = 128,000, L = 7)$, and $(W = 256,000, L = 8)$.
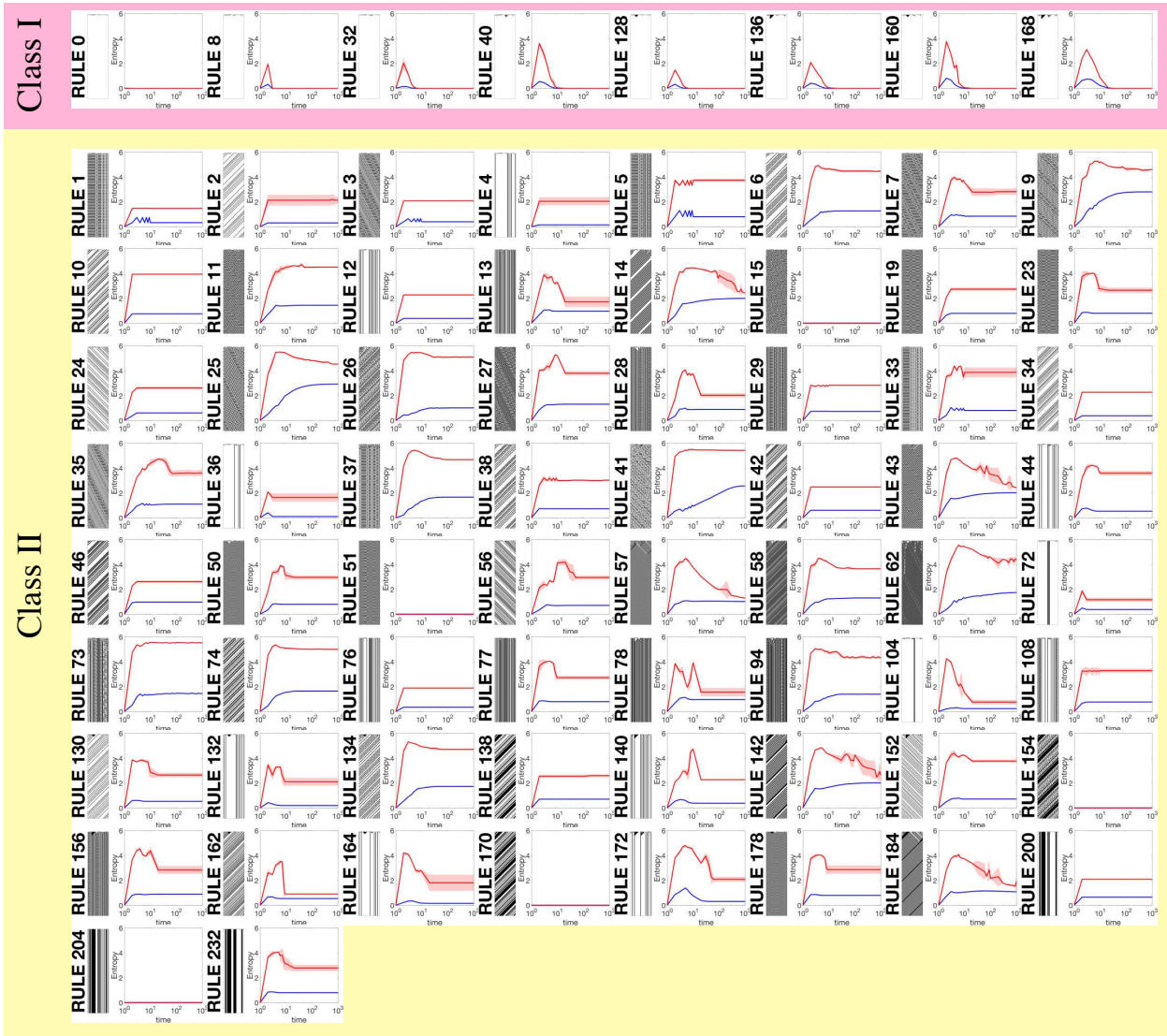

