Math. Stat. Learn. 7 (2024), 155–187
DOI 10.4171/MSL/45

Pattern reconstruction with restricted Boltzmann machines

Giuseppe Genovese

Abstract. Restricted Boltzmann machines are energy models made of a visible and a hidden layer. We identify an effective energy function describing the zero-temperature landscape on the visible units and depending only on the tail behaviour of the hidden layer prior distribution. Studying the location of the local minima of such an energy function, we show that the ability of a restricted Boltzmann machine to reconstruct a random pattern depends indeed only on the tail of the hidden prior distribution. We find that hidden priors with strictly super-Gaussian tails give only a logarithmic loss in pattern retrieval, while an efficient retrieval is much harder with hidden units with strictly sub-Gaussian tails; if the hidden prior has Gaussian tails, the retrieval capability is determined by the number of hidden units (as in the Hopfield model).

1. Introduction

Restricted Boltzmann machines (RBMs) are represented by probability distributions on the product space $\{-1,1\}^{N_1} \times \mathbb{R}^{N_2}$ whose density with respect to the uniform probability on $\{-1,1\}^{N_1}$ times some prior distribution on $\mathbb{R}^{N_2}$ depends on a matrix valued parameter W (so-called weight matrix), and it is proportional to

$$\exp\Bigg(\sum_{i\in[N_2]}\sum_{\mu\in[N_2]} \sigma_i W_i^{\mu} z_{\mu}\Bigg). \tag{1.1}$$

The vectors $(\sigma_1,\dots,\sigma_{N_1})$ and $(z_1,\dots,z_{N_2})$ are called, respectively, visible and hidden layers and their entries visible and hidden units. Typically, the units are i.i.d.

RBMs are widely studied generative models of machine learning, introduced long ago in [23]. Their mathematical relation with models of associative memory, such as the ones proposed by Little [18] or Hopfield [16], was noted at the very early stage of the theory, see [15]. Indeed, exploiting the product structure of the RBM distribution integrating out the hidden layer, one can analyse the corresponding model of associative memory, see [4, 9, 27]; the simplest example is the Hopfield model, obtained by an RBM with Gaussian hidden prior by a Hubbard–Stratonovic transformation. The present work aligns with this line of research.

Mathematics Subject Classification 2020: 82B44 (primary); 68T10 (secondary).
Keywords: restricted Boltzmann machines, associative memories, capacity bounds.

We will be interested in particular to the possibility of reconstructing a pattern just looking at the typical configurations of the visible layer of a RBM. This operation is also called pattern retrieval. The question is relevant for the understanding of how the hidden layer affects the configurations of the visible units in the RBM distribution. In the last years the main focus on RBMs has been on learning the unknown probability distribution underlying a given dataset [8, 10, 13, 14]. For the practitioner, learning a RBM amounts to fitting the true law of the data by tuning the weights in the density (1.1) and this is typically done by gradient ascent on the Kullback–Leibler divergence. After the learning process, deep local minima of the energy function are supposed to fall close to the datapoints. Therefore, the analysis of learning in RBMs consists of two tasks: understanding the complex landscape of the energy in the vicinity of the datapoints at given weights and devising good optimisation algorithms to fit the data. The investigation of each of these steps is a true mathematical challenge. In these respects pattern retrieval represents a simplified setting to study at first instance, as the roles of the datapoints and of the weights is undertaken by the same objects, the patterns. So, there is no optimisation, but one only has to look at the energy landscape in the vicinity of the patterns.

More precisely, we study the retrieval of i.i.d. binary patterns as the distribution of the hidden layer varies. We do it by looking at the local minima of the energy function, in what is called in statistical physics a zero temperature limit, in which retrieval is maximised (see, e.g., [2]). We prove that the tail of the hidden prior distribution determines the retrieval capability of RBMs. More precisely, for priors with tails ranging from exponential to Gaussian we prove that deep local minima are well localised about the patterns, while if the tails of the hidden prior decay faster than Gaussian, we show that the patterns cannot be retrieved well in any case. RBMs whose hidden priors have Gaussian tails (a class including the Hopfield model) represent special threshold cases which we treat separately in either the positive (Theorem 1.1) and the negative (Theorem 1.2) result below.

1.1. Setting

We consider RBMs with i.i.d. Bernoulli ± 1 visible units $\sigma_1, \dots, \sigma_{N_1}$ and symmetric i.i.d. hidden units $z_1, \dots, z_{N_2}$ distributed according to some prior π. We allow a certain freedom in the choice of the hidden prior π, for which we only require that

$$\pi(|z| \geq t) \simeq e^{-|t|^q} \quad \text{for some } q > 1. \tag{1.2}$$

Let $\xi^{(1)}, \dots, \xi^{(N_2)}$ denote independent random vectors, that we call patterns, with N_1 centred i.i.d. ± 1 components. We consider RBM probability distributions with (unnormalised) density (with respect to the priors)

$$p(\sigma, z; \xi) := \exp\left(\frac{\beta}{N_1^{\frac{1}{q_-}}} \sum_{\mu\in[N_2]} (\sigma, \xi^{(\mu)}) z_\mu\right), \tag{1.3}$$

where $\beta > 0$ is a parameter usually called inverse temperature and $q_- := \min(q, 2)$. We will consider the ratio between the number of visible and hidden units as follows:

$$\alpha := \frac{N_2}{N_1^{\frac{p_+}{2}}}, \tag{1.4}$$

where

$$\frac{1}{p_+} + \frac{1}{q_-} = 1.$$

This is a parameter which will be considered as a constant in the subsequent analysis. The normalisation factors in (1.3) and (1.4) are unusual. For instance, in (1.3), typically from a spin glass perspective one adopts a more familiar normalisation with $\sqrt{N_1}$, while for learning one leaves the energy unnormalised (as the best normalisation is learned with the weights). Our choice ensures that either the energy of the single pattern and the global maximum of (1.3) as $\beta \to \infty$ stay bounded as N_1 grows to infinity and scale linearly with α (with constants depending on q). The aim of Section 2 is to make this point more precise.

Integrating out the hidden layer in (1.3), we get a probability distribution over the visible units. Its density is written as

$$\int p(\sigma, z; \xi)\pi(dz_1)\cdots\pi(dz_{N_2}) = \exp\left(\sum_{\mu\in[N_2]} u\left(\frac{\beta}{N_1^{\frac{1}{q}}}(\xi^\mu, \sigma)\right)\right), \tag{1.5}$$

where

$$u(x) := \log E[e^{xz_1}].$$

We are interested in studying the local maxima of the right-hand side of (1.5) as β is very large, but N_1, N_2 finite. The main issue is that the dependency on β in the exponent is not multiplicative and it is not clear which function should be analysed in the limit $\beta \to \infty$ (compare it, for instance, with the easier cases of the Sherrington–Kirkpatrick model [1] or the Hopfield model [20], where β is just a multiplicative parameter).

Exploiting a reduction argument introduced in [4], we show how to single out an effective energy function which captures the RBM landscape at zero temperature. The following simple observation starts our considerations: for any z_1 such that (1.2) holds for some $q > 1$, we have

$$c(q)\|z_1\|_{\psi_q}^p \le \lim_{x\to\infty} \frac{u(x)}{|x|^p} \le C(q)\|z_1\|_{\psi_q}^p, \tag{1.6}$$

where $0 < c(q) \leq C(q) < \infty$ are universal constants depending only on q and p is the Hölder conjugate exponent of q. For a definition of the Orlicz norms $\|\cdot\|_{\psi_q}$, see (1.11) below. The proof of (1.6) is immediate, we write

$$
\begin{aligned}
E[e^{xz_1}] &= \int_0^\infty d\lambda P(z_1 \geq x^{-1}\log\lambda) \simeq \int_0^\infty d\lambda e^{-\frac{|\log\lambda|^q}{\|z_1\|_{\psi_q}^q |x|^q}} \\
&= \|z_1\|_{\psi_q}|x| \int_0^\infty d\lambda e^{\lambda\|z_1\|_{\psi_q}|x| - |\lambda|^q} \\
&= \|z_1\|_{\psi_q}|x| e^{\frac{\|z_1\|_{\psi_q}^p |x|^p}{p}} \int_0^\infty d\lambda e^{\lambda\|z_1\|_{\psi_q} x - |\lambda|^q - \frac{\|z_1\|_{\psi_q}^p |x|^p}{p}},
\end{aligned}
$$

and the last integral is finite uniformly in x by the Young inequality. Taking the log on both sides and passing to the limit, we get (1.6).

Thus, by (1.6), as $\beta \to \infty$, we are led to consider the following p-spin energy function [7, 12]:

$$
H^{(p)}(\sigma;\xi) := -\frac{1}{N_1^{\kappa(p)}} \sum_{\mu\in[N_2]} |(\xi^{(\mu)},\sigma)|^p, \quad \kappa(p) := 1 + p - \frac{p}{p_+} \tag{1.7}
$$

(here we include the usual normalisation factor $1/N_1$ of the internal energy directly in the definition of $H^{(p)}$). To fix the ideas, $\kappa(p) = p$ for $p \geq 2$ and $\kappa(p) = 1 + \frac{p}{2}$ for $p \leq 2$.

1.2. Main results

The focus of this paper is to study the location of the minima of (1.7) on $\{-1,1\}^{N_1}$ close to the pattern configurations in the limit $N_1, N_2 \to \infty$, while α remains constant. Hence, the main objects of interest will be the following two sets:

$$
\begin{aligned}
\mathtt{LM}_{N_1} &:= \{\text{local minima of (1.7)}\}, \\
\mathtt{dLM}_{N_1}^{(\mu)} &:= \{\text{local minima } \bar\sigma \text{ of (1.7) such that } H^{(p)}(\xi^{(\mu)};\xi) - H^{(p)}(\bar\sigma;\xi) > 0\}.
\end{aligned}
$$

Below, Hamming(a, b) denotes the Hamming distance between a, b (i.e., the number of different entries) and $\widehat{B}_{\mu,R}^{N_1}$ is the ball in this metric centred at the μth pattern with radius R. Throughout the paper, we will repeatedly use that two patterns are typically separated by $N_1/2$ flips so that $\widehat{B}_{\mu,\lfloor N_1/2\rfloor}^{N_1}$ and $\widehat{B}_{\mu',\lfloor N_1/2\rfloor}^{N_1}$ typically do not overlap. We say that the event A occurs with high probability (w.h.p.) if for all $x > 0$, for all sufficiently large N_1 it holds $P(A) \geq 1 - N_1^{-x}$. $S : [0,1] \mapsto \mathbb{R}$ denotes the coin tossing entropy

$$
S(r) := -r\log r - (1-r)\log(1-r). \tag{1.8}
$$

Our first result states that the error of reconstructing a given pattern is very small in terms of the number of visible units N_1 if the decay of the hidden prior (1.2) is slower than Gaussian, while a finite fraction of bits cannot be retrieved for $q = 2$.

Theorem 1.1. *Let $q \in (1,2)$. There exists $r_0 \in (0, \frac{1}{2}]$ such that w.h.p.*

$$\max_{\mu\in[N_2]} \max_{\sigma\in \mathtt{dLM}_{N_1}^{(\mu)}\cap \widehat{B}^{N_1}_{\mu,\lfloor r_0 N_1\rfloor}} \mathrm{Hamming}(\sigma,\xi^{(\mu)}) \le (\log N_1)^{\frac{q-1}{2-q}}. \tag{1.9}$$

Let $q = 2$. For any $r \in (0, 3/8)$ if $\alpha < \min\big(\frac{1}{3}\sqrt{\frac{r}{1-r}}, \frac{\sqrt{r}}{25S(r)}\big)$ then w.h.p.

$$\max_{\mu\in[N_2]} \max_{\sigma\in \mathtt{dLM}_{N_1}^{(\mu)}\cap \widehat{B}^{N_1}_{\mu,\lfloor 3N_1/8\rfloor}} \mathrm{Hamming}(\sigma,\xi^{(\mu)}) \le rN_1.$$

Theorem 1.1 identifies the models with $q < 2$ as excellent in pattern reconstruction: there are deep minima located few flips away from the patterns (in fact, polylog flips, see (1.9)) and no deeper minima appear in an extended region. We observe that albeit we formulate Theorem 1.1 in terms of local minima, we proved a stronger statement regarding all points in a Hamming ball about the pattern. Namely, we show that exploring all the points in a large Hamming ball centred at any pattern, to find a point with lower energy we need to go very close to the centre.

We do not attempt here at precisely characterising the radius r_0, the basin of attraction of the patterns, the maximal α allowing retrieval or any of the constants in the play. Indeed, the numbers appearing in the case $q = 2$ of the above Theorem carry no special meaning.

Local minima are not directly related to the typical configurations of (1.3). However, it is well known that any algorithmic search of typical configurations will finish to find a hopefully representative local minimum. This can be done by the usual FLIP algorithm, that is greedy flipping of one units at time decreasing the energy until no more decreasing is possible. Therefore, $\mathtt{dLM}_{N_1}^{(\mu)}$ has a direct interpretation in terms of retrieval. Take, for instance, $q < 2$. By the proof of Theorem 1.1, it follows that any FLIP search initialised, for instance, at $\xi^{(\mu)}$ will end up in a point of $\mathtt{dLM}_{N_1}^{(\mu)}$ falling only $(\log N_1)^{\frac{q-1}{2-q}}$ flips away from the pattern, which means that only few bits are misretrieved.

Somewhat in the opposite direction, the next result shows that for $q > 2$ in (1.2), the local minima of (1.7) are quite far from the patterns.

