Majority dynamics on trees and the dynamic cavity method

Yashodhan Kanoria* and Andrea Montanari†

May 13, 2010

Abstract

A voter sits on each vertex of an infinite tree of degree k, and has to decide between two alternative opinions. At each time step, each voter switches to the opinion of the majority of her neighbors. We analyze this majority process when opinions are initialized to independent and identically distributed random variables.

In particular, we bound the threshold value of the initial bias such that the process converges to consensus. In order to prove an upper bound, we characterize the process of a single node in the large k-limit. This approach is inspired by the theory of mean field spin-glass and can potentially be generalized to a wider class of models. We also derive a lower bound that is non-trivial for small, odd values of k.

1 Introduction

1.1 The majority process

Consider a graph $\mathcal{G}$ with vertex set $\mathcal{V}$, and edge set $\mathcal{E}$. In the following, we shall denote by ∂i the set of neighbors of $i \in \mathcal{V}$, and assume $|\partial i| < \infty$ (i.e. $\mathcal{G}$ is locally finite). To each vertex $i \in \mathcal{V}$ we assign an initial spin $\sigma_i(0) \in \{-1, +1\}$. The vector of all initial spins is denoted by $\underline{\sigma}(0)$. Configuration $\underline{\sigma}(t) = \{\sigma_i(t) : i \in \mathcal{V}\}$ at subsequent times $t = 1, 2, \dots$ are determined according to the following majority update rule. If ∂i is the set of neighbors of node $i \in \mathcal{V}$, we let

$$\sigma_i(t+1) = \text{sign}\Big(\sum_{j \in \partial i} \sigma_j(t)\Big) \tag{1}$$

when $\sum_{j\in\partial i} \sigma_j(t) \neq 0$. If $\sum_{j\in\partial i} \sigma_j(t) = 0$, then we let

$$\sigma_i(t+1) = \begin{cases} \sigma_i(t) & \text{with probability } 1/2, \\ -\sigma_i(t) & \text{with probability } 1/2. \end{cases} \tag{2}$$

In order to construct this process, we associate to each vertex $i \in \mathcal{V}$, a sequence of i.i.d. Bernoulli(1/2) random variables $\mathcal{A}_i = \{A_{i,0}, A_{i,1}, A_{i,2} \dots\}$, whereby $A_{i,t}$ is used to break the (eventual) tie at time t. A realization of the process is then determined by the triple $(\mathcal{G}, \mathcal{A}, \underline{\sigma}(0))$, with $\mathcal{A} = \{\mathcal{A}_i\}$.

*Department of Electrical Engineering, Stanford University

†Department of Electrical Engineering and Department of Statistics, Stanford University

In this work we will study the asymptotic dynamic of this process when $\mathcal{G}$ is an infinite regular tree of degree $k \geq 2$. Let $\mathbb{P}_\theta$ be the law of the majority process where, in the initial configuration, the spins $\sigma_i(0)$ are i.i.d. with $\mathbb{P}_\theta\{\sigma_i(0) = +1\} = (1+\theta)/2$. We define the *consensus threshold* as the smallest bias in the initial condition such that the dynamics converges to the all $+1$ configuration

$$\theta_*(k) \;=\; \inf\left\{\theta :\; \mathbb{P}_\theta\left(\lim_{t\to\infty} \underline{\sigma}(t) = \underline{+1}\,\right) = 1\right\}\,. \tag{3}$$

Here convergence to the all-(+1) configuration is understood to be point-wise. We shall call $\theta_*(k)$ the *consensus threshold* of the k-regular tree.

Two simple observations will be useful in stating our results:

Monotonicity. Denote by $\succeq$ the natural partial ordering between configurations (i.e. $\underline{\sigma} \succeq \underline{\sigma}'$ if and only if $\sigma_i \geq \sigma_i'$ for all $i \in \mathcal{V}$). Then the majority dynamics preserves this partial ordering. More precisely, given two copies of the process with initial conditions $\underline{\sigma}(0) \succeq \underline{\sigma}'(0)$, there exists a coupling between the two processes such that $\underline{\sigma}(t) \succeq \underline{\sigma}'(t)$ for all $t \geq 0$.

Symmetry. Let $-\underline{\sigma}$ denote the configuration obtained by inverting all the spin values in $\underline{\sigma}$. Then two copies of the process with initial conditions $\underline{\sigma}'(0) = -\underline{\sigma}(0)$ can be coupled in such a way that $\underline{\sigma}'(t) = -\underline{\sigma}(t)$ for all $t \geq 0$.

It immediately follows from these properties that

$$0 \leq \theta_*(k) \leq 1\,.$$

In this work we prove upper and lower bounds on θ_*. The upper bound follows from an analysis of the majority process using a new technique that we call the *dynamic cavity method.* This technique provides a precise characterization of the spin trajectory, i.e. of the process $\{\sigma_i(t)\}_{t\geq 0}$ for a given vertex i. In particular, in the limit of large degree k, this becomes a function of a well defined gaussian process. Among other things, this characterization will be used to prove that

$$\theta_*(k) = O(1/k^M)\,,\quad \text{for any } M > 0\,.$$

Thus, $\theta_*(k)$ rapidly approaches 0 with increasing degree k. This result is stated below as Theorem 2.2.

We also prove lower bounds on $\theta_*(k)$ based on the formation of stable structures of -1 spins at time T. Such structures, once formed, persist for all future times, and hence prevent convergence to $\underline{\sigma}(1)$. These lower bounds $\theta_{\rm lb}(k,T)$ are non-trivial i.e. strictly positive for small odd values of k (cf. Table 2.4 in Section 2.4). This result is stated below as Theorem 2.8.

A significant part of this paper is devoted to the rigorous development of the dynamic cavity method. We consider this a key contribution of this work. The cavity method has been successful in analysis of probabilistic models having locally tree structured graphs [MPV87, Ta06, DM10]. The basic idea of this method is to remove a node from the graph thus forming a 'cavity'. One then assumes that the behavior of the other nodes (surrounding the cavity) is known. The removed node is then put back in to derive a dynamic-programming type recursion.

Here, we show how to extend this method to the study of a stochastic process on a tree-like graph, specifically the majority process. In this setting the cavity recursion can be interpreted as an inductive procedure with respect to time t. We 'fix' the behavior of a selected vertex i up to time t, obtain a consistent characterization of its 'environment' up to the same time t. From this, we can compute the probability distribution of the trajectory $\{\sigma_i(t')\}_{0\leq t'\leq t+1}$ up to time $t+1$. The cavity recursion determines completely the distribution of the the spin trajectory at an arbitrary node, although in implicit form.

In order to analyze the cavity recursion, we consider the large k regime. However, since we want to study the decay of $\theta_*(k)$ with k, we cannot rely on generic tools and need to carry out an accurate probability calculation. In order to achieve this goal, we establish a convenient form of the local central limit theorem for binary random vectors. We use this central limit theorem to 'solve' the cavity recursion for large k. The solution is given by a 'cavity process' that can be defined explicitly in terms of an appropriate gaussian process.

1.2 Preliminary remarks

It is not too difficult to show that $\theta_*(k) < 1$ for all k. A simple quantitative estimate is provided by the next result.

Lemma 1.1. *For all $k \geq 3$, denote by $\rho_c(k)$ the threshold density for the appearance of an infinite cluster of occupied vertices in bootstrap percolation with threshold $\lfloor (k+1)/2 \rfloor$. Then*

$$\theta_*(k) \leq \theta_u(k) \equiv 1 - 2\rho_c(k) < 1 \,. \tag{4}$$

This result follows from the fact that if the initial -1's cannot form an infinite structure under bootstrap percolation, then they eventually all disappear under the majority dynamics. We defer a full proof of this lemma to the Appendix A.

A numerical evaluation of this upper bound [FS08] yields $\theta_u(5) \approx 0.670$, $\theta_u(6) \approx 0.774$, $\theta_u(7) \approx 0.600$. It is possible to show that $\theta_u(k) = O\left(\sqrt{(\log k)/k}\right)$. It turns out that this is far from being the correct k-dependence. We will prove a much tighter bound in Theorem 2.2.

The next Lemma simplifies the task of proving upper bounds on $\theta_*(k)$ for large k, by showing that it is sufficient to prove $\mathbb{E}_\theta\{\sigma_i(t)\} > 1 - \delta_*/k$ for some constant δ_* to conclude $\underline{\sigma}(t) \to \underline{+1}$.

Lemma 1.2. *Assume $\mathcal{G}$ to be the regular tree of degree k. There exists $k_*, \delta_* > 0$ such that for $k \geq k_*$, if $\mathbb{E}_\theta\{\sigma_i(t)\} > 1 - (\delta_*/k)$, then $\theta_*(k) \leq \theta$.*

We use a standard expansion argument to show that such convergence occurs for typical random graphs in the configuration model, and then extend the result to the infinite tree. Again the proof can be found in Appendix A.

1.3 Organization of the paper

We state our main results in Section 2. Section 3 surveys related work. We develop the cavity method and the resulting upper bound in Section 4.2. Our lower bound is proved in Section 5.

2 Results

We can now state our main results. They consist of the following:

(*i*) The exact cavity recursion (Lemma 2.1 in Sec. 2.1).

(*ii*) Convergence to the cavity process (Theorem 2.6 and Theorem 2.7 in Sec. 2.2).

(*iii*) Upper bound on θ_* (Theorem 2.2 in Sec. 2.2) as a consequence of convergence to the cavity process.

(*iv*) Lower bound on θ_* (Theorem 2.8 in Sec. 2.3) due to formation of blocking structures of -1's.

Section 2.4 contains numerical illustration of some of our results.

2.1 The exact cavity recursion

First, we state an exact recursive characterization of spin trajectories at nodes. This is the key tool we use in our development of the cavity process. Moreover, our lower bound is based on a very similar recursive analysis.

Let $\mathcal{G}_\emptyset = (\mathcal{V}_\emptyset, \mathcal{E}_\emptyset)$ be the tree rooted at vertex $\emptyset$ with degree $k-1$ at the root and k at all the other vertices, and let $u = \{u(0), u(1), u(2), \dots\}$ be an arbitrary sequence of real numbers. We define a modified Markov chain over spins $\{\sigma_i\}_{i\in\mathcal{V}_\emptyset}$ as follows. For $i \neq \emptyset$, $\sigma_i(t)$ is updated according to the rules (1) and (2). For the root spin we have instead

$$\sigma_\emptyset(t+1) = \text{sign}\Big(\sum_{i=1}^{k-1} \sigma_i(t) + u(t)\Big)\,, \tag{5}$$

where $1, \dots, k-1$ denote the neighbors of the root. In the case $\sum_{i=1}^{k-1} \sigma_i(t) + u(t) = 0$, $\sigma_\emptyset(t+1)$ is drawn as in Eq. (2), i.e. uniformly at random. We will call this the 'dynamics under external field'.

We will call the sequence $u = \{u(0), u(1), u(2), \dots\}$ 'external field applied at the root.' We denote by $\mathbb{P}((\sigma_\emptyset)_0^T || u_0^T)$ the probability of observing a trajectory $(\sigma_\emptyset)_0^T = (\sigma_\emptyset(0), \sigma_\emptyset(1), \dots, \sigma_\emptyset(T))$ for root spin under the above dynamics. Let us stress two elementary facts: (i) $\mathbb{P}((\sigma_\emptyset)_0^T || u_0^T)$ is not a conditional probability; (ii) As implied by the notation, the distribution of $(\sigma_\emptyset)_0^T$ does not depend on $u(t)$, $t > T$ (and indeed does not depend on $u(T)$ either, but we include it for notational convenience).

As before, we assume that in the initial configuration, the spins are i.i.d. Bernoulli random variables, and denote by $\mathbb{P}_0(\sigma_i(0))$ their common distribution.

Lemma 2.1. *The following recursion holds*

$$\mathbb{P}((\sigma_\emptyset)_0^{T+1} || u_0^{T+1}) = \mathbb{P}_0(\sigma_\emptyset(0)) \sum_{(\sigma_1)_0^T \dots (\sigma_{k-1})_0^T} \prod_{t=0}^{T} \mathsf{K}_{u(t)}(\sigma_\emptyset(t+1)|\sigma_{\partial\emptyset}(t)) \prod_{i=1}^{k-1} \mathbb{P}((\sigma_i)_0^T || (\sigma_\emptyset)_0^T)\,, \tag{6}$$

$$\mathsf{K}_{u(t)}(\cdots) \equiv \begin{cases} \mathbb{I}\Big\{\sigma_\emptyset(t+1) = \text{sign}\Big(\sum_{i=1}^{k-1} \sigma_i(t) + u(t)\Big)\Big\} & \textit{if } \sum_{i=1}^{k-1} \sigma_i(t) + u(t) \neq 0\,, \\ \frac{1}{2} & \textit{otherwise.} \end{cases} \tag{7}$$

The correctness of this recursion for the majority dynamics is fairly self evident. We provide a rigorous proof in Appendix B.

2.2 Upper bounds and the dynamic cavity method

While for small odd k the consensus threshold is strictly positive, our next result shows that it approaches 0 very rapidly as $k \to \infty$.

Theorem 2.2. *The consensus threshold on k regular trees converges to 0 as $k \to \infty$ faster than any polynomial. In other words, for any $M > 0$, there exists $C(M) > 0$ such that*

$$\theta_*(k) \le C(M)\, k^{-M}\,. \tag{8}$$

Fix a vertex $i \in \mathcal{V}$, and consider the process $\{\sigma_i(t)\}_{t\ge0}$. The proof of Theorem 2.2 is obtained by developing a pretty complete characterization of this process in the large k limit. We first consider the unbiased case (i.e. $\theta = 0$) and prove the convergence of this process $\{\sigma_i(t)\}_{t\ge0}$. to a well-defined limit as $k \to \infty$. We will call this limit the *cavity process*, for the case of unbiased initialization (i.e. for $\theta = 0$). We formally define the cavity process below and then state our result on convergence to the cavity process

Definition 2.3 (Effective process)**.** *Let $C = \{C(t,s)\}_{t,s\in\mathbb{Z}_+}$ be a positive definite symmetric matrix, and $R = \{R(t,s)\}_{t,s\in\mathbb{Z}_+,\, t>s}$, $h = \{h(t)\}_{t\in\mathbb{Z}_+}$ two arbitrary sets of real numbers.*

A sample path of the effective process *with parameters C, R, h is generated as follows: Let $\sigma(0)$ be a Bernoulli$(1/2)$ random variable and $\{\eta(t)\}_{t\in\mathbb{Z}_+}$ be jointly Gaussian zero mean random variables with covariance C, independent from $\sigma(0)$. For any $t \geq 0$ we let*

$$\sigma(t+1) = \mathrm{sign}\left(\eta(t) + \sum_{s=0}^{t-1} R(t,s)\sigma(s) + h(t)\right). \tag{9}$$

Notice that the distribution of the effective process depends on the three parameters C, R, h. We will denote expectation with respect to its distribution as $\mathbb{E}_{C,R,h}$. The functions $C(\,\cdot\,,\,\cdot\,)$ and $R(\,\cdot\,,\,\cdot\,)$ will be referred to as *correlation* and *response* functions. By convention, we let $R(t,s) = 0$ if $t \leq s$. Finally, h is a perturbation parameter needed to state our definition of the cavity process in terms of the effective process.

Definition 2.4 (Consistent parameters C, R)**.** *We say that C, R are consistent if they satisfy*

$$C(t,s) = \mathbb{E}_{C,R,0}\left[\sigma(t)\sigma(s)\right] \qquad \forall\ t,s \geq 0\,, \tag{10}$$

$$R(t,s) = \frac{\partial}{\partial h(s)}\mathbb{E}_{C,R,h}[\sigma(t)]\bigg|_{h=0} \qquad \forall\ 0 \leq s < t\,. \tag{11}$$

It is natural to ask whether consistent choices of C and R exist, and in that case, whether they are unique or not. This question is addressed in Lemma 4.1 below, which proves that there exist unique consistent R and C, i.e. unique solution of Eqs. (10) and (11). In fact, these values are determined recursively. One starts $C(0,0) = 1$ (and indeed $C(t,t) = 1$ for all t). This leads to uniquely determined values for $C(1,0)$ and $R(1,0)$, which then determines unique values for $C(2,s)$, $R(2,s)$ and so on.

Definition 2.5 (Cavity process)**.** *Let C, R be the unique consistent parameters (cf. Definition 2.4) as per Lemma 4.1. The* cavity process *$\{\sigma(t)\}_{t\in\mathbb{Z}_+}$ defined as the effective process with parameters C, R and with $h = 0$.*

In the following we will denote by $\mathbb{P}_{\rm cav}$ the law of the cavity process. Our next theorem establishes convergence of the majority process with unbiased initialization to the cavity process.

Theorem 2.6. *Consider the majority process on a regular tree of degree k with uniform initialization $\theta = 0$. Then for any $i \in \mathcal{V}$ and $T \geq 0$, we have*

$$\lim_{k\to\infty} \mathbb{P}_{\theta=0}\big\{\, (\sigma_i(0),\ldots,\sigma_i(T)) = (\sigma(0),\ldots,\sigma(T))\,\big\} = \mathbb{P}_{\rm cav}\big\{\sigma(0),\ldots,\sigma(T)\big\}\,. \tag{12}$$

Let us describe the intuitive image which is at the basis of the last theorem.

The trajectory at target node i follows the majority rule in Eq. (1). The study of this rule is complicated by the fact that the spins of the neighboring nodes ∂i at time $t > 0$ are not independent of each other. The past trajectory of target node i, i.e. $(\sigma_i)_0^{t-1}$ affects the spins of nodes in ∂i at time t. The exact recursion Eq. (6) allows an analytical treatment despite this dependence. We use a local central limit theorem (Theorem 4.4 in Section 4.3, proved in Appendix E) on the exact recursion Eq. (6), to show convergence to the cavity process inductively in T. The response term $\sum_{s=0}^{t-1} R(t,s)\sigma(s)$ captures the effect of the spin trajectory up to time $t-1$ at the target node, on its environment at time t. The key part of the proof is in Lemma 4.5.

We finally turn to the case of biased initialization $\mathbb{E}_\theta\{\sigma_i(0)\} = \theta$.

Theorem 2.7. *For T_* a non-negative integer and $\omega_0 \geq 0$, consider the majority process on a regular tree of degree k with i.i.d. initialization with bias $\theta = \omega_0/k^{(T_*+1)/2}$. Then for any $i \in V$ and $T \leq T_*$, we have*

$$(\sigma_i(0), \sigma_i(1), \dots, \sigma_i(T)) \xrightarrow{\mathrm{d}} (\sigma_{\rm cav}(0), \sigma_{\rm cav}(1), \dots, \sigma_{\rm cav}(T))$$

where $\{\sigma_{\rm cav}(0)\}_{t\geq 0}$ is distributed according to the cavity process and convergence is understood to be in distribution as $k \to \infty$.

Further, if $\omega_0 > 0$, then for any $i \in V$ and $T \geq T_* + 2$, we have

$$(\sigma_i(0), \sigma_i(1), \dots, \sigma_i(T)) \xrightarrow{\mathrm{d}} (\sigma_{\rm cav}(0), \sigma_{\rm cav}(1), \dots, \sigma_{\rm cav}(T_*), \sigma(T_*+1), +1, +1, \dots, +1) \tag{13}$$

where the random variable $\sigma(T_*+1)$ dominates stochastically $\sigma_{\rm cav}(T_*+1)$, and $\mathbb{P}\{\sigma(T_*+1) > \sigma_{\rm cav}(T_*+1)\}$ is strictly positive.

Finally, there exist $A = A(\omega_0)$, with $A(\omega_0) > 0$ for $\omega_0 > 0$ such that, for any $T \geq T_* + 2$,

$$\mathbb{E}_\theta\{\sigma_i(T)\} \geq 1 - e^{-A(\omega_0)k}\,. \tag{14}$$

Theorem 2.2 is an immediate corollary of the last general result.

Proof of Theorem 2.2. Choose $T_* = 2M$ and $\omega_0 = 1$ in Theorem 2.7 and use Eq. (14) to check the assumptions of Lemma 1.2, whereby for $t \geq T_* + 2$, id is sufficient to take $k \geq k_*$ such that $\delta_*/k \geq e^{-A(\omega_0)k}$. □

Clearly, Theorem 2.6 is a special case of Theorem 2.7 (just take T_* large enough and $\omega_0 = 0$). However our proof proceeds by first analyzing the unbiased case $\theta = 0$, and then turning to the biased one $\theta > 0$. The latter is treated by establishing a delicate relationship between processes with biased and unbiased initializations, derived in Lemmas 4.6 and 4.7. In the unbiased case, one has $\mathbb{E}\{\sigma_i(t)\} = 0$ by symmetry at all times. In the biased case, we will prove a quantitative estimate of how the fraction of $+1$ spins evolves with time. Let $\theta_t = \mathbb{E}[\sigma_i(t)]$ for an arbitrary node i. If $\theta_t = O(k^{-1})$ we obtain $\theta_{t+1} = \sqrt{k}\,\theta_t\, R(t+1,t)\,(1+o(1))$. In words, as long as the fraction of $+1$ spins is small enough, it gets multiplied at each step by a factor of order $\sqrt{k}$. By iterating this procedure with $\theta_0 = \omega_0/k^{(T_*+1)/2}$, we get $\theta_{T_*} = \Theta(k^{-1/2})$, and $\theta_{T_*+1} = \Theta(1)$. At the next iteration, the fraction of $+1$ spins approaches 1 and the bias saturates to $\theta_{T_*+2} = 1 - e^{-\Theta(k)}$.

The above theorem also implies that, for large degree trees, consensus to majority takes place very abruptly. Indeed the bias towards $+1$ passes from $k^{-1/2}$ (at $t = T_*$) to $1 - e^{-\Theta(k)}$ (at $t = T_* + 2$) in 2 iterations. Numerical illustrations of this phenomenon are provided in Section 2.4, specifically Figures 1 and 2.

2.3 Lower Bounds

We state a sequence of recursively computable lower bounds.

Theorem 2.8. *Consider any $T \geq 0$. For all σ_0^T, $u_0^T \in \{-1,+1\}^{T+1}$ define*

$$\begin{aligned}\Psi^0_{{\rm odd},T}((\sigma_\varnothing)_0^T||u_0^T)) &= \mathbb{P}((\sigma_\varnothing)_0^T||u_0^T)\,,\\ \Psi^0_{{\rm even},T}((\sigma_\varnothing)_0^T||u_0^T)) &= \mathbb{P}((\sigma_\varnothing)_0^T||u_0^T)\,\mathbb{I}(\sigma_\varnothing(T) = -1)\,.\end{aligned} \tag{15}$$

Define $\Psi^{d+1}_{\text{odd},T}((\sigma_\varnothing)_0^T||u_0^T)), \Psi^{d+1}_{\text{even},T}((\sigma_\varnothing)_0^T||u_0^T))$ *for* $d\geq 0$ *recursively as per*

$$\Psi^{d+1}_{\text{odd},T}((\sigma_\varnothing)_0^T||u_0^T) = \mathbb{P}_0(\sigma_\varnothing(0)) \sum_{r=\lceil\frac{k+1}{2}\rceil-1}^{k-1} \binom{k-1}{r} \sum_{(\sigma_1)_0^T\dots(\sigma_{k-1})_0^T} \prod_{t=0}^{T-1} \mathsf{K}_{u(t)}(\sigma_\varnothing(t+1)|\sigma_{\partial\varnothing}(t))$$
$$\prod_{i=1}^{r} \Psi^{d}_{\text{even},T}((\sigma_i)_0^T||(\sigma_\varnothing)_0^T) \prod_{i=r+1}^{k-1} \Big(\mathbb{P}((\sigma_i)_0^T||(\sigma_\varnothing)_0^T) - \Psi^{d}_{\text{even},T}((\sigma_i)_0^T||(\sigma_\varnothing)_0^T)\Big)\,, \tag{16}$$

$$\Psi^{d+1}_{\text{even},T}((\sigma_\varnothing)_0^T||u_0^T) = \mathbb{I}(\sigma_\varnothing(T)=-1)\mathbb{P}_0(\sigma_\varnothing(0)) \sum_{r=\lceil\frac{k+1}{2}\rceil-1}^{k-1} \binom{k-1}{r} \sum_{(\sigma_1)_0^T\dots(\sigma_{k-1})_0^T} \prod_{t=0}^{T-1} \mathsf{K}_{u(t)}(\sigma_\varnothing(t+1)|\sigma_{\partial\varnothing}(t))$$
$$\prod_{i=1}^{r} \Psi^{d}_{\text{odd},T}((\sigma_i)_0^T||(\sigma_\varnothing)_0^T) \prod_{i=r+1}^{k-1} \Big(\mathbb{P}((\sigma_i)_0^T||(\sigma_\varnothing)_0^T) - \Psi^{d}_{\text{odd},T}((\sigma_i)_0^T||(\sigma_\varnothing)_0^T)\Big)\,, \tag{17}$$

$$\mathsf{K}_{u(t)}(\cdots) \equiv \begin{cases} \mathbb{I}\Big\{\sigma_\varnothing(t+1) = \text{sign}\Big(\sum_{i=1}^{k-1}\sigma_i(t)+u(t)\Big)\Big\} & \textit{if } \sum_{i=1}^{k-1}\sigma_i(t)+u(t)\neq 0\,, \\ \frac{1}{2} & \textit{otherwise.}\end{cases} \tag{18}$$

Let $\Psi_{\text{odd},T}(\sigma_0^T||u_0^T) = \lim_{d\to\infty}\Psi^d_{\text{odd},T}(\sigma_0^T||u_0^T)$. *This limit exists.*

Define $\theta_{\text{lb}}(k,T) \equiv \sup\{\theta\in[0,1]:\ \Psi_{\text{odd},T}(\sigma_0^T||u_0^T)>0$ *for all* $\sigma_0^T, u_0^T\}$. *Then, for every* k,T

$$\theta_*(k) \geq \theta_{\text{lb}}(k,T)\,. \tag{19}$$

It is obvious that evaluating the lower bound $\theta_{\text{lb}}(k,T)$ analytically is quite challenging. A exception is provided by the case $k=3$, where is not too hard to show that $\theta_{\text{lb}}(k=3,T=1)>0$.

