Information Storage in the Stochastic Ising Model at Zero Temperature

Ziv Goldfeld
MIT
zivg@mit.edu

Guy Bresler
MIT
guy@mit.edu

Yury Polyanskiy
MIT
yp@mit.edu

Abstract*—Most information systems store data by modifying the local state of the matter, in the hope that atomic (or sub-atomic) local interactions would stabilize the state for sufficiently long time, thereby allowing later recovery. In this work we initiate the study of information retention properties of locally-interacting systems. We model the time-dependent interactions between the different particles via the stochastic Ising model (SIM). The initial spin configuration X_0 serves as the user-controlled input. The output configuration X_t is produced by running t steps of the Glauber chain. Our main goal is to evaluate the information capacity $I_n(t) \triangleq \max_{p_{X_0}} I(X_0; X_t)$ when the time t scales with the size of the system n according to various rates. For the zero-temperature SIM on the two-dimensional $\sqrt{n} \times \sqrt{n}$ grid and free boundary condition, it is easy to show that $I_n(t) = \Theta(n)$ as long as $t = O(n)$. In addition, we show that order of $\sqrt{n}$ bits can be stored for infinite time (and even with zero error). The $\sqrt{n}$ achievability is optimal when $t \to \infty$ and n is fixed. Our main result is in extending achievability to super-linear (in n) times via a coding scheme that reliably stores more than $\sqrt{n}$ bits (in orders of magnitude). The analysis of the scheme decomposes the system into $\Omega(\sqrt{n})$ independent Z-channels whose crossover probability is found via the (recently rigorously established) Lifshitz law of phase boundary movement. Finally, two order optimal characterizations of $I_n(t)$, for all t, are given for the grid dynamics with an external magnetic field and for the dynamics on the Honeycomb lattice. It shown that $I_n(t) = \Theta(n)$ in both cases, suggesting their superiority over the grid without an external field for storage purposes.

I. Introduction

A. Storing Information Inside Matter

The predominant technology for long-term storage of digital information is based on physical effects such as magnetization of domains, or changes of meta-stable states of organic molecules. Data is written to the system by perturbing the local state of matter, e.g., by magnetizing particles to take one of two possible spins (represented by $+1$ and 1). In the time between writing and reading, the stored data degrades due to quantum/thermal fluctuations. These effects inspire the particles to interact - a physical phenomena we aim to capture in order to understand its influence on the duration of (reliable) storage. We adopt the physical model of an Ising spin system to describe the interplay between particles and allow manipulating spins at the particle level. By doing so we aim set the ground for the study of the fundamental notion of data storage inside matter, isolated from any particular technology. More specifically, this paper is motivated by the following question: how much information can be stored in a planar medium for a given time-duration?

This work was supported in part by the National Science Foundation CAREER award under grant agreement CCF-12-53205 and by the Center for Science of Information (CSoI), an NSF Science and Technology Center, under grant agreement CCF-09-39370

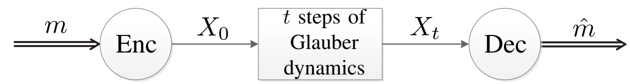


Fig. 1: A new model for storage inside matter: The encoder maps the data m into the initial configuration X_0. The channel progresses X_0 by t steps of Glauber dynamics. The decoder recovers m from X_t.

Typical models for the theoretical study of reliable data storage overlook the inherent time-dependent evolution of the system, and, in particular, the interactions between particles. Interactions are partially accounted for by run-length limited (RLL) codes [1], which mitigate unwanted patterns. On top of that, a certain (stationary, with respect to the sequence of input bits) error model is adopted and governs the relation between the input and output. Error-correcting codes are used to combat these errors. Thus, the storage problem is reduced to that of coding over a noisy channel. While the conversion to the well-understood channel coding problem is certainly helpful, this approach fails to capture central physical phenomena concerning the system's evolution in time.

We propose a new model for the study of storage in locally interacting systems, which accounts for the system's fluctuations in time. The time-evolution is modeled by Glauber dynamics [2] for the Ising model [3], also known as the *stochastic Ising model* (SIM). This widely accepted model for nonequilibrium ferromagnetic statistical mechanics [4] successfully describes interparticle interaction in real-life substances. In the proposed setup (Fig. 1), the encoder controls the initial configuration X_0. Once written, t steps of the Glauber chain are applied. This produces the output configuration X_t, from which the decoder tries to recover the message. The fundamental quantity we consider is the information capacity

$$\max_{p_{X_0}} I(X_0; X_t). \tag{1}$$

While mutual information (MI) estimation is of independent interest, the information capacity is shown to have the desired operational meaning: it characterizes (at least approximately) the number of bits that can be stored for t time-units in the storage medium. Our main focus is to evaluate (1) when t scales with the size of the system n according to various rates.

