An Instability in Variational Inference for Topic Models (Supplementary Material)

January 23, 2019

Abstract

Topic models are Bayesian models that are frequently used to capture the latent structure of certain corpora of documents or images. Each data element in such a corpus (for instance each item in a collection of scientific articles) is regarded as a convex combination of a small number of vectors corresponding to 'topics' or 'components'. The weights are assumed to have a Dirichlet prior distribution. The standard approach towards approximating the posterior is to use variational inference algorithms, and in particular a mean field approximation.

We show that this approach suffers from an instability that can produce misleading conclusions. Namely, for certain regimes of the model parameters, variational inference outputs a non-trivial decomposition into topics. However –for the same parameter values– the data contain no actual information about the true decomposition, and hence the output of the algorithm is uncorrelated with the true topic decomposition. Among other consequences, the estimated posterior mean is significantly wrong, and estimated Bayesian credible regions do not achieve the nominal coverage. We discuss how this instability is remedied by more accurate mean field approximations.

Keywords: Variational Inference; Topic Models; Approximate Message Passing; Mean Field Approximation; Credible Interval; TAP Free Energy.

1 Introduction

Topic modeling [12] aims at extracting the latent structure from a corpus of documents (either images or texts), that are represented as vectors $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_n \in \mathbb{R}^d$. The key assumption is that the n documents are (approximately) convex combinations of a small number k of topics $\tilde{\boldsymbol{h}}_1, \dots, \tilde{\boldsymbol{h}}_k \in \mathbb{R}^d$. Conditional on the topics, documents are generated independently by letting

$$\boldsymbol{x}_a = \frac{\sqrt{\beta}}{d} \sum_{\ell=1}^{k} w_{a,\ell} \tilde{\boldsymbol{h}}_\ell + \boldsymbol{z}_a \,, \tag{1.1}$$

where the weights $\boldsymbol{w}_a = (w_{a,\ell})_{1 \le \ell \le k}$ and noise vectors $\boldsymbol{z}_a$ are i.i.d. across $a \in \{1, \dots, n\}$. The scaling factor $\sqrt{\beta}/d$ is introduced for mathematical convenience (an equivalent parametrization would have been to scale $\boldsymbol{Z}$ by a noise-level parameter σ), and $\beta > 0$ can be interpreted as a signal-to-noise ratio. It is also useful to introduce the matrix $\boldsymbol{X} \in \mathbb{R}^{n \times d}$ whose i-th row is $\boldsymbol{x}_i$, and therefore

$$\boldsymbol{X} = \frac{\sqrt{\beta}}{d} \boldsymbol{W} \boldsymbol{H}^{\mathsf{T}} + \boldsymbol{Z} \,, \tag{1.2}$$

where $\boldsymbol{W} \in \mathbb{R}^{n \times k}$ and $\boldsymbol{H} \in \mathbb{R}^{d \times k}$. The a-th row of $\boldsymbol{W}$, is the vector of weights $\boldsymbol{w}_a$, while the rows of $\boldsymbol{H}$ will be denoted by $\boldsymbol{h}_i \in \mathbb{R}^k$.

Note that $\boldsymbol{w}_a$ belongs to the simplex $\mathsf{P}_1(k) = \{\boldsymbol{w} \in \mathbb{R}^k_{\ge 0} \; : \; \langle \boldsymbol{w}, \boldsymbol{1}_k \rangle = 1\}$. It is common to assume that its prior is Dirichlet: this class of models is known as *Latent Dirichlet Allocations*, or LDA [16]. Here we will take a particularly simple example of this type, and assume that the prior is Dirichlet in k dimensions with all parameters equal to ν (which we will denote by $\mathrm{Dir}(\nu; k)$). As for the topics $\boldsymbol{H}$, their prior distribution depends on the specific application. For instance, when applied to text corpora, the $\tilde{\boldsymbol{h}}_i$ are typically non-negative and represent normalized word count vectors. Here we will assume them to be standard Gaussian $(\tilde{\boldsymbol{h}}_i)_{i \le d} \sim_{iid} \mathsf{N}(0, \boldsymbol{I}_k)$. Finally, $\boldsymbol{Z}$ will be a noise matrix with entries $(Z_{ij})_{i \in [n], j \in [d]} \sim_{iid} \mathsf{N}(0, 1/d)$. We make these simplifying assumptions in order to keep analytical calculations manageable. It should be clear from our treatment that similar qualitative result should hold for more general models.

In fully Bayesian topic models, the parameters of the Dirichlet distribution, as well as the topic distributions are themselves unknown and to be learned from data. Here we will work in an idealized setting in which they are known. We will also assume that data are in fact distributed according to the postulated generative model. Since we are interested in studying some limitations of current approaches, our main point is only reinforced by assuming this idealized scenario.

As is common with Bayesian approaches, computing the posterior distribution of the factors $\boldsymbol{H}$, $\boldsymbol{W}$ given the data $\boldsymbol{X}$ is computationally challenging. Since the seminal work of Blei, Ng and Jordan [16], variational inference is the method of choice for addressing this problem within topic models. The term 'variational inference' refers to a broad class of methods that aim at approximating the posterior computation by solving an optimization problem, see [30, 45, 13] for background. A popular starting point is the Gibbs variational principle, namely the fact that the posterior solves the following convex optimization problem:

$$p_{\boldsymbol{W},\boldsymbol{H}|\boldsymbol{X}}(\,\cdot\,,\cdot\,|\boldsymbol{X}) = \arg\min_{q\in\mathcal{P}_{n,d,k}} \mathrm{KL}(q\|p_{\boldsymbol{W},\boldsymbol{H}|\boldsymbol{X}}) \tag{1.3}$$

$$= \arg\min_{q\in\mathcal{P}_{n,d,k}} \Big\{ -\mathbb{E}_q \log p_{\boldsymbol{X}|\boldsymbol{W},\boldsymbol{H}}(\boldsymbol{X}|\boldsymbol{H},\boldsymbol{W}) + \mathrm{KL}(q\|p_{\boldsymbol{W}}\times p_{\boldsymbol{H}})\Big\}\,, \tag{1.4}$$

where $\mathrm{KL}(\,\cdot\,\|\,\cdot\,)$ denotes the Kullback-Leibler divergence. The variational expression in Eq. (1.4) is also known as the Gibbs free energy. Optimization is within the space $\mathcal{P}_{n,d,k}$ of probability measures on $\boldsymbol{H},\boldsymbol{W}$. To be precise, we always assume that a dominating measure ν_0 over $\mathbb{R}^{n\times k}\times\mathbb{R}^{d\times k}$ is given for $\boldsymbol{W},\boldsymbol{H}$, and both $p_{\boldsymbol{W},\boldsymbol{H}|\boldsymbol{X}}$ and q have densities with respect to ν_0: we hence identify the measure with its density. Throughout the paper (with the exception of the example in Section 2) ν_0 can be taken to be the Lebesgue measure.

Even for $\boldsymbol{W},\boldsymbol{H}$ discrete, the Gibbs principle has exponentially many decision variables. Variational methods differ in the way the problem (1.3) is approximated. The main approach within topic modeling is *naive mean field*, which restricts the optimization problem to the space of probability measures that factorize over the rows of $\boldsymbol{W},\boldsymbol{H}$:

$$\hat{q}\,(\boldsymbol{W},\boldsymbol{H}) = q\,(\boldsymbol{H})\,\tilde{q}\,(\boldsymbol{W}) = \prod_{i=1}^{d} q_i\,(\boldsymbol{h}_i)\prod_{a=1}^{n}\tilde{q}_a\,(\boldsymbol{w}_a)\ . \tag{1.5}$$

By a suitable parametrization of the marginals q_i, $\tilde{q}_a$, this leads to an optimization problem of dimension $O((n+d)k)$, cf. Section 3. Despite being non-convex, this problem is separately convex in the $(q_i)_{i\le d}$ and $(\tilde{q}_a)_{a\le n}$, which naturally suggests the use of an alternating minimization algorithm which has been successfully deployed in a broad range of applications ranging from computer vision to genetics [25, 49, 42]. We will refer to this as to the *naive mean field iteration*. Following a common use in the topics models literature, we will use the terms 'variational inference' and 'naive mean field' interchangeably.

The main result of this paper is that naive mean field presents an instability for learning Latent Dirichlet Allocations. We will focus on the limit $n, d \to \infty$ with $n/d = \delta$ fixed. Hence, an LDA distribution is determined by the parameters (k, δ, ν, β). We will show that there are regions in this parameter space such that the following two findings hold simultaneously:

No non-trivial estimator. Any estimator $\widehat{\boldsymbol{H}}$, $\widehat{\boldsymbol{W}}$ of the topic or weight matrices is asymptotically uncorrelated with the real model parameters $\boldsymbol{H}, \boldsymbol{W}$. In other words, the data do not contain enough signal to perform any strong inference.

Variational inference is randomly biased. Given the above, one would hope the Bayesian posterior to be centered on an unbiased estimate. In particular, $p(\boldsymbol{w}_a|\boldsymbol{X})$ (the posterior distribution over weights of document a) should be centered around the uniform distribution $\boldsymbol{w}_a = (1/k, \dots, 1/k)$. In contrast, we will show that the posterior produced by naive mean field is centered around a random distribution that is uncorrelated with the actual weights. Similarly, the posterior over topic vectors is centered around random vectors uncorrelated with the true topics.

One key argument in support of Bayesian methods is the hope that they provide a measure of uncertainty of the estimated variables. In view of this, the failure just described is particularly dangerous because it suggests some measure of certainty, although the estimates are essentially random.

Is there a way to eliminate this instability by using a better mean field approximation? We show that a promising approach is provided by a classical idea in statistical physics, the Thouless-Anderson-Palmer (TAP) free energy [44, 38]. This suggests a variational principle that is analogous in form to naive mean field, but provides a more accurate approximation of the Gibbs principle:

Variational inference via the TAP free energy. We show that the instability of naive mean field is remedied by using the TAP free energy instead of the naive mean field free energy. The latter can be optimized using an iterative scheme that is analogous to the naive mean field iteration and is known as approximate message passing (AMP).

While the TAP approach is promising –at least for synthetic data– we believe that further work is needed to develop a reliable inference scheme.

The rest of the paper is organized as follows. Section 2 discusses a simpler example, $\mathbb{Z}_2$-synchronization, which shares important features with latent Dirichlet allocations. Since calculations are fairly straightforward, this example allows to explain the main mathematical points in a simple context. We then present our main results about instability of naive mean field in Section 3, and discuss the use of TAP free energy to overcome the instability in Section 4.

1.1 Related literature

Over the last fifteen years, topic models have been generalized to cover an impressive number of applications. A short list includes mixed membership models [24, 1], dynamic topic models [14], correlated topic models [33, 15], spatial LDA [50], relational topic models [19], Bayesian tensor models [53]. While other approaches have been used (e.g. Gibbs sampling), variational algorithms are among the most popular methods for Bayesian inference in these models. Variational methods provide a fairly complete and interpretable description of the posterior, while allowing to leverage advances in optimization algorithms and architectures towards this goal (see [28, 17]).

Despite this broad empirical success, little is rigorously known about the accuracy of variational inference in concrete statistical problems. Wang and Titterington [46, 48] prove local convergence of naive mean field estimate to the true parameters for exponential families with missing data and Gaussian mixture models. In the context of Gaussian mixtures, the same authors prove that the covariance of the variational posterior is asymptotically smaller (in the positive semidefinite order) than the inverse of the Fisher information matrix [47]. All of these results are established in the classical large sample asymptotics $n \to \infty$ with d fixed. In the present paper we focus instead on the high-dimensional limit $n = \Theta(d)$ and prove that also the mode (or mean) of the variational posterior is incorrect. Notice that the high-dimensional regime is particularly relevant for the analysis of Bayesian methods. Indeed, in the classical low-dimensional asymptotics Bayesian approaches do not outperform maximum likelihood.

In order to correct for the underestimation of covariances, [47] suggest replacing its variational estimate by the inverse Fisher information matrix. A different approach is developed in [27], building on linear response theory.

Naive mean field variational inference was used in [18, 10] to estimate the parameters of the stochastic block model. These works establish consistency and asymptotic normality of the variational estimates in a large signal-to-noise ratio regime. Our work focuses on estimating the latent factors: it would be interesting to consider implications on parameter estimation as well.

The recent paper [52] also studies variational inference in the context of the stochastic block model, but focuses on reconstructing the latent vertex labels. The authors prove that naive mean field achieves minimax optimal statistical rates. Let us emphasize that this problem is closely related to topic models: both are models for approximately low-rank matrices, with a probabilistic prior on the factors. The results of [52] are complementary to ours, in the sense that [52] establishes positive results at large signal-to-noise ratio (albeit for a different model), while we prove inconsistency at low signal-to-noise ratio.

General conditions for consistency of variational Bayes methods have been recently developed in [51] and [40]. The first paper establishes a Bernstein-von Mises type theorem for variational inference, while the latter proves consistency of point estimates derived from variational inference. Once more, these works focus on a high signal-to-noise regime.

Our work also builds on recent theoretical advances in high-dimensional low-rank models, that were mainly driven by techniques from mathematical statistical physics (more specifically, spin glass theory). An incomplete list of relevant references includes [31, 21, 20, 32, 6, 34, 36, 35, 2]. These papers prove asymptotically exact characterizations of the Bayes optimal estimation error in low-rank models, to an increasing degree of generality, under the high-dimensional scaling $n, d \to \infty$ with $n/d \to \delta \in (0, \infty)$.

Related ideas also suggest an iterative algorithm for Bayesian estimation, namely Bayes Approximate Message Passing [22, 23]. As mentioned above, Bayes AMP can be regarded as minimizing a different variational approximation known as the TAP free energy. An important advantage over naive mean field is that AMP can be rigorously analyzed using a method known as state evolution [7, 29, 9].

Let us finally mention that a parallel line of work develops polynomial-time algorithms to construct non-negative matrix factorizations under certain structural assumptions on the data matrix $\boldsymbol{X}$, such as separability [4, 3, 43]. It should be emphasized that the objective of these algorithms is different from the one of Bayesian methods: they return a factorization that is guaranteed to be unique under separability. In contrast, variational methods attempt to approximate the posterior distribution, when the data are generated according to the LDA model.

1.2 Notations

We denote by $\boldsymbol{I}_m$ the identity matrix, and by $\boldsymbol{J}_m$ the all-ones matrix in m dimensions (subscripts will be dropped when the number of dimensions is clear from the context). We use $\mathbf{1}_k \in \mathbb{R}^k$ for the all-ones vector.

We will use $\otimes$ for the tensor (outer) product. In particular, given vectors expressed in the canonical basis as $\boldsymbol{u} = \sum_{i=1}^{d_1} u_i \boldsymbol{e}_i \in \mathbb{R}^{d_1}$ and $\boldsymbol{v} = \sum_{i=j}^{d_2} v_j \boldsymbol{e}_j \in \mathbb{R}^{d_2}$, $\boldsymbol{u} \otimes \boldsymbol{v} \in \mathbb{R}^{d_1} \otimes \mathbb{R}^{d_2}$ is the tensor with coordinates $(\boldsymbol{u} \otimes \boldsymbol{v})_{ij} = u_i v_j$ in the basis $\boldsymbol{e}_i \otimes \boldsymbol{e}_j$. We will identify the space of matrices $\mathbb{R}^{d_1 \times d_2}$ with the tensor product $\mathbb{R}^{d_1} \otimes \mathbb{R}^{d_2}$. In particular, for $\boldsymbol{u} \in \mathbb{R}^{d_1}$, $\boldsymbol{v} \in \mathbb{R}^{d_2}$, we identify $\boldsymbol{u} \otimes \boldsymbol{v}$ with the matrix $\boldsymbol{u}\boldsymbol{v}^{\mathsf{T}}$.

Given a symmetric matrix $\boldsymbol{M} \in \mathbb{R}^{n \times n}$, we denote by $\lambda_1(\boldsymbol{M}) \geq \lambda_2(\boldsymbol{M}) \geq \cdots \geq \lambda_n(\boldsymbol{M})$ its eigenvalues in decreasing order. For a matrix (or vector) $\boldsymbol{A} \in \mathbb{R}^{d \times n}$ we denote the orthogonal projector operator onto the subspace spanned by the columns of $\boldsymbol{A}$ by $\boldsymbol{P}_{\boldsymbol{A}} \in \mathbb{R}^{d \times d}$, and its orthogonal complement by $\boldsymbol{P}_{\boldsymbol{A}}^{\perp} = \boldsymbol{I}_d - \boldsymbol{P}_{\boldsymbol{A}}$. When the subscript is omitted, this is understood to be the projector onto the space spanned by the all-ones vector: $\boldsymbol{P} = \mathbf{1}_d \mathbf{1}_d / d$ and $\boldsymbol{P}_{\perp} = \boldsymbol{I}_d - \boldsymbol{P}$.

2 A simple example: $\mathbb{Z}_2$-synchronization

In $\mathbb{Z}_2$ synchronization we are interested in estimating a vector $\boldsymbol{\sigma} \in \{+1, -1\}^n$ from observations $\boldsymbol{X} \in \mathbb{R}^{n\times n}$, generated according to

$$\boldsymbol{X} = \frac{\lambda}{n}\boldsymbol{\sigma}\boldsymbol{\sigma}^{\mathsf{T}} + \boldsymbol{Z}\,, \tag{2.1}$$

where $\boldsymbol{Z} = \boldsymbol{Z}^{\mathsf{T}} \in \mathbb{R}^{n\times n}$ is distributed according to the Gaussian Orthogonal Ensemble $\mathsf{GOE}(n)$, namely $(Z_{ij})_{i<j\le n} \sim_{iid} \mathsf{N}(0, 1/n)$ are independent of $(Z_{ii})_{i\le n} \sim_{iid} \mathsf{N}(0, 2/n)$. The parameter $\lambda \ge 0$ corresponds to the signal-to-noise ratio.

It is known that for $\lambda \le 1$ no algorithm can estimate $\boldsymbol{\sigma}$ from data $\boldsymbol{X}$ with positive correlation in the limit $n \to \infty$. The following is an immediate consequence of [31, 20], see Appendix C.1.

Lemma 2.1. *Under model (2.1), for $\lambda \le 1$ and any estimator $\hat{\boldsymbol{\sigma}} : \mathbb{R}^{n\times n} \to \mathbb{R}^n \setminus \{\mathbf{0}\}$, the following limit holds in probability:*

$$\limsup_{n\to\infty} \frac{|\langle \hat{\boldsymbol{\sigma}}(\boldsymbol{X}), \boldsymbol{\sigma}\rangle|}{\|\hat{\boldsymbol{\sigma}}(\boldsymbol{X})\|_2 \|\boldsymbol{\sigma}\|_2} = 0\,. \tag{2.2}$$

How does variational inference perform on this problem? Any product probability distribution $\hat{q}(\boldsymbol{\sigma}) = \prod_{i=1}^n q_i(\sigma_i)$ can be parametrized by the means $m_i = \sum_{\sigma_i\in\{+1,-1\}} q_i(\sigma_i)\,\sigma_i$, and it is immediate to get

$$\mathrm{KL}(\hat{q}\|p_{\boldsymbol{\sigma}|\boldsymbol{X}}) = \mathcal{F}(\boldsymbol{m}) + \text{const.}\,, \tag{2.3}$$

$$\mathcal{F}(\boldsymbol{m}) \equiv -\frac{\lambda}{2}\langle \boldsymbol{m}, \boldsymbol{X}_0\boldsymbol{m}\rangle - \sum_{i=1}^n \mathsf{h}(m_i)\,. \tag{2.4}$$

Here $\boldsymbol{X}_0$ is obtained from $\boldsymbol{X}$ by setting the diagonal entries to 0, and $\mathsf{h}(x) = -(1 + x)/2\log((1 + x)/2) - (1 - x)/2\log((1 - x)/2)$ is the binary entropy function. In view of Lemma 2.1, the correct posterior distribution should be essentially uniform, resulting in $\boldsymbol{m}$ vanishing. Indeed, $\boldsymbol{m}_* = 0$ is a stationary point of the mean field free energy $\mathcal{F}(\boldsymbol{m})$: $\nabla\mathcal{F}(\boldsymbol{m})|_{\boldsymbol{m}=\boldsymbol{m}_*} = 0$. We refer to this as the 'uninformative fixed point'.

Is $\boldsymbol{m}_*$ a local minimum? Computing the Hessian at the uninformative fixed point yields

$$\nabla^2 \mathcal{F}(\boldsymbol{m})\Big|_{\boldsymbol{m}=\boldsymbol{m}_*} = -\lambda \boldsymbol{X}_0 + \boldsymbol{I}\,. \tag{2.5}$$

The matrix $\boldsymbol{X}_0$ is a rank-one deformation of a Wigner matrix and its spectrum is well understood [5, 26, 8]. For $\lambda \le 1$, its eigenvalues are contained with high probability in the interval $[-2, 2]$, with $\lambda_{\min}(\boldsymbol{X}) \to -2$, $\lambda_{\max}(\boldsymbol{X}) \to 2$ as $n \to \infty$. For $\lambda > 1$, $\lambda_{\max}(\boldsymbol{X}) \to \lambda + \lambda^{-1}$, while the other eigenvalues are contained in $[-2, 2]$. This implies

$$\lim_{n\to\infty} \lambda_{\min}\big(\nabla^2 \mathcal{F}\Big|_{\boldsymbol{m}_*}\big) = \begin{cases} 1 - 2\lambda & \text{if } \lambda \le 1, \\ -\lambda^2 & \text{if } \lambda > 1. \end{cases} \tag{2.6}$$

In other words, $\boldsymbol{m}_* = 0$ is a local minimum for $\lambda < 1/2$, but becomes a saddle point for $\lambda > 1/2$. In particular, for $\lambda \in (1/2, 1)$, variational inference will produce an estimate $\hat{\boldsymbol{m}} \neq 0$, although the posterior should be essentially uniform. In fact, it is possible to make this conclusion more quantitative.

Proposition 2.2. *Let $\hat{\boldsymbol{m}} \in [-1, 1]^n$ be any local minimum of the mean field free energy $\mathcal{F}(\boldsymbol{m})$, under the $\mathbb{Z}_2$-synchronization model (2.1). Then there exists a numerical constant $c_0 > 0$ such that, with high probability, for $\lambda > 1/2$,*

$$\frac{1}{n}\|\hat{\boldsymbol{m}}\|_2^2 \ge c_0 \,\min\Big((2\lambda - 1)^2, 1\Big)\,. \tag{2.7}$$

In other words, although no estimator is positively correlated with the true signal $\boldsymbol{\sigma}$, variational inference outputs biases $\hat{m}_i$ that are non-zero (and indeed of order one, for a positive fraction of them).

The last statement immediately implies that naive mean field leads to incorrect inferential statements for $\lambda \in (1/2, 1)$. In order to formalize this point, given any estimators $\{\hat{q}_i(\,\cdot\,)\}_{i\le n}$ of the posterior marginals, we define the per-coordinate expected coverage as

$$\mathscr{Q}(\hat{q}) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{P}\Big(\sigma_i = \arg\max_{\tau_i \in \{+1,-1\}} \hat{q}_i(\tau_i)\Big)\,. \tag{2.8}$$

This is the expected fraction of coordinates that are estimated correctly by choosing $\boldsymbol{\sigma}$ according to the estimated posterior. Since the prior is assumed to be correct, it can be interpreted either as the expectation (with respect to the parameters) of the frequentist coverage, or as the expectation (with respect to the data) of the Bayesian coverage. On the other hand, if the $\hat{q}_i$ were accurate, Bayesian theory would suggest claiming the coverage

$$\widehat{\mathscr{Q}}(\hat{q}) \equiv \frac{1}{n}\sum_{i\le n}\max_{\tau_i}\hat{q}_i(\tau_i)\,. \tag{2.9}$$

The following corollary is a direct consequence of Proposition 2.2, and formalizes the claim that naive mean field leads to incorrect inferential statements. More precisely, it overestimates the coverage achieved.

Corollary 2.3. *Let $\hat{\boldsymbol{m}} \in [-1,1]^n$ be any local minimum of the mean field free energy $\mathcal{F}(\boldsymbol{m})$, under the $\mathbb{Z}_2$-synchronization model (2.1), and consider the corresponding posterior marginal estimates $\hat{q}_i(\sigma_i) = (1+\hat{m}_i\sigma_i)/2$. Then, there exists a numerical constant $c_0 > 0$ such that, with high probability, for $\lambda \in (1/2, 1)$,*

$$\mathscr{Q}(\hat{q}) \le \frac{1}{2} + o_n(1)\,, \qquad \widehat{\mathscr{Q}}(\hat{q}) \ge \frac{1}{2} + c_0\,\min\Big((2\lambda-1),1\Big)\,. \tag{2.10}$$

While similar formal coverage statements can be obtained also for the more complex case of topic models, we will not make them explicit, since they are relatively straightforward consequences of our analysis.

3 Instability of variational inference for topic models

3.1 Information-theoretic limit

As in the case of $\mathbb{Z}_2$ synchronization discussed in Section 2, we expect it to be impossible to estimate the factors $\boldsymbol{W}, \boldsymbol{H}$ with strictly positive correlation for small enough signal-to-noise ratio β (or small enough sample size δ). The exact threshold was characterized recently in [36] (but see also [21, 6, 34, 35] for closely related results). The characterization in [36] is given in terms of a variational principle over $k \times k$ matrices.

Theorem 1 (Special case of [36]). *Let* $\mathrm{I}_n(\boldsymbol{X};\boldsymbol{W},\boldsymbol{H})$ *denote the mutual information between the data* $\boldsymbol{X}$ *and the factors* $\boldsymbol{H},\boldsymbol{W}$ *under the LDA model (1.2). Then, the following limit holds almost surely*

$$\lim_{n,d\to\infty}\frac{1}{d}\mathrm{I}_n(\boldsymbol{X};\boldsymbol{W},\boldsymbol{H}) = \inf_{\boldsymbol{M}\in\mathbb{S}_k}\mathsf{RS}(\boldsymbol{M};k,\delta,\nu)\,, \tag{3.1}$$

where $\mathbb{S}_k$ *is the cone of* $k\times k$ *positive semidefinite matrices and* $\mathsf{RS}(\cdots)$ *is a function given explicitly in Appendix C.2.*

It is also shown in Appendix C.2 that $\boldsymbol{M}^* = (\delta\beta/k^2)\boldsymbol{J}_k$ is a stationary point of the free energy $\mathsf{RS}(\boldsymbol{M};k,\delta,\nu)$. We shall refer to $\boldsymbol{M}^*$ as the uninformative point. Let $\beta_{\text{Bayes}} = \beta_{\text{Bayes}}(k,\delta,\nu)$ be the supremum value of β such that the infimum in Eq. (3.1) is uniquely achieved at $\boldsymbol{M}^*$:

$$\beta_{\text{Bayes}}(k,\delta,\nu) = \sup\Big\{\beta\geq 0:\ \ \mathsf{RS}(\boldsymbol{M};k,\delta,\nu) > \mathsf{RS}(\boldsymbol{M}_*;k,\delta,\nu)\quad\text{for all}\quad \boldsymbol{M}\neq\boldsymbol{M}_*\Big\}. \tag{3.2}$$

As formalized below, for $\beta < \beta_{\text{Bayes}}$ the data $\boldsymbol{X}$ do not contain sufficient information for estimating $\boldsymbol{H}$, $\boldsymbol{W}$ in a non-trivial manner.

Proposition 3.1. *Let* $\boldsymbol{M}_* = \delta\beta\boldsymbol{J}_k/k^2$. *Then* $\boldsymbol{M}^*$ *is a stationary point of the function* $\boldsymbol{M}\mapsto\mathsf{RS}(\boldsymbol{M};\beta,k,\delta,\nu)$. *Further, it is a local minimum provided* $\beta < \beta_{\text{spect}}(k,\delta,\nu)$ *where the spectral threshold is given by*

$$\beta_{\text{spect}} \equiv \frac{k(k\nu+1)}{\sqrt{\delta}}. \tag{3.3}$$

Finally, if $\beta < \beta_{\text{Bayes}}(k,\delta,\nu)$, *for any estimator* $\boldsymbol{X}\mapsto\widehat{\boldsymbol{F}}_n(\boldsymbol{X})$, *we have*

$$\liminf_{n\to\infty}\mathbb{E}\Big\{\big\|\boldsymbol{W}\boldsymbol{H}^{\mathsf{T}}-\widehat{\boldsymbol{F}}_n(\boldsymbol{X})\big\|_F^2\Big\}\geq\lim_{n\to\infty}\mathbb{E}\Big\{\big\|\boldsymbol{W}\boldsymbol{H}^{\mathsf{T}}-c\boldsymbol{1}_n(\boldsymbol{X}^{\mathsf{T}}\boldsymbol{1}_n)^{\mathsf{T}}\big\|_F^2\Big\}\,, \tag{3.4}$$

for $c\equiv\sqrt{\beta}/(k+\beta\delta)$ *a constant.*

We refer to Appendix C for a proof of this statement.

Note that Eq. (3.4) compares the mean square error of an arbitrary estimator $\widehat{\boldsymbol{F}}_n$, to the mean square error of the trivial estimator that replaces each column of $\boldsymbol{X}$ by its average.

This is equivalent to estimating all the weights $\boldsymbol{w}_i$ by the uniform distribution $\mathbf{1}_k/k$. Of course, $\beta_{\text{Bayes}} \le \beta_{\text{spect}}$. However, this upper bound appears to be tight for small k.

Remark 3.1. Solving numerically the $k(k+1)/2$-dimensional problem (3.1) indicates that $\beta_{\text{Bayes}}(k,\nu,\delta) = \beta_{\text{spect}}(k,\nu,\delta)$ for $k \in \{2,3\}$ and $\nu = 1$.

3.2 Naive mean field free energy

We consider a trial joint distribution that factorizes according to rows of $\boldsymbol{W}$ and $\boldsymbol{H}$ according to Eq. (1.5). It turns out (see Appendix D.2) that, for any stationary point of $\mathrm{KL}(\hat{q}\|p_{\boldsymbol{H},\boldsymbol{W}|\boldsymbol{X}})$ over such product distributions, the marginals take the form

$$
\begin{aligned}
q_i(\boldsymbol{h}) &= \exp\left\{\langle \boldsymbol{m}_i, \boldsymbol{h}\rangle - \frac{1}{2}\langle \boldsymbol{h}, \boldsymbol{Q}_i \boldsymbol{h}\rangle - \phi(\boldsymbol{m}_i, \boldsymbol{Q}_i)\right\} q_0(\boldsymbol{h}) \,, \\
\tilde{q}_a(\boldsymbol{w}) &= \exp\left\{\langle \tilde{\boldsymbol{m}}_a, \boldsymbol{w}\rangle - \frac{1}{2}\left\langle \boldsymbol{w}, \widetilde{\boldsymbol{Q}}_a \boldsymbol{w}\right\rangle - \tilde{\phi}(\tilde{\boldsymbol{m}}_a, \widetilde{\boldsymbol{Q}}_a)\right\} \tilde{q}_0(\boldsymbol{w}) \,,
\end{aligned}
\tag{3.5}
$$

where $q_0(\,\cdot\,)$ is the density of $\mathsf{N}(0, \boldsymbol{I}_k)$, and $\tilde{q}_0(\,\cdot\,)$ is the density of $\mathrm{Dir}(\nu; k)$, and $\phi, \tilde{\phi} : \mathbb{R}^k \times \mathbb{R}^{k\times k} \to \mathbb{R}$ are defined implicitly by the normalization condition $\int q_i(\mathrm{d}\boldsymbol{h}_i) = \int \tilde{q}_a(\mathrm{d}\boldsymbol{w}_a) = 1$. In the following we let $\boldsymbol{m} = (\boldsymbol{m}_i)_{i\le d}$, $\tilde{\boldsymbol{m}} = (\tilde{\boldsymbol{m}}_a)_{a\le n}$ denote the set of parameters in these distributions; these can also be viewed as matrices $\boldsymbol{m} \in \mathbb{R}^{d\times k}$ and $\tilde{\boldsymbol{m}} \in \mathbb{R}^{d\times k}$ whose i-th row is $\boldsymbol{m}_i$ (in the former case) or $\tilde{\boldsymbol{m}}_i$ (in the latter).

It is useful to define the functions $\mathsf{F}, \widetilde{\mathsf{F}} : \mathbb{R}^k \times \mathbb{R}^{k\times k} \to \mathbb{R}^k$ and $\mathsf{G}, \widetilde{\mathsf{G}} : \mathbb{R}^k \times \mathbb{R}^{k\times k} \to \mathbb{R}^{k\times k}$ as (proportional to) expectations with respect to the approximate posteriors (3.5)

$$
\mathsf{F}(\boldsymbol{m}_i; \boldsymbol{Q}) \equiv \sqrt{\beta}\int \boldsymbol{h}\; q_i(\mathrm{d}\boldsymbol{h})\,, \quad \widetilde{\mathsf{F}}(\tilde{\boldsymbol{m}}_a; \widetilde{\boldsymbol{Q}}) \equiv \sqrt{\beta}\int \boldsymbol{w}\; \tilde{q}_a(\mathrm{d}\boldsymbol{w})\,, \tag{3.6}
$$

$$
\mathsf{G}(\boldsymbol{m}_i; \boldsymbol{Q}) \equiv \beta\int \boldsymbol{h}^{\otimes 2}\; q_i(\mathrm{d}\boldsymbol{h})\,, \quad \widetilde{\mathsf{G}}(\tilde{\boldsymbol{m}}_a; \widetilde{\boldsymbol{Q}}) \equiv \beta\int \boldsymbol{w}^{\otimes 2}\; \tilde{q}_a(\mathrm{d}\boldsymbol{w})\,. \tag{3.7}
$$

For $\boldsymbol{m} \in \mathbb{R}^{d\times k}$, we overload the notation and denote by $\mathsf{F}(\boldsymbol{m}; \boldsymbol{Q}) \in \mathbb{R}^{d\times k}$ the matrix whose i-th row is $\mathsf{F}(\boldsymbol{m}_i; \boldsymbol{Q})$ (and similarly for $\widetilde{\mathsf{F}}(\tilde{\boldsymbol{m}}; \widetilde{\boldsymbol{Q}})$).

When restricted to a product-form ansatz with parametrization (3.5), the mean field free energy takes the form (see Appendix D.3)

$$\mathrm{KL}(\hat{q}\|p_{\boldsymbol{W},\boldsymbol{H}|\boldsymbol{X}}) = \mathcal{F}(\boldsymbol{r},\tilde{\boldsymbol{r}},\boldsymbol{\Omega},\tilde{\boldsymbol{\Omega}}) + \frac{d}{2}\|\boldsymbol{X}\|_F^2 + \log p_{\boldsymbol{X}}(\boldsymbol{X})\,, \tag{3.8}$$

where

$$\mathcal{F}(\boldsymbol{r},\tilde{\boldsymbol{r}},\boldsymbol{\Omega},\tilde{\boldsymbol{\Omega}}) = \sum_{i=1}^{d}\psi_*(\boldsymbol{r}_i,\boldsymbol{\Omega}_i) + \sum_{a=1}^{n}\tilde{\psi}_*(\tilde{\boldsymbol{r}}_a,\tilde{\boldsymbol{\Omega}}) - \sqrt{\beta}\,\mathsf{Tr}\left(\boldsymbol{X}\boldsymbol{r}\tilde{\boldsymbol{r}}^{\mathsf{T}}\right) + \frac{\beta}{2d}\sum_{i=1}^{d}\sum_{a=1}^{n}\langle\boldsymbol{\Omega}_i,\tilde{\boldsymbol{\Omega}}_a\rangle\,, \tag{3.9}$$

$$\psi_*(\boldsymbol{r},\boldsymbol{\Omega}) \equiv \sup_{\boldsymbol{m},\boldsymbol{Q}}\left\{\langle\boldsymbol{r},\boldsymbol{m}\rangle - \frac{1}{2}\langle\boldsymbol{\Omega},\boldsymbol{Q}\rangle - \phi(\boldsymbol{m},\boldsymbol{Q})\right\}\,, \tag{3.10}$$

$$\tilde{\psi}_*(\tilde{\boldsymbol{r}},\tilde{\boldsymbol{\Omega}}) \equiv \sup_{\tilde{\boldsymbol{m}},\widetilde{\boldsymbol{Q}}}\left\{\langle\tilde{\boldsymbol{r}},\tilde{\boldsymbol{m}}\rangle - \frac{1}{2}\langle\tilde{\boldsymbol{\Omega}},\widetilde{\boldsymbol{Q}}\rangle - \tilde{\phi}(\tilde{\boldsymbol{m}},\widetilde{\boldsymbol{Q}})\right\}. \tag{3.11}$$

Note that Eq. (3.11) implies the following convex duality relation between $(\boldsymbol{r},\tilde{\boldsymbol{r}},\boldsymbol{\Omega},\tilde{\boldsymbol{\Omega}})$ and $(\boldsymbol{m},\tilde{\boldsymbol{m}},\boldsymbol{Q},\widetilde{\boldsymbol{Q}})$

$$\boldsymbol{r}_i \equiv \frac{1}{\sqrt{\beta}}\mathsf{F}(\boldsymbol{m}_i;\boldsymbol{Q}), \quad \tilde{\boldsymbol{r}}_a \equiv \frac{1}{\sqrt{\beta}}\widetilde{\mathsf{F}}(\tilde{\boldsymbol{m}}_a;\widetilde{\boldsymbol{Q}})\,, \tag{3.12}$$

$$\boldsymbol{\Omega}_i \equiv \frac{1}{\beta}\mathsf{G}(\boldsymbol{m}_i;\boldsymbol{Q}), \quad \tilde{\boldsymbol{\Omega}}_a \equiv \frac{1}{\beta}\widetilde{\mathsf{G}}(\tilde{\boldsymbol{m}}_a;\widetilde{\boldsymbol{Q}})\,. \tag{3.13}$$

By strict convexity of $\phi(\boldsymbol{m},\boldsymbol{Q})$, $\tilde{\phi}(\tilde{\boldsymbol{m}},\widetilde{\boldsymbol{Q}})$ (the latter is strongly convex on the hyperplane $\langle\boldsymbol{1},\tilde{\boldsymbol{m}}\rangle = 0$, $\langle\boldsymbol{1},\widetilde{\boldsymbol{Q}}\boldsymbol{1}\rangle = 0$) we can view $\mathcal{F}(\cdots)$ as a function of $(\boldsymbol{r},\tilde{\boldsymbol{r}},\boldsymbol{\Omega},\tilde{\boldsymbol{\Omega}})$ or $(\boldsymbol{m},\tilde{\boldsymbol{m}},\boldsymbol{Q},\widetilde{\boldsymbol{Q}})$. With an abuse of notation, we will write $\mathcal{F}(\boldsymbol{r},\tilde{\boldsymbol{r}},\boldsymbol{\Omega},\tilde{\boldsymbol{\Omega}})$ or $\mathcal{F}(\boldsymbol{m},\tilde{\boldsymbol{m}},\boldsymbol{Q},\widetilde{\boldsymbol{Q}})$ interchangeably.

A critical (stationary) point of the free energy (3.9) is a point at which $\nabla\mathcal{F}(\boldsymbol{m},\tilde{\boldsymbol{m}},\boldsymbol{Q},\widetilde{\boldsymbol{Q}}) = \boldsymbol{0}$. It turns out that the mean field free energy always admits a point that does not distinguish between the k latent factors, and in particular $\boldsymbol{m} = \boldsymbol{v}\boldsymbol{1}_k^{\mathsf{T}}$, $\tilde{\boldsymbol{m}} = \tilde{\boldsymbol{v}}\boldsymbol{1}_k^{\mathsf{T}}$, as stated in detail below. We will refer to this as the *uninformative critical point* (or *uninformative fixed point*).

Lemma 3.2. *Define*

$$\mathsf{E}(q;\nu) \equiv \frac{\int w_1^2\exp\{-q\|\boldsymbol{w}\|_2^2\}\,\tilde{q}_0(\mathrm{d}\boldsymbol{w})}{\int\exp\{-q\|\boldsymbol{w}\|_2^2\}\,\tilde{q}_0(\mathrm{d}\boldsymbol{w})}$$

and let q_1^* *be any solution of the following equation in* $[0,\infty)$

$$q_1^* = \frac{k\beta\delta}{k-1}\left\{\mathsf{E}\left(\frac{\beta}{1+q_1^*};\nu\right) - \frac{1}{k^2}\right\}\,. \tag{3.14}$$

(Such a solution always exists.) Further define

$$q_2^* = \frac{\beta\delta - kq_1^*}{k^2}\,, \qquad \tilde{q}_1^* = \frac{\beta}{1+q_1^*}, \tag{3.15}$$

$$\tilde{q}_2^* = \beta\left(\frac{\|\boldsymbol{X}^{\mathsf{T}}\mathbf{1}_n\|_2^2}{d(1+q_1^*+kq_2^*)^2} - \frac{q_2^*}{(1+q_1^*)(1+q_1^*+kq_2^*)}\right). \tag{3.16}$$

Then the naive mean field free energy of Eq. (3.9) admits a stationary point whereby, for all $i \in [d]$, $a \in [n]$,

$$\boldsymbol{m}_i^* = \frac{\sqrt{\beta}}{k}\,(\boldsymbol{X}^{\mathsf{T}}\mathbf{1}_n)_i\,\mathbf{1}_k\,, \tag{3.17}$$

$$\tilde{\boldsymbol{m}}_a^* = \frac{\beta}{k(1+q_1^*+kq_2^*)}\,(\boldsymbol{X}\boldsymbol{X}^{\mathsf{T}}\mathbf{1}_n)_a\,\mathbf{1}_k\,, \tag{3.18}$$

$$\boldsymbol{Q}_i^* = q_1^*\boldsymbol{I}_k + q_2^*\boldsymbol{J}_k\,, \quad \widetilde{\boldsymbol{Q}}_a^* = \tilde{q}_1^*\boldsymbol{I}_k + \tilde{q}_2^*\boldsymbol{J}_k\,. \tag{3.19}$$

The proof of this lemma is deferred to Appendix D.4. We note that Eq. (3.14) appears to always have a unique solution. Although we do not have a proof of uniqueness, in Appendix J we prove that the solution is unique conditional on a certain inequality that can be easily checked numerically.

3.3 Naive mean field iteration

As mentioned in the introduction, the variational approximation of the free energy is often minimized by alternating minimization over the marginals $(q_i)_{i\le d}$, $(\tilde{q}_a)_{a\le n}$ of Eq. (1.5). Using the parametrization (3.5), we obtain the following naive mean field iteration for $\boldsymbol{m}^t, \tilde{\boldsymbol{m}}^t, \boldsymbol{Q}^t, \widetilde{\boldsymbol{Q}}^t$ (see Appendix D.2):

$$\boldsymbol{m}^{t+1} = \boldsymbol{X}^{\mathsf{T}}\,\widetilde{\mathsf{F}}(\tilde{\boldsymbol{m}}^t;\widetilde{\boldsymbol{Q}}^t), \quad \boldsymbol{Q}^{t+1} = \frac{1}{d}\sum_{a=1}^{n}\widetilde{\mathsf{G}}(\tilde{\boldsymbol{m}}_a^t;\widetilde{\boldsymbol{Q}}^t)\,, \tag{3.20}$$

$$\tilde{\boldsymbol{m}}^t = \boldsymbol{X}\,\mathsf{F}(\boldsymbol{m}^t;\boldsymbol{Q}^t)\,, \qquad \widetilde{\boldsymbol{Q}}^t = \frac{1}{d}\sum_{i=1}^{d}\mathsf{G}(\boldsymbol{m}_i^t;\boldsymbol{Q}^t)\,. \tag{3.21}$$

Note that, while the free energy naturally depends on the $(\boldsymbol{Q}_i)_{i\le d}$, $(\widetilde{\boldsymbol{Q}}_a)_{a\le n}$, the iteration sets $\boldsymbol{Q}_i^t = \boldsymbol{Q}^t$, $\widetilde{\boldsymbol{Q}}_a^t = \widetilde{\boldsymbol{Q}}^t$, independent of the indices i, a. In fact, any stationary point of

$\mathcal{F}(\boldsymbol{m}, \tilde{\boldsymbol{m}}, \boldsymbol{Q}, \widetilde{\boldsymbol{Q}})$ can be shown to be of this form.

The state of the iteration in Eqs. (3.20), (3.21) is given by the pair $(\boldsymbol{m}^t, \boldsymbol{Q}^t) \in \mathbb{R}^{d\times k} \times \mathbb{R}^{k\times k}$, and $(\tilde{\boldsymbol{m}}^t, \widetilde{\boldsymbol{Q}}^t)$ can be viewed as derived variables. The iteration hence defines a mapping $\mathcal{M}_{\boldsymbol{X}} : \mathbb{R}^{d\times k} \times \mathbb{R}^{k\times k} \to \mathbb{R}^{d\times k} \times \mathbb{R}^{k\times k}$, and we can write it in the form

$$(\boldsymbol{m}^{t+1}, \boldsymbol{Q}^{t+1}) = \mathcal{M}_{\boldsymbol{X}}(\boldsymbol{m}^t, \boldsymbol{Q}^t)\,. \tag{3.22}$$

Any critical point of the free energy (3.9) is a fixed point of the naive mean field iteration and vice-versa, as follows from Appendix D.3. In particular, the uninformative critical point $(\boldsymbol{m}^*, \tilde{\boldsymbol{m}}^*, \boldsymbol{Q}^*, \widetilde{\boldsymbol{Q}}^*)$ is a fixed point of the naive mean field iteration.

3.4 Instability

In view of Section 3.1, for $\beta < \beta_{\rm Bayes}(k, \delta, \nu)$, the real posterior should be centered around a point symmetric under permutations of the topics. In particular, the posterior $\tilde{q}(\boldsymbol{w}_a)$ over the weights of document a should be centered around the symmetric distribution $\boldsymbol{w}_a = (1/k, \dots, 1/k)$. In other words, the uninformative fixed point should be a good approximation of the posterior for $\beta \le \beta_{\rm Bayes}$.

A minimum consistency condition for variational inference is that the uninformative stationary point is a local minimum of the posterior for $\beta < \beta_{\rm Bayes}$. The next theorem provides a necessary condition for stability of the uninformative point, which we expect to be tight. As discussed below, it implies that this point is a saddle in an interval of β below $\beta_{\rm Bayes}$. We recall that the index of a smooth function f at stationary point $\boldsymbol{x}_*$ is the number of the negative eigenvalues of the Hessian $\nabla^2 f(\boldsymbol{x}_*)$.

Theorem 2. *Define q_1^*, q_2^* as in Eqs. (3.14), (3.15), and let*

$$L(\beta, k, \delta, \nu) \equiv \frac{\beta(1+\sqrt{\delta})^2}{1+q_1^*}\left(\frac{q_1^*}{\delta\beta} + k\left[\frac{q_2^*}{1+q_1^*+kq_2^*}\left(\frac{1}{\delta\beta}+\frac{1}{k}\right) - \frac{1}{k^2}\right]_+\right)\,. \tag{3.23}$$

If $L(\beta, k, \delta, \nu) > 1$, then there exists $\varepsilon_1, \varepsilon_2 > 0$ such that the uninformative critical point of Lemma 3.2, $(\boldsymbol{m}^*, \tilde{\boldsymbol{m}}^*, \boldsymbol{Q}^*, \widetilde{\boldsymbol{Q}}^*)$ is, with high probability, a saddle point, with index at least $n\varepsilon_1$ and $\lambda_{\min}(\mathcal{F}|_{\boldsymbol{m}^*, \tilde{\boldsymbol{m}}^*, \boldsymbol{Q}^*, \widetilde{\boldsymbol{Q}}^*}) \le -\varepsilon_2$.

Correspondingly $(\boldsymbol{m}^*, \boldsymbol{Q}^*)$ *is an unstable critical point of the mapping* $\mathcal{M}_{\boldsymbol{X}}$ *in the sense that the Jacobian* $\boldsymbol{D}\mathcal{M}_{\boldsymbol{X}}$ *has spectral radius larger than one at* $(\boldsymbol{m}^*, \boldsymbol{Q}^*)$*.*

In the following, we will say that a fixed point $(\boldsymbol{m}^*, \boldsymbol{Q}^*)$ is stable if the linearization of $\mathcal{M}_{\boldsymbol{X}}(\,\cdot\,)$ at $(\boldsymbol{m}^*, \boldsymbol{Q}^*)$ (i.e. the Jacobian matrix $\boldsymbol{D}\mathcal{M}_{\boldsymbol{X}}(\boldsymbol{m}^*, \boldsymbol{Q}^*)$) has spectral radius smaller than one. By the Hartman-Grobman linearization theorem [41], this implies that $(\boldsymbol{m}^*, \boldsymbol{Q}^*)$ is an attractive fixed point. Namely, there exists a neighborhood $\mathcal{O}$ of $(\boldsymbol{m}^*, \boldsymbol{Q}^*)$ such that, initializing the naive mean field iteration within that neighborhood, results in $(\boldsymbol{m}^t, \boldsymbol{Q}^t) \to (\boldsymbol{m}^*, \boldsymbol{Q}^*)$ as $t \to \infty$. Vice-versa, we say that $(\boldsymbol{m}^*, \boldsymbol{Q}^*)$ is unstable if the Jacobian $\boldsymbol{D}\mathcal{M}_{\boldsymbol{X}}(\boldsymbol{m}^*, \boldsymbol{Q}^*)$ has spectral radius larger than one. In this case, for any neighborhood of $(\boldsymbol{m}^*, \boldsymbol{Q}^*)$, and a generic initialization in that neighborhood, $(\boldsymbol{m}^t, \boldsymbol{Q}^t)$ does not converge to the fixed point.

Motivated by Theorem 2, we define the instability threshold $\beta_{\rm inst} = \beta_{\rm inst}(k, \delta, \nu)$ by

$$\beta_{\rm inst}(k, \delta, \nu) \equiv \inf\Big\{\beta \geq 0 : \quad L(\beta, k, \delta, \nu) > 1\Big\}. \tag{3.24}$$

Let us emphasize that, while we discuss the consequences of the instability at $\beta_{\rm inst}$ on the naive mean field iteration, this is a problem of the variational free energy (3.9) and not of the specific optimization algorithm.

3.5 Numerical results for naive mean field

In order to investigate the impact of the instability described above, we carried out extensive numerical simulations with the variational algorithm (3.20), (3.21). After any number of iterations t, estimates of the factors $\boldsymbol{H}$, $\boldsymbol{W}$ are obtained by computing expectations with respect to the marginals (3.5). This results in

$$\widehat{\boldsymbol{H}}^t = \boldsymbol{r}^t = \frac{1}{\sqrt{\beta}}\mathsf{F}(\boldsymbol{m}^t; \boldsymbol{Q}_t)\,, \qquad \widehat{\boldsymbol{W}}^t = \tilde{\boldsymbol{r}}^t = \frac{1}{\sqrt{\beta}}\widetilde{\mathsf{F}}(\tilde{\boldsymbol{m}}^t; \widetilde{\boldsymbol{Q}}_t)\,. \tag{3.25}$$

Note that $(\widehat{\boldsymbol{H}}^t, \widehat{\boldsymbol{Q}}^t)$ can be used as the state of the naive mean-field iteration instead of $(\boldsymbol{m}^t, \boldsymbol{Q}^t)$.

We select a two-dimensional grid of (δ, β)'s and generate 400 different instances according to the LDA model for each grid point. We report various statistics of the estimates aggregated over the 400 instances. We have performed the simulations for $\nu = 1$ and $k \in \{2, 3\}$. For space considerations, we focus here on the case $\nu = 1$, $k = 2$, and discuss other results in Appendix E. (Simulations for other values of ν also yield similar results.)