We will instead evaluate the lower bounds $\theta_{\text{lb}}(k,T)$ numerically. The above recursion allows to do it through a number of operations (sums and multiplications) of order $2^{k(T+1)}T(T+k)$. As explained in Section 5, the recursion can be considerably simplified exploiting the symmetries of the problem, while remaining exponential in k and T. Evaluating the lower bound for $k=3$, 5, 7 and $T=3$ we get $\theta_*(3)>0.573$, $\theta_*(5)>0.052$, and $\theta_*(7)>0.0080$. This shows convincingly that $\theta_*(k)>0$ for $k\leq 7$, k odd.

2.4 Numerical illustration

The objective of this section is to provide illustrations of our results, and help to develop some intuition on the majority process.

It is obviously difficult to simulate the majority dynamics on infinite trees. On the other, hand the state of any node i after t iterations only depends on the state of its neighbors in the graph up to distance t. It is natural to consider sequences of finite graphs having an increasing number of vertices n, that converge locally to trees (in the sense of [AS03]). Random regular graphs drawn from the configuration model [Bol80] is a natural choice. A sequence of random k regular graphs does indeed converge to the regular tree of degree k almost surely [DM10].

Moreover, as demonstrated in Lemmas A.1 and A.2, the fraction of nodes that are $+1$ at time t in the configuration model converges to the probability in the infinite tree of an arbitrary node being $+1$.

It is worth emphasize that we are using random regular graphs as a tool for computing the evolution of the fraction of $(+1)$'s on the infinite tree. This approach is supported by Lemmas A.1 and A.2. On the

k	$\theta_{*,\mathrm{rgraph}}(k)$	Lower bd on $\theta_*(k)$ from Thm 2.8
3	0.58 ± 0.01	0.574
4	0.000 ± 0.001	0
5	0.054 ± 0.001	0.052
6	0.000 ± 0.001	0
7	0.010 ± 0.001	0.008

Table 1: Empirical thresholds $\theta_{*,\mathrm{rgraph}}(k)$ and computed lower bounds on $\theta_*(k)$

other hand, we will not attack the problem of defining a consensus threshold for finite graphs. This indeed requires some care as we briefly explain for clarity.

The consensus threshold θ_* is well defined for a general infinite graph $\mathcal{G}$. If $\mathcal{G}$ is finite, then trivially $\theta_*(\mathcal{G}) = 1$: indeed for any $\theta < 1$ there is a positive probability that $\underline{\sigma}(0)$ is the all -1 configurations. However, given a sequence of graphs with increasing number of vertices n, one can define a threshold function $\theta_{*,n}(\gamma)$ such that $\underline{\sigma}(t) \to \underline{+1}$ with probability γ for $\theta = \theta_{*,n}(\gamma)$. It is an open question to determine which graph sequences exhibit a sharp threshold (in the sense that $\theta_{*,n}(\gamma)$ has a limit independent of $\gamma \in (0,1)$ as $n \to \infty$). It is a natural conjecture that such a sharp threshold does indeed exist for sequences of random regular graphs.

We carried out numerical simulations with random regular graphs of degree k. [1] In this case, there appears empirically to be a sharp threshold bias that converges, as $n \to \infty$ to a limit $\theta_{*,\mathrm{rgraph}}(k)$. Above this threshold, the dynamics converges with high probability to all $+1$. Below this threshold, the dynamics converges instead to either a stationary point or to a length-two cycle [GO80]. Threshold biases found for small values of k are shown in Table 2.4.

The empirical threshold for the configuration model approaches 0 rapidly with increasing k, for k odd, and appears to be identically 0 for all even k. The origin of the odd-even difference lies in the fact that, for k odd, the majority dynamics is deterministic. For k even, the possibility of ties leads to random choices, cf. Eq. (2) thus reducing the chance of blocking structures. Getting a rigorous understanding of this phenomenon is an open problem.

For comparison, we have shown above the best lower bound value we could compute based on Theorem 2.8 (combined with the trivial lower bound of 0). The lower bounds we have obtained for the tree process are very close to the empirical thresholds $\theta_{*,\mathrm{rgraph}}(k)$. A full table of computed lower bound values is available in Table 5.3.

Figures 1 and 2 compare our predictions for the evolution of θ_t with the average observed values for finite values of k. Theorem 2.7 predicts almost complete consensus is reached sharply at iteration $T_* + 2$. We see that the prediction provided by our method is quite accurate already for $k \gtrsim 15$. In particular, consensus develops fairly rapidly between iteration T_* and $T_* + 2$.

[1] We used graphs of size up to $n = 5 \cdot 10^4$, generated according to a modified *configuration model* [Bol80] (with eventual self-edges and double edges rewired randomly). The initial bias was implemented by drawing a uniformly random configuration with $n(1+\theta)/2$ spins $\sigma_i = +1$.

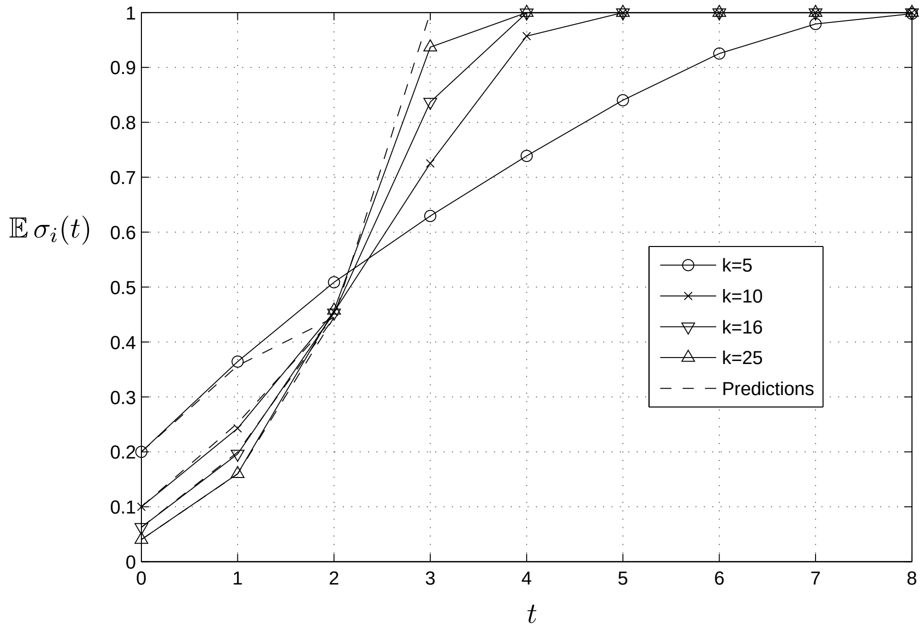

Figure 1: Change of bias $\mathbb{E}\sigma_i(t)$ over time t, with with initial bias $\mathbb{E}\sigma_i(0) \equiv \theta = 0.5/k$ (i.e., in our notation $T_* = 1$, $\omega_0 = 0.5$). The 'prediction' is based on $\omega_1, \ldots, \omega_{T_*}$ computed according to Eq. (39) and ω_{T_*+1} computed according to the modified cavity process (see Lemma 4.8 and Eq. (59)).

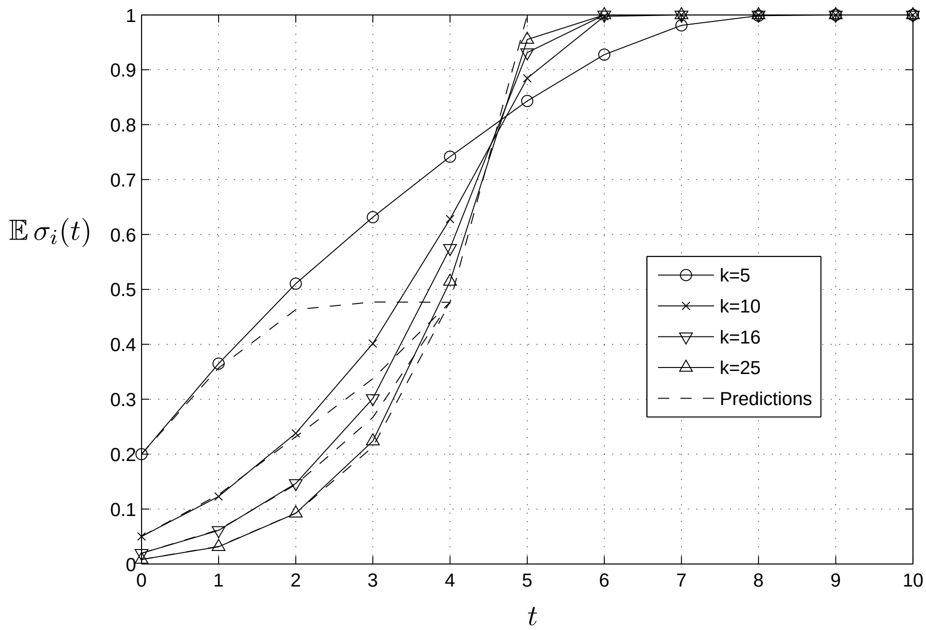

Figure 2: Change of bias $\mathbb{E}\sigma_i(t)$ over time t, with initial bias $\theta = \mathbb{E}\sigma_i(0) = 2.5/k^2$ (i.e. $T_* = 3$, $\omega_0 = 2.5$).

3 Related work

The majority process is a simple example of a stochastic dynamics evolving according to local rules on a graph. In the last few years, considerable effort has been devoted to the study of high-dimensional probability distributions with an underlying sparse graph structure [MM09]. Such distributions are referred to as Markov random fields, graphical models, spin models or constraint satisfaction problems, depending on the context. Common algorithmic and analytic tools were developed to address a number of questions ranging from statistical physics to computer science. Among such tools, we recall local weak convergence [AS03] and correlation decay [We06].

The objective of the present paper is to initiate a similar development in the context of stochastic dynamical processes that 'factor' according to a sparse graph structure. Rather than addressing a generic setting, we focus instead on a challenging concrete question, and try to develop tools that are amenable to generalization.

The majority process can be regarded as a example of interacting particle system [Lig85] or as a cellular automaton, two topics with a long record of important results. In particular, it bears some resemblance with the voter model. The latter is however considerably simpler because of the underlying martingale structure. Further, the voter model does not exhibit any sharp threshold for $\theta_*(k) < 1$.

More closely related to the model studied in this paper is the zero-temperature Glauber dynamics for the Ising model, which obeys the same update rule as in Eqs. (1), (2). Let us stress that Glauber dynamics is defined to be asynchronous: each spin is updated at the arrival times of an independent Poisson clock of rate 1. Fontes, Schonmann, and Sidoravicius [FSS02] studied this dynamics on d-dimensional grids, proving that the consensus threshold is $\theta_* < 1$ for all $d \geq 2$. Howard [H00] studied zero-temperature Glauber dynamics on 3-regular trees and found that infinite 'spin chains' of both signs are formed almost surely at positive times if we start with an unbiased initialization. In our notation, this implies $\theta_* > 0$ for this model. Positive-temperature Glauber dynamics on trees was the object of several recent papers [BK+05, MSW03]. While no 'complete consensus' can take place for positive temperature, at small enough temperature this model exhibits coarsening, namely the growth of a positively (or negatively) biased domain. In particular, Caputo and Martinelli [CM05] proved that the corresponding threshold $\theta_{*,\mathrm{coars}}(k) \to 0$ as $k \to \infty$.

As mentioned an important difference with respect to these studies lies in the fact that we focus on synchronous dynamics. Indeed our methods are somewhat simpler to apply to the synchronous case. Nevertheless we think that they can be generalized to the asynchronous setting as well. In particular, we expect that a limit theorem analogous to Theorem 2.6 (with a proper definition of the cavity process) can be proved for Glauber dynamics as well. More important is the difference between trees and grids. The methods developed in this paper are well-suited for analyzing stochastic processes on locally tree-like graphs, while a good part of the literature on Glauber dynamics focused on d-dimensional grids.

Variations of the majority dynamics on locally tree-like graphs have been studied recently within the statistical mechanics literature [HW+04, NB09]. In particular, the latter paper uses a non-rigorous version of the cavity method.

The main technical ideas developed in this paper are quite far from the ones within interacting particle systems. More precisely, we develop a dynamical analogue of the so-called 'cavity method' that has been successful in the analysis of probabilistic models on sparse random graphs. The basic idea in that context is to exploit the locally tree-like structure of such graphs to derive an approximate dynamic-programming type recursion. This idea was further developed mathematically in the local weak convergence framework of Aldous and Steele [AS03]. Adapting this framework to the study of a stochastic process is far from straightforward. First of all, one has to determine what quantity to write the recursion for. It turns out that an exact recursion can be proved for the probability distribution of the trajectory of the root spin in a modified majority process (see Section 2.1 for a precise definition). The next difficulty consists in extracting useful information from this recursion which is rather implicit and intricate. We demonstrate that this can be done for large k using an appropriate local central limit theorem proved in Appendix E. This allows to prove convergence to the cavity process, see Theorems 2.6 and 2.7.

The use of a dynamic cavity method for analyzing stochastic dynamics was pioneered in the statistical physics literature on mean field spin glasses. In that context one is typically interested in the asymptotic behavior of Langevin dynamics for large system sizes. The energy function is taken to be a spin-glass Hamiltonian, and the cavity method can be used to explore this asymptotics. A lucid (albeit non-rigorous) discussion can be found in [MPV87, Chapter VI]. This approach allows to derive limit deterministic equations for the covariance and the 'response function' of the process under study. The study of such equations lead to a deeper understanding of fascinating phenomena such as 'aging' in spin glasses [BC+97]. For some models, the limit equations were proved rigorously after a tour de force in stochastic processes theory [BDG06]. Theorem 2.6 presents remarkable structural similarities with these results. It suggests that this type of approach might be useful in analyzing a large array of stochastic dynamics on graphs.

Over the last couple of years, the cavity method has also been successfully applied in non-rigorous studies of quantum spin models on trees [KR+08, LSS08], a topic of interest in condensed matter physics. While this paper does treat quantum spin models, there are strong mathematical similarities similarities between the dynamic cavity method adopted here, and the cavity analysis of [KR+08, LSS08]. It would be interesting to adapt the rigorous methods developed here to the analysis of quantum models.

The majority process and similar models have been studied in the economic theory literature [Mor00, Kle07], within the general theme of 'learning in games'. In this context, each node corresponds to a strategic agent and each of the two states to a different strategy. The dynamics studied in this paper is just a best-response dynamics, whereby each agent plays a symmetric coordination game with each of its neighbors. It would be interesting to apply the present methodology to more general game-theoretic models.

4 The dynamic cavity method and proof of Theorem 2.7

The proof of Theorem 2.7 is organized as follows. Section 4.1 introduces some notations. We start by proving some basic properties of the cavity process in Section 4.2. We state a local central limit theorem for lattice random variables in Section 4.3. A proof of Theorem 2.6 follows in Section 4.4. Finally, in Section 4.5 we derive a delicate relationship between the biased and unbiased processes and prove Theorem 2.7.

4.1 Notations

Throughout this section we use the following notations. For a sequence $a(0), a(1), a(2), \dots$, and given $t \geq s$, we let $a_s^t \equiv (a(s), a(s+1), \dots, a(t))$. Further, given the correlation and response functions C and R, and an integer $T \geq 0$, we define the $(T+1)\times(T+1)$ matrices $C_T = \{C(t,s)\}_{t,s\leq T}$ and $R_T = \{R(t,s)\}_{s<t\leq T}$.

Given $m \in \mathbb{R}^d$ and $\Sigma \in \mathbb{R}^{d\times d}$, we let $\phi_{m,\Sigma}(x)$ be the density at $x \in \mathbb{R}^d$ of a Gaussian random variable with mean μ and covariance Σ. Finally, if $A \in \mathbb{R}^d$ is a rectangle, $A = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_d, b_d]$ (with

$a_i \leq b_i$), we let

$$\Phi_{m,\Sigma}(A) \equiv \int_A \phi_{m,\Sigma}(x) \prod_{i\in[d]:\, b_i>a_i} \mathrm{d}x_i\,. \tag{20}$$

Notice that those coordinates such that $a_i = b_i$ are not integrated over. For a partition $\{1,\dots,d\} = \mathcal{I}_0 \cup \mathcal{I}_+ \cup \mathcal{I}_-$, and a vector $a \in \mathbb{Z}^d$, define

$$A(a,\mathcal{I}) \equiv \{z \in \mathbb{Z}^d : z_i = a_i \;\forall\, i \in \mathcal{I}_0,\;\; z_i \geq a_i \;\forall\, i \in \mathcal{I}_+, z_i \leq a_i \;\forall\, i \in \mathcal{I}_-\}\,, \tag{21}$$

$$A_\infty(\mathcal{I}) \equiv \{z \in \mathbb{R}^d : z_i = 0 \;\forall\, i \in \mathcal{I}_0,\;\; z_i \geq 0 \;\forall\, i \in \mathcal{I}_+, z_i \leq 0 \;\forall\, i \in \mathcal{I}_-\}\,. \tag{22}$$

4.2 The cavity process

We start by checking that consistent R, C are uniquely defined, thus justifying the definition of the cavity process.

Lemma 4.1. *There exist unique consistent C, R (cf. Definition 2.4).*

Proof. For $T \geq 0$, let C_T, R_T denote the restriction of C, R to index values of at most T. Define

$$\mathcal{S}(T) \equiv \text{There exists unique } C_T, R_T, \text{ such that Eq. (10) is satisfied for all } s,t \leq T \text{ and Eq. (11) is satisfied for all } s < t \leq T$$

We want to show that $\mathcal{S}(T)$ holds for all $T \geq 0$. We proceed by induction. Clearly, $\mathcal{S}(0)$ holds with $C(0,0) = 1$.

Suppose $\mathcal{S}(T)$ holds. Denote by C_T and R_T the corresponding consistent covariance and response function, that exist and are unique by hypothesis. We construct consistent C_{T+1} and R_{T+1} by suitably extending C_T and R_T. Let $(\sigma(0), \sigma(1), \dots, \sigma(T+1))$ be a sample path of the uniquely defined effective process with parameters C_T, R_T and h as per Eq.(9). Define $C(s, T+1) = C(T+1, s)$ for all $s \leq T+1$ by Eq. (11) with $t = T+1$. Define $R(T+1, s)$ for all $s \leq T$ by Eq. (10) with $t = T+1$. The resulting C_{T+1}, R_{T+1} are clearly consistent up to $T+1$. Notice that C_{T+1} is positive semidefinite by construction. We now need to argue that there is no other consistent C_{T+1}, R_{T+1}. But this is clearly true since the restriction up to time T *must* match C_T, R_T for consistency up to T, and the extension to $T+1$ defined above is the only way to satisfy Eqs. (10) and (11) with $t = T+1$. Thus, $\mathcal{S}(T+1)$ holds. Induction completes the proof. □

Lemma 4.2. *Let $\{C(t,s)\}_{t,s\geq 0}$ be the correlation function of the cavity process. For any $T \geq 0$ the matrix C_T is strictly positive definite, and $\mathbb{P}(\sigma_0^T = \omega_0^T) > 0$ for each $\omega_0^T \in \{\pm 1\}^{T+1}$.*

Proof. As a preliminary remark notice that, by Lemma 4.1, $R(t,s)$ is well defined for all $s < t$. Moreover, it is easy to see that it is always finite.

We prove the lemma by induction. Clearly, C_0 is positive definite and $\mathbb{P}(\sigma(0) = \pm 1) = \frac{1}{2} > 0$. Suppose, C_T is positive definite. Now, from the definition of the cavity process, we have

$$\mathbb{P}(\sigma_0^{T+1} = \omega_0^{T+1}) = \frac{1}{2}\Phi_{\mu(\omega_0^T), C_T}(A_\infty(\mathcal{I}_C(\omega))), \tag{23}$$

where $\mathcal{I}_C(\omega_0^T)$ is the partition of $\{1, 2, \ldots, T\}$ defined as follows

$$\mathcal{I}_C(\omega) \equiv (\emptyset, \mathcal{I}_{C,+}, \mathcal{I}_{C,-}), \quad \mathcal{I}_{C,+} = \{i : \omega(i+1) = +1\}, \quad \mathcal{I}_{C,-} = \{i : \omega(i+1) = -1\}, \tag{24}$$

$$\mu(\omega_0^T) \equiv (\mu_0(\omega_0^T), \ldots, \mu_t(\omega_0^T)) \quad \text{with} \quad \mu_r(\omega_0^T) \equiv \sum_{s=0}^{r-1} R(r,s)\omega(s). \tag{25}$$

Since C_T is positive definite, we have $\Phi_{\mu(\omega_0^T),C_T}(x) > 0 \ \forall x \in \mathbb{R}^{T+1}$, whence $\mathbb{P}(\sigma_0^{T+1} = \omega_0^{T+1}) > 0$ for all $\omega_0^{T+1} \in \{-1,+1\}^{T+2}$. Notice that C_{T+1} is positive semidefinite by the definition of cavity process. If C_{T+1} is not strictly positive definite, there must be a linear combination of $(\sigma(0), \ldots, \sigma(T+1))$ that is equal to 0 with probability 1. Since the distribution of σ_0^{T+1} gives positive weight to each possible configuration, there must exist a non-trivial linear function in $\mathbb{R}^{T+2}$ that vanishes on very point of $\{\pm 1\}^{T+2}$, which is impossible. This proves that C_{T+1} is strictly positive definite. □

The above proof provides, in fact, a procedure to determine $C(t,s)$ and $R(t,s)$ by recursion over t. However, while the recursion for C consists just of a multi-dimensional integration over the Gaussian variables $\{\eta(t)\}$, the recursion for R, cf. Eq. (11) is a priori more complicated since it involves differentiation with respect to h. The next lemma provides more explicit expressions.

Lemma 4.3. *The correlation and response functions C and R of the cavity process are determined by the following recursion*

$$C(t+1,s) = \frac{1}{2} \sum_{\omega_0^{t+1} \in \{\pm 1\}^{(t+2)}} \omega(t+1)\omega(s)\ \Phi_{\mu(\omega_0^t),C_t}(A_\infty(\mathcal{I}_C(\omega))), \qquad \forall\, 0 \le s \le t, \tag{26}$$

$$R(t+1,s) = \frac{1}{2} \sum_{\omega_0^{t+1} \in \{\pm 1\}^{(t+2)}} \omega(t+1)\omega(s+1)\ \Phi_{\mu(\omega_0^t),C_t}(A_\infty(\mathcal{I}_R(\omega,s))), \qquad \forall\, 0 \le s \le t, \tag{27}$$

with boundary condition $R(t,s) = 0$ for $t \le s$, $C(t,t) = 1$ and $C(s,t) = C(t,s)$. Here, $\mathcal{I}_C$ and $\mathcal{I}_R$ are partitions of $\mathcal{T} = \{0, 1, \ldots, t\}$ of the form $\mathcal{I} \equiv (\mathcal{I}_0, \mathcal{I}_+, \mathcal{I}_-)$, with $\mathcal{I}_C$ and μ defined as per Eq. (25) with $T = t$ and $\mathcal{I}_R$ is defined by

$$\mathcal{I}_R(\omega,s) \equiv (\{s\}, \mathcal{I}_{R,+}, \mathcal{I}_{R,-}), \ \mathcal{I}_{R,+} = \{i : \omega(i+1) = +1\} \backslash \{s\}, \ \mathcal{I}_{R,-} = \{i : \omega(i+1) = -1\} \backslash \{s\}. \tag{28}$$

We provide a proof of this lemma in Appendix C.

Equation (27) yields in particular

$$R(t+1,t) = \sum_{\omega_0^t \in \{\pm 1\}^{t+1}} \Phi_{\mu(\omega_0^t),C_t}(A_\infty(\mathcal{I}_R(\omega,t))). \tag{29}$$

Note that $R(t+1,t) > 0 \ \forall\ t \ge 0$, since it is a sum of positive terms. These facts will be used later in Section 4.5.