FIG. 6. Evolution of $C_q^{(t)}$ (blue) and $C_\mu^{(t)}$ (red) for all Wolfram Class I and II rules. Lines indicate mean values over five different intial random states, and the translucent surrounding the standard deviation.
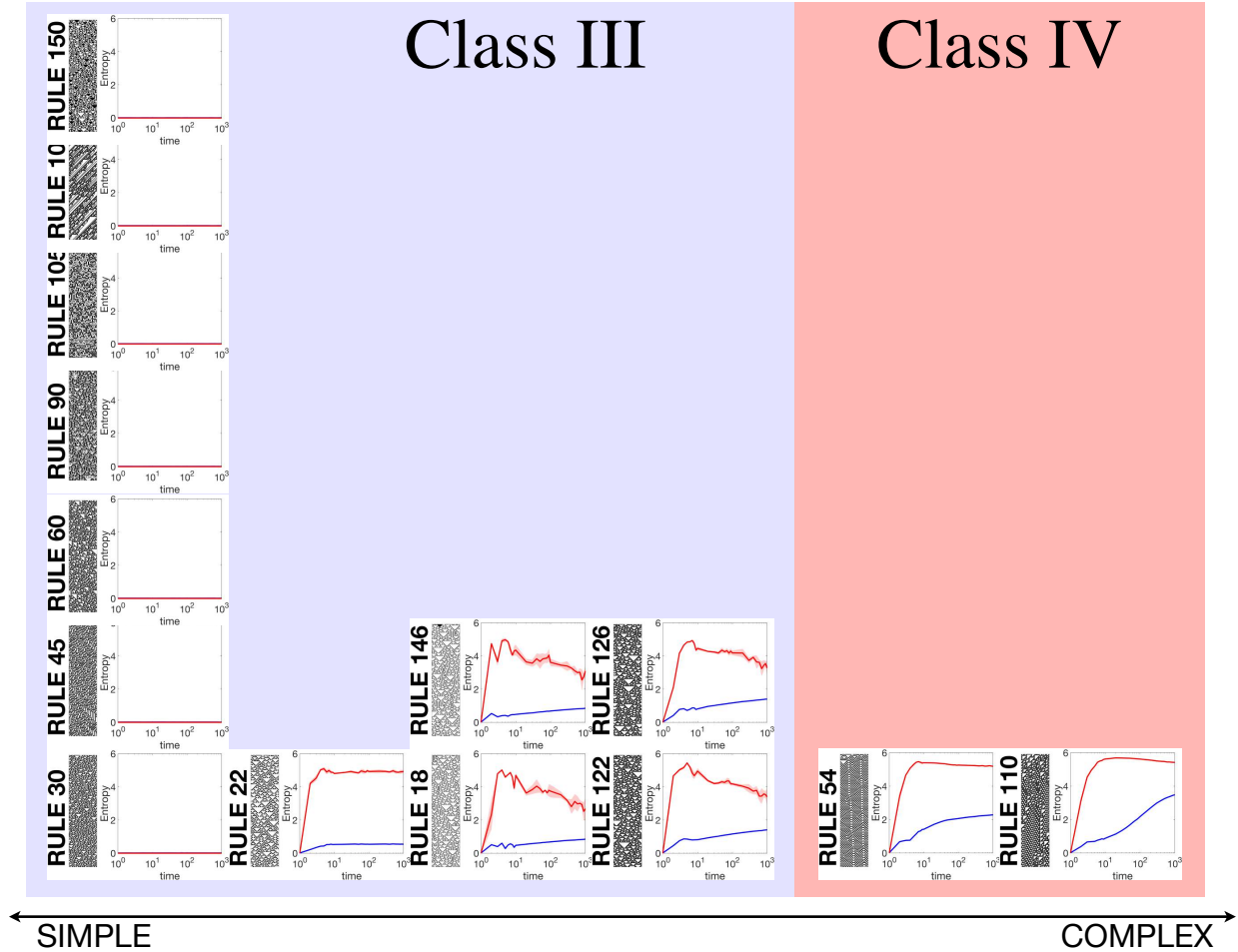
FIG. 7. Evolution of $C_q^{(t)}$ (blue) and $C_\mu^{(t)}$ (red) for all Wolfram Class III and IV rules. Rules are placed on a simplicity-complexity spectrum according to the growth of $C_q^{(t)}$ with Rule 30 (Class III, simple) and Rule 110 (Class IV, complex) at the extremes. Lines indicate mean values over five different intial random states, and the translucent surrounding the standard deviation.