We initialize both the naive mean field iteration near the uninformative fixed-point as follows:

$$\widehat{\boldsymbol{H}}^0 = (1-\epsilon)\,\boldsymbol{H}_* + \epsilon \frac{\boldsymbol{G}}{\|\boldsymbol{G}\|_F}\|\boldsymbol{H}_*\|_F, \tag{3.26}$$

$$\boldsymbol{Q}_0 = \boldsymbol{Q}_*\,. \tag{3.27}$$

Here $\boldsymbol{G}$ has entries $(G_{ij})_{i\le d, j\le k} \sim_{iid} \mathsf{N}(0,1)$ and $\epsilon = 0.01$ and $\boldsymbol{H}_* = \mathsf{F}(\boldsymbol{m}_*, \boldsymbol{Q}_*)/\sqrt{\beta}$ is the estimate at the uninformative fixed point. We run a maximum of 300 and a minimum of 40 iterations, and assess convergence at iteration t by evaluating

$$\Delta_t = \min_{\boldsymbol{\Pi}\in\mathsf{S}_k} \left\|\widehat{\boldsymbol{W}}^{t-1}\boldsymbol{\Pi} - \widehat{\boldsymbol{W}}^t\right\|_\infty, \tag{3.28}$$

where the minimum is over the set S_k of $k\times k$ permutation matrices. We declare convergence when $\Delta_t < 0.005$. We denote by $\widehat{\boldsymbol{H}}$, $\widehat{\boldsymbol{W}}$ the estimates obtained at convergence.

Recall the definition $\boldsymbol{P}_\perp = \boldsymbol{I}_k - \boldsymbol{1}_k\boldsymbol{1}_k^{\mathsf{T}}/k$. In order to investigate the instability of Theorem 2, we define the quantities

$$\mathsf{V}(\widehat{\boldsymbol{W}}) \equiv \frac{1}{\sqrt{n}}\|\widehat{\boldsymbol{W}}\boldsymbol{P}_\perp\|_F\,, \qquad \mathsf{V}(\widehat{\boldsymbol{H}}) \equiv \frac{1}{\sqrt{d}}\|\widehat{\boldsymbol{H}}\boldsymbol{P}_\perp\|_F \tag{3.29}$$

In Figure 1 we plot empirical results for the average $\mathsf{V}(\widehat{\boldsymbol{W}})$, $\mathsf{V}(\widehat{\boldsymbol{H}})$ for $k = 2$, $\nu = 1$ and four values of δ. In Figure 2, we plot the empirical probability that variational inference does not converge to the uninformative fixed point or, more precisely, $\widehat{\mathbb{P}}(\mathsf{V}(\widehat{\boldsymbol{W}}) \ge \varepsilon_0)$ with $\varepsilon_0 = 10^{-4}$, evaluated on a grid of (β, δ) values. We also plot the Bayes threshold β_{Bayes} (which we find numerically that it coincides with the spectral threshold β_{spect}) and the instability threshold β_{inst}.

It is clear from Figures 1, 2, that variational inference stops converging to the uninforma-

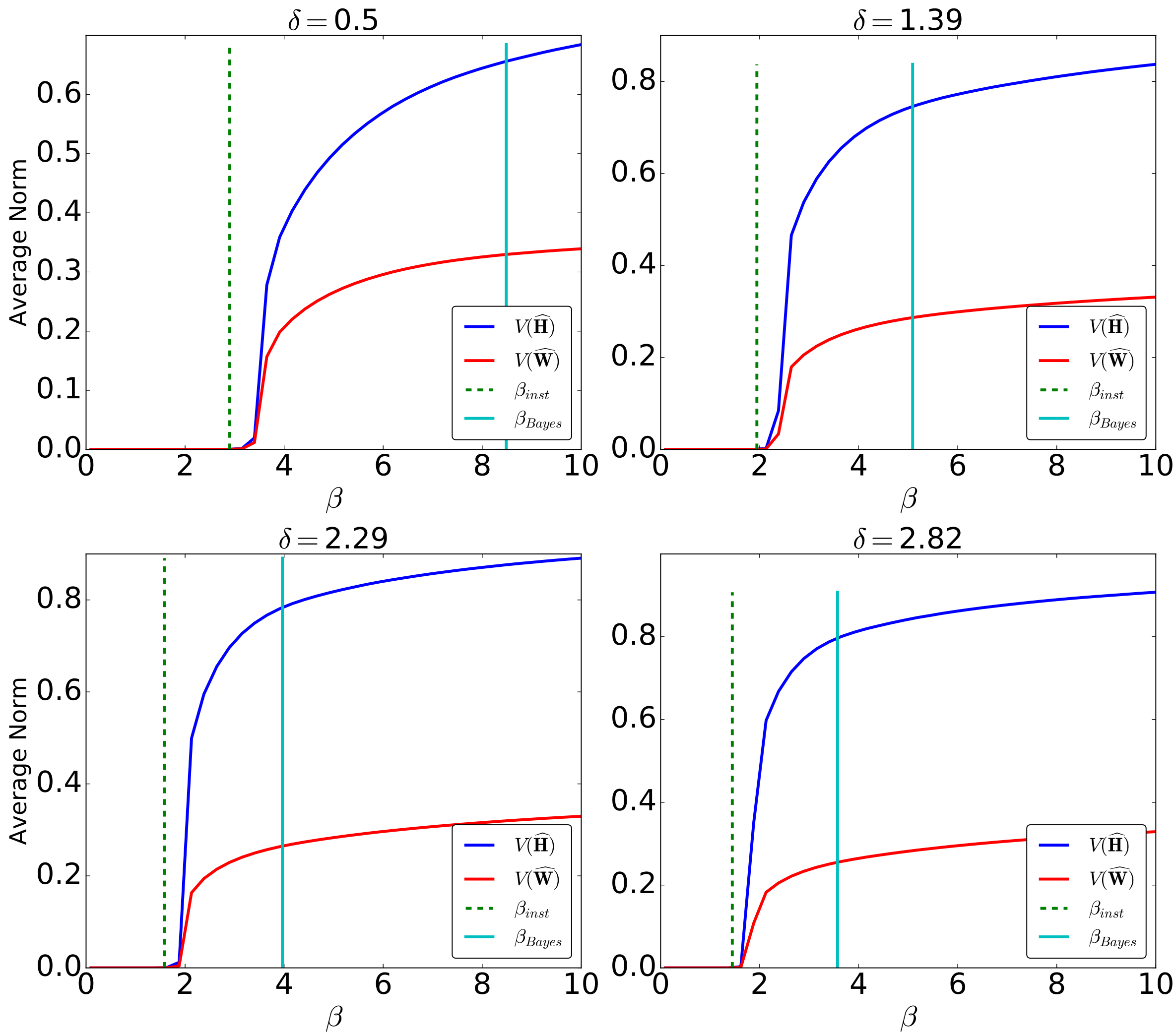

Figure 1: Normalized distances $\mathsf{V}(\widehat{\boldsymbol{H}})$, $\mathsf{V}(\widehat{\boldsymbol{W}})$ of the naive mean field estimates from the uninformative fixed point. Here $k = 2$, $d = 1000$ and $n = d\delta$: each data point corresponds to an average over 400 random realizations.

tive fixed point (although we initialize close to it) when β is still significantly smaller than the Bayes threshold $\beta_{\rm Bayes}$ (i.e. in a regime in which the uninformative fixed point would a reasonable output). The data are consistent with the hypothesis that variational inference becomes unstable at $\beta_{\rm inst}$, as predicted by Theorem 2.

Because of Proposition 3.1, we expect the estimates $\widehat{\boldsymbol{H}}, \widehat{\boldsymbol{W}}$ produced by variational inference to be asymptotically uncorrelated with the true factors for $\beta_{\rm inst} < \beta < \beta_{\rm Bayes}$. In order to test this hypothesis, we borrow a technique that has been developed in the study of phase transitions in statistical physics, and is known as the Binder cumulant [11]. For

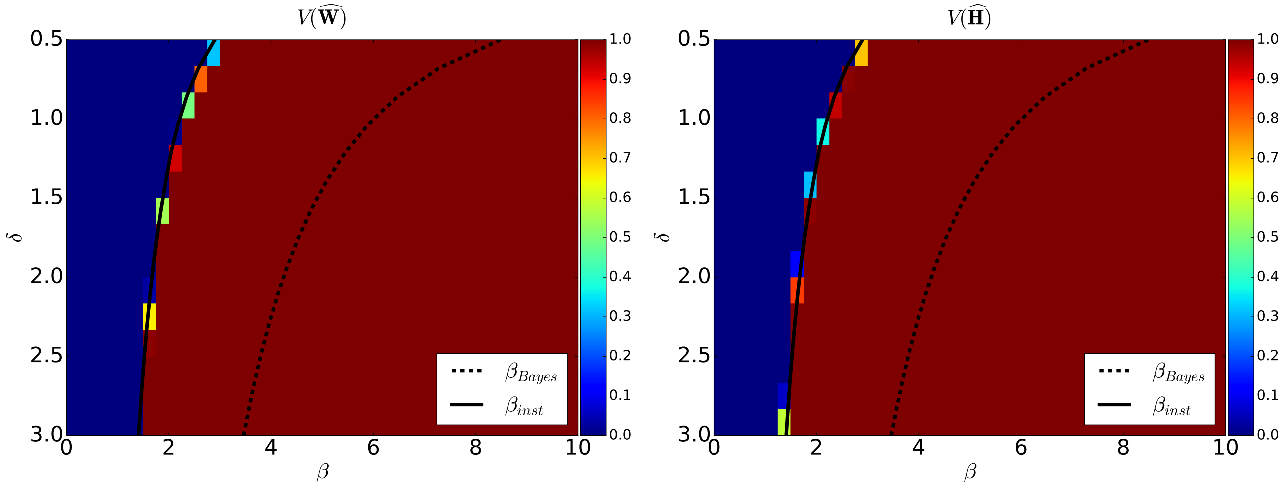

Figure 2: Empirical fraction of instances such that $\mathsf{V}(\widehat{\boldsymbol{W}}) \geq \varepsilon_0 = 10^{-4}$ (left frame) or $\mathsf{V}(\widehat{\boldsymbol{H}}) \geq \varepsilon_0$ (right frame), where $\widehat{\boldsymbol{W}}, \widehat{\boldsymbol{H}}$ are the naive mean field estimate. Here $k = 2$, $d = 1000$ and, for each (δ, β) point on a grid, we used 400 random realizations to estimate the probability of $\mathsf{V}(\widehat{\boldsymbol{W}}) \geq \varepsilon_0$.

the sake of simplicity, we focus here –again– on the case $k = 2$, deferring the general case to Appendix E. Since in this case $\widehat{\boldsymbol{H}}, \boldsymbol{H} \in \mathbb{R}^{d\times 2}$, $\widehat{\boldsymbol{W}}, \boldsymbol{W} \in \mathbb{R}^{n\times 2}$, we can encode the informative component of these matrices by taking the difference between their columns. For instance, we define $\widehat{\boldsymbol{h}}_\perp \equiv \widehat{\boldsymbol{H}}(\boldsymbol{e}_1 - \boldsymbol{e}_2)$, and analogously $\boldsymbol{h}_\perp$, $\widehat{\boldsymbol{w}}_\perp$, $\boldsymbol{w}_\perp$. We then define

$$\mathsf{C}_\eta(\boldsymbol{H}, \widehat{\boldsymbol{H}}) \equiv \langle \widehat{\boldsymbol{h}}_\perp + \eta \boldsymbol{g}, \boldsymbol{h}_\perp \rangle \,, \qquad \mathsf{B}_{\boldsymbol{H}} \equiv \frac{3}{2} - \frac{\widehat{\mathbb{E}}\{\mathsf{C}_\eta(\boldsymbol{H}, \widehat{\boldsymbol{H}})^4\}}{2\widehat{\mathbb{E}}\{\mathsf{C}_\eta(\boldsymbol{H}, \widehat{\boldsymbol{H}})^2\}^2} \,. \tag{3.30}$$

Here $\widehat{\mathbb{E}}$ denotes empirical average with respect to the sample, $\boldsymbol{g} \sim \mathsf{N}(0, \boldsymbol{I}_d)$, and we set $\eta = 10^{-4}$. An analogous definition holds for $\mathsf{C}_\eta(\widehat{\boldsymbol{W}})$, $\mathsf{B}_\eta(\widehat{\boldsymbol{W}})$.

The rationale for definition (3.30) is easy to explain. At small signal-to-noise ratio β, we expect $\widehat{\boldsymbol{h}}_\perp$ to be essentially uncorrelated from $\boldsymbol{h}_\perp$ and hence the correlation $\mathsf{C}_\eta(\boldsymbol{H}, \widehat{\boldsymbol{H}})$ to be roughly normal with mean zero and variance $\sigma^2_{\boldsymbol{H}}$. In particular $\mathbb{E}\{\mathsf{C}_\eta(\boldsymbol{H}, \widehat{\boldsymbol{H}})^4\} \approx 3\mathbb{E}\{\mathsf{C}_\eta(\boldsymbol{H}, \widehat{\boldsymbol{H}})^4\}$ and therefore $\mathsf{B}_{\boldsymbol{H}} \approx 0$. (Note that the term $\eta \boldsymbol{g}$ is added to avoid that empirical correlation vanishes, and hence $\mathsf{B}_{\boldsymbol{H}}$ is not defined.)

In contrast, for large β, we expect $\widehat{\boldsymbol{h}}_\perp$ to be positively correlated with $\boldsymbol{h}_\perp$, and $\mathsf{C}_\eta(\boldsymbol{H}, \widehat{\boldsymbol{H}})$ should concentrate around a non-random positive value. As a consequence, $\mathsf{B}_{\boldsymbol{H}} \approx 1$.

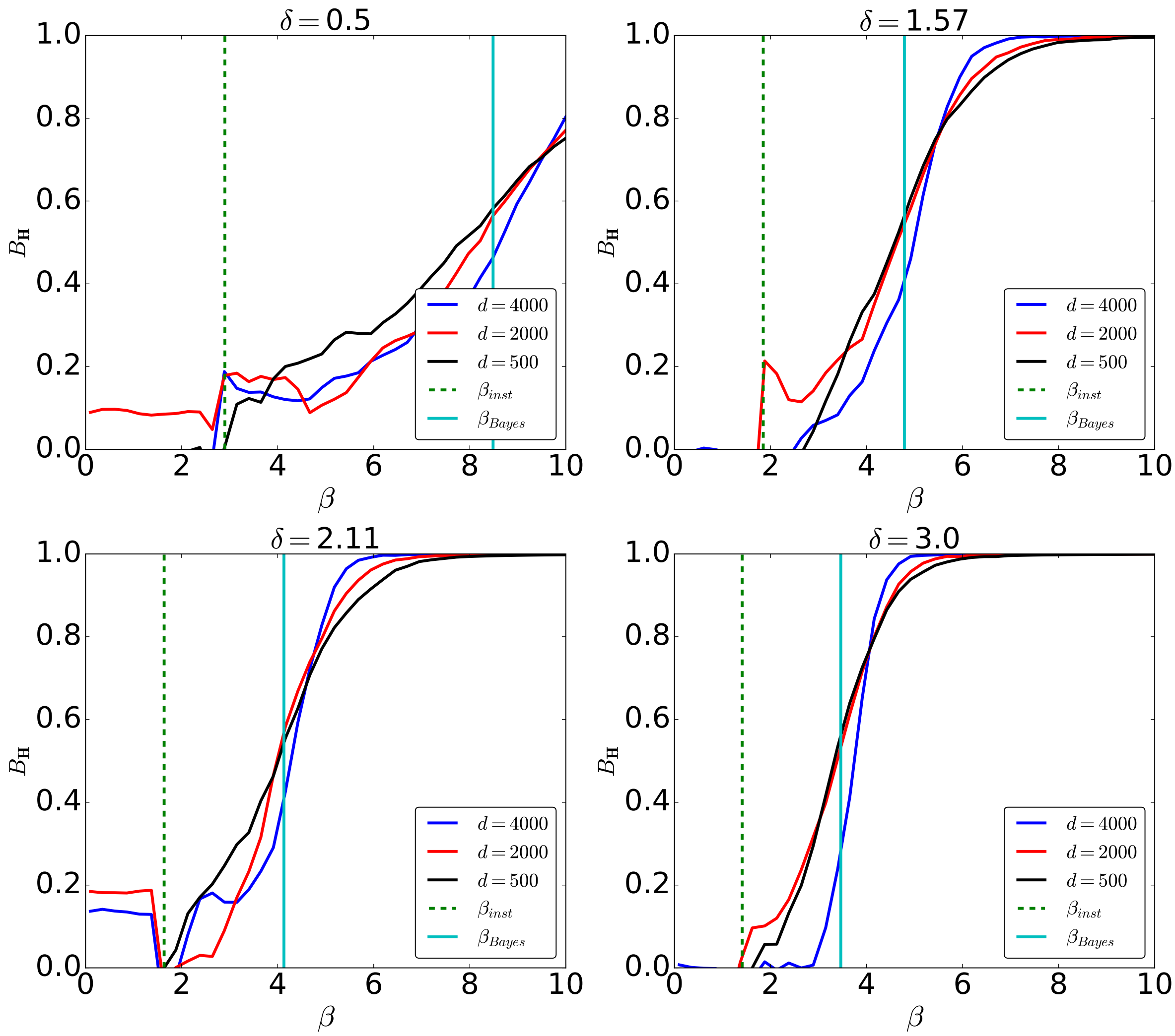

Figure 3: Binder cumulant for the correlation between the naive mean field estimates $\widehat{\boldsymbol{H}}$ and the true topics $\boldsymbol{H}$, see Eq. (3.30). Here we report results for $k = 2$, $d \in \{500, 2000, 4000\}$ and $n = d\delta$, obtained by averaging over 400 realizations. Note that for $\beta < \beta_{\rm Bayes}(k, \nu, \delta)$, $\mathsf{B}_{\boldsymbol{H}}$ decreases with increasing dimensions, suggesting asymptotically vanishing correlations.

In Figures 3 we report our empirical results for $\mathsf{B}_{\boldsymbol{H}}$ and $\mathsf{B}_{\boldsymbol{W}}$ for four different values of δ, and several values of d. As expected, these quantities grow from 0 to 1 as β grows, and the transition is centered around $\beta_{\rm Bayes}$. Figure 4 reports the results on a grid of (β, δ) values. Again, the transition is well predicted by the analytical curve $\beta_{\rm Bayes}$. These data support our claim that, for $\beta_{\rm inst} < \beta < \beta_{\rm Bayes}$, the output of variational inference is non-uniform but uncorrelated with the true signal.

Finally, in Figure 5 we plot the estimates obtained for 100 entries of the weights vector $w_{i,1}$ for three instances with $n = d = 5000$ and $\beta = 2 < \beta_{\rm inst}$, $\beta = 4.1 \in (\beta_{\rm inst}, \beta_{\rm Bayes})$ and

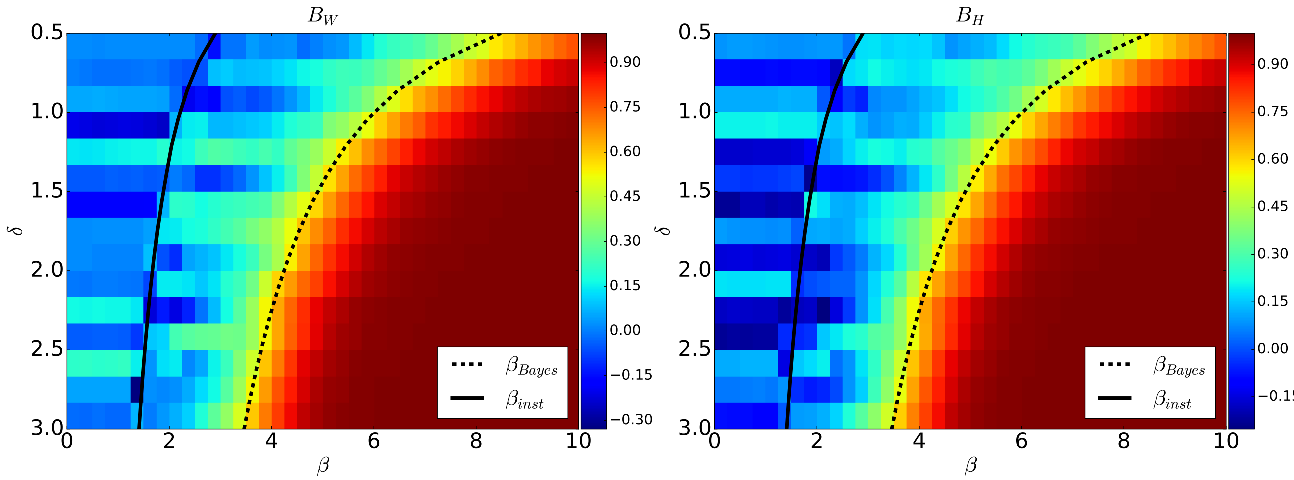

Figure 4: Binder cumulant for the correlation between the naive mean field estimates $\widehat{\boldsymbol{W}}$, $\widehat{\boldsymbol{H}}$ and the true weights and topics $\boldsymbol{W}$, $\boldsymbol{H}$. Here $k = 2$, $d = 1000$ and $n = d\delta$, and we averaged over 400 realizations.

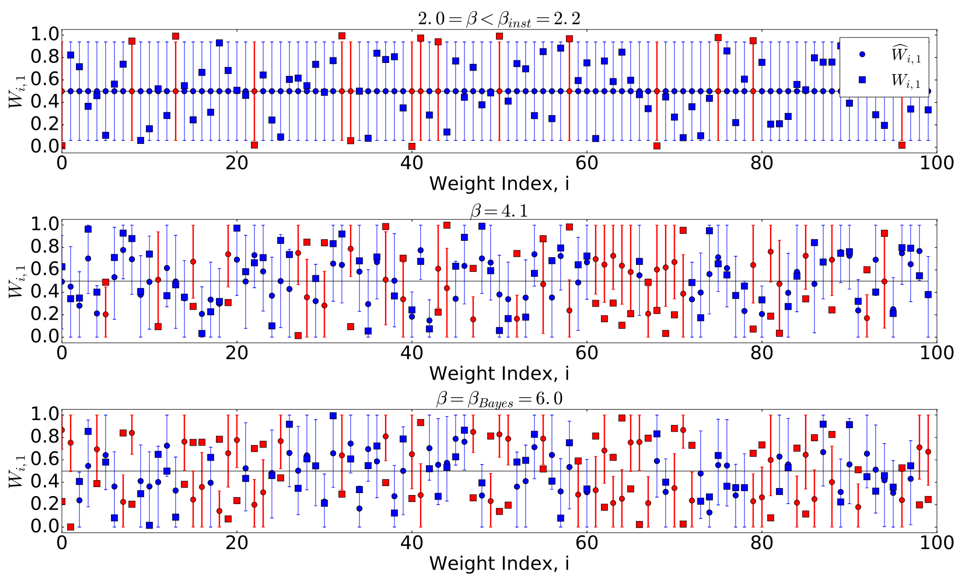

Figure 5: Bayesian credible intervals as computed by variational inference at nominal coverage level $1 - \alpha = 0.9$. Here $k = 2$, $n = d = 5000$, and we consider three values of β: $\beta \in \{2, 4.1, 6\}$ (for reference $\beta_{\rm inst} \approx 2.2, \beta_{\rm Bayes} = 6$). Circles correspond to the posterior mean, and squares to the actual weights. We use red for the coordinates on which the credible interval does not cover the actual value of $w_{i,1}$.

$\beta = 6 = \beta_{\rm Bayes}$. The interval for $w_{a,1}$ is the form $\{w_{a,1} \in [0,1] : \tilde{q}_a(w_{a,1}) \geq t_a(\alpha)\}$ and are constructed to achieve nominal coverage level $1 - \alpha = 0.9$. It is visually clear that the claimed coverage level is not verified in these simulations for $\beta > \beta_{\text{inst}}$, confirming our analytical results. Indeed, for the three simulations in Figure 5 we achieve coverage 0.87 (for $\beta = 2 < \beta_{\text{inst}}$), 0.65 (for $\beta = 4.1 \in (\beta_{\text{inst}}, \beta_{\text{Bayes}})$), and 0.51 (for $\beta = 6 = \beta_{\text{Bayes}}$). Further results of this type are reported in Appendix E.

4 Fixing the instability

The fact that naive mean field is not accurate for certain classes of random high-dimensional probability distributions is well understood within statistical physics. In particular, in the context of mean field spin glasses [39], naive mean field is known to lead to an asymptotically incorrect expression for the free energy. We expect the same mechanism to be relevant in the context of topic models.

Namely, the product-form expression (1.5) only holds asymptotically in the sense of finite-dimensional marginals. However, when computing the term $\mathbb{E}_{\hat{q}} \log p_{\boldsymbol{X}|\boldsymbol{W},\boldsymbol{H}}(\boldsymbol{X}|\boldsymbol{H},\boldsymbol{W})$ in the KL divergence (1.4), the error due to the product form approximation is non-negligible. Keeping track of this error leads to the so-called TAP free energy.

4.1 Revisiting $\mathbb{Z}_2$-synchronization

It is instructive to briefly discuss the $\mathbb{Z}_2$-synchronization example of Section 2, as the basic concepts can be explained more easily in this example. For this problem, the TAP approximation replaces the free energy (2.4) with

$$\mathcal{F}_{\text{TAP}}(\boldsymbol{m}) \equiv -\frac{\lambda}{2}\langle \boldsymbol{m}, \boldsymbol{X}_0 \boldsymbol{m}\rangle - \sum_{i=1}^{n} \mathsf{h}(m_i) - \frac{n\lambda^2}{4}\Big(1 - Q(\boldsymbol{m})\Big)^2 , \tag{4.1}$$

where $Q(\boldsymbol{m}) \equiv \|\boldsymbol{m}\|_2^2/n$.

We can now repeat the analysis of Section 2 with this new free energy approximation. It is easy to see that $\boldsymbol{m}_* = \boldsymbol{0}$ is again a stationary point. However, the Hessian is now

$$\nabla^2 \mathcal{F}(\boldsymbol{m})\Big|_{\boldsymbol{m}=\boldsymbol{m}_*} = -\lambda \boldsymbol{X}_0 + \Big(1 + \lambda^2\Big)\, \boldsymbol{I}\,. \tag{4.2}$$

In particular, for $\lambda < 1$, $\lambda_{\min}(\nabla^2 \mathcal{F}|_{\boldsymbol{m}=\boldsymbol{m}_*})$ converges to $(1-\lambda)^2 > 0$: the uninformative stationary point is (with high probability) a local minimum.

The stationarity condition for the TAP free energy are known as TAP equations, and the algorithm that corresponds to the naive mean field iteration is Bayesian approximate message passing (AMP). For the $\mathbb{Z}_2$ synchronization problem, Bayes AMP is known to achieve the Bayes optimal estimation error [20, 37].

4.2 TAP free energy for topic models

We now turn to topic models. The TAP approach replaces the free energy (3.9) with the following (see Appendix F.1 for a derivation)

$$\begin{aligned} \mathcal{F}_{\rm TAP}(\boldsymbol{r}, \tilde{\boldsymbol{r}}) = & \sum_{i=1}^{d} \psi\left(\boldsymbol{r}_i, \frac{\beta}{d}\sum_{a=1}^{n} \tilde{\boldsymbol{r}}_a^{\otimes 2}\right) + \sum_{a=1}^{n} \tilde{\psi}\left(\tilde{\boldsymbol{r}}_a, \frac{\beta}{d}\sum_{i=1}^{d} \boldsymbol{r}_i^{\otimes 2}\right) \\ & - \sqrt{\beta}\,\mathsf{Tr}\left(\boldsymbol{X}\boldsymbol{r}\tilde{\boldsymbol{r}}^{\mathsf{T}}\right) - \frac{\beta}{2d}\sum_{i=1}^{d}\sum_{a=1}^{n} \langle \boldsymbol{r}_i, \tilde{\boldsymbol{r}}_a\rangle^2 \,, \end{aligned} \tag{4.3}$$

where $\tilde{\boldsymbol{r}}\boldsymbol{1}_k = \boldsymbol{1}_n$, and we defined the partial Legendre transforms

$$\psi(\boldsymbol{r}, \boldsymbol{Q}) \equiv \sup_{\boldsymbol{m}} \left\{\langle \boldsymbol{r}, \boldsymbol{m}\rangle - \phi(\boldsymbol{m}, \boldsymbol{Q})\right\}\,, \quad \tilde{\psi}(\tilde{\boldsymbol{r}}, \widetilde{\boldsymbol{Q}}) \equiv \sup_{\tilde{\boldsymbol{m}}} \left\{\langle \tilde{\boldsymbol{r}}, \tilde{\boldsymbol{m}}\rangle - \tilde{\phi}(\tilde{\boldsymbol{m}}, \widetilde{\boldsymbol{Q}})\right\}\,. \tag{4.4}$$

Notice that $\tilde{\psi}(\tilde{\boldsymbol{r}}, \widetilde{\boldsymbol{Q}})$ is finite only if $\langle \boldsymbol{1}_k, \tilde{\boldsymbol{r}}\rangle = 1$.

When substituting in Eq. (4.3), the supremum of Eq. (4.4) is achieved at

$$\boldsymbol{r} = \frac{1}{\sqrt{\beta}}\mathsf{F}(\boldsymbol{m}; \boldsymbol{Q})\,, \quad \tilde{\boldsymbol{r}} = \frac{1}{\sqrt{\beta}}\widetilde{\mathsf{F}}(\tilde{\boldsymbol{m}}; \widetilde{\boldsymbol{Q}})\,, \tag{4.5}$$

$$\boldsymbol{Q} = \frac{\beta}{d}\sum_{a=1}^{n} \tilde{\boldsymbol{r}}_a^{\otimes 2}, \qquad \widetilde{\boldsymbol{Q}} = \frac{\beta}{d}\sum_{i=1}^{d} \boldsymbol{r}_i^{\otimes 2}\,. \tag{4.6}$$

Calculus shows that stationary points of this free energy are in one-to-one correspondence (via Eq. (4.5)) with the fixed points of the following iteration:

$$\boldsymbol{m}^{t+1} = \quad \boldsymbol{X}^{\mathsf{T}}\, \widetilde{\mathsf{F}}(\tilde{\boldsymbol{m}}^t; \widetilde{\boldsymbol{Q}}^t) - \mathsf{F}(\boldsymbol{m}^t; \boldsymbol{Q}^t)\tilde{\boldsymbol{\Omega}}_t\,, \tag{4.7}$$

$$\tilde{\boldsymbol{m}}^t = \quad \boldsymbol{X}\, \mathsf{F}(\boldsymbol{m}^t; \boldsymbol{Q}^t) \quad - \widetilde{\mathsf{F}}(\tilde{\boldsymbol{m}}^{t-1}; \widetilde{\boldsymbol{Q}}^{t-1})\boldsymbol{\Omega}_t\,, \tag{4.8}$$

$$\boldsymbol{Q}^{t+1} = \frac{1}{d}\sum_{a=1}^{n} \widetilde{\mathsf{F}}(\tilde{\boldsymbol{m}}_a^t; \widetilde{\boldsymbol{Q}}^t\,)^{\otimes 2} \tag{4.9}$$

$$\widetilde{\boldsymbol{Q}}^t = \frac{1}{d}\sum_{i=1}^{d} \mathsf{F}(\boldsymbol{m}_i^t; \boldsymbol{Q}^t\,\,)^{\otimes 2}. \tag{4.10}$$

where $\boldsymbol{\Omega}_t$, $\tilde{\boldsymbol{\Omega}}_t$ are defined as

$$\boldsymbol{\Omega}_t = \frac{1}{d\sqrt{\beta}}\sum_{i=1}^{d}[\mathsf{G}(\boldsymbol{m}_i^t, \boldsymbol{Q}^t) - \mathsf{F}(\boldsymbol{m}_i^t; \boldsymbol{Q}^t)^{\otimes 2}] = \frac{1}{d}\sum_{i=1}^{d}\frac{\partial \mathsf{F}}{\partial \boldsymbol{m}_i}(\boldsymbol{m}_i^t; \boldsymbol{Q}^t)\,, \tag{4.11}$$

$$\tilde{\boldsymbol{\Omega}}_t = \frac{1}{d\sqrt{\beta}}\sum_{a=1}^{n}[\widetilde{\mathsf{G}}(\tilde{\boldsymbol{m}}_a^t, \widetilde{\boldsymbol{Q}}) - \widetilde{\mathsf{F}}(\tilde{\boldsymbol{m}}_a^t; \widetilde{\boldsymbol{Q}}^t)^{\otimes 2}] = \frac{1}{d}\sum_{a=1}^{n}\frac{\partial \widetilde{\mathsf{F}}}{\partial \tilde{\boldsymbol{m}}_a}(\tilde{\boldsymbol{m}}_a^t; \widetilde{\boldsymbol{Q}}^t)\,. \tag{4.12}$$

The stationarity conditions for the TAP free energy (4.3) are known as TAP equations, and the corresponding iterative algorithm (4.7), (4.8) is a special case of approximate message passing (AMP), with Bayesian updates. Note that the specific choice of time indices in Eqs. (4.7), (4.8) is instrumental for the analysis in the next section to hold. We also note that the general AMP analysis of [7, 29] allows for quite general choices of the sequence of matrices $\boldsymbol{Q}_t, \widetilde{\boldsymbol{Q}}_t$. However, stationarity of the TAP free energy (4.3) requires that at convergence the condition (4.10) holds at the fixed point

Estimates of the factors $\boldsymbol{W}$, $\boldsymbol{H}$ are computed following the same recipe as for naive mean field, cf. Eq. (3.25), namely $\widehat{\boldsymbol{H}}^t = \boldsymbol{r}^t = \mathsf{F}(\boldsymbol{m}^t; \boldsymbol{Q}_t)/\sqrt{\beta}$, $\widehat{\boldsymbol{W}}^t = \tilde{\boldsymbol{r}}^t = \widetilde{\mathsf{F}}(\tilde{\boldsymbol{m}}^t; \widetilde{\boldsymbol{Q}}_t)/\sqrt{\beta}$.

It is not hard to see that the AMP iteration admits an uninformative fixed point, which is a stationary point of the TAP free energy, see proof in Appendix F.3.

Lemma 4.1. *Define* $q_0^* = \beta\delta/k^2$ *and* $\tilde{q}_0^* = \beta^2\|\boldsymbol{X}^{\mathsf{T}}\mathbf{1}_n\|_2^2/(dk^2(1+kq_0)^2)$. *Then, AMP iteration* *admits the following fixed point*

$$\boldsymbol{m}^* = \frac{\sqrt{\beta}}{k}(\boldsymbol{X}^{\mathsf{T}}\mathbf{1}_n) \otimes \mathbf{1}_k\,, \tag{4.13}$$

$$\tilde{\boldsymbol{m}}^* = \frac{\beta}{k(1+kq_0)}(\boldsymbol{X}\boldsymbol{X}^{\mathsf{T}}\mathbf{1}_n)\otimes\mathbf{1}_k \;-\; \frac{\beta}{k+\delta\beta}\,\mathbf{1}_n \otimes \mathbf{1}_k\,, \tag{4.14}$$

$$\boldsymbol{Q}^* = q_0^*\,\boldsymbol{J}_k\,, \qquad\qquad \widetilde{\boldsymbol{Q}}^* = \tilde{q}_0^*\,\boldsymbol{J}_k\,. \tag{4.15}$$

This corresponds to a stationary point of the TAP free energy (4.3), via Eq. (4.5):

$$\boldsymbol{r}_* = \frac{\sqrt{\beta}}{k(1+kq_0^*)}(\boldsymbol{X}^{\mathsf{T}}\mathbf{1}_n) \otimes \mathbf{1}_k\,, \qquad \tilde{\boldsymbol{r}}_* = \frac{1}{k}\mathbf{1}_n \otimes \mathbf{1}_k\,. \tag{4.16}$$

Further, this is the only stationary point that is unchanged under permutations of the topics.

4.3 State evolution analysis

State evolution is a recursion over matrices $\boldsymbol{M}_t$, $\widetilde{\boldsymbol{M}}_t \in \mathbb{R}^{k\times k}$, defined by

$$\boldsymbol{M}_{t+1} = \delta\,\mathbb{E}\Big\{\widetilde{\mathsf{F}}(\widetilde{\boldsymbol{M}}_t\boldsymbol{w} + \widetilde{\boldsymbol{M}}_t^{1/2}\boldsymbol{z}; \widetilde{\boldsymbol{M}}_t)^{\otimes 2}\Big\}\,, \tag{4.17}$$

$$\widetilde{\boldsymbol{M}}_t = \mathbb{E}\Big\{\mathsf{F}(\boldsymbol{M}_t\boldsymbol{h} + \boldsymbol{M}_t^{1/2}\boldsymbol{z}; \boldsymbol{M}_t)^{\otimes 2}\Big\}\,, \tag{4.18}$$

where expectation is with respect to $\boldsymbol{h} \sim q_0(\,\cdot\,)$, $\boldsymbol{w} \sim \tilde{q}_0(\,\cdot\,)$ and $\boldsymbol{z} \sim \mathsf{N}(0, \boldsymbol{I}_k)$ independent. Note that $\boldsymbol{M}_t, \widetilde{\boldsymbol{M}}_t$ are positive semidefinite symmetric matrices. Also, Eq. (4.18) can be written explicitly as

$$\widetilde{\boldsymbol{M}}_t = \beta(\boldsymbol{I}_k + \boldsymbol{M}_t)^{-1}\boldsymbol{M}_t\,. \tag{4.19}$$

State evolution provides an asymptotically exact characterization of the behavior of AMP, as formalized by the next theorem (which is a direct application of [29]).

Theorem 3. *Consider the AMP algorithm of Eqs. (4.7), with deterministic initialization* $\boldsymbol{m}^0, \boldsymbol{Q}^0$. *Assume* $\boldsymbol{G} \in \mathbb{R}^{d\times k}$ *to be independent of data* $\boldsymbol{X}$, *with entries* $(G_{ij})_{i\le d, j\le k} \sim_{iid} \mathsf{N}(0,1)$, *and let* $\boldsymbol{m}^0 = \boldsymbol{H}\boldsymbol{M}_0 + \boldsymbol{Z}\boldsymbol{M}_0^{1/2}$ *for* $\boldsymbol{M}_0 \in \mathbb{R}^{k\times k}$ *non-random,* $\boldsymbol{M}_0 \succeq 0$. *Let* $\{\boldsymbol{M}_t, \widetilde{\boldsymbol{M}}_t\}_{t\ge 1}$ *be defined by the state evolution recursion (4.17), (4.18). Then, for any* *pseudo-Lipschitz function* $g : \mathbb{R}^k \times \mathbb{R}^k \to \mathbb{R}$*, we have, almost surely,*

$$\lim_{n\to\infty} \frac{1}{d} \sum_{i=1}^{d} g(\boldsymbol{h}_i, \boldsymbol{m}_i^t) = \mathbb{E}\Big\{ g(\boldsymbol{h}, \boldsymbol{M}_t \boldsymbol{h} + \boldsymbol{M}_t^{1/2} \boldsymbol{z}) \Big\}, \tag{4.20}$$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{a=1}^{n} g(\boldsymbol{w}_a, \tilde{\boldsymbol{m}}_a^t) = \mathbb{E}\Big\{ g(\boldsymbol{w}, \widetilde{\boldsymbol{M}}_t \boldsymbol{w} + \widetilde{\boldsymbol{M}}_t^{1/2} \boldsymbol{z}) \Big\}, \tag{4.21}$$

where it is understood that $n, d \to \infty$ *with* $n/d \to \delta$*. In particular*

$$\lim_{n\to\infty} \frac{1}{d} \boldsymbol{H}^{\mathsf{T}} \widehat{\boldsymbol{H}}^t = \frac{1}{\sqrt{\beta}} \widetilde{\boldsymbol{M}}_t, \tag{4.22}$$

$$\lim_{n\to\infty} \frac{1}{n} \boldsymbol{W}^{\mathsf{T}} \widehat{\boldsymbol{W}}^t = \frac{1}{\sqrt{\beta}} \boldsymbol{M}_{t+1}. \tag{4.23}$$

Further $\lim_{n\to\infty} \boldsymbol{Q}^t = \boldsymbol{M}_t$*,* $\lim_{n\to\infty} \widetilde{\boldsymbol{Q}}^t = \widetilde{\boldsymbol{M}}_t$*.*

Using state evolution, we can establish a stability result for AMP. First of all, notice that the state evolution iteration (4.17), (4.18) admits a fixed point of the form $\boldsymbol{M}^* = (\delta\beta/k^2)\boldsymbol{J}_k$, $\widetilde{\boldsymbol{M}}^* = \rho_0 \boldsymbol{J}_k$, for $\rho_0 = \delta\beta^2/(k\delta\beta + k^2)$, see Appendix G.2. This is an uninformative fixed point, in the sense that the k topics are asymptotically identical. The next theorem is proved in Appendix G.3.

Theorem 4. *If* $\beta < \beta_{\text{spect}}(k, \nu, \delta)$*, then the uninformative fixed point is stable under the state evolution iteration (4.17), (4.18).*

In particular, for $\beta < \beta_{\text{spect}}(k, \nu, \delta)$*, there exists* $c_0 = c_0(\beta, k\nu, \delta)$ *such that, if we initialize AMP as in Theorem 3 with* $\|\boldsymbol{M}_0 - \boldsymbol{M}^*\|_F \le c_0$*, then (recalling* $\boldsymbol{P}_\perp = \boldsymbol{I}_k - \boldsymbol{1}_k \boldsymbol{1}_k / k$*)*

$$\lim_{t\to\infty} \lim_{n\to\infty} \frac{1}{n} \big\| \boldsymbol{m}^t \boldsymbol{P}_\perp \|_F^2 = 0, \qquad \lim_{t\to\infty} \lim_{n\to\infty} \frac{1}{n} \big\| \boldsymbol{m}^t \boldsymbol{P}_\perp \|_F^2 = 0. \tag{4.24}$$

4.4 Stability of the uninformative fixed point

The next theorem establishes that the uninformative fixed point of the TAP free energy is a local minimum for all β below the spectral threshold $\beta_{\text{spect}}(k, \nu, \delta)$. Since $\beta_{\text{Bayes}}(k, \nu, \delta) \le \beta_{\text{spect}}(k, \nu, \delta)$, this shows that the instability we discovered in the case of naive mean field is corrected by the TAP free energy.

Theorem 5. *Let $(\boldsymbol{r}_*, \tilde{\boldsymbol{r}}_*)$ be the uninformative stationary point of the TAP free energy, cf. Lemma 4.1. If $\beta < \beta_{\rm spect}(k, \nu, \delta)$, then there exists $\varepsilon > 0$ such that, with high probability*

$$\lambda_{\min}\left(\nabla^2 \mathcal{F}_{\rm TAP}\Big|_{(\boldsymbol{r}_*, \tilde{\boldsymbol{r}}_*)}\right) \geq \varepsilon \,. \tag{4.25}$$

Remark 4.1. Let us emphasize that this result is not implied by the state evolution result of Theorem 4, which only establishes stability in a certain asymptotic sense. Vice-versa, Theorem 5 does not directly imply Theorem 4.

4.5 Numerical results for TAP free energy

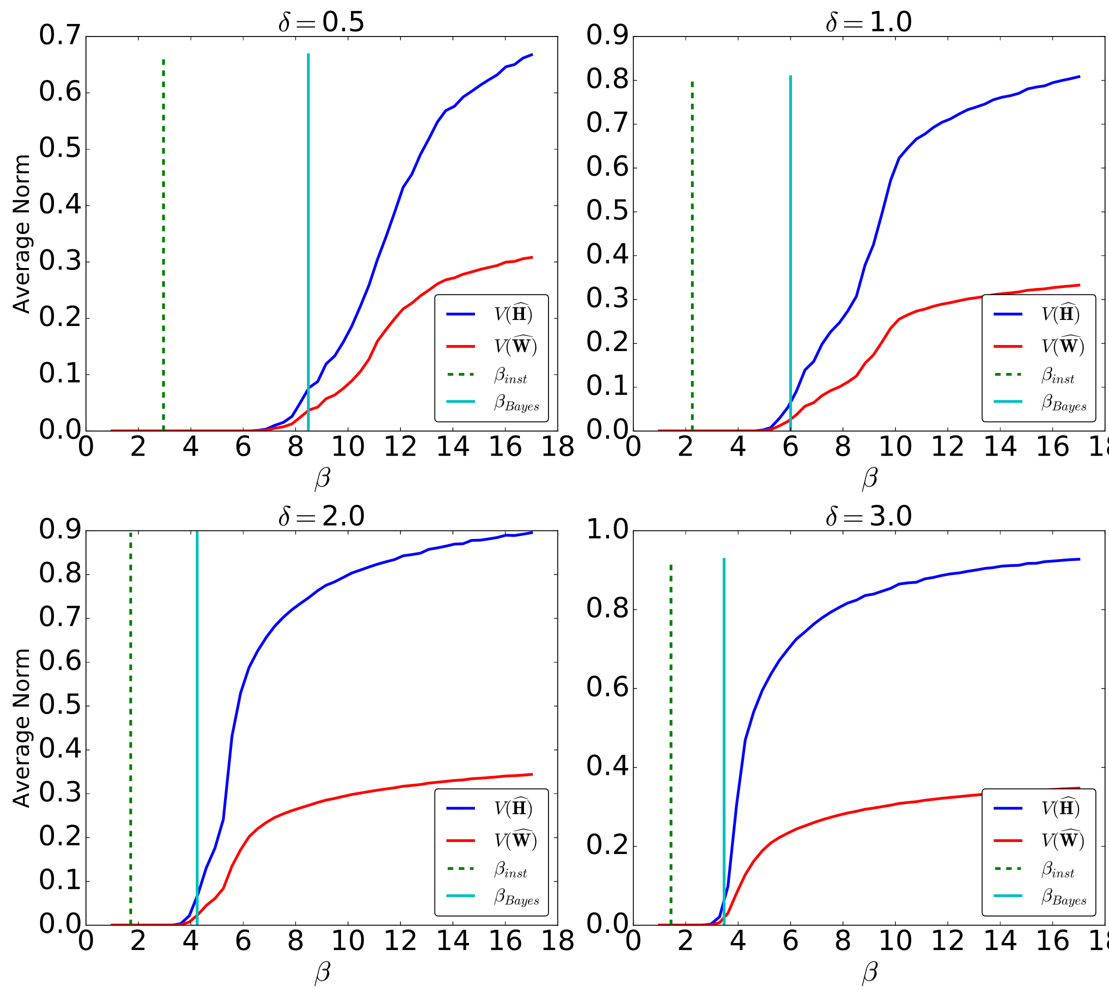

Figure 6: Normalized distances $\mathsf{V}(\widehat{\boldsymbol{H}})$, $\mathsf{V}(\widehat{\boldsymbol{W}})$ of the AMP estimates from the uninformative fixed point. Here, $k = 2$, $d = 1000$ and $n = d\delta$: each data point corresponds to an average over 400 random realizations.

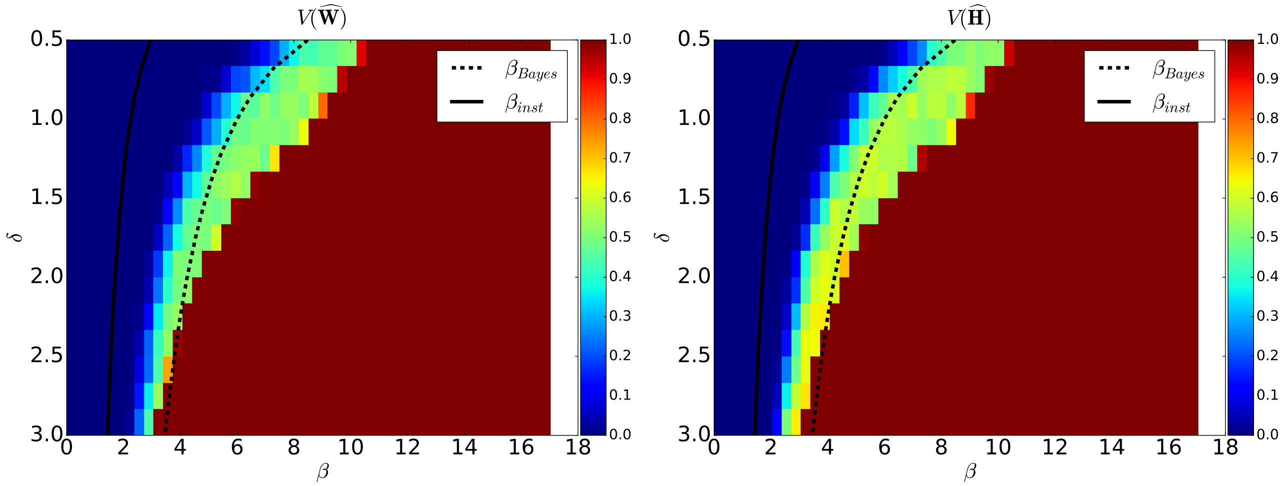

Figure 7: Empirical fraction of instances such that $\mathsf{V}(\widehat{\boldsymbol{W}}) \geq \varepsilon_0 = 5 \cdot 10^{-3}$, where $\widehat{\boldsymbol{W}}$ is the AMP estimate. Here $k = 2$, $d = 1000$, and for each (δ, β) point on the grid we ran AMP on 400 random realizations.

In order to confirm the stability analysis at the previous section, we carried out numerical simulations analogous to the ones of Section 3.5. We found that the AMP iteration of Eqs. (4.7), (4.8) is somewhat unstable when $\beta \approx \beta_{\text{spect}}$. In order to remedy this problem, we used a damped version of the same iteration, see Appendix H.1. Notice that damping does not change the stability of a local minimum or saddle, it merely reduces oscillations due to aggressive step sizes.

We initialize the iteration as for naive mean field, and monitor the same quantities, as in Section 3.5. In particular, here we report results on the distance from the uninformative subspace $\mathsf{V}(\widehat{\boldsymbol{H}})$, $\mathsf{V}(\widehat{\boldsymbol{W}})$, in Figures 6 and 7, and the Binder cumulants $\mathsf{B}_{\boldsymbol{H}}$ and $\mathsf{B}_{\boldsymbol{W}}$, measuring the correlation between AMP estimates and the true factors $\boldsymbol{W}$, $\boldsymbol{H}$, in Figures 8, 9. We focus on the case $k = 2$, deferring $k = 3$ to the appendices.

In the intermediate regime $\beta \in (\beta_{\text{inst}}, \beta_{\text{spect}})$, the behavior of AMP is strikingly different from the one of naive mean field. AMP remains close to the uninformative fixed point, confirming that this is a local minimum of the TAP free energy. The distance from the uninformative subspace starts growing only at the spectral threshold β_{spect} (which coincides, in the present cases, with the Bayes threshold β_{Bayes}). At the same point, the correlation with the true factors $\boldsymbol{W}$, $\boldsymbol{H}$ also becomes strictly positive.