Theorem 1.2. *Let $q \ge 2$, $r \in [0, \frac{1}{2}]$ and let $\alpha_q(r) := S(r)$ for $q \neq 2$ and $\alpha_2(r) := S(r)/(1-2r)^2$. There is a numerical constant $f(q) > 0$ such that every $r \in (0, \frac{1}{2})$ and for all $\alpha \ge f(q)\alpha_q(r)$ we have w.h.p.*

$$\mathrm{Hamming}\big(\mathtt{LM}_{N_1}, \xi^{(\mu)}\big) \ge \lfloor rN_1\rfloor. \tag{1.10}$$

For sake of brevity, the value of the numerical constant $f(q)$ is not specified in the statement of the previous theorem, but can be determined following its proof. Again, we stress that we did not aim at optimising the constants.

According to Theorem 1.2, if $q > 2$, one could still hope for retrieval with a very small amount of hidden variables, i.e., for α small enough (indeed for $\alpha = 0$ reconstruction is possible, see [4]), but for α larger than a given constant no recovery is allowed. For $q = 2$, the situation improves a bit in the sense that pattern reconstruction becomes less and less efficient as α grows.

The paper [4] showed that RBMs with hidden prior interpolating between a Gaussian and a bimodal symmetric distribution exhibit retrieval at finite $\alpha > 0$, which disappears in the degenerate case when the Gaussian part is switched off. It is also argued that such a lack of retrieval should persist at least for any compactly supported hidden prior. This is demonstrated using non-rigorous replica computations and numerics. We give here the first mathematical validation of these findings, as Theorem 1.1 (for $q = 2$) implies pattern retrieval if in the interpolating prior the Gaussian part is present, whatever small, and Theorem 1.2 is a strong indication for lack of retrieval for hidden prior with a Bernoulli ± 1 distribution (for which we should read $q = \infty$).

1.3. Related literature

The results presented here mark a neat difference in the retrieval capabilities of RBMs with hidden priors (1.2) with $q < 2$ (very good capabilities) and $q > 2$ (not so good) with a transition at the Gaussian tail case $q = 2$. As already remarked, a notable instance of the case $q = 2$ is the Hopfield model, for which a similar analysis at zero temperature was done in [20] (analog of Theorem 1.1), [19] (analog of Theorem 1.2) and [25] in the attempt of proving the picture of [2]. When comparing these papers to ours, we underline that we do not seek to characterise any of our estimates with the best possible constants, which was instead a relevant component of all these previous papers. In particular, by Theorem 1.1, it follows that in the case $q = 2$ we observe retrieval for $\alpha \leq 0.04$, much less than the threshold $\alpha \leq 0.14$ computed by Amit, Gutfreund, and Sompolinsky. However, from our analysis, it is clear that this critical threshold is not a specific of the Hopfield model, but it can be achieved universally for all the models whose hidden priors have Gaussian tails.

We exploit and make mathematically precise the heuristics of [4]. Namely, we use that the tail of the hidden prior determines the behaviour for large argument of the energy function of the associative network (around zero it is always quadratic). It is exactly this asymptotic that governs retrieval: the more convex the better. Mathematically speaking the introduction of the hidden layer is a way to linearise the energy function (over the visible units) and different prior distributions for the hidden layer correspond to different associative networks. Similar ideas have been used by [4, 9, 22, 26, 27] to study the performance of the RBMs with varying hidden unit statistics.

The FLIP algorithm is a very natural choice to explore the energy landscape of RBMs and indeed, it is essentially the original network dynamics proposed in [16]. This gives a nice connection with the local max-cut problem as analysed, for instance, in [11] and [3], even though here we exploit crucially the presence of the patterns, which constitute a special class of local minima. This is even more clear by comparing with the analysis for the Sherrington–Kirkpatrick model of [1].

Many other dynamics have been proposed alternative to the FLIP algorithm mainly for the Hopfield model and we will not give here an account on that (see the recent work [6] and the references therein). We just mention that the dynamics analysed in [5, 27], which is a zero-temperature version of the alternate Gibbs sampling typically used to train RBMs, is in spirit very close to our zero-temperature reduction.

1.4. Notations

Throughout the paper, $p, q \geq 1$ will always be Hölder conjugates, that is, $\frac{1}{p} + \frac{1}{q} = 1$, and similarly for q_-, p_+, with $q_- := \min(2, q)$, $p_+ := \max(2, p)$. C, c everywhere denote positive absolute constants which may change from formula to formula. We write $X \lesssim Y$ if $X \leq CY$ and $X \simeq Y$ if $Y \lesssim X \lesssim Y$. Sometimes, we write $\simeq_a$ or $\lesssim_a$ to stress the dependence of the constants C above on a parameter a. We indicate by $(\cdot,\cdot)$ the inner product in $\mathbb{R}^{N_1}$ or $\mathbb{R}^{N_2}$ and the meaning will be always clear from the context and by $\|\cdot\|_p$ the ℓ_p-norms. 1 may represent the vector in $\mathbb{R}^{N_1}$ or in $\mathbb{R}^{N_2}$ with all entries equal to 1. $B_N^{(q)}$ is the ℓ_q centred ball of radius one in $\mathbb{R}^N$. $\widehat{S}_{\mu,R}^{N_1-1}$, $\widehat{B}_{\mu,R}^{N_1}$ denote, respectively, the N_1-dimensional Hamming sphere and ball centred at $\xi^{(\mu)}$ of radius R. If $v \in \mathbb{R}^N$ and $J \subset [N]$, we denote by v_J a vector in $\mathbb{R}^{|J|}$ such that $(v_J)_i = v_{j_i}$ if $J = \{j_1, \dots, j_{|J|}\}$. To any $J \subset [N]$ we also associate a FLIP operator F_J defined by $(F_J v)_i = -v_i$ if $i \in J$ and $(F_J v)_i = v_i$ if $i \notin J$. We will use the following Orlicz norms:

$$\|Z\|_{\psi_r} := \inf\left\{\lambda > 0 : E\left[\psi_r\left(\frac{|Z|}{\lambda}\right)\right] < 2\right\}, \quad r > 0, \tag{1.11}$$

where $\psi_r(x) = e^{x^r}$ for any $x > 0$ for $r \geq 1$, while for $r \in (0, 1)$ there are $c(r), x(r)$ such that for $x \in (0, x(r))$ it is $\psi_r(x) = c(r)x$. We underline that, setting $q_s := \sup\{q' > 1 : \|Z\|_{\psi_{q'}} < \infty\}$, we have $P(|Z| \geq t) \simeq e^{-|t|_s^q}$ (we convey that bounded random variables have finite ψ_∞-norm). Bearing in mind the definition (1.8), we will often use the standard bound for $r \in [0, 1]$:

$$\operatorname{Card} \widehat{S}_{\mu,\lfloor rN_1\rfloor}^{N_1-1} = \binom{N_1}{\lfloor rN_1\rfloor} \leq e^{N_1 S(r)}. \tag{1.12}$$

We denote the transpose patterns $\tilde{\xi}^{(i)}$ by $\tilde{\xi}^{(i)}_\mu := \xi^{(\mu)}_i$, $i \in [N_1]$, $\mu \in [N_2]$. Sometimes, we write $\hat{\xi} := \xi/\sqrt{N_1}$. A^c is the complement of the set A. We say that the event A occurs with high probability (w.h.p.) if for all $x > 0$, for all sufficiently large N_1, it holds $P(A) \geq 1 - N_1^{-x}$.

2. Zero temperature reduction

In this section, which is in part independent on the rest of the paper, we study some interesting properties of the Hamiltonian (1.7).

First, we show that the single pattern energy is close to the ground state, so providing a motivation for the normalisation factors in (1.3) and (1.4). We give a lower bound for the ground state energy linear in α. To do so, we do not actually need binary patterns.

Proposition 2.1. *Let $\xi^{(1)}, \dots, \xi^{(N_2)}$ be independent vectors in $\mathbb{R}^{N_1}$ with i.i.d. centered sub-Gaussian entries. It holds for $p \geq 1$ that*

$$\inf_{\sigma \in \{-1,1\}^{N_1}} H^{(p)}(\sigma; \xi) \gtrsim_p -(1+\alpha) \tag{2.1}$$

with probability larger than $1 - e^{-c\alpha^{\frac{2}{p}} N_1}$.

To prove Proposition 2.1, we need the following auxiliary lemma.

Lemma 2.1. *Let $\xi^{(1)}, \dots, \xi^{(N_2)}$ be independent vectors in $\mathbb{R}^{N_1}$ with i.i.d. centred sub-Gaussian entries. Let $p \geq 1$. For all $t \gtrsim_p (1+\alpha)^{\frac{1}{p}}$, we have*

$$P\left(\frac{1}{N_1^{\max(0, \frac{q-2}{2q})}} \sup_{\sigma \in \frac{1}{\sqrt{N_1}}\{-1,1\}^{N_1}} \sup_{\tau \in B^{(q)}_{N_2}} \frac{(\xi^{(\mu)}, \sigma)\tau_\mu}{\sqrt{N_1}} \geq t \right) \leq 2e^{-ct^2 N_1},$$

where $c > 0$ depends only on the distribution of $\xi^{(1)}_1$.

Proof. We introduce the transpose patterns $\tilde{\xi}^{(i)}$ by $\tilde{\xi}^{(i)}_\mu := \xi^{(\mu)}_i$, $i \in [N_1]$, $\mu \in [N_2]$. First of all, we note that

$$\frac{1}{N_1^{\max(0, \frac{q-2}{2q})}} \sup_{\sigma \in \frac{1}{\sqrt{N}}\{-1,1\}^{N_1}} \sum_{i \in [N_1]} \sum_{\mu \in [N_2]} \frac{\xi^{(\mu)}_i \sigma_i \tau_\mu}{\sqrt{N_1}} = \frac{1}{N_1} \sum_{i \in [N_1]} \frac{|(\tilde{\xi}^{(i)}, \tau)|}{N_1^{\max(0, \frac{q-2}{2q})}}. \tag{2.2}$$

Moreover, since for all $\tau \in B^{(q)}_{N_2}$

$$\|\tau\|_2 \leq N_2^{\max(0, \frac{q-2}{2q})},$$

we have

$$P\left(\frac{1}{N_1^{\max(0,\frac{q-2}{2q})}}|(\tilde{\xi}^{(i)},\tau)| \geq t\right) \leq 2e^{-\frac{t^2 N_1^{\max(0,\frac{q-2}{q})}}{2N_2^{\max(0,\frac{q-2}{q})}\|\xi_1^{(1)}\|_{\psi_2}^2}}.$$

The right-hand side of (2.2) is the sum of independent sub-Gaussian random variables with

$$E[|(\tilde{\xi}^{(1)},\tau)|] \leq \sqrt{E[|(\tilde{\xi}^{(1)},\tau)|^2]} \lesssim N_2^{\max(0,\frac{q-2}{2q})},$$

thus for all $\tau \in B_{N_2}^{(q)}$ for $t \gtrsim 1$, we have

$$P\left(\frac{1}{N_1}\sum_{i\in[N_1]}\frac{|(\tilde{\xi}^{(i)},\tau)|}{N_1^{\max(0,\frac{q-2}{2q})}} \geq t\right) \leq 2e^{-\frac{t^2 N_1^{1+\max(0,\frac{q-2}{q})}}{2N_2^{\max(0,\frac{q-2}{q})}\|\xi_1^{(1)}\|_{\psi_2}^2}}$$

for some $c > 0$ depending only on the distribution of $\xi_1^{(1)}$.

Next, we cover $B_{N_2}^{(q)}$ with a number of balls in $\mathbb{R}^{N_2}$ with some small radius $\varepsilon > 0$. For $p \geq 2$, we can use Euclidean balls, and the Sudakov inequality gives a bound on the minimal number $N(B_q^{N_2}, \varepsilon B_2^{N_2})$ of such balls

$$N(B_q^{N_2}, \varepsilon B_2^{N_2}) \leq e^{cN_2^{\frac{2}{p}}}$$

(here we used that for a Gaussian vector g, $E[\max_{\tau\in B_q^{N_2}}(\tau, g)] = E[\|g\|_p] \simeq N_2^{\frac{1}{p}}$).

For $p \in (1,2)$ we use ℓ_q-balls and we have

$$N(B_q^{N_2}, \varepsilon B_q^{N_2}) \leq e^{cN_2}$$

(in the two estimates above the constants c depends on ε in a way we do not keep track of).

