We will present the proofs of Lemma 4.1, Lemma 4.3 and Theorem 4.2 here. For completeness, first recall Lyapunov’s theorem in probability theory.

Let $Y_n = \sum_{i=1}^{k_n} Y_{ni}$, for $n \in \mathbb{N}$, be a triangular array of random variables such that for each $n$, the random variables $Y_{ni}$, for $1 \leq i \leq k_n$ are independent with $E(Y_{ni}) = 0$ and $E(Y_{ni}^2) = \sigma_{ni}^2$, where $E(X)$ stands for the expected value of the random variable $X$. Let $s_n^2 = \sum_{i=1}^{k_n} \sigma_{ni}^2$. We use the notation $X \sim Y$ when the two random variables $X$ and $Y$ have the same distribution (for large $n$ if either or both of them depend on $n$).

Theorem (Lyapunov) [2, page 368] If for some $\delta > 0$, we have
\[
\frac{1}{s_n^{2+\delta}} E(|Y_n|^{2+\delta}) \to 0 \quad \text{as } n \to \infty
\]
then $\frac{1}{s_n} Y_n \xrightarrow{d} N(0, 1)$ as $n \to \infty$ where $\xrightarrow{d}$ denotes convergence in distribution, and we denote by $N(a, \sigma^2)$ the normal distribution with mean $a$ and variance $\sigma^2$. Thus, for large $n$ we have $Y_n \sim N(0, s_n^2)$. □

Lemma 4.1 Let $X$ be a random variable on $\mathbb{R}$ such that its probability distribution $F(x) = \Pr(X \leq x)$ is differentiable with density $F'(x) = f(x)$. If $g : \mathbb{R} \to \mathbb{R}$ is a bounded measurable function and $X_k$ ($k \geq 1$) is a sequence of independent and identically distributed random variables with distribution $X$, then
\[
\frac{1}{N} \sum_{i=1}^{N} g(X_i) \xrightarrow{a.s.} E_g(X) = \int_{-\infty}^{\infty} g(x)f(x)dx, \quad (1)
\]
and for all $\epsilon > 0$ and $t > 1$, we have:
\[
\Pr \left( \sup_{k \geq N} \left( \frac{1}{k} \sum_{i=1}^{k} (g(X_i) - kE(g(X))) \right) \geq \epsilon \right) = o(1/N^{t-1}) \quad (2)
\]

Proof Since $g$ is bounded, $E_g(X) = \int_{-\infty}^{\infty} g(x)f(x)dx$ is absolutely convergent and thus the expected value $E_g(X)$ is well-defined and $|E_g(X)| < \infty$. Equation (1) then follows from the Strong Law of Large Numbers [2 page 80] applied to the random variables $g(X_i)$, for $i \geq 1$, which are independent and identically distributed as $g(X)$ with expectation $E_g(X)$. We also have $E|g(X)|^t = \int_{-\infty}^{\infty} |g(x)|^t f(x)dx < \infty$ for all $t > 1$ and thus the convergence rate of the Strong Law of Large Numbers implies Equation (2), a consequence of Theorem 3 and the lemma in [1] pages 112 and 113. □
Assume \( p/N = \alpha > 0 \) with \( d_1 \ll p_0 \) and \( d_\mu = 1 \) for \( 1 < \mu \leq p_0 \). Consider the overlaps

\[
m_\nu = \frac{1}{N} \sum_{i=1}^{N} \xi_\nu^i \langle S_i \rangle
\]

and the mean field equations:

\[
m_\nu = \frac{1}{N} \sum_{i=1}^{N} \xi_\nu^i \tanh \left( \beta \sum_{\mu=1}^{p} d_\mu \xi_\mu^i m_\mu \right)
\]