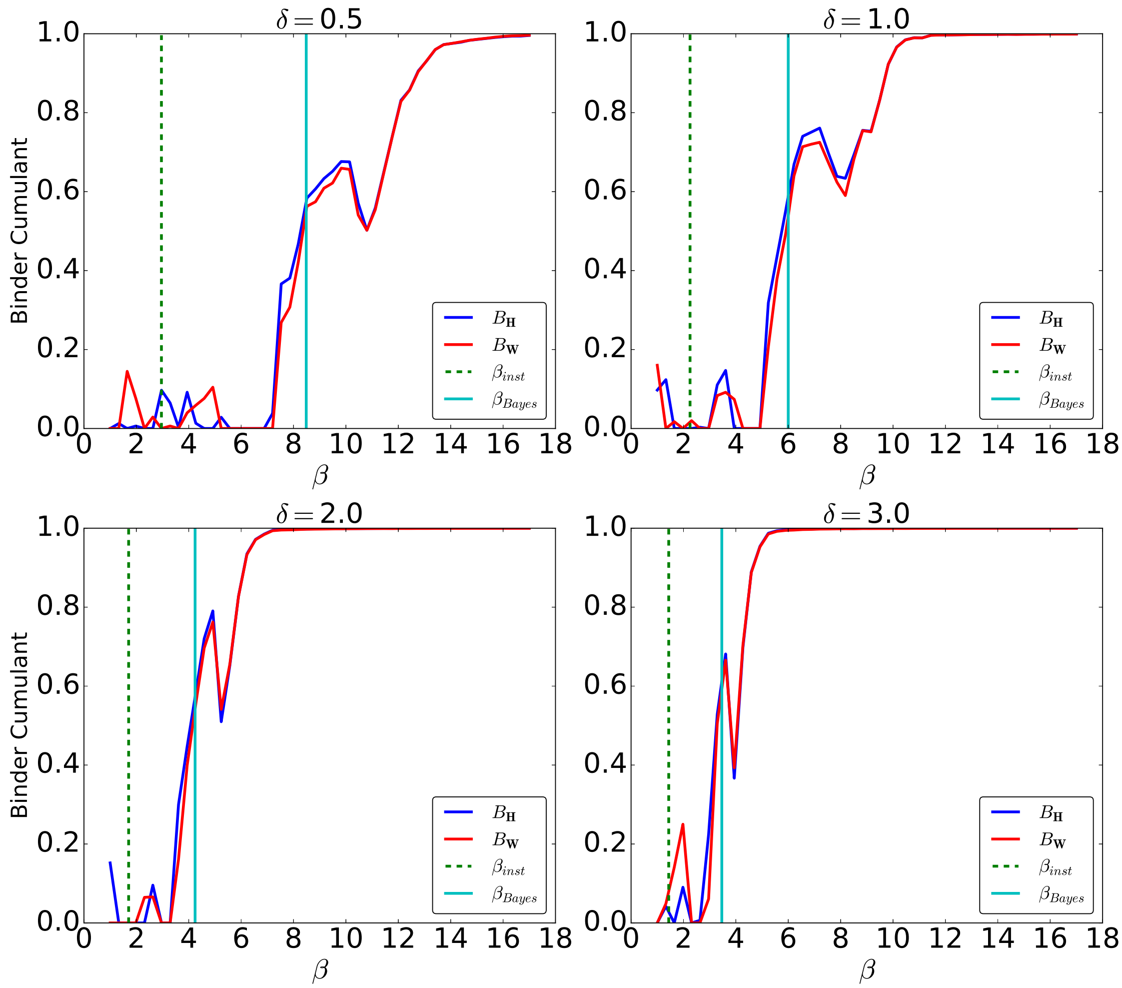

Figure 8: Binder cumulant for the correlation between AMP estimates $\widehat{\boldsymbol{H}}$ and the true topics $\boldsymbol{H}$, and between $\widehat{\boldsymbol{W}}$ and $\boldsymbol{W}$, see Eq. (3.30). Here $k = 2$, $d = 1000$, $n = d\delta$ and estimates are obtained by averaging over 400 realizations.

5 Discussion

Bayesian methods are particularly attractive in unsupervised learning problems such as topic modeling. Faced with a collection of documents $\boldsymbol{x}_1, \dots \boldsymbol{x}_n$, it is not clear a priori whether they should be modeled as convex combinations of topics, or how many topics should be used. Even after a low-rank factorization $\boldsymbol{X} \approx \boldsymbol{W}\boldsymbol{H}^{\mathsf{T}}$ is computed, it is still unclear how to evaluate it, or to which extent it should be trusted.

Bayesian approaches provide estimates of the factors $\boldsymbol{W}$, $\boldsymbol{H}$, but also a probabilistic measure of how much these estimates should be trusted. To the extent that the posterior concentrates around its mean, this can be considered as a good estimate of a true underlying

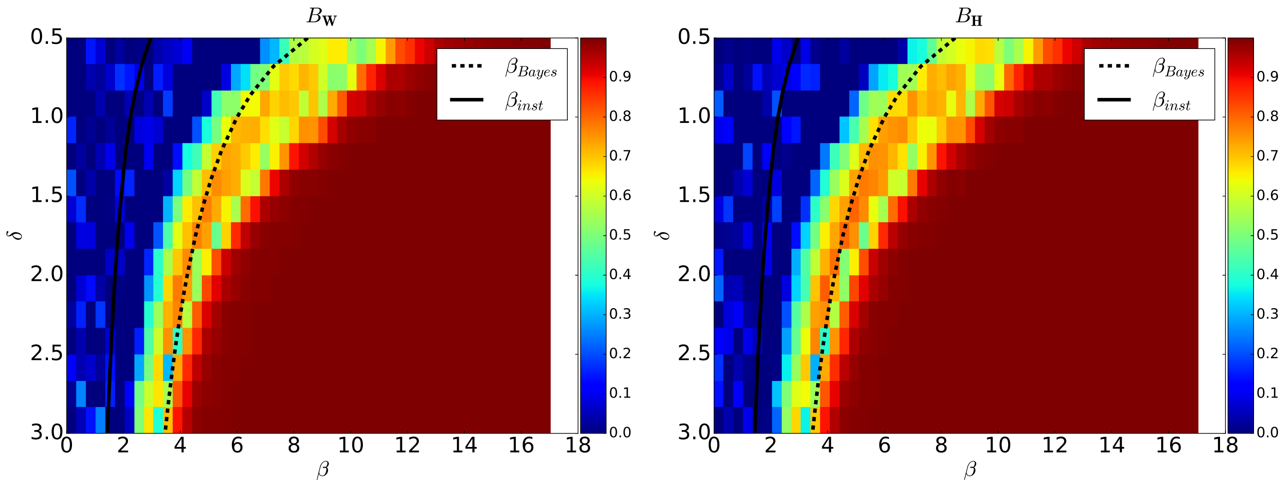

Figure 9: Binder cumulant for the correlation between AMP estimates $\widehat{\boldsymbol{W}}$, $\widehat{\boldsymbol{H}}$ and the true weights and topics $\boldsymbol{W}$, $\boldsymbol{H}$. Here $k = 2$, $d = 1000$ and estimates are obtained by averaging over 400 realizations.

signal.

It is well understood that Bayesian estimates can be unreliable if the prior is not chosen carefully. Our work points at a second reason for caution. When variational inference is used for approximating the posterior, the result can be incorrect even if the data are generated according to the prior. More precisely, we showed that for a certain regime of parameters, naive mean field 'believes' that there is a signal, even if it is information-theoretically impossible to extract any non-trivial estimate from the data.

Given that naive mean field is the method of choice for inference with topic models [16], it would be of great interest to remedy this instability. We showed that the TAP free energy provides a better mean field approximation, and in particular does not have the same instability. However, this approximation is also based on the correctness of the generative model, and further investigation is warranted on its robustness.

References

[1] Edoardo M Airoldi, David M Blei, Stephen E Fienberg, and Eric P Xing. Mixed membership stochastic blockmodels. *Journal of Machine Learning Research*, 9(Sep):1981–2014,

2008.

[2] Ahmed El Alaoui and Florent Krzakala. Estimation in the spiked wigner model: A short proof of the replica formula. *arXiv preprint arXiv:1801.01593*, 2018.

[3] Sanjeev Arora, Rong Ge, Ravindran Kannan, and Ankur Moitra. Computing a nonnegative matrix factorization–provably. In *Proceedings of the forty-fourth annual ACM symposium on Theory of computing*, pages 145–162. ACM, 2012.

[4] Sanjeev Arora, Rong Ge, and Ankur Moitra. Learning topic models–going beyond svd. In *Foundations of Computer Science (FOCS), 2012 IEEE 53rd Annual Symposium on*, pages 1–10. IEEE, 2012.

[5] Jinho Baik, Gérard Ben Arous, Sandrine Péché, et al. Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. *The Annals of Probability*, 33(5):1643–1697, 2005.

[6] Jean Barbier, Mohamad Dia, Nicolas Macris, Florent Krzakala, Thibault Lesieur, and Lenka Zdeborová. Mutual information for symmetric rank-one matrix estimation: A proof of the replica formula. In *Advances in Neural Information Processing Systems*, pages 424–432, 2016.

[7] Mohsen Bayati and Andrea Montanari. The dynamics of message passing on dense graphs, with applications to compressed sensing. *IEEE Transactions on Information Theory*, 57(2):764–785, 2011.

[8] Florent Benaych-Georges and Raj Rao Nadakuditi. The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices. *Advances in Mathematics*, 227(1):494–521, 2011.

[9] Raphael Berthier, Andrea Montanari, and Phan-Minh Nguyen. State evolution for approximate message passing with non-separable functions. *arXiv preprint arXiv:1708.03950*, 2017.

[10] Peter Bickel, David Choi, Xiangyu Chang, and Hai Zhang. Asymptotic normality of maximum likelihood and its variational approximation for stochastic blockmodels. *The Annals of Statistics*, pages 1922–1943, 2013.

[11] Kurt Binder. Finite size scaling analysis of ising model block distribution functions. *Zeitschrift für Physik B Condensed Matter*, 43(2):119–140, 1981.

[12] David M Blei. Probabilistic topic models. *Communications of the ACM*, 55(4):77–84, 2012.

[13] David M Blei, Alp Kucukelbir, and Jon D McAuliffe. Variational inference: A review for statisticians. *Journal of the American Statistical Association*, 112(518):859–877, 2017.

[14] David M Blei and John D Lafferty. Dynamic topic models. In *Proceedings of the 23rd international conference on Machine learning*, pages 113–120. ACM, 2006.

[15] David M Blei and John D Lafferty. A correlated topic model of science. *The Annals of Applied Statistics*, pages 17–35, 2007.

[16] David M Blei, Andrew Y Ng, and Michael I Jordan. Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan):993–1022, 2003.

[17] Tamara Broderick, Nicholas Boyd, Andre Wibisono, Ashia C Wilson, and Michael I Jordan. Streaming variational bayes. In *Advances in Neural Information Processing Systems*, pages 1727–1735, 2013.

[18] Alain Celisse, Jean-Jacques Daudin, Laurent Pierre, et al. Consistency of maximum-likelihood and variational estimators in the stochastic block model. *Electronic Journal of Statistics*, 6:1847–1899, 2012.

[19] Jonathan Chang and David Blei. Relational topic models for document networks. In *Artificial Intelligence and Statistics*, pages 81–88, 2009.

[20] Yash Deshpande, Emmanuel Abbe, and Emmanuel Abbe. Asymptotic mutual information for the balanced binary stochastic block model. *Information and Inference: A Journal of the IMA*, 6(2):125–170, 2016.

[21] Yash Deshpande and Andrea Montanari. Information-theoretically optimal sparse pca. In *Information Theory (ISIT), 2014 IEEE International Symposium on*, pages 2197–2201. IEEE, 2014.

[22] David L Donoho, Arian Maleki, and Andrea Montanari. Message-passing algorithms for compressed sensing. *Proceedings of the National Academy of Sciences*, 106(45):18914–18919, 2009.

[23] David L Donoho, Arian Maleki, and Andrea Montanari. Message passing algorithms for compressed sensing: I. motivation and construction. In *Information Theory (ITW 2010, Cairo), 2010 IEEE Information Theory Workshop on*, pages 1–5. IEEE, 2010.

[24] Elena Erosheva, Stephen Fienberg, and John Lafferty. Mixed-membership models of scientific publications. *Proceedings of the National Academy of Sciences*, 101(suppl 1):5220–5227, 2004.

[25] Li Fei-Fei and Pietro Perona. A bayesian hierarchical model for learning natural scene categories. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 2, pages 524–531. IEEE, 2005.

[26] Delphine Féral and Sandrine Péché. The largest eigenvalue of rank one deformation of large wigner matrices. *Communications in mathematical physics*, 272(1):185–228, 2007.

[27] Ryan J Giordano, Tamara Broderick, and Michael I Jordan. Linear response methods for accurate covariance estimates from mean field variational bayes. In *Advances in Neural Information Processing Systems*, pages 1441–1449, 2015.

[28] Matthew Hoffman, Francis R Bach, and David M Blei. Online learning for latent dirichlet allocation. In *advances in neural information processing systems*, pages 856–864, 2010.

[29] Adel Javanmard and Andrea Montanari. State evolution for general approximate message passing algorithms, with applications to spatial coupling. *Information and Inference: A Journal of the IMA*, 2(2):115–144, 2013.

[30] Michael I Jordan, Zoubin Ghahramani, Tommi S Jaakkola, and Lawrence K Saul. An introduction to variational methods for graphical models. *Machine learning*, 37(2):183–233, 1999.

[31] Satish Babu Korada and Nicolas Macris. Exact solution of the gauge symmetric p-spin glass model on a complete graph. *Journal of Statistical Physics*, 136(2):205–230, 2009.

[32] Florent Krzakala, Jiaming Xu, and Lenka Zdeborová. Mutual information in rank-one matrix estimation. In *Information Theory Workshop (ITW), 2016 IEEE*, pages 71–75. IEEE, 2016.

[33] John D Lafferty and David M Blei. Correlated topic models. In *Advances in neural information processing systems*, pages 147–154, 2006.

[34] Marc Lelarge and Léo Miolane. Fundamental limits of symmetric low-rank matrix estimation. *arXiv preprint arXiv:1611.03888*, 2016.

[35] Thibault Lesieur, Florent Krzakala, and Lenka Zdeborová. Constrained low-rank matrix estimation: Phase transitions, approximate message passing and applications. *Journal of Statistical Mechanics: Theory and Experiment*, 2017(7):073403, 2017.

[36] Léo Miolane. Fundamental limits of low-rank matrix estimation. *arXiv preprint arXiv:1702.00473*, 2017.

[37] Andrea Montanari and Ramji Venkataramanan. Estimation of low-rank matrices via approximate message passing. *arXiv preprint arXiv:1711.01682*, 2017.

[38] Manfred Opper and Ole Winther. Adaptive and self-averaging thouless-anderson-palmer mean-field theory for probabilistic modeling. *Physical Review E*, 64(5):056131, 2001.

[39] Giorgi Parisi, Marc Mézard, and Miguel Angel Virasoro. Spin glass theory and beyond. *World Scientific, Singapore*, 187:202, 1987.

[40] Debdeep Pati, Anirban Bhattacharya, and Yun Yang. On statistical optimality of variational bayes. *arXiv preprint arXiv:1712.08983*, 2017.

[41] Lawrence Perko. *Differential equations and dynamical systems*, volume 7. Springer Science & Business Media, 2013.

[42] Anil Raj, Matthew Stephens, and Jonathan K Pritchard. faststructure: variational inference of population structure in large snp data sets. *Genetics*, 197(2):573–589, 2014.

[43] Ben Recht, Christopher Re, Joel Tropp, and Victor Bittorf. Factoring nonnegative matrices with linear programs. In *Advances in Neural Information Processing Systems*, pages 1214–1222, 2012.

[44] David J Thouless, Philip W Anderson, and Robert G Palmer. Solution of'solvable model of a spin glass'. *Philosophical Magazine*, 35(3):593–601, 1977.

[45] Martin J Wainwright, Michael I Jordan, et al. Graphical models, exponential families, and variational inference. *Foundations and Trends® in Machine Learning*, 1(1–2):1–305, 2008.

[46] Bo Wang and DM Titterington. Convergence and asymptotic normality of variational bayesian approximations for exponential family models with missing values. In *Proceedings of the 20th conference on Uncertainty in artificial intelligence*, pages 577–584. AUAI Press, 2004.

[47] Bo Wang and DM Titterington. Inadequacy of interval estimates corresponding to variational bayesian approximations. In *AISTATS*, 2005.

[48] Bo Wang, DM Titterington, et al. Convergence properties of a general algorithm for calculating variational bayesian estimates for a normal mixture model. *Bayesian Analysis*, 1(3):625–650, 2006.

[49] Chong Wang and David M Blei. Collaborative topic modeling for recommending scientific articles. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 448–456. ACM, 2011.

[50] Xiaogang Wang and Eric Grimson. Spatial latent dirichlet allocation. In *Advances in neural information processing systems*, pages 1577–1584, 2008.

[51] Yixin Wang and David M Blei. Frequentist consistency of variational bayes. arXiv:1705.03439, 2017.

[52] Anderson Y Zhang and Harrison H Zhou. Theoretical and computational guarantees of mean field variational inference for community detection. *arXiv preprint arXiv:1710.11268*, 2017.

[53] Jing Zhou, Anirban Bhattacharya, Amy H Herring, and David B Dunson. Bayesian factorizations of big sparse tensors. *Journal of the American Statistical Association*, 110(512):1562–1576, 2015.

A Some remarks on alternating minimization

Let $f : \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}$ be twice continuously differentiable in an open neighborhood $\Omega_1 \times \Omega_2 \subseteq \mathbb{R}^n \times \mathbb{R}^d$ of a critical point $(\boldsymbol{x}^*, \boldsymbol{y}^*)$ (i.e. a point for which $\nabla_{(\boldsymbol{x},\boldsymbol{y})} f(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{0}$). Further assume that, fixing $\boldsymbol{x}_0 \in \Omega_1$, $f(\boldsymbol{x}_0, \,\cdot\,)$ is strongly convex with a minimizer in Ω_2, and fixing $\boldsymbol{y}_0 \in \Omega_2$, $f(\,\cdot\,, \boldsymbol{y}_0)$ is strongly convex with a minimizer in Ω_1. By taking Ω_1 and Ω_2 sufficiently small, these conditions follow by requiring that the partial Hessians satisfy $\nabla^2_{\boldsymbol{x}} f(\boldsymbol{x}^*, \boldsymbol{y}^*) \succ \boldsymbol{0}$ and $\nabla^2_{\boldsymbol{y}} f(\boldsymbol{x}^*, \boldsymbol{y}^*) \succ \boldsymbol{0}$ (i.e. they are strictly positive definite).

By strong convexity, the minimizers of $f(\boldsymbol{x}_0, \,\cdot\,)$ and $f(\,\cdot\,, \boldsymbol{y}_0)$ are unique, and we can define the functions $g : \mathbb{R}^d \to \mathbb{R}^n$ and $h : \mathbb{R}^n \to \mathbb{R}^d$ by

$$h(\boldsymbol{x}_0) = \arg\min_{\boldsymbol{y} \in \Omega_2} f(\boldsymbol{x}_0, \boldsymbol{y}) \,, \tag{A.1}$$

$$g(\boldsymbol{y}_0) = \arg\min_{\boldsymbol{x} \in \Omega_1} f(\boldsymbol{x}, \boldsymbol{y}_0) \,. \tag{A.2}$$

We then define the alternating minimization iteration

$$\boldsymbol{x}^{t+1} = h(\boldsymbol{y}^t), \qquad \boldsymbol{y}^t = g(\boldsymbol{x}^t) \,. \tag{A.3}$$

If $d = n$ and $h : \Omega_1 \to \Omega_2$, $g : \Omega_2 \to \Omega_1$ are bijective, we also define the dual iteration

$$\overline{\boldsymbol{x}}^{t+1} = g^{-1}(\overline{\boldsymbol{y}}^t), \qquad \overline{\boldsymbol{y}}^t = h^{-1}(\overline{\boldsymbol{x}}^t) \,. \tag{A.4}$$

Lemma A.1. *Let $f : \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}$ by twice continuously differentiable in $\Omega_1 \times \Omega_2$, satisfying the above assumptions. Then the following are equivalent:*

(A1) *The Hessian $\boldsymbol{H} = \nabla^2_{(\boldsymbol{x},\boldsymbol{y})} f\big|_{(\boldsymbol{x},\boldsymbol{y})=(\boldsymbol{x}^*,\boldsymbol{y}^*)}$ is strictly positive definite.*

(A2) *$(\boldsymbol{x}^*, \boldsymbol{y}^*)$ is a stable fixed point of the alternate minimization algorithm (A.3).*

(A3) *$f_1(\boldsymbol{x}) \equiv \min_{\boldsymbol{y} \in \Omega_2} f(\boldsymbol{x}, \boldsymbol{y})$ is strongly convex in a neighborhood of $\boldsymbol{x}^*$ (and in particular, $\boldsymbol{x}^*$ is a local minimum of f_1).*

Further, if $n = d$ and the matrix $\frac{\partial f}{\partial \boldsymbol{x} \partial \boldsymbol{y}}\big|_{\boldsymbol{x}^*,\boldsymbol{y}^*}$ is invertible, then the following are equivalent:

(B1) *$(\boldsymbol{x}^*, \boldsymbol{y}^*)$ is a stable fixed point of the dual algorithm (A.4).*

(B2) *$f_1(\boldsymbol{x}) \equiv \min_{\boldsymbol{y} \in \Omega_2} f(\boldsymbol{x}, \boldsymbol{y})$ is strongly concave in a neighborhood of $\boldsymbol{x}^*$ (and in particular, $\boldsymbol{x}^*$ is a local maximum).*

Proof. Let

$$\boldsymbol{H} = \begin{bmatrix} \boldsymbol{H}_{\boldsymbol{x}\boldsymbol{x}} & \boldsymbol{H}_{\boldsymbol{x}\boldsymbol{y}} \\ \boldsymbol{H}_{\boldsymbol{x}\boldsymbol{y}}^{\mathsf{T}} & \boldsymbol{H}_{\boldsymbol{y}\boldsymbol{y}} \end{bmatrix} = \nabla^2_{(\boldsymbol{x},\boldsymbol{y})} f\Big|_{(\boldsymbol{x},\boldsymbol{y})=(\boldsymbol{x}^*,\boldsymbol{y}^*)}. \tag{A.5}$$

(A1)≡(A2) We compute the linearization of the iterations in (A.3) around the fixed point $(\boldsymbol{x}^*, \boldsymbol{y}^*)$. Note that since $\boldsymbol{x}^*$ is a minimizer of $f(\,\cdot\,, \boldsymbol{y}^*)$, using the implicit function theorem for the Jacobian of the update rule for $\boldsymbol{x}$ in (A.3) we have

$$\left.\frac{\partial^2 f}{\partial \boldsymbol{x}\partial \boldsymbol{y}}\right|_{(\boldsymbol{x},\boldsymbol{y})=(\boldsymbol{x}^*,\boldsymbol{y}^*)} + \left[\left.\frac{\partial^2 f}{\partial \boldsymbol{x}^2}\right|_{(\boldsymbol{x},\boldsymbol{y})=(\boldsymbol{x}^*,\boldsymbol{y}^*)}\right][\boldsymbol{D}h(\boldsymbol{y}^*)] = 0. \tag{A.6}$$

Hence, we get

$$\boldsymbol{D}h(\boldsymbol{y}^*) = -\left[\left(\frac{\partial^2 f}{\partial \boldsymbol{x}^2}\right)^{-1}\left(\frac{\partial^2 f}{\partial \boldsymbol{x}\partial \boldsymbol{y}}\right)\right]_{(\boldsymbol{x},\boldsymbol{y})=(\boldsymbol{x}^*,\boldsymbol{y}^*)} = -\boldsymbol{H}_{\boldsymbol{xx}}^{-1}\boldsymbol{H}_{\boldsymbol{xy}}. \tag{A.7}$$

Similarly, for the Jacobian of the update rule for $\boldsymbol{y}$ in (A.3) we have

$$\boldsymbol{D}g(\boldsymbol{x}^*) = -\left[\left(\frac{\partial^2 f}{\partial \boldsymbol{y}^2}\right)^{-1}\left(\frac{\partial^2 f}{\partial \boldsymbol{y}\partial \boldsymbol{x}}\right)\right]_{(\boldsymbol{x},\boldsymbol{y})=(\boldsymbol{x}^*,\boldsymbol{y}^*)} = -\boldsymbol{H}_{\boldsymbol{yy}}^{-1}\boldsymbol{H}_{\boldsymbol{xy}}^{\top}. \tag{A.8}$$

Hence, $(\boldsymbol{x}^*, \boldsymbol{y}^*)$ is stable if and only if the operator

$$\boldsymbol{L} = \boldsymbol{D}h(\boldsymbol{x}^*)\cdot \boldsymbol{D}g(\boldsymbol{y}^*) = \boldsymbol{H}_{\boldsymbol{xx}}^{-1}\boldsymbol{H}_{\boldsymbol{xy}}\boldsymbol{H}_{\boldsymbol{yy}}^{-1}\boldsymbol{H}_{\boldsymbol{xy}}^{\top}, \tag{A.9}$$

has spectral radius

$$\sigma(\boldsymbol{L}) \equiv \max_i |\lambda_i(\boldsymbol{L})| < 1. \tag{A.10}$$

Since $f(\,\cdot\,, \boldsymbol{x}^*)$ is strongly convex, the matrices $\boldsymbol{H}_{\boldsymbol{xx}}, \boldsymbol{H}_{\boldsymbol{xx}}^{-1}$ are positive definite. Hence, the eigenvalues of $\boldsymbol{H}_{\boldsymbol{xx}}^{-1}\boldsymbol{H}_{\boldsymbol{xy}}\boldsymbol{H}_{\boldsymbol{yy}}^{-1}\boldsymbol{H}_{\boldsymbol{xy}}^{\top}$ are real and equal to the eigenvalues of the symmetric positive semi-definite matrix $\boldsymbol{H}_{\boldsymbol{xx}}^{-1/2}\boldsymbol{H}_{\boldsymbol{xy}}\boldsymbol{H}_{\boldsymbol{yy}}^{-1}\boldsymbol{H}_{\boldsymbol{xy}}^{\top}\boldsymbol{H}_{\boldsymbol{xx}}^{-1/2}$. Therefore, $\sigma(\boldsymbol{L}) < 1$ if and only if

$$\begin{aligned}\boldsymbol{H}_{\boldsymbol{xx}}^{-1/2}\boldsymbol{H}_{\boldsymbol{xy}}\boldsymbol{H}_{\boldsymbol{yy}}^{-1}\boldsymbol{H}_{\boldsymbol{xy}}^{\top}\boldsymbol{H}_{\boldsymbol{xx}}^{-1/2} \prec \boldsymbol{I}_n &\iff \boldsymbol{H}_{\boldsymbol{xy}}\boldsymbol{H}_{\boldsymbol{yy}}^{-1}\boldsymbol{H}_{\boldsymbol{xy}}^{\top} \prec \boldsymbol{H}_{\boldsymbol{xx}} \\ &\iff \boldsymbol{H}_{\boldsymbol{xx}} - \boldsymbol{H}_{\boldsymbol{xy}}\boldsymbol{H}_{\boldsymbol{yy}}^{-1}\boldsymbol{H}_{\boldsymbol{xy}}^{\top} \succ 0. \end{aligned} \tag{A.11}$$

Note that since $f(\boldsymbol{x}^*, \cdot\,)$ is convex, $\boldsymbol{H}_{\boldsymbol{yy}} \succ 0$. Therefore, $\boldsymbol{H}_{\boldsymbol{xx}} - \boldsymbol{H}_{\boldsymbol{xy}}\boldsymbol{H}_{\boldsymbol{yy}}^{-1}\boldsymbol{H}_{\boldsymbol{xy}}^{\top} \succ 0$ if and only if $\boldsymbol{H} \succ 0$. Hence, the fixed point is stable if and only if $\boldsymbol{H} \succ 0$ and this completes the proof.

(A1)≡ (A3) By differentiating $f_1(\boldsymbol{z}) = f(\boldsymbol{x}, g(\boldsymbol{x}))$, we obtain

$$\left.\frac{\partial^2 f_1}{\partial \boldsymbol{x}^2}\right|_{\boldsymbol{x}^*} = \left.\frac{\partial^2 f}{\partial \boldsymbol{x}^2}\right|_{\boldsymbol{x}^*,\boldsymbol{y}^*} + \left.\frac{\partial^2 f}{\partial \boldsymbol{x}\partial \boldsymbol{y}}\right|_{\boldsymbol{x}^*,\boldsymbol{y}^*}\cdot \boldsymbol{D}g(\boldsymbol{x}^*) \tag{A.12}$$

$$= \boldsymbol{H}_{\boldsymbol{xx}} - \boldsymbol{H}_{\boldsymbol{xy}}\boldsymbol{H}_{\boldsymbol{yy}}^{-1}\boldsymbol{H}_{\boldsymbol{xy}}^{\top}, \tag{A.13}$$

where in the last line we used Eq. (A.8). Hence $\left.\frac{\partial^2 f_1}{\partial \boldsymbol{x}^2}\right|_{\boldsymbol{x}^*} \succ \boldsymbol{0}$ if and only if $\boldsymbol{H}_{\boldsymbol{xx}} \succ \boldsymbol{H}_{\boldsymbol{xy}}\boldsymbol{H}_{\boldsymbol{yy}}^{-1}\boldsymbol{H}_{\boldsymbol{xy}}^{\top}$ which, by Schur's complement formula is equivalent to $\boldsymbol{H} \succ \boldsymbol{0}$. Further, since $f \in C^2(\mathbb{R}^{n+d})$, $\frac{\partial^2 f_1}{\partial \boldsymbol{x}^2}\Big|_{\boldsymbol{x}^*} \succ \boldsymbol{0}$ if and only if $\frac{\partial^2 f_1}{\partial \boldsymbol{x}^2} \succ \boldsymbol{0}$ in a neighborhood of $\boldsymbol{x}^*$.

(B1)≡ (B2) Linearizing the iteration (A.4), we get that $(\boldsymbol{x}^*, \boldsymbol{y}^*)$ is a stable fixed point if and only if the operator

$$\boldsymbol{L}^{-1} = \boldsymbol{D}g(\boldsymbol{x}^*)^{-1}\boldsymbol{D}h(\boldsymbol{y}^*)^{-1} = (\boldsymbol{H}_{\boldsymbol{xy}}^{\mathsf{T}})^{-1}\boldsymbol{H}_{\boldsymbol{yy}}\boldsymbol{H}_{\boldsymbol{xy}}^{-1}\boldsymbol{H}_{\boldsymbol{xx}} \tag{A.14}$$

has spectral radius

$$\sigma(\boldsymbol{L}^{-1}) \equiv \max_{i\le n}\left|\lambda_i\left(\boldsymbol{L}^{-1}\right)\right| < 1. \tag{A.15}$$

Using the fact that $\boldsymbol{H}_{\boldsymbol{xx}} \succ \boldsymbol{0}$, we have that $\sigma(\boldsymbol{L}^{-1}) < 1$ if and only if

$$\begin{aligned}\boldsymbol{H}_{\boldsymbol{xx}}^{1/2}(\boldsymbol{H}_{\boldsymbol{xy}}^{\mathsf{T}})^{-1}\boldsymbol{H}_{\boldsymbol{yy}}\boldsymbol{H}_{\boldsymbol{xy}}^{-1}\boldsymbol{H}_{\boldsymbol{xx}}^{1/2} \prec \boldsymbol{I}_n &\iff (\boldsymbol{H}_{\boldsymbol{xy}}^{\mathsf{T}})^{-1}\boldsymbol{H}_{\boldsymbol{yy}}\boldsymbol{H}_{\boldsymbol{xy}}^{-1} \prec \boldsymbol{H}_{\boldsymbol{xx}}^{-1}\\ &\iff \boldsymbol{H}_{\boldsymbol{xx}} - \boldsymbol{H}_{\boldsymbol{xy}}\boldsymbol{H}_{\boldsymbol{yy}}^{-1}\boldsymbol{H}_{\boldsymbol{xy}}^{\mathsf{T}} \prec \boldsymbol{0}. \end{aligned} \tag{A.16}$$

As shown above, the last condition is equivalent to $\frac{\partial^2 f_1}{\partial \boldsymbol{x}^2}\Big|_{\boldsymbol{x}^*} \prec \boldsymbol{0}$, and by continuity of the Hessian, this is equivalent to f_1 being strongly concave in a neighborhood of $\boldsymbol{x}^*$. □

B Proof of Proposition 2.2

It is useful to first prove a simple random matrix theory remark.

Lemma B.1. *For $S \subseteq [n]$, let $\boldsymbol{X}_{S,S}$ be the submatrix of $\boldsymbol{X}$ with rows and columns with index in S. Then, for any $\varepsilon \in [0,1)$, the following holds with high probability:*

$$\min\Big\{\lambda_{\max}(\boldsymbol{X}_{S,S}) : |S| \ge n(1-\varepsilon)\Big\} \ge 2\sqrt{1-\varepsilon} - o_n(1)\,. \tag{B.1}$$

Proof. Without loss of generality we can assume $\boldsymbol{X} \sim \mathsf{GOE}(n)$ (because the rank-one deformation cannot decrease the maximum eigenvalue), and $|S| = n(1-\varepsilon)$ (because $\lambda_{\max}(\boldsymbol{X}_{S,S})$ is non-decreasing in S). Note that $\boldsymbol{X}_{S,S}$ is distributed as $\sqrt{1-\varepsilon}$ times a $\mathsf{GOE}(n(1-\varepsilon))$ matrix. Large deviation bounds on the eigenvalues of GOE matrices imply that, for any $\delta > 0$, there exists $c(\delta) > 0$ such that

$$\mathbb{P}\Big(\lambda_{\max}(\boldsymbol{X}_{S,S}) \le 2\sqrt{1-\varepsilon} - \delta\Big) \le 2\, e^{-c(\delta)n^2}\,, \tag{B.2}$$

for all n large enough. The claim follows by union bound since there is at most 2^n such sets S. □

Proof of Proposition 2.2. First notice that Lemma B.1 continues to hold if $\boldsymbol{X}$ is replaced by $\boldsymbol{X}_0$ since $\|\boldsymbol{X}_{S,S} - (\boldsymbol{X}_0)_{S,S}\|_{\rm op} \le \max_{i\le n}|X_{ii}| \le 4\sqrt{\log n/n}$ (where the last bound holds with high probability since $(X_{ii})_{i\le n} \sim \mathsf{N}(0, 2/n)$.

Note that $\nabla\mathcal{F}(\boldsymbol{m})_i = \pm\infty$ if $m_i = \pm 1$, whence any local minimum must be in the interior of $[-1,+1]^n$. Let $\boldsymbol{m} \in (-1,-1)^n$ be a local minimum of $\mathcal{F}(\,\cdot\,)$. By the second-order minimality conditions, we must have

$$\nabla^2\mathcal{F}(\boldsymbol{m}) = -\lambda\boldsymbol{X}_0 + \mathrm{diag}\Big((1-m_i^2)^{-1}_{i\le n}\Big) \succeq \mathbf{0}\,. \tag{B.3}$$

Denote by $m_{(1)}$, $m_{(2)}$, ... the entries of $\boldsymbol{m}$ ordered by decreasing absolute value, and let S_ℓ be the set of indices corresponding to entries $m_{(\ell+1)},\dots,m_{(n)}$. Finally let $\boldsymbol{v}^{(\ell)}\in\mathbb{R}^n$ be the eigenvector corresponding to the largest eigenvalue of $(\boldsymbol{X}_0)_{S_\ell,S_\ell}$ (extended with zeros outside S_ℓ). We then have, for $\ell = n\varepsilon$

$$0 \le \langle \boldsymbol{v}^{(\ell)}, \nabla^2\mathcal{F}(\boldsymbol{m})\boldsymbol{v}^{(\ell)}\rangle \tag{B.4}$$

$$= -\lambda\cdot\lambda_{\max}\Big((\boldsymbol{X}_0)_{S_\ell,S_\ell}\Big) + \sum_{i\in S_\ell}\frac{(v_i^{(\ell)})^2}{1-m_i^2} \tag{B.5}$$

$$\le -2\lambda\sqrt{1-\varepsilon} + \frac{1}{1-m_{(n\varepsilon)}^2} + o_n(1)\,. \tag{B.6}$$

The last inequality holds with high probability by Lemma B.1. Inverting it, we get

$$m_{(n\varepsilon)}^2 \ge 1 - \frac{1}{2\lambda\sqrt{1-\varepsilon}} - o_n(1)\,, \tag{B.7}$$

and therefore

$$\frac{1}{n}\|\boldsymbol{m}\|_2^2 \ge \varepsilon\left(1-\frac{1}{2\lambda\sqrt{1-\varepsilon}}\right) - o_n(1). \tag{B.8}$$

The claim follows by taking $\varepsilon = c_1$ a small constant (for which the right-hand side is lower bounded by c_0 for all $\lambda \ge 1$), or $\varepsilon = c_2(2\lambda-1)$ (for which the right-hand side is lower bounded by $c_0(2\lambda-1)^2$). □

C Information-theoretic limits

C.1 Proof of Lemma 2.1

Let $\widehat{\boldsymbol{Q}}:\mathbb{R}^{n\times n}\mapsto\mathbb{R}^{n\times n}$, $\boldsymbol{X}\mapsto\widehat{\boldsymbol{Q}}(\boldsymbol{X})$ be any estimator of $\boldsymbol{\sigma}\boldsymbol{\sigma}^{\mathsf{T}}$. By [5, Theorem 1.6], for $\lambda\in[0,1]$,

$$\liminf_{n\to\infty}\frac{1}{n^2}\mathbb{E}\Big\{\Big\|\boldsymbol{\sigma}\boldsymbol{\sigma}^{\mathsf{T}} - \widehat{\boldsymbol{Q}}(\boldsymbol{X})\Big\|_F^2\Big\} \ge 1\,. \tag{C.1}$$

Given $\hat{\boldsymbol{\sigma}}:\mathbb{R}^{n\times n}\to\mathbb{R}^n\setminus\{\mathbf{0}\}$, set

$$\widehat{\boldsymbol{Q}}(\boldsymbol{X}) = c\,\frac{\hat{\boldsymbol{\sigma}}(\boldsymbol{X})\hat{\boldsymbol{\sigma}}(\boldsymbol{X})^{\mathsf{T}}}{\|\hat{\boldsymbol{\sigma}}(\boldsymbol{X})\|_2^2}\,,\qquad c = \mathbb{E}\left(\frac{\langle\hat{\boldsymbol{\sigma}}(\boldsymbol{X}),\boldsymbol{\sigma}\rangle^2}{\|\hat{\boldsymbol{\sigma}}(\boldsymbol{X})\|_2^2}\right)\,. \tag{C.2}$$

By a simple calculation

$$1 - o_n(1) \le \frac{1}{n^2}\mathbb{E}\Big\{\Big\|\boldsymbol{\sigma}\boldsymbol{\sigma}^{\mathsf{T}} - \widehat{\boldsymbol{Q}}(\boldsymbol{X})\Big\|_F^2\Big\} = 1 - \mathbb{E}\left(\frac{\langle\hat{\boldsymbol{\sigma}}(\boldsymbol{X}), \boldsymbol{\sigma}\rangle^2}{\|\hat{\boldsymbol{\sigma}}(\boldsymbol{X})\|_2^2}\right)^2, \tag{C.3}$$

which obviously implies the claim.

C.2 Proof of Proposition 3.1

We begin by providing the expression for the free energy functional $\mathsf{RS}(\boldsymbol{M}; k, \delta, \nu)$ of Theorem 1, which is obtained by specializing the expression in [8]. Recall the functions $\phi(\cdots)$, $\tilde{\phi}(\cdots)$, introduced in Eq. (3.5). We then define a function $\mathsf{RS}_0(\,\cdot\,,\,\cdot\,; k, \delta, \nu) : \mathbb{S}_k \times \mathbb{S}_k \to \mathbb{R}$ by

$$\begin{aligned}\mathsf{RS}_0(\boldsymbol{M}, \widetilde{\boldsymbol{M}}; k, \delta, \nu) =& \frac{\beta\delta(\nu+1)}{k\nu+1} + \frac{1}{2\beta}\langle\boldsymbol{M}, \widetilde{\boldsymbol{M}}\rangle \\ &- \mathbb{E}\,\phi(\boldsymbol{M}\boldsymbol{h} + \boldsymbol{M}^{1/2}\boldsymbol{z}; \boldsymbol{M}) - \delta\,\mathbb{E}\,\tilde{\phi}(\widetilde{\boldsymbol{M}}\boldsymbol{w} + \widetilde{\boldsymbol{M}}^{1/2}\boldsymbol{z}; \widetilde{\boldsymbol{M}}),\end{aligned} \tag{C.4}$$

where expectations are with respect to $\boldsymbol{z} \sim \mathsf{N}(0, \boldsymbol{I}_k)$ independent of $\boldsymbol{h} \sim \mathsf{N}(0, \boldsymbol{I}_k)$ and $\boldsymbol{w} \sim \mathrm{Dir}(\nu; k)$. We then have

$$\mathsf{RS}(\boldsymbol{M}; k, \delta, \nu) = \sup_{\widetilde{\boldsymbol{M}} \in \mathbb{S}_k} \mathsf{RS}_0(\boldsymbol{M}, \widetilde{\boldsymbol{M}}; k, \delta, \nu). \tag{C.5}$$

Further, the function $\mathsf{RS}_0(\boldsymbol{M}, \widetilde{\boldsymbol{M}}; k, \delta, \nu)$ on Eq. (C.4) is separately strictly concave in $\boldsymbol{M}$ and $\widetilde{\boldsymbol{M}}$, and in particular the last supremum is uniquely achieved at a point $\widetilde{\boldsymbol{M}} = \widetilde{\boldsymbol{M}}(\boldsymbol{M})$.

A simple calculation shows that

$$\frac{\partial \mathsf{RS}_0}{\partial \boldsymbol{M}}(\boldsymbol{M}, \widetilde{\boldsymbol{M}}; k, \delta, \nu) = \frac{1}{2\beta}\left\{\widetilde{\boldsymbol{M}} - \mathbb{E}\Big\{\mathsf{F}(\boldsymbol{M}\boldsymbol{h} + \boldsymbol{M}^{1/2}\boldsymbol{z}; \boldsymbol{M})^{\otimes 2}\Big\}\right\}, \tag{C.6}$$

$$\frac{\partial \mathsf{RS}_0}{\partial \widetilde{\boldsymbol{M}}}(\boldsymbol{M}, \widetilde{\boldsymbol{M}}; k, \delta, \nu) = \frac{1}{2\beta}\left\{\boldsymbol{M} - \delta\mathbb{E}\Big\{\tilde{\mathsf{F}}(\widetilde{\boldsymbol{M}}\boldsymbol{w} + \widetilde{\boldsymbol{M}}^{1/2}\boldsymbol{z}; \widetilde{\boldsymbol{M}})^{\otimes 2}\Big\}\right\}. \tag{C.7}$$

By Lemma D.1, for $\boldsymbol{M} = a\boldsymbol{J}_k$, $\widetilde{\boldsymbol{M}} = b\boldsymbol{J}_k$, we have

$$\frac{\partial \mathsf{RS}_0}{\partial \boldsymbol{M}}(\boldsymbol{M}, \widetilde{\boldsymbol{M}}; k, \delta, \nu) = \frac{1}{2\beta}\left\{b\boldsymbol{J}_k - \frac{\beta a}{1+ka}\boldsymbol{J}_k\right\}, \tag{C.8}$$

$$\frac{\partial \mathsf{RS}_0}{\partial \widetilde{\boldsymbol{M}}}(\boldsymbol{M}, \widetilde{\boldsymbol{M}}; k, \delta, \nu) = \frac{1}{2\beta}\left\{a\boldsymbol{J}_k - \frac{\beta\delta}{k^2}\boldsymbol{J}_k\right\}. \tag{C.9}$$

Therefore, this is a stationary point of RS_0 provided $a = \beta\delta/k^2$ and $b = \beta^2\delta/(k(k+\beta\delta))$ (in particular, $\boldsymbol{M} = \boldsymbol{M}^*$). Since $\mathsf{RS}(\boldsymbol{M}; k, \delta, \nu) = \mathsf{RS}_0(\boldsymbol{M}, \widetilde{\boldsymbol{M}}(\boldsymbol{M}); k, \delta, \nu)$, for $\widetilde{\boldsymbol{M}}(\,\cdot\,)$ a differentiable function, it also follows that $\boldsymbol{M}_*$ is a stationary point of RS.

In order to prove that $\boldsymbol{M}^*$ is a local minimum of RS for $\beta < \beta_{\text{spect}}$, we apply Lemma A.1 to the function $f(\boldsymbol{x}, \boldsymbol{y}) = -\mathsf{RS}_0(\boldsymbol{x}, \boldsymbol{y}; k, \delta, \nu)$, whence $f_1(\boldsymbol{x}) = -\mathsf{RS}(\boldsymbol{x}; k, \delta, \nu)$. It follows from Eqs. (C.6) and (C.7) that the dynamics (A.4) then coincides with the state evolution dynamics discussed in Section 4.3, namely

$$\boldsymbol{M}_{t+1} = \delta\,\mathbb{E}\Big\{\widetilde{\mathsf{F}}(\widetilde{\boldsymbol{M}}_t\boldsymbol{w} + \widetilde{\boldsymbol{M}}_t^{1/2}\boldsymbol{z}; \widetilde{\boldsymbol{M}}_t)^{\otimes 2}\Big\}\,, \tag{C.10}$$

$$\widetilde{\boldsymbol{M}}_t = \mathbb{E}\Big\{\mathsf{F}(\boldsymbol{M}_t\boldsymbol{h} + \boldsymbol{M}_t^{1/2}\boldsymbol{z}; \boldsymbol{M}_t)^{\otimes 2}\Big\}\,. \tag{C.11}$$

Hence, the claim follows immediately from Theorem 4 and Lemma A.1.

Finally, we prove that Eq. (3.4) holds for $\beta < \beta_{\text{Bayes}}$. Note that the estimator $\widehat{\boldsymbol{F}}_n(\boldsymbol{X})$ that minimizes the left-hand side is $\widehat{\boldsymbol{F}}_n(\boldsymbol{X}) = \mathbb{E}\{\boldsymbol{W}\boldsymbol{H}^{\mathsf{T}}|\boldsymbol{X}\}$. By [8, Proposition 29], for $\beta < \beta_{\text{Bayes}}$,

$$\begin{aligned}\lim_{n\to\infty}\frac{1}{nd}\mathbb{E}\left\{\left\|\boldsymbol{W}\boldsymbol{H}^{\mathsf{T}} - \mathbb{E}\{\boldsymbol{W}\boldsymbol{H}^{\mathsf{T}}|\boldsymbol{X}\}\right\|_F^2\right\} &= \lim_{n\to\infty}\frac{1}{nd}\mathbb{E}\left\{\left\|\boldsymbol{W}\boldsymbol{H}^{\mathsf{T}}\right\|_F^2\right\} - \frac{1}{\beta^2\delta}\mathsf{Tr}(\boldsymbol{M}^*\widetilde{\boldsymbol{M}}^*)\\ &= \lim_{n\to\infty}\frac{1}{nd}\mathbb{E}\left\{\left\|\boldsymbol{W}\boldsymbol{H}^{\mathsf{T}}\right\|_F^2\right\} - \frac{\beta\delta}{k(\beta\delta+k)}\,. \end{aligned}\tag{C.12}$$

On the other hand,

$$\lim_{n\to\infty}\frac{1}{nd}\mathbb{E}\left\{\left\|\boldsymbol{W}\boldsymbol{H}^{\mathsf{T}} - c\boldsymbol{1}_n(\boldsymbol{X}^{\mathsf{T}}\boldsymbol{1}_n)^{\mathsf{T}}\right\|_F^2\right\} = \lim_{n\to\infty}\frac{1}{nd}\mathbb{E}\left\{\left\|\boldsymbol{W}\boldsymbol{H}^{\mathsf{T}}\right\|_F^2\right\} - 2c\,A + c^2B\,. \tag{C.13}$$

Here, we defined A via

$$\begin{aligned}A &\equiv \lim_{n\to\infty}\frac{1}{nd}\mathbb{E}\mathsf{Tr}\Big(\boldsymbol{H}\boldsymbol{W}^{\mathsf{T}}\boldsymbol{1}_n(\boldsymbol{X}^{\mathsf{T}}\boldsymbol{1}_n)^{\mathsf{T}}\Big)\\ &= \lim_{n\to\infty}\frac{\sqrt{\beta}}{nd^2}\mathbb{E}\mathsf{Tr}\Big(\boldsymbol{W}^{\mathsf{T}}\boldsymbol{1}_n\boldsymbol{1}_n^{\mathsf{T}}\boldsymbol{W}\boldsymbol{H}^{\mathsf{T}}\boldsymbol{H}\Big)\\ &= \sqrt{\beta}\delta\mathsf{Tr}\Big(\frac{\boldsymbol{1}_k}{k}\frac{\boldsymbol{1}_k^{\mathsf{T}}}{k}\boldsymbol{I}_k\Big) = \frac{\sqrt{\beta}\delta}{k},\end{aligned}$$

(where we used $\boldsymbol{W}^{\mathsf{T}}\boldsymbol{1}_n/n \to \boldsymbol{1}_k/k$ and $\boldsymbol{H}^{\mathsf{T}}\boldsymbol{H}/d \to \boldsymbol{I}_k$ by the law of large numbers) and

$$\begin{aligned}B &\equiv \lim_{n\to\infty}\frac{1}{nd}\mathbb{E}\mathsf{Tr}\Big(\boldsymbol{1}_n(\boldsymbol{X}^{\mathsf{T}}\boldsymbol{1}_n)^{\mathsf{T}}(\boldsymbol{X}^{\mathsf{T}}\boldsymbol{1}_n)\boldsymbol{1}_n\Big)\\ &= \lim_{n\to\infty}\frac{1}{d}\mathbb{E}\Big\langle\boldsymbol{1}_n, \boldsymbol{X}\boldsymbol{X}^{\mathsf{T}}\boldsymbol{1}_n\Big\rangle\\ &= \lim_{n\to\infty}\frac{1}{d}\mathbb{E}\left\{\frac{\beta}{d^2}\mathsf{Tr}\Big((\boldsymbol{W}^{\mathsf{T}}\boldsymbol{1}_n)^{\mathsf{T}}\boldsymbol{H}^{\mathsf{T}}\boldsymbol{H}(\boldsymbol{W}^{\mathsf{T}}\boldsymbol{1}_n)\Big) + n\right\}\\ &= \beta\delta^2\mathsf{Tr}\Big(\frac{\boldsymbol{1}_k^{\mathsf{T}}}{k}\frac{\boldsymbol{1}_k}{k}\boldsymbol{I}_k\Big) + \delta = \frac{\beta\delta^2}{k} + \delta\,.\end{aligned}$$

Setting $c = A/B$, and substituting in Eq. (C.13), we obtain

$$\lim_{n\to\infty} \frac{1}{nd}\mathbb{E}\left\{\left\|\boldsymbol{W}\boldsymbol{H}^{\mathsf{T}} - c\mathbf{1}_n(\boldsymbol{X}^{\mathsf{T}}\mathbf{1}_n)^{\mathsf{T}}\right\|_F^2\right\} = \lim_{n\to\infty} \frac{1}{nd}\mathbb{E}\left\{\left\|\boldsymbol{W}\boldsymbol{H}^{\mathsf{T}}\right\|_F^2\right\} - \frac{\beta\delta}{k(\beta\delta+k)}\,, \qquad \text{(C.14)}$$

which coincides with Eq. (C.12) as claimed.

D Naive Mean Field: Analytical results

D.1 Preliminary definitions

The functions $\mathsf{F}, \widetilde{\mathsf{F}} : \mathbb{R}^k \times \mathbb{R}^{k\times k} \to \mathbb{R}^k$ are defined in Eq. (3.6). Explicitly

$$\mathsf{F}(\boldsymbol{y};\boldsymbol{Q}) \equiv \sqrt{\beta}\,\frac{\int \boldsymbol{h}\,\exp\{\langle \boldsymbol{y},\boldsymbol{h}\rangle - \langle \boldsymbol{h},\boldsymbol{Q}\boldsymbol{h}\rangle/2\}\,q_0(\mathrm{d}\boldsymbol{h})}{\int \exp\{\langle \boldsymbol{y},\boldsymbol{h}\rangle - \langle \boldsymbol{h},\boldsymbol{Q}\boldsymbol{h}\rangle/2\}\,q_0(\mathrm{d}\boldsymbol{h})}\,, \qquad \text{(D.1)}$$

$$\widetilde{\mathsf{F}}(\tilde{\boldsymbol{y}};\widetilde{\boldsymbol{Q}}) \equiv \sqrt{\beta}\,\frac{\int \boldsymbol{w}\,\exp\{\langle \tilde{\boldsymbol{y}},\boldsymbol{w}\rangle - \langle \boldsymbol{w},\widetilde{\boldsymbol{Q}}\boldsymbol{w}\rangle/2\}\,\tilde{q}_0(\mathrm{d}\boldsymbol{w})}{\int \exp\{\langle \tilde{\boldsymbol{y}},\boldsymbol{w}\rangle - \langle \boldsymbol{w},\widetilde{\boldsymbol{Q}}\boldsymbol{w}\rangle/2\}\,\tilde{q}_0(\mathrm{d}\boldsymbol{w})}\,, \qquad \text{(D.2)}$$

where $q_0(\,\cdot\,)$ is the prior distribution of the rows of $\boldsymbol{H}$, and $\tilde{q}_0(\,\cdot\,)$ is the prior distribution of the rows of $\boldsymbol{W}$.