C. Numerical Results

Our results are plotted in Fig. 6 and Fig. 7, charting the evolution of $C_q^{(t)}$ and $C_\mu^{(t)}$ for all unique ECA rules. Fig. 6 shows all ECA rules that belong to Wolfram Classes I and II; we see that $C_q^{(t)}$ indeed displays the features discussed above, i.e., that $\lim_{t\to\infty} C_q^{(t)} = 0$ for Class I ECA, and tends to a bounded value for Class II. Wolfram Class III and IV rules are displayed in Fig. 7, where they are ranked on a simplicity-complexity spectrum according to the rate of growth of $C_q^{(t)}$.

We first observe that our extremal cases indeed sit at the extremes of this spectrum:

Rule 30: Consistent with its role as a pseudo-random number generator, Rule 30 generates no discernable structure, yielding negligible $C_q^{(t)} \approx 0$.

Rule 110: Clearly exhibits the fastest growth in $C_q^{(t)}$. This aligns with its capability for universal computation; Rule 110 is able to propagate correlations over arbitrarily large distances.

More interesting are the rules with a more ambiguous classification. We make the following observations of illustrative examples, listed in order of increasing growth rates on $C_q^{(t)}$.

Rule 22: We see that $C_q^{(t)}$ stagnates after a short period of initial growth, and thus behaves as a simple CA according to our complexity criterion. Indeed, prior studies of this rule suggests that it behaves as a mix of a random and a periodic process, generating a finite amount of local structure that does not propagate with time, and is thus very unlikely to be capable of universal computation [57].

Rule 18: $C_q^{(t)}$ grows slowly (at around a quarter the rate of Rule 110), suggesting the presence of some complex dynamics within despite a Class III classification. Indeed, the presence of kinks in Rule 18's dynamics supports this observation.

Rule 122: $C_q^{(t)}$ shows similar growth to Rule 18, though slightly faster.

Rule 54: Aside from Rule 110, Rule 54 is the only other Wolfram Class IV rule. Consistent with this, we see that it exhibits a fast growth in $C_q^{(t)}$, likely due to its glider system propagating correlations.

Thus among these examples, all except Rule 22 still feature some degree of growth in $C_q^{(t)}$. We conclude that these ECA, despite their Class III classification, do feature some underlying complex dynamics. Moreover, we see that our spectrum of the relative growth rates of $C_q^{(t)}$ appears to offer a suitable means of ranking the relative complexity of each ECA. Particularly, our spectrum is consistent with the division into Class III and IV, but provides a further nuance than the traditional discrete classes.

We remark that on the other hand, the classical $C_\mu^{(t)}$ does not appear to provide such a clear indicator of complexity in the ECA, with some Class II-IV rules even showing a reduction in complexity over time. Moreover, the instability of the measure is evident. Together, this highlights the advantages of using the quantum measure $C_q^{(t)}$ to quantify complexity, over its classical counterpart.

We further remark that while $C_q^{(t)}$ has performed admirably for this purpose, it may not be unique in this regard. There exist a multitude of measures of complexity, of varying utility and areas of application. It is of course not possible for us to explore all such possibilities, but we will briefly

here comment on some of the more commonly-used measures. Perhaps the most famous measure of 'complexity' is the algorithmic information [58–60], that quantifies the shortest programme required to output a particular string. However, this is now recognised to be better suited as a measure of randomness rather than complexity, and is moreover, generically uncomputable. Similarly, the thermodynamic depth [61] has been linked with randomness rather than complexity, and suffers from ambiguities in its proper calculation [62]. However, one potentially suitable alternative is the excess entropy E, that quantifies the mutual information between the past and future outputs of a stochastic process [20], and is a lower bound on C_q and C_μ. We see from Fig. 9 that E may also work in place of C_q here as a suitable complexity measure to identify complex ECA. Indeed, qualitative similarities between C_q and E have previously been noted [41]. We leave a more complete analysis of the qualitative similarities and difference of C_q and E – in this context and beyond – as a question for future work.

