CHAPTER III

RESOLVENT AND SPECTRAL FUNCTIONS OF LARGE POOLED SAMPLE COVARIANCE MATRICES

The purpose of this chapter is to single out the leading parts of spectral functions of pooled sample covariance matrices which present the weighted sums of sample covariance matrices calculated over samples from different populations. We consider two populations $\mathfrak{S}_1$ and $\mathfrak{S}_2$ without assumptions on distributions having, in general, different true covariance matrices, and study relations between leading parts for spectral functions of true covariance matrices and for pooled sample covariance matrices under high dimension and large sample sizes. These relations, in particular, can be used for the improvement of the standard linear discriminant procedure when it is applied to a wide class of populations.

Problem Setting

Let $\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_n)$ be observation vectors from two populations $\mathfrak{S}_1$ and $\mathfrak{S}_2$. We restrict the populations with an only requirement that all four moments of all variables exist. For convenience let $\mathbf{E}\ \mathbf{x} = 0$ in the both populations. Define the parameters

$$M_\nu = \max_\nu \sup_{|\mathbf{e}|=1} \mathbf{E}\ (\mathbf{e}^T\mathbf{x})^4 \quad \text{for } \mathbf{x} \text{ in } \quad \mathfrak{S}_\nu, \quad \nu = 1, 2.$$

$$M = \max\ (M_1, M_2), \tag{1}$$

where (and in the following) non-random vectors $\mathbf{e}$ are of unit length (the absolute value of a vector means its length). For simplicity let $M > 0$. Denote $\Sigma_\nu = \text{cov}\ (\mathbf{x}, \mathbf{x})$ for $\mathbf{x}$ in population $\mathfrak{S}_\nu$, $\nu = 1, 2$.

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

Define

$$\gamma_\nu = \sup_{\|\Omega\|=1} \text{var } (\mathbf{x}^T \Omega \mathbf{x}/n)/M \quad \text{for } \mathbf{x} \text{ in } \mathfrak{S}_\nu, \quad \nu = 1,2,$$
$$\gamma = \max\ (\gamma_1, \gamma_2), \tag{2}$$

where Ω are non-random real symmetric positive semidefinite matrices of unit spectral norm (only the spectral norms of matrices are used). The values γ measure the variance of quadratic forms and restrict the dependence of variables (see Introduction).

Let $\mathfrak{X}_1$ and $\mathfrak{X}_2$ be two (independent) samples of size $N_1 > 1$ and $N_2 > 1$ from $\mathfrak{S}_1$ and $\mathfrak{S}_2$. Denote $N = N_1 + N_2$. Define

$$\bar{\mathbf{x}}_\nu = N^{-1} \sum_{\mathbf{x}_m \in \mathfrak{X}_\nu} \mathbf{x}_m, \quad S_\nu = N^{-1} \sum_{\mathbf{x}_m \in \mathfrak{X}_\nu} \mathbf{x}_m \mathbf{x}_m^T,$$
$$C_\nu = N^{-1} \sum_{\mathbf{x}_m \in \mathfrak{X}_\nu} (\mathbf{x}_m - \bar{\mathbf{x}}_\nu)(\mathbf{x}_m - \bar{\mathbf{x}}_\nu)^T,$$

where m runs over numbers of all vectors $\mathbf{x}_m$ from both samples $\mathfrak{X}_\nu$, $\nu = 1,2$. We consider pooled sample matrices of two forms

$$S = (N_1 S_1 + N_2 S_2)/N \quad \text{and} \quad C = (N_1 C_1 + N_2 C_2)/N,$$

the expectation matrix $\Sigma = (N_1 \Sigma_1 + N_2 \Sigma_2)/N$, and the resolvents

$$H_0 = H_0(t) = (I + tS)^{-1} \quad \text{and} \quad H = H(t) = (I + tC)^{-1}, \quad t \geq 0.$$

Note that $tH_0(t)$ and $tH(t)$ can be considered as regularized ridge estimators of the matrix Σ^{-1}.

We will be interested in functions

$$V_\nu = \mathbf{e}^T H_0 \bar{\mathbf{x}}_\nu, \quad \Phi_{\nu\mu} = \bar{\mathbf{x}}_\nu^T H_0 \bar{\mathbf{x}}_\mu, \quad U_\nu = \mathbf{e}^T H \bar{\mathbf{x}}_\nu,$$
$$\Psi_{\nu\mu} = \bar{\mathbf{x}}_\nu^T H \bar{\mathbf{x}}_\mu, \quad \nu, \mu = 1,2. \tag{3}$$

We also consider the functions

$$h_0(t) = \mathbf{E}\ n^{-1} \text{tr } H_0(t), \quad h(t) = \mathbf{E}\ n^{-1} \text{tr } H(t),$$
$$s_{0\nu} = s_{0\nu}(t) = 1 - t/N\ \mathbf{E} \text{ tr } (H_0(t) S_\nu),$$
$$s_\nu = s_\nu(t) = 1 - t/N\ \mathbf{E} \text{ tr } (H(t) C_\nu), \quad \nu = 1,2. \tag{4}$$

For brevity, denote $t_\nu = tN_\nu/N$, $\nu = 1,2$, and

$$y = n/N, \quad \tau = \sqrt{M} t, \quad \delta = \delta(t) = 2\tau^2 y^2 (\gamma + \tau^2/N). \tag{5}$$

Spectral Functions of Pooled Random Gram Matrices

The main tool of our proofs will be the method of alternating elimination of independent variables. For the convenience of notations, we enumerate sample vectors from both $\mathfrak{X}_1$ and $\mathfrak{X}_2$ in such a way that $\mathbf{x}_1 \in \mathfrak{X}_1$ and $\mathbf{x}_2 \in \mathfrak{X}_2$. Denote

$$S^\nu = S - N^{-1}\mathbf{x}_\nu \mathbf{x}_\nu^T, \quad H_0^\nu = (I + tS^\nu)^{-1},$$
$$\varphi_\nu = \mathbf{x}_\nu^T H_0^\nu \mathbf{x}_\nu / N, \quad \psi_\nu = \mathbf{x}_\nu^T H_0 \mathbf{x}_\nu / N, \quad \nu = 1, 2. \tag{6}$$

It is easy to verify the identities

$$H_0 = H_0^\nu - tH_0^\nu \mathbf{x}_\nu^T \mathbf{x}_\nu H_0 / N, \quad H_0 \mathbf{x}_\nu = (1 - t\psi_\nu) H_0^\nu \mathbf{x}_m,$$
$$(1 + t\varphi_\nu)(1 - t\psi_\nu) = 1, \quad m = 1, 2. \tag{7}$$

Let $\mathbf{e}$ be a non-random unit vector with n components. Denote

$$v_\nu = v_\nu(t) = \mathbf{e}^T H_0^\nu \mathbf{x}_\nu, \quad u_\nu = u_\nu(t) = \mathbf{e}^T H_0 \mathbf{x}_\nu, \quad \nu = 1, 2.$$

From (7) and (1) it is obvious that

$$u_\nu = (1 - t\psi_\nu) v_\nu, \quad 0 \le t\psi_\nu \le 1,$$
$$\mathbf{E}\ (1 - t\psi_\nu) = s_{0\nu}(t), \quad \mathbf{E}\ v_\nu^4 \le M, \quad \nu = 1, 2. \tag{8}$$

