The Annals of Applied Probability
1998, Vol. 8, No. 4, 1216–1250

ON THE STORAGE CAPACITY OF HOPFIELD MODELS WITH CORRELATED PATTERNS

BY MATTHIAS LÖWE

Universität Bielefeld

We analyze the storage capacity of the Hopfield model with correlated patterns (ξ_i^ν). We treat both the case of semantically and spatially correlated patterns (i.e., the patterns are either correlated in ν but independent in i or vice versa). We show that the standard Hopfield model of neural networks with N neurons can store $N/(\gamma \log N)$ or αN correlated patterns (depending on which notion of storage is used), provided that the correlation comes from a homogeneous Markov chain. This answers the open question whether the standard Hopfield model can store any increasing number of correlated patterns at all in the affirmative. While our bound on the critical value for α decreases with large correlations, the critical γ behaves differently for the different types of correlations.

1. Introduction. The study of neural networks has been a major topic of research in mathematics, physics and computer science during the past 15 years.

Basically a model of a neural network consists of a labeled and possibly oriented graph $G = (V, E)$ together with a set S with $\mathrm{card}(S) \geq 2$ to describe the set of neurons (by V), the synapses connecting these neurons (by E) and the activity of each of the neurons (by a variable $\sigma_i \in S$ for $i \in V$). Moreover, the information to be stored is encoded in so-called patterns ξ^μ, $\mu = 1, \ldots, M(N)$, each of the ξ^μ itself being a sequence of $\xi_i^\mu \in S$, $i = 1, \ldots, N$. Finally, to make the neural net capable of adapting to different sequences of patterns, we have to introduce a set of variables J_{ij}, $ij \in E$ called the synaptic efficacies and describing the strength of interaction between the neurons at sites i and j. It is commonly assumed that the variable J_{ij} is measurable with respect to the set $\{\xi_i^\mu, \xi_j^\mu,\ \mu = 1, \ldots, M(N)\}$ (so-called *locality* of the weights J_{ij}). The basic idea now behind this set-up is to define a Hamiltonian $H_N(\sigma)$ on S^N, such that the Monte Carlo dynamics governed by this Hamiltonian eventually converges to one of $M(N)$ states close to the original patterns.

One of the classical and best understood examples of a neural network is the so-called Hopfield model [14]. Although originally introduced by Pastur and Figotin [22] as a simplified model of a spin-glass, this model achieved much of its interest by its reinterpretation as an auto-associative memory by Hopfield and may therefore rightly be called the Hopfield model. Here the graph G is the complete graph K_N on N vertices, $S = \{-1, +1\}$, corresponding to the fact that a neuron may be either switched "on" or "off" and the weights J_{ij}

Received September 1997; revised December 1997.
AMS 1991 *subject classifications*. Primary 82C32; secondary 82B44, 60K35.
Key words and phrases. Hopfield model, neural networks, storage capacity, Markov chains, large deviations.

are given by "Hebb's learning rule," that is, the formula

$$J_{ij} := \frac{1}{N} \sum_{\mu=1}^{M(N)} \xi_i^\mu \xi_j^\mu.$$

One of the most important advances due to [14] has been to understand that this set-up corresponds to a Hamiltonian H_N given by

$$H_N(\sigma) := -\frac{1}{2} \sum_{i,\,j=1}^{N} J_{ij}\sigma_i\sigma_j = -\frac{1}{2N} \sum_{i,\,j=1}^{N} \sum_{\mu=1}^{M(N)} \xi_i^\mu \xi_j^\mu \sigma_i \sigma_j. \tag{1}$$

Note that (1) may be rewritten in the very convenient form

$$H_N(\sigma) = -\frac{N}{2} \sum_{\mu=1}^{M(N)} (m_N^\mu(\sigma))^2 = -\frac{N}{2} ||m_N(\sigma)||_2^2, \tag{2}$$

where $||m_N(\sigma)||_2^2$ is the l_2-norm of the so-called overlap vector

$$m_N(\sigma) := (m_N^\mu(\sigma))_{\mu=1,\dots,M(N)} = \left(\frac{1}{N} \sum_{i=1}^{N} \sigma_i \xi_i^\mu\right)_{\mu=1,\dots,M(N)}.$$

Note that $m_N(\sigma)$ may be regarded as an index for how much a configuration is correlated to one of the given patterns (large absolute values of a component of $m_N(\sigma)$ corresponding to large correlations). Also observe that (2) makes it plausible—at least for $M := M(N)$ small enough—that the minima of H_N are located close to the patterns ξ^μ (notice that this is trivially fulfilled, if the patterns are orthogonal, that is, if $m_N^\nu(\xi^\mu) = \delta_{\mu,\nu}$).

Indeed, much of the recent work on the Hopfield model can be summarized under the aspect of making precise and mathematically verifying this last statement. To this end, in most of the papers it has been assumed that the ξ_i^μ are unbiased i.i.d. random variables, that is, that $P(\xi_i^\mu = +1) = P(\xi_i^\mu = -1) = \frac{1}{2}$ for all i and μ independently of all the other ξ_j^ν.

Hopfield [14] had already discovered by computer simulations that in this case there is a value $\alpha_c > 0$ such that if $M(N) \leq \alpha_c N$, almost all patterns are memorized, whereas for $M(N) > \alpha_c N$ the Hopfield model tends to "forget" all of the patterns and therefore one may rightly speak about a storage capacity of the Hopfield model. The numerical value of α_c found by Hopfield was close to 0.14. This finding with a similar value for α_c has been supported by nonrigorous analytical computations (including the notorious replica-trick) by Amit, Gutfreund and Sompolinsky in [2] and [3] (for the current state of the art concerning the replica method, see [15]). The first rigorous results concerning the storage capacity of the Hopfield model were by McEliece, Posner, Rodemich and Venkatesh [18], showing that if one focuses on exact reproduction of the patterns, the memory capacity of the Hopfield model is only in the order of const.$(N/\log N)$ where the constant ranges between 1/2 and 1/6 and mainly depends on whether one is interested in "almost sure" results or in results holding with “probability converging to one” (see the corresponding results in Section 2; also see the survey paper [23]).

On the other hand, Newman [21] showed rigorously that, if small errors are tolerated, the Hopfield model is able to successfully retrieve a number of patterns M proportional to the number of neurons N, and his (lower) bound on the constant α_c was $\alpha_c \geq 0.056$. This bound recently has been improved by Loukianova [16] to $\alpha_c \geq 0.071$ and by Talagrand [25, 26] to $\alpha_c \geq 0.08$ by a refined use of Newman's technique. But to verify Hopfield's value for α_c rigorously and to give a good picture of what happens above this bound still remains a major open problem and only very few results on this subject are available (see [17] and [26]).

However, nowadays the Hopfield model is well understood in the regime where the patterns are chosen independently and their number is “small” compared to the number of neurons (also in the case of nonzero temperatures). The corresponding analysis has been carried out in a series of papers by Bovier and Gayrard, partially in collaboration with Picco [8, 9, 10, 5, 7]. Another milestone certainly has been set in a recent paper, [26], especially for what concerns the validity of the so-called replica–symmetric solution in the Hopfield model. For a comprehensive review over the rapid development in this area during the last few years and a particularly nice proof of the validity of the replica-symmetric solution, we refer the reader to [6]. In this context, of course, a result due to Gentz [12] should be mentioned where a central limit theorem for the overlap parameter in this model is proved.

Now in most realistic situations the patterns are not at all independent (see, e.g, [19]). Indeed, there are at least two sensible ways to introduce correlations among the patterns. One is to consider spatial correlation, that is, to choose the patterns correlated in i but independent in μ, which may be interesting when, for example, thinking about the patterns as images to be stored. The other way is to choose sequentially or semantically correlated patterns, which means that the dependency now enters via μ only. This situation might be useful as a very simple model for patterns with some sort of causal relations, as, for example, films (for a Hopfield model with deterministic sequences of i.i.d. patterns, see [24]).

In both situations the picture is less clear than in the “i.i.d. case.” It seems that the idea of encoding correlated patterns first was mentioned in [11] and [1]. Some investigations based on heuristic arguments claim that correlations among the patterns can only increase the storage capacity of the Hopfield model [20, 27]. This may be supported by the fact that in terms of information content a sequence of correlated patterns contains less information than a sequence of uncorrelated ones. On the other hand—based on a signal-noise-analysis for biased data—some other authors (see, e.g., [28]) argue that the Hopfield model cannot store any increasing (in N) amount of correlated patterns. The basic idea behind this argument is that in the presence of correlations among the patterns the cross-talk terms between two patterns become nonvanishing, and therefore the noise created by these cross-talk terms becomes large compared to the signal coming from a single pattern.

In this paper we will show that both effects may occur depending on the notion of storage and the type of dependencies we consider. In the case of semantically correlated data (which in our case come from a Markov chain), we will see that the numbers of patterns we can store such that all of them are fixed points of the retrieval dynamics is an increasing function of the strength of the correlation, while our bound on the storage capacity decreases in the strength of the correlation, if we define it in the sense of [21].

The situation of spatially correlated patterns may be considered to be even more interesting but definitely is also more challenging because the Hamiltonian for a fixed edge ij depends linearly on $\xi_i^\mu \xi_j^\mu$ (more precisely it depends on the sum of these variables), but for a fixed pattern μ it is a quadratic form in ξ_i^μ [see (1)]. We also will treat the case of spatially correlated patterns in this paper, again provided that the correlations stem from a homogeneous Markov chain (corresponding to one-dimensional images). We show that the Hopfield model can store $N/(\gamma \log N)$ or even $\alpha_c N$ spatially correlated patterns (depending on the notion of storage capacity). Here the lower bounds on the constants $1/\gamma$ and α_c are decreasing with an increasing strength of the correlation.

This paper has three further sections. Section 2 contains the basic set-up, including a definition of the two notions of storage capacity we use and a description of the homogeneous Markov chains we have in mind. Moreover, we state our main results. Section 3 will contain some auxiliary lemmata. Section 4 is devoted to the proofs.

2. The set-up and the main results. In this section we will mainly state our results on the storage capacity of the Hopfield model with correlated patterns.

First, let us briefly explain the two different concepts we are dealing with. The idea behind the first notion of storage capacity is that a possible retrieval dynamics is a Monte Carlo dynamics at zero temperature working as follows: Choose a site i at random. Flip the spin σ_i, if flipping lowers the energy (the Hamiltonian) and let the spin σ_i be unchanged otherwise. On a more formal level we define the gradient dynamics T on the energy landscape on $\{-1, +1\}$ induced by H_N via

$$T_i\colon \sigma_i \mapsto \operatorname{sgn}\Biggl(\sum_{\substack{j=1\\ j\neq i}}^{N} \sigma_j J_{ij}\Biggr),$$

where sgn is the sign function. The map T is then defined by $T(\sigma) := (T_i(\sigma_i))_i$. We will call a configuration $\sigma = (\sigma_i)_{i\le N}$ stable if it is a fixed point of T, that is,

$$\sigma_i = \operatorname{sgn}\Biggl(\sum_{\substack{j=1\\ j\neq i}}^{N} \sigma_j J_{ij}\Biggr) \quad \text{for all } i = 1, \dots, N,$$

which means that σ is a local minimum of the Hamiltonian. The storage capacity in this concept is defined as the greatest number of patterns $M := M(N)$ such that all the patterns ξ^ν are stable in the above sense almost surely or with probability converging to one. [Here and in the following the notion "almost surely" refers to the probability measure on the space of all sequences of patterns (of infinite length) while in "probability converging to one" the convergence is with $N \to \infty$.]

The other approach to storage capacity is due to [2] and has rigorously been analyzed in [21]. It takes into consideration that we are possibly willing to tolerate small errors in the restoration of the patterns. So we are satisfied if the retrieval dynamics converges to a configuration which is not too far away from the original patterns. Thus in this concept a pattern ξ^ν is called stable if it is close to a local minimum of the Hamiltonian or, in other words, if it is surrounded by a sufficiently high energy barrier. Technically speaking, we will call ξ^ν stable if there exist $\varepsilon > 0$ and $\delta > 0$ such that

$$\inf_{\sigma \in S_\delta(\xi^\nu)} H_N(\sigma) \geq H_N(\xi^\nu) + \varepsilon N. \tag{3}$$

Here the set $S_\delta(\xi^\nu)$ (over which the infimum is taken) is the Hamming sphere of radius δN centered in ξ^ν. Again we will use the notion of storage capacity for the maximal number $M(N)$ of patterns such that (3) holds true for all ξ^ν almost surely.

Before we can state our result, we have to describe the form of correlations we are going to study. We will differentiate between spatial and semantical correlations.

For the semantical correlations we will assume that the random variables $(\xi_i^\mu)_{i\in\mathbb{N},\,\mu\in\mathbb{N}}$ are independent for different i and for fixed i form a Markov chain in μ with initial distribution

$$P(\xi_i^1 = x_i^1,\ i = 1, \dots, N) = 2^{-N} \quad \text{for all } x_i^1 \in \{-1, 1\} \tag{4}$$

and transition probabilities

$$\begin{aligned} &P(\xi_i^\mu = x_i^\mu | \xi_j^\nu = x_j^\nu,\ j = 1, \dots, N,\ \nu = 1, \dots, \mu - 1) \\ &\qquad = P(\xi_i^\mu = x_i^\mu | \xi_i^{\mu-1} = x_i^{\mu-1}) = Q(x_i^{\mu-1}, x_i^\mu). \end{aligned} \tag{5}$$

Similarly, for the case of spatial correlations we will assume that the random variables $(\xi_i^\mu)_{i\in\mathbb{N},\,\mu\in\mathbb{N}}$ are independent for different μ and for fixed μ form a Markov chain in i with initial distribution

$$P(\xi_1^\mu = x_1^\mu,\ \mu = 1, \dots, M) = 2^{-M} \quad \text{for all } x_1^\mu \in \{-1, 1\} \tag{6}$$

and transition probabilities

$$\begin{aligned} &P(\xi_i^\mu = x_i^\mu | \xi_j^\nu = x_j^\nu,\ j = 1, \dots, i-1,\ \nu = 1, \dots, M) \\ &\qquad = P(\xi_i^\mu = x_i^\mu | \xi_{i-1}^\mu = x_{i-1}^\mu) = Q(x_{i-1}^\mu, x_i^\mu). \end{aligned} \tag{7}$$

In (5) and (7) Q denotes a symmetric 2×2 matrix with entries

$$Q = \begin{pmatrix} p & 1-p \\ 1-p & p \end{pmatrix},$$

where $0 < p < 1$ (note that $p = \frac{1}{2}$ is the case of independent patterns).