IV. DISCUSSION

Here, we introduced new methods for probing the complexity of ECA. By viewing the dynamics of a one-dimensional cellular automata at each time-step as a map from one stochastic process to another, we are able to quantify the structure of the ECA state at each timestep. Then, by initialising an ECA according to a random process with no structure, we can observe if the ECA's update rules are able to transduce to stochastic processes with increasing structure. In this picture an ECA is considered simple if the structure saturates to a bounded value over time, and is complex if it exhibits continued growth. To formalise this approach, we drew upon computational mechanics, which provides a formal means of quantifying structure within stochastic process as the memory cost needed to simulate them. We found that the memory cost associated with quantum models – the quantum statistical memory C_q – performed admirably at identifying complex ECA. It provides agreement with prior literature when there is concensus on the ECA's complexity, and is able to place the remaining ECA of ambiguous complexity on a simplicity-complexity spectrum.

One curiosity of this proposal is the unlikely juxtaposition of using a measure of complexity involving quantum information, when the systems involved (ECA) are purely classical objects. One rationale would be mathematical practicality – the continuity properties of C_q make it more stable, while its slower scaling makes it less susceptible to saturating numerical limitations. On the other hand, the use of quantum measures for classical objects is not entirely alien; quantum computers are expected to solve certain classical computational problems more efficiently than

classical computers, and one may thence argue that the true complexity of a process should account for all physical means of its simulation.

Our results open some interesting queries. Several of the ECA our methodology identified as complex lie within Wolfram Class III, suggesting that some ECA many considered to feature only randomness may actually be capable of more complex information processing tasks. However, for only one of these (Rule 18) was a means to do this previous identified [54, 55]. Could the other ECA with equal or higher $C_q^{(t)}$ growth, such as Rule 122, also feature such dynamics when properly filtered? More generally, it appears intuitive that universal computation and continual $C_q^{(t)}$ growth should be related, but can such a relation be formalised? Finally, while our studies here have focused on ECA, the methdology naturally generalises to any one-dimensional cellular automata – such as those with more cell states, longer-range update rules, asynchronous tuning [63], and memory [64]; it would certainly be interesting to see if our framework can provide new insight into what is complex in these less-extensively studied situations.

DATA AVAILABILITY

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

Appendix A: Sub-tree reconstruction algorithm

Here, inference of the classical statistical complexity C_μ is achieved through the sub-tree reconstruction algorithm [19]. It works by explicitly building an ε-machine of a stochastic process, from which C_μ may readily be deduced. The steps are detailed below.

1. Constructing a tree structure. The sub-tree construction begins by drawing a blank node to signify the start of the process with outputs $y \in \mathcal{A}$. A moving window of size $2L$ is chosen to parse through the process. Starting from the blank node, $2L$ successive nodes are created with a directed link for every y in each moving window $\{y_{0:2L}\}$. For any sequence starting from y_0 within $\{y_{0:2L}\}$ whose path can be traced with existing directed links and nodes, no new links and nodes are added. New nodes with directed links are added only when the $\{y_{0:2L}\}$ does not have an existing path. This is illustrated in Fig. 8

For example, suppose $y_{0:6} = 000000$, giving rise to six nodes that branch outwards in serial from the initial blank node. If $y_{1:7} = 000001$, the first five nodes gain no new branches, while the sixth node gains a new branch connecting to a new node with a directed link. Each different element of $|\mathcal{A}|^{2L}$ has its individual set of directed links and nodes, allowing a maximum of $|\mathcal{A}|^{2L}$ branches that originate from the blank node.
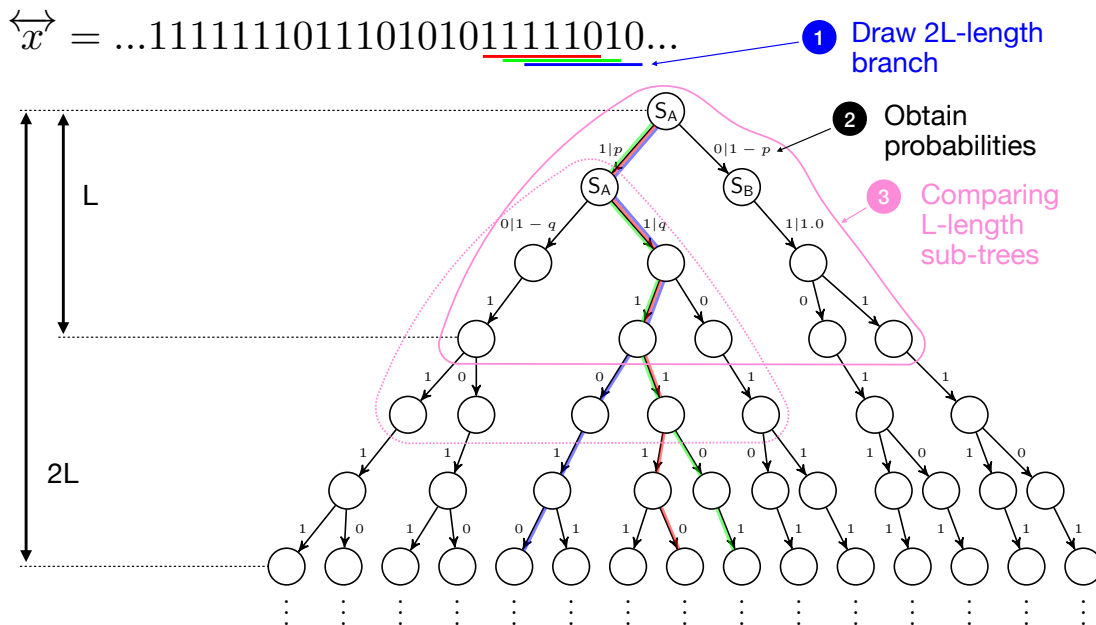