Assume now $(1+\alpha)^{\frac{1}{p}} \lesssim_p t$ (this is to take into account also the behaviour for small α). By the union bound for $p \geq 2$, we get

$$P\left(\frac{1}{N_1^{\max(0,\frac{q-2}{2q})}}\sup_{\sigma\in\frac{1}{\sqrt{N}}\{-1,1\}^{N_1}}\sup_{\tau\in B_{N_2}^{(q)}}(\hat{\xi}\sigma,\tau) \geq t\right) \leq 2e^{cN_2^{\frac{2}{p}}-cN_1t^2} \leq 2e^{-ct^2N_1}.$$

Similarly, for $p \in (1,2)$,

$$P\left(\frac{1}{N_1^{\frac{q-2}{2q}}}\sup_{\sigma\in\frac{1}{\sqrt{N}}\{-1,1\}^{N_1}}\sup_{\tau\in B_{N_2}^{(q)}}(\hat{\xi}\sigma,\tau) \geq t\right) \leq 2e^{cN_2-c\left(\frac{N_1}{N_2}\right)^{\frac{q-2}{q}}N_1t^2} \leq 2e^{-ct^2N_1}.$$

■

Proof of Proposition 2.1. The role of hidden variables at zero temperature is played by duality:

$$\sum_{\mu\in[N_2]}|(\xi^{(\mu)},\sigma)|^p=\left|\sup_{\tau\in B^{(q)}_{N_2}}\sum_{\mu\in[N_2]}(\xi^{(\mu)},\sigma)\tau_\mu\right|^p,\quad p\ge 1.$$

Therefore (here we shorten $\hat{\xi}:=\xi/\sqrt{N_1}$),

$$\begin{aligned}\inf_{\sigma\in\{-1,1\}^{N_1}}H^{(p)}(\sigma;\xi)&=-\sup_{\sigma\in\frac{1}{\sqrt{N_1}}\{-1,1\}^{N_1}}\left|\frac{1}{N_1^{\max(0,\frac{q-2}{2q})}}\sup_{\tau\in B^{(q)}_{N_2}}\sum_{\mu\in[N_2]}(\hat{\xi}^{(\mu)},\sigma)\tau_\mu\right|^p\\&=-\left|\frac{1}{N_1^{\max(0,\frac{q-2}{2q})}}\sup_{\sigma\in\frac{1}{\sqrt{N_1}}\{-1,1\}^{N_1}}\sup_{\tau\in B^{(q)}_{N_2}}\sum_{\mu\in[N_2]}(\hat{\xi}^{(\mu)},\sigma)\tau_\mu\right|^p\end{aligned}\tag{2.3}$$

by symmetry. It suffices to focus on the quantity inside the modulus above, which is dealt in Lemma 2.1. We have for all $t\gtrsim_p(1+\alpha)^{\frac{1}{p}}$

$$P\left(\frac{1}{N_1^{\max(0,\frac{q-2}{2q})}}\sup_{\sigma\in\frac{1}{\sqrt{N}}\{-1,1\}^{N_1}}\sup_{\tau\in B^{(q)}_{N_2}}(\hat{\xi}\sigma,\tau)\ge t\right)\le 2e^{-ct^2N_1}.$$

Combining (2.3) with the bound above we obtain the statement. ■

Now, we show that the patterns have energy of the same order in α of the global minimum, even though we can already observe a difference between the models with $p\ge 2$ and $p<2$. We deal with ±1 binary patterns for simplicity, but a similar argument can be easily repeated for symmetric patterns with minor modifications. We have

$$H^{(p)}(\xi^{(1)};\xi)=-\frac{\|\xi^{(1)}\|_2^p}{N_1^{1+p-p/p_+}}-\sum_{j=2}^{N_2}\frac{|(\xi^{(1)},\xi^{(j)})|^p}{N_1^{1+p-p/p_+}}.$$

This quantity concentrates around its average as N_1 grows.

Lemma 2.2. *Take $t>0$ uniformly in N_1, small enough. It holds*

$$P\left(\left|H^{(p)}(\xi^{(1)};\xi)-E[H^{(p)}(\xi^{(1)};\xi)]\right|\ge t\right)\le\begin{cases}\exp(-cN_1t^2), & p\in(1,2],\\ \exp(-cN_1t^{\frac{2}{p}}), & p>2.\end{cases}$$

Proof. We write

$$H^{(p)}(\xi^{(1)};\xi)=-\frac{\|\xi^{(1)}\|_2^p}{N_1^{1+p-p/p_+}}-\sum_{\mu=2}^{N_2}\frac{|(\xi^{(1)},\xi^{(\mu)})|^p}{N_1^{1+p-p/p_+}}.$$

It suffices to focus on the second summand on the right-hand side above. We have by independence

$$
\begin{aligned}
&P\left(\left|\sum_{\mu=2}^{N_2}\frac{|(\xi^{(1)},\xi^{(\mu)})|^p}{N_1^{\frac{p}{2}}}-E\left[\frac{|(\xi^{(1)},\xi^{(\mu)})|^p}{N_1^{\frac{p}{2}}}\right]\right|\geq tN_1^{1+p/2-p/p+}\right)\\
&\quad=P\left(\left|\sum_{\mu=2}^{N_2}\frac{|(1,\xi^{(\mu)})|^p}{N_1^{\frac{p}{2}}}-E\left[\frac{|(1,\xi^{(\mu)})|^p}{N_1^{\frac{p}{2}}}\right]\right|\geq tN_1^{1+p/2-p/p+}\right), \qquad (2.4)
\end{aligned}
$$

where 1 is the constant vector with all entries equal to 1. The random variables

$$
T_p^{(\mu)}:=\frac{|(1,\xi^{(\mu)})|^p}{N_1^{\frac{p}{2}}}-E\left[\frac{|(1,\xi^{(\mu)})|^p}{N_1^{\frac{p}{2}}}\right],\quad \mu=2,\dots,N_2
$$

are i.i.d. with

$$
\|T_p^{(\mu)}\|_{\psi_{\frac{2}{p}}}\simeq 1.
$$

Thus, by Proposition A.1 with $\ell=2/p$, we have

$$
(2.4)\leq\begin{cases}\exp\left(-cN_1\min(t^2,t^{\frac{2}{p}})\right), & p\in(1,2],\\ \exp\left(-cN_1\min(t^2N_1^{p-2},t^{\frac{2}{p}})\right), & p>2,\end{cases}
$$

and the proof is complete. ■

Moreover, we have

$$
|E[H^{(p)}(\xi^{(1)};\xi)]|\simeq-\alpha\frac{p}{2}\Gamma\left(\frac{p}{2}\right)-\frac{1}{N_1^{1-p/p+}}.
$$

In fact,

$$
\begin{aligned}
E[|(\xi^{(1)},\xi^{(\mu)})|^p]&=E\left[\int d\lambda P(\xi^{(\mu)}:\left|(\xi^{(1)},\xi^{(\mu)})\right|\geq\lambda^{1/p})\right]\\
&\simeq E[\|\xi^{(1)}\|_1^p]\int_0^{\infty}d\lambda e^{-\lambda^{\frac{2}{p}}}=\frac{p}{2}\Gamma\left(\frac{p}{2}\right)N_1^{\frac{p}{2}}. \qquad (2.5)
\end{aligned}
$$

Therefore,

$$
H^{(p)}(\xi^{(1)};\xi)\simeq-\frac{p}{2}\Gamma\left(\frac{p}{2}\right)\alpha-\frac{1}{N_1^{1-p/p+}}
$$

with very high probability. We see that if $p\geq 2$ this value is really of the same order of the ground state, while if $p\in(1,2)$ for α small and N_1 large the patterns have higher energy.

3. Retrieval for $p \geq 2$

In this section, we prove Theorem 1.1. We look at all configurations reachable from $\xi^{(\mu)}$ by $\lfloor rN_1 \rfloor$ flips and compare the energy of the pattern with the minimal energy of such configurations. By symmetry of the patterns, we can reduce to look at $\mu = 1$ and we may and will assume that 1, i.e., the vector with entries all equal to one, lies in $\widehat{S}^{N_1-1}_{1,\lfloor rN_1 \rfloor}$.

Without further explanation, we introduce some more notations. For any point $\sigma \in \{-1,1\}^{N_1}$ and subset of indices $J \subseteq [N_1]$, we set

$$X_J^{(\mu)}(\sigma) := \frac{1}{\sqrt{N_1}}(\xi_J^{(\mu)}, \sigma_J), \quad Y_J^{(\mu)}(\sigma) := \frac{1}{\sqrt{N_1}}(\xi_{J^c}^{(\mu)}, \sigma_{J^c}) \tag{3.1}$$

and $X_J^{(\mu)}(1) =: X_J^{(\mu)}$, $Y_J^{(\mu)}(1) =: Y_J^{(\mu)}$. We conveniently let

$$\Phi_p(x,y) := |x+y|^p - |x-y|^p, \quad \bar{\Phi}_p(r) := \Phi_p(r, 1-r) = 1-(1-2r)^p > 0 \tag{3.2}$$

(recall that we consider $r \in (0, 1/2)$). We have the following useful representation (recall the definition of the FLIP operator F_J in Section 1.4).

Lemma 3.1. *Let $r \in (0, \frac{1}{2})$, $J \subset [N_1]$ with $|J| = \lfloor rN_1 \rfloor$. It is*

$$\begin{aligned} &H^{(p)}(\xi^{(1)}) - H^{(p)}(F_J \xi^{(1)}) \\ &\quad = -\frac{1}{N_1^{\frac{p_+ - p}{2}}} \bar{\Phi}_p(r) - \frac{1}{N_1^{\frac{p_+}{2}}} \sum_{\mu=2}^{N_2} \Phi_p(X_J^{(\mu)}(\xi^{(1)}), Y_J^{(\mu)}(\xi^{(1)})). \end{aligned} \tag{3.3}$$

Proof. Compute

$$\begin{aligned} &H^{(p)}(\sigma) - H^{(p)}(F_J\sigma) \\ &\quad = -\frac{1}{N_1^{\kappa(p)}} \sum_{\mu \in [N_2]} \left(\left| (\xi_J^{(\mu)}, \sigma_J) + (\xi_{J^c}^{(\mu)}, \sigma_{J^c}) \right|^p - \left| -(\xi_J^{(\mu)}, \sigma_J) + (\xi_{J^c}^{(\mu)}, \sigma_{J^c}) \right|^p \right) \\ &\quad = -\frac{1}{N_1^{\kappa(p) - \frac{p}{2}}} \sum_{\mu \in [N_2]} \Phi_p(X_J^{(\mu)}(\sigma), Y_J^{(\mu)}(\sigma)) \end{aligned}$$

by the definitions (3.1), (3.2). We have $\kappa(p) - \frac{p}{2} = 1 + \frac{p}{2} - \frac{p}{p_+} = \frac{p_+}{2}$. Take now $v = \xi^{(1)}$. An easy computation gives

$$\begin{aligned} X_J^{(1)}(\xi^{(1)}) &= \frac{\|\xi_J^{(1)}\|_2^2}{\sqrt{N_1}} = \frac{|J|}{\sqrt{N_1}} = r\sqrt{N_1}, \\ Y_J^{(\mu)}(\xi^{(1)}) &= \frac{\|\xi_{J^c}^{(1)}\|_2^2}{\sqrt{N_1}} = \frac{|J^c|}{\sqrt{N_1}} = (1-r)\sqrt{N_1}. \end{aligned}$$

Thus,

$$\frac{1}{N_1^{\frac{p_+}{2}}}\Phi_p(X_J^{(1)}(\xi^{(1)}), Y_J^{(1)}(\xi^{(1)})) = \frac{1}{N_1^{\frac{p_+-p}{2}}}\Phi_p(r, 1-r),$$

and (3.3) follows. ∎

The necessary tail estimates in order to prove Theorem 1.1 are given in the next lemmas.

Lemma 3.2. *Let* $r \in (0, \frac{1}{2})$, $J \subset [N_1]$ *with*

$$|J| = \lfloor rN_1 \rfloor,$$

$\{X_J^{(\mu)}\}_{\mu\in[N_2]\setminus\{1\}}$ *and* $\{Y_J^{(\mu)}\}_{\mu\in[N_2]\setminus\{1\}}$ *are independent sub-Gaussian random variables, independent one from each other, with*

$$\|X_J^{(\mu)}\|_{\psi_2} \le \sqrt{\frac{3r}{2}}, \quad \|Y_J^{(\mu)}\|_{\psi_2} \le \sqrt{\frac{3(1-r)}{2}}. \tag{3.4}$$

Moreover, $\{\Phi_p(X_J^{(\mu)}, Y_J^{(\mu)})\}_{\mu\in[N_2]}$ *are i.i.d.* $\psi_{2/p}$ *r.vs. with*

$$\left\|\Phi_p(X_J^{(\mu)}, Y_J^{(\mu)})\right\|_{\psi_{2/p}}^{\frac{2}{p}} \le 3\sqrt{r(1-r)}. \tag{3.5}$$

Proof. The proof of (3.4) is standard. We proceed only for $X^{(\mu)}$, as for $Y^{(\mu)}$ is similar. We set

$$\tilde\lambda := \lambda\sqrt{N_1/2\lfloor rN_1\rfloor} \quad \text{and} \quad \tilde X^{(\mu)} := (\xi_J^{(\mu)}, 1)/\sqrt{\lfloor rN_1\rfloor}.$$

We let also $g \sim \mathcal{N}(0,1)$ and $\bar\xi$ be a symmetric Bernoulli ± 1 variable, whose expectation values are denoted by E_g and $E_{\bar\xi}$. We have

$$\begin{aligned} E\Big[e^{\frac{|X_J^{(\mu)}|^2}{\lambda^2}}\Big] &= E\Big[e^{\frac{|\tilde X^{(\mu)}|^2}{\tilde\lambda^2}}\Big] = EE_g\Big[e^{\frac{g\tilde X^{(\mu)}}{\tilde\lambda}}\Big] = E_g\Big[\Big(E_{\bar\xi}\Big[e^{\frac{g\bar\xi}{\tilde\lambda\sqrt{\lfloor rN_1\rfloor}}}\Big]\Big)^{\lfloor rN_1\rfloor}\Big] \\ &= E_g\Big[e^{\lfloor rN_1\rfloor \log\cosh(\frac{g}{\tilde\lambda\sqrt{\lfloor rN_1\rfloor}})}\Big] \le E_g[e^{\frac{g^2}{2\tilde\lambda^2}}] = (1-\tilde\lambda^2)^{-\frac{1}{2}}. \end{aligned}$$

Since

$$(1-\tilde\lambda^2)^{-\frac{1}{2}} = \left(1 - \frac{1}{2}\lambda^2\frac{N_1}{\lfloor rN_1\rfloor}\right)^{-\frac{1}{2}} < 2$$

for $\lambda < \sqrt{\frac{3r}{2}}$ we recover the first one of (3.4).

To prove (3.5), we bound

$$E\Big[e^{(\frac{\Phi_p}{t})^{\frac{2}{p}}}\Big] \le E\Big[e^{\frac{(|X_J^{(\mu)}|+|Y_J^{(\mu)}|)^2}{t^{\frac{2}{p}}}}\Big] \le \frac{1}{2}E\Big[e^{\frac{2|X_J^{(\mu)}|^2}{t^{\frac{2}{p}}}}\Big] + \frac{1}{2}E\Big[e^{\frac{2|Y_J^{(\mu)}|^2}{t^{\frac{2}{p}}}}\Big],$$

whence

$$\left\|\Phi_p(X_J^{(\mu)}, Y_J^{(\mu)})\right\|_{\psi_{2/p}}^{\frac{2}{p}} \le 2\|X_J^{(\mu)}\|_{\psi_2}\|Y_J^{(\mu)}\|_{\psi_2}.$$ ∎

We shorten in the next statement $P_\xi(\cdot) = P(\cdot \mid \xi)$.