The values of R and C evaluated for small values of s,t are as follows.

Note how $C(t,s) = 0$ when t and s have different parity, and $R(t,s) = 0$ when t and s have the same parity. This is a simple consequence of the fact that the dynamics is 'bipartite'. This also allows us to reduce the the dimensionality of integrals in Eqs. (26) and (27), making numerical computations easier.

	$s \rightarrow$			
	0	1	2	3
0	1			
t 1	0	1		
$\downarrow$ 2	0.5751	0	1	
3	0	0.7600	0	1

Table 2: Computed $C(t,s)$ values

	$s \rightarrow$			
	0	1	2	3
1	0.7979			
t 2	0	0.5804		
$\downarrow$ 3	0.4164	0	0.4607	
4	0	0.2920	0	0.3950

Table 3: Computed $R(t,s)$ values

4.3 A central limit theorem

In the following we will use repeatedly the following local central limit theorem for lattice random variables.

Theorem 4.4. *For any $B, d > 0$, there exist a finite constant $L = L(B,d)$ such that the following is true. Let $X_1, X_2, \dots, X_N$, be i.i.d. random vectors with $X_1 \in \{+1,-1\}^d$ and*

$$||\mathbb{E}X_1|| \;\; \leq \;\; \frac{B}{\sqrt{N}}\, , \qquad \min_{s\in\{+1,-1\}^d} \mathbb{P}(X_1 = s) \geq \frac{1}{B}\, .$$

Let p_N be the distribution of $S_N = \sum_{i=1}^N X_i$. For a partition $\{1,\dots,d\} = \mathcal{I}_0 \cup \mathcal{I}_+ \cup \mathcal{I}_-$, and a vector $a \in \mathbb{Z}^d$, with $||a||_\infty \leq B \log N$, define $A(a,\mathcal{I})$, $A_\infty(\mathcal{I})$ as in Eqs. (22), (21).

Assume the coordinates a_i to have the same parity as N. We then have

$$\sum_{y\in A(a,\mathcal{I})} p_N(y) = \frac{2^{|\mathcal{I}_0|}}{N^{|\mathcal{I}_0|/2}}\Phi_{\sqrt{N}\mathbb{E}X_1,\mathrm{Cov}(X_1)}(A_\infty(\mathcal{I}))\Big(1+\mathsf{Err}(a,\mathcal{I},N)\Big)\, , \tag{30}$$
$$|\mathsf{Err}(a,\mathcal{I},N)| \leq L(B,d)\, N^{-1/(2|\mathcal{I}_0|+2)}.$$

A simple proof of this result can be obtained using the Bernoulli decomposition method of [MD79, DMD94] and is reported in Appendix E. Indeed Appendix E proves a slightly stronger result.

4.4 Unbiased initialization: Proof of Theorem 2.6

Before passing to the details of the actual proof, we attempt to provide some intuition.

4.4.1 Theorem 2.6: Basic intuition

The central idea consists in studying the dynamics at the root of the rooted tree $\mathcal{G}_\varnothing = (\mathcal{V}_\varnothing, \mathcal{E}_\varnothing)$ with updates modified according to Eq. (5). The dynamics at the root is indeed completely characterized by the recursion (6). Let $y_0^T = \sum_{i=1}^{k-1}(\sigma_i)_0^T$, and write $\mathbb{P}(y_0^T||(\sigma_\varnothing)_0^T)$ for its distribution under the product measure $\prod_{i=1}^{k-1}\mathbb{P}((\sigma_i)_0^T||(\sigma_\varnothing)_0^T)$. Since $(\sigma_\varnothing)_0^T$ only depends on its neighbors through their sum y_0^T, all that matters is in fact the distribution $\mathbb{P}(y_0^T||(\sigma_\varnothing)_0^T)$. A further simplification arises in the large k limit because we can apply the central limit theorem to show that y_0^T converges to a gaussian random variable.

Two complications however arise: (i) The mean and variance of this gaussian depend in an *a priori* arbitrary way on $(\sigma_\varnothing)_0^T$ itself; (ii) In order to track this dependence, it is necessary to establish a central limit theorem for y_0^T.

In order to illustrate these points, it is useful to follow the first few steps of the dynamics. First take $T = 0$. We know that $\mathbb{P}(\sigma_i(0)||\sigma_\o(0)) = \mathbb{P}_0(\sigma_i(0)) = 1/2$. Thus, using Eq. (6) for $T = 0$, we get

$$\mathbb{P}((\sigma_\o)_0^1||u_0^1) = \frac{1}{2}^k \sum_{\sigma_1(0)\dots\sigma_{k-1}(0)} \mathsf{K}_{u(0)}(\sigma_\o(1)|\sigma_{\partial\o}(0)) \, .$$

This expression can be estimated by approximating $\mathbb{P}(y(0)||u(0))$ with the Gaussian distribution $\mathcal{N}(0, k-1)$. In particular using the expression (7) for $\mathsf{K}_{u(0)}(\sigma_\o(1)|\sigma_{\partial\o}(0))$ we get, for $(k-1)$ even and $u(0) \in \{+1, -1\}$

$$\begin{aligned}
\mathbb{E}(\sigma_\o(1)||u_0^1) &= \mathbb{E}\ \mathrm{sign}\Big(\sum_{i=1}^{k-1} \sigma_i(0) + u(0)\Big) \\
&= \mathbb{E}\ \mathrm{sign}\Big(y(0) + u(0)\Big) \\
&= u(0)\mathbb{P}\big(y(0) = 0\big) \\
&\approx u(0)\mathbb{P}\big(\sqrt{k-1}\, Z \in [-1, 1]\big) \approx \sqrt{\frac{2}{\pi k}}\, u(0) \, ,
\end{aligned}$$

where Z denotes a unit normal random variable. Using the fact that $\sigma_\o(0)$ is independent of $\sigma_\o(1)$ by the bipartite nature of the dynamics, we obtain the estimate

$$\mathbb{P}((\sigma_\o)_0^1||u_0^1) \approx \frac{1}{4}\left(1 + \frac{R(1,0)u(0)}{\sqrt{k}}\sigma_\o(1)\right) \, ,$$

where $R(1, 0) = \sqrt{2/\pi}$ as per Eq. (29). It follows that $\mathbb{E}_{\mathbb{P}(\cdot||u_0^1)}[\sigma_\o(1)] \approx R(1, 0)u(0)k^{-1/2}$. Also, $\mathbb{E}_{\mathbb{P}(\,\cdot\,||(u_0)_0^1)}[\sigma_\o(1)\sigma_\o(0)] = C(1, 0) = 0$ and $\mathbb{E}_{\mathbb{P}(\cdot||(u_0)_0^1)}[\sigma_\o(1)\sigma_\o(1)] = C(1, 1) = 1$. It follows that $\mathbb{P}(y_0^1||u_0^1)$ has a Gaussian approximation $\mathcal{N}\left(\sqrt{k}\,\mu((\sigma_\o)_0^1), k\,C_1\right)$, where $\mu((\sigma_\o)_0^1) = (0 \;\; R(1, 0)\sigma_\o(0))$. Note how the mean and standard deviation of y_0^1 are each of the same order $\Theta(\sqrt{k})$.

When passing to $T = 1$ in Eq. (6), we can make this normal approximation for the environment $y_0^1 = \sum_{i=1}^{k-1}(\sigma_i)_0^1$, up to time 1. We hence obtain a stochastic description of the root spin process up to time 2. Essentially the same argument is extended to any time T by induction, as is explained in detail in the following.

4.4.2 Theorem 2.6: The actual proof

The next Lemma rigorizes the above intuition and extends it to all times T by induction.

Lemma 4.5. *Let $T \geq 0$, and u_0^T with $u(t) \in \{+1, -1\}$ be given. Assume $(\sigma_\o)_0^T$ to be distributed according to $\mathbb{P}(\,\cdot\,||u_0^T)$. Then, as $k \to \infty$, we have*

$$|\mathbb{E}\{\sigma_\o(t)\sigma_\o(s)\} - C(t, s)| = o(1) \, , \qquad \left|\mathbb{E}\sigma_\o(t) - \frac{1}{\sqrt{k}}\sum_{s=0}^{t-1} R(t, s)\, u(s)\right| = o(k^{-1/2}) \, . \tag{31}$$

for all $t, s \leq T$. Further, for any u_0^T, $(\sigma_\o)_0^T \xrightarrow{\mathrm{d}} (\sigma_{\mathrm{ca}})_0^T$, with $(\sigma_{\mathrm{ca}})_0^T$ distributed according to the cavity process.

Proof. The proof is by induction on the number of steps T. Obviously the thesis holds for $T = 0$.

Assume that it holds up to time T. Consider the exact recursion Eq. (6) and fix a sequence $\sigma_\o(0), \cdots, \sigma_\o(T+1)$. Under the measure $\prod_{i=1}^{k-1} \mathbb{P}((\sigma_i)_0^T||(\sigma_\o)_0^T)$, the vectors $(\sigma_1)_0^T$, ..., $(\sigma_{k-1})_0^T$ are independent and identically distributed. Further, by the induction hypothesis

$$\mathbb{E}\sigma_1(t) = \frac{1}{\sqrt{k}} \sum_{s=0}^{t-1} R(t,s)\sigma_\o(s) + o(k^{-1/2}), \qquad \mathbb{E}\{\sigma_1(t)\sigma_1(s)\} = C(t,s) + o(1).$$

By central limit theorem $\{\frac{1}{\sqrt{k}}\sum_{i=1}^k \sigma_i(t)\}_{0\le t\le T}$ converge in distribution to

$$\left\{\eta(t) + \sum_{s=0}^{t-1} R(t,s)\sigma_0(s)\right\}_{0\le t\le T}, \tag{32}$$

where $\{\eta(t)\}_{0\le t\le T}$ is a centered Gaussian vector with covariance $\mathbb{E}\{\eta(t)\eta(s)\} = C(t,s)$. Since the product of indicator functions in Eq. (6) is a bounded function of the vector $\{\frac{1}{\sqrt{k}}\sum_{i=1}^k \sigma_i(t)\}_{0\le t\le T}$, and the normal distribution is everywhere continuous, we have

$$\lim_{k\to\infty} \mathbb{P}((\sigma_\o)_0^{T+1}||u_0^{T+1}) = \mathbb{P}_0(\sigma_\o(0))\mathbb{E}_\eta \left\{\prod_{t=0}^T \mathbb{I}\left(\sigma_\o(t+1) = \text{sign}\Big(\eta(t) + \sum_{s=0}^{t-1} R(t,s)\sigma_\o(s)\Big)\right)\right\} \tag{33}$$

i.e. $(\sigma_\o)_0^{T+1}$ converges in distribution to the first $T+1$ steps of the cavity process. This implies the first equation in (31). It is therefore sufficient to prove the second equation in (31), for $t = T+1$.

To get the estimate of the mean, we use again Eq. (6), and consider the distribution $\mathbb{P}((\sigma_0)_0^{T+1}||0_0^{T+1})$ whereby the root perturbation is set to 0. This satisfies the recursion Eq. (6), with $u(t) = 0$:

$$\mathbb{P}((\sigma_0)_0^{T+1}||0_0^{T+1}) = \mathbb{P}_0(\sigma_\o(0)) \sum_{(\sigma_1)_0^T\cdots(\sigma_{k-1})_0^T} \prod_{t=0}^T \mathsf{K}_0(\sigma_\o(t+1)|\sigma_{\partial\o}(t)) \prod_{i=1}^{k-1} \mathbb{P}((\sigma_i)_0^T||(\sigma_\o)_0^T). \tag{34}$$

Since $|u(t)| \le 1$, $K_{u(t)}(\cdots) = K_0(\cdots)$ for all values of t, except those in which $\sum_{i=1}^{k-1}\sigma_i(t) \in \{+1, 0, -1\}$. Let $\mathcal{I}_0 = \{t \ : \ |\sum_{i=1}^{k-1}\sigma_i(t)| \le 1\}$. Further, irrespective of $u(t)$, $\mathsf{K}_{u(t)}(\sigma_\o(t+1)|\sigma_{\partial\o}(t))$ is non vanishing only if $\sigma_0(t+1)\sum_{i=1}^{k-1}\sigma_i(t) \ge -1$. By taking the difference of Eq. (6) and (34), we get

$$\begin{aligned}
&\mathbb{P}((\sigma_\o)_0^{T+1}||u_0^{T+1}) - \mathbb{P}((\sigma_\o)_0^{T+1}||0_0^{T+1}) \\
&\quad = \mathbb{P}_0(\sigma_\o(0)) \sum_{(\sigma_1)_0^T\cdots(\sigma_{k-1})_0^T} \prod_{i=1}^{k-1} \mathbb{P}((\sigma_i)_0^T||(\sigma_\o)_0^T) \prod_{t=0}^T \mathbb{I}\left\{\sigma_\o(t+1)\sum_{i=1}^{k-1}\sigma_i(t) \ge -1\right\} \cdot \\
&\qquad \cdot \left(\prod_{t\in\mathcal{I}_0} \mathsf{K}_{u(t)}(\sigma_\o(t+1)|\sigma_{\partial\o}(t)) - \prod_{t\in\mathcal{I}_0} \mathsf{K}_0(\sigma_\o(t+1)|\sigma_{\partial\o}(t))\right).
\end{aligned} \tag{35}$$

Let $y_0^T = \sum_{i=1}^{k-1}(\sigma_i)_0^T$, and write $\mathbb{P}(y_0^T||(\sigma_\o)_0^T)$ for its distribution under the product measure $\prod_{i=1}^{k-1}\mathbb{P}((\sigma_i)_0^T||(\sigma_\o)_0^T)$. Further, let

$$\mathcal{I}_+ \equiv \{t : t < T, t \notin \mathcal{I}_0, \sigma_\o(t+1) = +1\}, \quad \mathcal{I}_- = \{t : t < T, t \notin \mathcal{I}_0, \sigma_0(t+1) = -1\}.$$

Then the above expression takes the form

$$\mathbb{P}((\sigma_\emptyset)_0^{T+1}||u_0^{T+1}) - \mathbb{P}((\sigma_\emptyset)_0^{T+1}||0_0^{T+1}) = \\ = \mathbb{P}_0(\sigma_\emptyset(0)) \sum_{y_0^T} \mathbb{P}(y_0^T||(\sigma_0)_0^T) \prod_{t\in\mathcal{I}_+} \mathbb{I}\{y(t) > 1\} \prod_{t\in\mathcal{I}_-} \mathbb{I}\{y(t) < -1\}\; f_{\mathcal{I}_0}(\{y(t)\}_{t\in\mathcal{I}_0})\,.$$

where we defined $f(\{y(t)\}_{t\in\mathcal{I}_0})$ to be the term in parentheses in Eq. (35).

Now, we can apply Theorem 4.4 for every possible $\mathcal{I}_0$, by letting $X_i = (\sigma_i)_0^T$, so that $d = T+1$, and $N = k-1$. Note that our induction hypothesis Eq. (31) on the mean implies that

$$\left| \mathbb{E}\sigma_i(t) - \frac{1}{\sqrt{k}} \sum_{s=0}^{t-1} R(t,s)\, \sigma_\emptyset(s) \right| = o(k^{-1/2}) \tag{36}$$

for all $t \le T$. In particular $||\mathbb{E}X_1|| \le B/\sqrt{k}$ as needed. Further, by Lemma 4.2, our induction hypothesis Eq. (31), and the convergence result (33), we have $\min_s \mathbb{P}\{X_1 = s\} \ge 1/B$ for all k large enough.

Now $f_{\mathcal{I}_0}(\{y(t)\}_{t\in\mathcal{I}_0}) = 0$ for $\mathcal{I}_0 = \emptyset$. From Theorem 4.4, the contribution for any $\mathcal{I}_0 \neq \emptyset$ is $\Theta\left(k^{-|\mathcal{I}_0|/2}\right)$. It follows that the dominating terms correspond to $\mathcal{I}_0 = \{t_0\}$. If we let $\mu'(\sigma_\emptyset) = \sqrt{k-1}\mathbb{E}[(\sigma_1)_0^T]$, $V(\sigma_\emptyset) = \mathrm{Cov}((\sigma_1)_0^T)$, then

$$\mathbb{P}((\sigma_\emptyset)_0^{T+1}||u_0^{T+1}) - \mathbb{P}((\sigma_\emptyset)_0^{T+1}||0_0^{T+1}) = \\ = \frac{2\mathbb{P}_0(\sigma_\emptyset(0))}{\sqrt{k-1}} \sum_{t_0=0}^{T} \Phi_{\mu'(\sigma_\emptyset),V(\sigma_\emptyset)}(A_\infty(\mathcal{I})) \sum_{|y(t_0)|\le 1} \left\{ \mathsf{K}_{u(t_0)}(\sigma_\emptyset(t_0+1)|y(t_0)) - \mathsf{K}_0(\sigma_\emptyset(t_0+1)|y(t_0)) \right\} \Big(1 + o(1)\Big)\,,$$

where, with an abuse of notation, we wrote $\mathsf{K}_\cdot(\sigma_\emptyset(t_0+1)|y(t_0))$ for $\mathsf{K}_\cdot(\sigma_\emptyset(t_0+1)|\sigma_{\partial\emptyset}(t_0))$ when $\sum_{i=1}^{k-1}\sigma_i(t_0) = y(t_0)$. Further, the rectangle $A_\infty(\mathcal{I})$ is defined as in Theorem 4.4.

If k is odd, then the only term in the above sum is $y(t_0) = 0$. An simple explicit calculation shows that

$$\mathsf{K}_{u(t_0)}(\sigma_\emptyset(t_0+1)|y(t_0) = 0) - \mathsf{K}_0(\sigma_\emptyset(t_0+1)|y(t_0) = 0) = u(t_0)\sigma_\emptyset(t_0+1)\,.$$

If k is even, two terms contribute to the sum: $y(t_0) = +1$ and $y(t_0) = -1$, with

$$\begin{aligned} \mathsf{K}_{u(t_0)}(\sigma_\emptyset(t_0+1)|y(t_0) = +1) - \mathsf{K}_0(\sigma_\emptyset(t_0+1)|y(t_0) = +1) &= -\sigma_\emptyset(t_0+1)\,\mathbb{I}(u(t_0) = -1)\,, \\ \mathsf{K}_{u(t_0)}(\sigma_\emptyset(t_0+1)|y(t_0) = -1) - \mathsf{K}_0(\sigma_\emptyset(t_0+1)|y(t_0) = -1) &= -\sigma_\emptyset(t_0+1)\,\mathbb{I}(u(t_0) = +1)\,. \end{aligned}$$

Also, by Eq. (36) we have $\lim_{k\to\infty}\mu'(\sigma_\emptyset) = \mu(\sigma_\emptyset)$ with $\mu(\,\cdot\,)$ defined as in Eq. (25). The induction hypothesis Eq. (31) on the covariance of $\sigma_\emptyset$ further implies $\lim_{k\to\infty} V(\sigma_\emptyset) = C_T$. By the continuity of Gaussian distribution we get

$$\lim_{k\to\infty} \Phi_{\mu'(\sigma_\emptyset),V(\sigma_\emptyset)}(A_\infty(\mathcal{I})) = \Phi_{\mu(\sigma_\emptyset),C_T}(A_\infty(\mathcal{I}))\,.$$

Applying these remarks to Eq. (42), and using the fact that $\mathbb{P}_0(\sigma_\emptyset(0)) = 1/2$, we finally get

$$\mathbb{P}((\sigma_\emptyset)_0^{T+1}||u_0^{T+1}) - \mathbb{P}((\sigma_\emptyset)_0^{T+1}||0_0^{T+1}) = \frac{1}{2\sqrt{k}} \sum_{t_0=0}^{T} \Phi_{\mu(\sigma_\emptyset),C_T}(A_\infty(\mathcal{I}))u(t_0)\sigma_\emptyset(t_0+1)\Big(1 + o(1)\Big)\,. \tag{37}$$

By symmetry, we have $\mathbb{E}_{\mathbb{P}(\cdot||0_0^{T+1})}[\sigma_\varnothing(T+1)] = 0$. By summing over $(\sigma_\varnothing)_0^T$ Eq. (37), we get

$$\mathbb{E}_{\mathbb{P}(\cdot||u_0^{T+1})}[\sigma_\varnothing(T+1)] = \frac{1}{\sqrt{k}} \sum_{t_0=0}^{T} u(t_0) \left(\frac{1}{2} \sum_{(\sigma_\varnothing)_0^{T+1}} \sigma_\varnothing(T+1)\sigma_\varnothing(t_0+1)\Phi_{\mu(\sigma_\varnothing),C_T}(A_\infty(\mathcal{I})) \right) \Big(1+o(1)\Big) .$$

It is easy to verify that the expression in parentheses matches the one for $R(T+1,t_0)$ from Lemma 4.3. Therefore we proved

$$\mathbb{E}_{\mathbb{P}(\cdot||u_0^{T+1})}[\sigma_\varnothing(T+1)] = \frac{1}{\sqrt{k}} \sum_{s\in\mathcal{I}} u(s)R(T+1,s) + o(1/\sqrt{k}) ,$$

which finishes the proof of the induction step. □

In the next section we will use this estimate to prove Theorem 2.7, which in particular implies Theorem 2.6. Let us notice however that Theorem 2.6 admits a direct proof as a consequence of the last lemma.

Proof. (Theorem 2.6) As part of Lemma 4.5, we have proved that $(\sigma_\varnothing)_0^T \xrightarrow{\text{d}} (\sigma_{\text{cav}})_0^T$ for each 'fixed' trajectory u_0^T, see Eq. (33). In particular, this holds for the extreme trajectories $(u_-)_0^T = (-1)_0^T$ and for $(u_+)_0^T = (+1)_0^T$. By monotonicity, the true trajectory of a spin σ_i in the regular tree $\mathcal{G}$ lies between the trajectories $(\sigma_{\varnothing,-})_0^T$ and $(\sigma_{\varnothing,+})_0^T$ distributed according to $\mathbb{P}(\,\cdot\,||(u_+)_0^T)$ and $\mathbb{P}(\,\cdot\,||(u_-)_0^T)$. Since both $(\sigma_{\varnothing,-})_0^T$ and $(\sigma_{\varnothing,+})_0^T$ converge in distribution to the cavity process $(\sigma_{\text{cav}})_0^T$, the original trajectory $(\sigma_i)_0^T$ converges to the cavity process as well. □

4.5 Biased initialization: Proof of Theorem 2.7

In this subsection we prove Theorem 2.7. The proof is based on Lemmas 4.6 and 4.7 that capture the asymptotic behavior of the recursion (6) as $k \to \infty$ in two different regimes.

Throughout this subsection, we adopt a special notation to simplify calculations. We reserve $\mathbb{P}((\sigma_\varnothing)_0^T||u_0^T)$ for the family of measures indexed by u_0^T and introduced in Section 2.1, in the case $\mathbb{P}_0(\sigma_\varnothing(0) = \pm 1) = 1/2$. We use instead $\mathbb{Q}((\sigma_\varnothing)_0^T||u_0^T)$ when the initialization is

$$\mathbb{Q}_0(\sigma_\varnothing(0) = \pm 1) = \frac{1}{2} \pm \frac{\omega_0}{k^{(T_*+1)/2}} ,$$

i.e. when in the initial configuration, the spins of $\mathcal{G}_\varnothing$ are i.i.d. Bernoulli with expectation $\mathbb{E}\sigma_i(0) = 2\omega_0 k^{-(T_*+1)/2}$.

Before providing the formal argument, we will describe the basic intuition.

4.5.1 Theorem 2.7: Basic intuition

The proof relies on a delicate comparison between the unbiased case (analyzed in the previous section) and the biased case treated here. An obvious but important fact is that, if the initialization is unbiased, then the distribution of $(\sigma_i)_0^T$ is exactly symmetric under inversion of all the spins.

How does the evolution with initial bias $\theta = 2\omega_0/k^{(T_*+1)/2}$, $\omega_0 > 0$ differ from the unbiased initialization trajectory? Consider the coupling between these two processes constructed as follows. Initialize the two evolution by drawing the initial spins at each vertex according to the optimal coupling between a

Bernoulli(1/2) and a Bernoulli(θ) random variable. At each subsequent step, the new spin values are either chosen deterministically, or according to a fair coin toss (in the case of a tie). In the former case, the coupling is obvious. In the latter –i.e. if a tie occurs in both processes at the same vertex– the coupling is constructed by using the same coin. Notice that this coupling is monotone: for $\theta > 0$ at each time the biased process dominates the unbiased one.

Denote by $\sigma_i(t)$ the spin at node i in the unbiased dynamics and by $\sigma_i'(t)$ the spin at node i in the biased dynamics. Since the coupling is monotone, the two processes can disagree at node i and time t only if $\sigma_i'(t) = +1$ and $\sigma_i(t) = -1$.