B. The Stochastic Ising Model

The SIM at inverse temperature $\beta > 0$ is a reversible Markov chain (MC) with the Gibbs measure as its stationary

Time	Information Capacity
$t = 0$	$I_n(t) = n$
$t = O(n)$	$I_n(t) = \Theta(n)$
$t = a(n) \cdot n$ where $a(n)$ is $o(n)$	$I_n(t) = \Omega\left(\frac{n}{a(n)}\right)$
$t \to \infty$ independent of n	$I_n(t) = \Theta(\sqrt{n})$

TABLE I: Main results for the zero-temperature SIM on the 2D grid

distribution [5, Chapter 15]. At each step, the discrete-time chain picks a site uniformly at random and generates a new spin for that site according to the Gibbs measure conditioned on all the other spins of the system. Accordingly, spins have a tendency to align, i.e., spins at adjacent sites favor having the same value. The lower the temperature is, the stronger is the influence of neighbouring spins on one another.

The literature on the SIM is too vast for this short paper to cover (see [5, Chapter 15] for a partial survey). Of particular interest here is the so-called zero-temperature SIM on the two-dimensional (2D) square lattice. Taking the limit of $\beta \to \infty$, the transition rule amounts to a majority update: the updated site takes the same spin as the majority of its neighbors, or, in case of a tie, draws is according to a fair coin toss. This process has been much studied in the physics literature as a model of "domain coarsening" (see, e.g., [6]). One of prominent rigorous results on this model concerns the disappearance time of an all-plus droplet in a sea of minuses. In [7] it was shown that a square droplet of side length L disappears in time $\Theta(L^2)$. This result, known as the Lifshitz law, plays a central role in analyzing one of the coding schemes in this work.

C. The Storage Problem and Contributions

The SIM's underlying graph describes the interactions between particles and corresponds to the topology of the modeled medium. This work assumes a planar topology, for which a natural choice is a 2D square grid of n vertices. At infinite temperature ($\beta = 0$), interactions are eliminated and, upon selection, particles flip with probability $\frac{1}{2}$, independently of their locality. Taking $t = cn$, the grid essentially becomes an n-fold binary-symmetric channel (BSC) with flip probability $\frac{1}{2}\left(1 - e^{-c/4}\right)$, which is arbitrarily close to $\frac{1}{2}$ for large c. Consequently, the per-site capacity is almost zero. Understanding whether introducing interactions (i.e., taking $\beta > 0$), enhances the capacity of the system is one of our main interests.

Classic results on the 2D Ising model phase transition and mixing times [8] imply the following: for $\beta < \beta_c$, where $\beta_c = \frac{1}{2}\log\left(1+\sqrt{2}\right)$, we have $I_n\big(\mathrm{poly}(n)\big) = 0$, while for $\beta > \beta_c$, $I_n\big(\exp(n)\big) \geq 1$.[1] A general analysis of the information capacity for $\beta > 0$, however, is an extremely challenging task.

As a first step towards the general $\beta > 0$ case, we endow the grid with the zero-temperature dynamics. The information capacity, denoted by $I_n(t)$, is first studied when n is fixed and $t \to \infty$. Beyond answering the question of 'how much information can be stored in the system forever?', this limiting quantity lower bounds $I_n(t)$, for all t. We characterize zero-temperature SIM as an absorbing MC, and identify the set of absorbing states (referred to as 'stable configurations'). A configuration is stable if and only if (iff) it has a (horizontal or vertical) striped pattern, with stripes of width at least two. As the number of stripes is of order $\sqrt{n}$, we obtain $\lim_{t\to\infty} I_n(t) = \Theta(\sqrt{n})$. Achievability follows by coding only over the stripes; the converse uses the MC's absorbing nature.

Next, Gilbert-Varshamov existence claim easily shows that up to linear times of approximately $\frac{n}{4}$, one can store a linear number of bits (which is order optimal). The main challenge, therefore, becomes understanding what happens in between these two regimes. A coding scheme that stores more than $\sqrt{n}$ bits (in order of magnitude) for super-linear times is of particular interest. We devise a scheme based on arranging monochromatic droplets (i.e., of all-plus or all-minus spins) in a sea of minuses. By growing the size of the droplets $a(n)$ as any $o(n)$ function, one can reliably store $\frac{n}{a(n)}$ bits for times up to $a(n) \cdot n$. We analyze the continuous-time version of the dynamics (easily shown to be equivalent for our purposes to the discrete-time version). The configurations in our codebook decouple into $\frac{n}{a(n)}$ independent MCs, and for each one, thresholding the number of pluses results in a Z-channel with positive capacity. The droplet survival time result from [7] and a tenzorization argument are used to conclude the analysis. Our main results are summarized in Table I.

