MATHEMATICAL TREATMENT OF THE CANONICAL FINITE STATE MACHINE FOR THE ISING MODEL

TRATAMIENTO MATEMÁTICO DE LA MÁQUINA CANÓNICA DE ESTADO FINITO PARA EL MODELO DE ISING

E. Estevez-Rams[a,b†], E. Rodríguez-Horta[a], R. Lora-Serrano[c]

a) Facultad de Física, Universidad de la Habana, San Lázaro y L. CP 10400. La Habana. Cuba
b) Instituto de Ciencias y Tecnología de Materiales, University of Havana (IMRE); estevez@fisica.uh.cu†
c) Universidade Federal de Uberlandia, AV. Joao Naves de Avila, 2121- Campus Santa Monica, Uberlandia, CEP 38408-144,Brazil
† corresponding author

Recibido 3/10/2024; Aceptado 26/5/2025

The complete framework for the minimal deterministic automata construction of the one-dimensional Ising model is presented. The approach follows the known treatment of the Ising model as a Markov random field, where the local characteristic is usually obtained from the stochastic matrix. The problem is the inverse relation or how to get the stochastic matrix from the local characteristics given via the transfer matrix treatment. The obtained expressions allow for performing complexity-entropy analysis of particular instances of the Ising model. Two examples are discussed: the 1/2-spin nearest neighbour and next nearest neighbours Ising model.

El marco teórico para la máquina mínima determinística del model de Ising en una dimensión es presentado. El tratamiento sigue el conocido modelo de Ising tratado como un campo aleatorio de Markov, donde las características locales son obtenidas de la matriz estocástica. El problema abordado necesita la relación inversa, o como obtener la matriz estocástica de las características locales, dadas a través del tratamiento de la matriz de transferencia. Las expresiones obtenidas permiten realizar el análisis de complejidad-entropía para instancias particulares del modelo de Ising. Dos ejemplos son discutidos: el spín-1/2 de vecinos más cercanos y el modelo de segundos vecinos más cercanos.

Keywords: complexity (complejidad); entropy (entropía); Ising model (Modelo de Ising).

I. INTRODUCTION

Minimal deterministic automata, introduced by Grassberger [1] and further developed by Crutchfield et al. [2, 3], is an approach to discovering and characterizing patterns in an information processing system. Building from information theory concepts, it has found applications in several fields and proved its value in several contexts [4–8]. For a stochastic process considered to be stationary, the minimal deterministic automaton is its optimal minimal description, understood as having the best (most accurate) predictive power while using the least possible resources (minimal forecasting complexity) [1,9]. Causality is taken in a general temporal sense: in a given context, cause-to-effect relations are established between past to future events [10].

The Ising model in one-dimension is the best known in Statistical Physics and has become a common topic in most Statistical Mechanics books (see, for example, [11]). Despite its intensive scrutiny in different settings, under different Hamiltonian and interaction ranges, its analysis, in terms of information theory, as a symbol production system is more or less recent. In fact, until 1998, this approach was not attempted when Feldman et al. undertook the task of casting the Ising model under such language [12–14]. Their work allowed the deduction of closed expressions for the entropy density, forecasting complexity, and effective measure complexity. However, these previous treatments did not consider the Ising model in the general framework of a Markov (Gibbs) random field. While the previous approaches are sufficient when nearest-neighbour interaction is considered, the more general framework is necessary, beyond nearest-neighbour interaction, to determine the general probability measure of the associated Markov process [15]. This is what it aims at in this contribution.

Despite being a well-studied system, working through all the mathematical details involved to solve the inverse problem, obtain the stochastic matrix from the Markov Field, and build the minimal deterministic automata in the most general setting in one dimension is worthwhile. Entropic magnitudes follow, which are usually not treated in the Ising model. In this light, we show the use of the developed framework via two examples.

II. THE TRANSFER MATRIX FORMALISM

The Ising (-Lenz) model is probably the most studied lattice-type model in statistical mechanics and is well covered in several statistical physics books for nearest 1/2 spin neighbour interactions [11]. Let us briefly recap, for completeness and notation purposes, the basic ideas of the transfer matrix formalism (we closely follow Dobson [16]) but in a general setting of a local type interaction Hamiltonian, which is usually not found in texts.

Consider a one-dimensional chain of discrete values of length L:

$$s^L = s_0 s_1 s_2 \dots s_{L-1},$$

where s_i can take values from a finite alphabet Θ of cardinality $\theta(=|\Theta|)$ (there will be θ^L possible sequences s^L). Each s_i is called a spin. The interaction between spins of the sequence has a finite range n such that it can be written as

$$E(s_i, s_{i\pm k}) = \begin{cases} \Lambda(s_i, s_{i\pm k}) & 0 < k \leq n \\ 0 & k > n \end{cases}. \tag{1}$$

The s^L sequence can be partitioned into blocks of length n

$$s^L = [s_0 s_1 \dots s_{n-1}] \dots [s_{(N-1)n} s_{(N-1)n+1} \dots s_{Nn-1}],$$

where it has been taken $L = Nn$. The expression can be relabeled as

$$\begin{aligned} s^L &= [s_0^{(0)} \dots s_{n-1}^{(0)}][s_0^{(1)} \dots s_{n-1}^{(1)}] \dots [s_0^{(N-1)} \dots s_{n-1}^{(N-1)}]. \\ &= \eta_0 \eta_1 \dots \eta_{N-1}, \end{aligned}$$

with

$$\eta_i = s_0^{(i)} s_1^{(i)} \dots s_{n-1}^{(i)}. \tag{2}$$

The set of all possible blocks η_i will be denoted by Υ with cardinality $\upsilon = \theta^n$. Υ will be taken as an ordered set (e.g., lexicographic order) where each η_i a natural number, between 0 and $\upsilon - 1$, will be assigned. In what follows, η_i should be understood not only as the configuration (2) but also as its corresponding order in the set Υ; context will eliminate any ambiguity.