FIG. 8. The sub-tree reconstruction algorithm, here illustrated for $L = 3$.

2. Assigning probabilities. The probability for each branch from the first node to occur can be determined by the ratio of the number of occurences the associated strings to the total number of strings. Correspondingly, this allows each link to be denoted with an output y with its respective transition probability p.

3. Sub-tree comparison. Next, starting from the initial node, the tree structure of L outputs is compared against all other nodes. Working through all reachable L nodes from the initial node,

any nodes with identical $y|p$ and branch structure of L size are given the same label. Because of finite data and finite L, a χ^2 test is used to account for statistical artefacts. The χ^2 test will merge nodes that have similar-enough tree structures. This step essentially enforces the causal equivalence relation on the nodes.

4. Constructing the ε-machine. It is now possible to analyse each individually-labelled node with their single output and transition probability to the next node. An edge-emitting hidden Markov model of the process can then be drawn up. This edge-emitting hidden Markov model represents the (inferred) ε-machine of the process.

5. Computing the statistical complexity. The hidden Markov model associated with the ε-machine has a transition matrix T_{kj}^y giving the probability of the next output being y given we are in causal state S_j, and S_k being the causal state of the updated past. The steady-state of this (i.e., the eigenvector π satisfying $\sum_y T^y \pi = \pi$) gives the steady-state probabilities of the causal states. Taking $P(S_j) = \pi_j$, we then have the Shannon entropy of this distribution gives the statistical complexity:

$$C_\mu := H[P(S_j)] = -\sum_j P(S_j) \log_2 P(S_j). \tag{A1}$$

For parity with the quantum inference protocol, we used $L = 6$ when inferring the $C_\mu^{(t)}$ of the ECA states, and set the tolerance of the χ^2 test to 0.05.