Lemma 3.3. *Let $r \in (0, \frac{1}{2})$, $p \geq 2$, $t = t(r) := 1 - (1-2r)^p - r^{\frac{p}{2}}$. Take any $\sigma \in \widehat{S}^{N_1-1}_{1,\lfloor rN_1 \rfloor}$. If*

$$\alpha \geq 3^{1-p} N_1^{\frac{p-2}{2}} \left(\frac{r}{1-r} \right)^{\frac{p-1}{2}}, \tag{3.6}$$

then

$$P_{\xi^{(1)}}\big(H^{(p)}(\xi^{(1)}) - H^{(p)}(\sigma) \geq -t\big) \leq \exp\left(- \frac{N_1^{\frac{p}{2}}}{24\alpha} \left(\frac{r}{1-r} \right)^{\frac{p}{2}} \right), \tag{3.7}$$

and otherwise,

$$P_{\xi^{(1)}}\big(H^{(p)}(\xi^{(1)}) - H^{(p)}(\sigma) \geq -t\big) \leq \exp\left(- \frac{1}{24} N_1 \sqrt{\frac{r}{1-r}} \right). \tag{3.8}$$

Remark 3.1. Thinking of N_1 very large, with an abuse of notation, we will say in the sequel that a property occurs for all $\alpha > 0$ in case it does for all $\alpha \gtrsim N_1^{-x}$ for some $x > 0$. Therefore, if $r > 0$ uniformly in N_1, i.e., we flip a number of bits proportional to N_1, we have for $p > 2$ the tail (3.8) for all $\alpha > 0$ and for $\alpha \lesssim \sqrt{r}$ for $p = 2$. A sub-linear number of flips corresponds to take $r \simeq N_1^{-x}$ for some $x \in [0, 1]$ (modulo log-corrections, see below). In this case, we see that if $x < \frac{p-2}{p-1}$, the estimate (3.8) still holds for any $\alpha > 0$, while otherwise we have (3.7).

Proof. In Lemma B.1, it is proven $t(r) > 0$ for any $r \in (0, 1/2)$. It is clear that any $\sigma \in \widehat{S}^{N_1-1}_{1,\lfloor rN_1 \rfloor}$ can be written as $F_J \xi^{(1)}$ for some index set J of $\lfloor rN_1 \rfloor$ elements (indeed $J = \{i \in [N_1] : \sigma_i \neq \xi^{(1)}_i\}$). Then, by (3.3), we have

$$\begin{aligned} &P_{\xi^{(1)}}(H^{(p)}(\xi^{(1)}) - H^{(p)}(\sigma) \geq -t) \\ &\quad = P_{\xi^{(1)}}\left(- \sum_{\mu=2}^{N_2} \Phi_p(X^{(\mu)}_J(\xi^{(1)}), Y^{(\mu)}_J(\xi^{(1)})) \geq N_1^{\frac{p}{2}} (\bar{\Phi}_p(r) - t) \right) \\ &\quad = P_{\xi^{(1)}}\left(- \sum_{\mu=2}^{N_2} \Phi_p(X^{(\mu)}_J, Y^{(\mu)}_J) \geq (N_1 r)^{\frac{p}{2}} \right) \\ &\quad = P\left(- \sum_{\mu=2}^{N_2} \Phi_p(X^{(\mu)}_J, Y^{(\mu)}_J) \geq (N_1 r)^{\frac{p}{2}} \right), \end{aligned} \tag{3.9}$$

because of independence of the patterns and $\Phi_p(r, 1-r) \geq 0$.

Note that $y > 0$ is equivalent to $0 \leq t < \bar{\Phi}_p(r)$. By Lemma 3.2, $\{\Phi_p(X^{(\mu)}_J, Y^{(\mu)}_J)\}_{\mu \in [N_2]}$ are centred i.i.d. r.vs. which fit the assumptions of Proposition A.1 below (with $\ell = 2/p \in (0, 1]$). Therefore,

$$(3.9) \leq \exp\left(-\frac{1}{24} \min\left(\frac{N_1^p r^{\frac{p}{2}}}{3^{p-1} N_2 (1-r)^{\frac{p}{2}}}, \frac{N_1\sqrt{r}}{\sqrt{1-r}} \right)\right).$$

The value of this minimum depends on α. We take the first term if (3.6) is fulfilled; otherwise, we take the second one. ■

Now, we are ready for the main proof.

Proof of Theorem 1.1. We shorten

$$\mathtt{D}_{\mu,N_1}(r_0) := \left\{\sigma \in \widehat{B}^{N_1}_{\mu,\lfloor r_0 N_1\rfloor} : H^{(p)}(\xi^{(\mu)};\xi) \geq H^{(p)}(\sigma;\xi)\right\}, \tag{3.10}$$

and note that since for any $\mu \in [N_2]$

$$\mathtt{dLM}^{(\mu)}_{N_1} \cap \widehat{B}^{N_1}_{\mu,\lfloor r_0 N_1\rfloor} \subseteq \mathtt{D}_{\mu,N_1}(r_0),$$

it is

$$P\left(\forall \mu \in [N_2]\mathtt{dLM}^{(\mu)}_{N_1} \cap \widehat{B}^{N_1}_{\mu,\lfloor r_0 N_1\rfloor} \subseteq \widehat{B}^{N_1}_{\mu,R}\right) \geq P\left(\forall \mu \in [N_2]\mathtt{D}_{\mu,N_1}(r_0) \subseteq \widehat{B}^{N_1}_{\mu,R}\right). \tag{3.11}$$

Let us introduce the sets

$$\mathrm{Bar}_{N_1,N_2,p}(n) := \left\{ \min_{\mu\in[N_2]} \min_{\sigma\in\widehat{S}^{N_1-1}_{\mu,n}} H^{(p)}(\sigma) - H^{(p)}(\xi^{(\mu)}) \geq t(n)\right\}$$

on which the minimal energy gap of n flips from the patterns is a given $t(n)$. Write now for $r \in (0, \frac{1}{2})$, $n = \lfloor rN_1\rfloor$ and $t(n) = t(r) = t$. We take some $r_0 \in (0, 1/2)$ to be specified later. Then, bearing in mind (3.11), the crux is

$$\begin{aligned} P(\forall \mu \in [N_2]\mathtt{D}_{\mu,N_1}(r_0) \subseteq \widehat{B}^{N_1}_{\mu,R}) &\geq P\left(\bigcap_{n=\lfloor R\rfloor}^{\lfloor r_0 N_1\rfloor} \mathrm{Bar}_{N_1,N_2,p}(n)\right) \\ &\geq 1 - \sum_{n=\lfloor R\rfloor}^{\lfloor r_0 N_1\rfloor} P(\mathrm{Bar}^c_{N_1,N_2,p}(n)) \\ &\geq 1 - N_1 \min_{\lfloor R\rfloor\leq n\leq\lfloor r_0 N_1\rfloor} P(\mathrm{Bar}^c_{N_1,N_2,p}(n)). \end{aligned} \tag{3.12}$$

By the standard estimate (1.12) and the union bound, we have

$$\begin{aligned} P(\mathrm{Bar}^c_{N_1,N_2,p}(n)) &= P\left(\min_{\mu\in[N_2]} \min_{\sigma\in\widehat{S}^{N_1-1}_{\mu,n}} H^{(p)}(\sigma) - H^{(p)}(\xi_\mu) \leq t\right) \\ &\leq N_2 \exp(N_1 S(r)) E\left[\sup_{\sigma\in\widehat{S}^{N_1-1}_{1,\lfloor rN_1\rfloor}} P_{\xi^{(1)}}(H^{(p)}(\xi^{(1)}) - H^{(p)}(\sigma) \geq -t)\right]. \end{aligned}$$

The probabilities appearing in the last line are evaluated using Lemma 3.3 with the same choice $t = 1-(1-2r)^p - r^{\frac{p}{2}}$.

Let us first deal with $p > 2$. We take $r_0 \in (0, \frac{1}{2}]$ such that for all $r \in [0, r_0]$ it is $25S(r) \le \sqrt{r/(1-r)}$ and $t(r) = 1-(1-2r)^p - r^{\frac{p}{2}}$ increases. Bearing in mind Remark 3.1, we let $x_p := \frac{p-2}{p-1}$ and consider different regimes. If $n > \lfloor N_1^{1-x_p} \rfloor$ then Lemma 3.3 yields for all $\alpha > 0$

$$\begin{aligned} P(\mathrm{Bar}^c_{N_1,N_2,p}(n)) &\le N_2 \exp\left(N_1\left(S(r) - \frac{\sqrt{r}}{24\sqrt{1-r}}\right)\right) \\ &\le N_2 e^{-c\sqrt{r}N_1} \simeq N_2 e^{-c\sqrt{nN_1}}. \end{aligned} \tag{3.13}$$

Thus,

$$\min_{\lfloor N_1^{1-x_p} \rfloor < n \le \lfloor r_0 N_1 \rfloor} P(\mathrm{Bar}^c_{N_1,N_2,p}(n)) \le N_2 e^{-cN_1^{1-\frac{x_p}{2}}}. \tag{3.14}$$

For $n < \lfloor N_1^{1-x_p} \rfloor$, Lemma 3.3 gives for all $\alpha > 0$

$$P(\mathrm{Bar}^c_{N_1,N_2,p}(n)) \le N_2 e^{N_1 S(r) - \frac{N_1^{\frac{p}{2}} r^{\frac{p}{2}}}{24\alpha}} \le N_2 e^{n|\log N_1| - \frac{n^{\frac{p}{2}}}{24\alpha}}. \tag{3.15}$$

Thus, for all $\varepsilon > 0$ sufficiently small,

$$\min_{\lfloor N_1^{2\varepsilon} \rfloor < n < \lfloor N_1^{1-x_p} \rfloor} P(\mathrm{Bar}^c_{N_1,N_2,p}(n)) \le N_2 e^{-cN_1^{\varepsilon p}}. \tag{3.16}$$

Moreover, by (3.15), we see that also a poly-log number of flips is allowed:

$$\min_{\lfloor (\log N_1)^{\frac{2}{p-2}} \rfloor < n \le \lfloor N_1^{2\varepsilon} \rfloor} P(\mathrm{Bar}^c_{N_1,N_2,p}(n)) \le N_2 e^{-c(\log N_1)^{1+\frac{2}{2-p}}}. \tag{3.17}$$

Finally, we look at $n \simeq N_1^{1-x_p}$. In this case, we have to fix some $\alpha_0 > 0$, and we use for $\alpha \le \alpha_0$ the bound (3.13) and for $\alpha > \alpha_0$ the bound (3.15). We have

$$\min_{n \simeq \lfloor N_1^{1-x_p} \rfloor} P(\mathrm{Bar}^c_{N_1,N_2,p}(n)) \le N_2 e^{-cN_1^{1-\frac{x_p}{2}}}. \tag{3.18}$$

Combining (3.12) with $R = (\log N_1)^{\frac{2}{p-2}}$ and (3.14), (3.16), (3.17), and (3.18), we get

$$P\left(\forall \mu \in [N_2] \mathtt{dLM}^{(\mu)}_{N_1} \cap \widehat{B}^{N_1}_{\mu, \lfloor r_0 N_1 \rfloor} \subseteq \widehat{B}^{N_1}_{\mu, \lfloor (\log N_1)^{\frac{2}{p-2}} \rfloor}\right) \ge 1 - N_1 N_2 e^{-c(\log N_1)^{1+\frac{2}{2-p}}} \tag{3.19}$$

whence (1.9) follows.

Now, we look at $p = 2$ and take $r_0 = \frac{3}{8}$, so that $t(r)$ is increasing for $r \in (0, \frac{3}{8})$ (the number 3/8 carries no special meaning). If $\alpha \le \min\big(\frac{1}{3}\sqrt{\frac{r}{1-r}}, \frac{\sqrt{r}}{25S(r)}\big)$ then Lemma 3.3 gives

$$P(B^c_{N_1,N_2,p}(n)) \le N_2 \exp\left(N_1\left(S(r) - \frac{\sqrt{r}}{24\alpha\sqrt{1-r}}\right)\right) \le N_2 e^{-c\sqrt{nN_1}}.$$

Thus, by (3.12) with $R = \lfloor rN_1 \rfloor$ and $r \in (0, \frac{3}{8})$,

$$P\big(\forall\mu \in [N_2]\mathtt{dLM}^{(\mu)}_{N_1} \cap \widehat{B}^{N_1}_{\mu,\lfloor r_0 N_1\rfloor} \subseteq \widehat{B}^{N_1}_{\mu,\lfloor rN_1\rfloor}\big) \ge 1 - N_1 N_2 e^{-\frac{1}{4}\sqrt{r}N_1}, \tag{3.20}$$

whence the $p = 2$ part of Theorem 1.1 follows. ∎

4. Absence of retrieval for $p \in (1, 2]$

In this section, we present the proof of Theorem 1.2.

We set for brevity for $\mu \in [N_2] \setminus \{1\}$, $p \in (1, 2]$, $k \in [N_1]$, $J \subseteq [N_1]$, $\sigma \in \{-1, 1\}^{N_1}$

$$W^{(\mu)}_{p,k,J}(\sigma) := \frac{2p}{\sqrt{N_1}} \xi^{(\mu)}_k v_k \operatorname{sign}(Z^{(\mu)}_J(\sigma))|Z^{(\mu)}_J(\sigma)|^{p-1}, \tag{4.1}$$

where (recall the definition of the FLIP operator F_J in Section 1.4)

$$Z^{(\mu)}_J(\sigma) := \frac{1}{\sqrt{N_1}}(\xi^{(\mu)}, F_J\sigma). \tag{4.2}$$

Next, we give the central technical lemma employed in the proof of Theorem 1.2. In the sequel, we shorten $J + k := J \cup \{k\}$ if $k \notin J$ and $J - k := J \setminus \{k\}$ if $k \in J$.