As the interaction has range n, one spin corresponding to the η_i block can only interact with all the spins within η_i (type I interaction), and at least one spin from the adjacent blocks $\eta_{i\pm1}$ (type II interaction).

Assuming the symmetry $\Lambda(s_i, s_j) = \Lambda(s_j, s_i)$, the interaction energy of type I for the η_p block, in the presence of an external field B, will be

$$x_{\eta_p} = -B \sum_{i=0}^{n-1} s_i^{(p)} + \sum_{i=0}^{n-2} \sum_{k=i+1}^{n-1} \Lambda(s_i^{(p)}, s_k^{(p)}), \tag{3}$$

which defines a vector $\langle X|$ of length υ. The contribution of type II will be denoted by $y_{\eta_p \eta_{p+1}}$, and will be given by

$$y_{\eta_p \eta_{p+1}} = \sum_{i=0}^{n-1} \sum_{k=0}^{i} \Lambda(s_i^{(p)}, s_k^{(p+1)}), \tag{4}$$

which defines a $\upsilon \times \upsilon$ matrix. In general $y_{\eta_i \eta_j} \neq y_{\eta_j \eta_i}$ which makes Y non-symmetric. The energy of the whole configuration S^L can then be written as

$$\Lambda(s^L) = x_{\eta_0} + y_{\eta_0 \eta_1} + x_{\eta_1} + y_{\eta_1 \eta_2} + \dots + y_{\eta_{N-2} \eta_{N-1}} + x_{\eta_{N-1}}. \tag{5}$$

The vector $\langle U|$ and the matrix V are then introduced as

$$u_{\eta_i} = \exp(-\frac{1}{2} \beta x_{\eta_i}) \tag{6}$$

$$v_{\eta_i \eta_j} = \exp[-\beta(\frac{1}{2} x_{\eta_i} + y_{\eta_i \eta_j} + \frac{1}{2} x_{\eta_j})]. \tag{7}$$

where $\beta \equiv (k_B T)^{-1}$ is the Boltzmann product. V is known as the transfer matrix.

The partition function follows

$$\begin{aligned} Z_{Nn} &= \sum_{\eta_0=0}^{\upsilon-1} \sum_{\eta_1=0}^{\upsilon-1} \dots \sum_{\eta_{N-1}=0}^{\upsilon-1} \exp[-\beta \Lambda(s^L)] \\ &= \langle U | V^{N-1} | U \rangle, \end{aligned} \tag{8}$$

for free boundary conditions. For periodic boundary conditions

$$Z_{Nn} = Tr(V^N). \tag{9}$$

$Tr(M)$ denotes the trace of the matrix M.

As the trace of a matrix is invariant to similarity transformations, from equation (9), for close boundary conditions,

$$Z_{Nn} = \sum \lambda_i^N. \tag{10}$$

λ_i are the eigenvalues of the matrix V. If λ_i is degenerate, then the term is added as many times as its multiplicity. If the eigenvalues are labeled in non increasing order ($|\lambda_i| \geq |\lambda_j| \rightarrow j \geq i$), then for $N \gg 1$

$$Z_{Nn} = \lambda_0^N \tag{11}$$

where λ_0 is the dominant eigenvalue; according to the Perron-Frobenius theorem, it is real, positive, and non-degenerate [15].

For open boundary conditions, again using the Perron-Frobenius theorem for a square positive defined matrix V, the following holds

$$\lim_{N \longrightarrow \infty} \frac{V^N}{\lambda_0^N} = |_r a_0\rangle \langle_l a_0|, \tag{12}$$

where $\langle_l a_0|$ and $|_r a_0\rangle$ are, respectively, the left and right eigenvectors corresponding to the dominant eigenvalue. The eigenvectors are normalized in the sense of $\langle_l a_0 | _r a_0 \rangle = 1$. The matrix $|_r a_0\rangle \langle_l a_0|$ is known as the Perron projection matrix. Using (12) and (8) we arrive at

$$Z_{Nn} = \langle U | _r a_0 \rangle \langle_l a_0 | U \rangle \lambda_0^{N-1} \tag{13}$$

which, in the particular case of a diagonalizable matrix, reduces to

$$Z_{Nn} = u_0^2 \langle a_0 | a_0 \rangle \lambda_0^{N-1} \tag{14}$$

and u_i are the components of the vector $\langle U|$ in the orthogonal base, defined by the eigenvectors $\langle a_i|$. It is well documented how the thermodynamic magnitudes can be obtained from the partition function [11].

The probability of a given spin chain will be given by

$$\begin{aligned} Pr(s^L) &= \tfrac{1}{Z_{Nn}} e^{-\beta \Lambda(s^L)} \\ &= \tfrac{1}{Z_{Nn}} \left(U_{\eta_0} V_{\eta_0 \eta_1} V_{\eta_1 \eta_2} \dots V_{\eta_{N-2} \eta_{N-1}} U_{\eta_{N-1}} \right) \\ &= \tfrac{U_{\eta_0} U_{\eta_{N-1}}}{M \lambda_0^{N-1}} \prod_{i=0}^{N-2} V_{\eta_i \eta_{i+1}} \end{aligned} \tag{15}$$

valid for free boundary conditions, and $M \equiv \langle U|_r a_0\rangle\langle {}_l a_0|U\rangle$. For periodic boundary conditions

$$Pr(s^L) = \frac{1}{\lambda_0^N}\prod_{i=0}^{N-2} V_{\eta_i\eta_{i+1}} \tag{16}$$

where in both cases, $N \gg 1$.

III. ISING MODEL AS A MARKOV (GIBBS) RANDOM FIELD

The Ising model is a particular case of a Gibbs random field [15], where a Markov random field can be defined

$$Pr(\eta_i = \eta|s^L - \eta_i) = Pr(\eta_i = \eta|\eta_t = b_t \,\text{for}\, t \in \mathcal{N}_i,\ b_t \in \Upsilon). \tag{17}$$

The η_i value is conditioned only in the neighbourhood and not on the entire spin configuration.The probabilities given by (17) are called the local characteristics associated with $\mathcal{P}$.

