

Non-negative Principal Component Analysis: Message Passing Algorithms and Sharp Asymptotics

Andrea Montanari* and Emile Richard†

April 16, 2015

Abstract

Principal component analysis (PCA) aims at estimating the direction of maximal variability of a high-dimensional dataset. A natural question is: does this task become easier, and estimation more accurate, when we exploit additional knowledge on the principal vector? We study the case in which the principal vector is known to lie in the positive orthant. Similar constraints arise in a number of applications, ranging from analysis of gene expression data to spike sorting in neural signal processing.

In the unconstrained case, the estimation performances of PCA has been precisely characterized using random matrix theory, under a statistical model known as the ‘spiked model.’ It is known that the estimation error undergoes a phase transition as the signal-to-noise ratio crosses a certain threshold. Unfortunately, tools from random matrix theory have no bearing on the constrained problem. Despite this challenge, we develop an analogous characterization in the constrained case, within a one-spike model.

In particular: *(i)* We prove that the estimation error undergoes a similar phase transition, albeit at a different threshold in signal-to-noise ratio that we determine exactly; *(ii)* We prove that –unlike in the unconstrained case– estimation error depends on the spike vector, and characterize the least favorable vectors; *(iii)* We show that a non-negative principal component can be approximately computed –under the spiked model– in nearly linear time. This despite the fact that the problem is non-convex and, in general, NP-hard to solve exactly.

1 Introduction

Principal Component Analysis (PCA) is arguably the most successful of dimensionality reduction techniques. Given samples $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ from a p -dimensional distribution, $\mathbf{x}_i \in \mathbb{R}^p$, PCA seeks the direction of maximum variability. Assuming for simplicity the \mathbf{x}_i ’s to be centered (i.e. $\mathbb{E}(\mathbf{x}_i) = 0$), and denoting by \mathbf{x} a random vector distributed as \mathbf{x}_i , the objective is to estimate the solution of

$$\begin{aligned} & \text{maximize} && \mathbb{E}(\langle \mathbf{x}, \mathbf{v} \rangle^2), \\ & \text{subject to} && \|\mathbf{v}\|_2 = 1. \end{aligned} \tag{1}$$

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

†Department of Electrical Engineering, Stanford University

The solution of this problem is the principal eigenvector of the covariance matrix $\mathbb{E}(\mathbf{x}\mathbf{x}^\top)$. This is normally estimated by replacing expectation above by the sample mean, i.e. solving

$$\begin{aligned} & \text{maximize} && \sum_{i=1}^n \langle \mathbf{x}_i, \mathbf{v} \rangle^2, && \text{Classical PCA} \\ & \text{subject to} && \|\mathbf{v}\|_2 = 1. \end{aligned}$$

Denoting by $\mathbf{X} \in \mathbb{R}^{n \times p}$ the matrix with rows $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$, the solution is of course given by the principal eigenvector of the sample covariance $\mathbf{X}^\top \mathbf{X} / (n-1) = \sum_{i=1}^n \mathbf{x}_i \mathbf{x}_i^\top / (n-1)$, that we will denote by $\mathbf{v}_1 = \mathbf{v}_1(\mathbf{X})$.

This approach is known to be consistent in low dimension. Let \mathbf{v}_0 be the solution of problem (1). If $n/p \rightarrow \infty$, then $\|\mathbf{v}_1 - \mathbf{v}_0\|_2 \rightarrow 0$ in probability [And63]. On the other hand, it is well understood that consistency can break dramatically in the high-dimensional regime $n = O(p)$. This phenomenon is crisply captured by the spiked covariance model [JL04, JL09], that postulates

$$\mathbf{x}_i = \sqrt{\beta} u_{0,i} \mathbf{v}_0 + \mathbf{z}_i, \quad (2)$$

where \mathbf{v}_0 has unit norm, $\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_p$ are i.i.d. p -dimensional standard normal vectors $\mathbf{z}_i \sim \mathcal{N}(0, \mathbf{I}_p/n)$, and $\mathbf{u}_0 = (u_{0,1}, \dots, u_{0,n})^\top$ is a unit-norm vector¹. The above model can also be written as

$$\mathbf{X} = \sqrt{\beta} \mathbf{u}_0 \mathbf{v}_0^\top + \mathbf{Z}, \quad \text{Spiked Model}$$

where $\mathbf{Z} \in \mathbb{R}^n$ has i.i.d. entries $\mathbf{Z}_{ij} \sim \mathcal{N}(0, 1/n)$.

The spectral properties of the random matrix \mathbf{X} defined by the Spiked Model have been studied in detail across statistics, signal processing and probability theory [BBAP05, BS06, BS06, Pau07, FP09, BGN12, CDMF12]. In the limit $n, p \rightarrow \infty$ with $p/n \rightarrow \alpha \in (0, \infty)$, the leading eigenvector \mathbf{v}_1 undergoes a phase transition:

$$\lim_{n \rightarrow \infty} |\langle \mathbf{v}_1, \mathbf{v}_0 \rangle| = \begin{cases} 0 & \text{if } \beta \leq \sqrt{\alpha}, \\ \sqrt{\frac{1 - \alpha/\beta^2}{1 + \alpha/\beta}} & \text{if } \beta > \sqrt{\alpha}, \end{cases} \quad (3)$$

In other words, Classical PCA contains information about the signal \mathbf{v}_0 if and only if the signal-to-noise ratio is above the threshold $\sqrt{\alpha}$. Below that threshold, the principal component is asymptotically orthogonal to the signal.

The failure of PCA has motivated significant effort aimed at developing better estimation methods. A recurring idea is to use additional structural information about the principal eigenvector \mathbf{v}_0 , such as its sparsity [JL04, ZHT06] or its distribution (within a Bayesian framework) [Bis99, LU09]. Here we focus on the simplest type of structural information, namely we assume \mathbf{v}_0 is known to be non-negative². It is then natural to replace the Classical PCA problem with the following one

¹The definition of [JL04] assumes $u_i \sim_{\text{i.i.d.}} \mathcal{N}(0, 1/n)$ but for our purposes it is more convenient to consider \mathbf{u} as a given deterministic vector. Equivalently, we can condition on \mathbf{u}_0 .

²Of course the case in which $\mathbf{v}_0 \in Q$ with Q an arbitrary, known, orthant, can be reduced to the present one.

(whereby we use the matrix \mathbf{X} to represent the data):

$$\begin{aligned} & \text{maximize} && \|\mathbf{X}\mathbf{v}\|_2^2, && \text{Non-negative PCA} \\ & \text{subject to} && \mathbf{v} \geq 0, \quad \|\mathbf{v}\|_2 = 1. \end{aligned}$$

Notice that this problem is non-convex and (despite its resemblance to spectral problems) cannot be solved by standard singular value decomposition. Indeed, from a worst case perspective, it is NP-hard by reduction from maximum independent set [dKP02]. Two questions are therefore natural. First –assuming that a solution is found– does the non-negativity constraint reduce the statistical error significantly compared to the standard PCA? Given that standard PCA enjoys computational tractability, it does not make sense to add the non-negativity constraint unless this yields significant statistical improvement. Second, are there efficient algorithms to solve the Non-negative PCA problem under our statistical model? Indeed, while the problem is hard in worst case, it might be solvable in polynomial time with high probability under the Spiked Model.

In this paper we answer *positively* to both questions within the spiked covariance model. Namely denoting by \mathbf{v}^+ the solution of the Non-negative PCA problem, we provide the following contributions:

- (i) We unveil a new phase transition phenomenon concerning \mathbf{v}^+ that is analogous to the classical one, see Eq. (3). Namely, for $\beta > \sqrt{\alpha/2}$, $\langle \mathbf{v}^+, \mathbf{v}_0 \rangle$ stays bounded away from 0, while, for $\beta < \sqrt{\alpha/2}$, there exists vectors \mathbf{v}_0 such that $\langle \mathbf{v}^+, \mathbf{v}_0 \rangle \rightarrow 0$ as $n, p \rightarrow \infty$.

Non-negative PCA is superior to classical PCA in this respect since $\sqrt{\alpha/2} < \sqrt{\alpha}$ strictly.

- (ii) We prove an explicit formula for the asymptotic scalar product $\lim_{n \rightarrow \infty} \langle \mathbf{v}^+, \mathbf{v}_0 \rangle$. Non-negative PCA is superior to Classical PCA also in this respect. Namely $\langle \mathbf{v}^+, \mathbf{v}_0 \rangle$ is strictly larger than $|\langle \mathbf{v}_1, \mathbf{v}_0 \rangle|$ with high probability as $n, p \rightarrow \infty$.

Note that the non-negativity constraint breaks the rotational invariance of classical PCA (under the spiked model). As a consequence, not all spikes \mathbf{v}_0 are equally hard –or easy– to estimate. We use our theory to characterize the least favorable vectors \mathbf{v}_0 .

- (iii) We prove that (for any fixed $\delta > 0$) a $(1 - \delta)$ approximation to the non-convex optimization Non-negative PCA problem can be found efficiently with high probability with respect to the noise realization. Our algorithm has complexity of order $T_{\text{mult}} \log(1/\delta)$, where T_{mult} is the maximum of the complexity of multiplying a vector by \mathbf{X} or by \mathbf{X}^\top .

Technically, our approach has two components. We use Sudakov-Fernique inequality to upper bound the expected value of the Non-negative PCA optimization problem. We then define an iterative algorithm to solve the optimization problem, and evaluate the value achieved by the algorithm after any number t of iterations. This provides a sequence of lower bounds which we prove converge to the upper bound as the number of iterations increase.

More precisely, we use an approximate message passing (AMP) algorithm of the type introduced in [DMM09, BM11]. Each iteration requires a multiplication by \mathbf{X} and a multiplication by \mathbf{X}^\top plus some lower complexity operations. While AMP is not guaranteed to solve the Non-negative PCA problem for arbitrary matrices \mathbf{X} , we establish the following properties:

1. After any number of iterations t , the algorithm produces a running estimate $\mathbf{v}^t \in \mathbb{R}^p$ that satisfies the constraints $\mathbf{v}^t \geq 0$ and $\|\mathbf{v}^t\|_2 = 1$.

Further the limit $\lim_{n,p \rightarrow \infty} \|\mathbf{X}\mathbf{v}^t\|_2^2 = r(t)$ exists almost surely, and $r(t)$ can be computed explicitly as a function of the empirical law of entries of \mathbf{v}_0 . Analogously, the asymptotic correlation $\lim_{n,p \rightarrow \infty} \langle \mathbf{v}^t, \mathbf{v}_0 \rangle = s(t)$ can be computed explicitly.

2. Denoting by r_* the upper bound on the value of the optimization Non-negative PCA problem implied by Sudakov-Fernique inequality, we prove that $r(t) \geq (1-\delta)r_*$ for all $t \geq t_0(\delta)$ for some dimension-independent $t_0(\delta)$. This implies that Sudakov-Fernique inequality is asymptotically tight in the high-dimensional limit.
3. The asymptotic correlation converges to a limit as the number of iteration tends to infinity $s_* = \lim_{t \rightarrow \infty} s(t)$ (the convergence is, again, exponentially fast). Further, if we add the constraint $|\langle \mathbf{v}, \mathbf{v}_0 \rangle - s_*| \geq \delta$ to the Non-negative PCA optimization problem, Sudakov-Fernique's upper bound on the resulting value is asymptotically smaller than r_* for any $\delta > 0$.

This implies that $\lim_{n \rightarrow \infty} \langle \mathbf{v}^+, \mathbf{v}_0 \rangle = s_*$.

Finally, we generalize our analysis to the case of symmetric matrices, namely assuming that data consist of a $n \times n$ symmetric matrix \mathbf{X} :

$$\mathbf{X} = \beta \mathbf{v}_0 \mathbf{v}_0^\top + \mathbf{Z} \quad \text{Symmetric Spiked Model}$$

with $\mathbf{v}_0 \geq 0$, $\|\mathbf{v}_0\|_2 = 1$. Here $\mathbf{Z} = \mathbf{Z}^\top$ is a noise matrix such that $(\mathbf{Z}_{ij})_{i \leq j}$ are independent with $\mathbf{Z}_{ij} \sim \mathcal{N}(0, 1/n)$ for $i < j$ and $\mathbf{Z}_{ii} \sim \mathcal{N}(0, 2/n)$.

In this case we study the analogue of the Non-negative PCA problem, namely

$$\begin{aligned} & \text{maximize} && \langle \mathbf{v}, \mathbf{X}\mathbf{v} \rangle, && \text{Symmetric non-negative PCA} \\ & \text{subject to} && \mathbf{v} \geq 0, \quad \|\mathbf{v}\|_2 = 1. \end{aligned}$$

Note that the problem of estimating \mathbf{v}_0 from the noisy observations \mathbf{X} –distributed according to the Symmetric Spiked Model– has a different interpretation with respect to the rectangular case, and the Spiked Model. Nevertheless, it fits naturally in the same paper because the underlying mathematical techniques are essentially the same, and the nature of results are closely related.

In the rectangular case, we are interested in estimating the direction of maximal variability of the data. In the symmetric case, the rows of \mathbf{X} are not i.i.d. and hence cannot be not interpreted as samples from a common distribution³ Instead, we are taking quadratic noisy measurements of the vector \mathbf{v}_0 , namely $\mathbf{X}_{ij} = \beta \mathbf{v}_{0,i} \mathbf{v}_{0,j} + \mathbf{Z}_{ij}$, and wish to estimate the unknown vector. This type of observations is relevant –for instance– in network estimation problems. In these problems, each vertex $i \in \{1, \dots, n\}$ carries a latent variable $\mathbf{v}_{0,i}$ and for each unordered pair $\{i, j\}$ we observe a random variable whose distribution depends on $\mathbf{v}_{0,i}$ and $\mathbf{v}_{0,j}$. The so-called ‘hidden-clique’ problem is an example of this class of models, see [DM13a] and references therein.

Let us conclude this introduction by mentioning that several result of the present paper are expected to hold in far greater generality than the specific models introduced above. In particular a natural generalization would be to consider the Spiked Model or Symmetric Spiked Model with non-Gaussian noise \mathbf{Z} . The results of [KM11, DM13a, BLM12] suggests that essentially all of our results should generalize to noise with subgaussian entries, without any change in the reconstruction algorithm. We defer a rigorous proof to future work.

³Also, the matrix \mathbf{X} is not positive-definite, with high probability.

1.1 Related literature

The non-negativity constraint on principal components arises naturally in many situations: we briefly discuss a few related areas. Let us emphasize that the theoretical understanding of the methods discussed below is much more limited than for Classical PCA.

Microarray data. Microarray measurements of gene expression result in a matrix $\mathbf{X} \in \mathbb{R}^{n \times p}$ whereby \mathbf{X}_{ij} denotes the expression level of gene j in sample i . Several authors [TSS02, KBCG03, MO04, SWPN09, SN13] seek for a subset of genes that are simultaneously over-expressed (or under-expressed) in a subset of samples. Lazzeroni and Owen [LO02] propose a model of the form

$$\mathbf{X}_{ij} = \mu_0 + \sum_{k=1}^K \mu_k \boldsymbol{\rho}_i^{(k)} \boldsymbol{\kappa}_j^{(k)}, \quad (4)$$

where k indexes such gene groups (or ‘layers’), and $\boldsymbol{\rho}^{(k)}$, $\boldsymbol{\kappa}^{(k)}$ indicate the level of participation of different samples or different genes in group k . These authors assume $\boldsymbol{\rho}_i^{(k)} \boldsymbol{\kappa}_i^{(k)} \in \{0, 1\}$, but it is natural to relax this condition allowing for partial participation in group k , i.e. $\boldsymbol{\rho}_i^{(k)} \boldsymbol{\kappa}_i^{(k)} \in [0, 1]$. By a change of normalization, this constraint can be simplified to $\boldsymbol{\rho}_i^{(k)} \boldsymbol{\kappa}_i^{(k)} \geq 0$. Note a few differences with respect to our work:

- (i) We study a model with only one non-negative component. While Eq. (4) corresponds to a model with multiple $K \geq 1$ components, in practice several authors fit one ‘layer’ at a time, hence effectively reducing the problem to a single-component case.
Extending our analysis to the multiple component case will be the object of future work.
- (ii) The non-negativity constraint is imposed in the model (4) on both components. This is a relatively straightforward modification of our setting.
- (iii) Several studies (e.g. [LO02]) fit models of the form Eq. (4) using greedy optimization methods. Their conclusions are based on the unproven belief that these methods approximately solve the optimization problem. Our results (establishing convergence, with high probability, of an iterative method) provide some mathematical justification for this approach.

Neural signal processing. Neurons’ activity can be recorded through thin implanted electrodes. The resulting signal is a superposition of localized effects of single neurons (spikes). In order to reconstruct the single neuron activity, it is necessary to assign each spike to a specific neuron that created it, a process known as ‘spike sorting’ [Lew98, QNBS04, QP09]. Once spikes are aligned, the resulting data can be viewed as a matrix $\mathbf{X} = (\mathbf{X}_{ij})_{i \in [n], j \in [p]}$, where i indexes the spikes and j time (or a transform domain, e.g. wavelet domain).

In this context, principal component analysis is often used to project each row of \mathbf{X} (i.e. each recorded spike) in a low dimensional space, or decomposing it as a sum of single neurons activity, see e.g. [BYS01, ZWZ⁺04, PMMP07]. Clustering may be carried out after dimensionality reduction. Note that each spike is a sum of single neuron activity *with non-negative* coefficients. In other words, the i -th row of \mathbf{X} reads

$$\mathbf{x}_i \approx \sum_{k=1}^K u_{0,ik} \mathbf{v}_0^{(k)}, \quad (5)$$

where $\mathbf{v}_0^{(1)}, \dots, \mathbf{v}_0^{(K)}$ are the signatures of K neurons and $u_{0,ik}$ are non-negative coefficients.

Again, this corresponds to a multiple component version of the problem we study here. To the best of our knowledge, the non-negativity constraint has not been exploited in this context.

Non-negative matrix factorization. Initially introduced in the context of chemometrics [PT94, Paa97], non-negative matrix factorization attracted considerable interest because of its applications in computer vision and topic modeling. In particular, Lee and Seung [LS99] demonstrated empirically that non-negative matrix factorization successfully identifies parts of images, or topics in documents' corpora.

A mathematical model to understand these findings was put forward in [DS03] and most recently studied, for instance, in [AGKM12]. Note that these results only apply under a no-noise or very-weak noise conditions, but for multiple components. Further, the aim is to approximate the original data matrix, rather than estimating the principal components.

In this sense, non-negative matrix factorization is the farther among all related areas to the scope of our work.

Approximate Message Passing. Approximate Message Passing algorithms proved successful as a fast first-order method for compressed sensing reconstruction [DMM09]. Their definition is inspired by ideas from statistical mechanics and coding theory [TAP77, MPV87, RU08], see also [Mon12] for further background. One attractive feature of AMP algorithms is that their high-dimensional asymptotics can be characterized exactly and in close form, through 'state-evolution' [BM11, JM13, BLM12]. Several applications and generalizations were developed by Rangan [Ran11], Schniter [VS11] and collaborators.

In particular Schniter and Cevher [SC11, PSC13] apply AMP the problem of reconstructing a vector from bilinear noisy observations, a problem that is mathematically equivalent to the one explored here. These authors consider however more complex Bayesian models, and evaluate performances through empirical simulations, while we characterize a fundamental threshold phenomenon in a worst case setting. Similar ideas were applied in [KMZ13] to the problem of dictionary learning, and in [VSM13] to hyperspectral imaging. Finally, Kabashima and collaborators [KKM⁺14] study low-rank matrix reconstruction using a similar approach, but focus on the case in which the rank scales linearly with the matrix dimensions.

1.2 Organization of the paper

In Section 2 we present formally our results, both for symmetric matrices and rectangular matrices. As mentioned above, the proof is obtained by establishing an upper bound on the value of the Non-negative PCA optimization problem using Sudakov-Fernique inequality, and a lower bound by analyzing an AMP algorithm. The upper bound is outlined in Section 3. Section 4 introduces formally AMP and its analysis, hence establishing the desired lower bound as well as the convergence properties of this algorithm. Section 5 presents a numerical illustration of the phase transition phenomenon, and of the behavior of our algorithm. Finally, Section 6 contains proofs, with some technical details deferred to the appendices.

2 Main results

In this section we present formally our results. For the sake of clarity, we consider first the case of symmetric (Wigner) matrices, and then the case of rectangular (or sample covariance, Wishart) matrices. Indeed formulæ for symmetric matrices are somewhat simpler. Before doing that, it is convenient to introduce some definitions. (For basic notations, we invite the reader to consult Section 2.4.)

2.1 Definitions

Our results concern sequences of matrices \mathbf{X} with diverging dimensions n, p , and are expressed in terms of the asymptotic empirical distribution of the entries of \mathbf{v}_0 . This is formalized through the following definition.

Definition 2.1. *Let $\{\mathbf{x}(n)\}_{n \geq 0}$ be a sequence of vectors with increasing dimensions. Namely, for each n , $\mathbf{x}(n)$ is an n -dimensional vector $\mathbf{x}(n) \in \mathbb{R}^n$. Further, let μ be a (Borel) probability measure on the real line \mathbb{R} . Then we say that $\mathbf{x}(n)$ converges in empirical distribution to μ if the probability measure*

$$\mu_{\mathbf{x}(n)} \equiv \frac{1}{n} \sum_{i=1}^n \delta_{\mathbf{x}(n)_i}, \quad (6)$$

converges weakly to μ and the second moment of $\mu_{\mathbf{x}(n)}$ converges as well or, equivalently, $\|\mathbf{x}(n)\|_2^2/n \rightarrow \int x^2 \mu(dx)$.

With an abuse of terminology, we will say that $\{\mathbf{x}(n)\}_{n \geq 0}$ converges in empirical distribution to X if X is a random variable with law μ .

Given a random variable X , we let μ_X denote its law. We next define a few functions of such a law. Throughout the paper the positive part function is represented by $(x)_+ = \max(x, 0)$.

Definition 2.2. *Let V be a real non-negative random variable independent of $G \sim \mathcal{N}(0, 1)$ and $x \in \mathbb{R}_{\geq 0}$ be a real number. We define the two functions*

$$\mathbf{F}_V(x) \equiv \frac{\mathbb{E} V (xV + G)_+}{\sqrt{\mathbb{E} (xV + G)_+^2}} \quad \text{and} \quad \mathbf{G}_V(x) \equiv \frac{\mathbb{E} G (xV + G)_+}{\sqrt{\mathbb{E} (xV + G)_+^2}}. \quad (7)$$

Using \mathbf{F}_V and \mathbf{G}_V define the following ‘Rayleigh functions’

$$\mathbf{R}_V^{\text{sym}}(x) \equiv \beta \mathbf{F}_V^2(x) + 2 \mathbf{G}_V(x) \quad (8)$$

$$\mathbf{R}_V^{\text{rec}}(x, \alpha) \equiv \sqrt{1 + \beta \mathbf{F}_V(x/\sqrt{\alpha})^2} + \sqrt{\alpha} \mathbf{G}_V(x/\sqrt{\alpha}) \quad (9)$$

For $\beta \geq 0$, we also define $\mathbf{T}_V(\beta)$ as the unique non-negative solution of $x = \beta \mathbf{F}_V(x)$ and $\mathbf{S}_V(\beta, \alpha)$ as the unique non-negative solution of $x^2(1 + \beta \mathbf{F}_V(x/\sqrt{\alpha})^2) = \beta^2 \mathbf{F}_V(x/\sqrt{\alpha})^2$. (Lemma 6.3 below proves that these solutions exist and are unique.)

Note that the above functions depend on the random variable V only through its law μ_V , but we prefer the notation –say– \mathbf{F}_V to the more indirect \mathbf{F}_{μ_V} . Existence and well-definedness of \mathbf{T}_V and \mathbf{S}_V (i.e. uniqueness of the solutions of the equations $x = \beta \mathbf{F}_V(x)$ and $x^2(1 + \beta \mathbf{F}_V(x/\sqrt{\alpha})^2) =$

$\beta^2 \mathbf{F}_V(x/\sqrt{\alpha})^2$) are proved in Lemma 6.3 below. Further in Lemma 6.5 we prove that the functions $\mathbf{R}_V^{\text{sym}}$ respectively $\mathbf{R}_V^{\text{rec}}(\cdot, \alpha)$ have a unique maximum reached respectively at $\mathbf{T}_V(\beta)$ and at $\mathbf{S}_V(\beta, \alpha)$.

Our results become particularly explicit in case \mathbf{v}_0 is sparse which (in the asymptotic setting) is equivalent to $\mathbb{P}(V \neq 0)$ small. We introduce some terminology to address this case.

Definition 2.3. *Given a real random variable V , we let $\varepsilon(V) \equiv \mathbb{P}(V \neq 0)$ denote its sparsity level. We let \mathcal{P} be the set of probability measures μ supported on $\mathbb{R}_{\geq 0}$, with second moment equal to one, and, for $\varepsilon \geq 0$, $\mathcal{P}_\varepsilon \equiv \{\mu \in \mathcal{P} : \mu(\{0\}) \geq 1 - \varepsilon\}$.*

Given a function $Q : \mathcal{P} \rightarrow \mathbb{R}$, $\mu_V \mapsto Q_V$, and a number $q \in \mathbb{R}$, we write that $\lim_{\varepsilon(V) \rightarrow 0} Q_V = q$ uniformly over \mathcal{P} if

$$\lim_{\varepsilon \rightarrow 0} \inf_{\mu_V \in \mathcal{P}_\varepsilon} Q_V = \lim_{\varepsilon \rightarrow 0} \sup_{\mu_V \in \mathcal{P}_\varepsilon} Q_V = q. \quad (10)$$

In the following, we will often state that an event holds almost surely as the dimensions of the random matrix \mathbf{X} tend to infinity. It is understood that such statements hold with respect to the law of a sequence $\{\mathbf{X}_n\}_{n \geq 1}$ of *independent* random matrices distributed according to the Spiked Model or the Symmetric Spiked Model.

2.2 Symmetric matrices

For the sake of comparison, we begin by recalling some asymptotic properties of Classical PCA. Given $\mathbf{X} \in \mathbb{R}^{n \times n}$ symmetric distributed according to the Symmetric Spiked Model, we denote by $\mathbf{v}_1 = \mathbf{v}_1(\mathbf{X})$ its principal eigenvector, and by $\lambda_1 = \lambda_1(\mathbf{X})$ the corresponding eigenvalue.

This model has been studied in probability theory under the name of ‘low rank deformation of a Wigner matrix’. The following is a simplified version of the main theorem in [CDMF09].

Theorem 1 ([CDMF09]). *Let $\mathbf{X} = \beta \mathbf{v}_0 \mathbf{v}_0^\top + \mathbf{Z}$ be a rank-one deformation of the Gaussian symmetric matrix \mathbf{Z} , with $\mathbf{Z}_{ij} \sim \mathcal{N}(0, 1/n)$ independent for $i < j$, and $\|\mathbf{v}_0\|_2 = 1$. Then we have, almost surely*

$$\lim_{n \rightarrow \infty} \lambda_1(\mathbf{X}) = \begin{cases} 2 & \text{if } \beta \leq 1, \\ \beta + 1/\beta & \text{if } \beta > 1. \end{cases} \quad (11)$$

Further

$$\lim_{n \rightarrow \infty} |\langle \mathbf{v}_1, \mathbf{v}_0 \rangle| = \begin{cases} 0 & \text{if } \beta \leq 1, \\ \sqrt{1 - \beta^{-2}} & \text{if } \beta > 1. \end{cases} \quad (12)$$

Numerous refinements exist on this basic result, see for instance [CDMF09, Péc09, BGN11, BGGM11, CDMF⁺11, BGGM12, KY13, PRS13].

