A Proof of Theorem 1.2

We now provide a rigorous proof of the result that was sketched in Section 2. This proof is essentially identical to [Deshpande et al., 2015, Lemma 4.4]. Recall the iterates (30) given by $\boldsymbol{s}^0 = \boldsymbol{x}^0 - \boldsymbol{\mu}^{\mathbf{0}} \odot \boldsymbol{x}^\star$ and

$$\boldsymbol{s}^{t+1} = \left(\frac{1}{\sqrt{N}\sqrt{\boldsymbol{\Delta}}} \odot \boldsymbol{A} \right) f_t \left(\boldsymbol{s}^t + \boldsymbol{\mu}^{\boldsymbol{t}} \odot \boldsymbol{x}^\star \right) - \mathbf{b}_t \odot f_{t-1} \left(\boldsymbol{s}^{t-1} + \boldsymbol{\mu}^{\boldsymbol{t-1}} \odot \boldsymbol{x}^\star \right), \tag{43}$$

where $\boldsymbol{\mu}^t = (\mu_i^t)_{i \le N}$ is given by the recursion

$$\mu_i^{t+1} = \mu_{g(i)}^{t+1} = \sum_{a \in [q]} \frac{c_a}{\tilde{\boldsymbol{\Delta}}_{ag(i)}} \mathbb{E}_{x_0^\star, Z} [x_0^\star f_t^a \left(\mu_a^t x_0^\star + \sigma_a^t Z \right)]$$

and $\boldsymbol{x}^\star = (x_i^\star)_{i \in [N]}$ is a vector with independent coordinates distributed according to $\mathbb{P}_0$. By Lemma 2.4, for each $a \in [q]$, and any pseudo-Lipschitz function $\phi : \mathbb{R} \to \mathbb{R} \mapsto \mathbb{R}$ we have that almost surely

$$\lim_{N \to \infty} \frac{1}{|C_a^N|} \sum_{i \in C_a^N} \phi(s_i^t, x_i^\star) = \mathbb{E}_{x_0^\star, Z} \phi(\sigma_a^t Z, x_0^\star) \tag{44}$$

where

$$(\sigma_b^{t+1})^2 := \sum_{a=1}^{q} \frac{c_a}{\tilde{\boldsymbol{\Delta}}_{ab}} \mathbb{E}_Z \left[(f_t^b(Z_b^t))^2 \right].$$

as was defined in (18). For any pseudo-Lipschitz function $\psi : \mathbb{R} \to \mathbb{R}$, we have $\phi(x, y) = \phi(x - \mu_a^t y)$ is also pseudo-Lipschitz, so (44) implies that

$$\lim_{N \to \infty} \frac{1}{|C_a^N|} \sum_{i \in C_a^N} \psi(s_i^t + \mu_a^t x_i^\star) = \mathbb{E}_{x_0^\star, Z} \psi(\sigma_a^t Z + \mu_a^t x_0^\star) \tag{45}$$

almost surely.

Now let $\boldsymbol{x}^t$ be the iterates from the spiked AMP iteration for the inhomogeneous Wigner matrix (24) we derived in (26)

$$\boldsymbol{x}^{t+1} = \left(\frac{1}{N \boldsymbol{\Delta}} \odot \boldsymbol{x}^\star (\boldsymbol{x}^\star)^{\boldsymbol{T}} \right) \hat{\boldsymbol{x}}^t + \left(\frac{1}{\sqrt{N}\sqrt{\boldsymbol{\Delta}}} \odot \boldsymbol{A} \right) \hat{\boldsymbol{x}}^t - \mathbf{b}_t \odot \hat{\boldsymbol{x}}^{t-1}. \tag{46}$$

It now suffices to show that for fixed t and all $a \in [q]$ that

$$\lim_{N \to \infty} \frac{1}{|C_a^N|} \sum_{i \in C_a^N} (\psi(s_i^t + \mu_a^t x_i^\star) - \psi(x_i^t)) = 0 \tag{47}$$

almost surely. This will imply that $s_i^t + \mu_a^t x_i^\star$ and x_i^t have the same asymptotic distribution which finish the proof of Theorem 1.2 by (45).