Theorem 3.1. *If $t \ge 0$, then*

$$\mathbf{E}\ H_0(t) = (I + t_1 s_{01}\Sigma_1 + t_2 s_{02}\Sigma_2)^{-1} + \Omega_0,$$
$$\mathrm{var}\ (\mathbf{e}^T H_0(t)\mathbf{e}) \le \tau^2/N. \tag{9}$$

Proof. We eliminate the vectors $\mathbf{x}_\nu$, $\nu = 1, 2$. By (7),

$$t_\nu H_0 \mathbf{x}_\nu \mathbf{x}_\nu^T = t_\nu (1 - t\psi_\nu) H_0^\nu \mathbf{x}_\nu \mathbf{x}_\nu^T, \quad \nu = 1, 2.$$

The expectation $\mathbf{E}\ H_0 \mathbf{x}_\nu \mathbf{x}_\nu^T = \mathbf{E}\ H_0 S_\nu$, $\nu = 1, 2$. Clearly, the sum $t_1 \mathbf{E}\ H_0 S_1 + t_2 \mathbf{E}\ H_0 S_2 = I - \mathbf{E}\ H_0$. On the right hand side we substitute $1 - t\psi_\nu = s_{0\nu} - \Delta_\nu$, where Δ_ν is a deviation of $t\psi_\nu$ from the expectation value, and notice that $\mathbf{E}\ H_0^\nu \mathbf{x}_\nu \mathbf{x}_\nu = \mathbf{E}\ H_0^\nu \Sigma_\nu$, $\nu = 1, 2$. It follows that

$$I - \mathbf{E}\ H_0 = t_1 s_{01} \mathbf{E}\ H_0^1 \Sigma_1 + t_2 s_{02} \mathbf{E}\ H_0^2 \Sigma_2 + \Omega_1 + \Omega_2, \tag{10}$$

where $\Omega_\nu = -\mathbf{E}\ t_\nu H_0^\nu \mathbf{x}_\nu \mathbf{x}_\nu^T \Delta_\nu$, $\nu = 1, 2$. Let us substitute the expressions for H_0^ν in terms of H_0 using the first equation from (7) with the transposed left and right hand sides, $\nu = 1, 2$. Equation (10) can be rewritten in the form

$$I = \mathbf{E}\ H_0 R + \Omega_1 + \Omega_2 + \tilde{\Omega}_1 + \tilde{\Omega}_2,$$

where $\tilde{\Omega}_\nu = t t_\nu s_{0\nu} N^{-1} \mathbf{E}\ H_0 \mathbf{x}_\nu \mathbf{x}_\nu^T H_0^\nu \Sigma_\nu$, $\nu = 1, 2$, and $R = I + t_1 s_{01} \Sigma_1 + t_2 s_{02} \Sigma_2$. We multiply both parts of this equation by R^{-1} from the right. It follows that $R^{-1} - \mathbf{E}\ H_0 = \Omega_0$, where $\Omega_0 = (\Omega_1 + \Omega_2 + \tilde{\Omega}_1 + \tilde{\Omega}_2) R^{-1}$. Let $\mathbf{e}$ be the eigenvector for the maximum eigenvalue of the symmetric matrix Ω_0. Using the Schwarz inequality, (8) and (1), we find

$$(\mathbf{e}^T \Omega_1 \mathbf{e})^2 \le t_1^2 \mathbf{E}\ \Delta_1^2 \mathbf{E}\ |(\mathbf{e}^T H_0^1 \mathbf{x}_1)^2 (\mathbf{x}_1^T \mathbf{e})^2| \le M t^2 \mathrm{var}\ (t\psi_1) \le \tau^2 \delta.$$

Similiarly, $(\mathbf{e}^T \Omega_1 \mathbf{e})^2 \le \tau^2 \delta$. Further, we have

$$|\mathbf{e}^T \tilde{\Omega}_1 \mathbf{e}| \le t t_1\ N^{-1} \mathbf{E}\ |(\mathbf{e}^T H_0 \mathbf{x}_1)(\mathbf{x}_1^T H_0^1 \Sigma_1 \mathbf{e})|.$$

Here, by (8), $|\mathbf{e}^T H_0 \mathbf{x}_1| \le |v_1|$, $\|\Sigma_1\| \le \sqrt{M}$ and by (1) the left hand side is not greater τ^2/N. Similiarly, $|\mathbf{e}^T \tilde{\Omega}_2 \mathbf{e}| \le \tau^2/N$. We obtain the statement of Theorem 3.1. □

THEOREM 3.2. *If* $t \ge 0$ *then:*

1^o $t\ (\mathbf{E}\ V_\nu)^2 \le \omega_{52}$, t var $V_\nu \le \omega_{20}$, $\nu = 1, 2$;

2^o $t_\nu \Phi_{\nu\nu} \le 1$, $t_\nu \mathbf{E}\ \Phi_{\nu\nu} = 1 - s_{0\nu} + o_\nu$, *where* $o_\nu^2 \le \omega_{52}$, var $(t_\nu \Phi_{\nu\nu}) \le \omega_{20}$, $\nu = 1, 2$;

3^o var $(\mathbf{e}^T H_0 \mathbf{e}) \le \tau^2$;

4^o $t_1 t_2 \mathbf{E}\ \Phi_{12} \le o$, *where* $o^2 \le \omega_{52}$, t^2var $\Phi_{12} \le \omega_{20}$;

5^o $s_{0\nu}(t) \ge (1 + \tau y)^{-1}$, $\nu = 1, 2$.

Proof. First, we notice that $S_1 = C_1 + \bar{\mathbf{x}}_1\bar{\mathbf{x}}_1^T$ and $S = A + t_1\bar{\mathbf{x}}_1\bar{\mathbf{x}}_1^T$, where A is a symmetric positive definite matrix. Denote $\mathbf{y} = A^{-1/2}\bar{\mathbf{x}}_1$. Then we can write:

$$t_1\Phi_{11} = t_1\mathbf{y}^T(I + t_1\mathbf{y}\mathbf{y}^T)^{-1}\mathbf{y} = t_1\mathbf{y}^2(1 + t_1\mathbf{y}^2)^{-1} \le 1.$$

Similiarly, $t_2\Phi_{22} \le 1$.

Now let $\nu = 1, 2$. Obviously, $\mathbf{E}\ v_\nu = 0$ and $\mathbf{E}\ V_\nu = \mathbf{E}\ u_\nu = \mathbf{E}\ (1 - t\psi_\nu)v_\nu = -\mathbf{E}\ \Delta_\nu v_\nu$, where Δ_ν is a deviation of $t\psi_\nu$ from the expectation value. Using (7) we have

$$t(\mathbf{E}\ V_\nu)^2 \le t\mathbf{E}\ v_\nu^2 \text{ var } t\psi_\nu \le t\sqrt{M}\,\delta = \tau\delta \le \omega_{52}.$$

This is the first statement of our theorem.