Finally, we highlight two modifications to the zero-temperature dynamics for which storage performance significantly improves. Introducing an external magnetic field to the grid dynamics, gives rise to a tie-braking rule for updating sites with a balanced neighbourhood. This increases the size of the stable set from order of $\sqrt{n}$ to $\Theta(n)$, which implies that $I_n(t) = \Theta(n)$, uniformly in t. The same holds, without an external field, when the grid is replaced with the Honeycomb lattice. This follows from a tiling-based achievability scheme that exploits the odd degree (namely, 3) of all the vertices in the interior of the lattice.

II. Zero-Temperature Dynamics and Preliminaries

For $k \in \mathbb{N}$, we set $[k] \triangleq \{1, \ldots, k\}$. Let $G_n = (\mathcal{V}_n, \mathcal{E}_n)$ be a square grid of side $\sqrt{n} \in \mathbb{N}$, where $\mathcal{V}_n = \big\{(i,j)\big\}_{i,j\in[\sqrt{n}]}$.[2] The neighborhood of v is $\mathcal{N}_v \triangleq \{w \in \mathcal{V}_n | \{v,w\} \in \mathcal{E}_n\}$.

Fix $n \in \mathbb{N}$ and let $\Omega_n \triangleq \{-1,+1\}^{\mathcal{V}_n}$. For every $\sigma \in \Omega_n$ and $v \in \mathcal{V}_n$, $\sigma(v)$ denotes the value of σ at v. Given a configuration $\sigma \in \Omega_n$ and $v \in \mathcal{V}_n$, σ^v denotes the configuration that agrees with σ everywhere except v, i.e., $\sigma^v(u) = \sigma(u)\left(\mathbb{1}_{\{u\neq v\}} - \mathbb{1}_{\{u=v\}}\right)$. The all-plus and the all-minus configurations are denoted by $\boxplus$ and $\boxminus$, respectively. Let $m_v(\sigma) \triangleq \big|\{w \in \mathcal{N}_v | \sigma(w) = \sigma(v)\}\big|$ be the number of v's neighbours whose spin is $\sigma(v)$. Also set $\ell_v(\sigma) \triangleq |\mathcal{N}_v| - m_v(\sigma)$.

[1] The 2D SIM on the grid mixes within $O(n \log n)$ time when $\beta < \beta_c$, and exhibits exponential mixing of $e^{\Omega(\sqrt{n})}$ time, when $\beta > \beta_c$ [8].

[2] For convenience, we assume $\sqrt{n} \in \mathbb{N}$; if $\sqrt{n} \notin \mathbb{N}$, simple modification of some of the subsequent statements using ceiling and/or floor operations are needed. Regardless, our focus is on the asymptotic regime as $n \to \infty$, and the assumption that $\sqrt{n} \in \mathbb{N}$ has no affect on the asymptotic behavior.

Given a configuration $\sigma \in \Omega_n$, the zero-temperature SIM on G_n, which amounts to a majority update, evolves as follows:

1) Pick a vertex $v \in \mathcal{V}$ uniformly at random.
2) Modify σ at v as:
   - If $m_v(\sigma) > \ell_v(\sigma)$, keep the value of $\sigma(v)$;
   - If $m_v(\sigma) < \ell_v(\sigma)$, flip the value of $\sigma(v)$;
   - Otherwise, draw $\sigma(v)$ uniformly over $\{-1, +1\}$.

Let P be the corresponding transition kernel and $(X_t)_{t\in\mathbb{N}_0}$ be the induced MC on Ω_n. We use $\mathbb{P}$ for the probability measure, while $\mathbb{P}_\sigma$ indicates a conditioning on $\{X_0 = \sigma\}$. If $X_0 \sim p_{X_0}$, the distribution of (X_0, X_t) is $p_{X_0,X_t}(\sigma, \eta) \triangleq \mathbb{P}(X_0 = \sigma, X_t = \eta) = p_{X_0}(\sigma)P^t(\sigma, \eta)$, where P^t is the t-step kernel. The MI $I(X_0; X_t)$ is taken with respect to p_{X_0,X_t}.

III. Operational versus Information Capacity

Our main focus is on the asymptotic behaviour of the information capacity $I_n(t) \triangleq \max_{p_{X_0}} I(X_0; X_t)$. While the study of $I_n(t)$ for the Ising model is of independent interest, we are also motivated by coding for storage. This section briefly describes the operational problem and establishes $I_n(t)$ as a fundamental quantity in the study thereof. The rest of the paper deals only with $I_n(t)$.