Let us also mention that for the spatial correlations we have used the notation and set-up of a Markov chain to stress the similarity with the case of semantically correlated patterns and since it is more convenient. However, let us mention that this set-up is equivalent to that of a one-dimensional homogeneous Markov random field (see, e.g., [13], Chapter 3), which maybe somewhat closer to the usual modeling of an image.

With these definitions, our results concerning the storage capacity for correlated patterns read as follows.

THEOREM 2.1. *Assume the random patterns ξ^μ fulfill* (4) *and* (5) *and suppose that $M(N) = N/(\gamma \log N)$.*

Then the following assertions hold true.

(i) *If $\gamma > 48p(1-p)(p^2 + (1-p)^2)$,*

$$P\biggl(\liminf_{N\to\infty}\biggl(\bigcap_{\mu=1}^{M(N)} T\xi^\mu = \xi^\mu\biggr)\biggr) = 1;$$

that is, the patterns are almost surely stable.

(ii) *If $\gamma > 32p(1-p)(p^2 + (1-p)^2)$,*

$$P\biggl(\biggl(\bigcap_{\mu=1}^{M(N)} T\xi^\mu = \xi^\mu\biggr)\biggr) = 1 - R_N$$

with $\lim_{N\to\infty} R_N = 0$; that is, all the patterns are stable with probability converging to one.

(iii) *If $\gamma > 16p(1-p)(p^2 + (1-p)^2)$ for every fixed $\mu = 1, \dots, M(N)$,*

$$P(T\xi^\mu = \xi^\mu) = 1 - R_N$$

with $\lim_{N\to\infty} R_N = 0$; that is, every fixed pattern is stable with probability converging to one.

THEOREM 2.2. *Assume the random patterns ξ^μ satisfy* (6) *and* (7) *and suppose that $M(N) = N/(\gamma \log N)$.*

Then the following assertions hold true.

(i) *If $\gamma > 3(p^2 + (1-p)^2)/p(1-p)$,*

$$P\biggl(\liminf_{N\to\infty}\biggl(\bigcap_{\mu=1}^{M(N)} T\xi^\mu = \xi^\mu\biggr)\biggr) = 1;$$

that is, the patterns are almost surely stable.

(ii) *If* $\gamma \geq 2(p^2 + (1-p)^2)/p(1-p)$,

$$P\left(\left(\bigcap_{\mu=1}^{M(N)} T\xi^\mu = \xi^\mu\right)\right) = 1 - R_N$$

with $\lim_{N\to\infty} R_N = 0$; *that is, all the patterns are stable with probability converging to one.*

(iii) *If* $\gamma > p^2 + (1-p)^2/p(1-p)$ *for every fixed* $\mu = 1, \dots, M(N)$,

$$P(T\xi^\mu = \xi^\mu) = 1 - R_N$$

with $\lim_{N\to\infty} R_N = 0$; *that is, every fixed pattern is stable with probability converging to one.*

Note that $1/\gamma \to \infty$ as $p \to 0$ or $p \to 1$ in the case of semantical correlations (i.e., the storage capacity increases with large correlations) while $1/\gamma \to 0$ as $p \to 0$ or $p \to 1$ in the case of spatial correlations (implying a decrease of the bound on the storage with large correlations). Also notice that for $p = 1/2$ (the case of independent patterns) the bounds coincide with the well-known bounds for the Hopfield model with i.i.d. patterns (see [18], [23]).

Let us now turn to the second notion of storage capacity. We will see that if small errors are tolerated, the Hopfield model indeed can store a number of spatially correlated patterns M proportional to the number of neurons N. The behavior of the critical value α_c will depend on the strength and type of the correlation and α_c decreases with an increasing correlation (this basically means that our bounds on α_c behave in such a way).

THEOREM 2.3. *Suppose that the random patterns fulfill* (4) *and* (5). *There exists an* $\alpha_c := \alpha_c^{\mathrm{SEM}} > 0$ *(depending on* p*) such that if* $M(N) \leq \alpha_c N$, *then there are* $\varepsilon > 0$ *and* $0 < \delta < 1/2$ *such that*

$$P\left(\liminf_{N\to\infty}\left(\bigcap_{\mu=1}^{M(N)} \bigcap_{\sigma\in S_\delta(\xi^\mu)} (H_N(\sigma) \geq H_N(\xi^\mu) + \varepsilon N)\right)\right) = 1,$$

where $S_\delta(\xi^\mu)$ *is the Hamming sphere of radius* δN *centered in* ξ^ν.

Similarly for the case of spatial correlations, we have the following theorems.

THEOREM 2.4. *Suppose that the random patterns fulfill* (6) *and* (7). *There exists an* $\alpha_c := \alpha_c^{\mathrm{SPA}} > 0$ *(depending on* p*) such that if* $M(N) \leq \alpha_c N$, *then there are* $\varepsilon > 0$ *and* $0 < \delta < 1/2$ *such that*

$$P\left(\liminf_{N\to\infty}\left(\bigcap_{\mu=1}^{M(N)} \bigcap_{\sigma\in S_\delta(\xi^\mu)} (H_N(\sigma) \geq H_N(\xi^\mu) + \varepsilon N)\right)\right) = 1,$$

where $S_\delta(\xi^\mu)$ *is the Hamming sphere of radius* δN *centered in* ξ^ν.

3. Preliminaries. This section contains a number of preparatory results for the proofs contained in the final section. The first collects two purely analytical statements.

LEMMA 3.1. (i) *For all* $0 < p < 1$ *and all* $t \in \mathbb{R}$,

$$p\exp(-2(1-p)t) + (1-p)\exp(2pt) \le \cosh((1+|2p-1|)t). \tag{8}$$

(ii) *Let* $x_i \in \mathbb{R}$, $i = 1, \dots, N$ *with* $\sum_{i=1}^N x_i = 0$ *and* $\sum_{i=1}^N x_i^2 \le 1$. *Then*

$$\prod_{i=1}^N (1 - x_i) \ge 1 - \sum_{i=1}^N x_i^2. \tag{9}$$

PROOF. Part (i) is a simple exercise. For (ii), note that $K := \{x \in \mathbb{R}^N: \sum_{i=1}^N x_i = 0 \sum_{i=1}^N x_i^2 \le 1\}$ is a compact subset of $\mathbb{R}^N$ and therefore the function

$$f(x) = \prod_{i=1}^N (1 - x_i) + \sum_{i=1}^N x_i^2$$

assumes its minimum in some $a \in K$. We have to show that $f(a) \ge 1$. If $a = 0$ we are done.

If $a \ne 0$ then, since $a \in K$, there are $j \ne k$ with $a_j \ne a_k$. Define a' by

$$a'_i = \begin{cases} a_i, & i \ne j, k, \\ \dfrac{a_j + a_k}{2}, & i = j, k. \end{cases}$$

Note that $a' \in K$ by convexity. Moreover,

$$\begin{aligned} f(a') &= \Biggl(\prod_{\substack{i=1\\ i\ne j,k}}^N (1 - a_i)\Biggr)\biggl(1 - \frac{a_j + a_k}{2}\biggr)^2 + \sum_{\substack{i=1\\ i\ne j,k}}^N a_i^2 + \frac{1}{2}(a_j + a_k)^2 \\ &= \prod_{i=1}^N (1 - a_i) + \frac{1}{4}\Biggl(\prod_{\substack{i=1\\ i\ne j,k}}^N (1 - a_i)\Biggr)(a_j - a_k)^2 + \sum_{i=1}^N a_i^2 - \frac{1}{2}(a_j - a_k)^2 \\ &= f(a) + \frac{1}{4}\Biggl(\prod_{\substack{i=1\\ i\ne j,k}}^N (1 - a_i)\Biggr)(a_j - a_k)^2 - \frac{1}{2}(a_j - a_k)^2. \end{aligned}$$

Since $f(a') \ge f(a)$,

$$\Biggl(\prod_{\substack{i=1\\ i\ne j,k}}^N (1 - a_i)\Biggr) \ge 2,$$

implying that

$$\begin{aligned} f(a) &\geq 2(1-a_j)(1-a_k) + a_j^2 + a_k^2 \\ &= 1 + (1-(a_j+a_k))^2 \geq 1. \end{aligned}$$ □

Note that (ii) of Lemma 3.1 can and actually will be used for giving a lower bound on the determinant of $Id - A$ where A is some symmetric matrix with trace 0.

Lemmas 3.2 and 3.3 show how to center quadratic forms of one or more homogeneous Markov chains on $\{-1, +1\}$. These lemmas are given here to make the proofs in the following section more transparent.

LEMMA 3.2. *Assume that $(Y_i)_{i\in\mathbb{N}}$ is a homogeneous Markov chain on $\{-1, +1\}$ with transition matrix*

$$R = \begin{pmatrix} q & 1-q \\ 1-q & q \end{pmatrix}$$

starting in equilibrium; that is, $P(Y_1 = +1) = P(Y_1 = -1) = \frac{1}{2}$. Let $A = (a_{i,j})$ be a symmetric $N \times N$ matrix with $a_{i,i} = 0$ for all $i = 1, \ldots, N$ Define

$$\overline{Y}_i = \begin{cases} Y_i - (2q-1)Y_{i-1}, & \text{if } i \geq 2, \\ Y_1, & \text{otherwise}. \end{cases} \tag{10}$$

Then

$$\sum_{i,j=1}^{N} a_{i,j} Y_i Y_j = \sum_{i,j=1}^{N} b_{i,j} \overline{Y}_i \overline{Y}_j + 2 \sum_{1\leq i<j\leq N} (2q-1)^{j-i} a_{i,j}, \tag{11}$$

where

$$b_{i,j} = \sum_{k=0}^{j-i-1} \sum_{l=0}^{N-j} (2q-1)^{k+l} a_{i+k,j+l} \tag{12}$$

for $i < j$, $b_{i,j} = b_{j,i}$ for $i > j$ and finally $b_{i,i} = 0$ for all i.

PROOF. First, note that due to the symmetry $a_{i,j} = a_{j,i}$ and $b_{i,j} = b_{j,i}$ together with $a_{i,i} = b_{i,i} = 0$ for all i we only have to show that

$$\sum_{1\leq i<j\leq N} a_{i,j} Y_i Y_j = \sum_{1\leq i<j\leq N} b_{i,j} \overline{Y}_i \overline{Y}_j + \sum_{1\leq i<j\leq N} (2q-1)^{j-i} a_{i,j}. \tag{13}$$

This will be done by induction. For $N = 2$ we have

$$\begin{aligned} \sum_{1\leq i<j\leq N} a_{i,j} Y_i Y_j &= a_{1,2} Y_1 Y_2 = a_{1,2} \overline{Y}_1 \overline{Y}_2 + (2q-1) a_{1,2} Y_1 Y_1 \\ &= b_{1,2} \overline{Y}_1 \overline{Y}_2 + (2q-1) a_{1,2} \\ &= \sum_{1\leq i<j\leq N} b_{i,j} \overline{Y}_i \overline{Y}_j + (2q-1) a_{1,2}. \end{aligned}$$

Now suppose (13) was already shown for $N-1$. To prove it for N we calculate

$$\begin{aligned}
\sum_{1\le i<j\le N} a_{i,j} Y_i Y_j &= \sum_{1\le i<j\le N-1} a_{i,j} Y_i Y_j + \sum_{1\le i<N} a_{i,N} Y_i Y_N \\
&= \sum_{1\le i<j\le N-1} a_{i,j} Y_i Y_j + \sum_{1\le i<N} a_{i,N} Y_i \overline{Y}_N \\
&\quad + (2q-1) \sum_{1\le i<N-1} a_{i,N} Y_i Y_{N-1} + (2q-1) a_{N-1,N} \qquad (14)\\
&= \sum_{1\le i<j\le N-1} a^{(1)}_{i,j} Y_i Y_j + \sum_{1\le i<N} b_{i,N} \overline{Y}_i \overline{Y}_N \\
&\quad + (2q-1) a_{N-1,N},
\end{aligned}$$

where we have set

$$a^{(1)}_{i,j} = \begin{cases} a_{i,j}, & \text{if } 1 \le i < j \le N-2, \\ a_{i,N-1} + (2q-1) a_{i,n}, & \text{if } 1 \le i < j = N-1. \end{cases}$$

Applying the induction hypotheses to the first summand in (14) leads to

$$\begin{aligned}
&\sum_{1\le i<j\le N-1} a^{(1)}_{i,j} Y_i Y_j \\
&\quad = \sum_{1\le i<j\le N-1} \sum_{k=0}^{j-i-1} \sum_{l=0}^{N-j-1} (2q-1)^{k+l} a^{(1)}_{i+k,\,j+l} \overline{Y}_i \overline{Y}_j \\
&\qquad + \sum_{1\le i<j\le N-1} (2q-1)^{j-i} a^{(1)}_{i,j} \\
&\quad = \sum_{1\le i<j\le N-1} \Bigg(\sum_{k=0}^{j-i-1} \sum_{l=0}^{N-j-2} (2q-1)^{k+l} a_{i+k,\,j+l} \\
&\qquad\qquad + \sum_{k=0}^{j-i-1} (2q-1)^{k+N-1} (a_{i+k,N-1} + (2q-1) a_{i+k,N}) \Bigg) \overline{Y}_i \overline{Y}_j \\
&\qquad + \sum_{1\le i<j\le N-2} (2q-1)^{j-i} a_{i,j} \\
&\qquad + \sum_{1\le i<N-1} (2q-1)^{N-i-1} (a_{i,N-1} + (2q-1) a_{i,N}) \\
&\quad = \sum_{1\le i<j\le N-1} b_{i,j} \overline{Y}_i \overline{Y}_j + \sum_{1\le i<j\le N-2} (2q-1)^{j-i} a_{i,j} \\
&\qquad + \sum_{1\le i<N-1} (2q-1)^{N-i-1} (a_{i,N-1} + (2q-1) a_{i,N})
\end{aligned}$$

such that together with the above calculations,

$$\sum_{1\le i<j\le N} a_{i,j} Y_i Y_j = \sum_{1\le i<j\le N} b_{i,j} \overline{Y}_i \overline{Y}_j + \sum_{1\le i<j\le N} (2q-1)^{j-i} a_{i,j}. \qquad \square$$

LEMMA 3.3. *Let $(Y^\mu)_{\mu\in\mathbb{N}_0}$ and $(Z^\mu)_{\mu\in\mathbb{N}_0}$ be independent and homogeneous Markov chains on $\{-1,+1\}$, both with transition matrix*

$$R = \begin{pmatrix} q & 1-q \\ 1-q & q \end{pmatrix},$$

starting in 1; *that is*, $P(Y^0 = +1) = P(Z^0 = +1) = 1$.