To estimate var V_ν, we use the martingale Lemma 2.2. Let $\nu = 1$. We eliminate the vector $\mathbf{x}_1$. Denote $\tilde{\mathbf{x}}_1 = \bar{\mathbf{x}}_1 - \mathbf{x}_1/N_1$, $S^1 = S - t\mathbf{x}_1\mathbf{x}_1^T/N$, and $w_1 = \mathbf{x}_1^T H_0^1\tilde{\mathbf{x}}_1$. Then

$$\mathbf{x}_1^T H_0\tilde{\mathbf{x}}_1(1 + t\mathbf{x}_1^T H_0^1\mathbf{x}_1/N) = \mathbf{x}_1^T H_0^1\tilde{\mathbf{x}}_1 = w_1.$$

It follows that $|\mathbf{x}_1^T H_0\tilde{\mathbf{x}}_1| \le |w_1|$. By (1), we have $t^2t_1^2\mathbf{E}\ w_1^4 \le Mt^2t_1^2\mathbf{E}\ (\tilde{\mathbf{x}}_1^T H_0^1\tilde{\mathbf{x}}_1)^2$. The expression in the parenthesis is the function Φ_{11} with one vector eliminated. Since $t_1\Phi_{11} \le 1$ we conclude that $t^2t_1^2\mathbf{E}\ w_1^4 \le Mt^2 = \tau^2$. First, we eliminate the dependence on $\mathbf{x}_1 \in \mathfrak{X}_1$ and then on $\mathbf{x}_2 \in \mathfrak{X}_2$ from V_1. Using (7) we have

$$\begin{aligned} V_1 = \mathbf{e}^T H_0\bar{\mathbf{x}}_1 &= \mathbf{e}^T H_0^1\tilde{\mathbf{x}}_1 + \mathbf{e}^T H_0^1\mathbf{x}_1/N - t\mathbf{e}^T H_0\mathbf{x}_1\mathbf{x}_1^T H_0^1\tilde{\mathbf{x}}_1/N \\ &= \mathbf{e}^T H_0^1\tilde{\mathbf{x}}_1 + u_1(1/N_1 - tw_1/N). \end{aligned}$$

On the other hand, $V_1 = \mathbf{e}^T H_0\bar{\mathbf{x}}_1 = \mathbf{e}^T H_0^2\bar{\mathbf{x}}_1 - tu_2\mathbf{x}_2^T H_0^2\bar{\mathbf{x}}_1/N$. Taking into account random vectors from both $\mathfrak{X}_1$ and $\mathfrak{X}_2$, by Lemma 2.2, we obtain

$$t \text{ var } (\mathbf{e}^T H_0\mathbf{x}_1) \le tN_1\mathbf{E}\ \sigma_1^2 + tN_2\mathbf{E}\ \sigma_2^2,$$

where $\sigma_1 = u_1(1/N_1 - tw_1/N)$, $\sigma_2 = u_2\mathbf{x}_2^T H_0^2\bar{\mathbf{x}}_1/N$. From the inequality $t_1\Phi_1 1 \le 1$ it follows that also $t_1\bar{\mathbf{x}}_1^T H_0^2\bar{\mathbf{x}}_1 \le 1$. We obtain

$$t_1\mathbf{E}\ \sigma_1^2 \le 2t_1\mathbf{E}\ u_1^2(N_1^{-2} + \mathbf{E}\ t^2w_1^2N^{-2}) \le 4\tau(1+\tau)/N_1^2,$$

$$t_2\mathbf{E}\ \sigma_2^2 \le t_2\mathbf{E}\ u_2^2 t^2 \mathbf{E}\ \bar{\mathbf{x}}_1^T H_0^2 \bar{\mathbf{x}}_1/N^2 \le \tau/NN_1.$$

Using these inequalities, we derive the required inequality for t var V_1. The symmetric estimate for $\nu = 2$ follows from assumptions. Statement 1 of the theorem is proved.

The first relation of statement 2 can be immediately deduced. Consider the expectation values. We have

$$\begin{aligned} t_1\mathbf{E}\ \Phi_{11} &= t_1\mathbf{E}\ \bar{\mathbf{x}}_1^T H_0 \bar{\mathbf{x}}_1 = t_1\mathbf{E}\ \mathbf{x}_1^T H_0 \bar{\mathbf{x}}_1 \\ &= t_1\mathbf{E}\ \mathbf{x}_1^T H_0 \mathbf{x}_1/N_1 + t_1\mathbf{E}\ \mathbf{x}_1^T H_0 \tilde{\mathbf{x}}_1. \end{aligned}$$

Here the first summand in the right hand side equals the product $t\mathbf{E}\ \psi_1 = 1 - s_{01}$. Using the second equation in (7), we can write the second summand as $\mathbf{E}\ t_1 \mathbf{x}_1^T H_0^1 \tilde{\mathbf{x}}_1(1 - t\psi_1)$. Let us replace the difference in the parenthesis by $1 - t\mathbf{E}\ \psi_1 - \Delta_1$, where Δ_1 is a deviation of $t\psi_1$ from its expectation. The contribution of the constant part of $1 - t\psi_1$ vanish. Since $t_1\Phi_{11} \le 1$, by eliminating the vector $\mathbf{x}_1$, we obtain that $t_1(\tilde{\mathbf{x}}_1^T H_0 \tilde{\mathbf{x}}_1) \le 1$ and

$$\begin{aligned} t_1|\mathbf{E}\ \mathbf{x}_1^T H_0 \tilde{\mathbf{x}}_1 \Delta_1| &\le [t_1^2\ \mathbf{E}\ (\mathbf{x}_1^T H_0^1 \tilde{\mathbf{x}}_1)^2 \text{ var } t\psi_1]^{1/2} \\ &\le [\sqrt{M}\,\mathbf{E}\ t_1^2(\tilde{\mathbf{x}}_1^T H_0^1 \tilde{\mathbf{x}}_1)\delta]^{1/2} \le \sqrt{\tau\delta} \le \sqrt{\omega_{52}}. \end{aligned}$$

To estimate var Φ_{11}, we again use Lemma 2.2. Let us eliminate the vector $\mathbf{x}_1$. We can rewrite $\Phi_{11} = \bar{\mathbf{x}}_1^T H_0 \bar{\mathbf{x}}_1$ in the form

$$\tilde{\mathbf{x}}_1^T H_0^1 \tilde{\mathbf{x}}_1 + 2\tilde{\mathbf{x}}_1^T H_0 \mathbf{x}_1/N_1 + \mathbf{x}_1^T H_0 \mathbf{x}_1/N_1^2 - t\tilde{\mathbf{x}}_1^T H_0 \mathbf{x}_1\, \mathbf{x}_1^T H_0^1 \tilde{\mathbf{x}}_1/N_1,$$

where $\tilde{\mathbf{x}}_1 = \bar{\mathbf{x}}_1 - \mathbf{x}_1/N_1$. On the other hand, excluding $\mathbf{x}_2$, we obtain

$$\Phi_{11} = \bar{\mathbf{x}}_1^T H_0^2 \bar{\mathbf{x}}_1 - t\bar{\mathbf{x}}_1^T H_0 \mathbf{x}_2 \mathbf{x}_2^T H_0^2 \bar{\mathbf{x}}_1/N.$$