Our analysis provides a version of this theorem that holds for non-negative PCA, and is intriguingly similar to the original one. Its proof can be found in Appendix D.

Theorem 2. *Let $\mathbf{X} = \beta \mathbf{v}_0 \mathbf{v}_0^\top + \mathbf{Z}$ be a rank-one deformation of the symmetric Gaussian matrix \mathbf{Z} with $\mathbf{Z}_{ij} \sim \mathcal{N}(0, 1/n)$ independent for $i < j$, and $\|\mathbf{v}_0\|_2 = 1$. Further let $\lambda^+ = \lambda^+(\mathbf{X})$ be the value of the Symmetric non-negative PCA problem, and $\mathbf{v}^+ = \mathbf{v}^+(\mathbf{X})$ be any of the optimizers. Finally assume that $\mathbf{v}_0 = \mathbf{v}_0(n) \in \mathbb{R}^n$ is such that $\{\sqrt{n} \mathbf{v}_0(n)\}$ converges in empirical distribution to μ_V .*

Then (with the notation introduced in Definition 2.2), we have almost surely

$$\lim_{n \rightarrow \infty} \lambda^+(\mathbf{X}) = R_V^{sym}(\mathbf{T}_V(\beta)), \quad (13)$$

$$\lim_{n \rightarrow \infty} \langle \mathbf{v}^+, \mathbf{v}_0 \rangle = F_V(\mathbf{T}_V(\beta)). \quad (14)$$

Further, uniformly over \mathcal{P} ,

$$\lim_{\varepsilon(V) \rightarrow 0} R_V^{sym}(\mathbf{T}_V(\beta)) = \begin{cases} \sqrt{2} & \text{if } \beta \leq 1/\sqrt{2}, \\ \beta + 1/(2\beta) & \text{otherwise.} \end{cases} \quad (15)$$

and

$$\lim_{\varepsilon(V) \rightarrow 0} F_V(\mathbf{T}_V(\beta)) = \begin{cases} 0 & \text{if } \beta \leq 1/\sqrt{2}, \\ \sqrt{1 - 1/(2\beta^2)} & \text{otherwise.} \end{cases} \quad (16)$$

The statement in Theorem 2 is dependent on the empirical distribution of the entries of \mathbf{v}_0 . It is of special interest to characterize the least favorable situation, i.e. the distribution corresponding to the smallest scalar product $\langle \mathbf{v}^+, \mathbf{v}_0 \rangle$. This has two motivations: (i) to guarantee the minimum value of $\langle \mathbf{v}^+, \mathbf{v}_0 \rangle$ achieved by a solution \mathbf{v}^+ of the optimization problem Symmetric non-negative PCA and (ii) to describe the least favorable signal \mathbf{v}_0 .

The worst-case scenario is realized for a particularly simple distribution, namely 2-atoms distribution, with an atom at 0. However, unlike in classical denoising [DJ94], the worst case mixture is not obtained by setting all the allowed coordinates to non-zero. In the following Theorem we are interested in the worst case among $\bar{\varepsilon}$ -sparse signals, or equivalently in vector sequences $\{\mathbf{v}_0(n)\}_{n \geq 0}$ such that $\lim_{n \rightarrow \infty} \|\mathbf{v}_0(n)\|_0/n \leq \bar{\varepsilon}$, or $V \in \mathcal{P}_{\bar{\varepsilon}}$ since sparse signals are naturally interesting for applications.

Theorem 3. *Consider the Symmetric Spiked Model with the Symmetric non-negative PCA estimator.*

If $\beta \leq 1/\sqrt{2}$, then there exists a sequence of vectors $\{\mathbf{v}_0(n)\}_{n \geq 0}$ such that $\lim_{n \rightarrow \infty} \|\mathbf{v}_0(n)\|_0/n = 0$ and, almost surely,

$$\lim_{n \rightarrow \infty} \langle \mathbf{v}^+, \mathbf{v}_0 \rangle = 0. \quad (17)$$

For any $\beta > 1/\sqrt{2}$, there exists $\varepsilon_(\beta, \bar{\varepsilon}) \in (0, \bar{\varepsilon}]$ such that the following is true. Let V_* be the random variable with law*

$$\mu_{V_*} = (1 - \varepsilon_*)\delta_0 + \varepsilon_*\delta_{1/\sqrt{\varepsilon_*}}. \quad (18)$$

Then for any sequence of vectors $\{\mathbf{v}_0(n)\}_{n \geq 0}$ such that $\|\mathbf{v}_0(n)\|_0 \leq n\bar{\varepsilon}$ we have, almost surely,

$$\liminf_{n \rightarrow \infty} \langle \mathbf{v}^+, \mathbf{v}_0 \rangle \geq F_{V_*}(\mathbf{T}_{V_*}(\beta)). \quad (19)$$

Equality holds if $\mathbf{v}_0(n)$ is the vector with $n\varepsilon_$ non-zero entries, all equal to $1/\sqrt{n\varepsilon_*}$.*

We defer this proof to Section 6.5. The worst case mixture $\varepsilon_{\#}(\beta)$ as well as the function $F_{V_*}(\mathbf{T}_{V_*}(\beta))$ can be expressed explicitly in terms of the Gaussian distribution function, see Section 6.5.

2.3 Rectangular matrices

We develop a very similar theory for the case of rectangular matrices. Our first result characterizes the value of the Non-negative PCA problem, and the estimation error, in analogy with Theorem 2. The proof can be found in Appendix D.

Theorem 4. *Let $\mathbf{X} = \sqrt{\beta} \mathbf{u}_0 \mathbf{v}_0^\top + \mathbf{Z}$ be a rank-one deformation of the Gaussian matrix \mathbf{Z} with $\mathbf{Z}_{ij} \sim \mathcal{N}(0, 1/n)$ independent, $\mathbf{v}_0 \geq 0$, and $\|\mathbf{u}_0\|_2 = \|\mathbf{v}_0\|_2 = 1$. Further let $\sigma^+ = \sigma^+(\mathbf{X})$ be the value of the Non-negative PCA problem, and $\mathbf{v}^+ = \mathbf{v}^+(\mathbf{X})$ be any of the optimizers.*

Assume that $n, p \rightarrow \infty$ with convergent aspect ratio $p/n \rightarrow \alpha \in (0, \infty)$, and that $\mathbf{v}_0 = \mathbf{v}_0(p) \in \mathbb{R}^p$ converges in empirical distribution to μ_V .

Then (with the notation introduced in Definition 2.2), we have almost surely

$$\lim_{n \rightarrow \infty} \sigma^+(\mathbf{X}) = R_V^{\text{rec}}(\mathbf{S}_V(\beta, \alpha), \alpha), \quad (20)$$

$$\lim_{n \rightarrow \infty} \langle \mathbf{v}^+, \mathbf{v}_0 \rangle = F_V(\mathbf{S}_V(\beta, \alpha)/\sqrt{\alpha}). \quad (21)$$

Further, uniformly over \mathcal{P} ,

$$\lim_{\varepsilon(V) \rightarrow 0} R_V^{\text{rec}}(\mathbf{S}_V(\beta, \alpha), \alpha) = \begin{cases} 1 + \sqrt{\alpha/2} & \text{if } \beta \leq \sqrt{\alpha/2}, \\ \sqrt{\left(\sqrt{\beta} + \frac{\alpha}{2\sqrt{\beta}}\right) \left(\sqrt{\beta} + \frac{1}{\sqrt{\beta}}\right)} & \text{otherwise,} \end{cases} \quad (22)$$

and

$$\lim_{\varepsilon(V) \rightarrow 0} F_V(\mathbf{S}_V(\beta, \alpha)/\sqrt{\alpha}) = \begin{cases} 0 & \text{if } \beta \leq \sqrt{\alpha/2}, \\ \sqrt{(\beta^2 - \alpha/2)(\beta^2 + \beta\alpha/2)^{-1}} & \text{otherwise.} \end{cases} \quad (23)$$

Finally, in the same fashion as Theorem 3, we can characterize the worst case signals \mathbf{v}_0 .

Theorem 5. *Consider the Spiked Model, with the Non-negative PCA estimator.*

If $\beta \leq \sqrt{\alpha/2}$, then there exists a sequence of vectors $\{\mathbf{v}_0(p)\}_{p \geq 1}$ such that $\lim_{p \rightarrow \infty} \|\mathbf{v}_0(p)\|_0/p = 0$ and, almost surely,

$$\lim_{p \rightarrow \infty} \langle \mathbf{v}^+, \mathbf{v}_0 \rangle = 0. \quad (24)$$

For any $\beta > \sqrt{\alpha/2}$, there exists $\varepsilon_{\text{rec},}(\alpha, \beta, \bar{\varepsilon}) \in (0, \bar{\varepsilon}]$ such that the following is true. Let V_* be the random variable with law $(1 - \varepsilon_{\text{rec},*})\delta_0 + \varepsilon_{\text{rec},*}\delta_{1/\sqrt{\varepsilon_{\text{rec},*}}}$. Then for any sequence of vectors $\{\mathbf{v}_0(p)\}_{p \geq 1}$, $\|\mathbf{v}_0(p)\|_0 \leq p\bar{\varepsilon}$, we have. almost surely,*

$$\liminf_{p \rightarrow \infty} \langle \mathbf{v}^+, \mathbf{v}_0 \rangle \geq F_{V_*}(\mathbf{S}_{V_*}(\beta, \alpha)/\sqrt{\alpha}). \quad (25)$$

Equality holds if $\mathbf{v}_0(p)$ is the vector with $p\varepsilon_$ non-zero entries, all equal to $1/\sqrt{p\varepsilon_*}$.*

For the proof we refer to Section 6.5 which also contains explicit expressions to compute $\varepsilon_{\text{rec},\#}$.

2.4 Additional notations

We use capital boldface for matrices, e.g. $\mathbf{X}, \mathbf{Z}, \dots$ and lowercase boldface for vectors, e.g. \mathbf{x} or \mathbf{y} . The ordinary scalar product between $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$ is denoted by $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^m \mathbf{x}_i \mathbf{y}_i$. The ℓ_p norm of a vector is denoted by $\|\mathbf{x}\|_p$, and we will occasionally omit the subscript for the case $p = 2$. The ℓ_2 operator norm of the matrix \mathbf{X} is denoted by $\|\mathbf{X}\|_2$.

As usual, we write $\phi(x) = e^{-x^2/2}/\sqrt{2\pi}$ for the standard Gaussian density, and $\Phi(x) = \int_{-\infty}^x \phi(z) dz$ for the Gaussian distribution function. Finally we will say that a function $\psi : \mathbb{R}^d \rightarrow \mathbb{R}$ is pseudo-Lipschitz if there exists a constant $L > 0$ such that

$$|\psi(\mathbf{x}) - \psi(\mathbf{y})| \leq L(1 + \|\mathbf{x}\|_2 + \|\mathbf{y}\|_2)\|\mathbf{x} - \mathbf{y}\|_2. \quad (26)$$

3 Upper bounds on non-negative PCA values

As mentioned above, Theorems 2 and 4 are proved in two steps. We establish an upper bound on the value of the optimization problem by using Sudakov-Fernique inequality and prove that the bound is tight by analyzing an iterative algorithm that solves the optimization problem.

The first statement concerns the Symmetric Spiked Model.

Lemma 3.1. *Consider the Symmetric Spiked Model, and let $\mathbf{v}^+ = \mathbf{v}^+(\mathbf{X})$ be the Symmetric non-negative PCA estimator, with $\lambda^+ = \lambda^+(\mathbf{X})$ the value of the corresponding optimization problem.*

Then, under the assumptions of Theorem 2, we have

$$\limsup_{n \rightarrow \infty} \mathbb{E} \lambda^+(\mathbf{X}) \leq R_V^{sym}(\mathbf{T}_V(\beta)). \quad (27)$$

Further, there exists a deterministic function $\Delta : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}$, with $\lim_{x \rightarrow 0} \Delta(x) = 0$ such that, almost surely,

$$\limsup_{n \rightarrow \infty} |\langle \mathbf{v}^+, \mathbf{v}_0 \rangle - F_V(\mathbf{T}_V(\beta))| \leq \Delta \left(R_V^{sym}(\mathbf{T}_V(\beta)) - \liminf_{n \rightarrow \infty} \lambda^+(\mathbf{X}) \right). \quad (28)$$

The second statement concern the (non-symmetric) Spiked Model.

Lemma 3.2. *Consider the Spiked Model and let $\mathbf{v}^+ = \mathbf{v}^+(\mathbf{X})$ be the Non-negative PCA estimator, with $\sigma^+ = \sigma^+(\mathbf{X})$ the value of the corresponding optimization problem.*

Then, under the assumptions of Theorem 4, we have

$$\limsup_{n \rightarrow \infty} \mathbb{E} \sigma^+(\mathbf{X}) \leq R_V^{rec}(\mathbf{S}_V(\beta, \alpha), \alpha). \quad (29)$$

Further, there exists a deterministic function $\Delta : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}$, with $\lim_{x \rightarrow 0} \Delta(x) = 0$ such that, almost surely,

$$\limsup_{n \rightarrow \infty} |\langle \mathbf{v}^+, \mathbf{v}_0 \rangle - F_V(\mathbf{S}_V(\beta, \alpha)/\sqrt{\alpha})| \leq \Delta \left(R_V^{rec}(\mathbf{S}_V(\beta, \alpha), \alpha) - \liminf_{n \rightarrow \infty} \sigma^+(\mathbf{X}) \right). \quad (30)$$

The proof of Lemma 3.2 can be found in Section 6.2. The proof for the case of symmetric matrices, cf. Lemma 3.1, is completely analogous and we omit it.

Remark 3.3. While the above upper bounds are stated in asymptotic form, the proofs in Section 6 imply non-asymptotic upper bounds. Roughly speaking, the above upper bounds hold non-asymptotically up to an additive correction of order $1/\sqrt{n}$.

4 Approximate message passing algorithm

We use an algorithmic approach to prove a lower bound that matches the upper bound in Lemmas 3.1, 3.2. The algorithm is close in spirit to the usual power method that computes the leading eigenvector of a symmetric matrix \mathbf{X} by iterating

$$\mathbf{v}^{t+1} = \mathbf{X} \mathbf{v}^t, \quad (31)$$

from an arbitrary initialization $\mathbf{v}^0 \in \mathbb{R}^n$. Of course the power method is not well suited for the present problem, since it does not enforce the non-negativity constraint $\mathbf{v} \geq 0$. We will enforce this constraint iteratively by projecting on the feasible set. Similar non-linear power methods were studied previously, for instance in the context of sparse PCA [JNRS10, YZ13] and a statistical analysis of a method of this type was developed in [Ma13].

Our approach differs substantially from this line of work. We develop an approximate message passing (AMP) algorithm that builds on ideas from statistical physics and graphical models [DMM09, Mon12]. Remarkably, exact high-dimensional asymptotics for these algorithms have been characterized in some generality using a method known as *state evolution* [BM11, BLM12]. We establish the desired lower bounds by applying this theory to our problem.

For further background, in Appendix A, we discuss the connection of this algorithm with a graphical model formulation of the problem (but see also [DMM10, Mon12]). In Appendix B we show that AMP fixed points are stationary points of the non-negative PCA optimization problem. Let us emphasize that –while this background is provided for the reader’s convenience– it is not necessary for our results or proofs.

As before, we will start by considering the case of symmetric matrices and then move to rectangular matrices.

4.1 Symmetric matrices

4.1.1 Algorithm definition

The AMP algorithm is iterative and, after t iterations, maintains a state $\mathbf{v}^t \in \mathbb{R}^n$. We initialize it with $\mathbf{v}^0 = (1, 1, \dots, 1)^\top$, $\mathbf{v}^{-1} = (0, 0, \dots, 0)^\top$, and use the update rule, for $t \geq 0$,

$$\mathbf{v}^{t+1} = \mathbf{X}f(\mathbf{v}^t) - \mathbf{b}_t f(\mathbf{v}^{t-1}), \quad \text{AMP-sym}$$

where $\mathbf{b}_t \equiv \|(\mathbf{v}^t)_+\|_0 / \{\sqrt{n}\|(\mathbf{v}^t)_+\|_2\}$ and $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is the normalized projection on the positive orthant:

$$f(\mathbf{x}) = \sqrt{n} \frac{(\mathbf{x})_+}{\|(\mathbf{x})_+\|_2}. \quad (32)$$

(The factor \sqrt{n} is introduced here for future convenience.)

If we neglect the memory term $-\mathbf{b}_t f(\mathbf{v}^{t-1})$, the algorithm AMP-sym is extremely simple: It alternates between a power iteration, and an orthogonal projection onto the constraint set $\{\mathbf{v} : \mathbf{v} \geq 0, \|\mathbf{v}\| \leq 1\}$. As proved in [BM11, BLM12] the memory term (‘Onsager term’) plays a crucial role in allowing for an exact high-dimensional characterization.

Note that \mathbf{v}^t does not satisfy –in general– the positivity constraint. Indeed it is not the algorithm estimate of \mathbf{v}_0 . After any number t of iteration we construct the estimate

$$\hat{\mathbf{v}}^t = \frac{(\mathbf{v}^t)_+}{\|(\mathbf{v}^t)_+\|_2}. \quad (33)$$

4.1.2 Asymptotic analysis

State evolution [DMM09, BM11, JM13, BLM12] is a mathematical technique that provides an exact distributional characterization of a class of algorithms that includes AMP-sym, under suitable probabilistic models for the matrix \mathbf{X} . In the present case, we will assume the Symmetric Spiked Model, with $\sqrt{n}\mathbf{v}_0$ converging in empirical distribution to a random variable V .

Informally, state evolution predicts that as $n \rightarrow \infty$, for any fixed $t \geq 1$, the state vector \mathbf{v}^t is approximately normal with mean $\sqrt{n}\tau_t \mathbf{v}_0$ and covariance $\mathbf{I}_{n \times n}$. In other words, it can be viewed as a noisy version of the signal \mathbf{v}_0 :

$$\mathbf{v}^t \approx \sqrt{n}\tau_t \mathbf{v}_0 + \mathbf{g}, \quad \mathbf{g} \sim \mathcal{N}(0, \mathbf{I}_{n \times n}). \quad (34)$$

The signal-to-noise ratio τ_t is determined recursively by letting $\tau_1 = \beta \mathbb{E}V$ and for all $t \geq 1$, $\tau_{t+1} = F_V(\tau_t)$. Explicitly:

$$\tau_{t+1} = \beta \frac{\mathbb{E} V (\tau_t V + G)_+}{\sqrt{\mathbb{E} (\tau_t V + G)_+^2}}, \quad (35)$$

with $G \sim \mathcal{N}(0, 1)$ independent of V . A formal statement is given below.

Proposition 4.1. *Consider the Symmetric Spiked Model, and assume that $\{\sqrt{n}\mathbf{v}_0(n)\}_{n \geq 0}$ converges in empirical distribution to a random variable V . Further, let $\{\tau_t\}_{t \geq 1}$ be defined by the state evolution recursion (35).*

Then, for any pseudo-Lipschitz function $\psi : \mathbb{R}^2 \rightarrow \mathbb{R}$ and any $t \geq 1$ we have, almost surely,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \psi(v_i^t, \sqrt{n}v_{0,i}) = \mathbb{E} \{ \psi(\tau_t V + G, V) \}, \quad (36)$$

where $G \sim \mathcal{N}(0, 1)$ is independent of V . Further, the convergence in Eq. (36) also holds for $\psi(x, y) = \mathbb{I}(x \leq a)$ and any $a \in \mathbb{R}$.

The proof of this result is a direct application of the results of [BM11, JM13] and can be found in Appendix C.1.

A second important result that follows from state evolution is that the sequence $\{\mathbf{v}^t\}_{t \geq 0}$ converges in the following asymptotic sense.

Proposition 4.2. *Under the assumptions of Proposition 4.1, fix any $\ell \geq 0$. Then, we have almost surely*

$$\lim_{t \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{1}{n} \|\mathbf{v}^t - \mathbf{v}^{t+\ell}\|_2^2 = 0. \quad (37)$$

The proof of this statement is deferred to Appendix C.2.

As $t \rightarrow \infty$, $\tau_t \rightarrow \tau$, with τ the unique positive solution of the fixed point equation $\tau = \beta F_V(\tau)$. By using the above two propositions, we then obtain the following lower bound, whose proof can be found in Section 6.3.

Theorem 6. *Consider the Symmetric Spiked Model, and assume that $\{\sqrt{n}\mathbf{v}_0(n)\}_{n \geq 0}$ converges in empirical distribution to a random variable V . Further, let $\{\hat{\mathbf{v}}^t\}_{t \geq 0}$ be the AMP iterates as defined by AMP-sym and Eq. (33). Finally, let τ be the unique positive solution of the fixed point equation $\tau = \beta F_V(\tau)$ (equivalently $\tau = T_V(\beta)$).*

Then we have, almost surely,

$$\lim_{t \rightarrow \infty} \lim_{n \rightarrow \infty} \langle \hat{\mathbf{v}}^t, \mathbf{X} \hat{\mathbf{v}}^t \rangle = R_V^{\text{sym}}(\tau), \quad (38)$$

$$\lim_{t \rightarrow \infty} \lim_{n \rightarrow \infty} \langle \hat{\mathbf{v}}^t, \mathbf{v}_0 \rangle = F_V(\tau). \quad (39)$$

This provides the necessary lower bound that complements the upper bound based on Sudakov-Fernique inequality, cf. Section 3.

4.2 Rectangular matrices

4.2.1 Algorithm definition

In this case the algorithm keeps track –after t iterations– of $\mathbf{u}^t \in \mathbb{R}^n$ and $\mathbf{v}^t \in \mathbb{R}^p$. These are initialized by setting $\mathbf{v}^0 = (1, 1, \dots, 1)^\top$, $\mathbf{u}^{-1} = 0$, and updated by letting, for $t \geq 0$,

$$\begin{cases} \mathbf{u}^t = \mathbf{X} f(\mathbf{v}^t) - \mathbf{b}_t g(\mathbf{u}^{t-1}), \\ \mathbf{v}^{t+1} = \mathbf{X}^\top g(\mathbf{u}^t) - \mathbf{d}_t f(\mathbf{v}^t), \end{cases} \quad \text{AMP-rec}$$

where $\mathbf{d}_t = \sqrt{n}/\|\mathbf{u}^t\|_2$ and $\mathbf{b}_t = \|(\mathbf{v}^t)_+\|_0/(\sqrt{n}\|(\mathbf{v}^t)_+\|_2)$ and $f : \mathbb{R}^p \rightarrow \mathbb{R}^p$ and $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ are defined by:

$$f(\mathbf{x}) = \sqrt{n} \frac{(\mathbf{x})_+}{\|(\mathbf{x})_+\|_2}, \quad g(\mathbf{x}) = \sqrt{n} \frac{\mathbf{x}}{\|\mathbf{x}\|_2}. \quad (40)$$

After any number t of iteration we construct the estimates

$$\hat{\mathbf{u}}^t = \frac{\mathbf{u}^t}{\|\mathbf{u}^t\|_2}, \quad \hat{\mathbf{v}}^t = \frac{(\mathbf{v}^t)_+}{\|(\mathbf{v}^t)_+\|_2}. \quad (41)$$

These satisfy the normalization and positivity constraints and are used as estimates of $\mathbf{u}_0, \mathbf{v}_0$.

4.2.2 Asymptotic analysis

We consider the high dimensional setup where $n \rightarrow \infty$, and $p = p(n) \rightarrow \infty$ with converging aspect ratio $p/n \rightarrow \alpha \in (0, 1)$. We assume that $\{\sqrt{n} \mathbf{u}_0(n)\}_{n \geq 0}$ converges in empirical distribution to U and $\{\sqrt{p} \mathbf{v}_0(p)\}_{p \geq 0}$ converges in empirical distribution to V .

The high dimensional asymptotics of $\mathbf{u}^t, \mathbf{v}^t$ is characterized –as in the symmetric case– through state evolution. We introduce the real-valued *state evolution sequences* $\{\vartheta_t\}_{t \geq 0}$ and $\{\mu_t\}_{t \geq 1}$ through the following recursion for $t \geq 0$

$$\begin{cases} \mu_t = \sqrt{\beta} F_V\left(\frac{\vartheta_t}{\sqrt{\alpha}}\right), \\ \vartheta_{t+1} = \sqrt{\beta} \frac{\mu_t}{\sqrt{1 + \mu_t^2}}, \end{cases} \quad \text{SE-rec}$$

with initial conditions $\mu_0 = \sqrt{\beta}\mathbb{E}V$. We refer to these as to the *state evolution equations*. Roughly speaking, state evolution establishes that \mathbf{u}^t is approximately normal with mean $\sqrt{n}\mu_t\mathbf{u}_0$ and unit covariance, and \mathbf{v}^t is approximately normal with mean $\sqrt{n}\vartheta_t\mathbf{v}_0$ and unit covariance. This is formalized below.

Proposition 4.3. *Consider the Spiked Model and assume that $\{\sqrt{n}\mathbf{u}_0(n)\}_{n \geq 0}$ converges in empirical distribution to a random variable U and $\{\sqrt{p}\mathbf{v}_0(p)\}_{p \geq 0}$ converges in empirical distribution to a random variable V . Further, let $\{\mu_t\}_{t \geq 0}$, $\{\vartheta_t\}_{t \geq 1}$ be defined by the state evolution recursion SE-rec.*

Then, for any pseudo-Lipshitz function $\psi : \mathbb{R}^2 \rightarrow \mathbb{R}$ and any $t \geq 1$ we have, almost surely

$$\begin{cases} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \psi(\mathbf{u}_i^t, \sqrt{n}(\mathbf{u}_0)_i) = \mathbb{E} \{ \psi(\mu_t U + G, U) \} \\ \lim_{p \rightarrow \infty} \frac{1}{p} \sum_{i=1}^p \psi(v_i^t, \sqrt{p}v_{0,i}) = \mathbb{E} \{ \psi(\vartheta_t / \sqrt{\alpha} V + G, V) \} \end{cases} \quad (42)$$

where $G \sim \mathcal{N}(0, 1)$ is independent of U and V . Further, the convergence in Eq. (42) also holds for $\psi(x, y) = \mathbb{I}(x \leq a)$ and any $a \in \mathbb{R}$.

The proof is very similar to the one of Proposition 4.1 and is again a direct application of the results of [BM11, JM13]. We omit it to avoid redundancy.

We also have an analogous of Proposition 4.2.

Proposition 4.4. *Under the assumptions of Proposition 4.3, let $\ell \geq 0$ be a fixed integer. Then we have, almost surely,*

$$\lim_{t \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{1}{n} \|\mathbf{v}^t - \mathbf{v}^{t+\ell}\|_2 = 0, \quad \lim_{t \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{1}{n} \|\mathbf{u}^t - \mathbf{u}^{t+\ell}\|_2 = 0. \quad (43)$$

We omit the proof, as it is very similar to the one of Proposition 4.4.

In the limit $t \rightarrow \infty$ (and assuming $\varepsilon > 0$), the sequence defined in SE-rec converges to a nonzero fixed point (μ, ϑ) satisfying the fixed point equations

$$\begin{cases} \mu = \sqrt{\beta} \mathbb{F}_V \left(\frac{\vartheta}{\sqrt{\alpha}} \right), \\ \vartheta = \sqrt{\beta} \frac{\mu}{\sqrt{1 + \mu^2}}. \end{cases} \quad (44)$$

We will prove that these equations admit a unique positive solution.