If $s^L - \eta_i$ is the configuration s^L without considering the block η_i, then the probability $Pr(\eta_i|s^L - \eta_i)$ that the i-block has value η_i when all the other spins (that is, excluding the η_i block) will have the configuration $s^L - \eta_i$ the product rule will given by,

$$Pr(\eta_i|s^L - \eta_i) = \frac{Pr(s^L)}{Pr(s^L - \eta_i)} = \frac{Pr(s^L)}{\sum\limits_{s^{L*}} Pr(s^{L*})} \tag{18}$$

where the sum s^{L*} is over all configurations identical to s^L except, possibly, for the block η_i.

Using equation (15), the probability of a configuration will be

$$\begin{aligned} Pr(s^L) = &\ \tfrac{1}{Z_{Nn}} e^{-\beta\Lambda(s^L)} \\ &= \frac{1}{Z_{Nn}} e^{-\beta x_{\eta_{N-1}}} \prod_{j=1}^{N-2} e^{-\beta x_{\eta_j}} e^{-\beta y_{\eta_j\eta_{j+1}}} \end{aligned} \tag{19}$$

and,

$$Pr(s^L - \eta_i) = \sum_{\eta_k} e^{-\beta y_{\eta_{i-1}\eta_k}} e^{-\beta x_{\eta_k}} e^{-\beta y_{\eta_k\eta_{i+1}}}. \tag{20}$$

The local characteristics equation (18) is then

$$Pr(\eta_i|s^L - \eta_i) = \frac{V_{\eta_{i-1}\eta_i} V_{\eta_i\eta_{i+1}}}{\sum\limits_{\eta_k} V_{\eta_{i-1}\eta_k} V_{\eta_k\eta_{i+1}}} \tag{21}$$

for blocks η_i not at the extremes.

For the first block

$$Pr_0(\eta_0|s^L - \eta_i) = \frac{U_{\eta_0} V_{\eta_0\eta_1}}{\sum\limits_{\eta_k} U_{\eta_k} V_{\eta_k\eta_1}} \tag{22}$$

A similar expression can be found for the last block. Expression (21) has the important consequence that

$$Pr(\eta_i|s^L - \eta_i) = Pr(\eta_i|\eta_{i-1}, \eta_{i+1}) \tag{23}$$

In the associated Markov process, the spin chain is considered a sequential process where blocks η are "emitted" sequentially. In this sense, the probability that, at a given moment, a block η_i is the output of the Markov process is conditioned only on the previously emitted block. In this sense, we can describe a transition probability from one emitted block to the next and associate a probability with it.

Consider the η_i blocks as describing the possible states of an arbitrary block of spins, then Υ is the set of all states. A stochastic matrix P can be defined as

$$P_{ij} = Pr(\eta_j|\eta_i). \tag{24}$$

which describes the transition probability from state η_i to state η_j. By definition $\sum_j P_{ij} = 1$. If $\langle p^\infty|$ is the vector of probabilities over the blocks η_i (the probability that a given block of spins is in a given state), then it is well known that the stationary distribution [15] is given by

$$\langle p^\infty| = \langle w_0| \tag{25}$$

where $\langle w_0|$ is the left dominant eigenvector of the matrix P. The vector $\langle p^\infty|$ allows to calculate $Pr(\eta_i)$ when the Markov process has been running for a sufficiently long time.

The local characteristics can be written in terms of the stochastic matrix P using Bayes theorem

$$\begin{aligned} Pr(\eta_i|\eta_{i-1}, \eta_{i+1}) &= \tfrac{Pr(\eta_i|\eta_{i-1})Pr(\eta_{i+1}|\eta_i\eta_{i-1})}{Pr(\eta_{i+1}|\eta_{i-1})} \\ &= \tfrac{Pr(\eta_i|\eta_{i-1})Pr(\eta_{i+1}|\eta_i)}{\sum\limits_l Pr(\eta_l|\eta_{i-1})Pr(\eta_{i+1}|\eta_l)} \end{aligned} \tag{26}$$

where the Markov character of the field has been used, and the total probability theorem justifies the last step.

Equation (26) can be rewritten as

$$\begin{aligned} &Pr(\eta_i|\eta_{i-1}, \eta_{i+1}) \sum_l Pr(\eta_l|\eta_{i-1}) Pr(\eta_{i+1}|\eta_l) = \\ &Pr(\eta_i|\eta_{i-1}) Pr(\eta_{i+1}|\eta_i) \end{aligned} \tag{27}$$

which forms, when written for each η_i, an homogeneous system of quadratic forms. Such a system can have a non-trivial solution if it is undetermined, which happens if the square of the number of unknowns is larger than the number of equations.

There are $\nu = \theta^n$ possible different blocks η. $Pr(\eta_i|\eta_{i-1}, \eta_{i+1})$ is known from (21). As each local characteristic is determined by three η's, there will be ν^3 equations. $Pr(\eta_i|\eta_k)$ only depends on the actual values of η_i and η_k and not on their position, therefore there will be ν^2 unknowns. The relations

$$\sum_i^\nu Pr(\eta_i|\eta_j) = 1 \qquad \forall j$$

must be added that eliminates ν unknowns. The total number of unknowns is $\nu(\nu-1)$, and the total number of equations ν^3.

$$\nu^3 < \nu^2(\nu-1)^2 \Longrightarrow \nu > 3.6$$

As we seek solutions for integer values of ν, the effective solution will be $\nu \leq 4$. Additional symmetry of the transfer matrix (e.g. $\nu = 2$) can lead to further reduction of the equations, and the system could also have a non-trivial solution for such cases.