Similarly to (10) *let $\overline{Y}^\mu$ and $\overline{Z}^\mu$ denote the centered versions of Y^μ and Z^μ; that is,*

$$\overline{Y}^\mu = Y_\mu - (2q-1) Y_{\mu-1} \quad \text{if } \mu \ge 1 \tag{15}$$

and

$$\overline{Z}^\mu = Z_\mu - (2q-1) Z_{\mu-1} \quad \text{if } \mu \ge 1. \tag{16}$$

Then

$$\sum_{\mu=1}^{k} Y^\mu Z^\mu = \sum_{\mu_1,\mu_2=1}^{k} a_{\mu_1,\mu_2} \overline{Y}^{\mu_1} \overline{Z}^{\mu_2} + \sum_{\mu=1}^{k} a_{\mu,0} (\overline{Y}^\mu + \overline{Z}^\mu) + \sum_{n=0}^{k-1} (2q-1)^{2n}, \tag{17}$$

where

$$a_{\mu_1,\mu_2} := \sum_{n=0}^{k-\max\{\mu_1,\mu_2\}} (2q-1)^{2n+|\mu_1-\mu_2|} \tag{18}$$

for $\mu_1, \mu_2 \ge 0$, $(\mu_1,\mu_2) \ne (0,0)$. Note that $a_{\mu_1,\mu_2} = a_{\mu_2,\mu_1}$.

PROOF. Observe that

$$Y^\mu Z^\mu = \overline{Y^\mu}\,\overline{Z^\mu} + (2q-1)\big(Y^{\mu-1}\overline{Z^\mu} + \overline{Y^\mu} Z^{\mu-1}\big) + (2q-1)^2 Y^{\mu-1} Z^{\mu-1}.$$

Hence

$$\begin{aligned}\sum_{\mu=1}^{k} Y^\mu Z^\mu &= \sum_{\mu=1}^{k} \sum_{n=0}^{k-\mu} (2q-1)^{2n} \overline{Y^\mu}\,\overline{Z^\mu} \\ &\quad + \sum_{\mu=1}^{k} \sum_{n=0}^{k-\mu} (2q-1)^{2n+1} \big(Y^{\mu-1}\overline{Z^\mu} + \overline{Y^\mu} Z^{\mu-1}\big) + \sum_{n=0}^{k-\nu-1} (2q-1)^{2n}.\end{aligned}$$

Centering also the variables in the second sum on the right-hand side above gives

$$\sum_{\mu=1}^{k} Y^\mu Z^\mu = \sum_{\mu_1,\mu_2=1}^{k} a_{\mu_1,\mu_2} \overline{Y^{\mu_1}}\,\overline{Z^{\mu_2}} + \sum_{\mu=1}^{k} a_{\mu,0}\big(\overline{Y^\mu} + \overline{Z^\mu}\big) + \sum_{n=0}^{k-\nu-1} (2q-1)^{2n}. \qquad \square$$

Finally, we show how to compute the moment generating function of a sum of random variables of the above type.

LEMMA 3.4. *Assume that* $(Y_i)_{i\in\mathbb{N}}$ *is a homogeneous Markov chain on* $\{-1,+1\}$ *with transition matrix*

$$R=\begin{pmatrix} q & 1-q\\ 1-q & q\end{pmatrix}$$

starting in equilibrium; that is, $P(Y_1=+1)=P(Y_1=-1)=\frac{1}{2}$. *Then for* $t\in[0,1]$ *and* $q\geq 1/2$ *and every* $1\leq i\leq k$ *there is a constant* C *such that*

$$\begin{aligned}
&E\Biggl(\exp\Biggl(-t\sum_{\substack{j=1\\ j\neq i}}^{k} Y_iY_j\Biggr)\Biggr)\\
&\quad\leq \exp\biggl(-t\frac{2q-1}{2(1-q)}[2-(2q-1)^{k-i}-(2q-1)^{i-1}]\biggr)\\
&\qquad\times\exp\biggl(t^2\frac{2(k-1)q(1-q)-(2q-1)[2-(2q-1)^{k-i}-(2q-1)^{i-1}]}{4(1-q)^2}\\
&\qquad\qquad + Ckt^3\biggr).
\end{aligned}$$

PROOF. By decomposing the exponential we obtain that

$$\begin{aligned}
&E\Biggl(\exp\Biggl(-t\sum_{\substack{j=1\\ j\neq i}}^{k} Y_iY_j\Biggr)\Biggr)\\
&\quad=\sum_{y_j=-1,+1} P(Y_1=y_1)P(Y_2=y_2|Y_1=y_1)\cdots P(Y_i=y_i|Y_{i-1}=y_{i-1})\\
&\qquad\qquad\times P(Y_{i+1}=y_{i+1}|Y_i=y_i)\cdots P(Y_k=y_k|Y_{k-1}=y_{k-1})\\
&\qquad\qquad\times\exp\Biggl(-t\sum_{\substack{j=1\\ j\neq i}}^{k} y_iy_j\Biggr)\\
&\quad=\frac{1}{2}\sum_{\substack{y_j=-1,+1,\\ j\neq i}} R(y_1,y_2)\exp(-ty_1)\cdots R(y_{i-1},1)\exp(-ty_{i-1})\\
&\qquad\qquad\times R(1,y_{i+1})\exp(-ty_{i+1})\cdots R(y_{k-1},y_k)\exp(-ty_k)\\
&\qquad+\frac{1}{2}\sum_{\substack{y_j=-1,+1,\\ j\neq i}} R(y_1,y_2)\exp(ty_1)\cdots R(y_{i-1},-1)\exp(ty_{i-1})\\
&\qquad\qquad\times R(-1,y_{i+1})\exp(ty_{i+1})\cdots R(y_{k-1},y_k)\exp(ty_k).
\end{aligned}\tag{19}$$

Due to the symmetry of R and the Y_j, the two sums on the right-hand side in (19) agree such that

$$E\left(\exp\left(-t\sum_{\substack{j=1,\\ j\neq i}}^{k} Y_i Y_j\right)\right)$$
$$= \sum_{\substack{y_j=-1,+1,\\ j\neq i}} R(y_1, y_2)\exp(-ty_1)\cdots R(y_{i-1}, 1)\exp(-ty_{i-1})$$
$$\times R(1, y_{i+1})\exp(-ty_{i+1})\cdots R(y_{k-1}, y_k)\exp(-ty_k)$$
$$= \sum_{\substack{y_1=-1,1,\\ y_k=-1,1}} \Pi_L^{i-1}(y_1, 1)\Pi_R^{k-i}(1, y_k),$$

where

$$\Pi_L := \begin{pmatrix} qe^{-t} & (1-q)e^{-t} \\ (1-q)e^{t} & qe^{t} \end{pmatrix}$$

and

$$\Pi_R := \begin{pmatrix} qe^{-t} & (1-q)e^{t} \\ (1-q)e^{-t} & qe^{t} \end{pmatrix}.$$

Observe that $(\Pi_L)^t = \Pi_R$, implying that Π_L and Π_R have the same eigenvalues, which are

$$\lambda_1 = q\cosh(t) + \sqrt{1 - 2q + q^2\cosh^2(t)} \tag{20}$$

and

$$\lambda_2 = q\cosh(t) - \sqrt{1 - 2q + q^2\cosh^2(t)}.$$

Now

$$\sum_{y_1=-1,1,\ y_N=-1,1} \Pi_L^{i-1}(y_1, 1)\Pi_R^{k-i}(1, y_k) = \left(\begin{pmatrix}1\\1\end{pmatrix}\Pi_L^{i-1}\begin{pmatrix}1\\0\end{pmatrix}\right)\left(\begin{pmatrix}1\\0\end{pmatrix}\Pi_R^{k-i}\begin{pmatrix}1\\1\end{pmatrix}\right).$$

By bringing Π_L and Π_R into diagonal form we see that

$$\begin{pmatrix}1\\1\end{pmatrix}\Pi_L^{i-1}\begin{pmatrix}1\\0\end{pmatrix} = \frac{1}{ad-bc}\left(a(d-b)\lambda_1^{i-1} + b(a-c)\lambda_2^{i-1}\right)$$

as well as

$$\begin{pmatrix}1\\0\end{pmatrix}\Pi_R^{k-i}\begin{pmatrix}1\\1\end{pmatrix} = \frac{1}{ad-bc}\left(a(d-b)\lambda_1^{k-i} + b(a-c)\lambda_2^{k-i}\right)$$

with

$$a = \sqrt{1 - 2q + q^2\cosh^2(t)} - q\sinh(t), \qquad b = (1-q)e^t,$$

$$c = (1-q)e^{-t}, \qquad d = q\sinh(t) - \sqrt{1 - 2q + q^2\cosh^2(t)}.$$

Put

$$f(t) = \frac{1}{ad - bc}\big(a(d-b)\lambda_1^n + b(a-c)\lambda_2^n\big)$$

for some n. Some calculations give that

$$\begin{aligned} f(0) &= 1,\\ f'(0) &= -\frac{2q-1}{2(1-q)}[1-(2q-1)^n],\\ f''(0) &= \frac{2(n-1)q(1-q) - (2q-1)[1-(2q-1)^n]}{(1-q)^2},\\ f'''(0) &= \mathscr{O}(n). \end{aligned}$$

Thus by Taylor expansion for $t \in [0,1]$,

$$\begin{aligned} f(t) &= 1 - t\frac{2q-1}{2(1-q)}[1-(2q-1)^n]\\ &\quad + t^2\frac{2(n-1)q(1-q) - (2q-1)[1-(2q-1)^n]}{4(1-q)^2} + \mathscr{O}(nt^3)\\ &\le \exp\bigg(-t\frac{2q-1}{2(1-q)}[1-(2q-1)^n]\\ &\qquad + t^2\frac{2(n-1)q(1-q) - (2q-1)[1-(2q-1)^n]}{4(1-q)^2} + \mathscr{O}(nt^3)\bigg). \end{aligned}$$

This implies the assertion of the lemma. □

4. Proofs. The proofs of the first two theorems are quite similar, so we will give the first in detail and only comment on the changes for the second.

PROOF OF THEOREM 2.1. Fix $1 \le \nu \le M(N)$. Following the definition of the dynamics T and the definition of stability introduced in Section 2 the pattern ξ^ν is stable if and only if

$$\xi_i^\nu = \operatorname{sgn}\bigg(\sum_{\substack{j=1\\ j\neq i}}^{N} \xi_j^\nu J_{ij}\bigg) = \operatorname{sgn}\bigg(\sum_{j=1}^{N}\sum_{\mu=1}^{M(N)} \xi_j^\nu \xi_i^\mu \xi_j^\mu\bigg)$$

for all $i = 1, \dots, N$, or, in other words, if

$$\sum_{\substack{j=1\\ j\neq i}}^{N}\sum_{\mu=1}^{M(N)} \xi_i^\nu \xi_j^\nu \xi_i^\mu \xi_j^\mu \ge 0$$

for all $i = 1, \dots, N$ (where we use the convention $\operatorname{sgn}(0) = 1$, which does not really influence our calculations since the probability that the above sum equals zero vanishes for large N). Hence by bounding the probability of a union of events by the sum of the corresponding probabilities and the exponential Chebyshev–Markov inequality we get for all $t \geq 0$,

$$
\begin{aligned}
P(\xi^\nu \text{ is not stable}) &\leq \sum_{i=1}^N P\left(\sum_{\substack{j=1\\ j\neq i}}^N \sum_{\substack{\mu=1\\ \mu\neq\nu}}^{M(N)} \xi_i^\nu \xi_j^\nu \xi_i^\mu \xi_j^\mu \leq -N \right) \\
&\leq \sum_{i=1}^N e^{-tN} E \exp\left(-t \sum_{\substack{j=1\\ j\neq i}}^N \sum_{\substack{\mu=1\\ \mu\neq\nu}}^{M(N)} \xi_i^\nu \xi_j^\nu \xi_i^\mu \xi_j^\mu \right) \qquad (21)\\
&= N e^{-tN} \left(E \exp\left(-t \sum_{\substack{\mu=1\\ \mu\neq\nu}}^{M(N)} \xi_1^\nu \xi_2^\nu \xi_1^\mu \xi_2^\mu \right) \right)^N .
\end{aligned}
$$

Now putting $Y_\mu := \xi_1^\mu \xi_2^\mu$, note that Y_μ is a Markov chain with transition matrix

$$
R = \begin{pmatrix} p^2 + (1-p)^2 & 2p(1-p) \\ 2p(1-p) & p^2 + (1-p)^2 \end{pmatrix} =: \begin{pmatrix} q & 1-q \\ 1-q & q \end{pmatrix},
$$

such that we can calculate

$$
E \exp\left(-t \sum_{\substack{\mu=1\\ \mu\neq\nu}}^{M(N)} \xi_1^\nu \xi_2^\nu \xi_1^\mu \xi_2^\mu \right) = E \exp\left(-t \sum_{\substack{\mu=1\\ \mu\neq\nu}}^{M(N)} Y_\nu Y_\mu \right)
$$

with the help of Lemma 3.4 (and $k = M$). Thus there is a constant C such that for all $t \leq 1$,

$$
\begin{aligned}
&E \exp\left(-t \sum_{\substack{\mu=1\\ \mu\neq\nu}}^{M(N)} Y_\nu Y_\mu \right) \\
&\quad \leq \exp\left(-t \frac{2q-1}{2(1-q)} [2 - (2q-1)^{M-\nu} - (2q-1)^{\nu-1}] \right) \\
&\qquad \times \exp\left(t^2 \frac{2(M-1)q(1-q) - (2q-1)[2-(2q-1)^{M-\nu} - (2q-1)^{M-1}]}{4(1-q)^2} \right. \\
&\qquad\qquad \left. + CMt^3 \right).
\end{aligned}
$$