By Lemma 2.2 we have var $\Phi_{11} \le N_1\mathbf{E}\ \sigma_1^2 + N_2\mathbf{E}\ \sigma_2^2$, where σ_1 is the sum of the last three summands of the first expression, and σ_2 is the last summand of the second expression for Φ_{11}. Using (7) and (1), we obtain that $t_1\tilde{\mathbf{x}}_1^T H_0 \tilde{\mathbf{x}}_1 \le 1$ and

$$\begin{gathered} t_1^2\mathbf{E}\ (\mathbf{x}_1^T H_0 \tilde{\mathbf{x}}_1)^2 \le t_1^2\mathbf{E}\ (\mathbf{x}_1^T H_0^1 \tilde{\mathbf{x}}_1)^2 \le \sqrt{M}\, t_1^2\ \mathbf{E}\ \tilde{\mathbf{x}}_1^T H_0^{12} \tilde{\mathbf{x}}_1 \le \tau, \\ t_1^2(\mathbf{x}_1^T H_0 \mathbf{x}_1)^2/N^2 = t_1^2\psi_1^2 \le 1, \\ t^2 t_1^2\mathbf{E}\ (\tilde{\mathbf{x}}_1^T H_0 \mathbf{x}_1 \mathbf{x}_1^T H_0^1 \tilde{\mathbf{x}}_1)^2 \le t^2 t_1^2\mathbf{E}\ (\tilde{\mathbf{x}}_1^T H_0^1 \mathbf{x}_1)^4 \le \tau^2 t^2\mathbf{E}\ w_1^4 \le \tau^2, \end{gathered}$$

where the second superscript denotes the square. If we put $\mathbf{x}_2 = 0$, then the relation $t_1\Phi_{11} \le 1$ implies $t_1\bar{\mathbf{x}}_1^T H_0^2 \bar{\mathbf{x}}_1 \le 1$. Hence,

$$t^2 t_1^2 \mathbf{E}\ (\bar{\mathbf{x}}_1^T H_0 \mathbf{x}_2 \mathbf{x}_2^T H_0^2 \bar{\mathbf{x}}_1)^2 \le t^2 t_1^2 \mathbf{E}\ (\bar{\mathbf{x}}_1^T H_0^2 \mathbf{x}_2)^4 \\ \le \tau^2 t_1^2 \mathbf{E}\ (\bar{\mathbf{x}}_1^T H_0^{22} \bar{\mathbf{x}}_1)^2 \le \tau^2.$$

Substituting these inequalities, we obtain the inequality $t_1 \mathrm{var}\ \Phi_{11} \le a(1+\tau^2)/N_1 \le \omega_{20}$, where a is a numerical constant. The symmetric relation for $\nu = 2$ follows from assumptions. Statement 2 is proved.

To prove statement 3 it suffices to repeat the arguments used in the proof of Lemma 2.3 with our resolvent H_0.

Further, from definition of Φ_{21} it follows that

$$t\mathbf{E}\ \Phi_{21} = t\mathbf{E}\ \bar{\mathbf{x}}_2^T H_0 \mathbf{x}_1 = t\mathbf{E}\ \bar{\mathbf{x}}_2^T H_0^1 \mathbf{x}_1 (1 - t\psi_1) = -t\mathbf{E}\ \bar{\mathbf{x}}_2^T H_0^1 \mathbf{x}_1 \Delta_1,$$

where Δ_1 is $t(\psi_1 - \mathbf{E}\ \psi_1)$. Estimating the right hand side by the Schwarz inequality and using (1) we obtain

$$(\mathbf{E}\ \Phi_{12})^2 \le \delta \mathbf{E}\ (\bar{\mathbf{x}}_1^T H_0^1 \bar{\mathbf{x}}_2)^2 \le \delta \mathbf{E}\ (\bar{\mathbf{x}}_2^T H_0^{12} \bar{\mathbf{x}}_2)\sqrt{M}.$$

Here $\|H_0^{12}\| \le \|H_0^1\|$, and the expression in the parenthesis is not greater Φ_{22} with $\mathbf{x}_1 = 0$. Since $t_2\Phi_{22} \le 1$, we can conclude that $t_1 t_2 (\mathbf{E}\ \Phi_{21})^2 \le \tau\delta \le \omega_{52}$.

To estimate the variance of Φ_{21} we rewrite this value in the form

$$\Phi_{12} = \bar{\mathbf{x}}_1^T H_0 \bar{\mathbf{x}}_2 = \tilde{\mathbf{x}}_1^T H_0^1 \mathbf{x}_2 + \mathbf{x}_1^T H_0 \bar{\mathbf{x}}_2 / N_1 - t\tilde{\mathbf{x}}_1^T H_0^1 \mathbf{x}_1\, \mathbf{x}_1^T H_0 \bar{\mathbf{x}}_2 / N_2,$$

where the first summand does not depend on $\mathbf{x}_1$. Using Lemma 2.2 we can state that var $\Phi_{12} \le N_1 \mathbf{E}\ \sigma_1^2 + N_2 \mathbf{E}\ \sigma_2^2$, where σ_1 is the sum of the last two terms and σ_2 is a symmetric expression. Using (7) we obtain

$$\mathbf{E}\ \sigma_1^2 \le [\mathbf{E}\ (\mathbf{x}_1^T H_0 \bar{\mathbf{x}}_2)^4 \mathbf{E}\ (1/N_1 - t\tilde{\mathbf{x}}_1^T H_0^1 \mathbf{x}_1 / N)^4]^{1/2},$$

$$t^2 t_2^2 \mathbf{E}\ (\mathbf{x}_1^T H_0 \bar{\mathbf{x}}_2)^4 \le t^2 t_2^2 \mathbf{E}\ (\mathbf{x}_1^T H_0^1 \bar{\mathbf{x}}_2)^4 \le M t^2 t_2^2 \mathbf{E}\ (\bar{\mathbf{x}}_2^T H_0^1 \bar{\mathbf{x}}_2)^2 \le \tau^2,$$

and

$$t^2 t_1^2 \mathbf{E}\ (\tilde{\mathbf{x}}_1^T H_0^1 \mathbf{x}_1)^4 \le M t^2 t_1^2 \mathbf{E}\ (\tilde{\mathbf{x}}^T H_0^{12} \tilde{\mathbf{x}}_1)^2 \le \tau^2,$$

where the second superscript 2 means the square. From these inequalities it follows that $N_1\mathbf{E}\,\sigma_1^2 \le 3\tau(1/N_1+\tau/N_1)$. The symmetric statement follows from assumptions. Combining these we obtain the inequality $t^2\text{var }\ \Phi_{12} \le 3\tau(1+\tau)/N_0 \le \omega_{20}$. Statement 4 is proved.

Now define $\varphi_{11} = \mathbf{x}_1 H_0^1 \mathbf{x}_1/N,\ \ \psi_{11} = \mathbf{x}_1 H_0 \mathbf{x}_1/N$. Using (7), we find that $s_{01}(t)$ equals

$$1 - \mathbf{E}\text{ tr } H_0 S_1/N = 1 - t\mathbf{E}\ \psi_{11} = \mathbf{E}\ (1+t\varphi_{11}) \ge (1+t\mathbf{E}\ \varphi_{11})^{-1}.$$

Here $t\mathbf{E}\ \varphi_{11} \le t\mathbf{E}\ \mathbf{x}_1^2/N \le \tau y$. The symmetric inequality for s_{02} follows from assumptions. We obtain the last statement. Theorem 3.2 is proved. □

Spectral Functions of Pooled Sample Covariance Matrices

THEOREM 3.3. *If $t \ge 0$, $N_1 > 1$ and $N_2 > 1$ then:*

1^o $\quad t\ \mathbf{E}\ U_\nu^2 \le \omega_{85}, \quad \nu = 1,2;$

2^o $\quad \mathbf{E}\ \|H(t) - H_0(t)\|^2 \le \omega_{63};$

3^o $\quad \mathbf{E}\ H(t) = (I + t_1 s_{01}(t)\Sigma_1 + t_2 s_{02}(t)\Sigma_2)^{-1} + \Omega, \quad \|\Omega\|^2 \le \omega_{63};$

