UNIVERSITY OF CAMBRIDGE
Faculty of Economics

Institute for
New Economic Thinking

CAMBRIDGE WORKING PAPERS IN ECONOMICS
CAMBRIDGE-INET WORKING PAPERS

Extreme canonical correlations and high-dimensional cointegration analysis

Alexei Onatski
University of Cambridge

Chen Wang
University of Cambridge

Abstract

The simplest version of Johansen's (1988) trace test for cointegration is based on the squared sample canonical correlations between a random walk and its own innovations. Onatski and Wang (2017) show that the empirical distribution of such squared canonical correlations weakly converges to the Wachter distribution as the sample size and the dimensionality of the random walk go to infinity proportionally. In this paper we prove that, in addition, the extreme squared correlations almost surely converge to the upper and lower boundaries of the support of the Wachter distribution. This result yields strong laws of large numbers for the averages of functions of the squared canonical correlations that may be discontinuous or unbounded outside the support of the Wachter distribution. In particular, we establish the a.s. limit of the scaled Johansen's trace statistic, which has a logarithmic singularity at unity. We use this limit to derive a previously unknown analytic expression for the Bartlett-type correction coefficient for Johansen's test in a high-dimensional environment.

Reference Details

1805	Cambridge Working Papers in Economics
2018/20	Cambridge-INET Working Paper Series
Published	25 January 2018
Key Words	High-dimensional random walk, cointegration, extreme canonical correlations, Wachter distribution, trace statistic
Websites	www.econ.cam.ac.uk/cwpe www.inet.econ.cam.ac.uk/working-papers

Extreme canonical correlations and high-dimensional cointegration analysis

Alexei Onatski and Chen Wang
Faculty of Economics, University of Cambridge

August 5, 2017

Abstract

The simplest version of Johansen's (1988) trace test for cointegration is based on the squared sample canonical correlations between a random walk and its own innovations. Onatski and Wang (2017) show that the empirical distribution of such squared canonical correlations weakly converges to the *Wachter distribution* as the sample size and the dimensionality of the random walk go to infinity proportionally. In this paper we prove that, in addition, the extreme squared correlations almost surely converge to the upper and lower boundaries of the support of the Wachter distribution. This result yields strong laws of large numbers for the averages of functions of the squared canonical correlations that may be discontinuous or unbounded outside the support of the Wachter distribution. In particular, we establish the a.s. limit of the scaled Johansen's trace statistic, which has a logarithmic singularity at unity. We use this limit to derive a previously unknown analytic expression for the Bartlett-type correction coefficient for Johansen's test in a high-dimensional environment.

Key words: High-dimensional random walk, cointegration, extreme canonical correlations, Wachter distribution, trace statistic.

1 Introduction and the main result

Analysis of cointegration between a large number of time series is a challenging but useful exercise. Its applications include high-dimensional vector error correction modelling for forecasting purposes (Engel et al. (2015)), inference in nonstationary panel data models (Banerjee et al. (2004), Pedroni et al. (2015)), and verification of the assumptions under which composite commodity price indexes satisfy microeconomic laws of demand (Lewbel (1993), Brown (2003)). With increasing availability of large datasets, the needs for high-dimensional cointegration research will multiply.

A central role in the likelihood-based cointegration analysis is played by the squared sample canonical correlation coefficients between a simple transformation of the levels and the first differences of the data. This paper and its companion Onatski and Wang (2017) study such canonical correlations under the *simultaneous asymptotic* regime, where the dimensionality of the data goes to infinity proportionally to the sample size.

Onatski and Wang (2017) (OW17) show that the empirical distribution of the squared sample canonical correlations weakly converges to the so-called *Wachter distribution*. They use this result to explain the severe over-rejection of the no cointegration hypothesis when the dimensionality of the data is relatively large. In this paper, we show that the extreme squared canonical correlations almost surely (a.s.) converge to the upper and lower boundaries of the support of the Wachter distribution.

Our finding yields strong laws of large numbers for the averages of functions of the squared canonical correlations that may be discontinuous or unbounded outside an open interval containing the support of the Wachter distribution. In particular, we establish the a.s. limit of the scaled Johansen's (1988) trace statistic, which has a logarithmic singularity at unity.

We use this limit to derive an explicit expression for the Bartlett-type correction coefficient for Johansen's test. Such an expression was previously unknown, and the value of the coefficient had to be obtained numerically (see Johansen et al. (2005)).

Our setting can be described in the context of the likelihood ratio testing for no cointegration in the model

$$\Delta X_t = \Pi \left(X_{t-1} - t\hat{\rho}_1 \right) + \gamma + \eta_t, \tag{1}$$

where X_t, $t = 1, ..., T+1$, are p-dimensional data, $\Delta X_t = X_t - X_{t-1}$ with $X_0 = 0$, η_t are i.i.d. $N(0, \Sigma)$ vectors, and $\hat{\rho}_1 = X_{T+1}/(T+1)$. This model is similar to Johansen's (1995, eq. 5.14) model H^* :

$$\Delta X_t = \Pi \left(X_{t-1} - t\rho_1 \right) + \gamma + \eta_t, \tag{2}$$

where the deterministic trend is introduced so that there is no quadratic trend in X_t. In (1) ρ_1 is replaced by a "preliminary estimate" $\hat{\rho}_1$. Such a replacement yields the simultaneous diagonalizability of matrices used in the computation of the squared canonical correlations, which makes our theoretical analysis possible. We explain this in more detail in Section 5.

As is well known, the LR statistic for testing the null hypothesis that $\Pi = 0$ against $\Pi \neq 0$ equals

$$LR = -(T+1)\sum_{j=1}^{p} \ln(1-\lambda_{pj}), \tag{3}$$

where λ_{pj} is the j-th largest squared sample canonical correlation between demeaned vectors ΔX_t and $X_{t-1} - t\hat{\rho}_1$. In what follows, we will always assume that the null hypothesis holds so that the true value of Π is zero. In addition, we will assume that the true value of γ in the data generating process (1) is zero as well.

Note that demeaning $X_{t-1} - t\hat{\rho}_1$ and $X_{t-1} - (t-1)\hat{\rho}_1$ yields the same result. On the other hand, $X_t - t\hat{\rho}_1$ is a p-dimensional random walk detrended so that its last values are tied down to zero. Hence, λ_{pj} can be interpreted as the squared sample canonical correlations between a lagged detrended and demeaned random walk and its demeaned innovations.

Consider the simultaneous asymptotic regime where $p, T \to \infty$ so that $p/T \to c_0$. We abbreviate such a regime as $p, T \to_{c_0} \infty$. Without loss of generality, we assume that p is strictly increasing along the sequence, so that T can be viewed as a function of p.

OW17 shows that as $p, T \to_{c_0} \infty$ with $c_0 \in (0, 1]$, the empirical distribution of $\lambda_{p1} \geq ... \geq \lambda_{pp}$,

$$F_p(\lambda) \equiv \frac{1}{p}\sum_{i=1}^{p} \mathbf{1}\{\lambda_{pi} \leq \lambda\},$$

a.s. weakly converges[1] to the Wachter distribution W_{c_0} with an atom of size $\max\{0, 2 - 1/c_0\}$ at unity, and density

$$f(\lambda; c_0) = \frac{1+c_0}{2\pi c_0 \lambda(1-\lambda)}\sqrt{(b_{0+}-\lambda)(\lambda-b_{0-})} \tag{4}$$

[1] OW17 establishes the weak convergence $F_p(\lambda) \Rightarrow W_{c_0}(\lambda)$ both for Gaussian and non-Gaussian η. When η is non-Gaussian and has two finite moments, OW17 establishes the weak convergence in probability. When η is Gaussian, the convergence is a.s.

supported on $[b_{0-}, b_{0+}] \subseteq (0,1]$, where

$$b_{0\pm} = c_0 \left(\sqrt{2} \mp \sqrt{1-c_0}\right)^{-2}.$$

The main result of this paper strengthens OW17's finding as follows.

Theorem 1 *For* $c_0 \in (0, 1/2)$, $\lambda_{p1} \overset{\text{a.s.}}{\to} b_{0+}$ *and* $\lambda_{pp} \overset{\text{a.s.}}{\to} b_{0-}$ *as* $p, T \to_{c_0} \infty$.

For $c_0 \in (0, 1/2)$, Theorem 1 implies that no squared canonical correlations lie outside any open interval covering $[b_{0-}, b_{0+}]$ for sufficiently large p, a.s. Since F_p a.s. weakly converges to W_{c_0}, this implies that any function $f(\cdot)$ that is continuous and bounded on the open interval covering $[b_{0-}, b_{0+}]$, but may have discontinuities or other singularities outside that interval, satisfies the strong law of large numbers

$$\frac{1}{p}\sum_{j=1}^{p} f(\lambda_{pj}) \overset{\text{a.s.}}{\to} \int f(\lambda)\, \mathrm{d}W_{c_0}(\lambda)$$

as $p, T \to_{c_0} \infty$. In particular, the likelihood ratio statistic (3), although defined in terms of an unbounded function $\ln(1-\lambda)$, a.s. converges to a constant because its singularity lies outside $[b_{0-}, b_{0+}]$ for $c_0 \in (0, 1/2)$.[2]

Corollary 2 *Suppose that* $c_0 \in (0, 1/2)$. *Then as* $p, T \to_{c_0} \infty$, $LR/p^2 \overset{\text{a.s.}}{\to} LR_{c_0}$, *where*

$$LR_{c_0} = \frac{1+c_0}{c_0^2}\ln(1+c_0) - \frac{1-c_0}{c_0^2}\ln(1-c_0) + \frac{1-2c_0}{c_0^2}\ln(1-2c_0).$$

Proof: OW17 shows that the expression on the right hand side of the above display equals $-\int \ln(1-\lambda)\, \mathrm{d}W_{c_0}(\lambda)$. Since by Theorem 1, λ_{p1} a.s. remains bounded away from unity, the a.s. weak convergence of F_p to W_{c_0} implies that this integral is the a.s. limit of LR/p^2.□

In the next section we use Corollary 2 to derive a previously unknown explicit expression for the Bartlett-type correction coefficient for Johansen's trace test. In Section 3, we describe the setup for the proof of Theorem 1. Section 4 contains the proof. In Section 5 we discuss reasons for working with model (1) rather than (2), and derive some results for (2). Section 6 discusses directions for future work and concludes. All technical proofs are given in the Supplementary Material (SM).

[2] For $c_0 > 1/2$, λ_{p1} equals 1 with probability 1. Therefore, LR statistic is not well defined. For $c_0 = 1/2$, $b_{0+} = 1$ so that the singularity of $\ln(1-\lambda)$ lies at the upper boundary of the support of W_{c_0}.

2 Bartlett-type correction

The standard Johansen's LR test is based on the asymptotic critical values that assume that p is fixed whereas $T \to \infty$. As is well known, the test performs poorly in finite samples where p is moderately large. Even relatively small p's, such as five or six, lead to substantial over-rejection of the null hypothesis (see Ho and Sorensen (1996) and Gonzalo and Pitarakis (1995, 1999)).

One of the partial solutions to the over-rejection problem is the Bartlett correction of the LR statistic (see Johansen (2002)). The idea is to scale the statistic so that its finite sample distribution better fits the asymptotic distribution of the unscaled statistic. Specifically, let $E_{p,\infty}$ be the mean of the asymptotic distribution under the fixed-p, large-T asymptotic regime. Then, if the finite sample mean, $E_{p,T}$, satisfies

$$E_{p,T} = E_{p,\infty}\left(1 + a(p)/T + o\left(1/T\right)\right), \tag{5}$$

the scaled statistic is defined as $LR/\left(1 + a(p)/T\right)$. By construction, the fit between the scaled mean and the original asymptotic mean is improved by an order of magnitude. Although, as shown by Jensen and Wood (1997) in the context of unit root testing, the fit between higher moments does not improve by an order of magnitude, it may become substantially better (see Nielsen (1997)).

Theoretical analysis of the adjustment factor $1 + a(p)/T$ is difficult. The exact expression for $a(p)$ is known only for $p = 1$ (see Larsson (1998)). Therefore, Johansen (2002) proposes to approximate the Bartlett correction factor $BC_{p,T} \equiv E_{p,T}/E_{p,\infty}$ numerically. Here, we propose an alternative correction factor, equal to the ratio of the limits of LR/p^2 under the simultaneous asymptotics $p, T \to_{c_0} \infty$ and under the sequential asymptotics, where first T and then p goes to infinity.

Monte Carlo analysis in OW17 suggests that the simultaneous asymptotic limit LR_{c_0}, derived in Corollary 2, provides a very good centering point for LR/p^2, for moderately large p. From a theoretical perspective, this can be explained by the fact that, in contrast to the standard asymptotics, the simultaneous asymptotics does not neglect terms $(p/T)^j$ of relatively high order, which results in an improved approximation quality. The sequential asymptotic limit is derived in the following Theorem (see SM for a proof).

Theorem 3 *Suppose that* $c_0 \in (0, 1/2)$. *Then, as first* T *and then* p *go to infinity,* $LR/p^2 \to 2$ *in probability.*

This theorem and Corollary 2 yield the following analytic expression for the proposed Bartlett-type correction factor

$$\widehat{BC}_{p,T} = \frac{1+c}{2c^2}\ln(1+c) - \frac{1-c}{2c^2}\ln(1-c) + \frac{1-2c}{2c^2}\ln(1-2c), \tag{6}$$

where $c \equiv p/T$.

It is interesting to compare $\widehat{BC}_{p,T}$ to the numerical approximation to $BC_{p,T} \equiv E_{p,T}/E_{p,\infty}$, obtained in Johansen et al. (2005). That paper simulates $BC_{p,T}$ for various values of $p \leq 10$ and $T \leq 3000$ and fits a function of the form

$$BC^*_{p,T} = \exp\left\{a_1 c + a_2 c^2 + \left[a_3 c^2 + b\right]/T\right\}$$

to the obtained results. For relatively large values of T, the term $\left[a_3 c^2 + b\right]/T$ in the above expression is small. When it is ignored, the fitted function becomes particularly simple:

$$\widetilde{BC}_{p,T} = \exp\left\{0.549c + 0.552c^2\right\}.$$

Figure 1 superimposes the graphs of $\widehat{BC}_{p,T}$ and $\widetilde{BC}_{p,T}$ as functions of c. For $c \leq 0.3$, there is a strikingly good fit between the two curves, with the maximum distance between them 0.0067. For $c > 0.3$ the quality of the fit quickly deteriorates. This can be explained by the fact that all (p, T)-pairs used in Johansen et al's (2005) simulations are such that $c < 0.3$, so their numerical approximation does not cover cases with $c > 0.3$.

To the best of our knowledge, analytical expressions, such as (6), for the Bartlett-type correction factors were previously unavailable. Although the expression is not simple, it certainly is elementary, and easy to compute and analyze. Since the expression is analytic, it does not depend on details of any numerical experiments, and the range of its applicability covers all $c < 1/2$.

3 Setup

In this section, we introduce the setup for the proof of Theorem 1. Let ΔX, X_{-1} and η be $p \times (T+1)$ matrices with columns ΔX_t, $X_{t-1} - t\hat{\rho}_1$, and η_t, respectively. Further, let l be a $(T+1)$-vector of ones, $M_l = I_{T+1} - ll'/(T+1)$ be the projection on the space orthogonal to l, and let U be the $(T+1) \times (T+1)$ upper triangular matrix with ones above the main diagonal and zeros on the diagonal. Then under

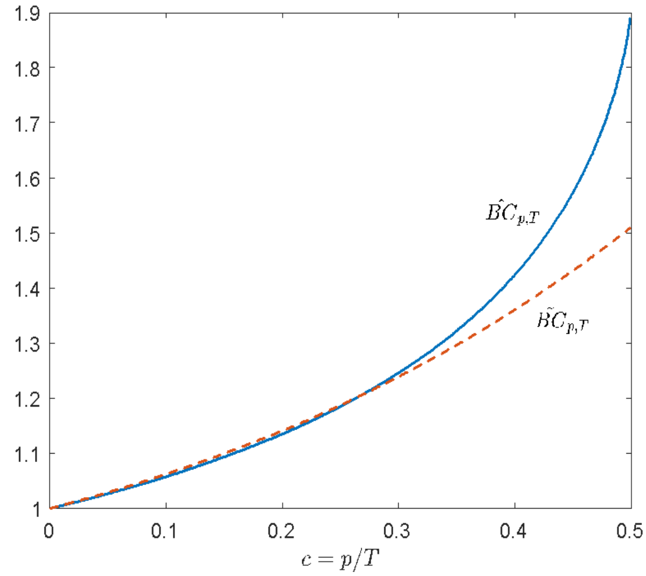

Figure 1: Bartlett correction factors as functions of p/T. Solid line: the factor based on the ratio of the simultaneous and seqeuntial limits of LR/p^2. Dashed line: numerical approximation from Johansen et al. (2005).

the null hypothesis

$$\Delta X M_l = \eta M_l \text{ and } X_{-1} M_l = \eta M_l U M_l, \tag{7}$$

where the second equality is derived as follows. Let $\tau = (1, 2, ..., T+1)'$. Note that $\tau' = l'U + l'$ and $\hat{\rho}_1 = \gamma + \eta l / (T+1)$. Therefore,

$$X_{-1} M_l = (\eta U - \hat{\rho}_1 \tau') M_l = \eta U M_l - \frac{1}{T+1} \eta l l' U M_l = \eta M_l U M_l.$$

Equations (7) imply that the squared sample canonical correlations λ_{pj}, $j = 1, ..., p$, between demeaned ΔX_t and demeaned $X_{t-1} - t\hat{\rho}_1$ can be interpreted as the eigenvalues of the product $P_1 P_2$, where P_1 and P_2 are projections on the column spaces of $M_l U' M_l \eta'$ and $M_l \eta'$, respectively. Clearly, λ_{pj}'s are invariant with respect to right-multiplication of η' by any invertible matrix. Hence, without loss of generality, we will assume that η_t are i.i.d. $N(0, I_p)$ vectors.

An equivalent interpretation of λ_{pj}, $j = 1, ..., p$, views them as the eigenvalues of matrix $S_{01} S_{11}^{-1} S_{10} S_{00}^{-1}$, where $S_{10} = S_{01}'$ and

$$S_{01} = \eta M_l U' M_l \eta', \; S_{11} = \eta M_l U M_l U' M_l \eta', \; S_{00} = \eta M_l \eta'. \tag{8}$$

As shown in OW17, $M_lU'M_l$, $M_lUM_lU'M_l$ and M_l are circulant matrices, that is, their (i_1, j_1)-th and (i_2, j_2)-th elements are equal as long as $i_1 - j_1$ equals $i_2 - j_2$ modulo $T+1$.

As is well known (e.g. Golub and Van Loan (1996), ch. 4.7.7), circulant matrices are simultaneously diagonalizable. Precisely, if V is a $(T+1) \times (T+1)$ circulant matrix with the first column v, then $V = \mathcal{F}^* \operatorname{diag}(\mathcal{F}v)\mathcal{F}/(T+1)$, where $\mathcal{F}$ is the Discrete Fourier Transform matrix with elements

$$\mathcal{F}_{st} = \exp\{-\mathrm{i}2\pi (s-1)(t-1)/(T+1)\},$$

and the superscript '$*$' denotes transposition and complex conjugation. This yields the following lemma.

Lemma 4 *Let* $\omega_s = 2\pi s/(T+1)$ *and*

$$\hat{\nabla} = \operatorname{diag}\left\{\left(e^{\mathrm{i}\omega_1} - 1\right)^{-1}, ..., \left(e^{\mathrm{i}\omega_T} - 1\right)^{-1}\right\}. \tag{9}$$

Further, let $\hat{\eta} = \eta\mathcal{F}^*$ *be a* $p \times (T+1)$ *matrix whose rows are the discrete Fourier transforms at frequencies* $0, \omega_1, ..., \omega_T$ *of the rows of* η*, and let* $\hat{\eta}_{-0}$ *be the* $p \times T$ *matrix obtained from* $\hat{\eta}$ *by removing its first column, corresponding to zero frequency. Then*

$$\begin{aligned} S_{01} &= \hat{\eta}_{-0}\hat{\nabla}\hat{\eta}_{-0}^*/(T+1), S_{11} = \hat{\eta}_{-0}\hat{\nabla}^*\hat{\nabla}\hat{\eta}_{-0}^*/(T+1), \\ S_{10} &= \hat{\eta}_{-0}\hat{\nabla}^*\hat{\eta}_{-0}^*/(T+1), \text{ and } S_{00} = \hat{\eta}_{-0}\hat{\eta}_{-0}^*/(T+1). \end{aligned}$$

The diagonal of $\hat{\nabla}$ consists of the reciprocals of the values of the transfer function (see e.g. Brillinger (1981) ch. 2.7) of the "leaded" first-difference filter

$$X_{t-1} \mapsto \Delta X_t \tag{10}$$

at frequencies ω_s, $s \neq 0$. Hence λ_{pj} can also be viewed as the sample squared canonical correlations between discrete Fourier transforms of $\hat{\eta}_t$ and their products with the inverse of the transfer function of filter (10). This yields a convenient frequency domain interpretation of Johansen's (1991) trace statistic (3). The strongly serially dependent time domain series $X_{t-1} - t\hat{\rho}_1$ are "replaced" by heteroskedastic frequency domain series $\left(1 - e^{-\mathrm{i}\omega_s}\right)^{-1}\hat{\eta}_s$ with $\hat{\eta}_{s_1}$ independent from $\hat{\eta}_{s_2}$ as long as $s_1 + s_2 \neq T+1$.

Below we will work with real-valued sin and cos Fourier transforms of η. In addition, we will interchange the order of frequencies so that ω_{s_1} and ω_{s_2} with $s_1 + s_2 = T + 1$ become adjacent pairs. Specifically, let T be even (the case of odd T can be analyzed similarly), let $P = \{p_{st}\}$ be a $T \times T$ permutation matrix with elements

$$p_{st} = \begin{cases} 1 \text{ if } s = 1, ..., T/2 \text{ and } t = 2s - 1 \\ 1 \text{ if } s = T/2 + 1, ..., T \text{ and } t = 2(T - s + 1) \\ 0 \text{ otherwise} \end{cases},$$

and let

$$W = I_{T/2} \otimes \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ \mathrm{i}/\sqrt{2} & -\mathrm{i}/\sqrt{2} \end{pmatrix},$$

where $\otimes$ denotes the Kronecker product. Further, let $\varepsilon = \hat{\eta}_{-0} P W^* / \sqrt{T(T+1)}$ and $\nabla = \operatorname{diag}\left\{\nabla_1, ..., \nabla_{T/2}\right\}$ with

$$\nabla_j = -\frac{1}{2} \begin{pmatrix} 1 & -\cot(\omega_j/2) \\ \cot(\omega_j/2) & 1 \end{pmatrix}.$$

A direct calculation shows that

$$\nabla \nabla' = \nabla' \nabla = \operatorname{diag}\left\{ r_1^{-1} I_2, ..., r_{T/2}^{-1} I_2 \right\} \text{ with } r_j = 4 \sin^2(\omega_j/2).$$

Lemma 5 *The columns of ε are i.i.d. $N(0, I_p/T)$ vectors. Matrix $S_{01} S_{11}^{-1} S_{10} S_{00}^{-1}$ equals $C D^{-1} C' A^{-1}$ where*

$$C = \varepsilon \nabla' \varepsilon', D = \varepsilon \nabla \nabla' \varepsilon', \text{ and } A = \varepsilon \varepsilon'.$$

This lemma yields yet another interpretation of λ_{pj}, $j = 1, ..., p$. They can be thought of as the eigenvalues of matrix

$$C D^{-1} C' A^{-1} \equiv (\varepsilon \nabla' \varepsilon') (\varepsilon \nabla \nabla' \varepsilon')^{-1} (\varepsilon \nabla \varepsilon') (\varepsilon \varepsilon')^{-1}.$$

The convenience of this interpretation stems from the block-diagonality of ∇ and the diagonality of $\nabla \nabla'$.

Let $\varepsilon_{(j)}$ be a $p \times 2$ matrix that consists of the $(2j-1)$-th and the $2j$-th columns of ε. In particular, $\varepsilon = \left[\varepsilon_{(1)}, ..., \varepsilon_{(T/2)}\right]$. The key advantage of studying C, D, A as opposed to S_{01}, S_{11}, and S_{00} is that C, D, A can be represented as sums of independent components of rank two. Specifically,

$$C=\sum\varepsilon_{(j)}\nabla_j'\varepsilon_{(j)}',\,D=\sum r_j^{-1}\varepsilon_{(j)}\varepsilon_{(j)}',\text{ and }A=\sum\varepsilon_{(j)}\varepsilon_{(j)}'.$$

OW17 exploits these representations to derive the limit of the empirical distribution F_p of the eigenvalues of $CD^{-1}C'A^{-1}$. That paper proves the convergence of F_p to W_{c_0} by establishing convergence of the Stieltjes transform of F_p, defined as

$$m_p(z)\equiv\int(\lambda-z)^{-1}\,\mathrm{d}F_p(\lambda)=\operatorname{tr}\left(CD^{-1}C'A^{-1}-zI_p\right)^{-1}/p.$$

Our proof of Theorem 1 relies on some of the results of OW17. Therefore, to complete the setup of the analysis below, we now briefly outline the relevant findings of that paper.

The first step in OW17's derivations is using the Sherman-Morrison-Woodbury formula for the inverse of a perturbed matrix V

$$(V+XWY)^{-1}=V^{-1}-V^{-1}X\left(W^{-1}+YV^{-1}X\right)^{-1}YV^{-1}$$

to derive identities

$$m_p(z) = \frac{T}{p}\frac{1}{1-z}-\frac{1}{p}\sum_{j=1}^{T/2}\frac{1}{(1-z)^2}\operatorname{tr}\left(\left[I_2,r_j\nabla_j'\right]\Omega_j^{(q)}\left[I_2,r_j\nabla_j'\right]'\right), \quad (11)$$

$$\frac{T}{p}+zm_p(z) = \frac{T}{p}\frac{1}{1-z}-\frac{1}{p}\sum_{j=1}^{T/2}\frac{1}{(1-z)^2}\operatorname{tr}\left(\left[I_2,r_jz\nabla_j'\right]\Omega_j^{(q)}\left[I_2,zr_j\nabla_j'\right]'\right) \quad (12)$$

$$1+zm_p(z) = \frac{T}{p}\frac{1}{1-z}-\frac{1}{p}\sum_{j=1}^{T/2}\frac{1}{(1-z)^2}\operatorname{tr}\left(\left[I_2,r_jz\nabla_j'\right]\Omega_j^{(q)}\left[I_2,r_j\nabla_j'\right]'\right), \quad (13)$$

$$0 = \frac{1}{p}\sum_{j=1}^{T/2}\frac{1}{1-z}\operatorname{tr}\left(\left[0,I_2\right]\Omega_j^{(q)}\left[I_2,r_j\nabla_j'\right]'\right), \quad (14)$$

where

$$\Omega_j^{(q)}\equiv\Omega_{pj}^{(q)}(z)=\begin{pmatrix}\frac{1}{1-z}I_2+v_j^{(q)}(z) & \frac{r_j}{1-z}\nabla_j'+u_j^{(q)\prime}(z)\\ \frac{r_j}{1-z}\nabla_j+u_j^{(q)}(z) & \frac{r_jz}{1-z}I_2+z\tilde{v}_j^{(q)}(z)\end{pmatrix}^{-1}. \quad (15)$$

The 2×2 matrices $v_j^{(q)}\equiv v_j^{(q)}(z)$, $u_j^{(q)}\equiv u_j^{(q)}(z)$, and $\tilde{v}_j^{(q)}\equiv\tilde{v}_j^{(q)}(z)$ are defined as follows. Let

$$\begin{aligned} A_j &= A - \varepsilon_{(j)}\varepsilon'_{(j)}, C_j = C - \varepsilon_{(j)}\nabla'_j\varepsilon'_{(j)},\ D_j = D - r_j^{-1}\varepsilon_{(j)}\varepsilon'_{(j)}, \\ M_j &= C_j D_j^{-1} C'_j - zA_j,\ \text{and}\ \tilde{M}_j = C'_j A_j^{-1} C_j - zD_j. \end{aligned}$$

Then,

$$v_j^{(q)} = \varepsilon'_{(j)} M_j^{-1} \varepsilon_{(j)},\ u_j^{(q)} = \varepsilon'_{(j)} D_j^{-1} C'_j M_j^{-1} \varepsilon_{(j)},\ \text{and}\ \tilde{v}_j^{(q)} = \varepsilon'_{(j)} \tilde{M}_j^{-1} \varepsilon_{(j)}.$$

The entries of these matrices are quadratic forms in the columns of $\varepsilon_{(j)}$. In what follows, we use superscript '(q)' to denote matrices that involve quadratic forms in the columns of $\varepsilon_{(j)}$ to distinguish them from similarly defined matrices that do not involve such quadratic forms.

The next step in OW17 is to replace $\Omega_j^{(q)}$ in equations (11-14) by matrix Ω_j, which is obtained from $\Omega_j^{(q)}$ by replacing $v_j^{(q)}(z)$, $u_j^{(q)}(z)$, and $\tilde{v}_j^{(q)}(z)$ in (15) with $v_p(z)I_2$, $u_p(z)I_2$, and $\tilde{v}_p(z)I_2$, respectively, where

$$v_p(z) = \operatorname{tr}\left(M^{-1}\right)/T,\ u_p(z) = \operatorname{tr}\left(D^{-1}C'M^{-1}\right)/T,\ \text{and}\ \tilde{v}_p(z) = \operatorname{tr}\left(\tilde{M}^{-1}\right)/T.$$

Here $M = CD^{-1}C' - zA$ and $\tilde{M} = C'A^{-1}C - zD$. To simplify notation, we will suppress the dependence of $v_p(z), u_p(z)$, and $\tilde{v}_p(z)$ on p and z. It is straightforward to verify that matrix Ω_j has the following explicit form

$$\Omega_j = \frac{1-z}{\delta_j}\begin{pmatrix} \frac{z}{1-z} r_j I_2 + z\tilde{v} I_2 & -\frac{1}{1-z} r_j \nabla'_j - uI_2 \\ -\frac{1}{1-z} r_j \nabla_j - uI_2 & \frac{1}{1-z} I_2 + vI_2 \end{pmatrix}, \tag{16}$$

where

$$\delta_j = z\tilde{v}\left(1 + v - zv\right) + r_j\left(u + zv - 1\right) - \left(1 - z\right)u^2. \tag{17}$$

Taking traces in equations (11-14), after replacing $\Omega_j^{(q)}$ by Ω_j, yields equations

$$m_p(z) = \frac{1}{c(1-z)} - \frac{2}{cT}\sum_{j=1}^{T/2} \frac{z\tilde{v} + r_j(u+v-1)}{(1-z)\delta_j} + e_1(z), \quad (18)$$

$$\frac{1}{c} + zm_p(z) = \frac{1}{c(1-z)} - \frac{2}{cT}\sum_{j=1}^{T/2} \frac{z\tilde{v} + r_j z(u+zv-1)}{(1-z)\delta_j} + e_2(z), \quad (19)$$

$$1 + zm_p(z) = \frac{1}{c(1-z)} - \frac{2}{cT}\sum_{j=1}^{T/2} \frac{z\tilde{v} + r_j(u(1+z)/2+zv-1)}{(1-z)\delta_j} + e_3(z), \quad (20)$$

$$0 = \frac{2}{cT}\sum_{j=1}^{T/2} \frac{-u - r_j v/2}{\delta_j} + e_4(z), \quad (21)$$

where $e_k(z)$, $k = 1, ..., 4$, are the approximation errors due to replacing $\Omega_j^{(q)}$ by Ω_j. Specifically,

$$e_1(z) = \frac{1}{p}\sum_{j=1}^{T/2} \frac{1}{(1-z)^2} \operatorname{tr}\left(\left[I_2, r_j\nabla_j'\right]\left(\Omega_j - \Omega_j^{(q)}\right)\left[I_2, r_j\nabla_j'\right]'\right), \quad (22)$$

$$e_2(z) = \frac{1}{p}\sum_{j=1}^{T/2} \frac{1}{(1-z)^2} \operatorname{tr}\left(\left[I_2, r_j z\nabla_j'\right]\left(\Omega_j - \Omega_j^{(q)}\right)\left[I_2, zr_j\nabla_j'\right]'\right), \quad (23)$$

$$e_3(z) = \frac{1}{p}\sum_{j=1}^{T/2} \frac{1}{(1-z)^2} \operatorname{tr}\left(\left[I_2, r_j z\nabla_j'\right]\left(\Omega_j - \Omega_j^{(q)}\right)\left[I_2, r_j\nabla_j'\right]'\right), \quad (24)$$

$$e_4(z) = -\frac{1}{p}\sum_{j=1}^{T/2} \frac{1}{1-z} \operatorname{tr}\left(\left[0, I_2\right]\left(\Omega_j - \Omega_j^{(q)}\right)\left[I_2, r_j\nabla_j'\right]'\right). \quad (25)$$

Finally, OW17 shows that the errors $e_k(z)$, $k = 1, ..., 4$, converge to zero point-wise over z from a compact subset of the upper half of the complex plane, $\mathbb{C}^+$. This allows OW17 to argue that $m_p(z)$ converges to $\bar{m}_0(z)$ uniformly over this compact subset, where $\bar{m}_0(z)$ satisfies the "limiting version" of system (18-21) that sets $e_k(z)$ to zeros. Solving the limiting system, OW17 shows that $\bar{m}_0(z)$ is the Stieltjes transform of W_{c_0}, which yields the convergence of F_p to W_{c_0}.

Our proof of Theorem 1 starts from the system (18-21). It amounts to establishing fast convergence of the errors $e_k(z)$, $k = 1, ..., 4$, to zero as z runs over a sequence z_p with $\operatorname{Im} z_p \to 0$ and $\operatorname{Re} z_p$ bounded away from the support of the Wachter distribution W_{c_0}.

4 Proof of Theorem 1

4.1 Outline of the proof

The general strategy of our proof is similar to that used in Bai and Silverstein's (1998) (BS98) study of the asymptotic behavior of the extreme eigenvalues of sample covariance matrices. The main ideas are as follows. Consider a sequence $\{z_p\}$ such that (s.t.)

$$x_p \equiv \operatorname{Re} z_p \in [0,1] \qquad \text{and} \qquad y_p \equiv \operatorname{Im} z_p = y_0 p^{-\alpha} \tag{26}$$

with $\alpha \geq 0$ and $y_0 \in (0,1]$ that are independent from p. We study the behavior of $m_p(z_p)$ as $p, T \to_{c_0} \infty$.

Let $m_0(z)$ be the Stieltjes transform of W_c, where W_c is obtained from the limiting distribution W_{c_0} by replacing c_0 with $c \equiv p/T$. Consider an interval $[a,b]$ outside the supports of W_c and W_{c_0} for all large p. Since F_p consists of masses $1/p$ at λ_{pj}, and since $W_c([a,b]) = 0$, we have the following decompositon

$$\operatorname{Im}(m_p(z_p) - m_0(z_p)) = \sum_{\lambda_{pj} \in [a,b]} \frac{1}{p} \frac{y_p}{(\lambda_{pj} - x_p)^2 + y_p^2} + \int_{[a,b]^c} \frac{y_p \mathrm{d}(F_p(\lambda) - W_c(\lambda))}{(\lambda - x_p)^2 + y_p^2}. \tag{27}$$

The existence of $\lambda_{pj} \in [a,b]$ puts an upper bound on the speed of convergence

$$\sup_{x_p \in [a,b]} |m_p(z_p) - m_0(z_p)| \to 0 \tag{28}$$

that is linked to the speed of convergence $y_p \to 0_+$ via the first term on the right hand side of (27). Proving that convergence (28) is faster than that bound shows that there are no λ_{pj} in $[a,b]$ for all sufficiently large p.

The analysis of the speed of convergence of (28) is done in several steps.

1. We show that the expected number of eigenvalues in $[a,b]$ cannot grow faster than p^β with $\beta < 1$ as $p \to \infty$.

2. We use 1. to derive an upper bound on the speed of convergence $m_p(z_p) - \mathbb{E}m_p(z_p) \to 0$ of the "stochastic part" of $m_p(z_p) - m_0(z_p)$.

3. We derive an upper bound on the speed of convergence $\mathbb{E}m_p(z_p) - m_0(z_p) \to 0$ of the "deterministic part" of $m_p(z_p) - m_0(z_p)$, and combine the results.

An implementation of these three steps requires a non-trivial extension of BS98. The fact that we have to deal with the product of four dependent stochastic matrices, $CD^{-1}C'A^{-1}$, presents substantial challenges, relative to the case of a sample covariance matrix, that we overcome. The key is to establish fast convergence of the errors $e_k(z_p)$ defined in (22-25) to zero, which requires detailed analysis of matrices Ω_j, $\Omega_j^{(q)}$, and their difference $\Omega_j - \Omega_j^{(q)}$.

4.2 Step 1: Speed of convergence of $\mathbb{E}F_p\left([a,b]\right)$

4.2.1 Rough bounds on the approximation errors

To establish bounds on the approximation errors $e_k(z_p)$, $k = 1, .., 4$, we will use the identity

$$\Omega_j - \Omega_j^{(q)} = \Omega_j^{(q)} \left((\Omega_j^{(q)})^{-1} - \Omega_j^{-1} \right) \Omega_j. \tag{29}$$

Definition 6 *(Tao and Vu (2011)) Let $\mathcal{E}$ be an event depending on p. Then $\mathcal{E}$ holds with overwhelming probability (w.ow.p.) if* $\Pr(\mathcal{E}) \geq 1 - O_C\left(p^{-C}\right)$ *for every constant $C > 0$. Here $O_C\left(p^{-C}\right)$ denotes a quantity that is smaller than Bp^{-C} with constant B that may depend on C.*

Lemma 7 *Suppose that $z = z_p$. Then for any $(C, d, \gamma) \in (0, \infty) \times (0, \infty) \times [0, 1/2)$ and any $\alpha \in [0, \alpha_{\gamma d})$ with $\alpha_{\gamma d} \equiv (1/2 - \gamma) / (1 + d)$, inequality*

$$\max_{j=1,\ldots,T/2} \left\| \Omega_j^{-1} - (\Omega_j^{(q)})^{-1} \right\| < Cp^{-\gamma} y_p^d$$

holds w.ow.p.

To prove the lemma, we use the convergence of quadratic forms $\xi_p' W_p \xi_p$ in Gaussian vectors ξ_p and the fact that the entries of $(\Omega_j^{(q)})^{-1}$ are such forms whereas the entries of Ω_j^{-1} are the points of concentration of these forms (see SM). Since $y_p = y_0 p^{-\alpha}$, the upper bound $Cp^{-\gamma} y_p^d$ on $||\Omega_j^{-1} - (\Omega_j^{(q)})^{-1}||$ converges to zero as fast as $p^{-\alpha d - \gamma}$. The rate $\alpha d + \gamma$ of such a convergence can be made arbitrarily close to $1/2$ by choosing α sufficiently close to $\alpha_{\gamma d}$, choosing d sufficiently large, and/or choosing γ sufficiently close to $1/2$. However, faster convergence rates for the bound are achieved at the expense of slower convergence of y_p to zero. The reason for such a trade-off is that the convergence of $\xi_p' W_p \xi_p$ is slowed down by large $\|W_p\|$, and quadratic forms appearing in the entries of $(\Omega_j^{(q)})^{-1}$ have $\|W_p\|$ that are proportional to $y_p^{-1} = y_0^{-1} p^{\alpha}$.

If we set $\alpha = 0$, y_p does not converge to zero as $p \to \infty$. However, since in such a case γ can be chosen arbitrarily close to 1/2, the upper bound on $||\Omega_j^{-1} - (\Omega_j^{(q)})^{-1}||$ derived by the lemma still may converge to zero at the rate arbitrarily close to 1/2.