Returning to equation (27) and rewriting for any local characteristic

$$Pr(\eta_i|\eta_j,\eta_m)\sum_l Pr(\eta_l|\eta_j)Pr(\eta_m|\eta_l) = \quad (28)$$
$$Pr(\eta_i|\eta_j)Pr(\eta_m|\eta_i),$$

introducing

$$Y_l(j,m) = Pr(\eta_l|\eta_j)Pr(\eta_m|\eta_l) \quad (29)$$

equation (28) can be written as

$$Pr(\eta_i|\eta_j,\eta_m)\sum_l Y_l(j,m) = Y_i(j,m). \quad (30)$$

The normalization condition (which can be derived from equation (26)) over the local characteristics determines

$$\sum_k^v Pr(\eta_k|\eta_j,\eta_m) = 1, \quad (31)$$

Equation (30) is linear and homogeneous over the $Y_i(j,m)$ which, upon solving for the non-trivial-case, leads to a system of simple homogeneous quadratic equations (29) which can be readily solved.

IV. ISING MODEL AS A MINIMAL DETERMINISTIC AUTOMATA

The Markov character of the system means that for the associated Markov process, the generation process can forget all the past except the last block η_{-1} (the last n spins) to determine, as certain as possible, the future. In other words, if the local characteristics imply a stochastic matrix as equation (28) implies, then all past configurations $\overleftarrow{s}^L$ with the same last block η_{-1} condition (statistically) the same future, this fact allows considering the Ising chain as a canonical finite state machine or minimal deterministic automata.

If two blocks η_{-1} and η'_{-1} give the same $Pr(\overrightarrow{s}^L|\overleftarrow{s}^L)$, for all possible futures $\overrightarrow{s}^L$, then η_{-1} y η'_{-1} are said to belong to the same causal state (C_p) and we write

$$\eta_{-1} \sim \eta'_{-1},$$

where $\eta_{-1}, \eta'_{-1} \in C_p$ [10]. Two blocks of the same causal state C_p define identical rows in the stochastic matrix.

The partition of the set Υ in classes of causal states is an equivalence relation.The set of causal states, denoted by C, uniquely determines the future of a sequence.

The probability of a causal state is directly deducible from equation (25),

$$Pr(C_p) = \sum_{\eta_j\in C_p} Pr(\eta_j) = \sum_{\eta_j\in C_p} p^{\infty}_{\eta_j} \quad (32)$$

As each causal state represents the set of past that determines (probabilistically) the same future, the set of causal states represents the memory the system has to keep to predict the future.

The forecasting complexity [1], also known as statistical complexity, has been defined as the Shannon entropy H over the causal states [9][1]

$$\begin{aligned} C_\mu &\equiv -\sum_{C_p\in C} Pr(C_p)\log Pr(C_p) \\ &= H[C]. \end{aligned} \quad (33)$$

The logarithm is usually taken in base two, and the units are then bit. Forecasting complexity measures how much memory (resources) the system needs to predict the future optimally. If the system has $|C|$ causal states, then the forecasting complexity has the upper bound

$$C_\mu \leq \log|C|,$$

corresponding to a uniform distribution of probabilities. The upper bound of the forecasting complexity is also known as topological entropy.

The probability of occurrence of block η_i conditional on the causal state C will be given by

$$\begin{aligned} Pr(\eta_i|C) &= \sum_{\eta_k} Pr(\eta_i|\eta_k\cap C)Pr(\eta_k|C) \\ &= Pr(\eta_i|\eta_j;\eta_j\in C)\sum_{\eta_k\in C} Pr(\eta_k|C) \\ &= Pr(\eta_i|\eta_j;\eta_j\in C) \end{aligned} \quad (34)$$

in the first step, the total probability theorem was used. In the second step, use has been made of the fact that conditioning in $\eta_k\cap C$ is equal to conditioning in η_j if the block belongs to the causal state C and, finally $\sum_{\eta_k\in C} Pr(\eta_k|C) = 1$.

On the other hand

$$\begin{aligned} Pr(\eta_j) &= \sum_{\eta_k} Pr(\eta_j|\eta_k)Pr(\eta_k) \\ &= \sum_{C_k\in C}\sum_{\eta_k\in C_k} Pr(\eta_j|\eta_k)Pr(\eta_k) \\ &= \sum_{C_k\in C} Pr(\eta_j|\eta_{k'},\eta_{k'}\in C_k)Pr(C_k) \end{aligned}$$

making use of equation (34) to get

$$Pr(\eta_j) = \sum_{C_k\in C} Pr(\eta_j|C_k)Pr(C_k). \quad (35)$$

which allows us to compute the occurrence of a block from the probabilities over the causal states.

[1] We prefer to say that it is defined in such a way, rather than state that it follows from its original definition because as pointed out by Grassberger [17], there is no proof that the minimal graph corresponds to the minimal forecasting complexity. However, in the kind of model we are studying, this seems to be the case in general.

The probability of a transition from one causal state to another will be given by

$$Pr(C_k \to C_p) \equiv Pr(C_p|C_k) = \sum_{\eta_j \in C_p} Pr(\eta_j|C_k) \tag{36}$$

We define the transition matrix $T^{(\eta)}$, whose elements are the probability of going from state C_k to state C_p upon emitting a block η:

$$T^{(\eta)}_{rq} = Pr(C_r \overset{\eta}{\to} C_q). \tag{37}$$

By construction, the emission of a block η determines the causal state uniquely to where the transition occurs (this is called the unifiliar property [10]). In this sense, the generation process is deterministic. Correspondingly, the connectivity matrix T is defined as

$$T_{rq} = \sum_{\eta \in \Sigma} T^{(\eta)}_{rq} = Pr(C_r \to C_q) \tag{38}$$

which connects causal states without regard to the emitted block.