**Theorem 4.2** There is a solution to the mean field equations \([4]\) for retrieving \( \xi^1 \) with independent random variables \( m_\nu \) (for \( 1 \leq \nu \leq p_0 \)), where \( m_1 \sim \mathcal{N}(m, s/N) \) and \( m_\nu \sim \mathcal{N}(0, r/N) \) (for \( \nu \neq 1 \)), if the real numbers \( m, s \) and \( r \) satisfy the four simultaneous equations:

\[
\begin{align*}
(i) \quad m &= \int_{-\infty}^{\infty} \frac{dz}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \tanh(\beta(d_1 m + \sqrt{\alpha r} z)) \\
(ii) \quad s &= q - m^2 \\
(iii) \quad q &= \int_{-\infty}^{\infty} \frac{dz}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \tanh^2(\beta(d_1 m + \sqrt{\alpha r} z)) \\
(iv) \quad r &= \frac{q}{(1-\beta(1-q))^2}
\end{align*}
\]

In the proof of this theorem, as given below, we seek a solution of the mean field equations assuming we have independent random variables \( m_\nu \) (for \( 1 \leq \nu \leq p_0 \)) such that for large \( N \) and \( p \) with \( p/N = \alpha \), we have \( m_1 \sim \mathcal{N}(m, s/N) \) and \( m_\nu \sim \mathcal{N}(0, r/N) \) (\( \nu \neq 1 \)), and then find conditions in terms of \( m, s \) and \( r \) to ensure that such a solution exists. Since by our assumption about the distribution of the overlaps \( m_\mu \), the standard deviation of each overlap is \( O(1/\sqrt{N}) \), we ignore terms of \( O(1/N) \) and more generally terms of \( o(1/\sqrt{N}) \) compared to terms of \( O(1/\sqrt{N}) \) in the proof including in the lemma below.

**Lemma 4.3** If \( m_\nu \sim \mathcal{N}(0, r/N) \) (for \( \nu \neq 1 \)), then we have the equivalence of distributions:

\[
\sum_{\mu \neq 1, \nu} \xi_\nu^i \xi_\mu^i m_\mu \sim \mathcal{N}(0, \alpha r) \approx \sum_{\mu \neq 1} \xi_\nu^i \xi_\mu^i m_\mu.
\]

**Proof** Recall that the sum \( \sum_{t=1}^{k} X_t \) of \( k \) independent random variables such that \( X_t \) has a normal distribution with mean \( a_t \) and variance \( \sigma_t^2 \) (for \( 1 \leq t \leq k \)) is itself normally distributed with mean \( \sum_{t=1}^{k} a_t \) and variance \( \sum_{t=1}^{k} \sigma_t^2 \). Consider the first equivalence. From \( -1 \leq \langle S_i \rangle \mid \leq 1 \), for \( 1 \leq i \leq N \), and Equation \([\nu]\), it follows that

\[
\mathbb{E}(m_\mu \xi_\nu^i) = \mathbb{E} \left( \frac{1}{N} \sum_{i=1}^{N} \xi_\nu^i \langle S_i \rangle \xi_\mu^j \right) \leq \mathbb{E} \left( \frac{1}{N} \sum_{i=1}^{N} \xi_\nu^i \xi_\mu^j \right) = \frac{1}{N}
\]

Similarly, \( \mathbb{E}(m_\mu \xi_\nu^i) \geq -1/N \), and thus \( \mathbb{E}(m_\mu \xi_\nu^i) = O(1/N) \). Therefore, for \( \mu \neq 1, \nu \), the three random variables \( \xi_\nu^i, \xi_\mu^i \) and \( m_\mu \) can be considered independent and it follows that the distribution of each product on the left hand side of the first equivalence is given by \( \mathcal{N}(0, r/N) \). Summing up the approximately \( p \) independent normal distributions \( \mu \neq 1, \nu \), we obtain the first equivalence. The second equivalence is proved in a similar way. \( \square \)