Lemma 8 *(i) For any $\alpha \in [0, 1/12)$ there exists $C > 0$ such that w.ow.p.*

$$\max_{j=1,\dots,T/2} ||\Omega_j^{(q)}|| \leq C y_p^{-5} \text{ and } \max_{j=1,\dots,T/2} \|\Omega_j\| \leq C y_p^{-5}.$$

(ii) For any $\alpha \in [0, 1/6)$ and any $\rho > 0$, there exists $C > 0$ such that

$$\mathbb{E}\left(\max_{j=1,\dots,T/2} ||\Omega_j^{(q)}||^\rho\right) \leq C y_p^{-5\rho}.$$

The constant C in Lemma 8 does not need to coincide with that in Lemma 7. In what follows, C denotes a constant whose value can change from one appearance to another. Identity (29), Lemmas 7-8, and the fact that $|1 - z_p|^{-2} \leq y_p^{-2}$ imply that for any $C \in (0, \infty)$, $d \in [5, \infty)$, and $\gamma \in [0, 1/2)$,

$$|1 - z_p|^{-2} \max_{j=1,\dots,T/2} \left\|\Omega_j - \Omega_j^{(q)}\right\| \leq C p^{-\gamma} y_p^{d-12} \text{ w.ow.p.} \tag{30}$$

as long as $0 \leq \alpha < \alpha_{\gamma d}$. The requirement $d \geq 5$ ensures that $\alpha_{\gamma d} \leq 1/12$ so that Lemma 8 applies. Combining (30) with equation (22) yields

$$|e_1(z_p)| \leq C p^{-\gamma} y_p^{d-12} \text{ w.ow.p.}$$

Similar inequalities hold for $e_k(z_p)$, $k = 2, 3, 4$. Hence, we have the following lemma.

Lemma 9 *For any $(C, d, \gamma) \in (0, \infty) \times [5, \infty) \times [0, 1/2)$, any $\alpha \in [0, \alpha_{\gamma d})$, and any $k = 1, \dots, 4$, $|e_k(z_p)| \leq C p^{-\gamma} y_p^{d-12}$ w.ow.p.*

4.2.2 System reduction

In the SM, we show that system of equations (18-21) can be reduced to the following simple form

$$\begin{cases} \tilde{v} + 2u = \tilde{e}_1, \\ zv + u + c/(1-c) = \tilde{e}_2, \\ m - v(1-c)/c = \tilde{e}_3, \\ m^2 cz(1-z) - m(c - z + cz) + 1 = \tilde{e}_4. \end{cases} \tag{31}$$

The transformed errors $\tilde{e}_k$ are non-linear functions of the original errors e_k and of the variables $\tilde{v}$, u, v, and m (we suppress the dependence of all these quantities on p and z for the brevity of notations). We use bounds on these variables and Lemma 9 to derive the following result.

Lemma 10 *For any* $(C, d, \gamma) \in (0, \infty) \times [30, \infty) \times [0, 1/2)$, *any* $\alpha \in [0, \alpha_{\gamma d})$, *and any* $k = 1, ..., 4$, $|\tilde{e}_k| \leq C y_p^{d-42}$ *w.ow.p.*

4.2.3 Analysis of $m - m_0$

Let us define $m_0 \equiv m_0(z)$ as the solution of equation

$$m_0^2 cz (1 - z) - m_0 (c - z + cz) + 1 = 0$$

equal to

$$m_0 = \frac{c - z + cz + \sqrt{(c - z + cz)^2 - 4cz(1 - z)}}{2cz(1 - z)}, \tag{32}$$

where the branch of the square root, with the cut along the positive real semi-axis, is chosen so that the square root has positive imaginary part. It follows from e.g. Theorem 1.6 of Bai et al. (2015) that such m_0 is the Stieltjes transform of the Wachter distribution W_c with density

$$f(\lambda; c) = \frac{1 + c}{2\pi c \lambda (1 - \lambda)} \sqrt{(b_+ - \lambda)(\lambda - b_-)}$$

supported on $[b_-, b_+] \subseteq (0, 1]$, where $b_\pm = c\left(\sqrt{2} \mp \sqrt{1 - c}\right)^{-2}$.

Note that the expression under the square root in (32) can be factorized as

$$(c - z + cz)^2 - 4cz(1 - z) = (1 + c)^2 (z - b_+)(z - b_-) \tag{33}$$

Since the linear factors $z - b_+$ and $z - b_-$ cannot be simultaneously small, (33) implies a useful inequality

$$\left|(c - z_p + cz_p)^2 - 4cz_p(1 - z_p)\right| > C y_p \tag{34}$$

for some $C > 0$ and all sufficiently large p.

From the last equation of system (31), we have

$$m=\frac{c-z+cz+\sqrt{(c-z+cz)^{2}-4cz\,(1-z)+4\tilde{e}_{4}cz\,(1-z)}}{2cz\,(1-z)},$$

which differs from (32) only by the term $4\tilde{e}_4cz\,(1-z)$ under the square root. By Lemma 10 and inequality (34), when $z=z_p$, this term can be made negligible relative to the rest of the expression under the square root by choosing $d\geq 42$.[3] Then, the difference $m-m_0$ is of order

$$\tilde{e}_4/\sqrt{(c-z_p+cz_p)^2-4cz_p\,(1-z_p)}.$$

In the SM, we use this fact to prove the following lemma.

Lemma 11 *For any $\alpha\in[0,1/90)$, $l>0$, and $\rho\geq 180l$, there exists a constant C that may depend on α, l, and ρ s.t. for any $\epsilon>0$*

$$\Pr\left(y_p^{-1}\sup_{x_p\in[0,1]}|m\,(z_p)-m_0(z_p)|>\epsilon\right)\leq C\epsilon^{-\rho}p^{-l}.$$

The inequality established in Lemma 11 is analogous to inequality (3.23) in BS98. In the SM, we use BS98's argument leading from (3.23) to (3.28) to obtain a bound on $\mathbb{E}F_p\left([a,b]\right)$. Let $\mathbb{E}_0$ denote the unconditional expectation and $\mathbb{E}_k$ denote conditional expectation given $\varepsilon_{(1)},...,\varepsilon_{(k)}$.

Proposition 12 *Let $[a,b]$ be an interval that lies outside the supports of W_c and W_{c_0} for all sufficiently large p. We have*

$$\max_{k=0,\dots,T/2}\mathbb{E}_k\left(F_p\left([a,b]\right)\right)^2=o_{\text{a.s.}}\left(p^{-2/91}\right)\ \textit{and}\ \max_{k=0,\dots,T/2}\mathbb{E}_kF_p\left([a,b]\right)=o_{\text{a.s.}}\left(p^{-1/91}\right).$$

For future reference, we similarly have

$$\begin{aligned}\max_{k=0,\dots,T/2}\mathbb{E}_k\left(F_p\left([a',b']\right)\right)^2 &= o_{\text{a.s.}}\left(p^{-2/91}\right)\text{ and}\\ \max_{k=0,\dots,T/2}\mathbb{E}_kF_p\left([a',b']\right) &= o_{\text{a.s.}}\left(p^{-1/91}\right),\end{aligned}\tag{35}$$

where $[a',b']=[a-\underline{\epsilon},b+\underline{\epsilon}]$ with $\underline{\epsilon}$ such that $[a-2\underline{\epsilon},b+2\underline{\epsilon}]$ lies outside the support

[3] Even the choice $d=42$ and $\gamma=0$ would lead to the negligibility of $\tilde{e}_4$ because the constant C in Lemma 10 can be chosen at will, that is, arbitrarily small.

of W_{c_0}. Indeed, for all sufficiently large p, $[a', b']$ lies outside the supports of both W_c and W_{c_0} so that the requirement of Proposition 12 is met.

4.3 Step 2: Convergence of $m - \mathbb{E}m$

We now consider behavior of $m - \mathbb{E}m \equiv m_p(z_p) - \mathbb{E}m_p(z_p)$ along the sequence $z_p \equiv x_p + \mathrm{i}y_p$ with $y_p = y_0 p^{-\alpha}$, $\alpha = 1/456$, and $y_0 \in \mathbb{R}^+$ an arbitrary fixed positive real number. We will show that, for such a choice of α,

$$\sup_{x_p \in [a,b]} p y_p \left| m_p(z_p) - \mathbb{E}m_p(z_p) \right| \xrightarrow{\text{a.s.}} 0.$$

Since $\left| m_p\left(x^{(1)} + \mathrm{i}y_p\right) - m_p\left(x^{(2)} + \mathrm{i}y_p\right) \right| \leq \left| x^{(1)} - x^{(2)} \right| y_p^{-2}$, it is sufficient to show that $\max_{x_p \in S_p} p y_p \left| m_p(z_p) - \mathbb{E}m_p(z_p) \right| \xrightarrow{\text{a.s.}} 0$, where S_p is the set of p^2 points uniformly spaced on $[a, b]$.

We use the following key representation of $m - \mathbb{E}m$ in the form of a sum of the martingale difference sequence

$$m - \mathbb{E}m = \sum_{j=1}^{T/2} \mathbb{E}_j m - \mathbb{E}_{j-1} m.$$

As shown in the SM, this representation can be rewritten in the following form

$$m - \mathbb{E}m = \frac{1}{p} \sum_{j=1}^{T/2} \left(\mathbb{E}_j - \mathbb{E}_{j-1}\right) \operatorname{tr}\left(\Gamma_j^{(q)} \Omega_j^{(q)}\right), \tag{36}$$

where $\Omega_j^{(q)}$ is as defined in (15) above and

$$\Gamma_j^{(q)} = \begin{pmatrix} \frac{1}{1-z} v_j^{(q)} - a_j^{(q)} & \frac{1}{1-z} v_j^{(q)} r_j \nabla_j' - b_j^{(q)\prime} \\ \frac{1}{1-z} u_j^{(q)} - b_j^{(q)} & \frac{1}{1-z} u_j^{(q)} r_j \nabla_j' - c_j^{(q)} \end{pmatrix}$$

with

$$\begin{aligned} a_j^{(q)} &= \varepsilon_{(j)}' M_j^{-1} A_j M_j^{-1} \varepsilon_{(j)}, \\ b_j^{(q)} &= \varepsilon_{(j)}' D_j^{-1} C_j' M_j^{-1} A_j M_j^{-1} \varepsilon_{(j)}, \text{ and} \\ c_j^{(q)} &= \varepsilon_{(j)}' D_j^{-1} C_j' M_j^{-1} A_j M_j^{-1} C_j D_j^{-1} \varepsilon_{(j)}. \end{aligned}$$

Consider the identity

$$
\begin{aligned}
\Omega_j^{(q)} &= \Omega_j^{(d)} + \Omega_j^{(d)} \left((\Omega_j^{(d)})^{-1} - (\Omega_j^{(q)})^{-1} \right) \Omega_j^{(q)} \qquad (37) \\
&= \Omega_j^{(d)} + \Omega_j^{(d)} \left((\Omega_j^{(d)})^{-1} - (\Omega_j^{(q)})^{-1} \right) \Omega_j^{(d)} + \left(\Omega_j^{(d)} \left((\Omega_j^{(d)})^{-1} - (\Omega_j^{(q)})^{-1} \right) \right)^2 \Omega_j^{(q)},
\end{aligned}
$$

where

$$
\Omega_j^{(d)} \equiv \Omega_{pj}^{(d)}(z) = \begin{pmatrix} \left(\frac{1}{1-z} + \mathbb{E}v\right) I_2 & \frac{r_j}{1-z}\nabla_j' + \mathbb{E}uI_2 \\ \frac{r_j}{1-z}\nabla_j + \mathbb{E}uI_2 & \left(\frac{r_j z}{1-z} + z\mathbb{E}\tilde{v}\right) I_2 \end{pmatrix}^{-1}. \qquad (38)
$$

In this definition, we use superscript '(d)' to emphasize the fact that $\Omega_j^{(d)}$ is a deterministic matrix.

Lemma 13 *There exists $C > 0$, such that $\sup_{x_p \in [a,b]} \max_{j=1,\ldots,T/2} \left\| \Omega_{pj}^{(d)}(z_p) \right\| \leq C$ for all sufficiently large p.*

Identity (37) implies the following decomposition

$$
\begin{aligned}
m - \mathbb{E}m &= \frac{1}{p} \sum_{j=1}^{T/2} (\mathbb{E}_j - \mathbb{E}_{j-1}) \operatorname{tr}\left(\Gamma_j^{(q)} \Omega_j^{(d)} \right) \\
&+ \frac{1}{p} \sum_{j=1}^{T/2} (\mathbb{E}_j - \mathbb{E}_{j-1}) \operatorname{tr}\left(\Gamma_j^{(q)} \Omega_j^{(d)} \left((\Omega_j^{(d)})^{-1} - (\Omega_j^{(q)})^{-1} \right) \Omega_j^{(d)} \right) \qquad (39) \\
&+ \frac{1}{p} \sum_{j=1}^{T/2} (\mathbb{E}_j - \mathbb{E}_{j-1}) \operatorname{tr}\left(\Gamma_j^{(q)} \left(\Omega_j^{(d)} \left((\Omega_j^{(d)})^{-1} - (\Omega_j^{(q)})^{-1} \right) \right)^2 \Omega_j^{(q)} \right).
\end{aligned}
$$

We further expand (39) as follows. Define

$$
\hat{\Gamma}_j = \begin{pmatrix} \left(\frac{1}{1-z}v_j - a_j\right) I_2 & \frac{1}{1-z}v_j r_j \nabla_j' - b_j I_2 \\ \left(\frac{1}{1-z}u_j - b_j\right) I_2 & \frac{1}{1-z}u_j r_j \nabla_j' - c_j I_2 \end{pmatrix}
$$

and

$$
\hat{\Omega}_j = \begin{pmatrix} \left(\frac{1}{1-z} + v_j\right) I_2 & \frac{r_j}{1-z}\nabla_j' + u_j I_2 \\ \frac{r_j}{1-z}\nabla_j + u_j I_2 & \left(\frac{r_j z}{1-z} + z\tilde{v}_j\right) I_2 \end{pmatrix}^{-1},
$$

where

$$
\begin{aligned}
v_j &= \frac{1}{T}\operatorname{tr} M_j^{-1}, u_j = \frac{1}{T}\operatorname{tr}(C_j D_j^{-1} M_j^{-1}), \tilde{v}_j = \frac{1}{T}\operatorname{tr}\tilde{M}_j^{-1}, \\
a_j &= \frac{1}{T}\operatorname{tr}(M_j^{-1} A_j M_j^{-1}), b_j = \frac{1}{T}\operatorname{tr}(D_j^{-1} C_j' M_j^{-1} A_j M_j^{-1}), \text{ and} \\
c_j &= \frac{1}{T}\operatorname{tr}(D_j^{-1} C_j' M_j^{-1} A_j M_j^{-1} C_j D_j^{-1}).
\end{aligned}
$$

Then as shown in the SM, we have

$$
m - \mathbb{E}m = W_1 + W_2 + W_3 + W_4, \tag{40}
$$

where

$$
\begin{aligned}
W_1 &= \frac{1}{p}\sum_{j=1}^{T/2} \mathbb{E}_j \operatorname{tr}\left(\left(\Gamma_j^{(q)} - \hat{\Gamma}_j\right)\Omega_j^{(d)}\right), \\
W_2 &= \frac{1}{p}\sum_{j=1}^{T/2} \mathbb{E}_j \operatorname{tr}\left(\hat{\Gamma}_j \Omega_j^{(d)}\left(\hat{\Omega}_j^{-1} - (\Omega_j^{(q)})^{-1}\right)\Omega_j^{(d)}\right), \\
W_3 &= \frac{1}{p}\sum_{j=1}^{T/2} (\mathbb{E}_j - \mathbb{E}_{j-1}) \operatorname{tr}\left(\left(\Gamma_j^{(q)} - \hat{\Gamma}_j\right)\Omega_j^{(d)}\left((\Omega_j^{(d)})^{-1} - (\Omega_j^{(q)})^{-1}\right)\Omega_j^{(d)}\right), \text{ and} \\
W_4 &= \frac{1}{p}\sum_{j=1}^{T/2} (\mathbb{E}_j - \mathbb{E}_{j-1}) \operatorname{tr}\left(\Gamma_j^{(q)}\left(\Omega_j^{(d)}\left((\Omega_j^{(d)})^{-1} - (\Omega_j^{(q)})^{-1}\right)\right)^2 \Omega_j^{(q)}\right).
\end{aligned}
$$

Terms W_k in the decomposition (40) are small in the sense that their moments quickly decay as $p \to \infty$. A general strategy of proving this uses the fact that all these terms can be viewed as sums of martingale difference sequences, and therefore Burkholder's moment inequalities (see Lemmas 2.1 and 2.2 in BS98) are applicable. The moments of the corresponding summands can be bounded using results on quadratic forms in Gaussian vectors, detailed in the SM.

The so-obtained bounds involve quantities such as $\mathbb{E}\operatorname{tr}(M^{-1})/T$. These quantities can be split into two parts, corresponding to the eigenvalues λ_{pj} that lie outside and inside the interval $[a', b']$. The "outside" components are bounded for $x_p \in [a, b]$ because the distance between $[a', b']^c$ and $[a, b]$ is fixed and positive. The "inside" components are bounded by products of powers of y_p^{-1} and $\mathbb{E}F_p([a', b'])$, the expected proportion of eigenvalues λ_{pj} that belong to $[a', b']$. Given the choice of y_p made in this section, such products are small by (35). Following this general strategy, we prove our next proposition (see SM), which is the main result of this subsection.

Proposition 14 *For any* $k = 1, ..., 4$, $\max_{x_p \in S_p} py_p |W_k| \overset{\text{a.s.}}{\to} 0$, *and hence,*

$$\max_{x_p \in S_p} py_p |m - \mathbb{E}m| \overset{\text{a.s.}}{\to} 0.$$

4.4 Step 3: Convergence of $\mathbb{E}m - m_0$

Taking expectations of both parts of equations (11-14) and replacing $\mathbb{E}\Omega_j^{(q)}$ by $\Omega_j^{(d)}$, we obtain an analog of the "approximate system" (18-21) for variables $\mathbb{E}m$, $\mathbb{E}v$, $\mathbb{E}u$, and $\mathbb{E}\tilde{v}$ instead of m, v, u, and $\tilde{v}$.

$$\mathbb{E}m = \frac{1}{c(1-z)} - \frac{2}{cT}\sum_{j=1}^{T/2}\frac{z\mathbb{E}\tilde{v} + r_j(\mathbb{E}u + \mathbb{E}v - 1)}{(1-z)\bar{\delta}_j} + \bar{e}_1, \tag{41}$$

$$\frac{1}{c} + z\mathbb{E}m = \frac{1}{c(1-z)} - \frac{2}{cT}\sum_{j=1}^{T/2}\frac{z\mathbb{E}\tilde{v} + r_j z(\mathbb{E}u + z\mathbb{E}v - 1)}{(1-z)\bar{\delta}_j} + \bar{e}_2, \tag{42}$$

$$1 + z\mathbb{E}m = \frac{1}{c(1-z)} - \frac{2}{cT}\sum_{j=1}^{T/2}\frac{z\mathbb{E}\tilde{v} + r_j(\mathbb{E}u(1+z)/2 + z\mathbb{E}v - 1)}{(1-z)\bar{\delta}_j} + \bar{e}_3, \tag{43}$$

$$0 = \frac{2}{cT}\sum_{j=1}^{T/2}\frac{-\mathbb{E}u - r_j\mathbb{E}v/2}{\bar{\delta}_j} + \bar{e}_4, \tag{44}$$

where

$$\bar{\delta}_j = z\mathbb{E}\tilde{v}(1 + \mathbb{E}v - z\mathbb{E}v) + r_j(\mathbb{E}u + z\mathbb{E}v - 1) - (1-z)(\mathbb{E}u)^2,$$

and

$$
\begin{aligned}
\bar{e}_1 &= \frac{1}{p}\sum_{j=1}^{T/2}\frac{1}{(1-z)^2}\operatorname{tr}\left(\left[I_2, r_j\nabla_j'\right]\left(\Omega_j^{(d)}-\mathbb{E}\Omega_j^{(q)}\right)\left[I_2, r_j\nabla_j'\right]'\right),\\
\bar{e}_2 &= \frac{1}{p}\sum_{j=1}^{T/2}\frac{1}{(1-z)^2}\operatorname{tr}\left(\left[I_2, r_jz\nabla_j'\right]\left(\Omega_j^{(d)}-\mathbb{E}\Omega_j^{(q)}\right)\left[I_2, zr_j\nabla_j'\right]'\right),\\
\bar{e}_3 &= \frac{1}{p}\sum_{j=1}^{T/2}\frac{1}{(1-z)^2}\operatorname{tr}\left(\left[I_2, r_jz\nabla_j'\right]\left(\Omega_j^{(d)}-\mathbb{E}\Omega_j^{(q)}\right)\left[I_2, r_j\nabla_j'\right]'\right),\\
\bar{e}_4 &= -\frac{1}{p}\sum_{j=1}^{T/2}\frac{1}{1-z}\operatorname{tr}\left(\left[0, I_2\right]\left(\Omega_j^{(d)}-\mathbb{E}\Omega_j^{(q)}\right)\left[I_2, r_j\nabla_j'\right]'\right).
\end{aligned}
$$

The identity $\Omega_j^{(d)}-\Omega_j^{(q)}=\Omega_j^{(d)}((\Omega_j^{(q)})^{-1}-(\Omega_j^{(d)})^{-1})\Omega_j^{(q)}$ yields a decomposition

$$
\Omega_j^{(d)}-\mathbb{E}\Omega_j^{(q)}=R_1+R_2+R_3, \tag{45}
$$

where

$$
\begin{aligned}
R_1 &= \Omega_j^{(d)}\mathbb{E}((\Omega_j^{(q)})^{-1}-(\Omega_j^{(d)})^{-1})\Omega_j^{(d)},\\
R_2 &= -\mathbb{E}\left(\left(\Omega_j^{(d)}((\Omega_j^{(q)})^{-1}-(\Omega_j^{(d)})^{-1})\right)^2\Omega_j^{(d)}\right), \text{ and}\\
R_3 &= \mathbb{E}\left(\left(\Omega_j^{(d)}((\Omega_j^{(q)})^{-1}-(\Omega_j^{(d)})^{-1})\right)^3\Omega_j^{(q)}\right).
\end{aligned}
$$

As we show in the SM, for any $x_p\in[a,b]$, $||\mathbb{E}((\Omega_j^{(q)})^{-1}-(\Omega_j^{(d)})^{-1})||$ and $\mathbb{E}||((\Omega_j^{(q)})^{-1}-(\Omega_j^{(d)})^{-1})||^2$ are of order p^{-1}, whereas $\mathbb{E}||((\Omega_j^{(q)})^{-1}-(\Omega_j^{(d)})^{-1})||^3$ is of an even smaller order, and $\left\|\Omega_j^{(d)}\right\|$ as well as $\mathbb{E}\left\|\Omega_j^{(q)}\right\|$ are bounded. These facts would have implied that $\bar{e}_k$ are of order p^{-1}, had there been no $(1-z)^{-2}$ multipliers in the definition of $\bar{e}_1,...,\bar{e}_3$, and $(1-z)^{-1}$ multiplier in the definition of $\bar{e}_4$. If $[a,b]$ includes unity, then these multipliers are not uniformly bounded over $\operatorname{Re} z\in[a,b]$. However, it turns out that the norms of $(1-z)^{-1}\left[I_2, r_j\nabla_j'\right]\Omega_j^{(d)}$ and of $(1-z)^{-1}\left[I_2, r_jz\nabla_j'\right]\Omega_j^{(d)}$ are uniformly bounded over $[a,b]$ (see the proof of Lemma 15 in the SM), which is sufficient to guarantee that $\bar{e}_k$ are of order p^{-1} notwithstanding the presence of the multipliers $(1-z)^{-2}$ and $(1-z)^{-1}$ in their definitions.

Lemma 15 *There exists $C>0$, s.t. for any $k=1,...,4$, $\sup_{x_p\in[a,b]}|\bar{e}_k(z_p)|\leq Cp^{-1}$.*

As explained in the proof given in the SM, the inequality for $\bar{e}_2$ can be slightly strengthened so that

$$\sup_{x_p \in [a,b]} \left| \bar{e}_2 \left(z_p \right) / z_p \right| \leq C p^{-1}. \tag{46}$$

Such a strengthened version is used in the proof of Lemma 16.

Similarly to the above reduction of the "approximate system" (18-21) to the simple form (31), we reduce the system of equations (41-44) to

$$\begin{cases} \mathbb{E}\tilde{v} + 2\mathbb{E}u = \hat{e}_1, \\ z\mathbb{E}v + \mathbb{E}u + c/\left(1-c\right) = \hat{e}_2, \\ \mathbb{E}m - \mathbb{E}v \left(1-c\right)/c = \hat{e}_3, \\ \left(\mathbb{E}m\right)^2 cz\left(1-z\right) - \mathbb{E}m\left(c - z + cz\right) + 1 = \hat{e}_4, \end{cases} \tag{47}$$

where $\hat{e}_k$, $k = 1, ..., 4$, are nonlinear functions of $\bar{e}_k$, $k = 1, ..., 4$, $\mathbb{E}v$, $\mathbb{E}u$, and $\mathbb{E}\tilde{v}$.

Lemma 16 *There exists $C > 0$, s.t. for any $k = 1, ..., 4$, $\sup_{x_p \in [a,b]} \left| \hat{e}_k \left(z_p \right) \right| \leq Cp^{-1}$.*

Now recall the explicit form (32) of m_0. The fourth equation of (47) yields a similar expression for $\mathbb{E}m$,

$$\mathbb{E}m(z) = \frac{c - z + cz + \sqrt{\left(c - z + cz\right)^2 - 4cz\left(1-z\right) + 4\hat{e}_4 cz\left(1-z\right)}}{2cz\left(1-z\right)}.$$

Hence, the difference $\left| \mathbb{E}m\left(z_p\right) - m_0\left(z_p\right) \right|$ is of the order of

$$\hat{e}_4\left(z_p\right) / \sqrt{\left(c - z_p + cz_p\right)^2 - 4cz_p\left(1 - z_p\right)}.$$

On the other hand, identity (33) implies that

$$\inf_{x_p \in [a,b]} \left| \left(c - z_p + cz_p\right)^2 - 4cz_p\left(1 - z_p\right) \right| > \underline{\epsilon}$$

for all sufficiently large p, where $\underline{\epsilon}$ is the positive number used in the definition of $[a', b']$. Therefore, the following Proposition follows from Lemma 16.

Proposition 17 *There exists $C > 0$, s.t. $\sup_{x_p \in [a,b]} \left| \mathbb{E}m\left(z_p\right) - m_0\left(z_p\right) \right| \leq Cp^{-1}$.*

Propositions 14 and 17 yield

$$\sup_{x_p \in [a,b]} \left| m\left(z_p\right) - m_0\left(z_p\right) \right| = o_{\text{a.s.}}\left(1/\left(py_p\right)\right) \tag{48}$$

with $y_p = y_0 p^{-\alpha}$, $\alpha = 1/456$, and y_0 an arbitrary fixed positive real number. This yields Theorem 1 via the following arguments. The main idea of these arguments is outlined in Section 4.1 above.

Using (48), we obtain

$$\max_{k\in\{1,2,\ldots,228\}} \sup_{x_p\in[a,b]} \left| m\left(x_p + \mathrm{i}\sqrt{k}p^{-\alpha}\right) - m_0\left(x_p + \mathrm{i}\sqrt{k}p^{-\alpha}\right)\right| = o_{\text{a.s.}}\left(p^{-1+\alpha}\right).$$

Taking imaginary parts, we obtain

$$\max_{k\in\{1,2,\ldots,228\}} \sup_{x_p\in[a,b]} \left| \int \frac{\mathrm{d}\left(F_p(\lambda) - W_c(\lambda)\right)}{\left(x_p-\lambda\right)^2 + kp^{-2\alpha}}\right| = o_{\text{a.s.}}\left(p^{-1+2\alpha}\right).$$

Taking differences of the integrals corresponding to different values of k yields

$$\begin{aligned}
\max_{k_1\neq k_2} \sup_{x_p\in[a,b]} \left| \int \frac{p^{-2\alpha}\mathrm{d}\left(F_p(\lambda) - W_c(\lambda)\right)}{\prod_{s=1}^{2}\left(\left(x_p-\lambda\right)^2 + k_s p^{-2\alpha}\right)}\right| &= o_{\text{a.s.}}\left(p^{-1+2\alpha}\right), \\
\max_{\substack{k_1,k_2,k_3\\ \text{distinct}}} \sup_{x_p\in[a,b]} \left| \int \frac{p^{-4\alpha}\mathrm{d}\left(F_p(\lambda) - W_c(\lambda)\right)}{\prod_{s=1}^{3}\left(\left(x_p-\lambda\right)^2 + k_s p^{-2\alpha}\right)}\right| &= o_{\text{a.s.}}\left(p^{-1+2\alpha}\right), \\
&\vdots \\
\sup_{x_p\in[a,b]} \left| \int \frac{p^{-454\alpha}\mathrm{d}\left(F_p(\lambda) - W_c(\lambda)\right)}{\prod_{s=1}^{228}\left(\left(x_p-\lambda\right)^2 + s p^{-2\alpha}\right)}\right| &= o_{\text{a.s.}}\left(p^{-1+2\alpha}\right),
\end{aligned}$$

so that

$$\sup_{x_p\in[a,b]} \left| \int \frac{\mathrm{d}\left(F_p(\lambda) - W_c(\lambda)\right)}{\prod_{s=1}^{228}\left(\left(x_p-\lambda\right)^2 + s p^{-2\alpha}\right)}\right| = o_{\text{a.s.}}(1).$$

Splitting up the integral, we obtain

$$\begin{aligned}
\sup_{x_p\in[a,b]} \Bigg| &\int \frac{\mathbf{1}_{\{[a',b']^c\}}(\lambda)\,\mathrm{d}\left(F_p(\lambda) - W_c(\lambda)\right)}{\prod_{s=1}^{228}\left(\left(x_p-\lambda\right)^2 + s p^{-2\alpha}\right)} \\
&+ \sum_{\lambda_{pj}\in[a',b']} \frac{p^{-1}}{\prod_{s=1}^{228}\left(\left(x_p-\lambda_{pj}\right)^2 + s p^{-2\alpha}\right)}\Bigg| = o_{\text{a.s.}}(1),
\end{aligned} \tag{49}$$

where $\mathbf{1}_{\{[a',b']^c\}}(\lambda)$ is the indicator function equal to unity iff $\lambda \notin [a', b']$.

Now suppose that there exists a subsequence $p_n \to \infty$ such that for each p_n, at least one eigenvalue $\lambda_{p_n j}$ belongs to $[a,b]$. Setting x_{p_n} equal to such an eigenvalue, we see that the sum on the left hand side of (49) is no smaller than $\prod_{s=1}^{228} s^{-1} > 0$ for all p_n. Therefore, at such x_{p_n}, the integral on the left hand side of (49) must be uniformly bounded away from zero over all p_n. But the integral must a.s. converge to zero because the integrand is uniformly bounded, and both F_{p_n} and W_c a.s. weakly converge to W_{c_0} that satisfies $W_{c_0}([a', b']) = 0$. Therefore, with probability one, no eigenvalues λ_{pj} will appear in $[a,b]$ for all sufficiently large p.

5 Johansen's H^* model

If the data generating process is described by Johansen's H^* model (2) rather than (1), the LR statistic for testing the null hypothesis that $\Pi = 0$ still has form (3). However now, λ_{pj}'s equal the eigenvalues of $\tilde{S}_{01}\tilde{S}_{11}^{-1}\tilde{S}_{10}\tilde{S}_{00}^{-1}$, where $\tilde{S}_{ij}$ are defined differently from S_{ij} given in (8). Specifically, they correspond to sample covariance and cross-covariance matrices of the demeaned processes ΔX_t and $\left(X_{t-1}', t\right)'$ (see Johansen (1995, ch. 6.2)). That is, in contrast to (8),

$$\tilde{S}_{01} = \begin{pmatrix} \eta M_l U'\eta', & \eta M_l \tau \end{pmatrix} \text{ and } \tilde{S}_{11} = \begin{pmatrix} \eta U M_l U'\eta' & \eta U M_l \tau \\ \tau' M_l U'\eta' & \tau' M_l \tau \end{pmatrix},$$

while similarly to above, $\tilde{S}_{10} = \tilde{S}_{01}'$ and $\tilde{S}_{00} = \eta M_l \eta'$. Here τ denotes the time trend, $\tau = (1, 2, ..., T+1)'$.

In contrast to matrices S_{01}, S_{11}, and S_{00} given in (8), matrices $\tilde{S}_{01}, \tilde{S}_{11}$, and $\tilde{S}_{00}$ cannot be simultaneously rotated to the form $\varepsilon' W \varepsilon$, where W is a block-diagonal matrix. Therefore, in the case of H^* model, there is no convenient frequency domain reformulation of Johansen's test, and the above analysis will not go through. It is however possible to show that at most one eigenvalue of $\tilde{S}_{01}\tilde{S}_{11}^{-1}\tilde{S}_{10}\tilde{S}_{00}^{-1}$ remains above and separated from b_{0+} and at most one eigenvalue remains below and separated from b_{0-}, asymptotically. Hence, the second largest and smallest eigenvalues of $\tilde{S}_{01}\tilde{S}_{11}^{-1}\tilde{S}_{10}\tilde{S}_{00}^{-1}$ a.s. converge to b_{0+} and b_{0-}.

Recall that the eigenvalues of $S_{01}S_{11}^{-1}S_{10}S_{00}^{-1}$ equal those of P_1P_2, where P_1 and P_2 are projections on the column spaces of $Y \equiv M_l U' M_l \eta'$ and $Z \equiv M_l \eta'$, respectively. Similarly, the eigenvalues of $\tilde{S}_{01}\tilde{S}_{11}^{-1}\tilde{S}_{10}\tilde{S}_{00}^{-1}$ equal those of $\tilde{P}_1 P_2$, where $\tilde{P}_1$ is the projection on the column space of $\tilde{Y} \equiv \begin{pmatrix} M_l U'\eta', & M_l \tau \end{pmatrix}$.

Note that $\tilde{Y}$ has $p+1$ columns whereas Y has p columns. Let us augment Y by a zero column to obtain $\bar{Y} \equiv \begin{pmatrix} M_l U' M_l \eta', & 0 \end{pmatrix}$. Obviously, projections on the columns of Y and $\bar{Y}$ coincide and equal P_1. Further,

$$\tilde{Y} - \bar{Y} = \left(\begin{array}{cc} M_l U' l l' \eta' / (T+1), & M_l \tau \end{array} \right) = \left(\begin{array}{cc} M_l \tau l' \eta' / (T+1), & M_l \tau \end{array} \right), \qquad (50)$$

and matrix $\left(\begin{array}{cc} M_l \tau l' \eta' / (T+1), & M_l \tau \end{array} \right)$ has rank one.

Lemma 18 *Let Y_1 and Y_2 be $n \times m$ matrices and let P_{Y_1} and P_{Y_2} be projections on the spaces spanned by the columns of Y_1 and Y_2, respectively. If* $\operatorname{rank}(Y_1 - Y_2) = r$, *then there exist $n \times r$ matrices y_1 and y_2 such that $P_{Y_1} - P_{Y_2} = P_{y_1} - P_{y_2}$, where P_{y_1} and P_{y_2} are projections on the spaces spanned by the columns of y_1 and y_2, respectively. In particular,* $\operatorname{rank}(P_{Y_1} - P_{Y_2}) \leq 2r$.

Proof: Assume that $Y_1 - Y_2 = ab$, where a is $n \times r$ and $b = \left(\begin{array}{cc} 0, & I_r \end{array} \right)$. This assumption does not lead to loss of generality because P_{Y_1} and P_{Y_2} are invariant with respect to multiplication of Y_1 and Y_2 from the right by arbitrary invertible $m \times m$ matrices. The above form of b can be achieved by such a multiplication. Let us partition Y_1 and Y_2 as $[Y_{11}, Y_{12}]$ and $[Y_{21}, Y_{22}]$, where Y_{12} and Y_{22} are the last r columns of Y_1 and Y_2, respectively. We have

$$Y_{21} = Y_{11} \text{ and } Y_{22} + a = Y_{12}.$$

Denote $I_m - P_{Y_{21}}$ as M_1, where $P_{Y_{21}}$ is the projection on the space spanned by the columns of Y_{21}, and let $y_2 = M_1 Y_{22}$. Note that

$$P_{Y_2} = P_{[Y_{21}, y_2]} = P_{Y_{21}} + P_{y_2},$$

where the second equality holds because Y_{21} is orthogonal to y_2. Similarly, we have

$$P_{Y_1} = P_{Y_{11}} + P_{y_1} = P_{Y_{21}} + P_{y_1},$$

where $y_1 = M_1 Y_{12}$. Therefore, $P_{Y_1} - P_{Y_2} = P_{y_1} - P_{y_2}$.$\square$

Lemma 18 and equality (50) imply that there exist no more than one eigenvalue of $\tilde{S}_{01}\tilde{S}_{11}^{-1}\tilde{S}_{10}\tilde{S}_{00}^{-1}$ that is larger than the largest eigenvalue of $S_{01}S_{11}^{-1}S_{10}S_{00}^{-1}$ and no more than one eigenvalue of $\tilde{S}_{01}\tilde{S}_{11}^{-1}\tilde{S}_{10}\tilde{S}_{00}^{-1}$ that is smaller than the smallest eigenvalue of $S_{01}S_{11}^{-1}S_{10}S_{00}^{-1}$. Indeed, note that the eigenvalues of $\tilde{S}_{01}\tilde{S}_{11}^{-1}\tilde{S}_{10}\tilde{S}_{00}^{-1}$, which equal those of $\tilde{P}_1 P_2$, coincide with the eigenvalues of a symmetric matrix $P_2 \tilde{P}_1 P_2$. Similarly, the eigenvalues of $S_{01}S_{11}^{-1}S_{10}S_{00}^{-1}$ coincide with the eigenvalues of a symmetric matrix $P_2P_1P_2$. By Lemma 18,

$$P_2\tilde{P}_1P_2 - P_2P_1P_2 = P_2P_{y_1}P_2 - P_2P_{y_2}P_2,$$

where P_{y_1} and P_{y_2} are projections on one-dimensional spaces. Hence, our statement concerning eigenvalues of $\tilde{S}_{01}\tilde{S}_{11}^{-1}\tilde{S}_{10}\tilde{S}_{00}^{-1}$ and $S_{01}S_{11}^{-1}S_{10}S_{00}^{-1}$ follows from Weyl's inequalities for eigenvalues of a sum of symmetric matrices (see e.g. Horn and Johnson (1985, Theorem 4.3.1)).

We have conducted a small-scale Monte Carlo study which suggests that, in fact, the largest eigenvalue of $\tilde{S}_{01}\tilde{S}_{11}^{-1}\tilde{S}_{10}\tilde{S}_{00}^{-1}$ converges to b_{0+} similarly to the largest eigenvalue of $S_{01}S_{11}^{-1}S_{10}S_{00}^{-1}$. However, the smallest eigenvalue of $\tilde{S}_{01}\tilde{S}_{11}^{-1}\tilde{S}_{10}\tilde{S}_{00}^{-1}$ is close to zero, whereas in accordance with our theoretical results, the smallest eigenvalue of $S_{01}S_{11}^{-1}S_{10}S_{00}^{-1}$ converges to b_{0-}.

A more formal analysis of the extreme eigenvalues of $\tilde{S}_{01}\tilde{S}_{11}^{-1}\tilde{S}_{10}\tilde{S}_{00}^{-1}$ would amount to studying low-rank perturbations of $S_{01}S_{11}^{-1}S_{10}S_{00}^{-1}$. There exists large literature on the low rank perturbations of classical random matrix ensembles (see e.g. Capitaine and Donati-Martin (2016) and references therein). However, this literature is not directly applicable to $S_{01}S_{11}^{-1}S_{10}S_{00}^{-1}$. We leave analysis of small rank perturbations of such a matrix for future research.