Now as for each $\varepsilon > 0$ we have that $[2-(2q-1)^{M-\nu}-(2q-1)^{\nu-1}] \geq 1-\varepsilon$ if N and therefore M is large enough. Thus,

$$\exp\biggl(-t\frac{2q-1}{2(1-q)}[2-(2q-1)^{M-\nu}-(2q-1)^{\nu-1}]\biggr)$$
$$\times \exp\biggl(t^2\frac{2(M-1)q(1-q)-(2q-1)[2-(2q-1)^{M-\nu}-(2q-1)^{M-1}]}{4(1-q)^2} + CMt^3\biggr)$$
$$\leq \exp\biggl(-t(1-\varepsilon)\frac{2q-1}{2(1-q)}+t^2\frac{(M-1)q}{2(1-q)}+CMt^3\biggr).$$

Hence we arrive at

$$P(\exists\ \nu\colon \xi^\nu \text{ is not stable}) \leq MN\exp\biggl(-tN\biggl(1+(1-\varepsilon)\frac{2q-1}{2(1-q)}\biggr)+t^2MN\frac{q}{2(1-q)}+CMNt^3\biggr),$$

where the last factor on the right-hand side does not contribute, if t goes to zero fast enough. Choosing $t = 1-\varepsilon(2q-1)/2qM$ gives

$$P(\exists\ \nu\colon \xi^\nu \text{ is not stable}) \leq MN\exp\biggl(-\frac{(1-\varepsilon(2q-1))^2}{8q(1-q)}\frac{N}{M}+C\frac{N}{M^2}\biggr).$$

So if $M = N/\gamma\log N$ the last factor on the right-hand side can be bounded by $\exp(\text{const.}((\log N)^2/N))$, which is converging to one. Therefore

$$P(\exists\ \nu\colon \xi^\nu \text{ is not stable}) \leq \frac{N^2}{\log N}N^{-(\gamma(1-\varepsilon(2q-1))^2)/8q(1-q)}(1+o(1))$$

and the choice $\gamma > 24q(1-q)/(1-\varepsilon(2q-1))^2 = 48(p^2+(1-p)^2)p(1-p)/(1-\varepsilon(2q-1))^2$ leads to the converging series $\sum(1/N^\kappa\log N)$ for a $\kappa > 1$. Since this is true for all $\varepsilon > 0$ this proves part (i) of the theorem by the Borel–Cantelli lemma.

The choice of $\gamma > 16q(1-q) = 32(p^2+(1-p)^2)p(1-p)$ yields

$$P(\exists\ \nu\colon \xi^\nu \text{ is not stable}) \to 0$$

and therefore part (ii) of the theorem.

Part (iii) of the theorem follows by observing that for any fixed ν,

$$P(\xi^\nu \text{ is not stable}) \leq N\exp\biggl(-tN\biggl(1+(1-\varepsilon)\frac{2q-1}{2(1-q)}\biggr)+t^2\frac{MNq}{2(1-q)}+CMt^3\biggr)$$

and then continuing just as above. □

PROOF OF THEOREM 2.2. Again fix $1 \le \nu \le M(N)$. Just as in the proof of Theorem 2.1 we use the exponential Chebyshev–Markov inequality together with the independence of the patterns to obtain, for all $t \ge 0$,

$$P(\xi^\nu \text{ is not stable}) \le \sum_{i=1}^{N} \exp(-tN)\left(E\exp\left(-t\sum_{\substack{j=1\\ j\neq i}}^{N} \xi_i^1\xi_j^1\xi_i^2\xi_j^2\right)\right)^{M-1}. \tag{22}$$

Putting $Y_i := \xi_i^1\xi_i^2$ this time and again,

$$R = \begin{pmatrix} p^2+(1-p)^2 & 2p(1-p) \\ 2p(1-p) & p^2+(1-p)^2 \end{pmatrix} =: \begin{pmatrix} q & 1-q \\ 1-q & q \end{pmatrix},$$

we can calculate the expectation in (22) with the help of Lemma 3.4 (with $k = N$), such that there exists a constant C such that for all $t \in [0,1]$,

$$\begin{aligned} E\exp\left(-t\sum_{\substack{j=1\\ j\neq i}}^{N} \xi_i^1\xi_i^2\xi_j^1\xi_j^2\right) &= E\exp\left(-t\sum_{\substack{j=1\\ j\neq i}}^{N} Y_iY_j\right) \\ &\le \exp\left(-t\frac{2q-1}{2(1-q)}[2-(2q-1)^{N-i}-(2q-1)^{i-1}]\right) \qquad (23)\\ &\quad\times \exp\left(t^2\frac{2(N-1)q(1-q)-(2q-1)[2-(2q-1)^{N-i}-(2q-1)^{N-1}]}{4(1-q)^2} + CNt^3\right) \\ &\le \exp\left(-t\frac{2q-1}{2(1-q)} + t^2\frac{(N-1)q}{2(1-q)} + CNt^3\right). \end{aligned}$$

Since the right-hand side of (23) is independent of the choice of i, we obtain

$$\begin{aligned} P(\exists\ \nu\colon \xi^\nu \text{ is not stable}) &\le MP(\xi^1 \text{ is not stable}) \\ &\le NM\exp(-tN)\exp\left(-tM\frac{2q-1}{2(1-q)} + t^2NM\frac{q}{2(1-q)} + CNt^3\right). \end{aligned} \tag{24}$$

Now the basic difference between the proofs of Theorems 2.1 and 2.2 (and the reason for the different qualitative behavior in q) occurs because the ansatz $M = N/\gamma\log N$ implies that $M = o(N)$ and therefore the term $\exp(-tM(2q-1/2(1-q)))$ becomes negligible in the limit for large N. More precisely, given $\varepsilon > 0$, we have that for all N large enough,

$$P(\exists\ \nu\colon \xi^\nu \text{ is not stable}) \le NM\exp\left(-tN(1+\varepsilon) + t^2NM\frac{q}{2(1-q)} + CNt^3\right).$$

Again the t^3-term for our choice of t does not contribute; namely, choosing $t = (1+\varepsilon)((1-q)/qM)$ gives

$$P\big(\exists\, \nu\colon \xi^\nu \text{ is not stable}\big) \le MN \exp\left(-\frac{1-q}{2q}(1+\varepsilon)^2 \frac{N}{M}\right)(1+o(1))$$

(provided that $M \le N$). So the choice of $M = N/\gamma \log N$ and $\gamma > 6q/1-q = 3(p^2+(1-p)^2)/p(1-p)$ again leads to a converging series and thus proves part (i) of the theorem by the Borel–Cantelli lemma.

The choice of $\gamma \ge 4q/1-q = 2(p^2+(1-p)^2)/p(1-p)$ yields

$$P\big(\exists\, \nu\colon \xi^\nu \text{ is not stable}\big) \to 0$$

and therefore part (ii) of the theorem. Part (iii) of the theorem follows similarly. □

REMARK 4.1. The attentive reader may have noticed that in the above proofs we only made use of the Markov property of our patterns in order to estimate the moment generating function, or more precisely to obtain a bound of the form

$$E\left(\exp\left(-t\sum_{\substack{j=1\\ j\neq i}}^{N}\sum_{\substack{\mu=1\\ \mu\neq\nu}}^{M(N)} \xi_i^\nu \xi_j^\nu \xi_i^\mu \xi_j^\mu\right)\right) \le \exp(\text{const.}MNt^2)$$

for t small enough, that is, that the moment generating function "looks Gaussian" in a neighborhood of zero. This is, of course, not only true for Markov chains (e.g., a similar bound might be shown for appropriate higher-dimensional random fields) and indeed it is an easy exercise to prove a corresponding theorem for sequences of patterns admitting a bound of the above form. However, the reason that we stick to Markov chains here is that we will need this assumption to prove Theorems 2.3 and 2.4 anyway. Moreover, in the case of Markov chains one can easily detect the qualitative behavior of the bound on the storage capacity.

We now give the proofs of the theorems treating the case of the second notion of storage capacity.

PROOF OF THEOREM 2.3. Again the proof uses large deviation estimates. These can be carried out by a centering of the variables (which already has been prepared in Lemma 3.3) together with the idea of replacing the ξ_i^μ by appropriate Gaussian random variables, which is a standard idea in the framework of the Hopfield model with independent patterns (see, e.g., [21] or [4]), but in our case technically quite involved. Finally, the resulting generating function of a quadratic form in Gaussian random variables has to be calculated.

Let us set

$$h_N(\sigma,\delta) := \inf_{\sigma' \in S_\delta(\sigma)} H_N(\sigma').$$

Then

$$P\Bigg(\Bigg\{\bigcap_{\nu=1}^{M(N)}(h_N(\xi^\nu,\delta)\geq H_N(\xi^\nu)+\varepsilon N)\Bigg\}^c\Bigg)$$

$$=P\Bigg(\bigcup_{\nu=1}^{M(N)}\bigcup_{J:|J|=\delta N}H_N(\xi^\nu_J)-H_N(\xi^\nu)\leq\varepsilon N\Bigg)$$

$$\leq\sum_{J:|J|=\delta N}\sum_{\nu=1}^{M(N)}P\big(H_N(\xi^\nu_J)-H_N(\xi^\nu)\leq\varepsilon N\big),$$

where ξ^ν_J denotes a configuration differing from ξ^ν exactly in the coordinates J and for convenience we have chosen δN to be an integer.

Estimating the probability on the right-hand side of the above inequality, again with the help of the exponential Chebyshev–Markov inequality, we obtain for a fixed $1\leq\nu\leq M(N)$ and all $t\geq 0$,

$$P\big(H_N(\xi^\nu_J)-H_N(\xi^\nu)\leq\varepsilon N\big)$$

$$=P\Bigg(-\frac{1}{2N}\sum_{\mu=1}^{M(N)}\sum_{i,\,j=1}^{N}\big(\xi^\nu_{J,\,i}\xi^\nu_{J,\,j}-\xi^\nu_i\xi^\nu_j)\big)\xi^\mu_i\xi^\mu_j\leq\varepsilon N\Bigg)$$

$$=P\Bigg(\frac{1}{N}\sum_{\mu=1}^{M(N)}\sum_{\substack{i\in J,\,j\notin J\\ i\notin J,\,j\in J}}\xi^\nu_i\xi^\nu_j\xi^\mu_i\xi^\mu_j\leq\varepsilon N\Bigg)$$

$$=P\Bigg(\frac{2}{N}\sum_{\mu=1}^{M(N)}\sum_{i\in J,\,j\notin J}\xi^\nu_i\xi^\nu_j\xi^\mu_i\xi^\mu_j\leq\varepsilon N\Bigg)$$

$$=P\Bigg(\frac{1}{N}\sum_{\mu\neq\nu}\sum_{i\in J,\,j\notin J}\xi^\nu_i\xi^\nu_j\xi^\mu_i\xi^\mu_j\leq(\varepsilon/2-\delta(1-\delta))N\Bigg)$$

$$\leq\exp(-t\varepsilon' N)E\exp\Bigg(-\frac{t}{N}\sum_{\mu\neq\nu}\sum_{i\in J,\,j\notin J}\xi^\nu_i\xi^\nu_j\xi^\mu_i\xi^\mu_j\Bigg),$$

where we have set $\varepsilon'=-\varepsilon/2+\delta(1-\delta)$ and we are again left with estimating the expectation of an exponential.

To this end we assume that $\xi^\nu_i=1$ for all $i=1,\dots,N$ (this can be done without loss of generality since the initial situation is completely symmetric) and apply Lemma 3.3. Thus,

$$\sum_{i\in J,\,j\notin J}\sum_{\mu=\nu+1}^{M}\xi^\mu_i\xi^\mu_j=\sum_{i\in J,\,j\notin J}\Bigg(\sum_{\mu_1,\,\mu_2=\nu+1}^{M}a_{\mu_1,\,\mu_2}\overline{\xi^{\mu_1}_i}\,\overline{\xi^{\mu_2}_j}$$

$$+\sum_{\mu=\nu+1}^{M}a_{\mu,\,\nu}(\overline{\xi^\mu_i}+\overline{\xi^\mu_j})+\sum_{n=0}^{M-\nu-1}(2p-1)^{2n}\Bigg),$$

where $\overline{\xi_i^\mu} = \xi_i^\mu + (2p-1)\xi_i^{\mu-1}$ and

$$a_{\mu_1,\mu_2} := \sum_{n=0}^{M-\max\{\mu_1,\mu_2\}} (2p-1)^{2n+|\mu_1-\mu_2|} \tag{25}$$

for $\mu_1, \mu_2 \geq \nu$, $(\mu_1, \mu_2) \neq (\nu, \nu)$. Note that $a_{\mu_1,\mu_2} = a_{\mu_2,\mu_1}$.