Lemma 4.1. *Let $r \in (0, \frac{1}{2}]$, $J \subseteq [N_1]$ with $|J| = \lfloor rN_1 \rfloor$. For any $p \in (1, 2]$, we have*

$$\begin{aligned} N_1(H^{(p)}(F_{J_{\pm k}}\xi^{(1)}) - H^{(p)}(F_J\xi^{(1)})) = \mp \sum_{\mu=2}^{N_2} W^{(\mu)}_{p,k,J}(\xi^{(1)}) - \alpha\varsigma N_1^{1-\frac{p}{2}} \\ \pm \frac{2p(1-2r)^{p-1}}{N_1^{1-\frac{p}{2}}} + O\left(\frac{1}{N_1^{2-\frac{p}{2}}}\right), \end{aligned} \tag{4.3}$$

where ς is a strictly positive and uniformly bounded random variable depending on $\{\xi^{(1)}_k \xi^{(\mu)}_k\}_{\mu=2,\dots,N_2}$ and $\{Z^{(\mu)}_J(\xi^{(1)})\}_{\mu=2,\dots,N_2}$. Setting

$$d(p) := 2^p\left(2p - 1 - 2^{p-1}\left(\frac{p-1}{p}\right)^{p-1}\frac{3p-2}{p}\right),$$

we have $C^p > \varsigma \ge d(p)$ for any realisation of ς and $C > 0$ an absolute constant.

In particular,

$$\frac{N_1}{4}(H^{(2)}(F_{J_{\pm k}}\xi^{(1)}) - H^{(2)}(F_J\xi^{(1)})) = \mp \sum_{\mu\geq 2} \frac{\xi_k^{(1)}\xi_k^{(\mu)}}{\sqrt{N_1}} Z_J^{(\mu)}(\xi^{(1)}) - \alpha \pm (1-2r) \mp \frac{1}{N_1}. \tag{4.4}$$

Proof. By Lemma 3.1, we have for any $k \notin J$

$$\begin{aligned} N_1(H^{(p)}(F_{J_{+k}}\xi^{(1)}) - H^{(p)}(F_J\xi^{(1)})) &= N_1\big(H(F_{J_{+k}}\xi^{(1)}) - H(\xi^{(1)}) - (H(F_J\xi^{(1)}) - H(\xi^{(1)}))\big) \\ &= N_1^{\frac{p}{2}}\left(\bar{\Phi}_p\left(r + \frac{1}{N_1}\right) - \bar{\Phi}_p(r)\right) \\ &\quad + \sum_{\mu=2}^{N_2} \Phi_p\big(X_{J_{+k}}^{(\mu)}(\xi^{(1)}), Y_{J_{+k}}^{(\mu)}(\xi^{(1)})\big) \\ &\quad - \sum_{\mu=2}^{N_2} \Phi_p\big(X_J^{(\mu)}(\xi^{(1)}), Y_J^{(\mu)}(\xi^{(1)})\big) \\ &= N_1^{\frac{p}{2}}\left(|1-2r|^p - \left|1-2r-\frac{2}{N_1}\right|^p\right) \\ &\quad + \sum_{\mu=2}^{N_2}\left(|Z_J^{(\mu)}(\xi^{(1)})|^p - \left|Z_J^{(\mu)}(\xi^{(1)}) - 2\frac{\xi_k^{(\mu)}\xi_k^{(1)}}{\sqrt{N_1}}\right|^p\right). \end{aligned} \tag{4.5}$$

Similarly, for all $k \in J$,

$$\begin{aligned} N_1(H^{(p)}(F_{J_{-k}}\xi^{(1)}) - H^{(p)}(F_J\xi^{(1)})) &= N_1^{\frac{p}{2}}\left(|1-2r|^p - \left|1-2r+\frac{2}{N_1}\right|^p\right) \\ &\quad + \sum_{\mu=2}^{N_2}\left(|Z_J^{(\mu)}(\xi^{(1)})|^p - \left|Z_J^{(\mu)}(\xi^{(1)}) + 2\frac{\xi_k^{(\mu)}\xi_k^{(1)}}{\sqrt{N_1}}\right|^p\right). \end{aligned} \tag{4.6}$$

For $p = 2$ a straightforward computation gives (4.4) from (4.5) and (4.6).

In general, for $p \in (1,2)$, we have to use Taylor expansion. Let $(p)_0 := 1$ and $(p)_k := \prod_{j=0}^{k-1}(p-j)$ for $k \geq 1$. Assuming $r \in (0, \frac{1}{2})$, N_1 large enough (i.e., $N_1(1-2r) > 2$), we have

$$\left|1-2r \pm \frac{2}{N_1}\right|^p = |1-2r|^p \pm \frac{2p}{N_1}|1-2r|^{p-1} + \frac{4}{N_1^2}\sum_{k\geq 2}\frac{(p)_k}{k!}\frac{2^{k-2}(1-2r)^{p-k}}{N_1^{k-2}},$$

and using $|(p)_k| \leq k!$, we get

$$\left|\frac{4}{N_1^2}\sum_{k\geq 2}\frac{(p)_k}{k!}\frac{2^{k-2}(1-2r)^{p-k}}{N_1^{k-2}}\right| \leq \frac{4}{N_1^2(1-2r)^{2-p}}\sum_{k\geq 0}\frac{2^k}{((1-2r)N_1)^k} \lesssim \frac{1}{N_1^2(1-2r)^{2-p}}.$$

Therefore,

$$N_1^{\frac{p}{2}}\left(|1-2r|^p - \left|1-2r \pm \frac{2}{N_1}\right|^p\right) = \mp\frac{2p}{N_1^{1-\frac{p}{2}}}|1-2r|^{p-1} + O\left(\frac{1}{N_1^{2-\frac{p}{2}}}\right). \quad (4.7)$$

On the other hand, for $r = \frac{1}{2}$, this correction term is trivially of order $N_1^{-\frac{p}{2}}$.

Recall now (4.1) and compute

$$\left|Z_J^{(\mu)}(\xi^{(1)}) \pm 2\frac{\xi_k^{(\mu)}\xi_k^{(1)}}{\sqrt{N_1}}\right|^p = 1_{\{|Z_J^{(\mu)}(\xi^{(1)})|\geq 2N_1^{-\frac{1}{2}}\}}\Bigg(|Z_J^{(\mu)}(\xi^{(1)})|^p \pm W_{p,k,J}^{(\mu)}(\xi^{(1)})$$
$$+\sum_{\ell\geq 2}\frac{(p)_\ell}{\ell!}\frac{|Z_J^{(\mu)}(\xi^{(1)})|^p}{(Z_J^{(\mu)}(\xi^{(1)}))^\ell}\frac{(\pm 2\xi_k^{(\mu)}\xi_k^{(1)})^\ell}{N_1^{\frac{\ell}{2}}}\Bigg)$$
$$+1_{\{|Z_J^{(\mu)}(\xi^{(1)})|< 2N_1^{-\frac{1}{2}}\}}\frac{2^p}{N_1^{\frac{p}{2}}}\Bigg(1 \pm \frac{p}{2}\xi_k^{(\mu)}\xi_k^{(1)}Z_J^{(\mu)}(\xi^{(1)})\sqrt{N_1}$$
$$+\sum_{\ell\geq 2}\frac{(p)_\ell}{\ell!}\left(\pm\frac{\xi_k^{(\mu)}\xi_k^{(1)}}{2}Z_J^{(\mu)}(\xi^{(1)})\sqrt{N_1}\right)^\ell\Bigg).$$

Thus,

$$\left(|Z_J^{(\mu)}(\xi^{(1)})|^p - \left|Z_J^{(\mu)}(\xi^{(1)}) \pm 2\frac{\xi_k^{(\mu)}\xi_k^{(1)}}{\sqrt{N_1}}\right|^p\right) = \mp W_{p,k,J}^{(\mu)}(\xi^{(1)})$$
$$-1_{\{|Z_J^{(\mu)}(\xi^{(1)})|\geq 2N_1^{-\frac{1}{2}}\}}\sum_{\ell\geq 2}\frac{(p)_\ell}{\ell!}\frac{|Z_J^{(\mu)}(\xi^{(1)})|^p}{(Z_J^{(\mu)}(\xi^{(1)}))^\ell}\frac{(\pm 2\xi_k^{(\mu)}\xi_k^{(1)})^\ell}{N_1^{\frac{\ell}{2}}} \quad (4.8)$$
$$-1_{\{|Z_J^{(\mu)}(\xi^{(1)})|< 2N_1^{-\frac{1}{2}}\}}$$
$$\times\left(\frac{2^p}{N_1^{\frac{p}{2}}} \pm \frac{2^p}{N_1^{\frac{p}{2}}}\frac{\xi_k^{(\mu)}\xi_k^{(1)}}{2}Z_J^{(\mu)}(\xi^{(1)})\sqrt{N_1} - |Z_J^{(\mu)}(\xi^{(1)})|^p \mp W_{p,k,J}^{(\mu)}(\xi^{(1)})\right) \quad (4.9)$$
$$-1_{\{|Z_J^{(\mu)}(\xi^{(1)})|< 2N_1^{-\frac{1}{2}}\}}\frac{2^p}{N_1^{\frac{p}{2}}}\sum_{\ell\geq 2}\frac{(p)_\ell}{\ell!}2^{-\ell}(\pm Z_J^{(\mu)}(\xi^{(1)})\xi_k^{(\mu)}\xi_k^{(1)}\sqrt{N_1})^\ell. \quad (4.10)$$

The contributions (4.8) and (4.10) are very similar and will be dealt together. Using that for $\ell \geq 2$ it is $\ell - p > 0$ and $\ell^2 (p)_\ell \leq 2p^2 \ell!$ we have

$$|(4.8)| \leq \frac{2p^2}{N_1^{\frac{p}{2}}} \sum_{\ell \geq 2} \ell^{-2}. \tag{4.11}$$

Similarly,

$$|(4.10)| \leq \frac{2^{p+1} p^2}{N_1^{\frac{p}{2}}} \sum_{\ell \geq 2} \ell^{-2}. \tag{4.12}$$

Furthermore, depending on the value of $\operatorname{sign}(Z_J^{(\mu)}(\xi^{(1)}))\xi_k^{(1)}\xi_k^{(\mu)}$, we have either

$$\begin{aligned}
(4.8) &= 1_{\{|Z_J^{(\mu)}(\xi^{(1)})| \geq 2N_1^{-\frac{1}{2}}\}} \frac{1}{N_1^{\frac{p}{2}}} \sum_{\ell \geq 2} \frac{(p)_\ell}{\ell!} \frac{|Z_J^{(\mu)}(\xi^{(1)})|^p}{|Z_J^{(\mu)}(\xi^{(1)})|^\ell} \frac{2^\ell}{N_1^{\frac{\ell}{2}}} \\
&= 1_{\{|Z_J^{(\mu)}(\xi^{(1)})| \geq 2N_1^{-\frac{1}{2}}\}} \frac{|Z_J^{(\mu)}(\xi^{(1)})|^p}{N_1^{\frac{p}{2}}} \\
&\quad \times \left[(1 + 2|Z_J^{(\mu)}(\xi^{(1)})\sqrt{N_1}|^{-1})^p - (1 + 2p|Z_J^{(\mu)}(\xi^{(1)})\sqrt{N_1}|^{-1}) \right] \geq 0,
\end{aligned}$$

where equality is achieved only if $p = 1$, or

$$\begin{aligned}
(4.8) &= 1_{\{|Z_J^{(\mu)}(\xi^{(1)})| \geq 2N_1^{-\frac{1}{2}}\}} \frac{1}{N_1^{\frac{p}{2}}} \sum_{\ell \geq 2} (-1)^\ell \frac{(p)_\ell}{\ell!} \frac{|Z_J^{(\mu)}(\xi^{(1)})|^p}{|Z_J^{(\mu)}(\xi^{(1)})|^\ell} \frac{2^\ell}{N_1^{\frac{\ell}{2}}} \\
&= 1_{\{|Z_J^{(\mu)}(\xi^{(1)})| \geq 2N_1^{-\frac{1}{2}}\}} \frac{|Z_J^{(\mu)}(\xi^{(1)})|^p}{N_1^{\frac{p}{2}}} \sum_{\ell \geq 1} \left(\frac{\sqrt{N_1} Z_J^{(\mu)}(\xi^{(1)})}{2} \right)^{-2\ell} \\
&\quad \times \left(\frac{(p)_{2\ell}}{2\ell!} - \frac{(p)_{2\ell+1}}{2\ell+1!} \left| \frac{\sqrt{N_1} Z_J^{(\mu)}(\xi^{(1)})}{2} \right|^{-1} \right),
\end{aligned} \tag{4.13}$$

where we split the sum over even and odd $\ell \geq 2$ and rename the indices to get the second identity. The quantity in (4.13) above is non-negative, since for $\ell \geq 1$, $(p)_{2\ell} \geq 0$ and $(p)_{2\ell+1} \leq 0$, which can be shown by observing that $(p)_{\ell \geq 3} = (-1)^\ell p(p-1) \prod_{j=2}^{\ell-1} (j - p)$ (again the equality is achieved only for $p = 1$).