6 Conclusion and discussion

This paper establishes the a.s. convergence of the largest and the smallest eigenvalues of $S_{01}S_{11}^{-1}S_{10}S_{00}^{-1}$ to the upper and lower boundaries of the support of the Wachter distribution W_{c_0}. This complements Onatski and Wang's (2017) result on the a.s. weak convergence of the empirical distribution of the eigenvalues to W_{c_0}. The strategy of our proofs is similar to that of the proof of the convergence of the extreme eigenvalues of the sample covariance matrix in BS98. However, the fact that we have to deal with the product of four dependent stochastic matrices, S_{01}, S_{11}^{-1}, S_{10}, and S_{00}^{-1}, presents non-trivial challenges that we overcome.

Eigenvalues of $S_{01}S_{11}^{-1}S_{10}S_{00}^{-1}$ can be interpreted as squared canonical correlations between demeaned innovations of high dimensional random walk and detrended and demeaned levels of this random walk. Such eigenvalues form the basis for the LR test of no cointegration in high-dimensional vector autoregression of order one. The LR statistic has a singularity at unity, hence Onatski and Wang's (2017) result cannot be used to establish its a.s. convergence.

The result of this paper shows that the singularity can be ignored because none of the eigenvalues of $S_{01}S_{11}^{-1}S_{10}S_{00}^{-1}$ are close to unity asymptotically. Thus, our Corollary 2 establishes the a.s. limit of the LR statistic. We use this result to obtain an analytic formula for a Bartlett-type correction coefficient for the LR test.

We establish Theorem 1 under Gaussianity of the errors η_t of model (1). We need the Gaussianity for two reasons. First, it allows us to reduce the analysis of $S_{01}S_{11}^{-1}S_{10}S_{00}^{-1}$ to that of $C'D^{-1}CA^{-1}$, where C, D, and A have form $\varepsilon' W\varepsilon$ with block-diagonal W, and ε has i.i.d. elements. Second, we use it to derive bounds on the expected value of the inverse of the smallest eigenvalue of A (in SM). In principle, the first reason can be circumvented by simply assuming that the matrix ε of the discrete Fourier transforms of η has i.i.d. (but not necessarily Gaussian) elements. This still leaves the second reason intact. Unfortunately even a seemingly innocuous assumption that the elements of ε are i.i.d. Bernoulli random variables leads to non-invertibility of A with small but positive probability, and hence, to nonexistence of the expected value of the inverse of the smallest eigenvalue of A. We leave removing the Gaussianity assumption as an important topic for future research.

Onatski and Wang (2017) establish the a.s. weak convergence of the empirical distribution of the eigenvalues of $S_{01}S_{11}^{-1}S_{10}S_{00}^{-1}$ to W_{c_0} under more general data generating processes than the one described by (1). Extension of the results of this paper to such more general processes would require analyzing the effect of small rank perturbations on the extreme eigenvalues of $S_{01}S_{11}^{-1}S_{10}S_{00}^{-1}$. As we discuss above, such an analysis is not straightforward and needs a substantial further research effort.

Another important research task is to study the asymptotic fluctuations of the functionals of F_p around their a.s. limits. This would allow one to derive an asymptotic distribution of the LR statistic under the simultaneous asymptotics. We are undertaking such a study as a separate project.

References

[1] Bai, Z. and Silverstein, J. (1998) “No Eigenvalues Outside the Support of the Limiting Spectral Distribution of Large-dimensional Sample Covariance Matrices,” *The Annals of Probability* 26, 316-345.

[2] Bai, Z., Hu, J., Pan, G., and Zhou, W. (2015) "Convergence of the empirical distribution function of Beta matrices," Bernoulli 21, 1538-74.

[3] Banerjee, A., Marcellino, M., and Osbat, C. (2004) "Some cautions on the use of panel methods for integrated series of macroeconomic data." *The Econometrics Journal* 7 (2), 322–340.

[4] Brillinger, D.R. (1981) *Time series: data analysis and theory.* Holden Day, Inc., San Francisco.

[5] Capitaine, M. and Donati-Martin, C. (2016) "Spectrum of deformed random matrices and free probability," *arXiv 1607.05560v*

[6] Davis, G.C. (2003) "The generalized composite commodity theorem: Stronger support in the presence of data limitations," *Review of Economics and Statistics* 85, 476-480.

[7] Engel, C., N.C. Mark, and K.D. West (2015) "Factor Model Forecasts of Exchange Rates," *Econometric Reviews* 34, 32-55.

[8] Golub, G.H. and C. F. Van Loan (1996) *Matrix Computations*, The John Hopkins University Press.

[9] Gonzalo, J. and J-Y Pitarakis (1995) "Comovements in Large Systems," Working Paper 95-38, Statistics and Econometrics Series 10, Universidad Carlos III de Madrid.

[10] Gonzalo, J. and J-Y Pitarakis (1999) "Dimensionality Effect in Cointegration Analysis," Cointegration, Causality, and Forecasting. A Festschrift in Honour of Clive WJ Granger, Oxford University Press, Oxford, pp. 212-229

[11] Ho, M.S., and B.E. Sorensen (1996) "Finding Cointegration Rank in High Dimensional Systems Using the Johansen Test: An Illustration Using Data Based Mote Carlo Simulations," *Review of Economics and Statistics* 78, 726-732.

[12] Horn, R.A. and C.R. Johnson (1985) *Matrix Analysis*, Cambridge University press.

[13] Jensen, J.L. and A.T.A. Wood (1997) "On the non-existence of a Bartlett correction for unit root tests," *Statistics and Probability Letters* 35, 181-187.

[14] Johansen, S. (1988) "Statistical Analysis of Cointegrating Vectors," *Journal of Economic Dynamics and Control* 12, 231-254

[15] Johansen, S. (1991) "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models," *Econometrica* 59, 1551-1580.

[16] Johansen, S. (1995) *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, Oxford University Press Inc., New York.

[17] Johansen, S. (2002) "A small sample correction for the test of cointegrating rank in the vector autoregressive model," *Econometrica* 70, 1929-1961.

[18] Johansen, S., H. Hansen, and S. Fachin (2005) "A simulation study of some functionals of random walk," manuscript available at http://www.math.ku.dk/~sjo/.

[19] Larsson, R. (1998) "Bartlett Corrections for Unit Root Test Statistics," *Journal of Time Series Analysis* 19, 426-238.

[20] Lewbel, A. (1996) "Aggregation without separability: a generalized composite commodity theorem," *American Economic Review* 86, 524-543.

[21] Nielsen, B. (1997). "Bartlett correction of the unit root test in autoregressive models," *Biometrika* 84 , 500–504.

[22] Onatski, A. and C. Wang (2017) "Alternative asymptotics for cointegration tests in large VARs." *manuscript*, University of Cambridge.

[23] Pedroni, P.L., Vogelsang, T.J., Wagner, M., and Westerlund, J. (2015) "Nonparametric rank tests for non-stationary panels," *Journal of Econometrics* 185, 378-391.

[24] Tao, T. and Vu, V. (2011) "Random Matrices: Universality of Local Eigenvalue Statistics," *Acta Mathematica* 206, 127-204.

Supplementary Material for “Extreme canonical correlations and high-dimensional cointegration analysis.”

Alexei Onatski and Chen Wang
Faculty of Economics, University of Cambridge

August 5, 2017

Abstract

This note contains supplementary material for Onatski and Wang (2017a) (OW in what follows). It is lined up with sections in the main text to make it easy to locate the required proofs.

Contents

1 Introduction and the main result

1.1 There is no supplementary material for this section.

2 Bartlett-type correction

2.1 Proof of Theorem OW3

Recall that

$$LR = -(T+1)\sum_{j=1}^{p}\log(1-\lambda_{pj}),$$

where λ_{pj} is the j-th largest squared sample canonical correlation between demeaned vectors ΔX_t and $X_{t-1} - t\hat{\rho}_1$. As explained in OW's Section 3, λ_{pj} can be equivalently interpreted as the j-th largest eigenvalue of $S_{01}S_{11}^{-1}S_{01}'S_{00}^{-1}$, where

$$S_{01} = \eta M_l U' M_l \eta', S_{11} = \eta M_l U M_l U' M_l \eta', S_{00} = \eta M_l \eta'.$$

Therefore, by standard arguments (see e.g. Johansen (1995, Appendix B)), as $T \to \infty$ while p is held fixed, the entries of matrix $(T/p)\, S_{01}S_{11}^{-1}S_{01}'S_{00}^{-1}$ jointly converge in distribution to those of matrix

$$\frac{1}{p}\int_0^1 \left(\mathrm{d}\bar{B}\right)\bar{F}'\left(\int_0^1 \bar{F}\bar{F}'\mathrm{d}u\right)^{-1}\int_0^1 \bar{F}\left(\mathrm{d}\bar{B}\right)', \tag{1}$$

where $\bar{B}$ is a p-dimensional *Brownian bridge* and $\bar{F}$ is its demeaned version. In particular, as $T\to\infty$ while p is held fixed, LR/p converges in distribution to the trace of (1).

Let us denote the eigenvalues of (1) as $\lambda_{0,j}$, and their empirical distribution function (d.f.) as $F_{0,p}(\lambda)$. Note that matrix (1) is a low-rank perturbation of matrix

$$\frac{1}{p}\int_0^1 (\mathrm{d}B)\, F'\left(\int_0^1 FF'\mathrm{d}u\right)^{-1}\int_0^1 F\,(\mathrm{d}B)',$$

where B is a p-dimensional *Brownian motion* and F is its demeaned version. Therefore, by Theorem 4 of Onatski and Wang (2017), $F_{0,p}(\lambda) \overset{\mathrm{P}}{\Rightarrow} F_0(\lambda)$ as $p\to\infty$, that is, $F_{0,p}(\lambda)$ weakly converges in probability to $F_0(\lambda)$, which corresponds to a distribution supported on $[a_-, a_+]$ with

$$a_\pm = \left(1 \pm \sqrt{2}\right)^2, \tag{2}$$

and having density

$$f(\lambda) = \frac{1}{2\pi}\frac{\sqrt{(a_+-\lambda)(\lambda-a_-)}}{\lambda}. \tag{3}$$

Moreover, by Theorem 5 of Onatski and Wang (2017), $\int \lambda \mathrm{d}F_0(\lambda) = 2$.

Since, as $T\to\infty$ while p is held fixed,

$$LR/p^2 \overset{\mathrm{d}}{\to} \frac{1}{p}\sum_{j=1}^{p}\lambda_{0,j} \equiv \int \lambda \mathrm{d}F_{0,p}(\lambda),$$

it remains to show that $\int \lambda\mathrm{d}F_{0,p}(\lambda) \overset{\mathrm{P}}{\to} \int \lambda \mathrm{d}F_0(\lambda)$ as $p\to\infty$. Unfortunately, such a convergence in probability does not follow from $F_{0,p}(\lambda) \overset{\mathrm{P}}{\Rightarrow} F_0(\lambda)$ because $f(\lambda)\equiv\lambda$ is not a bounded function of λ. To circumvent this caveat, we now prove that $\lambda_{0,1} \xrightarrow{\text{P}} a_+$ as $p \to \infty$. Hence, $\int \lambda \mathrm{d}F_{0,p}(\lambda)$ converges in probability to the same limit as $\int f_\delta(\lambda)\mathrm{d}F_{0,p}(\lambda)$, where δ is a fixed positive number and

$$f_\delta(\lambda) = \begin{cases} 0 & \text{if } \lambda < 0 \\ \min(\lambda, a_+ + \delta) & \text{if } \lambda \geq 0 \end{cases}$$

is a bounded and continuous function. On the other hand, $\int f_\delta(\lambda)\mathrm{d}F_{0,p}(\lambda) \xrightarrow{\text{P}} \int f_\delta(\lambda)\mathrm{d}F_0(\lambda) = 2$, where the latter equality holds because $f_\delta(\lambda)$ coincides with $f(\lambda) \equiv \lambda$ on the support of $F_0(\lambda)$.

2.1.1 Convergence of $\lambda_{0,1}$ (the largest eigenvalue of the T-limit of $(T/p)\, S_{01}S_{11}^{-1}S_{01}'S_{00}^{-1}$)

Without loss of generality, assume that η and (1) are defined on the common probability space so that the convergence of $(T/p)\, S_{01}S_{11}^{-1}S_{01}'S_{00}^{-1}$ to (1) is in probability. Lemma OW5 implies that $(T/p)\, S_{01}S_{11}^{-1}S_{01}'S_{00}^{-1}$ equals $(T/p)\, CD^{-1}C'A^{-1}$, where

$$C = \xi\nabla'\xi'/T, D = \xi\nabla\nabla'\xi'/T,\ A = \xi\xi'/T,$$

and ξ is a $p \times T$ matrix with i.i.d. $N(0,1)$ entries.

Let $\gamma < 1/2$ be a small positive number, T_γ be the smallest integer satisfying $p/T_\gamma \leq \gamma$, and let T be so large that $T > T_\gamma$. Further, let

$$C_\gamma = \xi_\gamma\nabla_\gamma'\xi_\gamma'/T_\gamma, D_\gamma = \xi_\gamma\nabla_\gamma\nabla_\gamma'\xi_\gamma'/T_\gamma, \text{ and } A_\gamma = \xi_\gamma\xi_\gamma'/T_\gamma,$$

where ξ_γ is the $p \times T_\gamma$ matrix from the partition $\xi = \left[\xi_\gamma, \xi_{-\gamma}\right]$, and ∇_γ is defined similarly to ∇ with T replaced by T_γ. Finally, let $\lambda_{p\gamma,1}$ be the largest eigenvalue of $C_\gamma D_\gamma^{-1}C_\gamma' A_\gamma^{-1}$. Note that, by Theorem OW1,

$$(T_\gamma/p)\,\lambda_{p\gamma,1} \xrightarrow{\text{a.s.}} a_{\gamma+} \equiv \left(\sqrt{2} - \sqrt{1-\gamma}\right)^{-2} \tag{4}$$

as $p \to \infty$.

We would like to show that for any $\delta > 0$ and all sufficiently large p

$$\Pr\left(|\lambda_{0,1} - a_+| < \delta\right) > 1 - \delta.$$

Suppose this is not so. Then, there exists $\delta > 0$ such that for any p_0, there exists $p > p_0$ yielding

$$\Pr\left(|\lambda_{0,1} - a_+| \geq \delta\right) > \delta. \tag{5}$$

By the triangle inequality,

$$|\lambda_{0,1} - a_+| \leq |\lambda_{0,1} - (T/p)\,\lambda_{p,1}| + |(T/p)\,\lambda_{p,1} - (T_\gamma/p)\,\lambda_{p\gamma,1}| + |(T_\gamma/p)\,\lambda_{p\gamma,1} - a_{\gamma+}| + |a_{\gamma+} - a_+|\,.$$

Choosing γ sufficiently small, we obtain

$$|a_{\gamma+} - a_+| < \delta/4. \tag{6}$$

Further, by (4) for all sufficiently large p, we have

$$\Pr\left(|(T_\gamma/p)\,\lambda_{p\gamma,1} - a_{\gamma+}| \geq \delta/4\right) < \delta/4. \tag{7}$$

Next, since $(T/p)\, CD^{-1}C'A^{-1}$ converges in probability to (1) as $T \to \infty$ while p is held fixed, for any p, we can choose T so large that

$$\Pr\left(|\lambda_{0,1} - (T/p)\,\lambda_{p,1}| \geq \delta/4\right) < \delta/4. \tag{8}$$

Finally, by inequality (114) proven in Section 3.1.1 of the Supplementary Material to Onatski and Wang (2017), for sufficiently small γ, all sufficiently large p, and all $T > \tilde{T}$, where $\tilde{T}$ may depend on p,

$$\Pr\left(|(T/p)\,\lambda_{p,1} - (T_\gamma/p)\,\lambda_{p\gamma,1}| \geq \delta/4\right) < \delta/2. \tag{9}$$

Combining (6-9) with the triangle inequality, we obtain a contradiction to (5), which completes the proof.

3 Setup

3.1 Proof of Lemma OW4 (diagonalization)

Note that $\mathcal{F}/\sqrt{T_1}$, where $T_1 = T+1$, is a unitary matrix with the first column and row equal to $l/\sqrt{T_1}$ and $l'/\sqrt{T_1}$. Therefore,

$$\mathcal{F}M_l\mathcal{F}^*/T_1 = I_{T_1} - e_1e_1' = \text{diag}\left\{0, I_T\right\}.$$

Next,

$$M_lUM_l = U - ll'U/T_1 - Ull'/T_1 + ll'Ull'/T_1^2$$

and thus, the first column of M_lUM_l equals

$$v \equiv \frac{1}{T_1}\tau - l\frac{T_1+1}{2T_1},$$

where $\tau = (1, 2, ..., T_1)'$. We have

$$(\mathcal{F}v)_s = \frac{1}{T_1}\sum_{t=1}^{T_1} te^{-\mathrm{i}\omega_{s-1}(t-1)} - \delta_{s=1}\frac{T_1+1}{2} \text{ with } \omega_s = \frac{2\pi}{T_1}s,$$

which yields $(\mathcal{F}v)_1 = 0$ and

$$\left(1-e^{-\mathrm{i}\omega_{s-1}}\right)(\mathcal{F}v)_s = \frac{1}{T_1}\sum_{t=1}^{T_1} e^{-\mathrm{i}\omega_{s-1}(t-1)} - 1 = -1 \text{ for } s > 1.$$

As is well known (see e.g. Golub and Van Loan (1996, ch. 4.7.7)), any $T_1 \times T_1$ circulant matrix V with the first column v admits the diagonalization $V = \frac{1}{T_1}\mathcal{F}^* \,\text{diag}\,(\mathcal{F}v)\,\mathcal{F}$. Hence,

$$M_lUM_l = \frac{1}{T_1}\mathcal{F}^* \,\text{diag}\left\{0, \hat{\nabla}^*\right\}\mathcal{F}$$

where

$$\hat{\nabla} = \text{diag}\left\{\left(e^{\mathrm{i}\omega_1} - 1\right)^{-1}, ..., \left(e^{\mathrm{i}\omega_T} - 1\right)^{-1}\right\}$$

and

$$\begin{aligned} M_lUM_lU'M_l &= \frac{1}{T_1}\mathcal{F}^* \,\text{diag}\left\{0, \hat{\nabla}^*\right\}\mathcal{F}\frac{1}{T_1}\mathcal{F}^* \,\text{diag}\left\{0, \hat{\nabla}\right\}\mathcal{F} \\ &= \frac{1}{T_1}\mathcal{F}^* \,\text{diag}\left\{0, \hat{\nabla}^*\hat{\nabla}\right\}\mathcal{F}. \end{aligned}$$

3.2 Proof of Lemma OW5 ($CD^{-1}C'A^{-1}$ form of $S_{01}S_{11}^{-1}S_{10}S_{00}^{-1}$)

Let $P_{+1} = \text{diag}\left\{1, P\right\}$, $W_{+1} = \text{diag}\left\{1, W\right\}$, and $\varepsilon_{+1} = \hat{\eta}P_{+1}W_{+1}^*/\sqrt{TT_1}$, where $T_1 = T+1$. Matrix ε can be obtained from ε_{+1} by deleting the first column of ε_{+1}. By the definition of $\hat{\eta}$,

$$\varepsilon = \eta\mathcal{F}^*P_{+1}W_{+1}^*/\sqrt{TT_1} \equiv \eta Q/\sqrt{T}.$$

Since $\mathcal{F}^*/\sqrt{T_1}$ and W_{+1}^* are unitary and P_{+1} is orthogonal, matrix Q is unitary. Moreover, it is orthogonal because it has real-valued entries. This implies that the columns of ε are i.i.d. $N\left(0, I_p/T\right)$. The rest of the lemma follows from the easy to verify fact that $WP'\hat{\nabla}PW^* = \nabla'$.

Table 1: Definitions of matrices, quadratic forms and traces that are used in the derivations below. Notations used in this table suppress the dependence of various quantities, such as M, v, u, w, etc., on z.

$p \times p$ matrices	2×2 matrices	scalars
$M = CD^{-1}C' - zA$	$v_j^{(q)} = \varepsilon'_{(j)} M_j^{-1} \varepsilon_{(j)}$	$v = \frac{1}{T} \operatorname{tr}\left\{M^{-1}\right\}$
$\tilde{M} = C'A^{-1}C - zD$	$u_j^{(q)} = \varepsilon'_{(j)} D_j^{-1} C_j' M_j^{-1} \varepsilon_{(j)}$	$u = \frac{1}{T} \operatorname{tr}\left\{D^{-1}C'M^{-1}\right\}$
$C_j = C - \varepsilon_{(j)} \nabla_j' \varepsilon_{(j)}'$,	$w_j^{(q)} = \varepsilon'_{(j)} D_j^{-1} C_j' M_j^{-1} C_j' D_j^{-1} \varepsilon_{(j)}$	$w = \frac{1}{T} \operatorname{tr}\left\{D^{-1}C'M^{-1}CD^{-1}\right\}$
$D_j = D - r_j^{-1} \varepsilon_{(j)} \varepsilon_{(j)}'$,	$s_j^{(q)} = \varepsilon'_{(j)} D_j^{-1} \varepsilon_{(j)}$	$s = \frac{1}{T} \operatorname{tr}\left\{D^{-1}\right\}$
$A_j = A - \varepsilon_{(j)} \varepsilon_{(j)}'$,	$\tilde{v}_j^{(q)} = \varepsilon'_{(j)} \tilde{M}_j^{-1} \varepsilon_{(j)}$	$\tilde{v} = \frac{1}{T} \operatorname{tr}\left\{\tilde{M}^{-1}\right\}$
$M_j = C_j D_j^{-1} C_j' - zA_j$	$\tilde{u}_j^{(q)} = \varepsilon'_{(j)} A_j^{-1} C_j \tilde{M}_j^{-1} \varepsilon_{(j)}$	$\tilde{u} = \frac{1}{T} \operatorname{tr}\left\{A^{-1}C\tilde{M}^{-1}\right\}$
$\tilde{M}_j = C_j' A_j^{-1} C_j - zD_j$	$\tilde{w}_j^{(q)} = \varepsilon'_{(j)} A_j^{-1} C_j \tilde{M}_j^{-1} C_j A_j^{-1} \varepsilon_{(j)}$	$\tilde{w} = \frac{1}{T} \operatorname{tr}\left\{A^{-1}C\tilde{M}^{-1}C'A^{-1}\right\}$
	$\tilde{s}_j^{(q)} = \varepsilon'_{(j)} A_j^{-1} \varepsilon_{(j)}$	$\tilde{s} = \frac{1}{T} \operatorname{tr}\left\{A^{-1}\right\}$
		$m = \frac{1}{p} \operatorname{tr}\left\{\left(CD^{-1}C'A^{-1} - zI_p\right)^{-1}\right\}$

3.3 Derivation of equations OW11-OW14

A detailed derivation of equations OW11-OW14 can be found in Section 2.1.4 of the Supplementary Material to Onatski and Wang (2017). However, since some equations and definitions from that derivation are used below, we reproduce the derivation here. For the reader's convenience, Table 1 below lists definitions of matrices and scalars used in our proofs.

We will need the following lemma, which is proven in the next section of this note.

Lemma 1 *The following identities hold*

$$u_j^{(q)} = \tilde{u}_j^{(q)\prime}, \ z\tilde{v}_j^{(q)} = w_j^{(q)} - s_j^{(q)}, \ \textit{and } zv_j^{(q)} = \tilde{w}_j^{(q)} - \tilde{s}_j^{(q)}. \tag{10}$$

Similarly,

$$u = \tilde{u}, \ z\tilde{v} = w - s, \ \textit{and } zv = \tilde{w} - \tilde{s}. \tag{11}$$

Derivation of identity (OW11) Applying the Sherman-Morrison-Woodbury (SMW) formula

$$(V + XWY)^{-1} = V^{-1} - V^{-1}X\left(W^{-1} + YV^{-1}X\right)^{-1} YV^{-1}$$

to the right hand side of

$$D^{-1} = \left(D_j + r_j^{-1}\varepsilon_{(j)}\varepsilon_{(j)}'\right)^{-1},$$

we obtain

$$D^{-1} = D_j^{-1} - D_j^{-1}\varepsilon_{(j)}\left(r_j I_2 + s_j\right)^{-1}\varepsilon_{(j)}' D_j^{-1}. \tag{12}$$

Using this and the identity

$$C = C_j + \varepsilon_{(j)}\nabla_j'\varepsilon_{(j)}', \tag{13}$$

we expand $CD^{-1}C'$ in the following form

$$\begin{aligned}
&C_j D_j^{-1} C_j' + \varepsilon_{(j)}\nabla_j'\varepsilon_{(j)}' D_j^{-1} C_j' - C_j D_j^{-1}\varepsilon_{(j)}\left(r_j I_2 + s_j^{(q)}\right)^{-1}\varepsilon_{(j)}' D_j^{-1} C_j' + C_j D_j^{-1}\varepsilon_{(j)}\nabla_j\varepsilon_{(j)}' \\
&-\varepsilon_{(j)}\nabla_j' s_j^{(q)}\left(r_j I_2 + s_j^{(q)}\right)^{-1}\varepsilon_{(j)}' D_j^{-1} C_j' - C_j D_j^{-1}\varepsilon_{(j)}\left(r_j I_2 + s_j^{(q)}\right)^{-1} s_j^{(q)}\nabla_j\varepsilon_{(j)}' + \varepsilon_{(j)}\nabla_j' s_j^{(q)}\nabla_j\varepsilon_{(j)}' \\
&-\varepsilon_{(j)}\nabla_j' s_j^{(q)}\left(r_j I_2 + s_j^{(q)}\right)^{-1} s_j^{(q)}\nabla_j\varepsilon_{(j)}'.
\end{aligned}$$

Simplifying this expression yields

$$\begin{aligned} CD^{-1}C' &= C_jD_j^{-1}C_j' - C_jD_j^{-1}\varepsilon_{(j)}\left(r_jI_2+s_j^{(q)}\right)^{-1}\varepsilon_{(j)}'D_j^{-1}C_j' + \varepsilon_{(j)}\nabla_j' r_j\left(r_jI_2+s_j^{(q)}\right)^{-1}\varepsilon_{(j)}'D_j^{-1}C_j' \\ &\quad + C_jD_j^{-1}\varepsilon_{(j)}\left(r_jI_2+s_j^{(q)}\right)^{-1}r_j\nabla_j\varepsilon_{(j)}' + \varepsilon_{(j)}\nabla_j's_j^{(q)}\left(r_jI_2+s_j^{(q)}\right)^{-1}r_j\nabla_j\varepsilon_{(j)}'. \end{aligned}$$

Since $M = CD^{-1}C' - zA$ and $A = A_j + \varepsilon_{(j)}\varepsilon_{(j)}'$, it follows that

$$M^{-1} = \left(M_j + \alpha_jK_j\alpha_j'\right)^{-1}, \tag{14}$$

where

$$\alpha_j = [\varepsilon_{(j)}, C_jD_j^{-1}\varepsilon_{(j)}]$$

and

$$K_j = \begin{pmatrix} \nabla_j's_j^{(q)}\left(r_jI_2+s_j^{(q)}\right)^{-1}r_j\nabla_j - zI_2 & \nabla_j'r_j\left(r_jI_2+s_j^{(q)}\right)^{-1} \\ \left(r_jI_2+s_j^{(q)}\right)^{-1}r_j\nabla_j & -\left(r_jI_2+s_j^{(q)}\right)^{-1} \end{pmatrix}.$$

Applying SMW formula to the right hand side of (14), we obtain

$$M^{-1} = M_j^{-1} - M_j^{-1}\alpha_j\left(K_j^{-1} + \alpha_j'M_j^{-1}\alpha_j\right)^{-1}\alpha_j'M_j^{-1}. \tag{15}$$

The identity $\nabla_j'\nabla_j = r_j^{-1}I_2$ yields

$$K_j = \begin{pmatrix} \nabla_j' & 0 \\ 0 & I_2 \end{pmatrix}\begin{pmatrix} s_j^{(q)}\left(r_jI_2+s_j^{(q)}\right)^{-1}r_j - zr_jI_2 & \left(r_jI_2+s_j^{(q)}\right)^{-1}r_j \\ \left(r_jI_2+s_j^{(q)}\right)^{-1}r_j & -\left(r_jI_2+s_j^{(q)}\right)^{-1} \end{pmatrix}\begin{pmatrix} \nabla_j & 0 \\ 0 & I_2 \end{pmatrix},$$

which implies that

$$K_j^{-1} = \frac{1}{1-z}\begin{pmatrix} \nabla_j^{-1} & 0 \\ 0 & I_2 \end{pmatrix}\begin{pmatrix} r_j^{-1}I_2 & I_2 \\ I_2 & z\left(r_jI_2+s_j^{(q)}\right) - s_j^{(q)} \end{pmatrix}\begin{pmatrix} \nabla_j'^{-1} & 0 \\ 0 & I_2 \end{pmatrix},$$

and therefore, using $\nabla_j'\nabla_j = r_j^{-1}I_2$ again, we obtain

$$K_j^{-1} = \begin{pmatrix} \frac{1}{1-z}I_2 & \frac{1}{1-z}r_j\nabla_j' \\ \frac{1}{1-z}r_j\nabla_j & \frac{z}{1-z}r_jI_2 - s_j^{(q)} \end{pmatrix}. \tag{16}$$

Further, the definitions of $v_j^{(q)}$, $u_j^{(q)}$ and $w_j^{(q)}$ yield

$$\alpha_j'M_j^{-1}\alpha_j = \begin{pmatrix} v_j^{(q)} & u_j^{(q)\prime} \\ u_j^{(q)} & w_j^{(q)} \end{pmatrix}. \tag{17}$$

Using (16) and (17) in (15), we obtain

$$M^{-1} = M_j^{-1} - M_j^{-1}\alpha_j\Omega_j^{(q)}\alpha_j'M_j^{-1}, \tag{18}$$

where

$$\Omega_j^{(q)} = \begin{pmatrix} \frac{1}{1-z}I_2 + v_j^{(q)} & \frac{1}{1-z}r_j\nabla_j' + u_j^{(q)\prime} \\ \frac{1}{1-z}r_j\nabla_j + u_j^{(q)} & \frac{z}{1-z}r_jI_2 - s_j^{(q)} + w_j^{(q)} \end{pmatrix}^{-1} = \begin{pmatrix} \frac{1}{1-z}I_2 + v_j^{(q)} & \frac{1}{1-z}r_j\nabla_j' + u_j^{(q)\prime} \\ \frac{1}{1-z}r_j\nabla_j + u_j^{(q)} & \frac{z}{1-z}r_jI_2 + z\tilde{v}_j^{(q)} \end{pmatrix}^{-1},$$

and the latter equality holds by Lemma 1.

Equation (18) yields

$$\varepsilon_{(j)}' M^{-1}\varepsilon_{(j)} = v_j^{(q)} - \left[v_j^{(q)}, u_j^{(q)\prime}\right]\Omega_j^{(q)}\left[v_j^{(q)}, u_j^{(q)\prime}\right]'. \tag{19}$$

Note that

$$v_j^{(q)} = \left[v_j^{(q)}, u_j^{(q)\prime}\right]\Omega_j^{(q)}(\Omega_j^{(q)})^{-1}\left[I_2, 0\right]' = \left[v_j^{(q)}, u_j^{(q)\prime}\right]\Omega_j^{(q)}\left(\frac{1}{1-z}\left[I_2, r_j\nabla_j'\right]' + \left[v_j^{(q)}, u_j^{(q)\prime}\right]'\right),$$

and thus, (19) can be rewritten as

$$\begin{aligned}
\varepsilon_{(j)}' M^{-1}\varepsilon_{(j)} &= \frac{1}{1-z}\left[v_j^{(q)}, u_j^{(q)\prime}\right]\Omega_j^{(q)}\left[I_2, r_j\nabla_j'\right]' \\
&= \frac{1}{1-z}\left(\left[\frac{1}{1-z}I_2 + v_j^{(q)}, \frac{1}{1-z}r_j\nabla_j' + u_j^{(q)\prime}\right] - \left[\frac{1}{1-z}I_2, \frac{1}{1-z}r_j\nabla_j'\right]\right)\Omega_j^{(q)}\left[I_2, r_j\nabla_j'\right]' \\
&= \frac{1}{1-z}\left(\left[I_2, 0\right]\left[I_2, r_j\nabla_j'\right]' - \left[\frac{1}{1-z}I_2, \frac{1}{1-z}r_j\nabla_j'\right]\Omega_j^{(q)}\left[I_2, r_j\nabla_j'\right]'\right) \\
&= \frac{1}{1-z}I_2 - \frac{1}{(1-z)^2}\left[I_2, r_j\nabla_j'\right]\Omega_j^{(q)}\left[I_2, r_j\nabla_{2j}'\right]'.
\end{aligned}$$

To summarize, we have the following identity

$$\varepsilon_{(j)}' M^{-1}\varepsilon_{(j)} = \frac{1}{1-z}I_2 - \frac{1}{(1-z)^2}\left[I_2, r_j\nabla_j'\right]\Omega_j^{(q)}\left[I_2, r_j\nabla_j'\right]'. \tag{20}$$

Recall that by definition,

$$m = \frac{1}{p}\operatorname{tr}\left[\left(CD^{-1}C'A^{-1} - zI_p\right)^{-1}\right] = \frac{1}{p}\operatorname{tr}\left[AM^{-1}\right] = \frac{1}{p}\sum_{j=1}^{T/2}\operatorname{tr}\left[\varepsilon_{(j)}' M^{-1}\varepsilon_{(j)}\right].$$

This equation and representation (20) yield identity (OW11)

$$m = \frac{T}{p}\frac{1}{1-z} - \frac{1}{p}\sum_{j=1}^{T/2}\frac{1}{(1-z)^2}\operatorname{tr}\left(\left[I_2, r_j\nabla_j'\right]\Omega_j^{(q)}\left[I_2, r_j\nabla_j'\right]'\right).$$

Derivation of identity (OW12) Since the eigenvalues of $CD^{-1}C'A^{-1}$ coincide with those of $C'A^{-1}CD^{-1}$, we have

$$m = \frac{1}{p}\operatorname{tr}\left[\left(C'A^{-1}CD^{-1} - zI_p\right)^{-1}\right] = \frac{1}{p}\operatorname{tr}\left[D\tilde{M}^{-1}\right] = \frac{1}{p}\sum_{j=1}^{T/2}\operatorname{tr}\left[r_j^{-1}\varepsilon_{(j)}'\tilde{M}^{-1}\varepsilon_{(j)}\right]. \tag{21}$$

Note that matrix $\tilde{M}$ can be obtained from M by swapping A for D and C for C'. Performing such a swap in the above derivations of (20) yields

$$\varepsilon_{(j)}'\tilde{M}^{-1}\varepsilon_{(j)} = \frac{r_j}{1-z}I_2 - \frac{r_j^2}{(1-z)^2}\left[I_2, \nabla_j\right]\tilde{\Omega}_j^{(q)}\left[I_2, \nabla_j\right]', \tag{22}$$

where

$$\tilde{\Omega}_j^{(q)} = \begin{pmatrix} \frac{r_j}{1-z}I_2 + \tilde{v}_j^{(q)} & \frac{r_j}{1-z}\nabla_j + \tilde{u}_j^{(q)\prime} \\ \frac{r_j}{1-z}\nabla_j' + \tilde{u}_j^{(q)} & \frac{z}{1-z}I_2 - \tilde{s}_j^{(q)} + \tilde{w}_j^{(q)} \end{pmatrix}^{-1}$$

(see Table 1 for the definitions of $\tilde{v}_j^{(q)}$, $\tilde{u}_j^{(q)}$, $\tilde{w}_j^{(q)}$, and $\tilde{s}_j^{(q)}$). Lemma 1 implies that

$$\tilde{\Omega}_j^{(q)} = \begin{pmatrix} \frac{r_j}{1-z}I_2 + z^{-1}\left(w_j^{(q)} - s_j^{(q)}\right) & \frac{r_j}{1-z}\nabla_j + u_j^{(q)} \\ \frac{r_j}{1-z}\nabla_j' + u_j^{(q)\prime} & \frac{z}{1-z}I_2 + zv_j^{(q)} \end{pmatrix}^{-1} = \begin{pmatrix} 0 & zI_2 \\ I_2 & 0 \end{pmatrix}\Omega_j^{(q)}\begin{pmatrix} 0 & z^{-1}I_2 \\ I_2 & 0 \end{pmatrix},$$

so that (22) yields

$$\varepsilon_{(j)}'\tilde{M}^{-1}\varepsilon_{(j)} = \frac{r_j}{1-z}I_2 - \frac{zr_j^2}{(1-z)^2}\left[z^{-1}\nabla_j, I_2\right]\Omega_j^{(q)}\left[z^{-1}\nabla_j, I_2\right]'. \tag{23}$$

Combining this with (21) gives us

$$m = \frac{T}{p}\frac{1}{1-z} - \frac{1}{p}\sum_{j=1}^{T/2}\mathrm{tr}\left[\frac{zr_j}{(1-z)^2}\left[z^{-1}\nabla_j, I_2\right]\Omega_j^{(q)}\left[z^{-1}\nabla_j, I_2\right]'\right].$$

Further, since $r_j\nabla_j\nabla_j' = I_2$, we have

$$\begin{aligned} &\frac{zr_j}{(1-z)^2}\mathrm{tr}\left[\left[z^{-1}\nabla_j, I_2\right]\Omega_j^{(q)}\left[z^{-1}\nabla_j, I_2\right]'\right] = \frac{zr_j}{(1-z)^2}\mathrm{tr}\left[r_j\nabla_j\nabla_j'\left[z^{-1}\nabla_j, I_2\right]\Omega_j^{(q)}\left[z^{-1}\nabla_j, I_2\right]'\right] \\ = &\frac{zr_j}{(1-z)^2}\mathrm{tr}\left[r_j\left[z^{-1}\nabla_j'\nabla_j, \nabla_j'\right]\Omega_j^{(q)}\left[z^{-1}\nabla_j\nabla_j', \nabla_j'\right]'\right] = \frac{z^{-1}}{(1-z)^2}\mathrm{tr}\left(\left[I_2, r_j z\nabla_j'\right]\Omega_j^{(q)}\left[I_2, zr_j\nabla_j'\right]'\right), \end{aligned}$$

and therefore,

$$m = \frac{T}{p}\frac{1}{1-z} - \frac{1}{p}\sum_{j=1}^{T/2}\frac{z^{-1}}{(1-z)^2}\mathrm{tr}\left(\left[I_2, r_j z\nabla_j'\right]\Omega_j^{(q)}\left[I_2, zr_j\nabla_j'\right]'\right),$$

which is equivalent to identity (OW12),

$$\frac{T}{p} + zm = \frac{T}{p}\frac{1}{1-z} - \frac{1}{p}\sum_{j=1}^{T/2}\frac{1}{(1-z)^2}\mathrm{tr}\left(\left[I_2, r_j z\nabla_j'\right]\Omega_j^{(q)}\left[I_2, zr_j\nabla_j'\right]'\right).$$

Derivation of identity (OW13) Multiplying both sides of the identity

$$MA^{-1} = CD^{-1}C'A^{-1} - zI_p$$

by AM^{-1}, taking trace, dividing by p, and rearranging yields

$$1 + zm = \frac{1}{p}\sum_{j=1}^{T/2}\mathrm{tr}\left[\nabla_j'\varepsilon_{(j)}'D^{-1}C'M^{-1}\varepsilon_{(j)}\right]. \tag{24}$$