For the summands with an index $\mu < \nu$ in $\sum_{i\in J,\, j\notin J}\sum_{\mu\neq\nu} \xi_i^\mu \xi_j^\mu$ we observe that the Markov chains $(\xi_i^\mu)_{\mu<\nu}$ and $(\xi_i^\mu)_{\mu\geq\nu+1}$, $i = 1, \dots, N$ conditioned on ξ_i^ν are independent. Applying the same transformation as above to the Markov chains $(\xi_i^\mu)_{\mu<\nu}$ $(i = 1, \dots, N)$ yields

$$\begin{aligned}
&E\biggl(\exp\biggl(-\frac{t}{N}\sum_{i\in J,\, j\notin J}\sum_{\mu\neq\nu}\xi_i^\mu\xi_j^\mu\biggr)\biggr)\\
&\quad= \exp\biggl(-\frac{t}{N}\sum_{i\in J,\, j\notin J}\biggl(\sum_{n=0}^{M-\nu-1}(2p-1)^{2n} + \sum_{n=0}^{\nu-1}(2p-1)^{2n}\biggr)\biggr)\\
&\qquad\times E\exp\biggl(-\frac{t}{N}\sum_{i\in J,\, j\notin J}\biggl(\sum_{\mu=\nu+1}^{M} a_{\mu,\nu}(\overline{\xi_i^\mu}+\overline{\xi_j^\mu}) + \sum_{\mu=1}^{\nu-1}\tilde a_{\mu,\nu}(\overline{\xi_i^\mu}+\overline{\xi_j^\mu})\\
&\qquad\qquad + \sum_{\mu_1,\mu_2=\nu+1}^{M} a_{\mu_1,\mu_2}\overline{\xi_i^{\mu_1}}\,\overline{\xi_j^{\mu_2}} + \sum_{\mu_1,\mu_2=1}^{\nu-1}\tilde a_{\mu_1,\mu_2}\overline{\xi_i^{\mu_1}}\,\overline{\xi_j^{\mu_2}}\biggr)\biggr),
\end{aligned}$$

where

$$\tilde a_{\mu_1,\mu_2} := \sum_{n=0}^{\nu-1-\min\{\nu-\mu_1,\,\nu-\mu_2\}} (2p-1)^{2n+|\mu_1-\mu_2|}. \tag{26}$$

Using the independence of the initial part and the tail part of the Markov chains mentioned above together with Hölder's inequality to split up the moment generating function of the linear part from the moment generating function of the genuine quadratic form, we obtain for all $\lambda > 1$,

$$\begin{aligned}
&E\exp\biggl(-\frac{t}{N}\sum_{\mu\neq\nu}\sum_{i\in J,\, j\notin J}\xi_i^\nu\xi_j^\nu\xi_i^\mu\xi_j^\mu\biggr)\\
&\quad\leq \exp\biggl(-tN\delta(1-\delta)\biggl(\sum_{n=0}^{M-\nu-1}(2p-1)^{2n} + \sum_{n=0}^{\nu-1}(2p-1)^{2n}\biggr)\biggr)\\
&\qquad\times\biggl(E\exp\biggl(-\frac{t}{N}\frac{\lambda}{\lambda-1}\sum_{\mu=\nu+1}^{M}a_{\mu,\nu}\sum_{i\in J,\, j\notin J}(\overline{\xi_i^\mu}+\overline{\xi_j^\mu})\biggr)\biggr)^{(\lambda-1)/\lambda}\\
&\qquad\times\biggl(E\exp\biggl(-\frac{t}{N}\frac{\lambda}{\lambda-1}\sum_{\mu=1}^{\nu-1}\tilde a_{\mu,\nu}\sum_{i\in J,\, j\notin J}(\overline{\xi_i^\mu}+\overline{\xi_j^\mu})\biggr)\biggr)^{(\lambda-1)/\lambda}\\
&\qquad\times\biggl(E\exp\biggl(-\frac{t}{N}\lambda\sum_{i\in J,\, j\notin J}\sum_{\mu_1,\mu_2=\nu+1}^{M}a_{\mu_1,\mu_2}\overline{\xi_i^{\mu_1}}\,\overline{\xi_j^{\mu_2}}\biggr)\biggr)^{1/\lambda}\\
&\qquad\times\biggl(E\exp\biggl(-\frac{t}{N}\lambda\sum_{i\in J,\, j\notin J}\sum_{\mu_1,\mu_2=1}^{\nu-1}\tilde a_{\mu_1,\mu_2}\overline{\xi_i^{\mu_1}}\,\overline{\xi_j^{\mu_2}}\biggr)\biggr)^{1/\lambda}.
\end{aligned} \tag{27}$$

Now we estimate the factors on the right-hand side of (27). First of all note that for M large enough (which is possible, since M is growing with N),

$$\sum_{n=0}^{M-\nu-1}(2p-1)^{2n}+\sum_{n=0}^{\nu-1}(2p-1)^{2n})\geq\frac{1}{C'(1-(2p-1)^2)}$$

for any $C'>1$.

To treat the other terms let us agree on the following notation: with $E_I^{I'}$ (where $I\subset\{1,\ldots,N\}$ and $I'\subset\{1,\ldots,M\}$) we denote the integration with respect to those random variables ξ_i^μ with $i\in I$ and $\mu\in I'$. Especially, if we drop the upper or lower indices we will usually mean the expectation with respect to all the random variables occuring in the argument of the integral. By the independence of the coordinate processes and the identical distribution of the ξ_i^μ, we obtain for the moment generating function of the linear part,

$$\begin{aligned}E\biggl(\exp\biggl(-\frac{t}{N}\frac{\lambda}{\lambda-1}\sum_{\mu=\nu+1}^{M}a_{\mu,\nu}\sum_{i\in J,\, j\notin J}(\overline{\xi_i^\mu}+\overline{\xi_j^\mu})\biggr)\biggr)\\ =\biggl[E\biggl(\exp\biggl(-\frac{t}{N}\frac{\lambda}{\lambda-1}\sum_{\mu=\nu+1}^{M}a_{\mu,\nu}\overline{\xi_1^\mu}\biggr)\biggr)\biggr]^{\delta(1-\delta)N^2}.\end{aligned}$$

Moreover,

$$\begin{aligned}&E\biggl(\exp\biggl(-t\frac{\lambda}{\lambda-1}\sum_{\mu=\nu+1}^{M}a_{\mu,\nu}\overline{\xi_1^\mu}\biggr)\biggr)\\
&\quad=E^{\nu<\mu\leq M-1}\biggl(\exp\biggl(-t\frac{\lambda}{\lambda-1}\sum_{\mu=\nu+1}^{M-1}a_{\mu,\nu}\overline{\xi_1^\mu}\biggr)\biggr)E^{M}\biggl(\exp\biggl(-t\frac{\lambda}{\lambda-1}a_{M,\nu}\overline{\xi_1^M}\biggr)\biggr)\\
&\quad=E^{\nu<\mu\leq M-1}\biggl(\exp\biggl(-t\frac{\lambda}{\lambda-1}\sum_{\mu=\nu+1}^{M-1}a_{\mu,\nu}\overline{\xi_1^\mu}\biggr)\biggr)\\
&\qquad\times\biggl(p\exp\biggl(-2t\frac{\lambda}{\lambda-1}a_{M,\nu}(1-p)\xi_1^{M-1}\biggr)\\
&\qquad\qquad+(1-p)\exp\biggl(2t\frac{\lambda}{\lambda-1}a_{M,\nu}p\xi_1^{M-1}\biggr)\biggr)\\
&\quad\leq E^{\nu<\mu\leq M-1}\biggl(\exp\biggl(-t\frac{\lambda}{\lambda-1}\sum_{\mu=\nu+1}^{M-1}a_{\mu,\nu}\overline{\xi_1^\mu}\biggr)\biggr)\\
&\qquad\times\cosh\biggl(t\frac{\lambda}{\lambda-1}a_{M,\nu}(1+|2p-1|)\xi_1^{M-1}\biggr)\\
&\quad\leq E^{\nu<\mu\leq M-1}\biggl(\exp\biggl(-t\frac{\lambda}{\lambda-1}\sum_{\mu=\nu+1}^{M-1}a_{\mu,\nu}\overline{\xi_1^\mu}\biggr)\biggr)\\
&\qquad\times\exp\biggl(\frac{1}{2}t^2\biggl(\frac{\lambda}{\lambda-1}\biggr)^2a_{M,\nu}^2(1+|2p-1|)^2\biggr),\end{aligned}$$

where we have used $|\xi_1^{M-1}| = 1$, Lemma 3.1, part (i), and finally

$$\cosh(x) \leq \exp(x^2/2).$$

Iterating these estimates gives

$$\begin{aligned}
&E\left(\exp\left(-t\frac{\lambda}{\lambda-1}\sum_{\mu=\nu+1}^{M} a_{\mu,\nu}\overline{\xi_1^{\mu}}\right)\right)\\
&\quad\leq \exp\left(\frac{1}{2}t^2\left(\frac{\lambda}{\lambda-1}\right)^2(1+|2p-1|)^2\sum_{\mu=\nu+1}^{M} a_{\mu,\nu}^2\right)\\
&\quad\leq \exp\left(\frac{1}{2}t^2\left(\frac{\lambda}{\lambda-1}\right)^2(1+|2p-1|)^2\frac{1}{(1-(2p-1)^2)^3}\right).
\end{aligned}$$

So altogether (using this estimate with t/N instead of t) we arrive at

$$\begin{aligned}
&E\left(\exp\left(-\frac{t}{N}\frac{\lambda}{\lambda-1}\sum_{\mu=\nu+1}^{M} a_{\mu,\nu}\sum_{i\in J,\, j\notin J}(\overline{\xi_i^{\mu}}+\overline{\xi_j^{\mu}})\right)\right)\\
&\quad\leq \exp\left(\frac{1}{2}t^2\delta(1-\delta)\left(\frac{\lambda}{\lambda-1}\right)^2(1+|2p-1|)^2\frac{1}{(1-(2p-1)^2)^3}\right)
\end{aligned}$$

Thus, applying the same techniques to the second linear term on the right-hand side of (27), we obtain

$$\begin{aligned}
&\left(E\left(\exp\left(-\frac{t}{N}\frac{\lambda}{\lambda-1}\sum_{\mu=\nu+1}^{M} a_{\mu,\nu}\sum_{i\in J,\, j\notin J}(\overline{\xi_i^{\mu}}+\overline{\xi_j^{\mu}})\right)\right)\right)^{(\lambda-1)/\lambda}\\
&\quad\times\left(E\left(\exp\left(-\frac{t}{N}\frac{\lambda}{\lambda-1}\sum_{\mu=1}^{\nu-1} \tilde{a}_{\mu,\nu}\sum_{i\in J,\, j\notin J}(\overline{\xi_i^{\mu}}+\overline{\xi_j^{\mu}})\right)\right)\right)^{(\lambda-1)/\lambda}\\
&\quad\leq \exp\left(t^2\delta(1-\delta)\left(\frac{\lambda}{\lambda-1}\right)(1+|2p-1|)^2\frac{1}{(1-(2p-1)^2)^3}\right).
\end{aligned}$$

We will see that due to our final choice of t, this factor will have a negligible contribution to the final estimate.

The moment generating function of the quadratic form is treated similarly, using the independence of the ξ_i^{μ} for different i to replace them by Gaussian random variables:

$$\begin{aligned}
&E\left(\exp\left(-\frac{t}{N}\lambda\sum_{i\in J,\, j\notin J}\sum_{\mu_1,\mu_2=\nu+1}^{M} a_{\mu_1,\mu_2}\overline{\xi_i^{\mu_1}}\,\overline{\xi_j^{\mu_2}}\right)\right)\\
&\quad= E^{\nu<\mu_1,\mu_2\leq M-1}E_{J^c}^{M}\left(\exp\left(-\frac{t}{N}\lambda\sum_{i\in J,\, j\notin J}\sum_{\mu_2=\nu+1}^{M}\sum_{\mu_1=\nu+1}^{M-1} a_{\mu_1,\mu_2}\overline{\xi_i^{\mu_1}}\,\overline{\xi_j^{\mu_2}}\right)\right.\\
&\qquad\qquad\left.\times E_J^M\left(\exp\left(-\frac{t}{N}\lambda\sum_{i\in J,\, j\notin J}\overline{\xi_i^{M}}\sum_{\mu_2=\nu+1}^{M} a_{M,\mu_2}\overline{\xi_j^{\mu_2}}\right)\right)\right)
\end{aligned}$$

$$= E^{\nu<\mu_1,\mu_2\le M-1} E^M_{J^c}\Bigg(\exp\Bigg(-\frac{t}{N}\lambda \sum_{i\in J,\, j\notin J}\sum_{\mu_2=\nu+1}^{M}\sum_{\mu_1=\nu+1}^{M-1} a_{\mu_1,\mu_2}\overline{\xi_i^{\mu_1}}\,\overline{\xi_j^{\mu_2}}\Bigg)$$

$$\times \prod_{i\in J} E^M_{\{i\}}\Bigg(\exp\Bigg(-\frac{t}{N}\lambda\overline{\xi_i^M}\sum_{j\notin J}\sum_{\mu_2=\nu+1}^{M} a_{M,\mu_2}\overline{\xi_j^{\mu_2}}\Bigg)\Bigg)\Bigg)$$

$$\le E^{\nu<\mu_1,\mu_2\le M-1} E^M_{J^c}\Bigg(\exp\Bigg(-\frac{t}{N}\lambda \sum_{i\in J,\, j\notin J}\sum_{\mu_2=\nu+1}^{M}\sum_{\mu_1=\nu+1}^{M-1} a_{\mu_1,\mu_2}\overline{\xi_i^{\mu_1}}\,\overline{\xi_j^{\mu_2}}\Bigg)\Bigg)$$

$$\times \prod_{i\in J}\exp\Bigg(\frac{1}{2}\frac{t^2}{N^2}\lambda^2(1+|2p-1|)^2\Bigg(\sum_{j\notin J}\sum_{\mu_2=\nu+1}^{M} a_{M,\mu_2}\overline{\xi_j^{\mu_2}}\Bigg)^2\Bigg)$$

$$= E^{\nu<\mu_1,\mu_2\le M-1} E^M_{J^c}\Bigg(\exp\Bigg(-\frac{t}{N}\lambda \sum_{i\in J,\, j\notin J}\sum_{\mu_2=\nu+1}^{M}\sum_{\mu_1=\nu+1}^{M-1} a_{\mu_1,\mu_2}\overline{\xi_i^{\mu_1}}\,\overline{\xi_j^{\mu_2}}\Bigg)\Bigg)$$

$$\times \prod_{i\in J} E_{z_i^M}\exp\Bigg(z_i^M\frac{t}{N}\lambda(1+|2p-1|)\sum_{i\in J,\, j\notin J}\sum_{\mu_2=\nu+1}^{M} a_{M,\mu_2}\overline{\xi_j^{\mu_2}}\Bigg)$$

$$= E^{\nu<\mu_1,\mu_2\le M-1} E^M_{J^c}\Bigg(\exp\Bigg(-\frac{t}{N}\lambda \sum_{i\in J,\, j\notin J}\sum_{\mu_2=\nu+1}^{M}\sum_{\mu_1=\nu+1}^{M-1} a_{\mu_1,\mu_2}\overline{\xi_i^{\mu_1}}\,\overline{\xi_j^{\mu_2}}\Bigg)\Bigg)$$

$$\times E_{z_J^M}\exp\Bigg(\frac{t}{N}\lambda(1+|2p-1|)\sum_{i\in J,\, j\notin J}\sum_{\mu_2=\nu+1}^{M} a_{M,\mu_2} z_i^M\overline{\xi_j^{\mu_2}}\Bigg),$$

where z_i^M are Gaussian random variables with expectation 0 and identity covariance matrix independent of the ξ_i^μ, $E_{z_i^M}$ denotes the expectation with respect to z_i^M and finally $E_{z_J^M}$ denotes the expectation with respect to the vector $(z_i^M)_{i\in J}$. Here we have used the well-known identity

$$\exp\left(\frac{1}{2}x^2\right) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp\left(xy-\frac{1}{2}y^2\right)dy. \tag{28}$$

