Detection limits in the high-dimensional spiked rectangular model

Ahmed El Alaoui* Michael I. Jordan†

Abstract

We study the problem of detecting the presence of a single unknown spike in a rectangular data matrix, in a high-dimensional regime where the spike has fixed strength and the aspect ratio of the matrix converges to a finite limit. This situation comprises Johnstone's spiked covariance model. We analyze the likelihood ratio of the spiked model against an "all noise" null model of reference, and show it has asymptotically Gaussian fluctuations in a region below—but in general not up to—the so-called BBP threshold from random matrix theory. Our result parallels earlier findings of Onatski et al. (2013) and Johnstone-Onatski (2015) for spherical spikes. We present a probabilistic approach capable of treating generic product priors. In particular, sparsity in the spike is allowed. Our approach operates through the general principle of the cavity method from spin-glass theory. The question of the maximal parameter region where asymptotic normality is expected to hold is left open. This region is shaped by the prior in a non-trivial way. We conjecture that this is the entire paramagnetic phase of an associated spin-glass model, and is defined by the vanishing of the replica-symmetric solution of Lesieur et al. (2015).

1 Introduction

The problem of detecting a signal of low-rank structure buried inside a large noise matrix has received enormous attention in the past decade. Prominent examples of this problem include the so-called *spiked* or *deformed ensembles* from random matrix theory (Péché, 2014). It is particularly interesting to study such problems in the high-dimensional setting where the signal strength is comparable to the noise. This models practical situations in modern data analysis where one wishes to make more complex inferences about a fainter signal as the amount of data accrues. In this paper we are concerned with the problem of testing the presence of a single weak spike in the data against an "all-noise" null hypothesis of reference.

Concretely we consider the observation of an $N \times M$ matrix of the form

$$\boldsymbol{Y} = \sqrt{\frac{\beta}{N}}\boldsymbol{u}\boldsymbol{v}^\top + \boldsymbol{W}, \tag{1}$$

where $\boldsymbol{u}$ and $\boldsymbol{v}$ are unknown factors and $\boldsymbol{W}$ is a matrix with i.i.d. noise entries, and we want to test whether $\beta > 0$ or $\beta = 0$. We will assume the noise is standard Gaussian. The parameter β represents the strength of the spike, and we assume a high-dimensional setting where $M/N \to \alpha$. The case $\boldsymbol{u} = \boldsymbol{v}$ and $\boldsymbol{W}$ symmetric is referred to as the *spiked Wigner model*. When the factors are independent, model (1) can be viewed as a linear model with additive noise and scalar random design:

$$\boldsymbol{y}_j = \bar{\beta} v_j \boldsymbol{u} + \boldsymbol{w}_j,$$

*Electrical Engineering & Computer Sciences, UC Berkeley, CA. Email: elalaoui@berkeley.edu
†Electrical Engineering & Computer Sciences, Statistics, UC Berkeley, CA.

with $1 \leq j \leq M$, $\bar{\beta} = \sqrt{\beta/N}$. Assuming v_j has zero mean and unit variance, this is a model of *spiked covariance*: the mean of the empirical covariance matrix $\widehat{\boldsymbol{\Sigma}} = \frac{1}{M}\sum_{j=1}^{M} \boldsymbol{y}_j \boldsymbol{y}_j^\top$ is a rank-one perturbation of the identity: $\boldsymbol{I}_N + \frac{\beta}{N}\boldsymbol{u}\boldsymbol{u}^\top$.

The introduction of a particular spiked covariance model by Johnstone (2001)—one corresponding to the special case $v_j \sim \mathcal{N}(0,1)$—has provided the foundations for a rich theory of Principal Component Analysis (PCA), in which the performance of several important tests and estimators is by now well understood (see, e.g., Ledoit and Wolf, 2002; Paul, 2007; Nadler, 2008; Johnstone and Lu, 2009; Amini and Wainwright, 2009; Berthet and Rigollet, 2013; Dobriban, 2017). Parallel developments in random matrix theory have unveiled the existence of sharp transition phenomena in the behavior of the spectrum of the data matrix, where for a spike of strength above a certain *spectral* threshold, the top eigenvalue separates from the remaining eigenvalues which are packed together in a "bulk" and thus indicates the presence of the spike; below this threshold, the top eigenvalue converges to the edge of the bulk. See Péché (2006); Féral and Péché (2007); Capitaine et al. (2009); Benaych-Georges and Nadakuditi (2011) for results on low-rank deformations of Wigner matrices, and Baik et al. (2005); Baik and Silverstein (2006); Bai and Yao (2012, 2008) for results on (real and complex) spiked covariance models. More recently, an intense research effort has been undertaken to pin down the fundamental limits for both estimating and detecting the spike.

In a series of papers (Korada and Macris, 2009; Krzakala et al., 2016; Barbier et al., 2016; Deshpande et al., 2016; Lelarge and Miolane, 2017; Miolane, 2017), the error of the Bayes-optimal estimator has been completely characterized for additive low-rank models with a separable (product) prior on the spike. In particular, these papers confirm an interesting phenomenon discovered by Lesieur et al. (2015a,b), based on plausible but non-rigorous arguments: for certain priors on the spike, estimation becomes possible—although computationally expensive—below the spectral threshold $\beta = 1$. More precisely, the posterior mean overlaps with the spike in regions where the top eigenvector is orthogonal to it. Lesieur et al. (2017) provides a full account of these phase transitions in a myriad of interesting situations, the majority of which still await rigorous treatment. As for the testing problem, Onatski et al. (2013, 2014) and Johnstone and Onatski (2015) considered the spiked covariance model for a uniformly distributed unit norm spike, and studied the asymptotics of the likelihood ratio (LR) of a spiked alternative against a spherical null. They showed that the log-LR is asymptotically Gaussian below the spectral threshold $\alpha\beta^2 = 1$ (which in this setting is known as the BBP threshold, after Baik et al., 2005), while it is divergent above it.

However their proof is intrinsically tied to the assumption of a spherical prior. Indeed, by rotational symmetry of the model, the LR depends only on the spectrum, the joint distribution of which is available in closed form. A representation of the LR in terms of a contour integral is then possible (in the single spike case), which can then be analyzed via the method of steepest descent. In a similar but unrelated effort, Baik and Lee (2016, 2017a,b) studied the fluctuations of the free energy of spherical, symmetric and bipartite versions of the Sherrington–Kirkpatrick (SK) model. This free energy coincides with the log-LR associated with the model (1) for a choice of parameters. The sphericity assumption is again key to their analysis, and both approaches require the execution of very delicate asymptotics and appeal to advanced results from random matrix theory.

In this paper we consider the case of separable priors: we assume that the entries of $\boldsymbol{u}$ and $\boldsymbol{v}$ are independent and identically distributed from base priors P_{u} and P_{v}, respectively, both having bounded support[1]. We prove fluctuation results for the log-LR in this setting with entirely different methods than used for spherical priors. In particular, our proof is more probabilistic and operates through general principles. The tools we use come from the mathematical theory of spin glasses (see Talagrand, 2011a,b).

Let us further mention that the region of parameters (α, β) we are able to cover with our proof method is optimal when (and only when) P_{u} and P_{v} are both symmetric Rademacher. In Section 6, we formulate a conjecture on the *maximal* region in which the log-LR has asymptotically Gaussian fluctuations. This region is of course below the BBP threshold, but does *not* extend up to it in general.

2 Main results

Throughout this paper, we assume that the priors P_{u} and P_{v} have zero mean, unit variance, and supports bounded in radius by K_{u} and K_{v} respectively. Let $\mathbb{P}_\beta$ be the probability distribution of the matrix $\boldsymbol{Y}$ as per (1). Define $L(\cdot;\beta)$ to be the likelihood ratio, or Radon-Nikodym derivative of $\mathbb{P}_\beta$ with respect to $\mathbb{P}_0$:

$$L(\cdot;\beta) \equiv \frac{\mathrm{d}\,\mathbb{P}_\beta}{\mathrm{d}\,\mathbb{P}_0}.$$

For a fixed $\boldsymbol{Y} \in \mathbb{R}^{N\times M}$, by conditioning on $\boldsymbol{u}$ and $\boldsymbol{v}$, we can write

$$L(\boldsymbol{Y};\beta) = \int \exp\Big(\sum_{i,j}\sqrt{\frac{\beta}{N}}Y_{ij}u_iv_j - \frac{\beta}{2N}u_i^2v_j^2\Big)\mathrm{d}P_{\mathrm{u}}^{\otimes N}(\boldsymbol{u})\mathrm{d}P_{\mathrm{v}}^{\otimes M}(\boldsymbol{v}).$$

Our main contribution is the following asymptotic distributional result.

Theorem 1. *Let $\alpha, \beta \geq 0$ such that $K_{\mathrm{u}}^4K_{\mathrm{v}}^4\alpha\beta^2 < 1$. Then in the limit $N \to \infty$ and $M/N \to \alpha$,*

$$\log L(\boldsymbol{Y};\beta) \rightsquigarrow \mathcal{N}\left(\pm\frac{1}{4}\log\left(1-\alpha\beta^2\right), -\frac{1}{2}\log\left(1-\alpha\beta^2\right)\right),$$

where "$\rightsquigarrow$" denotes convergence in distribution. The sign of the mean is $+$ under the null $\boldsymbol{Y} \sim \mathbb{P}_0$ and $-$ under the alternative $\boldsymbol{Y} \sim \mathbb{P}_\beta$.

We mention that fluctuations of this sort were first proved by Aizenman et al. (1987) in a seminal paper in the context of the SK model. A consequence of either one of the above statements and Le Cam's first lemma (Van der Vaart, 2000, Lemma 6.4) is the mutual contiguity[2] between the null and the spiked alternative:

Corollary 2. *For $K_{\mathrm{u}}^4K_{\mathrm{v}}^4\alpha\beta^2 < 1$, the families of distributions $\mathbb{P}_0$ and $\mathbb{P}_\beta$ (indexed by M, N) are mutually contiguous in the limit $N \to \infty$, $M/N \to \alpha$.*

Contiguity implies impossibility of strong detection: there exists no test that, upon observing a random matrix $\boldsymbol{Y}$ with the promise that it is sampled either from $\mathbb{P}_0$ or $\mathbb{P}_\beta$, can tell which is the case with asymptotic certainty in this regime. We also mention that contiguity

[1]Boundedness is required for technical reasons. This rules out the case where one factor is Gaussian.

[2]Two sequences of probability measures (P_n) and (Q_n) defined on the same (sequence of) measurable space(s) are said to be mutually contiguous if $P_n(A_n) \to 0$ is equivalent to $Q_n(A_n) \to 0$ as $n \to \infty$ for every sequence of measurable sets (A_n).

can be proved through the second-moment method and its conditional variants, as was done by Montanari et al. (2015); Perry et al. (2016); Banks et al. (2017) for closely related models. However, identifying the right event on which to condition in order to tame the second moment of L is a matter of a case-by-case deliberation. Study of the fluctuations of the log-LR appears to provide a more systematic route: the logarithm has a smoothing effect that kills the wild (but rare) events that otherwise dominate in the second moment. This being said, our result is optimal only in one special case:

When P_{u} and P_{v} are symmetric Rademacher, $K_{\mathrm{u}} = K_{\mathrm{v}} = 1$, and Theorem 1 covers the entire (α, β) region where such fluctuations hold. Indeed, for $\alpha\beta^2 > 1$, one can distinguish $\mathbb{P}_\beta$ from $\mathbb{P}_0$ by looking at the top eigenvalue of the empirical covariance matrix $\boldsymbol{Y}\boldsymbol{Y}^\top$. So the conclusion of Theorem 1 cannot hold in light of the above contiguity argument. Beyond this special case, our result is not expected to be optimal.

Limits of weak detection Since contiguity implies that testing errors are inevitable, it is natural to aim for tests $T : \mathbb{R}^{N\times M} \mapsto \{0, 1\}$ that minimize the sum of the Type-I and Type-II errors:

$$\mathsf{err}(T) = \mathbb{P}_0 \left(T(\boldsymbol{Y}) = 1\right) + \mathbb{P}_\beta \left(T(\boldsymbol{Y}) = 0\right).$$

By the Neyman-Pearson lemma, the test minimizing the above error is the likelihood ratio test that rejects the null iff $L(\boldsymbol{Y}; \beta) > 1$. The optimal error is thus

$$\mathsf{err}^*_{M,N}(\beta) = \mathbb{P}_0 \left(\log L(\boldsymbol{Y}; \beta) > 0\right) + \mathbb{P}_\beta \left(\log L(\boldsymbol{Y}; \beta) \le 0\right) = 1 - D_{\mathsf{TV}}(\mathbb{P}_\beta, \mathbb{P}_0).$$

The symmetry of the means under the null and the alternative in Theorem 1 implies that the above Type-I and Type-II errors are equal, and that the total error has a limit:

Corollary 3. *For $\alpha, \beta \ge 0$ such that $K_{\mathrm{u}}^4 K_{\mathrm{v}}^4 \alpha\beta^2 < 1$,*

$$\lim_{\substack{N\to\infty \\ M/N\to\alpha}} \mathsf{err}^*_{M,N}(\beta) = 1 - \lim_{\substack{N\to\infty \\ M/N\to\alpha}} D_{\mathsf{TV}}(\mathbb{P}_\beta, \mathbb{P}_0) = \mathsf{erfc}\left(\frac{1}{4}\sqrt{-\log\left(1 - \alpha\beta^2\right)}\right),$$

where $\mathsf{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} \mathrm{d}t$ is the complementary error function.