Equations (12), (13), and (18) imply that

$$\begin{aligned} D^{-1}C'M^{-1} &= \left(D_j^{-1} - D_j^{-1}\varepsilon_{(j)}\left(r_jI_2 + s_j^{(q)}\right)^{-1}\varepsilon_{(j)}'D_j^{-1}\right)\left(C_j' + \varepsilon_{(j)}\nabla_j\varepsilon_{(j)}'\right) \\ &\times\left(M_j^{-1} - M_j^{-1}\alpha_j\Omega_j^{(q)}\alpha_j'M_j^{-1}\right). \end{aligned}$$

Opening up brackets, we obtain

$$\begin{aligned} &D^{-1}C'M^{-1} \\ = \ &D_j^{-1}C_j'M_j^{-1} - D_j^{-1}\varepsilon_{(j)}\left(r_jI_2 + s_j^{(q)}\right)^{-1}\varepsilon_{(j)}'D_j^{-1}C_j'M_j^{-1} + D_j^{-1}\varepsilon_{(j)}\nabla_j\varepsilon_{(j)}'M_j^{-1} \\ &-D_j^{-1}C_j'M_j^{-1}\alpha_j\Omega_j^{(q)}\alpha_j'M_j^{-1} - D_j^{-1}\varepsilon_{(j)}\left(r_jI_2 + s_j^{(q)}\right)^{-1}\varepsilon_{(j)}'D_j^{-1}\varepsilon_{(j)}\nabla_j\varepsilon_{(j)}'M_j^{-1} \\ &+D_j^{-1}\varepsilon_{(j)}\left(r_jI_2 + s_j^{(q)}\right)^{-1}\varepsilon_{(j)}'D_j^{-1}C_j'M_j^{-1}\alpha_j\Omega_j^{(q)}\alpha_j'M_j^{-1} - D_j^{-1}\varepsilon_{(j)}\nabla_j\varepsilon_{(j)}'M_j^{-1}\alpha_j\Omega_j^{(q)}\alpha_j'M_j^{-1} \\ &+D_j^{-1}\varepsilon_{(j)}\left(r_jI_2 + s_j^{(q)}\right)^{-1}\varepsilon_{(j)}'D_j^{-1}\varepsilon_{(j)}\nabla_j\varepsilon_{(j)}'M_j^{-1}\alpha_j\Omega_j^{(q)}\alpha_j'M_j^{-1}. \end{aligned}$$

Multiplying from the left by $\varepsilon'_{(j)}$ and from the right by $\varepsilon_{(j)}$, and using the definitions of $u_j^{(q)}, v_j^{(q)}, s_j^{(q)}$, and $w_j^{(q)}$, we obtain

$$\begin{aligned}
&\varepsilon'_{(j)} D^{-1} C' M^{-1} \varepsilon_{(j)} \\
= \;& u_j^{(q)} - s_j^{(q)} \left(r_j I_2 + s_j^{(q)}\right)^{-1} u_j^{(q)} + s_j^{(q)} \nabla_j v_j^{(q)} - \left[u_j^{(q)}, w_j^{(q)}\right] \Omega_j^{(q)} \left[v_j^{(q)}, u_j^{(q)\prime}\right]' \\
&- s_j^{(q)} \left(r_j I_2 + s_j^{(q)}\right)^{-1} s_j^{(q)} \nabla_j v_j^{(q)} + s_j^{(q)} \left(r_j I_2 + s_j^{(q)}\right)^{-1} \left[u_j^{(q)}, w_j^{(q)}\right] \Omega_j^{(q)} \left[v_j^{(q)}, u_j^{(q)\prime}\right]' \\
&- s_j^{(q)} \nabla_j \left[v_j^{(q)}, u_j^{(q)\prime}\right] \Omega_j^{(q)} \left[v_j^{(q)}, u_j^{(q)\prime}\right]' + s_j^{(q)} \left(r_j I_2 + s_j^{(q)}\right)^{-1} s_j^{(q)} \nabla_j \left[v_j^{(q)}, u_j^{(q)\prime}\right] \Omega_j^{(q)} \left[v_j^{(q)}, u_j^{(q)\prime}\right]'.
\end{aligned}$$

Rearranging terms and simplifying gives us

$$\begin{aligned}
\varepsilon'_{(j)} D^{-1} C' M^{-1} \varepsilon_{(j)} \;=\;& r_j \left(r_j I_2 + s_j^{(q)}\right)^{-1} s_j^{(q)} \nabla_j \left(v_j^{(q)} - \left[v_j^{(q)}, u_j^{(q)\prime}\right] \Omega_j^{(q)} \left[v_j^{(q)}, u_j^{(q)\prime}\right]'\right) \\
&+ r_j \left(r_j I_2 + s_j^{(q)}\right)^{-1} \left(u_j^{(q)} - \left[u_j^{(q)}, w_j^{(q)}\right] \Omega_j^{(q)} \left[v_j^{(q)}, u_j^{(q)\prime}\right]'\right).
\end{aligned} \tag{25}$$

As follows from (19) and (20)

$$v_j^{(q)} - \left[v_j^{(q)}, u_j^{(q)\prime}\right] \Omega_j^{(q)} \left[v_j^{(q)}, u_j^{(q)\prime}\right]' = \frac{1}{1-z} I_2 - \frac{1}{(1-z)^2} \left[I_2, r_j \nabla'_j\right] \Omega_j^{(q)} \left[I_2, r_j \nabla'_j\right]'. \tag{26}$$

Further,

$$\begin{aligned}
u_j^{(q)} - \left[u_j^{(q)}, w_j^{(q)}\right] \Omega_j^{(q)} \left[v_j^{(q)}, u_j^{(q)\prime}\right]' \;=\;& \left[u_j^{(q)}, w_j^{(q)}\right] \Omega_j^{(q)} (\Omega_j^{(q)})^{-1} \left[I_2, 0\right]' - \left[u_j^{(q)}, w_j^{(q)}\right] \Omega_j^{(q)} \left[v_j^{(q)}, u_j^{(q)\prime}\right]' \\
=\;& \frac{1}{1-z} \left[u_j^{(q)}, w_j^{(q)}\right] \Omega_j^{(q)} \left[I_2, r_j \nabla'_j\right]'
\end{aligned}$$

Note that

$$\begin{aligned}
&\frac{1}{1-z} \left[u_j^{(q)}, w_j^{(q)}\right] \Omega_j^{(q)} \left[I_2, r_j \nabla'_j\right]' \\
=\;& \frac{1}{1-z} \left(\left[\frac{r_j}{1-z} \nabla_j + u_j^{(q)}, \frac{r_j z}{1-z} I_2 + w_j^{(q)} - s_j^{(q)}\right] - \left[\frac{r_j}{1-z} \nabla_j, \frac{r_j z}{1-z} I_2 - s_j^{(q)}\right]\right) \Omega_j^{(q)} \left[I_2, r_j \nabla'_j\right]' \\
=\;& \frac{1}{1-z} \left(\left[0, I_2\right] \left[I_2, r_j \nabla'_j\right]' - \frac{r_j z}{1-z} \left[z^{-1} \nabla_j, I_2\right] \Omega_j^{(q)} \left[I_2, r_j \nabla'_j\right]' + \left[0, s_j^{(q)}\right] \Omega_j^{(q)} \left[I_2, r_j \nabla'_j\right]'\right) \\
=\;& \frac{1}{1-z} r_j \nabla_j - \frac{r_j z}{(1-z)^2} \left[z^{-1} \nabla_j, I_2\right] \Omega_j^{(q)} \left[I_2, r_j \nabla'_j\right]' + \frac{1}{1-z} \left[0, s_j^{(q)}\right] \Omega_j^{(q)} \left[I_2, r_j \nabla'_j\right]'.
\end{aligned}$$

Therefore,

$$\begin{aligned}
u_j^{(q)} - \left[u_j^{(q)}, w_j^{(q)}\right] \Omega_j^{(q)} \left[v_j^{(q)}, u_j^{(q)\prime}\right]' \;=\;& \frac{1}{1-z} r_j \nabla_j - \frac{r_j z}{(1-z)^2} \left[z^{-1} \nabla_j, I_2\right] \Omega_j^{(q)} \left[I_2, r_j \nabla'_j\right]' \\
&+ \frac{1}{1-z} \left[0, s_j^{(q)}\right] \Omega_j^{(q)} \left[I_2, r_j \nabla'_j\right]'.
\end{aligned}$$

Using this and (26) in (25), we obtain

$$
\begin{aligned}
& \varepsilon_{(j)}' D^{-1} C' M^{-1} \varepsilon_{(j)} \\
= \; & r_j \left(r_j I_2 + s_j^{(q)}\right)^{-1} s_j^{(q)} \nabla_j \left(\frac{1}{1-z} I_2 - \frac{1}{(1-z)^2} \left[I_2, r_j \nabla_j'\right] \Omega_j^{(q)} \left[I_2, r_j \nabla_j'\right]' \right) + r_j \left(r_j I_2 + s_j^{(q)}\right)^{-1} \\
& \times \left(\frac{1}{1-z} r_j \nabla_j - \frac{r_j z}{(1-z)^2} \left[z^{-1}\nabla_j, I_2\right] \Omega_j^{(q)} \left[I_2, r_j \nabla_j'\right]' + \frac{1}{1-z} \left[0, s_j^{(q)}\right] \Omega_j^{(q)} \left[I_2, r_j \nabla_j'\right]' \right) \\
= \; & \frac{1}{1-z} r_j \nabla_j - \frac{r_j \left(r_j I_2 + s_j^{(q)}\right)^{-1}}{(1-z)^2} \left[s_j^{(q)} \nabla_j + r_j \nabla_j, z s_j^{(q)} + z r_j I_2\right] \Omega_j^{(q)} \left[I_2, r_j \nabla_j'\right]' \\
= \; & \frac{r_j}{1-z} \nabla_j - \frac{z r_j}{(1-z)^2} \left[z^{-1}\nabla_j, I_2\right] \Omega_j^{(q)} \left[I_2, r_j \nabla_j'\right]',
\end{aligned}
$$

that is,

$$
\varepsilon_{(j)}' D^{-1} C' M^{-1} \varepsilon_{(j)} = \frac{r_j}{1-z} \nabla_j - \frac{z r_j}{(1-z)^2} \left[z^{-1}\nabla_j, I_2\right] \Omega_j^{(q)} \left[I_2, r_j \nabla_j'\right]'. \tag{27}
$$

This identity together with (24) yield

$$
\begin{aligned}
1 + zm \; & = \; \frac{1}{p} \sum_{j=1}^{T/2} \operatorname{tr} \left[\nabla_j' \left(\frac{r_j}{1-z} \nabla_j - \frac{z r_j}{(1-z)^2} \left[z^{-1}\nabla_j, I_2\right] \Omega_j^{(q)} \left[I_2, r_j \nabla_j'\right]' \right) \right] \\
& = \; \frac{1}{p} \sum_{j=1}^{T/2} \operatorname{tr} \left[\left(\frac{1}{1-z} I_2 - \frac{1}{(1-z)^2} \left[I_2, r_j z \nabla_j'\right] \Omega_j^{(q)} \left[I_2, r_j \nabla_j'\right]' \right) \right],
\end{aligned}
$$

which is equivalent to identity (OW13),

$$
1 + zm = \frac{T}{p} \frac{1}{1-z} - \frac{1}{p} \sum_{j=1}^{T/2} \frac{1}{(1-z)^2} \operatorname{tr} \left(\left[I_2, r_j z \nabla_j'\right] \Omega_j^{(q)} \left[I_2, r_j \nabla_j'\right]' \right).
$$

Derivation of identity (OW14) An obvious identity

$$
\frac{1}{p} \operatorname{tr} \left[C' M^{-1}\right] = \frac{1}{p} \operatorname{tr} \left[D D^{-1} C' M^{-1}\right]
$$

and representations $C' = \sum_{j=1}^{T/2} \varepsilon_{(j)} \nabla_j \varepsilon_{(j)}'$ and $D = \sum_{j=1}^{T/2} r_j^{-1} \varepsilon_{(j)} \varepsilon_{(j)}'$ yield

$$
\frac{1}{p} \sum_{j=1}^{T/2} \operatorname{tr} \left[\nabla_j \varepsilon_{(j)}' M^{-1} \varepsilon_{(j)}\right] = \frac{1}{p} \sum_{j=1}^{T/2} \operatorname{tr} \left[r_j^{-1} \varepsilon_{(j)}' D^{-1} C' M^{-1} \varepsilon_{(j)}\right].
$$

Using (27) and (20) in this equation, we obtain

$$
\begin{aligned}
& \frac{1}{p} \sum_{j=1}^{T/2} \operatorname{tr} \left[\nabla_j \left(\frac{1}{1-z} I_2 - \frac{1}{(1-z)^2} \left[I_2, r_j \nabla_j'\right] \Omega_j^{(q)} \left[I_2, r_j \nabla_j'\right]' \right) \right] \\
= \; & \frac{1}{p} \sum_{j=1}^{T/2} \operatorname{tr} \left[r_j^{-1} \left(\frac{r_j}{1-z} \nabla_j - \frac{z r_j}{(1-z)^2} \left[z^{-1}\nabla_j, I_2\right] \Omega_j^{(q)} \left[I_2, r_j \nabla_j'\right]' \right) \right].
\end{aligned}
$$

Equivalently,

$$\begin{aligned} 0 &= \frac{1}{p}\sum_{j=1}^{T/2}\operatorname{tr}\left[\left(\frac{1}{(1-z)^2}\left[\nabla_j, I_2\right]\Omega_j^{(q)}\left[I_2, r_j\nabla_j'\right]' - \frac{1}{(1-z)^2}\left[\nabla_j, zI_2\right]\Omega_j^{(q)}\left[I_2, r_j\nabla_j'\right]'\right)\right] \\ &= \frac{1}{p}\sum_{j=1}^{T/2}\frac{1}{1-z}\operatorname{tr}\left(\left[0, I_2\right]\Omega_j^{(q)}\left[I_2, r_j\nabla_j'\right]'\right), \end{aligned}$$

which is the same as identity (OW14).

3.3.1 Proof of Lemma 1 (links between variables with and without tilde)

The identity $u_j^{(q)} = \tilde{u}_j^{(q)\prime}$ is established by the following sequence of equalities

$$\begin{aligned} u_j^{(q)} &= \varepsilon_{(j)}' D_j^{-1} C_j' M_j^{-1} \varepsilon_{(j)} = \varepsilon_{(j)}' D_j^{-1} C_j' \left(C_j D_j^{-1} C_j' - zA_j\right)^{-1} \varepsilon_{(j)} \\ &= \varepsilon_{(j)}' \left(C_j - zA_j\left(C_j'\right)^{-1} D_j\right)^{-1} \varepsilon_{(j)} = \left(\varepsilon_{(j)}'\left(C_j' - zD_j\left(C_j\right)^{-1} A_j\right)^{-1}\varepsilon_{(j)}\right)' \\ &= \left(\varepsilon_{(j)}' A_j^{-1} C_j\left(C_j' A_j^{-1} C_j - zD_j\right)^{-1}\varepsilon_{(j)}\right)' = \left(\varepsilon_{(j)}' A_j^{-1} C_j \tilde{M}_j^{-1}\varepsilon_{(j)}\right)' = \tilde{u}_j^{(q)\prime}. \end{aligned}$$

The relationship $z\tilde{v}_j^{(q)} = w_j^{(q)} - s_j^{(q)}$ is obtained as follows

$$\begin{aligned} z\tilde{v}_j^{(q)} + s_j^{(q)} &= \varepsilon_{(j)}'\left(z\tilde{M}_j^{-1} + D_j^{-1}\right)\varepsilon_{(j)} = \varepsilon_{(j)}' D_j^{-1}\left(zI_p\left(C_j' A_j^{-1} C_j D_j^{-1} - zI_p\right)^{-1} + I_p\right)\varepsilon_{(j)} \\ &= \varepsilon_{(j)}' D_j^{-1}\left(-I_p + C_j' A_j^{-1} C_j D_j^{-1}\left(C_j' A_j^{-1} C_j D_j^{-1} - zI_p\right)^{-1} + I_p\right)\varepsilon_{(j)} \\ &= \varepsilon_{(j)}' D_j^{-1} C_j'\left(C_j' - zD_j C_j^{-1} A_j\right)^{-1}\varepsilon_{(j)} = \varepsilon_{(j)}' D_j^{-1} C_j'\left(D_j^{-1} C_j' - zC_j^{-1} A_j\right)^{-1} D_j^{-1}\varepsilon_{(j)} \\ &= \varepsilon_{(j)}' D_j^{-1} C_j'\left(C_j D_j^{-1} C_j' - zA_j\right)^{-1} C_j D_j^{-1}\varepsilon_{(j)} = w_j^{(q)}. \end{aligned}$$

The relationship $zv_j^{(q)} = \tilde{w}_j^{(q)} - \tilde{s}_j^{(q)}$ is obtained as follows

$$\begin{aligned} zv_j^{(q)} + \tilde{s}_j^{(q)} &= \varepsilon_{(j)}'\left(zM_j^{-1} + A_j^{-1}\right)\varepsilon_{(j)} = \varepsilon_{(j)}' A_j^{-1}\left(zI_p\left(C_j D_j^{-1} C_j' A_j^{-1} - zI_p\right)^{-1} + I_p\right)\varepsilon_{(j)} \\ &= \varepsilon_{(j)}' A_j^{-1}\left(-I_p + C_j D_j^{-1} C_j' A_j^{-1}\left(C_j D_j^{-1} C_j' A_j^{-1} - zI_p\right)^{-1} + I_p\right)\varepsilon_{(j)} \\ &= \varepsilon_{(j)}' A_j^{-1} C_j\left(C_j - zA_j C_j'^{-1} D_j\right)^{-1}\varepsilon_{(j)} = \varepsilon_{(j)}' A_j^{-1} C_j\left(A_j^{-1} C_j - zC_j'^{-1} D_j\right)^{-1} A_j^{-1}\varepsilon_{(j)} \\ &= \varepsilon_{(j)}' A_j^{-1} C_j\left(C_j' A_j^{-1} C_j - zD_j\right)^{-1} C_j' A_j^{-1}\varepsilon_{(j)} = \tilde{w}_j^{(q)}. \end{aligned}$$

Identities (11) are established similarly. The only differences are that the matrices involved are not indexed by j, and instead of the quadratic forms in the columns of $\varepsilon_{(j)}$ we work with traces.

4 Proof of Theorem OW1

4.1 Outline of the proof

4.1.1 There is no supplementary material for this section of OW.

4.2 Step 1: Speed of convergence of $\mathbb{E}F_p\left([a,b]\right)$

4.2.1 Rough bounds. Proof of Lemma OW7 (bound on $||\Omega_j^{-1} - (\Omega_j^{(q)})^{-1}||$)

Assume that for all p and $T(p)$

$$c \equiv p/T(p) \in \left[c_0/2, 1/2\right). \tag{28}$$

There is no loss of generality in this assumption because for any sequence of pairs $p, T(p)$, such that $p, T \to_{c_0} \infty$ with $c_0 < 1/2$, (28) holds for *all* sufficiently large p and $T(p)$. Here and below, notation $p, T \to_{c_0} \infty$ is an abbreviation for $p, T(p) \to \infty$ so that $p/T(p) \to c_0$.

Let $\mu_{\min}$ be the minimum of the smallest eigenvalues of A_j, $j = 1, ..., T/2$, and let $\mu_{\max,0}$ be the maximum eigenvalue of A. Further, let $\underline{\mu}$ and $\bar{\mu}$ be positive numbers that are strictly less than $\left(1 - \sqrt{c_0/2}\right)^2$ and strictly larger than $\left(1 + \sqrt{1/2}\right)^2$, respectively. Consider the event

$$\mathcal{E}_0 = \left\{\underline{\mu} \leq \mu_{\min} \text{ and } \mu_{\max,0} \leq \bar{\mu}\right\}. \tag{29}$$

By Theorem II.13 of Davidson and Szarek (2001), the probability of the complementary event, $\mathcal{E}_0^c$, is exponentially small in p. Hence, event $\mathcal{E}_0$ holds w.ow.p.

By definition and by Lemma 1,

$$(\Omega_j^{(q)})^{-1} - \Omega_j^{-1} = \begin{pmatrix} v_j^{(q)} - vI_2 & u_j^{(q)\prime} - uI_2 \\ u_j^{(q)} - uI_2 & w_j^{(q)} - s_j^{(q)} - (w-s)\,I_2 \end{pmatrix}.$$

Therefore,

$$\left\|\Omega_j^{-1} - (\Omega_j^{(q)})^{-1}\right\| \leq \max\left\{\left\|v_j^{(q)} - vI_2\right\|, \left\|w_j^{(q)} - s_j^{(q)} - (w-s)\,I_2\right\|\right\} + \left\|u_j^{(q)} - uI_2\right\|,$$

and it is sufficient to establish bounds on the norms appearing on the right hand side of the above inequality. Here we establish such a bound only for $\left\|v_j^{(q)} - vI_2\right\|$. The other bounds can be obtained similarly.

By definition, the upper left element of the 2×2 matrix $v_j^{(q)} - vI_2$ equals

$$\varepsilon_{2j-1}' M_j^{-1} \varepsilon_{2j-1} - \frac{1}{T}\operatorname{tr} M^{-1} = V_{1j} + V_{2j},$$

where

$$V_{1j} = \varepsilon_{2j-1}' M_j^{-1} \varepsilon_{2j-1} - \frac{1}{T}\operatorname{tr} M_j^{-1} \text{ and } V_{2j} = \frac{1}{T}\operatorname{tr}\left(M_j^{-1} - M^{-1}\right).$$

The following lemma is a simple consequence of Lemma 2.7 in Bai and Silverstein (1998).

Lemma 2 *Let Ω be a $p \times p$ deterministic complex matrix, and $\xi \sim N_p(0, I_p/T)$. Then, for any $\rho \geq 2$*

$$\mathbb{E}\left|\xi'\Omega\xi - \operatorname{tr}\Omega/T\right|^\rho \leq C\left\|\Omega\right\|^\rho p^{\rho/2}T^{-\rho},$$

where C depends only on ρ.

In what follows, we will use C to denote a constant whose value may change from one appearance to another. By Lemma 2 and Markov's inequality, for any $\rho \geq 2$, we have

$$\Pr\left(|V_{1j}| > Cp^{-\gamma}y_p^d \mid ||M_j^{-1}|| \leq y_p^{-1}\underline{\mu}^{-1}\right) \leq Cc^\rho p^{-\rho/2}\left(Cp^{-\gamma}\underline{\mu}y_p^{d+1}\right)^{-\rho},$$

where $\Pr(\cdot \mid \cdot)$ denotes conditional probability. This inequality and our assumption that $y_p = y_0 p^{-\alpha}$ yield

$$\Pr\left(|V_{1j}| > Cp^{-\gamma}y_p^d \text{ and } ||M_j^{-1}|| \leq y_p^{-1}\underline{\mu}^{-1}\right) \leq Cp^{-\rho(1/2-\gamma-\alpha(d+1))}, \tag{30}$$

where C on the right hand side of (30) depends on ρ, d, $\underline{\mu}$, c, and y_0, but not on p.

Lemma 3 *Let $\mu_{\min,j}, \mu_{\max,j}$ and $\mu_{\min,0}, \mu_{\max,0}$ be the smallest and largest eigenvalues of A_j and of A, respectively. Then,*

$$\begin{aligned} \left\|M_j^{-1}\right\| &\leq 1/\left(y_p\mu_{\min,j}\right), \ \left\|D_j^{-1}\right\| \leq 4/\mu_{\min,j}, \ \left\|D_j^{-1}C_j'\right\|^2 \leq 4\mu_{\max,j}/\mu_{\min,j}, \\ \left\|M^{-1}\right\| &\leq 1/\left(y_p\mu_{\min,0}\right), \ \left\|D^{-1}\right\| \leq 4/\mu_{\min,0} \textit{ and } \left\|D^{-1}C'\right\|^2 \leq 4\mu_{\max,0}/\mu_{\min,0}. \end{aligned}$$

Further,

$$\left|\operatorname{tr}\left(M_j^{-1}-M^{-1}\right)\right| \leq 8/\left(y_p\mu_{\min,j}\right), \quad \left|\operatorname{tr}\left(D_j^{-1}C_j'M_j^{-1}-D^{-1}C'M^{-1}\right)\right| \leq 32\mu_{\max,0}^{1/2}/\left(y_p\mu_{\min,j}^{3/2}\right)$$

and

$$\left|\operatorname{tr}\left(D_j^{-1}C_j'M_j^{-1}C_jD_j^{-1}-D^{-1}C'M^{-1}CD^{-1}\right)\right| \leq 96\mu_{\max,0}/\left(y_p\mu_{\min,j}^2\right).$$

This lemma is equivalent to Lemma 13 from the Supplementary Material to Onatski and Wang (2017). For the reader's convenience, we provide its proof in the next section of this note. By Lemma 3,

$$\Pr\left(||M_j^{-1}|| > y_p^{-1}\underline{\mu}^{-1}\right) \leq \Pr\left(\mu_{\min,j} < \underline{\mu}\right) \leq \Pr\left(\mathcal{E}_0^c\right). \tag{31}$$

Combining (30) and (31), we obtain

$$\Pr\left(|V_{1j}| > Cp^{-\gamma}y_p^d\right) \leq Cp^{-\rho(1/2-\gamma-\alpha(d+1))} + \Pr\left(\mathcal{E}_0^c\right).$$

Since $T \equiv T(p)$ is proportional to p, this yields

$$\Pr\left(\max_{j=1,\dots,T/2}|V_{1j}| > Cp^{-\gamma}y_p^d\right) \leq Cp^{-\rho(1/2-\gamma-\alpha(d+1))+1} + Cp\Pr\left(\mathcal{E}_0^c\right).$$

Furthermore, since $\mathcal{E}_0$ holds w.ow.p. and since $\rho \geq 2$ can be chosen as large as we would like it to be, inequality

$$\max_{j=1,\dots,T/2}|V_{1j}| \leq Cp^{-\gamma}y_p^d \text{ holds w.ow.p.}$$

as long as $0 \leq \alpha < \alpha_{\gamma d}$ with $\alpha_{\gamma d} \equiv (1/2-\gamma)/(1+d)$.

Next, by Lemma 3, $|V_{2j}| \leq 8y_p^{-1}\mu_{\min,j}^{-1}T^{-1}$ so that

$$\Pr\left(|V_{2j}| > Cp^{-\gamma}y_p^d\right) \leq \Pr\left(\mu_{\min,j} < 8cp^{-1}C^{-1}p^{\gamma}y_p^{-1-d}\right).$$

The latter probability is no larger than $\Pr\left(\mathcal{E}_0^c\right)$ for all p such that $8C^{-1}cp^{-1+\gamma+\alpha(1+d)} < \underline{\mu}$. Hence,

$$\max_{j=1,\dots,T/2}|V_{2j}| \leq Cp^{-\gamma}y_p^d \text{ holds w.ow.p. and}$$

$$\max_{j=1,\dots,T/2}\left|\varepsilon_{2j-1}'M_j^{-1}\varepsilon_{2j-1} - \frac{1}{T}\operatorname{tr}M^{-1}\right| \leq Cp^{-\gamma}y_p^d \text{ holds w.ow.p.}$$

For the lower right element of $v_j^{(q)} - vI_2$, an inequality similar to that in the above display, can be established by replacing ε_{2j-1} by ε_{2j}. The upper right element of $v_j^{(q)} - vI_2$ equals

$$\varepsilon_{2j-1}'M_j^{-1}\varepsilon_{2j} = \frac{1}{2}\left(\varepsilon_{2j-1}',\varepsilon_{2j}'\right)\begin{pmatrix} 0 & M_j^{-1} \\ M_j^{-1} & 0 \end{pmatrix}\begin{pmatrix} \varepsilon_{2j-1} \\ \varepsilon_{2j} \end{pmatrix},$$

and arguments similar to those used in the above analysis of V_{1j} lead to the conclusion that

$$\max_{j=1,\dots,T/2}\left|\varepsilon_{2j-1}'M_j^{-1}\varepsilon_{2j}\right| \leq Cp^{-\gamma}y_p^d \text{ holds w.ow.p.}$$

Since for any $C > 0$, the maximums over $j = 1, ..., T/2$ of the absolute values of all elements of $v_j^{(q)} - vI_2$ are bounded by $Cp^{-\gamma}y_p^d$ w.ow.p.,

$$\max_{j=1,\dots,T/2}\left\|v_j^{(q)} - vI_2\right\| \leq Cp^{-\gamma}y_p^d \text{ holds w.ow.p.}$$

4.2.2 Rough bounds. Proof of Lemma 3 (bounds on $\left\|M^{-1}\right\|, \left\|M_j^{-1}\right\|$, etc.)

By definition of M_j, we have

$$\begin{aligned}
\left\|M_j^{-1}\right\| &= \left\|A_j^{-1/2}\left(A_j^{-1/2}C_jD_j^{-1}C_j'A_j^{-1/2}-zI_p\right)^{-1}A_j^{-1/2}\right\| \\
&\leq \left\|A_j^{-1}\right\|\left\|\left(A_j^{-1/2}C_jD_j^{-1}C_j'A_j^{-1/2}-zI_p\right)^{-1}\right\|.
\end{aligned}$$

On the other hand, $\left\|A_j^{-1}\right\| = \mu_{\min,j}^{-1}$ and

$$\left\|\left(A_j^{-1/2}C_jD_j^{-1}C_j'A_j^{-1/2}-zI_p\right)^{-1}\right\| \leq \max_{k=1,\ldots,p}\frac{1}{\left|\lambda_k\left(A_j^{-1/2}C_jD_j^{-1}C_j'A_j^{-1/2}\right)-z\right|}$$

where $\lambda_k(\cdot)$ is the k-th largest eigenvalue of a real symmetric matrix. For $z = z_p$, the above inequality implies that

$$\left\|\left(A_j^{-1/2}C_jD_j^{-1}C_j'A_j^{-1/2}-zI_p\right)^{-1}\right\| \leq y_p^{-1},$$

and therefore,

$$\left\|M_j^{-1}\right\| \leq 1/\left(y_p\mu_{\min,j}\right). \tag{32}$$

The required bound for $\left\|M^{-1}\right\|$ is established similarly.

Further, we have

$$\left\|D_j^{-1}\right\| = 1/\lambda_p\left(D_j\right) \leq 1/\left(\lambda_p\left(\nabla\nabla'\right)\mu_{\min,j}\right) \leq 4/\mu_{\min,j}. \tag{33}$$

The required bound for $\left\|D^{-1}\right\|$ is established similarly. Next,

$$\left\|D_j^{-1}C_j'\right\|^2 = \left\|D_j^{-1}C_j'C_jD_j^{-1}\right\| = \left\|D_j^{-1}\varepsilon_{-(j)}\nabla_{-j}'\varepsilon_{-(j)}'\varepsilon_{-(j)}\nabla_{-j}\varepsilon_{-(j)}'D_j^{-1}\right\|,$$

where ∇_{-j} is the block-diagonal matrix obtained from ∇ by removing its j-th 2×2 block, and $\varepsilon_{-(j)}$ is obtained from ε by removing the $2j-1$-th and $2j$-th columns. On the other hand,

$$\begin{aligned}
\left\|D_j^{-1}\varepsilon_{-(j)}\nabla_{-j}'\varepsilon_{-(j)}'\varepsilon_{-(j)}\nabla_{-j}\varepsilon_{-(j)}'D_j^{-1}\right\| &\leq \mu_{\max,j}\left\|D_j^{-1}\varepsilon_{-(j)}\nabla_{-j}'\nabla_{-j}\varepsilon_{-(j)}'D_j^{-1}\right\| \\
&= \mu_{\max,j}\left\|D_j^{-1}D_jD_j^{-1}\right\| = \mu_{\max,j}\left\|D_j^{-1}\right\|.
\end{aligned}$$

Using (33), we obtain

$$\left\|D_j^{-1}C_j'\right\|^2 \leq 4\mu_{\max,j}/\mu_{\min,j}. \tag{34}$$

The required bound for $\left\|D^{-1}C'\right\|$ is established similarly.

Now let us establish the bounds on the differences of traces. As follows from (18), M_j^{-1} differs from M^{-1} by a matrix of rank no larger than 4. Therefore,

$$\left|\operatorname{tr}\left(M_j^{-1}-M^{-1}\right)\right| \leq 4\left\|M_j^{-1}-M^{-1}\right\| \leq 4\left(\left\|M_j^{-1}\right\|+\left\|M^{-1}\right\|\right). \tag{35}$$

Therefore,

$$\left|\operatorname{tr}\left(M_j^{-1}-M^{-1}\right)\right| \leq 4/\left(y_p\mu_{\min,j}\right)+4/\left(y_p\mu_{\min,0}\right) \leq 8/\left(y_p\mu_{\min,j}\right), \tag{36}$$

where the last inequality holds because $A - A_j$ is a positive-semidefinite matrix and hence $\mu_{\min,j} \leq \mu_{\min,0}$.

Similarly, $D_j^{-1}C_j'M_j^{-1}$ differs from $D^{-1}C'M^{-1}$ by a matrix with rank no larger than 8. It is because

$$\begin{aligned}
D_j^{-1}C_j'M_j^{-1} - D^{-1}C'M^{-1} &= D_j^{-1}C_j'\left(M_j^{-1}-M^{-1}\right) + D_j^{-1}\left(C_j'-C'\right)M^{-1} \\
&\quad + \left(D_j^{-1}-D^{-1}\right)C'M^{-1},
\end{aligned}$$

where the rank of $M_j^{-1} - M^{-1}$ is no larger than 4, and the ranks of $C_j' - C'$ and $D_j^{-1} - D^{-1}$ are no larger than 2 each. Therefore,

$$\left|\operatorname{tr}\left(D_j^{-1}C_j'M_j^{-1} - D^{-1}C'M^{-1}\right)\right| \leq 8\left(\left\|D_j^{-1}C_j'\right\|\left\|M_j^{-1}\right\| + \left\|D^{-1}C'\right\|\left\|M^{-1}\right\|\right) \leq 32\mu_{\max,0}^{1/2}/\left(y_p\mu_{\min,j}^{3/2}\right),$$

where we used (32) and (34). Finally, $D_j^{-1}C_j'M_j^{-1}C_jD_j^{-1}$ differs from $D^{-1}C'M^{-1}CD^{-1}$ by a matrix with rank no larger than 12. Therefore,

$$\left|\operatorname{tr}\left(D_j^{-1}C_j'M_j^{-1}C_jD_j^{-1} - D^{-1}C'M^{-1}CD^{-1}\right)\right| \leq 96\mu_{\max,0}/\left(y_p\mu_{\min,j}^2\right).$$

4.2.3 Rough bounds. Bounds on $\mu_{\min,j}$ and $\mu_{\max,0}$ (Pr of tail events, and moments)

In this section we derive some bounds on $\mu_{\max,j}$ and $\mu_{\min,j}$ that we will refer later in this note. We will need the following lemma due to BS98.

Lemma 4 *(Bai and Silverstein, 1998) If, for all $t > 0$, $\Pr(|X| > t)\, t^q \leq K$ for some positive q, then, for any positive $g < q$,*

$$\mathbb{E}|X|^g \leq K^{g/q}\left(\frac{q}{q-g}\right).$$

Now, we are ready to prove the following result.

Lemma 5 *(i) For any $q > 0$, we have $\Pr\left(\mu_{\min,j}^{-1} > \epsilon\right) \leq \left(\epsilon/(2e)^2\right)^{-q}$ for all $\epsilon > 0$ and all sufficiently large p and T along a sequence $p, T \to_{c_0} \infty$ with $p/T \leq 1/2$.*
(ii) For any $\rho > 0$, there exists $C > 0$ that may depend on ρ, but does not depend on j, p, and T, such that $\mathbb{E}\mu_{\min,j}^{-\rho} \leq C$ for all sufficiently large p and T along a sequence $p, T \to_{c_0} \infty$ with $p/T \leq 1/2$.

Proof: It follows from Chen and Dongarra (2005, p. 610) that

$$\Pr\left(\mu_{\min,j} \leq \mu\right) < (T-2)^{T-p-1}\mu^{\frac{T-p-1}{2}}/\Gamma(T-p).$$

Their $\lambda_{\min}, n$, and m equal $(T-2)\,\mu_{\min,j}, T-2$, and p in our notation, respectively. By Stirling's formula (see e.g. 6.1.38 in Abramowitz and Stegun (1970)),

$$\Gamma(T-p) \geq \sqrt{2\pi}\,(T-p-1)^{T-p-1/2}\,e^{-(T-p-1)}.$$

Further, for $p/T \leq 1/2$, we have $(T-2)/2 \leq T-p-1$. Therefore, for any $\epsilon > 0$ we have

$$\Pr\left(\mu_{\min,j}^{-1} > \epsilon\right) < (T-2)^{T-p-1}\epsilon^{-\frac{T-p-1}{2}}/\Gamma(T-p) \leq (2\pi(T-p-1))^{-1/2}\left(\epsilon/(2e)^2\right)^{-\frac{T-p-1}{2}}$$

Since $\Pr\left(\mu_{\min,j}^{-1} > \epsilon\right) \leq 1$, we have for any $q > 0$ and sufficiently large p, T

$$\Pr\left(\mu_{\min,j}^{-1} > \epsilon\right) \leq \left(\epsilon/(2e)^2\right)^{-q}.$$

Part (ii) follows from part (i) and Lemma 4.□

Lemma 6 *(i) For any $q > 0$ there exists $C > 0$ that does not depend on p and T s.t. $\Pr\left(\mu_{\max,0} > \epsilon\right) \leq C\epsilon^{-q}$ for all $\epsilon > 0$ and all sufficiently large p and T along a sequence $p, T \to_{c_0} \infty$ with $p/T \leq 1/2$.*
(ii) For any $\rho > 0$, there exists $C > 0$ that may depend on ρ, but does not depend on p, and T, such that $\mathbb{E}\mu_{\max,0}^{\rho} \leq C$ for all sufficiently large p and T along a sequence $p, T \to_{c_0} \infty$ with $p/T \leq 1/2$.