<div align="center">

Appendix B: Convergence with L

</div>

We here provide an expanded form of Fig. 5 in Fig. 9 regarding the convergence of our measures of complexity with increasing probability window lengths L, now including $L = 3 : 8$, and also showing analogous plots for C_μ and E alongside C_q. We note that the excess entropy E is estimated for finite length windows using $E(L) = LH(X_{0:L-1}) - (L-1)H(X_{0:L})$.
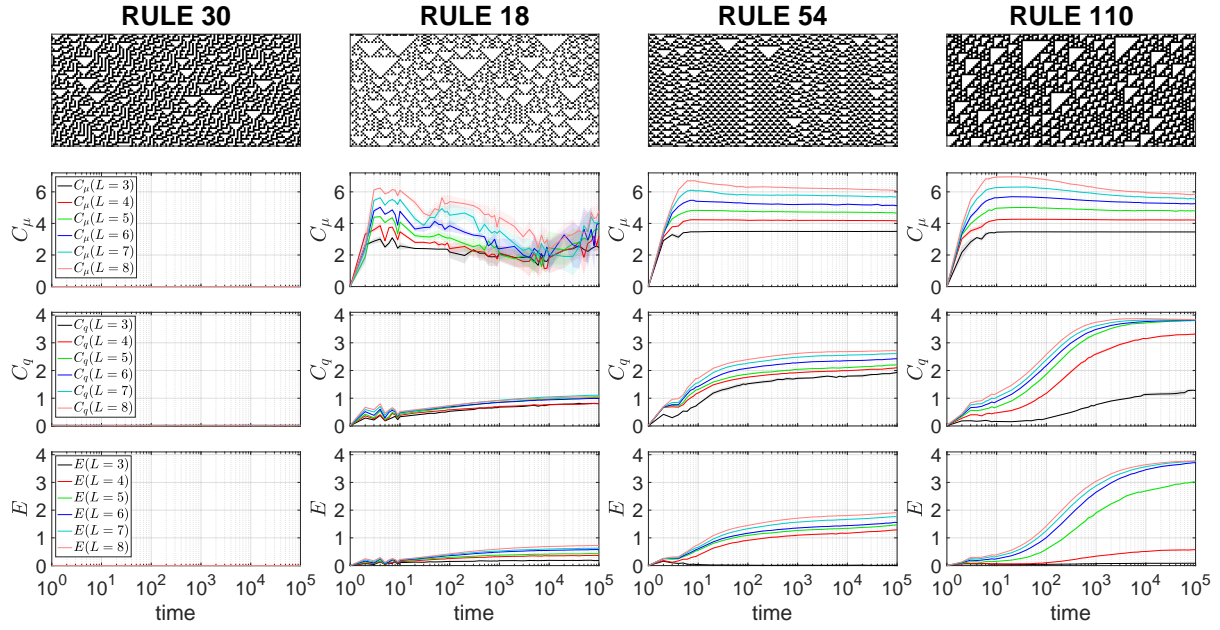
FIG. 9. Growth of L-length ($L = 3 : 8$) estimates of C_μ, C_q, and E with time, up to a maximum of $t_{\max} = 10^5$ time steps. Different L correspond to different widths $W = 10^3 * 2^L$.

[1] J. M. Deutch and I. Oppenheim, The Lennard-Jones lecture: The concept of Brownian motion in modern statistical mechanics, Faraday Discussions of the Chemical Society **83**, 1 (1987).

[2] L. Chew and C. Ting, Microscopic chaos and gaussian diffusion processes, Physica A: Statistical Mechanics and its Applications **307**, 275 (2002).

[3] J. P. Crutchfield, Between order and chaos, Nature Physics **8**, 17 (2011).

[4] S. Wolfram, Universality and complexity in cellular automata, Physica D: Nonlinear Phenomena **10**, 1 (1984).

[5] S. Wolfram, Cellular automata as models of complexity, Nature **311**, 419 (1984).

[6] M. Cook, Universality in Elementary Cellular Automata, Complex Systems **15**, 1 (2004).

[7] C. G. Langton, Studying artificial life with cellular automata, Physica D: Nonlinear Phenomena **22**, 120 (1986).

[8] W. Li, Power Spectra of Regular Languages and Cellular Automata, Complex Systems **1**, 107 (1987).

[9] H. A. Gutowitz, J. D. Victor, and B. W. Knight, Local structure theory for cellular automata, Physica D: Nonlinear Phenomena **28**, 18 (1987).

[10] W. Li and N. H. Packard, The Structure of the Elementary Cellular Automata Rule Space, Complex Systems **4**, 281 (1990).

[11] H. A. Gutowitz, A hierarchical classification of cellular automata, Physica D: Nonlinear Phenomena **45**, 136 (1990).

[12] W. Li, N. H. Packard, and C. G. Langton, Transition Phenomena In Cellular Automata Rule Space, Physica D: Nonlinear Phenomena **45**, 77 (1990).

[13] C. G. Langton, Computation at the edge of chaos: Phase transitions and emergent computation, Physica D: Nonlinear Phenomena **42**, 12 (1990).

[14] P. M. Binder, A phase diagram for elementary cellular automata, Complex Systems **7**, 241 (1993).

[15] A. Wuensche, Classifying cellular automata automatically: Finding gliders, filtering, and relating space-time patterns, attractor basins, and the Z parameter, Complexity **4**, 47 (1999).

[16] S. Ninagawa, Power Spectral Analysis of Elementary Cellular Automata, Complex Systems **17**, 399 (2008).