Furthermore, our proof of Theorem 1 allows us obtain the convergence of the mean (actually, all moments of $\log L$) under $\mathbb{P}_\beta$, which corresponds to the Kullback-Liebler divergence of $\mathbb{P}_\beta$ to $\mathbb{P}_0$:

Proposition 4. *For all $\alpha, \beta \ge 0$ such that $K_{\mathrm{u}}^4 K_{\mathrm{v}}^4 \alpha\beta^2 < 1$,*

$$\lim_{\substack{N\to\infty \\ M/N\to\alpha}} D_{\mathsf{KL}}(\mathbb{P}_\beta, \mathbb{P}_0) = -\frac{1}{4}\log\left(1 - \alpha\beta^2\right).$$

3 Replicas, overlaps, Gibbs measures and Nishimori

A crucial component of the proof involves understanding the convergence properties of certain overlaps between "replicas." To embark on the argument let us introduce some important notation and terminology. Let $H : \mathbb{R}^{N+M} \to \mathbb{R}$ be the (random) function, which we refer to as a *Hamiltonian*, defined as

$$-H(\boldsymbol{u}, \boldsymbol{v}) = \sum_{i,j} \sqrt{\frac{\beta}{N}} Y_{ij} u_i v_j - \frac{\beta}{2N} u_i^2 v_j^2, \tag{2}$$

where $\boldsymbol{Y} = (Y_{ij})$ comes from $\mathbb{P}_\beta$ or $\mathbb{P}_0$. Letting ρ denote the product measure $P_{\mathtt{u}}^{\otimes N} \otimes P_{\mathtt{v}}^{\otimes M}$, we have

$$L(\boldsymbol{Y};\beta) = \int e^{-H(\boldsymbol{u},\boldsymbol{v})} \mathrm{d}\rho(\boldsymbol{u},\boldsymbol{v}).$$

Let us define the Gibbs average of a function $f : (\mathbb{R}^{N+M})^n \mapsto \mathbb{R}$ of n replica pairs $(\boldsymbol{u}^{(l)}, \boldsymbol{v}^{(l)})_{l=1}^n$ with respect to the Hamiltonian H as

$$\langle f \rangle = \frac{\int f \prod_{l=1}^n e^{-H(\boldsymbol{u}^{(l)},\boldsymbol{v}^{(l)})} \mathrm{d}\rho(\boldsymbol{u}^{(l)},\boldsymbol{v}^{(l)})}{\left(\int e^{-H(\boldsymbol{u},\boldsymbol{v})} \mathrm{d}\rho(\boldsymbol{u},\boldsymbol{v})\right)^n}. \tag{3}$$

This is the mean of f with respect to the posterior distribution of $(\boldsymbol{u},\boldsymbol{v})$ given $\boldsymbol{Y}$: $\mathbb{P}_\beta(\cdot|\boldsymbol{Y})^{\otimes n}$. We interpret the replicas as random and independent draws from this posterior. When $\boldsymbol{Y} \sim \mathbb{P}_\beta$ we also allow f to depend on the spike pair $(\boldsymbol{u}^*, \boldsymbol{v}^*)$. For two different replicas $(\boldsymbol{u}^{(l)}, \boldsymbol{v}^{(l)})$ and $(\boldsymbol{u}^{(l')}, \boldsymbol{v}^{(l')})$ (l' is allowed to take the value $*$) we denote the overlaps of the $\mathtt{u}$ and $\mathtt{v}$ parts, both normalized by N, as

$$R^{\mathtt{u}}_{l,l'} = \frac{1}{N}\sum_{i=1}^N u_i^{(l)} u_i^{(l')} \quad \text{and} \quad R^{\mathtt{v}}_{l,l'} = \frac{1}{N}\sum_{j=1}^M v_j^{(l)} v_j^{(l')}.$$

3.1 The Nishimori property under $\mathbb{P}_\beta$

Let's perform the following experiment:

1. Construct $\boldsymbol{u}^* \in \mathbb{R}^N$ and $\boldsymbol{v}^* \in \mathbb{R}^M$ by independently drawing their coordinates from $P_{\mathtt{u}}$ and $P_{\mathtt{v}}$ respectively.
2. Construct $\boldsymbol{Y} = \sqrt{\frac{\beta}{N}}\boldsymbol{u}^*\boldsymbol{v}^{*\top} + \boldsymbol{W}$, where $W_{ij} \sim \mathcal{N}(0,1)$ are all independent. ($\boldsymbol{Y}$ is distributed according to $\mathbb{P}_\beta$.)
3. Draw $n+1$ independent random vector pairs, $(\boldsymbol{u}^{(l)}, \boldsymbol{v}^{(l)})_{l=1}^{n+1}$, from $\mathbb{P}_\beta((\boldsymbol{u},\boldsymbol{v}) \in \cdot|\boldsymbol{Y})$.

By the tower property of expectations, the following equality of joint laws holds

$$\Big(\boldsymbol{Y}, (\boldsymbol{u}^{(1)}, \boldsymbol{v}^{(1)}), \cdots, (\boldsymbol{u}^{(n+1)}, \boldsymbol{v}^{(n+1)})\Big) \stackrel{\mathrm{d}}{=} \Big(\boldsymbol{Y}, (\boldsymbol{u}^{(1)}, \boldsymbol{v}^{(1)}), \cdots, (\boldsymbol{u}^{(n)}, \boldsymbol{v}^{(n)}), (\boldsymbol{u}^*, \boldsymbol{v}^*)\Big). \tag{4}$$

This in particular implies that under the alternative $\mathbb{P}_\beta$, the overlaps $(R^{\mathtt{u}}_{1,*}, R^{\mathtt{v}}_{1,*})$ between replica and spike pairs have the same distribution as the overlaps $(R^{\mathtt{u}}_{1,2}, R^{\mathtt{v}}_{1,2})$ between two replica pairs. This is a very important property of the planted (spiked) model, which is usually named after Nishimori (2001). It allows for manipulations that are not possible under the null. For instance, to prove the convergence of the overlap between two replicas, $\mathbb{E}\langle (R^{\mathtt{u}}_{1,2})^2\rangle \to 0$, it suffices to prove $\mathbb{E}\langle (R^{\mathtt{u}}_{1,*})^2\rangle \to 0$ since the two quantities are equal. The latter turns out to be a much easier task.

3.2 Overlap decay implies super-concentration

Let us now explain how the behavior of the overlaps is related to the fluctuations of $\log L$. For concreteness we consider the null model as an example. Let $\boldsymbol{Y} \sim \mathbb{P}_0$, i.e., $Y_{ij} \sim \mathcal{N}(0,1)$ all independent. The log-likelihood ratio, seen as a function of $\boldsymbol{Y}$, is a differentiable function, and

$$\frac{\mathrm{d}}{\mathrm{d}Y_{ij}} \log L(\boldsymbol{Y};\beta) = \sqrt{\frac{\beta}{N}} \langle u_i v_j \rangle.$$

By the Gaussian Poincaré inequality, we can bound the variance by the norm of the gradient as

$$\mathbb{E}\left[(\log L - \mathbb{E}\log L)^2\right] \leq \mathbb{E}\left[\|\nabla \log L\|_{\ell_2}^2\right] = \beta N\, \mathbb{E}\langle R^{\mathtt{u}}_{1,2} R^{\mathtt{v}}_{1,2}\rangle.$$

The last equality follows from the fact $\langle u_i v_j\rangle^2 = \langle u_i^{(1)} v_j^{(1)} u_i^{(2)} v_j^{(2)}\rangle$. Since our priors have bounded support, we can already bound $R^{\mathtt{u}}_{1,2} R^{\mathtt{v}}_{1,2}$ by $\frac{M}{N} K_{\mathtt{u}}^2 K_{\mathtt{v}}^2$, and we deduce that the variance is $\mathcal{O}(N)$. In fact, by the Maurey-Pisier inequality (Pisier, 1986, Theorem 2.2), we can control the moment generating function of $\log L$ by that of $N\langle R^{\mathtt{u}}_{1,2} R^{\mathtt{v}}_{1,2}\rangle$. This implies sub-Gaussian concentration of the former. Observe now that if the quantity $\mathbb{E}\langle R^{\mathtt{u}}_{1,2} R^{\mathtt{v}}_{1,2}\rangle$ decays, then the much stronger result $\mathrm{var}(\log L) = o(N)$ holds. This behavior of unusually small variance is often referred to as "super-concentration." See Chatterjee (2014) for more on this topic. In our case, not only does $\mathbb{E}\langle R^{\mathtt{u}}_{1,2} R^{\mathtt{v}}_{1,2}\rangle$ decay when α and β are sufficiently small, but it does so at a rate of $1/N$ so that $N\,\mathbb{E}\langle R^{\mathtt{u}}_{1,2} R^{\mathtt{v}}_{1,2}\rangle$ converges to a finite limit, and $\mathrm{var}(\log L)$ is constant. This is a first reason why Theorem 1 should be expected: if anything, the fluctuations must be of constant order.

4 Proof of Theorem 1

It suffices to prove the fluctuations under one of the hypotheses. Fluctuations under the remaining one comes for free as a consequence of Le Cam's third lemma (or more specifically, the Portmanteau theorem Van der Vaart, 2000, Theorem 6.6). For the reader's convenience, we present this argument in Appendix A. We choose to treat the planted case $\boldsymbol{Y} \sim \mathbb{P}_\beta$. The reason is that we are able to achieve control on the overlaps and show their concentration under the alternative in a wider region of parameters (α, β) than under the null. This is ultimately due to the (generalized) Nishimori property (4).

We will show the convergence of the characteristic function of $\log L$ to that of a Gaussian. Let $\mu = -\frac{1}{4}\log(1-\alpha\beta^2)$, $\sigma^2 = -\frac{1}{2}\log(1-\alpha\beta^2)$, and let ϕ be the characteristic function of the Gaussian distribution $\mathcal{N}(\mu, \sigma^2)$: for $s \in \mathbb{R}$ and $\mathrm{i}^2 = -1$, let $\phi(s) = \exp\{\mathrm{i}s\mu - \frac{\sigma^2}{2}s^2\}$. The following is a more quantitative convergence result that implies Theroem 1.

Theorem 5. *Let $s \in \mathbb{R}$ and $\alpha, \beta \geq 0$. There exists $K = K(s, \alpha, \beta, K_{\mathtt{u}}, K_{\mathtt{v}}) < \infty$ such that for M, N sufficiently large and $M = \alpha N + \mathcal{O}(\sqrt{N})$, the following holds. If $\alpha\beta^2 K_{\mathtt{u}}^4 K_{\mathtt{v}}^4 < 1$, then*

$$\left|\mathbb{E}_{\mathbb{P}_\beta}\left[e^{\mathrm{i}s\log L(\boldsymbol{Y};\beta)}\right] - \phi(s)\right| \leq \frac{K}{\sqrt{N}}.$$

Our approach is to show that the function

$$\phi_N(\beta) = \mathbb{E}_{\mathbb{P}_\beta}\left[e^{\mathrm{i}s\log L(\boldsymbol{Y};\beta)}\right]$$

(for $s \in \mathbb{R}$ fixed) is an approximate solution to a differential equation whose solution is the characteristic function of the Gaussian.

Lemma 6. *For all $\beta \geq 0$, it holds that*

$$\frac{\mathrm{d}}{\mathrm{d}\beta}\phi_N(\beta) = \frac{\mathrm{i}s - s^2}{2} N\,\mathbb{E}\left[\left\langle R^{\mathtt{u}}_{1,2} R^{\mathtt{v}}_{1,2}\right\rangle e^{\mathrm{i}s\log L}\right]. \tag{5}$$

Proof. Since $\boldsymbol{Y} \sim \mathbb{P}_\beta$, we can rewrite the Hamiltonian (2) as

$$\begin{aligned}-H(\boldsymbol{u},\boldsymbol{v}) &= \sum_{i,j} \sqrt{\frac{\beta}{N}} Y_{ij} u_i v_j - \frac{\beta}{2N} u_i^2 v_j^2, \\ &= \sum_{i,j} \sqrt{\frac{\beta}{N}} W_{ij} u_i v_j + \frac{\beta}{N} u_i v_j u_i^* v_j^* - \frac{\beta}{2N} u_i^2 v_j^2.\end{aligned}$$

We take a derivative with respect to β:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\beta}\phi_N(\beta) &= \mathrm{i}s\,\mathbb{E}\left[\left\langle -\frac{\mathrm{d}H}{\mathrm{d}\beta}\right\rangle e^{\mathrm{i}s\log L}\right] \\ &= \mathrm{i}s \sum_{i,j}\left(\frac{1}{2\sqrt{\beta N}}\,\mathbb{E}\left[W_{ij}\left\langle u_i v_j\right\rangle e^{\mathrm{i}s\log L}\right] - \frac{1}{2N}\,\mathbb{E}\left[\left\langle u_i^2 v_j^2\right\rangle e^{\mathrm{i}s\log L}\right]\right) \\ &\quad + \mathrm{i}s\frac{1}{N}\sum_{i,j}\mathbb{E}\left[\left\langle u_i v_j u_i^* v_j^*\right\rangle e^{\mathrm{i}s\log L}\right].\end{aligned}$$

The last term is equal to $\mathrm{i}sN\,\mathbb{E}[\langle R^{\mathtt{u}}_{1,*}R^{\mathtt{v}}_{1,*}\rangle e^{\mathrm{i}s\log L}]$. As for the first term, since $W_{ij} \overset{\text{ind.}}{\sim} \mathcal{N}(0,1)$, we use Gaussian integration by parts to obtain

$$\begin{aligned}\mathbb{E}\left[W_{ij}\left\langle u_i v_j\right\rangle e^{\mathrm{i}s\log L}\right] &= \mathbb{E}\left[\frac{\mathrm{d}}{\mathrm{d}W_{ij}}\left(\left\langle u_i v_j\right\rangle e^{\mathrm{i}s\log L}\right)\right] \\ &= \sqrt{\frac{\beta}{N}}\left(\mathbb{E}\left[\left\langle u_i^2 v_j^2\right\rangle e^{\mathrm{i}s\log L}\right] - \mathbb{E}\left[\left\langle u_i v_j\right\rangle^2 e^{\mathrm{i}s\log L}\right] + \mathrm{i}s\,\mathbb{E}\left[\left\langle u_i v_j\right\rangle^2 e^{\mathrm{i}s\log L}\right]\right).\end{aligned}$$

Regrouping terms, we get

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\beta}\phi_N(\beta) &= -\mathrm{i}s\frac{N}{2}\,\mathbb{E}\left[\left\langle R^{\mathtt{u}}_{1,2}R^{\mathtt{v}}_{1,2}\right\rangle e^{\mathrm{i}s\log L}\right] + \mathrm{i}sN\,\mathbb{E}\left[\left\langle R^{\mathtt{u}}_{1,*}R^{\mathtt{v}}_{1,*}\right\rangle e^{\mathrm{i}s\log L}\right] \\ &\quad + (\mathrm{i}s)^2\frac{N}{2}\,\mathbb{E}\left[\left\langle R^{\mathtt{u}}_{1,2}R^{\mathtt{v}}_{1,2}\right\rangle e^{\mathrm{i}s\log L}\right].\end{aligned} \tag{6}$$

The first and third terms in (6) contain overlaps between two replicas while the middle term contains an overlap between one replica and the spike vectors. By the Nishimori property (4), we can replace the spike by a second replica in the overlaps appearing in the middle term, and this finishes the proof. ■

A heuristic argument Let us now heuristically consider what should happen. A rigorous argument will be presented shortly. If the quantity $N\langle R^{\mathtt{u}}_{1,2}R^{\mathtt{v}}_{1,2}\rangle$ concentrates very strongly about some deterministic value $\theta = \theta(\alpha,\beta)$, we would expect that the Gibbs averages in (5) would behave approximately independently from $\log L$, and we would obtain the following differential equation

$$\frac{\mathrm{d}}{\mathrm{d}\beta}\phi_N(\beta) \simeq \frac{1}{2}\left(\mathrm{i}s - s^2\right)\theta\phi_N(\beta).$$

Since $\phi_N(0) = 1$, one obtains $\phi_N(\beta) \simeq \exp\{\frac{1}{2}(\mathrm{i}s - s^2)\int_0^\beta \theta\,\mathrm{d}\beta'\}$ by integrating over β, and the result would follow. The concentration assumption we used is commonly referred to as *replica-symmetry* or *the replica-symmetric ansatz* in the statistical physics literature. Most of the difficulty of the proof lies in showing rigorously that replica symmetry indeed holds.

Sign symmetry between $\mathbb{P}_\beta$ and $\mathbb{P}_0$ One can execute the same argument under the null model. Since there is no planted term in the Hamiltonian, the analogue of (6) one obtains does not contain the middle term. Hence the differential equation one obtains is

$$\frac{\mathrm{d}}{\mathrm{d}\beta}\phi_N(\beta) \simeq \frac{1}{2}\left(-\mathrm{i}s - s^2\right)\theta\phi_N(\beta).$$

This is one way to interpret the sign symmetry of the means of the limiting Gaussians under the null and the alternative: the interaction of one replica with the planted spike under the planted model accounts for twice the contribution of the interaction between two independent replicas, and this flips the sign of the mean.

We now replace the above heuristic with a rigorous statement. Recall that $\boldsymbol{Y} \sim \mathbb{P}_\beta$.