Considering $t \rightarrow \infty$ (after $n \rightarrow \infty$) we can thus prove the following.

Theorem 7. *Consider the Spiked Model and assume that $\{\sqrt{n}\mathbf{u}_0(n)\}_{n \geq 0}$ converges in empirical distribution to a random variable U and $\{\sqrt{p}\mathbf{v}_0(p)\}_{p \geq 0}$ converges in empirical distribution to a random variable V . Further, let $\{\hat{\mathbf{u}}^t\}_{t \geq 0}$, $\{\hat{\mathbf{v}}^t\}_{t \geq 0}$ be the AMP estimates as defined by AMP-rec and Eq. (41). Finally, let (μ, ϑ) be the only positive solution of the fixed point equations (44).*

Then we have, almost surely,

$$\lim_{t \rightarrow \infty} \lim_{n \rightarrow \infty} \langle \hat{\mathbf{u}}^t, \mathbf{X} \hat{\mathbf{v}}^t \rangle = R_V^{\text{rec}}(\vartheta, \alpha), \quad (45)$$

$$\lim_{t \rightarrow \infty} \lim_{n \rightarrow \infty} \langle \hat{\mathbf{u}}^t, \mathbf{u}_0 \rangle = \frac{\mu}{\sqrt{1 + \mu^2}}, \quad (46)$$

$$\lim_{t \rightarrow \infty} \lim_{n \rightarrow \infty} \langle \hat{\mathbf{v}}^t, \mathbf{v}_0 \rangle = F_V \left(\frac{\vartheta}{\sqrt{\alpha}} \right). \quad (47)$$

The proof of this theorem can be found in Section 6.3.

4.3 Computational complexity

As a direct consequence of the characterization of AMP established in Propositions 4.1 and 4.3, we can upper bound the number of iterations needed for Algorithms AMP-rec and AMP-sym to converge. We point out that the cost of each step of the AMP algorithms is dominated by a matrix vector multiplication. This operation can easily be parallelized and performed efficiently.

To be definite, we state the next result in the case of symmetric matrices. A completely analogous statement holds for rectangular matrices.

Proposition 4.5. *For any law $\mu_V \in \mathcal{P}$ and any $\delta > 0$ there exists a constant $t_0(V, \delta) < \infty$ such that the following holds true. Under the assumptions of Proposition 4.1, let $\{\hat{\mathbf{v}}^t\}_{t \geq 0}$ be the sequence of estimates produced by AMP. Then, for all fixed $t \geq t_0$ we have*

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\langle \hat{\mathbf{v}}^t, \mathbf{X} \hat{\mathbf{v}}^t \rangle \geq (1 - \delta) \max_{\mathbf{v} \geq 0, \|\mathbf{v}\|=1} \langle \mathbf{v}, \mathbf{X} \mathbf{v} \rangle \right) = 1. \quad (48)$$

The proof of this statement follows immediately from Theorem 2 and 6. A more careful treatment of error terms in the latter can be used to show that –indeed– $t_0(V, \delta) \leq C(V) \log(1/\delta)$ for some finite constant $C(V)$.

Notice that the computational cost of AMP is dominated by the one of matrix vector multiplications, call it T_{mult} . The above discussion indicates that the average-case complexity of the algorithms AMP-rec and AMP-sym is $O(T_{\text{mult}} \log 1/\delta)$.

Let us also notice that Proposition 4.5 does not establish that $\hat{\mathbf{v}}^t$ converges to a solution of the non-convex optimization problem Symmetric non-negative PCA. It only proves that it converges to a near-optimizer, with essentially the same estimation error. Numerical simulations, as well as the analysis of AMP in similar problems [BM12, DM13b], suggest that indeed $\hat{\mathbf{v}}^t$ does converge to the maximizer of Symmetric non-negative PCA. We leave a proof of this fact to future work.

5 Numerical illustration

We carried out numerical simulations on synthetic data generated following Symmetric Spiked Model. We use a signal \mathbf{v}_0 that takes two values:

$$v_{0,i} = \begin{cases} \frac{1}{\sqrt{n\varepsilon}} & \text{if } i \in S, \\ 0 & \text{otherwise,} \end{cases} \quad (49)$$

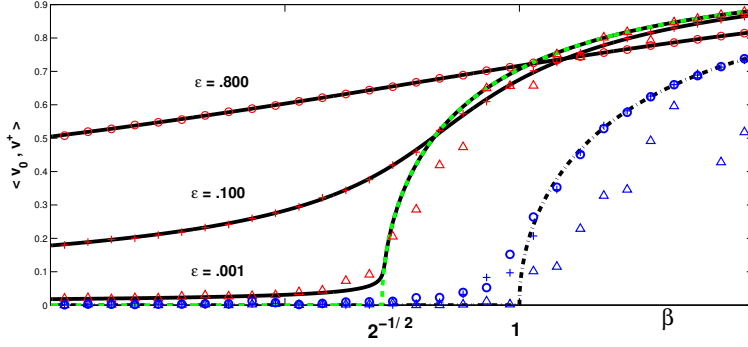


Figure 1: Numerical simulations with the Symmetric Spiked Model. Black lines represent the theoretical predictions of Theorem 2, and dots represent empirical values of $\langle \hat{\mathbf{v}}^t, \mathbf{v}_0 \rangle$ for the AMP estimator (in red) and $\langle \mathbf{v}_1, \mathbf{v}_0 \rangle$ for Classical PCA (in blue). The dashed red line corresponds to the limit behavior as $\varepsilon \rightarrow 0$. In the right hand side of the plot, blue dots and dashed black line correspond to $\langle \mathbf{v}_1, \mathbf{v}_0 \rangle$ and the theoretical prediction

where $S \subseteq [n]$ is of size $|S| = n\varepsilon$. It is immediate to see that the sequence $\{\sqrt{n}\mathbf{v}_0(n)\}_{n \geq 0}$ converges in empirical distribution to a random variable with distribution

$$V = \begin{cases} \varepsilon^{-1/2} & \text{with probability } \varepsilon, \\ 0 & \text{with probability } 1 - \varepsilon \end{cases} \quad (50)$$

In other words μ_V is the 2-points mixture $\mu_V = (1 - \varepsilon)\delta_0 + \varepsilon\delta_{1/\sqrt{\varepsilon}}$.

The predictions of Theorem 2 are stated in terms of the function $F_V \equiv F_\varepsilon$ that is rather explicit in this case. We have

$$F_\varepsilon(x) = \frac{\varepsilon B(x/\sqrt{\varepsilon})/x}{\sqrt{(1 - \varepsilon)/2 + \varepsilon(B(x/\sqrt{\varepsilon}) + \Phi(x/\sqrt{\varepsilon}))}}, \quad (51)$$

$$B(w) \equiv w^2\Phi(w) + w\phi(w). \quad (52)$$

5.1 Comparison with classical PCA

We implemented the algorithm AMP-sym, and report in Figure 5.1 the results of numerical simulations with $n = 10\,000$, sparsity level $\varepsilon \in \{0.001, 0.1, 0.8\}$, and signal-to-noise ratio $\beta \in \{0.05, 0.10, \dots, 1.5\}$. In each case we run AMP for $t = 50$ iterations and plot the empirical average of $\langle \hat{\mathbf{v}}^t, \mathbf{v}_0 \rangle$ over 32 instances. The algorithm convergence is fast and –for our purposes– this value of t is large enough so that $\tau_t \approx \mathsf{T}_V(\beta)$ and $\hat{\mathbf{v}}^t \approx \mathbf{v}^+$. (See below for further evidence of this point.)

The results agree well with the asymptotic predictions of Theorem 2, namely with the curves reporting $F_V(\mathsf{T}_V(\beta))$. The figure also illustrates that sparse vectors (small ε) correspond to the least favorable signal in small signal-to-noise ratio. The value $\beta = 1/\sqrt{2}$ corresponds to the phase transition.

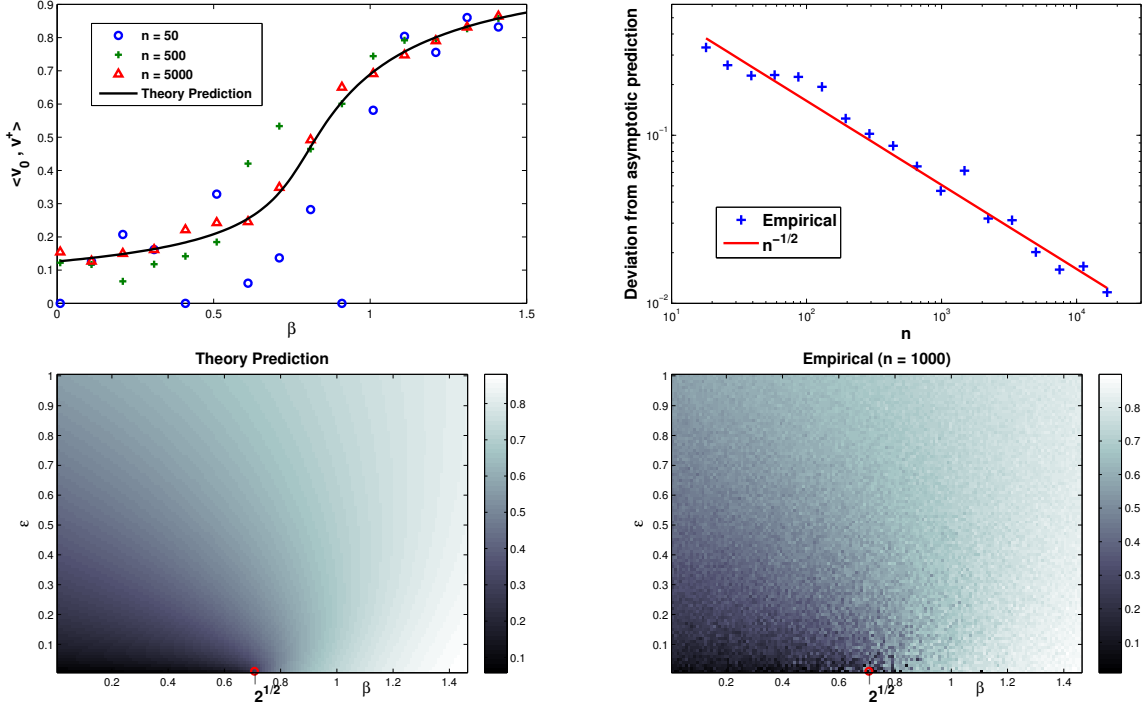


Figure 2: Comparison of theoretical prediction and empirical results for $\langle \hat{\mathbf{v}}^t, \mathbf{v}_0 \rangle \approx \langle \mathbf{v}^+, \mathbf{v}_0 \rangle$ for moderate values of n (see main text).

5.2 Deviation from the asymptotic behavior

Theorem 2 and Proposition 4.1 predict the value of $\langle \mathbf{v}_0, \mathbf{v}^+ \rangle$ and $\langle \mathbf{v}_0, \hat{\mathbf{v}}^t \rangle$ in the limit $n \rightarrow \infty$. It is natural to question the validity of the prediction for moderate values of n .

In order to investigate this point, we performed numerical experiments with AMP by generating instances of the problem for several values of n and compared the results with the asymptotic prediction of Eq. (51). The top left-hand frame in Figure 2 is obtained with $n = 50, 500, 5000$, $\varepsilon = 0.05$ and several value of β . For each point we plot the average of $\langle \hat{\mathbf{v}}^t, \mathbf{v}_0 \rangle$ after $t = 60$ iterations, over 32 instances. Already at $n = 500$ the agreement is good, and improving with n .

In the top-right plot we plot the deviation between the empirical averages of $\langle \hat{\mathbf{v}}^t, \hat{\mathbf{v}} \rangle \approx \langle \hat{\mathbf{v}}^+, \mathbf{v}_0 \rangle$ (over 32 instances) and the asymptotic prediction $F_V(\mathcal{T}_V(\beta))$. The data suggest

$$\langle \hat{\mathbf{v}}^+, \mathbf{v}_0 \rangle \approx F_V(\mathcal{T}_V(\beta)) + A n^{-b}, \quad (53)$$

with $b \approx 0.5$ and $A \approx 2$. Our theorems establish that $\langle \hat{\mathbf{v}}^+, \mathbf{v}_0 \rangle \rightarrow F_V(\mathcal{T}_V(\beta))$ in the high-dimensional limit. Developing rigorous bounds on the deviation $|\langle \hat{\mathbf{v}}^+, \mathbf{v}_0 \rangle - F_V(\mathcal{T}_V(\beta))|$ is an interesting open problem.

In the bottom frames we plot the theoretical and empirical (for $n = 1000$) values of $\langle \hat{\mathbf{v}}^t, \mathbf{v}_0 \rangle$ for a grid of parameters β, ε . The difference between the two has average $1 \cdot 10^{-3}$ and standard deviation $3 \cdot 10^{-2}$.

5.3 Comparison with a convex relaxation

A natural convex relaxation for the Symmetric non-negative PCA problem is the semi-definite program (here relation $\mathbf{W} \succeq 0$ means \mathbf{W} is positive semi-definite)

$$\begin{aligned} & \text{maximize } \langle \mathbf{X}, \mathbf{W} \rangle, \\ & \text{subject to } \mathbf{W} \succeq 0, \\ & \text{Trace}(\mathbf{W}) = 1, \\ & \mathbf{W} \geq 0. \end{aligned} \tag{SDP}$$

It is known [BAD09] that for $n \geq 5$ the completely positive cone is strictly included in the doubly non-negative cone

$$\text{conv} \left\{ \mathbf{v}\mathbf{v}^\top : \mathbf{v} \in \mathbb{R}_{\geq 0}^n \right\} \subsetneq \{ \mathbf{W} : \mathbf{W} \geq 0, \mathbf{W} \succeq 0 \}.$$

Hence in general this relaxation is not tight. The solution is a symmetric non-negative matrix $\hat{\mathbf{W}}$. We extract the leading eigenvector $\mathbf{v}_1(\hat{\mathbf{W}})$ and use its positive part as our approximation for \mathbf{v}^+ .

In simulations we use CVX [GB10] to solve SDP, and compare the result to the output of AMP stopped after $t = 50$ iterations. The interior point solver of CVX forces us to consider small problems. We use $n = 50$, $\beta = 1/\sqrt{2}$, $\varepsilon = 0.3$, and average over 50 instances.

On a 2.8 GHz Core 2 Duo with 8GB of RAM, CVX stops after about 40 seconds and a Matlab implementation of AMP after 2 ms. On average, the convex relaxation method achieves scalar product $\mathbb{E}\langle \mathbf{v}_0, \mathbf{v}_1(\hat{\mathbf{W}})_+ \rangle = 0.54 \pm 0.02$, while denoting by $\mathbf{v}_{\text{AMP}}^+$ the output of AMP, we obtain $\mathbb{E}\langle \mathbf{v}_0, \mathbf{v}_{\text{AMP}}^+ \rangle = 0.55 \pm 0.02$. In Figure 3 we compare the values reached by each algorithm over the 50 instances of the experiment with the predicted asymptotic value $F_V(\mathbf{T}_V(1/\sqrt{2})) \approx 0.53$. The plot suggests that indeed both methods solve to high accuracy the same problem.

6 Proofs

Given a random variable V , with $\mathbb{E}(V^2) < \infty$, it is useful to define the function $D_V : \mathbb{R} \rightarrow \mathbb{R}$, by

$$D_V(x) = \mathbb{E}\{(xV + G)_+^2\}. \tag{54}$$

6.1 Preliminaries

In this section we establish several useful properties of the functions F_V , G_V , R_V^{sym} , R_V^{rec} introduced in Definition 2.2. Throughout V is a random variable with law μ_V supported on $\mathbb{R}_{\geq 0}$ and such that $\mathbb{E}(V^2) = \int x^2 \mu_V(dx) = 1$. Note that, in particular, $V \neq 0$ with strictly positive probability. As before, we let $\varepsilon = \varepsilon(V) = \mathbb{P}(V \neq 0)$.

All statements concern these functions in their domain, namely F_V , G_V , $R_V^{\text{sym}} : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}$, $R_V^{\text{rec}} : \mathbb{R}_{\geq 0} \times \mathbb{R}_{> 0} \rightarrow \mathbb{R}$. Given a function $x \mapsto f(x)$, we will use $f'(x)$, $f''(x)$ to indicate its first and second derivatives.

Lemma 6.1. *Both F_V and G_V are strictly positive, differentiable and upper bounded by 1. Further F_V is strictly increasing on \mathbb{R} , with $F'_V(x) > 0$ for all $x \in \mathbb{R}$, G_V strictly decreasing on $\mathbb{R}_{\geq 0}$, with $G'_V(x) < 0$ for all $x \geq 0$, and D_V is strictly convex on \mathbb{R} .*

Finally $F_V(0) = \mathbb{E}V/\sqrt{\pi}$. and therefore $F_V(0) \in (0, \sqrt{(\varepsilon/\pi)}]$ and $G_V(0) = 1/\sqrt{2}$, $\lim_{x \rightarrow +\infty} F_V(x) = 1$, $\lim_{x \rightarrow -\infty} F_V(x) = 0$, and $\lim_{x \rightarrow \infty} G_V(x) = 0$.

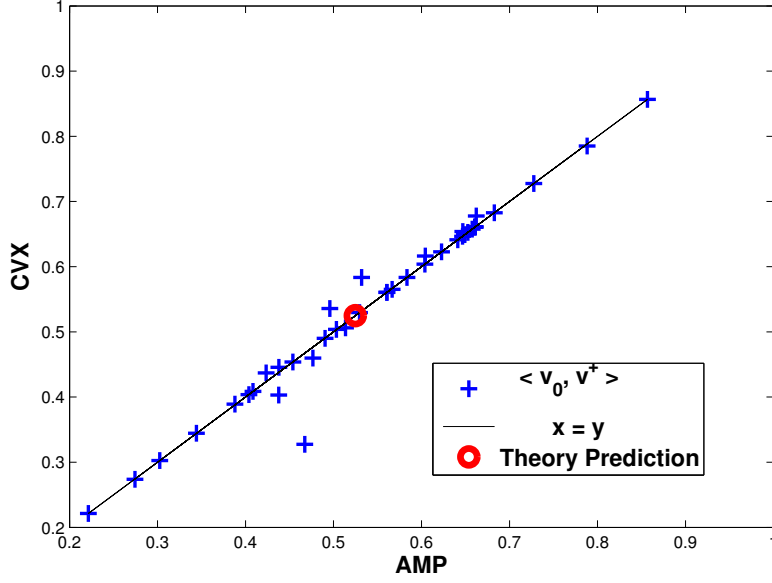


Figure 3: Comparing the AMP estimator with the estimator obtained by convex relaxation. We plot $\langle \mathbf{v}_0, \mathbf{v}_1(\hat{\mathbf{W}})_+ \rangle$ (the correlation achieved by convex optimization) versus $\langle \mathbf{v}_0, \mathbf{v}_{\text{AMP}}^+ \rangle$ (the correlation achieved by AMP), for 50 random instances.

Proof of Lemma 6.1. Positivity is immediate from the definition. The upper bound 1 follows Cauchy-Schwarz inequality. To prove differentiability, we write $F_V(x) = Y(x)/\sqrt{D_V(x)}$ with

$$Y(x) = \mathbb{E}V(xV + G)_+. \quad (55)$$

both differentiable (by dominated convergence) since V and G have bounded second moments, and strictly positive. Therefore F_V is differentiable.

A direct calculation yields the following relations

$$\frac{dD_V}{dx}(x) = 2Y(x), \quad (56)$$

$$F_V(x) = \frac{d}{dx} \sqrt{D_V(x)}, \quad (57)$$

$$\frac{dY}{dx}(x) = \mathbb{E}\{V^2 \mathbf{1}_{xV+G>0}\}, \quad (58)$$

$$D_V(x) \frac{dF_V}{dx}(x) = D_V(x) \frac{dY}{dx}(x) - \frac{1}{2}Y(x) \frac{d}{dx} D_V(x). \quad (59)$$

Using the last expression (and substituting the previous ones), we see that, to prove that F_V is increasing, it is sufficient to prove that

$$\{\mathbb{E}V(xV + G)_+\}^2 < (\mathbb{E}V^2 \mathbf{1}_{xV+G>0}) (\mathbb{E}(xV + G)_+^2) \quad , \quad (60)$$

which directly follows from Cauchy-Schwarz inequality, even for $x < 0$, and equality can not hold as V and G are independent.

In order to show that G_V is decreasing on $\mathbb{R}_{\geq 0}$ first observe that for any $x > 0$, $x F_V(x) + G_V(x) = \sqrt{D_V(x)}$. Differentiating with respect to x and using Eq. (57), we get

$$x \frac{dF_V}{dx}(x) = -\frac{dG_V}{dx}(x). \quad (61)$$

Since F_V is strictly increasing, it follows that G_V is strictly decreasing.

Finally, the values at $x = 0$ are obtained by simple calculus. The limits as $x \rightarrow \pm\infty$ follow by applying dominated convergence both to the numerator and to the denominator of $F_V(x)$ (or $G_V(x)$), after dividing both by x . \square

Lemma 6.2. *Let $n, p \in \mathbb{N}$, $\mathbf{g} \sim \mathcal{N}(0, \mathbf{I}_n)$, $\mathbf{h} \sim \mathcal{N}(0, \mathbf{I}_p)$ and, for each integer p , let $\mathbf{v}_0(p) \in \mathbb{R}^p$ be a deterministic vector with $\|\mathbf{v}_0(p)\|_2 = 1$ and such that $\{\sqrt{p} \mathbf{v}_0(p)\}_{p \geq 0}$ converges in empirical distribution to $V \in \mathcal{P}$. Similarly, for an integer n let $\mathbf{u}_0(n) \in \mathbb{R}^n$ be a deterministic vector such that $\|\mathbf{u}_0(n)\|_2 = 1$ and $\{\sqrt{n} \mathbf{u}_0(n)\}_{n \geq 0}$ converges in empirical distribution to U with $\mathbb{E} U^2 = 1$. Then, for any $b \in \mathbb{R}$ there exists a sequence $\{\delta_n(b)\}$, with $\delta_n(b) \rightarrow 0$ as $n \rightarrow \infty$ such that*

$$\mathbb{E} \left\{ \left\| \frac{1}{\sqrt{n}} \mathbf{g} + b \mathbf{u}_0 \right\|_2 \right\} \leq \sqrt{1 + b^2}, \quad (62)$$

$$\mathbb{E} \left\{ \left\| \left(\frac{1}{\sqrt{p}} \mathbf{h} + b \mathbf{v}_0 \right)_+ \right\|_2 \right\} \leq \sqrt{D_V(b)} + \delta_n. \quad (63)$$

Proof of Lemma 6.2. For Eq. (62) note that

$$\mathbb{E} \left\{ \left\| \frac{1}{\sqrt{n}} \mathbf{g} + b \mathbf{u}_0 \right\|_2 \right\} \leq \sqrt{\mathbb{E} \left\{ \left\| \frac{1}{\sqrt{n}} \mathbf{g} + b \mathbf{u}_0 \right\|_2^2 \right\}} \quad (64)$$

$$= \sqrt{1 + b^2 \|\mathbf{u}_0\|_2^2} = \sqrt{1 + b^2}. \quad (65)$$

In order to prove Eq. (63), first note that

$$\mathbb{E} \left\{ \left\| \left(\frac{1}{\sqrt{p}} \mathbf{h} + b \mathbf{v}_0 \right)_+ \right\|_2^2 \right\} \leq \mathbb{E} \left\{ \left\| \left(\frac{1}{\sqrt{p}} \mathbf{h} + b \mathbf{v}_0 \right)_+ \right\|_2^2 \right\}. \quad (66)$$

We then introduce the notation $K(x) = \mathbb{E}\{(x+G)_+^2\} = (1+x^2)\Phi(x) + x\phi(x)$ and $H(x) = K(x) - x^2$, and $\mu_p = \mu_{\mathbf{v}_0 \sqrt{p}}$ for the empirical distribution of $\{(\mathbf{v}_0)_i \sqrt{p}\}$. Note that we get

$$\mathbb{E} \left\{ \left\| \left(\frac{1}{\sqrt{p}} \mathbf{h} + b \mathbf{v}_0 \right)_+ \right\|_2^2 \right\} = \frac{1}{p} \sum_{i=1}^p K(b v_{0,i} \sqrt{p}) \quad (67)$$

$$= b^2 + \int H(b v) \mu_p(dv), \quad (68)$$

and

$$\left| \mathbb{E} \left\{ \left\| \left(\frac{1}{\sqrt{p}} \mathbf{h} + b \mathbf{v}_0 \right)_+ \right\|_2^2 \right\} - D_V(b) \right| = \left| \int H(b v) \mu_p(dv) - \int H(b v) \mu_V(dv) \right|. \quad (69)$$

Since $x \mapsto H(x)$ is bounded and Lipschitz continuous on \mathbb{R} , and by assumption μ_p converges weakly to μ , the last expression tends to 0 as $p \rightarrow \infty$, which proves our claim. \square

Lemma 6.3. *Each of the equations*

$$\beta = \frac{x}{F_V(x)}, \quad (70)$$

$$\beta = \frac{x\sqrt{1 + \beta F_V(x/\sqrt{\alpha})^2}}{F_V(x/\sqrt{\alpha})}, \quad (71)$$

admits a unique non-negative solution for each $\alpha, \beta > 0$, which we denote by $T_V(\beta)$ (for Eq. (70)) and $S_V(\beta, \alpha)$ (for Eq. (71)).

Further, we have

$$\frac{dF_V}{dx}(T_V(\beta)) \in (0, 1/\beta). \quad (72)$$

Proof of Lemma 6.3. Let us define the function $q : x \mapsto q(x) = F_V(x)/x$. We already know (by Lemma 6.1) that $F_V(0) > 0$, so $\lim_{x \rightarrow 0} q(x) = \infty$. Also, since $F_V(x) \leq 1$, we have $\lim_{x \rightarrow \infty} q(x) = 0$. Further F_V is differentiable and hence so is q on $(0, \infty)$. It is therefore sufficient to prove that q is strictly decreasing to prove existence and uniqueness of the solution of Eq. (70).

Recall that (cf. Eq. (57)):

$$F_V(x) = \frac{d}{dx} \sqrt{D_V(x)} \quad \text{where} \quad D_V(x) = \mathbb{E}\{(xV + G)_+^2\}. \quad (73)$$

We will prove that $z \mapsto D_V(\sqrt{z})$ is concave. This implies that q is decreasing: indeed, by the last equation we have

$$q(x) = 2 \frac{d}{d(x^2)} \sqrt{D_V(x)}.$$

Applying the change of variable $x = \sqrt{z}$, we get

$$\frac{d}{dx} q(x) = 4x \frac{d^2}{d(x^2)^2} \sqrt{D_V(x)} = 4\sqrt{z} \frac{d^2}{dz^2} \sqrt{D_V(\sqrt{z})} = 2\sqrt{z} \left(\frac{\frac{d^2}{dz^2} D_V(\sqrt{z})}{\sqrt{D_V(\sqrt{z})}} - \frac{(\frac{d}{dz} D_V(\sqrt{z}))^2}{2D_V(\sqrt{z})^3} \right). \quad (74)$$

This shows that the derivative of $q(x)$ is strictly negative provided that $\frac{d^2}{dz^2} D_V(\sqrt{z})$ is non-positive, or $z \mapsto D_V(\sqrt{z})$ concave. Indeed the second term in the last expression is strictly negative because $\frac{d}{dx} D_V(x) > 0$, cf. Lemma 6.1 and Eq. (57)

We can write $D_V(\sqrt{z})$ as

$$D_V(\sqrt{z}) = \int \mathbb{E}_G \{(\sqrt{z}v + G)_+^2\} d\mu_V(v),$$

(where \mathbb{E}_G denotes expectation with respect to $G \sim \mathcal{N}(0, 1)$) which, since $v \geq 0$ shows that our claim follows from concavity of $z \mapsto \mathbb{E}_G \{(\sqrt{z} + G)_+^2\} \equiv K(\sqrt{z})$, see Lemma 6.4.