For concreteness, let us consider $T_* = 2$, i.e. $\theta = 2\omega_0/k^{3/2}$. At time $t = 0$, of the k neighbors of an arbitrary node i, on average $\omega_0/k^{1/2}$ of them will be different in the two processes. Further each neighbor will disagree or not independently of the others. Since $\omega_0/k^{1/2}$ is much smaller than 1, most often no neighbors will differ, occasionally 1 neighbor will differ and very rarely more than 1 neighbor will differ. For simplicity assume k is odd. The main event leading to $\sigma_i(t) \neq \sigma_i'(t)$ at $t = 1$ will be the following: the two process disagree on one neighbor of i (call this event $\mathcal{E}_1$), and the spins that do not disagree across the two processes add up to 0 (call this $\mathcal{E}_2$). Since $\mathbb{E}[\sigma_i(1)] = 0$ exactly, we have

$$\mathbb{E}[\sigma_i'(1)] = 2\,\mathbb{P}\{\sigma_i'(1) \neq \sigma_i(1)\} \approx 2\,\mathbb{P}(\mathcal{E}_1)\,\mathbb{P}(\mathcal{E}_2)\,.$$

The probability of $\mathcal{E}_1$ is estimated by the expected number of disagreements $\mathbb{P}(\mathcal{E}_1) \approx \omega_0/k^{1/2}$. The probability of a near tie on the other spins is instead estimated through a gaussian approximation as at the beginning of Section 4.4, $\mathbb{P}(\mathcal{E}_1) \approx \mathbb{P}(\sqrt{k}Z \in [-1,1]) \approx \sqrt{2/(\pi k)}$. One gets therefore

$$\mathbb{E}[\sigma_i'(1)] \approx 2R(1,0)\frac{\omega_0}{k} \equiv \frac{2\omega_1}{k}\,.$$

where ω_1 is defined as in the statement of Theorem 2.7.

The next steps follow along the same lines. Consider the neighbors of i at time $t = 1$. The two processes have a probability close to ω_1/k of disagreeing. Assuming that disagreements are again roughly independent of each other, we expect $O(\log k)$ disagreements at most. This leads, by an argument similar to the above, to $\mathbb{E}[\sigma_i'(2)] \approx 2R(2,1)\omega_1/\sqrt{k} \equiv 2\omega_2/\sqrt{k}$.

Finally consider the neighbors of i at time $t = 2$. We expect the two processes to disagree –on average– on $\sqrt{k}\omega_2$ neighbors. The two processes still agree on most of the neighbors but the sum of these spins is also –by central limit theorem– of order $\sqrt{k}$. This leads to $\mathbb{E}[\sigma_i'(3)] = \Theta(1)$. Continuing for one more step, we get $\mathbb{E}[\sigma_i'(4)] \approx 1$.

The next section will make this calculation rigorous, the main challenge being of course a precise control of dependencies. The simple coupling argument above is insufficient to achieve this goal. We will instead use once more the exact cavity recursion (6), together with appropriate analytical arguments.

4.5.2 Theorem 2.7: The actual proof

This intuition is rigorized by Lemmas 4.6 and 4.7. We use the modified dynamics introduced in Section 2.1, so that the exact cavity recursion in Lemma 2.1 can be used.

Lemma 4.6. *For $\sigma_0^T \in \{\pm 1\}^{T+1}$, let $\mathcal{I}_+ = \{t : \sigma(t+1) = +1\}$, $\mathcal{I}_- = \{t : \sigma(t+1) = -1\}$, and $\mathcal{I}_0 = \{T\}$. Define*

$$I_T(\sigma_0^T) = \Phi_{\mu(\sigma),C_T}(A_\infty(\mathcal{I}))\,, \tag{38}$$

where $\mu(\sigma) = (\mu_0(\sigma), \dots, \mu_T(\sigma))$ *with* $\mu_r(\sigma) = \sum_{s=0}^{r-1} R(r,s)\sigma(s)$*. Set by definition* $I_{-1} = 1$*. Finally, for* $0 \le T < T_* - 1$*, define* ω_{T+1} *recursively by*

$$\omega_{T+1} = R(T+1, T)\omega_T \,. \tag{39}$$

Then, for $0 \le T < T_* - 1$*, and for all* $(\sigma_\varnothing)_0^{T+1}, u_0^{T+1} \in \{\pm 1\}^{T+2}$*, we have*

$$\mathbb{Q}((\sigma_\varnothing)_0^{T+1} || u_0^{T+1}) - \mathbb{P}((\sigma_\varnothing)_0^{T+1} || u_0^{T+1}) = \frac{\omega_T}{k^{(T_*-T)/2}} \, \sigma_\varnothing(T+1) \; I_T\big((\sigma_\varnothing)_0^T\big) \, \Big(1 + o(1)\Big) \,. \tag{40}$$

Further, for all $u_0^{T+1} \in \{\pm 1\}^{T+2}$*, we have*

$$\sum_{(\sigma_\varnothing)_0^{T+1}} \sigma_\varnothing(T+1) \Big\{ \mathbb{Q}((\sigma_\varnothing)_0^{T+1} || u_0^{T+1}) - \mathbb{P}((\sigma_\varnothing)_0^{T+1} || u_0^{T+1}) \Big\} = \frac{2\omega_{T+1}}{k^{(T_*-T)/2}} \, (1 + o(1)) \,. \tag{41}$$

Proof. The proof is by induction over T, for $0 \le T < T_* - 1$, whereby in the base case ($T + 1 = 0$), Eq. (40) corresponds to

$$\mathbb{Q}_0(\sigma_\varnothing(0)) - \mathbb{P}_0(\sigma_\varnothing(0)) = \frac{\omega_0}{k^{(T_*+1)/2}} \, \sigma_\varnothing(0) \, \Big(1 + o(1)\Big) \,, \tag{42}$$

and holds by definition. Making use of Eq. (6) for both $\mathbb{P}$ and $\mathbb{Q}$, we get

$$\begin{aligned}
&\mathbb{Q}((\sigma_\varnothing)_0^{T+1} || u_0^{T+1}) - \mathbb{P}((\sigma_\varnothing)_0^{T+1} || u_0^{T+1}) \\
&\quad = \mathbb{Q}_0(\sigma_\varnothing(0)) \sum_{(\sigma_1)_0^T \dots (\sigma_{k-1})_0^T} \prod_{t=0}^{T} \mathsf{K}_{u(t)}(\sigma_\varnothing(t+1) | \sigma_{\partial\varnothing}(t)) \prod_{i=1}^{k-1} \mathbb{Q}((\sigma_i)_0^T || (\sigma_\varnothing)_0^T) \\
&\qquad\quad - \mathbb{P}_0(\sigma_\varnothing(0)) \sum_{(\sigma_1)_0^T \dots (\sigma_{k-1})_0^T} \prod_{t=0}^{T} \mathsf{K}_{u(t)}(\sigma_\varnothing(t+1) | \sigma_{\partial\varnothing}(t)) \prod_{i=1}^{k-1} \mathbb{P}((\sigma_i)_0^T || (\sigma_\varnothing)_0^T) \\
&\quad = \frac{1}{2} \sum_{(\sigma_1)_0^T \dots (\sigma_{k-1})_0^T} \prod_{t=0}^{T} \mathsf{K}_{u(t)}(\sigma_\varnothing(t+1) | \sigma_{\partial\varnothing}(t)) \cdot \qquad (43) \\
&\qquad\quad \cdot \left\{ \prod_{i=1}^{k-1} \mathbb{Q}((\sigma_i)_0^T || (\sigma_\varnothing)_0^T) - \prod_{i=1}^{k-1} \mathbb{P}((\sigma_i)_0^T || (\sigma_\varnothing)_0^T) \right\} + O(k^{-(T_*+1)/2})
\end{aligned}$$

Now we use

$$\begin{aligned}
&\prod_{i=1}^{k-1} \mathbb{Q}((\sigma_i)_0^T || (\sigma_\varnothing)_0^T) - \prod_{i=1}^{k-1} \mathbb{P}((\sigma_i)_0^T || (\sigma_\varnothing)_0^T) = \\
&\qquad \sum_{r=1}^{k-1} \binom{k-1}{r} \prod_{i=1}^{r} \big(\mathbb{Q}((\sigma_i)_0^T || (\sigma_\varnothing)_0^T) - \mathbb{P}((\sigma_i)_0^T || (\sigma_\varnothing)_0^T) \big) \prod_{i=r+1}^{k-1} \mathbb{P}((\sigma_i)_0^T || (\sigma_\varnothing)_0^T)
\end{aligned}$$

in Eq. (43) to obtain

$$\mathbb{Q}((\sigma_\varnothing)_0^{T+1} || u_0^{T+1}) - \mathbb{P}((\sigma_\varnothing)_0^{T+1} || u_0^{T+1}) = \frac{1}{2} \sum_{r=1}^{k-1} \mathsf{D}(r,k) + O(k^{-(T_*+1)/2}) \,, \tag{44}$$

where

$$\mathsf{D}(r,k) \equiv \binom{k-1}{r} \sum_{(\sigma_1)_0^T \dots (\sigma_{k-1})_0^T} \prod_{t=0}^{T} \mathsf{K}_{u(t)}(\sigma_\emptyset(t+1)|\sigma_{\partial\emptyset}(t))\cdot \tag{45}$$

$$\cdot \prod_{i=1}^{r} \left(\mathbb{Q}((\sigma_i)_0^T||(\sigma_\emptyset)_0^T) - \mathbb{P}((\sigma_i)_0^T||(\sigma_\emptyset)_0^T)\right) \prod_{i=r+1}^{k-1} \mathbb{P}((\sigma_i)_0^T||(\sigma_\emptyset)_0^T) \tag{46}$$

We claim that only the term $r=1$ is relevant for large k:

$$\sum_{r=2}^{k-1} |\mathsf{D}(r,k)| = o(k^{-(T_*-T)/2})\,. \tag{47}$$

Before proving this claim, let us show that it implies the thesis. Set $r_0 = 1$ (we introduce this notation because the calculation below holds for larger values of r_0 and this fact will be exploited in the next lemma).

The $r=1$ term can be rewritten as

$$\mathsf{D}(1,k) = (k-1) \sum_{\{(\sigma_i)_0^T\}} \prod_{t=0}^{T} \mathsf{K}_{u(t)}(\sigma_\emptyset(t+1)|\sigma_{\partial\emptyset}(t+1)) \big\{\mathbb{Q}((\sigma_1)_0^T||(\sigma_\emptyset)_0^T) - \mathbb{P}((\sigma_1)_0^T||(\sigma_\emptyset)_0^T)\big\} \prod_{i=2}^{k-1} \mathbb{P}((\sigma_i)_0^T||(\sigma_\emptyset)_0^T)\,.$$

For $t \in \{0,1,\dots,T\}$, let

$$\mathcal{S}_t \equiv \Big\{(\sigma_2)_0^T \dots (\sigma_{k-1})_0^T : |\sigma_2(t) + \dots + \sigma_{k-1}(t) + u(t)| \le r_0\Big\}\,. \tag{48}$$

If $(\sigma_2)_0^T \dots (\sigma_{k-1})_0^T$ is not in $\cup_{t=0}^T \mathcal{S}_t$, then the sum over $(\sigma_1)_0^T$ can be evaluated immediately (as $\mathsf{K}_{u(t)}(\cdots)$ is independent of $(\sigma_1)_0^T$) and is equal to 0 due to the normalization of $\mathbb{Q}(\,\cdot\,||(\sigma_\emptyset)_0^T)$ and $\mathbb{P}(\,\cdot\,||(\sigma_\emptyset)_0^T)$. We can restrict the innermost sum to $(\sigma_2)_0^T \dots (\sigma_{k-1})_0^T$ in $\cup_{t=0}^T \mathcal{S}_t$, i.e. $|\sum_{i=2}^{k-1}\sigma_i(t) + u(t)| \le r_0$ for some $t \in \{0,\dots,T\}$. Let $\mathcal{I}_0 \subseteq \{0,\dots,T\}$ be the set of times such that this happens.

The expectation over $(\sigma_2)_0^T$, ..., $(\sigma_{k-1})_0^T$ can be estimated applying Theorem 4.4, with $N = k-2$, and using Lemmas 4.2 and 4.5 to check that the hypotheses 4.4 hold for all k large enough. Using the induction hypothesis $|\mathbb{Q}((\sigma_1)_0^T||(\sigma_\emptyset)_0^T) - \mathbb{P}((\sigma_1)_0^T||(\sigma_\emptyset)_0^T)| = O(k^{-(T_*-T+1)/2})$, this implies that the contribution of terms with $|\mathcal{I}_0| \ge 2$ is upper bounded as $kO(k^{-(T_*-T+1)/2})^2 = o(k^{-(T_*-T)/2})$ (for $T \le T_*-1$). Therefore we make a negligible error if we restrict ourselves to the case $|\mathcal{I}_0| = 1$.

If we let $\widehat{\mathcal{S}}_{t_0} \equiv \mathcal{S}_{t_o} \cap \{\cap_{t\neq t_0}\overline{\mathcal{S}_t}\}$, we then have

$$\mathsf{D}(1,k) \equiv (k-1) \sum_{t_0=0}^{T} \sum_{(\sigma_1)_0^T} \left(\mathbb{Q}((\sigma_1)_0^T||(\sigma_\emptyset)_0^T) - \mathbb{P}((\sigma_1)_0^T||(\sigma_\emptyset)_0^T)\right)\cdot \tag{49}$$

$$\cdot \sum_{((\sigma_2)_0^T \dots (\sigma_{k-1})_0^T) \in \widehat{\mathcal{S}}_{t_0}} \prod_{i=2}^{k-1} \mathbb{P}((\sigma_i)_0^T||(\sigma_\emptyset)_0^T)) \prod_{t=0}^{T} \mathsf{K}_{u(t)}(\sigma_\emptyset(t+1)|\sigma_{\partial\emptyset}(t)) \,+\, o(k^{-(T_*-T)/2})\,.$$

Consider the main term

$$J'_{t_0}((\sigma_\emptyset)_0^T, (\sigma_1)_0^T) \equiv \sum_{((\sigma_2)_0^T \dots (\sigma_{k-1})_0^T) \in \widehat{\mathcal{S}}_{t_0}} \prod_{i=2}^{k-1} \mathbb{P}((\sigma_i)_0^T||(\sigma_\emptyset)_0^T)) \prod_{t=0}^{T} \mathsf{K}_{u(t)}(\sigma_\emptyset(t+1)|\sigma_{\partial\emptyset}(t))\,. \tag{50}$$

The arguments of this function will often be dropped in what follows, and we will simply write J'_{t_0}. For $t \neq t_0$, the kernel $\mathsf{K}_{u(t)}(\sigma_\o(t+1)|\sigma_{\partial\o}(t))$ can be replaced by an indicator function, and the constraint $((\sigma_2)_0^T \dots (\sigma_{k-1})_0^T) \in \widehat{\mathcal{S}}_{t_0}$ can be removed. For $t = t_0$ we write

$$\mathsf{K}_{u(t_0)}(\sigma_\o(t_0+1)|\sigma_{\partial\o}(t_0)) = \widehat{\mathsf{K}}'_{\Omega(t_0)}\left\{\sigma_\o(t_0+1)\Big(u(t_0) + \sum_{i=2}^{k-1}\sigma_i(t_0)\Big)\right\}$$

where

$$\widehat{\mathsf{K}}'_a(x) = \begin{cases} 1 & \text{if } -a < x \le r_0, \\ 1/2 & \text{if } x = -a, \\ 0 & \text{otherwise,} \end{cases}$$

and $\Omega(t) = \sigma_\o(t+1)\sigma_1(t)$, $|\Omega(t)| \le r_0$. We thus have

$$J'_{t_0} = \sum_{(\sigma_2)_0^T\dots(\sigma_{k-1})_0^T}\prod_{i=2}^{k-1}\mathbb{P}((\sigma_i)_0^T||(\sigma_\o)_0^T))\,\widehat{\mathsf{K}}'_{\Omega(t_0)}\left\{\sigma_\o(t_0+1)\Big(u(t_0)+\sum_{i=2}^{k-1}\sigma_i(t_0)\Big)\right\}\cdot$$
$$\cdot\prod_{t=0}^{T}\mathbb{I}\left\{\sigma_\o(t_0+1)\Big(u(t)+\sum_{i=2}^{k-1}\sigma_i(t_0)\Big) > r_0\right\}.$$

Notice that the only dependence on $(\sigma_1)_0^T$ is through $\Omega(t_0)$. Therefore, we can replace $\widehat{\mathsf{K}}'_{\Omega(t_0)}\{\,\cdot\,\}$ by $\widehat{\mathsf{K}}_{\Omega(t_0)}\{\,\cdot\,\} = \widehat{\mathsf{K}}'_{\Omega(t_0)}\{\,\cdot\,\} - \widehat{\mathsf{K}}'_0\{\,\cdot\,\}$ because the difference, once integrated over $(\sigma_1)_0^T$ as in Eq. (49), vanishes by the normalization of $\mathbb{Q}(\,\cdot\,||(\sigma_\o)_0^T)$ and $\mathbb{P}(\,\cdot\,||(\sigma_\o)_0^T)$. We thus need to evaluate

$$J_{t_0} = \sum_{(\sigma_2)_0^T\dots(\sigma_{k-1})_0^T}\prod_{i=2}^{k-1}\mathbb{P}((\sigma_i)_0^T||(\sigma_\o)_0^T))\,\widehat{\mathsf{K}}_{\Omega(t_0)}\left\{\sigma_\o(t_0+1)\Big(u(t_0)+\sum_{i=2}^{k-1}\sigma_i(t_0)\Big)\right\}\cdot$$
$$\cdot\prod_{t=0}^{T}\mathbb{I}\left\{\sigma_\o(t+1)\Big(u(t)+\sum_{i=2}^{k-1}\sigma_i(t)\Big) > r_0\right\}.$$

where, for $a > 0, a \in \mathbb{Z}$

$$\widehat{\mathsf{K}}_a(x) = \begin{cases} 1 & \text{if } -a < x < 0, \\ 1/2 & \text{if } x = -a \text{ or } x = 0, \\ 0 & \text{otherwise,} \end{cases} \qquad \widehat{\mathsf{K}}_{-a}(x) = \begin{cases} -1 & \text{if } 0 < x < -a, \\ -1/2 & \text{if } x = 0 \text{ or } x = -a, \\ 0 & \text{otherwise.} \end{cases}$$

Notice that $\sum_{x\in\mathbb{Z}}\widehat{\mathsf{K}}_a(x) = a\ \forall a \ge -r_0$.

We apply Theorem 4.4 for any value of $s(t_0) \equiv \sum_{i=2}^{k-1}\sigma_i(t_0)$ such that $\widehat{\mathsf{K}}_{\Omega(t_0)}\{\,\cdot\,\}$ is non-vanishing, and then sum over these values. Notice that $|\sum_{i=2}^{k-1}\sigma_i(t_0)| \le r_0 + 1$ and therefore the central limit theorem 4.4 applies. The leading order terms are all independent of $s(t_0)$. The $O(1/k^{1/4})$ error term in Eq. (30) is multiplied by a factor r_0 and remains therefore negligible. We get

$$J_{t_0} = \frac{1}{\sqrt{k}}\,\sigma_\o(t_0+1)\sigma_1(t_0)\Phi_{\mu(\sigma_\o),C_T}(A_\infty(\mathcal{I}))(1+o(1)) \tag{51}$$
$$\equiv \frac{1}{\sqrt{k}}\,\sigma_\o(t_0+1)\sigma_1(t_0)J^*_{t_0}\,(1+o(1))\,, \tag{52}$$

where $\mu(\sigma) = (\mu_0(\sigma), \dots, \mu_T(\sigma))$ with $\mu_r(\sigma) = \sum_{s=0}^{r-1} R(r,s)\sigma(s)$, and $\mathcal{I}_+ = \{t : \sigma_\o(t) = +1\} \setminus \{t_0\}$, $\mathcal{I}_- = \{t : \sigma_\o(t) = -1\} \setminus \{t_0\}$, and $\mathcal{I}_0 = \{t_0\}$. Notice that, in particular $J^*_{t_0=T} = I_T((\sigma_\o)_0^T)$.

If we use this estimate in Eq. (49), we get

$$\begin{aligned}
\mathsf{D}(1,k) &= (k-1)\sum_{t_0=0}^{T} \sigma_1(t_0)\left(\mathbb{Q}((\sigma_i)_0^T||(\sigma_\o)_\o^T) - \mathbb{P}((\sigma_i)_0^T||(\sigma_\o)_0^T)\right)\frac{J^*_{t_0}}{\sqrt{k}}\sigma_\o(t_0+1)(1+\ o(1)) + o(k^{-(T_*-T)/2}) \\
&= k\sum_{t_0=0}^{T} \frac{2\omega_{t_0}}{k^{(T_*-t_0+1)/2}}\frac{J^*_{t_0}}{\sqrt{k}}\sigma_\o(t_0+1)(1+\ o(1)) + o(k^{-(T_*-T)/2}) \\
&= I((\sigma_\o)_0^T)\frac{2\omega_T}{k^{T_*-T}}\sigma_\o(t_0+1)(1+\ o(1))
\end{aligned}$$

which, along with Eq. (29) implies the thesis Eq. (40).

Let us now prove the claim (47). Recall that induction hypothesis we have $\mathbb{Q}((\sigma_i)_0^T||(\sigma_\o)_0^T) - \mathbb{P}((\sigma_i)_0^T||(\sigma_\o)_0^T) = O(k^{-(T_*-T+1)/2})$. Since $|\mathsf{K}_{u(t)}(\sigma_\o(t+1)|\sigma_{\partial\o}(t))| \leq 1$, this implies

$$|\mathsf{D}(r,k)| \leq\ k^r \sum_{(\sigma_1)_0^T \dots (\sigma_r)_0^T} \prod_{i=1}^{r} \left|\mathbb{Q}((\sigma_i)_0^T||(\sigma_\o)_0^T) - \mathbb{P}((\sigma_i)_0^T||(\sigma_\o)_0^T)\right| = O(k^{-r(T_*-T-1)/2}).$$

Since $T_* - T - 1 \geq 1$, we have

$$\sum_{r=3}^{k-1} |\mathsf{D}(r,k)| = O(k^{-3(T_*-T-1)/2}) = o(k^{-(T_*-T)/2}).$$

Further, $|\mathsf{D}(2,k)| = O(k^{-(T_*-T-1)}) = o(k^{-(T_*-T)/2})$ unless $T = T_* - 2$.

In order to argue in the $r = 2$, $T = T_* - 2$ case, we will proceed analogously to $r = 1$. Consider the definition of $\mathsf{D}(2,k)$ in Eq. (46). If $(\sigma_3)_0^T, \dots, (\sigma_{k-1})_0^T$ are such that $|\sum_{i=3}^{k-1}\sigma_i(t) + u(t)| > 2$ for all $t \in \{0, \dots, T\}$ then the factors $\mathsf{K}_{u(t)}(\sigma_\o(t+1)|\sigma_{\partial\o}(t))$ become independent of $(\sigma_1)_0^T$, $(\sigma_2)_0^T$. We can therefore carry out the sum over these variables obtaining:

$$\sum_{(\sigma_1)_0^T,(\sigma_2)_0^T}\prod_{i=1}^{2}\left\{\mathbb{Q}((\sigma_i)_0^T||(\sigma_\o)_0^T) - \mathbb{P}((\sigma_i)_0^T||(\sigma_\o)_0^T)\right\} = \prod_{i=1}^{r}\sum_{(\sigma_i)_0^T}\left\{\mathbb{Q}((\sigma_i)_0^T||(\sigma_\o)_0^T) - \mathbb{P}((\sigma_i)_0^T||(\sigma_\o)_0^T)\right\} = 0$$

because both $\mathbb{Q}(\,\cdot\,||(\sigma_\o)_0^T)$ and $\mathbb{Q}(\,\cdot\,||(\sigma_\o)_0^T)$ are normalized. Therefore, we can restrict the sum to those $(\sigma_3)_0^T, \dots, (\sigma_{k-1})_0^T$ such that $|\sum_{i=3}^{k-1}\sigma_i(t_0) + u(t_0)| \leq 2$ for at least one $t_0 \in \{0, \dots, T\}$. However, analogously to the case $r = 1$, the probability that this happens for the i.i.d. non-degenerate random vectors $(\sigma_3)_0^T \dots (\sigma_{k-1})_0^T$ is at most $O(k^{-1/2})$, using Theorem 4.4. Together with the induction hypothesis, this yields $|\mathsf{D}(2,k)| = O(k^{-1/2} \cdot k^{-(T_*-T-1)}) = o(k^{-(T_*-T)/2})$, which proves the claim.

Finally, Eq. (41) follows from (40) using the definitions (38), (39) and the identity (27). $\square$

We next show that Lemma 4.6 extends to $T = T_* - 1$. Since this case requires a different (more careful) calculation, we state it separately, although the conclusion is the same as for $T < T_* - 1$. The proof is in Appendix D.