For $\boldsymbol{Q}$ positive semidefinite and symmetric, $\mathsf{F}(\boldsymbol{y};\boldsymbol{Q})/\sqrt{\beta}$ can be interpreted as the posterior expectation of $\boldsymbol{h} \sim q_0(\,\cdot\,)$, given observations $\boldsymbol{y} = \boldsymbol{Q}\boldsymbol{h} + \boldsymbol{Q}^{1/2}\boldsymbol{z}$, where $\boldsymbol{z} \sim \mathsf{N}(0, \boldsymbol{I}_k)$, and analogously for $\widetilde{\mathsf{F}}(\tilde{\boldsymbol{y}};\widetilde{\boldsymbol{Q}})$. Explicitly

$$\mathsf{F}(\boldsymbol{y};\boldsymbol{Q}) = \sqrt{\beta}\,\mathbb{E}\Big\{\boldsymbol{h}\Big|\,\boldsymbol{Q}\boldsymbol{h} + \boldsymbol{Q}^{1/2}\boldsymbol{z} = \boldsymbol{y}\Big\}\,, \qquad \widetilde{\mathsf{F}}(\tilde{\boldsymbol{y}};\widetilde{\boldsymbol{Q}}) = \sqrt{\beta}\,\mathbb{E}\Big\{\boldsymbol{w}\Big|\,\widetilde{\boldsymbol{Q}}\boldsymbol{w} + \widetilde{\boldsymbol{Q}}^{1/2}\boldsymbol{z} = \tilde{\boldsymbol{y}}\Big\}\,.$$

In our specific application $q_0(\,\cdot\,)$ is $\mathsf{N}(0, \boldsymbol{I}_k)$, and $\tilde{q}_0(\,\cdot\,)$ is $\mathrm{Dir}(\nu; k)$, namely

$$q_0(\mathrm{d}\boldsymbol{h}) = \frac{1}{(2\pi)^{k/2}}\,\exp\{-\|\boldsymbol{h}\|_2^2/2\}\mathrm{d}\boldsymbol{h}\,, \qquad \tilde{q}_0(\mathrm{d}\boldsymbol{w}) = \frac{1}{Z(\nu;k)}\prod_{i=1}^k w_i^{\nu-1}\,\overline{q}(\mathrm{d}\boldsymbol{w})\,, \qquad \text{(D.3)}$$

where $\overline{q}(\,\cdot\,)$ is the uniform measure over the simplex $\mathsf{P}_1(k) = \{\boldsymbol{w} \in \mathbb{R}^k_{\geq 0}\ :\ \langle \boldsymbol{w}, \mathbf{1}_k\rangle = 1\}$. In particular, $\mathsf{F}(\boldsymbol{y};\boldsymbol{Q})$ can be computed explicitly, yielding

$$\mathsf{F}(\boldsymbol{y};\boldsymbol{Q}) = \sqrt{\beta}(\boldsymbol{I}_k + \boldsymbol{Q})^{-1}\boldsymbol{y}\,. \qquad \text{(D.4)}$$

We also define the second moment functions $\mathsf{G}, \widetilde{\mathsf{G}} : \mathbb{R}^k \times \mathbb{R}^{k\times k} \to \mathbb{R}^{k\times k}$ by

$$\mathsf{G}(\boldsymbol{y};\boldsymbol{Q}) \equiv \beta\,\frac{\int \boldsymbol{h}^{\otimes 2}\,\exp\{\langle \boldsymbol{y},\boldsymbol{h}\rangle - \langle \boldsymbol{h},\boldsymbol{Q}\boldsymbol{h}\rangle/2\}\,q_0(\mathrm{d}\boldsymbol{h})}{\int \exp\{\langle \boldsymbol{y},\boldsymbol{h}\rangle - \langle \boldsymbol{h},\boldsymbol{Q}\boldsymbol{h}\rangle/2\}\,q_0(\mathrm{d}\boldsymbol{h})}\,, \qquad \text{(D.5)}$$

$$\widetilde{\mathsf{G}}(\tilde{\boldsymbol{y}};\widetilde{\boldsymbol{Q}}) \equiv \beta\,\frac{\int \boldsymbol{w}^{\otimes 2}\,\exp\{\langle \tilde{\boldsymbol{y}},\boldsymbol{w}\rangle - \langle \boldsymbol{w},\widetilde{\boldsymbol{Q}}\boldsymbol{w}\rangle/2\}\,\tilde{q}_0(\mathrm{d}\boldsymbol{w})}{\int \exp\{\langle \tilde{\boldsymbol{y}},\boldsymbol{w}\rangle - \langle \boldsymbol{w},\widetilde{\boldsymbol{Q}}\boldsymbol{w}\rangle/2\}\,\tilde{q}_0(\mathrm{d}\boldsymbol{w})}\,. \qquad \text{(D.6)}$$

Again, $\mathsf{G}(\cdots)$ can be written explicitly as

$$\mathsf{G}(\boldsymbol{y};\boldsymbol{Q}) = \beta\Big\{(\boldsymbol{I}_k+\boldsymbol{Q})^{-1}\boldsymbol{y}\boldsymbol{y}^{\mathsf{T}}(\boldsymbol{I}_k+\boldsymbol{Q})^{-1} + (\boldsymbol{I}_k+\boldsymbol{Q})^{-1}\Big\}\,. \tag{D.7}$$

D.2 Derivation of the iteration (3.20), (3.21)

Let $\mathcal{D}$, the set of joint distributions $\hat{q}\,(\boldsymbol{W},\boldsymbol{H})$ that factorize over the rows of $\boldsymbol{W},\boldsymbol{H}$, namely

$$\hat{q}\,(\boldsymbol{W},\boldsymbol{H}) = q\,(\boldsymbol{H})\,\tilde{q}\,(\boldsymbol{W}) = \prod_{i=1}^{d} q_i\,(\boldsymbol{h}_i)\prod_{a=1}^{n}\tilde{q}_a\,(\boldsymbol{w}_a)\ . \tag{D.8}$$

The goal in variational inference is to find the distribution in $\mathcal{D}$ that minimizes the Kullback-Leibler (KL) divergence with respect to the actual posterior distribution of $\boldsymbol{X},\boldsymbol{W}$ given $\boldsymbol{X}$

$$\hat{q}^*\,(\,\cdot\,,\,\cdot\,) = \arg\min_{\hat{q}\in\mathcal{D}} \mathrm{KL}\,(\hat{q}\,(\,\cdot\,,\,\cdot\,)\,||\,p\,(\,\cdot\,,\,\cdot\,|\boldsymbol{X})) \tag{D.9}$$

The KL divergence can also be written as (denoting by $\mathbb{E}_{\hat{q}}$ expectation over $(\boldsymbol{W},\boldsymbol{H})\sim\hat{q}(\,\cdot\,,\,\cdot\,)$)

$$\begin{aligned}\mathrm{KL}\,(\hat{q}\,(\,\cdot\,,\,\cdot\,)\,||\,p\,(\,\cdot\,,\,\cdot\,|\boldsymbol{X})) &= \mathbb{E}_{\hat{q}}\,[\log\hat{q}\,(\boldsymbol{W},\boldsymbol{H})] - \mathbb{E}_{\hat{q}}\,[\log p\,(\boldsymbol{X},\boldsymbol{W},\boldsymbol{H})] + \log p\,(\boldsymbol{X}) &\text{(D.10)}\\ &\equiv \mathcal{F}(\hat{q}) + \log p\,(\boldsymbol{X})\ . &\text{(D.11)}\end{aligned}$$

The function $\mathcal{F}(\hat{q})$ is known as Gibbs free energy or –within the topic models literature– as the opposite of the evidence lower bound $\mathcal{F}(\hat{q}) = -\mathrm{ELBO}(\hat{q})$ [4]. Since $\log p\,(\boldsymbol{X})$ does not depend on $\hat{q}$, minimizing the KL divergence is equivalent to minimizing the Gibbs free energy.

In order to find $\hat{q}^*\,(\boldsymbol{W},\boldsymbol{H}) = q^*\,(\boldsymbol{H})\,\tilde{q}^*\,(\boldsymbol{W})$, the naive mean field iteration minimizes the Gibbs free energy by alternating minimization: we minimize the Gibbs free energy over $q\,(\boldsymbol{H})$ (while keeping $\tilde{q}\,(\boldsymbol{W})$ fixed), then minimize over $\tilde{q}\,(\boldsymbol{W})$ (while keeping $q\,(\boldsymbol{H})$ fixed), and repeat. With a slight abuse of notation, we will write $\mathcal{F}(\hat{q}) = \mathcal{F}(q,\tilde{q})$. Note that if we keep $\tilde{q}\,(\boldsymbol{W})$ fixed, we have

$$\begin{aligned}\arg\min_q \mathcal{F}(q,\tilde{q}) &= \arg\min_q \Big\{\mathbb{E}_{q(\boldsymbol{H})}\,[\log q\,(\boldsymbol{H})] - \mathbb{E}_{q(\boldsymbol{H})}\Big[\mathbb{E}_{\tilde{q}(\boldsymbol{W})}\,[\log p\,(\boldsymbol{X},\boldsymbol{W},\boldsymbol{H})]\Big]\Big\}\\ &= \arg\min_q \mathrm{KL}\Big(q\,(\boldsymbol{H})\,||\,C\exp\Big\{\mathbb{E}_{\tilde{q}(\boldsymbol{W})}\,[\log p\,(\boldsymbol{X},\boldsymbol{W},\boldsymbol{H})]\Big\}\Big)\\ &\propto \exp\Big\{\mathbb{E}_{\tilde{q}(\boldsymbol{W})}\,[\log p\,(\boldsymbol{X},\boldsymbol{W},\boldsymbol{H})]\Big\}\ . \end{aligned}\tag{D.12}$$

Similarly, by taking $q\,(\boldsymbol{H})$ fixed, we have

$$\arg\min_{\tilde{q}} \mathcal{F}(q,\tilde{q}) \propto \exp\Big\{\mathbb{E}_{q(\boldsymbol{H})}\,[\log p\,(\boldsymbol{X},\boldsymbol{W},\boldsymbol{H})]\Big\}\,. \tag{D.13}$$

Therefore, the naive mean field iterations have the form

$$
\begin{aligned}
q^{t+1}(\boldsymbol{H}) &= \prod_{i=1}^{d} q_i^{t+1}(\boldsymbol{h}_i) \propto \exp\left\{\mathbb{E}_{\tilde{q}^t(\boldsymbol{W})}\left[\log p(\boldsymbol{X}, \boldsymbol{W}, \boldsymbol{H})\right]\right\}, \\
\tilde{q}^t(\boldsymbol{W}) &= \prod_{a=1}^{n} \tilde{q}_a^t(\boldsymbol{w}_a) \ \propto \exp\left\{\mathbb{E}_{q^t(\boldsymbol{H})}\left[\log p(\boldsymbol{X}, \boldsymbol{W}, \boldsymbol{H})\right]\right\}.
\end{aligned} \tag{D.14}
$$

with initialization

$$
q^0(\boldsymbol{H}) = \prod_{i=1}^{d} q_0(\boldsymbol{h}_i)\,, \qquad \tilde{q}^0(\boldsymbol{W}) = \prod_{a=1}^{n} \tilde{q}_0(\boldsymbol{w}_a) \tag{D.15}
$$

where $q_0(\boldsymbol{h}_i)$, $\tilde{q}_0(\boldsymbol{w}_a)$ are the prior distributions on the rows of $\boldsymbol{H}$ and $\boldsymbol{W}$, cf. Eq. (D.3). Note that the iterations in (D.14) can be further simplified by noting that the densities q_i^t and $\tilde{q}_i^t$ have the form

$$
\begin{aligned}
q_i^t(\boldsymbol{h}) &\propto \exp\left\{\left\langle \boldsymbol{m}_i^t, \boldsymbol{h}\right\rangle - \frac{1}{2}\left\langle \boldsymbol{h}, \boldsymbol{Q}^t \boldsymbol{h}\right\rangle\right\} q_0(\boldsymbol{h})\,, \\
\tilde{q}_a^t(\boldsymbol{w}) &\propto \exp\left\{\left\langle \tilde{\boldsymbol{m}}_a^t, \boldsymbol{w}\right\rangle - \frac{1}{2}\left\langle \boldsymbol{w}, \widetilde{\boldsymbol{Q}}^t \boldsymbol{w}\right\rangle\right\} \tilde{q}_0(\boldsymbol{w})\,.
\end{aligned} \tag{D.16}
$$

In order to see this, note that the initial densities $q_0(\boldsymbol{h})$, $\tilde{q}_0(\boldsymbol{w})$ are in the form (D.16). Further, if we assume that $q_i^t(\boldsymbol{h})$, $\tilde{q}_a^t(\boldsymbol{w})$ are in the form (D.16), using the update equations (D.14), we have

$$
\begin{aligned}
q^{t+1}(\boldsymbol{H}) &= \prod_{i=1}^{d} q_i^{t+1}(\boldsymbol{h}_i) \\
&\propto \exp\left\{\mathbb{E}_{\tilde{q}^t(\boldsymbol{W})} \log p(\boldsymbol{X}, \boldsymbol{H}, \boldsymbol{W})\right\} \\
&\propto \exp\left\{\mathbb{E}_{\tilde{q}^t(\boldsymbol{W})} \log p(\boldsymbol{H}, \boldsymbol{X}|\boldsymbol{W})\right\} \\
&\propto q_0(\boldsymbol{H}) \exp\left\{\mathbb{E}_{\tilde{q}^t(\boldsymbol{W})} \log p(\boldsymbol{X}|\boldsymbol{H}, \boldsymbol{W})\right\} \\
&\propto q_0(\boldsymbol{H}) \exp\left\{-\mathbb{E}_{\tilde{q}^t(\boldsymbol{W})}\left[\frac{d}{2}\left\|\boldsymbol{X} - \frac{\sqrt{\beta}}{d}\boldsymbol{W}\boldsymbol{H}^{\mathsf{T}}\right\|_F^2\right]\right\} \\
&\propto q_0(\boldsymbol{H}) \exp\left\{\mathbb{E}_{\tilde{q}^t(\boldsymbol{W})} \mathsf{Tr}\left(\sqrt{\beta}\boldsymbol{X}\boldsymbol{H}\boldsymbol{W}^{\mathsf{T}} - \frac{\beta}{2d}\boldsymbol{W}\boldsymbol{H}^{\mathsf{T}}\boldsymbol{H}\boldsymbol{W}^{\mathsf{T}}\right)\right\} \\
&= q_0(\boldsymbol{H}) \exp\left\{\mathbb{E}_{\tilde{q}^t(\boldsymbol{W})} \sum_{a=1}^{n}\left(\sqrt{\beta}\left\langle \boldsymbol{x}_a, \boldsymbol{H}\boldsymbol{w}_a\right\rangle - \frac{\beta}{2d}\left\langle \boldsymbol{w}_a, \boldsymbol{H}^{\mathsf{T}}\boldsymbol{H}\boldsymbol{w}_a\right\rangle\right)\right\} \\
&= q_0(\boldsymbol{H}) \exp\left\{\sum_{a=1}^{n}\left\langle \boldsymbol{x}_a, \boldsymbol{H}\widetilde{\mathsf{F}}\left(\tilde{\boldsymbol{m}}_a^t; \widetilde{\boldsymbol{Q}}^t\right)\right\rangle - \frac{1}{2d}\left\langle \boldsymbol{H}^{\mathsf{T}}\boldsymbol{H}, \sum_{a=1}^{n}\widetilde{\mathsf{G}}\left(\tilde{\boldsymbol{m}}_a^t; \widetilde{\boldsymbol{Q}}_t\right)\right\rangle\right\} \\
&= \prod_{i=1}^{d}\left(q_0(\boldsymbol{h}_i) \exp\left\{\left\langle \boldsymbol{m}_i^{t+1}, \boldsymbol{h}_i\right\rangle - \frac{1}{2}\left\langle \boldsymbol{h}_i, \boldsymbol{Q}^{t+1}\boldsymbol{h}_i\right\rangle\right\}\right)
\end{aligned}
$$

where $\widetilde{\mathsf{F}}(\,\cdot\,;\,\cdot\,), \widetilde{\mathsf{G}}(\,\cdot\,;\,\cdot\,)$ are given in (D.2), (D.6) and

$$\boldsymbol{m}^{t+1} = \boldsymbol{X}^{\mathsf{T}}\widetilde{\mathsf{F}}\left(\tilde{\boldsymbol{m}}^t; \widetilde{\boldsymbol{Q}}^t\right), \qquad \boldsymbol{Q}^{t+1} = \frac{1}{d}\sum_{a=1}^{n} \widetilde{\mathsf{G}}\left(\tilde{\boldsymbol{m}}_a^t; \widetilde{\boldsymbol{Q}}^t\right). \tag{D.17}$$

Therefore, $q_i^{t+1}(\boldsymbol{h})$ has the form in (D.16) and the update formula for $\boldsymbol{m}^{t+1}$, $\boldsymbol{Q}^{t+1}$ are given in (D.17). Similarly, for $\tilde{q}^{t+1}(\boldsymbol{W})$ we have

$$\begin{aligned}
\tilde{q}^{t+1}(\boldsymbol{W}) &= \prod_{a=1}^{n} \tilde{q}_a^{t+1}(\boldsymbol{w}_a) \\
&\propto \exp\left\{\mathbb{E}_{q^{t+1}(\boldsymbol{H})} \log p(\boldsymbol{X}, \boldsymbol{H}, \boldsymbol{W})\right\} \\
&\propto \exp\left\{\mathbb{E}_{q^{t+1}(\boldsymbol{H})} \log p(\boldsymbol{W}, \boldsymbol{X}|\boldsymbol{H})\right\} \\
&= \tilde{q}_0(\boldsymbol{W}) \exp\left\{\mathbb{E}_{q^{t+1}(\boldsymbol{H})} \log p(\boldsymbol{X}|\boldsymbol{H}, \boldsymbol{W})\right\} \\
&\propto \tilde{q}_0(\boldsymbol{W}) \exp\left\{\mathbb{E}_{q^{t+1}(\boldsymbol{H})} \left[-\frac{d}{2}\left\|\boldsymbol{X} - \frac{\sqrt{\beta}}{d}\boldsymbol{W}\boldsymbol{H}^{\mathsf{T}}\right\|_F^2\right]\right\} \\
&\propto \tilde{q}_0(\boldsymbol{W}) \exp\left\{\mathbb{E}_{q^{t+1}(\boldsymbol{H})} \mathsf{Tr}\left(\sqrt{\beta}\boldsymbol{W}\boldsymbol{H}^{\mathsf{T}}\boldsymbol{X}^{\mathsf{T}} - \frac{\beta}{2d}\boldsymbol{W}\boldsymbol{H}^{\mathsf{T}}\boldsymbol{H}\boldsymbol{W}^{\mathsf{T}}\right)\right\}
\end{aligned}$$

Hence,

$$\begin{aligned}
\tilde{q}^{t+1}(\boldsymbol{W}) &\propto \tilde{q}_0(\boldsymbol{W}) \exp\left\{\mathbb{E}_{q^{t+1}(\boldsymbol{H})} \sum_{a=1}^{n}\left(\sqrt{\beta}\left\langle \boldsymbol{w}_a, \boldsymbol{x}_a\boldsymbol{H}\right\rangle - \frac{\beta}{2d}\left\langle \boldsymbol{w}_a, \boldsymbol{H}^{\mathsf{T}}\boldsymbol{H}\boldsymbol{w}_a\right\rangle\right)\right\} \\
&= \tilde{q}_0(\boldsymbol{W}) \exp\left\{\sum_{a=1}^{n}\left\langle \boldsymbol{w}_a, \boldsymbol{x}_a\mathsf{F}\left(\boldsymbol{m}^{t+1}; \boldsymbol{Q}^{t+1}\right)\right\rangle - \frac{1}{2d}\left\langle \boldsymbol{w}_a, \left(\sum_{i=1}^{d}\mathsf{G}\left(\boldsymbol{m}_i^{t+1}; \boldsymbol{Q}^{t+1}\right)\right)\boldsymbol{w}_a\right\rangle\right\} \\
&= \prod_{a=1}^{n}\left(\tilde{q}_0(\boldsymbol{w}_a) \exp\left\{\left\langle \boldsymbol{w}_a, \tilde{\boldsymbol{m}}_a^{t+1}\right\rangle - \frac{1}{2}\left\langle \boldsymbol{w}_a, \widetilde{\boldsymbol{Q}}^{t+1}\boldsymbol{w}_a\right\rangle\right\}\right)
\end{aligned}$$

where $\mathsf{F}(\,\cdot\,;\,\cdot\,), \mathsf{G}(\,\cdot\,;\,\cdot\,)$ are given in (D.1), (D.5) and

$$\tilde{\boldsymbol{m}}^{t+1} = \boldsymbol{X}\mathsf{F}\left(\boldsymbol{m}^{t+1}; \boldsymbol{Q}^{t+1}\right), \qquad \widetilde{\boldsymbol{Q}}^{t+1} = \frac{1}{d}\sum_{i=1}^{d}\mathsf{G}\left(\boldsymbol{m}_i^{t+1}; \boldsymbol{Q}^{t+1}\right). \tag{D.18}$$

Therefore, $\tilde{q}_a^{t+1}(\boldsymbol{w})$ has the form in (D.16) and the update formula for $\tilde{\boldsymbol{m}}^{t+1}$, $\widetilde{\boldsymbol{Q}}^{t+1}$ are given in (D.18).

D.3 Derivation of the variational free energy (3.9)

As already mentioned, naive mean field minimizes the KL divergence between a factorized distribution $\hat{q}(\boldsymbol{W}, \boldsymbol{H}) = \prod_{a=1}^{n}\tilde{q}(\boldsymbol{w}_a)\prod_{i=1}^{d}q(\boldsymbol{h}_i)$ and the real posterior $p(\boldsymbol{W}, \boldsymbol{H}|\boldsymbol{X})$. The KL divergence takes the form

$$\mathrm{KL}(\hat{q}(\,\cdot\,,\,\cdot\,)||p(\,\cdot\,,\,\cdot\,|\boldsymbol{X})) = \mathcal{F}(\hat{q}) + \log p(\boldsymbol{X}) + \frac{d}{2}\|\boldsymbol{X}\|_F^2\,, \tag{D.19}$$

where $\mathcal{F}(\hat{q})$ is the Gibbs free energy. In this appendix we derive an explicit form for $\mathcal{F}(\hat{q})$ when $\hat{q}$ is factorized. We have

$$\begin{aligned}
\mathcal{F}(\hat{q}) &= \mathbb{E}_{\hat{q}}[-\log p(\boldsymbol{W},\boldsymbol{H}|\boldsymbol{X})] + \mathbb{E}_{\hat{q}}[\log \hat{q}(\boldsymbol{W},\boldsymbol{H})] - \frac{d}{2}\|\boldsymbol{X}\|_F^2 \\
&= \mathbb{E}_{\hat{q}}[-\log p(\boldsymbol{W},\boldsymbol{H},\boldsymbol{X})] + \mathbb{E}_{\hat{q}}[\log \hat{q}(\boldsymbol{W},\boldsymbol{H})] - \frac{d}{2}\|\boldsymbol{X}\|_F^2 \\
&= \mathbb{E}_{\hat{q}}[-\log p(\boldsymbol{X}|\boldsymbol{W},\boldsymbol{H}) - \log p(\boldsymbol{W},\boldsymbol{H})] + \mathbb{E}_{\hat{q}}[\log \hat{q}(\boldsymbol{W},\boldsymbol{H})] - \frac{d}{2}\|\boldsymbol{X}\|_F^2 \\
&= \mathbb{E}_{\hat{q}}\left[\frac{d\|\boldsymbol{X} - \frac{\sqrt{\beta}}{d}\boldsymbol{W}\boldsymbol{H}^{\mathsf{T}}\|_F^2}{2} - \frac{d}{2}\|\boldsymbol{X}\|_F^2 - \log(p(\boldsymbol{W},\boldsymbol{H}))\right] + \mathbb{E}_{\hat{q}}[\log \hat{q}(\boldsymbol{W},\boldsymbol{H})] \\
&= \frac{d}{2}\mathbb{E}_{\hat{q}}\left[\|\boldsymbol{X} - \frac{\sqrt{\beta}}{d}\boldsymbol{W}\boldsymbol{H}^{\mathsf{T}}\|_F^2\right] - \frac{d}{2}\|\boldsymbol{X}\|_F^2 + \mathrm{KL}(\hat{q}(\,\cdot\,,\,\cdot\,)\|q_0(\,\cdot\,,\,\cdot\,))\,.
\end{aligned}$$

(The last term is the KL divergence between $\hat{q}$ and the prior.)

We can explicitly calculate each term. Let's denote by $\boldsymbol{r}_i, \boldsymbol{\Omega}_i$ the first and second moments of q_i and by $\tilde{\boldsymbol{r}}_a$, $\widetilde{\boldsymbol{Q}}_a$ the first and second moments of $\tilde{q}_a$:

$$\boldsymbol{r}_i = \int \boldsymbol{h}\; q_i(\mathrm{d}\boldsymbol{h})\,, \qquad \tilde{\boldsymbol{r}}_a = \int \boldsymbol{w}\; \tilde{q}_a(\mathrm{d}\boldsymbol{w})\,, \tag{D.20}$$

$$\boldsymbol{\Omega}_i = \int \boldsymbol{h}^{\otimes 2}\; q_i(\mathrm{d}\boldsymbol{h})\,, \qquad \tilde{\boldsymbol{\Omega}}_a = \int \boldsymbol{w}^{\otimes 2}\; \tilde{q}_a(\mathrm{d}\boldsymbol{w})\,. \tag{D.21}$$

We then have

$$\begin{aligned}
\frac{d}{2}\mathbb{E}_{\hat{q}}\|\boldsymbol{X} - \frac{\sqrt{\beta}}{d}\boldsymbol{W}\boldsymbol{H}^{\mathsf{T}}\|_F^2 - \frac{d}{2}\|\boldsymbol{X}\|_F^2 &= \frac{d}{2}\mathbb{E}_{\hat{q}}\left[\mathsf{Tr}\left(-\frac{2\sqrt{\beta}}{d}\boldsymbol{X}^{\mathsf{T}}\boldsymbol{W}\boldsymbol{H}^{\mathsf{T}}\right) + \mathsf{Tr}\left(\frac{\beta}{d^2}\boldsymbol{H}\boldsymbol{W}^{\mathsf{T}}\boldsymbol{W}\boldsymbol{H}^{\mathsf{T}}\right)\right] \\
&= -\sqrt{\beta}\,\mathsf{Tr}\left(\boldsymbol{X}^{\mathsf{T}}\mathbb{E}_{\hat{q}}[\boldsymbol{W}\boldsymbol{H}^{\mathsf{T}}]\right) + \frac{\beta}{2d}\mathsf{Tr}\left(\mathbb{E}_{\hat{q}}[\boldsymbol{H}\boldsymbol{W}^{\mathsf{T}}\boldsymbol{W}\boldsymbol{H}^{\mathsf{T}}]\right) \\
&= -\sqrt{\beta}\,\mathsf{Tr}\left(\boldsymbol{X}^{\mathsf{T}}\boldsymbol{r}\tilde{\boldsymbol{r}}^{\mathsf{T}}\right) + \frac{\beta}{2d}\sum_{i=1}^{d}\sum_{a=1}^{n}\langle\boldsymbol{\Omega}_i,\tilde{\boldsymbol{\Omega}}_a\rangle\,.
\end{aligned} \tag{D.22}$$

Since both $\hat{q}$ and q_0 have product form, their KL divergence is just a sum of KL divergences for each row of $\boldsymbol{W}$ and each row of $\boldsymbol{H}$:

$$\mathrm{KL}(\hat{q}(\,\cdot\,,\,\cdot\,)\|q_0(\,\cdot\,,\,\cdot\,))) = \sum_{i=1}^{d}\mathrm{KL}(q_i\|q_0) + \sum_{a=1}^{n}\mathrm{KL}(\tilde{q}_a\|q_0)\,. \tag{D.23}$$

Each of these terms is treated in the same manner: we minimize over q_i or $\tilde{q}_a$ subject to the moment constraints (D.20), and define

$$\psi_*(\boldsymbol{r}_i,\boldsymbol{\Omega}_i) = \min\left\{\mathrm{KL}(q_i\|q_0):\quad \int \boldsymbol{h}\; q_i(\mathrm{d}\boldsymbol{h}) = \boldsymbol{r}_i\,,\;\; \int \boldsymbol{h}^{\otimes 2}\; q_i(\mathrm{d}\boldsymbol{h}) = \boldsymbol{\Omega}_i\right\}\,, \tag{D.24}$$

$$\tilde{\psi}_*(\tilde{\boldsymbol{r}}_a,\tilde{\boldsymbol{\Omega}}_a) = \min\left\{\mathrm{KL}(\tilde{q}_a\|\tilde{q}_0):\quad \int \boldsymbol{w}\; \tilde{q}_a(\mathrm{d}\boldsymbol{w}) = \tilde{\boldsymbol{r}}_a\,,\;\; \int \boldsymbol{w}^{\otimes 2}\; \tilde{q}_a(\mathrm{d}\boldsymbol{w}) = \tilde{\boldsymbol{\Omega}}_a\right\}\,. \tag{D.25}$$

Standard duality between entropy and moment generating functions yields that ψ_*, $\tilde{\psi}_*$ are defined as per Eq. (3.11). We briefly recall the argument for the reader's convenience. Considering for instance $\tilde{\psi}_*(\tilde{\boldsymbol{r}},\tilde{\boldsymbol{\Omega}})$, we introduce the Lagrangian

$$\begin{aligned}\mathcal{L}(\tilde{q}_a,\tilde{\boldsymbol{m}}_a,\widetilde{\boldsymbol{Q}}_a) =& \mathrm{KL}(\tilde{q}_a\|\tilde{q}_0) + \langle\tilde{\boldsymbol{m}}_a,\tilde{\boldsymbol{r}}_a\rangle - \frac{1}{2}\langle\widetilde{\boldsymbol{Q}}_a,\tilde{\boldsymbol{\Omega}}_a\rangle\\ &- \int \exp\left\{\langle\tilde{\boldsymbol{m}}_a,\boldsymbol{w}\rangle - \frac{1}{2}\langle\boldsymbol{w},\widetilde{\boldsymbol{Q}}_a\boldsymbol{w}\rangle\right\}\tilde{q}_a(\mathrm{d}\boldsymbol{w})\,.\end{aligned}$$

This is minimized easily with respect to $\tilde{q}_a$. The minimum is achieved at the distribution (3.5), with

$$\min_{\tilde{q}_a}\mathcal{L}(\tilde{q}_a,\tilde{\boldsymbol{m}}_a,\widetilde{\boldsymbol{Q}}_a) = \langle\tilde{\boldsymbol{r}}_a,\tilde{\boldsymbol{m}}_a\rangle - \frac{1}{2}\langle\tilde{\boldsymbol{\Omega}}_a,\widetilde{\boldsymbol{Q}}_a\rangle - \tilde{\phi}(\tilde{\boldsymbol{m}}_a,\widetilde{\boldsymbol{Q}}_a)\,, \tag{D.26}$$

and the claim (3.11) follows by strong duality.

Putting together Eqs. (D.22), (D.23), and (D.24)-(D.25), we obtain the desired expression (3.9).

Using (3.11), we get the following expressions for the gradients of ψ_*

$$\frac{\partial\psi_*}{\partial\boldsymbol{r}}(\boldsymbol{r},\boldsymbol{\Omega}) = \boldsymbol{m}\,,\qquad \frac{\partial\psi_*}{\partial\boldsymbol{\Omega}}(\boldsymbol{r},\boldsymbol{\Omega}) = -\frac{1}{2}\boldsymbol{Q}\,, \tag{D.27}$$

and similarly for $\tilde{\psi}_*$ (where $\boldsymbol{m},\boldsymbol{Q}$ are related to $\boldsymbol{r},\boldsymbol{\Omega}$ via Eqs. (3.12), (3.12)). Hence, the gradients of $\mathcal{F}$ with respect to $\boldsymbol{r}_i$, $\boldsymbol{\Omega}_i$ read

$$\frac{\partial\mathcal{F}}{\partial\boldsymbol{r}_i}(\boldsymbol{r},\tilde{\boldsymbol{r}},\boldsymbol{\Omega},\tilde{\boldsymbol{\Omega}}) = -\sqrt{\beta}(\boldsymbol{X}^{\mathsf{T}}\tilde{\boldsymbol{r}})_{i,\cdot} + \boldsymbol{m}_i\,,\qquad \frac{\partial\mathcal{F}}{\partial\tilde{\boldsymbol{r}}_a}(\boldsymbol{r},\tilde{\boldsymbol{r}},\boldsymbol{\Omega},\tilde{\boldsymbol{\Omega}}) = -\sqrt{\beta}(\boldsymbol{X}\boldsymbol{r})_{a,\cdot} + \tilde{\boldsymbol{m}}_a\,, \tag{D.28}$$

$$\frac{\partial\mathcal{F}}{\partial\boldsymbol{\Omega}_i}(\boldsymbol{r},\tilde{\boldsymbol{r}},\boldsymbol{\Omega},\tilde{\boldsymbol{\Omega}}) = -\frac{1}{2}\boldsymbol{Q}_i + \frac{\beta}{2d}\sum_{a=1}^{n}\tilde{\boldsymbol{\Omega}}_a\,,\qquad \frac{\partial\mathcal{F}}{\partial\tilde{\boldsymbol{\Omega}}_a}(\boldsymbol{r},\tilde{\boldsymbol{r}},\boldsymbol{\Omega},\tilde{\boldsymbol{\Omega}}) = -\frac{1}{2}\boldsymbol{Q}_a + \frac{\beta}{2d}\sum_{i=1}^{d}\boldsymbol{\Omega}_i\,. \tag{D.29}$$

Notice that at stationarity points, we have $\boldsymbol{Q}_i = \boldsymbol{Q} = (\beta/d)\sum_{a=1}^{n}\tilde{\boldsymbol{\Omega}}_a$ independent of i.

D.4 Proof of Lemma 3.2

We start with some useful formulae.

Lemma D.1. *For $q \in \mathbb{R}$ define $\mathsf{E}(q)$ by*

$$\mathsf{E}(q;\nu) = \frac{\int w_1^2 \exp\{-q\|\boldsymbol{w}\|_2^2\}\, \tilde{q}_0(\mathrm{d}\boldsymbol{w})}{\int \exp\{-q\|\boldsymbol{w}\|_2^2\}\, \tilde{q}_0(\mathrm{d}\boldsymbol{w})}\,. \tag{D.30}$$

Then, we have

$$\mathsf{F}(\boldsymbol{y} = y\mathbf{1}_k; \boldsymbol{Q} = q_1\boldsymbol{I}_k + q_2\boldsymbol{J}_k) = \frac{\sqrt{\beta}\, y}{1+q_1+kq_2}\,\mathbf{1}_k\,, \tag{D.31}$$

$$\begin{aligned}\mathsf{G}(\boldsymbol{y} = y\mathbf{1}_k; \boldsymbol{Q} = q_1\boldsymbol{I}_k + q_2\boldsymbol{J}_k) = {}& \frac{\beta}{(1+q_1)}\boldsymbol{I}_k \\ &+ \beta\left\{\frac{y^2}{(1+q_1+kq_2)^2} - \frac{q_1}{(1+q_1)(1+q_1+kq_2)}\right\}\boldsymbol{J}_k\,,\end{aligned} \tag{D.32}$$

$$\widetilde{\mathsf{F}}(\tilde{\boldsymbol{y}} = \tilde{y}\mathbf{1}_k; \widetilde{\boldsymbol{Q}} = \tilde{q}_1\boldsymbol{I}_k + \tilde{q}_2\boldsymbol{J}_k) = \frac{\sqrt{\beta}}{k}\,\mathbf{1}_k\,, \tag{D.33}$$

$$\widetilde{\mathsf{G}}(\tilde{\boldsymbol{y}} = \tilde{y}\mathbf{1}_k; \widetilde{\boldsymbol{Q}} = \tilde{q}_1\boldsymbol{I}_k + \tilde{q}_2\boldsymbol{J}_k) = \beta\,\frac{k^2\mathsf{E}(\tilde{q}_1;\nu)-1}{k(k-1)}\,\boldsymbol{I}_k - \beta\,\frac{k\mathsf{E}(\tilde{q}_1;\nu)-1}{k(k-1)}\,\boldsymbol{J}_k\,. \tag{D.34}$$

In particular

$$\begin{aligned}\mathsf{F}(\boldsymbol{y} = y\mathbf{1}_k; \boldsymbol{Q} = q\boldsymbol{J}_k) &= \frac{\sqrt{\beta}\, y}{1+kq}\,\mathbf{1}_k\,,\\ \mathsf{G}(\boldsymbol{y} = y\mathbf{1}_k; \boldsymbol{Q} = q\boldsymbol{J}_k) &= \beta\,\boldsymbol{I}_k + \beta\,\frac{y^2}{(1+kq)^2}\,\boldsymbol{J}_k\,,\\ \widetilde{\mathsf{F}}(\tilde{\boldsymbol{y}} = \tilde{y}\mathbf{1}_k; \widetilde{\boldsymbol{Q}} = \tilde{q}\boldsymbol{J}_k) &= \frac{\sqrt{\beta}}{k}\,\mathbf{1}_k\,,\\ \widetilde{\mathsf{G}}(\tilde{\boldsymbol{y}} = \tilde{y}\mathbf{1}_k; \widetilde{\boldsymbol{Q}} = \tilde{q}\boldsymbol{J}_k) &= \frac{\beta}{k(k\nu+1)}\,\left(\boldsymbol{I}_k + \nu\boldsymbol{J}_k\right)\,.\end{aligned}$$

Proof. First note that

$$[(1+q_1)\,\boldsymbol{I}_k + q_2\boldsymbol{J}_k]^{-1} = \frac{1}{1+q_1}\boldsymbol{I}_k - \frac{q_2}{(1+q_1)(1+q_1+kq_2)}\boldsymbol{J}_k.$$

Hence, by (D.4) we have

$$\begin{aligned}\mathsf{F}(\boldsymbol{y} = y\mathbf{1}_k; \boldsymbol{Q} = q_1\boldsymbol{I}_k + q_2\boldsymbol{J}_k) &= \sqrt{\beta}y\,[(1+q_1)\,\boldsymbol{I}_k + q_2\boldsymbol{J}_k]^{-1}\,\mathbf{1}_k\\ &= \sqrt{\beta}y\left(\frac{1}{1+q_1}\boldsymbol{I}_k - \frac{q_2}{(1+q_1)(1+q_1+kq_2)}\boldsymbol{J}_k\right)\mathbf{1}_k\\ &= \sqrt{\beta}y\left(\frac{1}{1+q_1} - \frac{kq_2}{(1+q_1)(1+q_1+kq_2)}\right)\mathbf{1}_k\\ &= \frac{\sqrt{\beta}\, y}{1+q_1+kq_2}\,\mathbf{1}_k\,.\end{aligned}$$

Thus, by (D.7)

$$\begin{aligned}\mathsf{G}(\boldsymbol{y}=y\mathbf{1}_k;\boldsymbol{Q}=q_1\boldsymbol{I}_k+q_2\boldsymbol{J}_k)=&\frac{\beta\,y^2}{(1+q_1+kq_2)^2}\boldsymbol{J}_k\\&+\beta\left(\frac{1}{1+q_1}\boldsymbol{I}_k-\frac{q_2}{(1+q_1)(1+q_1+kq_2)}\boldsymbol{J}_k\right)\\=&\frac{\beta}{(1+q_1)}\,\boldsymbol{I}_k\\&+\beta\left\{\frac{y^2}{(1+q_1+kq_2)^2}-\frac{q_1}{(1+q_1)(1+q_1+kq_2)}\right\}\,\boldsymbol{J}_k\,.\end{aligned}$$

In addition, using (D.2), by symmetry, all entries of $\widetilde{\mathsf{F}}(\tilde{\boldsymbol{y}}=\tilde{y}\mathbf{1}_k;\widetilde{\boldsymbol{Q}}=\tilde{q}_1\boldsymbol{I}_k+\tilde{q}_2\boldsymbol{J}_k)$ are equal. Further,

$$\begin{aligned}\Big\langle\mathbf{1}_k\,,\widetilde{\mathsf{F}}(\tilde{\boldsymbol{y}}=\tilde{y}\mathbf{1}_k;\widetilde{\boldsymbol{Q}}=\tilde{q}_1\boldsymbol{I}_k+\tilde{q}_2\boldsymbol{J}_k)\Big\rangle&=\sqrt{\beta}\,\frac{\int\langle\mathbf{1}_k,\boldsymbol{w}\rangle\,\exp\{\langle\tilde{\boldsymbol{y}},\boldsymbol{w}\rangle-\langle\boldsymbol{w},\widetilde{\boldsymbol{Q}}\boldsymbol{w}\rangle/2\}\,\tilde{q}_0(\mathrm{d}\boldsymbol{w})}{\int\,\exp\{\langle\tilde{\boldsymbol{y}},\boldsymbol{w}\rangle-\langle\boldsymbol{w},\widetilde{\boldsymbol{Q}}\boldsymbol{w}\rangle/2\}\,\tilde{q}_0(\mathrm{d}\boldsymbol{w})}\\&=\sqrt{\beta}\frac{\int\exp\{\langle\tilde{\boldsymbol{y}},\boldsymbol{w}\rangle-\langle\boldsymbol{w},\widetilde{\boldsymbol{Q}}\boldsymbol{w}\rangle/2\}\tilde{q}_0(\mathrm{d}\boldsymbol{w})}{\int\exp\{\langle\tilde{\boldsymbol{y}},\boldsymbol{w}\rangle-\langle\boldsymbol{w},\widetilde{\boldsymbol{Q}}\boldsymbol{w}\rangle/2\}\tilde{q}_0(\mathrm{d}\boldsymbol{w})}\\&=\sqrt{\beta}.\end{aligned}$$

Therefore,

$$\widetilde{\mathsf{F}}(\tilde{\boldsymbol{y}}=\tilde{y}\mathbf{1}_k;\widetilde{\boldsymbol{Q}}=\tilde{q}_1\boldsymbol{I}_k+\tilde{q}_2\boldsymbol{J}_k)=\frac{\sqrt{\beta}}{k}\mathbf{1}_k.$$

Finally, again by symmetry, $\widetilde{\mathsf{G}}(\tilde{\boldsymbol{y}}=\tilde{y}\mathbf{1}_k;\widetilde{\boldsymbol{Q}}=\tilde{q}_1\boldsymbol{I}_k+\tilde{q}_2\boldsymbol{J}_k)$ has the same diagonal entries. Further, the off-diagonal entries of this matrix are equal. Thus, we have

$$\widetilde{\mathsf{G}}(\tilde{\boldsymbol{y}}=\tilde{y}\mathbf{1}_k;\widetilde{\boldsymbol{Q}}=\tilde{q}_1\boldsymbol{I}_k+\tilde{q}_2\boldsymbol{J}_k)=\Big(\widetilde{\mathsf{G}}_{11}-\widetilde{\mathsf{G}}_{12}\Big)\,\boldsymbol{I}_k+\widetilde{\mathsf{G}}_{12}\boldsymbol{J}_k.\tag{D.35}$$

Note that by (D.5), (D.30)

$$\begin{aligned}\widetilde{\mathsf{G}}_{1,1}&=\beta\,\frac{\int w_1^2\exp\{\tilde{y}\,\langle\boldsymbol{w},\mathbf{1}_k\rangle-\tilde{q}_1\|\boldsymbol{w}\|_2^2/2-\tilde{q}_2\,\langle\boldsymbol{w},\mathbf{1}_k\rangle^2/2\}\,\tilde{q}_0(\mathrm{d}\boldsymbol{w})}{\int\exp\{\tilde{y}\,\langle\boldsymbol{w},\mathbf{1}_k\rangle-\tilde{q}_1\|\boldsymbol{w}\|_2^2/2-\tilde{q}_2\,\langle\boldsymbol{w},\mathbf{1}_k\rangle^2/2\}\tilde{q}_0(\mathrm{d}\boldsymbol{w})}\\&=\beta\,\frac{\exp\{\tilde{y}-\tilde{q}_2/2\}\int w_1^2\exp\{-\tilde{q}_1\|\boldsymbol{w}\|_2^2/2\}\,\tilde{q}_0(\mathrm{d}\boldsymbol{w})}{\exp\{\tilde{y}-\tilde{q}_2/2\}\int\exp\{-\tilde{q}_1\|\boldsymbol{w}\|_2^2/2\}\tilde{q}_0(\mathrm{d}\boldsymbol{w})}=\beta\,\mathsf{E}(\tilde{q}_1;\nu).\end{aligned}$$

Further, by (D.5)

$$k\widetilde{\mathsf{G}}_{1,1} + k(k-1)\widetilde{\mathsf{G}}_{1,2} = \left\langle \widetilde{\mathsf{G}}(\tilde{\boldsymbol{y}} = \tilde{y}\mathbf{1}_k; \widetilde{\boldsymbol{Q}} = \tilde{q}_1 \boldsymbol{I}_k + \tilde{q}_2 \boldsymbol{J}_k), \boldsymbol{J}_k \right\rangle$$

$$= \beta \frac{\int \langle \boldsymbol{w}, \mathbf{1}_k \rangle^2 \exp\{\langle \tilde{\boldsymbol{y}}, \boldsymbol{w}\rangle - \langle \boldsymbol{w}, \widetilde{\boldsymbol{Q}}\boldsymbol{w}\rangle/2\}\, \tilde{q}_0(\mathrm{d}\boldsymbol{w})}{\int \exp\{\langle \tilde{\boldsymbol{y}}, \boldsymbol{w}\rangle - \langle \boldsymbol{w}, \widetilde{\boldsymbol{Q}}\boldsymbol{w}\rangle/2\} \tilde{q}_0(\mathrm{d}\boldsymbol{w})}, \tag{D.36}$$

$$= \beta \frac{\int \exp\{\langle \tilde{\boldsymbol{y}}, \boldsymbol{w}\rangle - \langle \boldsymbol{w}, \widetilde{\boldsymbol{Q}}\boldsymbol{w}\rangle/2\}\, \tilde{q}_0(\mathrm{d}\boldsymbol{w})}{\int \exp\{\langle \tilde{\boldsymbol{y}}, \boldsymbol{w}\rangle - \langle \boldsymbol{w}, \widetilde{\boldsymbol{Q}}\boldsymbol{w}\rangle/2\}\, \tilde{q}_0(\mathrm{d}\boldsymbol{w})} = \beta. \tag{D.37}$$

Therefore, by (D.37), (D.35), we get

$$\widetilde{\mathsf{G}}_{1,1} = \beta\, \mathsf{E}(\tilde{q}_1; \nu)\,, \qquad \widetilde{\mathsf{G}}_{1,2} = -\beta\, \frac{k\mathsf{E}(\tilde{q}_1; \nu) - 1}{k(k-1)}. \tag{D.38}$$

Hence,

$$\widetilde{\mathsf{G}}(\tilde{\boldsymbol{y}} = \tilde{y}\mathbf{1}_k; \widetilde{\boldsymbol{Q}} = \tilde{q}_1 \boldsymbol{I}_k + \tilde{q}_2 \boldsymbol{J}_k) = \beta\, \frac{k^2\mathsf{E}(\tilde{q}_1; \nu) - 1}{k(k-1)}\, \boldsymbol{I}_k - \beta\, \frac{k\mathsf{E}(\tilde{q}_1; \nu) - 1}{k(k-1)}\, \boldsymbol{J}_k. \tag{D.39}$$

In addition, note that

$$\mathsf{E}(0; \nu) = \int w_1^2 \tilde{q}_0(\mathrm{d}\boldsymbol{w}) = \frac{\nu + 1}{k(k\nu + 1)}. \tag{D.40}$$

Using this, and replacing $q_1, \tilde{q}_1 = 0$ in (D.31) - (D.34) will complete the proof. □

Proof of Lemma 3.2. Note that $q \geq 0$

$$k^2\mathsf{E}(q; \nu) = \frac{\int k^2 w_1^2 \exp\{-q\|\boldsymbol{w}\|_2^2\}\, \tilde{q}_0(\mathrm{d}\boldsymbol{w})}{\int \exp\{-q\|\boldsymbol{w}\|_2^2\}\, \tilde{q}_0(\mathrm{d}\boldsymbol{w})} = \frac{\int k\|\boldsymbol{w}\|_2^2 \exp\{-q\|\boldsymbol{w}\|_2^2\}\, \tilde{q}_0(\mathrm{d}\boldsymbol{w})}{\int \exp\{-q\|\boldsymbol{w}\|_2^2\}\, \tilde{q}_0(\mathrm{d}\boldsymbol{w})}$$

$$\geq \frac{\int \|\boldsymbol{w}\|_1^2 \exp\{-q\|\boldsymbol{w}\|_2^2\}\, \tilde{q}_0(\mathrm{d}\boldsymbol{w})}{\int \exp\{-q\|\boldsymbol{w}\|_2^2\}\, \tilde{q}_0(\mathrm{d}\boldsymbol{w})} = 1.$$

In addition, we have

$$\mathsf{E}(0; \nu) = \int w_1^2 \tilde{q}_0(\mathrm{d}\boldsymbol{w}) = \frac{\nu + 1}{k(k\nu + 1)}\,,$$

$$\lim_{q_1 \to \infty} \frac{k\beta\delta}{k-1} \left\{ \mathsf{E}\left(\frac{\beta}{1 + q_1}; \nu\right) - \frac{1}{k^2} \right\} = \frac{k\beta\delta}{k-1} \left\{ \mathsf{E}\left(0; \nu\right) - \frac{1}{k^2} \right\} = \frac{\beta\delta}{k(k\nu + 1)} < \infty.$$

Therefore, the right hand side of (3.14) is non-negative, continuous, bounded for $q_1^* \in [0, \infty)$. Hence, using intermediate value theorem, (3.14) has a solution in $[0, \infty)$.