We now prove (47). Since ψ is L-pseudo-Lipschitz we have

$$\begin{aligned} |\psi(s_i^t + \mu_a^t x_i^\star) - \psi(x_i^t)| &\le L(1 + |s_i^t + \mu_a^t x_i^\star| + |x_i^t|)|s_i^t + \mu_a^t x_i^\star - x_i^t| \\ &\le 2L|s_i^t + \mu_a^t x_i^\star - x_i^t|(1 + |s_i^t + \mu_a^t x_i^\star| + |s_i^t + \mu_a^t x_i^\star - x_i^t|). \end{aligned}$$

The Cauchy–Schwarz inequality implies that

$$\begin{aligned} &\left| \frac{1}{|C_a^N|} \sum_{i \in C_a^N} (\psi(s_i^t + \mu_a^t x_i^\star) - \psi(x_i^t)) \right| \\ &\le \frac{2L}{C_a^N} (\sqrt{C_a^N} \|\boldsymbol{s}_a^t + \mu_a^t \boldsymbol{x}_a^\star - \boldsymbol{x}_a^t\|_2 + \|\boldsymbol{s}_a^t + \mu_a^t \boldsymbol{x}_a^\star\|_2 \|\boldsymbol{s}_a^t + \mu_a^t \boldsymbol{x}_a^\star - \boldsymbol{x}_a^t\|_2 + \|\boldsymbol{s}_a^t + \mu_a^t \boldsymbol{x}_a^\star - \boldsymbol{x}_a^t\|_2^2) \end{aligned}$$

where $\boldsymbol{s}_a^t = (s_i^t)_{i \in C_a^N} \in \mathbb{R}^{|C_a^N|}$, $\boldsymbol{x}_a^t = (x_i^t)_{i \in C_a^N} \in \mathbb{R}^{|C_a^N|}$. Therefore, to prove (47) it suffices to prove that for all $t \ge 0$,

$$\lim_{N \to \infty} \frac{1}{|C_a^N|} \|\boldsymbol{s}_a^t + \mu_a^t \boldsymbol{x}_a^\star - \boldsymbol{x}_a^t\|_2^2 \to 0 \tag{48}$$

$$\limsup_{N \to \infty} \frac{1}{|C_a^N|} \|\boldsymbol{s}_a^t + \mu_a^t \boldsymbol{x}_a^\star\|_2^2 \to 0 \tag{49}$$

Clearly, if we initialize $\boldsymbol{x}^0$, $\boldsymbol{s}^0$ at zero then (48) and (49) are satisfied by our state evolution equations (5). Notice that (49) follows directly from (45) applied to the square function. We use here that we assumed that the second moment of $x^\star$ is finite.

We now focus on proving (48) through strong induction. By definition of the iterates (43) and (46),

$$
\begin{aligned}
&(\boldsymbol{s}_a^t + \mu_a^t \boldsymbol{x}_a^\star - \boldsymbol{x}_a^t) \\
&= \Bigg[\left(\frac{1}{\sqrt{N}\sqrt{\boldsymbol{\Delta}}} \odot \boldsymbol{A} \right) f_{t-1}\left(\boldsymbol{s}^{t-1} + \boldsymbol{\mu^{t-1}} \odot \boldsymbol{x}^\star \right) - \left(\frac{1}{\sqrt{N}\sqrt{\boldsymbol{\Delta}}} \odot \boldsymbol{A} \right) f_{t-1}(\boldsymbol{x}^{t-1}) \\
&\quad + \mu^t \odot \boldsymbol{x}^\star - \left(\frac{1}{N\boldsymbol{\Delta}} \odot \boldsymbol{x^\star (x^\star)^T} \right) f_{t-1}(\boldsymbol{x}^{t-1}) \\
&\quad + \mathbf{b}_{t-1}^x \odot f_{t-2}(\boldsymbol{x}^{t-2}) - \mathbf{b}_{t-1}^s \odot f_{t-2}\left(\boldsymbol{s}^{t-2} + \boldsymbol{\mu^{t-2}} \odot \boldsymbol{x}^\star \right) \Bigg]_{i \in C_a^N}
\end{aligned}
$$