Proposition 7. *For $s \in \mathbb{R}$ and $\alpha, \beta \geq 0$ such that $\alpha\beta^2 K_{\mathtt{u}}^4 K_{\mathtt{v}}^4 < 1$, there exist a constant $K = K(s, \alpha, \beta, K_{\mathtt{u}}, K_{\mathtt{v}}) < \infty$ such that*

$$N\,\mathbb{E}\left[\left\langle R_{1,2}^{\mathtt{u}} R_{1,2}^{\mathtt{v}}\right\rangle e^{\mathrm{i}s\log L}\right] = \frac{\alpha\beta}{1-\alpha\beta^2}\,\mathbb{E}\left[e^{\mathrm{i}s\log L}\right] + \delta,$$

where $|\delta| \leq K/\sqrt{N}$. Moreover, K, seen as a function of β, is bounded on any interval $[0, \beta']$ when $\alpha\beta'^2 K_{\mathtt{u}}^4 K_{\mathtt{v}}^4 < 1$.

Taking $s = 0$, we see that $\theta = \frac{\alpha\beta}{1-\alpha\beta^2}$. Proposition 7 vindicates replica symmetry, and its proof occupies the majority of the rest of the manuscript.

Proof of Theorem 5. Plugging the results of Proposition 7 in the derivative computed in Lemma 6, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}\beta}\phi_N(\beta) = \left(\frac{\mathrm{i}s - s^2}{2}\frac{\alpha\beta}{1-\alpha\beta^2}\right)\phi_N(\beta) + \delta,$$

where $|\delta| \leq \frac{K}{\sqrt{N}}\max\{|s|, s^2\}$, and K is the constant from Proposition 7. Integrating w.r.t. β we obtain

$$|\phi_N(\beta) - \phi(s)| \leq \frac{K'}{\sqrt{N}},$$

where K' depends on α, β, s and $K_{\mathtt{u}}, K_{\mathtt{v}}$, and $K' < \infty$ as long as $\alpha\beta^2 K_{\mathtt{u}}^4 K_{\mathtt{v}}^4 < 1$. ■

Let us prove in passing the convergence of the KL divergence between the null and alternative.

Proof of Proposition 4. Similarly to the computation of the derivative of ϕ_N, we can obtain

$$\frac{\mathrm{d}}{\mathrm{d}\beta}\mathbb{E}_{\mathbb{P}_\beta}\log L(\boldsymbol{Y};\beta) = -\frac{N}{2}\mathbb{E}\left\langle R_{1,2}^{\mathtt{u}} R_{1,2}^{\mathtt{v}}\right\rangle + N\,\mathbb{E}\left\langle R_{1,*}^{\mathtt{u}} R_{1,*}^{\mathtt{v}}\right\rangle = \frac{N}{2}\mathbb{E}\left\langle R_{1,2}^{\mathtt{u}} R_{1,2}^{\mathtt{v}}\right\rangle,$$

where we used the Nishimori property. By Proposition 7 with $s = 0$, this derivative is $K/\sqrt{N}$ away from $\frac{1}{2}\frac{\alpha\beta}{1-\alpha\beta^2}$. Integration and boundedness of K finishes the proof. ■

5 Overlap convergence

The question of overlap convergence is purely a spin glass problem. We will use the machinery developed by Talagrand to solve it. In particular, a crucial use is made of the cavity method and Guerra's interpolation scheme. In this section, we present the main underlying ideas. The arguments are technically involved (but conceptually simple) so we delay their full execution to the Appendix. We refer to Talagrand (2007) for a leisurely high-level introduction to these ideas.

5.1 Sketch of proof of Proposition 7

The basic idea is to show that the quantities of interest approximately obey a self-consistent (or self-bounding) property, the error terms of which can be controlled. This approach will be used at different stages of the proof. We will show that

$$N\,\mathbb{E}\left[\left\langle R^{\mathtt{u}}_{1,2}R^{\mathtt{v}}_{1,2}\right\rangle e^{\mathrm{i}s\log L}\right] = \alpha\beta\,\mathbb{E}\left[e^{\mathrm{i}s\log L}\right] + \alpha\beta^2 N\,\mathbb{E}\left[\left\langle R^{\mathtt{u}}_{1,2}R^{\mathtt{v}}_{1,2}\right\rangle e^{\mathrm{i}s\log L}\right] + \delta,$$

where δ is the error term. This will be achieved in two steps. We first prove

$$N\,\mathbb{E}\left[\left\langle R^{\mathtt{u}}_{1,2}R^{\mathtt{v}}_{1,2}\right\rangle e^{\mathrm{i}s\log L}\right] = N\beta\,\mathbb{E}\left[\left\langle (R^{\mathtt{v}}_{1,2})^2\right\rangle e^{\mathrm{i}s\log L}\right] + \delta, \tag{7}$$

via a cavity on N, i.e., by isolating the effect of the last variable u_N on the rest of the variables. We then show

$$N\,\mathbb{E}\left[\left\langle (R^{\mathtt{v}}_{1,2})^2\right\rangle e^{\mathrm{i}s\log L}\right] = \frac{M}{N}\,\mathbb{E}\left[e^{\mathrm{i}s\log L}\right] + M\beta\,\mathbb{E}\left[\left\langle R^{\mathtt{u}}_{1,2}R^{\mathtt{v}}_{1,2}\right\rangle e^{\mathrm{i}s\log L}\right] + \delta, \tag{8}$$

via a cavity on M, i.e., isolating the effect of v_M. In the arguments leading to (7) and (8), we accumulate error terms that are proportional to the third moments of the overlaps:

$$\delta \lesssim N\,\mathbb{E}\left\langle |R^{\mathtt{u}}_{1,2}|^3\right\rangle + N\,\mathbb{E}\left\langle |R^{\mathtt{v}}_{1,2}|^3\right\rangle, \tag{9}$$

where we hide constants depending on α and β. These cavity equations impose only a mild restriction on the parameters so that our bounds go in the right direction, namely that $\alpha\beta^2 < 1$. This is about to change. We prove that $\delta = \mathcal{O}(1/\sqrt{N})$ with methods that impose the stronger restrictions on (α,β) that ultimately appear in the final result.

5.2 Convergence in the planted model: from crude estimates to optimal rates

We prove overlap convergence under the alternative. Let $\boldsymbol{Y} \sim \mathbb{P}_\beta$.

Proposition 8. *For all $\alpha, \beta \geq 0$ such that $K_{\mathtt{u}}^4 K_{\mathtt{v}}^4 \alpha\beta^2 < 1$, there exists $K = K(\alpha,\beta) < \infty$ such that*

$$\mathbb{E}\left\langle (R^{\mathtt{u}}_{1,2})^4\right\rangle \vee \mathbb{E}\left\langle (R^{\mathtt{v}}_{1,2})^4\right\rangle \leq \frac{K}{N^2}.$$

The proof proceeds as follows. We use the cavity method to show the following self-consistency equations:

$$\mathbb{E}\left\langle (R^{\mathtt{u}}_{1,2})^4\right\rangle = \alpha\beta^2\,\mathbb{E}\left\langle (R^{\mathtt{u}}_{1,2})^4\right\rangle + \overline{M}_{\mathtt{u}} + \delta_{\mathtt{u}}, \tag{10}$$
$$\mathbb{E}\left\langle (R^{\mathtt{v}}_{1,2})^4\right\rangle = \alpha\beta^2\,\mathbb{E}\left\langle (R^{\mathtt{v}}_{1,2})^4\right\rangle + \overline{M}_{\mathtt{v}} + \delta_{\mathtt{v}}, \tag{11}$$

where $|\overline{M}_{\mathtt{u}}|, |\overline{M}_{\mathtt{v}}|$ are bounded by sums of expectations of monomials of degree five in the overlaps $R^{\mathtt{u}}$ and $R^{\mathtt{v}}$:

$$|\overline{M}_{\mathtt{u}}| \lesssim \sum_{a,b,c,d} \mathbb{E}\left\langle \left|(R^{\mathtt{u}}_{1,2})^3 R^{\mathtt{u}}_{a,b}R^{\mathtt{u}}_{c,d}\right|\right\rangle + \mathbb{E}\left\langle \left|(R^{\mathtt{u}}_{1,2})^3 R^{\mathtt{v}}_{a,b}R^{\mathtt{v}}_{c,d}\right|\right\rangle,$$

$$|\overline{M}_{\mathtt{v}}| \lesssim \sum_{a,b,c,d} \mathbb{E}\left\langle \left|(R^{\mathtt{v}}_{1,2})^3 R^{\mathtt{v}}_{a,b}R^{\mathtt{v}}_{c,d}\right|\right\rangle + \mathbb{E}\left\langle \left|(R^{\mathtt{v}}_{1,2})^3 R^{\mathtt{u}}_{a,b}R^{\mathtt{u}}_{c,d}\right|\right\rangle,$$

where the sum is over a finite number of combinations (a, b, c, d), and

$$\delta_{\mathtt{u}} \lesssim \frac{1}{N}\,\mathbb{E}\left\langle (R^{\mathtt{u}}_{1,2})^2 \right\rangle + \mathcal{O}\Big(\frac{1}{N^2}\Big), \qquad \delta_{\mathtt{v}} \lesssim \frac{1}{N}\,\mathbb{E}\left\langle (R^{\mathtt{v}}_{1,2})^2 \right\rangle + \mathcal{O}\Big(\frac{1}{N^2}\Big).$$

These results hold for *all* $\alpha, \beta \geq 0$. From here, further progress is unlikely unless one has *a priori* knowledge that the overlaps are unlikely to be large, so that the fifth-order terms do not overwhelm the main terms. More precisely, suppose that we are able to prove the following crude bound on the overlaps: for $\epsilon > 0$, there is $K = K(\epsilon, \alpha, \beta) > 0$ such that

$$\mathbb{E}\left\langle \mathbb{1}\left\{\left|R^{\mathtt{u}}_{1,2}\right| \geq \epsilon\right\}\right\rangle \vee \mathbb{E}\left\langle \mathbb{1}\left\{\left|R^{\mathtt{v}}_{1,2}\right| \geq \epsilon\right\}\right\rangle \leq K e^{-N/K}. \tag{12}$$

Then the fifth-order terms can be controlled by fourth-order terms as follows:

$$\begin{aligned}\mathbb{E}\left\langle \left|(R^{\mathtt{u}}_{1,2})^3 R^{\mathtt{v}}_{a,b} R^{\mathtt{v}}_{c,d}\right| \right\rangle &\leq \epsilon\,\mathbb{E}\left\langle \left|(R^{\mathtt{u}}_{1,2})^3 R^{\mathtt{v}}_{a,b}\right| \right\rangle + K_{\mathtt{u}}^6 K_{\mathtt{v}}^4 K e^{-N/K}\\ &\leq \epsilon M + K e^{-N/K},\end{aligned}$$

where $M = \mathbb{E}\langle (R^{\mathtt{u}}_{1,2})^4\rangle \vee \mathbb{E}\langle (R^{\mathtt{v}}_{1,2})^4\rangle$, and the last step is by Hölder's inequality. This way, $\overline{M}_{\mathtt{u}}$ and $\overline{M}_{\mathtt{v}}$ are controlled. Now it remains to control $\delta_{\mathtt{u}}$ and $\delta_{\mathtt{v}}$. We could re-execute the cavity argument on the second moment instead of the fourth, and this would allow us to obtain $\mathbb{E}\langle (R^{\mathtt{u}}_{1,2})^2\rangle \vee \mathbb{E}\langle (R^{\mathtt{v}}_{1,2})^2\rangle \leq K/N$. We instead use a shorter argument based on an elegant *quadratic replica coupling* technique of Guerra and Toninelli (2002) to prove this. This is presented in Appendix D.1. Plugging these estimates into (10) and (11), we obtain

$$\begin{aligned}\mathbb{E}\left\langle (R^{\mathtt{u}}_{1,2})^4 \right\rangle &\leq \alpha\beta^2\,\mathbb{E}\left\langle (R^{\mathtt{u}}_{1,2})^4 \right\rangle + K\epsilon M + \delta',\\ \mathbb{E}\left\langle (R^{\mathtt{v}}_{1,2})^4 \right\rangle &\leq \alpha\beta^2\,\mathbb{E}\left\langle (R^{\mathtt{v}}_{1,2})^4 \right\rangle + K\epsilon M + \delta',\end{aligned}$$

where $\delta' \leq K/N^2 + Ke^{-N/K}$, and this implies the desired result for ϵ sufficiently small.

The a priori bound (12) is proved via an interpolation argument at fixed overlap, combined with concentration of measure, and is presented in Appendices D.2 and D.3. These arguments impose a restriction on the parameters (α, β) that shows up in the final result. Finally, Proposition 8 allow us to conclude (via Jensen's inequality) that the error term δ displayed in (9) is bounded by $K/\sqrt{N}$.

6 Discussion

The limiting factor in our approach to prove LR fluctuations is the need for precise non-asymptotic control of moments of the overlaps $R^{\mathtt{u}}_{1,2}$ and $R^{\mathtt{v}}_{1,2}$ under the expected Gibbs measure $\mathbb{E}\langle\cdot\rangle$. We were able to reach this level of control only in a restricted regime. This is due to the failure of our approach to prove the crude estimate (12) in a larger region. In this section, we formulate a conjecture on the largest region where these fluctuations and overlap decay should occur. In one sentence, this should be the entire *annealed* or *paramagnetic* region of the model, as dictated by the vanishing of its replica-symmetric (RS) formula. We shall now be more precise.

Let $z \sim \mathcal{N}(0,1)$, $u^* \sim P_{\mathtt{u}}$ and $v^* \sim P_{\mathtt{v}}$ all independent. Define

$$\begin{aligned}\psi_{\mathtt{u}}(r) &:= \mathbb{E}_{u^*,z} \log \int \exp\Big(\sqrt{r} z u + r u u^* - \frac{r}{2} u^2\Big)\,\mathrm{d}P_{\mathtt{u}}(u),\\ \psi_{\mathtt{v}}(r) &:= \mathbb{E}_{v^*,z} \log \int \exp\Big(\sqrt{r} z v + r v v^* - \frac{r}{2} v^2\Big)\,\mathrm{d}P_{\mathtt{v}}(v).\end{aligned}$$

Moreover, define the RS potential as

$$F(\alpha, \beta, q_{\mathtt{u}}, q_{\mathtt{v}}) := \psi_{\mathtt{u}}(\beta q_{\mathtt{v}}) + \alpha\psi_{\mathtt{v}}(\beta q_{\mathtt{u}}) - \frac{\beta q_{\mathtt{u}} q_{\mathtt{v}}}{2}.$$

and finally define the RS formula as

$$\phi_{\mathsf{RS}}(\alpha, \beta) := \sup_{q_{\mathtt{v}} \geq 0} \inf_{q_{\mathtt{u}} \geq 0} F(\alpha, \beta, q_{\mathtt{u}}, q_{\mathtt{v}}).$$

It was argued by Lesieur et al. (2015a) based on the plausibility of the replica-symmetric ansatz, and then proved by Miolane (2017), that in the limit $N \to \infty$, $M/N \to \alpha$, $\frac{1}{N} \mathbb{E}_{\mathbb{P}_\beta} \log L(\boldsymbol{Y}; \beta) \to \phi_{\mathsf{RS}}(\alpha, \beta)$ for all $\alpha, \beta \geq 0$. (See also Barbier et al., 2017, for results in a more general setup.) Of course, by change of measure and Jensen's inequality,

$$\mathbb{E}_{\mathbb{P}_\beta} \log L(\boldsymbol{Y}; \beta) = \mathbb{E}_{\mathbb{P}_0} L(\boldsymbol{Y}; \beta) \log L(\boldsymbol{Y}; \beta) \geq 0,$$

for all M, N; therefore ϕ_{RS} is always nonnegative. Let

$$\Gamma = \{(\alpha, \beta) \in \mathbb{R}_+ \ : \ \phi_{\mathsf{RS}}(\alpha, \beta) = 0\}.$$

It is not hard to prove the following lemma by analyzing the stability of $(0, 0)$ as a stationary point of the RS potential:

Lemma 9. $\Gamma \subseteq \{(\alpha, \beta) \in \mathbb{R}_+ \ : \ \alpha\beta^2 \leq 1\}$.