Now we will check that equations (3.20) and (3.21) hold for $\boldsymbol{m}^{t+1} = \boldsymbol{m}^t = \boldsymbol{m}^*$, $\tilde{\boldsymbol{m}}^t = \tilde{\boldsymbol{m}}^*$, $\boldsymbol{Q}^t = \boldsymbol{Q}^{t+1} = \boldsymbol{Q}^*$, $\widetilde{\boldsymbol{Q}}^t = \widetilde{\boldsymbol{Q}}^*$. We start with the first equation in (3.20). Using Lemma D.1, we have

$$\widetilde{\mathsf{F}}(\tilde{\boldsymbol{m}}_a^*; \widetilde{\boldsymbol{Q}}^*) = \frac{\sqrt{\beta}}{k}\mathbf{1}_k.$$

Therefore,

$$\widetilde{\mathsf{F}}(\tilde{\boldsymbol{m}}^*; \widetilde{\boldsymbol{Q}}^*) = \frac{\sqrt{\beta}}{k}\mathbf{1}_n \otimes \mathbf{1}_k, \qquad \boldsymbol{X}^\mathsf{T}\widetilde{\mathsf{F}}(\tilde{\boldsymbol{m}}^*; \widetilde{\boldsymbol{Q}}^*) = \frac{\sqrt{\beta}}{k}\left(\boldsymbol{X}^\mathsf{T}\mathbf{1}_n\right) \otimes \mathbf{1}_k = \boldsymbol{m}^*.$$

Now we consider the first equation in (3.21). Using Lemma D.1, we have

$$\mathsf{F}(\boldsymbol{m}_i^*; \boldsymbol{Q}^*) = \frac{\beta}{k(1+q_1^*+kq_2^*)}\left\langle \boldsymbol{X}_{.,i}, \mathbf{1}_n \right\rangle \mathbf{1}_k.$$

Hence,

$$\begin{aligned}
\mathsf{F}(\boldsymbol{m}^*; \boldsymbol{Q}^*) &= \frac{\beta}{k(1+q_1^*+kq_2^*)}(\boldsymbol{X}^\mathsf{T}\mathbf{1}_n) \otimes \mathbf{1}_k, \\
\boldsymbol{X}\mathsf{F}(\boldsymbol{m}^*; \boldsymbol{Q}^*) &= \frac{\beta}{k(1+q_1^*+kq_2^*)}(\boldsymbol{X}\boldsymbol{X}^\mathsf{T}\mathbf{1}_n) \otimes \mathbf{1}_k = \tilde{\boldsymbol{m}}^*.
\end{aligned}$$

For the second equation in (3.20), note that using Lemma D.1, we have

$$\frac{1}{d}\sum_{a=1}^n \widetilde{\mathsf{G}}(\tilde{\boldsymbol{m}}_a^*; \widetilde{\boldsymbol{Q}}^*) = \delta\beta\left(\frac{k^2\mathsf{E}(\tilde{q}_1^*;\nu)-1}{k(k-1)}\boldsymbol{I}_k - \frac{k\mathsf{E}(\tilde{q}_1^*;\nu)-1}{k(k-1)}\boldsymbol{J}_k\right).$$

Note that using (3.14), (3.15)

$$\begin{aligned}
\frac{k^2\mathsf{E}(\tilde{q}_1^*;\nu)-1}{k(k-1)} &= \frac{1}{k(k-1)}\left[k^2\mathsf{E}\left(\frac{\beta}{1+q_1^*};\nu\right)-1\right] = \frac{q_1^*}{\delta\beta}, \\
\frac{-k\mathsf{E}(\tilde{q}_1^*;\nu)+1}{k(k-1)} &= \frac{-1}{k(k-1)}\left[k\mathsf{E}\left(\frac{\beta}{1+q_1^*};\nu\right)-1\right] \\
&= \frac{-1}{k(k-1)}\left[\frac{k-1}{\delta\beta}q_1^* + \frac{1-k}{k}\right] \\
&= \frac{1}{\delta\beta}\left(\frac{\beta\delta - kq_1^*}{k^2}\right) = \frac{q_2^*}{\delta\beta}.
\end{aligned}$$

Therefore,

$$\frac{1}{d}\sum_{a=1}^n \widetilde{\mathsf{G}}(\tilde{\boldsymbol{m}}_a^*; \widetilde{\boldsymbol{Q}}^*) = q_1^*\boldsymbol{I}_k + q_2^*\boldsymbol{J}_k = \boldsymbol{Q}^*. \tag{D.41}$$

Finally, we check the second equation in (3.21). Using Lemma D.1, we have

$$\mathsf{G}(\boldsymbol{m}_i^*;\boldsymbol{Q}^*) = \frac{\beta}{(1+q_1^*)}\boldsymbol{I}_k + \beta\left\{\frac{\langle \boldsymbol{X}_{\cdot,i},\boldsymbol{1}_n\rangle^2}{(1+q_1^*+kq_2^*)^2} - \frac{q_1^*}{(1+q_1^*)(1+q_1^*+kq_2^*)}\right\}\boldsymbol{J}_k. \tag{D.42}$$

Hence,

$$\begin{aligned}\frac{1}{d}\sum_{i=1}^{d}\mathsf{G}(\boldsymbol{m}_i^*;\boldsymbol{Q}^*) &= \frac{\beta}{(1+q_1^*)}\boldsymbol{I}_k + \beta\left\{\frac{\|\boldsymbol{X}^{\mathsf{T}}\boldsymbol{1}_n\|_2^2}{d(1+q_1^*+kq_2^*)^2} - \frac{q_1^*}{(1+q_1^*)(1+q_1^*+kq_2^*)}\right\}\boldsymbol{J}_k\\ &= \tilde{q}_1^*\boldsymbol{I}_k + \tilde{q}_2^*\boldsymbol{J}_k = \widetilde{\boldsymbol{Q}}^*,\end{aligned}$$

this completes the proof. □

D.5 Proof of Theorem 2

We will first prove that, if $L(\beta,k,\delta,\nu) > 1$, then the uninformative fixed point $(\boldsymbol{r}^*,\tilde{\boldsymbol{r}}^*,\boldsymbol{\Omega}^*,\tilde{\boldsymbol{\Omega}}^*)$ (or equivalently, its conjugate $(\boldsymbol{m}^*,\tilde{\boldsymbol{m}}^*,\boldsymbol{Q}^*,\widetilde{\boldsymbol{Q}}^*)$) is (with high probability) a saddle point of the naive mean field free energy (3.9). This implies immediately that the naive mean field iteration is unstable at that fixed point.

Note that the mapping $(\boldsymbol{r},\tilde{\boldsymbol{r}},\boldsymbol{\Omega},\tilde{\boldsymbol{\Omega}}) \to (\boldsymbol{r},\tilde{\boldsymbol{r}},\boldsymbol{Q},\widetilde{\boldsymbol{Q}})$ is a diffeomorphism (since the Jacobian is always invertible by strict convexity of ϕ, $\tilde{\phi}$). We define $\mathcal{F}_*$ to be the restriction of $\mathcal{F}$ to the submanifold defined by $\boldsymbol{Q} = \boldsymbol{Q}_*$, $\widetilde{\boldsymbol{Q}} = \widetilde{\boldsymbol{Q}}_*$. Explicitly, this can be written in terms of the partial Legendre transforms (we repeat the definition of Eq. (4.4) for the reader's convenience):

$$\psi(\boldsymbol{r},\boldsymbol{Q}) \equiv \sup_{\boldsymbol{m}}\left\{\langle\boldsymbol{r},\boldsymbol{m}\rangle - \phi(\boldsymbol{m},\boldsymbol{Q})\right\}, \quad \tilde{\psi}(\tilde{\boldsymbol{r}},\widetilde{\boldsymbol{Q}}) \equiv \sup_{\tilde{\boldsymbol{m}}}\left\{\langle\tilde{\boldsymbol{r}},\tilde{\boldsymbol{m}}\rangle - \tilde{\phi}(\tilde{\boldsymbol{m}},\widetilde{\boldsymbol{Q}})\right\}. \tag{D.43}$$

We then have

$$\begin{aligned}\mathcal{F}_*(\boldsymbol{r},\tilde{\boldsymbol{r}}) =& \sum_{i=1}^{d}\psi(\boldsymbol{r}_i,\boldsymbol{Q}_*) + \sum_{a=1}^{n}\tilde{\psi}(\tilde{\boldsymbol{r}}_a,\widetilde{\boldsymbol{Q}}_*) - \sqrt{\beta}\,\mathsf{Tr}\left(\boldsymbol{X}\boldsymbol{r}\tilde{\boldsymbol{r}}^{\mathsf{T}}\right)\\ &- \frac{d}{2}\langle\boldsymbol{Q}_*,\boldsymbol{\Omega}\rangle - \frac{n}{2}\langle\widetilde{\boldsymbol{Q}}_*,\tilde{\boldsymbol{\Omega}}\rangle + \frac{\beta n}{2}\langle\boldsymbol{\Omega},\tilde{\boldsymbol{\Omega}}\rangle, \end{aligned}\tag{D.44}$$

$$\boldsymbol{\Omega} \equiv \frac{1}{d\beta}\sum_{i=1}^{d}\mathsf{G}(\boldsymbol{m}_i;\boldsymbol{Q}^*), \qquad \tilde{\boldsymbol{\Omega}} \equiv \frac{1}{n\beta}\sum_{a=1}^{n}\widetilde{\mathsf{G}}(\tilde{\boldsymbol{m}}_a;\widetilde{\boldsymbol{Q}}^*), \tag{D.45}$$

$$\boldsymbol{r}_i \equiv \frac{1}{\sqrt{\beta}}\mathsf{F}(\boldsymbol{m}_i;\boldsymbol{Q}^*), \qquad \tilde{\boldsymbol{r}}_a \equiv \frac{1}{\sqrt{\beta}}\widetilde{\mathsf{F}}(\tilde{\boldsymbol{m}}_a;\widetilde{\boldsymbol{Q}}^*),\,. \tag{D.46}$$

In order to prove that $(\boldsymbol{r}_*,\tilde{\boldsymbol{r}}_*)$ is a saddle point of $\mathcal{F}$, it is sufficient to show that it is a saddle along a submanifold, and hence that the Hessian of $\mathcal{F}_*$ has a negative eigenvalue at $(\boldsymbol{r}_*,\tilde{\boldsymbol{r}}_*)$.

Next notice that

$$\mathcal{F}_*(\boldsymbol{r},\tilde{\boldsymbol{r}}) = \mathcal{G}_1(\boldsymbol{r},\tilde{\boldsymbol{r}}) + \mathcal{G}_2(\boldsymbol{r},\tilde{\boldsymbol{r}})\,,$$
$$\mathcal{G}_1(\boldsymbol{r},\tilde{\boldsymbol{r}}) \equiv \sum_{i=1}^d \psi(\boldsymbol{r}_i,\boldsymbol{Q}_*) + \sum_{a=1}^n \tilde{\psi}(\tilde{\boldsymbol{r}}_a,\widetilde{\boldsymbol{Q}}_*) - \sqrt{\beta}\mathsf{Tr}\left(\boldsymbol{X}\boldsymbol{r}\tilde{\boldsymbol{r}}^{\mathsf{T}}\right)\,,$$
$$\mathcal{G}_2(\boldsymbol{r},\tilde{\boldsymbol{r}}) \equiv -\frac{d}{2}\langle\boldsymbol{Q}_*,\boldsymbol{\Omega}\rangle - \frac{n}{2}\langle\widetilde{\boldsymbol{Q}}_*,\tilde{\boldsymbol{\Omega}}\rangle + \frac{\beta n}{2}\langle\boldsymbol{\Omega},\tilde{\boldsymbol{\Omega}}\rangle\,.$$

Consider deviations from the stationary point $\boldsymbol{r}_i = \boldsymbol{r}_i^* + \boldsymbol{\delta}_i$, $\tilde{\boldsymbol{r}}_a = \tilde{\boldsymbol{r}}_a^* + \tilde{\boldsymbol{\delta}}_a$. By Eqs. (D.45) and (D.46), we have (for some tensors $\boldsymbol{T},\tilde{\boldsymbol{T}} \in (\mathbb{R}^k)^{\otimes 3}$))

$$\boldsymbol{\Omega} = \boldsymbol{\Omega}^* + \frac{1}{d}\sum_{i=1}^d \boldsymbol{T}\boldsymbol{\delta}_i + \boldsymbol{\Delta}\,, \qquad \tilde{\boldsymbol{\Omega}} = \tilde{\boldsymbol{\Omega}}^* + \frac{1}{n}\sum_{a=1}^n \tilde{\boldsymbol{T}}\tilde{\boldsymbol{\delta}}_a + \widetilde{\boldsymbol{\Delta}}\,, \tag{D.47}$$

where $\boldsymbol{\Delta}$, $\widetilde{\boldsymbol{\Delta}}$ are of second order in $\boldsymbol{\delta},\tilde{\boldsymbol{\delta}}$. At the stationary point, by Eq. (D.29), we have $\boldsymbol{Q}^* = \beta\boldsymbol{\Omega}^*$, $\widetilde{\boldsymbol{Q}}^* = \beta\delta\tilde{\boldsymbol{\Omega}}^*$. Hence, substituting in $\mathcal{G}_2$, and letting $M_{ij} = \sum_{s,t} T_{st,i}\tilde{T}_{st,j}$, we obtain

$$\mathcal{G}_2(\boldsymbol{r},\tilde{\boldsymbol{r}}) = \mathcal{G}_2(\boldsymbol{r}_*,\tilde{\boldsymbol{r}}_*) + \frac{\beta}{2d}\sum_{i=1}^d\sum_{a=1}^n \langle\boldsymbol{\delta}_i,\boldsymbol{M}\tilde{\boldsymbol{\delta}}_a\rangle + o(\boldsymbol{\delta}^2) \tag{D.48}$$

Therefore, the Hessian $\nabla^2\boldsymbol{G}_2(\boldsymbol{r}_*,\tilde{\boldsymbol{r}}_*)$ has rank at most k.

Since $\psi(\,\cdot\,,\boldsymbol{Q}^*)$, $\tilde{\psi}(\,\cdot\,,\widetilde{\boldsymbol{Q}}^*)$ are Legendre transforms of $\phi(\,\cdot\,,\boldsymbol{Q}^*)$, $\tilde{\phi}(\,\cdot\,,\widetilde{\boldsymbol{Q}}^*)$, respectively, we have

$$\nabla^2_{\boldsymbol{r}\boldsymbol{r}}\psi(\boldsymbol{r},\boldsymbol{Q}^*) = \left(\nabla^2_{\boldsymbol{m}\boldsymbol{m}}\phi(\boldsymbol{m},\boldsymbol{Q}^*)\right)^{-1} = \boldsymbol{I}_k + \boldsymbol{Q}^*, \tag{D.49}$$
$$\nabla^2_{\tilde{\boldsymbol{r}}\tilde{\boldsymbol{r}}}\tilde{\psi}(\tilde{\boldsymbol{r}},\widetilde{\boldsymbol{Q}}^*) = \left(\nabla^2_{\tilde{\boldsymbol{m}}\tilde{\boldsymbol{m}}}\tilde{\phi}(\tilde{\boldsymbol{m}},\widetilde{\boldsymbol{Q}}^*)\right)^{-1} = \boldsymbol{D}^{-1} \tag{D.50}$$

where $\boldsymbol{D} \in \mathbb{R}^{k\times k}$ is as

$$D_{ij} = \frac{1}{\sqrt{\beta}}\frac{\partial\widetilde{\mathsf{F}}_i\left(\tilde{\boldsymbol{m}};\widetilde{\boldsymbol{Q}}\right)}{\partial\tilde{m}_j}\bigg|_{\tilde{\boldsymbol{m}}=0,\widetilde{\boldsymbol{Q}}=\widetilde{\boldsymbol{Q}}^*}. \tag{D.51}$$

Thus,

$$\begin{aligned}\boldsymbol{D} =& \frac{\left(\int \boldsymbol{w}^{\otimes 2}\exp\{-\tilde{q}_1^*\|\boldsymbol{w}\|_2^2/2\}\tilde{q}_0(\mathrm{d}\boldsymbol{w})\right)\left(\int \exp\{-\tilde{q}_1^*\|\boldsymbol{w}\|_2^2/2\}\tilde{q}_0(\mathrm{d}\boldsymbol{w})\right)}{\left(\int \exp\{-\tilde{q}_1^*\|\boldsymbol{w}\|_2^2/2\}\tilde{q}_0(\mathrm{d}\boldsymbol{w})\right)^2}\\ &- \frac{\left(\int \boldsymbol{w}\exp\{-\tilde{q}_1^*\|\boldsymbol{w}\|_2^2/2\}\tilde{q}_0(\mathrm{d}\boldsymbol{w})\right)^{\otimes 2}}{\left(\int \exp\{-\tilde{q}_1^*\|\boldsymbol{w}\|_2^2/2\}\tilde{q}_0(\mathrm{d}\boldsymbol{w})\right)^2}\\ =&\frac{\boldsymbol{Q}^*}{\delta\beta} - \frac{\boldsymbol{J}_k}{k^2}.\end{aligned}$$

Hence,

$$\nabla^2\mathcal{G}_1 = \begin{bmatrix} \boldsymbol{I}_d \otimes \left(\boldsymbol{I}_k + \widetilde{\boldsymbol{Q}}^*\right) & -\sqrt{\beta}\boldsymbol{X}^\mathsf{T} \otimes \boldsymbol{I}_k \\ -\sqrt{\beta}\boldsymbol{X} \otimes \boldsymbol{I}_k & \boldsymbol{I}_n \otimes \boldsymbol{D}^{-1} \end{bmatrix}. \tag{D.52}$$

Since $\boldsymbol{I}_k + \widetilde{\boldsymbol{Q}}^*$ is positive definite, $\nabla^2\mathcal{G} \succeq 0$ if and only if

$$\begin{aligned} &\boldsymbol{I}_n \otimes \boldsymbol{D}^{-1} \succeq \beta\left(\boldsymbol{X} \otimes \boldsymbol{I}_k\right)\left(\boldsymbol{I}_d \otimes \left(\boldsymbol{I}_k + \widetilde{\boldsymbol{Q}}^*\right)\right)^{-1}\left(\boldsymbol{X}^\mathsf{T} \otimes \boldsymbol{I}_k\right) \\ \iff\ &\boldsymbol{I}_n \otimes \boldsymbol{D}^{-1} \succeq \beta\left(\boldsymbol{X}\boldsymbol{X}^\mathsf{T}\right) \otimes \left(\boldsymbol{I}_k + \widetilde{\boldsymbol{Q}}^*\right)^{-1} \\ \iff\ &\boldsymbol{I}_n \otimes \boldsymbol{I}_k \succeq \beta\left(\boldsymbol{X}\boldsymbol{X}^\mathsf{T}\right) \otimes \left(\boldsymbol{I}_k + \widetilde{\boldsymbol{Q}}^*\right)^{-1}\boldsymbol{D}. \end{aligned}$$

Hence, $\nabla^2\mathcal{G}_1$ has a negative eigenvalue if and only if

$$\beta\lambda_{\max}\left(\boldsymbol{X}\boldsymbol{X}^\mathsf{T}\right)\lambda_{\max}\left(\left(\boldsymbol{I}_k + \widetilde{\boldsymbol{Q}}^*\right)^{-1}\boldsymbol{D}\right) > 1.$$

Further, by the same argument, if $\beta\lambda_\ell(\boldsymbol{X}\boldsymbol{X}^\mathsf{T})\lambda_{\max}((\boldsymbol{I}_k + \widetilde{\boldsymbol{Q}}^*)^{-1}\boldsymbol{D}) > 1$, then $\nabla^2\mathcal{G}_1$ has at least ℓ negative eigenvalues (recall that $\lambda_\ell(\boldsymbol{M})$ denotes the ℓ-th eigenvalue of $\boldsymbol{M}$ in decreasing order).

Note that

$$\begin{aligned} \left(\boldsymbol{I}_k + \boldsymbol{Q}^*\right)^{-1}\boldsymbol{D} &= \left(\frac{\boldsymbol{I}_k}{1+q_1^*} - \frac{q_2^*}{(1+q_1^*)(1+q_1^*+kq_2^*)}\boldsymbol{J}_k\right)\left(\frac{q_1^*}{\delta\beta}\boldsymbol{I}_k + \left(\frac{q_2^*}{\delta\beta} - \frac{1}{k^2}\right)\boldsymbol{J}_k\right) \\ &= \frac{1}{1+q_1^*}\left(\frac{q_1^*}{\delta\beta}\boldsymbol{I}_k + \left(\frac{q_2^*}{1+q_1^*+kq_2^*}\left(\frac{1}{\delta\beta} + \frac{1}{k}\right) - \frac{1}{k^2}\right)\boldsymbol{J}_k\right), \\ \mu(\beta,\delta) &\equiv \lambda_{\max}\left(\left(\boldsymbol{I}_k + \boldsymbol{Q}^*\right)^{-1}\boldsymbol{D}\right) \\ &= \frac{1}{1+q_1^*}\left(\frac{q_1^*}{\delta\beta} + k\left[\frac{q_2^*}{1+q_1^*+kq_2^*}\left(\frac{1}{\delta\beta} + \frac{1}{k}\right) - \frac{1}{k^2}\right]_+\right). \end{aligned}$$

where $q_1^*, \tilde{q}_1^*, q_2^*$ are given in (3.14), (3.15), (3.16). Further $\boldsymbol{X}\boldsymbol{X}^\mathsf{T}$ is a low-rank deformation of a Wishart matrix. Hence, for any fixed ℓ, we have, almost surely

$$\lim\inf_{n,d\to\infty} \lambda_\ell(\boldsymbol{X}\boldsymbol{X}^\mathsf{T}) \geq \left(1 + \frac{1}{\sqrt{\delta}}\right)^2.$$

Thus, if

$$L(\beta,\delta) = \beta\lambda_{\max}\left(1 + \frac{1}{\sqrt{\delta}}\right)^2\mu(\beta,\delta) > 1,$$

we have $\lambda_n(\nabla^2\mathcal{G}_1) \leq \cdots \leq \lambda_{n-\ell}(\nabla^2\mathcal{G}_1) < 0$ with high probability for any fixed ℓ.

As explained above, $\nabla^2\mathcal{G}_2$ has rank at most k. Therefore, by Cauchy's interlacing inequality, if $L(\beta, k, \delta, \nu) > 1$,

$$\lambda_{\min}\left(\nabla^2 \mathcal{F}_*\right) \le \lambda_{n+k}\left(\nabla^2 \mathcal{G}_1 + \nabla^2 \mathcal{G}_2\right) < 0.$$

Hence, for $L(\beta, \delta) > 1$, $\nabla^2 \mathcal{F}_*$ has a negative eigenvalue.

Note that the mapping $(\boldsymbol{r}, \tilde{\boldsymbol{r}}, \boldsymbol{\Omega}, \tilde{\boldsymbol{\Omega}}) \to (\boldsymbol{m}, \tilde{\boldsymbol{m}}, \boldsymbol{Q}, \widetilde{\boldsymbol{Q}})$ is a diffeomorphism, and therefore, uninformative fixed point $(\boldsymbol{m}^*, \tilde{\boldsymbol{m}}^*, \boldsymbol{Q}^*, \widetilde{\boldsymbol{Q}}^*)$ is a saddle also when we consider the free energy as a function of the parameters $(\boldsymbol{m}, \tilde{\boldsymbol{m}}, \boldsymbol{Q}, \widetilde{\boldsymbol{Q}})$. The claim that $(\boldsymbol{m}^*, \boldsymbol{Q}^*)$ is unstable under the naive mean field iteration follows immediately from the above, by using Lemma A.1, applied to $f(\boldsymbol{x}, \boldsymbol{y}) = \mathcal{F}(\boldsymbol{m}, \tilde{\boldsymbol{m}}, \boldsymbol{Q}, \widetilde{\boldsymbol{Q}})$, whereby $\boldsymbol{x} = (\boldsymbol{m}, \boldsymbol{Q})$, $\boldsymbol{y} = (\tilde{\boldsymbol{m}}, \widetilde{\boldsymbol{Q}})$.

E Naive Mean Field: Further numerical results

In this section we report on additional numerical simulations using the alternate minimization to minimize the naive mean field free energy. These results confirm the one presented in the main text in Section 3.5.

E.1 Credible intervals

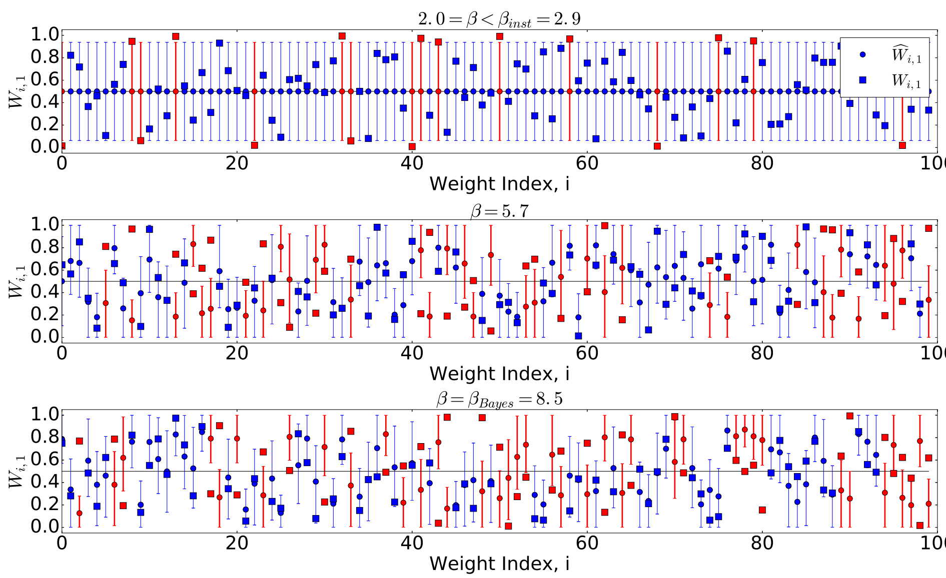

Figure 10: Bayesian credible intervals as computed by variational inference at nominal coverage level $1 - \alpha = 0.9$. Here $k = 2$, $d = 5000$, $n = 2500$ and we consider three values of β: $\beta \in \{2, 5.7, 8.5\}$ (for reference $\beta_{\text{inst}} \approx 2.9$, $\beta_{\text{Bayes}} \approx 8.5$. Circles correspond to the posterior mean, and squares to the actual weights. We use red for the coordinates on which the credible interval does not cover the actual value of $w_{i,1}$.

In Figures 10 and 11 we plot Bayesian credible intervals for the weights $w_{i,1}$ as computed within naive mean field, for $k = 2$, $d = 5000$. These simulations are analogous to the one

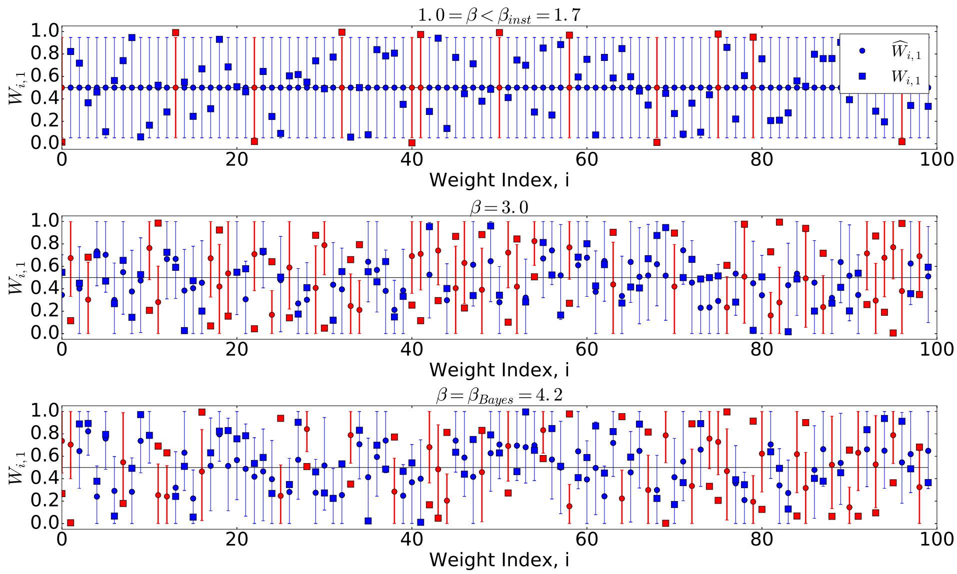

Figure 11: Bayesian credible intervals as computed by variational inference at nominal coverage level $1-\alpha = 0.9$. Here $k = 2$, $d = 5000$, $n = 10000$ and we consider three values of β: $\beta \in \{1, 3, 4.2\}$ (for reference $\beta_{\rm inst} \approx 1.7$, $\beta_{\rm Bayes} \approx 4.2$. Circles correspond to the posterior mean, and squares to the actual weights. We use red for the coordinates on which the credible interval does not cover the actual value of $w_{i,1}$.

reported in the main text in Figure 5, but we use $n = 2500$ ($\delta = 0.5$) in Figure 10 and $n = 10000$ ($\delta = 2$) in Figure 10.

The nominal coverage of these intervals is 0.9, but we obtain a smaller empirical coverage. For $\delta = 0.5$, the empirical coverage was 0.87 (for $\beta = 2 < \beta_{\rm inst}$), 0.61 (for $\beta = 5.7 \in (\beta_{\rm inst}, \beta_{\rm Bayes})$), and 0.64 (for $\beta = 8.5 \approx \beta_{\rm Bayes}$). For $\delta = 2$, the empirical coverage was 0.89 (for $\beta = 1 < \beta_{\rm inst}$), 0.69 (for $\beta = 3 \in (\beta_{\rm inst}, \beta_{\rm Bayes})$), and 0.65 (for $\beta = 4.2 \approx \beta_{\rm Bayes}$).

E.2 Results for $k = 3$ topics

In Figures 12 to 15 we report our results using alternating minimization to minimize the naive mean field free energy for $k = 3$.

In Figures 12, 13 we plot (respectively) the normalized distances $\mathsf{V}(\widehat{\boldsymbol{H}})$, $\mathsf{V}(\widehat{\boldsymbol{W}})$ from the uninformative subspaces $\{\boldsymbol{H} = \boldsymbol{v} \otimes \mathbf{1}_k : \boldsymbol{v} \in \mathbb{R}^d\}$ and $\{\boldsymbol{W} = \boldsymbol{v} \otimes \mathbf{1}_k : \boldsymbol{v} \in \mathbb{R}^d\}$. Data are consistent with the claim that this distance becomes significant when $\beta \gtrsim \beta_{\rm inst}(k, \nu, \delta)$.

In Figures 14, 15 we consider the correlation between the estimates $\widehat{\boldsymbol{H}}, \widehat{\boldsymbol{W}}$ and the true factorization $\boldsymbol{H}, \boldsymbol{W}$, and define a Binder cumulant as follows for $k \geq 3$. Let $\mathsf{C}_\eta(\boldsymbol{H}, \widehat{\boldsymbol{H}})$ be the $k \times k$ matrix with entries

$$\mathsf{C}_\eta(\boldsymbol{H}, \widehat{\boldsymbol{H}})_{i,j} = \frac{\langle (\widehat{\boldsymbol{H}}_\perp)_i + \eta \boldsymbol{g}, (\boldsymbol{H}_\perp)_j \rangle}{\|(\widehat{\boldsymbol{H}}_\perp)_i + \eta \boldsymbol{g}\|_2 \|(\boldsymbol{H}_\perp)_j\|_2} \tag{E.1}$$

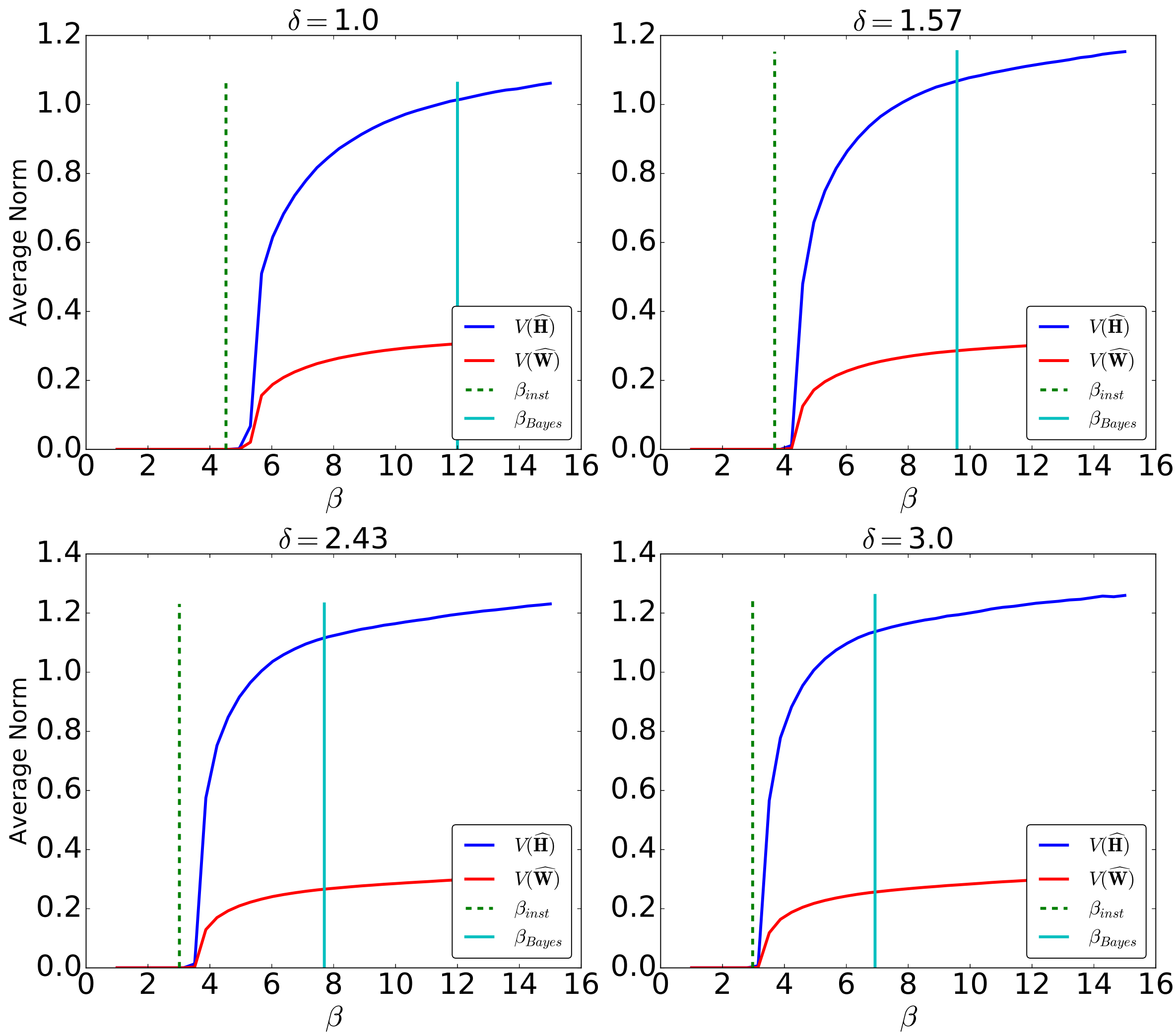

Figure 12: Normalized distances $\mathsf{V}(\widehat{\boldsymbol{H}})$, $\mathsf{V}(\widehat{\boldsymbol{W}})$ of the naive mean field estimates from the uninformative fixed point. Here $d = 1000$ and changed $n = d\delta$: each data point corresponds to an average over 400 random realizations.

We then define

$$\hat{\boldsymbol{R}} \equiv \frac{\widehat{\mathbb{E}}\left\{\sum_{i,j\le k} \mathsf{C}_\eta(\boldsymbol{H}, \widehat{\boldsymbol{H}})_{i,j}^4\right\}}{\widehat{\mathbb{E}}\left\{\sum_{i,j\le k} \mathsf{C}_\eta(\boldsymbol{H}, \widehat{\boldsymbol{H}})_{i,j}^2\right\}^2} \tag{E.2}$$

$$\mathsf{B}_{\boldsymbol{H}} \equiv \begin{cases} 6\left(\max\left\{\frac{2}{3} - \hat{\boldsymbol{R}}\right\} - \frac{1}{3}\right) & \text{if } \widehat{\mathbb{E}}\left\{\sum_{i,j\le k} \mathsf{C}_\eta(\boldsymbol{H}, \widehat{\boldsymbol{H}})_{i,j}^2\right\} > 0.01\,, \\ 0 & \text{otherwise.} \end{cases} \tag{E.3}$$

Here $\widehat{\mathbb{E}}$ denotes empirical average with respect to the sample and $\boldsymbol{g} \sim \mathsf{N}(0, \boldsymbol{I}_d)$. We set $\eta = 10^{-4}$. An analogous definition holds for $\mathsf{C}_\eta(\widehat{\boldsymbol{W}})$, $\mathsf{B}_\eta(\widehat{\boldsymbol{W}})$. In equation (E.2) we introduced a max thresholding step and a threshold on the denominator. These are added to ensure

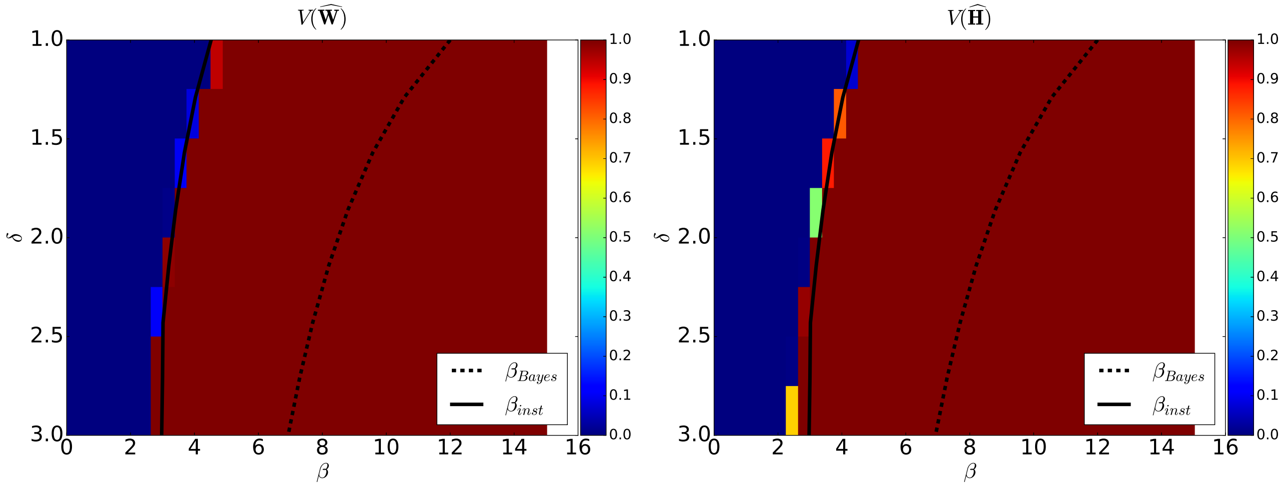

Figure 13: Empirical fraction of instances such that $\mathsf{V}(\widehat{\boldsymbol{W}}) \geq \varepsilon_0 = 5 \cdot 10^{-3}$ (left) or $\mathsf{V}(\widehat{\boldsymbol{H}}) \geq \varepsilon_0$ (right), where $\widehat{\boldsymbol{W}}, \widehat{\boldsymbol{H}}$ are the naive mean field estimate. Here $k = 3$, $d = 1000$ and, for each (δ, β) point on a grid, we used 400 random realizations to estimate the probability of $\mathsf{V}(\widehat{\boldsymbol{W}}) \geq \varepsilon_0$.

the stability of the fraction below the phase transition region where the denominator of $\hat{\boldsymbol{R}}$ vanishes.

Figures 14, 15 are consistent with the prediction that the correlation between the AMP estimates and the true factors $\boldsymbol{W}, \boldsymbol{H}$ starts to be non-negligible at the Bayes threshold.

F TAP free energy and approximate message passing

F.1 Heuristic derivation of the TAP free energy

Several heuristic approaches exist to construct the TAP free energy. Here we will derive the expression (4.3) of the TAP free energy for topic models as an approximation of the Bethe free energy for the same problem: we refer to [10, 7, 6] for background on the latter. Let us emphasize that our derivation will be only heuristic, since our rigorous results are obtained by analyzing the resulting expression $\mathcal{F}_{\text{TAP}}(\boldsymbol{r}, \tilde{\boldsymbol{r}})$ and do not require a rigorous justification of Eq. (4.3).

The posterior $p_{\boldsymbol{H},\boldsymbol{W}|\boldsymbol{X}}$ takes the form

$$p_{\boldsymbol{H},\boldsymbol{W}|\boldsymbol{X}}(\boldsymbol{H}, \boldsymbol{W}|\boldsymbol{X}) = \frac{1}{Z(\boldsymbol{X})} \prod_{(a,i)\in[n]\times[d]} \exp\left\{\sqrt{\beta} X_{ai} \langle \boldsymbol{w}_a, \boldsymbol{h}_i\rangle - \frac{\beta}{2d} \langle \boldsymbol{w}_a, \boldsymbol{h}_i\rangle^2\right\} \prod_{a=1}^{d} \tilde{q}_0(\boldsymbol{w}_a) \prod_{i=1}^{d} q_0(\boldsymbol{h}_i)\,.$$

This can be regarded as a pairwise graphical model whose underlying graph is the complete bipartite graph over vertex sets $[n]$ (associated to variables $\boldsymbol{w}_1, \dots \boldsymbol{w}_n$) and $[d]$ (associated to

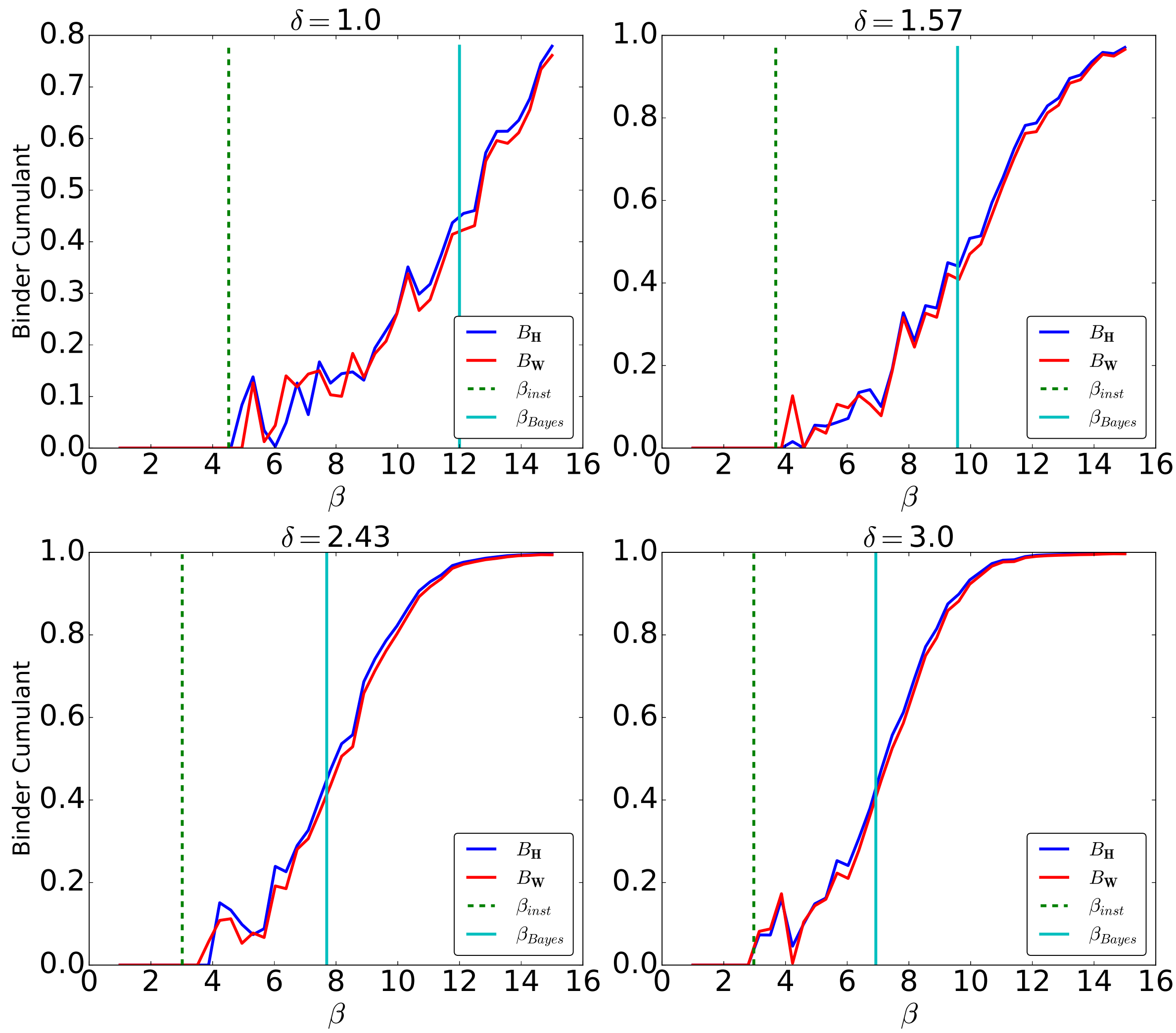

Figure 14: Binder cumulant for the correlation between the naive mean field estimates $\widehat{\boldsymbol{H}}$ and the true topics $\boldsymbol{H}$. Here we report results for $k = 3$, $d = 1000$ and $n = d\delta$, obtained by averaging over 400 realizations. Note that for $\beta < \beta_{\rm Bayes}(k, \nu, \delta)$, $\mathsf{B}_{\boldsymbol{H}}$ decreases with the dimensions, suggesting asymptotically vanishing correlations.

variables $\boldsymbol{h}_1, \dots \boldsymbol{h}_d$). The Bethe free energy $\mathcal{F}_{\rm Bethe}$ takes as input messages $\boldsymbol{q} \equiv (q_{i\to a})_{i\in[d],a\in[n]}$, $\widetilde{\boldsymbol{q}} = (\tilde{q}_{a\to i})_{i\in[d],a\in[n]}$. Messages are probability densities over the $\boldsymbol{h}_i$'s (for $q_{i\to a}$) or the $\boldsymbol{w}_a$'s (for $\tilde{q}_{a\to i}$), indexed by the directed edges in this graph (each pair (a, i), $a \in [n]$, $i \in [d]$ gives

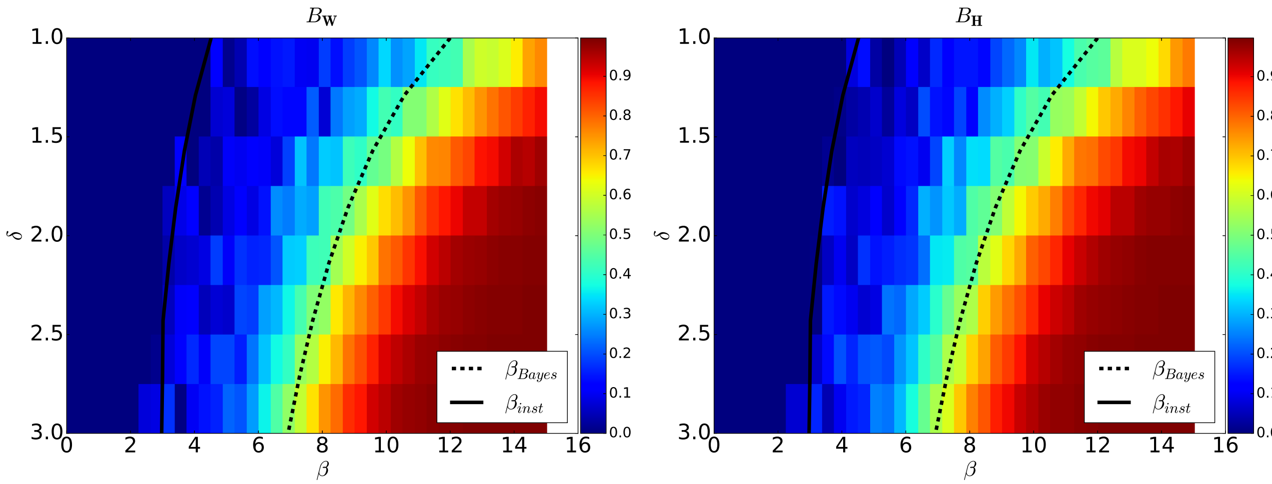

Figure 15: Binder cumulant for the correlation between the naive mean field estimates $\widehat{\boldsymbol{W}}$, $\widehat{\boldsymbol{H}}$ and the true weights and topics $\boldsymbol{W}$, $\boldsymbol{H}$. Here $k = 3$, $d = 1000$ and $n = d\delta$, and we averaged over 400 realizations.

rise to two directed edges). The free energy takes the form [7]

$$\mathcal{F}_{\text{Bethe}}(\boldsymbol{q}, \tilde{\boldsymbol{q}}) = \sum_{a=1}^{n}\sum_{i=1}^{d} \log Z_{ai} - \sum_{i=1}^{d} \log Z_i - \sum_{a=1}^{n} \log \tilde{Z}_a,$$

$$Z_i = \int \prod_{a=1}^{n} \exp\{\sqrt{\beta} X_{ai}\langle \boldsymbol{w}_a, \boldsymbol{h}_i\rangle - \frac{\beta}{2d}\langle \boldsymbol{w}_a, \boldsymbol{h}_i\rangle^2\} \mathrm{d}q_0(\boldsymbol{h}_i) \prod_{a=1}^{n} \mathrm{d}\tilde{q}_{a\to i}(\boldsymbol{w}_a)\,, \tag{F.1}$$

$$\tilde{Z}_a = \int \prod_{i=1}^{d} \exp\{\sqrt{\beta} X_{ai}\langle \boldsymbol{w}_a, \boldsymbol{h}_i\rangle - \frac{\beta}{2d}\langle \boldsymbol{w}_a, \boldsymbol{h}_i\rangle^2\} \mathrm{d}\tilde{q}_0(\boldsymbol{w}_a) \prod_{i=1}^{d} \mathrm{d}q_{i\to a}(\boldsymbol{h}_i)\,, \tag{F.2}$$

$$Z_{ai} = \int \exp\{\sqrt{\beta} X_{ai}\langle \boldsymbol{w}_a, \boldsymbol{h}_i\rangle - \frac{\beta}{2d}\langle \boldsymbol{w}_a, \boldsymbol{h}_i\rangle^2\}\ \mathrm{d}q_{i\to a}(\boldsymbol{h}_i)\, \mathrm{d}\tilde{q}_{a\to i}(\boldsymbol{w}_a)\,. \tag{F.3}$$

The stationarity conditions for $\mathcal{F}_{\text{Bethe}}(\boldsymbol{q}, \tilde{\boldsymbol{q}})$ correspond to the belief propagation fixed point equations

$$q_{i\to b}(\boldsymbol{h}_i) = \frac{1}{C_{i\to b}}\, q_0(\boldsymbol{h}_i) \prod_{a\in[n]\setminus b} \int \exp\{\sqrt{\beta} X_{ai}\langle \boldsymbol{w}_a, \boldsymbol{h}_i\rangle - \frac{\beta}{2d}\langle \boldsymbol{w}_a, \boldsymbol{h}_i\rangle^2\} \mathrm{d}\tilde{q}_{a\to i}(\boldsymbol{w}_a)\,, \tag{F.4}$$

$$\tilde{q}_{a\to j}(\boldsymbol{w}_i) = \frac{1}{\tilde{C}_{a\to j}}\, \tilde{q}_0(\boldsymbol{w}_i) \prod_{i\in[d]\setminus j} \int \exp\{\sqrt{\beta} X_{ai}\langle \boldsymbol{w}_a, \boldsymbol{h}_i\rangle - \frac{\beta}{2d}\langle \boldsymbol{w}_a, \boldsymbol{h}_i\rangle^2\} \mathrm{d}q_{i\to a}(\boldsymbol{h}_i)\,. \tag{F.5}$$

We define $\boldsymbol{f}_{i\to a} = \int \boldsymbol{h}_i \mathrm{d}q_{i\to a}(\boldsymbol{h}_i)$, $\tilde{\boldsymbol{f}}_{a\to i} = \int \boldsymbol{w}_a \mathrm{d}\tilde{q}_{a\to i}(\boldsymbol{w}_a)$, and $\boldsymbol{g}_{i\to a} = \int \boldsymbol{h}_i^{\otimes 2} \mathrm{d}q_{i\to a}(\boldsymbol{h}_i)$, $\tilde{\boldsymbol{g}}_{a\to i} =$

$\int \boldsymbol{w}_a^{\otimes 2} \mathrm{d}\tilde{q}_{a\to i}(\boldsymbol{w}_a)$. Since $X_{ai} = O(1/\sqrt{n})$, we have

$$
\begin{aligned}
&\prod_{i=1}^{d} \int \exp\{\sqrt{\beta} X_{ai} \langle \boldsymbol{w}_a, \boldsymbol{h}_i\rangle - \frac{\beta}{2d}\langle \boldsymbol{w}_a, \boldsymbol{h}_i\rangle^2\} \mathrm{d}q_{i\to a}(\boldsymbol{h}_i) \\
&= \prod_{i=1}^{d} \exp\left\{\sqrt{\beta} X_{ai} \langle \boldsymbol{f}_{i\to a}, \boldsymbol{w}_a\rangle - \frac{\beta}{2d}\langle \boldsymbol{f}_{i\to a}, \boldsymbol{w}_a\rangle^2 + \frac{\beta}{2}\Big(X_{ai}^2 - \frac{1}{d}\Big)\langle \boldsymbol{g}_{i\to a} - \boldsymbol{f}_{i\to a}^{\otimes 2}, \boldsymbol{w}_a^{\otimes 2}\rangle + O(n^{-3/2})\right\} \\
&= \exp\left\{\sum_{i=1}^{d} \sqrt{\beta} X_{ai} \langle \boldsymbol{f}_{i\to a}, \boldsymbol{w}_a\rangle - \frac{\beta}{2d}\sum_{i=1}^{d}\langle \boldsymbol{f}_{i\to a}, \boldsymbol{w}_a\rangle^2 + O(n^{-1/2})\right\}, \qquad \text{(F.6)}
\end{aligned}
$$

where in the last step we used the fact that $\mathbb{E}\{X_{ai}^2 - d^{-1}\} = O(n^{-3/2})$ and applied the central limit theorem.