Lemma 6.4. *The function $z \mapsto K(\sqrt{z})$ is concave on $\mathbb{R}_{\geq 0}$.*

Proof. We have $\mathbb{E}\{(G + x)_+^2\} = (x^2 + 1)\Phi(x) + x\phi(x)$ and $K(\sqrt{z}) = (z + 1)\Phi(\sqrt{z}) + \sqrt{z}\phi(\sqrt{z})$ so

$$\frac{d}{dz} K(\sqrt{z}) = \Phi(\sqrt{z}) + \frac{1}{\sqrt{z}} \phi(\sqrt{z}) \quad \Rightarrow \quad \frac{d^2}{dz^2} K(\sqrt{z}) = -\frac{1}{2} z^{-3/2} \phi(\sqrt{z}) < 0.$$

□

This concludes the proof that Eq. (70) admits a unique positive solution.

Consider now existence and uniqueness of solutions of Eq. (71). Note that this is equivalent to proving that for every $\beta > 0$ there exists a unique $x > 0$ such that

$$\frac{\sqrt{\alpha}}{\beta} = \mathbf{q}(x/\sqrt{\alpha}) \frac{1}{\sqrt{1 + \beta F_V(x/\sqrt{\alpha})^2}} .$$

We know that F_V is an increasing function, so $x \mapsto 1/\sqrt{1 + \beta F_V(x/\sqrt{\alpha})^2}$ is a decreasing function taking positive values. The result follows by using monotonicity of \mathbf{q} .

In order to prove Eq. (72), notice that the lower bound follows from Lemma 6.1. For the upper bound, observe that $\mathbf{q}'(x) = (F'_V(x) - \mathbf{q}(x))/x$. By evaluating it at $T_V(\beta)$, and using $\mathbf{q}(T_V(\beta)) = 1/\beta$, $\mathbf{q}'(x) \leq 0$, we get $\beta F'_V(T_V(\beta)) < 1$. \square

Lemma 6.5. *Let $T_V(\beta)$ and $S_V(\beta, \alpha)$ be defined as per Eq. (6.3).*

Then the function $x \mapsto R_V^{\text{sym}}(x)$ is strictly increasing on $(0, T_V(\beta))$ and strictly decreasing on $(T_V(\beta), +\infty)$. Similarly $x \mapsto R_V^{\text{rec}}(x, \alpha)$ is strictly increasing on $(0, S_V(\beta, \alpha))$ and strictly decreasing on $(S_V(\beta, \alpha), +\infty)$.

Proof of Lemma 6.5. Recall that letting $D_V(x) \equiv \mathbb{E}\{(xV + G)_+^2\}$ we have, for any $x \geq 0$, $x F_V(x) + G_V(x) = \sqrt{D_V(x)}$ and $F_V(x) = \frac{d}{dx} \sqrt{D_V(x)}$. As a consequence $x F'_V(x) = -G'_V(x)$, and therefore, for all $\beta, x > 0$,

$$\frac{d}{dx} R_V^{\text{sym}}(x) = 2x \left(\beta \frac{F_V(x)}{x} - 1 \right) \frac{d}{dx} F_V(x) . \quad (75)$$

Recall that, by Lemma 6.1, $F'_V(x) > 0$. Further, as per the proof of Lemma 6.3, $x \mapsto \mathbf{q}(x) = F_V(x)/x$ is strictly decreasing with $\mathbf{q}(T_V(\beta)) = 1/\beta$. This immediately implies the claim for R_V^{sym} .

The argument for $R_V^{\text{rec}}(\cdot, \alpha)$ is completely analogous. We write the derivative of $R_V^{\text{rec}}(x, \alpha)$ with respect to x :

$$\frac{\partial}{\partial x} R_V^{\text{rec}}(x, \alpha) = x \left(\beta \frac{F_V(x/\sqrt{\alpha})}{x \sqrt{1 + \beta F_V(x/\sqrt{\alpha})^2}} - 1 \right) \frac{d}{dx} F_V(x/\sqrt{\alpha}) .$$

The claim follows again from $F'_V > 0$, and using the properties of $x \mapsto F_V(x/\sqrt{\alpha})/\{x \sqrt{1 + \beta F_V(x/\sqrt{\alpha})^2}\}$ already discussed in the proof of Lemma 6.3. \square

Lemma 6.6. *Let the state evolution sequence $\{\tau_t\}_{t \geq 0}$ defined by $\tau_1 = \beta \mathbb{E}V$ and $\tau_{t+1} = \beta F_V(\tau_t)$ for all $t \geq 1$. Then, for any law μ_V , there exist constants $c_0, c_1 > 0, \gamma_0 \in (0, 1)$ such that, for all $t \geq 1$*

$$|T_V(\beta) - \tau_t| \leq c_0 \gamma_0^t \quad \text{and} \quad |R_V^{\text{sym}}(T_V(\beta)) - R_V^{\text{sym}}(\tau_t)| \leq c_1 \gamma_0^{2t} . \quad (76)$$

Proof of Lemma 6.7. We proved in Lemma 6.3 that $x \mapsto \beta F_V(x)$ is monotone increasing with $\beta F_V(x) > x$ if $x < T_V(\beta)$ and $\beta F_V(x) < x$ if $x > T_V(\beta)$. It follows that $\tau_{t+1} > \tau_t$ if $\tau_t < T_V(\beta)$ and $\tau_{t+1} < \tau_t$ if $\tau_t > T_V(\beta)$. Hence $\lim_{t \rightarrow \infty} \tau_t = T_V(\beta)$. Convergence is exponentially fast, i.e. $|T_V(\beta) - \tau_t| \leq c_0 \gamma_0^t$, since, by Lemma 6.3 $\beta F'_V(T_V(\beta)) \in (0, 1)$.

This proves the first inequality. Note that $T_V(\beta)$ is the global maximum of $x \mapsto R_V^{\text{sym}}(x)$ and hence, in a neighborhood of $T_V(\beta)$, $|R_V^{\text{sym}}(T_V(\beta)) - R_V^{\text{sym}}(\tau_t)| \leq c_*(\tau_t - T_V(\beta))^2$ \square

We state without proof the analogous result for the rectangular case. The argument is exactly the same as for the symmetric case.

Lemma 6.7. *Let the state evolution sequence $\{\mu_t, \vartheta_t\}_{t \geq 0}$ be defined by the recursion SE-rec with the initial condition $\mu_0 = \sqrt{\beta} \mathbb{E}V$. For any law μ_V there exist constants $k_0, k_1 > 0, \kappa_0 \in (0, 1)$ such that, for all $t \geq 1$,*

$$\forall t \geq 0, \quad |\mathbf{S}_V(\beta, \alpha) - \vartheta_t| \leq k_0 \kappa_0^t \quad \text{and} \quad |\mathbf{R}_V^{\text{rec}}(\mathbf{S}_V(\beta, \alpha), \alpha) - \mathbf{R}_V^{\text{rec}}(\vartheta_t, \alpha)| \leq k_1 \kappa_0^{2t}. \quad (77)$$

Our results are stated in terms of \mathbf{F}_V and \mathbf{G}_V , and depend on the law of V . However, when $\varepsilon(V) \rightarrow 0$, interestingly, two different phenomena occur. First, our results can be stated independently of law of V . Second, a phase transition occurs for a specific value of the signal-to-noise ratio β . This is stated formally below using the notion of uniform convergence introduced in Definition 2.3.

Lemma 6.8. *The following limits hold uniformly over the class \mathcal{P} of probability distributions on $\mathbb{R}_{\geq 0}$ with second moment equal to 1, and over $x \in [0, M]$ for any $M < \infty$:*

$$\lim_{\varepsilon(V) \rightarrow 0} \mathbf{D}_V(x) = \frac{1}{2} + x^2, \quad (78)$$

$$\lim_{\varepsilon(V) \rightarrow 0} \mathbf{F}_V(x) = \frac{x}{\sqrt{1/2 + x^2}}, \quad (79)$$

$$\lim_{\varepsilon(V) \rightarrow 0} \mathbf{G}_V(x) = \frac{1/2}{\sqrt{1/2 + x^2}}. \quad (80)$$

Further, again uniformly over \mathcal{P} , for any $\beta, \alpha \in \mathbb{R}_{\geq 0}$

$$\lim_{\varepsilon(V) \rightarrow 0} \mathbf{T}_V(\beta) = \begin{cases} 0 & \text{if } \beta \leq 1/\sqrt{2}, \\ \sqrt{\beta^2 - (1/2)} & \text{otherwise.} \end{cases} \quad (81)$$

$$\lim_{\varepsilon(V) \rightarrow 0} \mathbf{S}_V(\beta, \alpha) = \begin{cases} 0 & \text{if } \beta \leq \sqrt{\alpha/2}, \\ \sqrt{(\beta^2 - \alpha/2) / (1 + \beta)} & \text{otherwise.} \end{cases} \quad (82)$$

Proof. In order to prove Eq. (78) note that, by taking first the expectation over G in $\mathbf{D}_V(x) \equiv \mathbb{E}\{(xV + G)_+^2\}$, we get

$$\mathbf{D}_V(x) - \left(\frac{1}{2} + x^2\right) = \mathbb{E}\left\{(1 + x^2 V^2) \Phi(xV) + xV \phi(xV)\right\} - \left(\frac{1}{2} + x^2\right) \quad (83)$$

$$= \mathbb{E}\left\{[\Phi(xV) - \Phi(0)] + x^2 V^2 [\Phi(xV) - 1] + xV \phi(xV)\right\} \equiv \mathbb{E}\{f(xV)\}, \quad (84)$$

where $f(z) \equiv [\Phi(z) - \Phi(0)] + z^2 [\Phi(z) - 1] + z\phi(z)$. Note that $f(0) = 0$ and $f(z)$ is bounded, whence

$$\left|\mathbf{D}_V(x) - \frac{1}{2} - x^2\right| \leq \mathbb{E}\{|f(xV)| \mathbf{1}_{\{V \neq 0\}}\} \leq \|f\|_{\infty} \varepsilon, \quad (85)$$

which yields the desired uniform convergence of \mathbf{D}_V .

Next recall that $\mathbf{F}_V(x) = \frac{d}{dx} \sqrt{\mathbf{D}_V(x)}$, cf. Eq. (57) and, by Lemma 6.1, $\sqrt{\mathbf{D}_V(x)}$ is strictly convex. We hence have, for all $\delta > 0$,

$$\frac{1}{\delta} \inf_{\mu_V \in \mathcal{P}_{\varepsilon}} [\sqrt{\mathbf{D}_V(x)} - \sqrt{\mathbf{D}_V(x - \delta)}] \leq \inf_{\mu_V \in \mathcal{P}_{\varepsilon}} \mathbf{F}_V(x) \leq \sup_{\mu_V \in \mathcal{P}_{\varepsilon}} \mathbf{F}_V(x) \leq \frac{1}{\delta} \sup_{\mu_V \in \mathcal{P}_{\varepsilon}} [\sqrt{\mathbf{D}_V(x + \delta)} - \sqrt{\mathbf{D}_V(x)}]. \quad (86)$$

The claim (79) follows by taking the limit $\varepsilon \rightarrow 0$ (using Eq. (78)) followed by $\delta \rightarrow 0$. The expression of $\lim_{\varepsilon(V) \rightarrow 0} \mathbf{G}_V(x)$ follows by taking the limit on the identity $x \mathbf{F}_V(x) + \mathbf{G}_V(x) = \sqrt{\mathbf{D}_V(x)}$.

In order to prove Eq. (81), let $\mathbf{T}_0(\beta)$ denote the function on the right-hand side and assume by contradiction that there exists a sequence $\varepsilon_n \rightarrow 0$, probability measures $\mu_{V_n} \in \mathcal{P}_{\varepsilon_n}$ such that $\lim_{n \rightarrow \infty} \mathbf{T}_{V_n}(\beta) > x_* = \mathbf{T}_0(\beta) + \delta$ for some $\delta > 0$. As shown in the proof of Lemma 6.3, $x \mapsto x/\mathbf{F}_V(x)$ is monotone increasing. Using the definition we have, for all n large enough

$$\beta = \frac{\mathbf{T}_{V_n}(\beta)}{\mathbf{F}_{V_n}(\mathbf{T}_{V_n}(\beta))} \geq \frac{x_*}{\mathbf{F}_{V_n}(x_*)} \geq \frac{x_*}{\sup_{\mu_V \in \mathcal{P}_{\varepsilon_n}} \mathbf{F}_V(x_*)}. \quad (87)$$

Taking the limit $n \rightarrow \infty$, and using Eq. (79), we get

$$\beta \geq \sqrt{\frac{1}{2} + (\mathbf{T}_0(\beta) + \delta)^2}, \quad (88)$$

that yields a contradiction by the definition of \mathbf{T}_0 . Hence $\limsup_{\mu_V \in \mathcal{P}_\varepsilon} \mathbf{T}_V(\beta) \leq \mathbf{T}_0(\beta)$. The matching lower bound is proved in the same way.

Finally, the proof of Eq. (82) follows along the same lines. \square

6.2 Upper bounds: Proof of Lemma 3.2

In this section we prove Lemma 3.2. As mentioned before, the proof of Lemma 3.1 is completely analogous and omitted.

For $\mu \in [0, 1]$, we define

$$\mathcal{W}_\mu \equiv \{(\mathbf{u}, \mathbf{v}) \in \mathbb{R}^n \times \mathbb{R}^p : \|\mathbf{u}\|_2 = 1, \|\mathbf{v}\|_2 = 1, \mathbf{v} \geq 0, \langle \mathbf{v}, \mathbf{v}_0 \rangle = \mu\}, \quad (89)$$

$$M_{\mathbf{X}}(\mu) \equiv \max \{ \langle \mathbf{u}, \mathbf{X} \mathbf{v} \rangle : (\mathbf{u}, \mathbf{v}) \in \mathcal{W}_\mu \}, \quad (90)$$

$$\overline{M}(\mu) \equiv \mathbb{E} M_{\mathbf{X}}(\mu) = \mathbb{E} \max \{ \langle \mathbf{u}, \mathbf{X} \mathbf{v} \rangle : (\mathbf{u}, \mathbf{v}) \in \mathcal{W}_\mu \}. \quad (91)$$

Note that

$$\sigma^+(\mathbf{X}) = \max_{\mu \in [0, 1]} M_{\mathbf{X}}(\mu) = M_{\mathbf{X}}(\langle \mathbf{v}^+, \mathbf{v}_0 \rangle). \quad (92)$$

The function $\mathbf{X} \mapsto M_{\mathbf{X}}(\mu)$ is Lipschitz continuous with Lipschitz constant 1 (namely $|M_{\mathbf{X}}(\mu) - M_{\mathbf{X}'}(\mu)| \leq \|\mathbf{X} - \mathbf{X}'\|_F$). Hence, by Gaussian isoperimetry, we have

$$\mathbb{P} \left\{ |M_{\mathbf{X}}(\mu) - \overline{M}(\mu)| \geq t \right\} \leq 2 e^{-nt^2/2}. \quad (93)$$

Further we claim that $\mu \mapsto M_{\mathbf{X}}(\mu)$ is uniformly continuous for $\mu \in [0, 1]$. Indeed if $\mathbf{v}(\mu) = \sqrt{1 - \mu^2} \mathbf{v}_\perp(\mu) + \mu \mathbf{v}_0$ realizes the maximum over \mathcal{W}_μ (with $\langle \mathbf{v}_\perp(\mu), \mathbf{v}_0 \rangle = 0$), we have

$$M_{\mathbf{X}}(\mu_1) = \|\mathbf{X} \mathbf{v}(\mu_1)\|_2 \geq \left\| \mathbf{X} \left(\sqrt{1 - \mu_1^2} \mathbf{v}_\perp(\mu_0) + \mu_1 \mathbf{v}_0 \right) \right\|_2 \quad (94)$$

$$\geq M_{\mathbf{X}}(\mu_0) - C \|\mathbf{X}\|_2 (\mu_1 - \mu_0)^{1/2}. \quad (95)$$

Recall that $\mathbb{P}\{\|\mathbf{X}\| \geq C_1\} \leq C_2 e^{-n/C_2}$ for some constants $C_1(\alpha), C_2(\alpha)$ [AGZ09]. Define the event

$$\mathcal{E} \equiv \{\mathbf{X} : \|\mathbf{X}\| \leq C_1\}. \quad (96)$$

Hence, $\mathbb{P}(\mathcal{E}) \geq 1 - C_2 e^{-n/C_2}$. Further for $\mathbf{X} \in \mathcal{E}$ we have that for all $\mu_0, \mu_1 \in [0, 1]$

$$|M_{\mathbf{X}}(\mu_1) - M_{\mathbf{X}}(\mu_0)| \leq C' |\mu_1 - \mu_0|^{1/2}, |\overline{M}(\mu_1) - \overline{M}(\mu_0)| \quad (97)$$

$$\leq C' |\mu_1 - \mu_0|^{1/2}. \quad (98)$$

Let $\mathcal{I}_n \equiv \{0, 1/n, 2/n, \dots\} \cap [0, 1]$ be a grid. Always on \mathcal{E} , we have that Eq. (98) holds and therefore

$$\sup_{\mu \in [0, 1]} |M_{\mathbf{X}}(\mu) - \overline{M}(\mu)| \leq \max_{\mu \in \mathcal{I}_n} |M_{\mathbf{X}}(\mu) - \overline{M}(\mu)| + C'' n^{-1/2}. \quad (99)$$

Using Eq. (93) and union bound over \mathcal{I}_n , we conclude that

$$\mathbb{P}\left\{\sup_{\mu \in [0, 1]} |M_{\mathbf{X}}(\mu) - \overline{M}(\mu)| \geq t\right\} \leq \mathbb{P}\left\{\max_{\mu \in [0, 1]} |M_{\mathbf{X}}(\mu) - \overline{M}(\mu)| \geq t; \mathbf{X} \in \mathcal{E}\right\} + \mathbb{P}(\mathcal{E}^c) \quad (100)$$

$$\leq \mathbb{P}\left\{\max_{\mu \in \mathcal{I}_n} |M_{\mathbf{X}}(\mu) - \overline{M}(\mu)| \geq t - C'' n^{-1/2}; \mathbf{X} \in \mathcal{E}\right\} + \mathbb{P}(\mathcal{E}^c) \quad (101)$$

$$\leq 2n \exp\left\{-\frac{n}{2}(t - C'' n^{-1/2})^2\right\} + C_2 e^{-n/C_2} \quad (102)$$

$$\leq Cn e^{-nt^2/4}, \quad (103)$$

where the last inequality holds for all $t \leq t_0$ with t_0 a suitable constant. In particular, by Borel-Cantelli we have, almost surely and in expectation,

$$\lim_{n \rightarrow \infty} \max_{\mu \in [0, 1]} |M_{\mathbf{X}}(\mu) - \overline{M}(\mu)| = 0. \quad (104)$$

In order to upper bound $\overline{M}(\mu)$, we apply Vitale's extension of Sudakov-Fernique inequality (see e.g. [Vit00, Theorem 1] and [Cha05, Theorem 1] for a quantitative version) to the two processes $\{\mathcal{X}(\mathbf{u}, \mathbf{v})\}, \{\mathcal{Y}(\mathbf{u}, \mathbf{v})\}$ indexed by $(\mathbf{u}, \mathbf{v}) \in \mathcal{W}_\mu$ defined as follows:

$$\mathcal{X}(\mathbf{u}, \mathbf{v}) \equiv \langle \mathbf{u}, \mathbf{X}\mathbf{v} \rangle = \sqrt{\beta} \langle \mathbf{u}_0, \mathbf{u} \rangle \langle \mathbf{v}_0, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{Z}\mathbf{v} \rangle, \quad (105)$$

$$\mathcal{Y}(\mathbf{u}, \mathbf{v}) \equiv \sqrt{\beta} \langle \mathbf{u}_0, \mathbf{u} \rangle \langle \mathbf{v}_0, \mathbf{v} \rangle + \frac{1}{\sqrt{n}} (\langle \mathbf{g}, \mathbf{u} \rangle + \langle \mathbf{h}, \mathbf{v} \rangle), \quad (106)$$

for independent random vectors $\mathbf{g} \sim \mathcal{N}(0, \mathbf{I}_n), \mathbf{h} \sim \mathcal{N}(0, \mathbf{I}_p)$. It is easy to see that $\mathbb{E}\mathcal{X}(\mathbf{u}, \mathbf{v}) = \mathbb{E}\mathcal{Y}(\mathbf{u}, \mathbf{v}) = \sqrt{\beta} \langle \mathbf{u}_0, \mathbf{u} \rangle \langle \mathbf{v}_0, \mathbf{v} \rangle$ and

$$\mathbb{E}\left\{\left[\mathcal{X}(\mathbf{u}_1, \mathbf{v}_1) - \mathcal{X}(\mathbf{u}_2, \mathbf{v}_2)\right]^2\right\} = \beta (\langle \mathbf{u}_0, \mathbf{u}_1 \rangle \langle \mathbf{v}_0, \mathbf{v}_1 \rangle - \langle \mathbf{u}_0, \mathbf{u}_2 \rangle \langle \mathbf{v}_0, \mathbf{v}_2 \rangle)^2 \quad (107)$$

$$+ \frac{1}{n} \|\mathbf{u}_1 \mathbf{v}_1 - \mathbf{u}_2 \mathbf{v}_2\|_F^2 \quad (108)$$

$$= \{\mathbb{E}\mathcal{X}(\mathbf{u}_1, \mathbf{v}_1) - \mathcal{X}(\mathbf{u}_2, \mathbf{v}_2)\}^2 + \frac{2}{n} (1 - \langle \mathbf{u}_1, \mathbf{u}_2 \rangle \langle \mathbf{v}_1, \mathbf{v}_2 \rangle), \quad (109)$$

$$\mathbb{E}\left\{\left[\mathcal{Y}(\mathbf{u}_1, \mathbf{v}_1) - \mathcal{Y}(\mathbf{u}_2, \mathbf{v}_2)\right]^2\right\} = \beta (\langle \mathbf{u}_0, \mathbf{u}_1 \rangle \langle \mathbf{v}_0, \mathbf{v}_1 \rangle - \langle \mathbf{u}_0, \mathbf{u}_2 \rangle \langle \mathbf{v}_0, \mathbf{v}_2 \rangle)^2 \quad (110)$$

$$+ \frac{1}{n} \|\mathbf{u}_1 - \mathbf{u}_2\|_2^2 + \frac{1}{n} \|\mathbf{v}_1 - \mathbf{v}_2\|_2^2 \quad (111)$$

$$= \{\mathbb{E}\mathcal{Y}(\mathbf{u}_1, \mathbf{v}_1) - \mathcal{Y}(\mathbf{u}_2, \mathbf{v}_2)\}^2 + \frac{2}{n} (2 - \langle \mathbf{u}_1, \mathbf{u}_2 \rangle - \langle \mathbf{v}_1, \mathbf{v}_2 \rangle). \quad (112)$$

Hence $\mathbb{E}\{[\mathcal{X}(\mathbf{u}_1, \mathbf{v}_1) - \mathcal{X}(\mathbf{u}_2, \mathbf{v}_2)]^2\} \leq \mathbb{E}\{[\mathcal{Y}(\mathbf{u}_1, \mathbf{v}_1) - \mathcal{Y}(\mathbf{u}_2, \mathbf{v}_2)]^2\}$ (this follows from $1 - ab \leq 2 - a - b$ for $a, b \in [-1, 1]$). We conclude that

$$\overline{M}(\mu) \leq \mathbb{E} \max \left\{ \sqrt{\beta} \mu \langle \mathbf{u}_0, \mathbf{u} \rangle + \frac{1}{\sqrt{n}} (\langle \mathbf{g}, \mathbf{u} \rangle + \langle \mathbf{h}, \mathbf{v} \rangle) : (\mathbf{u}, \mathbf{v}) \in \mathcal{W}_\mu \right\} \quad (113)$$

$$\leq \mathbb{E} \max \left\{ \left\langle \frac{1}{\sqrt{n}} \mathbf{g} + \sqrt{\beta} \mu \mathbf{u}_0, \mathbf{u} \right\rangle + \left\langle \frac{1}{\sqrt{n}} \mathbf{h} + \vartheta \mathbf{v}_0, \mathbf{v} \right\rangle - \vartheta \mu : (\mathbf{u}, \mathbf{v}) \in \mathcal{W} \right\}. \quad (114)$$

where last inequality holds for any $\vartheta \in \mathbb{R}$, setting $\mathcal{W} \equiv \cup_\mu \mathcal{W}_\mu$.

The maximum in the last expression is achieved for

$$\mathbf{u} = \frac{\mathbf{g} + \mu \sqrt{\beta n} \mathbf{u}_0}{\|\mathbf{g} + \mu \sqrt{\beta n} \mathbf{u}_0\|_2}, \quad \mathbf{v} = \frac{(\mathbf{h} + \vartheta \sqrt{n} \mathbf{v}_0)_+}{\|(\mathbf{h} + \vartheta \sqrt{n} \mathbf{v}_0)_+\|_2}. \quad (115)$$

Hence, by Lemma 6.2, there exists a deterministic sequence $\delta_n = \delta_n(\alpha, \beta, \vartheta)$ independent of $\mu \in [0, 1]$, such that $\lim_{n \rightarrow \infty} \delta_n = 0$ for any $\vartheta \in \mathbb{R}$ and

$$\overline{M}(\mu) \leq \mathbb{E} \left\{ \left\| \frac{1}{\sqrt{n}} \mathbf{g} + \sqrt{\beta} \mu \mathbf{u}_0 \right\|_2 + \left\| \left(\frac{1}{\sqrt{n}} \mathbf{h} + \vartheta \mathbf{v}_0 \right)_+ \right\|_2 - \vartheta \mu \right\} \quad (116)$$

$$\leq \sqrt{1 + \beta \mu^2} + \sqrt{\alpha D_V(\vartheta/\sqrt{\alpha})} - \vartheta \mu + \delta_n, \quad (117)$$

where we recall that $D_V(x) \equiv \mathbb{E}\{(xV + G)_+^2\}$.

We next fix $\vartheta = \vartheta_*(\alpha, \beta) = S_V(\beta, \alpha)$, which is also the unique maximizer of $x \mapsto R_V^{\text{rec}}(x, \alpha)$, as shown in Lemma 6.5. Note that Eq. (117) is strictly concave in $\mu \in [0, 1]$, with unique maximum at $\mu_* = (\vartheta_*/\beta)(1 - \vartheta_*^2/\beta)^{-1/2}$. Substituting in Eq. (117), we get

$$\max_{\mu \in [0, 1]} \overline{M}(\mu) \leq \sqrt{1 - \vartheta_*^2/\beta} + \sqrt{\alpha D_V(\vartheta_*/\sqrt{\alpha})} \quad (118)$$

$$= R_V^{\text{rec}}(\vartheta_*, \alpha) + \delta_n, \quad (119)$$

where the last equality follows from the identity $\sqrt{D_V(x)} = xF_V(x) + G_V(x)$, and from the equation $\vartheta_* = \beta F_V(1 + \beta F_V)^{-1/2}$ with $F_V = F_V(\vartheta_*/\sqrt{\alpha})$ that holds by definition of $\vartheta_* = S_V(\beta, \alpha)$.