Interchanging the order of integration and using the above technique on every ξ_i^μ, we are now able to consecutively replace all the variables ξ_i^μ by Gaussian random variables z_i^μ with expectation zero and identity covariance matrix. This leads to

$$E\Bigg(\exp\Bigg(-\frac{t}{N}\lambda\sum_{i\in J,\, j\notin J}\sum_{\mu_1,\mu_2=\nu+1}^{M} a_{\mu_1,\mu_2}\overline{\xi_i^{\mu_1}}\,\overline{\xi_j^{\mu_2}}\Bigg)\Bigg)$$

$$\le E_z\Bigg(\exp\Bigg(\frac{t}{N}\lambda(1+|2p-1|)^2\sum_{i\in J,\, j\notin J}\sum_{\mu_1,\mu_2=\nu+1}^{M} a_{\mu_1,\mu_2} z_i^{\mu_1} z_j^{\mu_2}\Bigg)\Bigg)$$

$$= E_z\Bigg(\exp\Bigg(t\lambda(1+|2p-1|)^2\sqrt{\delta(1-\delta)}\sum_{\mu_1,\mu_2=\nu+1} a_{\mu_1,\mu_2} z^{\mu_1} z^{\mu_2}\Bigg)\Bigg)$$

$$= E_z\Bigg(\exp\frac{1}{2}\Big(t\lambda(1+|2p-1|)^2\sqrt{\delta(1-\delta)}\langle z,\hat{A}z\rangle\Big)\Bigg),$$

where (by normalizing) the $(z^\mu)_{\mu=\nu+1,\dots,M}$ are now Gaussian random variables with expectation 0 and identity covariance matrix, z denotes the vector of the z_i^μ and E_z is integration with respect to z. Finally $\hat{A}$ is an $2(M-\nu)\times 2(M-\nu)$-matrix with entries

$$\hat{A} = \left(\begin{array}{c|c} 0 & A \\ \hline A & 0 \end{array}\right)$$

and the $(M-\nu)\times(M-\nu)$-matrix A is given by

$$A = (A_{\mu_1,\mu_2}) = (a_{\mu_1-\nu,\mu_2-\nu}).$$

Observe that the above integral only exists if t is small enough [i.e., if $Id - t\lambda(1+|2p-1|)^2\sqrt{\delta(1-\delta)}\hat{A}$ is positive definite] and in this case it equals the inverse of the square root of the determinant of $Id - t\lambda(1+|2p-1|)^2\sqrt{\delta(1-\delta)}\hat{A}$. On the other hand, since trivially Id and $\hat{A}$ are simultaneously diagonalizable,

$$\begin{aligned}\det&\Big(Id - t\lambda(1+|2p-1|)\sqrt{(\delta(1-\delta)}\hat{A}\Big)\\ &= \prod_{k=1}^{2(M-\nu)} \varrho_k\\ &= \prod_{i=k}^{2(M-\nu)} \Big(1 - t\lambda(1+|2p-1|)^2\sqrt{\delta(1-\delta)}\alpha_k\Big),\end{aligned}$$

where the ϱ_k are the eigenvalues of $Id - t\lambda(1+|2p-1|)^2\sqrt{\delta(1-\delta)}\hat{A}$ and the α_k are the eigenvalues of $\hat{A}$. Moreover note that $\hat{A}$ has a symmetric spectrum; that is, if α_k is an eigenvalue of $\hat{A}$ then so is $-\alpha_k$. Thus

$$\begin{aligned}\det&\Big(Id - t\lambda(1+|2p-1|)\sqrt{(\delta(1-\delta)}\hat{A}\Big)\\ &= \prod_{k=1}^{M-\nu} (1 - t^2\lambda^2(1+|2p-1|)^4\delta(1-\delta)\alpha_k^2),\end{aligned}$$

where the product is taken over all nonnegative eigenvalues.

Thus

$$\begin{aligned}E_z&\bigg(\exp\bigg(\frac{t}{N}\lambda(1+|2p-1|)\sqrt{\delta(1-\delta)}\sum_{\mu_1,\mu_2=\nu+1} a_{\mu_1,\mu_2} z^{\mu_1} z^{\mu_2}\bigg)\bigg)\\ &\le \prod_{k=1}^{M-\nu}\bigg(\frac{1}{\sqrt{1 - t^2\lambda^2(1+|2p-1|)^4\delta(1-\delta)\alpha_k^2}}\bigg)\\ &= \exp\bigg(-\frac{1}{2}\sum_{k=1}^{M-\nu}\log\big(1 - t^2\lambda^2(1+|2p-1|)^4\delta(1-\delta)\alpha_k^2\big)\bigg),\end{aligned}$$

where we have assumed that t is so small that the latter quantity is real [e.g., $t^2\lambda^2(\delta(1-\delta)/(1-(2p-1)^2)^2)(4(1+|2p-1|)^4/(1-|2p-1|)^2) \le 1$ suffices as by Gershgorin's theorem,

$$|\alpha_k| \le |\alpha_{\max}| \le \max_{\mu_1} \sum_{\mu_2} |a_{\mu_1,\mu_2}| \le \frac{1}{1-(2p-1)^2}\frac{2}{1-|2p-1|},$$

where $\alpha_{\max}$ denotes the maximal eigenvalue of $\hat{A}$].

Thus, repeating the estimate for the moment generating function of the second quadratic form,

$$\begin{aligned} &P\big(H_N(\xi_J^\nu) - H_N(\xi^\nu) \le \varepsilon N\big) \\ &\quad \le \inf_{t^* \ge t \ge 0} \exp\bigg(-t\varepsilon' N - tN\delta(1-\delta)\frac{1}{C'(1-(2p-1)^2)}\bigg) \\ &\qquad \times \exp\bigg(-\frac{1}{2}\sum_{k=1}^{M-\nu} \log\big(1 - t^2\lambda^2(1+|2p-1|)^4\delta(1-\delta)\alpha_k^2\big)\bigg) \\ &\qquad \times \exp\bigg(-\frac{1}{2}\sum_{k=1}^{\nu} \log\big(1 - t^2\lambda^2(1+|2p-1|)^4\delta(1-\delta)\tilde{\alpha}_k^2\big)\bigg) \\ &\qquad \times \exp\bigg(t^2\delta(1-\delta)\frac{\lambda}{\lambda-1}(1+|2p-1|)^2\frac{1}{(1-(2p-1)^2)^3}\bigg), \end{aligned}$$

where $t^* = ((1-(2p-1)^2)(1-|2p-1|)/2\lambda(1+|2p-1|)^2)\sqrt{1/\delta(1-\delta)}$, the $\tilde{\alpha}_k$'s are the positive eigenvalues of the matrix

$$\tilde{\tilde{A}} = \left(\begin{array}{c|c} 0 & \tilde{A} \\ \hline \tilde{A} & 0 \end{array}\right)$$

and the $\nu \times \nu$-matrix $\tilde{A}$ is given by $\tilde{A} = (\tilde{A}_{\mu_1,\mu_2}) = (\tilde{a}_{\mu_1,\mu_2})$.

Finally, by Stirling's formula (to bound the binomial coefficient), the ansatz $M = \alpha N$ and the above estimate,

$$\begin{aligned} &\sum_{J:|J|=\delta N} \sum_{\nu=1}^{M(N)} P\big(H_N(\xi_J^\nu) - H_N(\xi^\nu) \le \varepsilon N\big) \\ &\quad \le M(N)\binom{N}{\delta N} \exp\bigg(-t\varepsilon' N - tN\delta(1-\delta)\frac{1}{C'(1-(2p-1)^2)}\bigg) \\ &\qquad \times \exp\bigg(-\frac{1}{2}\sum_{k=1}^{M-\nu} \log\big(1 - t^2\lambda^2(1+|2p-1|)^4\delta(1-\delta)\alpha_k^2\big)\bigg) \\ &\qquad \times \exp\bigg(-\frac{1}{2}\sum_{k=1}^{\nu} \log\big(1 - t^2\lambda^2(1+|2p-1|)^4\delta(1-\delta)\tilde{\alpha}_k^2\big)\bigg) \\ &\qquad \times \exp\bigg(t^2\delta(1-\delta)\bigg(\frac{\lambda}{\lambda-1}\bigg)(1+|2p-1|)^2\frac{1}{(1-(2p-1)^2)^3}\bigg) \end{aligned}$$

$$
\begin{aligned}
&\le \alpha N \inf_{t^*\ge t\ge 0} \exp\big((-\delta\log\delta-(1-\delta)\log(1-\delta))N\big)\\
&\quad\times \exp\bigg(-t\varepsilon' N - tN\delta(1-\delta)\frac{1}{C'(1-(2p-1)^2)}\bigg)\\
&\quad\times \exp\bigg(-\frac{1}{2}\sum_{k=1}^{M-\nu}\log\big(1-t^2\lambda^2(1+|2p-1|)^4\delta(1-\delta)\alpha_k^2\big)\bigg)\\
&\quad\times \exp\bigg(-\frac{1}{2}\sum_{k=1}^{\nu}\log\big(1-t^2\lambda^2(1+|2p-1|)^4\delta(1-\delta)\tilde{\alpha}_k^2\big)\bigg)\\
&\quad\times \exp\bigg(t^2\delta(1-\delta)\frac{\lambda}{\lambda-1}(1+|2p-1|)^2\frac{1}{(1-(2p-1)^2)^3}\bigg)
\end{aligned}
$$

and we have to find an admissible t (i.e., $0\le t\le t^*$) and values of δ and α such that the above exponent becomes negative. To this end, first note that for all admissible t,

$$
\exp\bigg(t^2\delta(1-\delta)\bigg(\frac{\lambda}{\lambda-1}\bigg)(1+|2p-1|)^2\frac{1}{(1-(2p-1)^2)^3}\bigg)=\mathscr{O}(1)
$$

and therefore this term does not influence the convergence (as promised above).

Moreover, if $t^2\lambda^2(\delta(1-\delta)/(1-(2p-1)^2)^2)(4(1+|2p-1|)^4/(1-|2p-1|)^2)\le 3/4$,

$$
\frac{1}{\sqrt{1-t^2\lambda^2(1+|2p-1|)^4\delta(1-\delta)\alpha_k^2}}\le \exp\big(t^2\lambda^2\delta(1-\delta)(1+|2p-1|)^4\alpha_k^2\big)
$$

as well as the same inequality for the $\tilde{\alpha}_k$-terms.

Hence, up to factors of order one $\sum_{J:|J|=\delta N}\sum_{\nu=1}^{M(N)}P(H_N(\xi_J^\nu)-H_N(\xi^\nu)\le \varepsilon N)$ can be bounded by

$$
\begin{aligned}
&\exp\bigg(\big(-\delta\log\delta-(1-\delta)\log(1-\delta)\big)N-t\varepsilon' N-tN\delta(1-\delta)\frac{1}{C'(1-(2p-1)^2)}\bigg)\\
&\quad\times \exp\bigg(-\frac{1}{2}\sum_{k=1}^{\alpha N-\nu}\log\bigg(1-t^2\lambda^2(1+|2p-1|)^4\delta(1-\delta)\alpha_k^2\sum_{\mu=\nu+1}^{M-1}a_{\mu,\nu}\sum_{\mu=\nu+1}^{M-1}a_{\mu,\nu}\bigg)\bigg)\\
&\quad\times \exp\bigg(-\frac{1}{2}\sum_{k=1}^{\nu}\log\big(1-t^2\lambda^2(1+|2p-1|)^4\delta(1-\delta)\tilde{\alpha}_k^2\big)\bigg)\\
&\quad\le \exp\bigg(\big(-\delta\log\delta-(1-\delta)\log(1-\delta)\big)N-t\varepsilon' N\\
&\qquad\qquad - tN\delta(1-\delta)\frac{1}{C'(1-(2p-1)^2)}\bigg)\\
&\qquad\times \exp\bigg(t^2\lambda^2(1+|2p-1|)^4\delta(1-\delta)\bigg(\sum_{k=1}^{\alpha N-\nu}\alpha_k^2+\sum_{k=1}^{\nu}\tilde{\alpha}_k^2\bigg)\bigg)
\end{aligned}
$$

if

$$t \le t^{**} := \frac{(1-(2p-1)^2)(1-|2p-1|)}{4\lambda(1+|2p-1|)^2}\sqrt{\frac{3}{\delta(1-\delta)}}.$$