For a fixed $\sqrt{n} \in \mathbb{N}$ and $t \in \mathbb{N}_0^+$, P^t is a channel from X_0 to X_t, referred to as the t-th Stochastic Ising Channel on G_n ($\mathsf{SIC}_n(t)$). The encoder controls X_0 and the decoder observes X_t (Fig. 1). The goal is to maintain reliable communication of a message $m \in [M]$ with the largest possible alphabet size.

Definition 1 (Code) *An (M, n, t, ϵ)-code for the $\mathsf{SIC}_n(t)$ is a pair of maps: the encoder $f_n^{(t)} : [M] \to \Omega_n$ and the decoder $\phi_n^{(t)} : \Omega_n \to [M]$, satisfying* $\frac{1}{M}\sum_{m\in[M]} \mathbb{P}_{f_n^{(t)}(m)}\left(\phi_n^{(t)}(X_t) \neq m\right) \leq \epsilon$.

Let $M^\star(n,t,\epsilon) = \max\left\{M \middle| \exists \text{ an } (M, n, t, \epsilon)\text{-code for } \mathsf{SIC}_n(t)\right\}$ be the largest code size with error probability at most ϵ. The next proposition relates $M^\star(n, t, \epsilon)$ and $I_n(t)$. Due to space limitations, the reader is referred to Appendix A of the full version of this work [9].

Proposition 1 *The following bounds on $M^\star(n, t, \epsilon)$ hold:*

1) Upper Bound: For any $n \in \mathbb{N}$, $t \geq 0$ and $\epsilon > 0$, we have

$$\log M^\star(n, t, \epsilon) \leq \frac{1}{1-\epsilon}\Big(I_n(t) + h(\epsilon)\Big),$$

where $h : [0, 1] \to [0, 1]$ is the binary entropy function.

2) Lower Bound: For $n_1 = o(n)$, $t \geq 0$ and $\epsilon > 0$, we have

$$\frac{1}{n}\log M^\star\left(n + o\left(\frac{n}{\sqrt{n_1}}\right), t, \epsilon\right) \geq \frac{1}{n_1} I_{n_1}(t) - \sqrt{\frac{n_1}{n(1-\epsilon)}}.$$

To interpret item (2), let $\alpha \in (0, 1)$ and $n_1 = n^{1-\alpha}$. This gives

$$\frac{1}{n}\log M^\star\left(n + o\left(n^{\frac{1+\alpha}{2}}\right), t, \epsilon\right) \geq \frac{1}{n^{1-\alpha}} I_{n^{1-\alpha}}(t) - \sqrt{\frac{1}{n^\alpha(1-\epsilon)}},$$

and approximates the normalized largest code size by the normalized information capacity of $G_{n^{1-\alpha}}$, for α however small.

IV. Infinite-Time Capacity

We focus now on $I_n^{(\infty)} \triangleq \lim_{t\to\infty} I_n(t)$, for fixed $n \in \mathbb{N}$. This is motivated by storage for infinite-time. Furthermore, thought the Data Processing Inequality (DPI), $I_n^{(\infty)}$ is a uniform (in t) lower bound on $I_n(t)$.

To characterize $I_n^{(\infty)}$, we identify $(X_t)_{t\in\mathbb{N}_0}$ as an absorbing MC. Because the chain inevitably lands in an absorbing state, the number of possible output configurations is equal to the number of absorbing states. To get the right dependence of $I_n^{(\infty)}$ on n, we first study the absorbing or *stable* set.

Definition 2 (Stable Configurations) *The set of stable configurations is defined as $\mathcal{S}_n \triangleq \left\{\sigma \in \Omega_n \middle| P(\sigma, \sigma) = 1\right\}$.*

We next show that $\mathcal{S}_n$ equals to the set of all striped configurations. A striped configuration partitions the grid into horizontal or vertical monochromatic stripes of width at least 2 (Fig. 2(a)); see [9, Definition 3]) for a formal definition. We use $\mathcal{A}_n$ for the set of all striped configurations.

The size of $\mathcal{A}_n$ can be found by mapping it to the number of binary sequences in a certain class and then relating the latter to a Fibonacci sequence (see [9, Proposition 2]). For our purposes here, it suffices to note that $|\mathcal{A}_n| = 2^{\Theta(\sqrt{n})}$.