Similarly, we have

$$\begin{aligned}
(4.10) &= 1_{\{|Z_J^{(\mu)}(\xi^{(1)})| < 2N_1^{-\frac{1}{2}}\}} \frac{2^p}{N_1^{\frac{p}{2}}} \sum_{\ell \geq 2} \frac{(p)_\ell}{\ell!} 2^{-\ell} |Z_J^{(\mu)}(\xi^{(1)})\sqrt{N_1}|^\ell \\
&= 1_{\{|Z_J^{(\mu)}(\xi^{(1)})| < 2N_1^{-\frac{1}{2}}\}} \frac{2^p}{N_1^{\frac{p}{2}}} \\
&\quad \times \left[(1 + p|2^{-1} Z_J^{(\mu)}(\xi^{(1)})\sqrt{N_1}|)^p - (1 + |2^{-1} Z_J^{(\mu)}(\xi^{(1)})\sqrt{N_1}|) \right] \geq 0
\end{aligned} \tag{4.14}$$

(equality is achieved only if $p = 1$) or

$$
\begin{aligned}
(4.10) &= 1_{\{|Z_J^{(\mu)}(\xi^{(1)})|<2N_1^{-\frac{1}{2}}\}} \frac{2^p}{N_1^{\frac{p}{2}}} \sum_{\ell\geq 2} (-1)^\ell \frac{(p)_\ell}{\ell!} 2^{-\ell} |Z_J^{(\mu)}(\xi^{(1)})\sqrt{N_1}|^\ell \\
&= 1_{\{|Z_J^{(\mu)}(\xi^{(1)})|<2N_1^{-\frac{1}{2}}\}} \frac{2^p}{N_1^{\frac{p}{2}}} \sum_{\ell\geq 1} \left(\frac{\sqrt{N_1} Z_J^{(\mu)}(\xi^{(1)})}{2} \right)^{2\ell} \\
&\times \left(\frac{(p)_{2\ell}}{2\ell!} - \frac{(p)_{2\ell+1}}{2\ell+1!} \frac{|\sqrt{N_1} Z_J^{(\mu)}(\xi^{(1)})|}{2} \right) \geq 0 \qquad (4.15)
\end{aligned}
$$

by the same argument used for the (4.13).

We conclude that

$$
(4.8) \leq 0, \quad (4.10) \leq 0 \quad \text{for } p \in (1, 2).
$$

Moreover, for any $a \in (0, 1)$,

$$
(4.14) 1_{\{|Z_J^{(\mu)}(\xi^{(1)})|\geq 2aN_1^{-\frac{1}{2}}\}} \geq \frac{(2a)^p}{aN_1^{\frac{p}{2}}}, \qquad (4.16)
$$

$$
(4.15) 1_{\{|Z_J^{(\mu)}(\xi^{(1)})|\geq 2aN_1^{-\frac{1}{2}}\}} \geq \frac{2^p}{N_1^{\frac{p}{2}}} a^2 \left(\frac{(p)_2}{2} + a\frac{|(p)_3|}{6} \right). \qquad (4.17)
$$

With a bit of algebra, we rewrite the term in (4.9) as follows:

$$
\begin{aligned}
&1_{\{|Z_J^{(\mu)}(\xi^{(1)})|<2N_1^{-\frac{1}{2}}\}} \frac{2^p}{N_1^{\frac{p}{2}}} \Bigg(\left(1 - \frac{|\sqrt{N_1} Z_J^{(\mu)}(\xi^{(1)})|^p}{2^p} \right) \\
&\quad \mp p \xi_k^{(\mu)} \xi_k^{(1)} \operatorname{sign}(Z_J^{(\mu)}(\xi^{(1)})) \left(\frac{|\sqrt{N_1} Z_J^{(\mu)}(\xi^{(1)})|^{p-1}}{2^{p-1}} - \frac{|\sqrt{N_1} Z_J^{(\mu)}(\xi^{(1)})|}{2} \right) \Bigg). \qquad (4.18)
\end{aligned}
$$

According to the value of $\xi_k^{(\mu)} \xi_k^{(1)} \operatorname{sign}(Z_J^{(\mu)}(\xi^{(1)}))$, the term inside the parenthesis can be either

$$
1 - |x|^p + p(|x|^{p-1} - |x|) \quad \text{or} \quad 1 - |x|^p - p(|x|^{p-1} - |x|),
$$

where we shortened

$$
|x| := \frac{|2Z_J^{(\mu)}(\xi^{(1)})|}{N_1^{\frac{1}{2}}} < 1.
$$

The first expression above is clearly positive, while the second one is positive thanks to Lemma B.1. More precisely, for any $a \in (0, 1)$,

$$
(4.18) 1_{\{|Z_J^{(\mu)}(\xi^{(1)})|<2aN_1^{-\frac{1}{2}}\}} \geq 1 - a^p - p(a^{p-1} - a) > 0. \qquad (4.19)
$$

From the representation (4.18), we also get the bound

$$(4.9) \leq \frac{C}{N_1^{\frac{p}{2}}}. \tag{4.20}$$

Now, we pick $a = 2(p-1)/p$ into (4.16), (4.17), and (4.19). Combining with (4.11), (4.12), and (4.20), we conclude that the lines (4.8), (4.9), and (4.10) define a random variable

$$\varsigma := \varsigma(\{\xi_k^{(1)}\xi_k^{(\mu)}, Z_J^{(\mu)}(\xi^{(1)})\}_{\mu=2,\dots,N_2})$$

lying in a uniformly bounded interval away from the origin such that

$$\begin{aligned}\sum_{\mu=2}^{N_2}\left(|Z_J^{(\mu)}(\xi^{(1)})|^p - \left|Z_J^{(\mu)}(\xi^{(1)}) \pm 2\frac{\xi_k^{(\mu)}\xi_k^{(1)}}{\sqrt{N_1}}\right|^p\right) \\ = \mp\sum_{\mu=2}^{N_2} W_{p,k,J}^{(\mu)}(\xi^{(1)}) + \alpha\varsigma N_1^{1-\frac{p}{2}}.\end{aligned}$$

Precisely, we have

$$\varsigma \geq 2^p\left(2p - 1 - 2^{p-1}\left(\frac{p-1}{p}\right)^{p-1}\frac{3p-2}{p}\right).$$

This and (4.7) give (4.3). ∎

Now, we turn to the proof of Theorem 1.2, which we conveniently split into several steps.

Step 1: reduction. Due to the exchangeability of the patterns and their entries, we have

$$\begin{aligned}&P\left(\mathtt{LM}_{N_1}^{(\mu)} \cap \widehat{B}_{\mu,\lfloor rN_1\rfloor}^{N_1} \neq \emptyset\right) \\ &\leq N_2\sum_{\ell=1}^{\lfloor rN_1\rfloor}\sum_{\substack{J\subset[N_1]\\|J|=\ell}} P\left(\bigcap_{k\notin J}\{H(F_{J+k}\xi^{(1)}) - H(F_J\xi^{(1)}) > 0\},\right. \\ &\qquad\qquad \left.\bigcap_{k\in J}\{H(F_{J-k}\xi^{(1)}) - H(F_J\xi^{(1)}) > 0\}\right) \\ &= N_2\sum_{\ell=1}^{\lfloor rN_1\rfloor}\binom{N_1}{\ell} P\left(\bigcap_{k>\ell}\{H(F_{[\ell]+k}\xi^{(1)}) - H(F_{[\ell]}\xi^{(1)}) > 0\},\right. \\ &\qquad\qquad \left.\bigcap_{k\leq\ell}\{H(F_{[\ell]-k}\xi^{(1)}) - H(F_{[\ell]}\xi^{(1)}) > 0\}\right).\end{aligned} \tag{4.21}$$

Recalling (4.1) and (4.2), we set for brevity

$$W^{(\mu)}_{p,k,\ell}(\xi^{(1)}) := W^{(\mu)}_{p,k,[\ell]}(\xi^{(1)}), \quad W^{(\mu)}_{p,k,\ell} := W^{(\mu)}_{p,k,\ell}(1),$$
$$M^{(\mu)} := \frac{1}{N_1}\sum_{i\in[N_1]} \xi^{(\mu)}_i, \quad Q^{(\mu)} := \mathrm{sign}(M^{(\mu)})|M^{(\mu)}|^{p-1}.$$

By Lemma 4.1, we have for N_1 large enough ($r' := \ell/N_1$)

$$\begin{aligned}
(4.21) \leq N_2 \sum_{\ell=1}^{\lfloor rN_1\rfloor} \binom{N_1}{\ell} P\Bigg(&\forall k > \ell - \sum_{\mu=2}^{N_2} W^{(\mu)}_{p,k,\ell}(\xi^{(1)}) \geq d(p)\alpha N_1^{1-\frac{p}{2}} \\
&- \frac{2p(1-2r')^{p-1}}{N_1^{1-\frac{p}{2}}}, \forall k \in [\ell] \sum_{\mu=2}^{N_2} W^{(\mu)}_{p,k,\ell}(\xi^{(1)}) \\
&\geq d(p)\alpha N_1^{1-\frac{p}{2}} + \frac{2p(1-2r')^{p-1}}{N_1^{1-\frac{p}{2}}}\Bigg) \\
= N_2 \sum_{\ell=1}^{\lfloor rN_1\rfloor} \binom{N_1}{\ell} P\Bigg(&\forall k > \ell - \sum_{\mu=2}^{N_2} W^{(\mu)}_{p,k,\ell} \geq d(p)\alpha N_1^{1-\frac{p}{2}} \\
&- \frac{2p(1-2r')^{p-1}}{N_1^{1-\frac{p}{2}}}, \forall k \in [\ell] \sum_{\mu=2}^{N_2} W^{(\mu)}_{p,k,\ell} \\
&\geq d(p)\alpha N_1^{1-\frac{p}{2}} + \frac{2p(1-2r')^{p-1}}{N_1^{1-\frac{p}{2}}}\Bigg) \\
= N_2 \sum_{\ell=1}^{\lfloor rN_1\rfloor} \binom{N_1}{\ell} P\Bigg(&\forall k > \ell (Q, \tilde{\xi}^{(k)}) \geq d(p)\alpha N_1^{1-\frac{p}{2}} - \frac{2p(1-2r')^{p-1}}{N_1^{1-\frac{p}{2}}}, \\
&\forall k \in [\ell] (Q, \tilde{\xi}^{(k)}) \geq d(p)\alpha N_1^{1-\frac{p}{2}} + \frac{2p(1-2r')^{p-1}}{N_1^{1-\frac{p}{2}}}\Bigg)
\end{aligned}$$

(recall that $\tilde{\xi}^{(k)}$ denotes the kth transposed pattern). In the second identity above we have exploited independence of the pattern $\xi^{(1)}$ to replace $W^{(\mu)}_{p,k,\ell}(\xi^{(1)})$ by $W^{(\mu)}_{p,k,\ell}$ and in the third one the independence of the first ℓ entries from all the others and the flip-symmetry to replace $Z^{(\mu)}_\ell$ by $M^{(\mu)}$. We also used that the independent random variables $Q^{(2)} \dots Q^{(N_2)}$ are symmetric. We denote by Q (respectively, M) the vector whose μth component is $Q^{(\mu)}$ (respectively, $M^{(\mu)}$).

Step 2: disentangling by the FKG inequality. Let us denote by $\mathcal{M}$ the σ-field generated by $M^{(2)}, \dots, M^{(N_2)}$ and notice that $Q^{(2)} \dots Q^{(N_2)}$ are $\mathcal{M}$-measurable. We shorten $P_{\mathcal{M}}(\cdot) := P(\cdot \mid \mathcal{M})$. The crucial observation here (first remarked in [19] for the Hopfield model) is that for each $\mu \in [N_2]$ the law of $\xi^{(\mu)}$ conditionally on $\mathcal{M}$ is a permutation distribution (as the increments of a simple random walk given the position). Therefore, the FKG inequality applies (see, for instance, [21]), and we have

$$\begin{aligned} P_{\mathcal{M}}\Bigg(&\forall k > \ell (Q, \tilde\xi^{(k)}) \ge d(p)\alpha N_1^{1-\frac{p}{2}} - \frac{2p(1-2r')^{p-1}}{N_1^{1-\frac{p}{2}}}, \\ &\forall k \in [\ell] (Q, \tilde\xi^{(k)}) \ge d(p)\alpha N_1^{1-\frac{p}{2}} + \frac{2p(1-2r')^{p-1}}{N_1^{1-\frac{p}{2}}} \Bigg) \\ &\le \prod_{k>\ell} P_{\mathcal{M}}\Bigg((Q, \tilde\xi^{(k)}) \ge d(p)\alpha N_1^{1-\frac{p}{2}} - \frac{2p(1-2r')^{p-1}}{N_1^{1-\frac{p}{2}}} \Bigg) \\ &\times \prod_{k\in[\ell]} P_{\mathcal{M}}\Bigg((Q, \tilde\xi^{(k)}) \ge d(p)\alpha N_1^{1-\frac{p}{2}} + \frac{2p(1-2r')^{p-1}}{N_1^{1-\frac{p}{2}}} \Bigg). \end{aligned} \tag{4.22}$$

Note that

$$E[\xi_k^{(\mu)} \mid \mathcal{M}] = M^{(\mu)}$$

and

$$\mathrm{Var}[\xi_k^{(\mu)} \mid \mathcal{M}] = 1 - (M^{(\mu)})^2.$$

Therefore, introducing $\bar Q \in \mathbb{R}^{N_2}$ with components $\bar Q^{(\mu)} := Q^{(\mu)} \sqrt{1-(M^{(\mu)})^2}$, we get

$$P_{\mathcal{M}}\big((Q, \tilde\xi^{(k)}) \ge t\big) \le \exp\left(- \frac{(t - \|M\|_p^p)^2}{2\|\bar Q\|_2^2} \right)$$

by the Hoeffding inequality. Note that this quantity is independent on k and also, we have the simple bound $\|\bar Q\|_2^2 \le \|Q\|_2^2 = \|M\|_{2p-2}^{2p-2}$. Then, (note $d(2) = 1$)

$$(4.22)_{p\in(1,2)} \le \exp\left(- N_1 \frac{(\alpha d(p) N_1^{1-\frac{p}{2}} - \|M\|_p^p)^2}{2\|M\|_{2p-2}^{2p-2}} \right), \tag{4.23}$$

$$\begin{aligned} (4.22)_{p=2} \le \exp\Bigg(&- \frac{N_1}{2\|M\|_2^2} \big(r'(\alpha - (1-2r') - \|M\|_2^2)^2 \\ &+ (1-r')(\alpha + (1-2r') - \|M\|_2^2)^2 \big) \Bigg). \end{aligned} \tag{4.24}$$

Step 3: concentration. Now, we have to take the global expectations of the right-hand side above. We notice that $E[\|M\|_p^p] \simeq \alpha N_1^{1-\frac{p}{2}}$ (this is an identity for $p = 2$) and $\|(M^{(\mu)})^p - E[(M^{(\mu)})^p]\|_{\psi_{\frac{2}{p}}} \simeq N_1^{-\frac{p}{2}}$. We can give precise upper bounds for these quantities. It holds for any $p > 0$

$$E[\|M\|_p^p] \le \frac{\alpha N_1}{N_1^{\frac{p}{2}}} \int e^{-\frac{x^{\frac{2}{p}}}{2}} dx =: \alpha N_1^{1-\frac{p}{2}} e(p) \tag{4.25}$$

(to prove it, proceed as in the computation giving (2.5)) and

$$\|(M^{(\mu)})^p - E[(M^{(\mu)})^p]\|_{\psi_{\frac{2}{p}}} \le \left(\frac{3}{2N_1}\right)^{\frac{p}{2}}$$

(to prove it, proceed as in the proof of Lemma 3.2).