This lemma tells us (unsurprisingly) that Γ is entirely below the BBP threshold. The inclusion may or may not be strict depending on the priors $P_{\mathtt{u}}$ and $P_{\mathtt{v}}$. For instance, there is equality of the above sets if $P_{\mathtt{u}}$ and $P_{\mathtt{v}}$ are symmetric Rademacher and/or Gaussian respectively. One case of strict inclusion is when $P_{\mathtt{v}}$ is Gaussian $\mathcal{N}(0, 1)$ and $P_{\mathtt{u}}$ is a sparse Rademacher prior, $\frac{\rho}{2}\delta_{1/\sqrt{\rho}} + (1 - \rho)\delta_0 + \frac{\rho}{2}\delta_{-1/\sqrt{\rho}}$, for sufficiently small ρ (e.g., $\rho = .04$). This is a canonical model for sparse principal component analysis. In this case, there is a region of parameters below the BBP threshold where the posterior mean $\mathbb{E}[\boldsymbol{u}^*|\boldsymbol{Y}]$ ($= \langle \boldsymbol{u} \rangle$ in our notation) has a non-trivial overlap with the spike $\boldsymbol{u}^*$, while the top eigenvector of the empirical covariance matrix $\boldsymbol{Y}\boldsymbol{Y}^\top$ is orthogonal to it. Estimation becomes impossible only in the region Γ, so the following conjecture is highly plausible:

Conjecture 1. *Let Γ' be the interior of Γ. For all $(\alpha, \beta) \in \Gamma'$,*

$$\log L(\boldsymbol{Y}, \beta) \rightsquigarrow \mathcal{N}\left(\pm\frac{1}{4}\log(1 - \alpha\beta^2), -\frac{1}{2}\log(1 - \alpha\beta^2)\right),$$

where the plus sign holds under the null $\mathbb{P}_0$ and the minus sign under the alternative $\mathbb{P}_\beta$.

Our conjecture is formulated only in the interior of Γ; this is not a superfluous condition since diverging behavior may appear at the boundary. Moreover, this conjecture is about the *maximal* region in which such fluctuations can take place. This is not difficult to show. By (sub-Gaussian) concentration of the normalized likelihood ratio, we have for $\epsilon > 0$

$$\mathbb{P}_\beta\left(\frac{1}{N}\log L(\boldsymbol{Y}; \beta) - \phi_{\mathsf{RS}}(\alpha, \beta) \leq -\epsilon\right) \longrightarrow 0,$$

where $K = K(\alpha, \beta) < \infty$. This already shows that $\log L$ must grow with N under the alternative if $\phi_{\mathsf{RS}} > 0$. As for the behavior under the null, the same sub-Gaussian concentration holds, although the expectation is not known (see Question 1):

$$\mathbb{P}_0\left(\frac{1}{N}\log L(\boldsymbol{Y}; \beta) - \frac{1}{N}\mathbb{E}_{\mathbb{P}_0}\log L(\boldsymbol{Y}; \beta) \geq \epsilon\right) \longrightarrow 0.$$

We do know that the above expectation is non-positive, by Jensen's inequality. Therefore if (α, β) are such that $\phi_{\mathsf{RS}} > 0$, one can distinguish $\mathbb{P}_\beta$ from $\mathbb{P}_0$ with asymptotic certainty by testing whether $\frac{1}{N} \log L(\boldsymbol{Y}; \beta)$ is above or below (say) $\frac{1}{2}\phi_{\mathsf{RS}}(\alpha, \beta)$. This implies that $\mathbb{P}_\beta$ and $\mathbb{P}_0$ are not contiguous outside Γ. This—short of proving that $\log L$ grows in the negative direction with N—shows that the fluctuations cannot be of the above form under the null, since this would contradict Le Cam's first lemma.

The difficulty we encountered in our attempts to prove the above conjecture is a loss of control over the overlaps $R^{\mathtt{u}}_{1,2}$ and $R^{\mathtt{v}}_{1,2}$ near the boundary of the set Γ. The interpolation bound at fixed overlap (between a replica and the spike) we used under the alternative $\mathbb{P}_\beta$ is vacuous beyond the region $\alpha\beta^2 < (K_{\mathtt{u}}K_{\mathtt{v}})^{-4}$. It is possible that the latter bound could be marginally improved by more careful analysis, but this is unlikely to yield the optimal result since no information about ϕ_{RS} is used in the proof. One can imagine refining this technique by constraining two replicas and using an interpolation with broken replica-symmetry, in the spirit of the "2D" Guerra-Talagrand bound (Guerra, 2003; Talagrand, 2011b). Although this strategy is successful in the symmetric model where $\boldsymbol{u} = \boldsymbol{v}$ it is not at all obvious why such an interpolation bound should be true in the bipartite case: in the analysis, certain terms that are hard to control have a sign in the symmetric case, hence they can be dropped to obtain a bound. This is no longer true (or at least not obviously so) in the bipartite case.

Another interesting question concerns the LR asymptotics under the null, outside Γ. While under the alternative $\mathbb{P}_\beta$, the normalized log-likelihood ratio converges to the RS formula ϕ_{RS} for all (α, β), no such simple formula is expected to hold under the null. Even the existence of a limit seems to be unknown.

Question 1. *Does $\frac{1}{N}\mathbb{E}_{\mathbb{P}_0} \log L(\boldsymbol{Y}; \beta)$ have a limit for all (α, β)? If so, what is its value?*

We refer to Barra et al. (2011, 2014) and Auffinger and Chen (2014) for some progress on the replica-symmetric phase, and Panchenko (2015) for progress on the related problem of the "multispecies" SK model at all temperatures.

References

Aizenman, M., Lebowitz, J. L., and Ruelle, D. (1987). Some rigorous results on the Sherrington–Kirkpatrick spin glass model. *Communications in Mathematical Physics*, 112(1):3–20.

Amini, A. A. and Wainwright, M. J. (2009). High-dimensional analysis of semidefinite relaxations for sparse principal components. *Annals of Statistics*, 37(5B):2877–2921.

Auffinger, A. and Chen, W.-K. (2014). Free energy and complexity of spherical bipartite models. *Journal of Statistical Physics*, 157(1):40–59.

Bai, Z. and Yao, J. (2012). On sample eigenvalues in a generalized spiked population model. *Journal of Multivariate Analysis*, 106:167–177.

Bai, Z. and Yao, J.-f. (2008). Central limit theorems for eigenvalues in a spiked population model. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 44(3):447–474.

Baik, J., Arous, G. B., and Péché, S. (2005). Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. *Annals of Probability*, 33(5):1643–1697.

Baik, J. and Lee, J. O. (2016). Fluctuations of the free energy of the spherical Sherrington–Kirkpatrick model. *Journal of Statistical Physics*, 165(2):185–224.

Baik, J. and Lee, J. O. (2017a). Fluctuations of the free energy of the spherical Sherrington–Kirkpatrick model with ferromagnetic interaction. In *Annales Henri Poincaré*, volume 18, pages 1867–1917. Springer.

Baik, J. and Lee, J. O. (2017b). Free energy of bipartite spherical Sherrington–Kirkpatrick model. *arXiv preprint arXiv:1711.06364*.

Baik, J. and Silverstein, J. W. (2006). Eigenvalues of large sample covariance matrices of spiked population models. *Journal of Multivariate Analysis*, 97(6):1382–1408.

Banks, J., Moore, C., Vershynin, R., Verzelen, N., and Xu, J. (2017). Information-theoretic bounds and phase transitions in clustering, sparse PCA, and submatrix localization. In *IEEE International Symposium on Information Theory (ISIT)*, pages 1137–1141. IEEE.

Barbier, J., Dia, M., Macris, N., Krzakala, F., Lesieur, T., and Zdeborová, L. (2016). Mutual information for symmetric rank-one matrix estimation: A proof of the replica formula. In *Advances in Neural Information Processing Systems (NIPS)*, pages 424–432.

Barbier, J., Krzakala, F., Macris, N., Miolane, L., and Zdeborová, L. (2017). Phase transitions, optimal errors and optimality of message-passing in generalized linear models. *arXiv preprint arXiv:1708.03395*.

Barra, A., Galluzzi, A., Guerra, F., Pizzoferrato, A., and Tantari, D. (2014). Mean field bipartite spin models treated with mechanical techniques. *The European Physical Journal B*, 87(3):74.

Barra, A., Genovese, G., and Guerra, F. (2011). Equilibrium statistical mechanics of bipartite spin systems. *Journal of Physics A: Mathematical and Theoretical*, 44(24):245002.

Benaych-Georges, F. and Nadakuditi, R. R. (2011). The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices. *Advances in Mathematics*, 227(1):494–521.

Berthet, Q. and Rigollet, P. (2013). Optimal detection of sparse principal components in high dimension. *Annals of Statistics*, 41(4):1780–1815.

Boucheron, S., Lugosi, G., and Massart, P. (2013). *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford University Press.

Capitaine, M., Donati-Martin, C., and Féral, D. (2009). The largest eigenvalues of finite rank deformation of large Wigner matrices: convergence and nonuniversality of the fluctuations. *Annals of Probability*, pages 1–47.

Chatterjee, S. (2014). *Superconcentration and Related Topics*. Springer.

Deshpande, Y., Abbé, E., and Montanari, A. (2016). Asymptotic mutual information for the binary stochastic block model. In *IEEE International Symposium on Information Theory (ISIT)*, pages 185–189. IEEE.

Dobriban, E. (2017). Sharp detection in PCA under correlations: all eigenvalues matter. *Annals of Statistics*, 45(4):1810–1833.

Féral, D. and Péché, S. (2007). The largest eigenvalue of rank one deformation of large Wigner matrices. *Communications in Mathematical Physics*, 272(1):185–228.

Guerra, F. (2003). Broken replica symmetry bounds in the mean field spin glass model. *Communications in Mathematical Physics*, 233(1):1–12.

Guerra, F. and Toninelli, F. L. (2002). Quadratic replica coupling in the Sherrington–Kirkpatrick mean field spin glass model. *Journal of Mathematical Physics*, 43(7):3704–3716.

Johnstone, I. M. (2001). On the distribution of the largest eigenvalue in principal components analysis. *Annals of Statistics*, pages 295–327.

Johnstone, I. M. and Lu, A. Y. (2009). On consistency and sparsity for principal components analysis in high dimensions. *Journal of the American Statistical Association*, 104(486):682–693.

Johnstone, I. M. and Onatski, A. (2015). Testing in high-dimensional spiked models. *arXiv preprint arXiv:1509.07269*.

Korada, S. B. and Macris, N. (2009). Exact solution of the gauge symmetric p-spin glass model on a complete graph. *Journal of Statistical Physics*, 136(2):205–230.

Krzakala, F., Xu, J., and Zdeborová, L. (2016). Mutual information in rank-one matrix estimation. In *Information Theory Workshop (ITW)*, pages 71–75. IEEE.

Ledoit, O. and Wolf, M. (2002). Some hypothesis tests for the covariance matrix when the dimension is large compared to the sample size. *Annals of Statistics*, pages 1081–1102.

Lelarge, M. and Miolane, L. (2017). Fundamental limits of symmetric low-rank matrix estimation. In Kale, S. and Shamir, O., editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 1297–1301, Amsterdam, Netherlands. PMLR (arXiv:1611.03888).

Lesieur, T., Krzakala, F., and Zdeborová, L. (2015a). MMSE of probabilistic low-rank matrix estimation: Universality with respect to the output channel. In *Communication, Control, and Computing (Allerton), 2015 53rd Annual Allerton Conference on*, pages 680–687. IEEE.

Lesieur, T., Krzakala, F., and Zdeborová, L. (2015b). Phase transitions in sparse PCA. In *IEEE International Symposium on Information Theory (ISIT)*, pages 1635–1639. IEEE.

Lesieur, T., Krzakala, F., and Zdeborová, L. (2017). Constrained low-rank matrix estimation: Phase transitions, approximate message passing and applications. *Journal of Statistical Mechanics: Theory and Experiment*, 2017(7).

Miolane, L. (2017). Fundamental limits of low-rank matrix estimation. *arXiv preprint arXiv:1702.00473*.

Montanari, A., Reichman, D., and Zeitouni, O. (2015). On the limitation of spectral methods: From the gaussian hidden clique problem to rank-one perturbations of gaussian tensors. In *Advances in Neural Information Processing Systems*, pages 217–225.

Nadler, B. (2008). Finite sample approximation results for principal component analysis: A matrix perturbation approach. *Annals of Statistics*, pages 2791–2817.

Nishimori, H. (2001). *Statistical physics of spin glasses and information processing: an introduction*, volume 111. Clarendon Press.

Onatski, A., Moreira, M. J., and Hallin, M. (2013). Asymptotic power of sphericity tests for high-dimensional data. *Annals of Statistics*, 41(3):1204–1231.

Onatski, A., Moreira, M. J., and Hallin, M. (2014). Signal detection in high dimension: The multispiked case. *Annals of Statistics*, 42(1):225–254.

Panchenko, D. (2015). The free energy in a multi-species Sherrington–Kirkpatrick model. *Annals of Probability*, 43(6):3494–3513.

Paul, D. (2007). Asymptotics of sample eigenstructure for a large dimensional spiked covariance model. *Statistica Sinica*, pages 1617–1642.

Péché, S. (2006). The largest eigenvalue of small rank perturbations of Hermitian random matrices. *Probability Theory and Related Fields*, 134(1):127–173.

Péché, S. (2014). Deformed ensembles of random matrices. In *Proceedings of the International Congress of Mathematicians, Seoul*, volume III, pages 1059–1174. ICM.

Perry, A., Wein, A. S., Bandeira, A. S., and Moitra, A. (2016). On the optimality and sub-optimality of PCA for spiked random matrix models. *Annals of Statistics (to appear). arXiv preprint arXiv:1609.05573*.

Pisier, G. (1986). *Probabilistic methods in the geometry of Banach spaces*, pages 167–241. Springer, Berlin, Heidelberg.

Talagrand, M. (2007). Mean field models for spin glasses: some obnoxious problems. In *Spin Glasses*, pages 63–80. Springer.

Talagrand, M. (2011a). *Mean field models for spin glasses. Volume I: Basic examples*, volume 54. Springer Science & Business Media.

Talagrand, M. (2011b). *Mean field models for spin glasses. Volume II: Advanced replica-symmetry and low temperature*, volume 55. Springer Science & Business Media.

Van der Vaart, A. W. (2000). *Asymptotic Statistics*. Cambridge University Press.