**Proof of Theorem 4.2** First consider Equation \([\nu]\) for \( \nu = 1 \), which, by separating the contributions of \( \mu = 1 \) and \( \mu \neq 1 \) on the right hand side, we write as

\[
m_1 = Y_N := \frac{1}{N} \sum_{i=1}^{N} \xi_1^i \tanh(\beta(d_1 m_1 \xi_1^i + \sum_{\mu \neq 1} \xi_\mu^i m_\mu)).
\]

Multiplying the odd function \( \tanh \) and its argument by \( \xi_1^i \), we obtain:

\[
\begin{align*}
Y_N &= \frac{1}{N} \sum_{i=1}^{N} \xi_1^i \xi_1^i \tanh(\beta(d_1 m_1 \xi_1^i + \sum_{\mu \neq 1} \xi_\mu^i m_\mu)) \\
&= \frac{1}{N} \sum_{i=1}^{N} \tanh(\beta(d_1 m_1 + \sum_{\mu \neq 1} \xi_\mu^i m_\mu)) \\
&\xrightarrow{\text{as}} \int_{-\infty}^{\infty} \frac{dz}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \tanh(\beta(m d_1 + \sqrt{\alpha r} z)),
\end{align*}
\]
implies convergence in distribution, it follows that as $N \to \infty$ using $g \sim N$ distribution $X$. Therefore, the random variables

$$m \sim N(m, s/N)$$

Since, by assumption, $m_1$ has distribution $N(m, s/N)$ and is independent of $\sum_{\mu \neq 1} \xi^\mu_i \xi^\mu_1 m_\mu$, it follows that $d_1 m_1 + \sum_{\mu \neq 1} \xi^\mu_i \xi^\mu_1 m_\mu$ is the sum of two normal distributions and thus has itself normal distribution $N(d_1 m_1 + \sum_{\mu \neq 1} \xi^\mu_i \xi^\mu_1 m_\mu, \sigma^2)$ where the last step is justified as follows. By Lemma 4.3

$$\sum_{\mu \neq 1} \xi^\mu_1 m_\mu \sim N(0, \alpha r)$$

(8)

Since, by assumption, $m_1$ has distribution $N(m, s/N)$ and is independent of $\sum_{\mu \neq 1} \xi^\mu_i \xi^\mu_1 m_\mu$, it follows that $d_1 m_1 + \sum_{\mu \neq 1} \xi^\mu_i \xi^\mu_1 m_\mu$ is the sum of two normal distributions and thus has itself normal distribution $N(d_1 m_1 + \sum_{\mu \neq 1} \xi^\mu_i \xi^\mu_1 m_\mu, \sigma^2)$.

$$X_i := d_1 m_1 + \sum_{\mu \neq 1} \xi^\mu_i \xi^\mu_1 m_\mu \sim N(d_1 m_1 + \sum_{\mu \neq 1} \xi^\mu_i \xi^\mu_1 m_\mu, \sigma^2)$$

(9)

Therefore, the random variables $X_i$, for $i \geq 1$, are independent and identically distributed with distribution $\sim N(d_1 m_1 + \sum_{\mu \neq 1} \xi^\mu_i \xi^\mu_1 m_\mu, \sigma^2)$ and the last step in Equation (7) then follows by applying Lemma 4.1 using $g(x) = \tanh(\beta x)$, which is a bounded continuous function. Since almost sure convergence implies convergence in distribution, it follows that as $N \to \infty$,

$$Y_N \overset{d}{\to} \int_{-\infty}^{\infty} \frac{dz}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \tanh(\beta \sqrt{2\pi} z),$$

(10)

where the latter is the degenerate (point) distribution with the integral on the right hand side as its value. On the other hand, by the assumption about $m_1$, we have

$$m_1 \sim N(m, s/N) \overset{d}{\to} m_1$$

(11)

as $N \to \infty$. Therefore, from Equations (6), (11) and (10), we can now obtain

$$m = \int_{-\infty}^{\infty} \frac{dz}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \tanh(\beta \sqrt{2\pi} z),$$

(12)

which gives Equation (5(iii)).