Proof: By Proposition 2.4 of Rudelson and Vershynin (2010), there exists $C > 0$ such that

$$\Pr\left(\mu_{\max,0} > (1+t)^2\right) \leq 2e^{-CTt^2}$$

for all sufficiently large p and T along a sequence $p, T \to_{c_0} \infty$ with $p/T \leq 1/2$. Since for any $q > 0$, there exists $C_q > 0$ such that, for all $t \geq 0$ we have $e^{-CTt^2} \leq C_q(1+t)^{-2q}$, and since $\Pr\left(\mu_{\max,0} > (1+t)^2\right) \leq 1$, we have $\Pr\left(\mu_{\max,0} > \epsilon\right) \leq C\epsilon^{-q}$. This completes the proof of part (i). Part (ii) follows from part (i) and Lemma 4.□

4.2.4 Rough bounds. Proof of Lemma OW8 (bounds on $\left\|\Omega_j^{(q)}\right\|$)

We start our proof from establishing a useful identity (see eq. (42) below). By definition of $\Omega_j^{(q)}$ and by Lemma 1, we have

$$\begin{aligned}\left(\frac{1}{1-z}\left[I_2, r_j\nabla_j'\right] + \left[v_j^{(q)}, u_j^{(q)\prime}\right]\right)\Omega_j^{(q)} &= [I_2, 0] \text{ and} &(37)\\ \left(\left[\frac{r_j}{1-z}\nabla_j, \frac{r_j z}{1-z}I_2\right] + \left[u_j^{(q)}, w_j^{(q)} - s_j^{(q)}\right]\right)\Omega_j^{(q)} &= [0, I_2]\,. &(38)\end{aligned}$$

Using the transposed of (37) in (27), we obtain

$$\varepsilon_{(j)}' D^{-1}C'M^{-1}\varepsilon_{(j)} = \left[\frac{r_j}{1-z}\nabla_j, \frac{r_j z}{1-z}I_2\right]\Omega_j^{(q)}\left[v_j^{(q)}, u_j^{(q)\prime}\right]'.$$

Using (38) in the above equation yields

$$\varepsilon_{(j)}' D^{-1}C'M^{-1}\varepsilon_{(j)} = u_j^{(q)} - \left[u_j^{(q)}, w_j^{(q)} - s_j^{(q)}\right]\Omega_j^{(q)}\left[v_j^{(q)}, u_j^{(q)\prime}\right]'. \tag{39}$$

Further, multiplying equation (23) by z, we get

$$\varepsilon_{(j)}' z\tilde{M}^{-1}\varepsilon_{(j)} = \frac{r_j z}{1-z}I_2 - \left[\frac{r_j}{1-z}\nabla_j, \frac{z r_j}{1-z}I_2\right]\Omega_j^{(q)}\left[\frac{r_j}{1-z}\nabla_j, \frac{r_j z}{1-z}I_2\right]'.$$

Using (38) and its transpose in the latter equation, we obtain

$$\varepsilon_{(j)}' z\tilde{M}^{-1}\varepsilon_{(j)} = w_j^{(q)} - s_j^{(q)} - \left[u_j^{(q)}, w_j^{(q)} - s_j^{(q)}\right]\Omega_j^{(q)}\left[u_j^{(q)}, w_j^{(q)} - s_j^{(q)}\right]'.$$

The identity $z\tilde{M}^{-1} = D^{-1}C'M^{-1}CD^{-1} - D^{-1}$ yields

$$\varepsilon_{(j)}' D^{-1}C'M^{-1}CD^{-1}\varepsilon_{(j)} - \varepsilon_{(j)}' D^{-1}\varepsilon_{(j)} = w_j^{(q)} - s_j^{(q)} - \left[u_j^{(q)}, w_j^{(q)} - s_j^{(q)}\right]\Omega_j^{(q)}\left[u_j^{(q)}, w_j^{(q)} - s_j^{(q)}\right]'. \tag{40}$$

By (19), we have

$$\varepsilon_{(j)}' M^{-1}\varepsilon_{(j)} = v_j^{(q)} - \left[v_j^{(q)}, u_j^{(q)\prime}\right]\Omega_j^{(q)}\left[v_j^{(q)}, u_j^{(q)\prime}\right]'. \tag{41}$$

Combining (39-41), we obtain identity

$$\mathcal{A}_j\Omega_j^{(q)}\mathcal{A}_j = \mathcal{M}_j, \tag{42}$$

where

$$\mathcal{A}_j = \begin{pmatrix} v_j^{(q)} & u_j^{(q)\prime} \\ u_j^{(q)} & w_j^{(q)} - s_j^{(q)} \end{pmatrix} \tag{43}$$

and

$$\begin{aligned}\mathcal{M}_{j,11} &= v_j^{(q)} - \varepsilon_{(j)}' M^{-1}\varepsilon_{(j)},\\ \mathcal{M}_{j,21} &= \mathcal{M}_{j,12}' = u_j^{(q)} - \varepsilon_{(j)}' D^{-1}C'M^{-1}\varepsilon_{(j)}, \text{ and}\\ \mathcal{M}_{j,22} &= w_j^{(q)} - s_j^{(q)} - \varepsilon_{(j)}' D^{-1}C'M^{-1}CD^{-1}\varepsilon_{(j)} + \varepsilon_{(j)}' D^{-1}\varepsilon_{(j)}.\end{aligned}$$

Lemma 7 $\max_{j=1,\ldots,T/2}\|\mathcal{M}_j\| \le y_p^{-1}20\mu_{\max,0}^2/\mu_{\min}^2$.

Proof: By definitions of $v_j^{(q)}$, $u_j^{(q)}$, $w_j^{(q)}$, and $s_j^{(q)}$, we have

$$\begin{aligned}\|\mathcal{M}_{j,11}\| &= \left\|\varepsilon_{(j)}'\left(M_j^{-1}-M^{-1}\right)\varepsilon_{(j)}\right\| \leq \left\|\varepsilon_{(j)}\varepsilon_{(j)}'\right\| \left\|M_j^{-1}-M^{-1}\right\| \leq \|\varepsilon\varepsilon'\| \left\|M_j^{-1}-M^{-1}\right\|, \\ \|\mathcal{M}_{j,21}\| &\leq \left\|\varepsilon_{(j)}\varepsilon_{(j)}'\right\| \left\|D_j^{-1}C_j'M_j^{-1}-D^{-1}C'M^{-1}\right\| \leq \|\varepsilon\varepsilon'\| \left(\left\|D_j^{-1}C_j'\right\|\left\|M_j^{-1}\right\|+\left\|D^{-1}C'\right\|\left\|M^{-1}\right\|\right), \text{ and} \\ \|\mathcal{M}_{j,22}\| &\leq \left\|\varepsilon_{(j)}\varepsilon_{(j)}'\right\| \left\|D_j^{-1}C_j'M_j^{-1}C_jD_j^{-1}-D^{-1}C'M^{-1}CD^{-1}\right\| + \left\|\varepsilon_{(j)}\varepsilon_{(j)}'\right\| \left\|D_j^{-1}-D^{-1}\right\|.\end{aligned}$$

Therefore, using Lemma 3 and the fact that $\mu_{\min} = \min\left\{\mu_{\min,0}, \mu_{\min,j}, j=1,...,T/2\right\}$, we obtain

$$\begin{aligned}\max_{j=1,...,T/2} \|\mathcal{M}_{j,11}\| &\leq y_p^{-1}2\mu_{\max,0}/\mu_{\min}, \\ \max_{j=1,...,T/2} \|\mathcal{M}_{j,21}\| &\leq y_p^{-1}4\left(\mu_{\max,0}/\mu_{\min}\right)^{3/2}, \text{ and} \\ \max_{j=1,...,T/2} \|\mathcal{M}_{j,22}\| &\leq y_p^{-1}8\left(\mu_{\max,0}/\mu_{\min}\right)^2 + 8\mu_{\max,0}/\mu_{\min} \leq y_p^{-1}16\left(\mu_{\max,0}/\mu_{\min}\right)^2.\end{aligned}$$

Finally, since $\|\mathcal{M}_j\| \leq \max\{\|\mathcal{M}_{j,11}\|, \|\mathcal{M}_{j,22}\|\} + \|\mathcal{M}_{j,21}\|$, we have

$$\max_{j=1,...,T/2} \|\mathcal{M}_j\| \leq y_p^{-1}\left(2\mu_{\max,0}/\mu_{\min}\right)\left(8\left(\mu_{\max,0}/\mu_{\min}\right) + 2\left(\mu_{\max,0}/\mu_{\min}\right)^{1/2}\right) \leq y_p^{-1}20\mu_{\max,0}^2/\mu_{\min}^2. \square$$

If $\max_{j=1,...,T/2}\left\|\mathcal{A}_j^{-1}\right\|$ is bounded, then Lemma 7 and identity (42) yield the boundedness of $\Omega_j^{(q)}$. However, the boundedness of $\max_{j=1,...,T/2}\left\|\mathcal{A}_j^{-1}\right\|$ is far from being obvious. To deal with the issue, let us multiply (18) by $\left[\varepsilon_{(j)}, C_jD_j^{-1}\varepsilon_{(j)}\right]'$ from the left and by $\left[\varepsilon_{(j)}, C_jD_j^{-1}\varepsilon_{(j)}\right]$ from the right. Rearranging the result, we obtain

$$\mathcal{B}_j\Omega_j^{(q)}\mathcal{B}_j = \mathcal{W}_j, \tag{44}$$

where

$$\mathcal{B}_j = \begin{pmatrix} v_j^{(q)} & u_j^{(q)\prime} \\ u_j^{(q)} & w_j^{(q)} \end{pmatrix},$$

and

$$\mathcal{W}_j = \begin{pmatrix} v_j^{(q)} - \varepsilon_{(j)}'M^{-1}\varepsilon_{(j)} & u_j^{(q)\prime} - \varepsilon_{(j)}'D_j^{-1}C_j'M^{-1}\varepsilon_{(j)} \\ u_j^{(q)} - \varepsilon_{(j)}'M^{-1}C_jD_j^{-1}\varepsilon_{(j)} & w_j^{(q)} - \varepsilon_{(j)}'D_j^{-1}C_j'M^{-1}C_jD_j^{-1}\varepsilon_{(j)} \end{pmatrix}.$$

The following lemma can be proven similarly to Lemma 7.

Lemma 8 $\max_{j=1,...,T/2}\|\mathcal{W}_j\| < y_p^{-1}12\mu_{\max,0}^2/\mu_{\min}^2$.

Now, it is sufficient to find a uniform over $j=1,...,T/2$ bound on $\min\left\{\left\|\mathcal{A}_j^{-1}\right\|, \left\|\mathcal{B}_j^{-1}\right\|\right\}$. Let

$$\hat{A} = \begin{pmatrix} v & u \\ u & w-s \end{pmatrix} \text{ and } \hat{B} = \begin{pmatrix} v & u \\ u & w \end{pmatrix}.$$

Lemma 9 $\min\left\{\left\|\hat{A}^{-1}\right\|, \left\|\hat{B}^{-1}\right\|\right\} \leq 100y_p^{-2}\mu_{\max,0}^3/\left(c^3\mu_{\min}^2\right)$ *for sufficiently large p.*

Proof: Note that, by the definition of v and by Lemma 3,

$$|v| \leq \frac{p}{T}\left\|M^{-1}\right\| \leq y_p^{-1}c/\mu_{\min}. \tag{45}$$

Further, since each of the entries of $D^{-1}C'M^{-1}$ can not be larger than $\left\|D^{-1}C'M^{-1}\right\|$ by absolute value, using the definition of u and Lemma 3, we obtain

$$|u| \leq y_p^{-1}2c\mu_{\max,0}^{1/2}/\mu_{\min}^{3/2}. \tag{46}$$

Similarly,

$$|w| \leq y_p^{-1} 4c\mu_{\max,0}/\mu_{\min}^2 \tag{47}$$

and

$$|w - s| \leq y_p^{-1} 4c\mu_{\max,0}/\mu_{\min}^2 + 4c/\mu_{\min} \leq y_p^{-1} 8c\mu_{\max,0}/\mu_{\min}^2. \tag{48}$$

Therefore

$$\max\left\{\left\|\hat{A}\right\|, \left\|\hat{B}\right\|\right\} \leq \max\{|v|, |w|, |w-s|\} + |u| \leq y_p^{-1} 10c\mu_{\max,0}/\mu_{\min}^2.$$

Note that since $\hat{A}$ and $\hat{B}$ are 2×2 matrices,

$$\left|\det \hat{A}\right| = \left\|\hat{A}\right\| \left\|\hat{A}^{-1}\right\|^{-1} \text{ and } \left|\det \hat{B}\right| = \left\|\hat{B}\right\| \left\|\hat{B}^{-1}\right\|^{-1}.$$

Using the latter two displays, we obtain

$$\left|\det \hat{B} - \det \hat{A}\right| \leq \left|\det \hat{B}\right| + \left|\det \hat{A}\right| \leq \left(y_p^{-1} 10c\mu_{\max,0}/\mu_{\min}^2\right) 2\max\left\{\left\|\hat{A}^{-1}\right\|^{-1}, \left\|\hat{B}^{-1}\right\|^{-1}\right\},$$

or equivalently,

$$\min\left\{\left\|\hat{A}^{-1}\right\|, \left\|\hat{B}^{-1}\right\|\right\} \leq y_p^{-1} \frac{20c\mu_{\max,0}/\mu_{\min}^2}{\left|\det \hat{B} - \det \hat{A}\right|}. \tag{49}$$

On the other hand,

$$\det \hat{B} - \det \hat{A} = vs. \tag{50}$$

For v, we have

$$v = \frac{1}{T} \operatorname{tr} A^{-1/2}\left(A^{-1/2}CD^{-1}C'A^{-1/2} - z_p I_p\right)^{-1} A^{-1/2} = \frac{1}{T}\sum_{j=1}^{p} \frac{h_j' A^{-1} h_j}{\lambda_{pj} - z_p},$$

where λ_{pj} and h_j are the j-th largest eigenvalue of $A^{-1/2}CD^{-1}C'A^{-1/2}$ (necessarily belonging to $[0,1]$) and a corresponding eigenvector. We have

$$\operatorname{Im} v = \frac{1}{T}\sum_{j=1}^{p} h_j' A^{-1} h_j y_p / \left|\lambda_{pj} - z_p\right|^2. \tag{51}$$

But $|\lambda_{pj} - z_p|^2$ is bounded from above by 2 (see assumption (OW26)). Therefore,

$$|v| \geq |\operatorname{Im} v| \geq \sum_{j=1}^{p} h_j' A^{-1} h_j y_p / (2T) = y_p \operatorname{tr} A^{-1} / (2T) \geq y_p c / \left(2\mu_{\max,0}\right). \tag{52}$$

As to s, let $\theta_1 \geq ... \geq \theta_p > 0$ be the eigenvalues of D. Then,

$$s = \frac{1}{T}\sum_{i=1}^{p} \theta_i^{-1} \geq \frac{p-k}{T}\theta_{k+1}^{-1}$$

with $k = 2\lceil p/4 \rceil$, where $\lceil p/4 \rceil$ denotes the smallest integer that is no smaller than $p/4$. Let us now decompose D into the sum $\varepsilon^{(1)}\Delta^{(1)}\varepsilon^{(1)\prime} + \varepsilon^{(2)}\Delta^{(2)}\varepsilon^{(2)\prime}$, where $\varepsilon^{(1)}$ is the $p \times k$ matrix that consists of the first k columns of ε, $\varepsilon^{(2)}$ is the $p \times (T-k)$ matrix that consists of last $T-k$ columns of ε, $\Delta^{(1)} = \operatorname{diag}\left\{r_1^{-1} I_2, ..., r_{k/2}^{-1} I_2\right\}$, and $\Delta^{(2)} = \operatorname{diag}\left\{r_{k/2+1}^{-1} I_2, ..., r_{T/2}^{-1} I_2\right\}$. Further, let $\beta_1 \geq ... \geq \beta_p > 0$ be the eigenvalues of $\varepsilon^{(2)}\Delta^{(2)}\varepsilon^{(2)\prime}$. By Theorem 4.3.6 of Horn and Johnson (1985), $\theta_{k+1} \leq \beta_1$, and therefore,

$$s \geq \frac{p-k}{T}\beta_1^{-1}.$$

On the other hand,

$$||\Delta^{(2)}|| = \left[2 - 2\cos\left(\frac{2\pi\left(k/2+1\right)}{T+1}\right)\right]^{-1} \leq \left[2 - 2\cos\left(\pi c/2\right)\right]^{-1}.$$

Since $1 - \cos x \geq x^2/4$ for $x \in [0, \pi/2]$, we have

$$||\Delta^{(2)}|| \leq 8/\left(c^2\pi^2\right).$$

Combining this with inequality $\left\|\varepsilon^{(2)}\varepsilon^{(2)\prime}\right\| \leq \mu_{\max,0}$, we obtain

$$\|\beta_1\| \leq 8\mu_{\max,0}/\left(c^2\pi^2\right),$$

and therefore

$$s \geq \frac{p - 2\lceil p/4\rceil}{T}\frac{c^2\pi^2}{8\mu_{\max,0}} \geq \frac{c^3\pi^2}{24\mu_{\max,0}}, \tag{53}$$

where the latter inequality holds because $\left(p - 2\lceil p/4\rceil\right)/T \geq c/3$ for sufficiently large p.

Using (52) and (53) in (50), we get

$$\left|\det\hat{B} - \det\hat{A}\right| \geq y_p c^4\pi^2/\left(48\mu^2_{\max,0}\right).$$

Combining this with (49), we obtain

$$\min\left\{\left\|\hat{A}^{-1}\right\|, \left\|\hat{B}^{-1}\right\|\right\} \leq y_p^{-2}960\mu^3_{\max,0}/\left(c^3\pi^2\mu^2_{\min}\right) \leq y^{-2}100\mu^3_{\max,0}/\left(c^3\mu^2_{\min}\right).\square$$

Lemma 10 *Suppose that* $0 \leq \alpha < 1/6$. *Then, for any* $C > 0$, *the inequalities* $\max_{j=1,\ldots,T/2}\left\|\mathcal{A}_j - \hat{A}\otimes I_2\right\| \leq Cy_p^2$ *and* $\max_{j=1,\ldots,T/2}\left\|\mathcal{B}_j - \hat{B}\otimes I_2\right\| \leq Cy_p^2$ *are satisfied w.ow.p.*

The proof of this lemma is very similar to that of Lemma OW7, and we omit it. Weyl's inequalities for singular values of a sum of two matrices (e.g. Horn and Johnson (1985), exerc. 16 on p.423) imply that

$$\begin{aligned} \left\|\mathcal{A}_j^{-1}\right\|^{-1} &\geq \left\|\hat{A}^{-1}\right\|^{-1} - \left\|\mathcal{A}_j - \hat{A}\otimes I_2\right\| \text{ and} \\ \left\|\mathcal{B}_j^{-1}\right\|^{-1} &\geq \left\|\hat{B}^{-1}\right\|^{-1} - \left\|\mathcal{B}_j - \hat{B}\otimes I_2\right\|. \end{aligned}$$

(To see this, note that $\left\|\mathcal{A}_j^{-1}\right\|^{-1}$ and $\left\|\hat{A}^{-1}\right\|^{-1}$ equal the smallest singular values of $\mathcal{A}_j$ and $\hat{A}\otimes I_2$, respectively.) Therefore, Lemmas 9 and 10 and the fact that event $\mathcal{E}_0$ holds w.ow.p. guarantee that, for any non-negative $\alpha < 1/6$, there exists $C > 0$ such that

$$\min_{j=1,\ldots,T/2}\max\left\{\left\|\mathcal{A}_j^{-1}\right\|^{-1}, \left\|\mathcal{B}_j^{-1}\right\|^{-1}\right\} \geq Cy_p^2 \text{ w.ow.p.}$$

Hence, as long as $0 \leq \alpha < 1/6$, there exists $C > 0$ such that

$$\max_{j=1,\ldots,T/2}\min\left\{\left\|\mathcal{A}_j^{-1}\right\|, \left\|\mathcal{B}_j^{-1}\right\|\right\} \leq Cy_p^{-2} \text{ w.ow.p.}$$

This fact taken together with Lemmas 7, 8, and equations (42), (44) imply that

$$\max_{j=1,\ldots,T/2}\left\|\Omega_j^{(q)}\right\| \leq Cy_p^{-5} \text{ holds w.ow.p.} \tag{54}$$

as long as $0 \leq \alpha < 1/6$.

Finally, note that

$$\|\Omega_j\|^{-1} \geq \left\|\Omega_j^{(q)}\right\|^{-1} - \left\|\Omega_j^{-1} - (\Omega_j^{(q)})^{-1}\right\|. \tag{55}$$

By inequality (54) and Lemma OW7, the term $\left\|\Omega_j^{-1} - (\Omega_j^{(q)})^{-1}\right\|$ in (55) is dominated by $\left\|\Omega_j^{(q)}\right\|^{-1}$ as long as $\alpha d + \gamma \geq 5\alpha$ for some $d \in (0, \infty)$, $\gamma \in [0, 1/2)$ and $0 \leq \alpha < \min\{1/6, (1/2 - \gamma)/(1+d)\}$. Inequalities $\alpha < (1/2-\gamma)/(1+d)$ and $\alpha d + \gamma \geq 5\alpha$ certainly hold if $d = 5$ and $\gamma = 0$ as long as $\alpha < 1/12$. Therefore, we conclude that

$$\max_{j=1,\ldots,T/2} \|\Omega_j\| \leq C y_p^{-5} \text{ holds w.ow.p.}$$

as long as $0 \leq \alpha < 1/12$.

To establish part (ii) of Lemma OW8, we need to rewrite identity (42) in a different form. For this, note that by definition of $\Omega_j^{(q)}$

$$\begin{aligned}
& \begin{pmatrix} v_j^{(q)} & u_j^{(q)\prime} \end{pmatrix} \Omega_j^{(q)} \begin{pmatrix} v_j^{(q)} \\ u_j^{(q)} \end{pmatrix} \\
= & \left(\begin{bmatrix} I_2 & 0 \end{bmatrix} - \frac{1}{1-z} \begin{bmatrix} I_2 & r_j \nabla_j' \end{bmatrix} \Omega_j^{(q)} \right) \begin{pmatrix} v_j^{(q)} \\ u_j^{(q)} \end{pmatrix} \\
= & v_j^{(q)} - \frac{1}{1-z} \begin{bmatrix} I_2 & r_j \nabla_j' \end{bmatrix} \left(\begin{bmatrix} I_2 \\ 0 \end{bmatrix} - \frac{1}{1-z} \Omega_j^{(q)} \begin{bmatrix} I_2 \\ r_j \nabla_j \end{bmatrix} \right) \\
= & v_j^{(q)} - \frac{1}{1-z} I_2 + \frac{1}{(1-z)^2} \begin{bmatrix} I_2 & r_j \nabla_j' \end{bmatrix} \Omega_j^{(q)} \begin{bmatrix} I_2 \\ r_j \nabla_j \end{bmatrix}.
\end{aligned}$$

Similarly,

$$\begin{aligned}
& \begin{pmatrix} v_j^{(q)} & u_j^{(q)\prime} \end{pmatrix} \Omega_j^{(q)} \begin{pmatrix} u_j^{(q)\prime} \\ w_j^{(q)} - s_j^{(q)} \end{pmatrix} \\
= & u_j^{(q)\prime} - \frac{1}{1-z} \begin{bmatrix} I_2 & r_j \nabla_j' \end{bmatrix} \left(\begin{bmatrix} 0 \\ I_2 \end{bmatrix} - \frac{1}{1-z} \Omega_j^{(q)} \begin{bmatrix} r_j \nabla_j' \\ r_j z I_2 \end{bmatrix} \right) \\
= & u_j^{(q)\prime} - \frac{1}{1-z} r_j \nabla_j' + \frac{1}{(1-z)^2} \begin{bmatrix} I_2 & r_j \nabla_j' \end{bmatrix} \Omega_j^{(q)} \begin{bmatrix} r_j \nabla_j' \\ r_j z I_2 \end{bmatrix}
\end{aligned}$$

and

$$\begin{aligned}
& \begin{pmatrix} u_j^{(q)} & w_j^{(q)} - s_j^{(q)} \end{pmatrix} \Omega_j^{(q)} \begin{pmatrix} u_j^{(q)\prime} \\ w_j^{(q)} - s_j^{(q)} \end{pmatrix} \\
= & \left(\begin{bmatrix} 0 & I_2 \end{bmatrix} - \frac{1}{1-z} \begin{bmatrix} r_j \nabla_j & r_j z I_2 \end{bmatrix} \Omega_j^{(q)} \right) \begin{pmatrix} u_j^{(q)\prime} \\ w_j^{(q)} - s_j^{(q)} \end{pmatrix} \\
= & w_j^{(q)} - s_j^{(q)} - \frac{1}{1-z} \begin{bmatrix} r_j \nabla_j & r_j z I_2 \end{bmatrix} \left(\begin{bmatrix} 0 \\ I_2 \end{bmatrix} - \frac{1}{1-z} \Omega_j^{(q)} \begin{bmatrix} r_j \nabla_j' \\ r_j z I_2 \end{bmatrix} \right) \\
= & w_j^{(q)} - s_j^{(q)} - \frac{1}{1-z} r_j z I_2 + \frac{1}{(1-z)^2} \begin{bmatrix} r_j \nabla_j & r_j z I_2 \end{bmatrix} \Omega_j^{(q)} \begin{bmatrix} r_j \nabla_j' \\ r_j z I_2 \end{bmatrix}.
\end{aligned}$$

Therefore, (42) can be written as

$$\mathcal{K}_j \Omega_j \mathcal{K}_j = \mathcal{K}_j - \begin{bmatrix} \varepsilon_{(j)}' M^{-1} \varepsilon_{(j)} & \varepsilon_{(j)}' M^{-1} C D^{-1} \varepsilon_{(j)} \\ \varepsilon_{(j)}' D^{-1} C' M^{-1} \varepsilon_{(j)} & z \varepsilon_{(j)}' \tilde{M}^{-1} \varepsilon_{(j)} \end{bmatrix}, \tag{56}$$

where

$$\mathcal{K}_j = \frac{1}{1-z} \begin{bmatrix} I_2 & r_j \nabla_j' \\ r_j \nabla_j & r_j z I_2 \end{bmatrix}.$$

The inverse of $\mathcal{K}_j$ equals

$$\mathcal{K}_j^{-1} = \left[\begin{array}{cc} -zI_2 & \nabla_j' \\ \nabla_j & -r_j^{-1} I_2 \end{array} \right].$$

By triangle inequality,

$$\left\| \mathcal{K}_j^{-1} \right\| \leq |z| + r_j^{-1} + 2 \left\| \nabla_j \right\| = |z| + r_j^{-1} + 2 r_j^{-1/2} \leq |z| - 1 + \left(1 + \frac{1}{2 \sin (\omega_1 / 2)} \right)^2.$$

Thus, for z with $\operatorname{Re} z \in [0,1]$ and $\operatorname{Im} z \in [0,1]$,

$$\left\| \mathcal{K}_j^{-1} \right\| \leq 1 + \left(1 + \frac{1}{2 \sin (\pi / (T+1))} \right)^2 \leq C T^2$$

for some absolute constant C. On the other hand, similarly to Lemma 7, we can show that

$$\left\| \left[\begin{array}{cc} \varepsilon_{(j)}' M^{-1} \varepsilon_{(j)} & \varepsilon_{(j)}' M^{-1} C D^{-1} \varepsilon_{(j)} \\ \varepsilon_{(j)}' D^{-1} C' M^{-1} \varepsilon_{(j)} & z \varepsilon_{(j)}' \tilde{M}^{-1} \varepsilon_{(j)} \end{array} \right] \right\| \leq y_p^{-1} 10 \mu_{\max,0}^2 / \mu_{\min,0}^2$$

Hence, from (56) and the above two displays,

$$\left\| \Omega_j^{(q)} \right\| \leq C T^2 + C T^4 y_p^{-1} \mu_{\max,0}^2 / \mu_{\min,0}^2.$$

Finally, let $\mathcal{E}_\Omega$ be event $\max_{j=1,\ldots,T/2} \left\| \Omega_j^{(q)} \right\| \leq C y_p^{-5}$. According to (54) $\mathcal{E}_\Omega$ holds w.ow.p. as long as $0 \leq \alpha < 1/6$. We have

$$\mathbb{E} \max_{j=1,\ldots,T/2} \left\| \Omega_j^{(q)} \right\|^\rho \leq C y_p^{-5\rho} + \mathbb{E} \mathbf{1} \left\{ \mathcal{E}_\Omega^c \right\} \left(C T^2 + C T^4 y_p^{-1} \mu_{\max,0}^2 / \mu_{\min,0}^2 \right)^\rho \leq C y_p^{-5\rho}$$

for sufficiently large p, T.

4.2.5 System reduction. Derivation of system (OW31) and proof of Lemma OW10

To simplify reference, let us reproduce here the original system of equations

$$m = \frac{1}{c(1-z)} - \frac{2}{cT} \sum_{j=1}^{T/2} \frac{z\tilde{v} + r_j (u + v - 1)}{(1-z) \delta_j} + e_1, \tag{57}$$

$$\frac{1}{c} + zm = \frac{1}{c(1-z)} - \frac{2}{cT} \sum_{j=1}^{T/2} \frac{z\tilde{v} + r_j z (u + zv - 1)}{(1-z) \delta_j} + e_2, \tag{58}$$

$$1 + zm = \frac{1}{c(1-z)} - \frac{2}{cT} \sum_{j=1}^{T/2} \frac{z\tilde{v} + r_j (u(1+z)/2 + zv - 1)}{(1-z) \delta_j} + e_3, \tag{59}$$

$$0 = \frac{2}{cT} \sum_{j=1}^{T/2} \frac{-u - r_j v/2}{\delta_j} + e_4. \tag{60}$$

We will assume that $z = z_p$, where z_p satisfies (OW26), that is,

$$x_p \equiv \operatorname{Re} z_p \in [0,1] \qquad \text{and} \qquad y_p \equiv \operatorname{Im} z_p = y_0 p^{-\alpha} \tag{61}$$

with $\alpha \geq 0$ and $y_0 \in (0,1]$ that are independent from p.

We begin from establishing some bounds on u, v, $\tilde{v}$, and on

$$\theta \equiv \frac{2}{cT} \sum_{j=1}^{T/2} \delta_j^{-1}.$$

Note that, by definition, $|\tilde{v}| \le \frac{p}{T}\left\|\tilde{M}^{-1}\right\|$. Using a similar argument to the one that yielded inequality $\left\|M^{-1}\right\| \le 1/\left(y_p\mu_{\min,0}\right)$ in Lemma 3, we obtain $\left\|\tilde{M}^{-1}\right\| \le 4/\left(y_p\mu_{\min,0}\right)$. Hence,

$$|\tilde{v}| \le \frac{4c}{y_p\mu_{\min,0}}. \tag{62}$$

Collecting inequalities (62), (45), (46), (52), and recalling the definition (29) of event $\mathcal{E}_0$, we obtain the following result.

Lemma 11 *There exists $C > 0$ such that each of the following events*

$$|u| \le Cy_p^{-1}, |v| \le Cy_p^{-1}, \text{ and } |\tilde{v}| \le Cy_p^{-1} \tag{63}$$

holds w.ow.p. Furthermore, there exists $C > 0$ such that

$$|v| \ge Cy_p \text{ holds w.ow.p.} \tag{64}$$

Subtracting (58) from (59) and then adding (60) multiplied by u/v yields

$$1 - \frac{1}{c} = \frac{2}{cT}\sum_{j=1}^{T/2}\frac{r_j\left(-u - zv + 1\right) - u^2/v}{\delta_j} + ue_4/v + e_3 - e_2.$$

Adding $1/c$ to both sides of this equation and recalling that

$$\delta_j = z\tilde{v}\left(1 + v - zv\right) + r_j\left(u + zv - 1\right) - \left(1 - z\right)u^2,$$

we obtain

$$1 = \left(z\tilde{v}\left(1 + v - zv\right) - \left(1 - z\right)u^2 - u^2/v\right)\theta + \left(ue_4/v + e_3 - e_2\right). \tag{65}$$

By Lemma OW9 and Lemma 11, for any $C \in (0, \infty)$, $d \in (0, \infty)$, and $\gamma \in [0, 1/2)$ s.t. $\alpha_{\gamma d} < 1/12$

$$|ue_4/v + e_3 - e_2| \le Cp^{-\gamma}y_p^{d-14} \text{ holds w.ow.p.}$$

as long as $0 \le \alpha < \alpha_{\gamma d}$. Choosing $d = 14$, $\gamma = 0$, (so that $\alpha_{\gamma d} \equiv (1/2 - \gamma)/(1 + d) = 1/30$) and setting $C = 1/2$, we obtain

$$|1 - (ue_4/v + e_3 - e_2)| \ge 1/2 \text{ holds w.ow.p.}$$

Hence, equation (65) and Lemma 11 imply that there exists $C > 0$ such that

$$\theta \ge Cy_p^3 \text{ holds w.ow.p.} \tag{66}$$

as long as $0 \le \alpha < 1/30$.

Derivation of the first equation of system (OW31). Subtracting $1/c$ from both sides of equation (58) and dividing it by z, then subtracting the resulting equation from equation (57), and, finally, subtracting equation (60) multiplied by two yields

$$0 = (\tilde{v} + 2u)\theta + e_1 - e_2/z - 2e_4.$$

Equivalently,

$$\tilde{v} + 2u = \tilde{e}_1, \tag{67}$$

where

$$\tilde{e}_1 = (-e_1 + e_2/z + 2e_4)\theta^{-1}. \tag{68}$$

Lemma 12 *For any $(C, d, \gamma) \in (0, \infty) \times [14, \infty) \times [0, 1/2)$ and any $\alpha \in [0, \alpha_{\gamma d})$, inequality $|\tilde{e}_1(z_p)| \le Cp^{-\gamma}y_p^{d-16}$ holds w.ow.p.*

Proof: The lemma follows from Lemma OW9 and equations (66) and (68).□

An intermediate version of the second equation of system (OW31). Replacing $\tilde{v}$ by $\tilde{e}_1 - 2u$ in (65), rearranging terms, and using (68) yields

$$1 = -\frac{u}{v}(2zv + u)(1 + v - zv)\theta + (-ze_1 + e_2 + 2ze_4)(1 + v - zv) + (ue_4/v + e_3 - e_2). \tag{69}$$

Further, multiplying equation (60) by $2c(u + zv - 1)/v$ gives us

$$0 = \frac{2}{T}\sum_{j=1}^{T/2} \frac{-2u(u + zv - 1)/v - r_j(u + zv - 1)}{\delta_j} + 2e_4c(u + zv - 1)/v.$$

Since the numerator of the summands can be written in the form

$$-2u(u + zv - 1)/v + z\tilde{v}(1 + v - zv) - (1 - z)u^2 - \delta_j,$$

we have

$$1 = c\left(-2u(u + zv - 1)/v + z\tilde{v}(1 + v - zv) - (1 - z)u^2\right)\theta + 2e_4c(u + zv - 1)/v.$$

Using $\tilde{v} = -2u + \tilde{e}_1$ in this equation and rearranging terms yields

$$1 = \frac{u}{v}c(2 - (2 + v - zv)(2zv + u))\theta + c(-ze_1 + e_2 + 2ze_4)(1 + v - zv) + 2e_4c(u + zv - 1)/v. \tag{70}$$

Subtracting equation (70) from (69), we obtain

$$0 = -\frac{u}{v}((2zv + u)((1 + v - zv)(1 - c) - c) + 2c)\theta + \xi_1.$$

with

$$\xi_1 \equiv (1 - c)(-ze_1 + e_2 + 2ze_4)(1 + v - zv) + (ue_4/v + e_3 - e_2) - 2e_4c(u + zv - 1)/v. \tag{71}$$

Equivalently,

$$(2zv + u)((1 + v - zv)(1 - c) - c) + 2c = \xi_2, \tag{72}$$

where

$$\xi_2 = -\frac{v}{u}\theta^{-1}\xi_1. \tag{73}$$

Note that the right hand side of (72) is linear in u. We will call equation (72) the intermediate version of the second equation of system (OW31). At the end of our derivations of (OW31), we show how one can obtain the final version of the second equation from the intermediate version. To bound the right hand side, ξ_2, of equation (72), we need to establish a bound on u^{-1}.

Lemma 13 *For any $(d, \gamma) \in [17, \infty) \times [0, 1/2)$ and any $\alpha \in [0, \alpha_{\gamma d})$, there exists $C > 0$ s.t. inequality $|u| > Cy_p$ holds w.ow.p.*

Proof: Recall that

$$\tilde{v} = \frac{1}{T}\operatorname{tr}\tilde{M}^{-1} = \frac{1}{T}\operatorname{tr}D^{-1/2}\left(D^{-1/2}C'A^{-1}CD^{-1/2} - z_pI_p\right)^{-1}D^{-1/2}.$$

This definition and the fact that the eigenvalues λ_{pj} of $D^{-1/2}C'A^{-1}CD^{-1/2}$ belong to $[0, 1]$, so that $|\lambda_{pj} - z_p|^{-1} \geq y_p/(1 + |z_p|)^2$, imply that

$$\operatorname{Im}\tilde{v} > \frac{y_p}{(1 + |z_p|)^2}\frac{1}{T}\operatorname{tr}D^{-1} = \frac{y_p}{(1 + |z_p|)^2}s.$$

As follows from inequality (53), $|s|$ is bounded away from zero w.ow.p. Therefore, there exists $C > 0$ such that

$$|\tilde{v}| > Cy_p \text{ w.ow.p.}$$

But by (67), $u = -\tilde{v}/2 + \tilde{e}_1/2$. On the other hand, Lemma 12 implies that, as long as $d \geq 17$, $|-\tilde{v}/2|$ dominates $|\tilde{e}_1/2|$ w.ow.p. We conclude that there exists $C > 0$ such that $|u| > Cy_p$ w.ow.p.□

Using inequality (66) as well as results of Lemma 13, Lemma 11, and Lemma OW9, we obtain the following bounds on $|v\xi_1|$ and $|\xi_2|$.

Lemma 14 *(i) For any* $(C,d,\gamma) \in (0,\infty) \times [5,\infty) \times [0,1/2)$ *and any* $\alpha \in [0,\alpha_{\gamma d})$, $|v\xi_1| \le Cp^{-\gamma} y_p^{d-14}$ *w.ow.p.*
(ii) For any $(C,d,\gamma) \in (0,\infty) \times [17,\infty) \times [0,1/2)$ *and any* $\alpha \in [0,\alpha_{\gamma d})$, $|\xi_2| \le Cp^{-\gamma} y_p^{d-18}$ *w.ow.p.*

Derivation of the third equation of system (OW31). Subtracting $1/c$ from both sides of equation (58) and dividing it by z yields

$$m = \frac{1}{c(1-z)} - \frac{2}{cT}\sum_{j=1}^{T/2} \frac{\tilde{v} + r_j(u + zv - 1)}{(1-z)\delta_j} + e_2/z. \tag{74}$$

On the other hand,

$$\frac{1}{c(1-z)} = \frac{2}{cT}\sum_{j=1}^{T/2} \frac{\delta_j}{(1-z)\delta_j} = \frac{2}{cT}\sum_{j=1}^{T/2} \frac{z\tilde{v}(1+v-zv) + r_j(u+zv-1) - (1-z)u^2}{(1-z)\delta_j}.$$

Using this in (74), we obtain

$$m = -\left(\tilde{v} - z\tilde{v}v + u^2\right)\theta + e_2/z.$$

Replacing $\tilde{v}$ by $-2u + \tilde{e}_1$ and using (68) yields

$$m = u(2 - 2zv - u)\theta + ve_2 + (e_1 - 2e_4)(1 - zv). \tag{75}$$

But from (72) and (73),

$$2zv + u = \frac{2c - \xi_2}{c - (1+v-zv)(1-c)} = \frac{2c + v\theta^{-1}\xi_1/u}{c - (1+v-zv)(1-c)}. \tag{76}$$

Using this in (75), we get

$$m = \frac{2(1+v-zv)(1-c)}{(1+v-zv)(1-c) - c}u\theta + \xi_3, \tag{77}$$

where

$$\xi_3 \equiv ve_2 + (e_1 - 2e_4)(1 - zv) - \frac{v\xi_1}{(1+v-zv)(1-c) - c}.$$

In the above identity, replacing ξ_1 by the right hand side of (71) and simplifying, we get

$$\xi_3 = e_1 - 2e_4 + \frac{czve_1 - e_3v + e_2v(1-c) - ue_4 + 2e_4c(u-1)}{(1+v-zv)(1-c) - c}.$$

Lemma 15 *(i) For any* $(d,\gamma) \in [17,\infty) \times [0,1/2)$ *and any* $\alpha \in [0,\alpha_{\gamma d})$*, there exists* $C > 0$ *such that* $|(1+v-zv)(1-c) - c| > Cy_p$ *w.ow.p.*
(ii) For any $(C,d,\gamma) \in (0,\infty) \times [17,\infty) \times [0,1/2)$ *and any* $\alpha \in [0,\alpha_{\gamma d})$, $|\xi_3| \le Cp^{-\gamma} y_p^{d-14}$ *w.ow.p.*

Proof: According to (72) and Lemma 14, for any $d \ge 17$, there exists $C > 0$ such that

$$|(2zv + u)((1+v-zv)(1-c) - c)| = |\xi_2 - 2c| > C \text{ w.ow.p.}$$

On the other hand, by Lemma 11, $|2zv + u| < Cy_p^{-1}$ w.ow.p. Hence, there must exist $C > 0$ such that

$$|(1+v-zv)(1-c) - c| > Cy_p \text{ w.ow.p.}$$

Part (ii) follows from part (i), Lemma 11 and Lemma OW9.□

Further, using (76) in (69), we obtain

$$\begin{aligned} 1 \;=\; & -\frac{u}{v}\frac{2c(1+v-zv)}{c - (1+v-zv)(1-c)}\theta + (-ze_1 + e_2 + 2ze_4)(1+v-zv) \\ & + (ue_4/v + e_3 - e_2) - \frac{\xi_1(1+v-zv)}{c - (1+v-zv)(1-c)}, \end{aligned} \tag{78}$$

or equivalently,

$$u\theta = \frac{v\left(\left(1+v-zv\right)\left(1-c\right)-c\right)}{2c\left(1+v-zv\right)} + \tilde{\xi}_3 \tag{79}$$

with

$$\tilde{\xi}_3 = \frac{\left(\left(1+v-zv\right)\left(1-c\right)-c\right)}{2c}\left(v\left(ze_1-e_2-2ze_4\right)-\frac{\left(ue_4+ve_3-ve_2\right)}{\left(1+v-zv\right)}\right)-\frac{v\xi_1}{2c}.$$

Lemma 16 *(i) There exists $C>0$ such that $|1+v-zv|>Cy_p^2$ holds w.ow.p.*
(ii) For any $(C,d,\gamma)\in(0,\infty)\times[5,\infty)\times[0,1/2)$ and any $\alpha\in[0,\alpha_{\gamma d})$, inequality $\left|\tilde{\xi}_3\right|\le Cp^{-\gamma}y_p^{d-16}$ holds w.ow.p.