where $[\cdot]_i$ corresponds to the ith row of a vector and $\mathbf{b}_{t-1}^x$ and $\mathbf{b}_t^s$ are the Onsager terms defined in (14) with respect to $\boldsymbol{x}^{t-1}$ and $\boldsymbol{s}^{t-1}$ respectively. The Cauchy–Schwarz inequality and Jensen's inequality imply that there exists some universal constant C such that

$$
\begin{aligned}
&\frac{1}{|C_a^N|} \| \boldsymbol{s}_a^t + \mu_a^t \boldsymbol{x}_a^\star - \boldsymbol{x}_a^t \|_2^2 \\
&\leq \frac{C}{|C_a^N|} \sum_{i \in C_a^N} \frac{1}{N} \left\| \left[\frac{1}{\sqrt{\boldsymbol{\Delta}}} \odot \boldsymbol{A} \right]_i \right\|_2^2 \| [f_{t-1}\left(\boldsymbol{s}^{t-1} + \boldsymbol{\mu^{t-1}} \odot \boldsymbol{x}^\star \right) - f_{t-1}(\boldsymbol{x}^{t-1})]_i \|_2^2 \\
&\quad + \frac{C}{|C_a^N|} \sum_{i \in C_a^N} \left(\mu_a^t - \left[\frac{1}{N\boldsymbol{\Delta}} (f_{t-1}(\boldsymbol{x}^{t-1}) \odot \boldsymbol{x}^\star) \right]_i \right)^2 (x_i^\star)^2 \\
&\quad + \frac{C}{|C_a^N|} \sum_{i \in C_a^N} ([\mathrm{b}_{t-1}^x]_i - [\mathrm{b}_{t-1}^s]_i)^2 [f_{t-2}(\boldsymbol{s}^{t-2} + \boldsymbol{\mu^{t-2}} \odot \boldsymbol{x}^\star)]_i^2 \\
&\quad + \frac{C}{|C_a^N|} \sum_{i \in C_a^N} [\mathbf{b}_{t-1}^x]_i^2 ([f_{t-2}(\boldsymbol{x}^{t-2})]_i - [f_{t-2}(\boldsymbol{s}^{t-2} + \boldsymbol{\mu^{t-2}} \odot \boldsymbol{x}^\star)]_i)^2
\end{aligned}
$$

We now control each term separately.

1. To control the first term, notice that the matrix $\frac{1}{N}\left[\frac{1}{\sqrt{\boldsymbol{\Delta}}} \odot \boldsymbol{A} \right]$ has iid entries within blocks and the sizes of the blocks diverge with the dimension, so we can control the sums of the squares of within each block using standard operator norm bounds Anderson et al. [2010]. The first term vanishes in the limit because f is pseudo-Lipschitz so we can apply the induction hypothesis bound which controls (48) at time $t-1$.
2. To control the second term, notice that for $i \in C_a^N$ by Lemma 2.4 applied to the pseudo-Lipschitz function $y f_{t-1}(x)$ that
$$\left[\frac{1}{N\boldsymbol{\Delta}} (f_{t-1}(\boldsymbol{x}^{t-1}) \odot \boldsymbol{x}^\star) \right]_i \to \mu_a$$
almost surely. This implies that the average of such terms vanishes since we assumed that the second moment $\mathbb{E}[x_0^\star]^2$ is finite.
3. To control the third and fourth terms, we can expand the definition of the Onsager terms and use the assumption that f' is pseudo-Lipschitz and almost surely bounded. Both terms vanish because our strong induction hypothesis gives us control of (48) at time $t-2$.

Since all terms vanish in the limit, we have proven (48) for all $a \in [q]$, which finishes the proof of statement (47) and the proof of Theorem 1.2.