Next, write $Y_N = \sum_{i=1}^{N} Y_{N_i}$ with $Y_{N_i} = \frac{1}{N} \tanh(\beta X_i)$. We have a triangular array of random variables with $E(Y_{N_i}) = m/N$. By Equation (9), the equality in Equation (11), using $g(x) = \frac{1}{\sqrt{2\pi}} \tanh(\beta x)$ and $f$ as the Gaussian distribution $N(d_1 m_1, \sigma^2)$, and Equation (12), furthermore,

$$E(Y_{N_i}^2) = \frac{q}{N^2},$$

(13)

by Equation (6), where $q$ is given in Equation (5(iii)). This gives

$$\sigma_{N_i}^2 := E(Y_{N_i}^2) - (E(Y_{N_i}))^2 = (q - m^2)/N^2, \quad \sigma_{N_i}^2 := \sum_{i=1}^{N} \sigma_{N_i}^2 = (q - m^2)/N.$$

Moreover, it is easy to see that $E(|Y_{N_i}|^3) \leq 1/N^3$ since $\tanh$ is bounded by 1. Thus,

$$\frac{1}{N} \sum_{i=1}^{N} E(|Y_{N_i}|^3) = O(1/N^{1/2})$$

(14)

and it follows that the Lyapunov condition holds for $\delta = 1$. Therefore, by Lyapunov’s theorem $Y_N - m/\sigma_N \sim N(0, 1)$, as $N \to \infty$, and thus $m_1 = Y_N \sim N(m, (q - m^2)/N)$, as $N \to \infty$. Since by assumption $m_1 \sim N(m, s/N)$, we obtain the value of $s = q - m^2$ as in Equation (5(iii)).

Now fix $\nu \neq 1$ in Equation (4), take a sample point $\omega \in \Omega$, separate the three terms for $\mu = 1, \mu = \nu$ and $\mu \neq 1, \nu$ on the right hand side of the equation, as before multiply $\tanh$ and its argument by $\xi_1^\nu$ and write the equation as $m_\nu(\omega) = h(m_\nu(\omega))$, where $h : \mathbb{R} \to \mathbb{R}$ with

$$h(x) = \frac{1}{N} \sum_{i=1}^{N} \xi^\nu_i(\omega) \xi^\nu_1(\omega) \tanh \beta \left( d_1 m_1(\omega) + \xi^\nu_i(\omega) \xi^\nu_1(\omega) x + \sum_{\mu \neq 1, \nu} \xi^\nu_i(\omega) \xi^\nu_1(\omega) m_\mu(\omega) \right)$$

(15)
By assumption $m_\nu$ is normally distributed with mean zero and standard deviation $\sqrt{r}/\sqrt{N}$. Therefore, in contrast to the case for $m_1$ treated earlier, here $m_\nu(\omega)$ is small and of order $O(1/\sqrt{N})$. Since $m_\nu$ appears in two terms on both sides of $m_\nu(\omega) = h(m_\nu(\omega))$, we need to collect together on one side of the equation the contributions of these two terms. To this end, we regard $m_\nu(\omega)$ as small compared with the term $m_1(\omega)d_1$ and the term $\sum_{\mu \neq 1, \nu} \xi^1_\mu(\omega)\xi^1_\nu(\omega)m_\mu(\omega)$ which are both of order $O(1)$, and we employ the Taylor expansion of $h$ near the origin $x = 0$:

$$h(m_\nu(\omega)) = \frac{1}{N} \sum_{i=1}^N \xi^\nu_i(\omega) \xi^1_i(\omega) \tanh(\beta(d_1m_1(\omega) + \sum_{\mu \neq 1, \nu} \xi^\mu_i(\omega)\xi^1_i(\omega)m_\mu(\omega))) + \frac{c}{2} \left( \sum_{i=1}^N (1 - \tanh^2(\beta(d_1m_1(\omega) + \sum_{\mu \neq 1, \nu} \xi^\mu_i(\omega)\xi^1_i(\omega)m_\mu(\omega))) \right) m_\nu(\omega) + c(m_\nu(\omega))^2$$