Theorem 1 *Any $\sigma \in \Omega_n$ is stable iff it is striped, i.e., $\mathcal{S}_n = \mathcal{A}_n$.*

Proof Outline: The inclusion $\mathcal{A}_n \subseteq \mathcal{S}_n$ is trivial. We thus focus on the opposite inclusion. View the $\sqrt{n} \times \sqrt{n}$ grid as a board of n squares (each of side 1), such that each square is associated with a site $v \in \mathcal{V}_n$. Assigning a $+1$ (respectively, -1) spin to a site corresponds to coloring the appropriate square in, e.g., blue (respectively, red). Now, fix a stable configuration $\sigma \in \mathcal{S}_n$ and consider the corresponding $\sqrt{n} \times \sqrt{n}$ board of blue and red squares. The striped pattern of σ essentially follows by noting that while traveling along the borders between two monochromatic regions, one can never turn. Indeed, each change of orientation (from vertical to horizontal or vice versa) would reveal an unstable site that, if selected, would have a probability of at least $\frac{1}{2}$ to flip. The stripes' width must be at least 2 because any stripe of width 1 (even on the borders) contains unstable sites. Appendix C of [9] formulates this idea based on Peierl's contours. ■

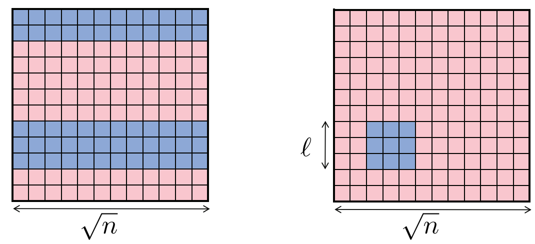
Next, we claim that $(X_t)_{t\in\mathbb{N}_0}$ is an absorbing MC. Namely, setting $\mathcal{T}_n \triangleq \Omega_n \setminus \mathcal{S}_n$, $(X_t)_{t\in\mathbb{N}_0}$ is an absorbing MC if for any $\rho \in \mathcal{T}_n$ there exist $t(\rho) \in \mathbb{N}$ such that $P^{t(\rho)}(\rho, \mathcal{S}_n) = \sum_{\sigma\in\mathcal{S}} P^{t(\rho)}(\rho, \sigma) > 0$. While $(X_t)_{t\in\mathbb{N}_0}$ being absorbing seems very intuitive, proving this relies on a rather intricate connectivity argument that is found in [9, Appendix D]).

Lemma 1 (Absorbing MC) *$(X_t)_{t\in\mathbb{N}_0}$ is an absorbing MC, and consequently,* $\lim_{t\to\infty} \max_{\sigma\in\Omega_n} \mathbb{P}_\sigma\big(X_t \notin \mathcal{S}_n\big) = 0$.

The convergence in probability above is a well-known property of absorbing MCs (see, e.g., [10, Chapter 11]). We are now ready to state and prove the infinite-time capacity result.

Theorem 2 (Infinite Time) *For any $n \in \mathbb{N}$, we have*

$$\log|\mathcal{S}_n| \leq I_n^{(\infty)} \leq \log|\mathcal{S}_n| + 1, \tag{2}$$



(a) (b)

Fig. 2: (a) A horizontally striped configuration; (b) A droplet configuration of size $\ell \times \ell$. Positive spins are represented by blue squares, while negative spins are depicted in red.

and, in particular, $I_n^{(\infty)} = \Theta(\sqrt{n})$.

Proof: For achievability, let $X_0 \sim \mathsf{Unif}(\mathcal{S}_n)$. Since $P(\sigma, \sigma) = 1$ for all $\sigma \in \mathcal{S}_n$, we have

$$I_n(t) \geq H(X_0) = \log |\mathcal{S}_n|, \quad \forall t \in \mathbb{N}_0. \tag{3}$$

For the upper bound, define $E_t \triangleq \mathbb{1}_{\{X_t \in \mathcal{S}_n\}}$. For any p_{X_0} and $t \in \mathbb{N}_0$, we have $I(X_0; X_t) = I(X_0; E_t) + I(X_0; X_t | E_t)$, and by expanding over E_t, one may verify that

$$I(X_0; X_t) \leq 1 + \log |\mathcal{S}_n| + n \cdot \max_{\sigma \in \Omega} \mathbb{P}_\sigma(X_t \notin \mathcal{S}_n). \tag{4}$$

Combining (3)-(4) and taking $t \to \infty$ concludes the proof. ■

V. Storing for Super-Linear Time

The previous section showed that for any $t \in \mathbb{N}_0$, one can achieve $I_n(t) = \Theta(\sqrt{n})$. The question, therefore, becomes: can we achieve a higher information capacity and, if yes, at what time scales? A simple initial observation is as follows.