$\quad \text{var }(\mathbf{e}^T H(t)\mathbf{e}) \le a\tau^2/N,$ *where a is a numerical coefficient;*

4^o $\quad t_\nu s_{0\nu}\mathbf{E}\ \Psi_{\nu\nu} = 1 - s_{0\nu} + o_\nu, \quad o_\nu^2 \le \omega_{74};$

$\quad t_\nu^2 \text{var }\Psi_{\nu\nu} \le \omega_{96}, \quad \nu = 1,2;$

5^o $\quad t_1 t_2 \mathbf{E}\ \Psi_{12}^2 \le \omega_{96}.$

Proof. We start from the identities

$$\begin{aligned} H &= H_0 + t_1 H\bar{\mathbf{x}}_1\bar{\mathbf{x}}_1^T H_0 + t_2 H\bar{\mathbf{x}}_2\bar{\mathbf{x}}_2^T H_0, \\ U_1 &= V_1 + t_1 U_1\Phi_{11} + t_2 U_2\Phi_{12}, \\ \Psi_{11} &= \Phi_{11} + t_1\Psi_{11}\Phi_{11} + t_2\Psi_{12}\Phi_{21}, \\ \Psi_{12} &= \Phi_{12} + t_1\Psi_{11}\Phi_{12} + t_2\Psi_{12}\Phi_{22}. \end{aligned} \tag{11}$$

Using the relation $1 - t_1\Phi_{11} = s_{01} - o_1$ from Theorem 3.2, we have

$$s_{01}U_1 = V_1 - U_1 o_1 + t_2 U_2 \Phi_{21}.$$

We square the both parts of this equation, multiply by t_1 and calculate the expectation. Using the Schwarz inequality we obtain

$$t_1 s_{01}^2 (\mathbf{E}\ U_1)^2/3 \le t_1(\mathbf{E}\ V_1)^2 + t_1 \mathbf{E}\ U_1^2 o_1^2 + t_2 \mathbf{E}\ U_2^2 t_1 t_2 \mathbf{E}\ \Phi_{21}^2. \quad (12)$$

Here, by Theorem 3.2, the first summand is not greater ω_{52}. In the second summand, $t_1 \mathbf{E}\, U_1^2 \le t_1 \mathbf{E}\ \bar{\mathbf{x}}_1^2 \le t_1 \sqrt{M}\, n/N_1 = \tau y, \quad o_1^2 \le \omega_{52}$. In the third summand, $t_2 \mathbf{E}\ U_2^2 \le \tau y$, and by Theorem 3.2, we have $t_1 t_2 \mathbf{E}\ \Phi_{21}^2 \le \omega_{52}$. Thus the right hand side of (12) is not greater ω_{63}. But $s_{01} \ge 1 + \tau y$. It follows that $t_1(\mathbf{E}\ U_1)^2 \le \omega_{85}$. The similiar estimate for $\nu = 2$ follows from assumptions. Statement 1 is proved.

Further, by (11), for any non-random vector $\mathbf{e}$ of unit length $\mathbf{e}^T H_0 \mathbf{e} - \mathbf{e}^T H \mathbf{e} = t_1 U_1 V_1 + t_2 U_2 V_2$, and consequently

$$\|\mathbf{E}\ H_0 - \mathbf{E}\ H\|^2 \le 2\mathbf{E}\ t_1 U_1^2 \mathbf{E}\ t_1 V_1^2 + 2\mathbf{E}\ t_2 U_2^2 \mathbf{E}\ t_2 V_2^2.$$

But $t_\nu U_{\nu\nu} \le 1, \quad \nu = 1, 2$. From Theorem 3.2, statement 2 follows.

Consider the expectation in the formulation of statement 3. We have

$$\mathbf{E}\ H = (I + t_1 s_{01} \Sigma_1 + t_2 s_{02} \Sigma_2)^{-1} + \Omega,$$

where $\Omega = \Omega_0 + \mathbf{E}\ H - \mathbf{E}\ H_0$. By Theorem 3.2, we obtain $\|\Omega\|^2 \le \omega_{63}$.

To estimate the variance of $\mathbf{e}^T H \mathbf{e}$, we apply Lemma 2.2. Let us eliminate vectors $\mathbf{x}_1 \in \mathfrak{X}_1$ and $\mathbf{x}_2 \in \mathfrak{X}_2$. Define $C^\nu = C - \mathbf{x}_\nu \mathbf{x}_\nu^T/N + \bar{\mathbf{x}}\mathbf{x}_\nu/N + \mathbf{x}_\nu \bar{\mathbf{x}}/N$, which does not depend on $\mathbf{x}_\nu$, $\nu = 1, 2$. Let $H^\nu = (I + tC^\nu)^{-1}$, $\nu = 1, 2$. Then

$$H^\nu = H - tH^\nu \mathbf{x}_\nu \mathbf{x}_\nu H(1 + 1/N)/N + tH^\nu \bar{\mathbf{x}}_\nu \mathbf{x}_\nu^T H/N + tH^\nu \mathbf{x}_\nu \bar{\mathbf{x}}_\nu^T H/N,$$

$\nu = 1, 2$. We apply Lemma 2.2 eliminating the dependence first on $\mathbf{x}_1$ and then on $\mathbf{x}_2$. We obtain

$$\begin{aligned}\text{var}\ (\mathbf{e}^T H \mathbf{e}) \le at^2 N^{-2} \sum_{\nu=1,2} \mathbf{E}\ &[(\mathbf{e}^T H^\nu \mathbf{x}_\nu)^2 (\mathbf{x}_\nu^T H \mathbf{e})^2 \\ &+ (\mathbf{e}^T H^\nu \mathbf{x}_\nu)^2 (\bar{\mathbf{x}}_\nu^T H \mathbf{e})^2 + (\mathbf{e}^T H^\nu \bar{\mathbf{x}}_\nu)^2 (\bar{\mathbf{x}}_\nu^T H \mathbf{e})^2].\end{aligned}$$

where a is a number. Here $|\mathbf{e}^T H \mathbf{x}_\nu| \le |\mathbf{e}^T H^\nu \mathbf{x}_\nu|$, $\mathbf{E}\ (\mathbf{e}^T H^\nu \mathbf{x}_\nu)^4 \le M$, and $\mathbf{E}\ \bar{\mathbf{x}}_\nu^2 \le \sqrt{M} y, \quad \nu = 1, 2$. Using this relation we obtain var $(\mathbf{e}^T H \mathbf{e}) \le a\tau^2/N$. Statement 3 is proved.