Lemma 4.7. *Let $I_T(\sigma_0^T)$ be defined as in Lemma 4.6, and define ω_{T_*} by*

$$\omega_{T_*} = R(T_*, T_* - 1)\omega_{T_*-1} \,. \tag{53}$$

Then, for all $(\sigma_\emptyset)_0^{T_*}, u_0^{T_*} \in \{\pm 1\}^{T_*+1}$, we have

$$\mathbb{Q}((\sigma_\emptyset)_0^{T_*}||u_0^{T_*}) - \mathbb{P}((\sigma_\emptyset)_0^{T_*}||u_0^{T_*}) = \frac{\omega_{T_*-1}}{k^{1/2}}\,\sigma_\emptyset(T_*)\; I\big((\sigma_\emptyset)_0^{T_*-1}\big)\,\Big(1+o(1)\Big) \tag{54}$$

Further, for all $u_0^{T_*} \in \{\pm 1\}^{T_*+1}$, we have

$$\sum_{(\sigma_\emptyset)_0^{T_*}} \sigma_\emptyset(T_*)\Big\{\mathbb{Q}((\sigma_\emptyset)_0^{T_*}||u_0^{T_*}) - \mathbb{P}((\sigma_\emptyset)_0^{T_*}||u_0^{T_*})\Big\} = \frac{2\omega_{T_*}}{k^{1/2}}\,(1+o(1)) \;. \tag{55}$$

We now show that, for the dynamics under external field, the process of the root spin $\{\sigma_\emptyset(t)\}_{t\geq 0}$ converges as in Theorem 2.7.

Lemma 4.8. *For T_* a non-negative integer, $\omega_0 > 0$, and $\{u(t)\}_{t\geq 0} \in \{\pm 1\}^{\mathbb{N}}$, consider the majority process under external field u, on the rooted tree $\mathcal{G}_\emptyset = (\mathcal{V}_\emptyset, \mathcal{E}_\emptyset)$, with i.i.d. initialization with bias $\theta = \omega_0/k^{(T_*+1)/2}$. Then for any $T \geq T_* + 2$, we have*

$$(\sigma_\emptyset(0), \sigma_\emptyset(1), \dots, \sigma_\emptyset(T)) \xrightarrow{\mathrm{d}} (\sigma_{\rm cav}(0), \sigma_{\rm cav}(1), \dots, \sigma_{\rm cav}(T_*), \sigma(T_*+1), +1, +1, \dots, +1) \;,$$

where the random variable $\sigma(T_*+1)$ dominates stochastically $\sigma_{\rm cav}(T_*+1)$, and $\mathbb{P}\{\sigma(T_*+1) > \sigma_{\rm cav}(T_*+1)\}$ is strictly positive. Finally, there exist $A(\omega_0) > 0$ such that, for any $T \geq T_* + 2$,

$$\mathbb{E}_\theta\{\sigma_\emptyset(T)\} \geq 1 - e^{-A(\omega_0)k} \,.$$

Proof. An immediate consequence of Eqs. (55) and (41) is that, for all T, $0 \leq T \leq T_*$

$$\mathbb{E}_{\mathbb{Q}(\cdot||u_0^T)}[\sigma_\emptyset(T)] - \mathbb{E}_{\mathbb{P}(\cdot||u_0^T)}[\sigma_\emptyset(T)] = \frac{2\omega_T}{k^{(T_*-T+1)/2}}\,(1+o(1)) \;. \tag{56}$$

Further Lemmas 4.5 and 4.6 imply that

$$|\mathbb{E}_{\mathbb{Q}}\{\sigma_\emptyset(t)\sigma_\emptyset(s)\} - C(t,s)| = o(1)\,, \qquad \left|\mathbb{E}_{\mathbb{Q}}\sigma_\emptyset(t) - \frac{1}{\sqrt{k}}\sum_{s=0}^{t-1} R(t,s)\,u(s)\right| = o(k^{-1/2})\,, \tag{57}$$

for $t, s \leq T_* - 1$. At T_*, using Lemma 4.7 and Eq. (56) with $T = T_*$ we obtain

$$|\mathbb{E}_{\mathbb{Q}}\{\sigma_\emptyset(T_*)\sigma_\emptyset(s)\} - C(t,s)| = o(1)\,, \quad \left|\mathbb{E}_{\mathbb{Q}}\left[\sigma_\emptyset(T_*+1) - \frac{1}{\sqrt{k}}\left(\sum_{s=0}^{T_*-1} R(t,s)\,u(s) + 2\omega_{T_*}\right)\right]\right| = o(k^{-1/2})\,, \tag{58}$$

which holds for all $s \leq T_*$.

Now, repeating the CLT-based argument as in the proof of lemma 4.5, we can show that with a biased initialization, $(\sigma_\emptyset(0), \sigma_\emptyset(1), \dots, \sigma_\emptyset(T_* + 1))$ converges to a modified cavity process, where the governing equation at T_* is

$$\sigma(T_* + 1) = \text{sign}\left(\eta(T_*) + \sum_{s=0}^{T_*-1} R(t,s)\sigma_{\text{cav}}(s) + 2\omega_{T_*}\right). \tag{59}$$

Convergence to this process occurs for all $u_0^{T_*+1}$. Clearly, since $\omega_{T_*} > 0$, this process dominates the unmodified cavity process. Further, we have $B(\omega_0) = \mathbb{E}[\sigma'(T_* + 1)] > 0$. We know $\lim_{k\to\infty} \mathbb{E}[\sigma_\emptyset(T_* + 1)] = \mathbb{E}[\sigma(T_* + 1)]$, and therefore there exists k_0, such that for all $k > k_0$, $\mathbb{E}[\sigma_\emptyset(T_* + 1)] > B(\omega_0)/2$. Plugging this back into the recursion Eq. (6) applied to $\mathbb{Q}$, and using Azuma's inequality, we see that at $T = T_* + 2$.

$$\mathbb{E}_\theta\{\sigma_\emptyset(T)\} \geq 1 - e^{-(B(\omega_0))^2 k/8}, \qquad \forall\ k > k_0$$

Clearly, the same continues to hold for for $T > T_* + 2$, for sufficiently large k. □

Finally, we can prove Theorem 2.7.

Proof. (Theorem 2.7) As in the proof of Theorem 2.6, we consider the dynamics on the rooted tree $\mathcal{G}_\emptyset$ under external fields $u_- = (-1, -1, \dots)$ and $u_+ = (+1, +1, \dots)$, and we denote by $(\sigma_{\emptyset,-})_0^T$, $(\sigma_{\emptyset,+})_0^T$ be the corresponding trajectories. By monotonicity of the dynamics, the process $(\sigma_i)_0^T$ at any vertex of the regular tree $\mathcal{G}$ is dominated by $(\sigma_{\emptyset,+})_0^T$ and dominates $(\sigma_{\emptyset,-})_0^T$. Since by Lemma 4.8 both $(\sigma_{\emptyset,+})_0^T$ and $(\sigma_{\emptyset,-})_0^T$ converge to the same limit, the same holds for $(\sigma_i)_0^T$ as well. □

5 Lower bound: Proof of Theorem 2.8

In this section we prove Theorem 2.8, that provides a sequence of lower bounds on the consensus threshold $\theta_*(k)$.

Our lower bounds are based on the formation of 'stable' structures of -1 spins, i.e. once such a structure is formed, it continues to exist at all future times, hence preventing consensus from being reached.

Consider $k = 3$. Clearly, if there is an infinite path of -1 spins, spins along the path remain unchanged for all future times. In fact, it is sufficient to have an infinite path having alternate vertices with -1 spins, due to the 'bipartite' nature of the dynamics. To see this, label an arbitrary node on the path 0. Choose an arbitrary direction on the path and hence label nodes $\dots, -2, -1, 0, 1, 2, \dots$. Suppose that at time t_0, nodes with even labels $\dots, -2, 0, 2, \dots$ all have spin -1. At time $t_0 + 1$, all nodes with with odd labels will have spin -1. At time $t_0 + 2$, all nodes with with even labels will again have spin -1 and so on.Note that any value for t_0 suffices. $t_0 = 0$ corresponds to an alternating core existing initially, but it is sufficient for such structure to be formed, say, at $t_0 = 3$.

This idea can be generalized to any k. A $\lceil \frac{k+1}{2} \rceil$-core of -1 spins is clearly stable. In fact, an alternating $\lceil \frac{k+1}{2} \rceil$-core of -1 spins (having alternate 'levels' of -1 spins) is stable. We formally define such structures below. The key point to note is that though such structures exist in abundance at $T = 0$ with small positive θ for $k = 3$, this is not the case for larger values of k. We do not obtain a positive lower bound for $k > 3$ based on analysis of the initial configuration only. Thus, we need a means to show that such structures form in abundance at $t_0(k) > 0$ for positive bias. We develop a set of iterative equations (see Theorem 2.8) whose fixed point corresponds (roughly) to the probability of formation of an alternating core at time t_0 at an arbitrary node. A non-trivial fixed point implies that alternating cores are formed in abundance. Such iterative equations are in the spirit of the exact cavity recursion (Lemma 2.1) though the analysis here is more intricate.

5.1 Notations and Preliminaries

Let $\mathcal{H} = (\mathcal{V}_{\mathcal{H}}, \mathcal{E}_{\mathcal{H}})$ be an induced subgraph of $\mathcal{G}$ with vertex set $\mathcal{V}_{\mathcal{H}}$ and edge set $\mathcal{E}_{\mathcal{H}}$. We denote by $\partial_{\mathcal{H}} i$ the set of neighbors in $\mathcal{H}$ of a node $i \in \mathcal{H}$. Since $\mathcal{H}$ is an induced subgraph of $\mathcal{G}$, we have $\mathcal{V}_{\mathcal{H}} \subseteq \mathcal{V}$ and, for all $i \in \mathcal{V}_{\mathcal{H}}$, $\partial_{\mathcal{H}} i = \{j : j \in \partial i, j \in \mathcal{V}_{\mathcal{H}}\}$. Given the graph $\mathcal{G}$, $\mathcal{V}_{\mathcal{H}}$ uniquely determines the induced subgraph $\mathcal{H}$.

Definition 5.1. *The subgraph $\mathcal{H}$ is an r-*core *of $\mathcal{G}$ with respect to spins $\sigma : \mathcal{V} \to \{-1, +1\}$ if $\mathcal{H}$ is an induced subgraph of $\mathcal{G}$ such that $|\partial_{\mathcal{H}} i| \geq r$ and $\sigma_i = -1$ for all $i \in \mathcal{V}_{\mathcal{H}}$.*

Clearly, this definition is useful only for $r \leq k$. Now, it is easy to see that if $\mathcal{H}$ is an $\lceil \frac{k+1}{2} \rceil$-core with respect to $\underline{\sigma}(T)$, then it is also an $\lceil \frac{k+1}{2} \rceil$-core with respect to $\underline{\sigma}(T')$ for all $T' > T$, by definition of majority dynamics. In fact, a less stringent requirement suffices for persistence of negative spins.

Definition 5.2. *$\mathcal{H}$ is an* alternating r-core *of a graph $\mathcal{G}$ with respect to spins $\sigma : \mathcal{V} \to \{-1, +1\}$, if $\mathcal{H}$ is an induced subgraph of $\mathcal{G}$ such that:*

1. *$|\partial_{\mathcal{H}} i| \geq r \ \forall \ i \in \mathcal{V}_H$*
2. *There is a partition $(\mathcal{V}_{-,\mathcal{H}}, \mathcal{V}_{*,\mathcal{H}})$ of $\mathcal{V}_{\mathcal{H}}$ such that:*
   (a) *$\sigma_i = -1$ for all $i \in \mathcal{V}_{-,\mathcal{H}}$*
   (b) *$\partial_{\mathcal{H}} i \subseteq \mathcal{V}_{-,\mathcal{H}}$ for all $i \in \mathcal{V}_{*,\mathcal{H}}$ and $\partial_{\mathcal{H}} i \subseteq \mathcal{V}_{*,\mathcal{H}}$ for all $i \in \mathcal{V}_{-,\mathcal{H}}$ i.e. $\mathcal{H}$ is bipartite with respect to the vertex partition $(\mathcal{V}_{-,\mathcal{H}}, \mathcal{V}_{*,\mathcal{H}})$. We call $\mathcal{V}_{-,\mathcal{H}}$ the even vertices and $\mathcal{V}_{*,\mathcal{H}}$ the odd vertices.*

Lemma 5.3. *If $\mathcal{H}$ is an alternating $\lceil \frac{k+1}{2} \rceil$-core with respect to $\underline{\sigma}(T)$, then it is also an alternating $\lceil \frac{k+1}{2} \rceil$-core with respect to $\underline{\sigma}(T')$ for all $T' > T$.*

Proof. We prove the lemma by induction over T'. Let

$$\mathcal{S}_{T'} \equiv \text{`}\mathcal{H} \text{ is an alternating } \left\lceil \frac{k+1}{2} \right\rceil \text{-core with respect to } \underline{\sigma}(T').\text{'}$$

Clearly, $\mathcal{S}_T$ holds. Suppose $\mathcal{S}_{T'}$ holds. Let $(\mathcal{V}_{-,\mathcal{H}}, \mathcal{V}_{*,H}) = (\mathcal{V}_1, \mathcal{V}_2)$ be a partition of $\mathcal{H}$ as in the definition 5.2. In particular $\sigma_i(T') = -1$ for all $i \in V_1$. By the definition of majority dynamics we know that $\sigma_i(T'+1) = -1$ for all $i \in \mathcal{V}_2$. As a consequence $\mathcal{H}$ is an alternating $\lfloor (k+1)/2 \rfloor$-core with respect to $\sigma(T'+1)$ with partition $(\mathcal{V}_{-,\mathcal{H}}, \mathcal{V}_{*,H}) = (\mathcal{V}_2, \mathcal{V}_1)$, and therefore $\mathcal{S}_{T'+1}$ holds. $\square$

We now proceed in a manner similar to Section 2.1. We consider the rooted tree $\mathcal{G}_{\o} = (\mathcal{V}_{\o}, \mathcal{E}_{\o})$, with a root vertex $\o$ having $k-1$ 'children'. The root spin $\sigma_{\o}$ evolves under as external field $\{u(t)\}_{t \geq 0}$ as in Eq. (5) and we denote by $\mathbb{P}((\sigma_{\o})_0^T || u_0^T)$ its distribution. We use $\tilde{\partial} i$ to denote the 'children' of node $i \in \mathcal{G}_{\o}$. In this section we will assume $u_0^T \in \{-1, +1\}^{T+1}$.

Definition 5.4. *$\mathcal{H}$ is a* rooted alternating r-core *of $\mathcal{G}_{\o}$ with respect to spins $\sigma : \mathcal{V}_{\o} \to \{-1, +1\}$, if $\mathcal{H}$ is a connected induced subgraph of $\mathcal{G}_{\o}$ such that:*

1. *$\o \in \mathcal{V}_{\mathcal{H}}$.*

2. $|\tilde{\partial}_{\mathcal{H}} i| \geq r-1$ *for all* $i \in \mathcal{V}_{\mathcal{H}}$.

3. *There is a partition* $(\mathcal{V}_{-,\mathcal{H}}, \mathcal{V}_{*,\mathcal{H}})$ *of* $\mathcal{V}_{\mathcal{H}}$ *such that:*

   (a) $\sigma_i = -1$ *for all* $i \in \mathcal{V}_{-,\mathcal{H}}$.

   (b) $\partial_{\mathcal{H}} i \subseteq \mathcal{V}_{-,\mathcal{H}}$ *for all* $i \in \mathcal{V}_{*,\mathcal{H}}$ *and* $\partial_{\mathcal{H}} i \subseteq \mathcal{V}_{*,\mathcal{H}}$ *for all* $i \in \mathcal{V}_{-,\mathcal{H}}$ *i.e.* $\mathcal{H}$ *is bipartite with respect to the vertex partition* $(\mathcal{V}_{-,\mathcal{H}}, \mathcal{V}_{*,\mathcal{H}})$. *We call* $\mathcal{V}_{-,\mathcal{H}}$ *the even vertices and* $\mathcal{V}_{*,\mathcal{H}}$ *the odd vertices.*

Let $\mathcal{G}^d_{ø} = (\mathcal{V}^d_{ø}, \mathcal{E}^d_{ø})$, be the induced subgraph of $\mathcal{G}_{ø}$ containing all vertices that are at a depth less than or equal to d from ø, the depth of ø itself being 0. For example, $\mathcal{G}^0_{ø}$ contains ø alone. Denote by $\tilde{\partial}\mathcal{G}^d_{ø}$, the set of leaves of $\mathcal{G}^d_{ø}$. For example, $\tilde{\partial}\mathcal{G}^0_{ø} = \{ø\}$.

Definition 5.5. $\mathcal{H}$ *is a* depth-d partial rooted alternating r-core *of* $\mathcal{G}_{ø}$ *with respect to spins* $\sigma : \mathcal{V}^d_{ø} \to \{-1,+1\}$, *if* $\mathcal{H}$ *is an connected induced subgraph of* $\mathcal{G}^d_{ø}$ *such that:*

1. $ø \in \mathcal{V}_{\mathcal{H}}$.

2. $|\tilde{\partial}_{\mathcal{H}} i| \geq r-1$ *for all* $i \in \mathcal{V}_{\mathcal{H}} \setminus \tilde{\partial}\mathcal{G}^d_{ø}$.

3. *There is a partition* $(\mathcal{V}_{-,\mathcal{H}}, \mathcal{V}_{*,\mathcal{H}})$ *of* $\mathcal{V}_{\mathcal{H}}$ *such that:*

   (a) $\sigma_i = -1$ *for all* $i \in \mathcal{V}_{-,\mathcal{H}}$.

   (b) $\partial_{\mathcal{H}} i \subseteq \mathcal{V}_{-,\mathcal{H}}$ *for all* $i \in \mathcal{V}_{*,\mathcal{H}}$ *and* $\partial_{\mathcal{H}} i \subseteq \mathcal{V}_{*,\mathcal{H}}$ *for all* $i \in \mathcal{V}_{-,\mathcal{H}}$ *i.e.* $\mathcal{H}$ *is bipartite with respect to the vertex partition* $(\mathcal{V}_{-,\mathcal{H}}, \mathcal{V}_{*,\mathcal{H}})$. *We call* $\mathcal{V}_{-,\mathcal{H}}$ *the even vertices and* $\mathcal{V}_{*,\mathcal{H}}$ *the odd vertices.*

We define $\mathcal{H}_{ø,\text{even}}(T)$ to be the maximal rooted alternating $\lceil \frac{k+1}{2} \rceil$-core of $\mathcal{G}_{ø}$ with respect to $\underline{\sigma}(T)$, such that ø is an even vertex. For all $d \geq 0$, we define $\mathcal{H}^d_{ø,\text{even}}(T)$ to be the maximal depth d partial rooted alternating $\lceil \frac{k+1}{2} \rceil$-core of $\mathcal{G}_{ø}$ with respect to $\underline{\sigma}^d(T)$, such that ø is even. Here $\underline{\sigma}^d(T)$ is the restriction of $\underline{\sigma}(T)$ to $\mathcal{V}^d_{ø}$. We similarly define $\mathcal{H}_{ø,\text{odd}}(T)$ and $\mathcal{H}^d_{ø,\text{odd}}(T)$.

We define $C_{\text{even}}(T) = \{ø \in \mathcal{V}_{\mathcal{H}_{ø,\text{even}}(T)}\}$, i.e. $C_{\text{even}}(T)$ is the event of $\mathcal{H}_{ø,\text{even}}(T)$ being non-empty. Define $C^d_{\text{even}}(T) = \{ø \in \mathcal{V}^d_{\mathcal{H}_{ø,\text{even}}(T)}\}$. We similarly define $C_{\text{odd}}(T)$ and $C^d_{\text{odd}}(T)$. It is easy to see that $C^d_{\text{even}}(T) \subseteq C^{d'}_{\text{even}}(T)$, $\forall\ d' < d$. Also, $C_{\text{even}}(T) = \bigcap_{d \geq 0} C^d_{\text{even}}(T)$. Similarly, $C^d_{\text{odd}}(T) \subseteq C^{d'}_{\text{odd}}(T)$, $\forall\ d' < d$ and $C_{\text{odd}}(T) = \bigcap_{d \geq 0} C^d_{\text{odd}}(T)$. We thus have the following remark.

Lemma 5.6. $C^d_{\text{even}}(T), d \geq 0$ *form a monotonic non-increasing sequence of events in d with limit* $\bigcap_{d \geq 0} C^d_{\text{even}}(T) = C_{\text{even}}(T)$, *for all* $T \geq 0$. *Similarly for the 'odd' quantities.*

Let $A(\omega^T_0) \equiv \{\sigma_{ø}(t) = \omega(t), 0 \leq t \leq T\}$ and define the events

$$
\begin{aligned}
B_{\text{even}}(T, \omega^T_0) &= C_{\text{even}}(T) \cap A(\omega^T_0)\,, \\
B^d_{\text{even}}(T, \omega^T_0) &= C^d_{\text{even}}(T) \cap A(\omega^T_0), \qquad d \geq 0\,.
\end{aligned}
$$

We similarly define $B_{\text{odd}}, B^d_{\text{odd}}$.

We now proceed to define $\Psi^d_{\text{even},T}((\sigma_{ø})^T_0 || u^T_0)$ and $\Psi^d_{\text{odd},T}((\sigma_{ø})^T_0 || u^T_0)$ as probabilities. *Immediately after the new definitions, we show that they are consistent with the recursive definitions in Theorem 2.8.*

Definition 5.7.

$$\Psi_{\text{even},T}((\sigma_\emptyset)_0^T||u_0^T) \equiv \mathbb{P}(B_{\text{even}}(T,(\sigma_\emptyset)_0^T)||u_0^T)\,,$$
$$\Psi^d_{\text{even},T}((\sigma_\emptyset)_0^T||u_0^T) \equiv \mathbb{P}(B^d_{\text{even}}(T,(\sigma_\emptyset)_0^T)||u_0^T)\,, \qquad d \geq 0\,.$$

We similarly define $\Psi_{\text{odd},T}((\sigma_\emptyset)_0^T||u_0^T))$, $\Psi^d_{\text{odd},T}((\sigma_\emptyset)_0^T||u_0^T))$.

It follows from Lemma 5.6 that $B_{\text{even}}(T) = \bigcap_{d\geq 0} B^d_{\text{even}}(T)$. Therefore $\Psi^d_{\text{even},T}((\sigma_\emptyset)_0^T||u_0^T)$ is non-increasing in d and by the monotone convergence theorem

$$\Psi_{\text{even},T}((\sigma_\emptyset)_0^T||u_0^T) = \lim_{d\to\infty} \Psi^d_{\text{even},T}((\sigma_\emptyset)_0^T||u_0^T)\,. \tag{60}$$

Similarly, we have $\Psi_{\text{odd},T}((\sigma_\emptyset)_0^T||u_0^T)) = \lim_{d\to\infty} \Psi^d_{\text{odd},T}((\sigma_\emptyset)_0^T||u_0^T))$. This is consistent with the definition of $\Psi_{\text{odd},T}((\sigma_\emptyset)_0^T||u_0^T))$ in Theorem 2.8.

The values for $d = 0$ follow from Definition 5.7,

$$\Psi^0_{\text{odd},T}((\sigma_\emptyset)_0^T||u_0^T)) = \mathbb{P}((\sigma_\emptyset)_0^T||u_0^T)\,,$$
$$\Psi^0_{\text{even},T}((\sigma_\emptyset)_0^T||u_0^T)) = \mathbb{P}((\sigma_\emptyset)_0^T||u_0^T)\,\mathbb{I}(\sigma_\emptyset(T) = -1)\,. \tag{61}$$

Note consistency with Eq. (15).

Next, in Lemma 5.8, we show that $\Psi^d_{\text{even},T}((\sigma_\emptyset)_0^T||u_0^T)$ and $\Psi^d_{\text{odd},T}((\sigma_\emptyset)_0^T||u_0^T)$ –as per Definition 5.7– satisfy Eqs. (16), (17) (repeated as Eqs. (62), (63) below).