A Fluctuation equivalence

We explain in this appendix how the fluctuation result under $\mathbb{P}_\beta$ implies the corresponding fluctuation result under $\mathbb{P}_0$. This is a consequence of the Portmanteau characterization of convergence in distribution. The argument can be made in the other direction as well. Assume that

$$\log L(\boldsymbol{Y};\beta) \rightsquigarrow \mathcal{N}(\mu, \sigma^2),$$

for $\boldsymbol{Y} \sim \mathbb{P}_\beta$, where $\mu = \frac{1}{2}\sigma^2$. By the Portmanteau theorem (Van der Vaart, 2000, Lemma 2.2), this is equivalent to the assertion

$$\liminf \mathbb{E}_{\mathbb{P}_\beta}\left[f(\log L)\right] \geq \mathbb{E}\left[f(Z)\right], \tag{13}$$

where $Z \sim \mathcal{N}(\mu, \sigma^2)$ for all nonnegative continuous functions $f : \mathbb{R} \mapsto \mathbb{R}_+$. On the other hand, by a change of measure (and absolute continuity of $\mathbb{P}_0$ w.r.t $\mathbb{P}_\beta$), we have that for such an f,

$$\mathbb{E}_{\mathbb{P}_0}\left[f(\log L)\right] = \mathbb{E}_{\mathbb{P}_\beta}\left[\frac{\mathrm{d}\,\mathbb{P}_0}{\mathrm{d}\,\mathbb{P}_\beta} f(\log L)\right] = \mathbb{E}_{\mathbb{P}_\beta}\left[e^{-\log L} f(\log L)\right].$$

The function $g : x \mapsto e^{-x} f(x)$ is still nonnegative continuous, so by (13), we have

$$\liminf \mathbb{E}_{\mathbb{P}_0}\left[f(\log L)\right] \geq \mathbb{E}\left[e^{-Z} f(Z)\right]. \tag{14}$$

Since $\mu = \frac{1}{2}\sigma^2$,

$$\mathbb{E}\left[e^{-Z} f(Z)\right] = \int f(x) e^{-x} e^{-(x-\mu)^2/2\sigma^2} \frac{\mathrm{d}x}{\sqrt{2\pi\sigma^2}} = \int f(x) e^{-(x+\mu)^2/2\sigma^2} \frac{\mathrm{d}x}{\sqrt{2\pi\sigma^2}} = \mathbb{E}\left[f(Z')\right],$$

where $Z' \sim \mathcal{N}(-\mu, \sigma^2)$. Since (14) is valid for every nonnegative continuous f, the result

$$\log L(\boldsymbol{Y};\beta) \rightsquigarrow \mathcal{N}(-\mu, \sigma^2)$$

under $\mathbb{P}_0$ follows.

B Notation and useful lemmas

We make repeated use of interpolation arguments in our proofs. In this section, we state a few elementary lemmas we subsequently invoke several times. We denote the overlaps between replicas when the last variables are deleted by a superscript " $-$ " :

$$R^{\mathtt{u}-}_{l,l'} = \frac{1}{N}\sum_{i=1}^{N-1} u_i^{(l)} u_i^{(l')} \quad \text{and} \quad R^{\mathtt{v}-}_{l,l'} = \frac{1}{N}\sum_{j=1}^{M-1} v_j^{(l)} v_j^{(l')}.$$

If $\{H_t : t \in [0,1]\}$ is a generic family of random Hamiltonians, we let $\langle \cdot \rangle_t$ be the corresponding Gibbs average, and $\nu_t(f) = \mathbb{E}\langle f \rangle_t$, where the expectation is over the randomness of H_t. We will often write ν for ν_1.

In our executions of the cavity method, we use interpolations that isolate one last variable (either u_N or v_M) from the rest of the system. Taking the first case an example, we consider

$$\begin{aligned} -H_t(\boldsymbol{u},\boldsymbol{v}) = &\sum_{i=1}^{N-1}\sum_{j=1}^{M} \sqrt{\frac{\beta}{N}} W_{ij} u_i v_j + \frac{\beta}{N} u_i u_i^* v_j v_j^* - \frac{\beta}{2N} u_i^2 v_j^2 \\ &+ \sum_{j=1}^{M} \sqrt{\frac{\beta t}{N}} W_{Nj} u_N v_j + \frac{\beta t}{N} u_N u_N^* v_j v_j^* - \frac{\beta t}{2N} u_N^2 v_j^2. \end{aligned}$$

Lemma 10. *Let f be a function of n replicas $(\boldsymbol{u}^{(l)}, \boldsymbol{v}^{(l)})_{1\le l\le n}$. Then*

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\nu_t(f) = &\frac{\beta}{2}\sum_{1\le l\neq l'\le n} \nu_t(R^{\mathtt{v}}_{l,l'} u^{(l)} u^{(l')} f) - \frac{\beta}{2} n \sum_{l=1}^{n} \nu_t(R^{\mathtt{v}}_{l,n+1} u^{(l)} u^{(n+1)} f) \\ &+ \beta n \sum_{l=1}^{n} \nu_t(R^{\mathtt{v}}_{l,*} u^{(l)} u^* f) - \beta n \nu_t(R^{\mathtt{v}}_{n+1,*} u^{(n+1)} u^* f) \\ &+ \beta \frac{n(n+1)}{2} \nu_t(R^{\mathtt{v}}_{n+1,n+2} u^{(n+1)} u^{(n+2)} f). \end{aligned}$$

Proof. This is a simple computation based on Gaussian integration by parts, similarly to Lemma 5. ∎

The next lemma allows us to control interpolated averages by averages at time 1.

Lemma 11. *Let f be a nonnegative function of n replicas $(\boldsymbol{u}^{(l)}, \boldsymbol{v}^{(l)})_{1\le l\le n}$. Then for all $t \in [0,1]$*

$$\nu_t(f) \le K(n,\alpha,\beta)\nu(f).$$

Proof. This is a consequence of Lemma 10, boundedness of the variables u_i and v_j, and Grönwall's lemma. ∎

It is clear that Lemma 11 also holds if we switch the roles of $\boldsymbol{u}$ and $\boldsymbol{v}$ and extract v_M instead (so that ν_t is defined accordingly).

C Proof of Proposition 7

We make use of two interpolation arguments; the first one extracts the last variable u_N from the system, and the second one extracts v_M. This allows to establish the self-consistency equations (7) and (8). We will assume decay of the forth moments of the overlaps, i.e., we assume Proposition 8 (which we prove in Appendix D), and this allows us the prove that the error terms emerging from the cavity method converge to zero. Recall that the Nishimori property implies

$$\mathbb{E}\left[\left\langle R^{\mathtt{u}}_{1,2} R^{\mathtt{v}}_{1,2}\right\rangle e^{\mathrm{i}s\log L}\right] = \mathbb{E}\left[\left\langle R^{\mathtt{u}}_{1,*} R^{\mathtt{v}}_{1,*}\right\rangle e^{\mathrm{i}s\log L}\right].$$

As it turns out, it is more convenient to work with the right-hand side.

C.1 Cavity on N

By symmetry of the $\mathtt{u}$ variables, we have

$$\mathbb{E}\left[\left\langle R^{\mathtt{u}}_{1,*} R^{\mathtt{v}}_{1,*}\right\rangle e^{\mathrm{i}s\log L}\right] = \mathbb{E}\left[\left\langle u_N^{(1)} u_N^* R^{\mathtt{v}}_{1,*}\right\rangle e^{\mathrm{i}s\log L}\right].$$

Now we consider the interpolating Hamiltonian

$$\begin{aligned} -H_t(\boldsymbol{u},\boldsymbol{v}) = &\sum_{i=1}^{N-1}\sum_{j=1}^{M} \sqrt{\frac{\beta}{N}} W_{ij} u_i v_j + \frac{\beta}{N} u_i u_i^* v_j v_j^* - \frac{\beta}{2N} u_i^2 v_j^2 \\ &+ \sum_{j=1}^{M} \sqrt{\frac{\beta t}{N}} W_{Nj} u_N v_j + \frac{\beta t}{N} u_N u_N^* v_j v_j^* - \frac{\beta t}{2N} u_N^2 v_j^2, \end{aligned}$$

and let $\langle\cdot\rangle_t$ be the associated Gibbs average. We let

$$X(t) = \exp\Big(\mathrm{i}s\log\int e^{-H_t(\boldsymbol{u},\boldsymbol{v})}\mathrm{d}\rho(\boldsymbol{u},\boldsymbol{v})\Big),$$

and

$$\varphi(t) = N\,\mathbb{E}\left[\left\langle u_N^{(1)} u_N^* R^{\mathtt{v}}_{1,*}\right\rangle_t X(t)\right].$$

Observe that $\varphi(1)$ is the quantity we seek to analyze. We will use the following error bound on Taylor's expansion:

$$\big|\varphi(1) - \varphi(0) - \varphi'(0)\big| \le \sup_{0\le t\le 1} |\varphi''(t)|,$$

to approximate $\varphi(1)$ by $\varphi(0) + \varphi'(0)$. Since $P_{\mathtt{u}}$ is centered, we have $\varphi(0) = 0$. With a computation similar to the one leading to Lemma 6, the time derivative $\varphi'(t)$ is a sum of terms of the form

$$N\beta\,\mathbb{E}\left[\left\langle u_N^{(1)} u_N^* u_N^{(a)} u_N^{(b)} R^{\mathtt{v}}_{1,*} R^{\mathtt{v}}_{a,b}\right\rangle_t X(t)\right],$$

for $(a,b)\in\{(1,*),(2,*),(1,2),(2,3)\}$. At $t=0$ all terms vanish expect when $(a,b) = (1,*)$ and we get

$$\varphi'(0) = N\beta\,\mathbb{E}\left[\left\langle (R^{\mathtt{v}}_{1,*})^2\right\rangle_0 X(0)\right].$$

Now we wish to replace the time index $t=0$ in the above quantity by the time index $t=1$. Similarly to φ, the derivative of the function $t\mapsto N\beta\,\mathbb{E}[\langle (R^{\mathtt{v}}_{1,*})^2\rangle_t X(t)]$, is a sum of terms of the form

$$N\beta^2\,\mathbb{E}\left[\left\langle u_N^{(a)} u_N^{(b)} (R^{\mathtt{v}}_{1,*})^2 R^{\mathtt{v}}_{a,b}\right\rangle_t X(t)\right].$$

By boundedness of the $\mathtt{u}$ variables and Hölder's inequality, this is bounded by

$$\begin{aligned} N\beta^2 K_{\mathtt{u}}^4 \,\mathbb{E}\left[\left\langle \left|(R^{\mathtt{v}}_{1,*})^2 R^{\mathtt{v}}_{a,b}\right|\right\rangle_t\right] &\le N\beta^2 K_{\mathtt{u}}^4 \,\mathbb{E}\left[\left\langle |R^{\mathtt{v}}_{1,*}|^3\right\rangle_t\right] \\ &\le N\beta^2 K_{\mathtt{u}}^4 K \,\mathbb{E}\left[\left\langle |R^{\mathtt{v}}_{1,*}|^3\right\rangle\right] \\ &\le \frac{K\beta^2}{\sqrt{N}}, \end{aligned}$$

where the second bound is by Lemma 11, and the last bound is a consequence of Proposition 8 (and Jensen's inequality). Therefore

$$\left|\varphi'(0) - N\beta\,\mathbb{E}\left[\left\langle (R^{\mathtt{v}}_{1,*})^2\right\rangle X(1)\right]\right| \le \frac{K}{\sqrt{N}}.$$

Similarly, we control the second derivative φ''. This can be written as a finite sum of terms of the form

$$N\beta^2\,\mathbb{E}\left[\left\langle u_N^{(1)} u_N^* u_N^{(a)} u_N^{(b)} u_N^{(c)} u_N^{(d)} R^{\mathtt{v}}_{1,*} R^{\mathtt{v}}_{a,b} R^{\mathtt{v}}_{c,d}\right\rangle_t X(t)\right],$$

which are bounded in the same way by

$$N\beta^2 K_{\mathtt{u}}^6 \,\mathbb{E}\left[\left\langle \left|R^{\mathtt{v}}_{1,*} R^{\mathtt{v}}_{a,b} R^{\mathtt{v}}_{c,d}\right|\right\rangle_t\right] \le \beta^2 K_{\mathtt{u}}^6 \frac{K}{\sqrt{N}}.$$

Therefore $|\varphi''| \le K/\sqrt{N}$. We end up with

$$N\,\mathbb{E}\left[\left\langle R^{\mathtt{u}}_{1,*} R^{\mathtt{v}}_{1,*}\right\rangle e^{\mathrm{i}s\log L}\right] = N\beta\,\mathbb{E}\left[\left\langle (R^{\mathtt{v}}_{1,*})^2\right\rangle e^{\mathrm{i}s\log L}\right] + \delta, \tag{15}$$

where $|\delta| \le K/\sqrt{N}$ whenever (α,β) satisfy the conditions of Proposition 8.

C.2 Cavity on M

By symmetry of the $\mathtt{v}$ variables,

$$\begin{aligned} N\,\mathbb{E}\left[\left\langle (R^{\mathtt{v}}_{1,*})^2\right\rangle e^{\mathrm{i}s\log L}\right] &= M\,\mathbb{E}\left[\left\langle v_M^{(1)} v_M^* R^{\mathtt{v}}_{1,*}\right\rangle e^{\mathrm{i}s\log L}\right] \\ &= \frac{M}{N}\,\mathbb{E}\left[\left\langle (v_M^{(1)} v_M^*)^2\right\rangle e^{\mathrm{i}s\log L}\right] + M\,\mathbb{E}\left[\left\langle v_M^{(1)} v_M^* R^{\mathtt{v}-}_{1,*}\right\rangle e^{\mathrm{i}s\log L}\right]. \end{aligned}$$

Now we execute the same argument as above with the roles of $\mathtt{u}$ and $\mathtt{v}$ flipped to prove that

$$\mathbb{E}\left[\left\langle (v_M^{(1)} v_M^*)^2\right\rangle e^{\mathrm{i}s\log L}\right] = \mathbb{E}\left[e^{\mathrm{i}s\log L}\right] + \delta,$$

and

$$M\,\mathbb{E}\left[\left\langle v_M^{(1)} v_M^* R^{\mathtt{v}-}_{1,*}\right\rangle e^{\mathrm{i}s\log L}\right] = M\beta\,\mathbb{E}\left[\left\langle R^{\mathtt{u}}_{1,*} R^{\mathtt{v}}_{1,*}\right\rangle e^{\mathrm{i}s\log L}\right] + \delta,$$

where $|\delta| \le K(M/N^{3/2} \vee 1/\sqrt{N})$. Here we use the interpolating Hamiltonian

$$\begin{aligned} -H_t(\boldsymbol{u},\boldsymbol{v}) = &\sum_{j=1}^{M-1}\sum_{i=1}^{N} \sqrt{\frac{\beta}{N}} W_{ij} u_i v_j + \frac{\beta}{N} u_i u_i^* v_j v_j^* - \frac{\beta}{2N} u_i^2 v_j^2 \\ &+ \sum_{i=1}^{N} \sqrt{\frac{\beta t}{N}} W_{iM} u_i v_M + \frac{\beta t}{N} u_i u_i^* v_M v_M^* - \frac{\beta t}{2N} u_i^2 v_M^2, \end{aligned}$$

and similarly define the random variable $X(t) = \exp\left(\mathrm{i}s \log \int e^{-H_t(\boldsymbol{u},\boldsymbol{v})} \mathrm{d}\rho(\boldsymbol{u},\boldsymbol{v})\right)$. After executing the argument, we obtain

$$N\,\mathbb{E}\left[\left\langle (R^{\mathtt{v}}_{1,*})^2 \right\rangle e^{\mathrm{i}s\log L}\right] = \frac{M}{N}\,\mathbb{E}\left[e^{\mathrm{i}s\log L}\right] + M\beta\,\mathbb{E}\left[\left\langle R^{\mathtt{u}}_{1,*} R^{\mathtt{v}}_{1,*}\right\rangle e^{\mathrm{i}s\log L}\right] + \delta. \tag{16}$$

From (15) and (16), we obtain

$$N\,\mathbb{E}\left[\left\langle R^{\mathtt{u}}_{1,*} R^{\mathtt{v}}_{1,*}\right\rangle e^{\mathrm{i}s\log L}\right] = \frac{M}{N}\beta\,\mathbb{E}\left[e^{\mathrm{i}s\log L}\right] + M\beta^2\,\mathbb{E}\left[\left\langle R^{\mathtt{u}}_{1,*} R^{\mathtt{v}}_{1,*}\right\rangle e^{\mathrm{i}s\log L}\right] + \delta,$$

where $|\delta| \le K(M/N^{3/2} \vee 1/\sqrt{N})$. For $M = \alpha N + \mathcal{O}(\sqrt{N})$, we arrive at

$$N\,\mathbb{E}\left[\left\langle R^{\mathtt{u}}_{1,*} R^{\mathtt{v}}_{1,*}\right\rangle e^{\mathrm{i}s\log L}\right] = \frac{\alpha\beta}{1-\alpha\beta^2}\,\mathbb{E}\left[e^{\mathrm{i}s\log L}\right] + \delta,$$

with $|\delta| \le K/\sqrt{N}$, and this finishes the proof.