B Comparison with a naive PCA spectral method

In this appendix, we wish to show how the spectral method we propose differs, in practice, from a naive PCA. We provide an example of the spectrums of $\boldsymbol{Y}$ and $\tilde{\boldsymbol{Y}}$ before and after the transition at

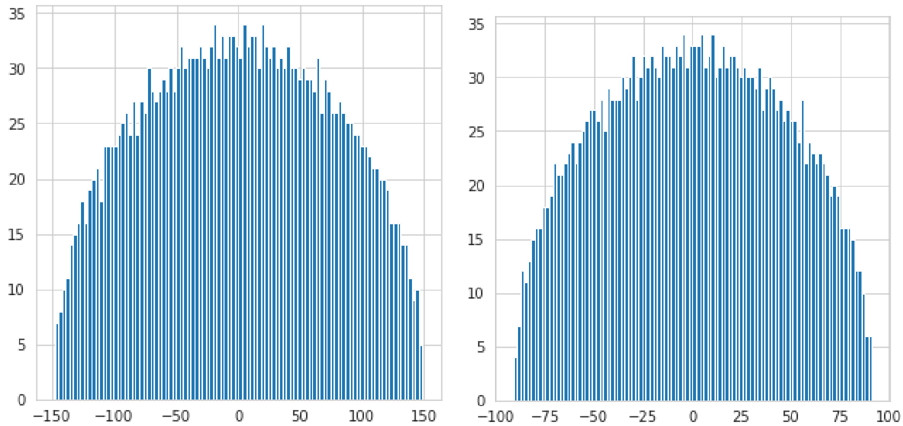

Figure 3: Illustration of the spectrum of $Y \in \mathbb{R}^{2500\times 2500}$ evaluated at noise profiles with snr $\lambda(\boldsymbol{\Delta}) = 0.7$ (left, before the transition) and on the left and 1.8 on the right (after the transition). There is no outlying eigenvalue in contrast to the transformed matrix: the transition for a naive spectral method is sub-optimal.

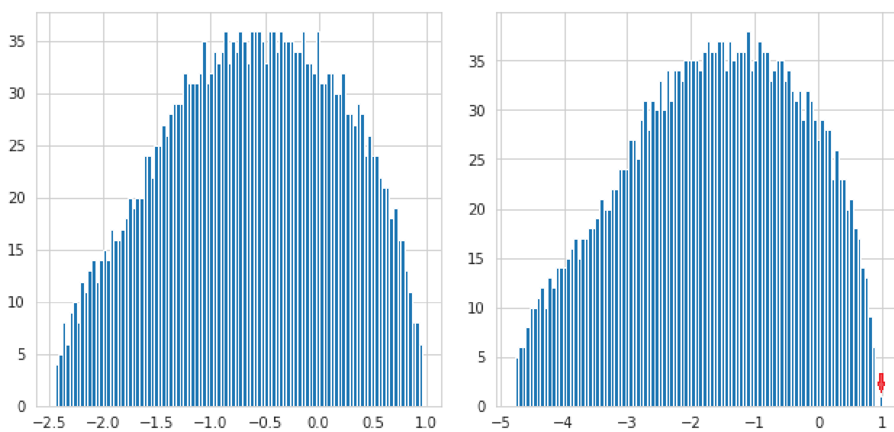

Figure 4: Illustration of the spectrum of $\tilde{Y} \in \mathbb{R}^{2500\times 2500}$ evaluated at noise profiles with snr $\lambda(\boldsymbol{\Delta}) = 0.7$ (left, before the transition) and on the left and 1.8 on the right (after the transition), with the outlying eigenvector correlated with the spike arises at eigenvalue one. This is at variance with the results of the naive method in Fig.3

$\mathrm{SNR}(\Delta) = 1$. In Figure 3 there is no clear separation of the extremal eigenvalue of $\boldsymbol{Y}$ from the bulk around this transition. This is in contrast to Figure 4 where there is an extremal eigenvalue of $\tilde{\boldsymbol{Y}}$ appearing at value one.