Further, we start from (11). Substituting $1 - t_1\Phi_{11}$ by Theorem 3.2, we find

$$s_{01}\Psi_{11} = \Phi_{11} + \Psi_{11}\Delta_1 + t_2\Psi_{12}\Phi_{21}, \tag{13}$$

where $\Delta_1 = t_1(\Phi_{11} - \mathbf{E}\ \Phi_{11})$. Let us multiply the left hand side by t_1 and calculate the expectation value. Using the Schwarz inequality we obtain the relation $t_1 s_{01}\mathbf{E}\ \Psi_{11} = 1 - s_{01} + o$, where

$$o^2/3 \le \omega_{52} + t_1^2\mathbf{E}\ \Psi_{11}^2 \text{var}\ (t_1\Phi_{11}) + t_1t_2\mathbf{E}\ \Phi_{12}^2 t_1t_2\mathbf{E}\ \Psi_{12}^2.$$

Here $t_1^2\mathbf{E}\ \Psi_{11} \le t_1^2\mathbf{E}\ (\bar{\mathbf{x}}_1^2)^2 \le Mt_1^2y_1^2 = \tau^2y^2$. By Theorem 3.2, we have that $t_1t_2\mathbf{E}\ \Phi_{12}^2 \le \omega_{52}$ and $t_1t_2\mathbf{E}\ \Psi_{12}^2 \le t_1t_2\mathbf{E}\ \bar{\mathbf{x}}_1^2\bar{\mathbf{x}}_2^2 \le \tau^2y^2$. We conclude that $o^2 \le \omega_{74}$. The symmetric statement for Ψ_{22} follows from assumptions.

To estimate the variance of Ψ_{11} we also start from (11). It suffices to estimate variances of the summands. Multiplying by t_1 and using the Schwarz inequality and the equality $t_1t_2\Phi_{12} \le 1$, we obtain that $s_{01}^2\text{var}\ (t_1\Psi_{11})/3$ is not larger than

$$\text{var}\ (t_1\Phi_{11}) + \mathbf{E}\ (t_1\Psi_{11})^2\text{var}\ (t_1\Phi_{11}) + t_1t_2\mathbf{E}\ \Phi_{12}^2\ t_1t_2\mathbf{E}\ \Psi_{12}^2.$$

Here, in the right hand side,

$$\mathbf{E}\ (t_1\Psi_{11})^2 \le t_1^2\mathbf{E}\ (\bar{\mathbf{x}}_1^2)^2 \le \tau^2y^2,$$
$$\mathbf{E}\ t_1t_2\Psi_{12}^2 \le t_1t_2\mathbf{E}\ \bar{\mathbf{x}}_1^2\bar{\mathbf{x}}_2^2 \le \tau^2y^2.$$

Using Theorem 3.2, we find that the right hand side is not greater ω_{74}. Since $s_{01} \ge 1 + \tau y$, the second part of statement 4 follows.

Further, (11) implies that $s_{02}\Psi_{12} = \Phi_{12} + \Psi_{12}\Delta_2 + t_1\Psi_{11}\Phi_{12}$, where $\Delta_2 = t_2(\Phi_{22} - \mathbf{E}\ \Phi_{22})$. We square both parts of this equality and multiply by t_1t_2. It follows that $s_{02}^2t_1t_2\mathbf{E}\ \Psi_{12}^2/3$ is not greater than

$$t_1t_2\mathbf{E}\ \Phi_{12}^2 + t_1t_2\mathbf{E}\ \Psi_{12}^2\ \text{var}\ (t_2\Phi_{22}) + t_1^2\mathbf{E}\ \Psi_{11}^2\ t_1t_2\mathbf{E}\ \Phi_{12}^2.$$

Here the first summand does not exceed ω_{52}; in the second summand $t_1t_2\mathbf{E}\ \Psi_{12}^2 \le \tau^2y^2$, $\text{var}\ t_2\Phi_{22} \le \omega_{20}$, and in the third summand we have $t_1^2\mathbf{E}\ \Psi_{11}^2 \le \tau^2y^2$. We obtain ω_{74} in the right hand side. Taking into account that $s_{01} \ge 1 + \tau y$, we come to statement 5. The proof of Theorem 3.3 is complete. $\square$

We now find the relations between $s_{0\nu}$ and s_ν and consider the unbiased estimators of these. Define

$$\widehat{s}_{0\nu} = \widehat{s}_{0\nu}(t) = 1 - t/N \text{ tr } H_0 S_\nu, \quad \widehat{s}_\nu = \widehat{s}_\nu(t) = 1 - t/N \text{ tr } HC_\nu,$$
$$\widehat{\Psi}_{\nu\nu} = (1 - \widehat{s}_\nu)/(t_\nu \widehat{s}_\nu), \quad \nu = 1, 2.$$

THEOREM 3.4. *If $t \geq 0$, then for $\nu = 1, 2$:*

1^o $\widehat{s}_{0\nu}(t) \geq 1 - n/N_\nu$, $\widehat{s}_\nu(t) \geq 1 - n/N_\nu$;

2^o $\mathbf{E}\ |\widehat{s}_\nu(t) - s_{0\nu}|^2 \leq \omega_{11}$;

3^o $\mathbf{E}\ |\widehat{s}_{0\nu}(t) - s_{0\nu}(t)|^2 \leq \delta \leq \omega_{42}$;

4^o $1 - s_{0\nu}(t)$
$= t/N\ s_{0\nu}(t) \text{ tr } [\Sigma_\nu (I + t_1 s_{01}(t)\Sigma_1 + t_2 s_{02}(t)\Sigma_2)^{-1} + o_\nu$,
where $o_\nu^2 \leq \omega_{64}$;

5^o $t^2(1-y)^2 \mathbf{E}\ (\widehat{\Psi}_{\nu\nu} - \Psi_{\nu\nu})^2 \leq \omega_{96}$.

Proof. We note that $t_\nu/N \text{ tr } (H_0 S_\nu) = 1/N \text{tr } (I - t_2 H_0 S_\nu) \leq y$. Obviously, $\widehat{s}_{0\nu} \geq 1 - n/N_\nu$, $\nu = 1, 2$. The inequality for s_1 and s_2 follows similiarly.

Next, we have $S_\nu = C_\nu + \bar{\mathbf{x}}_\nu \bar{\mathbf{x}}_\nu^T$, $\nu = 1, 2$, and the relation $H = H_0 + t_1 H_0 \bar{\mathbf{x}}_1 \bar{\mathbf{x}}_1^T H + t_2 H_0 \bar{\mathbf{x}}_2 \bar{\mathbf{x}}_2^T H$. We can write

$$s_1 - s_{01} = t/N\ \mathbf{E} \text{ tr } (H_0 S_1 - HC_1)$$
$$= t/N\ \bar{\mathbf{x}}_1^T H_0 \bar{\mathbf{x}}_1 - tt_1/N\ \mathbf{E}\ \bar{\mathbf{x}}_1^T H_0 C_1 H \bar{\mathbf{x}}_1 - tt_2/N\ \mathbf{E}\ \bar{\mathbf{x}}_2^T H_0 C_1 H \bar{\mathbf{x}}_2.$$

Here, in the right hand side, the first term is $t_1 \Phi_{11}/N_1 \leq 1/N_1$. We reduce the estimation of the second term to the estimation of two symmetric terms by the Schwarz inequality. Since $t_1 \|HC_1\| \leq 1$ and $t_1 \|H_0 C_1\| \leq 1$ we can state that the second term is not greater than $t\mathbf{E}\ \bar{\mathbf{x}}_1^2/N \leq \tau y/N_1$. Similiarly, the third term also is not greater $\tau y/N_2$. We conclude that $|\widehat{s}_1 - s_{01}| \leq 1/N_1 + \tau y/N \leq \omega_{11}$. The symmetric inequality follows from assumptions. Statement 2 is proved.