Proof: Obviously, $|1+v-zv|=|1-z|\left|(1-z)^{-1}+v\right|$. On the other hand, both $(1-z)^{-1}$ and v have positive imaginary parts (to see this for v, recall (51)). Now using $z=z_p$ and equation (52), we obtain $|1+v-zv|>Cy_p^2$ w.ow.p. Part (ii) follows from part (i), Lemmas 11 and 14 and Lemma OW9.□

Using (79) in (77) yields

$$m=\left(c^{-1}-1\right)v+\tilde{e}_3, \tag{80}$$

where

$$\tilde{e}_3=\xi_3+2\tilde{\xi}_3+\frac{2c}{\left(1+v-zv\right)\left(1-c\right)-c}\tilde{\xi}_3.$$

Lemma 17 *For any $(C,d,\gamma)\in(0,\infty)\times[17,\infty)\times[0,1/2)$ and any $\alpha\in[0,\alpha_{\gamma d})$, inequality $|\tilde{e}_3(z_p)|\le Cp^{-\gamma}y_p^{d-17}$ holds w.ow.p.*

Proof: The lemma is a direct consequence of the definition of $\tilde{e}_3$, Lemma 15, and Lemma 16.□

Derivation of the fourth equation of system (OW31). Define

$$\delta\left(\varphi\right)=z\tilde{v}\left(1+v-zv\right)-\left(1-z\right)u^2+4\sin^2\varphi\left(u+zv-1\right)$$

and let

$$\xi_4=\frac{2}{T}\sum_{j=1}^{T/2}\delta_j^{-1}-\frac{2}{\pi}\int_0^{\pi/2}\delta\left(\varphi\right)^{-1}\mathrm{d}\varphi, \tag{81}$$

and

$$\tilde{\xi}_4=\frac{2}{T}\sum_{j=1}^{T/2}\frac{-u-r_jv/2}{\delta_j}-\frac{2}{\pi}\int_0^{\pi/2}\frac{-u-2\sin^2\varphi v}{\delta\left(\varphi\right)}\mathrm{d}\varphi. \tag{82}$$

Lemma 18 *For any $\alpha\in[0,1/12)$, there exists $C>0$ s.t. (i) $\min_{j=1,\dots,T/2}|\delta_j|\ge Cy_p^6$ w.ow.p., (ii) $\min_{\varphi\in[0,2\pi]}|\delta(\varphi)|\ge Cy_p^6$ w.ow.p., (iii) $|\xi_4|\le Cp^{-1}y_p^{-13}$ w.ow.p., (iv) $\left|\tilde{\xi}_4\right|\le Cp^{-1}y_p^{-14}$ w.ow.p.*

Proof: Using the definition of Ω_j, it is straightforward to verify that $\left(\delta_j/\left(1-z\right)\right)^2=\det\left(\Omega_j^{-1}\right)$. This implies that

$$\left|\frac{1}{1-z}\delta_j\right|^2=\prod_{i=1}^4\sigma_i\left(\Omega_j^{-1}\right), \tag{83}$$

where $\sigma_i\left(\mathcal{M}\right)$ denotes the i-th largest singular value of matrix $\mathcal{M}$.

By the inclusion principle (see Theorem 4.3.15 of Horn and Johnson (1985)), the first and second largest eigenvalues of $\Omega_j^{-1}\left(\Omega_j^{-1}\right)^*$ are no smaller than the first and the second largest eigenvalues of the upper left 2×2 block of $\Omega_j^{-1}\left(\Omega_j^{-1}\right)^*$, respectively. Such a block equals

$$\left|\frac{1}{1-z}+v\right|^2I_2+\left(\frac{r_j}{1-z}\nabla_j'+uI_2\right)\left(\frac{r_j}{1-z^*}\nabla_j+u^*I_2\right)\ge\left|\frac{1}{1-z}+v\right|^2I_2.$$

Therefore,

$$\min_{j=1,\ldots,T/2}\sigma_1\left(\Omega_j^{-1}\right)\geq\min_{j=1,\ldots,T/2}\sigma_2\left(\Omega_j^{-1}\right)\geq\left|\frac{1}{1-z}+v\right|\geq\operatorname{Im}\frac{1}{1-z}=\frac{y_p}{\left|1-z\right|^2},$$

where the latter inequality follows from the fact that both the imaginary part of v and that of $1/\left(1-z\right)$ are positive (for v, this follows from (51)).

On the other hand, by Lemma OW8, for any $\alpha\in[0,1/12)$, there exists $C>0$ such that

$$\min_{j=1,\ldots,T/2}\sigma_3\left(\Omega_j^{-1}\right)\geq\min_{j=1,\ldots,T/2}\sigma_4\left(\Omega_j^{-1}\right)\geq Cy_p^5\text{ w.ow.p.}$$

Combining the latter two displays with (83), we obtain $\left|\delta_j/\left(1-z\right)\right|^2\geq Cy_p^{12}/\left|1-z\right|^4$. Therefore,

$$\min_{j=1,\ldots,T/2}\left|\delta_j\right|\geq Cy_p^6/\left|1-z\right|\geq Cy_p^6\text{ w.ow.p.},$$

which establishes part (i).

Now, recall that $r_j=4\sin^2\left(\pi j/\left(T+1\right)\right)$. Therefore, for any $\varphi\in[0,2\pi]$, there exists $j\in\{1,...,T/2\}$ s.t.

$$\left|4\sin^2\varphi-r_j\right|\leq4\pi/T=4\pi c/p.$$

For such j,

$$\left|\delta\left(\varphi\right)-\delta_j\right|=\left|4\sin^2\varphi-r_j\right|\left|u+zv-1\right|\leq\frac{4\pi c}{p}\left|u+zv-1\right|,$$

so that, by Lemma 11, there exists $C>0$ s.t. $\left|\delta\left(\varphi\right)-\delta_j\right|\leq C/\left(py_p\right)$ w.ow.p. For $y_p=y_0p^{-\alpha}$ with $\alpha\in[0,1/12)$ and $y_0\in(0,1]$, quantity $1/\left(py_p\right)$ is clearly dominated by y_p^6. Therefore, using the result of part (i) of the lemma, we conclude that there exists $C>0$ s.t.

$$\min_{\varphi\in[0,2\pi]}\left|\delta\left(\varphi\right)\right|\geq Cy_p^6\text{ w.ow.p.},$$

which establishes part (ii).

To see that part (iii) holds, note that ξ_4 can be interpreted as the error due to replacing $\delta\left(\varphi\right)$ in the integral $\frac{2}{\pi}\int_0^{\pi/2}\delta\left(\varphi\right)^{-1}\mathrm{d}\varphi$ by a step function

$$\bar{\delta}\left(\varphi\right)=\delta_j\text{ for }\varphi\in\left[\left(j-1\right)\pi/T,j\pi/T\right).$$

We have

$$\left|\delta\left(\varphi\right)^{-1}-\bar{\delta}\left(\varphi\right)^{-1}\right|=\left|\bar{\delta}\left(\varphi\right)-\delta\left(\varphi\right)\right|/\left|\delta\left(\varphi\right)\bar{\delta}\left(\varphi\right)\right|.$$

On the other hand, similar arguments to those used in the proof of part (ii) show that for $\alpha\in[0,1/12)$, there exists $C>0$ such that

$$\left|\bar{\delta}\left(\varphi\right)-\delta\left(\varphi\right)\right|/\left|\delta\left(\varphi\right)\bar{\delta}\left(\varphi\right)\right|\leq\frac{C}{py_p}y_p^{-12}\text{ w.ow.p.}$$

Hence, $\left|\xi_4\right|\leq Cp^{-1}y_p^{-13}$ w.ow.p.

Similarly for part (iv), $\tilde{\xi}_4$ can be interpreted as the error due to replacing $\left(-u-2\sin^2\varphi v\right)/\delta\left(\varphi\right)$ in the integral $\frac{2}{\pi}\int_0^{\pi/2}\left(-u-2\sin^2\varphi v\right)\delta^{-1}\left(\varphi\right)\mathrm{d}\varphi$ by a step function

$$f\left(\varphi\right)=\frac{-u-r_jv/2}{\delta_j}\text{ for }\varphi\in\left[\left(j-1\right)\pi/T,j\pi/T\right).$$

For $\varphi\in\left[\left(j-1\right)\pi/T,j\pi/T\right)$, we have

$$\begin{aligned}\left|\frac{-u-4\sin^2\varphi v/2}{\delta\left(\varphi\right)}-f\left(\varphi\right)\right| &= \left|\frac{-u-2\sin^2\varphi v}{\delta\left(\varphi\right)}-\frac{-u-r_jv/2}{\delta_j}\right|\\ &\leq \left|\frac{\left(4\sin^2\varphi-r_j\right)v/2}{\delta_j}\right|+\frac{\left|-u-2\sin^2\varphi v\right|\left|\delta_j-\delta\left(\varphi\right)\right|}{\left|\delta\left(\varphi\right)\delta_j\right|}\\ &\leq C\left(p^{-1}y_p^{-7}+p^{-1}y_p^{-14}\right)\leq Cp^{-1}y_p^{-14}\text{ w.ow.p.}\end{aligned}$$

Hence, $\left|\tilde{\xi}_4\right| \leq Cp^{-1}y_p^{-14}$ w.ow.p.□

In (78), replacing $c\theta$ by $\frac{2}{\pi}\int_0^{\pi/2}\delta(\varphi)^{-1}\,\mathrm{d}\varphi+\xi_4$, and then dividing the resulting equation by $\frac{2}{\pi}\int_0^{\pi/2}\delta(\varphi)^{-1}\,\mathrm{d}\varphi$ yields

$$\left(\frac{2}{\pi}\int_0^{\pi/2}\delta(\varphi)^{-1}\,\mathrm{d}\varphi\right)^{-1} = \frac{u}{v}\frac{2(1+v-zv)}{(1+v-zv)(1-c)-c}+\xi_5, \tag{84}$$

where

$$\begin{aligned}\xi_5 \;=\;& \left(\frac{2}{\pi}\int_0^{\pi/2}\delta(\varphi)^{-1}\,\mathrm{d}\varphi\right)^{-1}\left(\frac{u}{v}\frac{2(1+v-zv)}{(1+v-zv)(1-c)-c}\xi_4\right.\\ &\left.+\frac{\xi_1(1+v-zv)}{(1+v-zv)(1-c)-c}+(-ze_1+e_2+2ze_4)(1+v-zv)+(ue_4/v+e_3-e_2)\right).\end{aligned} \tag{85}$$

Since

$$\delta(\varphi)=2(u+zv-1)\left(x+2\sin^2\varphi\right)\ \text{ with } x=\frac{z\tilde{v}(1+v-zv)-(1-z)u^2}{2(u+zv-1)}, \tag{86}$$

equation (84) and the fact that

$$\left(\frac{2}{\pi}\int_0^{\pi/2}\frac{1}{x+2\sin^2\varphi}\mathrm{d}\varphi\right)^2=\frac{1}{x(x+2)}$$

for any $x\in\mathbb{C}\backslash[-2,0]$ imply that

$$4(u+zv-1)^2x(x+2)=\left(\frac{u}{v}\frac{2(1+v-zv)}{(1+v-zv)(1-c)-c}\right)^2+\frac{u}{v}\frac{4(1+v-zv)}{(1+v-zv)(1-c)-c}\xi_5+\xi_5^2. \tag{87}$$

Note that $x\in\mathbb{C}\backslash[-2,0]$ is satisfied because otherwise $\delta(\varphi)=0$ for some $\varphi\in[0,2\pi]$, which contradicts Lemma 18. Also, we show below (see the proof of Lemma 19) that $u+zv-1$ is bounded away from zero w.ow.p. so that x is well defined by (86).

Using the definition of x in (87) and multiplying both sides of the equation by $v^2/(1+v-zv)^2$ yields

$$\begin{aligned}&v^2\left(z\tilde{v}-\frac{(1-z)u^2}{1+v-zv}\right)\left(z\tilde{v}+\frac{-(1-z)u^2+4(u+zv-1)}{1+v-zv}\right)=\left(\frac{2u}{(1+v-zv)(1-c)-c}\right)^2\\&+\frac{4uv\xi_5}{(1+v-zv)((1+v-zv)(1-c)-c)}+\frac{v^2\xi_5^2}{(1+v-zv)^2}.\end{aligned}$$

Next, using (67) in the above equation and rearranging, we obtain

$$\frac{v^2(2z+(2zv+u)(1-z))(u-uz-2)(2zv+u-2)}{(1+v-zv)^2}=u\left(\frac{2}{(1+v-zv)(1-c)-c}\right)^2+\xi_6, \tag{88}$$

where

$$\begin{aligned}\xi_6 \;=\;& \frac{4v\xi_5}{(1+v-zv)((1+v-zv)(1-c)-c)}+\frac{v^2\xi_5^2}{u(1+v-zv)^2}\\&-\frac{z^2v^2}{u}\tilde{e}_1^2-\frac{2zv^2}{u}\tilde{e}_1\left(-2uz+\frac{-(1-z)u^2+2(u+zv-1)}{(1+v-zv)}\right).\end{aligned} \tag{89}$$

Now our goal is to use (76) to eliminate u from equation (88). We have

$$\begin{aligned}2z+(2zv+u)(1-z) \;=\;& \frac{2(z(1+v-zv)(1-c)-c)}{(1+v-zv)(1-c)-c}+\frac{(1-z)\xi_2}{(1+v-zv)(1-c)-c}\\ \equiv\;& \frac{a_1}{b}+\frac{(1-z)\xi_2}{b},\end{aligned} \tag{90}$$

$$\begin{aligned} 2zv+u-2 &= \frac{(1+v-zv)(-2(1-c))}{(1+v-zv)(1-c)-c}+\frac{\xi_2}{(1+v-zv)(1-c)-c} \\ &\equiv \frac{(1+v-zv)\,a_2}{b}+\frac{\xi_2}{b}, \end{aligned}$$

$$\begin{aligned} u-uz-2 &= \frac{(1+v-zv)(-2(zv(1-z)(1-c)+1-c-zc))}{(1+v-zv)(1-c)-c}+\frac{(1-z)\,\xi_2}{(1+v-zv)(1-c)-c} \\ &\equiv \frac{(1+v-zv)\,a_3}{b}+\frac{(1-z)\,\xi_2}{b}, \end{aligned}$$

and

$$\begin{aligned} u &= \frac{-2\left(zv^2(1-z)(1-c)+zv(1-2c)+c\right)}{(1+v-zv)(1-c)-c}+\frac{\xi_2}{(1+v-zv)(1-c)-c} \\ &\equiv \frac{a_4}{b}+\frac{\xi_2}{b}. \end{aligned} \tag{91}$$

Using these identities in (88), and simplifying, we obtain

$$\begin{aligned} & v^2a_1a_2a_3+\xi_2v^2\left((1-z)\,a_2\left(\frac{a_1}{1+v-zv}+a_3\right)+\frac{a_1a_3}{1+v-zv}\right) \\ & +\xi_2^2\frac{v^2(1-z)}{1+v-zv}\left(a_3+(1-z)\,a_2+\frac{a_1}{1+v-zv}\right)+\xi_2^3\frac{v^2(1-z)^2}{(1+v-zv)^2} = (a_4+\xi_2)\,4+b^3\xi_6, \end{aligned}$$

or more explicitly,

$$\begin{aligned} & 8v^2(zv(1-z)(1-c)+z-c-zc)(1-c)(zv(1-z)(1-c)+1-c-zc) \\ = \; & -8\left(zv^2(1-z)(1-c)+zv(1-2c)+c\right) \\ & +b^3\xi_6+\xi_2\left(4-(1-z)\,a_2\left(\frac{a_1}{1+v-zv}+a_3\right)-\frac{a_1a_3}{1+v-zv}\right) \\ & -\xi_2^2\frac{v^2(1-z)}{1+v-zv}\left(a_3+(1-z)\,a_2+\frac{a_1}{1+v-zv}\right)-\xi_2^3\frac{v^2(1-z)^2}{(1+v-zv)^2}. \end{aligned}$$

It turns out that the difference between the left hand side of the latter equation and the first term on its right hand side can be factorized. Specifically, it is straightforward although laborious to verify that

$$\begin{aligned} & 8v^2(zv(1-z)(1-c)+z-c-zc)(1-c)(zv(1-z)(1-c)+1-c-zc) \\ & +8\left(zv^2(1-z)(1-c)+zv(1-2c)+c\right) \\ = \; & 8(1-c)\left(z(1-c)(1-z)v^2+(1-zc-c)v+1\right)\left(z(1-c)(1-z)v^2-(-z+c+zc)v+\frac{c}{1-c}\right). \end{aligned}$$

Therefore, we have

$$v^2z(1-c)(1-z)-v(c-z+zc)+\frac{c}{1-c}=\xi_7, \tag{92}$$

where

$$\begin{aligned} \xi_7 \equiv \; & \left(8(1-c)\left(z(1-c)(1-z)v^2+(1-zc-c)v+1\right)\right)^{-1} \\ & \times\left(b^3\xi_6+\xi_2\left(4-(1-z)\,a_2\left(\frac{a_1}{1+v-zv}+a_3\right)-\frac{a_1a_3}{1+v-zv}\right)\right. \\ & \left.-\xi_2^2\frac{v^2(1-z)}{1+v-zv}\left(a_3+(1-z)\,a_2+\frac{a_1}{1+v-zv}\right)-\xi_2^3\frac{v^2(1-z)^2}{(1+v-zv)^2}\right). \end{aligned} \tag{93}$$

Finally, using (80) in (92), we obtain the fourth equation of system (OW31):

$$m^2 z(1-z)c - m(c-z+zc) + 1 = \tilde{e}_4, \tag{94}$$

where

$$\tilde{e}_4 \equiv \frac{1-c}{c}\xi_7 + \tilde{e}_3(2mz(1-z)c - (c-z+zc)) - \tilde{e}_3^2 z(1-z)c. \tag{95}$$

Lemma 19 *For any* $(C,d,\gamma) \in (0,\infty)\times[30,\infty)\times[0,1/2)$ *and any* $\alpha \in [0,\alpha_{\gamma d})$, $|\tilde{e}_4(z_p)| \le Cp^{-\gamma}y_p^{d-41}$ *w.ow.p.*

Proof: Recall that by (66), $\theta \ge Cy_p^3$ holds w.ow.p. By definition of ξ_4,

$$\frac{2}{\pi}\int_0^{\pi/2}\delta(\varphi)^{-1}\,\mathrm{d}\varphi = c\theta - \xi_4.$$

On the other hand, by Lemma 18, $|\xi_4| \le Cp^{-1}y_p^{-13}$ w.ow.p. Quantity $c\theta$ dominates $|\xi_4|$ as long as $py_p^{16} = p^{1-16\alpha} \to \infty$, which is certainly true for $\alpha < 1/16$. Of course, for $(d,\gamma) \in [30,\infty)\times[0,1/2)$, $\alpha_{\gamma d} \le 1/62 < 1/16$. Hence, there exists $C>0$ s.t.

$$\left|\frac{2}{\pi}\int_0^{\pi/2}\delta(\varphi)^{-1}\,\mathrm{d}\varphi\right| \ge Cy_p^3 \text{ w.ow.p.} \tag{96}$$

Using inequality (96) and Lemmas 11, 14(i), 15(i), and 18(iii) in the definition (85) of ξ_5, we obtain

$$|\xi_5| \le Cp^{-\gamma}y_p^{d-20} \text{ w.ow.p.} \tag{97}$$

Next, using inequality (97) and Lemmas 11, 15(i), and 16(i) we conclude that the first term on the right hand side of equation (89) defining ξ_6 is bounded by $Cp^{-\gamma}y_p^{d-24}$ w.ow.p. The last term is bounded by $Cp^{-\gamma}y_p^{d-23}$ w.ow.p. This follows from Lemmas 11, 12, 13, and 16(i). The second term has form $\left(\frac{v\xi_5}{u^{1/2}(1+v-zv)}\right)^2$. Inequality (97) and Lemmas 11, 13, and 16(i) yield

$$\left|\frac{v\xi_5}{u^{1/2}(1+v-zv)}\right| \le Cp^{-\gamma}y_p^{d-23.5} \text{ w.ow.p.}$$

For $d \ge 30$ and sufficiently small C this implies that

$$\left|\frac{v\xi_5}{u^{1/2}(1+v-zv)}\right|^2 \le \left|\frac{v\xi_5}{u^{1/2}(1+v-zv)}\right| \text{ w.ow.p.}$$

Therefore, the second term on the right hand side of (89) is bounded by $Cp^{-\gamma}y_p^{d-23.5}$ w.ow.p. By a similar argument, the third term is bounded by $Cp^{-\gamma}y_p^{d-17.5}$ w.ow.p. Summing up, since the bound $Cp^{-\gamma}y_p^{d-24}$ on the first term on the right hand side of (89) is the largest, we conclude that

$$|\xi_6| \le Cp^{-\gamma}y_p^{d-24} \text{ w.ow.p.} \tag{98}$$

Now consider definition (93) of ξ_7. Let us show that $z(1-c)(1-z)v^2 + (1-zc-c)v + 1$ is bounded away from zero w.ow.p. From (91), we have

$$u + zv - 1 = \frac{-v^2 z(1-c)(1-z) + v(c+zc-1) - 1}{(1+v-zv)(1-c)-c} + \frac{\xi_2}{(1+v-zv)(1-c)-c}.$$

Hence,

$$v^2 z(1-z)(1-c) - v(c+cz-1) + 1 = -(u+zv-1)((1+v-zv)(1-c)-c) + \xi_2 \tag{99}$$

By Lemma 15(i),

$$|(1+v-zv)(1-c)-c|>Cy_p \text{ w.ow.p.}$$

Furthermore, by Lemma 14(ii), $|\xi_2|\le Cp^{-\gamma}y_p^{d-18}$ w.ow.p. The latter bound can be made arbitrarily small by choosing d and γ sufficiently large. Therefore, it remains to bound $u+zv-1$ away from zero.

Consider event

$$|u+zv-1|\le Cy_p^{13} \tag{100}$$

for some $C>0$. Recall that

$$\delta(\varphi)=z\tilde{v}(1+v-zv)-(1-z)u^2+4\sin^2\varphi(u+zv-1).$$

Our plan is to show that $z\tilde{v}(1+v-zv)-(1-z)u^2$ is bounded away from zero so that if (100) holds, then $\delta(\varphi)$ is nearly constant for $\varphi\in[0,2\pi]$. This will lead to a contradiction.

Using (67), rewrite $\delta(\varphi)$ as

$$\delta(\varphi)=-u(2z+(2zv+u)(1-z))+4\sin^2\varphi(u+zv-1)+z\tilde{e}_1(1+v-zv). \tag{101}$$

Focus on the first term of this expression. Slightly rearranging terms on the right hand side of (90) yields

$$2z+(2zv+u)(1-z)=\frac{2((z+z(1-z)v)(1-c)-c)}{(1+v-zv)(1-c)-c}+\frac{(1-z)\xi_2}{(1+v-zv)(1-c)-c}. \tag{102}$$

Let us show that

$$|(z+z(1-z)v)(1-c)-c|\ge(1-c)y_p. \tag{103}$$

For this, it is sufficient to prove that $\operatorname{Im}\{z(1-z)v\}\ge 0$. Note that v is a weighted sum of the form $\sum_{i=1}^p W_i(\lambda_{pi}-z)^{-1}$, where λ_{pi} are eigenvalues of $CD^{-1}C'A^{-1}$, and thus, belong to $[0,1]$, and W_i are non-negative weights. Therefore, for $z=z_p=x_p+\mathrm{i}y_p$, we have

$$z(1-z)v=\sum_{i=1}^p W_i\frac{(x_p+\mathrm{i}y_p)(1-x_p-\mathrm{i}y_p)(\lambda_{pi}-x_p+\mathrm{i}y_p)}{|\lambda_{pi}-z|^2}.$$

Hence, it is sufficient to show that $\operatorname{Im}\{(x_p+\mathrm{i}y_p)(1-x_p-\mathrm{i}y_p)(\lambda_{pi}-x_p+\mathrm{i}y_p)\}\ge 0$ for all i. We have

$$\begin{aligned}
&\operatorname{Im}\{(x_p+\mathrm{i}y_p)(1-x_p-\mathrm{i}y_p)(\lambda_{pi}-x_p+\mathrm{i}y_p)\}\\
=\;& y_p^3+y_p((1-x_p)(\lambda_{pi}-x_p)-x_p(\lambda_{pi}-x_p)+x_p(1-x_p))\\
=\;& y_p^3+y_p\left(\lambda_{pi}-2x_p\lambda_{pi}+x_p^2\right)\\
=\;& y_p^3+y_p\left(\lambda_{pi}(1-\lambda_{pi})+(x_p-\lambda_{pi})^2\right)\ge 0
\end{aligned}$$

Since (103) holds, equation (102) and Lemma 11 imply that for some $C>0$

$$|2z+(2zv+u)(1-z)|\ge Cy_p^2 \tag{104}$$

w.ow.p. It is because by Lemmas 14(ii) and 15(i)

$$\left|\frac{(1-z)\xi_2}{(1+v-zv)(1-c)-c}\right|\le Cp^{-\gamma}y_p^{d-19}<y_p^3$$

w.ow.p. for $d\ge 30$ and sufficiently small C. Inequality (104) and Lemma 13 imply that

$$|-u(2z+(2zv+u)(1-z))|\ge Cy_p^3 \tag{105}$$

w.ow.p.

Now let us again consider equation (101). By Lemmas 11 and 12, the last term in that equation satisfies

$$|z\tilde{e}_1(1+v-zv)|\le Cp^{-\gamma}y_p^{d-17} \text{ w.ow.p.}$$

for any $C > 0$. Since, by assumption, $d \geq 30$, we have

$$|z\tilde{e}_1 (1 + v - zv)| \leq Cp^{-\gamma} y_p^{13} \tag{106}$$

w.ow.p.

Now, if inequalities (100), (105) and (106) hold, we have

$$\frac{1}{\delta(\varphi)} = \frac{1 + \kappa(\varphi)}{-u(2z + (2zv + u)(1 - z))} \tag{107}$$

with

$$\max_{\varphi \in [0, 2\pi]} |\kappa(\varphi)| \leq Cy_p^{10} \tag{108}$$

for some $C > 0$.

On the other hand, equations (60) and (82) yield

$$0 = \frac{2}{c\pi} \int_0^{\pi/2} \frac{-u - 2\sin^2\varphi v}{\delta(\varphi)} \mathrm{d}\varphi + \tilde{\xi}_4/c + e_4.$$

Using (107) in this equation, we obtain

$$\begin{aligned} 0 &= \frac{2}{c\pi} \int_0^{\pi/2} \frac{\left(-u - 2\sin^2\varphi v\right)(1 + \kappa(\varphi))}{-u(2z + (2zv + u)(1 - z))} \mathrm{d}\varphi + \tilde{\xi}_4/c + e_4 \\ &= \frac{u + v}{cu(2z + (2zv + u)(1 - z))} + \frac{2}{c\pi} \int_0^{\pi/2} \frac{\left(-u - 2\sin^2\varphi v\right)\kappa(\varphi)}{-u(2z + (2zv + u)(1 - z))} \mathrm{d}\varphi + \tilde{\xi}_4/c + e_4. \end{aligned}$$

By inequalities (105), (108) and

$$\max_{\varphi \in [0, 2\pi]} \left|-u - 2\sin^2\varphi v\right| \leq Cy_p^{-1}, \tag{109}$$

which holds w.ow.p. according to Lemma 11, we have

$$\left| \frac{2}{\pi} \int_0^{\pi/2} \frac{\left(-u - 2\sin^2\varphi v\right)\kappa(\varphi)}{-u(2z + (2zv + u)(1 - z))} \mathrm{d}\varphi \right| \leq Cy_p^6$$

for some $C > 0$. Furthermore, by Lemmas OW9 and 18(iv), we have, for some $C > 0$,

$$\left|\tilde{\xi}_4/c + e_4\right| \leq Cy_p^6 \tag{110}$$

w.ow.p. Hence, as long as inequalities (100), (105), (106), (109), and (110) hold, we have

$$|u + v| \leq Cy_p^3$$

for some $C > 0$. Taken together with (100), this means that

$$|v - zv + 1| \leq |-u - zv + 1| + |u + v| \leq Cy_p^3.$$

But by Lemma 16(i), there exists $C > 0$ such that

$$|v - zv + 1| > Cy_p^2 \text{ w.ow.p.}$$

Besides, inequalities (105), (106), (109), and (110) hold w.ow.p. This implies that the complementary event to (100) holds w.ow.p. That is, for some $C > 0$,

$$|u + zv - 1| \geq Cy_p^{13} \text{ w.ow.p.} \tag{111}$$

Using this in (99), we obtain

$$\left|z(1 - c)(1 - z)v^2 + (1 - zc - c)v + 1\right| \geq Cy_p^{14} \text{ w.ow.p.} \tag{112}$$

for some $C > 0$.

Return now to definition (93) of ξ_7. By Lemma 11 and inequality (98), we have, for any $C > 0$,

$$\left|b^3 \xi_6\right| \leq C p^{-\gamma} y_p^{d-27} \text{ w.ow.p.}$$

By Lemmas 11, 14(ii), and 16(i), for any $C > 0$,

$$\left|\xi_2 \left(4 - (1-z)\, a_2 \left(\frac{a_1}{1+v-zv} + a_3\right) - \frac{a_1 a_3}{1+v-zv}\right)\right| \leq C p^{-\gamma} y_p^{d-22} \text{ w.ow.p.}$$

Note that by Lemma 14(ii), $\left|\xi_2^2\right| \leq |\xi_2|$ w.ow.p. as long as $d > 18$. Hence, for any $C > 0$,

$$\left|\xi_2^2 \frac{v^2 (1-z)}{1+v-zv} \left(a_3 + (1-z)\, a_2 + \frac{a_1}{1+v-zv}\right)\right| \leq C p^{-\gamma} y_p^{d-25} \text{ w.ow.p.}$$

and similarly,

$$\left|\xi_2^3 \frac{v^2 (1-z)^2}{(1+v-zv)^2}\right| \leq C p^{-\gamma} y_p^{d-24} \text{ w.ow.p.}$$

Combining these inequalities with (112) we obtain, for any $C > 0$,

$$|\xi_7| \leq C p^{-\gamma} y_p^{d-41} \text{ w.ow.p.} \tag{113}$$

Finally, using definition (95) of $\tilde{e}_4$, Lemma 17, equation (80) and Lemma 11, we conclude that, for any $(C, d, \gamma) \in (0, \infty) \times [30, \infty) \times [0, 1/2)$ and any $\alpha \in [0, \alpha_{\gamma d})$,

$$|\tilde{e}_4| \leq C p^{-\gamma} y_p^{d-41} \text{ w.ow.p.}\square$$

Derivation of the second equation of system (OW31) and conclusion. Subtracting equation (92) from (72) and factorizing the left hand side of the result, we obtain

$$((1+v-zv)(1-c) - c)\,(u + vz + c/(1-c)) = \xi_2 - \xi_7.$$

Dividing both sides of this equation by $((1+v-zv)(1-c) - c)$ yields

$$u + vz + c/(1-c) = \tilde{e}_2 \tag{114}$$

with

$$\tilde{e}_2 = (\xi_2 - \xi_7) / ((1+v-zv)(1-c) - c)\,.$$

Inequality (113) together with Lemmas 14(ii) and 15(i) imply the following result.

Lemma 20 *For any* $(C, d, \gamma) \in (0, \infty) \times [30, \infty) \times [0, 1/2)$ *and any* $\alpha \in [0, \alpha_{\gamma d})$, $|\tilde{e}_2(z_p)| \leq C y_p^{d-42}$ *w.ow.p.*

To conclude, equations (67), (114), (80), and (94) derived above are the equations of system (OW31). Lemma OW10 is a direct consequence of Lemmas 12, 20, 17, and 19.

4.2.6 Analysis of $m - m_0$. Proof of Lemma OW11

As explained in OW, by choosing $d \geq 42$ we can make $4\tilde{e}_4 c z_p (1 - z_p)$ negligible relative to $(c - z_p + c z_p)^2 - 4 c z_p (1 - z_p)$ so that the difference $m(z_p) - m_0(z_p)$ is of order

$$\tilde{e}_4 / \sqrt{(c - z_p + c z_p)^2 - 4 c z_p (1 - z_p)}.$$

Then by Lemma OW10 and by inequality (OW34), event

$$\mathcal{E}_z = \left\{|m(z_p) - m_0(z_p)| \leq C p^{-\gamma} y_p^{d-42.5}\right\} \tag{115}$$

holds w.ow.p. for any $(C,d,\gamma)\in(0,\infty)\times[42,\infty)\times[0,1/2)$ and any $\alpha\in[0,\alpha_{\gamma d})$. We use index z in the notation $\mathcal{E}_z$ to emphasize the dependence of the event on $\{z_p\}$. We will set $C=1$ to shorten notation.

Since $m(z_p)$ and $m_0(z_p)$ are both Stieltjes transforms, their absolute values are always bounded by y_p^{-1}, and hence the inequality

$$|m(z_p)-m_0(z_p)|\leq 2y_p^{-1}$$

is always valid. Therefore, we always have

$$y_p^{-1}|m(z_p)-m_0(z_p)|\leq p^{-\gamma}y_p^{d-43.5}+2y_p^{-2}\mathbf{1}_{\mathcal{E}_z^c}$$

where $\mathbf{1}_{\mathcal{E}_z^c}$ is the indicator of the event $\mathcal{E}_z^c$ complementary to $\mathcal{E}_z$.

Let S_p be a subset of $[0,1]$ containing at most p elements. Then,

$$y_p^{-1}\max_{x_p\in S_p}|m(z_p)-m_0(z_p)|\leq p^{-\gamma}y_p^{d-43.5}+2y_p^{-2}\max_{x_p\in S_p}\mathbf{1}_{\mathcal{E}_z^c} \tag{116}$$

and

$$\Pr\left(y_p^{-1}\max_{x_p\in S_p}|m(z_p)-m_0(z_p)|>\epsilon\right)\leq\Pr\left(p^{-\gamma}y_p^{d-43.5}>\epsilon/2\right)+\Pr\left(2y_p^{-2}\max_{x_p\in S_p}\mathbf{1}_{\mathcal{E}_z^c}>\epsilon/2\right). \tag{117}$$

Let us choose $d=43.5$ and $\gamma=1/180$ so that $\alpha_{\gamma d}\equiv(1/2-\gamma)/(1+d)=1/90$. Note that for any $l>0$ and $\rho\geq 180l$, and any $\epsilon>0$, the following inequalities trivially hold

$$\Pr\left(p^{-1/180}>\epsilon/2\right)\leq\left(2p^{-1/180}\epsilon^{-1}\right)^{\rho}\leq 2^{\rho}\epsilon^{-\rho}p^{-l}. \tag{118}$$

Similarly,

$$\Pr\left(2y_p^{-2}\max_{x_p\in S_p}\mathbf{1}_{\mathcal{E}_z^c}>\epsilon/2\right)\leq\Pr\left(\max_{x_p\in S_p}\mathbf{1}_{\mathcal{E}_z^c}=1\right)\left(4y_p^{-2}\epsilon^{-1}\right)^{\rho}.$$

Since $\mathcal{E}_z$ holds w.ow.p., we must have

$$\Pr\left(\mathbf{1}_{\mathcal{E}_z^c}=1\right)\leq Cp^{-2\rho\alpha-l-1}$$

for any $\rho>0$ and $l>0$, where constant C may depend on ρ, l, and α but not on p. Therefore,

$$\Pr\left(2y_p^{-2}\max_{x_p\in S_p}\mathbf{1}_{\mathcal{E}_z^c}>\epsilon/2\right)\leq\sum_{x_p\in S_p}\Pr\left(\mathbf{1}_{\mathcal{E}_z^c}=1\right)\left(4y_p^{-2}\epsilon^{-1}\right)^{\rho}\leq C\epsilon^{-\rho}p^{-l}. \tag{119}$$

Using (118) and (119) in (117) we conclude that for any $l>0$ and $\rho\geq 180l$, and any $\epsilon>0$,

$$\Pr\left(y_p^{-1}\max_{x_p\in S_p}|m(z_p)-m_0(z_p)|>\epsilon\right)\leq C\epsilon^{-\rho}p^{-l}, \tag{120}$$

where C may depend on l, ρ, and α.

Let the p elements of S_p be equally spaced between 0 and 1. Then, for any $x^{(1)},x^{(2)}\in[0,1]$ s.t. $\left|x^{(1)}-x^{(2)}\right|\leq p^{-1}$,

$$\left|m\left(x^{(1)}+\mathrm{i}y_p\right)-m\left(x^{(2)}+\mathrm{i}y_p\right)\right|\leq y_p^{-2}\left|x^{(1)}-x^{(2)}\right|\leq y_p^{-2}p^{-1}$$

and similarly,

$$\left|m_0\left(x^{(1)}+\mathrm{i}y_p\right)-m_0\left(x^{(2)}+\mathrm{i}y_p\right)\right|\leq y_p^{-2}p^{-1}.$$

Therefore,

$$\begin{aligned}
&\Pr\left(y_p^{-1}\sup_{x_p\in[0,1]}|m(z_p)-m_0(z_p)|>\epsilon\right)\\
\leq\;&\Pr\left(y_p^{-1}\max_{x_p\in S_p}|m(z_p)-m_0(z_p)|+2y_p^{-3}p^{-1}>\epsilon\right)\\
\leq\;&\Pr\left(y_p^{-1}\max_{x_p\in S_p}|m(z_p)-m_0(z_p)|>\epsilon/2\right)+\Pr\left(2y_p^{-3}p^{-1}>\epsilon/2\right)\\
\leq\;&C\epsilon^{-\rho}p^{-l}+C4^{\rho}\epsilon^{-\rho}p^{-(1-3\alpha)\rho}\leq C\epsilon^{-\rho}p^{-l}.
\end{aligned}$$

To summarize, for any $\alpha < \alpha_{\gamma d} = 1/90$, any $l > 0$ and $\rho \geq 180l$, and any $\epsilon > 0$, we have

$$\Pr\left(y_p^{-1} \sup_{x_p\in[0,1]} |m(z_p) - m_0(z_p)| > \epsilon\right) \leq C\epsilon^{-\rho}p^{-l},$$

where C is a constant that may depend on l, ρ, and α.

4.2.7 Analysis of $m - m_0$. Proof of Proposition OW12 (bound on $\mathbb{E}F_p([a,b])$)

First, we prove the following lemma.