(16)

where $|c| \leq \beta^2$, which is obtained by using the Lagrange form of remainder $c(m_\nu(\omega))^2$ to estimate the second derivative $h''(0)$ and by noting that $|\tanh(x)| \leq 1$ for all $x \in \mathbb{R}$. Thus, the Taylor series remainder is of order $O(1/N)$, which we ignore compared to the standard deviation of $m_\nu$, namely $\sqrt{r}/\sqrt{N}$. By Lemma 4.1, the last summation in Equation (16), containing the bounded continuous function $\tanh^2$, converges almost surely to $q$ as $N \to \infty$. Moreover, by using $t = 3/2$ in the second part of Lemma 4.1, it follows that for large $N$, while retaining $m_\nu$ which is of order $1/\sqrt{N}$, we can replace the sum in the equation with $q$ by ignoring the error which, for any degree of certainty, is of order $o(1/\sqrt{N})$. Thus, by using $m_\nu(\omega) = h(m_\nu(\omega))$ from Equation (2), we now obtain the following reduced stochastic equation for $\nu \neq 1$:

$$(1 - \beta(1-q))m_\nu(\omega) = \frac{1}{N} \sum_{i=1}^N \xi^\nu_i(\omega)\xi^1_i(\omega) \tanh(\beta(d_1m_1(\omega) + \sum_{\mu \neq 1, \nu} \xi^\mu_i(\omega)\xi^1_i(\omega)m_\mu(\omega)))$$

(17)

Now we drop $\omega$ everywhere and let the right hand side of Equation (17) be written as $Z_N = \sum_{i=1}^N Z_{Ni}$ with $Z_{Ni} = \frac{1}{N} \xi^\nu_i\xi^1_i \tanh(\beta(X'_i))$, where $X'_i = d_1m_1 + \sum_{\mu \neq 1, \nu} \xi^\mu_i\xi^1_i m_\mu$. By Lemma 4.3 and Equation (2), $X'_i \sim X_i \sim \mathcal{N}(d_1m, r\alpha)$ and the three random variables $\xi^\nu_i$, $\xi^1_i$ and $X'_i$ are independent.

We again have an array $Z_{Ni}$ of random variables $1 \leq i \leq N$ for each $N$, and by the independence of the above three random variables we have: $E(Z_{Ni}) = 0$ and

$$E(Z^2_{Ni}) = \frac{1}{N} E(\xi^\nu_i)^2E(\xi^1_i)^2 = E(\tanh^2(\beta(X'_i))) = E(Y^2_{N,i}) = \frac{q}{N},$$

(18)

as in Equation (13). Therefore, $\sigma^2_{Z_N} = \sum_{i=1}^N E(Z^2_{Ni}) = q/N$. Moreover, it is easy to see that $E(|Z_{Ni}|^3) \leq 1/N$ since $|\tanh(x)|$ is bounded by 1 for all $x \in \mathbb{R}$. Thus,

$$\frac{1}{\sigma_N^2} \sum_{i=1}^N E(|Z_{Ni}|^3) = O(1/N^{1/2})$$

(19)

and it follows that the Lyapunov condition holds for $\delta = 1$. We conclude by Lyapunov’s theorem that $Z_N/\sigma_N \sim \mathcal{N}(0, 1)$ and thus $Z_N \sim \mathcal{N}(0, q/N)$. From this and Equation (17), we deduce that

$$m_\nu \sim \mathcal{N} \left( 0, \frac{q}{(1 - \beta(1-q))^2N} \right)$$

(20)

and obtain $r = q/(1 - \beta(1-q))^2$ in Equation (5)iv). This completes the proof of the theorem. □

References