From Eq. (92), (104) and (119) we finally get

$$\limsup_{n \rightarrow \infty} \mathbb{E} \sigma^+(\mathbf{X}) \leq \limsup_{n \rightarrow \infty} \max_{\mu \in [0, 1]} \overline{M}(\mu) \quad (120)$$

$$\leq R_V^{\text{rec}}(\vartheta_*, \alpha) = \max_{\vartheta \in \mathbb{R}} R_V^{\text{rec}}(\vartheta, \alpha), \quad (121)$$

which coincides with claim (29).

Next reconsidering Eq. (117) with $\vartheta = \vartheta_*$, we see that since the right-hand side is strictly concave in $\mu \in [0, 1]$, we can strengthen Eq. (119) to

$$\overline{M}(\mu) \leq R_V^{\text{rec}}(\vartheta_*, \alpha) - c_*(\mu - \mu_*)^2 + \delta_n, \quad (122)$$

for some $c_* > 0$. We call $H(x) = c_* x^2$.

By Eq. (92) and (104) we have, almost surely,

$$\liminf_{n \rightarrow \infty} \sigma^+(\mathbf{X}) = \liminf_{n \rightarrow \infty} M_{\mathbf{X}}(\langle \mathbf{v}^+, \mathbf{v}_0 \rangle) \leq \liminf_{n \rightarrow \infty} \overline{M}(\langle \mathbf{v}^+, \mathbf{v}_0 \rangle). \quad (123)$$

We then use Eq. (122) to deduce that

$$\liminf_{n \rightarrow \infty} \sigma^+(\mathbf{X}) \leq R_V^{\text{rec}}(\vartheta_*, \alpha) - \limsup_{n \rightarrow \infty} H(|\langle \mathbf{v}^+, \mathbf{v}_0 \rangle - \mu_*|) \quad (124)$$

$$= R_V^{\text{rec}}(\vartheta_*, \alpha) - H\left(\limsup_{n \rightarrow \infty} |\langle \mathbf{v}^+, \mathbf{v}_0 \rangle - \mu_*|\right). \quad (125)$$

This implies immediately Eq. (30) with $\Delta = H^{-1}$, since (as shown above) $\vartheta_* = S_V(\beta, \alpha)$, and $\mu_* = F_V(S_V(\beta, \alpha)/\sqrt{\alpha})$.

6.3 Lower bounds: Proofs of Theorem 6 and Theorem 7

In this section we prove lower bounds on the non-negative eigenvalue (singular value) that follows from the analysis of the AMP algorithm, namely Theorem 6 for symmetric matrices and Theorem 7 for rectangular matrices. The proofs are very similar in the two cases, hence we will provide details only in the case of rectangular matrices, and limit ourselves to pointing out differences arising in the symmetric setting.

6.4 Proof of Theorem 7

Define

$$r_t(n) \equiv \langle \hat{\mathbf{u}}^t, \mathbf{X} \hat{\mathbf{v}}^t \rangle = \frac{1}{n} \langle g(\mathbf{u}^t), \mathbf{X} f(\mathbf{v}^t) \rangle, \quad (126)$$

and observe, using AMP-rec,

$$r_t(n) = \frac{1}{n} \langle g(\mathbf{u}^t), \mathbf{u}^t + \mathbf{b}_t g(\mathbf{u}^{t-1}) \rangle \quad (127)$$

$$= \frac{1}{n} \langle g(\mathbf{u}^t), \mathbf{u}^t \rangle + \frac{\mathbf{b}_t}{n} \langle g(\mathbf{u}^t), g(\mathbf{u}^t) \rangle + \frac{\mathbf{b}_t}{n} \langle g(\mathbf{u}^t), (g(\mathbf{u}^{t-1}) - g(\mathbf{u}^t)) \rangle \quad (128)$$

$$= \frac{1}{n} \langle g(\mathbf{u}^t), \mathbf{u}^t \rangle + \mathbf{b}_t + E_t(n) \quad (129)$$

where

$$|E_t(n)| = \frac{\mathbf{b}_t}{n} \left| \langle g(\mathbf{u}^t), (g(\mathbf{u}^{t-1}) - g(\mathbf{u}^t)) \rangle \right| \quad (130)$$

$$\leq \frac{\mathbf{b}_t}{\sqrt{n}} \|g(\mathbf{u}^{t-1}) - g(\mathbf{u}^t)\|_2 \quad (131)$$

$$\leq 4\mathbf{b}_t \frac{\|\mathbf{u}^{t-1} - \mathbf{u}^t\|_2}{\|\mathbf{u}^{t-1}\|_2 + \|\mathbf{u}^t\|_2}. \quad (132)$$

The last step follows from triangular inequality.

By Proposition 4.3 applied to $\psi(x, y) = x^2$, we have, almost surely,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \|\mathbf{u}^t\|_2^2 = \mathbb{E}\{(\mu_t U + G)^2\} = 1 + \mu_t^2, \quad (133)$$

and therefore

$$\lim_{n \rightarrow \infty} \frac{1}{n} \langle g(\mathbf{u}^t), \mathbf{u}^t \rangle = \lim_{n \rightarrow \infty} \frac{1}{\sqrt{n}} \|\mathbf{u}^t\|_2 = \sqrt{1 + \mu_t^2} \quad (134)$$

$$= \sqrt{1 + \beta F_V(\vartheta_{t-1}/\sqrt{\alpha})^2}. \quad (135)$$

By applying the same proposition to $\psi(x, y) = x_+^2$ we have

$$\lim_{n \rightarrow \infty} \frac{1}{p} \|(\mathbf{v}^t)_+\|_2^2 = \mathbb{E}\{(\vartheta_t/\sqrt{\alpha} V + G)_+^2\}. \quad (136)$$

Further, letting $\psi(x, y) = \mathbb{I}(x > 0) = 1 - \mathbb{I}(x \leq 0)$ we get

$$\lim_{n \rightarrow \infty} \frac{1}{p} \|(\mathbf{v}^t)_+\|_0 = \lim_{n \rightarrow \infty} \frac{1}{p} \sum_{i=1}^p \mathbb{I}(\mathbf{v}_i^t > 0) = \mathbb{P}(\vartheta_t/\sqrt{\alpha} V + G > 0) \quad (137)$$

$$= \mathbb{E}\{G(\vartheta_t/\sqrt{\alpha} V + G)_+\}, \quad (138)$$

where the last equality follows from Stein's lemma [Ste72] (see Lemma E.1, whereby we consider the weakly differentiable function $f(G) = (c + G)_+$ for $c = \vartheta_t/\sqrt{\alpha} V$, and subsequently take expectation over V .)

Using together Eq. (136) and Eq. (138), we get

$$\lim_{n \rightarrow \infty} \mathbf{b}_t(n) = \lim_{n \rightarrow \infty} \sqrt{\alpha} \frac{\|(\mathbf{v}^t)_+\|_0/p}{\|(\mathbf{v}^t)_+\|_2/p} = \sqrt{\alpha} \mathbf{G}_V \left(\frac{\vartheta_t}{\sqrt{\alpha}} \right). \quad (139)$$

Using Eq. (133) and Eq. (139) in the upper bound (132), we get

$$\lim_{t \rightarrow \infty} \lim_{n \rightarrow \infty} |E_t(n)| = 0. \quad (140)$$

Finally, substituting this result together with Eq. (135) and (139) in Eq. (129), we obtain

$$\lim_{t \rightarrow \infty} \lim_{n \rightarrow \infty} r_t(n) = \lim_{t \rightarrow \infty} \mathbf{R}_V^{\text{rec}}(\vartheta_t, \alpha). \quad (141)$$

The claim (45) follows by Lemma 6.7.

Consider next Eq. (46). We have

$$\lim_{n \rightarrow \infty} \langle \hat{\mathbf{u}}^t, \mathbf{u}_0 \rangle = \lim_{n \rightarrow \infty} \frac{\langle \mathbf{u}^t, \mathbf{u}_0 \rangle}{\|\mathbf{u}^t\|_2} \quad (142)$$

$$= \frac{\mathbb{E}\{U(\mu_t U + G)\}}{\sqrt{\mathbb{E}\{(\mu_t U + G)^2\}}} = \frac{\mu_t}{\sqrt{1 + \mu_t^2}}, \quad (143)$$

where the second equality follows by applying Proposition 4.3 to $\psi(x, y) = xy$ (for the numerator) and using Eq. (133) (for the denominator). Finally, the claim (46) follows by taking $t \rightarrow \infty$, and using Lemma 6.7.

The proof of claim (47) follows by the same argument and we omit it.

6.4.1 Proof of Theorem 6

The proof in the symmetric case is very similar to the one for rectangular matrices, see Theorem 7. We limit ourselves to sketching the first steps. We have, using AMP-sym,

$$\rho_t(n) \equiv \langle \hat{\mathbf{v}}^t, \mathbf{X} \hat{\mathbf{v}}^t \rangle \quad (144)$$

$$= \frac{1}{n} \langle f(\mathbf{v}^t), \mathbf{X} f(\mathbf{v}^t) \rangle \quad (145)$$

$$= \frac{1}{n} \langle f(\mathbf{v}^t), \mathbf{v}^{t+1} + \mathbf{b}_t f(\mathbf{v}^{t-1}) \rangle \quad (146)$$

$$= \frac{1}{n} \langle f(\mathbf{v}^t), \mathbf{v}^t \rangle + \frac{\mathbf{b}_t}{n} \langle f(\mathbf{v}^t), f(\mathbf{v}^t) \rangle + \frac{1}{n} \langle f(\mathbf{v}^t), \mathbf{v}^{t+1} - \mathbf{v}^t \rangle + \frac{\mathbf{b}_t}{n} \langle f(\mathbf{v}^t), (f(\mathbf{v}^{t-1}) - f(\mathbf{v}^t)) \rangle \quad (147)$$

$$= \frac{1}{n} \langle f(\mathbf{v}^t), \mathbf{v}^t \rangle + \mathbf{b}_t + \tilde{E}_t(n) \quad (148)$$

and we are left with a term $\tilde{E}_t = \frac{1}{n} \langle f(\mathbf{v}^t), \mathbf{v}^{t+1} - \mathbf{v}^t \rangle + \frac{\mathbf{b}_t}{n} \langle f(\mathbf{v}^t), (f(\mathbf{v}^{t-1}) - f(\mathbf{v}^t)) \rangle$ which we treat similarly to $E_t(n)$ of Theorem 7. Namely, by using Proposition 4.3 and Proposition 4.4, we prove that

$$\lim_{t \rightarrow \infty} \lim_{n \rightarrow \infty} \tilde{E}_t(n) = 0 \quad (149)$$

In addition, it follows from Proposition 4.3 that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \langle f(\mathbf{v}^t), \mathbf{v}^t \rangle = \tau_t F_V(\tau_t) + G_V(\tau_t) = \beta F_V(\tau_{t-1}) F_V(\tau_t) + G_V(\tau_t) \quad (150)$$

$$\lim_{n \rightarrow \infty} \mathbf{b}_t = G_V(\tau_t) \quad (151)$$

This terminates the proof sketch.

6.5 Minimax analysis: proof of Theorems 3 and 5

In this section we prove that the least favorable vectors \mathbf{v}_0 are –asymptotically– of the following form: $v_{0,i} = 1/\sqrt{|S|}$ for all $i \in S$, and $v_{0,i} = 0$ otherwise, for some support $S \subseteq [p]$. Further, we characterize the least favorable size of the support $|S|$.

The proofs proceed by analyzing the expression in Theorem 2 and applying strong duality to a certain linear program over probability distributions, that is related to the function $\mu_V \mapsto F_V(x)$. We start with some preliminary facts and definitions in Section 6.5.1. The key step is to reduce ourselves to two points mixtures: this is achieved in Section 6.5.2. Finally, in Sections 6.5.3 and 6.5.4, we use these results to prove Theorems 3 and 5. Since the proof of Theorem 5 is completely analogous to the one of Theorem 3, we will limit ourselves to mentioning the main differences.

6.5.1 Preliminary definitions

For $\varepsilon \in (0, 1]$ and $v \in \mathbb{R}_{\geq 0}$, we define the 2-points mixture

$$\mu_{\varepsilon, v} \equiv (1 - \varepsilon) \delta_0 + \varepsilon \delta_v \quad (152)$$

In particular, when $v = 1/\sqrt{\varepsilon}$ (and hence the above distribution has second moment equal to 1), we write $\mu_\varepsilon = \mu_{\varepsilon, 1/\sqrt{\varepsilon}}$. We also write –with a slight abuse of notation– $F_\varepsilon(x)$ instead of $F_V(x)$ when $V \sim \mu_\varepsilon$. Explicitly

$$F_\varepsilon(x) = \frac{\mathbb{E}\{(x + \sqrt{\varepsilon}G)_+\}}{\sqrt{(1 - \varepsilon)/2 + \mathbb{E}\{(x + \sqrt{\varepsilon}G)_+^2\}}}. \quad (153)$$

Even more explicitly

$$F_\varepsilon(x) = \frac{\varepsilon B(x/\sqrt{\varepsilon})/x}{\sqrt{(1 - \varepsilon)/2 + \varepsilon(B(x/\sqrt{\varepsilon}) + \Phi(x/\sqrt{\varepsilon}))}}, \quad (154)$$

$$B(w) \equiv w^2\Phi(w) + w\phi(w). \quad (155)$$

We will also adopt the shorthand $T_\varepsilon(\beta) = T_V(\beta)$ when $V \sim \mu_\varepsilon$.

We will next establish two calculus lemmas that are useful for the following.

Lemma 6.9. *For any given $a, b \in \mathbb{R}$, the equation*

$$\frac{\phi(v)}{v} + b\Phi(v) = a, \quad (156)$$

in the unknown $v \in \mathbb{R}_{>0}$ has at most two solutions v_1, v_2 .

Proof. Let $h_b(v) = (\phi(v)/v) + b\Phi(v)$ denote the left-hand side of Eq. (156). Then

$$h'_b(v) = -\left(1 - b + \frac{1}{v^2}\right)\phi(v). \quad (157)$$

If $b \leq 1$, then we conclude that $h'_b(v) < 0$ for all $v > 0$ and hence the equation $h_b(v) = a$ has at most one positive solution. If –on the other hand– $b > 1$, then $h'_b(v) < 0$ for $v < v_* \equiv (b - 1)^{-1/2}$ and $h'_b(v) > 0$ for $v > v_*$. It follows that the equation $h_b(v) = a$ has at most one solution in $(0, v_*]$ and at most one in (v_*, ∞) . \square

Lemma 6.10. *Let $b : \mathbb{R} \rightarrow \mathbb{R}$ be defined as $b(x) = x^2(\Phi(x) - 1) + x\phi(x)$. Then, for every $x \in (0, \infty)$, we have*

$$\phi(x)b(x) > \left(\Phi(x) - \frac{1}{2}\right)b'(x). \quad (158)$$

Proof. By simple calculus, we get $b(0) = 0$, and the derivatives

$$b'(x) = \phi(x) - 2x(1 - \Phi(x)), \quad b'(0) = \phi(0), \quad (159)$$

$$b''(x) = x\phi(x) - 2(1 - \Phi(x)), \quad b''(0) = -1, \quad (160)$$

$$b'''(x) = (3 - x^2)\phi(x), \quad b'''(0) = 3\phi(0). \quad (161)$$

Let us further recall the inequalities (valid for $x > 0$)

$$\frac{\phi(x)}{x} \left(1 - \frac{1}{x^2}\right) < 1 - \Phi(x) < \frac{\phi(x)}{x}, \quad (162)$$

which immediately imply for all $x > 0$

$$0 < b(x) < \frac{\phi(x)}{x}. \quad (163)$$

Therefore the left-hand side of Eq. (158) is always strictly positive. Consider the right-hand side. By consulting special values of the normal distribution, we see that $b'(1) = \phi(1) - 2(1 - \Phi(1)) < -0.07 < 0$. By a change of variables we know that $x(1 - \Phi(x))/\phi(x) = \int_0^\infty \exp(-z - z^2/(2x^2))dz$ or, equivalently

$$b'(x) = \phi(x) \left\{ 1 - 2\mathbb{E}[e^{-Z^2/(2x^2)}] \right\}. \quad (164)$$

Since the term in curly brackets is decreasing in x , and is negative at $x = 1$, we have $b'(x) < 0$ for all $x \geq 1$. Therefore the right-hand side of Eq. (158) is non-positive for $x \geq 1$. This proves the claim for $x \geq 1$, and we will assume hereafter $x \in (0, 1)$.

Next notice that $0 \leq b'''(x) \leq 3\phi(0)$ for $x \in (0, 1)$. Therefore, by Taylor expansion and intermediate value theorem, we get, for $x \in (0, 1)$,

$$b'(x) \leq \phi(0) - x + \frac{3\phi(0)}{2} x^2. \quad (165)$$

The right-hand side is negative for $x \in (x_0, x_1)$ where

$$x_{1/0} = \frac{1 \pm \sqrt{1 - 6\phi(0)^2}}{3\phi(0)}. \quad (166)$$

In particular $x_0 \leq 2/3$, and $x_1 > 1$. It follows that the right-hand side of Eq. (158) is non-positive for $x \geq x_0$.

We will therefore restrict ourselves to considering $x \in (0, x_0) \subseteq (0, 2/3)$. Note that our claim can be equivalently written as

$$b(x) \geq \left(\Phi(x) - \frac{1}{2} \right) \left(1 - 2x \frac{1 - \Phi(x)}{\phi(x)} \right). \quad (167)$$

We will next develop, for $x \in (0, x_0)$, a lower bound on the left-hand side, to be denoted by $l(x)$, and an upper bound on the right-hand side, to be denoted by $u(x)$ and prove that $l(x) \geq u(x)$. For the left hand side note that $b'''(x) \geq 0$ for $x \in (0, x_0)$ and hence, again by Taylor expansion

$$b(x) \geq \phi(0)x - \frac{1}{2}x^2 \equiv l(x). \quad (168)$$

For the right hand side note that $\Phi(x) - (1/2) \leq \phi(0)x$. Further $x \mapsto (1 - \Phi(x))/\phi(x)$ is monotone decreasing. We therefore define

$$d_0 \equiv \frac{2(1 - \Phi(x_0))}{\phi(x_0)}, \quad (169)$$

and obtain the upper bound $u(x) = \phi(0)x(1 - d_0x)$. Hence

$$l(x) - u(x) = \left(d_0\phi(0) - \frac{1}{2} \right) x^2, \quad (170)$$

It is a straightforward exercise to check that indeed $d_0\phi(0) > (1/2)$ thus completing the proof. \square

6.5.2 Reduction to two points mixtures

The main theorem of this Section shows that $F_V(x)$ is minimized by probability measures μ_V that are mixture of at most two point masses.

Theorem 8. *Fix $x \geq 0$. Then for any random variable V with probability distribution $\mu_V \in \mathcal{P}_{\bar{\varepsilon}}$, we have*

$$F_V(x) \geq \min_{\varepsilon \in (0, \bar{\varepsilon}]} F_{\varepsilon}(x). \quad (171)$$

The proof of this theorem is presented at the end of the section. Before getting to it, we'll introduce a related problem. Note that

$$F_V(x) \geq \inf_{y \in \mathbb{R}_{>0}} \frac{\mathcal{F}(x, y)}{y}, \quad (172)$$

where $\mathcal{F}(x, y)$ is the value of a constrained optimization problem:

$$\begin{aligned} & \text{minimize} && \mathbb{E}\{V(xV + G)_+\}, \\ & \text{subject to} && \mu_V \in \mathcal{P}_{\bar{\varepsilon}}, \\ & && \mathbb{E}\{(xV + G)_+^2\} = y^2. \end{aligned} \quad (173)$$

Here it is understood that $\mathcal{F}(x, y) = \infty$ if this problem is unfeasible.

Lemma 6.11. *Let $x, y \in \mathbb{R}_{>0}$ be such that the problem (173) is feasible. Then there exist $\varepsilon \leq \bar{\varepsilon}$ such that μ_{ε} is feasible and $q \in [0, 1]$, such that, letting $v_*^2 = (1 - q)/\varepsilon$, we have*

$$\mathcal{F}(x, y) = qx + \int \mathbb{E}\{v(xv + G)_+\} \mu_{\varepsilon, v_*}(\mathrm{d}v), \quad (174)$$

$$y^2 = qx^2 + \int \mathbb{E}\{(xv + G)_+^2\} \mu_{\varepsilon, v_*}(\mathrm{d}v). \quad (175)$$

Proof. By a rescaling of the objective function, and letting $W = xV$, we can rewrite the problem (173) as

$$\begin{aligned} & \text{minimize} && \mathbb{E}\{W(W + G)_+\}, \\ & \text{subject to} && \mu_W(\{0\}) \geq 1 - \bar{\varepsilon}, \\ & && \mathbb{E}\{W^2\} = x^2, \\ & && \mathbb{E}\{(W + G)_+^2\} = y^2. \end{aligned} \quad (176)$$

Now, for fixed $w \in \mathbb{R}$, let

$$f(w) \equiv \mathbb{E}\{w(w + G)_+\} = w^2\Phi(w) + w\phi(w), \quad (177)$$

$$g(w) \equiv \mathbb{E}\{(w + G)_+^2\} = (1 + w^2)\Phi(w) + w\phi(w), \quad (178)$$

and write $\mu_W = (1 - \bar{\varepsilon})\delta_0 + (1/g(w))\mu$ with μ a measure on $\mathbb{R}_{\geq 0}$. Then we can rewrite the optimization problem as the following (with decision variable μ)

$$\begin{aligned} & \text{minimize} && \int \frac{f(w)}{g(w)} \mu(dw), \\ & \text{subject to} && \int \frac{1}{g(w)} \mu(dw) = \bar{\varepsilon}, \\ & && \int \frac{w^2}{g(w)} \mu(dw) = x^2, \\ & && \int \mu(dw) = y^2. \end{aligned} \tag{179}$$

The corresponding value is $x\mathcal{F}(x, y)$. Note that each of the functions $(f(w)/g(w))$, $1/g(w)$, $w^2/g(w)$ is bounded and Lipschitz continuous, with a finite limit as $w \rightarrow \infty$. This implies that the value $x\mathcal{F}(x, y)$ is achieved by a measure μ_* on the completed real line $[0, \infty]$, with total mass y^2 . Indeed the family of normalized distributions on $[0, \infty]$ is tight and both the objective and the constraints are continuous in the weak topology. Hereafter, we shall assume this holds. Functions on $[0, \infty)$ are extended by continuity to $+\infty$.

By introducing Lagrange multipliers, we obviously have, for any $\alpha, \beta, \gamma \in \mathbb{R}$

$$x\mathcal{F}(x, y) \geq \bar{\varepsilon}\alpha + \beta x^2 + \gamma y^2 + \inf_{\mu} \int \left\{ \frac{f(w) - \alpha - \beta w^2 - \gamma g(w)}{g(w)} \right\} \mu(dw), \tag{180}$$

where the infimum is over measures μ on $\mathbb{R}_{\geq 0}$. By strong duality (which follows, for instance, from the Kneser-Kuhn minimax theorem [Kne52], see also [Joh11, Theorem A.1]), there exists⁴ $\alpha_*, \beta_*, \gamma_* \in \mathbb{R}$ such that the above holds with equality. Note that for such choice $f(w) - \alpha_* - \beta_* w^2 - \gamma_* g(w) \geq 0$ for all $w \in [0, \infty]$, because otherwise the infimum is $-\infty$. Under this condition, the infimum term in Eq. (180) is zero, and hence we must have

$$\alpha_* = \inf_{w \in \mathbb{R}_{\geq 0}} \left[f(w) - \beta_* w^2 - \gamma_* g(w) \right], \tag{181}$$

because otherwise we could increase the lower bound Eq. (180) by increasing α . Further, since the right-hand side is an analytic function of w , the infimum in Eq. (181) is achieved on a finite set $S_* \in [0, \infty]$, and the minimizer μ_* of problem (179) has support $\text{supp}(\mu_*) \subseteq S_*$ because otherwise the infimum in the lower bound (180) would not be achieved.

Next we claim that $S_* \subseteq \{0, w_*, \infty\}$ for some finite $a \in \mathbb{R}_{>0}$. Indeed, let $h(w) \equiv f(w) - \beta_* w^2 - \gamma_* g(w)$. It follows from Eqs. (177) and (178) that

$$h'(w) = (1 - 2\gamma_*)\phi(w) + 2(1 - \gamma_*)w\Phi(w) - 2\beta_* w. \tag{182}$$

Assume $\gamma_* \neq 1/2$. We then have $h'(w) = 0$ for some finite $w \in \mathbb{R}_{>0}$ if and only if

$$\frac{\phi(w)}{w} + \left(\frac{2 - 2\gamma_*}{1 - 2\gamma_*} \right) \Phi(w) = \left(\frac{2\beta_*}{1 - 2\gamma_*} \right) w. \tag{183}$$

⁴In general, the mentioned theorem only imply that equality is achieved asymptotically, along a sequence $\alpha_n, \beta_n, \gamma_n$. In the present case, it is not had to show that, letting $P(\alpha, \beta, \gamma; \mu)$ denote the right hand side of Eq. (180), the $\sup_{\alpha, \beta, \gamma} [\inf_{\mu} P]$ is actually achieved at finite $\alpha_*, \beta_*, \gamma_*$, by showing that the sequence must remain bounded and using standard compactness arguments.

By Lemma 6.9 this has at most two solutions w_1, w_2 . If on the other hand $\gamma_* = 1/2$, then the above equation reduces to $\Phi(w) = 2\beta_*$ which has at most one solution. In both cases, at most one solution –call it w_* – is a local minimum of h .

We conclude that $S_* \subseteq \{0, w_*, \infty\}$, and therefore the value of the problem (179) is achieved by a measure of the form

$$\mu_* = p_0 \delta_0 + p_1 \delta_{w_*} + p_2 \delta_\infty \quad (184)$$

The three constraints imply the following relations

$$2p_0 + p_1 \frac{1}{g(w_*)} = \bar{\varepsilon}, \quad (185)$$

$$p_1 \frac{w_*^2}{g(w_*)} + p_2 = x^2, \quad (186)$$

$$p_0 + p_1 + p_2 = y^2, \quad (187)$$

and the value is

$$x\mathcal{F}(x, y) = p_1 \frac{f(w_*)}{g(w_*)} + p_2. \quad (188)$$

The proof is completed by the change of variables $p_1 = g(w_*) \varepsilon$, $p_0 = (\bar{\varepsilon} - \varepsilon)/2$, $p_2 = qx^2$, $w_* = v_*x$, $p_2 = qx^2$. With these substitutions Eq. (187) yields (175), and Eq. (188) yields (174). \square

We are now in position to prove Theorem 8, that is the main result in this section.

Proof of Theorem 8. By Lemma 6.11 and Eq. (172), we have, for any $\mu_V \in \mathcal{F}_{\bar{\varepsilon}}$,

$$\mathbf{F}_V(x) \geq \inf_{q, v, \varepsilon} \frac{qx + \int \mathbb{E}\{v(xv + G)_+\} \mu_{\varepsilon, v_*}(\mathrm{d}v)}{\sqrt{qx^2 + \int \mathbb{E}\{(xv + G)_+^2\} \mu_{\varepsilon, v_*}(\mathrm{d}v)}}. \quad (189)$$

where the infimum is over $q \in [0, 1]$, $\varepsilon \in (0, \bar{\varepsilon}]$, $v_* = \sqrt{(1-q)/\varepsilon}$. Our claim is equivalent to saying that the infimum on the right hand side is achieved when $q = 0$.