Lemma 21 *For any* $\alpha \in [0, 1/90)$, $\max_{k=0,\ldots,T/2} y_p^{-2}\mathbb{E}_k\left(\sup_{x_p\in[0,1]} |m(z_p) - m_0(z_p)|^2\right) \overset{\text{a.s.}}{\to} 0.$

Proof: The arguments used in the proof of this lemma closely follow the arguments of BS98 that lead from their inequality (3.23) to equation (3.24). Let $l, l_1 > 0$ be arbitrary and let $l_0 > l$ be s.t.

$$\Pr\left(y_p^{-1} \sup_{x_p\in[0,1]} |m(z_p) - m_0(z_p)| > \epsilon\right) \leq C\epsilon^{-\rho}p^{-l_0} \tag{121}$$

for a $\rho \geq 180l_0$ s.t. $r \equiv l\rho/(l_0l_1) > 1$. Constant C in (121) may depend on α, ρ, and l_0. Inequality (121) and Lemma 4 imply that

$$y_p^{-l\rho/l_0}\mathbb{E} \sup_{x_p\in[0,1]} |m(z_p) - m_0(z_p)|^{l\rho/l_0} \leq C^{l/l_0}p^{-l}l_0/(l_0 - l). \tag{122}$$

Further, for any $\epsilon > 0$, we have by Jensen's inequality

$$\Pr\left(\max_{k=0,\ldots,T/2} \mathbb{E}_k\left(y_p^{-l_1} \sup_{x_p\in[0,1]} |m - m_0|^{l_1}\right) > \epsilon\right) \leq \Pr\left(\max_{k=0,\ldots,T/2} \mathbb{E}_k\left(y_p^{-rl_1} \sup_{x_p\in[0,1]} |m - m_0|^{rl_1}\right) > \epsilon^r\right).$$

Note that $\mathbb{E}_k\left(y_p^{-rl_1} \sup_{x_p\in[0,1]} |m - m_0|^{rl_1}\right)$, $k = 0, 1, \ldots, T/2$ forms a martingale. By Kolmogorov's inequality for sub-martingales (Lemma 2.5 of BS98), we have

$$\begin{aligned}\Pr\left(\max_{k=0,\ldots,T/2} \mathbb{E}_k\left(y_p^{-rl_1} \sup_{x_p\in[0,1]} |m - m_0|^{rl_1}\right) > \epsilon^r\right) &\leq \epsilon^{-r}y_p^{-rl_1}\mathbb{E} \sup_{x_p\in[0,1]} |m - m_0|^{rl_1} \\ &= \epsilon^{-r}y_p^{-l\rho/l_0}\mathbb{E} \sup_{x_p\in[0,1]} |m - m_0|^{l\rho/l_0}.\end{aligned}$$

Hence,

$$\Pr\left(\max_{k=0,\ldots,T/2} \mathbb{E}_k\left(y_p^{-l_1} \sup_{x_p\in[0,1]} |m - m_0|^{l_1}\right) > \epsilon\right) \leq \epsilon^{-r}y_p^{-l\rho/l_0}\mathbb{E} \sup_{x_p\in[0,1]} |m - m_0|^{l\rho/l_0}.$$

Using (122) in the above inequality yields

$$\Pr\left(\max_{k=0,\ldots,T/2} \mathbb{E}_k\left(y_p^{-l_1} \sup_{x_p\in[0,1]} |m - m_0|^{l_1}\right) > \epsilon\right) \leq \epsilon^{-r}p^{-l}C^{l/l_0}l_0/(l_0 - l).$$

Setting $l_1 = 2$, and noting that the right hand side is summable in p for $l > 1$, by Borel-Cantelli lemma, we have

$$\max_{k=0,\ldots,T/2} y_p^{-2}\mathbb{E}_k\left(\sup_{x_p\in[0,1]} |m - m_0|^2\right) \overset{\text{a.s.}}{\to} 0.\square$$

Now let us turn to the proof of Proposition OW12. It will follow arguments on pp. 330-331 of BS98. Let $G_{pk}(x_1, x_2)$, $k = 0, ..., T(p)/2$ be the following functions on $\mathbb{R}^2$

$$G_{pk}(x_1, x_2) = \mathbb{E}_k F_p(x_1) F_p(x_2),$$

where $F_p(x)$ denotes the cumulative distribution function corresponding to F_p. Any integer $q \geq 1$ can be represented in the form

$$q = \sum_{j=1}^{p-1} (T(j)/2 + 1) + k$$

with $0 \leq k \leq T(p)/2$. Using this representation, define a sequence of probability distribution functions $\{G_q\}_{q=1}^{\infty}$ on $\mathbb{R}^2$ as

$$G_q(x_1, x_2) = G_{pk}(x_1, x_2).$$

The two-dimensional Stieltjes transform of G_q, $m_q^{(G)}(x_1 + \mathrm{i}y_1, x_2 + \mathrm{i}y_2)$, equals $\mathbb{E}_k m_p(x_1 + \mathrm{i}y_1) m_p(x_2 + \mathrm{i}y_2)$. When $\alpha = 0$, Lemma OW11 implies that, with probability 1,

$$\sup_{(x_1,x_2)\in[0,1]^2} \left| m_q^{(G)}(x_1 + \mathrm{i}y_1, x_2 + \mathrm{i}y_2) - m_0(x_1 + \mathrm{i}y_1) m_0(x_2 + \mathrm{i}y_2) \right| \to 0$$

as $q \to \infty$ for countably many (y_1, y_2) forming a dense subset of an open set in $[0,1]^2$, uniformly bounded away from the axes. Therefore, with probability 1, $G_q(x_1, x_2)$ weakly converges to $W_{c_0}(x_1) W_{c_0}(x_2)$.

Let $[a', b'] = [a - \underline{\epsilon}, b + \underline{\epsilon}]$ with $\underline{\epsilon}$ such that $[a - 2\underline{\epsilon}, b + 2\underline{\epsilon}]$ lies outside the support of W_{c_0}. Clearly, $[a', b']$ will lie outside the support of W_{c_0}, and it will lie outside the support of W_c for sufficiently large p. Let

$$\operatorname{Re} m = m_1^{\text{out}} + m_1^{\text{in}} \text{ and } \operatorname{Im} m = m_2^{\text{out}} + m_2^{\text{in}},$$

where

$$m_1^{\text{out}}(x + \mathrm{i}y) = \frac{1}{p} \sum_{\lambda_{pj}\in[a',b']} \frac{x - \lambda_{pi}}{(x - \lambda_{pi})^2 + y^2} \text{ and } m_2^{\text{out}}(x + \mathrm{i}y) = \frac{1}{p} \sum_{\lambda_{pj}\in[a',b']} \frac{y}{(x - \lambda_{pi})^2 + y^2}.$$

Now let $\bar{m}_0(z)$ be the Stieltjes transform of W_{c_0}. We have

$$\begin{aligned} \mathbb{E}_k \left| m_2^{\text{in}}(x + \mathrm{i}y)/y - \frac{\mathrm{d}}{\mathrm{d}x} \bar{m}_0(x) \right|^2 \leq & \left| \mathbb{E}_k \left(m_2^{\text{in}}(x + \mathrm{i}y)/y \right)^2 - \left(\frac{\mathrm{d}}{\mathrm{d}x} \bar{m}_0(x) \right)^2 \right| \\ & + 2 \left| \frac{\mathrm{d}}{\mathrm{d}x} \bar{m}_0(x) \right| \left| \mathbb{E}_k m_2^{\text{in}}(x + \mathrm{i}y)/y - \frac{\mathrm{d}}{\mathrm{d}x} \bar{m}_0(x) \right|. \end{aligned} \quad (123)$$

Since $G_q(x_1, x_2)$ a.s. weakly converges to $W_{c_0}(x_1) W_{c_0}(x_2)$ and function

$$\left((x - x_1)^2 + y^2 \right)^{-1} \left((x - x_2)^2 + y^2 \right)^{-1}$$

of $(x_1, x_2) \in [a', b']^c \times [a', b']^c$ is uniformly bounded and equicontinuous for $(x, y) \in [a, b] \times [0, 1]$, we have

$$\max_{k=0,...,T/2} \sup_{(x,y)\in[a,b]\times[0,1]} \left| \mathbb{E}_k \left(m_2^{\text{in}}(x + \mathrm{i}y)/y \right)^2 - \left(\operatorname{Im} \bar{m}_0(x + \mathrm{i}y)/y \right)^2 \right| \overset{\text{a.s.}}{\to} 0. \quad (124)$$

On the other hand,

$$\sup_{x\in[a,b]} \left| \operatorname{Im} \bar{m}_0(x + \mathrm{i}y_p)/y_p - \frac{\mathrm{d}}{\mathrm{d}x} \bar{m}_0(x) \right| \to 0$$

for any $y_p \to 0$. Therefore, (124) yields

$$\max_{k=0,...,T/2} \sup_{x\in[a,b]} \left| \mathbb{E}_k \left(m_2^{\text{in}}(x + \mathrm{i}y_p)/y_p \right)^2 - \left(\frac{\mathrm{d}}{\mathrm{d}x} \bar{m}_0(x) \right)^2 \right| \overset{\text{a.s.}}{\to} 0$$

for any $y_p \to 0$. Similarly, we have

$$\max_{k=0,\ldots,T/2} \sup_{x\in[a,b]} \left| \mathbb{E}_k m_2^{\text{in}}\left(x+\mathrm{i}y_p\right)/y_p - \frac{\mathrm{d}}{\mathrm{d}x}\bar{m}_0\left(x\right)\right| \overset{\text{a.s.}}{\to} 0.$$

Combining the latter two displays with (123) and noting that $\sup_{x\in[a,b]}\left|\frac{\mathrm{d}}{\mathrm{d}x}\bar{m}_0\left(x\right)\right|$ is bounded, we obtain

$$\max_{k=0,\ldots,T/2} \sup_{x\in[a,b]} \mathbb{E}_k \left| m_2^{\text{in}}\left(x+\mathrm{i}y_p\right)/y_p - \frac{\mathrm{d}}{\mathrm{d}x}\bar{m}_0\left(x\right)\right|^2 \overset{\text{a.s.}}{\to} 0 \tag{125}$$

for any $y_p \to 0$. On the other hand, by Lemma 21,

$$\max_{k=0,\ldots,T/2} \mathbb{E}_k \left(\sup_{x_p\in[0,1]} \left| m\left(z_p\right)/y_p - m_0\left(z_p\right)/y_p\right|^2 \right) \overset{\text{a.s.}}{\to} 0,$$

which implies that

$$\max_{k=0,\ldots,T/2} \mathbb{E}_k \left(\sup_{x_p\in[0,1]} \left| \operatorname{Im} m\left(z_p\right)/y_p - \frac{\mathrm{d}}{\mathrm{d}x_p}\bar{m}_0\left(z_p\right)\right|^2 \right) \overset{\text{a.s.}}{\to} 0. \tag{126}$$

Since

$$\operatorname{Im} m\left(z_p\right)/y_p = m_2^{\text{in}}\left(x+\mathrm{i}y_p\right)/y_p + m_2^{\text{out}}\left(x+\mathrm{i}y_p\right)/y_p,$$

convergences (125) and (126) yield

$$\max_{k=0,\ldots,T/2} \sup_{x\in[a,b]} \mathbb{E}_k \left| m_2^{\text{out}}\left(x+\mathrm{i}y_p\right)/y_p\right|^2 \overset{\text{a.s.}}{\to} 0. \tag{127}$$

Finally, for any $x \in [a,b]$, we have

$$\max_{k=0,\ldots,T/2} \mathbb{E}_k \left| m_2^{\text{out}}\left(x+\mathrm{i}y_p\right)/y_p\right|^2 \geq \frac{1}{p^2} \max_{k=0,\ldots,T/2} \mathbb{E}_k \left(\sum_{\lambda_{pj}\in[a,b]\cap[x-y_p,x+y_p]} \frac{1}{\left(x-\lambda_{pj}\right)^2 + y_p^2} \right)^2.$$

Since $pF_p\left([a,b]\cap[x-y_p,x+y_p]\right)$ equals the number of λ_{pj} that belong to $[a,b]\cap[x-y_p,x+y_p]$, and since $\left(x-\lambda_{pj}\right)^2+y_p^2 \leq 2y_p^2$ for any such λ_{pj}, the above inequality yields

$$\max_{k=0,\ldots,T/2} \mathbb{E}_k \left| m_2^{\text{out}}\left(x+\mathrm{i}y_p\right)/y_p\right|^2 \geq \max_{k=0,\ldots,T/2} \mathbb{E}_k \left(\frac{\left(F_p\left([a,b]\cap[x-y_p,x+y_p]\right)\right)^2}{4y_p^4} \right). \tag{128}$$

Let J be the smallest integer larger than $(b-a)/(2y_p)$, and let $x_1,\ldots,x_J$ be such that

$$[a,b] \subseteq \cup_{j=1}^J \left[x_j - y_p, x_j + y_p\right].$$

Then,

$$\left(F_p\left([a,b]\right)\right)^2 \leq \left(\sum_{j=1}^J F_p\left([a,b]\cap[x_j-y_p,x_j+y_p]\right) \right)^2 \leq J \sum_{j=1}^J \left(F_p\left([a,b]\cap[x_j-y_p,x_j+y_p]\right)\right)^2,$$

and equations (127) and (128) yield

$$\max_{k=0,\ldots,T/2} \mathbb{E}_k \left(F_p\left([a,b]\right)\right)^2 = o_{\text{a.s.}}\left(y_p^2\right) \text{ and } \max_{k=0,\ldots,T/2} \mathbb{E}_k F_p\left([a,b]\right) = o_{\text{a.s.}}\left(y_p\right).$$

Recall that $y_p = y_0 p^{-\alpha}$, where $\alpha \in [0, 1/90)$. Choosing $\alpha = 1/91$, we obtain

$$\max_{k=0,\ldots,T/2} \mathbb{E}_k \left(F_p\left([a,b]\right)\right)^2 = o_{\text{a.s.}}\left(p^{-2/91}\right) \text{ and } \max_{k=0,\ldots,T/2} \mathbb{E}_k F_p\left([a,b]\right) = o_{\text{a.s.}}\left(p^{-1/91}\right).$$

4.3 Step 2: Convergence of $m - \mathbb{E}m$

4.3.1 Proof of equation (OW36) (initial representation of $m - \mathbb{E}m$)

By definition, $m = \mathrm{tr}\left(AM^{-1}\right)/p$. On the other hand, by (18),

$$\begin{aligned} AM^{-1} &= \left(A_j + \varepsilon_{(j)}\varepsilon_{(j)}'\right)\left(M_j^{-1} - M_j^{-1}\alpha_j\Omega_j^{(q)}\alpha_j' M_j^{-1}\right) \\ &= A_j M_j^{-1} + \varepsilon_{(j)}\varepsilon_{(j)}' M_j^{-1} - A_j M_j^{-1}\alpha_j\Omega_j^{(q)}\alpha_j' M_j^{-1} - \varepsilon_{(j)}\varepsilon_{(j)}' M_j^{-1}\alpha_j\Omega_j^{(q)}\alpha_j' M_j^{-1}, \end{aligned}$$

so that

$$\mathrm{tr}\left(AM^{-1}\right) - \mathrm{tr}\left(A_j M_j^{-1}\right) = \mathrm{tr}\, v_j^{(q)} - \mathrm{tr}\left[v_j^{(q)}, u_j^{(q)\prime}\right]\Omega_j^{(q)}\left[v_j^{(q)}, u_j^{(q)\prime}\right]' - \mathrm{tr}\left[\begin{pmatrix} a_j^{(q)} & b_j^{(q)\prime} \\ b_j^{(q)} & c_j^{(q)} \end{pmatrix}\Omega_j^{(q)}\right]. \tag{129}$$

This can be written in a more compact form by noting that

$$v_j^{(q)} - \left[v_j^{(q)}, u_j^{(q)\prime}\right]\Omega_j^{(q)}\left[v_j^{(q)}, u_j^{(q)\prime}\right]' = \frac{1}{1-z}\left[I_2, r_j\nabla_j'\right]\Omega_j^{(q)}\left[v_j, u_j'\right]'.$$

Using this identity and (129), we conclude that

$$\mathrm{tr}\left(AM^{-1}\right) - \mathrm{tr}\left(A_j M_j^{-1}\right) = \mathrm{tr}\left(\Gamma_j^{(q)}\Omega_j^{(q)}\right),$$

where

$$\Gamma_j^{(q)} = \begin{pmatrix} \frac{1}{1-z} v_j^{(q)} - a_j^{(q)} & \frac{1}{1-z} r_j v_j^{(q)}\nabla_j' - b_j^{(q)\prime} \\ \frac{1}{1-z} u_j^{(q)} - b_j^{(q)} & \frac{1}{1-z} r_j u_j^{(q)}\nabla_j' - c_j^{(q)} \end{pmatrix}.$$

On the other hand, since $\left(\mathbb{E}_j - \mathbb{E}_{j-1}\right)\mathrm{tr}\left(A_j M_j^{-1}\right) = 0$,

$$m - \mathbb{E}m = \frac{1}{p}\sum_{j=1}^{T/2}\left(\mathbb{E}_j - \mathbb{E}_{j-1}\right)\mathrm{tr}\left(AM^{-1}\right) = \frac{1}{p}\sum_{j=1}^{T/2}\left(\mathbb{E}_j - \mathbb{E}_{j-1}\right)\left(\mathrm{tr}\left(AM^{-1}\right) - \mathrm{tr}\left(A_j M_j^{-1}\right)\right).$$

Hence,

$$m - \mathbb{E}m = \frac{1}{p}\sum_{j=1}^{T/2}\left(\mathbb{E}_j - \mathbb{E}_{j-1}\right)\mathrm{tr}\left(\Gamma_j^{(q)}\Omega_j^{(q)}\right).$$

4.3.2 Proof of Lemma OW13 (boundedness of $||\Omega_j^{(d)}||$)

Recall that $m_0(z)$ was defined as a solution to the fourth equation of system (31) after $\tilde{e}_4$ is replaced by 0. Let $v_0(z), u_0(z)$, and $\tilde{v}_0(z)$ be the corresponding solutions to the first three equations after $\tilde{e}_1$, $\tilde{e}_2$, and $\tilde{e}_3$ are replaced by zeros. That is,

$$\begin{aligned} v_0(z) &= \frac{c}{1-c} m_0(z), \\ u_0(z) &= -\frac{c}{1-c}\left(z m_0(z) + 1\right), \text{ and} \\ \tilde{v}_0(z) &= \frac{2c}{1-c}\left(z m_0(z) + 1\right). \end{aligned}$$

Following a similar strategy to that used in the above proofs of Lemma OW11 and Lemma 21, we can show that

$$\sup_{x_p \in [0,1]} \left|\mathbb{E}v(z_p) - v_0(z_p)\right| = o_{\text{a.s.}}(y_p) \tag{130}$$

$$\sup_{x_p \in [0,1]} \left|\mathbb{E}u(z_p) - u_0(z_p)\right| = o_{\text{a.s.}}(y_p), \text{ and} \tag{131}$$

$$\sup_{x_p \in [0,1]} \left|\mathbb{E}\tilde{v}(z_p) - \tilde{v}_0(z_p)\right| = o_{\text{a.s.}}(y_p). \tag{132}$$

We give most important details of the derivation of the first of the above three equations in the next section of this note. The other two equations can be derived similarly.

Equations (130-132) imply that the boundedness of $\max_{j=1,\ldots,T/2}\sup_{x_p\in[a,b]}\left\|\Omega_j^{(d)}\right\|$ would follow from the boundedness of $\max_{j=1,\ldots,T/2}\sup_{x_p\in[a,b]}\|\Omega_{j0}(z_p)\|$, where

$$\begin{aligned}\Omega_{j0}(z) &= \begin{pmatrix}\left(\frac{1}{1-z}+v_0\right)I_2 & \frac{r_j}{1-z}\nabla_j'+u_0I_2\\ \frac{r_j}{1-z}\nabla_j+u_0I_2 & \left(\frac{r_jz}{1-z}+z\tilde{v}_0\right)I_2\end{pmatrix}^{-1}\\ &= \frac{1}{\delta_{j0}}\begin{pmatrix}zr_jI_2+z(1-z)\tilde{v}_0I_2 & -r_j\nabla_j'-(1-z)u_0I_2\\ -r_j\nabla_j-(1-z)u_0I_2 & I_2+(1-z)v_0I_2\end{pmatrix}\end{aligned}$$

with

$$\delta_{j0}\equiv\delta_{j0}(z)=(1-z)\left(z\tilde{v}_0v_0-u_0^2\right)+z\tilde{v}_0+r_j\left(u_0+zv_0-1\right).$$

Since $\sup_{x_p\in[a,b]}|v_0|$, $\sup_{x_p\in[a,b]}|u_0|$, and $\sup_{x_p\in[a,b]}|\tilde{v}_0|$ are bounded, it is sufficient to show that $\inf_{x_p\in[a,b]}\min_j|\delta_{j0}(z_p)|$ is bounded away from zero.

Lemma 22 *For any $c<1/2$, there exists a positive ϵ such that* $\inf_{z\in\mathbb{C}^+}\min_j|\delta_{j0}(z)|>\epsilon$.

Proof: Denote $c/(1-c)$ as $\hat{C}$. Then by (OW31) we have

$$\left(1+\hat{C}\right)\hat{C}+\left(z-\hat{C}\right)v_0+z(1-z)v_0^2=0. \tag{133}$$

Let us define $V_0=zv_0$. Then, by (133), for any $z\in\mathbb{C}^+$, we have

$$\left(1+\hat{C}\right)\hat{C}z+\left(z-\hat{C}\right)V_0+(1-z)V_0^2=0 \tag{134}$$

and thus,

$$z=\frac{\left(\hat{C}-V_0\right)V_0}{\left(\hat{C}-V_0+1\right)\left(\hat{C}+V_0\right)} \tag{135}$$

and

$$\delta_{j0}=-\frac{\left(\hat{C}-V_0\right)^2}{\hat{C}-V_0+1}\left(\hat{C}+1\right)-r_j\left(\hat{C}+1\right).$$

If V_0 is a zero of δ_{j0} for some $r_j\in[0,4]$, then we must have $-\frac{(\hat{C}-V_0)^2}{\hat{C}-V_0+1}\in[0,4]$, or, equivalently,

$$b+b^2\in(-\infty,-1/4],$$

where $b\equiv\left(\hat{C}-V_0\right)^{-1}$. This implies that $\operatorname{Re}b=-1/2$. On the other hand, (135) yields

$$z=\frac{\hat{C}b-1}{(1+b)\left(2\hat{C}b-1\right)}.$$

For such a z to belong to $\mathbb{C}^+$, we must have $\operatorname{Im}b\le 0$, and therefore, $\operatorname{Im}V_0\le 0$. But $V_0=zv_0=\hat{C}zm_0$, where

$$V_0=zv_0=\int\frac{\hat{C}z}{\lambda-z}\mathrm{d}W_c(\lambda),$$

Let $z=x+\mathrm{i}y$ with $y>0$. Then,

$$V_0=\int\frac{\hat{C}(x+\mathrm{i}y)(\lambda-x+\mathrm{i}y)}{|\lambda-z|^2}\mathrm{d}W_c(\lambda)$$

so that

$$\operatorname{Im} V_0 = \hat{C} \int \frac{\lambda y}{|\lambda - z|^2} \mathrm{d}W_c(\lambda) > 0$$

and therefore, $\delta_{0j} \neq 0$ for any $z \in \mathbb{C}^+$ and $r_j \in [0, 4]$.

It remains a possibility that there is a sequence $\{z_k\} \in \mathbb{C}^+$ such that the corresponding $\delta_{0j}(z_k)$ converge to zero. Let us show that this is not the case. Indeed, by (OW32) and since $v_0 = \hat{C} m_0$,

$$V_0(z) = \frac{-(z - c - cz) + \sqrt{(z - c - cz)^2 - 4c(1 - z)z}}{2(1 - z)(1 - c)}, \tag{136}$$

where we choose the branch of the square root, with the cut along the positive real semi-axis, which has positive imaginary part. This implies that $V_0(z)$, and $\delta_{0j}(z)$, can be extended to a continuous function over $z \in \mathbb{C}^+ \cup \mathbb{R}$. Note that $z = 1$ is not causing problems for $c < 1/2$ because the support of $W_c(\lambda)$ is bounded away from 1 for such c. It is, thus, sufficient to show that $\delta_{0j}(z) \neq 0$ for $z \in \mathbb{R}$. Note that, by continuity, $V_0(z)$ still satisfies (134) for $z \in \mathbb{R}$. Hence, the only possible way to have $\delta_{0j}(z) = 0$ for $z \in \mathbb{R}$ is to have $\left(\hat{C} - V_0\right)^{-1} = -1/2$, that is, $V_0 = \hat{C} + 2$. Using (135), we find that $z = \left(\hat{C} + 2\right)/(\hat{C} + 1) = 2 - c$. But then, (136) implies that $V_0 = -\hat{C}\left(1 + \hat{C}\right) \neq \hat{C} + 2$, that is, $V_0 = \hat{C} + 2$ is the "wrong" root of the quadratic equation (134). Therefore, $\delta_{0j}(z) \neq 0$ for $z \in \mathbb{R}$.□

4.3.3 Details of a proof of (130) (about the a.s. convergence of $y_p^{-1}(\mathbb{E}v(z_p) - v_0(z_p))$)

Our proof closely follows the logic of the proofs of Lemma OW10 and Lemma 21 above. We skip most of the details, and emphasize the differences. First, note that

$$v - v_0 = \frac{c}{1 - c}(m - m_0) - \frac{c}{1 - c}\tilde{e}_3.$$

Recall that (see equation (115) above) $|m - m_0| \leq C p^{-\gamma} y_p^{d - 42.5}$ w.ow.p. for any $(C, d, \gamma) \in (0, \infty) \times [42, \infty) \times [0, 1/2)$ and any $\alpha \in [0, \alpha_{\gamma d})$. Combining this with Lemma OW10, we conclude that event

$$\mathcal{E}_{vz} = \left\{|v - v_0| \leq p^{-\gamma} y_p^{d - 42.5}\right\} \text{ holds w.ow.p.}$$

Since $v(z_p) = \frac{1}{T} \operatorname{tr}\left(A^{-1}\left(C' D^{-1} C A^{-1} - z_p I_p\right)^{-1}\right)$ and $\left\|\left(C' D^{-1} C A^{-1} - z_p I_p\right)^{-1}\right\| \leq 1/y_p$, we always have

$$|v - v_0| \leq |v| + |v_0| = |v| + \frac{c}{1 - c}|m_0| \leq 1/\left(\mu_{\min,0} y_p\right) + 1/y_p,$$

where $\mu_{\min,0}$ is the smallest eigenvalue of A. Therefore, always,

$$y_p^{-1}|v - v_0| \leq p^{-\gamma} y_p^{d - 43.5} + \left(1 + \mu_{\min,0}^{-1}\right) y_p^{-2} \mathbf{1}_{\mathcal{E}_{vz}^c},$$

and

$$\Pr\left(y_p^{-1} \max_{x_p \in S_p} |v - v_0| > \epsilon\right) \leq \Pr\left(p^{-\gamma} y_p^{d - 43.5} > \epsilon/2\right) + \Pr\left(\left(1 + \mu_{\min,0}^{-1}\right) y_p^{-2} \max_{x_p \in S_p} \mathbf{1}_{\mathcal{E}_{vz}^c} > \epsilon/2\right).$$

In contrast to the upper bound on $\Pr\left(y_p^{-1} \max_{x_p \in S_p} |m - m_0| > \epsilon\right)$ derived in (117), the above upper bound on $\Pr\left(y_p^{-1} \max_{x_p \in S_p} |v - v_0| > \epsilon\right)$ depend on $\mu_{\min,0}$.

Let us choose $d = 43.5$ and $\gamma = 1/180$ so that $\alpha_{\gamma d} \equiv (1/2 - \gamma)/(1 + d) = 1/90$. Then, for any $l > 0$ and $\rho \geq 180l$, and any $\epsilon > 0$,

$$\Pr\left(p^{-\gamma} y_p^{d - 43.5} > \epsilon/2\right) \leq 2^\rho \epsilon^{-\rho} p^{-l},$$

and we have

$$\begin{aligned} \Pr\left(y_p^{-1} \max_{x_p \in S_p} |v - v_0| > \epsilon\right) &\leq 2^\rho \epsilon^{-\rho} p^{-l} + \Pr\left(\left(1 + \mu_{\min,0}^{-1}\right) y_p^{-2} \max_{x_p \in S_p} \mathbf{1}_{\mathcal{E}_{vz}^c} > \epsilon/2\right) \\ &= 2^\rho \epsilon^{-\rho} p^{-l} + \Pr\left(\left(1 + \mu_{\min,0}^{-1}\right) y_p^{-2} > \epsilon/2 \text{ and } \max_{x_p \in S_p} \mathbf{1}_{\mathcal{E}_{vz}^c} = 1\right). \end{aligned} \tag{137}$$

Following the logic of the proofs of Lemma OW11 and Lemma 21, we would like to use (137) to show that

$$\Pr\left(y_p^{-1}\sup_{x_p\in[0,1]}|v-v_0|>\epsilon\right)\le C\epsilon^{-\rho}p^{-l}$$

and then 'convert' this into a bound on the expectation by applying Lemma 4. For such a strategy to work, the latter probability bound must be established *for all* $\epsilon>0$. This necessitates an analysis of the lower tail of the distribution of $\mu_{\min,0}$ (note how $\mu_{\min,0}$ enters the right hand side of (137)).

By Lemma 5(i)

$$\Pr\left(\mu_{\min,0}\le\mu\right)\le\Pr\left(\mu_{\min,j}\le\mu\right)<\mu^{2\rho}\left(2e\right)^{4\rho}$$

for all $\mu>0$ and all sufficiently large p,T. Therefore, when $\epsilon y_p^2/4>1$,

$$\Pr\left(\left(1+\mu_{\min,0}^{-1}\right)y_p^{-2}>\epsilon/2\right)\le\Pr\left(\mu_{\min,0}\le\left(\epsilon y_p^2/2-1\right)^{-1}\right)<\left(\frac{\epsilon y_p^2/2-1}{4e^2}\right)^{-2\rho}<\left(\frac{\epsilon y_p^2}{16e^2}\right)^{-2\rho}$$

When $\epsilon y_p^2/4\le1$, we obviously have

$$\Pr\left(\left(1+\mu_{\min,0}^{-1}\right)y_p^{-2}>\epsilon/2\right)\le\left(\epsilon y_p^2/4\right)^{-2\rho}.$$

Hence, in any case,

$$\Pr\left(\left(1+\mu_{\min,0}^{-1}\right)y_p^{-2}>\epsilon/2\right)\le\left(16e^2\right)^{2\rho}\left(\epsilon y_p^2\right)^{-2\rho}$$

Using this, we obtain

$$\begin{aligned}
&\Pr\left(\left(1+\mu_{\min,0}^{-1}\right)y_p^{-2}>\epsilon/2\text{ and }\max_{x_p\in S_p}\mathbf{1}_{\mathcal{E}_{vz}^c}=1\right)\\
\le\;&\left(\Pr\left(\left(1+\mu_{\min,0}^{-1}\right)y_p^{-2}>\epsilon/2\right)\right)^{1/2}\left(\Pr\left(\max_{x_p\in S_p}\mathbf{1}_{\mathcal{E}_{vz}^c}=1\right)\right)^{1/2}\\
\le\;&C\left(\epsilon y_p^2\right)^{-\rho}\left(\Pr\left(\max_{x_p\in S_p}\mathbf{1}_{\mathcal{E}_{vz}^c}=1\right)\right)^{1/2},
\end{aligned}$$

and similarly to (119),

$$\Pr\left(\left(1+\mu_{\min,0}^{-1}\right)y_p^{-2}>\epsilon/2\text{ and }\max_{x_p\in S_p}\mathbf{1}_{\mathcal{E}_{vz}^c}=1\right)\le C\epsilon^{-\rho}p^{-l}$$

for any $l>0$, $\rho>0$, and all sufficiently large p,T, where C may depend on α,l, and ρ.

Recalling inequality (137), we conclude that for any $l>0$ and $\rho\ge180l$, and any $\epsilon>0$,

$$\Pr\left(y_p^{-1}\max_{x_p\in S_p}|v-v_0|>\epsilon\right)\le C\epsilon^{-\rho}p^{-l}$$

for all sufficiently large p,T, where C may depend on α,l, and ρ. This inequality is an equivalent of inequality (120) in the proof of Lemma OW11. Using similar ideas and following the logic of the proofs of Lemma OW11 and Lemma 21, we arrive at

$$\sup_{x_p\in[0,1]}\left|\mathbb{E}v\left(z_p\right)-v_0\left(z_p\right)\right|=o_{\text{a.s.}}\left(y_p\right).$$

4.3.4 Proof of the decomposition (OW40) ($m - \mathbb{E}m = W_1 + W_2 + W_3 + W_4$)

From equation (OW39), we have

$$\begin{aligned} m - \mathbb{E}m &= \frac{1}{p}\sum_{j=1}^{T/2} (\mathbb{E}_j - \mathbb{E}_{j-1}) \operatorname{tr}\left(\Gamma_j^{(q)}\Omega_j^{(d)}\right) \\ &+ \frac{1}{p}\sum_{j=1}^{T/2} (\mathbb{E}_j - \mathbb{E}_{j-1}) \operatorname{tr}\left(\Gamma_j^{(q)}\Omega_j^{(d)}\left((\Omega_j^{(d)})^{-1} - (\Omega_j^{(q)})^{-1}\right)\Omega_j^{(d)}\right) + W_4. \end{aligned} \tag{138}$$

Focus on the first term on the right hand side of (138). Observe that

$$(\mathbb{E}_j - \mathbb{E}_{j-1})\,\hat{\Gamma}_j = 0 \tag{139}$$

because $\hat{\Gamma}_j$ does not depend on $\varepsilon_{(j)}$. Let $\mathbb{E}_{-j}$ be the expectation conditional on $\varepsilon_{(i)}, i \neq j, i = 1, ..., T/2$. Then

$$\mathbb{E}_{j-1}\left(v_j^{(q)} - v_j I_2\right) = \mathbb{E}_{j-1}\mathbb{E}_{-j}\left(v_j^{(q)} - v_j I_2\right) = 0,$$

and similarly $\mathbb{E}_{j-1}\left(u_j^{(q)} - u_j I_2\right) = 0$, $\mathbb{E}_{j-1}\left(\tilde{v}_j^{(q)} - \tilde{v}_j I_2\right) = 0$, $\mathbb{E}_{j-1}\left(a_j^{(q)} - a_j I_2\right) = 0$, etc. These equalities together with (139) yield

$$\frac{1}{p}\sum_{j=1}^{T/2} (\mathbb{E}_j - \mathbb{E}_{j-1}) \operatorname{tr}\left(\Gamma_j^{(q)}\Omega_j^{(d)}\right) = \frac{1}{p}\sum_{j=1}^{T/2} \mathbb{E}_j \operatorname{tr}\left(\left(\Gamma_j^{(q)} - \hat{\Gamma}_j\right)\Omega_j^{(d)}\right) = W_1.$$

Next, the second term on the right hand side of (138) can be decomposed into the following sum

$$\frac{1}{p}\sum_{j=1}^{T/2} (\mathbb{E}_j - \mathbb{E}_{j-1}) \operatorname{tr}\left(\hat{\Gamma}_j\Omega_j^{(d)}\left((\Omega_j^{(d)})^{-1} - (\Omega_j^{(q)})^{-1}\right)\Omega_j^{(d)}\right) + W_3.$$

Since $\hat{\Gamma}_j\Omega_j^{(d)}\left((\Omega_j^{(d)})^{-1} - (\hat{\Omega}_j)^{-1}\right)\Omega_j^{(d)}$ does not depend on $\varepsilon_{(j)}$, we have

$$(\mathbb{E}_j - \mathbb{E}_{j-1}) \operatorname{tr}\left(\hat{\Gamma}_j\Omega_j^{(d)}\left((\Omega_j^{(d)})^{-1} - (\hat{\Omega}_j)^{-1}\right)\Omega_j^{(d)}\right) = 0$$

and

$$(\mathbb{E}_j - \mathbb{E}_{j-1}) \operatorname{tr}\left(\hat{\Gamma}_j\Omega_j^{(d)}\left((\Omega_j^{(d)})^{-1} - (\Omega_j^{(q)})^{-1}\right)\Omega_j^{(d)}\right) = (\mathbb{E}_j - \mathbb{E}_{j-1}) \operatorname{tr}\left(\hat{\Gamma}_j\Omega_j^{(d)}\left((\hat{\Omega}_j)^{-1} - (\Omega_j^{(q)})^{-1}\right)\Omega_j^{(d)}\right).$$

On the other hand, since $\hat{\Gamma}_j$ does not depend on $\varepsilon_{(j)}$ and $\Omega_j^{(d)}$ is deterministic, we have

$$\begin{aligned} \mathbb{E}_{j-1} \operatorname{tr}\left(\hat{\Gamma}_j\Omega_j^{(d)}\left((\hat{\Omega}_j)^{-1} - (\Omega_j^{(q)})^{-1}\right)\Omega_j^{(d)}\right) &= \mathbb{E}_{j-1}\mathbb{E}_{-j} \operatorname{tr}\left(\hat{\Gamma}_j\Omega_j^{(d)}\left((\hat{\Omega}_j)^{-1} - (\Omega_j^{(q)})^{-1}\right)\Omega_j^{(d)}\right) \\ &= \mathbb{E}_{j-1} \operatorname{tr}\left(\hat{\Gamma}_j\Omega_j^{(d)}\mathbb{E}_{-j}\left((\hat{\Omega}_j)^{-1} - (\Omega_j^{(q)})^{-1}\right)\Omega_j^{(d)}\right) = 0. \end{aligned}$$

Hence,

$$\begin{aligned} &\frac{1}{p}\sum_{j=1}^{T/2} (\mathbb{E}_j - \mathbb{E}_{j-1}) \operatorname{tr}\left(\hat{\Gamma}_j\Omega_j^{(d)}\left((\Omega_j^{(d)})^{-1} - (\Omega_j^{(q)})^{-1}\right)\Omega_j^{(d)}\right) \\ = &\frac{1}{p}\sum_{j=1}^{T/2} \mathbb{E}_j \operatorname{tr}\left(\hat{\Gamma}_j\Omega_j^{(d)}\left((\hat{\Omega}_j)^{-1} - (\Omega_j^{(q)})^{-1}\right)\Omega_j^{(d)}\right) = W_2. \end{aligned}$$

4.3.5 Proof of Proposition OW14 (a.s. convergence of $py_p |m - \mathbb{E}m|$)

We split the proof in four parts corresponding to the terms $W_1, ..., W_4$ in the decomposition (OW40) of $m - \mathbb{E}m$.