To account for the irreducible randomness, the entropy density is defined as [18]

$$\begin{aligned} h &= \lim_{L\to\infty} H[\eta_0|\eta_{-L}\eta_{-L+1}\dots\eta_{-1}] \\ &= \lim_{L\to\infty} H[\eta_0|\overleftarrow{s}^L] \end{aligned} \tag{39}$$

h is the uncertainty on the next emitted block η_0 conditional on having seen infinite previous blocks (spins). By definition, $h \geq 0$.

$$\begin{aligned} H[\eta_0|\overleftarrow{s}^L] &= -\sum_{\eta_0}\sum_{\eta_{-n}}\dots\sum_{\eta_{-1}} Pr(\overleftarrow{s}^L\eta_0)\log Pr(\eta_0|\overleftarrow{s}^L) \\ &= -\sum_{\eta_0}\sum_{\eta_{-N}}\dots\sum_{\eta_{-1}} Pr(\eta_0|\overleftarrow{s}^L)Pr(\overleftarrow{s}^L)\log Pr(\eta_0|\overleftarrow{s}^L) \end{aligned} \tag{40}$$

where $\overleftarrow{s}^L = \eta_{-N}\eta_{-N+1}\dots\eta_{-1}$, , on the other hand, using $\overleftarrow{s}^{L-n} = \eta_{-N}\dots\eta_{-2}$

$$Pr(s^L) = Pr(s^{L-n}|\eta_{-1})Pr(\eta_{-1}). \tag{41}$$

use has been made of Bayes theorem. Substituting equation (41) on equation (40) and reordering terms

$$\begin{aligned} H[\eta_0|\overleftarrow{s}^L] = -\sum_{\eta_0}\Bigg[\sum_{\eta_{-1}} Pr(\eta_0|\eta_{-1})Pr(\eta_{-1})\log Pr(\eta_0|\eta_{-1}) \\ \left\{\sum_{\eta_{-N}}\dots\sum_{\eta_{-2}} Pr(\overleftarrow{s}^{L-n}|\eta_{-1})\right\}\Bigg]. \end{aligned} \tag{42}$$

$$\sum_{\eta_{-N}}\dots\sum_{\eta_{-2}} Pr(\overleftarrow{s}^{L-n}|\eta_{-1})$$

is the probability that from η_{-1} any configuration is conditioned and this probability is 1. Equation (42) then reduces to

$$\begin{aligned} h &= \lim_{L\to\infty} H[\eta_0|\overleftarrow{s}^L] = H[\eta_0|\eta_{-1}] \\ &= -\sum_{C_\alpha \in \mathcal{C}} Pr(C_\alpha) \sum_{\eta_k \in \Sigma} Pr(\eta_k|C_\alpha)\log Pr(\eta_k|C_\alpha) \end{aligned} \tag{43}$$

The mutual information between past and future is called the effective measure complexity [1], also known as excess entropy [18]

$$E \equiv I[\overleftarrow{s} : \overrightarrow{s}], \tag{44}$$

where $I[X : Y]$ is the mutual information between X and Y. From the finite range character of the interaction in the Ising model

$$\begin{aligned} E = I[\overleftarrow{s} : \overrightarrow{s}] &= I[\eta_{-1} : \eta_0] \\ &= H[\eta_{-1}] - H[\eta_0|\eta_{-1}] \end{aligned} \tag{45}$$

where

$$H[\eta_i] = \sum_{\eta_i \in \Sigma} Pr(\eta_i)\log Pr(\eta_i), \tag{46}$$

and

$$H[\eta_0|\eta_{-1}] = H[\eta_0|\overleftarrow{s}^L] = h$$

given by equation (43).

From equation (33) and (43) we arrive to the expression

$$E = C_\mu - h. \tag{47}$$

The effective measure complexity measures the resources the system needs to optimally predict the future once the irreducible randomness has been subtracted [18]. As E is mutual information, it will always be a non-negative value, which implies