D Proof of Proposition 8

This section is about overlap convergence in the planted model. As explained in the main text, the proof is in several steps. We first present a proof of convergence of the second moment of the overlaps that does not rely on the cavity method, but on a *quadratic replica coupling* scheme of Guerra and Toninelli (2002). Then we present the interpolation argument as a fixed overlap that will allow us to prove the crude convergence bound (12). Finally we execute a round of the cavity method to prove convergence of the fourth moment.

D.1 Convergence of the second moment

Proposition 12. *For all α, β such that $K_{\mathtt{u}}^4 K_{\mathtt{v}}^4 \alpha\beta^2 < 1$, there exists $K = K(\alpha,\beta) < \infty$ such that*

$$\mathbb{E}\left\langle (R^{\mathtt{u}}_{1,*})^2\right\rangle \vee \mathbb{E}\left\langle (R^{\mathtt{v}}_{1,*})^2\right\rangle \le \frac{K}{N^2}.$$

Of course, by the Nishimori property, this is also a statement about the overlaps between two independent replicas.

Proof. Let $\sigma_{\mathtt{u}}$ and $\sigma_{\mathtt{v}}$ be the sub-Gaussian parameters of $P_{\mathtt{u}}$ and $P_{\mathtt{v}}$ respectively. We since $P_{\mathtt{u}}$ and $P_{\mathtt{v}}$ have unit variance, we have $1 \le \sigma_{\mathtt{u}}^2 \le K_{\mathtt{u}}^2$ and similarly for $P_{\mathtt{v}}$.

We start with the $\mathtt{u}$-overlap. Let us define the function

$$\Phi_{\mathtt{u}}(\lambda) = \frac{1}{N}\,\mathbb{E}\log\int \exp\left(-H(\boldsymbol{u},\boldsymbol{v}) + \frac{\lambda}{2}N(R^{\mathtt{u}}_{1,*})^2\right)\mathrm{d}\rho(\boldsymbol{u},\boldsymbol{v}).$$

The outer expectation is on $\boldsymbol{Y} \sim \mathbb{P}_\beta$ (or equivalently on $\boldsymbol{u}^*$, $\boldsymbol{v}^*$ and $\boldsymbol{W}$ independently). A simple inspection shows that the above function is convex and increasing in λ, and

$$\Phi'_{\mathtt{u}}(0) = \frac{1}{2}\,\mathbb{E}\left\langle (R^{\mathtt{u}}_{1,*})^2\right\rangle.$$

The convexity then implies for all $\lambda \ge 0$,

$$\frac{\lambda}{2}\,\mathbb{E}\left\langle (R^{\mathtt{u}}_{1,*})^2\right\rangle \le \Phi_{\mathtt{u}}(\lambda) - \Phi_{\mathtt{u}}(0).$$

Of course $\Phi_{\mathtt{u}}(0) = \frac{1}{N}\,\mathbb{E}_{\mathbb{P}_\beta} \log L(\boldsymbol{Y};\beta) \geq 0$ by Jensen's inequality, so it remains to upper bound $\Phi_{\mathtt{u}}(\lambda)$. To this end we consider the interpolation

$$\Phi_{\mathtt{u}}(\lambda,t) = \frac{1}{N}\,\mathbb{E}\log\int \exp\left(-H_t(\boldsymbol{u},\boldsymbol{v}) + \frac{\lambda}{2}N(R^{\mathtt{u}}_{1,*})^2\right)\mathrm{d}\rho(\boldsymbol{u},\boldsymbol{v}),$$

where

$$-H_t(\boldsymbol{u},\boldsymbol{v}) = \sum_{i,j}\sqrt{\frac{\beta t}{N}}W_{ij}u_iv_j + \frac{\beta}{N}u_iu_i^*v_jv_j^* - \frac{\beta t}{2N}u_i^2v_j^2.$$

Notice that the planted (middle) term in the Hamiltonian is left unaltered. The time derivative is

$$\partial_t\Phi_{\mathtt{u}}(\lambda,t) = -\frac{\beta}{2}\,\mathbb{E}\left\langle (R^{\mathtt{u}}_{1,2})^2\right\rangle_{\lambda,t} \leq 0,$$

where $\langle\cdot\rangle_{\lambda,t}$ is the Gibbs average w.r.t $-H_t(\boldsymbol{u},\boldsymbol{v}) + \frac{\lambda}{2}N(R^{\mathtt{u}}_{1,*})^2$. Therefore

$$\begin{aligned}\Phi_{\mathtt{u}}(\lambda) \leq \Phi_{\mathtt{u}}(\lambda,0) &= \frac{1}{N}\,\mathbb{E}\log\int\exp\left(\beta N R^{\mathtt{u}}_{1,*}R^{\mathtt{v}}_{1,*} + \frac{\lambda}{2}N(R^{\mathtt{u}}_{1,*})^2\right)\mathrm{d}\rho(\boldsymbol{u},\boldsymbol{v})\\ &\leq \frac{1}{N}\,\mathbb{E}\log\int\exp\left(\frac{\alpha\beta^2\sigma_{\mathtt{v}}^2\widehat{v}+\lambda}{2}N(R^{\mathtt{u}}_{1,*})^2\right)\mathrm{d}P_{\mathtt{u}}^{\otimes N}(\boldsymbol{u}),\end{aligned}$$

where we have used the sub-Gaussianity of $P_{\mathtt{v}}$, and let $\widehat{v} = \frac{1}{M}\sum_{j=1}^M v_j^{*2}$. (Here, we have abused notation and let $\alpha = \frac{M}{N}$. This will not cause any problems.) Next we introduce an independent r.v. $g\sim\mathcal{N}(0,1)$, exchange integrals by Fubini's theorem, and continue:

$$\begin{aligned}&\frac{1}{N}\,\mathbb{E}\log\mathbb{E}_g\left[\int\exp\left(\sqrt{(\alpha\beta^2\sigma_{\mathtt{v}}^2\widehat{v}+\lambda)N}R^{\mathtt{u}}_{1,*}g\right)\mathrm{d}P_{\mathtt{u}}^{\otimes N}(\boldsymbol{u})\right]\\ &\leq \frac{1}{N}\,\mathbb{E}\log\mathbb{E}_g\left[\exp\left(\frac{\alpha\beta^2\sigma_{\mathtt{v}}^2\widehat{v}+\lambda}{2}\sigma_{\mathtt{u}}^2\widehat{u}g^2\right)\right],\end{aligned}$$

where we use the sub-Gaussianity of $P_{\mathtt{u}}$, and let $\widehat{u} = \frac{1}{N}\sum_{i=1}^N u_i^{*2}$. We bound $\widehat{u}$ and $\widehat{v}$ by $K_{\mathtt{u}}^2$ and $K_{\mathtt{v}}^2$ respectively and integrate on g to obtain the upper bound

$$\Phi_{\mathtt{u}}(\lambda) \leq -\frac{1}{2N}\log\left(1-(\alpha\beta^2\sigma_{\mathtt{v}}^2K_{\mathtt{v}}^2+\lambda)\sigma_{\mathtt{u}}^2K_{\mathtt{u}}^2\right),$$

valid as long as $(\alpha\beta^2\sigma_{\mathtt{v}}^2K_{\mathtt{v}}^2+\lambda)\sigma_{\mathtt{u}}^2K_{\mathtt{u}}^2 < 1$. Letting $\lambda = (1-\alpha\beta^2\sigma_{\mathtt{v}}^2K_{\mathtt{v}}^2\sigma_{\mathtt{u}}^2K_{\mathtt{u}}^2)/(2\sigma_{\mathtt{u}}^2K_{\mathtt{u}}^2) > 0$, we obtain

$$\mathbb{E}\left\langle (R^{\mathtt{u}}_{1,*})^2\right\rangle \leq \frac{K(\alpha,\beta)}{N},$$

with $K(\alpha,\beta) = \frac{2\sigma_{\mathtt{u}}^2K_{\mathtt{u}}^2\log((1-\alpha\beta^2\sigma_{\mathtt{v}}^2K_{\mathtt{v}}^2\sigma_{\mathtt{u}}^2K_{\mathtt{u}}^2)/2)}{(1-\alpha\beta^2\sigma_{\mathtt{v}}^2K_{\mathtt{v}}^2\sigma_{\mathtt{u}}^2K_{\mathtt{u}}^2)}$.

We use the exact same argument for the v-overlaps. We define $\Phi_{\mathtt{v}}(\lambda)$ in the same way by replacing the quadratic term $\frac{\lambda}{2}N(R^{\mathtt{u}}_{1,*})^2$ by $\frac{\lambda}{2}N(R^{\mathtt{v}}_{1,*})^2$ and obtain

$$\Phi_{\mathtt{v}}(\lambda) \leq -\frac{1}{2N}\log\left(1-(\beta^2\sigma_{\mathtt{u}}^2K_{\mathtt{u}}^2+\lambda)\alpha\sigma_{\mathtt{v}}^2K_{\mathtt{v}}^2\right).$$

We choose $\lambda = (1-\alpha\beta^2\sigma_{\mathtt{v}}^2K_{\mathtt{v}}^2\sigma_{\mathtt{u}}^2K_{\mathtt{u}}^2)/(2\alpha\sigma_{\mathtt{v}}^2K_{\mathtt{v}}^2)$ and use the same convexity argument to obtain

$$\mathbb{E}\left\langle (R^{\mathtt{v}}_{1,*})^2\right\rangle \leq \frac{K'(\alpha,\beta)}{N},$$

with $K'(\alpha,\beta) = \frac{2\alpha\sigma_{\mathtt{v}}^2K_{\mathtt{v}}^2\log((1-\alpha\beta^2\sigma_{\mathtt{v}}^2K_{\mathtt{v}}^2\sigma_{\mathtt{u}}^2K_{\mathtt{u}}^2)/2)}{(1-\alpha\beta^2\sigma_{\mathtt{v}}^2K_{\mathtt{v}}^2\sigma_{\mathtt{u}}^2K_{\mathtt{u}}^2)}$. ■

D.2 Interpolation bound at fixed overlap

In this section we present and prove an interpolation bound on the free energy of a subpopulation of configurations having a fixed overlap with the planted spike $(\boldsymbol{u}^*, \boldsymbol{v}^*)$. This is a key step in proving the crude bound (12).

Proposition 13. *Fix $\boldsymbol{u}^* \in \mathbb{R}^N, \boldsymbol{v}^* \in \mathbb{R}^M$ with $\|\boldsymbol{u}^*\|_{\ell_2}^2 /N \leq K_{\mathtt{u}}^2$ and $\|\boldsymbol{v}^*\|_{\ell_2}^2 /M \leq K_{\mathtt{v}}^2$. Let $\alpha = \frac{M}{N}$ and $\Delta = \alpha\beta^2\sigma_{\mathtt{u}}^2\sigma_{\mathtt{v}}^2 K_{\mathtt{u}}^2 K_{\mathtt{v}}^2 - 1$. For $m \in \mathbb{R} \setminus \{0\}$, $\epsilon \geq 0$, let $A_{\mathtt{u}}$ be the event*

$$A_{\mathtt{u}} = \begin{cases} R_{1,*}^{\mathtt{u}} \in [m, m+\epsilon) & \text{if } m > 0, \\ R_{1,*}^{\mathtt{u}} \in (m-\epsilon, m] & \text{if } m < 0. \end{cases}$$

Define $A_{\mathtt{v}}$ similarly. We have

$$\frac{1}{N}\mathbb{E}\log\int \mathbb{1}(A_{\mathtt{u}}) e^{-H(\boldsymbol{u},\boldsymbol{v})} \mathrm{d}\rho(\boldsymbol{u},\boldsymbol{v}) \leq \frac{\Delta}{2\sigma_{\mathtt{u}}^2 K_{\mathtt{u}}^2} m^2 + \alpha\beta K_{\mathtt{v}}^2 \epsilon, \tag{17}$$

and

$$\frac{1}{N}\mathbb{E}\log\int \mathbb{1}(A_{\mathtt{v}}) e^{-H(\boldsymbol{u},\boldsymbol{v})} \mathrm{d}\rho(\boldsymbol{u},\boldsymbol{v}) \leq \frac{\Delta}{2\alpha\sigma_{\mathtt{v}}^2 K_{\mathtt{v}}^2} m^2 + \beta K_{\mathtt{u}}^2 \epsilon. \tag{18}$$

The expectation $\mathbb{E}$ is over the Gaussian disorder $\boldsymbol{W}$.

Proof. We only prove (17). The bound (18) follows by flipping the roles of $\boldsymbol{u}$ and $\boldsymbol{v}$. We consider the interpolating Hamiltonian

$$-H_t(\boldsymbol{u},\boldsymbol{v}) = \sum_{i,j} \sqrt{\frac{\beta t}{N}} W_{ij} u_i v_j + \frac{\beta t}{N} u_i u_i^* v_j v_j^* - \frac{\beta t}{2N} u_i^2 v_j^2 + \sum_{j=1}^{M} (1-t)\beta m v_j v_j^*,$$

and let

$$\varphi(t) = \frac{1}{N}\mathbb{E}\log\int \mathbb{1}\{R_{1,*}^{\mathtt{u}} \in [m, m+\epsilon)\} e^{-H_t(\boldsymbol{u},\boldsymbol{v})} \mathrm{d}\rho(\boldsymbol{u},\boldsymbol{v}).$$

We have

$$\varphi'(t) = -\frac{\beta}{2}\mathbb{E}\left\langle R_{1,2}^{\mathtt{u}} R_{1,2}^{\mathtt{v}} \right\rangle_t + \beta\,\mathbb{E}\left\langle R_{1,*}^{\mathtt{u}} R_{1,*}^{\mathtt{v}} \right\rangle_t - \beta m\,\mathbb{E}\left\langle R_{1,*}^{\mathtt{v}} \right\rangle_t.$$

The first term in the above expression is ≤ 0, and since the overlap $R_{1,*}^{\mathtt{u}}$ is constrained to be close to m we have $\left|\mathbb{E}\left\langle (R_{1,*}^{\mathtt{u}} - m) R_{1,*}^{\mathtt{v}} \right\rangle_t\right| \leq \alpha K_{\mathtt{v}}^2 \epsilon$. So $\varphi'(t) \leq \alpha K_{\mathtt{v}}^2 \epsilon$. Moreover, the variables $\boldsymbol{u}$ and $\boldsymbol{v}$ decouple at $t = 0$ and one can write

$$\varphi(1) \leq \frac{1}{N}\log\Pr(A_{\mathtt{u}}) + \frac{1}{N}\sum_{j=1}^{M}\log\mathbb{E}_v\left[e^{\beta m v v_j^*}\right] + K_{\mathtt{v}}^2\epsilon.$$

By sub-Gaussianity of the prior $P_{\mathtt{v}}$ we have $\mathbb{E}_v\left[e^{\beta m v v_j^*}\right] \leq e^{\beta^2\sigma_{\mathtt{v}}^2 m^2 v_j^{*2}/2}$. On the other hand, for a fixed parameter γ of the same sign as m, we have

$$\frac{1}{N}\log\Pr(A_{\mathtt{u}}) \leq -\gamma m + \frac{1}{N}\sum_{i=1}^{N}\log\mathbb{E}_u[e^{\gamma u u_i^*}] \leq -\gamma m + \frac{1}{2N}\sum_{i=1}^{N} u_i^{*2}\sigma_{\mathtt{u}}^2\gamma^2.$$

The last inequality uses sub-Gaussianity of $P_{\mathtt{u}}$. We minimize this quadratic w.r.t γ and obtain

$$\varphi(1) \leq -\frac{m^2}{2\sigma_{\mathtt{u}}^2\widehat{u}} + \frac{M}{2N}\beta^2\sigma_{\mathtt{v}}^2\widehat{v}m^2 + \alpha K_{\mathtt{v}}^2\epsilon,$$

where $\widehat{u} = \frac{1}{N}\sum_{i=1}^{N} u_i^{*2}$ and $\widehat{v} = \frac{1}{M}\sum_{j=1}^{M} v_j^{*2}$. We upper bound the latter two numbers by $K_{\mathtt{u}}^2$ and K_v^2 respectively. ∎

D.3 Overlap concentration (proof of (12))

Here we prove convergence of the overlaps to zero in probability. We first state a useful and standard result of concentration of measure.