Using the expression (F.6) in Eq. (F.2), and repeating a similar calculation for (F.1), we get

$$
\log Z_i = \phi\left(\sqrt{\beta}\sum_{a=1}^{n} X_{ai}\tilde{\boldsymbol{f}}_{a\to i}, \frac{\beta}{d}\sum_{a=1}^{n}\tilde{\boldsymbol{f}}_{a\to i}^{\otimes 2}\right) + O(n^{-1/2}), \qquad \text{(F.7)}
$$

$$
\log \tilde{Z}_a = \tilde{\phi}\left(\sqrt{\beta}\sum_{i=1}^{d} X_{ai}\boldsymbol{f}_{i\to a}, \frac{\beta}{d}\sum_{i=1}^{d}\boldsymbol{f}_{i\to a}^{\otimes 2}\right) + O(n^{-1/2}), \qquad \text{(F.8)}
$$

where the functions ϕ, $\tilde{\phi}$ are defined implicitly in Eq. (3.5).

We can similarly expand Z_{ai} for large n, d:

$$
\begin{aligned}
Z_{ai} &= 1 + \sqrt{\beta} X_{ai}\langle \tilde{\boldsymbol{f}}_{a\to i}, \boldsymbol{f}_{i\to a}\rangle + \frac{\beta}{2}\Big(X_{ai}^2 - \frac{1}{d}\Big)\langle \tilde{\boldsymbol{g}}_{a\to i}, \boldsymbol{g}_{i\to a}\rangle + O(n^{-3/2}) \\
&= \exp\left\{\sqrt{\beta} X_{ai}\langle \tilde{\boldsymbol{f}}_{a\to i}, \boldsymbol{f}_{i\to a}\rangle - \frac{\beta}{2} X_{ai}^2\langle \tilde{\boldsymbol{f}}_{a\to i}, \boldsymbol{f}_{i\to a}\rangle^2 + \frac{\beta}{2}\Big(X_{ai}^2 - \frac{1}{d}\Big)\langle \tilde{\boldsymbol{g}}_{a\to i}, \boldsymbol{g}_{i\to a}\rangle + O(n^{-3/2})\right\}.
\end{aligned}
$$

Therefore, using again the central limit theorem,

$$
\sum_{a\le n, i\le d}\log Z_{ai} = \sqrt{\beta}\sum_{a\le n, i\le d} X_{ai}\langle \tilde{\boldsymbol{f}}_{a\to i}, \boldsymbol{f}_{i\to a}\rangle - \frac{\beta}{2d}\sum_{a\le n, i\le d}\langle \tilde{\boldsymbol{f}}_{a\to i}, \boldsymbol{f}_{i\to a}\rangle^2 + O(n^{1/2}). \qquad \text{(F.9)}
$$

Putting together Eqs. (F.7), (F.8), and (F.9), we obtain

$$
\begin{aligned}
\mathcal{F}_{\text{Bethe}}(\boldsymbol{q}, \tilde{\boldsymbol{q}}) = &-\sum_{i=1}^{d}\phi\left(\sqrt{\beta}\sum_{a=1}^{n} X_{ai}\tilde{\boldsymbol{f}}_{a\to i}, \frac{\beta}{d}\sum_{a=1}^{n}\tilde{\boldsymbol{f}}_{a\to i}^{\otimes 2}\right) - \sum_{a=1}^{n}\tilde{\phi}\left(\sqrt{\beta}\sum_{i=1}^{d} X_{ai}\boldsymbol{f}_{i\to a}, \frac{\beta}{d}\sum_{i=1}^{d}\boldsymbol{f}_{i\to a}^{\otimes 2}\right) \\
&+ \sqrt{\beta}\sum_{a\le n, i\le d} X_{ai}\langle \tilde{\boldsymbol{f}}_{a\to i}, \boldsymbol{f}_{i\to a}\rangle - \frac{\beta}{2d}\sum_{a\le n, i\le d}\langle \tilde{\boldsymbol{f}}_{a\to i}, \boldsymbol{f}_{i\to a}\rangle^2 + O(n^{1/2}).
\end{aligned}
$$

Close to the solution of the stationarity conditions (F.4), (F.5), the message $\boldsymbol{f}_{i\to a}$ should be roughly independent of $a \in [n]$ and $\tilde{\boldsymbol{f}}_{a\to i}$ should be roughly independent of $i \in [d]$. Hence, we can approximate

$$-\frac{\beta}{2d}\sum_{a\le n,i\le d}\langle\tilde{\boldsymbol{f}}_{a\to i},\boldsymbol{f}_{i\to a}\rangle^2=-\frac{\beta}{2nd^2}\sum_{a\le n,i\le d}\sum_{b\le n,j\le d}\langle\tilde{\boldsymbol{f}}_{a\to j},\boldsymbol{f}_{i\to b}\rangle^2+o(n)\,. \tag{F.10}$$

In order to obtain the expression of Eq. (4.3) we add auxiliary variables $\boldsymbol{m}_i,\tilde{\boldsymbol{m}}_a\in\mathbb{R}^k$, and $\boldsymbol{Q}_i,\widetilde{\boldsymbol{Q}}_a\in\mathbb{R}^{k\times k}$, alongside Lagrange multipliers $\boldsymbol{r}_i$, $\tilde{\boldsymbol{r}}_a$, $\boldsymbol{\Omega}_i$, $\tilde{\boldsymbol{\Omega}}_a$ to enforce the constraints

$$\boldsymbol{m}_i=\sqrt{\beta}\sum_{a=1}^{n}X_{ai}\tilde{\boldsymbol{f}}_{a\to i}\,,\qquad \boldsymbol{Q}_i=\frac{\beta}{d}\sum_{a=1}^{n}\tilde{\boldsymbol{f}}_{a\to i}^{\otimes 2}\,, \tag{F.11}$$

$$\boldsymbol{m}_a=\sqrt{\beta}\sum_{i=1}^{d}X_{ai}\boldsymbol{f}_{i\to a}\,,\qquad \widetilde{\boldsymbol{Q}}_a=\frac{\beta}{d}\sum_{i=1}^{d}\boldsymbol{f}_{i\to a}^{\otimes 2}\,. \tag{F.12}$$

Denoting by $\boldsymbol{m}\in\mathbb{R}^{d\times k}$ the matrix whose i-th row is $\boldsymbol{m}_i$ (and analogously for $\tilde{\boldsymbol{m}}$, $\boldsymbol{f}$, $\tilde{\boldsymbol{f}}$ and the Lagrange multipliers $\boldsymbol{r}$, $\tilde{\boldsymbol{r}}$), and using Eq. (F.10) we obtain the Lagrangian (here all sums run over $a\in[n]$ and $i\in[d]$)

$$\begin{aligned}\mathcal{L}=&\langle\boldsymbol{r},\boldsymbol{m}\rangle-\sqrt{\beta}\sum_{a,i}X_{ai}\langle\boldsymbol{r}_i,\tilde{\boldsymbol{f}}_{a\to i}\rangle+\langle\tilde{\boldsymbol{r}},\tilde{\boldsymbol{m}}\rangle-\sqrt{\beta}\sum_{a,i}X_{ai}\langle\tilde{\boldsymbol{r}}_a,\boldsymbol{f}_{i\to a}\rangle+\sqrt{\beta}\sum_{a,i}X_{ai}\langle\tilde{\boldsymbol{f}}_{a\to i},\boldsymbol{f}_{i\to a}\rangle\\&+\frac{\sqrt{\beta}}{2n}\sum_{a,i}\langle\tilde{\boldsymbol{\Omega}}_a,\boldsymbol{f}_{i\to a}^{\otimes 2}\rangle-\frac{d}{2n\sqrt{\beta}}\sum_{a}\langle\tilde{\boldsymbol{\Omega}}_a,\widetilde{\boldsymbol{Q}}_a\rangle+\frac{\sqrt{\beta}}{2d}\sum_{a,i}\langle\tilde{\boldsymbol{\Omega}}_i,\tilde{\boldsymbol{f}}_{a\to i}^{\otimes 2}\rangle-\frac{d}{2d\sqrt{\beta}}\sum_{a}\langle\boldsymbol{\Omega}_i,\widetilde{\boldsymbol{Q}}_i\rangle\\&-\sum_{i}\phi(\boldsymbol{m}_i,\boldsymbol{Q}_i)-\sum_{a}\tilde{\phi}(\tilde{\boldsymbol{m}}_i,\widetilde{\boldsymbol{Q}}_i)-\frac{d}{2\beta dn}\sum_{a,i}\langle\widetilde{\boldsymbol{Q}}_a,\boldsymbol{Q}_i\rangle\,.\end{aligned} \tag{F.13}$$

We next minimize with respect to the message variables $(\boldsymbol{f}_{i\to a})$, $(\tilde{\boldsymbol{f}}_{a\to i})$. The first order stationarity conditions read

$$X_{ai}\tilde{\boldsymbol{f}}_{a\to i}=X_{ai}\tilde{\boldsymbol{r}}_a-\frac{1}{n}\tilde{\boldsymbol{\Omega}}_a\boldsymbol{f}_{i\to a}\,, \tag{F.14}$$

$$X_{ai}\boldsymbol{f}_{i\to a}=X_{ai}\boldsymbol{r}_i-\frac{1}{d}\boldsymbol{\Omega}_i\tilde{\boldsymbol{f}}_{a\to i}\,. \tag{F.15}$$

In particular these imply that $\tilde{\boldsymbol{f}}_{a\to i}=\tilde{\boldsymbol{r}}_a+O(1/\sqrt{n})$ and $\tilde{\boldsymbol{f}}_{a\to i}=\tilde{\boldsymbol{r}}_a+O(1/\sqrt{n})$. Multiplying the first of these equations by $\boldsymbol{f}_{i\to a}$ and the second by $\tilde{\boldsymbol{f}}_{a\to i}$, and summing over i,a we obtain

$$\begin{aligned}&\sum_{a,i}X_{ai}\langle\tilde{\boldsymbol{f}}_{a\to i},\boldsymbol{f}_{i\to a}\rangle\\&\quad=\frac{1}{2}\sum_{i,a}X_{ai}\Big(\langle\boldsymbol{f}_{i\to a},\tilde{\boldsymbol{r}}_a\rangle+\langle\tilde{\boldsymbol{f}}_{a\to i},\boldsymbol{r}_i\rangle\Big)-\frac{1}{2n}\sum_{i,a}\langle\tilde{\boldsymbol{\Omega}}_a,\boldsymbol{f}_{i\to a}^{\otimes 2}\rangle-\frac{1}{2d}\sum_{i,a}\langle\boldsymbol{\Omega}_i,\tilde{\boldsymbol{f}}_{a\to i}^{\otimes 2}\rangle\\&\quad=\frac{1}{2}\sum_{i,a}X_{ai}\Big(\langle\boldsymbol{f}_{i\to a},\tilde{\boldsymbol{r}}_a\rangle+\langle\tilde{\boldsymbol{f}}_{a\to i},\boldsymbol{r}_i\rangle\Big)-\frac{1}{2n}\sum_{i,a}\langle\tilde{\boldsymbol{\Omega}}_a,\boldsymbol{r}_i^{\otimes 2}\rangle-\frac{1}{2d}\sum_{i,a}\langle\boldsymbol{\Omega}_i,\tilde{\boldsymbol{r}}_a^{\otimes 2}\rangle+O(n^{1/2})\,.\end{aligned}$$

Further, multiplying Eqs. (F.14), (F.15) respectively by $\boldsymbol{r}_i$ and $\tilde{\boldsymbol{r}}_a$, we get

$$
\begin{aligned}
&\frac{1}{2}\sum_{i,a} X_{ai}\Big(\langle \boldsymbol{f}_{i\to a}, \tilde{\boldsymbol{r}}_a\rangle + \langle \tilde{\boldsymbol{f}}_{a\to i}, \boldsymbol{r}_i\rangle\Big)\\
&\quad= \sum_{a,i} X_{ai}\langle \tilde{\boldsymbol{r}}_a, \boldsymbol{r}_i\rangle - \frac{1}{2n}\sum_{a,i}\langle \boldsymbol{r}_i, \tilde{\boldsymbol{\Omega}}_a \boldsymbol{f}_{i\to a}\rangle - \frac{1}{2d}\sum_{a,i}\langle \tilde{\boldsymbol{r}}_a, \boldsymbol{\Omega}_i \tilde{\boldsymbol{f}}_{a\to i}\rangle\\
&\quad= \sum_{a,i} X_{ai}\langle \tilde{\boldsymbol{r}}_a, \boldsymbol{r}_i\rangle - \frac{1}{2n}\sum_{a,i}\langle \tilde{\boldsymbol{\Omega}}_a, \boldsymbol{r}_i^{\otimes 2}\rangle - \frac{1}{2d}\sum_{a,i}\langle \boldsymbol{\Omega}_i, \tilde{\boldsymbol{r}}_a^{\otimes 2}\rangle + O(n^{1/2})\,.
\end{aligned}
$$

Substituting the last two expressions in Eq. (F.13), we obtain

$$
\begin{aligned}
\mathcal{L} =&\langle \boldsymbol{r}, \boldsymbol{m}\rangle + \langle \tilde{\boldsymbol{r}}, \tilde{\boldsymbol{m}}\rangle - \sqrt{\beta}\langle \tilde{\boldsymbol{r}}, \boldsymbol{X}\boldsymbol{r}\rangle + \frac{\sqrt{\beta}}{2n}\sum_{a,i}\langle \tilde{\boldsymbol{\Omega}}_a, \boldsymbol{r}_i^{\otimes 2}\rangle + \frac{\sqrt{\beta}}{2d}\sum_{a,i}\langle \boldsymbol{\Omega}_i, \tilde{\boldsymbol{r}}_a^{\otimes 2}\rangle\\
&- \frac{d}{2n\sqrt{\beta}}\sum_{a}\langle \tilde{\boldsymbol{\Omega}}_a, \widetilde{\boldsymbol{Q}}_a\rangle - \frac{d}{2d\sqrt{\beta}}\sum_{i}\langle \boldsymbol{\Omega}_i, \boldsymbol{Q}_i\rangle - \sum_i \phi(\boldsymbol{m}_i, \boldsymbol{Q}_i) - \sum_a \tilde{\phi}(\tilde{\boldsymbol{m}}_i, \widetilde{\boldsymbol{Q}}_i)\\
&- \frac{d}{2\beta dn}\sum_{a,i}\langle \widetilde{\boldsymbol{Q}}_a, \boldsymbol{Q}_i\rangle + O(n^{1/2}).
\end{aligned}
\tag{F.16}
$$

Setting $\boldsymbol{Q}_i = \boldsymbol{Q}$ independent of i, $\widetilde{\boldsymbol{Q}}_a = \widetilde{\boldsymbol{Q}}$ independent of a, defining $\boldsymbol{\Omega} = d^{-1}\sum_{i=1}^d \boldsymbol{\Omega}_i$, $\tilde{\boldsymbol{\Omega}} = n^{-1}\sum_{a=1}^n \tilde{\boldsymbol{\Omega}}_a$, and neglecting $o(n)$ terms, we get

$$
\begin{aligned}
\tilde{\mathcal{F}}_{\text{TAP}} =&\frac{d}{2}\|\boldsymbol{X}\|_F - \sqrt{\beta}\mathsf{Tr}\left(\boldsymbol{X}\boldsymbol{r}\tilde{\boldsymbol{r}}^{\mathsf{T}}\right) + \mathsf{Tr}(\boldsymbol{r}^{\mathsf{T}}\boldsymbol{m}) + \mathsf{Tr}(\tilde{\boldsymbol{r}}^{\mathsf{T}}\tilde{\boldsymbol{m}}) - \frac{d}{2\sqrt{\beta}}\mathsf{Tr}(\boldsymbol{Q}\boldsymbol{\Omega}) - \frac{d}{2\sqrt{\beta}}\mathsf{Tr}(\widetilde{\boldsymbol{Q}}\tilde{\boldsymbol{\Omega}})\\
&- \sum_{a=1}^n \tilde{\phi}(\tilde{\boldsymbol{m}}_a, \widetilde{\boldsymbol{Q}}) - \sum_{i=1}^d \phi(\boldsymbol{m}_i, \boldsymbol{Q}) + \frac{\sqrt{\beta}}{2}\sum_{i=1}^d\langle \tilde{\boldsymbol{\Omega}}, \boldsymbol{r}_i^{\otimes 2}\rangle + \frac{\sqrt{\beta}}{2}\sum_{a=1}^n\langle \boldsymbol{\Omega}, \tilde{\boldsymbol{r}}_a^{\otimes 2}\rangle\\
&- \frac{d}{2\beta}\langle \boldsymbol{Q}, \widetilde{\boldsymbol{Q}}\rangle\,.
\end{aligned}
$$

Finally, the expression (4.3) is recovered by using the stationarity conditions with respect to $\boldsymbol{\Omega}$ and $\tilde{\boldsymbol{\Omega}}$, which imply $\boldsymbol{Q} = (\sqrt{\beta}/d)\sum_{a=1}^n \tilde{\boldsymbol{r}}_a^{\otimes 2}$ and $\widetilde{\boldsymbol{Q}} = (\sqrt{\beta}/d)\sum_{i=1}^d \boldsymbol{r}_i^{\otimes 2}$, and maximizing with respect to $\boldsymbol{m}$, $\tilde{\boldsymbol{m}}$.

F.2 Gradient of the TAP free energy

From the definition of the partial Legendre transforms $\psi(\boldsymbol{r}, \boldsymbol{Q})$, $\tilde{\psi}(\tilde{\boldsymbol{r}}, \widetilde{\boldsymbol{Q}})$, the following derivatives hold

$$
\frac{\partial \psi}{\partial \boldsymbol{r}}(\boldsymbol{r}, \boldsymbol{Q}) = \boldsymbol{m}(\boldsymbol{r}, \boldsymbol{Q})\,, \qquad \frac{\partial \psi}{\partial \boldsymbol{Q}}(\boldsymbol{r}, \boldsymbol{Q}) = -\frac{1}{2\beta}\mathsf{G}\Big(\boldsymbol{m}(\boldsymbol{r}, \boldsymbol{Q}), \boldsymbol{Q}\Big)\,, \tag{F.17}
$$

where $\boldsymbol{m}(\boldsymbol{r},\boldsymbol{Q})\in\mathbb{R}^k$ is the unique solution of

$$\boldsymbol{r}=\frac{1}{\sqrt{\beta}}\,\mathsf{F}(\boldsymbol{m};\boldsymbol{Q})\,. \tag{F.18}$$

Using these derivatives we can compute the gradient of the free energy

$$\begin{aligned}
\frac{\partial\mathcal{F}_{\rm TAP}}{\partial\boldsymbol{r}_i}(\boldsymbol{r},\tilde{\boldsymbol{r}}) &= -\sqrt{\beta}(\boldsymbol{X}^{\mathsf{T}}\tilde{\boldsymbol{r}})_i+\boldsymbol{m}_i-\frac{\beta}{d}\sum_{a=1}^{n}\langle\tilde{\boldsymbol{r}}_a,\boldsymbol{r}_i\rangle\,\tilde{\boldsymbol{r}}_a+\frac{1}{d}\sum_{a=1}^{n}\widetilde{\mathsf{G}}(\tilde{\boldsymbol{m}}_a,\widetilde{\boldsymbol{Q}})\boldsymbol{r}_i\\
&= -\sqrt{\beta}(\boldsymbol{X}^{\mathsf{T}}\tilde{\boldsymbol{r}})_i+\boldsymbol{m}_i+\sqrt{\beta}\tilde{\boldsymbol{\Omega}}\boldsymbol{r}_i\,, \qquad\qquad (\mathrm{F.19})\\
\frac{\partial\mathcal{F}_{\rm TAP}}{\partial\tilde{\boldsymbol{r}}_a}(\boldsymbol{r},\tilde{\boldsymbol{r}}) &= -\sqrt{\beta}(\boldsymbol{X}\boldsymbol{r})_a+\tilde{\boldsymbol{m}}_a-\frac{\beta}{d}\sum_{i=1}^{d}\langle\tilde{\boldsymbol{r}}_a,\boldsymbol{r}_i\rangle\,\boldsymbol{r}_i+\frac{1}{d}\sum_{i=1}^{d}\mathsf{G}(\boldsymbol{m}_i,\boldsymbol{Q})\tilde{\boldsymbol{r}}_a\\
&= -\sqrt{\beta}(\boldsymbol{X}\boldsymbol{r})_a+\tilde{\boldsymbol{m}}_a+\sqrt{\beta}\boldsymbol{\Omega}\tilde{\boldsymbol{r}}_a\,, \qquad\qquad (\mathrm{F.20})
\end{aligned}$$

where $\boldsymbol{m}_i=\boldsymbol{m}(\boldsymbol{r}_i,(\beta/d)\sum_{a\le n}\tilde{\boldsymbol{r}}_a^{\otimes 2})$, $\tilde{\boldsymbol{m}}_a=\tilde{\boldsymbol{m}}(\tilde{\boldsymbol{r}}_a,(\beta/d)\sum_{i\le d}\boldsymbol{r}_i^{\otimes 2})$, are defined as above, $\boldsymbol{Q}=(\beta/d)\sum_{a\le n}\tilde{\boldsymbol{r}}_a^{\otimes 2}$, $\boldsymbol{Q}=(\beta/d)\sum_{i\le d}\tilde{\boldsymbol{r}}_i^{\otimes 2}$, and

$$\boldsymbol{\Omega}=\frac{1}{d\sqrt{\beta}}\sum_{i=1}^{d}\Big\{\mathsf{G}(\boldsymbol{m}_i,\boldsymbol{Q})-\mathsf{F}(\boldsymbol{m}_i,\boldsymbol{Q})^{\otimes 2}\Big\}\,, \tag{F.21}$$

$$\tilde{\boldsymbol{\Omega}}=\frac{1}{d\sqrt{\beta}}\sum_{a=1}^{n}\Big\{\widetilde{\mathsf{G}}(\tilde{\boldsymbol{m}}_a,\widetilde{\boldsymbol{Q}})-\widetilde{\mathsf{F}}(\boldsymbol{m}_a,\widetilde{\boldsymbol{Q}})^{\otimes 2}\Big\}\,. \tag{F.22}$$

Remark F.1. We can express $\boldsymbol{r}$, $\tilde{\boldsymbol{r}}$ in terms of $\boldsymbol{m}$, $\tilde{\boldsymbol{m}}$ in Eqs. (F.19), (F.20) by using Eq. (F.18)

$$\boldsymbol{m}=\boldsymbol{X}^{\mathsf{T}}\,\widetilde{\mathsf{F}}(\tilde{\boldsymbol{m}};\widetilde{\boldsymbol{Q}})-\mathsf{F}(\boldsymbol{m};\boldsymbol{Q})\tilde{\boldsymbol{\Omega}}\,,\qquad \tilde{\boldsymbol{m}}=\boldsymbol{X}\,\mathsf{F}(\boldsymbol{m};\boldsymbol{Q})-\widetilde{\mathsf{F}}(\tilde{\boldsymbol{m}};\widetilde{\boldsymbol{Q}})\boldsymbol{\Omega}\,, \tag{F.23}$$

$$\boldsymbol{Q}=\frac{1}{d}\sum_{a=1}^{n}\widetilde{\mathsf{F}}(\tilde{\boldsymbol{m}}_a;\widetilde{\boldsymbol{Q}})^{\otimes 2}\,,\qquad \widetilde{\boldsymbol{Q}}=\frac{1}{d}\sum_{i=1}^{d}\mathsf{F}(\boldsymbol{m}_i;\boldsymbol{Q})^{\otimes 2}\,. \tag{F.24}$$

These coincide with the fixed point of the AMP algorithm in Section 4.2.

F.3 Uninformative critical point: Proof of Lemma 4.1

Consider the stationarity conditions (F.23) and (F.24), together with the definitions of Eqs. (4.11), (4.12). Since these are invariant under permutations of the topics, they admit a solution of the form $\boldsymbol{m}=\boldsymbol{v}\boldsymbol{1}_k^{\mathsf{T}}$, $\tilde{\boldsymbol{m}}=\tilde{\boldsymbol{v}}\boldsymbol{1}_k^{\mathsf{T}}$, $\boldsymbol{Q}=q_0\boldsymbol{J}_k+q_0'\boldsymbol{I}_k$, $\widetilde{\boldsymbol{Q}}=\tilde{q}_0\boldsymbol{J}_k+\tilde{q}_0'\boldsymbol{I}_k$. Using Eq. (F.24) and Lemma D.1, Eqs. (D.31), (D.33), we get $q_0'=\tilde{q}_0'=0$.

Substituting this in Eqs. (4.11), (4.12), and using again Lemma D.1, we get

$$\boldsymbol{\Omega}=\sqrt{\beta}\,\boldsymbol{I}_k\,,\qquad \tilde{\boldsymbol{\Omega}}=\frac{\sqrt{\beta}\delta}{k(k\nu+1)}\boldsymbol{P}_{\perp}\,, \tag{F.25}$$

where we recall that $\boldsymbol{P}_{\perp}=\boldsymbol{I}_k-\boldsymbol{1}_k\boldsymbol{1}_k/k$. Substituting these in Eq. (F.23), we obtained that this is satisfied provided $\boldsymbol{v}, \tilde{\boldsymbol{v}}$ are given as in Eqs. (4.13), (4.14). Finally, q_0, $\tilde{q}_0$ are fixed by substituting in Eq. (F.24).

G State evolution analysis

G.1 State evolution equations

Note that there is an alternative way to express the state evolution recursion in Eqs. (4.17), (4.18). Given a probability measure p on $\mathbb{R}^k$ and a matrix $\boldsymbol{M} \succeq 0$, $\boldsymbol{M} \in \mathbb{R}^{k\times k}$, we define the minimum mean square error

$$\mathsf{mmse}(\boldsymbol{M}; p) \equiv \inf_{\hat{\boldsymbol{x}}(\,\cdot\,)} \mathbb{E}\Big\{[\boldsymbol{x} - \hat{\boldsymbol{x}}(\boldsymbol{y})][\boldsymbol{x} - \hat{\boldsymbol{x}}(\boldsymbol{y})]^{\mathsf{T}}\Big\}\,, \tag{G.1}$$

where the expectation is with respect to $\boldsymbol{x} \sim p(\,\cdot\,)$ and $\boldsymbol{y} = \boldsymbol{M}^{1/2}\boldsymbol{x} + \boldsymbol{z}$ for $\boldsymbol{z} \sim \mathsf{N}(0, \boldsymbol{I}_k)$. The infimum is understood in the positive semidefinite order, and it is achieved by $\hat{\boldsymbol{x}}(\boldsymbol{y}) = \mathbb{E}\{\boldsymbol{x}|\boldsymbol{y}\}$. We then rewrite Eqs. (4.17), (4.18) as

$$\boldsymbol{M}_{t+1} = \beta\delta \left\{\mathsf{mmse}(0; \tilde{q}_0) - \mathsf{mmse}(\widetilde{\boldsymbol{M}}_t; \tilde{q}_0)\right\}, \tag{G.2}$$

$$\widetilde{\boldsymbol{M}}_t = \beta \left\{\mathsf{mmse}(0; q_0) - \mathsf{mmse}(\boldsymbol{M}_t; q_0)\right\}. \tag{G.3}$$

G.2 Uninformative fixed point

Lemma G.1. *The state evolution recursion in* (4.17), (4.18) *admit uninformative fixed point of the form*

$$\begin{aligned} \widetilde{\boldsymbol{M}}^* &= \rho_0 \boldsymbol{J}_k, \qquad \rho_0 = \frac{\delta\beta^2}{k\delta\beta + k^2}, \\ \boldsymbol{M}^* &= \frac{\delta\beta}{k^2}\boldsymbol{J}_k. \end{aligned} \tag{G.4}$$

Proof. First note that for this value of $\widetilde{\boldsymbol{M}}^*$, $\widetilde{\boldsymbol{M}}^*\boldsymbol{w} + \widetilde{\boldsymbol{M}}^{*^{1/2}}\boldsymbol{z} = y\boldsymbol{1}_k$ for some (random) y. Hence, using Eq. (D.33)

$$\delta\, \mathbb{E}\Big\{\widetilde{\mathsf{F}}(\widetilde{\boldsymbol{M}}^*\boldsymbol{w} + \widetilde{\boldsymbol{M}}^{*^{1/2}}\boldsymbol{z}; \widetilde{\boldsymbol{M}}^*)^{\otimes 2}\Big\} = \frac{\delta\beta}{k^2}\boldsymbol{J}_k = \boldsymbol{M}^*.$$

In addition, using the explicit form (D.4)

$$
\begin{aligned}
\mathbb{E}\Big\{\mathsf{F}(\boldsymbol{M}^*\boldsymbol{h}+\boldsymbol{M}^{*1/2}\boldsymbol{z};\boldsymbol{M}^*)^{\otimes 2}\Big\} &= \beta(\boldsymbol{I}_k+\boldsymbol{M}^*)^{-1}\boldsymbol{M}^* \\
&= \frac{\beta^2\delta}{k^2}\left(\boldsymbol{I}_k+\frac{\delta\beta}{k^2}\boldsymbol{J}_k\right)^{-1}\boldsymbol{J}_k \\
&= \rho_0\boldsymbol{J}_k = \widetilde{\boldsymbol{M}}^*.
\end{aligned}
$$

Hence, the pair $\boldsymbol{M}^*, \widetilde{\boldsymbol{M}}^*$ in (G.4) is a fixed point for the iterations in (4.17), (4.18). □

G.3 Stability of state evolution and proof of Theorem 4

The following theorem characterizes the region of parameters in which the uninformative fixed point of the state evolution iterations in Lemma G.1 is stable.

Theorem 6. *Consider the state evolution equations in* (4.17), (4.18). *The uninformative symmetric fixed point of these equations is stable if and only if*

$$
\beta < \beta_{\rm spect} = \frac{k(k\nu+1)}{\sqrt{\delta}}. \tag{G.5}
$$

Proof. We linearize Eqs. (4.17), (4.18) around the fixed point in (G.4) by setting $\boldsymbol{M}_t = \boldsymbol{M}_* + \boldsymbol{\Delta}_t$, $\widetilde{\boldsymbol{M}}_t = \widetilde{\boldsymbol{M}}_* + \widetilde{\boldsymbol{\Delta}}_t$ and expanding Eqs. (4.17), (4.18) to first order in $\boldsymbol{\Delta}, \widetilde{\boldsymbol{\Delta}}_t$. First note that Eq. (4.18) takes the explicit form

$$
\widetilde{\boldsymbol{M}}_t = \beta(\boldsymbol{I}_k+\boldsymbol{M}_t)^{-1}\boldsymbol{M}_t\,. \tag{G.6}
$$

Hence, expanding to linear order we get

$$
\widetilde{\boldsymbol{\Delta}}_t = \beta\left(\boldsymbol{I}_k+\frac{\delta\beta}{k^2}\boldsymbol{J}_k\right)^{-1}\boldsymbol{\Delta}_t\left(\boldsymbol{I}_k+\frac{\delta\beta}{k^2}\boldsymbol{J}_k\right)^{-1} + o(\boldsymbol{\Delta}_t)\,. \tag{G.7}
$$

In the following, we shall decompose $\boldsymbol{\Delta}_t$ and $\widetilde{\boldsymbol{\Delta}}_t$ in the components along $\mathbf{1}_k$ and the ones orthogonal

$$
\begin{aligned}
\boldsymbol{\Delta}_t &= \delta_t\,\boldsymbol{P} + \boldsymbol{\Delta}_t^{(1)} + \boldsymbol{\Delta}_t^{(2)}\,,\\
\boldsymbol{\Delta}_t^{(1)} &= \boldsymbol{P}\boldsymbol{\Delta}_t\boldsymbol{P}_\perp + \boldsymbol{P}_\perp\boldsymbol{\Delta}_t\boldsymbol{P}\,,\\
\boldsymbol{\Delta}_t^{(2)} &= \boldsymbol{P}_\perp\boldsymbol{\Delta}_t\boldsymbol{P}_\perp\,,
\end{aligned} \tag{G.8}
$$

and similarly for $\widetilde{\boldsymbol{\Delta}}_t$. Note that the linearization (G.7) preserves these subspaces

$$\tilde{\delta}_t = \beta \left(1 + \frac{\delta\beta}{k}\right)^{-2} \delta_t + o(\boldsymbol{\Delta}_t)\,, \tag{G.9}$$

$$\widetilde{\boldsymbol{\Delta}}_t^{(1)} = \beta \left(1 + \frac{\delta\beta}{k}\right)^{-1} \boldsymbol{\Delta}_t^{(1)} + o(\boldsymbol{\Delta}_t)\,, \tag{G.10}$$

$$\widetilde{\boldsymbol{\Delta}}_t^{(2)} = \beta\, \boldsymbol{\Delta}_t^{(2)} + o(\boldsymbol{\Delta}_t)\,. \tag{G.11}$$

Next we consider Eq. (4.17). We compute the value of

$$\begin{aligned} f_{\boldsymbol{w},\boldsymbol{z}} &= \widetilde{\mathsf{F}}(\widetilde{\boldsymbol{M}}_t \boldsymbol{w} + \widetilde{\boldsymbol{M}}_t^{1/2} \boldsymbol{z}; \widetilde{\boldsymbol{M}}_t) \\ &= \sqrt{\beta} \frac{\int \boldsymbol{w}_1 \exp\left\{\left\langle \widetilde{\boldsymbol{M}}_t \boldsymbol{w} + \widetilde{\boldsymbol{M}}_t^{1/2} \boldsymbol{z}, \boldsymbol{w}_1 \right\rangle - \frac{1}{2}\left\langle \boldsymbol{w}_1, \widetilde{\boldsymbol{M}}_t \boldsymbol{w}_1 \right\rangle\right\} \tilde{q}_0(\mathrm{d}\boldsymbol{w}_1)}{\int \exp\left\{\left\langle \widetilde{\boldsymbol{M}}_t \boldsymbol{w} + \widetilde{\boldsymbol{M}}_t^{1/2} \boldsymbol{z}, \boldsymbol{w}_1 \right\rangle - \frac{1}{2}\left\langle \boldsymbol{w}_1, \widetilde{\boldsymbol{M}}_t \boldsymbol{w}_1 \right\rangle\right\} \tilde{q}_0(\mathrm{d}\boldsymbol{w}_1)} \\ &= \sqrt{\beta} \frac{A_{\boldsymbol{w},\boldsymbol{z}}}{B_{\boldsymbol{w},\boldsymbol{z}}}. \end{aligned}$$

for $\boldsymbol{w} \in \mathsf{P}_1(k)$. We have

$$\widetilde{\boldsymbol{M}}_t \boldsymbol{w} = \rho_0 \mathbf{1}_k + \widetilde{\boldsymbol{\Delta}}^t \boldsymbol{w}, \tag{G.12}$$

$$\left\langle \boldsymbol{w}_1, \widetilde{\boldsymbol{M}}_t \boldsymbol{w}_1 \right\rangle = \rho_0 + \left\langle \boldsymbol{w}_1, \widetilde{\boldsymbol{\Delta}}^t \boldsymbol{w}_1 \right\rangle. \tag{G.13}$$

Hence,

$$\begin{aligned} A_{\boldsymbol{w},\boldsymbol{z}} &= \int \boldsymbol{w}_1 \exp\left\{\left\langle \rho_0 \mathbf{1}_k + \widetilde{\boldsymbol{\Delta}}^t \boldsymbol{w} + \left(\rho_0 \boldsymbol{J}_k + \widetilde{\boldsymbol{\Delta}}^t\right)^{1/2} \boldsymbol{z}, \boldsymbol{w}_1 \right\rangle - \frac{\rho_0}{2} - \frac{1}{2}\left\langle \boldsymbol{w}_1, \widetilde{\boldsymbol{\Delta}}^t \boldsymbol{w}_1 \right\rangle\right\} \tilde{q}_0(\mathrm{d}\boldsymbol{w}_1) \\ &= \int \boldsymbol{w}_1 \exp\left\{\frac{\rho_0}{2} + \left\langle \boldsymbol{w}_1, \widetilde{\boldsymbol{\Delta}}^t \boldsymbol{w} \right\rangle - \frac{1}{2}\left\langle \boldsymbol{w}_1, \widetilde{\boldsymbol{\Delta}}^t \boldsymbol{w}_1 \right\rangle + \sqrt{\frac{\rho_0}{k}} \left\langle \boldsymbol{J}_k \boldsymbol{z}, \boldsymbol{w}_1 \right\rangle + \left\langle \boldsymbol{C}_{\boldsymbol{\Delta}}^t \boldsymbol{z}, \boldsymbol{w}_1 \right\rangle\right\} \tilde{q}_0(\mathrm{d}\boldsymbol{w}_1) \end{aligned}$$

where $\boldsymbol{C}_{\boldsymbol{\Delta}}^t \equiv \left(\rho_0 \boldsymbol{J}_k + \widetilde{\boldsymbol{\Delta}}^t\right)^{1/2} - (\rho_0/k)^{1/2} \boldsymbol{J}_k$. Therefore, we have

$$A_{\boldsymbol{w},\boldsymbol{z}} = a \int \boldsymbol{w}_1 \exp\left\{\left\langle \boldsymbol{w}_1, \widetilde{\boldsymbol{\Delta}}^t \boldsymbol{w} \right\rangle - \frac{1}{2}\left\langle \boldsymbol{w}_1, \widetilde{\boldsymbol{\Delta}}^t \boldsymbol{w}_1 \right\rangle + \left\langle \boldsymbol{C}_{\boldsymbol{\Delta}}^t \boldsymbol{z}, \boldsymbol{w}_1 \right\rangle\right\} \tilde{q}_0(\mathrm{d}\boldsymbol{w}_1)$$

where $a = \exp\left\{\rho_0/2 + \sqrt{\rho_0/k}\, \langle \boldsymbol{z}, \mathbf{1}_k \rangle\right\}$. Expanding the exponential, we get

$$A_{\boldsymbol{w},\boldsymbol{z}} = a \int \boldsymbol{w}_1 \left\{1 + \left\langle \boldsymbol{w}_1, \widetilde{\boldsymbol{\Delta}}^t \boldsymbol{w} \right\rangle - \frac{1}{2}\left\langle \boldsymbol{w}_1, \widetilde{\boldsymbol{\Delta}}^t \boldsymbol{w}_1 \right\rangle + \left\langle \boldsymbol{z}, \boldsymbol{C}_{\boldsymbol{\Delta}}^t \boldsymbol{w}_1 \right\rangle + \frac{1}{2}\left\langle \boldsymbol{z}, \boldsymbol{C}_{\boldsymbol{\Delta}}^t \boldsymbol{w}_1 \right\rangle^2 + o\left(\widetilde{\boldsymbol{\Delta}}^t\right)\right\} \tilde{q}_0(\mathrm{d}\boldsymbol{w}_1).$$

Thus,

$$A_{\boldsymbol{w},\boldsymbol{z}} = a\left(\frac{1}{k}\mathbf{1}_k + \boldsymbol{S}\widetilde{\boldsymbol{\Delta}}^t\boldsymbol{w} - \frac{1}{2}\begin{pmatrix}\left\langle\widetilde{\boldsymbol{\Delta}}^t, \boldsymbol{T}_1\right\rangle\\ \left\langle\widetilde{\boldsymbol{\Delta}}^t, \boldsymbol{T}_2\right\rangle\\ \vdots\\ \left\langle\widetilde{\boldsymbol{\Delta}}^t, \boldsymbol{T}_k\right\rangle\end{pmatrix} + \boldsymbol{S}\boldsymbol{C}^t_{\boldsymbol{\Delta}}\boldsymbol{z} + \frac{1}{2}\begin{pmatrix}\left\langle\boldsymbol{C}^t_{\boldsymbol{\Delta}}\boldsymbol{z}^{\otimes 2}\boldsymbol{C}^t_{\boldsymbol{\Delta}}, \boldsymbol{T}_1\right\rangle\\ \left\langle\boldsymbol{C}^t_{\boldsymbol{\Delta}}\boldsymbol{z}^{\otimes 2}\boldsymbol{C}^t_{\boldsymbol{\Delta}}, \boldsymbol{\Omega}'_2\right\rangle\\ \vdots\\ \left\langle\boldsymbol{C}^t_{\boldsymbol{\Delta}}\boldsymbol{z}^{\otimes 2}\boldsymbol{C}^t_{\boldsymbol{\Delta}}, \boldsymbol{T}_k\right\rangle\end{pmatrix} + o\left(\widetilde{\boldsymbol{\Delta}}^t\right)\right)$$

where $\boldsymbol{S}, \boldsymbol{T} \in \mathbb{R}^{k\times k}$ are the moment tensors

$$\boldsymbol{S} = \int \boldsymbol{w}_1^{\otimes 2}\tilde{q}_0(\mathrm{d}\boldsymbol{w}_1) = \frac{\nu}{k\nu(k\nu+1)}(\boldsymbol{I}_k + \nu\boldsymbol{J}_k) = \frac{1}{k(k\nu+1)}\boldsymbol{P}_\perp + \frac{1}{k}\boldsymbol{P}\,, \tag{G.14}$$

$$\boldsymbol{T} = \int \boldsymbol{w}_1^{\otimes 3}\tilde{q}_0(\mathrm{d}\boldsymbol{w}_1), \tag{G.15}$$

$$(T_i)_{jl} = \frac{1}{k\nu(k\nu+1)(k\nu+2)}.\begin{cases}\nu(\nu+1)(\nu+2) & \text{if } j=l=i,\\ \nu^2(\nu+1) & \text{if } j=i,\ l\neq i \text{ or } l=i,\ j\neq i \text{ or } l=j, j\neq i,\\ \nu^3 & \text{otherwise.}\end{cases} \tag{G.16}$$

Similarly, we have

$$B_{\boldsymbol{w},\boldsymbol{z}} = a\int\left\{1 + \left\langle\boldsymbol{w}_1, \widetilde{\boldsymbol{\Delta}}^t\boldsymbol{w}\right\rangle - \frac{1}{2}\left\langle\boldsymbol{w}_1, \widetilde{\boldsymbol{\Delta}}^t\boldsymbol{w}_1\right\rangle + \left\langle\boldsymbol{z}, \boldsymbol{C}^t_{\boldsymbol{\Delta}}\boldsymbol{w}_1\right\rangle + \frac{1}{2}\left\langle\boldsymbol{z}, \boldsymbol{C}^t_{\boldsymbol{\Delta}}\boldsymbol{w}_1\right\rangle^2 + o\left(\widetilde{\boldsymbol{\Delta}}^t\right)\right\}\tilde{q}_0(\mathrm{d}\boldsymbol{w}_1).$$

Therefore,

$$B_{\boldsymbol{w},\boldsymbol{z}} = a\left(1 + \frac{1}{k}\left\langle\mathbf{1}_k\otimes\boldsymbol{w}, \widetilde{\boldsymbol{\Delta}}^t\right\rangle - \frac{1}{2}\left\langle\boldsymbol{S}, \widetilde{\boldsymbol{\Delta}}^t\right\rangle + \frac{1}{k}\left\langle\mathbf{1}_k\otimes\boldsymbol{z}, \boldsymbol{C}^t_{\boldsymbol{\Delta}}\right\rangle + \frac{1}{2}\left\langle\boldsymbol{z}, \boldsymbol{C}^t_{\boldsymbol{\Delta}}\boldsymbol{S}\boldsymbol{C}^t_{\boldsymbol{\Delta}}\boldsymbol{z}\right\rangle + o\left(\widetilde{\boldsymbol{\Delta}}^t\right)\right).$$

Hence, we can write

$$\begin{aligned}f_{\boldsymbol{w},\boldsymbol{z}} = \sqrt{\beta}\frac{A_{\boldsymbol{w},\boldsymbol{z}}}{B_{\boldsymbol{w},\boldsymbol{z}}} =& \sqrt{\beta}\Bigg(\frac{1}{k}\mathbf{1}_k + \boldsymbol{S}\widetilde{\boldsymbol{\Delta}}^t\boldsymbol{w} - \frac{1}{2}\begin{pmatrix}\left\langle\widetilde{\boldsymbol{\Delta}}^t, \boldsymbol{T}_1\right\rangle\\ \left\langle\widetilde{\boldsymbol{\Delta}}^t, \boldsymbol{T}_2\right\rangle\\ \vdots\\ \left\langle\widetilde{\boldsymbol{\Delta}}^t, \boldsymbol{T}_k\right\rangle\end{pmatrix} + \boldsymbol{S}\boldsymbol{C}^t_{\boldsymbol{\Delta}}\boldsymbol{z} + \frac{1}{2}\begin{pmatrix}\left\langle\boldsymbol{C}^t_{\boldsymbol{\Delta}}\boldsymbol{z}^{\otimes 2}\boldsymbol{C}^t_{\boldsymbol{\Delta}}, \boldsymbol{T}_1\right\rangle\\ \left\langle\boldsymbol{C}^t_{\boldsymbol{\Delta}}\boldsymbol{z}^{\otimes 2}\boldsymbol{C}^t_{\boldsymbol{\Delta}}, \boldsymbol{T}_2\right\rangle\\ \vdots\\ \left\langle\boldsymbol{C}^t_{\boldsymbol{\Delta}}\boldsymbol{z}^{\otimes 2}\boldsymbol{C}^t_{\boldsymbol{\Delta}}, \boldsymbol{T}_k\right\rangle\end{pmatrix}\\ &- \frac{1}{k^2}\left\langle\mathbf{1}_k\otimes\boldsymbol{w}, \widetilde{\boldsymbol{\Delta}}^t\right\rangle\mathbf{1}_k + \frac{1}{2k}\left\langle\boldsymbol{S}, \widetilde{\boldsymbol{\Delta}}^t\right\rangle\mathbf{1}_k\\ &- \frac{1}{k^2}\left\langle\mathbf{1}_k\otimes\boldsymbol{z}, \boldsymbol{C}^t_{\boldsymbol{\Delta}}\right\rangle\mathbf{1}_k - \frac{1}{2k}\left\langle\boldsymbol{z}, \boldsymbol{C}^t_{\boldsymbol{\Delta}}\boldsymbol{S}\boldsymbol{C}^t_{\boldsymbol{\Delta}}\boldsymbol{z}\right\rangle\mathbf{1}_k\\ &- \frac{1}{k}\left\langle\mathbf{1}_k\otimes\boldsymbol{z}, \boldsymbol{C}^t_{\boldsymbol{\Delta}}\right\rangle\boldsymbol{S}\boldsymbol{C}^t_{\boldsymbol{\Delta}}\boldsymbol{z} - \frac{1}{k^3}\left\langle\mathbf{1}_k\otimes\boldsymbol{z}, \boldsymbol{C}^t_{\boldsymbol{\Delta}}\right\rangle^2\mathbf{1}_k + o\left(\widetilde{\boldsymbol{\Delta}}^t\right)\Bigg).\end{aligned}$$

Therefore, linearizing Eq. ((4.17)), we get (below, we denote by $[\boldsymbol{A}]_s$ the symmetric part of matrix $\boldsymbol{A}$, namely $[\boldsymbol{A}]_s = (\boldsymbol{A} + \boldsymbol{A}^\mathsf{T})/2)$

$$
\begin{aligned}
\boldsymbol{\Delta}_{t+1} =& \delta \mathbb{E}_{\boldsymbol{w},\boldsymbol{z}} \left(f_{\boldsymbol{w},\boldsymbol{z}}^{\otimes 2} \right) - \frac{\delta\beta}{k^2} \boldsymbol{J}_k && \text{(G.17)} \\
=& \delta\beta \Bigg(\frac{2}{k^2} \left[\boldsymbol{S} (\widetilde{\boldsymbol{\Delta}}^t - (\boldsymbol{C}_{\boldsymbol{\Delta}}^t)^2) \boldsymbol{J}_k \right]_s - \frac{1}{2k} \begin{pmatrix} \left\langle \widetilde{\boldsymbol{\Delta}}^t - (\boldsymbol{C}_{\boldsymbol{\Delta}}^t)^2, \boldsymbol{T}_1 \right\rangle \\ \left\langle \widetilde{\boldsymbol{\Delta}}^t - (\boldsymbol{C}_{\boldsymbol{\Delta}}^t)^2, \boldsymbol{T}_2 \right\rangle \\ \vdots \\ \left\langle \widetilde{\boldsymbol{\Delta}}^t - (\boldsymbol{C}_{\boldsymbol{\Delta}}^t)^2, \boldsymbol{T}_k \right\rangle \end{pmatrix} \otimes \boldsymbol{1}_k && \text{(G.18)} \\
& - \frac{1}{2k} \boldsymbol{1}_k \otimes \begin{pmatrix} \left\langle \widetilde{\boldsymbol{\Delta}}^t - (\boldsymbol{C}_{\boldsymbol{\Delta}}^t)^2, \boldsymbol{T}_1 \right\rangle \\ \left\langle \widetilde{\boldsymbol{\Delta}}^t - (\boldsymbol{C}_{\boldsymbol{\Delta}}^t)^2, \boldsymbol{T}_2 \right\rangle \\ \vdots \\ \left\langle \widetilde{\boldsymbol{\Delta}}^t - (\boldsymbol{C}_{\boldsymbol{\Delta}}^t)^2, \boldsymbol{T}_k \right\rangle \end{pmatrix} - \frac{2}{k^4} \left\langle \boldsymbol{J}_k, \widetilde{\boldsymbol{\Delta}}^t \right\rangle \boldsymbol{J}_k + \frac{1}{k^2} \left\langle \boldsymbol{S}, \widetilde{\boldsymbol{\Delta}}^t - (\boldsymbol{C}_{\boldsymbol{\Delta}}^t)^2 \right\rangle \boldsymbol{J}_k \\
& - \frac{2}{k^4} \left\langle \boldsymbol{J}_k, (\boldsymbol{C}_{\boldsymbol{\Delta}}^t)^2 \right\rangle \boldsymbol{J}_k + \boldsymbol{S} (\boldsymbol{C}_{\boldsymbol{\Delta}}^t)^2 \boldsymbol{S} - \frac{2}{k^2} \left[\boldsymbol{S} (\boldsymbol{C}_{\boldsymbol{\Delta}}^t)^2 \boldsymbol{J}_k \right]_s + \frac{1}{k^4} \left\langle \boldsymbol{J}_k, (\boldsymbol{C}_{\boldsymbol{\Delta}}^t)^2 \right\rangle \boldsymbol{J}_k + o(\widetilde{\boldsymbol{\Delta}}_t) \Bigg).
\end{aligned}
$$

We next decompose $\widetilde{\boldsymbol{\Delta}}_t$ in the component along $\boldsymbol{J}_k$ and the one orthogonal, as per Eq. (G.8), and note that

$$
\begin{aligned}
\boldsymbol{C}_{\boldsymbol{\Delta}}^t &= \left((k\rho_0 + \tilde{\delta}_t) \boldsymbol{P} + \widetilde{\boldsymbol{\Delta}}_t^{(1)} + \widetilde{\boldsymbol{\Delta}}_t^{(2)} \right)^{1/2} - (k\rho_0)^{1/2} \boldsymbol{P} \\
&= \sqrt{k\rho_0 + \tilde{\delta}_t}\, \boldsymbol{P} + \left(\widetilde{\boldsymbol{\Delta}}_t^{(2)} \right)^{1/2} - \sqrt{k\rho_0}\, \boldsymbol{P} + O(\widetilde{\boldsymbol{\Delta}}_t) - (k\rho_0)^{1/2} \boldsymbol{P} = \left(\widetilde{\boldsymbol{\Delta}}_t^{(2)} \right)^{1/2} + O(\widetilde{\boldsymbol{\Delta}}_t) \,,
\end{aligned}
$$

whence

$$
(\boldsymbol{C}_{\boldsymbol{\Delta}}^t)^2 = \widetilde{\boldsymbol{\Delta}}_t^{(2)} + o(\boldsymbol{\Delta}) \,. \tag{G.19}
$$

Using this identity together with Eqs. (G.15), (G.16) in Eq. (G.19) we get

$$
\begin{aligned}
\delta_{t+1} &= o(\widetilde{\boldsymbol{\Delta}}_t) \,, && \text{(G.20)} \\
\boldsymbol{\Delta}_{t+1}^{(1)} &= o(\widetilde{\boldsymbol{\Delta}}_t) \,, && \text{(G.21)} \\
\boldsymbol{\Delta}_{t+1}^{(2)} &= \frac{\beta\delta}{k^2 (k\nu + 1)^2} \widetilde{\boldsymbol{\Delta}}_t^{(2)} + o(\widetilde{\boldsymbol{\Delta}}_t) \,. && \text{(G.22)}
\end{aligned}
$$

Together with Eqs. (G.9) to (G.11), these yield

$$\delta_{t+1} = o(\boldsymbol{\Delta}_t)\,, \tag{G.23}$$

$$\boldsymbol{\Delta}^{(1)}_{t+1} = o(\boldsymbol{\Delta}_t)\,, \tag{G.24}$$

$$\boldsymbol{\Delta}^{(2)}_{t+1} = \frac{\beta^2\delta}{k^2(k\nu+1)^2}\,\boldsymbol{\Delta}^{(2)}_t + o(\widetilde{\boldsymbol{\Delta}}_t)\,. \tag{G.25}$$

Hence the uninformative fixed point is stable if and only if

$$\beta \le \frac{k(k\nu+1)}{\sqrt{\delta}}. \tag{G.26}$$

Note that this is the same condition as the spectral threshold. □

G.4 Stability of the uninformative point: Proof of Theorem 5

In this section we compute the Hessian of the TAP free energy around the uninformative stationary point. We will establish a second order approximation of $\tilde{\mathcal{F}}_{\text{TAP}}(\boldsymbol{r}, \tilde{\boldsymbol{r}})$ near the stationary point. Namely, we denote by $\boldsymbol{r}^*_i = r^*_i\mathbf{1}_k$, $\tilde{\boldsymbol{r}}^*_a = \tilde{r}^*_a\mathbf{1}_k$ the uninformative stationary point, and by $\boldsymbol{m}^*_i = m^*_i\mathbf{1}_k$, $\tilde{\boldsymbol{m}}^*_a = \tilde{m}^*_a\mathbf{1}_k$ the dual variables, where

$$m^*_i = \frac{\sqrt{\beta}}{k}(\boldsymbol{X}^{\mathsf{T}}\mathbf{1}_n)_i\,, \qquad \tilde{m}^*_a = \frac{\beta}{k(1+kq_0)}(\boldsymbol{X}\boldsymbol{X}^{\mathsf{T}}\mathbf{1}_n)_a - \frac{\beta}{k+\delta\beta}\,, \tag{G.27}$$

$$r^*_i = \frac{\sqrt{\beta}}{k(1+kq^*_0)}(\boldsymbol{X}^{\mathsf{T}}\mathbf{1}_n)_i\,, \qquad \tilde{r}^*_a = \frac{1}{k}\,. \tag{G.28}$$

For any other assignment of the variables, $\boldsymbol{r}, \tilde{\boldsymbol{r}}$, $\boldsymbol{m}, \tilde{\boldsymbol{m}}$, we introduce the decomposition

$$\boldsymbol{r}_i = r^s_i\mathbf{1}_k + \boldsymbol{\delta}_i\,, \qquad \tilde{\boldsymbol{r}}_a = \tilde{r}^s_a\mathbf{1}_k + \tilde{\boldsymbol{\delta}}_a\,, \tag{G.29}$$

$$r^s_i = r^*_i + \delta^s_i\,, \qquad \tilde{r}^s_a = \tilde{r}^*_a + \tilde{\delta}^s_a\,, \tag{G.30}$$

$$\boldsymbol{m}_i = m^s_i\mathbf{1}_k + \boldsymbol{\eta}_i\,, \qquad \tilde{\boldsymbol{m}}_a = \tilde{m}^s_a\mathbf{1}_k + \tilde{\boldsymbol{\eta}}_a\,, \tag{G.31}$$

$$m^s_i = m^*_i + \eta^s_i\,, \qquad \tilde{m}^s_a = \tilde{m}^*_a + \tilde{\eta}^s_a\,, \tag{G.32}$$

where $\langle\boldsymbol{\delta}_i, \mathbf{1}_k\rangle = \langle\tilde{\boldsymbol{\delta}}_a, \mathbf{1}_k\rangle = \langle\boldsymbol{\eta}_i, \mathbf{1}_k\rangle = \langle\tilde{\boldsymbol{\eta}}_a, \mathbf{1}_k\rangle = 0$. Note that, by construction $\tilde{r}^s_a = 1/k$.