5.2 Proof of Theorem 2.8

Lemma 5.8. *The following iterative equations are satisfied for all* $d \geq 0$*:*

$$\Psi^{d+1}_{\text{odd},T}((\sigma_\emptyset)_0^T||u_0^T) = \mathbb{P}_0(\sigma_\emptyset(0)) \sum_{r=\lceil\frac{k+1}{2}\rceil-1}^{k-1} \binom{k-1}{r} \sum_{(\sigma_1)_0^T\cdots(\sigma_{k-1})_0^T} \prod_{t=0}^{T-1} \mathsf{K}_{u(t)}(\sigma_\emptyset(t+1)|\sigma_{\partial\emptyset}(t))$$
$$\prod_{i=1}^{r} \Psi^d_{\text{even},T}((\sigma_i)_0^T||(\sigma_\emptyset)_0^T) \prod_{i=r+1}^{k-1} \Big(\mathbb{P}((\sigma_i)_0^T||(\sigma_\emptyset)_0^T) - \Psi^d_{\text{even},T}((\sigma_i)_0^T||(\sigma_\emptyset)_0^T)\Big)\,, \tag{62}$$

$$\Psi^{d+1}_{\text{even},T}((\sigma_\emptyset)_0^T||u_0^T) = \mathbb{I}(\sigma_\emptyset(T) = -1)\mathbb{P}_0(\sigma_\emptyset(0)) \sum_{r=\lceil\frac{k+1}{2}\rceil-1}^{k-1} \binom{k-1}{r} \sum_{(\sigma_1)_0^T\cdots(\sigma_{k-1})_0^T} \prod_{t=0}^{T-1} \mathsf{K}_{u(t)}(\sigma_\emptyset(t+1)|\sigma_{\partial\emptyset}(t))$$
$$\prod_{i=1}^{r} \Psi^d_{\text{odd},T}((\sigma_i)_0^T||(\sigma_\emptyset)_0^T) \prod_{i=r+1}^{k-1} \Big(\mathbb{P}((\sigma_i)_0^T||(\sigma_\emptyset)_0^T) - \Psi^d_{\text{odd},T}((\sigma_i)_0^T||(\sigma_\emptyset)_0^T)\Big)\,, \tag{63}$$

$$\mathsf{K}_{u(t)}(\cdots) \equiv \begin{cases} \mathbb{I}\Big\{\sigma_\emptyset(t+1) = \text{sign}\Big(\sum_{i=1}^{k-1}\sigma_i(t) + u(t)\Big)\Big\} & \textit{if } \sum_{i=1}^{k-1}\sigma_i(t) + u(t) \neq 0\,, \\ \frac{1}{2} & \textit{otherwise.} \end{cases}$$

Proof. Consider any $d \geq 0$. We denote the neighbors of the root as $\{1,\ldots,k-1\}$. We reuse the definitions of $\underline{\sigma}(0)$ and $\underline{\sigma}_i(0)$ $1 \leq i \leq (k-1)$ from Lemma 2.1, with depth T replaced with depth $(T+d+1)$. We denote by $\mathcal{A}_{T-1}$ the set of coin flips $\{A_{i,t}\}$ with $t \leq T-1$, and i at distance at most $T+d+1$

from the root. We have $\mathcal{A}_{T-1} = ((A_\o)_0^{T-1}, \mathcal{A}_{1,T-1}, \dots, \mathcal{A}_{k-1,T-1})$, where $\mathcal{A}_{i,T-1}$ is the subset of coin flips in the subtree rooted at $i \in \{1, \dots, k-1\}$. Let $\mathcal{G}_i$ be the subtree rooted at i. Define $\mathcal{H}^d_{i,\text{even}}(T)$, as the maximal depth d partial rooted alternating $\lceil \frac{k+1}{2} \rceil$-core of $\mathcal{G}_i$ with respect to $\underline{\sigma}_i^d(T)$, such that i is even. Define $C^d_{i,\text{even}}(T) = \{\o \in \mathcal{V}^d_{\mathcal{H}_{i,\text{even}}(T)}\}$. Let $A_i(\omega_0^T) \equiv \{\sigma_i(t) = \omega(t), 0 \le t \le T\}$. We define $B^d_{i,\text{even}}(T, (\omega)_0^T) = C^d_{i,\text{even}}(T) \cap A_i(\omega_0^T)$. Hence, we have mirrored the definitions for the root $\o$ at the child i.

Let $\mathcal{C} = \{1, 2, \dots, k-1\}$. By Definition 5.5, it follows that (here $A^\complement$ denotes the complement of an event A)

$$
\begin{aligned}
C^{d+1}_{\text{odd}}(T) &= \bigcup_{\substack{\mathcal{S} \subseteq \mathcal{C} \\ |\mathcal{S}| \ge \lceil (k-1)/2 \rceil}} \bigcap_{i \in \mathcal{S}} C^d_{i,\text{even}}(T) \bigcap_{j \in \mathcal{C}-\mathcal{S}} \left(C^d_{j,\text{even}}(T) \right)^\complement \\
C^{d+1}_{\text{even}}(T) &= \mathbb{I}(\sigma_\o(T) = -1) \bigcup_{\substack{\mathcal{S} \subseteq \mathcal{C} \\ |\mathcal{S}| \ge \lceil (k-1)/2 \rceil}} \bigcap_{i \in \mathcal{S}} C^d_{i,\text{odd}}(T) \bigcap_{j \in \mathcal{C}-\mathcal{S}} \left(C^d_{j,\text{odd}}(T) \right)^\complement
\end{aligned} \tag{64}
$$

Let $\mathcal{J}^d_{\text{odd}}(\underline{\sigma}^{T+d}, u_0^T, \mathcal{A}_T) \equiv \mathbb{I}(C^d_{\text{odd}}(T))$. Note that $\mathcal{J}^d_{\text{odd}}$ is a deterministic function. Similarly define $\mathcal{J}^d_{\text{even}}$. From Eq. (64), we have

$$
\mathcal{J}^{d+1}_{\text{odd}}(\underline{\sigma}^{T+d+1}, u_0^T, \mathcal{A}_T) = \sum_{\substack{\mathcal{S} \subseteq \mathcal{C} \\ |\mathcal{S}| \ge \lceil (k-1)/2 \rceil}} \prod_{i \in \mathcal{S}} \mathcal{J}^d_{i,\text{even}}(\underline{\sigma}_i^{T+d}, (\sigma_\o)_0^T, \mathcal{A}_{i,T}) \prod_{j \in \mathcal{C}-\mathcal{S}} \left(1 - \mathcal{J}^d_{j,\text{even}}(\underline{\sigma}_j^{T+d}, (\sigma_\o)_0^T, \mathcal{A}_{j,T})\right) \tag{65}
$$

Define $f(\cdot,\cdot,\cdot,\cdot)$ and $\mathcal{F}(\cdot,\cdot,\cdot)$ as in Lemma 2.1. We have $\mathbb{I}(B^{d+1}_{\text{odd}}(T, \omega_0^T)) = \mathbb{I}(A(\omega_0^T))\mathbb{I}(C^{d+1}_{\text{odd}}(T))$, leading to

$$
\Psi^{d+1}_{\text{odd}}(\omega_0^T || u_0^T) = \mathbb{E}_{\mathcal{A}_{T-1}} \textstyle\sum_{\underline{\sigma}(0)} \mathbb{P}(\underline{\sigma}(0)) \mathbb{I}\left(\omega_0^T = \mathcal{F}_0^T(\underline{\sigma}(0), u_0^T, \mathcal{A}_{T-1})\right) \mathcal{J}^{d+1}_{\text{odd}}(\underline{\sigma}(0), u_0^T, \mathcal{A}_{T+d}) \,. \tag{66}
$$

Subtracting Eq. (66) from Eq. (82) after replacing $T+1$ by T, we get

$$
\begin{aligned}
\mathbb{P}(\omega_0^T || u_0^T) - \Psi^{d+1}_{\text{odd}}(\omega_0^T || u_0^T) = \mathbb{E}_{\mathcal{A}_{T-1}} \sum_{\underline{\sigma}(0)} \mathbb{P}(\underline{\sigma}(0)) \\
\cdot \mathbb{I}\left(\omega_0^T = \mathcal{F}_0^T(\underline{\sigma}(0), u_0^T, \mathcal{A}_{T-1})\right) (1 - \mathcal{J}^{d+1}_{\text{odd}}(\underline{\sigma}(0), u_0^T, \mathcal{A}_{T+d})) \,.
\end{aligned} \tag{67}
$$

Equations (83) and (84) (with T replaced by $T-1$) continue to hold. Using Eq. (65), we have the following decomposition, similar to Eq. (85):

$$
\begin{aligned}
&\mathbb{I}\left(\omega_1^T = \mathcal{F}_1^T(\underline{\sigma}(0), u_0^T, \mathcal{A}_{T-1})\right) \mathcal{J}^{d+1}_{\text{odd}}(\underline{\sigma}(0), u_0^T, \mathcal{A}_{T+d}) \\
=&\mathbb{I}\Big(\sigma_\o(0) = \omega(0)\Big) \sum_{(\sigma_1)_0^T \dots (\sigma_{k-1})_0^T} \prod_{t=0}^{T-1} \mathbb{I}\left(\omega(t+1) = f(\sigma_\o(t), \underline{\sigma}_{\partial\o}(t), u(t), A_{\o,t})\right) \\
&\cdot \sum_{\substack{\mathcal{S} \subseteq \mathcal{C} \\ |\mathcal{S}| \ge \lceil (k-1)/2 \rceil}} \prod_{i \in \mathcal{S}} \mathbb{I}\left((\sigma_i)_0^T = \mathcal{F}_0^T(\underline{\sigma}_i(0), \omega_0^T, \mathcal{A}_{i,T-1})\right) \mathcal{J}^d_{i,\text{even}}(\underline{\sigma}_i^{T+d}, (\sigma_\o)_0^T, \mathcal{A}_{i,T}) \\
&\cdot \prod_{j \in \mathcal{C}-\mathcal{S}} \mathbb{I}\left((\sigma_j)_0^T = \mathcal{F}_0^T(\underline{\sigma}_j(0), \omega_0^T, \mathcal{A}_{j,T-1})\right) \left(1 - \mathcal{J}^d_{j,\text{even}}(\underline{\sigma}_j^{T+d}, (\sigma_\o)_0^T, \mathcal{A}_{j,T})\right) .
\end{aligned} \tag{68}
$$

Using Eqs. (83), (84) and (68) in Eq. (66) and separating terms that depend only on $\underline{\sigma}_i(0)$, we get

$$\Psi_{\text{odd}}^{d+1}(\omega_0^T||u_0^T) = \mathbb{P}(\omega(0)) \sum_{(\sigma_1)_0^T \cdots (\sigma_{k-1})_0^T} \prod_{t=0}^{T-1} \mathbb{I}\left\{\omega(t+1) = f(\sigma_\varnothing(t), \underline{\sigma}_{\partial\varnothing}(t), u(t), A_{\varnothing,t})\right\}$$

$$\cdot \sum_{\substack{\mathcal{S}\subseteq\mathcal{C} \\ |\mathcal{S}|\geq\lceil(k-1)/2\rceil}} \left\{ \prod_{i\in\mathcal{S}} \sum_{\underline{\sigma}_i(0)} \mathbb{P}(\underline{\sigma}_i(0)) \mathbb{I}\left((\sigma_i)_0^T = \mathcal{F}_0^T(\underline{\sigma}_i(0), \omega_0^T, \mathcal{A}_{i,T-1})\right) \mathcal{J}_{i,\text{even}}^d(\underline{\sigma}_i^{T+d}, (\sigma_\varnothing)_0^T, \mathcal{A}_{i,T}) \right.$$

$$\left. \cdot \prod_{j\in\mathcal{C}-\mathcal{S}} \sum_{\underline{\sigma}_j(0)} \mathbb{P}(\underline{\sigma}_j(0)) \mathbb{I}\left((\sigma_j)_0^T = \mathcal{F}_0^T(\underline{\sigma}_j(0), \omega_0^T, \mathcal{A}_{j,T-1})\right) \left(1 - \mathcal{J}_{j,\text{even}}^d(\underline{\sigma}_j^{T+d}, (\sigma_\varnothing)_0^T, \mathcal{A}_{j,T})\right) \right\}.$$

Using the 'even' versions of Eqs. (66) and (67), and noticing the symmetry in the expression between the $k-1$ children, we recover Eq. (62).

Eq. (63) follows similarly, with the additional $\mathbb{I}(\sigma_\varnothing(T) = -1)$ term appearing due to the modification in Eq. (64). □

Let the vector of values taken by $\Psi_{\text{odd},T}(\cdot\,||\,\cdot)$ be denoted by $\bar{\Psi}_{\text{odd},T}$. Similarly define $\bar{\Psi}_{\text{even},T}$. Define $\bar{\Psi}_T = \left(\bar{\Psi}_{\text{odd},T}, \bar{\Psi}_{\text{even},T}\right)$.

As before, $\mathbb{P}_0(-1) = \frac{1-\theta}{2}$ and $\mathbb{P}_0(+1) = \frac{1+\theta}{2}$. Define $\theta_{\text{lb}}(k,T) = \sup\{\theta : \bar{\Psi}_{\text{odd},T} \succ 0\}$, where $\bar{v} \succ 0$, denotes that every component of the vector $\bar{v}$ is strictly positive.

Finally, we relate quantities on the process on the rooted graph $\mathcal{G}_\varnothing$ to the process on the infinite k-ary tree $\mathcal{G}$. Pick an arbitrary node $v \in \mathcal{V}$. Let $\mathcal{G}^d = (\mathcal{V}^d, \mathcal{E}^d)$, be the induced subgraph of $\mathcal{G}$ containing all vertices that are at a distance less than or equal to d from v. For example, $\mathcal{G}^0$ contains v alone. Denote by $\tilde{\partial}\mathcal{G}^d$, the set of leaves of $\mathcal{G}^d$. For example, $\tilde{\partial}\mathcal{G}^0 = \{v\}$.

Definition 5.9. *$\mathcal{H}$ is a* depth-d partial alternating r-core *of $\mathcal{G}$ with respect to spins $\sigma : \mathcal{V}^d \to \{-1,+1\}$, if $\mathcal{H}$ is an connected induced subgraph of $\mathcal{G}^d$ such that:*

1. *$v \in \mathcal{V}_\mathcal{H}$.*
2. *$|\tilde{\partial}_\mathcal{H} i| \geq r-1$ for all $i \in \mathcal{V}_\mathcal{H} \setminus \tilde{\partial}\mathcal{G}^d$.*
3. *There is a partition $(\mathcal{V}_{-,\mathcal{H}}, \mathcal{V}_{*,\mathcal{H}})$ of $\mathcal{V}_\mathcal{H}$ such that:*
   - *(a) $\sigma_i = -1$ for all $i \in \mathcal{V}_{-,\mathcal{H}}$.*
   - *(b) $\partial_\mathcal{H} i \subseteq \mathcal{V}_{-,\mathcal{H}}$ for all $i \in \mathcal{V}_{*,\mathcal{H}}$ and $\partial_\mathcal{H} i \subseteq \mathcal{V}_{*,\mathcal{H}}$ for all $i \in \mathcal{V}_{-,\mathcal{H}}$ i.e. $\mathcal{H}$ is bipartite with respect to the vertex partition $(\mathcal{V}_{-,\mathcal{H}}, \mathcal{V}_{*,\mathcal{H}})$. We call $\mathcal{V}_{-,\mathcal{H}}$ the even vertices and $\mathcal{V}_{*,\mathcal{H}}$ the odd vertices.*

We define $\widehat{\mathcal{H}}_{\text{even}}(T)$, as the maximal alternating $\lceil\frac{k+1}{2}\rceil$-core of $\mathcal{G}$ with respect to $\underline{\sigma}(T)$, such that v is an even vertex. For all $d \geq 0$, we define $\widehat{\mathcal{H}}_{\text{even}}^d(T)$, as the maximal depth d partial alternating $\lceil\frac{k+1}{2}\rceil$-core of $\mathcal{G}$ with respect to $\underline{\sigma}^d(T)$, such that v is even. Here, $\underline{\sigma}^d(T)$ is the restriction of $\underline{\sigma}(T)$, to $\mathcal{V}^d$. We similarly define $\widehat{\mathcal{H}}_{\text{odd}}(T)$ and $\widehat{\mathcal{H}}_{\text{odd}}^d(T)$.

We now proceed to define $\widehat{C}_{\text{even}}(T)$, $\widehat{C}_{\text{even}}^d(T)$, $\widehat{C}_{\text{odd}}(T)$, $\widehat{C}_{\text{odd}}^d(T)$, $\widehat{A}(\omega_0^T)$, and $\widehat{B}_{\text{even}}(T,\omega_0^T)$, $\widehat{B}_{\text{even}}^d(T,\omega_0^T)$, $\widehat{B}_{\text{odd}}(T,\omega_0^T)$, $\widehat{B}_{\text{odd}}^d(T,\omega_0^T)$ for $\mathcal{G}$, analogously to the definitions of $C_{\text{even}}(T)$ etc. for $\mathcal{G}_\varnothing$. An analog of Lemma 5.6 holds.

Define the probabilities

$$\widehat{\Psi}_{\text{even},T}(\sigma_0^T) = \mathbb{P}(B_{\text{even}}(T,\sigma_0^T))\,,$$
$$\widehat{\Psi}^d_{\text{even},T}(\sigma_0^T) = \mathbb{P}(B^d_{\text{even}}(T,\sigma_0^T))\,, \qquad d\geq 0\,.$$

As before, we have $\widehat{\Psi}^d_{\text{even},T}(\sigma_0^T)$ is non-increasing in d and

$$\widehat{\Psi}_{\text{even},T}(\sigma_0^T) = \lim_{d\to\infty} \widehat{\Psi}^d_{\text{even},T}(\sigma_0^T)\,. \tag{69}$$

We similarly define $\widehat{\Psi}_{\text{odd},T}(\sigma_0^T)$, $\widehat{\Psi}^d_{\text{odd},T}(\sigma_0^T)$ and have $\widehat{\Psi}^d_{\text{odd},T}(\sigma_0^T)$ converging to $\widehat{\Psi}_{\text{odd},T}(\sigma_0^T)$ as $d\to\infty$.

Lemma 5.10. *The following identities are satisfied for all $d\geq 0$:*

$$\widehat{\Psi}^{d+1}_{\text{odd},T}(\sigma_0^T) = \mathbb{P}_0(\sigma(0)) \sum_{r=\lceil\frac{k+1}{2}\rceil}^{k} \binom{k}{r} \sum_{(\sigma_1)_0^T\dots(\sigma_k)_0^T} \prod_{t=0}^{T-1} \widehat{\mathsf{K}}(\sigma(t+1)|\sigma_{\partial v}(t))$$
$$\prod_{i=1}^{r} \Psi^d_{\text{even},T}((\sigma_i)_0^T||\sigma_0^T) \prod_{i=r+1}^{k} \Big(\mathbb{P}((\sigma_i)_0^T||\sigma_0^T) - \Psi^d_{\text{even},T}((\sigma_i)_0^T||\sigma_0^T)\Big)\,, \tag{70}$$

$$\widehat{\Psi}^{d+1}_{\text{even},T}(\sigma_0^T) = \mathbb{I}(\sigma(T)=-1)\mathbb{P}_0(\sigma(0)) \sum_{r=\lceil\frac{k+1}{2}\rceil}^{k} \binom{k}{r} \sum_{(\sigma_1)_0^T\dots(\sigma_k)_0^T} \prod_{t=0}^{T-1} \widehat{\mathsf{K}}(\sigma(t+1)|\sigma_{\partial v}(t))$$
$$\prod_{i=1}^{r} \Psi^d_{\text{odd},T}((\sigma_i)_0^T||\sigma_0^T) \prod_{i=r+1}^{k} \Big(\mathbb{P}((\sigma_i)_0^T||\sigma_0^T) - \Psi^d_{\text{odd},T}((\sigma_i)_0^T||\sigma_0^T)\Big)\,, \tag{71}$$

$$\widehat{\mathsf{K}}(\cdots) \equiv \begin{cases} \mathbb{I}\Big\{\sigma(t+1) = \text{sign}\Big(\sum_{i=1}^k \sigma_i(t)\Big)\Big\} & \textit{if } \sum_{i=1}^k \sigma_i(t)\neq 0\,,\\ \frac{1}{2} & \textit{otherwise.}\end{cases} \tag{72}$$

Proof. The proof is very similar to the one of Lemma 5.8, and we omit it for the sake of space. ☐

Lemma 5.11. *Assume that $\bar{\Psi}_{\text{odd},T}\succ 0$ for some $T\geq 0$ and $\theta\in[0,1]$. Then for the same θ and T, there exists an alternating $\lceil\frac{k+1}{2}\rceil$-core of $\mathcal{G}$ with positive probability with respect to $\underline{\sigma}(T)$.*

Proof. Take the limit $d\to\infty$ in Eq. (71). We have,

$$\widehat{\Psi}_{\text{even},T}(\sigma_0^T) = \mathbb{I}(\sigma(T)=-1)\mathbb{P}_0(\sigma(0)) \sum_{r=\lceil\frac{k+1}{2}\rceil}^{k} \binom{k}{r} \sum_{(\sigma_1)_0^T\dots(\sigma_k)_0^T} \prod_{t=0}^{T-1} \widehat{\mathsf{K}}(\sigma(t+1)|\sigma_{\partial v}(t))$$
$$\prod_{i=1}^{r} \Psi_{\text{odd},T}((\sigma_i)_0^T||\sigma_0^T) \prod_{i=r+1}^{k} \big(\mathbb{P}((\sigma_i)_0^T||\sigma_0^T) - \Psi_{\text{odd},T}((\sigma_i)_0^T||\sigma_0^T)\big)\,, \tag{73}$$

Now, consider any θ such that $\bar{\Psi}_{\text{odd},T}\succ 0$. Consider $\widehat{\Psi}_{\text{even},T}(\sigma_0^T)$ for any σ_0^T with $\sigma(T)=-1$. Note that every term in the summation over r in Eq. (73) is non-negative, and, in fact, positive when $\bar{\Psi}_{\text{odd},T}\succ 0$ holds. Hence, $\widehat{\Psi}_{\text{even},T}(\sigma_0^T)>0 \Rightarrow \mathbb{P}_\theta(\exists$ alternating $\lceil\frac{k+1}{2}\rceil$-core $\mathcal{H}$ of $\mathcal{G}$ with respect to $\underline{\sigma}(T)$ s.t. $v\in\mathcal{H})>0$. ☐

The lower bound on $\theta_*(k)$ is an immediate consequence of the above lemmas.

Proof. (Theorem 2.8). The thesis follows Lemmas 5.3 and 5.11 and the definition of θ_* in Eq. (3). ☐

Table 4: Computed lower bound values

k	$T=0$	$T=1$	$T=2$	$T=3$	*Simulation*
3	$\mathbf{+0.508}$	$\mathbf{+0.568}$	$\mathbf{+0.572}$	$\mathbf{+0.574}$	*0.58*
5	-0.084	$\mathbf{+0.026}$	$\mathbf{+0.048}$	$\mathbf{+0.052}$	*0.054*
7	-0.14	-0.020	$\mathbf{+0.002}$	$\mathbf{+0.008}$	*0.010*
9	-0.14	-0.030	-0.006	-0.0008	
11	-0.12	-0.028	-0.010	-0.0028	
15	-0.12	-0.024	-0.008	-0.0028	
21	-0.084	-0.018	-0.0054	-0.0018	
31	-0.080	-0.014	-0.0032	-0.0010	
51	-0.046	-0.0070	-0.0014	-0.00038	
101	-0.026	-0.0032	-0.00048		
201	-0.016	-0.0014	-0.00014		
401	-0.0084	-0.00048	-0.000040		
1001	-0.0035	-0.00012	-0.000008		
Asymptotics	$-\Theta\left(\frac{\sqrt{\log k}}{k}\right)$	$-\Theta\left(\frac{\sqrt{\log k}}{k^{3/2}}\right)$	$-\Theta\left(\frac{\sqrt{\log k}}{k^{2}}\right)$		

5.3 Evaluating the lower bound

Equations (16) and (17) can be iterated with initial values given by Eq. (15) to compute $\theta_{\rm lb}(k,T)$. To simplify the recursion, we notice that the dynamics is 'bipartite': each of $\mathcal{A}$ and $\underline{\sigma}(0)$ can be partitioned $\mathcal{A} = (\hat{\mathcal{A}}, \tilde{\mathcal{A}}), \underline{\sigma}(0) = (\hat{\underline{\sigma}}(0), \tilde{\underline{\sigma}}(0))$ such that $(\hat{\mathcal{A}}, \hat{\underline{\sigma}}(0)$ and $(\tilde{\mathcal{A}}, \tilde{\underline{\sigma}}(0))$ never 'interact' in the majority dynamics on an infinite tree. This remark reduces the number of variables in the recursions Eqs. (16) and (17). Further, for small values of T, instead of summing over all possible trajectories of children, it is faster to sum over all possibilities for the histogram of the trajectories followed by children.

In Table 5.3, we present some of the lower bounds $\theta_{\rm lb}(k,T)$ computed through this approach, and compare them with the empirical threshold $\theta_*(k)$ deduced from numerical simulations.

As observed in the introduction $\theta_*(k) \geq 0$ by symmetry and monotonicity. Therefore the lower bounds are non-trivial only if $\theta_{\rm lb}(k,T) > 0$. It turns out that for any fixed T, $\theta_{\rm lb}(k,T)$ becomes negative at large k. We present in the same table the asymptotic behaviors. Nevertheless, for $k \leq 7$, our lower bounds provide good estimates of the actual threshold.

The values of $\theta_{\rm lb}(k,T)$ are much lower for even values of k. For example, for $k = 4$, 6, 8, $\theta_{\rm lb}(k,3) \approx -0.22$, -0.09, -0.05 respectively. This is as expected, since our requirement of an alternating $\lceil \frac{k+1}{2} \rceil$-core is more stringent for even k. On the other hand, numerical simulations suggest that $\theta_*(k) = 0$ for small even values of k.