Now

$$\sum_{k=1}^{\alpha N-\nu} \alpha_k^2 = \tfrac{1}{2}\operatorname{tr}(\hat{A}) = \tfrac{1}{2}\times 2\sum_{\mu_1=1}^{\alpha N-\nu}\sum_{\mu_2=1}^{\alpha N-\nu} \alpha_{\mu_1,\mu_2}\alpha_{\mu_2,\mu_1} = \sum_{\mu_1=1}^{\alpha N-\nu}\sum_{\mu_2=1}^{\alpha N-\nu} \alpha_{\mu_1,\mu_2}^2$$

and with the definition of α_{μ_1,μ_2} one therefore obtains

$$\sum_{k=1}^{\alpha N-\nu} \alpha_k^2 = \frac{(\alpha N-\nu)(1+(2p-1)^2)}{(1-(2p-1)^2)^3} + \mathscr{O}(1),$$

where the $\mathscr{O}(1)$ refers to the N that will tend to infinity. As also $\sum_{k=1}^{\nu}\tilde{\alpha}_k^2 = (\nu(1+(2p-1)^2)/(1-(2p-1)^2)^3) + \mathscr{O}(1)$, we obtain—again for $t \le t^{**}$,

$$\begin{aligned}
&\sum_{J:\,|J|=\delta N}\sum_{\nu=1}^{M(N)} P\big(H_N(\xi_J^\nu) - H_N(\xi^\nu) \le \varepsilon N\big)\\
&\qquad \le \exp\bigg(\big(-\delta\log\delta-(1-\delta)\log(1-\delta)\big)N - t\varepsilon' N\\
&\qquad\qquad\qquad - tN\delta(1-\delta)\frac{1}{C'(1-(2p-1)^2)}\bigg)\\
&\qquad\quad \times \exp\bigg(t^2\lambda^2(1+|2p-1|)^4\delta(1-\delta)\alpha N\frac{1+(2p-1)^2}{(1-(2p-1)^2)^3}\bigg)\times\mathscr{O}(1).
\end{aligned}\tag{29}$$

Choosing ε very small, the exponent is minimized by a t which is close to

$$t_{\min} = \frac{1}{\alpha}\,\frac{1}{2\lambda^2(1+|2p-1|)^4}\bigg((1-(2p-1)^2)+\frac{1}{C'}\bigg)\frac{(1-(2p-1)^2)^2}{1+(2p-1)^2}.$$

Observe that $t_{\min} \le t^{**}$ if

$$\alpha \ge \sqrt{\delta(1-\delta)}\,\frac{2(1-(2p-1)^2+(1/C'))(1-(2p-1)^2)}{\sqrt{3}\lambda(1+|2p-1|)^2(1+(2p-1)^2)(1-|2p-1|)}.\tag{30}$$

On the other hand, inserting $t_{\min}$ into the essential part of the exponent and choosing ε sufficiently small gives (for the exponent)

$$\begin{aligned}
&(-\delta\log\delta-(1-\delta)\log(1-\delta))N - t_{\min}\varepsilon' N\\
&\quad - t_{\min}N\delta(1-\delta)\frac{1}{C'(1-(2p-1)^2)}\\
&\quad + t_{\min}^2\lambda^2(1+|2p-1|)^4\delta(1-\delta)\alpha N\frac{1+(2p-1)^2}{(1-(2p-1)^2)^3}\\
&\qquad \le (-\delta\log\delta-(1-\delta)\log(1-\delta))N\\
&\qquad\quad - \frac{\gamma}{4\alpha}\delta(1-\delta)N\frac{(1-(2p-1)^2)(1-(2p-1)^2+(1/C'))^2}{\lambda^2(1+|2p-1|)^4(1+(2p-1)^2)}
\end{aligned}\tag{31}$$

with $\gamma < 1$ and close to 1 (as ε becomes small). The right-hand side of this inequality becomes negative when δ and α are chosen appropriately. To check whether this can be done in agreement with (30), we insert

$$\alpha = \sqrt{\delta(1-\delta)}\frac{2(1-(2p-1)^2+(1/C'))(1-(2p-1)^2)}{\sqrt{3}\lambda(1+|2p-1|)^2(1+(2p-1)^2)(1-|2p-1|)}$$

into the right-hand side of (31) and obtain

$$\begin{aligned}\Bigl(&-\frac{\sqrt{3}\gamma(1-(2p-1)^2+(1/C'))}{2\lambda(1+|2p-1|)^2}(1-|2p-1|)\sqrt{\delta(1-\delta)}\\ &\qquad\qquad -\delta\log\delta-(1-\delta)\log(1-\delta)\Bigr)N.\end{aligned} \tag{32}$$

As it is quickly checked that for each positive constant C there is an interval $[0, r]$ (depending on C, of course) such that

$$C\sqrt{\delta(1-\delta)} \geq -\delta\log\delta-(1-\delta)\log(1-\delta)$$

for all $\delta \in [0, r]$, the above exponent becomes negative if we choose δ small enough and, for example, α as the right-hand side of (30). This completes the proof of the theorem. □

REMARK 4.2. Observe that the bound on the moment generating function in (31) depends on p mainly via the factor $(1-|2p-1|)$ (the other terms containing p are bounded from above and away from 0), which converges to zero for p close to one or close to zero and therefore can only deteriorate the bounds for δ or α (allowing smaller $\alpha's$ or $\delta's$ only) for large correlations. The bound on the admissible α in (30) shows the interplay between α and δ. As $(1-(2p-1)^2/1-|2p-1|) \to 1$ for $p \to 0$ or $p \to 1$, (30) seems to indicate that one might formally choose α independent of p, but then (32) shows that in this case δ shrinks to 0 when the correlations become large and therefore (as $\alpha \sim \sqrt{\delta}$) so does α. Indeed, such a behavior can already be expected from (29), when substituting t by $t(1-(2p-1)^2)$ and noticing that the t^2-term then is still multiplied by a factor $1/(1-(2p-1)^2)$, that is, the quadratic term grows faster for $p \to 0, 1$ than the linear term. This behavior of course seems to be in contradiction to the result of Theorem 2.1. On the other hand, this contradiction might well be a result of the different notions of storage capacity. Indeed, although the calculations in the proof of Theorem 2.3 are rather lengthy, there seems to be only one inequality (Lemma 3.1) where we possibly could lose essential factors for the qualitative behavior of the storage capacity with large correlations.

PROOF OF THEOREM 2.4. The central idea of the proof is the same as in the proof of Theorem 2.3: a centering of the patterns as prepared in Lemma 3.2 and their replacement by appropriate Gaussian random variables. To be able to evaluate the resulting integral we will make use of Lemma 3.1(ii).

With the notation of the proof of Theorem 2.3, we first of all observe that

$$P\Bigg(\Bigg\{\bigcap_{\nu=1}^{M(N)}(h_N(\xi^\nu,\delta)\ge H_N(\xi^\nu)+\varepsilon N)\Bigg\}^c\Bigg)$$
$$\le \sum_{J:|J|=\delta N}\sum_{\nu=1}^{M(N)} P\big(H_N(\xi^\nu_J)-H_N(\xi^\nu)\le \varepsilon N\big).$$

Again let us keep ν fixed in the sequel (without loss of generality we choose $\nu=1$).

By the exponential Chebyshev–Markov inequality for any $t\ge 0$,

$$\begin{aligned}(33)\qquad &P\big(H_N(\xi^1_J)-H_N(\xi^1)\le\varepsilon N\big)\\ &\quad\le \exp(-t\varepsilon' N)E\Bigg(\exp\Bigg(-\frac{t}{N}\sum_{\mu\ne 1}\sum_{\substack{(i\in J,\, j\notin J)\\ (i\notin J,\, j\in J)}}\xi_i^1\xi_j^1\xi_i^\mu\xi_j^\mu\Bigg)\Bigg)\\ &\quad= \exp(-t\varepsilon' N)E\Bigg(\exp\Bigg(-\frac{t}{N}\sum_{\substack{(i\in J,\, j\notin J)\\ (i\notin J,\, j\in J)}}\xi_i^1\xi_j^1\xi_i^2\xi_j^2\Bigg)\Bigg)^{M-1},\end{aligned}$$

where we have set $\varepsilon'=-\varepsilon+2\delta(1-\delta)$.

Our main goal is now to estimate the expectation on the right-hand side of (33). To this end, note that the exponent is a quadratic form in Markovian random variables as treated in Lemma 3.2. Indeed, putting

$$Y_i := \xi_i^1\xi_i^2$$

and

$$a_{i,j} := \begin{cases}1, & \text{if } i\in J,\ j\notin J,\\ 1, & \text{if } i\notin J,\ j\in J,\\ 0, & \text{otherwise,}\end{cases}$$

(note that this especially implies that $a_{i,i}=0$) we obtain that

$$E\Bigg(\exp\Bigg(-\frac{t}{N}\sum_{\substack{(i\in J,\, j\notin J)\\ (i\notin J,\, j\in J)}}\xi_i^1\xi_j^1\xi_i^2\xi_j^2\Bigg)\Bigg)=E\Bigg(\exp\Bigg(-\frac{t}{N}\sum_{i,j}a_{i,j}Y_iY_j\Bigg)\Bigg).$$

Observe that Y_i is a Markov chain on the set $\{-1,+1\}$ (in i) with transition matrix R given by

$$R=\begin{pmatrix}p^2+(1-p)^2 & 2p(1-p)\\ 2p(1-p) & p^2+(1-p)^2\end{pmatrix}=:\begin{pmatrix}q & 1-q\\ 1-q & q\end{pmatrix}$$

(notice that $q\ge 1/2$). Hence we are in the situation of Lemma 3.2. Therefore

$$\begin{aligned}(34)\qquad &E\Bigg(\exp\Bigg(-\frac{t}{N}\sum_{i,j}a_{i,j}Y_iY_j\Bigg)\Bigg)\\ &\quad=\exp\Bigg(-\frac{2t}{N}\sum_{1\le i<j\le N}(2q-1)^{j-i}a_{i,j}\Bigg)E\Bigg(\exp\Bigg(-\frac{t}{N}\sum_{i,j}b_{i,j}\overline{Y}_i\overline{Y}_j\Bigg)\Bigg),\end{aligned}$$

where the $\overline{Y}_i$ and $b_{i,j}$ are defined as in (10) and (12), respectively. To estimate the right-hand side of (34), observe that due to the symmetry of the $a_{i,j}$ we have that $2\sum_{1\le i<j\le N}(2q-1)^{j-i}a_{i,j} = \sum_{1\le i,j\le N}(2q-1)^{|j-i|}a_{i,j}$ and that

$$
\begin{aligned}
&2\left(\frac{2q-1}{2(1-q)}\right)^2(1-(2q-1)^{\delta N})(1-(2q-1)^{(1-\delta)N}) \\
(35)\qquad &\le \sum_{1\le i,j\le N}(2q-1)^{|j-i|}a_{i,j} \\
&\le \frac{N}{1-q}.
\end{aligned}
$$

The upper bound in (35) is derived by simply setting all the $a_{i,j}=1$ and is just to show that the term $\exp(-(2t/N)\sum_{1\le i<j\le N}(2q-1)^{j-i}a_{i,j})$ is at most of the same order as the expectation on the right-hand side of (34) (which, for example, may fail in the Hopfield model with biased patterns if the Hamiltonian is not appropriately normalized; this would be a sign that the model is not chosen correctly). The lower bound in (35) is because the number of $a_{i,j}$ that are equal to one is independent of the choice of J such that $\sum_{1\le i,j\le N}(2q-1)^{|j-i|}a_{i,j}$ becomes minimal for sets J which are maximally connected, for example, $J=\{1,\dots,\delta N\}$. Hence for N large enough

$$\exp\left(-\frac{2t}{N}\sum_{1\le i<j\le N}(2q-1)^{j-i}a_{i,j}\right) \le \exp\left(-\frac{t}{N}\left(\frac{2q-1}{2(1-q)}\right)^2\right).$$

To estimate the expectation on the right-hand side of (34), we employ Lemma 3.1(ii). First, let us agree on the following notation: with E_I (where $I\subset\{1,\dots,N\}$) we denote the integration with respect to those random variables Y_i (resp. $\overline{Y}_i$) with $i\in I$. With these notations we will be able to prove that

$$(36)\qquad E\left(\exp\left(-\frac{t}{N}\sum_{1\le i,j\le N}b_{i,j}\overline{Y}_i\overline{Y}_j\right)\right) \le E\left(\exp\left(\frac{2qt}{N}\sum_{1\le i,j\le N}b_{i,j}z_iz_j\right)\right),$$

where the z_i are i.i.d. standard Gaussian random variables.

Indeed,

$$
\begin{aligned}
&E\left(\exp\left(-\frac{t}{N}\sum_{i,j}b_{i,j}\overline{Y}_i\overline{Y}_j\right)\right) \\
&\quad = E\left(\exp\left(-\frac{2t}{N}\sum_{1\le i<j\le N}b_{i,j}\overline{Y}_i\overline{Y}_j\right)\right) \\
&\quad = E_{\{1,\dots,N-1\}}\left(\exp\left(-\frac{2t}{N}\sum_{1\le i<j\le N-1}b_{i,j}\overline{Y}_i\overline{Y}_j\right)\right. \\
&\quad\qquad \times \left. E_{\{N\}}\left(\exp\left(-\frac{2t}{N}\sum_{i=1}^{N-1}b_{i,N}\overline{Y}_i\overline{Y}_N\right)\right)\right)
\end{aligned}
$$

$$
\begin{aligned}
&= E_{\{1,\dots,N-1\}}\biggl(\exp\biggl(-\frac{2t}{N}\sum_{1\le i<j\le N-1} b_{i,j}\overline{Y}_i\overline{Y}_j\biggr)\biggr)\\
&\quad\times\biggl(q\exp\biggl(-4(1-q)\frac{t}{N}\sum_{i=1}^{N-1} b_{i,N}\overline{Y}_i\biggr)+(1-q)\exp\biggl(4q\frac{t}{N}\sum_{i=1}^{N-1} b_{i,N}\overline{Y}_i\biggr)\biggr)\\
&\le E_{\{1,\dots,N-1\}}\biggl(\exp\biggl(-\frac{2t}{N}\sum_{1\le i<j\le N-1} b_{i,j}\overline{Y}_i\overline{Y}_j\biggr)\biggr)\times\cosh\biggl(\frac{4tq}{N}\sum_{i=1}^{N-1} b_{i,N}\overline{Y}_i\biggr)\\
&\le E_{\{1,\dots,N-1\}}\biggl(\exp\biggl(-\frac{2t}{N}\sum_{1\le i<j\le N-1} b_{i,j}\overline{Y}_i\overline{Y}_j\biggr)\biggr)\\
&\quad\times\exp\biggl(2\frac{q^2t^2}{N^2}\biggl(\sum_{i=1}^{N-1} b_{i,N}\overline{Y}_i\biggr)^2\biggr),
\end{aligned}
$$

where we have used the inequality (8).