Since $x > 0$ is given, we can regard the right-hand side as a function of $w = v_*x$ and ε , and substitute $qx^2 = x^2 - \varepsilon w^2$. We then define the function

$$G(w, \varepsilon) = \frac{x^2 - \varepsilon w^2 + \varepsilon \mathbb{E}\{w(w + G)_+\}}{\sqrt{x^2 - \varepsilon w^2 + (1 - \varepsilon)/2 + \varepsilon \mathbb{E}\{(w + G)_+^2\}}}. \quad (190)$$

More explicitly

$$G(w, \varepsilon) = \frac{x^2 + \varepsilon b(w)}{\sqrt{x^2 + (1/2) + \varepsilon(b(w) + \Phi(w) - (1/2))}}, \quad (191)$$

$$b(w) \equiv w^2(\Phi(w) - 1) + w\phi(w), \quad (192)$$

which needs to be optimized over $\varepsilon \in (0, \bar{\varepsilon}]$, and $w \in [0, x/\sqrt{\varepsilon}]$. Our claim is equivalent to saying that the minimum cannot be in the interior of this domain.

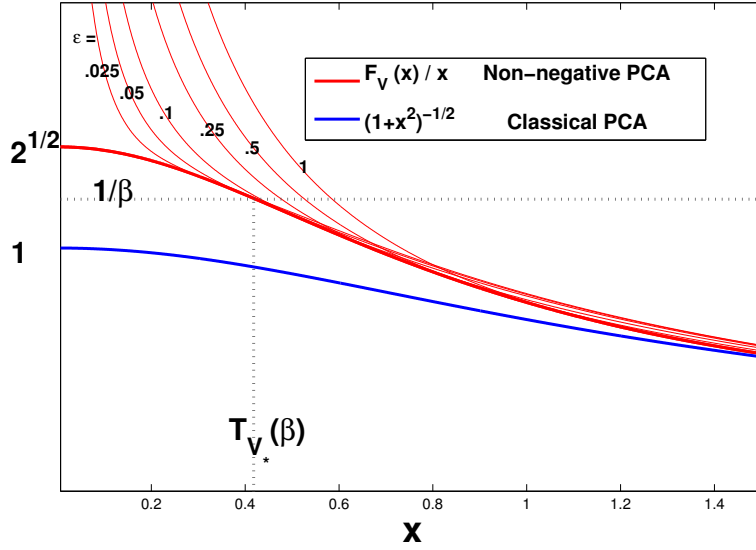


Figure 4: The function $F_V(x)/x$ where V is a mixture of two Dirac δ s at 0 and at $1/\sqrt{\varepsilon}$ with various values of ε , and the worst case curve $F_*(x)/x$. The analogue curve in the case of classical PCA is drawn in blue and the construction of $T_*(\beta)$ for $\beta \in (1/\sqrt{2}, 1)$ is illustrated with dashed lines.

Since G is analytic in the mentioned domain, a minimum in the interior must satisfy $\partial_w G(w, \varepsilon) = \partial_\varepsilon G(w, \varepsilon) = 0$. Simple calculus shows that these two conditions are equivalent –respectively– to:

$$2\varepsilon b'(w) \left[x^2 + \frac{1}{2} + \varepsilon \left(b(w) + \Phi(w) - \frac{1}{2} \right) \right] = [x^2 + \varepsilon b(w)] \varepsilon [b'(w) + \phi(w)], \quad (193)$$

$$2b(w) \left[x^2 + \frac{1}{2} + \varepsilon \left(b(w) + \Phi(w) - \frac{1}{2} \right) \right] = [x^2 + \varepsilon b(w)] \left[b(w) + \Phi(w) - \frac{1}{2} \right]. \quad (194)$$

Taking the ratio of these equations, we obtain the necessary condition

$$\frac{b'(w)}{b(w)} = \frac{b'(w) + \phi(w)}{b(w) + \Phi(w) - (1/2)}, \quad (195)$$

or equivalently

$$\left(\Phi(x) - \frac{1}{2} \right) b'(x) = \phi(x) b(x). \quad (196)$$

Lemma 6.10 establishes that this equation does not have any solution in $\mathbb{R}_{>0}$ and hence $G(w, \varepsilon)$ does not have stationary points in domain $\varepsilon \in (0, \bar{\varepsilon}]$, $w \in [0, x/\sqrt{\varepsilon}]$. This finishes our proof. \square

We conclude with a Corollary of Theorem 8. (Figure 4 provides an illustration of the argument used in the proof.)

Corollary 6.12. *Fix $\beta \in (0, \infty)$. Then for any random variable V with probability distribution $\mu_V \in \mathcal{P}_{\bar{\varepsilon}}$, we have*

$$T_V(\beta) \geq \inf_{\varepsilon \in (0, \bar{\varepsilon}]} T_\varepsilon(\beta). \quad (197)$$

Further, for any $\beta > 1/\sqrt{2}$, the infimum on the right-hand side is achieved at some $\varepsilon_* \in (0, \bar{\varepsilon}]$.

Proof. Assume the claim (197) does not hold. Then there exists $\mu_V \in \mathcal{P}_{\bar{\varepsilon}}$ such that $\mathsf{T}_V(\beta) < \mathsf{T}_{\varepsilon}(\beta)$ for all $\varepsilon \in (0, \bar{\varepsilon}]$. Now, on the one hand, by definition we have

$$\frac{1}{\beta} = \frac{\mathsf{F}_V(\mathsf{T}_V(\beta))}{\mathsf{T}_V(\beta)} = \frac{\mathsf{F}_{\varepsilon}(\mathsf{T}_{\varepsilon}(\beta))}{\mathsf{T}_{\varepsilon}(\beta)}. \quad (198)$$

On the other hand, by Theorem 8, there exists $\varepsilon_0 \in (0, \bar{\varepsilon}]$ such that $\mathsf{F}_V(x) \geq \mathsf{F}_{\varepsilon_0}(x)$ for $x = \mathsf{T}_V(\beta) \in \mathbb{R}_{>0}$. Using this fact, the contradiction assumption $\mathsf{T}_V(\beta) < \mathsf{T}_{\varepsilon_0}(\beta)$, and the fact that $x \mapsto \mathsf{F}_{\varepsilon_0}(x)/x$ is strictly decreasing on $\mathbb{R}_{>0}$ as shown in the proof of Lemma 6.3, we get

$$\frac{1}{\beta} = \frac{\mathsf{F}_V(\mathsf{T}_V(\beta))}{\mathsf{T}_V(\beta)} \geq \frac{\mathsf{F}_{\varepsilon_0}(\mathsf{T}_V(\beta))}{\mathsf{T}_V(\beta)} > \frac{\mathsf{F}_{\varepsilon_0}(\mathsf{T}_{\varepsilon_0}(\beta))}{\mathsf{T}_{\varepsilon_0}(\beta)} = \frac{1}{\beta}. \quad (199)$$

We therefore reached a contradiction, which proves the claim (197).

In order to prove that the infimum is achieved at some $\varepsilon_* \in (0, \bar{\varepsilon}]$, note that $\varepsilon \mapsto \mathsf{T}_{\varepsilon}(\beta)$ is clearly continuous and, by Lemma 6.8,

$$\lim_{\varepsilon \rightarrow 0} \mathsf{T}_{\varepsilon}(\beta) = \sqrt{\beta^2 - (1/2)}. \quad (200)$$

It is therefore sufficient to show that $\varepsilon \mapsto \mathsf{T}_{\varepsilon}(\beta)$ is decreasing for ε small enough. By an argument similar to the above, this follows if we show that $\varepsilon \mapsto \mathsf{F}_{\varepsilon}(x)$ is decreasing for $x = \mathsf{T}_0(\beta) = \sqrt{\beta^2 - (1/2)}$ and ε small enough. Indeed using the definition (154) and recalling that $\Phi(w) = 1 - O(\phi(w))$ as $w \rightarrow \infty$, we get, for every fixed $x > 0$

$$\mathsf{F}_{\varepsilon}(x) = \frac{x}{\sqrt{\frac{1+\varepsilon}{2} + x^2}} + O(\phi(x/\sqrt{\varepsilon})), \quad (201)$$

which is of course decreasing in ε for $\varepsilon \in [0, c(x)]$ with $c(x) > 0$. □

6.5.3 Proof of Theorems 3

First let $\beta \in [0, 1/\sqrt{2}]$. We then set $\ell = \lfloor n\varepsilon \rfloor$ and

$$v_{0,i} = \begin{cases} 1/\sqrt{\ell} & \text{for } i \in \{1, 2, \dots, \ell\}, \\ 0 & \text{for } i \in \{\ell + 1, \dots, n\}. \end{cases} \quad (202)$$

Then of course $\{\mathbf{v}_0(n)\}_{n \geq 0}$ converges in empirical distribution to μ_{ε} and, by Theorem 2

$$\lim_{n \rightarrow \infty} \langle \mathbf{v}^+, \mathbf{v}_0 \rangle = \mathsf{F}_{\varepsilon}(\mathsf{T}_{\varepsilon}(\beta)), \quad (203)$$

with $\mathsf{T}_{\varepsilon}(\beta)$ the only non-negative solution of $x = \beta \mathsf{F}_{\varepsilon}(x)$. By Lemma 6.8 (cf. Eqs. (79), (81)), we have $\lim_{\varepsilon \rightarrow 0} \mathsf{F}_{\varepsilon}(\mathsf{T}_{\varepsilon}(\beta)) = 0$, and hence

$$\lim_{\varepsilon \rightarrow 0} \lim_{n \rightarrow \infty} \langle \mathbf{v}^+, \mathbf{v}_0 \rangle = 0. \quad (204)$$

The claim (17) then follows by replacing ε , by sequence $\{\varepsilon_n\}_{n \geq 1}$ with $\varepsilon_n \downarrow 0$ sufficiently slowly. The limit vanishes in this case as well by a standard argument.

Next consider the claim (19). We let ε_* be the value achieving the infimum in Eq. (197), which exists by Corollary 6.12. It is obvious (by another application of Theorem 2) that equality holds for the stated choice of $\mathbf{v}_0(n)$. Assume by contradiction that the inequality (19) does not hold for some sequence $\{\mathbf{v}_0(n)\}$. Then, by tightness, there exists a subsequence along which the limit on the left hand side exists, and that converges in empirical distribution to a certain probability measure $\mu_V \in \mathcal{P}_\varepsilon$. Hence, using Theorem 2, it follows that (using the definition of $T_V(\beta)$)

$$T_V(\beta) < T_{\varepsilon_*}(\beta) = \inf_{\varepsilon \in (0, \bar{\varepsilon}]} T_\varepsilon(\beta). \quad (205)$$

This contradicts corollary 6.12, hence proving our claim.

6.5.4 Proof of Theorem 5

The proof of Theorem 5 is very similar to the proof of Theorem 3, and therefore we will only sketch the first steps.

First if $\beta \in [0, \sqrt{\alpha/2})$, we set $p = \lfloor n\varepsilon \rfloor$

$$v_{0,i} = \begin{cases} 1/\sqrt{\ell} & \text{for } i \in \{1, 2, \dots, \ell\}, \\ 0 & \text{for } i \in \{\ell + 1, \dots, p\}. \end{cases} \quad (206)$$

Then $\{\mathbf{v}_0(p)\}_{p \geq 0}$ converges in empirical distribution to μ_ε and, by Theorem 4,

$$\lim_{p \rightarrow \infty} \langle \mathbf{v}^+, \mathbf{v}_0 \rangle = F_\varepsilon(S_\varepsilon(\beta, \alpha)/\sqrt{\alpha}). \quad (207)$$

with $S_\varepsilon(\beta, \alpha) \equiv S_V(\beta, \alpha)$ for $V \sim \mu_\varepsilon$ is given by Definition 2.2, i.e. is the only positive solution x of

$$x = \frac{\beta F_\varepsilon(x/\sqrt{\alpha})}{\sqrt{1 + \beta F_\varepsilon(x/\sqrt{\alpha})^2}}. \quad (208)$$

By Lemma 6.8 (cf. Eqs. (79), (82)), we have $\lim_{\varepsilon \rightarrow 0} F_\varepsilon(S_\varepsilon(\beta, \alpha)/\sqrt{\alpha}) = 0$, and hence

$$\lim_{\varepsilon \rightarrow 0} \lim_{p \rightarrow \infty} \langle \mathbf{v}^+, \mathbf{v}_0 \rangle = 0. \quad (209)$$

The claim follows by taking $\varepsilon = \varepsilon(p) \rightarrow 0$ slowly enough.

Next consider $\beta > \sqrt{\alpha/2}$. By the same argument as in Corollary 6.12, we have, for any $\mu_V \in \mathcal{P}_\varepsilon$,

$$S_V(\beta, \alpha) \geq \inf_{\varepsilon \in (0, \bar{\varepsilon}]} S_\varepsilon(\beta, \alpha). \quad (210)$$

Further, for any $\beta > 1/\sqrt{2}$, the infimum on the right-hand side is achieved at some $\varepsilon_* \in (0, \bar{\varepsilon}]$. We then take $V_* \sim \mu_{\varepsilon_*}$.

Assuming that the claim (25) is false, we can construct by the same tightness argument used in the previous section, a probability distribution μ_V , such that $S_V(\beta, \alpha) < S_{\varepsilon_*}(\beta, \alpha)$. This contradicts Eq. (210), which proves our claim.

Acknowledgements

This work was partially supported by the NSF grant CCF-1319979 and the grants AFOSR/DARPA FA9550-12-1-0411 and FA9550-13-1-0036.

A Derivation of AMP

In this appendix we sketch a derivation of the approximate message passing algorithm (AMP) algorithm analyzed in the main text. Let us emphasize two important points:

1. The derivation developed in this appendix is not really necessary: none of the rigorous results of our analysis relies on this derivation.

This appendix is only meant to provide intuition to some of the readers. As a consequence, our treatment here will be heuristic.

2. A similar argument was presented in several other papers, including tutorials and reviews [DMM10, Mon12, DM13a].

As a consequence, we will sketch the main steps, highlighting differences from earlier work. Also, to be definite we will focus on the Symmetric Spiked Model. The asymmetric Spiked Model can be treated analogously.

According to the Symmetric Spiked Model, we are given observations $\mathbf{X} \in \mathbb{R}^{n \times n}$ generated according to

$$\mathbf{X} = \beta \mathbf{v}_0 \mathbf{v}_0^\top + \mathbf{Z}. \quad (211)$$

A possible estimation approach is to minimize the following cost function $\mathcal{H} : \mathbb{R}^n \rightarrow \mathbb{R}$

$$\mathcal{H}_{\mathbf{X}}(\mathbf{v}) \equiv -\frac{1}{2} \langle \mathbf{v}, \mathbf{X} \mathbf{v} \rangle + \sum_{i=1}^n \mathcal{I}_+(\mathbf{v}_i), \quad (212)$$

with $\mathcal{I}_+ : \mathbb{R} \rightarrow \mathbb{R}$ a penalty function that enforces $v_i \geq 0$. Namely $\mathcal{I}_+(x) = 0$ if $x \geq 0$ and $\mathcal{I}_+(x) = +\infty$ if $x < 0$.

Note that this cost function can be represented as a pairwise graphical model, with underlying graph $G = K_n$ (the complete graph over n vertices). Explicitly:

$$\mathcal{H}_{\mathbf{X}}(\mathbf{v}) = \sum_{(i,j) \in E(K_n)} \theta_{ij}(\mathbf{v}_i, \mathbf{v}_j) + \sum_{i=1}^n \theta_i(\mathbf{v}_i), \quad (213)$$

$$\theta_{ij}(\mathbf{v}_i, \mathbf{v}_j) = -\mathbf{X}_{ij} \mathbf{v}_i \mathbf{v}_j, \quad (214)$$

$$\theta_i(\mathbf{v}_i) = -\frac{1}{2} \mathbf{X}_{ii} \mathbf{v}_i^2 + \mathcal{I}_+(\mathbf{v}_i). \quad (215)$$

Hereafter, for simplicity, we shall neglect the diagonal terms $\mathbf{X}_{ii} = O(1/\sqrt{n})$ (these have lower order effect for large n).

Note that the cost function (212) is –in general– not bounded below on \mathbb{R}^n . A more careful argument would require to regularize the cost function, or to constrain the minimization on $\|\mathbf{v}\|_2 = \text{const.}$. However, since we are only interested in the optimum up to a global scaling, and in an iterative algorithm, we can omit such a regularization in our heuristic derivation.

Min-sum belief propagation can be used to approximate the minimizer of $\mathcal{H}_{\mathbf{X}}(\cdot)$, see e.g. [MM09]. It uses messages $J_{i \rightarrow j}^{(t)}(\mathbf{v}_i)$ that are a collection of functions of $\mathbf{v}_i \in \mathbb{R}$, indexed by ordered pairs of vertices $i \neq j$. The index t denotes the iteration number. Messages are updated as

follows

$$J_{i \rightarrow j}^{(t+1)}(\mathbf{v}_i) \cong I_+(\mathbf{v}_i) + \sum_{l \in [n] \setminus \{i, j\}} \min_{\mathbf{v}_l \in \mathbb{R}} \{ -\mathbf{X}_{il} \mathbf{v}_i \mathbf{v}_l + J_{l \rightarrow i}^{(t)}(\mathbf{v}_l) \}. \quad (216)$$

(Here and below \cong denotes equality of two functions—in this case, functions of \mathbf{v}_i —up to a global additive constant.)

At this point we follow the argument presented in [Mon12] (in the context of compressed sensing reconstruction). Namely, we use a second order approximation for $J_{i \rightarrow j}^{(t)}(\mathbf{v}_i)$:

$$J_{i \rightarrow j}^{(t)}(\mathbf{v}_i) \approx I_+(\mathbf{v}_i) - \alpha_{i \rightarrow j}^t \mathbf{v}_i + \frac{1}{2} \beta_{i \rightarrow j}^t \mathbf{v}_i^2 \quad (217)$$

Note that we added a term $I_+(\mathbf{v}_i)$ as suggested by Eq. (216). Substituting this term, a straightforward calculation yields

$$\min_{\mathbf{v}_l \in \mathbb{R}} \{ -\mathbf{X}_{il} \mathbf{v}_i \mathbf{v}_l + J_{l \rightarrow i}^{(t)}(\mathbf{v}_l) \} = -\frac{1}{2\beta_{l \rightarrow i}^t} (\alpha_{l \rightarrow i}^t + \mathbf{X}_{il} \mathbf{v}_i)^2 \mathbb{I}(\alpha_{l \rightarrow i}^t + \mathbf{X}_{il} \mathbf{v}_i \geq 0) \quad (218)$$

$$\approx -\frac{1}{2\beta_{l \rightarrow i}^t} (\alpha_{l \rightarrow i}^t + \mathbf{X}_{il} \mathbf{v}_i)^2 \mathbb{I}(\alpha_{l \rightarrow i}^t \geq 0) \quad (219)$$

$$= \text{const} - \mathbf{X}_{il} \frac{(\alpha_{l \rightarrow i}^t)_+}{\beta_{l \rightarrow i}^t} \mathbf{v}_i - \mathbf{X}_{il}^2 \frac{\mathbb{I}(\alpha_{l \rightarrow i}^t \geq 0)}{2\beta_{l \rightarrow i}^t} \mathbf{v}_i^2. \quad (220)$$

Substituting this in Eq. (216), we obtain equations for $\alpha_{i \rightarrow j}^t$ and $\beta_{i \rightarrow j}^t$:

$$\alpha_{i \rightarrow j}^{t+1} = \sum_{l \in [n] \setminus \{i, j\}} \mathbf{X}_{il} \frac{(\alpha_{l \rightarrow i}^t)_+}{\beta_{l \rightarrow i}^t}, \quad (221)$$

$$\beta_{i \rightarrow j}^{t+1} = - \sum_{l \in [n] \setminus \{i, j\}} \mathbf{X}_{il}^2 \frac{\mathbb{I}(\alpha_{l \rightarrow i}^t \geq 0)}{\beta_{l \rightarrow i}^t}. \quad (222)$$

Notice that, by the second equation and the law of large numbers, $\beta_{i \rightarrow j}^{t+1}$ concentrates around a value β^t independent of the vertices $i \rightarrow j$. A similar argument is made in the applications of AMP to compressed sensing, cf. for instance [Mon12]. We can therefore absorb this into an overall constants over the messages $\alpha_{i \rightarrow j}^t$ in Eq. (221):

$$\alpha_{i \rightarrow j}^{t+1} = c(t) \sum_{l \in [n] \setminus \{i, j\}} \mathbf{X}_{il} (\alpha_{l \rightarrow i}^t)_+. \quad (223)$$

Rather than computing the constant $c(t)$, we will fix it later by normalization requirements.

Again following [DMM10, Mon12, DM13a], the message $\alpha_{i \rightarrow j}^t$ depends weakly on the index j , and hence we will write

$$\alpha_{i \rightarrow j}^t = \mathbf{v}_i^t + \delta_{i \rightarrow j}^t, \quad (224)$$

and expand to first order in the variables $\{\delta_{i \rightarrow j}^t\}$. This calculation was described in a general setting in [BM11], and yields (defining $\mathbf{v}^t = (\mathbf{v}_1^t, \mathbf{v}_2^t, \dots, \mathbf{v}_n^t)^\top$):

$$\mathbf{v}^t = c(t) \mathbf{X}(\mathbf{v}^t)_+ - \tilde{\mathbf{b}}_t c(t-1) (\mathbf{v}^{t-1})_+, \quad (225)$$

where $\tilde{\mathbf{b}}_t \equiv c(t) \|(\mathbf{v}^t)_+\|_0/n$. Setting $c(t) = \sqrt{n}/\|(\mathbf{v}^t)_+\|_2$ yields the symmetric AMP updates introduced in Section 4.1.1.

B AMP fixed points are stationary points

In this appendix we prove the claim that AMP fixed points are stationary points of the non-negative PCA problem. For the sake of simplicity, we will focus on the symmetric problem that we reproduce here for the reader's convenience:

$$\begin{aligned} & \text{maximize} \quad \langle \mathbf{v}, \mathbf{X}\mathbf{v} \rangle, \\ & \text{subject to} \quad \mathbf{v} \geq 0, \quad \|\mathbf{v}\|_2 = 1. \end{aligned} \quad (226)$$

(The general case can be treated similarly.)

It is not hard to derive the stationarity (KKT) conditions for this point. The vector $\hat{\mathbf{v}} \in \mathbb{R}^n$ is a stationary point of the non-negative PCA problem if there exist $\mu \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$ such that

$$\mathbf{X}\hat{\mathbf{v}} = \lambda\hat{\mathbf{v}} - \mu, \quad (227)$$

$$\hat{\mathbf{v}} \geq 0, \quad (228)$$

$$\mu \geq 0, \quad (229)$$

$$\|\hat{\mathbf{v}}\|_2 = 1, \quad (230)$$

$$\langle \hat{\mathbf{v}}, \mu \rangle = 0. \quad (231)$$

Let \mathbf{v} a fixed point of AMP. We claim that $\hat{\mathbf{v}} = (\mathbf{v})_+ / \|(\mathbf{v})_+\|_2$ satisfies the above KKT conditions. Indeed, the fixed point condition reads

$$\frac{1}{\sqrt{n}}\mathbf{v} = \mathbf{X}\hat{\mathbf{v}} - \mathbf{b}_\infty \hat{\mathbf{v}}. \quad (232)$$

with $\mathbf{b}_\infty = \|(\mathbf{v})_+\|_0 / (\sqrt{n}\|(\mathbf{v})_+\|_2)$. By rearranging terms, we see that this is equivalent to the KKT condition if we identify

$$\lambda = \frac{1}{\sqrt{n}} \|\mathbf{v}\|_2 + \frac{\|(\mathbf{v})_+\|_0}{\sqrt{n} \|\mathbf{v}\|_2}, \quad (233)$$

$$\mu = \frac{1}{\sqrt{n}} \|(\mathbf{v})_-\|_2, \quad (234)$$

where $(\cdot)_-$ denotes the negative part.

C State evolution: Proofs of Proposition 4.1 and Proposition 4.2

In this appendix we characterize the high-dimensional behavior of AMP as per Proposition 4.1 and Proposition 4.2. The analogous results for rectangular matrices (namely, Propositions 4.3 and 4.4) follow from very similar arguments which we omit here.

It is convenient to first state two simple facts. The first one allows to control small perturbations of a given iterative scheme.

Lemma C.1. *Let \mathbf{X} be as in the statement of Proposition 4.1, and the sequences $\{\mathbf{u}^t\}_{t \geq 0}$, $\{\tilde{\mathbf{u}}^t\}_{t \geq 0}$ be defined by the recursions*

$$\mathbf{u}^{t+1} = \mathbf{X}g_t(\mathbf{u}^t) - \mathbf{a}_t g_{t-1}(\mathbf{u}^{t-1}), \quad (235)$$

$$\tilde{\mathbf{u}}^{t+1} = \mathbf{X}g_t(\tilde{\mathbf{u}}^t) - \mathbf{a}_t g_{t-1}(\tilde{\mathbf{u}}^{t-1}) + \Delta^t, \quad (236)$$

where $\mathbf{a}_t \in \mathbb{R}$ and $g_t : \mathbb{R}^n \rightarrow \mathbb{R}^n$.

Assume that $\lim_{n \rightarrow \infty} \|\mathbf{u}^0 - \tilde{\mathbf{u}}^0\|_2 / \sqrt{n} = 0$, $\limsup_{n \rightarrow \infty} \|\mathbf{u}^0\|_2^2 / n < \infty$ and, for every $t \in \{0, \dots, T\}$, we have the following, almost surely

1. $\lim_{n \rightarrow \infty} \|\Delta^t\|_2^2 / n = 0$.

2. $\limsup_{n \rightarrow \infty} |\mathbf{a}^t| < \infty$.

3. g_t is Lipschitz continuous with bounded Lipschitz constant. Namely, there exists $L_t \in \mathbb{R}$ independent of n such that $\|g_t(\mathbf{u}) - g_t(\mathbf{u}')\|_2 \leq L_t \|\mathbf{u} - \mathbf{u}'\|_2$ for all $\mathbf{u}, \mathbf{u}' \in \mathbb{R}^n$.

Then, for all $t \in \{0, 1, \dots, T+1\}$ we have

$$\lim_{n \rightarrow \infty} \frac{1}{n} \|\mathbf{u}^t - \tilde{\mathbf{u}}^t\|_2^2 = 0, \quad (237)$$

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \|\mathbf{u}^t\|_2^2 < \infty. \quad (238)$$

Proof. The proof is immediate by induction over t . We will prove Eq. (237): Eq. (238) follows by a similar argument. The case $t = 0$ holds by assumption. In order to prove the induction step, note that $\|\mathbf{X}\|_2 \leq \beta + \|\mathbf{Z}\|_2 \leq \beta + 3$ with probability larger than $1 - c^{-1}e^{-cn}$ for some $c > 0$ [AGZ09]. By triangular inequality

$$\|\mathbf{u}^{t+1} - \tilde{\mathbf{u}}^{t+1}\|_2 \leq \|\mathbf{X}\|_2 \|g_t(\mathbf{u}^t) - g_t(\tilde{\mathbf{u}}^t)\|_2 + |\mathbf{a}_t| \|g_{t-1}(\mathbf{u}^{t-1}) - g_{t-1}(\tilde{\mathbf{u}}^{t-1})\|_2 + \|\Delta^t\|_2 \quad (239)$$

$$\leq L(\beta + 3) \|\mathbf{u}^t - \tilde{\mathbf{u}}^t\|_2 + |\mathbf{a}_t| \|\mathbf{u}^{t-1} - \tilde{\mathbf{u}}^{t-1}\|_2 + \|\Delta^t\|_2, \quad (240)$$

where the second inequality holds with probability at least $1 - c^{-1}e^{-cn}$. The induction claim follows by dividing the above inequality by \sqrt{n} . \square

The second remark allows to establish limit results as in Proposition 4.1, once they have been established for a perturbed sequence.