We will establish an expansion of the form

$$\mathcal{F}_{\text{TAP}}(\boldsymbol{r}, \tilde{\boldsymbol{r}}) = \tilde{\mathcal{F}}_{\text{TAP}}(\boldsymbol{r}^*, \tilde{\boldsymbol{r}}^*) + \mathcal{F}^{(2)}_{\text{TAP}}(\boldsymbol{\delta}, \tilde{\boldsymbol{\delta}}, \delta^s, \tilde{\delta}^s) + o(\delta^2)\,, \tag{G.33}$$

where $\mathcal{F}^{(2)}_{\text{TAP}}$ is a quadratic function, and when using the $O(\,\cdot\,)$ notation, we implicitly consider all δ, η parameters to be of the same order and use δ for denoting that order. Notice that the first-order term is missing from this expansion since $(\boldsymbol{r}^*, \tilde{\boldsymbol{r}}^*)$ is a stationary point.

The crucial step in obtaining the expansion (G.33) is to derive a second order expansion for the logarithmic moment generating functions ϕ, $\tilde{\phi}$, and subsequently for the entropy functions ψ, $\tilde{\psi}$.

Lemma G.2. *Setting variables as per Eq. (G.29), we have*

$$\phi\left(\boldsymbol{m}_i, \frac{\beta}{d}\sum_{a=1}^{n}\tilde{\boldsymbol{r}}_a^{\otimes 2}\right) = -\frac{1}{2}\log(1+ka_0) + \frac{\beta^2(1+\beta\delta/k+k(m_i^*)^2)}{2d^2k(1+\beta\delta/k)^2}\left\|\sum_{a=1}^{n}\tilde{\boldsymbol{\delta}}_a\right\|_2^2 \tag{G.34}$$

$$+\frac{k(m_i^s)^2}{2(1+ka_0)} - \frac{\beta m_i^*}{d(1+\beta\delta/k)}\sum_{a=1}^{n}\langle\boldsymbol{\eta}_i, \tilde{\boldsymbol{\delta}}_a\rangle + \frac{1}{2}\|\boldsymbol{\eta}_i\|_2^2 \tag{G.35}$$

$$-\frac{\beta}{2d}\sum_{a=1}^{n}\|\tilde{\boldsymbol{\delta}}_a\|_2^2 + o(\delta^2)\,,$$

where $a_0 = (\beta/d)\sum_{a=1}^{n}(\tilde{r}_a^s)^2$.

Proof. Let $\boldsymbol{Q} = (\beta/d)\sum_{a=1}^{n}\tilde{\boldsymbol{r}}_a^{\otimes 2}$, and define the orthogonal decomposition $\boldsymbol{Q} = \boldsymbol{Q}_0 + \boldsymbol{Q}_1 + \boldsymbol{Q}_2$, where $\boldsymbol{Q}_0 = \boldsymbol{P}\boldsymbol{Q}\boldsymbol{P}$, $\boldsymbol{Q}_1 = \boldsymbol{P}\boldsymbol{Q}\boldsymbol{P}_\perp + \boldsymbol{P}_\perp\boldsymbol{Q}\boldsymbol{P}$, $\boldsymbol{Q}_2 = \boldsymbol{P}_\perp\boldsymbol{Q}\boldsymbol{P}_\perp$. Using the representation (G.29), we get

$$\boldsymbol{Q}_0 = a_0\boldsymbol{1}_k\boldsymbol{1}_k^\mathsf{T}\,, \qquad a_0 = \frac{\beta}{d}\sum_{a=1}^{n}(\tilde{r}_a^s)^2\,, \tag{G.36}$$

$$\boldsymbol{Q}_1 = \boldsymbol{1}_k\boldsymbol{a}_1^\mathsf{T} + \boldsymbol{a}_1\boldsymbol{1}_k^\mathsf{T}\,, \qquad \boldsymbol{a}_1 = \frac{\beta}{d}\sum_{a=1}^{n}\tilde{r}_a^s\tilde{\boldsymbol{\delta}}_a\,, \tag{G.37}$$

$$\boldsymbol{Q}_2 = \frac{\beta}{d}\sum_{a=1}^{n}\tilde{\boldsymbol{\delta}}_a\tilde{\boldsymbol{\delta}}_a^\mathsf{T}\,. \tag{G.38}$$

By Gaussian integration, we have

$$\phi(\boldsymbol{m}_i, \boldsymbol{Q}) = -\frac{1}{2}\mathsf{Tr}\log\left(\boldsymbol{I}+\boldsymbol{Q}\right) + \frac{1}{2}\langle\boldsymbol{m}_i, (\boldsymbol{I}+\boldsymbol{Q})^{-1}\boldsymbol{m}_i\rangle\,. \tag{G.39}$$

Expanding the logarithm, we get

$$\begin{aligned}
\mathsf{Tr}\log\left(\boldsymbol{I}+\boldsymbol{Q}\right) =& \mathsf{Tr}\log\left(\boldsymbol{I}+\boldsymbol{Q}_0\right) + \mathsf{Tr}\Big\{(\boldsymbol{I}+\boldsymbol{Q}_0)^{-1}(\boldsymbol{Q}_1+\boldsymbol{Q}_2)\Big\} \\
&-\frac{1}{2}\mathsf{Tr}\Big\{(\boldsymbol{I}+\boldsymbol{Q}_0)^{-1}\boldsymbol{Q}_1(\boldsymbol{I}+\boldsymbol{Q}_0)^{-1}\boldsymbol{Q}_1\Big\} + o(\delta^2) \\
=& \mathsf{Tr}\log\left(\boldsymbol{I}+\boldsymbol{Q}_0\right) + \mathsf{Tr}(\boldsymbol{Q}_2) - \langle\boldsymbol{a}_1, (\boldsymbol{I}+\boldsymbol{Q}_0)^{-1}\boldsymbol{a}_1\rangle\,\langle\boldsymbol{1}, (\boldsymbol{I}+\boldsymbol{Q}_0)^{-1}\boldsymbol{1}\rangle + o(\delta^2) \\
=& \log(1+ka_0) + \frac{\beta}{d}\sum_{a=1}^{n}\|\tilde{\boldsymbol{\delta}}_a\|_2^2 - \frac{k}{1+ka_0}\left\|\frac{\beta}{d}\sum_{a=1}^{n}\tilde{r}_a^s\tilde{\boldsymbol{\delta}}_a\right\|_2^2 + o(\delta^2) \\
=& \log(1+ka_0) + \frac{\beta}{d}\sum_{a=1}^{n}\|\tilde{\boldsymbol{\delta}}_a\|_2^2 - \frac{\beta^2}{kd^2(1+kq_0^*)}\left\|\sum_{a=1}^{n}\tilde{\boldsymbol{\delta}}_a\right\|_2^2 + o(\delta^2)
\end{aligned} \tag{G.40}$$

Considering next the second term in Eq. (G.39), we get

$$\begin{aligned}\langle \boldsymbol{m}_i, (\boldsymbol{I}+\boldsymbol{Q})^{-1}\boldsymbol{m}_i\rangle =&(m_i^s)^2\langle \mathbf{1}, (\boldsymbol{I}+\boldsymbol{Q}_0+\boldsymbol{Q}_1+\boldsymbol{Q}_2)^{-1}\mathbf{1}\rangle + 2m_i^s\langle \boldsymbol{\eta}_i, (\boldsymbol{I}+\boldsymbol{Q}_0+\boldsymbol{Q}_1)^{-1}\mathbf{1}\rangle \\ &+ \langle \boldsymbol{\eta}_i, (\boldsymbol{I}+\boldsymbol{Q}_0)^{-1}\boldsymbol{\eta}_i\rangle + o(\delta^2) \\ =&(m_i^s)^2\langle \mathbf{1}, (\boldsymbol{I}+\boldsymbol{Q}_0)^{-1}\mathbf{1}\rangle \\ &+ (m_i^s)^2\langle \mathbf{1}, (\boldsymbol{I}+\boldsymbol{Q}_0)^{-1}\boldsymbol{Q}_1(\boldsymbol{I}+\boldsymbol{Q}_0)^{-1}\boldsymbol{Q}_1(\boldsymbol{I}+\boldsymbol{Q}_0)^{-1}\mathbf{1}\rangle \\ &- 2m_i^s\langle \boldsymbol{\eta}_i, (\boldsymbol{I}+\boldsymbol{Q}_0)^{-1}\boldsymbol{Q}_1(\boldsymbol{I}+\boldsymbol{Q}_0)^{-1}\mathbf{1}\rangle + \|\boldsymbol{\eta}_i\|_2^2 + o(\delta^2) \\ =&\frac{k(m_i^s)^2}{1+ka_0} + \frac{(km_i^s)^2}{(1+ka_0)^2}\|\boldsymbol{a}_1\|_2^2 - \frac{2km_i^s}{(1+ka_0)}\langle \boldsymbol{\eta}_i, \boldsymbol{a}_1\rangle + \|\boldsymbol{\eta}_i\|_2^2 + o(\delta^2) \\ =&\frac{k(m_i^s)^2}{1+ka_0} + \frac{(\beta m_i^s)^2}{d^2(1+kq_0^*)^2}\left\|\sum_{a=1}^n \tilde{\boldsymbol{\delta}}_a\right\|_2^2 - \frac{2\beta m_i^s}{d(1+kq_0^*)}\sum_{a=1}^n \langle \boldsymbol{\eta}_i, \tilde{\boldsymbol{\delta}}_a\rangle \\ &+ \|\boldsymbol{\eta}_i\|_2^2 + o(\delta^2)\,.\end{aligned}$$

□

Lemma G.3. *Setting variables as per Eq. (G.29), we have*

$$\begin{aligned}\tilde{\phi}\left(\tilde{\boldsymbol{m}}_a, \frac{\beta}{d}\sum_{i=1}^d \boldsymbol{r}_i^{\otimes 2}\right) =&\tilde{m}_a^s - \frac{1}{2}b_0 + \frac{1}{2k(k\nu+1)}\left\|\tilde{\boldsymbol{\eta}}_a - \frac{\beta}{d}\sum_{i=1}^d r_i^*\boldsymbol{\delta}_i\right\|_2^2 \\ &- \frac{\beta}{2dk(k\nu+1)}\sum_{i=1}^d \|\boldsymbol{\delta}_i\|_2^2 + o(\delta^2)\,,\end{aligned}$$

where $b_0 = (\beta/d)\sum_{i=1}^d (r_i^s)^2$.

Proof. Let $\widetilde{\boldsymbol{Q}} = (\beta/d)\sum_{i=1}^d \boldsymbol{r}_i^{\otimes 2}$ and, as in the previous proof, define the orthogonal decomposition $\widetilde{\boldsymbol{Q}} = \widetilde{\boldsymbol{Q}}_0 + \widetilde{\boldsymbol{Q}}_1 + \widetilde{\boldsymbol{Q}}_2$, where $\widetilde{\boldsymbol{Q}}_0 = \boldsymbol{P}\widetilde{\boldsymbol{Q}}\boldsymbol{P}$, $\widetilde{\boldsymbol{Q}}_1 = \boldsymbol{P}\widetilde{\boldsymbol{Q}}\boldsymbol{P}_\perp + \boldsymbol{P}_\perp\widetilde{\boldsymbol{Q}}\boldsymbol{P}$, $\widetilde{\boldsymbol{Q}}_2 = \boldsymbol{P}_\perp\widetilde{\boldsymbol{Q}}\boldsymbol{P}_\perp$. Using the representation (G.29), we get

$$\begin{aligned}\widetilde{\boldsymbol{Q}}_0 &= b_0\mathbf{1}_k\mathbf{1}_k^{\mathsf{T}}\,, & b_0 &= \frac{\beta}{d}\sum_{i=1}^d (r_i^s)^2\,, \\ \widetilde{\boldsymbol{Q}}_1 &= \mathbf{1}_k\boldsymbol{b}_1^{\mathsf{T}} + \boldsymbol{b}_1\mathbf{1}_k^{\mathsf{T}}\,, & \boldsymbol{b}_1 &= \frac{\beta}{d}\sum_{i=1}^d r_i^s\boldsymbol{\delta}_i\,, \\ \widetilde{\boldsymbol{Q}}_2 &= \frac{\beta}{d}\sum_{i=1}^d \boldsymbol{\delta}_i\boldsymbol{\delta}_i^{\mathsf{T}}\,.\end{aligned}$$

For $\boldsymbol{w} \in \mathrm{supp}(\tilde{q}_0)$, we have $\langle \mathbf{1}, \boldsymbol{w}\rangle = 1$ and therefore

$$\begin{aligned}\tilde{\phi}(\tilde{\boldsymbol{m}}_a, \widetilde{\boldsymbol{Q}}) &= \log\left\{\int e^{\langle \tilde{\boldsymbol{m}}, \boldsymbol{w}\rangle - \frac{1}{2}\langle \boldsymbol{w}, \widetilde{\boldsymbol{Q}}\boldsymbol{w}\rangle}\tilde{q}_0(\mathrm{d}\boldsymbol{w})\right\}\\ &= \tilde{m}_a^s - \frac{1}{2}b_0 + \log\left\{\int e^{\langle \tilde{\boldsymbol{\eta}}_a - \boldsymbol{b}_1, \boldsymbol{w}\rangle - \frac{1}{2}\langle \boldsymbol{w}, \widetilde{\boldsymbol{Q}}_2\boldsymbol{w}\rangle}\tilde{q}_0(\mathrm{d}\boldsymbol{w})\right\}\\ &= \tilde{m}_a^s - \frac{1}{2}b_0 + \frac{1}{2}\langle (\tilde{\boldsymbol{\eta}}_a - \boldsymbol{b}_1)(\tilde{\boldsymbol{\eta}}_a - \boldsymbol{b}_1)^{\mathsf{T}} - \widetilde{\boldsymbol{Q}}_1, \boldsymbol{S}_\perp\rangle + o(\delta^2)\,,\end{aligned}$$

where, cf. Eq. (G.14),

$$\boldsymbol{S}_\perp = \int (\boldsymbol{P}_\perp \boldsymbol{w})^{\otimes 2}\tilde{q}_0(\mathrm{d}\boldsymbol{w}) = \frac{1}{k(k\nu+1)}\boldsymbol{P}_\perp\,.$$

Hence, we obtain immediately the claim. □

We next transfer the above results on the moment generating functions ϕ, $\tilde{\phi}$, to analogous results on the entropy functions ψ, $\tilde{\psi}$.

Lemma G.4. *Setting variables as per Eq. (G.29), we have*

$$\begin{aligned}\psi\left(\boldsymbol{r}_i, \frac{\beta}{d}\sum_{a=1}^n \tilde{\boldsymbol{r}}_a^{\otimes 2}\right) =& \frac{1}{2}\log(1+ka_0) + \frac{1}{2}k(1+ka_0)(r_i^s)^2 - \frac{\beta^2(1+\beta\delta/k + k(m_i^*)^2)}{2d^2k(1+\beta\delta/k)^2}\left\|\sum_{a=1}^n \tilde{\boldsymbol{\delta}}_a\right\|_2^2\\ &+ \frac{1}{2}\left\|\boldsymbol{\delta}_i + \frac{\beta m_i^*}{d(1+\beta\delta/k)}\sum_{a=1}^n \tilde{\boldsymbol{\delta}}_a\right\|_2^2 + \frac{\beta}{2d}\sum_{a=1}^n \|\tilde{\boldsymbol{\delta}}_a\|_2^2 + o(\delta^2)\,,\end{aligned}$$

where $a_0 = (\beta/d)\sum_{a=1}^n (\tilde{r}_a^s)^2$.

Proof. By definition

$$\psi(\boldsymbol{r}_i, \boldsymbol{Q}) = \max_{m_i^s, \boldsymbol{\eta}_i}\left\{k m_i^s r_i^s + \langle \boldsymbol{\eta}_i, \boldsymbol{\delta}_i\rangle - \phi(\boldsymbol{m}_i, \boldsymbol{Q})\right\}.$$

Since $\phi(\,\cdot\,, \boldsymbol{Q})$ is strongly convex, the maximum is realized when $\eta_i^s, \boldsymbol{\eta}_i = O(\delta)$ and can be computed order-by-order in δ. Hence, substituting (G.34) we obtain the claim. □

Lemma G.5. *Setting variables as per Eq. (G.29), we have*

$$\begin{aligned}\tilde{\psi}\left(\tilde{\boldsymbol{r}}_a, \frac{\beta}{d}\sum_{i=1}^d \boldsymbol{r}_i^{\otimes 2}\right) &= \frac{1}{2}b_0 + \frac{1}{2}k(k\nu+1)\|\tilde{\boldsymbol{\delta}}_a\|_2^2 + \frac{\beta}{d}\sum_{i=1}^d r_i^*\langle \boldsymbol{\delta}_i, \tilde{\boldsymbol{\delta}}_a\rangle\\ &\quad + \frac{\beta}{2dk(k\nu+1)}\sum_{i=1}^d \|\boldsymbol{\delta}_i\|_2^2 + o(\delta^2)\,,\end{aligned}$$

where $b_0 = (\beta/d)\sum_{i=1}^d (r_i^s)^2$.

Proof. By definition

$$\tilde{\psi}(\tilde{\boldsymbol{r}}_i, \widetilde{\boldsymbol{Q}}) = \max_{\tilde{m}_i^s, \tilde{\boldsymbol{\eta}}_i} \left\{ k\tilde{m}_i^s \tilde{r}_i^s + \langle \tilde{\boldsymbol{\eta}}_i, \tilde{\boldsymbol{\delta}}_i \rangle - \tilde{\phi}(\tilde{\boldsymbol{m}}_i, \widetilde{\boldsymbol{Q}}) \right\}.$$

The proof is again obtained by maximizing order by order in δ, and using $\tilde{r}_a^s = 1/k$. □

Lemma G.6. *Setting variables as per Eq. (G.29), and introducing the vectors* $\boldsymbol{r}^s = (r_i^s)_{i \le d} \in \mathbb{R}^d$, $\tilde{\boldsymbol{r}}^s = (\tilde{r}_a^s)_{a \le n} \in \mathbb{R}^n$, *we obtain*

$$\mathcal{F}_{\text{TAP}}(\boldsymbol{r}, \tilde{\boldsymbol{r}}) = \mathcal{F}_{\text{TAP}}^{(s)}(\boldsymbol{r}^s, \tilde{\boldsymbol{r}}^s) + \mathcal{F}_{\text{TAP}}^{(a)}(\boldsymbol{\delta}, \tilde{\boldsymbol{\delta}}) + o(\delta^2), \tag{G.41}$$

$$\mathcal{F}_{\text{TAP}}^{(s)}(\boldsymbol{r}^s, \tilde{\boldsymbol{r}}^s) = \frac{d}{2} \log\left(1 + \frac{\beta\delta}{k}\right) + \frac{1}{2} k \Big(1 + \frac{\beta\delta}{k}\Big) \|\boldsymbol{r}^s\|_2^2 - k\sqrt{\beta} \langle \mathbf{1}, \boldsymbol{X}\boldsymbol{r}^s \rangle, \tag{G.42}$$

$$\begin{aligned}\mathcal{F}_{\text{TAP}}^{(a)}(\boldsymbol{\delta}, \tilde{\boldsymbol{\delta}}) = &\frac{1}{2}\left(1 + \frac{\beta\delta}{k(k\nu+1)}\right) \|\boldsymbol{\delta}\|_F^2 + \frac{1}{2}\Big(\beta + k(k\nu+1)\Big) \|\tilde{\boldsymbol{\delta}}\|_F^2 \\ &- \frac{\beta^2}{2dk(1+\beta\delta/k)} \left\| \sum_{a \le n} \tilde{\boldsymbol{\delta}}_a \right\|_2^2 - \sqrt{\beta}\mathsf{Tr}(\boldsymbol{X}\boldsymbol{\delta}\tilde{\boldsymbol{\delta}}^{\mathsf{T}}) + \frac{\beta}{d(1+\beta\delta/k)} \sum_{i \le d, a \le n} m_i^* \langle \boldsymbol{\delta}_i, \tilde{\boldsymbol{\delta}}_a \rangle.\end{aligned} \tag{G.43}$$

Proof. Using the decomposition (G.29), we get

$$\begin{aligned}\mathsf{Tr}(\boldsymbol{X}\boldsymbol{r}\tilde{\boldsymbol{r}}^{\mathsf{T}}) &= k\mathsf{Tr}\Big(\boldsymbol{X}\boldsymbol{r}^s(\tilde{\boldsymbol{r}}^s)^{\mathsf{T}}\Big) + \mathsf{Tr}(\boldsymbol{X}\boldsymbol{\delta}\tilde{\boldsymbol{\delta}}^{\mathsf{T}}), \\ \sum_{i \le d, a \le n} \langle \boldsymbol{r}_i, \tilde{\boldsymbol{r}}_a \rangle^2 &= k^2 \sum_{i \le d, a \le n} (r_i^s)^2 (\tilde{r}_a^s)^2 + 2k \sum_{i \le d, a \le n} (r_i^s \tilde{r}_a^s) \langle \boldsymbol{\delta}_i, \tilde{\boldsymbol{\delta}}_a \rangle + o(\delta^2) \\ &= k^2 \sum_{i \le d, a \le n} (r_i^s)^2 (\tilde{r}_a^s)^2 + 2 \sum_{i \le d, a \le n} r_i^s \langle \boldsymbol{\delta}_i, \tilde{\boldsymbol{\delta}}_a \rangle + o(\delta^2),\end{aligned}$$

where we used the fact that $\tilde{r}_a^s = 1/k$. Using these, together with Lemma G.4, G.5 in Eq. (4.3), we get the decomposition (G.41) where

$$\begin{aligned}\mathcal{F}_{\text{TAP}}^{(s)}(\boldsymbol{r}^s, \tilde{\boldsymbol{r}}^s) = &\frac{d}{2} \log\left(1 + \frac{\beta k}{d} \|\tilde{\boldsymbol{r}}^s\|_2^2\right) + \frac{1}{2} k^2 \Big(1 + \frac{\beta k}{d} \|\tilde{\boldsymbol{r}}^s\|_2^2\Big) \|\boldsymbol{r}^s\|_2^2 + \frac{1}{2} \beta\delta \|\boldsymbol{r}^s\|_2^2 \\ &- k\sqrt{\beta} \mathsf{Tr}(\boldsymbol{X}\boldsymbol{r}^s(\tilde{\boldsymbol{r}}^s)^{\mathsf{T}}) - \frac{\beta k^2}{2d} \|\boldsymbol{r}^s\|_2^2 \|\tilde{\boldsymbol{r}}^s\|_2^2,\end{aligned}$$

Substituting $\tilde{\boldsymbol{r}}^s = \mathbf{1}_n/k$, we obtain Eq. (G.42). □

Notice that $\mathcal{F}_{\text{TAP}}^{(s)}(\boldsymbol{r}^s, \tilde{\boldsymbol{r}}^s)$ is a positive definite quadratic function in $\boldsymbol{r}^s$, minimized at $\boldsymbol{r}^s = \boldsymbol{r}^*$. Hence, in order to establish the stability of the uninformative stationary point, it is sufficient to check that the quadratic form $\mathcal{F}_{\text{TAP}}^{(a)}(\boldsymbol{\delta}, \tilde{\boldsymbol{\delta}})$ is positive definite. The matrix representation of this quadratic form yields

$$\boldsymbol{\Omega} = \begin{bmatrix} \Big(1 + \frac{\delta\beta}{k(k\nu+1)}\Big) \boldsymbol{I}_d & -\sqrt{\beta}\boldsymbol{X}^{\mathsf{T}}\Big(\boldsymbol{I}_n - \frac{\beta}{d(k+\delta\beta)} \mathbf{J}_n\Big) \\ -\sqrt{\beta}\Big(\boldsymbol{I}_n - \frac{\beta}{d(k+\delta\beta)} \mathbf{J}_n\Big)\boldsymbol{X} & \Big(\beta + k(k\nu+1)\Big)\boldsymbol{I}_n - \frac{\beta^2}{d(k+\delta\beta)} \mathbf{J}_n \end{bmatrix}. \tag{G.44}$$

We are left with the task of proving that $\mathbf{\Omega} \succ \mathbf{0}$ for $\beta < \beta_{\rm spect}(k,\delta,\nu)$. We will use the following random matrix theory lemma.

Lemma G.7. *Let $\boldsymbol{u} \in \mathbb{R}^n$, $\boldsymbol{v} \in \mathbb{R}^d$ be vectors with $\|\boldsymbol{u}\|_2 = \|\boldsymbol{v}\|_2 = 1$, $\gamma, \alpha_{\|}$, $\alpha_{\perp}, \overline{\lambda} \in \mathbb{R}$ be numbers, and let $\boldsymbol{P}_{\boldsymbol{u}} = \boldsymbol{u}\boldsymbol{u}^{\sf T}$ be the orthogonal projector onto $\boldsymbol{u}$, and $\boldsymbol{P}_{\boldsymbol{u}}^{\perp} = \boldsymbol{I} - \boldsymbol{u}\boldsymbol{u}^{\sf T}$ be its orthogonal complement. Denote by $\boldsymbol{Z} \in \mathbb{R}^{n\times d}$ random matrices with $(Z_{ij})_{i\le n, j\le d} \sim \mathsf{N}(0, 1/d)$, with $n/d \to \delta \in (0,\infty)$ as $n \to \infty$, and define the matrix*

$$\boldsymbol{M} = \gamma \boldsymbol{u}\boldsymbol{v}^{\sf T} + \alpha_{\|}\boldsymbol{P}_{\boldsymbol{u}}\boldsymbol{Z} + \alpha_{\perp}\boldsymbol{P}_{\boldsymbol{u}}^{\perp}\boldsymbol{Z}\,. \tag{G.45}$$

Finally define $\gamma_*^2 \equiv (1+\sqrt{\delta})\alpha_{\perp}^2 - \alpha_{\|}^2$, and

$$\lambda_*^2 \equiv \begin{cases} \frac{(\gamma^2+\alpha_{\|}^2)(\gamma^2+\alpha_{\|}^2-\alpha_{\perp}^2(1-\delta))}{\gamma^2+\alpha_{\|}^2-\alpha_{\perp}^2} & \text{if } \gamma^2 > \gamma_*^2, \\ \alpha_{\perp}^2(1+\sqrt{\delta})^2 & \text{otherwise.} \end{cases} \tag{G.46}$$

Then, denoting by $s_{\max}(\boldsymbol{M})$ the largest singular value of $\boldsymbol{M}$, we have $\lim_{n\to\infty} s_{\max}(\boldsymbol{M}) = \lambda_*$ *in probability.*

Proof. By rotational invariance of $\boldsymbol{Z}$, we can and will assume $\boldsymbol{u} = \boldsymbol{e}_1$, and will denote by $\widetilde{\boldsymbol{Z}} \in \mathbb{R}^{(n-1)\times d}$ the matrix containing the last $(n-1)$ rows of $\boldsymbol{Z}$. We further let $\boldsymbol{w} = \gamma\boldsymbol{v} + \alpha_{\|}\boldsymbol{Z}^{\sf T}\boldsymbol{u}$. With these definitions,

$$\boldsymbol{M}\boldsymbol{M}^{\sf T} = \begin{bmatrix} \|\boldsymbol{w}\|_2^2 & \alpha_{\perp}(\widetilde{\boldsymbol{Z}}\boldsymbol{w})^{\sf T} \\ \alpha_{\perp}(\widetilde{\boldsymbol{Z}}\boldsymbol{w}) & \alpha_{\perp}^2\widetilde{\boldsymbol{Z}}\widetilde{\boldsymbol{Z}}^{\sf T} \end{bmatrix}.$$

Note that, almost surely, $\lim_{n\to\infty}\lambda_{\max}(\widetilde{\boldsymbol{Z}}\widetilde{\boldsymbol{Z}}^{\sf T}) = (1+\sqrt{\delta})^2$ [2], and therefore

$$\liminf_{n\to\infty} s_{\max}(\boldsymbol{M})^2 \ge \alpha_{\perp}^2(1+\sqrt{\delta})^2$$

almost surely.

Recall that, as long as s_n^2 is not an eigenvalue of $\alpha_{\perp}^2\widetilde{\boldsymbol{Z}}\widetilde{\boldsymbol{Z}}^{\sf T}$, we have

$$\det(s_n^2\boldsymbol{I} - \boldsymbol{M}\boldsymbol{M}^{\sf T}) = \det(s_n^2\boldsymbol{I} - \alpha_{\perp}^2\widetilde{\boldsymbol{Z}}\widetilde{\boldsymbol{Z}}^{\sf T})\left\{s_n^2 - \|\boldsymbol{w}\|_2^2 - \alpha_{\perp}^2\langle\boldsymbol{w}, \widetilde{\boldsymbol{Z}}^{\sf T}(s_n^2\boldsymbol{I} - \alpha_{\perp}^2\widetilde{\boldsymbol{Z}}\widetilde{\boldsymbol{Z}}^{\sf T})^{-1}\widetilde{\boldsymbol{Z}}\boldsymbol{w}\rangle\right\}$$

It is immediate to see that (unless $\alpha_{\perp} = 0$ or $\boldsymbol{v} = 0$), $s_n^2 > \lambda_{\max}(\alpha_{\perp}^2\widetilde{\boldsymbol{Z}}\widetilde{\boldsymbol{Z}}^{\sf T})$ almost surely, and therefore s_n is given by the largest solution of the equation

$$s_n^2 = \|\boldsymbol{w}\|_2^2 + \alpha_{\perp}^2\langle\boldsymbol{w}, \widetilde{\boldsymbol{Z}}^{\sf T}(s_n^2\boldsymbol{I} - \alpha_{\perp}^2\widetilde{\boldsymbol{Z}}\widetilde{\boldsymbol{Z}}^{\sf T})^{-1}\widetilde{\boldsymbol{Z}}\boldsymbol{w}\rangle\,. \tag{G.47}$$

Note that, almost surely, $\lim_{n\to\infty}\|\boldsymbol{w}\|_2^2 = \gamma^2 + \alpha_{\|}^2 \equiv \tilde{\gamma}^2$. Further, $\boldsymbol{w}$ is independent of $\widetilde{\boldsymbol{Z}}$. Hence, by a standard random matrix theory argument [1, 2], for any $s^2 > \alpha_{\perp}^2(1+\sqrt{\delta})^2$, the following limits hold almost surely

$$\lim_{n\to\infty}\frac{\alpha_\perp^2}{\|\boldsymbol{w}\|_2^2}\langle \boldsymbol{w},\widetilde{\boldsymbol{Z}}^{\mathsf{T}}(s^2\boldsymbol{I}-\alpha_\perp^2\widetilde{\boldsymbol{Z}}\widetilde{\boldsymbol{Z}}^{\mathsf{T}})^{-1}\widetilde{\boldsymbol{Z}}\boldsymbol{w}\rangle = \lim_{n\to\infty}\frac{1}{d}\mathsf{Tr}\Big[\widetilde{\boldsymbol{Z}}^{\mathsf{T}}\Big((s^2/\alpha_\perp^2)\boldsymbol{I}-\widetilde{\boldsymbol{Z}}\widetilde{\boldsymbol{Z}}^{\mathsf{T}}\Big)^{-1}\widetilde{\boldsymbol{Z}}\Big]$$

$$= -\delta-\frac{s^2\delta}{\alpha_\perp^2}\lim_{n\to\infty}\frac{1}{n}\mathsf{Tr}\Big[\Big(\widetilde{\boldsymbol{Z}}\widetilde{\boldsymbol{Z}}^{\mathsf{T}}-(s^2/\alpha_\perp^2)\boldsymbol{I}\Big)^{-1}\Big]$$

$$= -\delta-\frac{s^2\delta}{\alpha_\perp^2}R\Big(\frac{s^2}{\alpha_\perp^2}\Big),$$

where $R(t)$ is the Stieltjes transform of the limit eigenvalues distribution of a Wishart matrix, which is given by the Marchenko-Pastur law [2]

$$R(z)=\frac{-z-\delta+1+\sqrt{(z+\delta-1)-4\delta z}}{2\delta z}.$$

Recall that $z\mapsto R(z)$ is increasing on $[z_v,\infty)$, $z_c\equiv(1+\sqrt{\delta})^2$, with $R(z_c+u)=R(z_c)-c\sqrt{u}+O(u)$ (for a constant $c>0$) as $u\downarrow 0$, and $R(z)=-1/z+O(1/z^2)$ as $z\to\infty$. We therefore can consider the following asymptotic version of Eq. (G.47):

$$\frac{s^2}{\tilde{\gamma}^2}=\widehat{R}\left(\frac{s^2}{\alpha_\perp^2}\right), \qquad \widehat{R}(z)=1-\delta-\delta z\,R(z)\,. \tag{G.48}$$

Note that $\widehat{R}(z)$ is monotone decreasing on $[z_c,\infty)$ with $\widehat{R}(z_c)=(1+\sqrt{\delta})$, $\widehat{R}(z_c+u)=\widehat{R}(z_c)-c\sqrt{u}+O(u)$, and $\widehat{R}(z)=1+O(1/z)$ as $z\to\infty$. For $\tilde{\gamma}^2>(1+\sqrt{\delta})\alpha_\perp^2$, this equation has a unique solution s_*^2 with $s^2/\tilde{\gamma}^2<\widehat{R}(s^2\alpha_\perp^2)$ for $s^2\in[\alpha_\perp^2(1+\sqrt{\delta})^2,s_*^2)$ and $s^2/\tilde{\gamma}^2>\widehat{R}(s^2\alpha_\perp^2)$ for $s^2>s_*^2$. Hence, the largest solution s_n^2 of (G.47) converges almost surely to s_*^2 as $n\to\infty$.

For $\tilde{\gamma}^2>(1+\sqrt{\delta})\alpha_\perp^2$, we have $s^2/\tilde{\gamma}^2>\widehat{R}(s^2\alpha_\perp^2)$ for all $s^2>\alpha_\perp^2(1+\sqrt{\delta})^2$ and therefore $\limsup_{n\to\infty}s_n^2\leq\alpha_\perp^2(1+\sqrt{\delta})^2$ almost surely. Since we have a matching lower bound, we conclude that $\lim_{n\to\infty}s_n^2\leq\alpha_\perp^2(1+\sqrt{\delta})^2$ in this case.

Finally, the expression (G.46) follows by solving rewriting Eq. (G.48) as $\widehat{R}^{-1}(s^2/\tilde{\gamma}^2)=s^2/\alpha_\perp^2$, whereby the inverse of $\widehat{R}$ in $(1,1+\sqrt{\delta})$ is given by

$$\widehat{R}^{-1}(x)=\frac{x(x+\delta-1)}{x-1}.$$

□

We next state a general lemma that can be used to check whether a matrix of the form (G.44) is positive semidefinite.

Lemma G.8. *Let $\boldsymbol{Z}\in\mathbb{R}^{n\times d}$ be random matrices with $(Z_{ij})_{i\leq n,j\leq d}\sim\mathsf{N}(0,1/d)$, and $\boldsymbol{u}\in\mathbb{R}^n$, $\boldsymbol{v}\in\mathbb{R}^d$ be unit vectors, with $n/d\to\delta$ as $n\to\infty$. Define the projectors $\boldsymbol{P}_{\boldsymbol{u}}=\boldsymbol{u}\boldsymbol{u}^{\mathsf{T}}$ and*

$\boldsymbol{P}_{\boldsymbol{u}}^{\perp} = \boldsymbol{I} - \boldsymbol{u}\boldsymbol{u}^{\mathsf{T}}$. *For* $a, br, s, \beta, \xi \in \mathbb{R}$ *with* $\beta \geq 0$ *and* $r > s$*, let*

$$\bar{\boldsymbol{X}} = \xi\, \boldsymbol{u}\boldsymbol{v}^{\mathsf{T}} + \boldsymbol{Z}\,, \tag{G.49}$$

$$\bar{\boldsymbol{\Omega}} = \begin{bmatrix} a\,\boldsymbol{I}_d & -\sqrt{\beta}\,\bar{\boldsymbol{X}}^{\mathsf{T}}(\boldsymbol{I}_n - b\boldsymbol{P}_{\boldsymbol{u}}) \\ -\sqrt{\beta}\,(\boldsymbol{I}_n - b\boldsymbol{P}_{\boldsymbol{u}})\bar{\boldsymbol{X}} & (r\boldsymbol{I}_n - s\boldsymbol{P}_{\boldsymbol{u}}) \end{bmatrix}. \tag{G.50}$$

Assume that one of the following two conditions holds:

1. $(1-b)^2(1+\xi^2)/(r-s) \geq (1+\sqrt{\delta})/r$ *and*

$$a(r-s) > \beta\frac{(1-b)^2(1+\xi^2)\big[(1-b)^2(1+\xi^2)r - (1-\delta)(r-s)\big]}{(1-b)^2(1+\xi^2)r - (r-s)}\,. \tag{G.51}$$

2. $(1-b)^2(1+\xi^2)/(r-s) < (1+\sqrt{\delta})/r$ *and*

$$a > \frac{\beta}{r}(1+\sqrt{\delta})^2\,. \tag{G.52}$$

Then, there exists a constant $\varepsilon > 0$ *such that, almost surely,* $\boldsymbol{\Omega} \succeq \varepsilon\boldsymbol{I}$ *for all* n *large enough.*

Proof. Let us first prove that, under the stated conditions, $\boldsymbol{\Omega} \succeq \boldsymbol{0}$. Since $r\boldsymbol{I}_n - s\boldsymbol{P}_{\boldsymbol{u}} \succ \boldsymbol{0}$, we have $\boldsymbol{\Omega} \succ \boldsymbol{0}$ if and only if

$$a\boldsymbol{I}_d \succ \beta\bar{\boldsymbol{X}}^{\mathsf{T}}(\boldsymbol{I} - b\boldsymbol{P}_{\boldsymbol{u}})(r - s\boldsymbol{P}_{\boldsymbol{u}})^{-1}(\boldsymbol{I} - b\boldsymbol{P}_{\boldsymbol{u}})\bar{\boldsymbol{X}}\,. \tag{G.53}$$

Notice that

$$(\boldsymbol{I} - b\boldsymbol{P}_{\boldsymbol{u}})(r - s\boldsymbol{P}_{\boldsymbol{u}})^{-1}(\boldsymbol{I} - b\boldsymbol{P}_{\boldsymbol{u}}) = \frac{1}{r}\,\boldsymbol{P}_{\boldsymbol{u}}^{\perp} + \frac{(1-b)^2}{r-s}\,\boldsymbol{P}_{\boldsymbol{u}}\,.$$

Hence, condition (G.53) is equivalent to $a > \lambda_{\max}(\boldsymbol{M}^{\mathsf{T}}\boldsymbol{M}) = s_{\max}(\boldsymbol{M})^2$, where

$$\boldsymbol{M} = \sqrt{\beta}\left[\frac{1-b}{\sqrt{r-s}}\boldsymbol{P}_{\boldsymbol{u}} + \frac{1}{\sqrt{r}}\boldsymbol{P}_{\boldsymbol{u}}^{\perp}\right]\bar{\boldsymbol{X}}\,.$$

Note that $\boldsymbol{M}$ is of the form of Lemma G.7, with

$$\gamma = \sqrt{\frac{\beta\xi^2(1-b)^2}{r-s}}\,, \qquad \alpha_{\parallel} = \sqrt{\frac{\beta(1-b)^2}{r-s}}\,, \qquad \alpha_{\perp} = \sqrt{\frac{\beta}{r}}\,.$$

The claim that $\boldsymbol{\Omega} \succ \boldsymbol{0}$ then follows by using the asymptotic characterization of $s_{\max}(\boldsymbol{M})$ in Lemma G.7.

We next prove that in fact $\boldsymbol{\Omega} \succeq \varepsilon\boldsymbol{I}$. If the stated conditions hold, there exists ε small enough such that they hold also after replacing a with $a' = a - \varepsilon$ and r with $r' = r - \varepsilon$. Let us write $\boldsymbol{\Omega}(a, r)$ for the matrix of Eq. (G.50), where we emphasized the dependence on the parameters a, r. We have $\boldsymbol{\Omega}(a,r) = \boldsymbol{\Omega}(a',r') + \varepsilon \boldsymbol{I}$, and hence the thesis follows since $\boldsymbol{\Omega}(a',b') \succeq \mathbf{0}$. □

In order to apply the last lemma, we will show that, for $\beta < \beta_{\rm spect}$, the LDA model of Eq. (1.2) is equivalent for our purposes to a simpler model.

Lemma G.9. *Let $\boldsymbol{X} \in \mathbb{R}^{n\times d}$ be distributed according to the LDA model (1.2) and let $\boldsymbol{R}_1 \in \mathbb{R}^{n\times n}$, $\boldsymbol{R}_2 \in \mathbb{R}^{d\times d}$ be uniformly random (Haar distributed) orthogonal matrices conditional to $\boldsymbol{R}_1\mathbf{1} = \mathbf{1}$, with $\{\boldsymbol{X}, \boldsymbol{R}_1, \boldsymbol{R}_2\}$ mutually independent. Denote by $\mathbb{P}_{1,n}$ the law of $\boldsymbol{X}_R \equiv \boldsymbol{R}_1\boldsymbol{X}\boldsymbol{R}_2$.*

Define $\bar{\boldsymbol{X}} = \xi\, \boldsymbol{u}\boldsymbol{v}^{\mathsf{T}} + \boldsymbol{Z}$ as per Eq. (G.49), with $\boldsymbol{u} = \mathbf{1}_n/\sqrt{n}$, $\boldsymbol{v}$ be a vector with i.i.d. entries $v_i \sim \mathsf{N}(0, 1/d)$, independent of $\boldsymbol{Z}$, and $\xi = \sqrt{\beta\delta/k}$, and denote by $\mathbb{P}_{0,n}$ the law of $\bar{\boldsymbol{X}}$.

If $\beta < \beta_{\rm spect}(k,\nu,\delta)$, then $\mathbb{P}_{1,n}$ is contiguous to $\mathbb{P}_{0,n}$.

Proof. Recalling that $\boldsymbol{P} = \mathbf{1}_k\mathbf{1}_k^{\mathsf{T}}/k$, $\boldsymbol{P}_\perp = \boldsymbol{I}_k\boldsymbol{P}$, and letting $\boldsymbol{v}_0 = \boldsymbol{H}\mathbf{1}_k/\sqrt{dk}$, we have

$$\boldsymbol{X} = \xi\, \boldsymbol{u}\boldsymbol{v}_0^{\mathsf{T}} + \frac{\sqrt{\beta}}{d}\boldsymbol{W}_\perp\boldsymbol{H}_\perp^{\mathsf{T}} + \boldsymbol{Z} \equiv \xi\, \boldsymbol{u}\boldsymbol{v}_0^{\mathsf{T}} + \widetilde{\boldsymbol{Z}}\,,$$

where $\boldsymbol{W}_\perp = \boldsymbol{W}\boldsymbol{P}_\perp$ and $\boldsymbol{H}_\perp = \boldsymbol{H}\boldsymbol{P}_\perp$. Since $\boldsymbol{v}_0$ is distributes as $\boldsymbol{v}$, and independent of $\widetilde{\boldsymbol{Z}}$, it is sufficient to prove that the law of $\widetilde{\boldsymbol{Z}}_R = \boldsymbol{R}_1\widetilde{\boldsymbol{Z}}\boldsymbol{R}_2$ is contiguous to the law of $\boldsymbol{Z}$.

Note that by the law of large numbers, almost surely (see Eq. (G.14))

$$\begin{aligned}
&\lim_{n\to\infty}\frac{1}{n}\|\boldsymbol{W}_\perp\|_{\rm op}^2 = \lim_{n\to\infty}\frac{1}{n}\|\boldsymbol{W}_\perp^{\mathsf{T}}\boldsymbol{W}_\perp\|_{\rm op} = \Big\|\int(\boldsymbol{P}_\perp\boldsymbol{w})^{\otimes 2}\tilde{q}_0({\rm d}\boldsymbol{w})\Big\|_{\rm op} = \frac{1}{k(k\nu+1)}\,,\\
&\lim_{d\to\infty}\frac{1}{d}\|\boldsymbol{H}_\perp\|_{\rm op}^2 = \lim_{d\to\infty}\frac{1}{d}\|\boldsymbol{H}_\perp^{\mathsf{T}}\boldsymbol{H}_\perp\|_{\rm op} = 1\,.
\end{aligned}$$

Hence

$$\limsup_{n\to\infty}\Big\|\frac{\sqrt{\beta}}{d}\boldsymbol{W}_\perp\boldsymbol{H}_\perp^{\mathsf{T}}\Big\|_{\rm op} \le \sqrt{\frac{\beta\delta}{k(k\nu+1)}} \equiv \sqrt{\beta_\perp}\,.$$

For $\beta < \beta_{\rm spect}$, we have $\beta_\perp < \sqrt{\delta}$, and therefore the rank-k perturbation in $\widetilde{\boldsymbol{Z}}$ does not produce an outlier eigenvalue [3].

In order to prove that the law of $\widetilde{\boldsymbol{Z}}_R = \boldsymbol{R}_1\widetilde{\boldsymbol{Z}}\boldsymbol{R}_2$ is contiguous to the law of $\boldsymbol{Z}$, note that $\widetilde{\boldsymbol{Z}}_R \stackrel{\rm d}{=} (\sqrt{\beta}/d)\boldsymbol{R}_1\boldsymbol{W}_\perp\boldsymbol{H}_\perp\boldsymbol{R}_2 + \boldsymbol{Z}$. Let $\mathbb{Q}_{1,n}$ be the law of $\boldsymbol{W}_1 = \boldsymbol{R}_1\boldsymbol{W}_\perp$ and $\mathbb{Q}_{2,n}$ the law of $\boldsymbol{W}_2 = \tilde{\boldsymbol{R}}_1\boldsymbol{W}_\perp$, where $\tilde{\boldsymbol{R}}_1$ is a uniformly random orthogonal matrix (not Haar distributed). We claim that $\lim_{n\to\infty}\|\mathbb{Q}_{1,n} - \mathbb{Q}_2\|_{\rm TV} = 0$. Indeed both $\mathbb{Q}_1$ and $\mathbb{Q}_1$ are uniform conditional on $\boldsymbol{W}^{\mathsf{T}}\boldsymbol{W}/\sqrt{n} = \boldsymbol{Q}$ and $\boldsymbol{W}^{\mathsf{T}}\mathbf{1}/\sqrt{n} = \boldsymbol{b}$. However, the joint laws of $(\boldsymbol{Q}, \boldsymbol{b})$ converge in total variation to the same Gaussian limit by the local central limit theorem.