Acknowledgments

This work was partially supported by a Terman fellowship, the NSF CAREER award CCF-0743978 and the NSF grant DMS-0806211. Yashodhan Kanoria is supported by a 3Com Corporation Stanford Graduate Fellowship.

References

[AS03] D. Aldous and J. M. Steele, “The Objective Method: Probabilistic Combinatorial Optimization and Local Weak Convergence,” in *Probability on discrete structures*, H. Kesten ed., pp. 1-72, Springer Verlag (2003)

[BPP06] J. Balogh, Y. Peres and G. Pete, “Bootstrap Percolation on Infinite Trees and Non-Amenable Groups,” Combinatorics, Probability and Computing, 15 (2006) 715-730

[BG07] A. Bandyopadhyay, D. Gamarnik, “Counting without sampling. New algorithms for enumeration problems using statistical physics”, Proceedings of the 17th ACM-SIAM Symposium on Discrete Algorithm, Miami, Flo., 2007.

[BR76] R. N. Bhattacharya, and R. Ranga Rao, *Normal approximation and asymptotic expansions*, Wiley, New York, (1976).

[BK+05] N. Berger, C. Kenyon, E. Mossel and Y. Peres, “Glauber dynamics on trees and hyperbolic graphs,” *Prob. Theory Rel. Fields* 131, 311 (2005)

[Bol80] B. Bollobás, “A probabilistic proof of an asymptotic formula for the Number of labeled regular graphs,” Eur. J. Combinatorics, **1** (1980), 311–316.

[H00] C. Douglas Howard, “Zero-Temperature Ising Spin Dynamics on the Homogeneous Tree of Degree Three,” J. App. Probab. 37, 736-747 (2000).

[BC+97] J. P. Bouchaud, L. F. Cugliandolo, J. Kurchan, and M. Mezard, “Out of equilibrium dynamics in spin-glasses and other glassy systems,” in A. P. Young, ed., *Spin glass dynamics and Random Fields*, World Scientific (1997)

[CM05] P. Caputo and F. Martinelli, “Phase ordering after a deep quench: the stochastic Ising and hard core gas models on a tree,” Probab. Theor. Rel. Fields, 136, 37-80 (2005)

[DMD94] B. Davis and D. McDonald, “An Elementary Proof of the Local Central Limit Theorem”, J. Theor. Probab. 8 (1995), 693-701

[DM10] A. Dembo and A. Montanari, “Gibbs Measures and Phase Transitions on Sparse Random Graphs,” Brazilian J. of Probab. and Stat. (2010)

[BDG06] G. Ben Arous, A. Dembo and A. Guionnet, “Cugliandolo-Kurchan equations for dynamics of Spin-Glasses,” Probab. Theor. Rel. Fields, 136, 619-660 (2006)

[FS08] L.R.G. Fontes and R.H. Schonmann, “Bootstrap Percolation on Homogeneous Trees Has 2 Phase Transitions,” J. Stat. Phys. 132 (2008) 839-861

[FSS02] L. R. Fontes, R. H. Schonmann and V. Sidoravicius, “Stretched Exponential Fixation in Stochastic Ising Models at Zero Temperature,” Commun. Math. Phys. 228, 495-518 (2002)

[GO80] E. Goles and J. Olivos, “Periodic behavior of generalized threshold functions,” Disc. Math. 30, 187-189 (1980)

[HW+04] J.P.L. Hatchett, B. Wemmenhove, I. Perez Castillo, T. Nikoletopoulos, N.S. Skantzos, A.C.C. Coolen, "Parallel dynamics of disordered Ising spin systems on finitely connected random graphs," J. Phys. A: Math. Gen. 37 (2004) 6201

[HLW06] S. Hoory, N. Linial and A. Widgerson, "Expander Graphs and Their Applications," Bull. Amer. Math. Soc. 43, 439-561 (2006)

[Kle07] J. Kleinberg, "Cascading Behavior in Networks: Algorithmic and Economic Issues," in *Algorithmic Game Theory*, N. Nisan et al. eds., Cambridge University Press (2007)

[KM+07] F. Krzakala, A. Montanari, F. Ricci-Tersenghi, G. Semerjian, and L. Zdeborova, "Gibbs States and the Set of Solutions of Random Constraint Satisfaction Problems," Proc. Natl. Acad. Sci., 104, 10318-10323 (2007)

[KR+08] F. Krzakala, A. Rosso, G. Semerjian, and F. Zamponi, "On the path integral representation for quantum spin models and its application to the quantum cavity method and to Monte Carlo simulations," Phys. Rev. B 78, 134428 (2008)

[LSS08] C. Laumann, A. Scardicchio, and S. Sondhi, "Cavity method for quantum spin glasses on the Bethe lattice," Phys. Rev. B 78, 134424 (2008)

[Lig85] T. J. Liggett, *Interacting Particle System*, Springer, Berlin, (1985)

[MSW03] F. Martinelli, A. Sinclair, and D. Weitz, "The Ising model on trees: boundary conditions and mixing time," Proc. of IEEE Symposium on Found. of Computer Science 2003

[MD79] D. McDonald, "On local limit theorems for integer valued random variables", Theor. Prob. and its Appl. 24 (1979), 613-619

[MM09] M. Mézard and A. Montanari, *Information, Physics, and Computation*, Oxford University Press (2009)

[MPV87] M. Mézard, G. Parisi, and M. A. Virasoro, *Spin Glass Theory and Beyond*, World Scientific, Singapore, (1987)

[Mor00] S. Morris, "Contagion", Review of Econ. Studies 67, 57-78 (2000)

[MPZ03] M. Mézard, G. Parisi and R. Zecchina, "Analytic and Algorithmic Solution of Random Satisfiability Problems," Science, 297, 812-815 (2003)

[NB09] I. Neri, D. Bollé, "The Cavity Approach to Parallel Dynamics of Ising Spins on a Graph," J. Stat. Mech. (2009) P08009

[Ta06] M. Talagrand, *Spin Glasses: A Challenge for Mathematicians*, Springer-Verlag, New York, 2003

[We06] D. Weitz, "Counting independent sets up to the tree threshold," Proceedings of the 38th annual ACM Symposium on Theory of Computing, 140-149 (2006)

A Proof of preliminary results

This Section presents the proofs of Lemmas 1.1 and 1.2, with some auxiliary results proved in the second subsection.

A.1 Proofs

Proof. (Lemma 1.1) Consider the subgraph $\mathcal{G}_+$ of $\mathcal{G}$ induced by vertices $i \in \mathcal{V}$, such that $\sigma_i(0) = +1$: each vertex belongs to this subgraph independently with probability $(1+\theta)/2$. Let $\mathcal{G}_{+,q}$ be the maximal subgraph of $\mathcal{G}_+$ with minimum degree $q = k - \lfloor (k+1)/2 \rfloor + 1$. It is clear that no vertex in $\mathcal{G}_{+,q}$ ever flips to -1 under the majority process. Consider a modified initial condition such that $\sigma_i(0) = +1$ for $i \in \mathcal{G}_{+,q}$, and $\sigma_i(0) = -1$ otherwise. By monotonicity of the dynamics, it is sufficient to show that such a modified initial condition converges to $\underline{+1}$ under the majority process.

Notice that $\mathcal{H} = \mathcal{G} \setminus \mathcal{G}_{+,q}$ is the subgraph induced by the final set of a bootstrap percolation process with initial density $\rho = (1-\theta)/2$ and threshold $\lfloor (k+1)/2 \rfloor$ (a vertex joins if at least $\lfloor (k+1)/2 \rfloor$ of its neighbors have joined). It is proven in [FS08, Theorem 1.1] that there exists $\rho_{\rm c}(k) > 0$ such that, for $\rho < \rho_{\rm c}(k)$, $\mathcal{H}$ is almost surely the disjoint union of a of countable number of finite trees. This implies the thesis. Indeed we can restrict our attention to any such finite tree occupied by -1, and surrounded by $+1$ elsewhere. On such a tree, the set of vertices such that of $\sigma_i(t) = -1$ never increases, and at least one vertex quits the set at each iteration. Therefore, any such tree turns to $+1$ in finitely many iterations. □

Proof. (Lemma 1.2) Let $\mathcal{G}_n = ([n], \mathcal{E}_n)$ be a random graph of degree k over n vertices distributed according to the configuration model. We recall that a graph is generated with this distribution by attaching k labeled half-edges to each vertex $i \in [n]$ and pairing them according to a uniformly random matching among nk objects.

The proof of Lemma 1.2 is based on the analysis of the majority process on the graph $\mathcal{G}_n$. We will denote by $\mathbb{P}_{\theta,n}$ the law of this process when the spins $\{\sigma_i(0)\}_{i\in[n]}$ are initialized to i.i.d. random variables with $\mathbb{E}_{\theta,n}\{\sigma_i(0)\} = \theta$. We use the following auxiliary results.

Lemma A.1. *For any fixed $i \in \mathbb{N}$, $j \in \mathcal{V}$ and $t \geq 0$ we have*

$$\lim_{n\to\infty} \mathbb{E}_{\theta,n}\{\sigma_i(t)\} = \mathbb{E}_{\theta}\{\sigma_j(t)\}\,. \tag{74}$$

Lemma A.2. *Let $\{\sigma_i(t)\}_{i\in[n],t\in\mathbb{Z}_+}$ be distributed according to the majority process on $\mathcal{G}_n$, and define $B(k,t) \equiv 4(t+1)(k^{t+1}-1)^2/(k-1)^2$. Then*

$$\mathbb{P}_{\theta,n}\left\{ \left| \sum_{i=1}^{n} \sigma_i(t) - n\mathbb{E}_{\theta,n}\sigma_1(t) \right| \geq n\varepsilon \;\middle|\; \mathcal{G}_n \right\} \leq 2\,\exp\left\{ -\frac{n\varepsilon^2}{2B(k,t)} \right\}\,. \tag{75}$$

Lemma A.3. *There exists δ_*, $k_* > 0$ such that for any $k \geq k_*$ there is a set $S_{k,n}$ of 'good graphs' such that $\mathbb{P}\{\mathcal{G}_n \in S_{k,n}\} \to 1$, and the following happens. For any $\mathcal{G}_n \in S_{k,n}$ and any initial condition $\{\sigma_i(0)\}_{i\in[n]}$ on the vertices of $\mathcal{G}_n$ with $\sum_{i=1}^n \sigma_i(0) \geq n(1-2\delta_*/k)$, we have*

$$\sum_{i=1}^{n}(1-\sigma_i(1)) \leq \frac{3}{4}\sum_{i=1}^{n}(1-\sigma_i(0))\,. \tag{76}$$

Let us now turn to the actual proof. Choose δ_* and k_* as per Lemma A.3 and assume $k \geq k_*$. By assumption there exists a time t_* such that $\mathbb{E}_\theta\{\sigma_i(t_*)\} \geq 1-\delta_*/k$. By Lemmas A.1 and A.2, for all n large enough we have

$$\mathbb{P}_{\theta,n}\left\{\sum_{i=1}^{n}\sigma_i(t_*) \geq n\left(1-2\frac{\delta}{k}\right)\right\} \geq 1-e^{-Cn}\,. \tag{77}$$

Assume $\sum_{i=1}^{n}\sigma_i(t_*) \geq n\left(1-2\frac{\delta_*}{k}\right)$ and $\mathcal{G}_n \in S_{k,n}$. Then, by Lemma A.3, and any $t \geq t_*$ we have

$$\sum_{i=1}^{n}(1-\sigma_i(t)) \leq n\,(3/4)^{t-t_*}\,. \tag{78}$$

Combining this with the above remarks, and using the symmetry of the graph distribution with respect to permutation of the vertices, we get

$$\mathbb{P}_{\theta,n}\{\sigma_1(t) \neq +1\} \leq 2(3/4)^{t-t_*} + \mathbb{P}\{\mathcal{G}_n \not\in S_{k,n}\} + e^{-Cn}\,. \tag{79}$$

By Lemma A.1, this implies $\mathbb{P}_\theta\{\sigma_i(t) \neq +1\} \leq 5(3/4)^{t-t_*}$ which, by Borel-Cantelli implies $\sigma_i(t) \to +1$ almost surely, whence the thesis follows. □

Proof. (Lemma A.1) Fix a vertex i in $\mathcal{G}_n$, and denote by $\mathsf{B}_i(t)$ the subgraph induced by vertices whose distance from i is at most t. The value of $\sigma_i(t)$ only depends on $\mathcal{G}_n$ through the $\mathsf{B}_i(t)$. If $\mathsf{B}_i(t)$ is a k-regular tree of depth t (to be denoted by $\mathsf{T}(t)$) then the distribution of $\sigma_j(t)$ is the same that would be obtained on $\mathcal{G}$, whence

$$|\mathbb{E}_{\theta,n}\{\sigma_i(t)\} - \mathbb{E}_\theta\{\sigma_j(t)\}| \leq 2\,\mathbb{P}_{\theta,n}\{\mathsf{B}_i(t) \not\simeq \mathsf{T}(t)\}\,.$$

The thesis follows since $\mathbb{P}_{\theta,n}\{\mathsf{B}_i(t) \not\simeq \mathsf{T}(t)\} \leq A^t/n$ for some constant A (dependent only on k. □

Proof. (Lemma A.2). Let $X_n(t) \equiv \sum_{i=1}^{n}\sigma_i(t)$. This is a deterministic function of the $n(t+1)$ bounded random variables $\{\sigma_i(0)\}_{i\in[n]}$ and of $\{A_{i,s}\}_{i\in[n],s\leq t}$. Further, it is a Lipschitz function with constant $L(k,t) \leq 2(k^{t+1}-1)/(k-1)$, because any change in $\sigma_i(0)$, or $A_{i,s}$ only influences the values $\sigma_j(t)$ within a ball of radius t around i. By Azuma-Hoeffding inequality

$$\mathbb{P}_{\theta,n}\left\{|X_n(t) - \mathbb{E}_{\theta,n}X_n(t)| \geq \Delta\right\} \leq 2\,\exp\left\{-\frac{\Delta^2}{2n(t+1)L(k,t)^2}\right\} \tag{80}$$

which implies the thesis. □

Proof. (Lemma A.3) Although the proof follows from a standard expansion argument, we reproduce it here for the convenience of the reader.

Recall that a graph $\mathcal{G}_n$ over n vertices is a $(k(1-\varepsilon),\delta/k)$ (vertex) expander if each subset $\mathcal{W}$ of at most $n\delta/k$ vertices is connected to at least $k(1-\varepsilon)|\mathcal{W}|$ vertices in the rest of the graph. It is known that there exists $\delta_* > 0$ such that, for all k large enough, a random k regular graph is, with high probability, a $(3k/4,\delta_*/k)$ expander [HLW06]. We let $S_{k,n}$ be the set of k-regular graphs $\mathcal{G}_n$ that are $(3k/4,\delta_*/k)$ expanders.

Let $\mathcal{W}$ be the set of vertices $i \in [n]$ such that $\sigma_i(0) = -1$. By hypothesis $|\mathcal{W}| \leq n\delta/k$. Denote by n_- the number of vertices in $[n]\setminus\mathcal{W}$ that have at least $\lceil k/2\rceil$ neighbors in $\mathcal{W}$ (and hence such that potentially

$\sigma_i(1) = -1$), and by n_+ the set of vertices that have between 1 and $\lceil k/2 \rceil - 1$ neighbors in $\mathcal{W}$. Further, let l be the number of edges between vertices in $\mathcal{W}$. Then

$$\left\lceil \frac{k}{2} \right\rceil n_- + n_+ + 2l \leq k\,|\mathcal{W}|\,, \qquad\qquad n_- + n_+ \geq \frac{3}{4} k\,|\mathcal{W}|\,,$$

where the first inequality follows by edge-counting and the second by the expansion property. By taking the difference of these inequalities, we get

$$\left(\left\lceil \frac{k}{2} \right\rceil - 1\right) n_- + 2l \quad \leq \quad \frac{k}{4}\,|\mathcal{W}|\,.$$

Let $\mathcal{W}'$ be the set of vertices such that $\sigma_i(1) = -1$. It is easy to see that $|\mathcal{W}'| \leq n_- + (2l)/\lceil k/2 \rceil$, and therefore

$$|\mathcal{W}'| \leq \frac{k}{4(\lceil k/2 \rceil - 1)}\,|\mathcal{W}|\,.$$

which yields the thesis. □

B Proof of the exact cavity recursion

Proof of Lemma 2.1. Throughout the proof we denote the neighbors of the root as $\{1, \ldots, k-1\}$. Let $\underline{\sigma}(0)$ be the vector of initial spins of the root and all the vertices up to a distance T from the root. For each $i \in \{1, \ldots, k-1\}$, let $\underline{\sigma}_i(0)$ be the vector of initial spins of the sub-tree rooted at i, and not including the root, and up to the same distance T from the root. Clearly, if we choose an appropriate ordering, we have $\underline{\sigma}(0) = (\sigma_\o(0), \underline{\sigma}_1(0), \underline{\sigma}_2(0), \ldots, \underline{\sigma}_{k-1}(0))$. Finally, we denote by $\mathcal{A}_T$ the set of coin flips $\{A_{i,t}\}$ with $t \leq T$, and i at distance at most T from the root. As above, we have $\mathcal{A}_T = ((A_\o)_0^T, \mathcal{A}_{1,T}, \ldots, \mathcal{A}_{k-1,T})$, where $\mathcal{A}_{i,T}$ is the subset of coin flips in the subtree rooted at $i \in \{1, \ldots, k-1\}$. g By definition, the trajectory $(\sigma_\o)_0^{T+1}$ is a deterministic function of $\underline{\sigma}(0)$, u_0^{T+1} and $\mathcal{A}_T$. We shall denote this function by $\mathcal{F}$ and write $(\sigma_\o)_s^t = \mathcal{F}_s^t(\underline{\sigma}(0), u_0^{T+1}, \mathcal{A}_T)$. This function is uniquely determined by the update rules. We shall write the latter as

$$\sigma_\o(t+1) = f(\sigma_\o(t), \underline{\sigma}_{\partial\o}(t), u(t), A_{\o,t})\,. \tag{81}$$

We have therefore

$$\mathbb{P}((\sigma_0)_0^{T+1} = \omega_0^{T+1} || u_0^{T+1}) \quad = \mathbb{E}_{\mathcal{A}_T} \textstyle\sum_{\underline{\sigma}(0)} \mathbb{P}(\underline{\sigma}(0)) \mathbb{I}\left(\omega_0^{T+1} = \mathcal{F}_0^{T+1}(\underline{\sigma}(0), u_0^{T+1}, \mathcal{A}_T)\right)\,. \tag{82}$$

Now we analyze each of the terms appearing in this sum. Since the initialization is i.i.d., we have

$$\mathbb{P}(\underline{\sigma}(0)) = \mathbb{P}_0(\sigma_\o(0)) \mathbb{P}(\underline{\sigma}_1(0)) \mathbb{P}(\underline{\sigma}_2(0)) \ldots \mathbb{P}(\underline{\sigma}_{k-1}(0))\,. \tag{83}$$

Further since the coin flips $A_{i,t}$ and $A_{j,t'}$ are independent for $i \neq j$, we have

$$\mathbb{E}_{\mathcal{A}_T}\{\cdots\} = \mathbb{E}_{(A_\o)_0^T} \mathbb{E}_{\mathcal{A}_{1,T}} \ldots \mathbb{E}_{\mathcal{A}_{k-1,T}}\{\cdots\}\,. \tag{84}$$

Finally, the function $\mathcal{F}_0^{T+1}(\cdots)$ can be decomposed as follows

$$\begin{aligned}\mathbb{I}\left(\omega_0^{T+1} = \mathcal{F}_0^{T+1}(\underline{\sigma}(0), u_0^{T+1}, \mathcal{A}_T)\right) &= \mathbb{I}\Big(\sigma_0(0) = \omega(0)\Big)\mathbb{I}\left(\omega_1^{T+1} = \mathcal{F}_1^{T+1}(\underline{\sigma}(0), u_0^{T+1}, \mathcal{A}_T)\right) \\ &= \mathbb{I}\Big(\sigma_0(0) = \omega(0)\Big) \sum_{(\sigma_1)_0^T \dots (\sigma_{k-1})_0^T} \prod_{t=0}^{T} \mathbb{I}\left(\omega(t+1) = f(\sigma_{\o}(t), \underline{\sigma}_{\partial\o}(t), u(t), A_{\o,t})\right) \\ &\qquad \cdot \prod_{i=1}^{k-1} \mathbb{I}\left((\sigma_i)_0^T = \mathcal{F}_0^T(\underline{\sigma}_i(0), \omega_0^T, \mathcal{A}_{i,T-1})\right) . \end{aligned} \tag{85}$$

Using Eqs. (83), (84) and (85) in Eq. (82) and separating terms that depend only on $\underline{\sigma}_i(0)$, we get

$$\begin{aligned}\mathbb{P}((\sigma_{\o})_0^{T+1} = \omega_0^{T+1} || u_0^{T+1}) = \mathbb{P}(\omega(0)) &\sum_{(\sigma_1)_0^T \dots (\sigma_{k-1})_0^T} \prod_{t=0}^{T} \mathbb{I}\left\{\omega(t+1) = f(\sigma_{\o}(t), \underline{\sigma}_{\partial\o}(t), u(t), A_{\o,t})\right\} \\ &\prod_{i=1}^{k-1} \sum_{\bar{\sigma}_i(0)} \mathbb{P}(\underline{\sigma}_i(0))\, \mathbb{I}\left((\sigma_i)_0^T = \mathcal{F}_0^T(\underline{\sigma}_i(0), \omega_0^T, \mathcal{A}_{i,T-1})\right)\end{aligned}$$

□

Notice that the same proof applies to a quite general class of processes on the regular rooted tree $\mathcal{G}_{\o}$. More precisely, we consider a model with spins taking value in a finite domain $\sigma_i(t) \in \mathcal{X}$, and are updated in parallel according to the rule (for $i \neq \o$):

$$\sigma_i(t+1) = f(\sigma_i(t), \underline{\sigma}_{\partial i \setminus \pi(i)}(t), \sigma_{\pi(i)}(t), A_{i,t}) \tag{86}$$

where $\pi(i)$ is the parent of node i (i.e. the only neighbor of i that is closer to the root), and $\{A_{i,t}\}$ are a collection of i.i.d. random variables. For the root $\o$ the above rule is modified by replacing $\sigma_{\pi(i)}(t)$ by the arbitrary quantity $u(t)$.

The next remark follows from a verbatim repetition of our proof.

Remark B.1. *For a model with general update rule (86), the distribution of the root trajectory satisfies Eq. (6) with the kernel*

$$\mathsf{K}_t(\sigma_{\o}(t+1)|\sigma_{\o}(t), \sigma_{\partial\o}(t)) \equiv \mathbb{E}_{A_{\o,t}}\Big\{\mathbb{I}(\sigma_{\o}(t+1) = f\big(\sigma_{\o}(t), \underline{\sigma}_{\partial\o}(t), u(t), A_{\o,t}\big)\Big\} . \tag{87}$$

C Proof of Lemma 4.3

Proof. Equation (26) follows directly from Eq. (10) and Eq. (23). We only need to prove Eq. (27).