Now we estimate var $(\widehat{s}_{01})$. We notice that it is not greater than t^2 var $[(\text{tr } H_0 S_1)/N]$ which is equal to

$$t^2[\mathbf{E} \text{ tr } (H_0 S_1)/N(\mathbf{x}_1^T H_0 \mathbf{x}_1/N) - \mathbf{E} \text{ tr } [(H_0 S_1)/N]\ \mathbf{E}\ \mathbf{x}_1^T H_0 \mathbf{x}_1/N]$$
$$\leq [\text{var } (\widehat{s}_{01}) \text{ var } (t\psi_1)]^{1/2}.$$

We have var $(\widehat{s}_{01}) \leq$ var $(t\psi_1) \leq \delta \leq \omega_{42}$. The symmetric inequality also holds. Statement 3 follows.

To prove the fourth statement we notice that $\psi_1 = (1 - t\psi_1)\varphi_1$ and $1 - s_{01} = t\mathbf{E}\,\psi_1 = ts_{01}\mathbf{E}\,\varphi_1 - t\mathbf{E}\,\varphi_1\Delta_1$, where $\Delta_1 = t(\psi_1 - \mathbf{E}\,\psi_1)$. Using (7) we obtain that $\mathbf{E}\,\varphi_1$ is equal to

$$\mathbf{E}\ \mathrm{tr}\ (H_0^1\Sigma_1)/N = \mathrm{tr}\ (\mathbf{E}\ H_0\Sigma_1)/N - t\mathbf{E}\ \mathbf{x}_1^T H_0^1\Sigma_1 H_0^1\mathbf{x}_1/N^2.$$

Substitute $\mathbf{E}\ H_0$ from Theorem 3.1. It follows that $1 - s_{01}$ is equal to

$$ts_{01}\ \mathrm{tr}\ (\Sigma_1 R^{-1})/N - t^2 s_{01}\mathbf{E}\ \mathbf{x}_1^T H_0^1\Sigma_1 H_0\mathbf{x}_1/N^2 - t\mathbf{E}\ \varphi_1\Delta_1,$$

where $R = I + t_1 s_{01}\Sigma_1 + t_2 s_{02}\Sigma_2$. The first term in the right hand side is the main term of the fifth statement of the theorem. The other two terms constitute the correction o_1. We find

$$|o_1| \leq t^2\|\Sigma_1\|\ \mathbf{E}\ \mathbf{x}_1^2/N^2 + t\mathbf{E}\ \mathbf{x}_1^2|\Delta_1|/N \leq \tau^2 y/N + \tau y\sqrt{\delta}.$$

Since $\delta \leq \omega_{42}$ we obtain $o_1^2 \leq \omega_{64}$. By symmetry we have $o_2^2 \leq \omega_{64}$.

Let us prove the last statement of our theorem. Denote $y_\nu = n/N_\nu,\ \ \nu = 1, 2$. Since $\widehat{s}_\nu \geq 1 - y_\nu$ we have

$$\begin{aligned} &t_\nu^2(1 - y_\nu)^2\mathbf{E}\ (\widehat{\Psi}_{\nu\nu} - \Psi_{\nu\nu})^2 \leq \\ &\leq \mathbf{E}\ (1 - \widehat{s}_\nu - t_\nu\widehat{s}_\nu\Psi_{\nu\nu})^2 \\ &= [\mathbf{E}\ (1 - s_\nu - t_\nu s_\nu\Psi_{\nu\nu} + t_\nu(s_\nu - \widehat{s}_\nu)\Delta_\nu]^2 + \mathrm{var}\ [\widehat{s}_\nu(1 + t_\nu\Psi_{\nu\nu})], \end{aligned} \tag{14}$$

where $\Delta_\nu = \Psi_{\nu\nu} - \mathbf{E}\ \Psi_{\nu\nu},\ \ \nu = 1, 2$. By Theorem 3.3, the first summand is not greater than $2\omega_{74} + 2t_\nu$ var $\Psi_{\nu\nu} \leq \omega_{96}$. The second summand is not greater than

$$\mathbf{E}\ (1 + t_\nu\Psi_{\nu\nu})^2 \mathrm{var}\ \widehat{s}_\nu + \mathrm{var}\ (t_\nu\Psi_{\nu\nu}), \quad \nu = 1, 2.$$

Here $t_\nu^2\mathbf{E}\ \Psi_{\nu\nu}^2 \leq t_\nu^2\mathbf{E}\ (\bar{\mathbf{x}}_\nu^2)^2 \leq \tau^2 y^2$, and by statement 3, var $\widehat{s}_\nu \leq \omega_{42}$, $\nu = 1, 2$. Thus the second summand also is not greater ω_{96}. We conclude that the left hand side of (14) is not greater than ω_{96}. Since $t_\nu(1 - y_\nu) \geq t(\rho - y),\ \ \nu = 1, 2$, we obtain statement 5. The proof of Theorem 3.4 is complete. □

Limit Spectral Functions of the Increasing Pooled Sample Covariance Matrices

Let us perform the limit transition as $n \to \infty$. We consider a sequence $\mathfrak{P} = \{\mathfrak{P}_n\}$ of problems

$$\mathfrak{P}_n = (\mathfrak{S}_1, \mathfrak{S}_2, \Sigma_1, \Sigma_2, N_1, N_2, \mathfrak{X}_1, \mathfrak{X}_2, S, C)_n, \quad n = 1, 2, \ldots,$$

of the investigation of spectral functions of the pooled sample covariance matrices S and C calculated over two samples $\mathfrak{X}_1$ and $\mathfrak{X}_2$ of sizes N_1 and N_2 from different populations $\mathfrak{S}_1$ and $\mathfrak{S}_2$ with the parameters M, γ, and covariance matrices Σ_1 and Σ_2 (we do not write out the subscript n for the arguments of $\mathfrak{P}_n$). Assume that for some c:

A. $M < c$ for each $n = 1, 2, \ldots$.

B. $\lim\limits_{n\to\infty} \gamma = 0.$

C. $\lim\limits_{n\to\infty} n/N_\nu = \lambda_\nu, \quad \nu = 1, 2.$

D. $\lim\limits_{n\to\infty} N_\nu/N = \pi_\nu, \quad \nu = 1, 2.$

An additional assumption is required to provide the convergence of spectral functions. Denote

$$\varphi_n(x_0, x_1, x_2) = n^{-1} \ln\det\,(x_0 I + x_1\Sigma_1 + x_2\Sigma_2).$$

Assume that

E. $\lim\limits_{n\to\infty} \varphi_n(x_0, x_1, x_2) = \varphi_n(x_0, x_1, x_2)$, where the convergence is uniform with respect to $x_0 \geq 1$, $x_1, x_2 \geq 0$.

Denote $\lambda = \lambda_1\lambda_2/(\lambda_1 + \lambda_2)$. Under assumptions A-D, for each $t \geq 0$, we have $y = n/N \to \lambda$, and $\omega_{kl} = \omega_{kl}(t) \to 0$ for all $k, l \geq 0$. Denote the convergence in the square mean by $\xrightarrow{2}$.