$$C_\mu \geq h$$

If the system is perfectly periodic, then $h = 0$ and

$$C_\mu = E.$$

V. EXAMPLES

V.1. 1/2 nearest neighbors spin chain

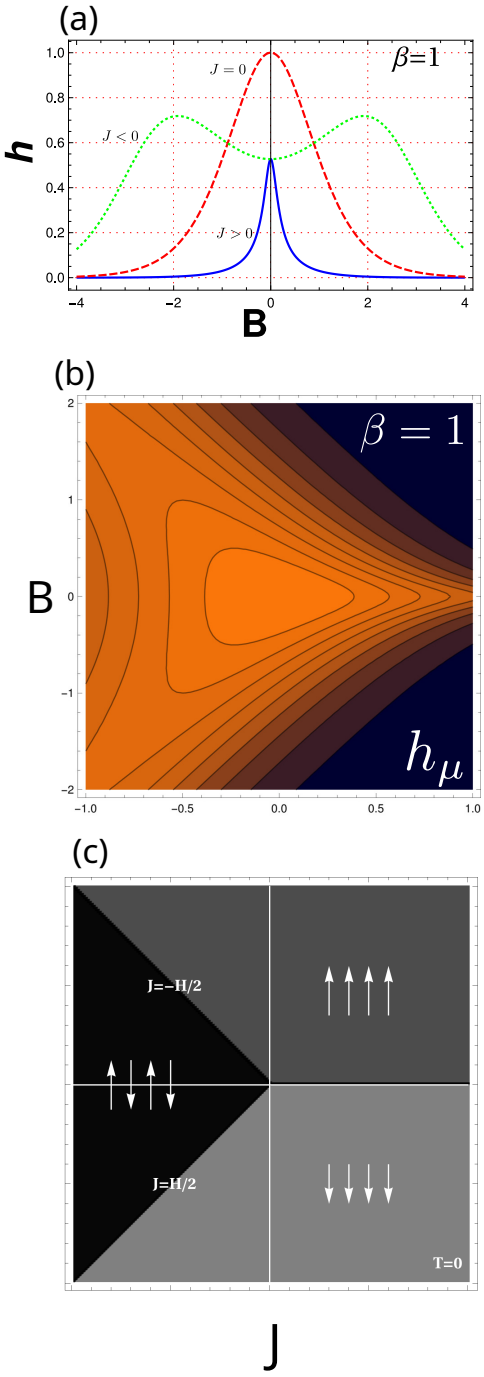

Figure 1. 1/2 nearest neighbors spin chain. (a) The entropy density as a function of the applied field B for ferromagnetic ($J > 0$), antiferromagnetic ($J < 0$), and paramagnetic ($J = 0$) interaction. The Boltzmann factor in all cases is taken as $\beta = 1$. (b) The entropy density h map as a function of the system control parameter J and the applied field B. Black corresponds to $h = 0$, and full orange to $h = 1$. Maximum disorder values happen along the line with no applied field $B = 0$, with the paramagnetic state's maximum at $J = 0$. Boltzmann factor as in (a). (c) The diagram shows the phase diagram for zero absolute temperature as a function of the parameters J and B. Two phases can be identified: ferro- and antiferromagnetic.

The 1/2 nearest neighbour spin chain is defined by the interaction Hamiltonian [20]

$$E = -B\sum_i s_i - J\sum_j s_j s_{j+1}, \tag{48}$$

where B is the external field, and J is the interaction parameter. The η blocks set will be

$$\eta = \downarrow, \uparrow.$$

The local characteristics derived from equation (21) reduce to

$$\begin{aligned}
Pr(\downarrow \mid \downarrow,\downarrow) &= \tfrac{e^{4\beta J}}{e^{2\beta B}+e^{4\beta J}}\\
Pr(\downarrow \mid \downarrow,\uparrow) &= \tfrac{1}{e^{2\beta B}+1}\\
Pr(\uparrow \mid \downarrow,\downarrow) &= \tfrac{1}{e^{4\beta J-2\beta B}+1}\\
Pr(\uparrow \mid \downarrow,\uparrow) &= \tfrac{e^{2\beta B}}{e^{2\beta B}+1}\\
Pr(\downarrow \mid \uparrow,\downarrow) &= \tfrac{1}{e^{2\beta B}+1}\\
Pr(\downarrow \mid \uparrow,\uparrow) &= \tfrac{1}{e^{2\beta(B+2J)}+1}\\
Pr(\uparrow \mid \uparrow,\downarrow) &= \tfrac{e^{2\beta B}}{e^{2\beta B}+1}\\
Pr(\uparrow \mid \uparrow,\uparrow) &= \tfrac{e^{2\beta(B+2J)}}{e^{2\beta(B+2J)}+1}.
\end{aligned} \tag{49}$$

Solving the linear system of equation (30), results in the system of quadratic equations

$$\begin{aligned}
Pr(\downarrow \mid \uparrow)Pr(\uparrow \mid \downarrow) &= e^{2\beta B-4\beta J}Pr(\downarrow \mid \downarrow)\\
Pr(\uparrow \mid \uparrow) &= e^{2B\beta}P(\downarrow \mid \downarrow),
\end{aligned} \tag{50}$$

together with the normalization conditions

$$\begin{aligned}
Pr(\downarrow \mid \uparrow) + Pr(\uparrow \mid \uparrow) &= 1\\
Pr(\downarrow \mid \downarrow) + Pr(\uparrow \mid \downarrow) &= 1
\end{aligned} \tag{51}$$

lead to the solution for[2]

$$Pr(\downarrow \mid \downarrow) = \frac{2e^{2\beta J}}{\sqrt{4e^{2\beta B}+e^{4\beta(B+J)}-2e^{2\beta(B+2J)}+e^{4\beta J}}+e^{2\beta(B+J)}+e^{2\beta J}}. \tag{52}$$

The other entries of the stochastic matrix follow.

[2]This result is equivalent to the one reported as equation (7.31) in [12], but, if one calculates all matrix entries from equation (7.15), the row normalization condition is violated. Therefore, the correctness of the entries in (7.31) in [12] is accidental due to forcing row normalization.

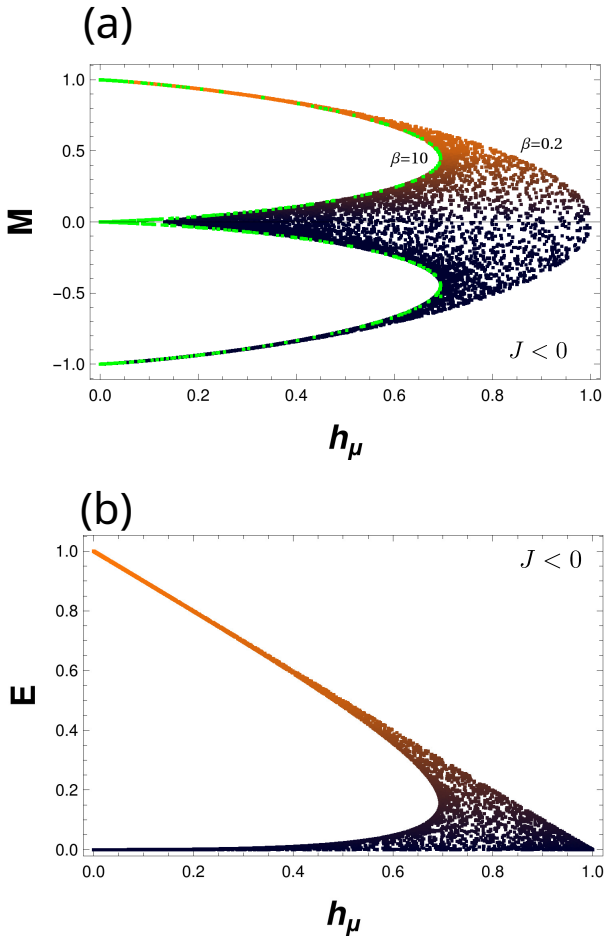

Figure 2. 1/2 nearest neighbors spin chain. (a) The dependence of the magnetization M with the disorder measured by the entropy density B. 10^5 points were used with random parameters in the range $\beta \in [10^{-4}, 10^2]$, $J \in [-1.5, 0]$ and $B \in [-3, 3]$. The green points correspond to $\beta = 10$, while the rust colors are for $\beta = 0.2$, as labeled in the figure. The rust color degradation grows lighter the larger the M value and is used as a visual aid. For all points, the spin coupling is antiferromagnetic. (b) The complexity-entropy diagram for the $1/2$-nearest neighbor Ising model. Simulation conditions follow the same parameter range as in (a). The rust color degradation grows lighter the larger the M value and is used as a visual aid.