Again using (28) yields

$$
\begin{aligned}
&E\biggl(\exp\biggl(-\frac{t}{N}\sum_{i,j} b_{i,j}\overline{Y}_i\overline{Y}_j\biggr)\biggr)\\
&\qquad\le E_{\{1,\dots,N-1\}}\biggl(\exp\biggl(-\frac{2t}{N}\sum_{1\le i<j\le N-1} b_{i,j}\overline{Y}_i\biggr)\biggr)E_{z_N}\exp\biggl(\frac{4qt}{N}\sum_{i=1}^{N-1} b_{i,N}\overline{Y}_i z_N\biggr),
\end{aligned}
$$

where E_{z_N} denotes integration with respect to the Gaussian random variable z_N. Note that we have replaced $\overline{Y}_N$ by z_N and also observe that this replacement was independent of all the other $\overline{Y}_i$, $i \neq N$ and just relied on the fact that $\overline{Y}_N$ was appropriately centered such that part one of Lemma 3.1 could be applied. Since this is true for all the $\overline{Y}_i$, too, the other replacements can be carried out along the same lines just by rearranging the order of integration appropriately, such that (36) indeed is true.

So we have arrived at

$$
\begin{aligned}
&E\biggl(\exp\biggl(-\frac{t}{N}\sum_{1\le i,j\le N} a_{i,j}Y_iY_j\biggr)\biggr)\\
&\qquad\le\exp\biggl(-\frac{t}{N}\biggl(\frac{2q-1}{2(1-q)}\biggr)^2\biggr)E\biggl(\exp\biggl(\frac{2qt}{N}\sum_{1\le i,j\le N} b_{i,j}z_iz_j\biggr)\biggr).
\end{aligned}
\tag{37}
$$

Note that the expectation on the right-hand side of (37) is only defined if t is small enough, that is, if $Id-(2qt/N)B$ is a positive definite matrix (here B is the quadratic form given by $B=(b_{i,j})_{1\le i,j\le N}$). In this case

$$
E\biggl(\exp\biggl(\frac{2qt}{N}\sum_{i,j} b_{i,j}z_iz_j\biggr)\biggr)=\frac{1}{\sqrt{\det(Id-(2qt/N)B)}}.
$$

To estimate this determinant, denote by $(\beta_i)_{1\le i\le N}$ the (real) eigenvalues of B. Then, on one hand, since Id and B are trivially simultaneously diagonalizable,

$$
\det\biggl(Id-\frac{2qt}{N}B\biggr)=\prod_{i=1}^{N}\biggl(1-\frac{2qt}{N}\beta_i\biggr).
$$

On the other hand $\sum_{i=1}^N \beta_i = 0$, since all the diagonal entries of B are zero. Now let us assume that t is so small that even

$$\frac{4t^2q^2}{N^2}\sum_{i=1}^N \beta_i^2 \le \frac{3}{4}. \tag{38}$$

An immediate application of (9) yields that

$$\prod_{i=1}^N \left(1-\frac{2qt}{N}\beta_i\right) \ge 1-\frac{4q^2t^2}{N^2}\sum_{i=1}^N \beta_i^2,$$

implying that

$$\frac{1}{\sqrt{\det(Id-(2qt/N)B)}} \le \frac{1}{\sqrt{1-(4q^2t^2/N^2)\sum_{i=1}^N \beta_i^2}}.$$

Since $1/\sqrt{1-x} \le e^x$ for $0 \le x \le 3/4$, we obtain

$$\frac{1}{\sqrt{\det(Id-(2qt/N)B)}} \le \exp\left(\frac{4q^2t^2}{N^2}\sum_{i=1}^N \beta_i^2\right).$$

Finally, observe that due to the symmetry of B,

$$\sum_{i=1}^N \beta_i^2 = \operatorname{tr}(B^2) = \sum_{i=1}^N\sum_{j=1}^N b_{i,j}^2.$$

Since

$$\sum_{i,j=1}^N a_{i,j} = 2\delta(1-\delta)N^2,$$

we arrive at

$$\begin{aligned}\sum_{i=1}^N\sum_{j=1}^N b_{i,j}^2 &= \sum_{i=1}^N\sum_{j=1}^N\sum_{k=0}^{j-i-1}\sum_{l=0}^{N-j}(2q-1)^{k+l}a_{i+k,j+l}\\ &\le 2\left(\frac{1}{2(1-q)}\right)^2\delta(1-\delta)N^2.\end{aligned} \tag{39}$$

So altogether for $t \ge 0$ and N large enough,

$$\begin{aligned}&P\big(H_N(\xi_J^1)-H_N(\xi^1)\le \varepsilon N\big)\\ &\quad\le \exp(-t\varepsilon' N)\exp\left(-\frac{tM}{N}\left(\frac{2q-1}{2(1-q)}\right)^2\right)\\ &\qquad\times\exp\left(8q^2t^2\left(\frac{1}{2(1-q)}\right)^2\delta(1-\delta)M\right).\end{aligned} \tag{40}$$

Now note that the above bound (40) is uniform in J with $|J| = \delta N$. Hence by setting $M = \alpha N$ and estimating the term $\binom{N}{\delta N}$ by Stirling's formula we arrive at

$$\begin{aligned}\sum_{J:|J|=\delta N} \sum_{\nu=1}^{M(N)} & P\big(H_N(\xi_J^\nu) - H_N(\xi^\nu) \leq \varepsilon N\big) \\ &\leq \alpha N \inf_{t^* \geq t \geq 0} \exp\big((-\delta \log \delta - (1-\delta)\log(1-\delta))N\big) \\ &\quad \times \exp\left(-t\left(\varepsilon' N + \alpha\left(\frac{2q-1}{2(1-q)}\right)^2\right) + 8t^2 q^2 \left(\frac{1}{2(1-q)}\right)^2 \delta(1-\delta)\alpha N\right),\end{aligned}$$

where t^* is the maximum t fulfilling (38) [e.g., choosing $t^* := (1-q)^2/4\alpha q^2$ and

$$\alpha \geq \sqrt{\frac{\delta(1-\delta)}{6}}\,\frac{1-q}{q} \tag{41}$$

suffices, since then t^* is admissible in the sense that (38) is fulfilled. Indeed, together with (39) we obtain

$$\begin{aligned}\frac{4t^2q^2}{N^2}\sum_{i=1}^N \beta_i^2 &\leq 2\frac{4q^2(1-q)^4}{16q^4\alpha^2N^2}\frac{1}{4(1-q)^2}\delta(1-\delta)N^2 \\ &= \frac{(1-q)^2\delta(1-\delta)}{8q^2\alpha^2} \leq \frac{3}{4}.\end{aligned}$$

Now the proof can be completed along the lines of the proof of Theorem 2.3. □

Finally, let us comment a little on the result of Theorem 2.4. Observe that the bound on the moment generating function in (40) as well as the bound on α in (41) depends on p in a way that implies that large correlations again lead to smaller α's (and δ's).

This might, of course, originate from the fact that due to the very crude lower bound in (35), our bound on

$$P\big(H_N(\xi_J^\nu) - H_N(\xi^\nu) \leq \varepsilon N\big)$$

is independent of the special choice of J and just depends on its size. Indeed, for most choices of J, the sum $\sum_{1\leq i,\, j\leq N}(2q-1)^{|j-i|}a_{i,\,j}$ is not of order constant but of some order growing with N (maybe even of order N). However, to see whether such a refined estimate really can change the qualitative behavior in q (which in a way would be the only interesting change), that is, if the above sum is of order $N/(1-q)$ [i.e., the order of the upper bound in (36)] for "almost all" choices of J requires a large deviation analysis we have not been able to do. On the other hand, in view of Theorems 2.2 and 2.3, one may also conjecture that improving the estimates in this point may, of course, increase the numerical value of the storage capacity, but would not change the qualitative behavior in p.

Moreover, the reader may have noticed that for the proofs of Theorem 2.3 and 2.4 we have invented techniques to be able to adapt the ideas of Newman [21] to our situation rather than the recent improvements by Loukianova [16] and Talagrand [26]. The reason is that, while establishing some extra difficulties on the one side, these improvements are basically useful to give a better numerical value for α_c on the other (but not its behavior in p). Since we think that such a numerical constant would only be interesting if we could describe the exact borderline between memory and loss of memory or at least if we could give a bound for α_c with a uniform error in p (and we cannot see how to do that), we have tried to avoid these unnecessary complications.

Finally, let us mention that we assume that the storage abilities of the Hopfield model are not limited to the case where the correlations come from a Markov chain (at least the cases of exponentially decaying correlations or even summable correlations should be tractable). Our method, however, seems to be thus limited. The reader may notice that even for the (interesting) case of a two-dimensional Markov random field (Ising model) the centered variables do not depend only on the two-point correlations any longer, so that new methods to treat the corresponding moment generating function have to be invented.

Acknowledgments. The author thanks the referee for some very thoughtful remarks, which helped to improve some bounds on the storage capacities. He also thanks Holger Knöpfel for discussions on the analytical problems.

REFERENCES

[1] AMIT, D. J. (1987). The properties of models of simple neural networks. *Heidelberg Colloquium on Glassy Dynamics. Lecture Notes in Phys.* **275**. Springer, Berlin.

[2] AMIT, D. J., GUTFREUND, G. and SOMPOLINSKY, H. (1985). Spin-glass models of neural networks. *Phys. Rev. A* **32** 1007–1018.

[3] AMIT, D. J., GUTFREUND, G. and SOMPOLINSKY, H. (1987). Statistical mechanics of neural networks near saturation. *Ann. Physics* **173** 30–67.

[4] BOVIER, A. and GAYRARD, V. (1992). Rigorous bounds on the storage capacity of the dilute Hopfield model. *J. Statist. Phys.* **69** 597–627.

[5] BOVIER, A. and GAYRARD, V. (1996). An almost sure large deviation principle for the Hopfield model. *Ann. Probab.* **24** 1444–1475.

[6] BOVIER, A. and GAYRARD, V. (1997). Hopfield models as a generalized mean field model. In *Mathematics of Spin Glasses and Neural Networks* (A. Bovier and P. Picco, eds.). Birkhäuser, Boston. To appear.

[7] BOVIER, A. and GAYRARD, V. (1997). The retrieval phase of the Hopfield model. *Probab. Theory Related Fields* **107** 61–98.

[8] BOVIER, A., GAYRARD, V. and PICCO, P. (1994). Gibbs states for the Hopfield model in the regime of perfect memory. *Probab. Theory Related Fields* **100** 329–363.

[9] BOVIER, A., GAYRARD, V. and PICCO, P. (1995). Large deviation principles for the Hopfield model and the Kac-Hopfield model. *Probab. Theory Related Fields* **101** 511–546.

[10] BOVIER, A., GAYRARD, V. and PICCO, P. (1995). Gibbs states for the Hopfield model with extensively many patterns. *J. Statist. Phys.* **79** 395–414.

[11] DREYFUS, G., GUYON, I. and PERSONNAZ, L. (1986). Neural network design for efficient information retrieval. Disordered systems and biological organization (Les Houches, 1985). *NATO Adv. Sci. Inst. Ser. F Comput. Systems Sci.* **20** 227–231.

[12] GENTZ, B. (1996). A central limit theorem for the overlap parameter in the Hopfield model. *Ann. Probab.* **24** 1809–1841.

[13] GEORGII, H.-O. (1988). Gibbs measures and phase transition. In *Studies in Mathematics* **9** (H. V. Bauer, J. Heinz-Kazden and E. Zehnder, eds.). de Gruyter, Berlin.

[14] HOPFIELD, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proc. Nat. Acad. Sci. U.S.A.* **79** 2554–2558.

[15] KÜHN, R. and STEFFAN, H. (1994). Replica symmetry breaking in attractor neural network models. *Z. Phys. B* **95** 249–260.

[16] LOUKIANOVA, D. (1994). Capacité de mémoire dans le modèle de Hopfield. *C.R. Acad. Sci. Paris* **318** 157–160.

[17] LOUKIANOVA, D. (1997). Lower bounds on the restitution error in the Hopfield model. *Probab. Theory Related Fields* **107** 161–176.

[18] MCELIECE, R., POSNER, E., RODEMICH, E. and VENKATESH, S. (1987). The capacity of the Hopfield associative memory. *IEEE Trans. Inform. Theory* **33** 461–482.

[19] MIYASHITA, Y. (1988). Neuronal correlate of visual associative long term memory in the primate temporal cortex. *Nature* **335** 817–819.

[20] MONASSON, R. (1992). Properties of neural networks storing spatially correlated patterns. *J. Phys. A Math. Gen.* **335** 3701–3720.

[21] NEWMAN, C. (1988). Memory capacity in neural networks. *Neural Networks* **1** 223–238.

[22] PASTUR, L. A. and FIGOTIN, A. L. (1977). Exactly soluble model of a spin-glass. *Soviet J. of Low Temperature Phys.* **3** 378–383.

[23] PETRITIS, D. (1995). Thermodynamic formalism of neural computing. Univ. Rennes I. Preprint.

[24] SANDMEIER, M. (1997). On the storage capacity of neural networks with temporal association. Ph.D. thesis. Univ. Bielefeld.

[25] TALAGRAND, M. (1995). Résultats rigoureux pour le modèle de Hopfield. *C.R. Acad. Sci. Paris Ser. I* **321** 309–312.

[26] TALAGRAND, M. (1996). Rigorous results of the Hopfield model with many patterns. Preprint.

[27] TARKOWSKI, W. and LEWENSTEIN, M. (1993). Storage of sets of correlated data in neural network memories. *J. Phys. A Math. Gen.* **26** 2453–2469.

[28] VAN HEMMEN, L. and KÜHN, R. (1991). Collective phenomena in neural networks. In *Models of Neural Networks* (E. Domany, L. v. Hemmen and R. Schulte, eds.). Springer, Berlin.

UNIVERSITÄT BIELEFELD
FAKULTÄT FÜR MATHEMATIK
POSTFACH 100131
33501 BIELEFELD
GERMANY
E-MAIL: loewe@mathematik.uni-bielefeld.de