Lemma C.2. Assume that the sequences of vectors $\mathbf{u} = \mathbf{u}(n)$, $\tilde{\mathbf{u}} = \tilde{\mathbf{u}}(n)$ satisfy

$$\lim_{n \rightarrow \infty} \frac{1}{n} \|\mathbf{u}(n) - \tilde{\mathbf{u}}(n)\|_2^2 = 0, \quad (241)$$

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \|\tilde{\mathbf{u}}(n)\|_2^2 < \infty, \quad (242)$$

and further assume $\mathbf{u}_0 = \mathbf{u}_0(n)$ be such that $\sup_n \|\mathbf{u}_0(n)\|_2 < \infty$. If $\lim_{n \rightarrow \infty} n^{-1} \sum_{i=1}^n \psi(\tilde{\mathbf{u}}_i, \sqrt{n}(\mathbf{u}_0)_i)$ exists for some pseudo-Lipschitz function ψ , then

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \psi(u_i, \sqrt{n} u_{0,i}) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \psi(\tilde{u}_i, \sqrt{n} u_{0,i}). \quad (243)$$

Proof. Using the pseudo-Lipschitz property of ψ , and Cauchy-Schwartz, we get

$$\frac{1}{n} \sum_{i=1}^n \left| \psi(u_i, \sqrt{n} u_{0,i}) - \psi(\tilde{u}_i, \sqrt{n} u_{0,i}) \right| \leq \frac{L}{n} \sum_{i=1}^n (1 + 2\sqrt{n}|(\mathbf{u}_0)_i| + |u_i| + |\tilde{\mathbf{u}}_i|) |u_i - \tilde{u}_i| \quad (244)$$

$$\leq \frac{L}{n} (\sqrt{n} + 2\sqrt{n}\|\mathbf{u}_0\|_2 + \|\mathbf{u}\|_2 + \|\tilde{\mathbf{u}}\|_2) \|\mathbf{u} - \tilde{\mathbf{u}}\|. \quad (245)$$

By Eqs. (241) and (242), $\limsup_{n \rightarrow \infty} \frac{1}{n} \|\tilde{\mathbf{u}}^t\|_2^2 < \infty$. Using this fact together with the other assumptions, we get from the last inequality

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \left| \psi(u_i, \sqrt{n} u_{0,i}) - \psi(\tilde{u}_i, \sqrt{n} u_{0,i}) \right| = 0, \quad (246)$$

which proves our claim. \square

C.1 Proof of Proposition 4.1

The proof consists in modifying the AMP sequence $\{\mathbf{v}^t\}_{t \geq 0}$ as to reduce ourselves to the setting of [BM11]. The first step consists in introducing a sequence $\{\mathbf{w}^t\}_{t \geq 0}$ defined by $\mathbf{w}^0 = (1, 1, \dots, 1)^\top$, $\mathbf{w}^{-1} = 0$ and letting, for all $t \geq 0$,

$$\mathbf{w}^{t+1} = \mathbf{X}(\mathbf{w}^t)_+ - \tilde{\mathbf{b}}_t(\mathbf{w}^{t-1})_+, \quad (247)$$

where $\tilde{\mathbf{b}}_t = \|(\mathbf{w}^t)_+\|_0/n$. The relation between this recursion and the original one is quite direct: they differ only by a normalization factor.

Lemma C.3. *Let $\{\mathbf{w}^t\}_{t \geq 0}$ be defined per Eq. (247) and $\{\mathbf{v}^t\}_{t \geq 0}$ be the AMP sequence, as per (AMP-sym). Then, for all $t \geq 1$ we have*

$$\mathbf{v}^t = \sqrt{n} \frac{\mathbf{w}^t}{\|(\mathbf{w}^{t-1})_+\|_2}. \quad (248)$$

Proof. The proof is by induction over the number of iterations. Let us first assume that it holds for all iterations until t , and prove it for iteration $t+1$. Multiplying Eq. (247) by $\sqrt{n}/\|(\mathbf{w}^t)_+\|_2$, we get

$$\sqrt{n} \frac{\mathbf{w}^{t+1}}{\|(\mathbf{w}^{t-1})_+\|_2} = \mathbf{X} \frac{(\mathbf{w}^t)_+ \sqrt{n}}{\|(\mathbf{w}^t)_+\|_2} - \frac{1}{\sqrt{n}} \frac{\|(\mathbf{w}^t)_+\|_0}{\|(\mathbf{w}^t)_+\|_2} (\mathbf{w}^{t-1})_+. \quad (249)$$

Note that the induction hypothesis implies $\mathbf{v}^t = c \mathbf{w}^t$ for some constant c , and hence

$$\sqrt{n} \frac{(\mathbf{w}^t)_+}{\|(\mathbf{w}^t)_+\|_2} = \sqrt{n} \frac{(\mathbf{v}^t)_+}{\|(\mathbf{v}^t)_+\|_2} = f(\mathbf{v}^t). \quad (250)$$

By the same argument and using $\|(\mathbf{v}^t)_+\|_2 = \sqrt{n} \|(\mathbf{w}^t)_+\|_2 / \|(\mathbf{w}^{t-1})_+\|_2$, we get

$$\frac{1}{\sqrt{n}} \frac{\|(\mathbf{w}^t)_+\|_0}{\|(\mathbf{w}^t)_+\|_2} (\mathbf{w}^{t-1})_+ = \frac{1}{n} \frac{\|(\mathbf{w}^t)_+\|_0 \|(\mathbf{w}^{t-1})_+\|_2}{\|(\mathbf{w}^t)_+\|_2} f(\mathbf{v}^{t-1}) \quad (251)$$

$$= \frac{1}{\sqrt{n}} \frac{\|(\mathbf{v}^t)_+\|_0}{\|(\mathbf{v}^t)_+\|_2} f(\mathbf{v}^{t-1}) = \mathbf{b}_t f(\mathbf{v}^{t-1}). \quad (252)$$

Using Eqs. (250) and (252) in Eq. (249), we obtain

$$\sqrt{n} \frac{\mathbf{w}^{t+1}}{\|(\mathbf{w}^{t-1})_+\|_2} = \mathbf{X} f(\mathbf{v}^t) - \mathbf{b}_t f(\mathbf{v}^{t-1}). \quad (253)$$

The induction step is completed by comparing this with (AMP-sym). The base case follow easily by a similar argument. \square

As a second step, we introduce a sequence $\{\mathbf{s}^t\}_{t \geq 0}$ defined as follows. First, we let μ_t, σ_t be scalars given by

$$\mu_{t+1} = \beta \mathbb{E}\{V(\mu_t V + \sigma_t G)_+\}, \quad (254)$$

$$\sigma_{t+1}^2 = \mathbb{E}\{(\mu_t V + \sigma_t G)_+^2\}. \quad (255)$$

with initial conditions $\mu_1 = \beta \mathbb{E}(V)$ and $\sigma_1 = 1$. Note that by Cauchy-Schwartz $\mu_{t+1} \leq \beta \sqrt{\mu_t^2 + \sigma_t^2}$ and $\sigma_{t+1}^2 \leq \mu_t^2 + \sigma_t^2$, whence $\mu_t, \sigma_t < \infty$ for all t . Further, since $G \geq 0$ with probability $1/2$, we also have $\mu_{t+1} \geq \beta \mu_t/2$, $\sigma_{t+1}^2 \geq \mu_t^2/2$, whence $\mu_t, \sigma_t \in (0, \infty)$ for all t .

Using these quantities (and recalling that $\mathbf{X} = \beta \mathbf{v}_0 \mathbf{v}_0^\top + \mathbf{Z}$ with $(\mathbf{Z})_{ij} \sim \mathcal{N}(0, 1)$, i.i.d. for $i < j$), we define

$$\mathbf{s}^{t+1} = \mathbf{Z} h_t(\mathbf{s}^t, \mathbf{v}_0 \sqrt{n}) - \mathbf{d}_t h_{t-1}(\mathbf{s}^{t-1}, \mathbf{v}_0 \sqrt{n}), \quad (256)$$

$$h_t(x, y) \equiv (x + \mu_t y)_+, \quad (257)$$

$$\mathbf{d}_t \equiv \frac{1}{n} \|(\mathbf{s}^t + \mu_t \mathbf{v}_0 \sqrt{n})_+\|_0, \quad (258)$$

As usual, here $h_t(\mathbf{s}^t, \mathbf{v}_0 \sqrt{n})$ is interpreted as the component-wise application of h_t . The initial condition is $\mathbf{s}^1 = \mathbf{w}^1 - \mu_1 \mathbf{v}_0$. This iteration is in the form of [JM13, Theorem 1] (and analogous to [BM11, Theorem 4]), which implies immediately the following.

Lemma C.4. *For any $t \geq 1$ and any pseudo-Lipshitz function $\psi : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$, we have, almost surely*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \psi(\mathbf{s}_i^t, \sqrt{n} v_{0,i}) = \mathbb{E}\{\psi(\sigma_t G, V)\}, \quad (259)$$

where expectation is with respect to $G \sim \mathcal{N}(0, 1)$ independent of V .

The sequences $\{\mathbf{s}^t\}_{t \geq 0}$ and $\{\mathbf{w}^t\}_{t \geq 0}$ are in fact closely related as we show next.

Lemma C.5. *For any $t \geq 1$ and any pseudo-Lipshitz function $\psi : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$, we have*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \|\mathbf{w}^t - \mu_t \mathbf{v}_0 - \mathbf{s}^t\|_2 = 0, \quad (260)$$

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \psi(\mathbf{w}_i^t, \sqrt{n} v_{0,i}) = \mathbb{E}\{\psi(\mu_t V + \sigma_t G, V)\}, \quad (261)$$

where expectation is with respect to $G \sim \mathcal{N}(0, 1)$ independent of V .

Proof. Define $\tilde{\mathbf{s}}^t = \mathbf{w}^t - \mu_t \mathbf{v}_0 \sqrt{n}$. Then Eq. (247) implies immediately

$$\tilde{\mathbf{s}}^{t+1} = \mathbf{Z} h_t(\mathbf{s}^t, \mathbf{v}_0 \sqrt{n}) - \mathbf{d}_t h_{t-1}(\mathbf{s}^{t-1}, \mathbf{v}_0 \sqrt{n}) + \mathbf{\Delta}^t, \quad (262)$$

$$\mathbf{\Delta}^t \equiv (\beta \langle \mathbf{v}_0, (\mu_t \mathbf{v}_0 \sqrt{n} + \tilde{\mathbf{s}}^t)_+ \rangle - \mu_{t+1} \sqrt{n}) \mathbf{v}_0 + (\mathbf{d}_t - \tilde{\mathbf{b}}_t) (\mu_{t-1} \mathbf{v}_0 \sqrt{n} + \tilde{\mathbf{s}}^{t-1})_+. \quad (263)$$

Next note that our claim is equivalent to the following holding for every pseudo-Lipshitz ψ and every iteration number ℓ :

$$\lim_{n \rightarrow \infty} \frac{1}{n} \|\tilde{\mathbf{s}}^\ell - \mathbf{s}^\ell\|_2^2 = 0, \quad (264)$$

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \psi(\tilde{\mathbf{s}}_i^\ell, \sqrt{n} v_{0,i}) = \mathbb{E}\{\psi(\sigma_\ell G, V)\}. \quad (265)$$

We prove this by induction over the iteration number. Assume that the claim indeed holds for all $\ell \in \{1, \dots, t\}$. Then comparing Eq. (257) and Eq. (262) we obtain –by Lemma C.1– that Eq. (264) holds for all $\ell \in \{1, \dots, t, t+1\}$ provided we can prove that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \|\Delta^\ell\|_2^2 = 0 \quad (266)$$

for all $\ell \in \{1, \dots, t\}$. Now we have

$$\frac{1}{n} \|\Delta^\ell\|_2^2 \leq D_1^\ell + D_2^\ell, \quad (267)$$

$$D_1^\ell \equiv 2(\beta \langle \mathbf{v}_0, (\mu_\ell \mathbf{v}_0 + \tilde{\mathbf{s}}_\ell n^{-1/2})_+ \rangle - \mu_{\ell+1})^2, \quad (268)$$

$$D_2^\ell \equiv 4(\mathbf{d}_\ell - \tilde{\mathbf{b}}_\ell)^2 \left(\mu_{\ell-1}^2 + \frac{1}{n} \|\tilde{\mathbf{s}}^{\ell-1}\|_2^2 \right). \quad (269)$$

Now, using $\psi(x, y) = (\mu_\ell y + x)_+ y$ in Eq. (265) we get for all $\ell \in \{1, \dots, t\}$, almost surely

$$\lim_{n \rightarrow \infty} \beta \langle \mathbf{v}_0, (\mu_\ell \mathbf{v}_0 + \tilde{\mathbf{s}}_\ell n^{-1/2})_+ \rangle = \beta \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \psi(\tilde{\mathbf{s}}_i^\ell, \sqrt{n} v_{0,i}) \quad (270)$$

$$= \beta \mathbb{E} \{ V(\mu_\ell V + \sigma_\ell G)_+ \} = \mu_{\ell+1}. \quad (271)$$

In other words $D_1^\ell \rightarrow 0$ almost surely.

Using again Eq. (265) we have $n^{-1} \|\tilde{\mathbf{s}}^{\ell-1}\|_2^2 \rightarrow \sigma_{\ell-1}^2$ almost surely. Therefore, since μ_t is finite for all t , we get $|D_2^\ell| \leq C |\mathbf{d}_\ell - \tilde{\mathbf{b}}_\ell|$ for some constant C bounded uniformly in n . Finally, fix $\delta > 0$ and let

$$\psi_\delta(x) = \begin{cases} 1 & \text{if } x > \delta, \\ x/\delta & \text{if } x \in (0, \delta), \\ 0 & \text{otherwise.} \end{cases} \quad (272)$$

Then

$$\mathbf{d}_\ell - \tilde{\mathbf{b}}_\ell = \frac{1}{n} \sum_{i=1}^n \left[\mathbb{I}(\mathbf{s}_i^\ell + \mu_\ell v_{0,i} \sqrt{n} > 0) - \mathbb{I}(\tilde{\mathbf{s}}_i^\ell + \mu_\ell v_{0,i} \sqrt{n} > 0) \right] \quad (273)$$

$$\leq \frac{1}{n} \sum_{i=1}^n \left[\psi_\delta(\mathbf{s}_i^\ell + \mu_\ell v_{0,i} \sqrt{n} + \delta) - \psi_\delta(\tilde{\mathbf{s}}_i^\ell + \mu_\ell v_{0,i} \sqrt{n}) \right] \quad (274)$$

$$\leq \frac{1}{n} \sum_{i=1}^n \left[\psi_\delta(\mathbf{s}_i^\ell + \mu_\ell v_{0,i} \sqrt{n} + \delta) - \psi_\delta(\mathbf{s}_i^\ell + \mu_\ell v_{0,i} \sqrt{n}) + \frac{1}{\delta} \|\mathbf{s}_2^\ell - \tilde{\mathbf{s}}_i^\ell\| \right] \quad (275)$$

$$\leq \frac{1}{n} \sum_{i=1}^n \left[\psi_\delta(\mathbf{s}_i^\ell + \mu_\ell v_{0,i} \sqrt{n} + \delta) - \psi_\delta(\mathbf{s}_i^\ell + \mu_\ell v_{0,i} \sqrt{n}) \right] + \frac{1}{\delta \sqrt{n}} \|\mathbf{s}^\ell - \tilde{\mathbf{s}}^\ell\|_2. \quad (276)$$

Taking $n \rightarrow \infty$ and using Eqs. (264), (265), we conclude that

$$\limsup_{n \rightarrow \infty} (\mathbf{d}_\ell - \tilde{\mathbf{b}}_\ell) \leq \mathbb{E} \left\{ \psi_\delta(\mu_\ell V + \sigma_\ell G + \delta) - \psi_\delta(\mu_\ell V + \sigma_\ell G) \right\}, \quad (277)$$

And since G has a density with respect to Lebesgue measure, this implies, by letting $\delta \rightarrow 0$, $\limsup_{n \rightarrow \infty} (d_\ell - \tilde{b}_\ell) \leq 0$. A lower bound is obtained by a similar argument yielding

$$\lim_{n \rightarrow \infty} (d_\ell - \tilde{b}_\ell) = 0. \quad (278)$$

and hence $D_2^\ell \rightarrow 0$. By Eq. (267) we have $\|\Delta^\ell\|_2^2/n \rightarrow 0$ and hence Eq. (264) holds for $\ell = t+1$. Finally, Eq. (265) follows for $\ell = t+1$ by Lemma C.2, Lemma C.4 and Eq. (264) (for $\ell = t+1$). \square

Finally, the proof of Proposition 4.1 follows immediately from Lemma C.5, using Lemma C.3. Indeed, by applying C.5 to $\psi(x, y) = (x)_+^2$, we get, almost surely

$$\lim_{n \rightarrow \infty} \frac{1}{\sqrt{n}} \|(\mathbf{w}^t)_+\|_2 = \sqrt{\mathbb{E}\{(\mu_t V + \sigma_t G)_+^2\}} = \sigma_{t+1} \in (0, \infty). \quad (279)$$

Hence, for any pseudo-Lipshitz function ψ , almost surely

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \psi(v_i^t, \sqrt{n}v_{0,i}) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \psi\left(\sqrt{n} \frac{\mathbf{w}_i^t}{\|(\mathbf{w}^{t-1})_+\|_2}, \sqrt{n}v_{0,i}\right) \quad (280)$$

$$= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \psi\left(\frac{\mathbf{w}_i^t}{\sigma_{t+1}}, \sqrt{n}v_{0,i}\right) \quad (281)$$

$$\mathbb{E} \left\{ \psi\left(\frac{\mu_t}{\sigma_t} V + G, V\right) \right\}. \quad (282)$$

We conclude by noting that –by comparison of Eq. (35) with Eqs. (254) and (255)– it follows that $\tau_t = \mu_t/\sigma_t$ for all t .

Finally the claim $\psi(x, y) = \mathbb{I}(x \leq a)$, follows by a standard argument already used above. Namely, we use the bounds $\psi_\delta(x) \leq \mathbb{I}(x \leq a) \leq \psi_\delta(x + \delta)$, with the definition in Eq. (272), apply the previous result to the Lipschitz functions $\psi_\delta(x)$, $\psi_\delta(x + \delta)$ and eventually let $\delta \rightarrow 0$.

C.2 Proof of Proposition 4.2

The proof is based on a version of state evolution that describes the asymptotic joint distribution of $\mathbf{v}^t, \mathbf{v}^s$ for two distinct times t, s . For this purpose, we define a function $H_V : [-1, 1] \times \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}$ as follows

$$H_V(Q; \tau_1, \tau_2) = \frac{\mathbb{E}\{(\tau_1 V + G_1)_+(\tau_2 V + G_2)_+\}}{\sqrt{\mathbb{E}\{(\tau_1 V + G_1)_+^2\} \mathbb{E}\{(\tau_2 V + G_2)_+^2\}}}, \quad (283)$$

where expectation is with respect to the centered Gaussian vector (G_1, G_2) with $\mathbb{E}\{G_1^2\} = \mathbb{E}\{G_2^2\} = 1$, $\mathbb{E}\{G_1 G_2\} = Q$, independent of V .

Let the state evolution sequence $\{\tau_t\}_{t \geq 0}$ be given as per Eq. (35), and define recursively $\{Q_{t,s}\}_{t,s \geq 0}$ by letting

$$Q_{t+1,s+1} = H_V(Q_{t,s}; \tau_t, \tau_s). \quad (284)$$

with initial condition $Q_{1,1} = 1$ and, for $t \geq 2$,

$$Q_{t,1} = \frac{\mathbb{E}\{(\tau_{t-1} V + G)_+\}}{\sqrt{\mathbb{E}\{(\tau_{t-1} V + G)_+^2\}}}. \quad (285)$$

Then we have the following extension of state evolution.

Lemma C.6. *With the above definitions, let $\psi : \mathbb{R}^3 \rightarrow \mathbb{R}$ be a pseudo-Lipschitz function. Then, for any $t, s \geq 1$, we have, almost surely*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \psi(\mathbf{v}_i^t, \mathbf{v}_i^s, \sqrt{n}v_{0,i}) = \mathbb{E}\{\psi(\tau_t V + G_t, \tau_s V + G_s, V)\}, \quad (286)$$

where expectation is with respect to the centered Gaussian vector (G_t, G_s) with $\mathbb{E}\{G_t^2\} = \mathbb{E}\{G_s^2\} = 1$, $\mathbb{E}\{G_t G_s\} = Q_{t,s}$, independent of V .

Proof. Very similar statements were proven, for instance in [BM11, Theorem 4.2] or [DM13b, Lemma C1]. The construction is always the same, and we will only sketch the first steps. Thanks to Lemma C.5, it is sufficient to prove that, for $\{\mathbf{s}^t\}_{t \geq 0}$ defined per Eq. (257), we have

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \psi(\mathbf{s}_i^t, \mathbf{s}_i^s, \sqrt{n}v_{0,i}) = \mathbb{E}\{\psi(\Gamma_t, \Gamma_s, V)\}, \quad (287)$$

where (Γ_t, Γ_s) is a centered Gaussian vector with $\mathbb{E}\{\Gamma_t^2\} = \sigma_t^2$, $\mathbb{E}\{\Gamma_s^2\} = \sigma_s^2$, $\mathbb{E}\{\Gamma_t \Gamma_s\} = \sigma_t \sigma_s Q_{t,s}$.

In order to prove the last claim, we fix a maximum time T , and consider all $t, s \in \{0, 1, \dots, T-1\}$. We then define $\mathbf{r}^t \in (\mathbb{R}^T)^n$, $t \in \{0, 1, \dots, T-1\}$ that we can think of either as a vector of length n , with entries in \mathbb{R}^T , or as a matrix with dimensions $n \times T$. With the last interpretation in mind, \mathbf{r}^t is defined as a matrix whose first $t+1$ columns are $\mathbf{s}^0, \mathbf{s}^1, \dots, \mathbf{s}^t$, and the others vanish, namely

$$\mathbf{r}^t = \begin{bmatrix} \mathbf{s}^0 & \mathbf{s}^1 & \dots & \mathbf{s}^t & 0 & \dots & 0 \end{bmatrix}. \quad (288)$$

Define $\mathbf{h}_t : \mathbb{R}^{T+1} \rightarrow \mathbb{R}^T$ by letting

$$\mathbf{h}_t(s_0, s_1, \dots, s_{T-1}; v) \equiv (s_0, h_0(s_0; v), \dots, h_{t-1}(s_{t-1}; v), h_t(s_t; v), 0, \dots, 0); \quad (289)$$

Then it is easy to see that Eq. (257) implies

$$\mathbf{r}^{t+1} = \mathbf{Z} \mathbf{h}_t(\mathbf{r}^t; \sqrt{n} \mathbf{v}_0) - \mathbf{h}_t(\mathbf{r}^t; \sqrt{n} \mathbf{v}_0) \mathbf{D}_t, \quad (290)$$

for a certain sequence of matrices $\mathbf{D}_t \in \mathbb{R}^{T \times T}$. The proof then follows by applying [JM13, Theorem 1] to $\{\mathbf{r}^t\}_{t \geq 0}$. \square

The next lemma provides the basic tool for applying the state evolution method to prove our claim.

Lemma C.7. *Let $\{Q_{t,s}\}_{t,s \geq 1}$ be defined as above using the two times state evolution recursion (284). Then*

$$\lim_{t \rightarrow \infty} Q_{t,t+1} = 1. \quad (291)$$

Before proving this Lemma, we state a useful general fact (which appeared already in specific forms in [BM12, DM13b]).

Lemma C.8. Let $h : \mathbb{R}^2 \rightarrow \mathbb{R}$ be a Borel function, W, Z_1, Z_2 random variables, and \mathbb{P}_q a probability distribution such that –under \mathbb{P}_q – (Z_1, Z_2) is a centered Gaussian vector independent of W , with covariance given $\mathbb{E}_q(Z_1^2) = \mathbb{E}_q(Z_2^2) = 1$ and $\mathbb{E}_q\{Z_1 Z_2\} = q$. Assume $\mathbb{E}\{h(Z_1, W)^2\} < \infty$ and define

$$\mathcal{H}(q) \equiv \mathbb{E}_q\{h(Z_1, W)h(Z_2, W)\}. \quad (292)$$

Then $q \mapsto \mathcal{H}(q)$ is non-decreasing and convex on $[0, 1]$. Further, unless $h(x, y)$ is affine in x , it is strictly convex. Finally assuming h is weakly differentiable, and denoting by $\partial_1 h$ its derivative with respect to the first argument, we have

$$\left. \frac{d\mathcal{H}}{dq} \right|_{q=1} = \mathbb{E}\{[\partial_1 h(Z, W)]^2\}. \quad (293)$$

Proof. First consider the case of $h(x, y) = h(x)$ independent of the second argument. Let $\{X_t\}_{t \geq 0}$ be the stationary Ornstein–Uhlenbeck process. We remind that this is a Gaussian stationary process with vanishing expectation $\mathbb{E}(X_t) = 0$ and covariance $\mathbb{E}(X_t X_s) = e^{-|t-s|}$. Then, by substitution we have

$$\mathcal{H}(q) = \mathbb{E}\{h(X_0)h(X_t)\} \Big|_{t=\log(1/q)}, \quad (294)$$

We recall that the Ornstein–Uhlenbeck is a reversible Markov process with stationary distribution given by the standad Gaussian measure. It follows (see for instance [BGL13], Section 2.7.6 and in particular Eq. (2.7.1) and following) that we have the spectral representation (for $t = \log(1/q)$ and $c_\ell = \langle \phi_\ell, h \rangle$, ϕ_ℓ the ℓ -th eigenfunction of the Ornstein–Uhlenbeck generator)

$$\mathcal{H}(q) = \sum_{\ell=0}^{\infty} c_\ell^2 e^{-\ell t} = \sum_{\ell=0}^{\infty} c_\ell^2 q^\ell, \quad (295)$$

whence the \mathcal{H} is non-decreasing and convex. Strict convexity follows since $c_\ell \neq 0$ for some $\ell \geq 2$ as long as $h(x)$ is non-linear.

Finally, if h depends on its second argument as well, we have $\mathcal{H}(q) = \mathbb{E}\{\mathcal{H}_W(q)\}$, with $\mathcal{H}_W(q) \equiv \mathbb{E}_q\{h(Z_1, W)h(Z_2, W)|W\}$. Using independence of (Z_1, Z_2) and W , the previous proof applies to \mathcal{H}_W for almost every W and, by linearity, to $\mathcal{H}(q)$.

Equation (293) follows by writing

$$\mathcal{H}(q) = \mathbb{E}_q\{h(Z, W)^2\} - \frac{1}{2}\mathbb{E}_q\{[h(Z_1, W) - h(Z_2, W)]^2\}, \quad (296)$$

with $Z \sim \mathcal{N}(0, 1)$. The claim follows by using the representation $Z_1 = aX + bY$, $Z_2 = aX - bY$, with X, Y independent standard normal, $a = \sqrt{(1+q)/2}$, $b = \sqrt{(1-q)/2}$, and Taylor expanding the right hand side in b . \square

We are now in position to prove Lemma C.7.