The behavior of the entropy density h with the applied field for the three signs of the interaction term is shown in Figure 1a for $\beta = 1$. The results are consistent with the usual graphic of the Boltzmann entropy for this model. The applied field B generally lowers the system's entropy as it tends to align the spins along the field. In the case of the antiferromagnetic coupling, with an increasing value of B, first entropy increases due to random spins, initially contrary to the external field flips, leading to increasing disorder. For the larger field, B, the order starts prevailing as further production of aligned spins overcomes the initial disordering process. For $B = 0$, the system starts with an initial amount of randomness ($h > 0$) for all signs of J as temperature introduces disorder.

In Figure 1(b), the amount of disorder as a function of the applied field B and interaction parameter J. The maximum value of h, for a given value J, is attained at the line $B = 0$ and is the result of $\beta > 0$. The absolute maximum value of h is taken for $B = 0$ and $J = 0$. Increasing the applied field and interaction parameter decreases the spin system's disorder. Compare the entropy map with the phase diagram of the spin system at zero temperature (1(c)). The map of effective measure complexity for $\beta \to \infty$ reproduces the phase diagram (not shown). However, for $\beta < \infty$, the entropy density map merely indicates the mapping of randomness regardless of the underlying pattern, whether ferro- or antiferromagnetic.

The behavior of the magnetization with the disorder as measured through h can be seen in figure 2a. From the stochastic matrix, the two-state minimal deterministic automata were built. The plot was calculated for 10^5 points, randomly taking the parameters' value but keeping $J < 0$. First, we notice that three magnetization values are possible at zero disorder $h = 0$, two at the extremes, corresponding to the spin alignment forced by a sufficiently strong applied field B, and a zero magnetization at $B = 0$. At lower temperatures ($\beta = 10$), for a given disorder value of h, the amount of magnetization that can be accommodated around the vanishing magnetization has fewer values than for larger temperatures $\beta = 0.2$. Also, increasing temperature makes more disorder available for the system, which is seen for the larger possible values of h.

Finally, (h, E) was calculated using the same procedure for the magnetization plot and is shown as a complexity map in figure 2b. This type of complexity map has been discussed before [18]. Small values of the disorder can accommodate a large range of effective measure complexity values, which means varying probability between the two possible causal states. As disorder increases, the system loses structure, tending towards a single-state process that, although increasingly random, is also increasingly less complex.

V.2. 1/2 next nearest neighbors spin chain

If a second coordination is added, the 1/2 next nearest neighbors spin chain, the interaction Hamiltonian [20] now is given by

$$E = -B\sum_i s_i - J_1\sum_j s_j s_{j+1} - J_2\sum_k s_k s_{k+2}. \tag{53}$$

where B is the external field, and J_1, J_2 are spin coupling parameters. The η blocks set will be

$$\eta = \downarrow\downarrow, \downarrow\uparrow, \uparrow\downarrow, \uparrow\uparrow,$$

again, the value 1 corresponds to spin up, whereas -1 corresponds to spin down.

Figure 3 above shows the effective measure complexity as a function of J_1 and J_2 for the ground state ($\beta \to \infty$) at zero field ($B = 0$). While in the nearest neighbor case, only the ferromagnetic and antiferromagnetic case was found in the phase diagram, here, besides those two, a third ordered phase comes into existence, with a periodicity of 4, which can be regarded as a higher coordination anti-ferromagnetic phase. $\uparrow\uparrow\downarrow\downarrow$. This phase results from the larger range of interaction between the spins and competing interactions. For $J_2 \geq 0$, the second range interaction, governed by J_2, favors the ferromagnetic phase, and its balance with the strength and sign of J_1 determines the ferromagnetic or antiferromagnetic

ordering similar to the nearest neighbor coupling. When J_2 becomes negative, it strengthens an antiferromagnetic coupling between second range spin now depending on the sign of J_1, this ordering competes with the first range coupling in two different ways, which will always be at odds with J_2. For certain values of J_1 and J_2, the solution to this balance is the period 4 antiferromagnetic ordering.

When the applied field is different from zero, as shown in the Effective Complexity Measure plot 3 below, as a function of J_1/B and J_2/B, four orderings of periodicity 1, 2, 3, and 4, identified as the ordered sequences:

(1) $\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\cdots$

(2) $\uparrow\downarrow\uparrow\downarrow\uparrow\downarrow\uparrow\downarrow\cdots$

(3) $\uparrow\uparrow\downarrow\uparrow\uparrow\downarrow\uparrow\uparrow\cdots$

(4) $\uparrow\uparrow\downarrow\downarrow\uparrow\uparrow\downarrow\downarrow\cdots$