Lemma 14. *Let* $\boldsymbol{Y} = \sqrt{\frac{\beta}{N}}\boldsymbol{u}^*\boldsymbol{v}^{*\top} + \boldsymbol{W}$*, where the planted vectors* $\boldsymbol{u}^*$ *and* $\boldsymbol{v}^*$ *are fixed, and* $W_{ij} \sim \mathcal{N}(0,1)$*. For a Borel set* $A \subset \mathbb{R}^{M+N}$*, let*

$$Z = \int_A e^{-H(\boldsymbol{u},\boldsymbol{v})}\mathrm{d}\rho(\boldsymbol{u},\boldsymbol{v}).$$

We have for every $t \geq 0$*,*

$$\Pr\left(|\log Z - \mathbb{E}\log Z| \geq Nt\right) \leq 2e^{-\frac{Nt^2}{2\beta K_{\mathtt{u}}^2 K_{\mathtt{v}}^2}}.$$

(Here $\Pr$ *and* $\mathbb{E}$ *are conditional on* $\boldsymbol{u}^*$ *and* $\boldsymbol{v}^*$*.)*

Proof. We simply observe that the function $\boldsymbol{W} \mapsto \log Z$ is Lipschitz with constant $\sqrt{N\beta\alpha K_{\mathtt{u}}^2 K_{\mathtt{v}}^2}$. The result follows from concentration of Lipschitz functions of Gaussian r.v.'s (this is the Borell-Tsirelson-Ibragimov-Sudakov inequality; see Boucheron et al., 2013, Theorem 5.6). ■

Proposition 15. *Let* α, β *such that* $\alpha\beta^2\sigma_{\mathtt{u}}^2\sigma_{\mathtt{v}}^2K_{\mathtt{u}}^2K_{\mathtt{v}}^2 < 1$*, and* $\epsilon > 0$*. There exist constants* $c = c(\epsilon, \alpha, \beta, K_{\mathtt{u}}, K_{\mathtt{v}}) > 0$ *and* $K = K(K_{\mathtt{u}}, K_{\mathtt{v}}) > 0$ *such that*

$$\mathbb{E}\left\langle \mathbb{1}\{|R^{\mathtt{u}}_{1,*}| \geq \epsilon\}\right\rangle \vee \mathbb{E}\left\langle \mathbb{1}\{|R^{\mathtt{v}}_{1,*}| \geq \epsilon\}\right\rangle \leq \frac{K}{\epsilon^2}e^{-cN}.$$

Proof. We only prove the assertion for the $\mathtt{u}$-overlap since the argument is strictly the same for the $\mathtt{v}$-overlap.

For $\epsilon, \epsilon' > 0$, we can write the decomposition

$$\begin{aligned}\mathbb{E}\left\langle \mathbb{1}\{\left|R^{\mathtt{u}}_{1,*}\right| \geq \epsilon'\}\right\rangle &= \sum_{l\geq 0}\mathbb{E}\left\langle \mathbb{1}\{R^{\mathtt{u}}_{1,*} - \epsilon' \in [l\epsilon, (l+1)\epsilon)\}\right\rangle \\ &\quad + \sum_{l\geq 0}\mathbb{E}\left\langle \mathbb{1}\{-R^{\mathtt{u}}_{1,*} + \epsilon' \in [l\epsilon, (l+1)\epsilon)\}\right\rangle,\end{aligned}$$

where the integer index l ranges over a finite set of size $\leq K/\epsilon$. We only treat the generic term in the first sum; the second sum can be handled similarly. Fix $m > 0, \epsilon > 0$. We have

$$\mathbb{E}\left\langle \mathbb{1}\{R^{\mathtt{u}}_{1,*} \in [m, m+\epsilon)\}\right\rangle = \mathbb{E}\left[\frac{\int \mathbb{1}\{R^{\mathtt{u}}_{1,*} \in [m, m+\epsilon)\}e^{-H(\boldsymbol{u},\boldsymbol{v})}\mathrm{d}\rho(\boldsymbol{u},\boldsymbol{v})}{\int e^{-H(\boldsymbol{u},\boldsymbol{v})}\mathrm{d}\rho(\boldsymbol{u},\boldsymbol{v})}\right]. \tag{19}$$

Let

$$A = \frac{1}{N}\mathbb{E}_{\boldsymbol{W}}\log\int \mathbb{1}\{R^{\mathtt{u}}_{1,*} \in [m, m+\epsilon)\}e^{-H(\boldsymbol{u},\boldsymbol{v})}\mathrm{d}\rho(\boldsymbol{u},\boldsymbol{v}),$$

and

$$B = \frac{1}{N}\mathbb{E}_{\boldsymbol{W}}\log\int e^{-H(\boldsymbol{u},\boldsymbol{v})}\mathrm{d}\rho(\boldsymbol{u},\boldsymbol{v}).$$

By concentration over the Gaussian disorder, Lemma 14, for any $u \geq 0$, we simultaneously have

$$\frac{1}{N}\log\int \mathbb{1}\{R^{\mathtt{u}}_{1,*} \in [m, m+\epsilon)\}e^{-H(\boldsymbol{u},\boldsymbol{v})}\mathrm{d}\rho(\boldsymbol{u},\boldsymbol{v}) - A \leq u,$$

and

$$\frac{1}{N}\log\int e^{-H(\boldsymbol{u},\boldsymbol{v})}\mathrm{d}\rho(\boldsymbol{u},\boldsymbol{v}) - B \geq -u,$$

with probability at least $1-4e^{-Nu^2/(2\beta K_{\mathtt{u}}^2K_{\mathtt{v}}^2)}$. On the complement event we simply upper bound the fraction (19) by 1. Therefore, we have

$$\mathbb{E}\left\langle \mathbb{1}\{R_{1,*}^{\mathtt{u}} \in [m,m+\epsilon)\}\right\rangle \leq \mathbb{E}_{\boldsymbol{u}^*,\boldsymbol{v}^*}\left[e^{N(A-B+2u)}\right] + 4e^{-Nu^2/(2\beta K_{\mathtt{u}}^2K_{\mathtt{v}}^2)}.$$

By Proposition 13 we have $A \leq \frac{\Delta}{2\sigma_{\mathtt{u}}^2K_{\mathtt{u}}^2}m^2 + \alpha\beta K_{\mathtt{v}}^2\epsilon$ deterministically over $\boldsymbol{u}^*$ and $\boldsymbol{v}^*$. Now it remains to control $\mathbb{E}_{\boldsymbol{u}^*,\boldsymbol{v}^*}\left[e^{-NB}\right]$.

Lemma 16. *We have* $\mathbb{E}_{\boldsymbol{u}^*,\boldsymbol{v}^*}\left[e^{-NB}\right] \leq 2e^{-N\,\mathbb{E}_{\boldsymbol{u}^*,\boldsymbol{v}^*}[B]}$.

Moreover, observe that

$$\begin{aligned}\mathbb{E}_{\boldsymbol{u}^*,\boldsymbol{v}^*}[B] &= \frac{1}{N}\,\mathbb{E}\log\int e^{-H(\boldsymbol{u},\boldsymbol{v})}\mathrm{d}\rho(\boldsymbol{u},\boldsymbol{v})\\ &= \frac{1}{N}\,\mathbb{E}_{\mathbb{P}_\beta}\log L(\boldsymbol{Y};\beta)\\ &= \frac{1}{N}\,\mathbb{E}_{\mathbb{P}_0}\,L(\boldsymbol{Y};\beta)\log L(\boldsymbol{Y};\beta) \geq 0.\end{aligned}$$

Positivity is obtained by Jensen's inequality and convexity of $x \mapsto x\log x$. In view of the above, Lemma 16 means that the random variable B is "essentially" positive. Therefore,

$$\mathbb{E}\left\langle \mathbb{1}\{R_{1,*}^{\mathtt{u}} \in [m,m+\epsilon)\}\right\rangle \leq 2e^{N(\delta+2u)} + 4e^{-Nu^2/(2\beta K_{\mathtt{u}}^2K_{\mathtt{v}}^2)},$$

where $\delta = \frac{\Delta}{2\sigma_{\mathtt{u}}^2K_{\mathtt{u}}^2}m^2 + \alpha\beta K_{\mathtt{v}}^2\epsilon$. We let $u = -\delta/3 \geq 0$, and $m = \epsilon' + l\epsilon$. Since $\Delta < 0$, $\Delta m^2 \leq \Delta\epsilon'^2$. Now we let $\epsilon = -\frac{\Delta}{4\alpha\beta\sigma_{\mathtt{u}}^2K_{\mathtt{u}}^2K_{\mathtt{v}}^2}\epsilon'^2$ so that $\delta \leq \frac{3\Delta}{4\sigma_{\mathtt{u}}^2K_{\mathtt{u}}^2}\epsilon'^2 < 0$. ■

Proof of Lemma 16. We abbreviate $\mathbb{E}_{\boldsymbol{u}^*,\boldsymbol{v}^*}$ by $\mathbb{E}$. We have

$$\mathbb{E}\left[e^{N(\mathbb{E}[B]-B)}\right] = \int_{-\infty}^{+\infty} e^t\,\Pr\left(N(\mathbb{E}[B]-B)\geq t\right)\mathrm{d}t \leq 1 + \int_0^{+\infty} e^t\,\Pr\left(N(\mathbb{E}[B]-B)\geq t\right)\mathrm{d}t.$$

Now we bound the lower tail probability. The r.v. B, seen as a function of the vector $[\boldsymbol{u}^*|\boldsymbol{v}^*] \in \mathbb{R}^{N+M}$ is jointly convex (the Hessian can be easily shown to be positive semi-definite), and Lipschitz with constant $\beta K_{\mathtt{u}}K_{\mathtt{v}}\sqrt{\frac{\alpha K_{\mathtt{u}}^2+\alpha^2K_{\mathtt{v}}^2}{N}}$ with respect to the ℓ_2 norm. Under the above conditions, a bound on the lower tail of deviation of B is available; this is (one side of) Talagrand's inequality (see Boucheron et al., 2013, Theorem 7.12). Therefore, we have for all $t \geq 0$

$$\Pr\left(B-\mathbb{E}[B]\leq -t\right) \leq e^{-Nt^2/2K^2},$$

where $K^2 = \alpha\beta^2K_{\mathtt{u}}^2K_{\mathtt{v}}^2(K_{\mathtt{u}}^2+\alpha K_{\mathtt{v}}^2)$. Thus,

$$\begin{aligned}\mathbb{E}\left[e^{N(\mathbb{E}[B]-B)}\right] &\leq 1 + \int_0^{+\infty} e^t e^{-t^2/(2NK^2)}\mathrm{d}t\\ &= 1 + K\sqrt{N}e^{NK^2/2}\int_{K\sqrt{N}}^{+\infty} e^{-t^2/2}\mathrm{d}t\\ &\leq 2.\end{aligned}$$

The last inequality is a restatement of the fact $\Pr(g \geq t) \leq \frac{e^{-t^2/2}}{\sqrt{2\pi}t}$ where $g \sim \mathcal{N}(0,1)$. ■

D.4 Convergence of the fourth moment

In this section we prove that for all α, β such that $\alpha\beta^2\sigma_{\mathtt{u}}^2\sigma_{\mathtt{v}}^2 K_{\mathtt{u}}^2 K_{\mathtt{v}}^2 < 1$, we have

$$\mathbb{E}\left\langle (R^{\mathtt{u}}_{1,2})^4 \right\rangle \vee \mathbb{E}\left\langle (R^{\mathtt{v}}_{1,2})^4 \right\rangle \leq \frac{K(\alpha,\beta)}{N^2}.$$

We proceed as follows. Let

$$M = \max\left\{ \mathbb{E}\left\langle (R^{\mathtt{u}}_{1,2})^4 \right\rangle, \mathbb{E}\left\langle (R^{\mathtt{v}}_{1,2})^4 \right\rangle \right\}.$$

We prove that for $\epsilon > 0$, the following self-boundedness properties hold:

$$\mathbb{E}\left\langle (R^{\mathtt{u}}_{1,2})^4 \right\rangle \leq \alpha\beta^2\, \mathbb{E}\left\langle (R^{\mathtt{u}}_{1,2})^4 \right\rangle + K\epsilon M + \delta, \tag{20}$$
$$\mathbb{E}\left\langle (R^{\mathtt{v}}_{1,2})^4 \right\rangle \leq \alpha\beta^2\, \mathbb{E}\left\langle (R^{\mathtt{v}}_{1,2})^4 \right\rangle + K\epsilon M + \delta, \tag{21}$$

where $\delta \leq K/N^2 + K/\epsilon^2 e^{-c(\epsilon)N}$. This implies the desired result by letting ϵ be sufficiently small (e.g., $\epsilon = (1-\alpha\beta^2)/2$). We prove (20) and (21) using the cavity method, i.e. by isolating the effect of the last variables u_N and v_M, one at a time. We prove (20) in full detail, then briefly highlight how (21) is obtained in a similar way.

By symmetry between the $\mathtt{u}$ variables, we have

$$\begin{aligned}\mathbb{E}\left\langle (R^{\mathtt{u}}_{1,*})^4 \right\rangle &= \mathbb{E}\left\langle u_N^{(1)} u_N^* (R^{\mathtt{u}}_{1,*})^3 \right\rangle \\ &= \mathbb{E}\left\langle u_N^{(1)} u_N^* \Big(R^{\mathtt{u}-}_{1,*} + \frac{1}{N} u_N^{(1)} u_N^* \Big)^3 \right\rangle.\end{aligned}$$

Expanding the term $\big(R^{\mathtt{u}-}_{1,*} + \frac{1}{N}u_N^{(1)}u_N^*\big)^3$ we obtain

$$\mathbb{E}\left\langle (R^{\mathtt{u}}_{1,*})^4 \right\rangle \leq \mathbb{E}\left\langle u_N^{(1)} u_N^* (R^{\mathtt{u}-}_{1,*})^3 \right\rangle + \frac{K_{\mathtt{u}}^4}{N}\, \mathbb{E}\left\langle (R^{\mathtt{u}-}_{1,*})^2 \right\rangle + \frac{K_{\mathtt{u}}^6}{N^2}\, \mathbb{E}\left\langle \left| R^{\mathtt{u}-}_{1,*} \right| \right\rangle + \frac{K_{\mathtt{u}}^8}{N^3}. \tag{22}$$

We have already proved convergence of the second moment (Proposition 12), hence $\mathbb{E}\langle (R^{\mathtt{u}-}_{1,*})^2 \rangle \leq K/N$ and $\mathbb{E}\langle |R^{\mathtt{u}-}_{1,*}| \rangle \leq K/\sqrt{N}$. Now we need to control the leading term involving $(R^{\mathtt{u}-}_{1,*})^3$. The next proposition shows that this quantity can be related back to $(R^{\mathtt{u}}_{1,*})^4$, plus additional higher-order terms. This is is achieved through the cavity method.