Proof of Lemma C.7. Recall that $\lim_{t \rightarrow \infty} \tau_t = \mathsf{T}_V(\beta) \in (0, \infty)$, cf. Lemma 6.6. Letting $\tau_* \equiv \mathsf{T}_V(\beta)$, we define $H_V^*(Q) \equiv H_V(Q; \tau_*, \tau_*)$, i.e.

$$H_V^*(Q) = \frac{\mathbb{E}\{(\tau_* V + G_1)_+ (\tau_* V + G_2)_+\}}{\sqrt{\mathbb{E}\{(\tau_* V + G_1)_+^2\} \mathbb{E}\{(\tau_* V + G_2)_+^2\}}}, \quad (297)$$

with (G_1, G_2) a centered Gaussian vector with $\mathbb{E}(G_1^2) = \mathbb{E}(G_2^2) = 1$ and $\mathbb{E}\{G_1 G_2\} = Q$. By Lemma C.8, the function $Q \mapsto H_V^*(Q)$ is strictly convex and monotone increasing in $[0, 1]$. Further we have $H_V^*(1) = 1$ and, for $G \sim \mathcal{N}(0, 1)$,

$$\left. \frac{d}{dQ} H_V^*(Q) \right|_{Q=1} = \frac{\mathbb{P}(\tau_* V + G \geq 0)}{\mathbb{E}\{(\tau_* V + G)_+^2\}}. \quad (298)$$

Note that

$$\mathbb{E}\{(\tau_* V + G)_+^2\} = \mathbb{E}\{(\tau_* V + G)(\tau_* V + G)_+\} \quad (299)$$

$$= \tau_* \mathbb{E}\{V(\tau_* V + G)_+\} + \mathbb{E}\{G(\tau_* V + G)_+\} \quad (300)$$

$$\geq \mathbb{P}(\tau_* V + G \geq 0), \quad (301)$$

where the last inequality follows since $V \geq 0$, and applying Stein's Lemma to the second term (see Lemma E.1, whereby we consider the weakly differentiable function $f(G) = (c + G)_+$ for $c = \tau_* V$, and subsequently take expectation over V .) . We therefore have

$$\left. \frac{d}{dQ} H_V^*(Q) \right|_{Q=1} \leq 1, \quad (302)$$

and therefore, by convexity, $H_V^*(Q) > Q$ for all $Q \in [0, 1)$.

Now, for ease of notation, let $Q_t \equiv Q_{t,t+1}$. Note that $H_V(Q; \tau_1, \tau_2) \in [0, 1]$ for all $Q \in [0, 1]$: indeed, for $Q \geq 0$, the random variables $(\tau_1 V + G_1)_+$ and $(\tau_2 V + G_2)_+$ are non-decreasing functions of positively correlated ones, and hence are positively correlated. Therefore $Q_t \in [0, 1]$ for all t . Assume by contradiction that Q_t does not converge to 1, and let $Q_* \equiv \liminf_{t \rightarrow \infty} Q_t$. Let $\{t(k)\}_{k \in \mathbb{N}}$ be a subsequence with $\lim_{k \rightarrow \infty} Q_{t(k)} = Q_*$. Since $\tau_t \rightarrow \tau_*$, H_V is continuous and H_V^* is non-decreasing, we have

$$Q_* = \lim_{k \rightarrow \infty} Q_{t(k)} \quad (303)$$

$$= \lim_{k \rightarrow \infty} \inf H_V(Q_{t(k)-1}; \tau_{t(k)}, \tau_{t(k)-1}) \quad (304)$$

$$= \lim_{k \rightarrow \infty} \inf H_V^*(Q_{t(k)-1}) \quad (305)$$

$$\geq H_V^*(Q_*). \quad (306)$$

This contradicts the previous remark that $H_V^*(Q) > Q$ for all $Q \in [0, 1)$, and hence proves the claim that $Q_t \rightarrow 1$. \square

We are now in position to prove our claim (4.2). First note that –by triangular inequality– it is sufficient to consider the case $\ell = 1$. Using Lemma C.6 for $\psi(x, y, z) = (x - y)^2$ and $s = t + 1$, we get, almost surely

$$\lim_{n \rightarrow \infty} \frac{1}{n} \|\mathbf{v}^t - \mathbf{v}^{t+1}\|_2^2 = \mathbb{E}\{(\tau_t V + G_t - \tau_{t+1} V - G_{t+1})^2\} = (\tau_t - \tau_{t+1})^2 + 2(1 - Q_{t,t+1}). \quad (307)$$

Since by Lemma 6.6 the sequence τ_t converges to a finite limit as $t \rightarrow \infty$, we have $\lim_{t \rightarrow \infty} (\tau_t - \tau_{t+1}) = 0$. Hence taking the limit $t \rightarrow \infty$ in the last expression and using Lemma C.7, we obtain the desired result.

D Proof of Theorems 2 and 4

Proof of Theorem 2. For the sake of clarity, we will note the dimension index n for a matrix $\mathbf{X}_n \in \mathbb{R}^{n \times n}$, distributed according to the Symmetric Spiked Model. In order to prove Eq. (13) (i.e. $\lim_{n \rightarrow \infty} \lambda^+(\mathbf{X}_n) = \mathbf{R}_V^{\text{sym}}(\mathbf{T}_V(\beta))$ almost surely), we need to prove:

$$\mathbb{P} \left[\liminf_{n \rightarrow \infty} \lambda^+(\mathbf{X}_n) \geq \mathbf{R}_V^{\text{sym}}(\mathbf{T}_V(\beta)) \right] = 1 \quad \text{and} \quad \mathbb{P} \left[\limsup_{n \rightarrow \infty} \lambda^+(\mathbf{X}_n) \leq \mathbf{R}_V^{\text{sym}}(\mathbf{T}_V(\beta)) \right] = 1. \quad (308)$$

- Theorem 6 states that there exists a deterministic sequence $\{\delta_t\}_t$ such that $\lim_t \delta_t = 0$ and

$$\mathbb{P} \left[\lim_{n \rightarrow \infty} \langle \hat{\mathbf{v}}^t, \mathbf{X}_n \hat{\mathbf{v}}^t \rangle \geq \mathbf{R}_V^{\text{sym}}(\mathbf{T}_V(\beta)) - \delta_t \right] = 1.$$

It follows, using $\lambda^+(\mathbf{X}_n) \geq \langle \hat{\mathbf{v}}^t, \mathbf{X}_n \hat{\mathbf{v}}^t \rangle$, and taking the intersection of these events for $t \in \mathbb{N}$, that

$$\mathbb{P} \left[\liminf_{n \rightarrow \infty} \lambda^+(\mathbf{X}_n) \geq \mathbf{R}_V^{\text{sym}}(\mathbf{T}_V(\beta)) \right] = 1.$$

- Since the function $\mathbf{X}_n \mapsto \max \{ \langle \mathbf{v}, \mathbf{X}_n \mathbf{v} \rangle : \mathbf{v} \geq 0, \|\mathbf{v}\|_2 \leq 1 \}$ is 1-Lipschitz continuous, then using the upper bound of Lemma 3.1 and Gaussian isoperimetry, for any $s > 0$ we have, with probability at least $1 - \exp\{-ns^2/2\}$,

$$\lambda^+(\mathbf{X}_n) \leq \mathbf{R}_V^{\text{sym}}(\mathbf{T}_V(\beta)) + s.$$

Taking $s = \sqrt{(4 \log n)/n}$, with probability at least $1 - n^{-2}$,

$$\lambda^+(\mathbf{X}_n) \leq \mathbf{R}_V^{\text{sym}}(\mathbf{T}_V(\beta)) + \frac{4 \log n}{n}.$$

Hence $\limsup_{n \rightarrow \infty} \lambda^+(\mathbf{X}_n) \leq \mathbf{R}_V^{\text{sym}}(\mathbf{T}_V(\beta))$ almost surely by Borel-Cantelli.

This concludes the proof of Eq. (13). Equation (14) follows immediately from Lemma 3.1 since $\lim_{x \rightarrow 0} \Delta(x) = 0$, and we know that the sequence $\lambda^+(\mathbf{X})$ converges almost surely to $\mathbf{R}_V^{\text{sym}}(\mathbf{T}_V(\beta))$.

In order to prove the limit behavior as $\varepsilon \rightarrow 0$ of Eqs. (15) and (16) we refer to Lemma 6.8 in Section 6.1 that establish the limit behavior of functions of interest $\mathbf{T}_V, \mathbf{F}_V, \mathbf{G}_V$ uniformly over the class of probability distributions \mathcal{P} . We know, thanks to Definition 2.3, Lemma 6.8, and uniform continuity on the interval $[0, 1]$ of the square function $x \mapsto x^2$, and on $\mathbb{R}_{\geq 0}$ of

$$\mathbf{F}_0 : x \mapsto \frac{x}{\sqrt{1/2 + x^2}}, \quad \mathbf{G}_0 : x \mapsto \frac{1/2}{\sqrt{1/2 + x^2}} \quad \text{and} \quad \mathbf{T}_0 : \beta \mapsto \begin{cases} 0 & \text{if } \beta \leq 1/\sqrt{2}, \\ \sqrt{\beta^2 - (1/2)} & \text{otherwise,} \end{cases}$$

that for any $\kappa > 0$, one can find $\varepsilon_0 = \varepsilon_0(\kappa)$ such that for any $\varepsilon < \varepsilon_0$ and $\mu_V \in \mathcal{P}_\varepsilon$, we have, for $\beta \geq 0$, $|\mathbf{F}_V(\mathbf{T}_V(\beta)) - \mathbf{F}_0(\mathbf{T}_0(\beta))| \leq \kappa$ and

$$|\beta \mathbf{F}_V(\mathbf{T}_V(\beta))^2 + 2 \mathbf{G}_V(\mathbf{T}_V(\beta)) - [\beta \mathbf{F}_0(\mathbf{T}_0(\beta))^2 + 2 \mathbf{G}_0(\mathbf{T}_0(\beta))]| \leq \kappa.$$

This proves uniform convergence of $\mathbf{F}_V(\mathbf{T}_V(\cdot))$ to $\mathbf{F}_0(\mathbf{T}_0(\cdot))$ and of $\beta \mathbf{F}_V(\mathbf{T}_V(\cdot))^2 + 2 \mathbf{G}_V(\mathbf{T}_V(\cdot))$ to $\beta \mathbf{F}_0(\mathbf{T}_0(\cdot))^2 + 2 \mathbf{G}_0(\mathbf{T}_0(\cdot))$. Since we have

$$\beta \mathbf{F}_0(\mathbf{T}_0(\beta))^2 + 2 \mathbf{G}_0(\mathbf{T}_0(\beta)) = \begin{cases} \sqrt{2} & \text{if } \beta \leq 1/\sqrt{2}, \\ \beta + 1/(2\beta) & \text{otherwise,} \end{cases}$$

and

$$F_0(T_0(\beta)) = \begin{cases} 0 & \text{if } \beta \leq 1/\sqrt{2}, \\ \sqrt{1 - 1/(2\beta^2)} & \text{otherwise,} \end{cases}$$

the result is proved. \square

Proof of Theorem 4. In order to prove Theorem 4 we proceed as for Theorem 2. We consider a sequence of random matrices $\{\mathbf{X}_n\}_{n \geq 1}$ of size $n \times p$, generated according to the Spiked Model. We use Lemma 3.2 and Gaussian isoperimetry for the 1-Lipschitz function

$$\max \{ \langle \mathbf{u}, \mathbf{X}_n \mathbf{v} \rangle : \|\mathbf{u}\|_2 \leq 1, \|\mathbf{v}\|_2 \leq 1, \mathbf{v} \geq 0 \} ,$$

to conclude that with probability at least $1 - n^{-2}$,

$$\sigma^+(\mathbf{X}_n) \leq R_V^{\text{rec}}(S_V(\beta, \alpha)/\sqrt{\alpha}) + \frac{\log n}{n} . \quad (309)$$

This proves, using Borel-Cantelli Lemma, that $\limsup_{n \rightarrow \infty} \sigma^+(\mathbf{x}_n) \leq R_V^{\text{rec}}(S_V(\beta)/\sqrt{\alpha})$ almost surely.

By Theorem 7 there exists a deterministic sequence $\{\delta_t\}_t$ such that

$$\mathbb{P} [\langle \hat{\mathbf{u}}^t, \mathbf{X}_n \hat{\mathbf{v}}^t \rangle \geq R_V^{\text{rec}}(S_V(\beta, \alpha)) - \delta_t] = 1 . \quad (310)$$

Since $\sigma^+(\mathbf{X}_n) \geq \langle \hat{\mathbf{u}}^t, \mathbf{X}_n \hat{\mathbf{v}}^t \rangle$, and by taking the intersection over $t \in \mathbb{N}$, we get $\liminf_{n \rightarrow \infty} \sigma^+(\mathbf{X}_n) \geq R_V^{\text{rec}}(S_V(\beta)) = 1$ almost surely. This concludes the proof of Eq. (20), i.e. $\lim_{n \rightarrow \infty} \sigma^+(\mathbf{X}_n) = R_V^{\text{rec}}(S_V(\beta, \alpha))$.

Together with Lemma 3.2, and using $\lim_{x \rightarrow 0} \Delta(x) = 0$, this implies Eq. (21), i.e. $\lim_{n \rightarrow \infty} \langle \mathbf{v}_0, \mathbf{v}^+ \rangle = F_V(S_V(\beta, \alpha)/\sqrt{\alpha})$ almost surely.

Finally the proof of Eqs. (22) and (23) follows from Lemma 6.8 as in the symmetric case. \square

E Stein's lemma

We use several times Stein's lemma [Ste72] in its scalar form.

Lemma E.1 ([Ste72]). *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be weakly differentiable, with derivative f' . If $G \sim \mathcal{N}(0, 1)$, then*

$$\mathbb{E}[G f(G)] = \mathbb{E}[f'(G)] , \quad (311)$$

Provided the two expectations are finite.

References

- [AGKM12] Sanjeev Arora, Rong Ge, Ravindran Kannan, and Ankur Moitra, *Computing a non-negative matrix factorization—provably*, Proceedings of the 44th symposium on Theory of Computing, ACM, 2012, pp. 145–162.

- [AGZ09] G. W. Anderson, A. Guionnet, and O. Zeitouni, *An introduction to random matrices*, Cambridge University Press, 2009.
- [And63] Theodore Wilbur Anderson, *Asymptotic theory for principal component analysis*, Annals of Mathematical Statistics (1963), 122–148.
- [BAD09] S. Burer, K. M. Anstreicher, and M. Dur, *The difference between 5×5 doubly nonnegative and completely positive matrices*, Linear Algebra and its Applications **431**(9) (2009), 1539–1552.
- [BBAP05] Jinho Baik, Gérard Ben Arous, and Sandrine Péché, *Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices*, Annals of Probability (2005), 1643–1697.
- [BGGM11] Florent Benaych-Georges, Alice Guionnet, and Mylène Maida, *Fluctuations of the extreme eigenvalues of finite rank deformations of random matrices*, Electron. J. Probab **16** (2011), 1621–1662.
- [BGGM12] Florent Benaych-Georges, Alice Guionnet, and Mylène Maïda, *Large deviations of the extreme eigenvalues of random deformations of matrices*, Probability Theory and Related Fields **154** (2012), no. 3-4, 703–751.
- [BGL13] Dominique Bakry, Ivan Gentil, and Michel Ledoux, *Analysis and geometry of markov diffusion operators*, vol. 348, Springer Science & Business Media, 2013.
- [BGN11] Florent Benaych-Georges and Raj Rao Nadakuditi, *The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices*, Advances in Mathematics **227** (2011), no. 1, 494–521.
- [BGN12] ———, *The singular values and vectors of low rank perturbations of large rectangular random matrices*, Journal of Multivariate Analysis **111** (2012), 120–135.
- [Bis99] Christopher M Bishop, *Bayesian pca*, Advances in neural information processing systems (1999), 382–388.
- [BLM12] Mohsen Bayati, Marc Lelarge, and Andrea Montanari, *Universality in polytope phase transitions and message passing algorithms*, arXiv:1207.7321 (2012).
- [BM11] M. Bayati and A. Montanari, *The dynamics of message passing on dense graphs, with applications to compressed sensing*, IEEE Trans. on Inform. Theory **57** (2011), 764–785.
- [BM12] ———, *The LASSO risk for gaussian matrices*, IEEE Trans. on Inform. Theory **58** (2012), 1997–2017.
- [BS06] Jinho Baik and Jack W Silverstein, *Eigenvalues of large sample covariance matrices of spiked population models*, Journal of Multivariate Analysis **97** (2006), no. 6, 1382–1408.
- [BYS01] Glen D Brown, Satoshi Yamada, and Terrence J Sejnowski, *Independent component analysis at the neural cocktail party*, Trends in neurosciences **24** (2001), no. 1, 54–63.

- [CDMF09] Mireille Capitaine, Catherine Donati-Martin, and Delphine Féral, *The largest eigenvalues of finite rank deformation of large wigner matrices: convergence and nonuniversality of the fluctuations*, The Annals of Probability **37** (2009), no. 1, 1–47.
- [CDMF⁺11] Mireille Capitaine, Catherine Donati-Martin, Delphine Féral, Maxime Février, et al., *Free convolution with a semicircular distribution and eigenvalues of spiked deformations of wigner matrices*, Electron. J. Probab **16** (2011), no. 64, 1750–1792.
- [CDMF12] M Capitaine, C Donati-Martin, and D Féral, *Central limit theorems for eigenvalues of deformations of wigner matrices*, Annales de l’Institut Henri Poincaré, Probabilités et Statistiques, vol. 48, Institut Henri Poincaré, 2012, pp. 107–133.
- [Cha05] S. Chatterjee, *An error bound in the sudakov-ferniqe inequality*, <http://arxiv.org/pdf/math/0510424v1.pdf>, 2005.
- [DJ94] D. L. Donoho and I. M. Johnstone, *Neo-classical minimax problems, thresholding, and adaptation*, Bernoulli (1994), 39–62.
- [dKP02] Etienne de Klerk and Dmitrii V Pasechnik, *Approximation of the stability number of a graph via copositive programming*, SIAM Journal on Optimization **12** (2002), no. 4, 875–892.
- [DM13a] Yash Deshpande and Andrea Montanari, *Finding hidden cliques of size $\sqrt{N/e}$ in nearly linear time*, Foundations of Computational Mathematics (2013), 1–60.
- [DM13b] David Donoho and Andrea Montanari, *High dimensional robust M-estimation: Asymptotic variance via approximate message passing*, arXiv preprint arXiv:1310.7320 (2013).
- [DMM09] D. L. Donoho, A. Maleki, and A. Montanari, *Message Passing Algorithms for Compressed Sensing*, Proceedings of the National Academy of Sciences **106** (2009), 18914–18919.
- [DMM10] ———, *Message Passing Algorithms for Compressed Sensing: I. Motivation and Construction*, Proceedings of IEEE Inform. Theory Workshop (Cairo), 2010.
- [DS03] David L Donoho and Victoria Stodden, *When does non-negative matrix factorization give a correct decomposition into parts?*, NIPS, vol. 16, 2003.
- [FP09] Delphine Féral and Sandrine Péché, *The largest eigenvalues of sample covariance matrices for a spiked population: diagonal case*, Journal of Mathematical Physics **50** (2009), 073302.
- [GB10] M. Grant and S. Boyd, *CVX: Matlab software for disciplined convex programming, version 1.21*, <http://cvxr.com/cvx>, May 2010.
- [JL04] Iain M Johnstone and Arthur Yu Lu, *Sparse principal components analysis*, Unpublished manuscript (2004).
- [JL09] ———, *On consistency and sparsity for principal components analysis in high dimensions*, Journal of the American Statistical Association **104** (2009), no. 486.

- [JM13] Adel Javanmard and Andrea Montanari, *State evolution for general approximate message passing algorithms, with applications to spatial coupling*, Information and Inference **2** (2013), no. 2, 115–144.
- [JNRS10] Michel Journée, Yurii Nesterov, Peter Richtárik, and Rodolphe Sepulchre, *Generalized power method for sparse principal component analysis*, The Journal of Machine Learning Research **11** (2010), 517–553.
- [Joh11] Iain M. Johnstone, *Gaussian estimation: Sequence and wavelet models*, 2011, Draft version, December 27, 2011, available at <http://statweb.stanford.edu/~imj/GE12-27-11.pdf>.
- [KBCG03] Yuval Kluger, Ronen Basri, Joseph T Chang, and Mark Gerstein, *Spectral biclustering of microarray data: coclustering genes and conditions*, Genome research **13** (2003), no. 4, 703–716.
- [KKM⁺14] Yoshiyuki Kabashima, Florent Krzakala, Marc Mézard, Ayaka Sakata, and Lenka Zdeborová, *Phase transitions and sample complexity in bayes-optimal matrix factorization*, arXiv preprint arXiv:1402.1298 (2014).
- [KM11] S. Korada and A. Montanari, *Applications of Lindeberg Principle in Communications and Statistical Learning*, IEEE Trans. on Inform. Theory **57** (2011), 2440–2450.
- [KMZ13] Florent Krzakala, Marc Mézard, and Lenka Zdeborová, *Phase diagram and approximate message passing for blind calibration and dictionary learning*, arXiv preprint arXiv:1301.5898 (2013).
- [Kne52] Hellmuth Kneser, *Sur un théoreme fondamental de la théorie des jeux*, Comptes Rendus de l’Academie des Sciences **234** (1952), no. 25, 2418–2420.
- [KY13] Antti Knowles and Jun Yin, *The isotropic semicircle law and deformation of wigner matrices*, Communications on Pure and Applied Mathematics (2013).
- [Lew98] Michael S Lewicki, *A review of methods for spike sorting: the detection and classification of neural action potentials*, Network: Computation in Neural Systems **9** (1998), no. 4, R53–R78.
- [LO02] Laura Lazzeroni and Art Owen, *Plaid models for gene expression data*, Statistica sinica **12** (2002), no. 1, 61–86.
- [LS99] Daniel D Lee and H Sebastian Seung, *Learning the parts of objects by non-negative matrix factorization*, Nature **401** (1999), no. 6755, 788–791.
- [LU09] Neil D Lawrence and Raquel Urtasun, *Non-linear matrix factorization with gaussian processes*, Proceedings of the 26th International Conference on Machine Learning, 2009, pp. 601–608.
- [Ma13] Zongming Ma, *Sparse principal component analysis and iterative thresholding*, The Annals of Statistics **41** (2013), no. 2, 772–801.
- [MM09] M. Mézard and A. Montanari, *Information, Physics and Computation*, Oxford, 2009.

- [MO04] Sara C Madeira and Arlindo L Oliveira, *Biclustering algorithms for biological data analysis: a survey*, Computational Biology and Bioinformatics, IEEE/ACM Transactions on **1** (2004), no. 1, 24–45.
- [Mon12] A. Montanari, *Graphical Models Concepts in Compressed Sensing*, Compressed Sensing: Theory and Applications (Y.C. Eldar and G. Kutyniok, eds.), Cambridge University Press, 2012.
- [MPV87] M. Mézard, G. Parisi, and M. A. Virasoro, *Spin glass theory and beyond*, World Scientific, 1987.
- [Paa97] Pentti Paatero, *Least squares formulation of robust non-negative factor analysis*, Chemometrics and intelligent laboratory systems **37** (1997), no. 1, 23–35.
- [Pau07] Debashis Paul, *Asymptotics of sample eigenstructure for a large dimensional spiked covariance model*, Statistica Sinica **17** (2007), no. 4, 1617.
- [Péc09] Sandrine Péché, *Universality results for the largest eigenvalues of some sample covariance matrix ensembles*, Probability Theory and Related Fields **143** (2009), no. 3-4, 481–516.
- [PMMP07] Alexey Pavlov, Valeri A Makarov, Ioulia Makarova, and Fivos Panetsos, *Sorting of neural spikes: when wavelet based methods outperform principal component analysis*, Natural Computing **6** (2007), no. 3, 269–281.
- [PRS13] Alessandro Pizzo, David Renfrew, and Alexander Soshnikov, *On finite rank deformations of wigner matrices*, Annales de l’Institut Henri Poincaré, Probabilités et Statistiques, vol. 49, Institut Henri Poincaré, 2013, pp. 64–94.
- [PSC13] Jason T Parker, Philip Schniter, and Volkan Cevher, *Bilinear generalized approximate message passing*, arXiv preprint arXiv:1310.2632 (2013).
- [PT94] Pentti Paatero and Unto Tapper, *Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values*, Environmetrics **5** (1994), no. 2, 111–126.
- [QNBS04] R Quian Quiroga, Zoltan Nadasdy, and Yoram Ben-Shaul, *Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering*, Neural computation **16** (2004), no. 8, 1661–1687.
- [QP09] Rodrigo Quian Quiroga and Stefano Panzeri, *Extracting information from neuronal populations: information theory and decoding approaches*, Nature Reviews Neuroscience **10** (2009), no. 3, 173–185.
- [Ran11] S. Rangan, *Generalized Approximate Message Passing for Estimation with Random Linear Mixing*, IEEE Intl. Symp. on Inform. Theory (St. Petersburg), August 2011.
- [RU08] T.J. Richardson and R. Urbanke, *Modern Coding Theory*, Cambridge University Press, Cambridge, 2008.

- [SC11] Philip Schniter and Volkan Cevher, *Approximate message passing for bilinear models*, Proc. Workshop Signal Process. Adaptive Sparse Struct. Repr.(SPARS), 2011, p. 68.
- [SN13] Xing Sun and Andrew B Nobel, *On the maximal size of large-average and anova-fit submatrices in a gaussian random matrix*, Bernoulli: official journal of the Bernoulli Society for Mathematical Statistics and Probability **19** (2013), no. 1, 275.
- [Ste72] C. Stein, *A bound for the error in the normal approximation to the distribution of a sum of dependent random variables*, Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, 1972.
- [SWPN09] Andrey A Shabalin, Victor J Weigman, Charles M Perou, and Andrew B Nobel, *Finding large average submatrices in high dimensional data*, The Annals of Applied Statistics (2009), 985–1012.
- [TAP77] DJ Thouless, PW Anderson, and RG Palmer, *Solution of solvable model of a spin glass*, Philosophical Magazine **35** (1977), no. 3, 593–601.
- [TSS02] Amos Tanay, Roded Sharan, and Ron Shamir, *Discovering statistically significant biclusters in gene expression data*, Bioinformatics **18** (2002), no. suppl 1, S136–S144.
- [Vit00] Richard A Vitale, *Some comparisons for gaussian processes*, Proceedings of the American Mathematical Society (2000), 3043–3046.
- [VS11] J. Vila and P. Schniter, *Expectation-maximization bernoulli-gaussian approximate message passing*, Proc. Asilomar Conf. on Signals, Systems, and Computers (Pacific Grove, CA), 2011.
- [VSM13] Jeremy Vila, Philip Schniter, and Joseph Meola, *Hyperspectral image unmixing via bilinear generalized approximate message passing*, SPIE Defense, Security, and Sensing, International Society for Optics and Photonics, 2013, pp. 87430Y–87430Y.
- [YZ13] Xiao-Tong Yuan and Tong Zhang, *Truncated power method for sparse eigenvalue problems*, The Journal of Machine Learning Research **14** (2013), no. 1, 899–925.
- [ZHT06] Hui Zou, Trevor Hastie, and Robert Tibshirani, *Sparse principal component analysis*, Journal of computational and graphical statistics **15** (2006), no. 2, 265–286.
- [ZWZ⁺04] Pu-Ming Zhang, Jin-Yong Wu, Yi Zhou, Pei-Ji Liang, and Jing-Qi Yuan, *Spike sorting based on automatic template reconstruction with a partial solution to the overlapping problem*, Journal of neuroscience methods **135** (2004), no. 1, 55–65.