Let $\mathcal{I} = \{0, \dots, t\}$ and, for $S \subset \mathcal{I}$, define the rectangle $\mathcal{R}(\omega, S, R, h) \subseteq \mathbb{R}^{t+1}$ as the set of vectors $\eta_0^t = (\eta(0), \dots, \eta(t))$ such that

$$\eta(r) + \sum_{s=0}^{r-1} R(r,s)\omega(s) + h(r) = 0 \qquad \text{for all } r \in S\,, \tag{88}$$

$$\text{sign}\Big(\eta(r) + \sum_{s=0}^{r-1} R(r,s)\omega(s) + h(r)\Big) = \omega(r+1) \qquad \text{for all } r \in \mathcal{I}\setminus S\,. \tag{89}$$

Equation (9) defines $\sigma(t+1)$ as a function of $\sigma(0)$, η_0^t and h. Let us denote this function by writing $\sigma(t+1) = \mathsf{F}_{\sigma(t+1)}(\sigma(0), \eta_0^t; h)$.

$$\begin{aligned}
\mathbb{E}_{C,R,h}[\sigma(t+1)] &= \frac{1}{2} \sum_{\omega(0)\in\{\pm 1\}} \int_{\mathbb{R}^{t+1}} \mathsf{F}_{\sigma(t+1)}(\omega(0), \eta_0^t; h)\, \phi_{0,C_t}(\eta_0^t) \prod_{i=0}^{t} \mathrm{d}\eta(i) \\
&= \frac{1}{2} \sum_{\omega_0^t\in\{\pm 1\}^{t+1}} \int_{\mathbb{R}^{t+1}} \omega(t+1)\, \phi_{0,C_t}(\eta_0^t) \prod_{i=0}^{t} \mathbb{I}\Big\{\mathrm{sign}\Big(\eta(i) + \sum_{s=0}^{i-1} R(i,s)\omega(s) + h(i)\Big) = \omega(i+1)\Big\} \mathrm{d}\eta(i) \\
&= \frac{1}{2} \sum_{\omega_0^t\in\{\pm 1\}^{t+1}} \omega(t+1) \int_{\mathcal{R}(\omega,\emptyset,R,h)} \phi_{0,C_t}(\eta_0^t) \prod_{i=0}^{t} \mathrm{d}\eta(i) \\
&= \frac{1}{2} \sum_{\omega_0^t\in\{\pm 1\}^{t+1}} \omega(t+1)\, \Phi_{0,C_t}(\mathcal{R}(\omega,\emptyset,R,h))\,.
\end{aligned}$$

Since C_t is strictly positive definite by Lemma 4.2, $x \mapsto \phi_{0,C_t}(x)$ is a continuous function. By the fundamental theorem of calculus, we have

$$\left.\frac{\partial \Phi_{0,C_t}(\mathcal{R}(\omega,\emptyset,R,h))}{\partial h(s)}\right|_{h=0} = \begin{cases} \Phi_{0,C_t}(\mathcal{R}(\omega,\{s\},R,0)) & \text{if } \omega(s+1) = +1, \\ -\Phi_{0,C_t}(\mathcal{R}(\omega,\{s\},R,0)) & \text{if } \omega(s+1) = -1. \end{cases}$$

The definition of $R(t,s)$ for a cavity process in Eq.(11) now leads to

$$R(t+1,s) = \frac{1}{2} \sum_{\omega_0^{t+1}\in\{\pm 1\}^{t+2}} \omega(t+1)\omega(s+1)\Phi_{0,C_t}(\mathcal{R}(\omega,\{s\},R,0))$$

for all $t \geq s \geq 0$. The result follows by the change $z'(i) = z(i) + \mu_i(\omega_0^t)$ in the Gaussian integral defining Φ. □

D Proof of Lemma 4.7

Proof. Throughout the proof we let $T = T_* - 1$. Equation (43) continues to hold. We rewrite it as

$$\mathbb{Q}((\sigma_\varnothing)_0^{T+1}||u_0^{T+1}) - \mathbb{P}((\sigma_\varnothing)_0^{T+1}||u_0^{T+1}) = \frac{1}{2}\sum_{r=1}^{k-1} \mathsf{D}(r,k) + O(k^{-(T_*+1)/2})\,, \tag{90}$$

$$\mathsf{D}(r,k) \equiv \binom{k-1}{r} \sum_{(\sigma_1)_0^T\ldots(\sigma_r)_0^T} \prod_{i=1}^{r} \big(\mathbb{Q}((\sigma_i)_0^T||(\sigma_\varnothing)_0^T) - \mathbb{P}((\sigma_i)_0^T||(\sigma_\varnothing)_0^T)\big) \cdot \tag{91}$$

$$\cdot \sum_{(\sigma_{r+1})_0^T\ldots(\sigma_{k-1})_0^T} \prod_{i=r+1}^{k-1} \mathbb{P}((\sigma_i)_0^T||(\sigma_\varnothing)_0^T)) \prod_{t=0}^{T} \mathsf{K}_{u(t)}(\sigma_\varnothing(t+1)|\sigma_{\partial\varnothing}(t))\,.$$

Let $r_0 = \lfloor \log k \rfloor$. Split the summation over r in Eq. (90) into two parts: the first for $1 \leq r \leq r_0$, the second for $r_0 < r \leq k-1$. We will first show that the second part is of order $o(k^{-1/2})$. Indeed, by Lemma

4.6, we know that $\mathbb{Q}((\sigma_i)_0^T||(\sigma_\o)_0^T) - \mathbb{P}((\sigma_i)_0^T||(\sigma_\o)_0^T) \le B/k$ for some constant B and all $(\sigma_i)_0^T \in \{\pm 1\}^{T+1}$. Using the fact that the innermost sum in Eq. (91) is bounded by 1, we get

$$\left|\sum_{r=r_0+1}^{k-1} \mathsf{D}(r,k)\right| \le \sum_{r=r_0+1}^{k-1} \binom{k-1}{r} \sum_{(\sigma_1)_0^T \dots (\sigma_r)_0^T} \prod_{i=1}^{r} \left|\mathbb{Q}((\sigma_i)_0^T||(\sigma_\o)_0^T) - \mathbb{P}((\sigma_i)_0^T||(\sigma_\o)_0^T)\right| \tag{92}$$

$$\le \sum_{r=r_0+1}^{k-1} \binom{k-1}{r} \left(\frac{2^{T+1}B}{k}\right)^r \le \sum_{r \ge \log(k)} \frac{1}{r!}(2^{T+1}B)^r = o(k^{-1/2}), \tag{93}$$

where the last estimate follows from standard tail bounds on Poisson random variables.

We are left with the sum of $\mathsf{D}(r,k)$ over $r \in \{0, \dots, r_0\}$. As in Lemma 4.6, let

$$\mathcal{S}_t \equiv \Big\{(\sigma_{r+1})_0^T \dots (\sigma_{k-1})_0^T : |\sigma_{r+1}(t) + \dots + \sigma_{k-1}(t) + u(t)| \le r_0\Big\}.$$

If $(\sigma_{r+1})_0^T \dots (\sigma_{k-1})_0^T$ is not in $\cup_{t=0}^T \mathcal{S}_t$, then the sum over $(\sigma_1)_0^T \dots (\sigma_r)_0^T$ is 0 due to the normalization of $\mathbb{Q}(\,\cdot\,||(\sigma_\o)_0^T)$ and $\mathbb{P}(\,\cdot\,||(\sigma_\o)_0^T)$ (the same argument was already used in the proof of Lemma 4.6). Restricting the innermost sum and letting as before $\widehat{\mathcal{S}}_{t_0} \equiv \mathcal{S}_{t_o} \cap \{\cap_{t \ne t_0} \overline{\mathcal{S}_t}\}$ with $\mathcal{S}_t$ defined as in Eq. (48), we then have

$$\mathsf{D}(r,k) = \binom{k-1}{r} \sum_{t_0=0}^{T} \sum_{(\sigma_1)_0^T \dots (\sigma_r)_0^T} \prod_{i=1}^{r} \left(\mathbb{Q}((\sigma_i)_0^T||(\sigma_\o)_0^T) - \mathbb{P}((\sigma_i)_0^T||(\sigma_\o)_0^T)\right) \cdot \tag{94}$$

$$\cdot \sum_{((\sigma_{r+1})_0^T \dots (\sigma_{k-1})_0^T) \in \widehat{\mathcal{S}}_{t_0}} \prod_{i=r+1}^{k-1} \mathbb{P}((\sigma_i)_0^T||(\sigma_\o)_0^T)) \prod_{t=0}^{T} \mathsf{K}_{u(t)}(\sigma_\o(t+1)|\sigma_{\partial\o}(t)) \,+\, \mathsf{R}(r,k).$$

By inclusion-exclusion, the error term is bounded as

$$|\mathsf{R}(r,k)| \le \binom{k-1}{r} \sum_{t_1 \ne t_2} \sum_{(\sigma_1)_0^T \dots (\sigma_r)_0^T} \prod_{i=1}^{r} \left|\mathbb{Q}((\sigma_i)_0^T||(\sigma_\o)_0^T) - \mathbb{P}((\sigma_i)_0^T||(\sigma_\o)_0^T)\right| \cdot$$

$$\cdot \sum_{((\sigma_{r+1})_0^T \dots (\sigma_{k-1})_0^T) \in \mathcal{S}_{t_1} \cap \mathcal{S}_{t_2}} \prod_{i=r+1}^{k-1} \mathbb{P}((\sigma_i)_0^T||(\sigma_\o)_0^T)) \prod_{t=0}^{T} \mathsf{K}_{u(t)}(\sigma_\o(t+1)|\sigma_{\partial\o}(t))$$

$$\le \binom{k-1}{r} \sum_{t_1 \ne t_2} \sum_{(\sigma_1)_0^T \dots (\sigma_r)_0^T} \prod_{i=1}^{r} \left|\mathbb{Q}((\sigma_i)_0^T||(\sigma_\o)_0^T) - \mathbb{P}((\sigma_i)_0^T||(\sigma_\o)_0^T)\right| \frac{Br_0^2}{k}$$

$$\le \binom{k-1}{r} T^2 2^{Tr} \left(\frac{B}{k}\right)^r \frac{Br_0^2}{k}.$$

The first inequality follows by applying Lemma 4.4 to the $N = k - r - 1 \ge k - \log(k) - 1$ i.i.d. random vectors $(\sigma_{r+1})_0^T, \dots, (\sigma_{k-1})_0^T$, which are non-degenerate for all k large enough by Lemma 4.5, and summing over the values of $a_{t_1} = \sum_{i=r+1}^{k-1} \sigma_i(t_1) + u(t_1)$ and $a_{t_2} = \sum_{i=r+1}^{k-1} \sigma_i(t_2) + u(t_2)$, with $|a_{t_1}|, |a_{t_2}| \le r_0$. The second inequality is instead implied by Lemma 4.6. It is now easy to sum over r to get

$$\left|\sum_{r=1}^{r_0} \mathsf{R}(r,k)\right| \le \sum_{r=0}^{\infty} \frac{1}{r!} T^2 (2^T B)^r B \,\frac{(\log k)^2}{k} = o(k^{-1/2}).$$

Therefore the error terms $\mathsf{R}(r,k)$ can be neglected.

Let us now consider the main term in Eq. (94), and define

$$J'_{t_0}((\sigma_\o)_0^T,(\sigma_1)_0^T,\dots,(\sigma_r)_0^T) \equiv \sum_{((\sigma_{r+1})_0^T\dots(\sigma_{k-1})_0^T)\in\widehat{\mathcal{S}}_{t_0}} \prod_{i=r+1}^{k-1}\mathbb{P}((\sigma_i)_0^T||(\sigma_\o)_0^T))\prod_{t=0}^{T}\mathsf{K}_{u(t)}(\sigma_\o(t+1)|\sigma_{\partial\o}(t))\,.$$

We now proceed exactly as in the proof of Lemma 4.6, cf. Eq. (50) to (52) with $\Omega(t)=\sigma_\o(t+1)(\sum_{i=1}^r\sigma_i(t))$ and $r_0=\log(k)$. Notice Theorem 4.4 continues to hold and r_0 times the $O(k^{-1/4})$ error is still $o(1)$. We arrive at

$$J_{t_0} \;=\; \frac{1}{\sqrt{k}}\,\sigma_\o(t_0+1)\Big(\sum_{i=1}^r\sigma_i(t_0)\Big)\,J^*_{t_0}\,(1+\tilde{R}_{t_0}(k))\,,$$

where $\tilde{R}_{t_0}(k)\to 0$ as $k\to\infty$ for any fixed t_0.

If we use this estimate in Eq. (94), we get

$$\begin{aligned}
&\mathsf{D}(r,k) = \\
&= \binom{k'}{r}\sum_{t_0=0}^{T}\sum_{\{(\sigma_i)_0^T\}}\prod_{i=1}^{r}\left(\mathbb{Q}((\sigma_i)_0^T||(\sigma_\o)_0^T)-\mathbb{P}((\sigma_i)_0^T||(\sigma_\o)_0^T)\right)\frac{J^*_{t_0}}{\sqrt{k}}\sigma_\o(t_0+1)\sum_{i=1}^{r}\sigma_i(t_0)(1+\;\tilde{R}_{t_0}(k))+o(k^{-1/2})\\
&= r\binom{k'}{r}\sum_{t_0=0}^{T}\sum_{\{(\sigma_i)_0^T\}}\prod_{i=1}^{r}\left(\mathbb{Q}((\sigma_i)_0^T||(\sigma_\o)_0^T)-\mathbb{P}((\sigma_i)_0^T||(\sigma_\o)_0^T)\right)\frac{J^*_{t_0}}{\sqrt{k}}\,\sigma_\o(t_0+1)\sigma_1(t_0)(1+\;\tilde{R}_{t_0}(k))+o(k^{-1/2})\,,
\end{aligned}$$

where $k'\equiv k-1$ and we used the symmetry among the vertices $\{1,\dots,r\}$ to replace $(\sum_{i=1}^r\sigma_i(t))$ by $r\sigma_1(t)$. If $r\geq 2$, the sums over $(\sigma_2)_0^T,\dots,(\sigma_r)_0^T$ vanish except for the error terms $\tilde{R}_{t_0}(k)$ (once more by the normalization of $\mathbb{P}(\,\cdot\,||(\sigma_\o)_0^T)$ and $\mathbb{Q}(\,\cdot\,||(\sigma_\o)_0^T)$). We need to bound contribution of such error terms. Find M such that $\left|\mathbb{Q}((\sigma_i)_0^T||(\sigma_\o)_\o^T)-\mathbb{P}((\sigma_i)_0^T||(\sigma_\o)_0^T)\right|\leq M/k$. We have

$$\begin{aligned}
&\left|r\binom{k-1}{r}\sum_{(\sigma_1)_0^T\dots(\sigma_r)_0^T}\left(\mathbb{Q}((\sigma_i)_0^T||(\sigma_\o)_\o^T)-\mathbb{P}((\sigma_i)_0^T||(\sigma_\o)_0^T)\right)\tilde{R}_{t_0}(k)\right|\\
&\leq r\left(\frac{(k-1)e}{r}\right)^r 2^T\left(\frac{M}{k}\right)^r|\tilde{R}_{t_0}(k)|\\
&\leq r\left(\frac{2^TeM}{r}\right)^r|\tilde{R}_{t_0}(k)|\\
&\leq \left(\frac{M'}{2^r}\right)|\tilde{R}_{t_0}(k)| \qquad (95)
\end{aligned}$$

for suitable M'. Here we have used the standard bound $\binom{n}{m}\leq\left(\frac{ne}{m}\right)^m$. Summing (95) over t_0 and r, we see that $\sum_{r=2}^{r_0}|\mathsf{D}(k,r)|\leq C|J^*_{t_0}\tilde{R}_{t_0}(k)|/sqrtk=o(k^{-1/2})$.

Further,

$$
\begin{aligned}
&\sum_{(\sigma_1)_0^T} \sigma_1(t)\big\{\mathbb{Q}((\sigma_1)_0^T||(\sigma_\emptyset)_0^T) - \mathbb{P}((\sigma_1)_0^T||(\sigma_\emptyset)_0^T)\big\} \\
=&\sum_{(\sigma_1)_0^t} \sigma_1(t)\big\{\mathbb{Q}((\sigma_1)_0^t||(\sigma_\emptyset)_0^t) - \mathbb{P}((\sigma_1)_0^t||(\sigma_\emptyset)_0^t)\big\} \\
=&2\frac{\omega_t}{k^{(T_*-t+1)/2}}\,(1+o(1))\,.
\end{aligned}
$$

where the second equality follows by Lemma 4.6. Note that for $t < T_* - 1$, this sum is $o(k^{-1})$. As a consequence, only the $t_0 = T$ term is relevant in the sum over t_0.

Using these two remarks we finally obtain

$$
\begin{aligned}
\sum_{r=1}^{r_0} \mathsf{D}(k,r) =&\mathsf{D}(k,1) + o(k^{-1/2}) \\
=&k\sum_{t_0=0}^{T}\frac{J^*_{t_0}}{\sqrt{k}}\;2\frac{\omega_{t_0}}{k^{(T_*-t_0+1)/2}}\;\sigma_\emptyset(t_0+1)(1+o(1)) + o(k^{-1/2}) \\
=&2\frac{\omega_{T_*-1}}{k^{1/2}}\;\sigma_\emptyset(T_*)\;I\left((\sigma_\emptyset)_0^{T_*-1}\right)\;\Big(1+o(1)\Big)\,,
\end{aligned}
$$

which, together with Eq. (93) and Eq. (90), proves our thesis. Equation (55) follows as in the previous lemma. □

E Proof of the local central limit theorem

The proof repeats the arguments of [DMD94], while keeping track explicitly of error terms. We will therefore focus on the differences with respect to [DMD94]. We will indeed prove a result that is slightly stronger than Theorem 4.4. Apart from a trivial rescaling, the statement below differs from Theorem 4.4 in that we allow for larger deviations from the mean.

Theorem E.1. *Let $X_1, \dots, X_N$ be i.i.d. vectors $X_i = (X_{i,1}, X_{i,2}, \dots, X_{i,d}) \in \{0,1\}^d$ with*

$$
\Big|\mathbb{P}\{X_{1,\ell} = 1\} - \frac{1}{2}\Big| \le \frac{B}{\sqrt{N}}\,, \tag{96}
$$

for $\ell \in \{1, \dots, d\}$. Further assume $\mathbb{P}\{X_i = s\} \ge \mathbb{P}\{X_i = 0\} \ge 1/B$ for all $s \in \{0,1\}^d$.

Let $a \in \mathbb{Z}^d$ be such that $\sup_i |a_i - N/2| \le B\sqrt{N}$, and define, for a partition $\{1, \dots, d\} = \mathcal{I}_0 \cup \mathcal{I}_+$,

$$
\begin{aligned}
A(a,\mathcal{I}) &\equiv \{z \in \mathbb{Z}^d : z_i = a_i \ \forall\, i \in \mathcal{I}_0,\ \ z_i \ge a_i \ \forall\, i \in \mathcal{I}_+\}\,,\\
A_\infty(a,\mathcal{I}) &\equiv \{z \in \mathbb{R}^d : z_i = a_i/\sqrt{N} \ \forall\, i \in \mathcal{I}_0,\ \ z_i \ge a_i/\sqrt{N} \ \forall\, i \in \mathcal{I}_+\}\,.
\end{aligned}
$$

Then, for $K = |\mathcal{I}_0|$,

$$
\begin{aligned}
&\Big|F(a,\mathcal{I}) - \frac{1}{N^{K/2}}\Phi_{\sqrt{N}\mathbb{E}X_1,\mathrm{Cov}(X_1)}(\mathcal{A}_\infty(a,\mathcal{I}))\Big| \le \frac{L(B,d)}{N^{(K+(K+1)^{-1})/2}} \\
&F(a,\mathcal{I}) \equiv \sum_{y \in A(a,\mathcal{I})} p_N(y)\,.
\end{aligned}
\tag{97}
$$

Since $\Phi_{\sqrt{N}\mathbb{E}X_1,\mathrm{Cov}(X_1)}(\mathcal{A}_\infty(\mathcal{I}))$ is bounded away from 0 for B bounded, the error estimate in the last statement is equivalent to the one in Theorem 4.4. For $K = 0$ our claim is implied by the multi-dimensional Berry-Esseen theorem [BR76], and we will therefore focus on $K \geq 1$.

Recall that the Bernoulli decomposition of [DMD94] allows to write, for $S_N = (S_{N,1}, \dots, S_{N,d})$ and $r \in \{1, \dots, d\}$

$$S_{N,r} = Z_{N,r} + \sum_{i=1}^{M_{N,r}} L_{i,r} \tag{98}$$

where Z_N is a lattice random variable, $M_{N,r}$ ∼Binom(N, q_r) for $r = 1, \dots, d$, and $\{L_{i,r}\}$ is a collection of i.i.d. Bernoulli(1/2) random variables independent from Z_N and M_N. Finally, it is easy to check that $q_r \geq 1/(Bd)$.

We have the following key estimate.

Lemma E.2. *There exists a numerical constant C such that, for any $a, b \in \mathbb{Z}^d$*

$$\big|F(a,\mathcal{I}) - F(b,\mathcal{I})\big| \leq C \left(\frac{Bd}{N}\right)^{(K+1)/2} ||a-b|| \,. \tag{99}$$

where $||\cdot||$ denotes the L^1 norm.

Proof. As in [DMD94], we let, for $x, m \in \mathbb{Z}^d$,

$$r_m(x) \equiv \prod_{i=1}^{d} \frac{1}{2^{m_i}} \binom{m_i}{x_i}, \tag{100}$$

be the probability mass function of the vector $\Lambda_m \equiv (\sum_{i=1}^{m_1} L_{i,1}, \dots, \sum_{i=1}^{m_d} L_{i,d})$. It then follows immediately that

$$\Big|\sum_{x\in A(a,\mathcal{I})} r_m(x) - \sum_{y\in A(b,\mathcal{I})} r_m(y)\Big| \leq \frac{C_*}{\min_i(m_i)^{(K+1)/2}} ||a-b|| \,, \tag{101}$$

for some numerical constant C. This is is a slight generalization of Lemma 2.2 of [DMD94], and follows again immediately from the same estimates on the combinatorial coefficients used in [DMD94].

We then proceed analogously to the proof of Theorem 2.1 of [DMD94], namely, for $h \in \mathbb{Z}^d$,

$$\begin{aligned}
&\sup_{a\in\mathbb{Z}^d} |F(a+h,\mathcal{I}) - F(a,\mathcal{I})| \\
&\leq \sup_{a\in\mathbb{Z}^d} \sum_{m\in\mathbb{Z}^d} \mathbb{P}\{M_N = m\} \big|\mathbb{P}\{S_N \in A(a,\mathcal{I}) | M_N = m\} - \mathbb{P}\{S_N \in A(a+h,\mathcal{I}) | M_N = m\}\big| \\
&= \sup_{a\in\mathbb{Z}^d} \sum_{m\in\mathbb{Z}^d} \mathbb{P}\{M_N = m\} \Big|\mathbb{P}\big\{Z_N + \Lambda_m \in A(a,\mathcal{I}) \big| M_N = m\big\} - \mathbb{P}\big\{Z_N + \Lambda_m \in A(a+h,\mathcal{I}) \big| M_N = m\big\}\Big| \\
&\leq \sup_{a\in\mathbb{Z}^d} \sum_{m\in\mathbb{Z}^d} \mathbb{P}\{M_N = m\} \sum_{l\in\mathbb{Z}^d} \mathbb{P}\{Z_N = l\} \Big|\mathbb{P}\big\{\Lambda_m \in A(a-l,\mathcal{I}) \big| M_N = m\big\} - \mathbb{P}\big\{\Lambda_m \in A(a+h-l,\mathcal{I}) \big| M_N = m\big\}\Big| \\
&\leq \sum_{m\in\mathbb{Z}^d} \frac{C_*}{\min_i(m_i)^{(K+1)/2}} ||h||
\end{aligned}$$

which is bounded as in the statement by the same argument used in [DMD94]. □

We are now in a position to prove Theorem E.1.

Proof. (Theorem E.1) For a as in the statement and $\ell > 0$, let

$$\begin{aligned} R(a,\ell) &= \left\{z \in \mathbb{Z}^d : |z_i - a_i| \le \ell \ \forall i \in \mathcal{I}_0, \ \ z_i = a_i \ \forall i \in \mathcal{I}_+\right\}, \\ R_\infty(a,\ell) &= \left\{z \in \mathbb{R}^d : |z_i - a_i/\sqrt{N}| \le \ell/\sqrt{N} \ \forall i \in \mathcal{I}_0, \ \ z_i = a_i/\sqrt{N} \ \forall i \in \mathcal{I}_+\right\}. \end{aligned}$$

Then, by Lemma E.2, there exists a constant $C_1(B,d)$ such that

$$\left| F(a,\mathcal{I}) - \frac{1}{|R(a,\ell)|} \sum_{z \in R(a,\ell)} F(z,\mathcal{I}) \right| \le \frac{C_1(B,d)\ell}{N^{(K+1)/2}}. \tag{102}$$

On the other hand, by the Berry-Esseen theorem

$$\left| \sum_{z \in R(a,\ell)} F(x,\mathcal{I}) - \int_{R_\infty(a,\ell)} \Phi_{\sqrt{N}\mathbb{E}X_1,\mathrm{Cov}(X_1)}(\mathcal{A}_\infty(z,\mathcal{I}))\,\mathrm{d}z \right| \le \frac{C_2(d)}{N^{1/2}}. \tag{103}$$

Finally, it is easy to see that $\Phi_{\sqrt{N}\mathbb{E}X_1,\mathrm{Cov}(X_1)}(\mathcal{A}_\infty(z,\mathcal{I}))$ is Lipschitz continuous in z with Lipschitz constant bounded uniformly in N, whence

$$\left| \Phi_{\sqrt{N}\mathbb{E}X_1,\mathrm{Cov}(X_1)}(\mathcal{A}_\infty(a,\mathcal{I})) - \frac{1}{|R_\infty(a,\ell)|} \int_{R_\infty(a,\ell)} \Phi_{\sqrt{N}\mathbb{E}X_1,\mathrm{Cov}(X_1)}(\mathcal{A}_\infty(z,\mathcal{I}))\mathrm{d}z \right| \le \frac{C_3\ell}{\sqrt{N}}. \tag{104}$$

The proof is completed by putting together Eqs. (102), (103) and (104), using $|R(a,\ell)| = \Theta(\ell^K)$, $|R_\infty(a,\ell)| = \Theta(\ell^K N^{-K/2})$, and setting $\ell = N^{K/(2K+2)}$. □