THEOREM 3.5. *Let conditions A-E be satisfied. If* $t \geq 0$ *and* $\lambda_1 + \lambda_2 < 1$ *then in* $\{\mathfrak{P}_n\}$ *the convergence holds:*

$$1^o \quad s_{0\nu}(t) \to s^*_{0\nu}(t) = 1 - t\lambda \left.\frac{\partial\varphi(t_0, t_1, t_2)}{\partial t_\nu}\right|_{t_0=1},$$

$$\textit{for}\ \ t_\nu = t\pi_\nu s^*_{0\nu}(t), \quad \nu = 1, 2;$$

$$2^o \quad t_\nu \Phi_{\nu\nu} \xrightarrow{2} 1 - s^*_{0\nu}(t), \qquad \nu = 1, 2;$$

$$3^o \quad n^{-1}\mathrm{tr}\ H_0(t) \xrightarrow{2} h^*(t), \quad n^{-1}\mathrm{tr}\ H(t) \to h^*(t),$$

$$\textit{where}\ \ h^*(t) = \frac{\partial \varphi(t_0, t_1, t_2)}{\partial t_0}\Big|_{t_0=1} = 1, \quad t_\nu = t\pi_\nu s^*_{0\nu}(t), \quad \nu = 1, 2.$$

Proof. We first prove that the sequence $\{s_{0\nu}(t)\}$ converges in $\mathfrak{P}$ as $n \to \infty$, $\nu = 1, 2$.

Define $g_{0\nu} = g_{0\nu}(n) = (1 - s_{0\nu})/s_{0\nu}$, $\nu = 1, 2$. Note that $s_{0\nu} \geq (1 + \tau y)^{-1}$, $\nu = 1, 2$, and (20) implies that

$$\begin{aligned} yt\ \frac{\partial \varphi_n(1, x_1, x_2)}{\partial x_\nu} &= t/N\ \text{tr}\ [\Sigma_\nu (I + t_1 s_{01}\Sigma_1 + t_2 s_{02}\Sigma_2)^{-1}] \\ &= g_{0\nu} + o(1), \end{aligned}$$

where $x_\nu = t_\nu s_{0\nu}$, $\nu = 1, 2$. Under conditions A-E, $y \to \lambda$ and the functions $\varphi_n(1, x_1, x_2)$ converge to $\varphi(1, x_1, x_2)$ uniformly. Since the partial derivatives of φ_n of the second order are bounded from above (by $M < c$), we have

$$g_{0\nu} = \lambda t\ \frac{\partial \varphi(1, x_1, x_2)}{\partial x_\nu} + o(1) \tag{15}$$

with the same x_ν, $\nu = 1, 2$. Since $s_{0\nu} \geq (1 + \tau y)^{-1}$, to prove the convergence of $s_{0\nu}$, it suffices to show that $\{g_{0\nu}(n)\}$ converges as $n \to \infty$, $\nu = 1, 2$. Define $\Delta g_{0\nu} = g_{0\nu}(n+1) - g_{0\nu}(n)$, $\nu = 1, 2$. Let us estimate the change of the right hand side of (15) using derivatives at an intermediate point. We obtain the system of equations

$$\Delta g_{0\nu} = a_{\nu 1}\Delta g_{01} + a_{\nu 2}\Delta g_{02} + o(1), \quad \nu = 1, 2,$$

with coefficients that can be written in the form

$$a_{\nu\mu} = \lambda t \varphi_{n\nu\mu}\pi_\mu/(1 + \xi_\mu)^2,$$

where

$$\varphi_{n\nu\mu} = t_\mu n^{-1} \text{tr}\ (\Sigma_\nu\, \Sigma_\mu \Omega_\mu^{-2}), \quad \nu, \mu = 1, 2,$$

the matrix $\Omega_1 = I + t\Sigma_1\pi_1/(1 + \xi_1) + t\Sigma_2\pi_2/(1 + \xi_2)$, the matrix Ω_2 being defined similarly, and ξ_ν are magnitudes between $g_{0\nu}(n)$ and $g_{0\nu}(n + 1)$, $\nu = 1, 2$. To prove the convergence, it suffices to show that the determinant $\Delta_n = (1 - a_{11})(1 - a_{22}) - a_{12}a_{21}$ remains large enough in absolute value as $n \to \infty$. By the Cauchy–Bunyakovskii inequality, $\varphi_{n11}^2 \leq \varphi_{n11}\varphi_{n22}$, and we find that $a_{12}a_{21} \leq a_{11}a_{22}$. It follows that $\Delta_n \geq 1 - a_{11} - a_{22}$. But $\pi_1 a_{11}$ does not exceed λ multiplied by an expression of the form $n^{-1}\text{tr}\,[A(A+B)^{-1}A(A+B)^{-1}]$, where A and B are symmetric positive semidefinite matrices (matrix B being non-degenerate). It can be readily seen that this expression is not greater 1. Therefore, $a_{11} \leq \lambda_1$, and similarly, $a_{22} \leq \lambda_2$. We obtain

$$\Delta_n \geq 1 - \lambda_1 - \lambda_2 + o(1).$$

The existence of limits of $\lim s_{0\nu} = s_{0\nu}^*$ as $n \to \infty$, $\nu = 1, 2$, is shown. Now we notice that the arguments $x_\nu = t_\nu s_{0\nu} \to t\pi_\nu s_{0\nu}^*$, $\nu = 1, 2$, and (15) gives statement 1 of our theorem. The second statement follows immediately. We also notice that $h_0(t) = \partial\varphi_n(x_0, x_1, x_2)/\partial x_0$, where $x_0 = 1$, $x_\nu = t_\nu s_{0\nu}$, $\nu = 1, 2$. Since the functions φ_n are twice differentiable, we can perform the limit transition and obtain the last statement of our theorem. Theorem 3.5 is proved. □

Example. Let $\Sigma_1 = \Sigma_2 = \Sigma$ for $n = 1, 2, \ldots$. We have $s^*(t) \overset{\text{def}}{=} s_{01}^*(t) = s_{02}^*(t)$ and $h^*(t) - h^{-1}\text{tr}\,(I + ts^*(t)\Sigma)^{-1} \to 0$, where $s^*(t) = 1 - \lambda(1 - h^*(t))$ for each $t \geq 0$. If the matrices Σ have a limit spectrum and all eigenvalues of Σ are located on a segment $[c_1, c_2]$, where $c_1 > 0$ and c_2 do not depend on n, then the limit spectrum of matrices S and C exists and, if $\lambda < 1$, it lies within the segment $[(1-\sqrt{\lambda})^2 c_1,\ (1+\sqrt{\lambda})^2 c_2]$.

Now let the matrices Σ_1 and Σ_2 be of a special form with pairwise identical eigenvectors $\mathbf{e}_i$ with different eigenvalues $\{\lambda_i\}$:

$$\Sigma_1\mathbf{e}_i = d_1\mathbf{e}_i, \quad \Sigma_2\mathbf{e}_i = 0, \quad i = 1, \ldots, m,$$
$$\Sigma_1\mathbf{e}_i = 0, \quad \Sigma_2\mathbf{e}_i = d_2 e_i, \quad i = m+1, \ldots, n, \qquad n = 2, 3 \ldots.$$

For simplicity let d_1 and d_2 do not depend on n and $m/n \to \beta_1 > 0$, $\beta_2 = 1 - \beta_1$. Then the well-known routine of the Stieltjes transform inversion shows that the limit spectrum of S and C is located on the union of segments $[a_\nu(1-\sqrt{\lambda})^2,\ a_\nu(1+\sqrt{\lambda})^2]$, where $a_\nu = \beta_\nu d_\nu$, $\nu = 1, 2$. Thus we obtain a superposition of two well known 'semi-circle' distributions as expected. If $a_1 < a_2$ the inequality $a_1(1+\sqrt{\lambda})^2 < a_2(1-\sqrt{\lambda})^2$ is necessary and sufficient for the limit spectrum of S and C to be split into two disconnected parts.