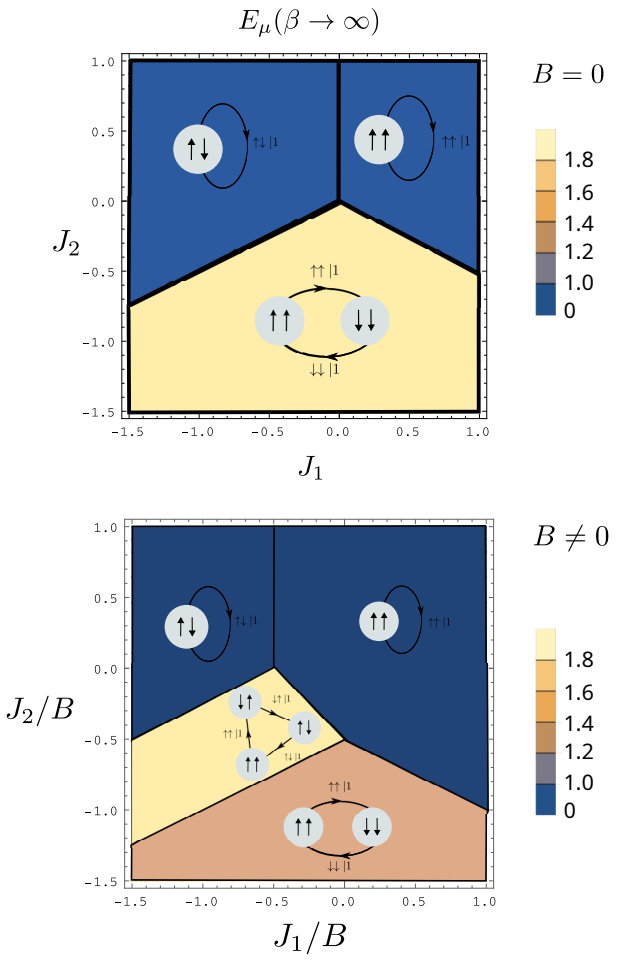

Figure 3. 1/2 next nearest neighbors spin chain. Effective measure complexity (E) at the ground state ($\beta \to \infty$) for zero field ($B = 0$) and non-zero field ($B \neq 0$). Four possible orderings are found $\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\cdots$, $\uparrow\downarrow\uparrow\downarrow\uparrow\downarrow\uparrow\downarrow\cdots$, $\uparrow\uparrow\downarrow\uparrow\uparrow\downarrow\uparrow\uparrow\cdots$ and $\uparrow\uparrow\downarrow\downarrow\uparrow\uparrow\downarrow\downarrow\cdots$. The third configuration is only possible at $B \neq 0$.

This has been reported before [20]. The new phase $\uparrow\uparrow\downarrow\uparrow\uparrow\downarrow\uparrow\uparrow\cdots$ appears as a wedge in the phase diagram when B is zero, and it is the result of the new competing factor when the applied field favors one sense in space. The applied field has an effect equivalent to a dipolar average field pointing, in this case, in the up direction, therefore implying a flip of a spin to the upper direction that would otherwise have an antiferromagnetic coupling. This phase, therefore, appears between the simple antiferromagnetic phase and the periodicity 4 antiferromagnetic phase.

VI. CONCLUSION

In this article, we aimed to develop the mathematical treatment of minimal deterministic automata to fully model one-dimensional Ising models, going beyond previous approaches. The deduced expressions can be used to model specific instances of the interaction Hamiltonian. Furthermore, the detailed deduction can be a route map to similar deductions for other common statistical mechanics models, such as Pott or Heisenberg.

VII. ACKNOWLEDGEMENT

FAPEMIG partially financed this work under project BPV-00047-13. EER to thank PVE/CAPES for financial support under grant 1149-14-8. Infrastructure support was given under project FAPEMIG APQ-02256-12.

We want to thank Prof. Crutchfield and coworkers for pointing out that Feldman's treatment still gives the correct quantitative results for the NN model and for further discussions on the subject as a whole.

REFERENCES

[1] P. Grassberger, *Int. J. Theor. Phys.* **25**, 907 (1986).
[2] J. P. Crutchfield, *Knowledge and meaning... chaos and complexity*, in *Modeling complex phenomena*, edited by L. Lam and V. Naroditsky, pp. 66-101, Springer, Berlin (1992).
[3] J. P. Crutchfield, *Nature Physics* **8**, 17 (2012).
[4] D. P. Varn, G. S. Canright, J. P. Crutchfield, *Acta Cryst. A* **A69**, 413 (2013).
[5] V. Ryabov, D. Neroth, *Chaos* **21**, 037113 (2011).
[6] R. Haslinger, K. L. Klinker, C. R. Shalizi, *Neural Computation* **22**, 121 (2010).
[7] E. Rodriguez-Horta, E. Estevez-Rams, R. Neder, R. Lora-Serrano, *Acta. Cryst. A* **A73**, 357 (2017).
[8] E. Rodriguez-Horta, E. Estevez-Rams, R. Lora-Serrano, R. Neder, *Acta. Cryst. A* **A73**, 377 (2017).
[9] J. P. Crutchfield, *Physica D* **75**, 11 (1994).
[10] C. R. Shalizi, J. P. Crutchfield, *J. Stat. Phys.* **104**, 817 (2001).
[11] D. A. Lavis, G. M. Bell, *Statistical Mechanics of Lattice Systems*, Springer, Berlin (1999).
[12] D. P. Feldman, *Computational mechanics of classical spin systems*, Ph.D. Dissertation, Physics Department, University of California, Davis, (September, 1998).
[13] J. P. Crutchfield, D. P. Feldman, *Phys. Rev. E* **55**, R1239(R) (1997).

[14] D. P. Feldman, J. P. Crutchfield, Santa Fe Institute Working Paper 98-04-026 (2008).
[15] E. Behrends, *Introduction to Markov chain*, Springer, Heidelberg (2000).
[16] J. F. Dobson, *J. Math. Phys.* **10**, 40 (1969).
[17] P. Grassberger, arXiv preprint arXiv:1708.03197v3 (2018).
[18] D. P. Feldman, C. S. McTeague, J. P. Crutchfield, *Chaos* **18**, 043106 (2008).
[19] J. Crutchfield, D. P. Feldman, *Chaos* **13**, 25 (2003).
[20] T. Morita, T. Horiguchi, *Phys. Lett.* **A41**, 043106 (1972).