Proposition 17. *For $\alpha, \beta \geq 0$, there exists a constant $K = K(\alpha, \beta, K_{\mathtt{u}}, K_{\mathtt{v}}) > 0$ such that*

$$\mathbb{E}\left\langle u_N^{(1)} u_N^* (R^{\mathtt{u}-}_{1,*})^3 \right\rangle = \beta\, \mathbb{E}\left\langle (R^{\mathtt{u}}_{1,*})^3 R^{\mathtt{v}}_{1,*} \right\rangle + \delta_1, \tag{23}$$

where

$$|\delta_1| \leq K \sum_{a,b,c,d} \mathbb{E}\left\langle \left| (R^{\mathtt{u}-}_{1,*})^3 R^{\mathtt{v}}_{a,b} R^{\mathtt{v}}_{c,d} \right| \right\rangle.$$

Moreover,

$$\mathbb{E}\left\langle (R^{\mathtt{u}}_{1,*})^3 R^{\mathtt{v}}_{1,*} \right\rangle = \alpha\beta\, \mathbb{E}\left\langle (R^{\mathtt{u}}_{1,*})^4 \right\rangle + \delta_2, \tag{24}$$

where

$$|\delta_2| \leq K \sum_{a,b,c,d} \mathbb{E}\left\langle \left| (R^{\mathtt{u}}_{1,*})^3 R^{\mathtt{u}}_{a,b} R^{\mathtt{u}}_{c,d} \right| \right\rangle.$$

From Proposition 17 we deduce

$$\mathbb{E}\left\langle u_N^{(1)} u_N^* (R_{1,*}^{\mathtt{u}-})^3 \right\rangle = \alpha\beta^2\, \mathbb{E}\left\langle (R_{1,*}^{\mathtt{u}})^4 \right\rangle + \delta,$$

where $\delta = \delta_1 + \delta_2$. Plugging into (22), we obtain

$$\mathbb{E}\left\langle (R_{1,*}^{\mathtt{u}})^4 \right\rangle \leq \alpha\beta^2\, \mathbb{E}\left\langle (R_{1,*}^{\mathtt{u}})^4 \right\rangle + \frac{K}{N^2} + \delta.$$

Now we need to control the error term δ, which involves monomials of degree 5 in the overlaps $R^{\mathtt{u}}$ and $R^{\mathtt{v}}$. This is where the a priori bound on the convergence of the overlaps, Proposition 15, is useful. Since the overlaps are bounded, we can write for any $\epsilon > 0$,

$$\begin{aligned}\mathbb{E}\left\langle \left|(R_{1,*}^{\mathtt{u}})^3 R_{a,b}^{\mathtt{v}} R_{c,d}^{\mathtt{v}}\right| \right\rangle &\leq \epsilon\, \mathbb{E}\left\langle \left|(R_{1,*}^{\mathtt{u}})^3 R_{a,b}^{\mathtt{v}}\right| \right\rangle + K_{\mathtt{u}}^6 K_{\mathtt{v}}^4\, \mathbb{E}\left\langle \mathbb{1}\{\left|R_{c,d}^{\mathtt{v}}\right| \geq \epsilon\} \right\rangle \\ &= \epsilon\, \mathbb{E}\left\langle \left|(R_{1,*}^{\mathtt{u}})^3 R_{a,b}^{\mathtt{v}}\right| \right\rangle + K_{\mathtt{u}}^6 K_{\mathtt{v}}^4\, \mathbb{E}\left\langle \mathbb{1}\{\left|R_{1,*}^{\mathtt{v}}\right| \geq \epsilon\} \right\rangle,\end{aligned}$$

where the last line is a consequence of the Nishimori property. Now we use Hölder's inequality on the first term:

$$\begin{aligned}\mathbb{E}\left\langle \left|(R_{1,*}^{\mathtt{u}})^3 R_{a,b}^{\mathtt{v}}\right| \right\rangle &\leq \left(\mathbb{E}\left\langle \left|(R_{1,*}^{\mathtt{u}})^4\right| \right\rangle\right)^{3/4} \left(\mathbb{E}\left\langle \left|(R_{a,b}^{\mathtt{v}}\right|)^4 \right\rangle\right)^{1/4} \\ &= \left(\mathbb{E}\left\langle \left|(R_{1,*}^{\mathtt{u}})^4\right| \right\rangle\right)^{3/4} \left(\mathbb{E}\left\langle \left|(R_{1,*}^{\mathtt{v}}\right|)^4 \right\rangle\right)^{1/4} \\ &\leq M.\end{aligned}$$

Using Proposition 15, we have $\mathbb{E}\langle \mathbb{1}\{|R_{1,*}^{\mathtt{v}}| \geq \epsilon\}\rangle \leq K e^{-cN}/\epsilon^2$. Therefore,

$$|\delta_1| \leq K\epsilon M + \frac{K}{\epsilon^2} e^{-cN}.$$

It is clear that we can use the same argument to bound δ_2, so we end up with

$$\mathbb{E}\left\langle (R_{1,*}^{\mathtt{u}})^4 \right\rangle \leq \alpha\beta^2\, \mathbb{E}\left\langle (R_{1,*}^{\mathtt{u}})^4 \right\rangle + \frac{K}{N^2} + K\epsilon M + \frac{K}{\epsilon^2} e^{-cN},$$

thereby proving (20). To prove (21) we use the same approach. We write

$$\begin{aligned}\mathbb{E}\left\langle (R_{1,*}^{\mathtt{v}})^4 \right\rangle &= \frac{M}{N}\, \mathbb{E}\left\langle v_M^{(1)} v_M^* (R_{1,*}^{\mathtt{v}})^3 \right\rangle \\ &= \alpha\, \mathbb{E}\left\langle v_M^{(1)} v_M^* \Big(R_{1,*}^{\mathtt{v}-} + \frac{1}{N} v_M^{(1)} v_M^*\Big)^3 \right\rangle.\end{aligned}$$

Then use an equivalent of Proposition 17 in this case, which is obtained by flipping the role of the $\mathtt{u}$ and $\mathtt{v}$ variables:

$$\mathbb{E}\left\langle v_N^{(1)} v_N^* (R_{1,*}^{\mathtt{v}-})^3 \right\rangle = \beta\, \mathbb{E}\left\langle (R_{1,*}^{\mathtt{v}})^3 R_{1,*}^{\mathtt{u}} \right\rangle + \delta_1,$$

and

$$\mathbb{E}\left\langle (R_{1,*}^{\mathtt{v}})^3 R_{1,*}^{\mathtt{u}} \right\rangle = \beta\, \mathbb{E}\left\langle (R_{1,*}^{\mathtt{v}})^4 \right\rangle + \delta_2,$$

where δ_1 and δ_2 are similarly bounded by expectations of monomials of degree 5 in the overlaps $R^{\mathtt{u}}$ and $R^{\mathtt{v}}$. These two quantities are then bounded in exactly the same way.

Proof of Proposition 17. The proof uses two interpolations; the first one decouples the variable u_N from the rest of the system and allows to obtain (23), and the second one decouples the variable v_M and allows to obtain (24). We start with the former.

Proof of (23)**.** Consider the interpolating Hamiltonian

$$-H_t(\boldsymbol{u},\boldsymbol{v}) = \sum_{i=1}^{N-1}\sum_{j=1}^{M}\sqrt{\frac{\beta}{N}}W_{ij}u_iv_j + \frac{\beta}{N}u_iu_i^*v_jv_j^* - \frac{\beta}{2N}u_i^2v_j^2$$
$$+\sum_{j=1}^{M}\sqrt{\frac{\beta t}{N}}W_{Nj}u_Nv_j + \frac{\beta t}{N}u_Nu_N^*v_jv_j^* - \frac{\beta t}{2N}u_N^2v_j^2,$$

and let $\langle\cdot\rangle_t$ be the associated Gibbs average and $\nu_t(\cdot) = \mathbb{E}\langle\cdot\rangle_t$. The idea is to approximate $\nu_1(f)$ where $f \equiv u_N^{(1)}u_N^*(R_{1,*}^{\mathtt{u}-})^3$ by $\nu_0(f)+\nu_0'(f)$. Of course one then has to control the second derivative, as dictated by Taylor's approximation

$$\left|\nu_1(f)-\nu_0(f)-\nu_0'(f)\right| \le \sup_{0\le t\le 1}\left|\nu_t''(f)\right|. \tag{25}$$

We see that at time $t=0$, the variables u_N and u_N^* decouple the Hamiltonian, so

$$\nu_0(u_N^{(1)}u_N^*(R_{1,*}^{\mathtt{u}-})^3) = \mathbb{E}[u_N]\,\mathbb{E}[u_N^*]\nu_0((R_{1,*}^{\mathtt{u}-})^3) = 0. \tag{26}$$

On the other hand, by applying Lemma 10 with $n=1$, we see that $\nu_0'(u_N^{(1)}u_N^*(R_{1,*}^{\mathtt{u}-})^3)$ is a sum of a few terms of the form

$$\nu_0\big(u_N^{(1)}u_N^*u_N^{(a)}u_N^{(b)}(R_{1,*}^{\mathtt{u}-})^3R_{a,b}^{\mathtt{v}}\big).$$

Since $P_{\mathtt{u}}$ has zero mean, all terms in which a variable $u_N^{(a)}$ (for any a) appears with degree 1 vanish. We are thus left with one term where $a=1, b=*$, and we get

$$\nu_0'(u_N^{(1)}u_N^*(R_{1,*}^{\mathtt{u}-})^3) = \beta\,\mathbb{E}[(u_N^{(1)})^2]\,\mathbb{E}[(u_N^*)^2]\nu_0((R_{1,*}^{\mathtt{u}-})^3R_{1,*}^{\mathtt{v}}) = \beta\nu_0((R_{1,*}^{\mathtt{u}-})^3R_{1,*}^{\mathtt{v}}). \tag{27}$$

Moreover, we see that $\nu_0((R_{1,*}^{\mathtt{u}-})^3R_{1,*}^{\mathtt{v}}) = \nu_0((R_{1,*}^{\mathtt{u}})^3R_{1,*}^{\mathtt{v}})$ since the last variable u_N has no contribution under ν_0. Now we are tempted to replace the average at time $t=0$ by an average at time $t=1$ in the last quantity. We use Lemmas 10 and 11 to justify this. Indeed these lemmas and boundedness of the variables u_N imply

$$\left|\nu_0((R_{1,*}^{\mathtt{u}})^3R_{1,*}^{\mathtt{v}}) - \nu_1((R_{1,*}^{\mathtt{u}})^3R_{1,*}^{\mathtt{v}})\right| \le K(\alpha,\beta)\sum_{a,b}\nu(\left|(R_{1,*}^{\mathtt{u}})^3R_{1,*}^{\mathtt{v}}R_{a,b}^{\mathtt{v}})\right|), \tag{28}$$

where $(a,b)\in\{(1,2),(1,*),(2,*),(2,3)\}$. Now we control the second derivative $\sup_t \nu_t''(\cdot)$. In view of Lemma 10, we see that taking two derivative of $\nu_t(u_N^{(1)}u_N^*(R_{1,*}^{\mathtt{u}-})^3)$ creates terms of the form

$$\nu_t\left(u_N^{(1)}u_N^*u_N^{(a)}u_N^{(b)}u_N^{(c)}u_N^{(d)}(R_{1,*}^{\mathtt{u}-})^3R_{a,b}^{\mathtt{v}}R_{c,d}^{\mathtt{v}}\right),$$

with a larger (but finite) set of combinations (a,b,c,d). We use Lemma 11 to replace ν_t by ν_1 and use boundedness of variables u_N to obtain the bound

$$\left|\sup_{0\le t\le 1}\nu_t''\Big(u_N^{(1)}u_N^*(R_{1,*}^{\mathtt{u}-})^3\Big)\right| \le K(\alpha,\beta)\sum_{a,b,c,d}\nu\Big(\Big|(R_{1,*}^{\mathtt{u}-})^3R_{a,b}^{\mathtt{v}}R_{c,d}^{\mathtt{v}}\Big|\Big). \tag{29}$$

Now putting the bounds and estimates (25), (26), (27), (28), and (29), we obtain the desired bound (23):

$$\left|\nu\big(u_N^{(1)} u_N^* (R_{1,*}^{\mathtt{u}-})^3\big) - \beta\nu((R_{1,*}^{\mathtt{u}})^3 R_{1,*}^{\mathtt{v}})\right| \le K(\alpha,\beta) \sum_{a,b,c,d} \nu\Big(\Big|(R_{1,*}^{\mathtt{u}-})^3 R_{a,b}^{\mathtt{v}} R_{c,d}^{\mathtt{v}}\Big|\Big).$$

Proof of (24). By symmetry of the $\mathtt{v}$ variables we have

$$\mathbb{E}\left\langle (R_{1,*}^{\mathtt{u}})^3 R_{1,*}^{\mathtt{v}} \right\rangle = \frac{M}{N}\,\mathbb{E}\left\langle (R_{1,*}^{\mathtt{u}})^3 v_M^{(1)} v_M^* \right\rangle.$$

Now we apply the same machinery. Consider the interpolating Hamiltonian

$$\begin{aligned} -H_t(\boldsymbol{u},\boldsymbol{v}) = &\sum_{j=1}^{M-1}\sum_{i=1}^{N} \sqrt{\frac{\beta}{N}} W_{ij} u_i v_j + \frac{\beta}{N} u_i u_i^* v_j v_j^* - \frac{\beta}{2N} u_i^2 v_j^2 \\ &+ \sum_{i=1}^{N} \sqrt{\frac{\beta t}{N}} W_{iM} u_i v_M + \frac{\beta t}{N} u_i u_i^* v_M v_M^* - \frac{\beta t}{2N} u_i^2 v_M^2, \end{aligned}$$

and let $\langle\cdot\rangle_t$ be the associated Gibbs average and $\nu_t(\cdot) = \mathbb{E}\langle\cdot\rangle_t$. The exact same argument goes through with the roles of $\mathtt{u}$ and $\mathtt{v}$ flipped. For instance, when one takes time derivatives, terms of the form $v_M^{(a)} v_M^{(b)} R_{a,b}^{\mathtt{u}}$ arise from the Hamiltonian, and one sees that

$$\nu_0'\big((R_{1,*}^{\mathtt{u}})^3 v_M^{(1)} v_M^*\big) = \beta\nu_0\big((R_{1,*}^{\mathtt{u}})^4\big).$$

Thus we similarly obtain

$$\left|\nu\big((R_{1,*}^{\mathtt{u}})^3 v_M^{(1)} v_M^*\big) - \beta\nu\big((R_{1,*}^{\mathtt{u}})^4\big)\right| \le K(\alpha,\beta) \sum_{a,b,c,d} \nu\Big(\Big|(R_{1,*}^{\mathtt{u}})^3 R_{a,b}^{\mathtt{u}} R_{c,d}^{\mathtt{u}}\Big|\Big).$$

■