It is therefore sufficient to show that the law of $\widetilde{\boldsymbol{Z}}_{RR} = \tilde{\boldsymbol{R}}_1\widetilde{\boldsymbol{Z}}\boldsymbol{R}_2$ is contiguous to the law of $\boldsymbol{Z}$. This follows by second moment method and follows exactly as in [9]. □

Lemma G.10. *Let $\bar{\boldsymbol{X}}$ as per Eq. (G.49), with $\boldsymbol{u} = \mathbf{1}_n/\sqrt{n}$, $\boldsymbol{v}$ be a vector with i.i.d. entries $v_i \sim \mathsf{N}(0, 1/d)$, independent of $\boldsymbol{Z}$, and $\xi = \sqrt{\beta\delta/k}$, and define*

$$\bar{\boldsymbol{\Omega}} = \begin{bmatrix} \Big(1 + \frac{\delta\beta}{k(k\nu+1)}\Big)\boldsymbol{I}_d & -\sqrt{\beta}\bar{\boldsymbol{X}}^\mathsf{T}\Big(\boldsymbol{I}_n - \frac{\beta}{d(k+\delta\beta)}\mathbf{J}_n\Big) \\ -\sqrt{\beta}\Big(\boldsymbol{I}_n - \frac{\beta}{d(k+\delta\beta)}\mathbf{J}_n\Big)\bar{\boldsymbol{X}} & \Big(\beta + k(k\nu+1)\Big)\boldsymbol{I}_n - \frac{\beta^2}{d(k+\delta\beta)}\mathbf{J}_n \end{bmatrix}. \tag{G.54}$$

If $\beta < \beta_{\rm spect}(k, \nu, \delta)$, then the law of the eigenvalues of the Hessian $\boldsymbol{\Omega}$ defined in Eq. (G.44) is contiguous to the law of the eigenvalues of $\bar{\boldsymbol{\Omega}}$.

Proof. Consider the random orthogonal matrix $\boldsymbol{R} \in \mathbb{R}^{(n+d)\times(n+d)}$

$$\boldsymbol{R} = \begin{bmatrix} \boldsymbol{R}_2^\mathsf{T} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{R}_1 \end{bmatrix}$$

where $\boldsymbol{R}_1 \in \mathbb{R}^{n\times n}$, $\boldsymbol{R}_2 \in \mathbb{R}^{d\times d}$ be uniformly random (Haar distributed) orthogonal matrices conditional to $\boldsymbol{R}_1\mathbf{1} = \mathbf{1}$. Notice that the eigenvalues of $\boldsymbol{\Omega}$ are the same as the ones of $\boldsymbol{R}\boldsymbol{\Omega}\boldsymbol{R}^\mathsf{T}$. Further, we have

$$\boldsymbol{R}\boldsymbol{\Omega}\boldsymbol{R}^\mathsf{T} = \begin{bmatrix} \Big(1 + \frac{\delta\beta}{k(k\nu+1)}\Big)\boldsymbol{I}_d & -\sqrt{\beta}\boldsymbol{X}_R^\mathsf{T}\Big(\boldsymbol{I}_n - \frac{\beta}{d(k+\delta\beta)}\mathbf{J}_n\Big) \\ -\sqrt{\beta}\Big(\boldsymbol{I}_n - \frac{\beta}{d(k+\delta\beta)}\mathbf{J}_n\Big)\boldsymbol{X}_R & \Big(\beta + k(k\nu+1)\Big)\boldsymbol{I}_n - \frac{\beta^2}{d(k+\delta\beta)}\mathbf{J}_n \end{bmatrix},$$

where $\boldsymbol{X}_R = \boldsymbol{R}_1\boldsymbol{X}\boldsymbol{R}_2$ is defined as in the statement of Lemma G.9. Applying that lemma, we obtain that the law of $\boldsymbol{R}\boldsymbol{\Omega}\boldsymbol{R}^\mathsf{T}$ is contiguous to the one of $\bar{\boldsymbol{\Omega}}$, and therefore we obtain the desired contiguity for the laws of eigenvalues. □

The next lemma establishes that the simplified Hessian $\bar{\boldsymbol{\Omega}}$ is positive semidefinite.

Lemma G.11. *Let $\bar{\boldsymbol{\Omega}}$ be defined as per Eq. (G.54) where $\bar{\boldsymbol{X}} = \xi\,\boldsymbol{u}\boldsymbol{v}^\mathsf{T} + \boldsymbol{Z}$ with $\boldsymbol{u} = \mathbf{1}_n/\sqrt{n}$, $\boldsymbol{v}$ be a vector with i.i.d. entries $v_i \sim \mathsf{N}(0, 1/d)$, independent of $(Z_{ij})_{i\le n, j\le d} \sim_{i.i.d.} \mathsf{N}(0, 1/d)$, and $\xi = \sqrt{\beta\delta/k}$.*

If $\beta < \beta_{\rm spect}(k, \delta, \nu)$, then there exists $\varepsilon > 0$ such that, almost surely, $\bar{\boldsymbol{\Omega}} \succeq \varepsilon\boldsymbol{I}$ for all n large enough.

Proof. The matrix $\bar{\boldsymbol{X}}$ fits the setting of Lemma G.8 with

$$a = 1 + \frac{\delta\beta}{k(k\nu+1)}, \qquad b = \frac{\beta\delta}{k+\delta\beta},$$
$$r = \beta + k(k\nu+1), \qquad s = \frac{\beta^2\delta}{k+\delta\beta}.$$

The claim follows by checking that condition 2 in Lemma G.8 holds. Indeed we have

$$A \equiv \frac{(1-b)^2(1+\xi^2)}{r-s} = \frac{1}{\beta + (k\nu+1)(k+\beta\delta)}.$$

Hence $A < (1+\sqrt{\delta}/r)$. Further, setting $q = k(k\nu+1)$, we have

$$
\begin{aligned}
a - \frac{\beta}{r}(1+\sqrt{\delta})^2 &= 1 + \frac{\delta\beta}{q} - \frac{\beta(1+\sqrt{\delta})^2}{\beta+q} \\
&= \frac{1}{\beta+q}\Big(\frac{\delta\beta^2}{q} - 2\sqrt{\delta}\beta + q\Big) \\
&= \frac{\delta}{q(\beta+q)}\Big(\beta - \frac{q}{\sqrt{\delta}}\Big) > 0\,.
\end{aligned}
$$

(The last inequality follows since $\beta_{\text{spect}} = q/\sqrt{\delta}$.) This completes the proof. □

The proof of Theorem 5 follows immediately from the above lemmas. Since the law of the eigenvalues of $\boldsymbol{\Omega}$ is contiguous to the law of the eigenvalues of $\bar{\boldsymbol{\Omega}}$ (by Lemma G.10), and $\bar{\boldsymbol{\Omega}} \succeq \varepsilon \boldsymbol{I}$ with high probability, we have

$$
\lim_{n\to\infty} \mathbb{P}(\lambda_{\min}(\boldsymbol{\Omega}) < \varepsilon/2) = 0\,.
$$

H TAP free energy: Numerical results

H.1 Damped AMP

AMP turns out to converge poorly near the spectral threshold, i.e. for $\beta \approx \beta_{\text{spect}}$. Note that this appears to be an algorithmic problem, rather than a problem related to the free energy approximation. To alleviate this issue, we used damped AMP for our numerical simulations. Damped AMP iterations are as follows

$$
\begin{aligned}
\boldsymbol{m}^{t+1} &= (1-\gamma)\boldsymbol{m}^t + \gamma \boldsymbol{X}^{\mathsf{T}}\,\widetilde{\mathsf{F}}(\tilde{\boldsymbol{m}}^t;\widetilde{\boldsymbol{Q}}^t) - \gamma^2 \mathsf{F}(\boldsymbol{m}^t;\boldsymbol{Q}^t)\boldsymbol{K}_W^t\,, \\
\tilde{\boldsymbol{m}}^t &= (1-\gamma)\tilde{\boldsymbol{m}}^{t-1} + \gamma \boldsymbol{X}\,\mathsf{F}(\boldsymbol{m}^t;\boldsymbol{Q}^t) - \gamma^2 \widetilde{\mathsf{F}}(\tilde{\boldsymbol{m}}^{t-1};\widetilde{\boldsymbol{Q}}_{t-1})\boldsymbol{K}_H^t\,, \\
\boldsymbol{Q}^{t+1} &= \frac{1}{d}\sum_{a=1}^{n} \widetilde{\mathsf{F}}(\tilde{\boldsymbol{m}}_a^t;\widetilde{\boldsymbol{Q}}^t)^{\otimes 2}\,, \\
\widetilde{\boldsymbol{Q}}^t &= \frac{1}{d}\sum_{i=1}^{d} \mathsf{F}(\boldsymbol{m}_i^t;\boldsymbol{Q}^t)^{\otimes 2}\,.
\end{aligned}
$$

The matrices $\boldsymbol{K}_H^t$ and $\boldsymbol{K}_W^t$ are smoothed sum of Jacobian matrices and are computed as

$$
\begin{aligned}
\boldsymbol{K}_H^{t+1} &= \sum_{i=1}^{t+1}(1-\gamma)^{t-i+1}\mathsf{B}_t\,, \\
\boldsymbol{K}_W^t &= \sum_{i=1}^{t}(1-\gamma)^{t-i}\mathsf{C}_t
\end{aligned}
$$

where

$$
\begin{aligned}
(\mathsf{B}_t)_{rs} &= \frac{1}{d}\sum_{i=1}^{d} \frac{\partial \mathsf{F}_s}{\partial(\boldsymbol{m}_i^t)_r}(\boldsymbol{m}_i^t;\boldsymbol{Q}^t)\,,\\
(\mathsf{C}_t)_{rs} &= \frac{1}{d}\sum_{a=1}^{n} \frac{\partial \widetilde{\mathsf{F}}_s}{\partial(\tilde{\boldsymbol{m}}_i^t)_r}(\tilde{\boldsymbol{m}}_a^t;\widetilde{\boldsymbol{Q}}^t)\,.
\end{aligned}
$$

In these calculations, γ is the smoothing parameter that throughout our simulations is fixed to $\gamma = 0.8$.

The specific choice of this damping scheme (and –in particular– the construction of matrices $\boldsymbol{K}_H^{t+1}$, $\boldsymbol{K}_W^{t+1}$) is dictated by the fact that this specific choice admits a state evolution analysis, analogous to the one holding on the undamped case.

I Approximate Message Passing: Numerical results for $k = 3$

In Figures 16 to 19 we report our numerical results using damped AMP for the case of $k = 3$ topics. These simulations are analogous to the one presented in the main text for $k = 2$, cf. Section 4.5.

Figures 16 and 17 report results on the normalized distance from the uninformative subspace $\mathsf{V}(\widehat{\boldsymbol{H}})$, $\mathsf{V}(\widehat{\boldsymbol{W}})$. These are consistent with the claim that AMP converges to a fixed point that is significantly distant from this subspace only if $\beta > \beta_{\text{Bayes}}(k,\nu,\delta) = \beta_{\text{spect}}(k,\nu,\delta)$. In Figures 18 and 19 we present our results on the correlation between the AMP estimates $\widehat{\boldsymbol{H}}$, $\widehat{\boldsymbol{W}}$ and the true factors $\boldsymbol{H}$, $\boldsymbol{W}$. We measure this correlation through the same Binder parameter introduced in Section E.2.

J Uniqueness of the solution to (3.14)

In this appendix, we prove that the solution to (3.14) is unique under the following conjecture

Conjecture J.1. *Let $q > 0$ and $\boldsymbol{w} \in \mathbb{R}^k$ be a random variable with density $p(\boldsymbol{w}) \propto \exp\left\{-q\left\|\boldsymbol{w}\right\|_2^2\right\}\tilde{q}_0(\boldsymbol{w})$. Then*

$$
\sigma(q)\gamma(q) \le \frac{2}{q}
$$

where $\sigma(q)$ and $\gamma(q)$ are the standard deviation and skewness of $\|\boldsymbol{w}\|_2^2$.

Remark J.1. For a Gaussian random vector $\boldsymbol{z} \sim \mathcal{N}(0,(2q)^{-1}\boldsymbol{I}_k)$ so that $p(\boldsymbol{z}) \propto \exp\left\{-q\left\|\boldsymbol{z}\right\|_2^2\right\}$,

$$
\tilde{\sigma}(q)\tilde{\gamma}(q) = \frac{2}{q}
$$

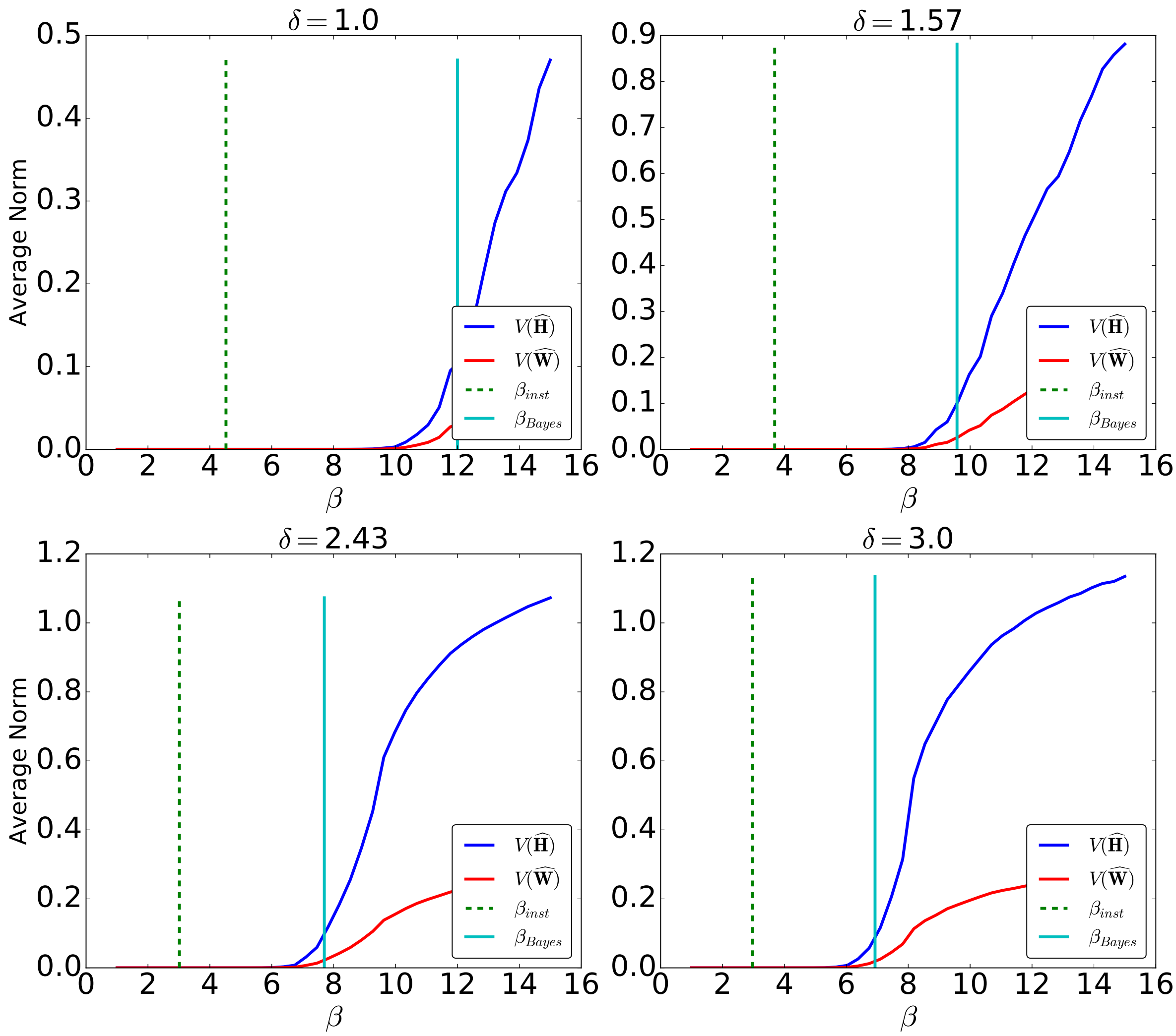

Figure 16: Normalized distances $\mathsf{V}(\widehat{\boldsymbol{H}})$, $\mathsf{V}(\widehat{\boldsymbol{W}})$ of the AMP estimates from the uninformative fixed point. Here $k = 3$, $d = 1000$ and $n = d\delta$: each data point corresponds to an average over 400 random realizations.

where $\tilde{\sigma}(q), \tilde{\gamma}_1(q)$ are the standard deviation and the skewness of $\|\boldsymbol{z}\|_2^2$.

Using the above conjecture, it can be shown that the solution to (3.14) is unique.

Let $V(q)$ be the variance of $X = \|\boldsymbol{w}\|_2^2$, when $\boldsymbol{w}$ is distributed with density $p(\boldsymbol{w}) \propto \exp\left\{-q\,\|\boldsymbol{w}\|_2^2\right\}\tilde{q}_0(\boldsymbol{w})$. Define

$$f(q) = \frac{k\beta\delta}{k-1}\left\{\mathsf{E}\left(\frac{\beta}{1+q};\nu\right) - \frac{1}{k^2}\right\}.$$

Note that using the proof of Lemma (3.2), $f(q)$ is non-negative, continuous and monotone

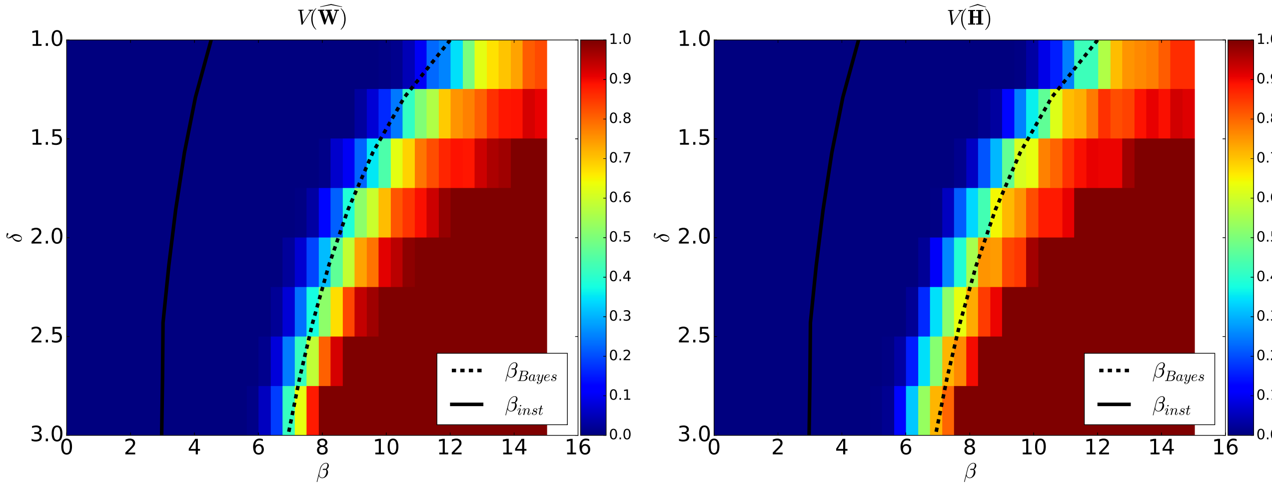

Figure 17: Empirical fraction of instances such that $\mathsf{V}(\widehat{\boldsymbol{W}}) \geq \varepsilon_0 = 5 \cdot 10^{-3}$ (left) or $\mathsf{V}(\widehat{\boldsymbol{H}}) \geq \varepsilon_0$ (right), where $\widehat{\boldsymbol{W}}, \widehat{\boldsymbol{H}}$ are the AMP estimates. Here $k = 3$, $d = 1000$, and for each (δ, β) point on a grid we ran AMP on 400 random realizations.

increasing for $q > 0$. Further,

$$f'(q) = \frac{\beta^2 \delta}{(k-1)(1+q)^2} V\left(\frac{\beta}{1+q}\right).$$

Since $f(0) > 0$, if we show that $f'(q)$ is decreasing, then for $q > q^*$ where q^* is the smallest solution to $f(q) = q$, $f'(q) < 1$. This will imply that $f(q) < q$ for $q > q^*$ that proves the uniqueness. We have

$$f''(q) = \frac{\beta^2 \delta}{(k-1)(1+q)^4} \left[-\frac{\beta}{(1+q)^2} V'\left(\frac{\beta}{1+q}\right)(1+q)^2 - 2(1+q) V\left(\frac{\beta}{1+q}\right) \right]$$

Hence, $f'(q)$ is decreasing if and only if

$$-V'\left(\frac{\beta}{1+q}\right) \leq 2\left(\frac{1+q}{\beta}\right) V\left(\frac{\beta}{1+q}\right).$$

Therefore, it is sufficient to show that for $q > 0$,

$$\frac{-V'(q)}{V(q)} \leq \frac{2}{q}.$$

Note that if we let $X = \|\boldsymbol{w}\|_2^2$ where $\boldsymbol{w}$ is as in Conjecture J.1, we have

$$V(q) = \mathbb{E}(X^2) - (\mathbb{E}X)^2.$$

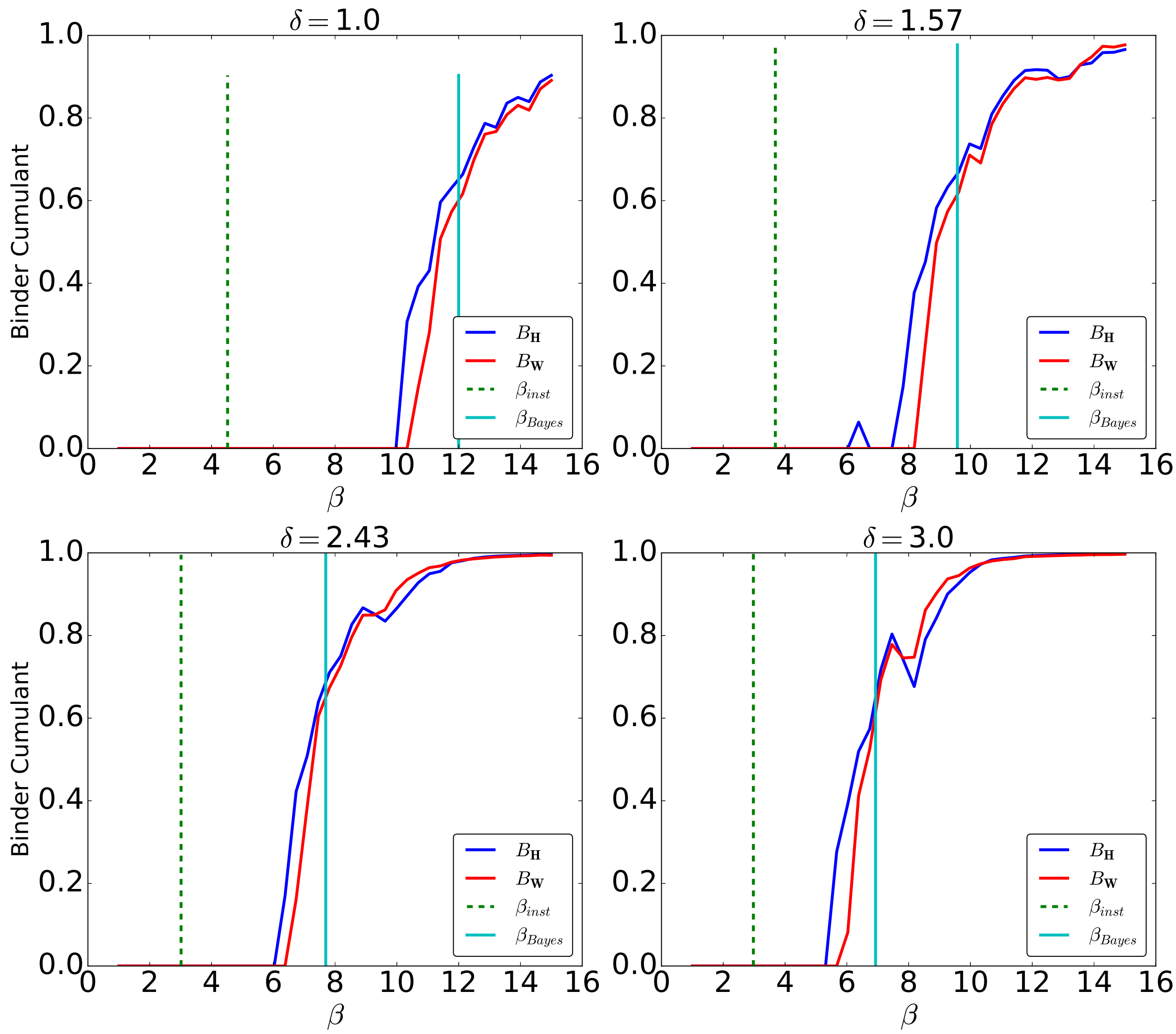

Figure 18: Binder cumulant for the correlation between AMP estimates $\widehat{\boldsymbol{W}}, \widehat{\boldsymbol{H}}$ and the true weights and topics $\boldsymbol{W}, \boldsymbol{H}$. Here $k = 3$, $d = 1000$, $n = d\delta$ and estimates are obtained by averaging over 400 realizations.

Further,

$$\begin{aligned} V'(q) &= -\mathbb{E}X^3 + (\mathbb{E}X)(\mathbb{E}X^2) - 2(\mathbb{E}X)\left[-\mathbb{E}X^2 + (\mathbb{E}X)^2\right] \\ &= -\mathbb{E}X^3 + 3(\mathbb{E}X^2)(\mathbb{E}X) - 2(\mathbb{E}X)^3. \end{aligned}$$

Hence,

$$\frac{-V'(q)}{V(q)} = \frac{\mathbb{E}(X^3) - 3(\mathbb{E}X^2)(\mathbb{E}X) + 2(\mathbb{E}X)^3}{-\mathbb{E}X^2 + (\mathbb{E}X)^2} = \sigma(q)\gamma(q) \le \frac{2}{q}$$

using Conjecture J.1. Therefore, $f(q)$ is concave and (3.14) has a unique solution in $q \in (0, \infty)$.

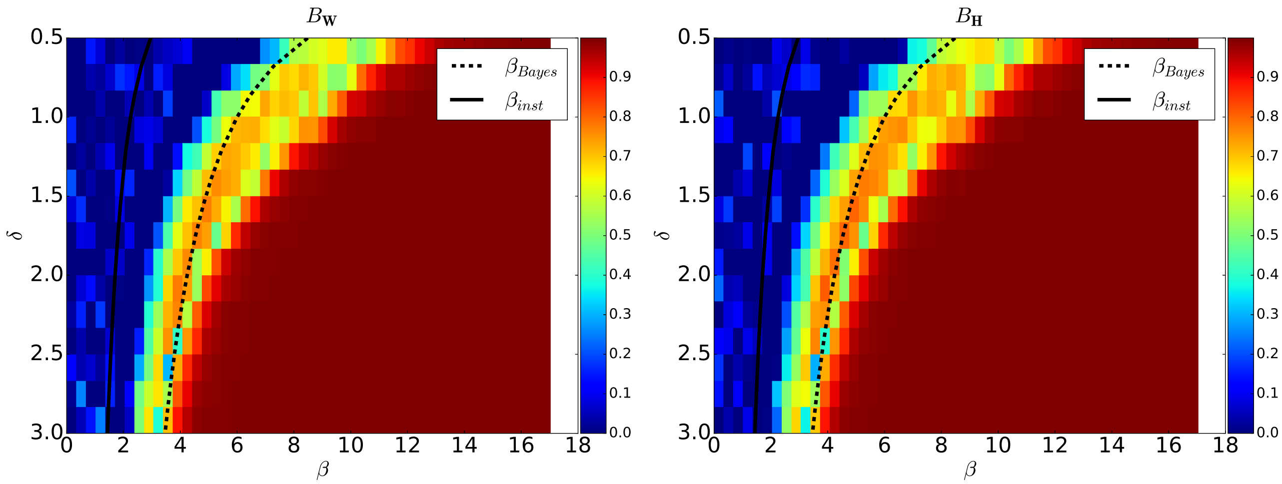

Figure 19: Binder cumulant for the correlation between AMP estimates $\widehat{\boldsymbol{W}}$, $\widehat{\boldsymbol{H}}$ and the true weights and topics $\boldsymbol{W}, \boldsymbol{H}$. Here $k = 3$, $d = 1000$ and estimates are obtained by averaging over 400 realizations.

K Extension to the Case of Dirichlet $\boldsymbol{H}$

In this appendix, we consider the model (1.2) but instead with the modification that $\boldsymbol{w}_a \sim_{iid} \mathrm{Dir}(\nu_1; k)$ and $\boldsymbol{h}_i \sim_{iid} \mathrm{Dir}(\nu_2; k)$. We briefly provide the formulae used for our experiments in this modified model.

In the modified model, the mean field iterations are still of the form presented in (3.20) and (3.21). Similarly, the AMP iterations are of the form stated in equations (4.7) and (4.8). However, Dirichlet prior should be used as $q_0(\boldsymbol{h})$ when computing the quantities $\mathsf{F}(\boldsymbol{m}_i; \boldsymbol{Q})$ and $\mathsf{G}(\boldsymbol{m}_i; \boldsymbol{Q})$.

Lemma K.1 (Counterpart to Lemma 3.2)**.** *Let q_1^* be the solution of the following equation in $(0, \infty)$*

$$q_1^* = \frac{\delta\beta}{k(k-1)}\left\{k^2\mathsf{E}\Big\{\frac{\beta}{k(k-1)}(k^2\mathsf{E}\Big(q_1^*;\nu_2\Big)-1);\nu_1\Big\}-1\right\}. \tag{K.1}$$

Further define

$$q_2^* = \delta\beta\frac{-k\mathsf{E}\Big(\tilde{q}_1^*;\nu_1\Big)+1}{k(k-1)}, \qquad \tilde{q}_1^* = \beta\frac{k^2\mathsf{E}\Big(q)1^*;\nu_2\Big)-1}{k(k-1)}, \qquad \tilde{q}_2^* = \beta\frac{-k\mathsf{E}\Big(q_1^*;\nu_2\Big)+1}{k(k-1)} \tag{K.2}$$

Then the naive mean field free energy of Eq. (3.9) admits a stationary point whereby, for all

$i \in [d]$, $a \in [n]$,

$$\boldsymbol{m}_i^* = \frac{\sqrt{\beta}}{k} (\boldsymbol{X}^\mathsf{T} \mathbf{1}_n)_i \, \mathbf{1}_k \,, \tag{K.3}$$

$$\tilde{\boldsymbol{m}}_a^* = \frac{\sqrt{\beta}}{k} (\boldsymbol{X} \mathbf{1}_d)_a \, \mathbf{1}_k \,, \tag{K.4}$$

$$\boldsymbol{Q}_i^* = q_1^* \boldsymbol{I}_k + q_2^* \boldsymbol{J}_k \,, \quad \widetilde{\boldsymbol{Q}}_a^* = \tilde{q}_1^* \boldsymbol{I}_k + \tilde{q}_2^* \boldsymbol{J}_k \,. \tag{K.5}$$

Theorem 7. *Define q_i^*, $\tilde{q}_i^*$ as in Eqs. (K.1), (K.2), and let*

$$L(\beta, k, \delta, \nu_1, \nu_2) \equiv \beta(1 + \frac{1}{\sqrt{\delta}})^2 \lambda_{\max}(D\tilde{D}) \tag{K.6}$$

where

$$D \equiv \frac{q_1^*}{\delta\beta} \boldsymbol{I}_k + \left(\frac{q_2^*}{\delta\beta} - \frac{1}{k^2}\right) \boldsymbol{J}_k, \qquad \tilde{D} \equiv \frac{\tilde{q}_1^*}{\beta} \boldsymbol{I}_k + \left(\frac{\tilde{q}_2^*}{\beta} - \frac{1}{k^2}\right) \boldsymbol{J}_k. \tag{K.7}$$

If $L(\beta, k, \delta, \nu) > 1$, then there exists $\varepsilon_1, \varepsilon_2 > 0$ such that the uninformative critical point of Lemma K.1, $(\boldsymbol{m}^*, \tilde{\boldsymbol{m}}^*, \boldsymbol{Q}^*, \widetilde{\boldsymbol{Q}}^*)$ is, with high probability, a saddle point, with index at least $n\varepsilon_1$ and $\lambda_{\min}(\mathcal{F}|_{\boldsymbol{m}^*, \tilde{\boldsymbol{m}}^*, \boldsymbol{Q}^*, \widetilde{\boldsymbol{Q}}^*}) \leq -\varepsilon_2$.

Moreover, $\lambda_{\max}(D\tilde{D})$ can be written explicitly as

$$\frac{q_1^* \tilde{q}_1^*}{\delta\beta^2} + k\left[\frac{\tilde{q}_1^* q_2^* + q_1^* \tilde{q}_2^*}{\delta\beta^2} - \frac{1}{k^2}(\frac{\tilde{q}_1^*}{\beta} + \frac{q_1^*}{\delta\beta}) + \frac{k q_2^* \tilde{q}_2^*}{\delta\beta^2} - \frac{1}{k}(\frac{q_2^*}{\delta\beta} + \frac{\tilde{q}_2^*}{\beta}) + \frac{1}{k^3}\right]_+ \tag{K.8}$$

Under the Dirichlet distributed H model, the spectral threshold has the form

$$\beta_{spect} = \frac{(k^2(k\nu_1 + 1)(k\nu_2 + 1))}{\sqrt{\delta}}. \tag{K.9}$$

For our experiments, we use this quantity as our theoretical prediction for the instability threshold of AMP.

To initialize our experiments, we need to calculate the uninformative fixed point of AMP. Similar to the case of the Gaussian $\boldsymbol{H}$, we can leverage the symmetry of the solution and get the fixed point quantities:

$$\mathsf{F}(\boldsymbol{m}; \boldsymbol{Q}) = \frac{\sqrt{\beta}}{k} \boldsymbol{J}_{d\times k}, \qquad \widetilde{\mathsf{F}}(\tilde{\boldsymbol{m}}; \widetilde{\boldsymbol{Q}}) = \frac{\sqrt{\beta}}{k} \boldsymbol{J}_{n\times k}, \tag{K.10}$$

$$\tilde{\boldsymbol{\Omega}} = \frac{\sqrt{\beta}\delta}{k(k\nu_1 + 1)} \boldsymbol{P}_\perp, \qquad \boldsymbol{\Omega} = \frac{\sqrt{\beta}}{k(k\nu_2 + 1)} \boldsymbol{P}_\perp \tag{K.11}$$

Which uniquely define the fixed point.

K.1 Additional Figures For Dirichlet $\boldsymbol{H}$ Model

In this subsection, we present additional figures describing our experimental results in the Dirichlet $\boldsymbol{H}$ case. Due to space considerations, these figures were omitted from the main text.

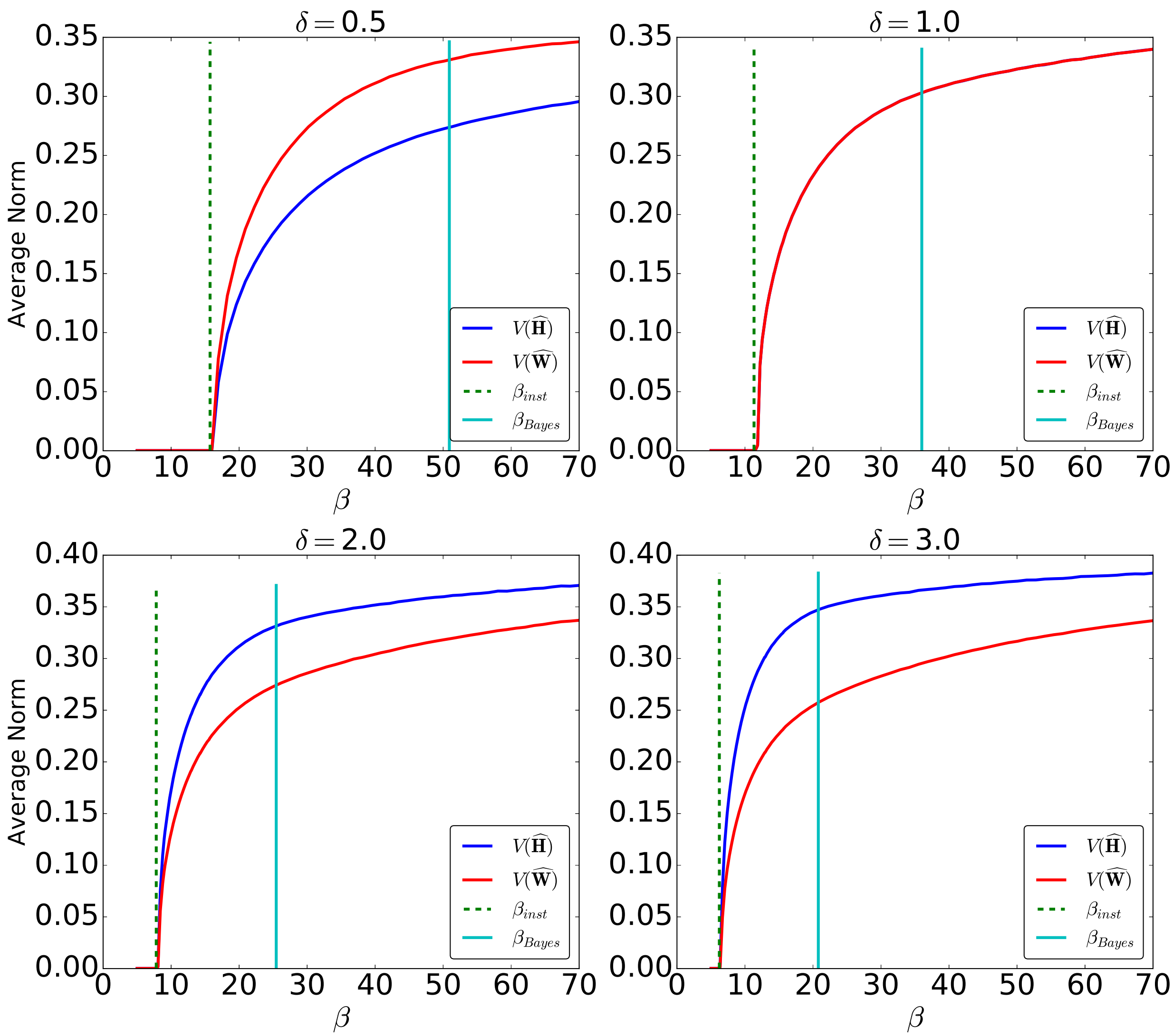

Figure 20: Normalized distances $\mathsf{V}(\widehat{\boldsymbol{H}})$, $\mathsf{V}(\widehat{\boldsymbol{W}})$ of the naive mean field estimates from the uninformative fixed point. Here $k = 2$, $\nu_1 = \nu_2 = 1$, $d = 1000$ and $n = d\delta$: each data point corresponds to an average over 400 random realizations.

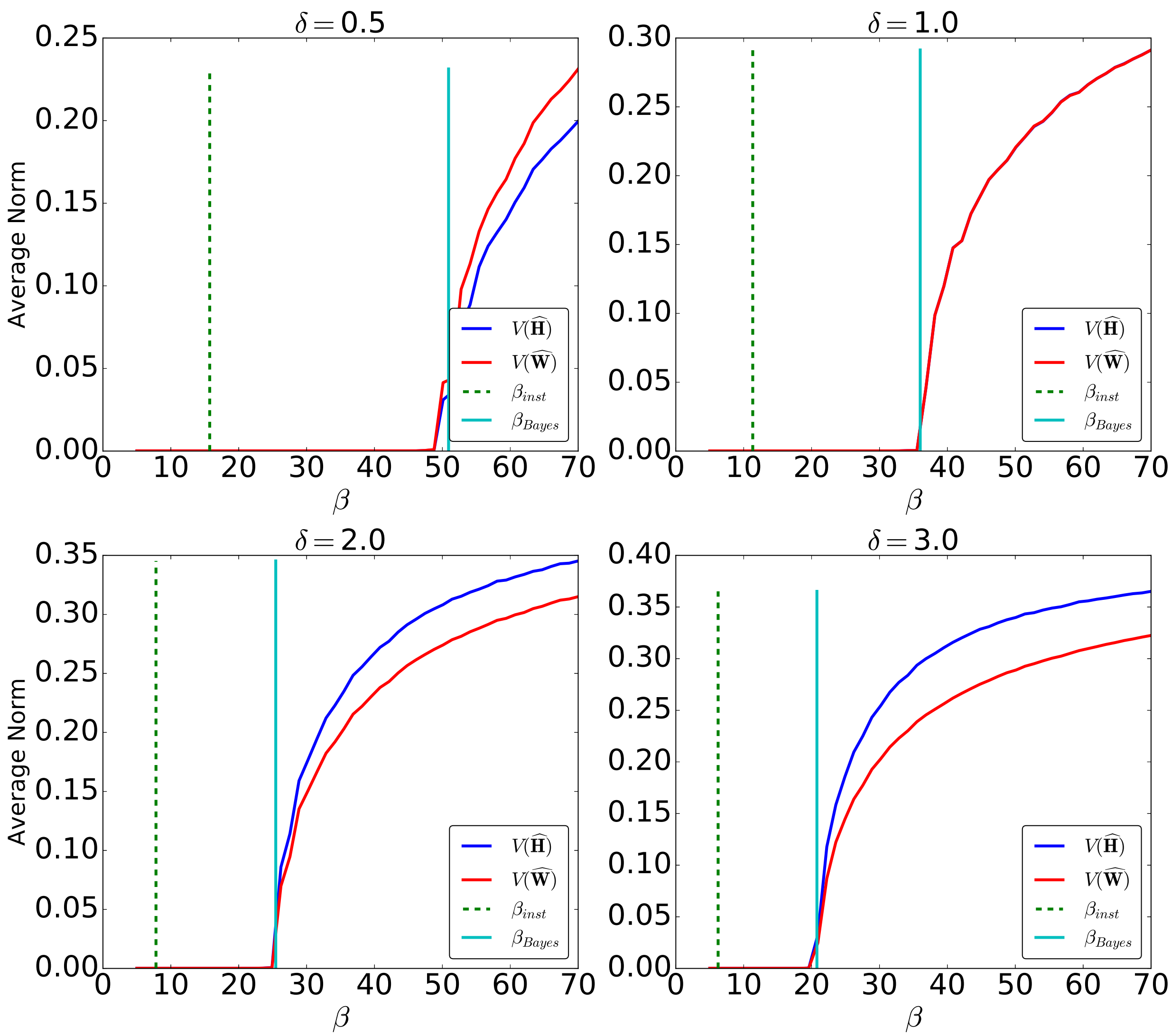

Figure 21: Normalized distances $\mathsf{V}(\widehat{\boldsymbol{H}})$, $\mathsf{V}(\widehat{\boldsymbol{W}})$ of the AMP estimates from the uninformative fixed point. Here $k = 2$, $\nu_1 = \nu_2 = 1$, $d = 1000$ and $n = d\delta$: each data point corresponds to an average over 400 random realizations.

References

[1] Greg W Anderson, Alice Guionnet, and Ofer Zeitouni. An introduction to random matrices, volume 118 of cambridge studies in advanced mathematics, 2010.

[2] Zhidong Bai and Jack W Silverstein. *Spectral analysis of large dimensional random matrices*, volume 20. Springer, 2010.

[3] Florent Benaych-Georges and Raj Rao Nadakuditi. The singular values and vectors of low rank perturbations of large rectangular random matrices. *Journal of Multivariate Analysis*, 111:120–135, 2012.

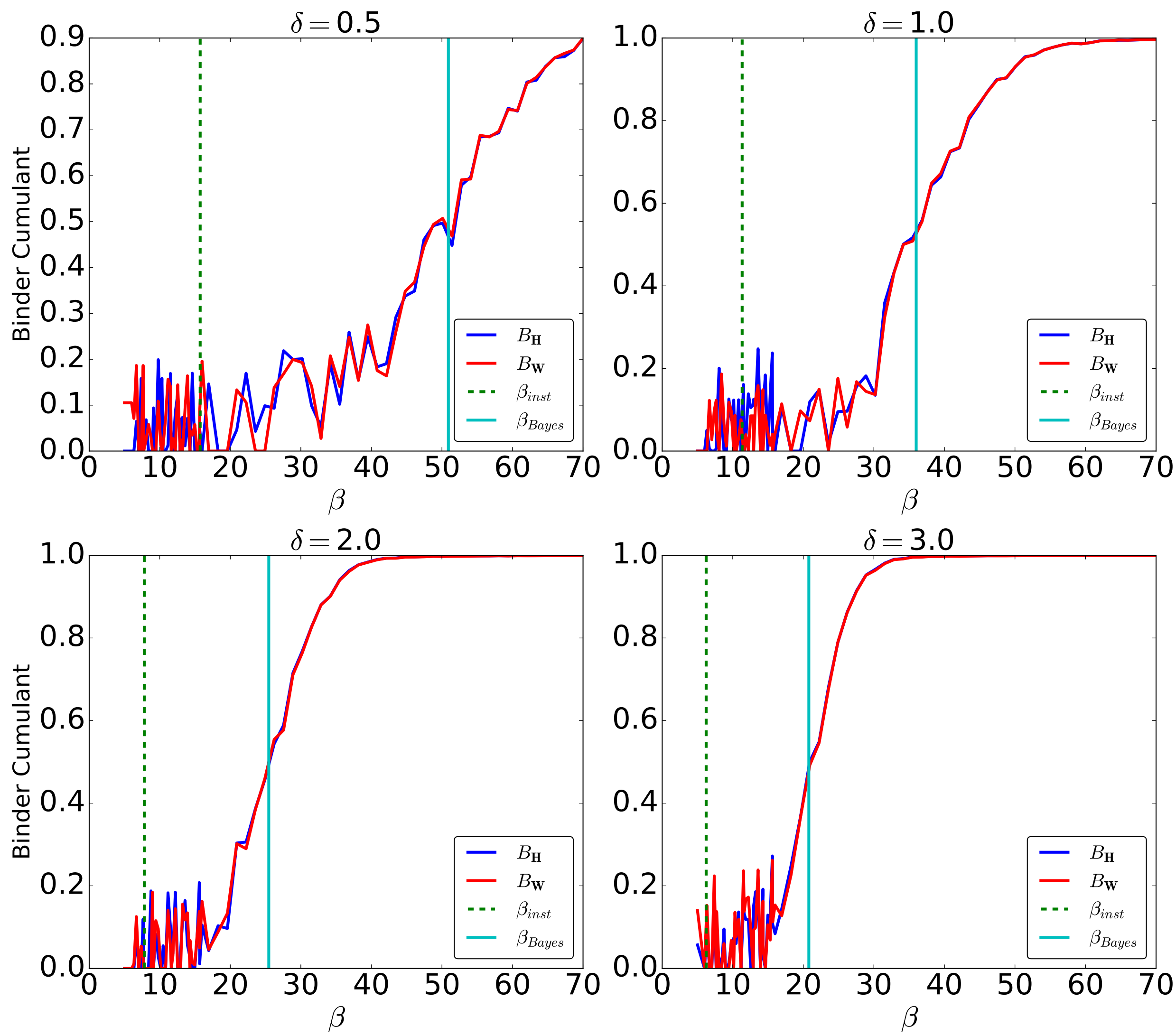

Figure 22: Binder cumulant for the correlation between the naive mean field estimates $\widehat{\boldsymbol{W}}$, $\widehat{\boldsymbol{H}}$ and the true weights and topics $\boldsymbol{W}, \boldsymbol{H}$. Here $k = 2$, $\nu_1 = \nu_2 = 1$, $d = 1000$ and estimates are obtained by averaging over 400 realizations.

[4] David M Blei, Alp Kucukelbir, and Jon D McAuliffe. Variational inference: A review for statisticians. *Journal of the American Statistical Association*, 112(518):859–877, 2017.

[5] Yash Deshpande, Emmanuel Abbe, and Emmanuel Abbe. Asymptotic mutual information for the balanced binary stochastic block model. *Information and Inference: A Journal of the IMA*, 6(2):125–170, 2016.

[6] Daphne Koller and Nir Friedman. *Probabilistic graphical models: principles and techniques*. MIT press, 2009.

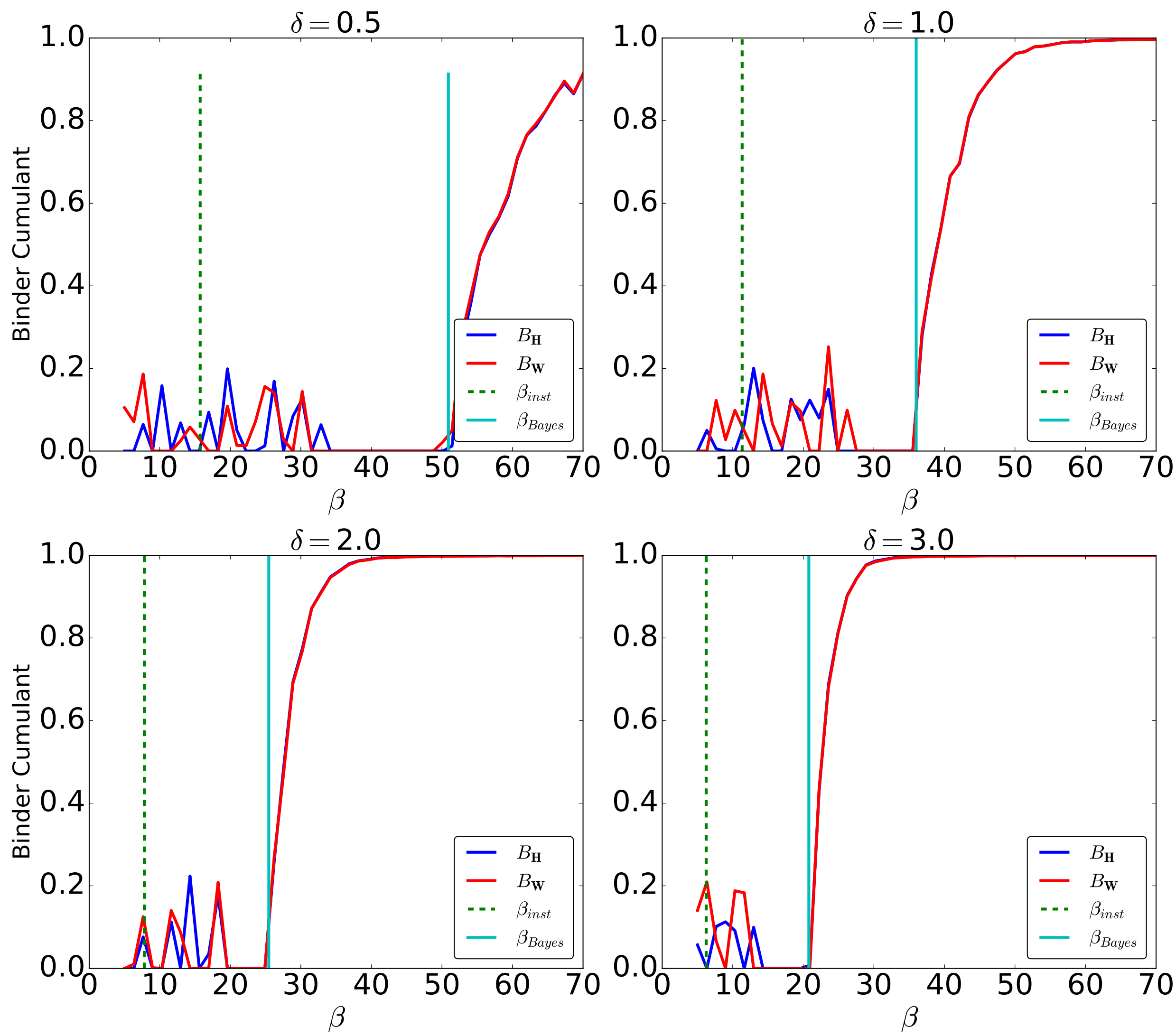

Figure 23: Binder cumulant for the correlation between AMP estimates $\widehat{\boldsymbol{W}}$, $\widehat{\boldsymbol{H}}$ and the true weights and topics $\boldsymbol{W}, \boldsymbol{H}$. Here $k = 2$, $\nu_1 = \nu_2 = 1$, $d = 1000$ and estimates are obtained by averaging over 400 realizations.

[7] Marc Mezard and Andrea Montanari. *Information, physics, and computation*. Oxford University Press, 2009.

[8] Léo Miolane. Fundamental limits of low-rank matrix estimation. *arXiv preprint arXiv:1702.00473*, 2017.

[9] Andrea Montanari, Daniel Reichman, and Ofer Zeitouni. 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, 2015.

[10] Martin J Wainwright, Michael I Jordan, et al. Graphical models, exponential families, and variational inference. *Foundations and Trends® in Machine Learning*, 1(1–2):1–305, 2008.