Hence, by Proposition A.1 in Appendix A (with $\ell = 2/p$),

$$P\left(\left|\sum_{\mu\ge 2} |M^{(\mu)}|^p - E[|M^{(\mu)}|^p]\right| \ge t\right) \le 2\exp\left(-\frac{1}{8}\min\left(\frac{2^p t^2 N_1^{p-1}}{3^p\alpha}, \frac{2t^{\frac{2}{p}} N_1^{2-\frac{2}{p}}}{3\alpha^{\frac{2}{p}-1}}\right)\right). \tag{4.26}$$

We will also use the following sub-Gaussian estimate, which follows from [24, Corollary 2.8] (there the constant was not specified, but our choice is however not the optimal one). For any $p \in (1,2)$ there is a number $h > 0$ such that for any $t < 2\alpha h N_1^{2-p}$

$$P\left(\left|\sum_{\mu\ge 2} |M^{(\mu)}|^{2p-2} - E[|M^{(\mu)}|^{2p-2}]\right| \ge t\right) \le 2\exp\left(-\frac{t^2}{4\alpha h N_1^{3-2p}}\right). \tag{4.27}$$

A sketch of the proof of (4.27) is given at the end of this section. Note that for $p = 2$ (4.27) reduces to the standard Gaussian estimate in the Bernstein inequality $(4.26)|_{p=2}$ (however, numerical constants may change a bit).

Step 4: finalising the argument for $p \in (1,2)$. Using (4.26) with

$$t = \frac{1}{2}\alpha N_1^{1-\frac{p}{2}} |d(p) - e(p)| =: \tau_p,$$

we obtain

$$\begin{aligned} &E[\text{right-hand side of (4.23)}] \le E\big[\text{right-hand side of (4.23)} 1_{\{|\|M\|_p^p - E[\|M\|_p^p]| \le \tau_p\}}\big] \\ &\quad + P\big(|\|M\|_p^p - E[\|M\|_p^p]| \ge \tau_p\big) \\ &\le E\left[\exp\left(-\frac{\alpha^2 N_1^{3-p}(d(p)-e(p))^2}{\|M\|_{2p-2}^{2p-2}}\right)\right] \\ &\quad + 2\exp\left(-\frac{1}{8}\min\left(\frac{2^p \tau_p^2 N_1^{p-1}}{3^p\alpha}, \frac{2\tau_p^{\frac{2}{p}} N_1^{2-\frac{2}{p}}}{3\alpha^{\frac{2}{p}-1}}\right)\right) \\ &= E\left[\exp\left(-\frac{\alpha^2 N_1^{3-p}(d(p)-e(p))^2}{\|M\|_{2p-2}^{2p-2}}\right)\right] + 2e^{-\alpha N_1 \kappa_1(p)}, \end{aligned} \tag{4.28}$$

where

$$\kappa_1(p) := \frac{1}{8}\min\left(\frac{1}{4}\left(\frac{2}{3}\right)^p (d(p)-e(p))^2, \frac{2}{3}2^{-\frac{2}{p}}(d(p)-e(p))^{\frac{2}{p}}\right).$$

Using (4.27) with $t = \frac{1}{2}E[\|M\|_{2p-2}^{2p-2}]$ and (4.25) (with $p \to 2p-2$), we obtain

$$\begin{aligned}
&E\left[\exp\left(-\frac{\alpha^2 N_1^{3-p}(d(p)-e(p))^2}{\|M\|_{2p-2}^{2p-2}}\right)\right] \\
&\le E\left[\exp\left(-\frac{2\alpha^2 N_1^{3-p}(d(p)-e(p))^2}{3\alpha N^{2-p}e(2p-2)}\right)\right] + 2\exp\left(-\frac{\alpha N_1}{16h}\right) \\
&= 2\exp\left(-\frac{2\alpha N_1(d(p)-e(p))^2}{2e(2p-2)}\right) + 2\exp\left(-\frac{\alpha N_1}{16h}\right).
\end{aligned}$$

Combining the display above with (4.21), (4.22), (4.23), (4.28) and using (1.12), we have

$$\begin{aligned}
&P\left(\mathtt{LM}_{N_1}^{(\mu)} \cap \widehat{B}_{\mu,\lfloor rN_1\rfloor}^{N_1} \neq \emptyset\right) \\
&\le 2N_2 \sum_{\ell=1}^{\lfloor rN_1\rfloor}\binom{N_1}{\ell}\left(e^{-\alpha N_1\kappa_1(p)} + 2e^{-\alpha N_1\frac{2(d(p)-e(p))^2}{2e(2p-2)}} + e^{-\frac{\alpha N_1}{16h}}\right) \\
&\le 6rN_2N_1 e^{N_1\left(S(r)-\alpha\min(\kappa_1(p),\frac{2(d(p)-e(p))^2}{2e(2p-2)},\frac{1}{16h})\right)},
\end{aligned}$$

which yields the assertion for $p \in (1,2)$.

Step 5: finalising the argument for $p = 2$. Using $(4.26)_{p=2}$ with $t = (1+\alpha)(1-2r') =: \tau$ in the first line of the display below and $(4.27)_{p=2}$ with $t = \frac{1}{2}E[\|M\|_2^2]$ in the third line, we have

$$\begin{aligned}
&E[\text{right-hand side of (4.24)}] \\
&\le E\left[\exp\left(-\frac{N_1}{2}\frac{((1-2r')-\tau)^2}{\|M\|_2^2}\right)\right] + P\left(|\|M\|_2^2 - E[\|M\|_2^2]| \ge \tau\right) \\
&\le E\left[\exp\left(-\frac{N_1}{2}\frac{\alpha^2(1-2r')^2}{\|M\|_2^2}\right)\right] + 2e^{-\frac{2}{3}N_1(1+\alpha)(1-2r')\min\left(1,\frac{2(1+\alpha)(1-2r')}{3\alpha}\right)} \\
&\le e^{-\frac{1}{3}\alpha N_1(1-2r')^2} + 2e^{-\frac{1}{3}\alpha N_1} + 2e^{-\frac{2}{3}N_1(1+\alpha)(1-2r')\min\left(1,\frac{2(1+\alpha)(1-2r')}{3\alpha}\right)} \\
&\le e^{-\frac{1}{3}\alpha N_1(1-2r)^2} + 2e^{-\frac{1}{3}\alpha N_1} + 2e^{-\frac{2}{3}N_1(1+\alpha)(1-2r)\min\left(1,\frac{2(1+\alpha)(1-2r)}{3\alpha}\right)},
\end{aligned} \tag{4.29}$$

as r' ranges from $1/N_1$ to r. So, combining (4.21), (4.22), (4.23), (4.29) and using again (1.12), we obtain

$$P\left(\mathtt{LM}_{N_1}^{(\mu)} \cap \widehat{B}^{N_1}_{\mu,\lfloor rN_1\rfloor} \neq \emptyset\right) \leq N_2 \sum_{\ell=1}^{\lfloor rN_1\rfloor} \binom{N_1}{\ell} \quad \text{(terms in (4.29))}$$
$$\leq 2rN_2N_1\Big(e^{N_1(S(r)-\frac{1}{3}\alpha(1-2r)^2)} + e^{N_1(S(r)-\frac{1}{3}\alpha)}$$
$$+ e^{N_1\left(S(r)-\frac{2}{3}(1+\alpha)(1-2r)\min(1,\frac{2(1+\alpha)(1-2r)}{3\alpha})\right)}\Big).$$

The first two summands are negative if $\alpha > 3S(r)/(1-2r)^2$. The fact that also the third one is so is verified in Lemma B.2, Appendix B. The proof is complete.

Proof of (4.27) *(sketch).* If $p < 2$, it is $2p-2 < 2$. We bound

$$E\left[e^{\left|\frac{1}{\sqrt{N_1}}\sum_{i\in[N_1]}\xi_i\right|^{2p-2}}\right]$$
$$\leq 2\sum_{n\geq 1} E\left[1_{\{n-1\leq \frac{1}{\sqrt{N_1}}\sum_{i\in[N_1]}\xi_i\leq n\}} e^{\left|\frac{1}{\sqrt{N_1}}\sum_{i\in[N_1]}\xi_i\right|^{2p-2}}\right]$$
$$\leq 2\sum_{n\geq 1} e^{n^{2p-2}-\frac{(n-1)^2}{4}} =: h < \infty.$$

We can write

$$|M^{(\mu)}|^{2p-2} = \frac{1}{N_1^{p-1}}\left(\frac{1}{\sqrt{N_1}}\sum_{i\in[N_1]}\xi_i^{(\mu)}\right)^{2p-2}.$$

If $s \in (0,1)$, we have (see, for instance, [24, Lemma 2.6])

$$E[e^{s|M^{(\mu)}|^{2p-2}}] \leq e^{s^2h}.$$

Thus, by the Markov inequality and optimisation over s,

$$P\left(N_1^{p-1}\left|\sum_{\mu\geq 2}|M^{(\mu)}|^{2p-2} - E[|M^{(\mu)}|^{2p-2}]\right| \geq t\right) \leq e^{s^2\alpha N_1 h - st} \leq e^{-\frac{t^2}{4\alpha h N_1}}.$$

provided $t \leq 2h\alpha N_1$. Changing variables implies the assertion. ■

A. Tail estimates

In this appendix, we present tail estimates for sums of i.i.d. r.vs. used in the main text. The following statement is not new, and we give the proof here mainly for the reader's convenience. In fact, the proof of the subsequent formula (A.1) for $\ell \in [1,2]$

is classical and can be found, for instance, in [24, Corollaries 2.9 and 2.10] (though the formulation is slightly different there). So, we focus on the case $\ell \in (0,1)$. For similar statements, see [20, Proposition 3.2] and [17, Theorem 6.21].

Proposition A.1. *Let $\ell \in (0,2]$, $X_1,\dots,X_N$ i.i.d. r.vs. Then, for N large enough,*

$$P\Big(\Big|\sum_{i\in[N]} X_i\Big| \ge t\Big) \le 2\exp\Big(-\frac{1}{8}\min\Big(\frac{t^2}{\|X_1\|_{\psi_\ell}^2 N}, \frac{t^\ell}{\|X_1\|_{\psi_\ell}^\ell N^{\max(\ell-1,0)}}\Big)\Big). \tag{A.1}$$

Proof (only for $\ell \in (0,1)$). We have by assumption

$$P(|X_1| \ge t) \le e^{-\frac{t^\ell}{2\|X_1\|_{\psi_\ell}^\ell}}.$$

Let now $s := \|X_1\|_{\psi_\ell} N^{\frac{1}{2-\ell}}$ and set $X_i^s := X_i 1_{\{|X_i|<s\}}$. Then, we have

$$\begin{aligned} P\Big(\Big|\sum_{i\in[N]} X_i\Big| \ge t\Big) &\le P\Big(\sum_{i\in[N]} X_i \ge t, \sup_{i\in[N]} |X_i| < s\Big) + P\Big(\sup_{i\in[N]} |X_i| \ge s\Big) \\ &\le P\Big(\Big|\sum_{i\in[N]} X_i^s\Big| \ge t\Big) + e^{-\frac{s^\ell}{4\|X_1\|_{\psi_\ell}^\ell}}. \end{aligned} \tag{A.2}$$

Set now

$$\bar\mu := \frac{1}{4s^{1-\ell}\|X_1\|_{\psi_\ell}^\ell}.$$

We note that for any $0 \le \mu \le \bar\mu$ (and N large enough) it is

$$\mu X_i^s \le \frac{|X_i^s|^\ell}{A\|X_i\|_{\psi_\ell}^\ell}.$$

Using the bound $x^2 \le e^{\frac{|x|^\ell}{A}}$, we compute

$$\begin{aligned} E[e^{\mu X_i^s}] &= 1 + \mu^2\|X_i^s\|_{\psi_\ell}^2 \sum_{n\ge 0} \mu^n \frac{E[(X_i^s)^{n+2}]}{\|X_i\|_{\psi_\ell}^2 (n+2)!} \\ &\le 1 + \mu^2\|X_i\|_{\psi_\ell}^2 E\Big[e^{\frac{|X_i^s|^\ell}{10\|X_i\|_{\psi_\ell}^\ell}} \sum_{n\ge 0}\frac{1}{n!}\Big(\frac{|X_i^s|^\ell}{10\|X_i\|_{\psi_\ell}^\ell}\Big)^n\Big] \\ &\le 1 + \mu^2\|X_i\|_{\psi_\ell}^2 E\Big[e^{\frac{|X_i|^\ell}{5\|X_i\|_{\psi_\ell}^\ell}}\Big] \\ &\le \exp\Big(\mu^2\|X_i\|_{\psi_\ell}^2 E\Big[e^{\frac{|X_i|^\ell}{5\|X_i\|_{\psi_\ell}^\ell}}\Big]\Big) \le \exp\big(2\mu^2\|X_i\|_{\psi_\ell}^2\big). \end{aligned}$$

It follows that

$$P\left(\left|\sum_{i\in[N]} X_i^s\right| \geq t\right) \leq 2e^{-\mu t+2N\mu^2\|X_1\|_{\psi_\ell}^2}$$

$$\leq \begin{cases} 2\exp\left(-\frac{t^2}{8\|X_1\|_{\psi_\ell}^2 N}\right), & 0 < t < 4N\bar{\mu}\|X_1\|_{\psi_\ell}^2, \\ 2\exp(-\bar{\mu}t + 2N\bar{\mu}^2\|X_1\|_{\psi_\ell}^2), & t \geq 4N\bar{\mu}\|X_1\|_{\psi_\ell}^2. \end{cases}$$

With our choice of parameters, the above formula rewrites as

$$P\left(\left|\sum_{i\in[N]} X_i^s\right| \geq t\right) \leq \begin{cases} 2\exp\left(-\frac{t^2}{8\|X_1\|_{\psi_\ell}^2 N}\right), & 0 < t < s, \\ 2\exp\left(-\frac{t^\ell}{8\|X_1\|_{\psi_\ell}^\ell}\right), & t \geq s. \end{cases} \tag{A.3}$$

Combining (A.2) and (A.3) gives the assertion. ■

B. Two technical lemmas

The following two results are basically calculus.