Proposition 2 (Linear Time Order Optimality) *Fix $\epsilon > 0$. For $n \in \mathbb{N}$ and any $t < \left(\frac{1}{4} - \epsilon\right) n$, we have $I_n(t) = \Theta(n)$.*

The converse $I_n(t) \leq n$ is trivial, while achievability follows by the Gilbert-Varshamov existence claim, which implies that there exist error-correcting codes with minimum distance $d > \left(\frac{1}{4} + \epsilon\right) n$. Since until the aforementioned time there can occur at most $\left(\frac{1}{4} - \epsilon\right) n$ flips, the code is able to correct them. Thus, $I_n(t) > n \cdot \left[1 - h\left(\frac{1}{4} - \epsilon\right) - o(1)\right] = \Omega(n)$, for all $t < \left(\frac{1}{4} - \epsilon\right)$.

A. Beyond Linear Times - A Droplet-Based Scheme

To get an estimate on $I_n(t)$ beyond linear times, we propose a coding scheme that decomposes G_n into independent sub-squares, each capable of reliably storing a bit for $\omega(n)$ time. The decomposition relies on separating the sub-squares by all-minus stripes of width 2. This disentangles the dynamics inside the sub-squares and enables a tenzorization-based analysis.

Thus, we now focus on the evolution of a square droplet of positive spins surrounded by a sea of minuses. Fix $\ell \leq \sqrt{n} - 2$, and let $\mathcal{R}_\ell \subset \Omega_n$ be the set of configurations with all spins -1 except for those inside an $\ell \times \ell$ square, which are $+1$, such that the graph distance between the square and the boundary of the grid is at least one (see Fig. 2(b) for an illustration).

Consider a system initiated at $X_0 = \rho \in \mathcal{R}_\ell$ and let τ be the hitting time of the all-minus configuration $\boxminus$, i.e., $\tau \triangleq \inf\{t \in \mathbb{N}_0 | X_t = \boxminus\}$. The gaps between the borders of the square and the boundary of the grid ensure that the only stable configuration reachable from $\mathcal{R}_\ell$ is $\boxminus$. Thus, the original square shrinks with time until its disappearance.

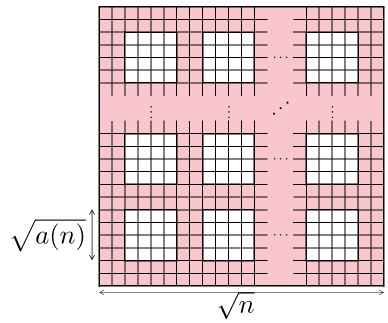


Fig. 3: Partitioning the grid into sub-squares of side length $\sqrt{a(n)}$. Red squares represent -1 spins; white squares are unspecified spins.

To approximate τ, it is convenient to consider the continuous-time version of the zero-temperature dynamics, which is effectively equivalent to the discrete-time chain via Poisson approximation. Although this equivalence is not proven in this paper (see [9, Proposition 4]), it is subsequently used for proving Theorem 4. The continuous-time dynamics are described as follows: When in state $\sigma \in \Omega_n$, each site $v \in \mathcal{V}_n$ is assigned with an independent Poisson clock of rate

$$c_{v,\sigma} = P(\sigma, \sigma^v) = \begin{cases} \frac{1}{n}, & m_v(\sigma) < \ell_v(\sigma) \\ \frac{1}{2n}, & m_v(\sigma) = \ell_v(\sigma) \\ 0, & m_v(\sigma) > \ell_v(\sigma). \end{cases} \tag{5}$$

When the clock at site v rings, the spin at v is flipped. Thus, if $\sigma(v)$ agrees with the majority/half/minority or the spins of its neighbors, it flips with rate $\frac{1}{n}$ / $\frac{1}{2n}$ a/ 0. The benefit from moving to the continuous-time dynamics is the decorrelation it induces between non-interacting portions of the G_n.

A landmark result from the zero-temperature SIM literature [7, Theorem 2] is that, with high probability, $\tau = \Theta\left(n\ell^2\right)$.

Theorem 3 (Erosion Time of a Square Droplet [7]) *For any $\rho \in \mathcal{R}_\ell$, there exist constants $c, C, \gamma > 0$, such that*

$$\mathbb{P}_\rho\left(cn\ell^2 \leq \tau \leq Cn\ell^2\right) \geq 1 - e^{-\gamma n \ell}, \quad \forall \ell \geq 1. \tag{6}$$

Our main result for super-linear time storage is given next.