[17] E. L. P. Ruivo and P. P. B. de Oliveira, A Spectral Portrait of the Elementary Cellular Automata Rule Space, in *Irreducibility and Computational Equivalence: 10 Years After Wolfram's A New Kind of Science*, edited by H. Zenil (Springer Berlin Heidelberg, Berlin, Heidelberg, 2013) pp. 211–235.

[18] G. J. Martinez, A note on elementary cellular automata classification, Journal of Cellular Automata **8**, 233 (2013).

[19] J. P. Crutchfield and K. Young, Inferring statistical complexity, Physical Review Letters **63**, 105 (1989).

[20] C. R. Shalizi and J. P. Crutchfield, Computational mechanics: Pattern and prediction, structure and simplicity, Journal of Statistical Physics **104**, 817 (2001).

[21] J. E. Hanson and J. P. Crutchfield, Computational mechanics of cellular automata: An example, Physica D: Nonlinear Phenomena **103**, 169 (1997).

[22] W. M. Gonçalves, R. D. Pinto, J. C. Sartorelli, and M. J. De Oliveira, Inferring statistical complexity in the dripping faucet experiment, Physica A **257**, 385 (1998).

[23] J. B. Park, J. Won Lee, J. S. Yang, H. H. Jo, and H. T. Moon, Complexity analysis of the stock market, Physica A **379**, 179 (2007).

[24] R. Haslinger, K. L. Klinkner, and C. R. Shalizi, The computational structure of spike trains, Neural Computation **22**, 121 (2010).

[25] H. N. Huynh, A. Pradana, and L. Y. Chew, The complexity of sequences generated by the arc-fractal system, PLOS ONE **10**, 1 (2015).

[26] M. Gu, K. Wiesner, E. Rieper, and V. Vedral, Quantum mechanics can reduce the complexity of classical models, Nature Communications **3**, 762 (2012).

[27] J. R. Mahoney, C. Aghamohammadi, and J. P. Crutchfield, Occam's Quantum Strop: Synchronizing and Compressing Classical Cryptic Processes via a Quantum Channel, Scientific Reports **6**, 20495 (2016).

[28] T. J. Elliott and M. Gu, Superior memory efficiency of quantum devices for the simulation of continuous-time stochastic processes, npj Quantum Information **4**, 18 (2018).

[29] F. C. Binder, J. Thompson, and M. Gu, A Practical Unitary Simulator for Non-Markovian Complex Processes, Physical Review Letters **120**, 240502 (2018).

[30] T. J. Elliott, A. J. P. Garner, and M. Gu, Memory-efficient tracking of complex temporal and symbolic dynamics with quantum simulators, New Journal of Physics **21**, 013021 (2019).

[31] Q. Liu, T. J. Elliott, F. C. Binder, C. Di Franco, and M. Gu, Optimal stochastic modeling with unitary quantum dynamics, Physical Review A **99**, 1 (2019).

[32] S. P. Loomis and J. P. Crutchfield, Strong and weak optimizations in classical and quantum models of stochastic processes, Journal of Statistical Physics **176**, 1317 (2019).

[33] W. Y. Suen, J. Thompson, A. J. P. Garner, V. Vedral, and M. Gu, The classical-quantum divergence of complexity in modelling spin chains, Quantum **1**, 25 (2017).

[34] C. Aghamohammadi, J. R. Mahoney, and J. P. Crutchfield, The ambiguity of simplicity in quantum and classical simulation, Physics Letters A **381**, 1223 (2017).

[35] A. J. P. Garner, Q. Liu, J. Thompson, V. Vedral, and M. Gu, Provably unbounded memory advantage in stochastic simulation using quantum mechanics, New Journal of Physics **19** (2017).

[36] C. Aghamohammadi, J. R. Mahoney, and J. P. Crutchfield, Extreme Quantum Advantage when Simulating Classical Systems with Long-Range Interaction, Scientific Reports **7**, 6735 (2017).

[37] J. Thompson, A. J. Garner, J. R. Mahoney, J. P. Crutchfield, V. Vedral, and M. Gu, Causal Asymmetry in a Quantum World, Physical Review X **8**, 31013 (2018).

[38] T. J. Elliott, C. Yang, F. C. Binder, A. J. P. Garner, J. Thompson, and M. Gu, Extreme Dimensionality Reduction with Quantum Modeling, Physical Review Letters **125**, 260501 (2020).