Lemma B.1. *Let* $g(x,p) := 1-(1-2x)^p - x^{\frac{p}{2}}$ *and* $f(x,p) := 1-x^p - px^{p-1} + px$. *It is* $g > 0$ *for all* $p \geq 2$ *and* $x \in [0,1/2]$. *Moreover, for any* $a \in (0,1)$, *it is* $f(p,x) \geq f(p,a) > 0$ *for all* $p \in (1,2]$ *and* $x \in [0,a]$.

Proof. For $x \in [0,1/2]$ the function $(1-2x)^p + x^{\frac{p}{2}}$ is decreasing in p, so it suffices to study $g(x,2)$ for which one verifies explicitly $g(x,2) > 0$ for all $x \in [0,1/2]$.

Now, we pass to f. First, we note that

$$1 + (x+p)\log x \leq x \quad \forall x \in [0,1]. \tag{B.1}$$

The proof is simple: we compare the function $\log x$ with $\frac{x-1}{x+p}$ for $x \in [0,1]$, and since

$$\frac{1}{x} = \frac{d}{dx}\log x \geq \frac{d}{dx}\frac{x-1}{x+p} = \frac{p+1}{(x+p)^2} \quad \forall x \in [0,1]$$

and in $x = 1$ the two functions intersect, (B.1) follows.

Next, we note that $f(x,1) = 0$ and $f(x,2) \geq 0$ for all $x \in [0,1]$. Then, we show that f is non-decreasing in p uniformly in $x \in [0,1]$. We compute

$$\frac{\partial}{\partial p} f(x,p) = x(1 - x^{p-2}(1 + (x+p)\log x)) \geq x(1-x^{p-1}) \geq 0,$$

thanks to (B.1). This tells us $f \geq 0$. Moreover, we compute

$$\frac{\partial}{\partial x} f(x,p) = p\left(1 - x^{p-1}\left(\frac{p-1}{x} + 1\right)\right).$$

We have for all $x \in [0,1]$

$$\frac{p-1}{x} + 1 \geq \frac{1}{x^{p-1}}.$$

The above inequality is clearly true if x is near the origin and at $x = 1$. Indeed, it must hold in the whole interval $[0,1]$ since the functions on both sides are decreasing.

It follows that f is decreasing in $[0,1]$ uniformly in $p \in (1,2]$, whence the assertion follows. ■

Lemma B.2. *Let $r \in [0, \frac{1}{2}]$, $c_1 > 0$. Let also $\bar{r} = \bar{r}(c_1) \in [0, 1/2]$ defined implicitly by*

$$\frac{S(\bar{r})}{1-2\bar{r}} = c_1,$$

and set

$$c_2 := \max\left(\frac{1}{c_1}, \frac{(1-2\bar{r})^2}{2c_1\bar{r}}\right).$$

For all $\alpha \geq c_2 S(r)/(1-2r)^2$, it holds

$$S(r) \leq c_1(1+\alpha)(1-2r)\min\left(1, \frac{(1+\alpha)(1-2r)}{\alpha}\right). \tag{B.2}$$

Proof. (B.2) selects two conditions, namely, either

$$\alpha \leq \frac{1-2r}{2r}, \quad c_1\alpha \geq \frac{S(r)}{1-2r} - c_1 \quad \text{or} \quad \alpha > \frac{1-2r}{2r}, \quad c_1\frac{(1+\alpha)^2}{\alpha} \geq \frac{S(r)}{(1-2r)^2}. \tag{B.3}$$

For $r \in [0, 1/2]$, the function $S(r)$ increases and $c_1(1-2r)$ decreases. Let us denote by $\bar{r}$ their unique intersection point in $[0, 1/2]$. Clearly, $\bar{r}$ depends on c_1 and $\bar{r} \to 0$ as $c_1 \to 0$. If $r \in [0, \bar{r}]$ then for every $\alpha < (1-2r)/2r$, it holds

$$c_1\alpha \geq \frac{S(r)}{1-2r} - c_1.$$

Moreover, there is $C > 0$ such that

$$\frac{1-2r}{2r} \geq C\frac{S(r)}{(1-2r)^2} \quad \forall r \in [0, \bar{r}].$$

Indeed, by definition of $\bar{r}$, the above condition is implied by

$$\frac{1-2r}{2r} \geq \frac{c_1 C}{1-2r} \quad \forall r \in [0, \bar{r}];$$

therefore, it suffices to take

$$c_2 := \frac{(1-2\bar{r})^2}{2c_1\bar{r}},$$

and we have the statement for $r \in [0, \bar{r}]$.

For $r \in [\bar{r}, 1/2]$, we use the second condition in (B.3). First, we observe that, since $(1+\alpha)^2/\alpha > \alpha$, the condition $c_1 \frac{(1+\alpha)^2}{\alpha} \geq \frac{S(r)}{(1-2r)^2}$ is implied by $\alpha \geq CS(r)/(1-2r)^2$ for all $C > c_1^{-1}$. It remains to show that there is $C > c_1^{-1}$ such that

$$\frac{1-2r}{2r} \leq C \frac{S(r)}{(1-2r)^2} \quad \text{or equivalently} \quad \frac{(1-2r)^2}{2r} \leq C \frac{S(r)}{(1-2r)}. \tag{B.4}$$

By definition of $\bar{r}$,

$$\frac{S(r)}{(1-2r)} \geq c_1 \quad \forall r \in [\bar{r}, 1/2].$$

The left-hand side of the second inequality in (B.4) is decreasing and its right-hand side is increasing, whence it suffices to require

$$\frac{(1-2\bar{r})^2}{2\bar{r}} \leq C c_1.$$

Thus, taking

$$c_2 := \max\left(\frac{1}{c_1}, \frac{(1-2\bar{r})^2}{2c_1\bar{r}}\right),$$

we have proved the statement also for $r \in [\bar{r}, 1/2]$. ■

Acknowledgements. The author thanks David Belius and two anonymous referees for many helpful comments and suggestions.

References

[1] L. Addario-Berry, L. Devroye, G. Lugosi, and R. I. Oliveira, Local optima of the Sherrington–Kirkpatrick Hamiltonian. *J. Math. Phys.* **60** (2019), no. 4, article no. 043301, 13 pp. Zbl 1426.82066 MR 3940914

[2] D. J. Amit, H. Gutfreund, and H. Sompolinsky, Storing infinite numbers of patterns in a spin-glass model of neural networks. *Phys. Rev. Lett.* **55** (1985), 1530–1533

[3] O. Angel, S. Bubeck, Y. Peres, and F. Wei, Local max-cut in smoothed polynomial time. In *STOC'17—Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, pp. 429–437, Association for Computing Machinery, New York, 2017 Zbl 1369.68226 MR 3678200

[4] A. Barra, G. Genovese, P. Sollich, and D. Tantari, Phase diagram of restricted Boltzmann machines and generalized Hopfield networks with arbitrary priors. *Phys. Rev. E* **97** (2018), no. 2, article no. 022310

[5] M. Benedetti, V. Dotsenko, G. Fischetti, E. Marinari, and G. Oshanin, Recognition capabilities of a Hopfield model with auxiliary hidden neurons. *Phys. Rev. E* **103** (2021), no. 6, article no. L060401 MR 4286623

[6] M. Benedetti, E. Ventura, E. Marinari, G. Ruocco, and F. Zamponi, Supervised perceptron learning vs. unsupervised hebbian unlearning: Approaching optimal memory retrieval in hopfield-like networks. *J. Chem. Phys.* **156** (2022), no. 10, article no. 0084219

[7] A. Bovier and B. Niederhauser, The spin-glass phase-transition in the Hopfield model with p-spin interactions. *Adv. Theor. Math. Phys.* **5** (2001), no. 6, 1001–1046 Zbl 1011.82008 MR 1926663

[8] G. Bresler, F. Koehler, and A. Moitra, Learning restricted Boltzmann machines via influence maximization. In *STOC'19—Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, pp. 828–839, Association for Computing Machinery, New York, 2019 Zbl 1433.68327 MR 4003387

[9] N. Bulso and Y. Roudi, Restricted Boltzmann machines as models of interacting variables. *Neural Comput.* **33** (2021), no. 10, 2646–2681 Zbl 1483.92011 MR 4323721

[10] L. Dabelow and M. Ueda, Three learning stages and accuracy–efficiency tradeoff of restricted Boltzmann machines. *Nat. Commun.* **13** (2022), no. 1, article no. 5474, 11 pp.

[11] M. Etscheid and H. Röglin, Smoothed analysis of local search for the maximum-cut problem. *ACM Trans. Algorithms* **13** (2017), no. 2, article no. 3011870, 25 pp. Zbl 1421.68152 MR 3659247

[12] E. Gardner, Multiconnected neural network models. *J. Phys. A* **20** (1987), no. 11, 3453–3464 MR 0914065

[13] S. Goel, Learning ising and potts models with latent variables. In *International conference on artificial intelligence and statistics, PMLR*, pp. 3557–3566, The University of Texas, Austin, 2020

[14] S. Goel, A. Klivans, and F. Koehler, From Boltzmann machines to neural networks and back again. In *Advances in neural information processing systems*, pp. 6354–6365, 33, The NeurIPS, San Diego, CA, 2020

[15] G. E. Hinton and T. J. Sejnowski, Optimal perceptual inference. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 448, IEEE Transactions on Pattern Analysis and Machine, Washington, D. C., 1983

[16] J. J. Hopfield, Neural networks and physical systems with emergent collective computational abilities. *Proc. Nat. Acad. Sci. U.S.A.* **79** (1982), no. 8, 2554–2558 Zbl 1369.92007 MR 0652033

[17] M. Ledoux and M. Talagrand, *Probability in Banach spaces. Isoperimetry and processes*. Ergeb. Math. Grenzgeb. (3) 23, Springer, Berlin, 1991 Zbl 0748.60004 MR 1102015

[18] W. A. Little, The existence of persistent states in the brain. *Math. Biosci.* **19** (1974), no. 1, 101–120 Zbl 0272.92011

[19] D. Loukianova, Lower bounds on the restitution error in the Hopfield model. *Probab. Theory Related Fields* **107** (1997), no. 2, 161–176 Zbl 0870.60021 MR 1431217

[20] C. M. Newman, Memory capacity in neural network models: Rigorous lower bounds. *Neural Netw.* **1** (1988), no. 3, 223–238

[21] D. Ranjan, D. Dubhashi, and V. Priebe, Negative dependence through the FKG inequality. Tech. Report RS-96-27, University of Aarhus, Denmark, 1996

[22] F. Schönsberg, Y. Roudi, and A. Treves, Efficiency of local learning rules in threshold-linear associative networks. *Phys. Rev. Lett.* **126** (2021), article no. 018301

[23] P. Smolensky, *Information processing in dynamical systems: Foundations of harmony theory*, chap. Parallel Distributed Processing: Explorations in the Microstructure of Cognition: Foundations, pp. 194–281. MIT Press, Cambridge, MA, 1986

[24] M. Talagrand, The supremum of some canonical processes. *Amer. J. Math.* **116** (1994), no. 2, 283–325 Zbl 0798.60040 MR 1269606

[25] M. Talagrand, Rigorous results for the Hopfield model with many patterns. *Probab. Theory Relat. Fields* **110** (1998), no. 2, 177–275 Zbl 0897.60041

[26] J. Tubiana, S. Cocco, and R. Monasson, Learning compositional representations of interacting systems with restricted Boltzmann machines: comparative study of lattice proteins. *Neural Comput.* **31** (2019), no. 8, 1671–1717 Zbl 1429.92113 MR 4002808

[27] J. Tubiana and R. Monasson, Emergence of compositional representations in restricted Boltzmann machines. *Phys. Rev. Lett.* **118** (2017), article no. 138301

Received 4 October 2022; revised 6 September 2023.

Giuseppe Genovese
Institute of Mathematics, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland; giuseppe.genovese@math.uzh.ch