Theorem 4 (Storing Beyond Linear Time) *For any $n \in \mathbb{N}$ let $a(n) = o\left(\sqrt{n}\right)$. Then there exists $C > 0$, such that for all $t \leq C \cdot a(n) \cdot n$, we have $I_n(t) = \Omega\left(\frac{n}{a(n)}\right)$.*

Proof: We move from discrete-time to continuous-time[3] and construct the input distribution of X_0 as follows. Tile the $\sqrt{n} \times \sqrt{n}$ grid with monochromatic sub-squares of side $\sqrt{a(n)}$ (whose spins are to be specified later) separated by all-minus stripes of width 2. The partitioning contains a total number of K^2 sub-squares, where $K = \left\lfloor \frac{\sqrt{n} - 3}{\sqrt{a(n)} - 2} \right\rfloor$ (Fig. 3).

Let $\mathcal{C}_n$ be the collection of configurations whose topology corresponds to Fig. 3 with monochromatic spin assignments to each of the $\Theta\left(\frac{n}{a(n)}\right)$ sub-squares (see [9] for the technical

[3] Abusing notation, we still use $(X_t)_{t \geq 0}$ to denote the corresponding MC.

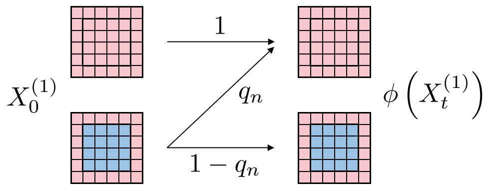


Fig. 4: The Z-channel between $X_0^{(1)}$ and $\phi\left(X_t^{(1)}\right)$ with crossover probability $q_n \triangleq \mathbb{P}(\tau_1 \leq C \cdot a(n) \cdot n)$.

definition). The collection $\mathcal{C}_n$ can be thought of as our codebook, where each droplet stores a bit. Accordingly, one may encode $K^2 = \Theta\left(\frac{n}{a(n)}\right)$ bits into the initial configuration by mapping bits to droplets with, say, a zero bit mapped to a -1 droplet. We next show that $I_n(t) \geq K^2$, for $t \leq C \cdot a(n) \cdot n$.

Let $X_0 \sim p_{X_0}$ with $\text{supp}(p_{X_0}) = \mathcal{C}_n$ (the exact distribution will be understood later), and note that by the very nature of the continuous-time dynamics, X_0 and X_t can be decomposed into K^2 independent components. Each is a sub-grid of side $\sqrt{a(n)} + 2$ corresponding to a different $\sqrt{a(n)} \times \sqrt{a(n)}$ square surrounded by an all-minus strip of width 1. This follows by the independent Poisson clocks that define the flip rates of $\left(X_t\right)_{t\geq 0}$ and because when $X_0 = \sigma \in \mathcal{C}_n$, interactions are confined to the original sub-squares (namely, the white regions in Fig. 3). Let $\left(X_0^{(i)}\right)_{i\in[K^2]}$ and $\left(X_t^{(i)}\right)_{i\in[K^2]}$ be the decomposition, respectively, of X_0 and X_t to the aforementioned independent components. We have,

$$I_n(t) \geq K^2 \cdot \max_{p_1} I\left(X_0^{(1)}; X_t^{(1)}\right) \triangleq K^2 \cdot I_1(t), \qquad (7)$$

where $X_0^{(1)}$ is a binary random variable that sets the 1st sub-square to $\boxminus$ with probability (w.p.) p_1, and sets it to an all-plus droplet (surrounded by a strip of minuses) w.p. $1 - p_1$.

Based on (7), to prove Theorem 4 it suffices to show that there exists $C > 0$ such that $I_1(t) > 0$, for all $t \leq C \cdot a(n) \cdot n$. Interestingly, this follows because the relation between $X_0^{(1)}$ and $X_t^{(1)}$ can be described in terms of a Z-channel. To see this denote the two possible values of $X_0^{(1)}$ by σ_- and σ_+, where $\sigma_- = \boxminus \in \Omega_{\sqrt{a(n)}+2}$, while σ_+ is an all-plus droplet of side $\sqrt{a(n)}$ surrounded by a strip of width 1 of minuses (see the right-hand side of Fig. 4). Define $\phi : \{\sigma_-, \sigma_+\} \to \{\sigma_-, \sigma_+\}$ as $\phi(\sigma) = \sigma_+ \cdot \mathbb{1}_{\{\sigma \neq \sigma_-\}} + \sigma_- \cdot \mathbb{1}_{\{\sigma = \sigma_-\}}$, and note that $X_0^{(1)}$ and $\phi\left(X_t^{(1)}\right)$ are related through the Z-channel from Fig. 4 with crossover probability $q_n \triangleq \mathbb{P}\left(\tau_1 \leq C \cdot a(n) \cdot n\right)$. By Theorem 3, there exist $C, \gamma > 0$, $q_n \leq e^{-\gamma \cdot a(n)}$, and consequently, $I_1(t) > 0$, for all $t \leq C \cdot a(n) \cdot n$. ∎