[39] T. J. Elliott, Quantum coarse graining for extreme dimension reduction in modeling stochastic temporal dynamics, PRX Quantum **2**, 020342 (2021).

[40] T. J. Elliott, M. Gu, A. J. P. Garner, and J. Thompson, Quantum adaptive agents with efficient long-term memories, Physical Review X **12**, 011007 (2022).

[41] W. Y. Suen, T. J. Elliott, J. Thompson, A. J. P. Garner, J. R. Mahoney, V. Vedral, and M. Gu, Surveying structural complexity in quantum many-body systems, Journal of Statistical Physics **187**, 1 (2022).

[42] M. Ho, M. Gu, and T. J. Elliott, Robust inference of memory structure for efficient quantum modeling of stochastic processes, Physical Review A **101**, 32327 (2020).

[43] The term 'quantum statistical memory' is used in place of 'quantum statistical complexity' as such quantum machines may not necessarily be memory-minimal among all quantum models [31, 33].

[44] R. Tan, D. R. Terno, J. Thompson, V. Vedral, and M. Gu, Towards quantifying complexity with quantum mechanics, European Physical Journal Plus **129:191** (2014).

[45] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd ed. (Cambridge University Press, 2012).

[46] S. Wolfram, Random sequence generation by cellular automata, Advances in Applied Mathematics **7**, 123 (1986).

[47] Random number generation, `https://reference.wolfram.com/language/tutorial/RandomNumberGeneration.html`, Last Accessed: 2022-08-17.

[48] M. Schüle and R. Stoop, A full computation-relevant topological dynamics classification of elementary cellular automata, Chaos: An Interdisciplinary Journal of Nonlinear Science **22**, 043143 (2012).

[49] P. Kurka, Languages, equicontinuity and attractors in cellular automata, Ergodic Theory and Dynamical Systems **17**, 417–433 (1997).

[50] P. Kurka, Topological dynamics of cellular automata, in *Codes, Systems, and Graphical Models* (Springer, 2001) pp. 447–485.

[51] R. H. Gilman, Classes of linear automata, Ergodic theory and dynamical systems **7**, 105 (1987).

[52] B. Martin, A group interpretation of particles generated by one-dimensional cellular automaton, Wolfram's rule 54, International Journal of Modern Physics C **11**, 101 (2000).

[53] G. J. Martínez, A. Adamatzky, and H. V. McIntosh, Complete characterization of structure of rule 54, Complex Systems **23**, 259 (2014).

[54] P. Grassberger, Chaos and Diffusion in Deterministic Cellular Automata, Physica D: Nonlinear Phenomena , 52 (1984).

[55] K. Eloranta and E. Nummelin, The kink of cellular automaton rule 18 performs a random walk, Journal of Statistical Physics **69**, 1131 (1992).

[56] Note that this is many, many orders of magnitude smaller than the time for which a typical finite-width ECA is guaranteed to cycle through already-visited states ($\mathcal{O}(2^W)$) [57].

[57] P. Grassberger, Long-range effects in an elementary cellular automaton, Journal of Statistical Physics **45**, 27 (1986).

[58] A. N. Kolmogorov, On tables of random numbers, Sankhyā: The Indian Journal of Statistics, Series A , 369 (1963).

[59] R. J. Solomonoff, A formal theory of inductive inference. part i, Information and control **7**, 1 (1964).

[60] G. J. Chaitin, On the simplicity and speed of programs for computing infinite sets of natural numbers, Journal of the ACM (JACM) **16**, 407 (1969).

[61] S. Lloyd and H. Pagels, Complexity as thermodynamic depth, Annals of physics **188**, 186 (1988).

[62] J. P. Crutchfield and C. R. Shalizi, Thermodynamic depth of causal states: Objective complexity via minimal representations, Physical Review E **59**, 275 (1999).

[63] D. Uragami and Y. P. Gunji, Universal emergence of 1/f noise in asynchronously tuned elementary cellular automata, Complex Systems **27**, 399 (2018).

[64] G. J. Martínez, A. Adamatzky, J. C. Seck-Tuoh-Mora, and R. Alonso-Sanz, How to make dull cellular automata complex by adding memory: Rule 126 case study, Complexity **15**, 34 (2010).