VI. Tight Information Capacity Results

A. 2D Grid with External Field

Introducing an (arbitrarily small) positive external magnetic field to the zero-temperature SIM on G_n serves as a tie-breaker. Namely, when in configuration $\sigma \in \Omega_n$ and updating $v \in \mathcal{V}_n$ with $m_v(\sigma) = \ell_v(\sigma)$, the spin at v is set to $+1$ w.p 1.

Theorem 5 (Information Capacity with External Field) *For the zero-temperature SIM on G_n with a positive external field, we have $I_n(t) = \Theta(n)$, for all $t \geq 0$.*

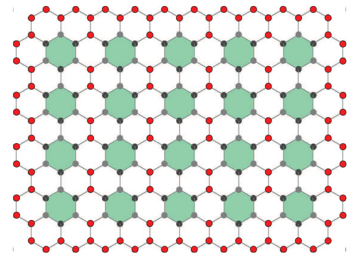


Fig. 5: Storing $\Omega(n)$ stable bits in the Honeycomb lattice: Each green region corresponds to a bit (written by assigning the same spin to all the vertices on the border of that region); the rest of the vertices are assigned with negative spins, shown by the red circles in the figure.

Theorem 5 essentially follows because, under the aforementioned tie-braking rule, the droplet configurations in $\mathcal{C}_n$ from the proof of Theorem 4, with $a(n) = 2$, are all stable. Thus, an (however small) external field is highly beneficial for storage purposes, as compared to the case where its absent.

B. The Honeycomb Lattice

Another interesting instance is when the underlying graph is the Honeycomb Lattice on n vertices H_n (Fig. 5; ignore the coloring for now) without an external field.

Theorem 6 (Honeycomb Lattice Information Capacity) *For the zero-temperature stochastic Ising model on the Honeycomb lattice H_n, we have $I_n(t) = \Theta(n)$, for all $t \geq 0$.*

The achievability proof of Theorem 6 uses the tiling of the Honeycomb lattice shown in 5. A stable bit can be stored in each colored region. To write a bit, all the sites at the border of that region are assigned with the same spin (say, negative for '0' and positive for '1'). The rest of the spins is the systems are set to -1. Any such configuration is stable. It is readily verified by doing so, $\Omega(n)$ bits can be stored in the system indefinitely. Thus, when external field is applied, the Honeycomb lattice is preferable (over the grid) for storage purposes.

References

[1] P. A. Franaszek. Sequence-state methods for run-length-limited coding. *IBM J. Res. Dev.*, 14(4):376–383, Jul. 1970.

[2] R. J. Glauber. Time-dependent statistics of the Ising model. *J. Mathematical Phys.*, 4(2):294–307, Feb. 1963.

[3] E. Ising. Beitrag zur theorie des ferromagnetismus. *Zeitschrift für Physik A Hadrons and Nuclei*, 31(1):253–258, Feb. 1925.

[4] S. Friedli and Y. Velenik. *Statistical mechanics of lattice systems: a concrete mathematical introduction*. Cambridge University Press, 2017.

[5] D. A. Levin and Y. Peres; with contributions by E. L. Wilmer. *Markov Chains and Mixing Times*. Princeton University Press, Providence, RI, 2nd edition edition, 2017.

[6] V. Spirin, P. L. Krapivsky, and S. Redner. Fate of zero-temperature Ising ferromagnets. *Phys. Rev. E*, 63(3):036118, Feb. 2001.

[7] P. Caputo, F. Martinelli, F. Simenhaus, and F. L. Toninelli. "*zero*" temperature stochastic 3D Ising model and dimer covering fluctuations: a first step towards interface mean curvature motion. *Comm. Pure Appl. Math.*, 64(6):778–831, Jun. 2011.

[8] F. Martinelli. Lectures on Glauber dynamics for discrete spin models. In *Lectures on prob. theory and stat.*, pages 93–191. Springer, 1999.

[9] Z. Goldfeld, Y. Polyanskiy, and G. Bresler. Information storage in the stochastic Ising model at zero temperature. *Submitted to IEEE Trans. Inf. Theory*, 2018. Available at https://arxiv.org/abs/1805.03027.

[10] C. M. Grinstead and J. L. Snell. *Introduction to probability*. American Math. Soc., 2012.