Analysis of W_1. Recall the definition of $\Omega_j^{(d)}$

$$\Omega_j^{(d)} = \begin{pmatrix} \left(\frac{1}{1-z} + \mathbb{E}v\right) I_2 & \frac{r_j}{1-z}\nabla_j' + \mathbb{E}uI_2 \\ \frac{r_j}{1-z}\nabla_j + \mathbb{E}uI_2 & \left(\frac{r_j z}{1-z} + z\mathbb{E}\tilde{v}\right) I_2 \end{pmatrix}^{-1}.$$

A more explicit form of this matrix is

$$\Omega_j^{(d)} = \frac{1}{\delta_j^{(d)}} \begin{pmatrix} zr_j I_2 + z(1-z)\mathbb{E}\tilde{v}I_2 & -r_j\nabla_j' - (1-z)\mathbb{E}uI_2 \\ -r_j\nabla_j - (1-z)\mathbb{E}uI_2 & I_2 + (1-z)\mathbb{E}vI_2 \end{pmatrix} \tag{140}$$

with

$$\delta_j^{(d)} \equiv (1-z)\left(z\mathbb{E}\tilde{v}\mathbb{E}v - (\mathbb{E}u)^2\right) + z\mathbb{E}\tilde{v} + r_j(\mathbb{E}u + z\mathbb{E}v - 1).$$

Equation (140) and the definitions of $\Gamma_j^{(q)}$ and $\hat{\Gamma}_j$ yield

$$\left(\Gamma_j^{(q)} - \hat{\Gamma}_j\right)\Omega_j^{(d)} = \begin{pmatrix} v_j^{(q)} - v_j I_2 \\ u_j^{(q)} - u_j I_2 \end{pmatrix}\Psi_j^{(d)} - \begin{pmatrix} a_j^{(q)} - a_j I_2 & b_j^{(q)\prime} - b_j I_2 \\ b_j^{(q)} - b_j I_2 & c_j^{(q)} - c_j I_2 \end{pmatrix}\Omega_j^{(d)}$$

where

$$\Psi_j^{(d)} = (\delta_j^{(d)})^{-1}\begin{pmatrix} (z\mathbb{E}\tilde{v} - r_j) I_2 - r_j\mathbb{E}u\nabla_j' & r_j\mathbb{E}v\nabla_j' - \mathbb{E}uI_2 \end{pmatrix}.$$

Both $\Psi_j^{(d)}$ and $\Omega_j^{(d)}$ are deterministic matrices. Furthermore, as follows from the proof of Lemma OW13, their entries are bounded by absolute value. Hence, the elements of $\left(\Gamma_j^{(q)} - \hat{\Gamma}_j\right)\Omega_j^{(d)}$ are linear combinations with bounded weights of random variables $\mathcal{M}_j$ that may be equal to any entry of any of the matrices $v_j^{(q)} - v_j I_2$, $u_j^{(q)} - u_j I_2$, $a_j^{(q)} - a_j I_2$, $b_j^{(q)} - b_j I_2$, or $c_j^{(q)} - c_j I_2$. Therefore, to show that $\max_{x_p \in S_p} |py_p W_1| \overset{\text{a.s.}}{\to} 0$, it is sufficient to prove that

$$\max_{x_p \in S_p}\left|y_p \sum_{j=1}^{T/2}\mathbb{E}_j\mathcal{M}_j\right| \overset{\text{a.s.}}{\to} 0. \tag{141}$$

Note that all $\mathcal{M}_j$ have form $\xi'\mathcal{W}\xi - \frac{1}{T}\operatorname{tr}\mathcal{W}$, where ξ is a high-dimensional Gaussian vector, independent from $\mathcal{W}$ and having i.i.d. entries with variance $1/T$. For example, the first row and first column entry of $c_j^{(q)} - c_j I_2$ has form

$$c_{j,11}^{(q)} - c_j \equiv \varepsilon_{2j-1}' D_j^{-1} C_j' M_j^{-1} A_j M_j^{-1} C_j D_j^{-1}\varepsilon_{2j-1} - \frac{1}{T}\operatorname{tr} D_j^{-1} C_j' M_j^{-1} A_j M_j^{-1} C_j D_j^{-1} \tag{142}$$

and its first row and second column entry has form

$$c_{j,12}^{(q)} \equiv \left(\varepsilon_{2j-1}', \varepsilon_{2j}'\right)\begin{pmatrix} 0 & \frac{1}{2}D_j^{-1}C_j'M_j^{-1}A_jM_j^{-1}C_jD_j^{-1} \\ \frac{1}{2}D_j^{-1}C_j'M_j^{-1}A_jM_j^{-1}C_jD_j^{-1} & 0 \end{pmatrix}\begin{pmatrix} \varepsilon_{2j-1} \\ \varepsilon_{2j} \end{pmatrix}.$$

By Lemma 2.7 from BS98, we have for any $\rho > 0$

$$\mathbb{E}_{\mathcal{W}}\left|\xi'\mathcal{W}\xi - \frac{1}{T}\operatorname{tr}\mathcal{W}\right|^{\rho} \le C\left(\operatorname{tr}\mathcal{W}\mathcal{W}^*\right)^{\rho/2} T^{-\rho}, \tag{143}$$

where $\mathbb{E}_{\mathcal{W}}$ denotes expectation conditional on $\mathcal{W}$, ξ is independent from $\mathcal{W}$ and distributed as $N(0, I_p/T)$, and constant C may depend on ρ.

Let us prove (141) for $\mathcal{M}_j$ equal to the expression in (142). Proofs for other possible $\mathcal{M}_j$ are very similar and we omit them. Let F_{pj} be the empirical distribution of the eigenvalues of $C_j D_j^{-1} C_j' A_j$. Then, by the so-called rank inequality (see e.g. Lemma 2.12 of BS98) and by equations (OW35), we have

$$\max_{k=0,\dots,T/2} \max_{j=1,\dots,T/2} \mathbb{E}_k \left(F_{pj}\left([a',b']\right)\right)^2 = o_{\text{a.s.}}\left(p^{-2/91}\right) = o_{\text{a.s.}}\left(y_p^{10}\right), \tag{144}$$

where the last equality holds because we assume in this section that $y_p = y_0 p^{-1/456}$. Define functions

$$\mathcal{B}_j = \mathbf{1}\left\{\mathbb{E}_{j-1}\left(F_{pj}\left([a',b']\right)\right)^2 \le y_p^{10}\right\} = \mathbf{1}\left\{\mathbb{E}_j\left(F_{pj}\left([a',b']\right)\right)^2 \le y_p^{10}\right\},$$

where $\mathbf{1}\{\cdot\}$ is the indicator function. Note that $\mathbb{E}_{j-1}\mathcal{B}_j = \mathbb{E}_j\mathcal{B}_j = \mathcal{B}_j$.

Let i.o. abbreviate "infinitely often". Equation (144) implies that $\Pr\left(\bigcup_{j=1}^{T/2}[\mathcal{B}_j = 0] \text{ i.o.}\right) = 0$ and therefore, for any $\epsilon > 0$,

$$\begin{aligned}
&\Pr\left(\max_{x_p\in S_p}\left|y_p\sum_{j=1}^{T/2}\mathbb{E}_j\left(c_{j,11}^{(q)} - c_j\right)\right| > \epsilon \text{ i.o.}\right)\\
\le\ &\Pr\left(\left(\left[\max_{x_p\in S_p}\left|y_p\sum_{j=1}^{T/2}\mathbb{E}_j\left(c_{j,11}^{(q)} - c_j\right)\right| > \epsilon\right] \cap \bigcap_{j=1}^{T/2}[\mathcal{B}_j = 1]\right) \cup \bigcup_{j=1}^{T/2}[\mathcal{B}_j = 0] \text{ i.o.}\right)\\
\le\ &\Pr\left(\max_{x_p\in S_p}\left|y_p\sum_{j=1}^{T/2}\mathbb{E}_j\mathcal{B}_j\left(c_{j,11}^{(q)} - c_j\right)\right| > \epsilon \text{ i.o.}\right).
\end{aligned}$$

Note that, for each $x_p \in \mathbb{R}$, $\mathbb{E}_j\mathcal{B}_j\left(c_{j,11}^{(q)} - c_j\right)$ forms a martingale difference sequence. Therefore, by Burkholder's lemma (e.g. Lemma 2.1 in BS98), for any $x_p \in [a,b]$ and $\rho \ge 2$

$$\begin{aligned}
\mathbb{E}\left|y_p\sum_{j=1}^{T/2}\mathbb{E}_j\mathcal{B}_j\left(c_{j,11}^{(q)} - c_j\right)\right|^\rho &\le C\left(\mathbb{E}\left(\sum_{j=1}^{T/2}\mathbb{E}_{j-1}\left|y_p\mathbb{E}_j\mathcal{B}_j\left(c_{j,11}^{(q)} - c_j\right)\right|^2\right)\right)^{\rho/2}\\
&\quad + C\mathbb{E}\left(\sum_{j=1}^{T/2}\left|y_p\mathbb{E}_j\mathcal{B}_j\left(c_{j,11}^{(q)} - c_j\right)\right|^\rho\right).
\end{aligned} \tag{145}$$

Using (143) and Lemma 3, we then have

$$\mathbb{E}_{j-1}\left|y_p\mathbb{E}_j\mathcal{B}_j\left(c_{j,11}^{(q)} - c_j\right)\right|^2 \le Cy_p^2p^{-2}\mathbb{E}_{j-1}\left\{\mathcal{B}_j\mu_{\max,j}^2\mu_{\min,j}^{-4}\operatorname{tr}\left(\left(A_jM_j^{-1}\right)^2\left(M_j^{*-1}A_j\right)^2\right)\right\}.$$

Denote the eigenvalues of $C_jD_j^{-1}C_j'A_j^{-1}$ as $\lambda_{1,-j} \ge \dots \ge \lambda_{p,-j}$. Then, we have

$$\begin{aligned}
&\sum_{j=1}^{T/2}\mathbb{E}_{j-1}\left\{\mathcal{B}_j\mu_{\max,j}^2\mu_{\min,j}^{-4}\operatorname{tr}\left(\left(A_jM_j^{-1}\right)^2\left(M_j^{*-1}A_j\right)^2\right)\right\}\\
=\ &\sum_{j=1}^{T/2}\mathcal{B}_j\mathbb{E}_{j-1}\mu_{\max,j}^2\mu_{\min,j}^{-4}\left(\sum_{\lambda_{k,-j}\notin[a',b']}\left(\left(\lambda_{k,-j} - x_p\right)^2 + y_p^2\right)^{-2} + \sum_{\lambda_{k,-j}\in[a',b']}\left(\left(\lambda_{k,-j} - x_p\right)^2 + y_p^2\right)^{-2}\right)\\
\le\ &\sum_{j=1}^{T/2}\left(\mathbb{E}_{j-1}\mu_{\max,j}^2\mu_{\min,j}^{-4}p\underline{\epsilon}^{-4} + \mathcal{B}_j\left(\mathbb{E}_{j-1}\mu_{\max,j}^4\mu_{\min,j}^{-8}\right)^{1/2}y_p^{-4}\left(\mathbb{E}_{j-1}\left(pF_{pj}\left([a',b']\right)\right)^2\right)^{1/2}\right)\\
\le\ &\sum_{j=1}^{T/2}\left(\mathbb{E}_{j-1}\mu_{\max,j}^2\mu_{\min,j}^{-4}p\underline{\epsilon}^{-4} + \left(\mathbb{E}_{j-1}\mu_{\max,j}^4\mu_{\min,j}^{-8}\right)^{1/2}py_p\right) \le Cp\sum_{j=1}^{T/2}\left(\mathbb{E}_{j-1}\mu_{\max,j}^4\mu_{\min,j}^{-8}\right)^{1/2}.
\end{aligned}$$

On the other hand, by Hölder's inequality, for any $\tau \geq 2$ and non-negative x_j,

$$\left(\sum_{j=1}^{T/2} x_j\right)^{\tau/2} \leq \sum_{j=1}^{T/2} x_j^{\tau/2} (T/2)^{\tau/2-1}, \tag{146}$$

so that

$$\left(\sum_{j=1}^{T/2}\left(\mathbb{E}_{j-1}\mu_{\max,j}^4\mu_{\min,j}^{-8}\right)^{1/2}\right)^{\rho/2} \leq Cp^{\rho/2-1}\sum_{j=1}^{T/2}\left(\mathbb{E}_{j-1}\mu_{\max,j}^4\mu_{\min,j}^{-8}\right)^{\rho/4} \leq Cp^{\rho/2-1}\sum_{j=1}^{T/2}\mathbb{E}_{j-1}\mu_{\max,j}^{\rho}\mu_{\min,j}^{-2\rho}$$

for any $\rho \geq 4$. Therefore, for the first term on the right hand side of (145), we have

$$\begin{aligned}\left(\mathbb{E}\left(\sum_{j=1}^{T/2}\mathbb{E}_{j-1}\left|y_p\mathbb{E}_j\mathcal{B}_j\left(c_{j,11}^{(q)}-c_j\right)\right|^2\right)\right)^{\rho/2} &\leq Cy_p^{\rho}p^{-1}\sum_{j=1}^{T/2}\mathbb{E}\mu_{\max,j}^{\rho}\mu_{\min,j}^{-2\rho}\\ \leq\ Cy_p^{\rho}p^{-1}\sum_{j=1}^{T/2}\left(\mathbb{E}\mu_{\max,j}^{2\rho}\right)^{1/2}\left(\mathbb{E}\mu_{\min,j}^{-4\rho}\right)^{1/2} &\leq Cy_p^{\rho},\end{aligned}$$

where, as always, the value of C may be different from one appearance to another, and here C may depend on ρ. The last inequality follows from the boundedness of $\mathbb{E}\mu_{\max,j}^{2\rho}$ and $\mathbb{E}\mu_{\min,j}^{-4\rho}$, which is implied by Lemmas 5 and 6.

For the second term on the right hand side of (145), using (143) and Lemma 3, we obtain

$$\begin{aligned}\mathbb{E}\left(\sum_{j=1}^{T/2}\left|y_p\mathbb{E}_j\mathcal{B}_j\left(c_{j,11}^{(q)}-c_j\right)\right|^{\rho}\right) &\leq Cy_p^{\rho}p^{-\rho}\sum_{j=1}^{T/2}\mathbb{E}\mu_{\max,j}^{\rho}\mu_{\min,j}^{-2\rho}\left[\operatorname{tr}\left(\left(A_jM_j^{-1}\right)^2\left(M_j^{*-1}A_j\right)^2\right)\right]^{\rho/2}\\ &\leq Cy_p^{-\rho}p^{-\rho/2}\sum_{j=1}^{T/2}\mathbb{E}\mu_{\max,j}^{\rho}\mu_{\min,j}^{-2\rho} \leq Cy_p^{-\rho}p^{1-\rho/2}.\end{aligned}$$

Combining this with the previous display and recalling that $y_p = y_0p^{-1/456}$, we have for any $\rho \geq 4$

$$\mathbb{E}\left|y_p\sum_{j=1}^{T/2}\mathbb{E}_j\mathcal{B}_j\left(c_{j,11}^{(q)}-c_j\right)\right|^{\rho} \leq Cy_p^{\rho}.$$

By Markov's inequality,

$$\Pr\left(\max_{x_p\in S_p}\left|y_p\sum_{j=1}^{T/2}\mathbb{E}_j\mathcal{B}_j\left(c_{j,11}^{(q)}-c_j\right)\right| > \epsilon\right) \leq C\epsilon^{-\rho}p^2y_p^{\rho} = C\epsilon^{-\rho}p^{2-\rho/456},$$

which is summable for sufficiently large ρ. Therefore, $\max_{x_p\in S_p}|py_pW_1| \overset{\text{a.s.}}{\to} 0$.

Analysis of W_2. Analysis of W_2 is similar to that of W_1. Using the definition of $\hat{\Gamma}_j$ and equation (140), we obtain

$$\hat{\Gamma}_j\Omega_j^{(d)} = \begin{pmatrix} v_jI_2 \\ u_jI_2\end{pmatrix}\Psi_j^{(d)} - \begin{pmatrix} a_jI_2 & b_jI_2 \\ b_jI_2 & c_jI_2\end{pmatrix}\Omega_j^{(d)}.$$

Thus, the entries of $\hat{\Gamma}_j\Omega_j^{(d)}$ can be viewed as linear combinations with bounded weights of random variables γ_j that may be equal to any of the quantities v_j, u_j, a_j, b_j, and c_j. Further, the component $\hat{\Omega}_j^{-1}-(\Omega_j^{(q)})^{-1}$ of

W_2 depend only on random variables $\mathcal{M}_j$ that may be equal to any entry of any of the matrices $v_j^{(q)} - v_j I_2$, $u_j^{(q)} - u_j I_2$, $\tilde{v}_j^{(q)} - \tilde{v}_j I_2$. Hence, to show that $\max_{x_p \in S_p} |p y_p W_2| \xrightarrow{\text{a.s.}} 0$, it is sufficient to prove that

$$\max_{x_p \in S_p} \left| y_p \sum_{j=1}^{T/2} \mathbb{E}_j \left(\gamma_j \mathcal{M}_j \right) \right| \xrightarrow{\text{a.s.}} 0.$$

Take, for example, $\gamma_j = v_j$ and $\mathcal{M}_j = c_{j,11}^{(q)} - c_j$. Similarly to the above analysis of W_1, it is sufficient to prove that

$$\Pr \left(\max_{x_p \in S_p} \left| y_p \sum_{j=1}^{T/2} \mathbb{E}_j \mathcal{B}_j v_j \left(c_{j,11}^{(q)} - c_j \right) \right| > \epsilon \text{ i.o.} \right) = 0.$$

Again, by Burkholder's lemma, for any $\rho \geq 2$

$$\begin{aligned} \mathbb{E} \left| y_p \sum_{j=1}^{T/2} \mathbb{E}_j \mathcal{B}_j v_j \left(c_{j,11}^{(q)} - c_j \right) \right|^{\rho} \quad \leq \quad & C \left(\mathbb{E} \left(\sum_{j=1}^{T/2} \mathbb{E}_{j-1} \left| y_p \mathbb{E}_j \mathcal{B}_j v_j \left(c_{j,11}^{(q)} - c_j \right) \right|^2 \right) \right)^{\rho/2} \\ & + C \mathbb{E} \left(\sum_{j=1}^{T/2} \left| y_p \mathbb{E}_j \mathcal{B}_j v_j \left(c_{j,11}^{(q)} - c_j \right) \right|^{\rho} \right). \end{aligned} \tag{147}$$

Using (143) and the definition of v_j, we have

$$\begin{aligned} & \sum_{j=1}^{T/2} \mathbb{E}_{j-1} \left| y_p \mathbb{E}_j \mathcal{B}_j v_j \left(c_{j,11}^{(q)} - c_j \right) \right|^2 \\ \leq \quad & \sum_{j=1}^{T/2} C y_p^2 p^{-4} \mathbb{E}_{j-1} \left\{ \mathcal{B}_j \left| \operatorname{tr} M_j^{-1} \right|^2 \mu_{\max,j}^2 \mu_{\min,j}^{-4} \operatorname{tr} \left(\left(A_j M_j^{-1} \right)^2 \left(M_j^{*-1} A_j \right)^2 \right) \right\} \\ \leq \quad & \sum_{j=1}^{T/2} C y_p^2 p^{-3} \mathcal{B}_j \mathbb{E}_{j-1} \mu_{\max,j}^2 \mu_{\min,j}^{-6} \sum_k \left((\lambda_{k,-j} - x_p)^2 + y_p^2 \right)^{-1} \sum_k \left((\lambda_{k,-j} - x_p)^2 + y_p^2 \right)^{-2} \\ \leq \quad & \sum_{j=1}^{T/2} C y_p^2 p^{-3} \mathcal{B}_j \mathbb{E}_{j-1} \mu_{\max,j}^2 \mu_{\min,j}^{-6} \left(p \underline{\epsilon}^{-2} + y_p^{-2} p F_{pj} \left([a', b'] \right) \right) \left(p \underline{\epsilon}^{-4} + y_p^{-4} p F_{pj} \left([a', b'] \right) \right) \\ \leq \quad & \sum_{j=1}^{T/2} C y_p^2 p^{-1} \mathbb{E}_{j-1} \mu_{\max,j}^2 \mu_{\min,j}^{-6} + \sum_{j=1}^{T/2} C y_p^2 p^{-1} \mathcal{B}_j y_p^{-8} \mathbb{E}_{j-1} \mu_{\max,j}^2 \mu_{\min,j}^{-6} \left(F_{pj} \left([a', b'] \right) \right)^2. \end{aligned}$$

For the last inequality, we used the fact that

$$\left(p \underline{\epsilon}^{-2} + y_p^{-2} p F_{pj} \left([a', b'] \right) \right) \left(p \underline{\epsilon}^{-4} + y_p^{-4} p F_{pj} \left([a', b'] \right) \right) \leq C p^2 + C p^2 y_p^{-8} \left(F_{pj} \left([a', b'] \right) \right)^2.$$

Let $\delta = 1/3$. By Hölder's inequality,

$$\begin{aligned} \mathbb{E}_{j-1} \mu_{\max,j}^2 \mu_{\min,j}^{-6} \left(F_{pj} \left([a', b'] \right) \right)^2 \quad \leq \quad & \left(\mathbb{E}_{j-1} \left(\mu_{\max,j}^2 \mu_{\min,j}^{-6} \right)^7 \right)^{1/7} \left(\mathbb{E}_{j-1} \left(F_{pj} \left([a', b'] \right) \right)^{7/3} \right)^{6/7} \\ \leq \quad & \left(\mathbb{E}_{j-1} \left(\mu_{\max,j}^2 \mu_{\min,j}^{-6} \right)^7 \right)^{1/7} \left(\mathbb{E}_{j-1} \left(F_{pj} \left([a', b'] \right) \right)^2 \right)^{6/7}, \end{aligned}$$

where the last inequality follows from the fact that $F_{pj} \left([a', b'] \right) \leq 1$. Therefore

$$y_p^{-8} \mathcal{B}_j \mathbb{E}_{j-1} \mu_{\max,j}^2 \mu_{\min,j}^{-6} \left(F_{pj} \left([a', b'] \right) \right)^2 \leq y_p^{-8} \left(\mathbb{E}_{j-1} \left(\mu_{\max,j}^2 \mu_{\min,j}^{-6} \right)^7 \right)^{1/7} y_p^{60/7} \leq \left(\mathbb{E}_{j-1} \left(\mu_{\max,j}^2 \mu_{\min,j}^{-6} \right)^7 \right)^{1/7}.$$

Hence,

$$\sum_{j=1}^{T/2}\mathbb{E}_{j-1}\left|y_p\mathbb{E}_j\mathcal{B}_jv_j\left(c_{j,11}^{(q)}-c_j\right)\right|^2\leq\sum_{j=1}^{T/2}Cy_p^2p^{-1}\left(\mathbb{E}_{j-1}\left(\mu_{\max,j}^2\mu_{\min,j}^{-6}\right)^7\right)^{1/7}.$$

Using inequality (146) and the proportionality of p and T, we get

$$\begin{aligned}\left(\sum_{j=1}^{T/2}\left(\mathbb{E}_{j-1}\left(\mu_{\max,j}^2\mu_{\min,j}^{-6}\right)^7\right)^{1/7}\right)^{\rho/2} &\leq Cp^{\rho/2-1}\sum_{j=1}^{T/2}\left(\mathbb{E}_{j-1}\mu_{\max,j}^{14}\mu_{\min,j}^{-42}\right)^{\rho/14}\\ &\leq Cp^{\rho/2-1}\sum_{j=1}^{T/2}\mathbb{E}_{j-1}\mu_{\max,j}^{\rho}\mu_{\min,j}^{-3\rho}\end{aligned}$$

for any $\rho>14$. Therefore, for the first term on the right hand side of (147), we have for $\rho>14$

$$\left(\mathbb{E}\left(\sum_{j=1}^{T/2}\mathbb{E}_{j-1}\left|y_p\mathbb{E}_j\mathcal{B}_jv_j\left(c_{j,11}^{(q)}-c_j\right)\right|^2\right)\right)^{\rho/2}\leq Cy_p^{\rho}p^{-1}\sum_{j=1}^{T/2}\mathbb{E}\mu_{\max,j}^{\rho}\mu_{\min,j}^{-3\rho}\leq Cy_p^{\rho},$$

where the last inequality is implied by Lemmas 5 and 6.

For the second term on the right hand side of (147),

$$\begin{aligned}&\mathbb{E}\left(\sum_{j=1}^{T/2}\left|y_p\mathbb{E}_j\mathcal{B}_jv_j\left(c_{j,11}^{(q)}-c_j\right)\right|^{\rho}\right)\\ \leq\ & Cy_p^{\rho}p^{-2\rho}\sum_{j=1}^{T/2}\mathbb{E}\mu_{\max,j}^{\rho}\mu_{\min,j}^{-2\rho}\left[\left|\operatorname{tr}\left(M_j^{-1}\right)\right|^2\operatorname{tr}\left(\left(A_jM_j^{-1}\right)^2\left(M_j^{*-1}A_j\right)^2\right)\right]^{\rho/2}\\ \leq\ & Cy_p^{-2\rho}p^{-\rho/2}\sum_{j=1}^{T/2}\mathbb{E}\mu_{\max,j}^{\rho}\mu_{\min,j}^{-3\rho}\leq Cy_p^{-2\rho}p^{1-\rho/2}.\end{aligned}$$

Combining this with the previous display and recalling that $y_p=y_0p^{-1/456}$, we have for any $\rho>14$,

$$\mathbb{E}\left|y_p\sum_{j=1}^{T/2}\mathbb{E}_j\mathcal{B}_jv_j\left(c_{j,11}^{(q)}-c_j\right)\right|^{\rho}\leq Cy_p^{\rho}.$$

By Markov's inequality,

$$\Pr\left(\max_{x_p\in S_p}\left|y_p\sum_{j=1}^{T/2}\mathbb{E}_j\mathcal{B}_jv_j\left(c_{j,11}^{(q)}-c_j\right)\right|>\epsilon\right)\leq C\epsilon^{-\rho}p^2y_p^{\rho}=C\epsilon^{-\rho}p^{2-\rho/456},$$

which is clearly summable for sufficiently large ρ. Therefore, $\max_{x_p\in S_p}|py_pW_2|\xrightarrow{\text{a.s.}}0$.

Analysis of W_3. We need the following lemma, which is proven in the next section of this note.

Lemma 23 *For any $\rho\geq 2$ there exists $C>0$, s.t.* $\max_j\sup_{x_p\in[a,b]}\mathbb{E}\left(\left\|(\Omega_j^{(d)})^{-1}-(\Omega_j^{(q)})^{-1}\right\|^{\rho}\right)\leq Cp^{-\rho/2}y_p^{-\rho}$.

Similarly to the cases of W_1 and W_2, to establish convergence $\max_{x_p\in S_p}|py_pW_3|\xrightarrow{\text{a.s.}}0$, it is sufficient to prove that

$$\max_{x_p\in S_p}\left|y_p\sum_{j=1}^{T/2}\left(\mathbb{E}_j-\mathbb{E}_{j-1}\right)\left(\gamma_j\mathcal{M}_j\right)\right|\xrightarrow{\text{a.s.}}0,\tag{148}$$

where γ_j may be equal to any entry of any of the matrices $v_j^{(q)} - v_j I_2$, $u_j^{(q)} - u_j I_2$, $a_j^{(q)} - a_j I_2$, $b_j^{(q)} - b_j I_2$, or $c_j^{(q)} - c_j I_2$, and $\mathcal{M}_j$ may be equal to any entry of $(\Omega_j^{(d)})^{-1} - (\Omega_j^{(q)})^{-1}$. Take for example

$$\gamma_j = v_{j,11}^{(q)} - v_j \text{ and } \mathcal{M}_j = \mathbb{E}v - v_{j,11}^{(q)}$$

Since $(\mathbb{E}_j - \mathbb{E}_{j-1})\left(\gamma_j \mathcal{M}_j\right)$ is a martingale difference sequence, for any $\rho \geq 2$ and $x_p \in S_p$, we have by Burkholder inequality (see Lemma 2.2 in BS98), Hölder inequality, (143), and Lemmas 23 and 5

$$\mathbb{E}\left|y_p \sum_{j=1}^{T/2} (\mathbb{E}_j - \mathbb{E}_{j-1})\left(\gamma_j \mathcal{M}_j\right)\right|^\rho \leq C y_p^\rho \mathbb{E}\left(\sum_{j=1}^{T/2} \left|(\mathbb{E}_j - \mathbb{E}_{j-1})\left(\gamma_j \mathcal{M}_j\right)\right|^2\right)^{\rho/2}$$

$$\begin{aligned} &\leq C y_p^\rho p^{\rho/2-1} \sum_{j=1}^{T/2} \left(\mathbb{E}\left|\gamma_j\right|^{2\rho}\right)^{1/2} \left(\mathbb{E}\left|\mathcal{M}_j\right|^{2\rho}\right)^{1/2} \leq C p^{-1} \sum_{j=1}^{T/2} \left(\mathbb{E}\left|v_{j,11}^{(q)} - v_j\right|^{2\rho}\right)^{1/2} \\ &\leq C p^{-1} T^{-\rho} \sum_{j=1}^{T/2} \left(\mathbb{E}\left(\operatorname{tr} M_j^{-1} M_j^{*-1}\right)^\rho\right)^{1/2} \leq C p^{-1} T^{-\rho} p^{\rho/2} y_p^{-\rho} \sum_{j=1}^{T/2} \left(\mathbb{E}\mu_{\min,j}^{-2\rho}\right)^{1/2} \leq C p^{-\rho/2} y_p^{-\rho}, \end{aligned}$$

which implies (148).

Analysis of W_4. Let us define an event $\mathcal{E}_{\Gamma\Omega}$ as follows

$$\mathcal{E}_{\Gamma\Omega} = \left\{\max_{j=1,\ldots,T/2} \left\|\Gamma_j^{(q)}\right\| \leq C y_p^{-2} \text{ and } \max_{j=1,\ldots,T/2} \left\|\Omega_j^{(q)}\right\| \leq C y_p^{-5}\right\}$$

for some $C > 0$. As follows from Lemma OW8, the definition of $\Gamma_j^{(q)}$ and Lemma 3, $\mathcal{E}_{\Gamma\Omega}$ holds w.ow.p. Therefore, to establish convergence $\max_{x_p \in S_p} |p y_p W_4| \overset{\text{a.s.}}{\to} 0$, it is sufficient to prove that $\max_{x_p \in S_p} \left|p y_p \tilde{W}_4\right| \overset{\text{a.s.}}{\to} 0$, where

$$\tilde{W}_4 = \frac{1}{p} \sum_{j=1}^{T/2} (\mathbb{E}_j - \mathbb{E}_{j-1}) \operatorname{tr}\left(\mathbf{1}\left\{\mathcal{E}_{\Gamma\Omega}\right\} \Gamma_j^{(q)} \left(\Omega_j^{(d)} \left((\Omega_j^{(d)})^{-1} - (\Omega_j^{(q)})^{-1}\right)\right)^2 \Omega_j^{(q)}\right).$$

By Burkholder inequality (see Lemma 2.2 in BS98) for any $\rho \geq 2$

$$\mathbb{E}\left|p y_p \tilde{W}_4\right|^\rho \leq y_p^\rho \mathbb{E}\left(\sum_{j=1}^{T/2} \left|(\mathbb{E}_j - \mathbb{E}_{j-1}) \operatorname{tr}\left(\mathbf{1}\left\{\mathcal{E}_{\Gamma\Omega}\right\} \Gamma_j^{(q)} \left(\Omega_j^{(d)} \left((\Omega_j^{(d)})^{-1} - (\Omega_j^{(q)})^{-1}\right)\right)^2 \Omega_j^{(q)}\right)\right|^2\right)^{\rho/2}.$$

Recall that $\left\|\Omega_j^{(d)}\right\|$ is bounded. Therefore, by Lemma 23 and Hölder's inequality

$$\mathbb{E}\left|p y_p \tilde{W}_4\right|^\rho \leq C y_p^\rho (T/2)^{\rho/2-1} \sum_{j=1}^{T/2} y_p^{-7\rho} p^{-\rho} y_p^{-2\rho} \leq C p^{-\rho/2} y_p^{-8\rho},$$

which yields $\max_{x_p \in S_p} \left|p y_p \tilde{W}_4\right| \overset{\text{a.s.}}{\to} 0$.

4.3.6 Proof of Lemma 23 (bound on $\max_j \sup_{x_p \in [a,b]} \mathbb{E}\left\|(\Omega_j^{(d)})^{-1} - (\Omega_j^{(q)})^{-1}\right\|^\rho$)

Consider the decomposition

$$(\Omega_j^{(d)})^{-1} - (\Omega_j^{(q)})^{-1} = \left(\tilde{\Omega}_j^{-1} - (\Omega_j^{(q)})^{-1}\right) + \left((\Omega_j^{(d)})^{-1} - \tilde{\Omega}_j^{-1}\right), \tag{149}$$

where

$$\tilde{\Omega}_j^{-1} = \begin{pmatrix} \left(\frac{1}{1-z} + \mathbb{E}v_j\right) I_2 & \frac{r_j}{1-z}\nabla_j' + \mathbb{E}u_j I_2 \\ \frac{r_j}{1-z}\nabla_j + \mathbb{E}u_j I_2 & \left(\frac{r_j z}{1-z} + z\mathbb{E}\tilde{v}_j\right) I_2 \end{pmatrix}.$$

Lemma 24 *For any* $\rho \geq 2$ *there exists* $C > 0$, *s.t.* $\max_j \sup_{x_p \in [0,1]} \mathbb{E} \left\| \tilde{\Omega}_j^{-1} - (\Omega_j^{(q)})^{-1} \right\|^\rho \leq C p^{-\rho/2} y_p^{-\rho}$.

Proof: Split $\tilde{\Omega}_j^{-1} - (\Omega_j^{(q)})^{-1}$ into the following sum

$$\tilde{\Omega}_j^{-1} - (\Omega_j^{(q)})^{-1} = \left(\tilde{\Omega}_j^{-1} - \hat{\Omega}_j^{-1} \right) + \left(\hat{\Omega}_j^{-1} - (\Omega_j^{(q)})^{-1} \right).$$

Let us show that

$$\max_j \sup_{x_p \in [0,1]} \mathbb{E} \left\| \tilde{\Omega}_j^{-1} - \hat{\Omega}_j^{-1} \right\|^\rho \leq C p^{-\rho/2} y_p^{-\rho}. \tag{150}$$

It is sufficient to establish analogous bounds for each entry of $\tilde{\Omega}_j^{-1} - \hat{\Omega}_j^{-1}$. Consider, for example, the upper left entry, $\mathbb{E} v_j - v_j$. We have

$$\mathbb{E} \left| v_j - \mathbb{E} v_j \right|^\rho = \mathbb{E} \left| \sum_{i:\, i \neq j} \left(\mathbb{E}_i - \mathbb{E}_{i-1} \right) v_j \right|^\rho = \mathbb{E} \left| \frac{1}{T} \sum_{i:\, i \neq j} \left(\mathbb{E}_i - \mathbb{E}_{i-1} \right) \operatorname{tr} \left(M_j^{-1} - M_{ji}^{-1} \right) \right|^\rho,$$

where $M_{ji} = C_{ji} D_{ji}^{-1} C_{ji}' - z A_{ji}$ and

$$A_{ji} = A_j - \varepsilon_{(i)} \varepsilon_{(i)}', C_{ji} = C_j - \varepsilon_{(i)} \nabla_i' \varepsilon_{(i)}', \text{ and } D_{ji} = D_j - r_i^{-1} \varepsilon_{(i)} \varepsilon_{(i)}'.$$

Since $\left(\mathbb{E}_i - \mathbb{E}_{i-1} \right) \operatorname{tr} \left(M_j^{-1} - M_{ji}^{-1} \right)$ is a martingale difference sequence, by Burkholder inequality (see Lemma 2.2 in BS98) for any $\rho \geq 2$,

$$\mathbb{E} \left| \frac{1}{T} \sum_{i \neq j} \left(\mathbb{E}_i - \mathbb{E}_{i-1} \right) \operatorname{tr} \left(M_j^{-1} - M_{ji}^{-1} \right) \right|^\rho \leq C T^{-\rho} \mathbb{E} \left(\sum_{i \neq j} \left| \left(\mathbb{E}_i - \mathbb{E}_{i-1} \right) \operatorname{tr} \left(M_j^{-1} - M_{ji}^{-1} \right) \right|^2 \right)^{\rho/2}.$$

Further, similarly to inequality (36), we have

$$\left| \operatorname{tr} \left(M_j^{-1} - M_{ji}^{-1} \right) \right| \leq 8 / \left(y_p \mu_{\min,ji} \right)$$

where $\mu_{\min,ji}$ is the smallest eigenvalue of A_{ji}. Hence, by the Hölder inequality

$$\begin{aligned} \mathbb{E} \left| \frac{1}{T} \sum_{i \neq j} \left(\mathbb{E}_i - \mathbb{E}_{i-1} \right) \operatorname{tr} \left(M_j^{-1} - M_{ji}^{-1} \right) \right|^\rho &\leq C T^{-\rho} y_p^{-\rho} \mathbb{E} \left(\sum_{i \neq j} \mu_{\min,ji}^{-2} \right)^{\rho/2} \\ \leq \quad C T^{-\rho} y_p^{-\rho} p^{\rho/2 - 1} \sum_{i \neq j} \mathbb{E} \mu_{\min,ji}^{-\rho} &\leq C p^{-\rho/2} y_p^{-\rho}, \end{aligned}$$

where the boundedness of $\mathbb{E} \mu_{\min,ji}^{-\rho}$ can be established similarly to the boundedness of $\mathbb{E} \mu_{\min,j}^{-\rho}$ (see Lemma 5). Hence, for any $\rho \geq 2$,

$$\mathbb{E} \left| v_j - \mathbb{E} v_j \right|^\rho \leq C p^{-\rho/2} y_p^{-\rho}, \tag{151}$$

where C does not depend on j or x_p. Similar inequalities hold for the other corresponding entries of $\tilde{\Omega}_j^{-1} - \hat{\Omega}_j^{-1}$ and therefore, (150) holds.

Now, let us consider $\hat{\Omega}_j^{-1} - (\Omega_j^{(q)})^{-1}$. All entries of this matrix are "small". Take, for example its upper left entry $v_j - v_{j,11}^{(q)}$. By (143), Hölder's inequality, and Lemmas 5 and 6, for any $x_p \in [0,1]$,

$$\begin{aligned} \mathbb{E} \left| v_j - v_{j,11}^{(q)} \right|^\rho &= \mathbb{E} \mathbb{E}_{-j} \left| v_j - v_{j,11}^{(q)} \right|^\rho \leq \mathbb{E} C p^{-\rho} \left(\operatorname{tr} M_j^{-1} M_j^{*-1} \right)^{\rho/2} \\ &\leq \mathbb{E} C p^{-\rho} \mu_{\min,j}^{-\rho} p^{\rho/2 - 1} \sum_k \left(\left(\lambda_{k,-j} - x_p \right)^2 + y_p^2 \right)^{-\rho/2} \\ &\leq C p^{-\rho/2} y_p^{-\rho} \mathbb{E} \mu_{\min,j}^{-\rho} \leq C p^{-\rho/2} y_p^{-\rho}. \end{aligned} \tag{152}$$

Similar inequalities hold for the other entries of $\hat{\Omega}_j^{-1} - (\Omega_j^{(q)})^{-1}$. Hence,

$$\max_j \sup_{x_p \in [0,1]} \mathbb{E} \left\| \hat{\Omega}_j^{-1} - (\Omega_j^{(q)})^{-1} \right\|^\rho \leq C p^{-\rho/2} y_p^{-\rho}. \tag{153}$$

Combining (150) and (153) finishes the proof.□

Now, let us turn to the second term on the right hand side of (149).

Lemma 25 $\max_j \sup_{x_p \in [a,b]} \left\| (\Omega_j^{(d)})^{-1} - \tilde{\Omega}_j^{-1} \right\| \leq C p^{-1}$.

Proof: We have

$$(\Omega_j^{(d)})^{-1} - \tilde{\Omega}_j^{-1} = \begin{pmatrix} (\mathbb{E}v - \mathbb{E}v_j)\, I_2 & (\mathbb{E}u - \mathbb{E}u_j)\, I_2 \\ (\mathbb{E}u - \mathbb{E}u_j)\, I_2 & z\,(\mathbb{E}\tilde{v} - \mathbb{E}\tilde{v}_j)\, I_2 \end{pmatrix}.$$

All entries of this matrix are of order p^{-1}. Indeed, consider for example $\mathbb{E}v - \mathbb{E}v_j$. We have

$$\mathbb{E}v - \mathbb{E}v_j = -\frac{1}{T} \mathbb{E}\,\mathrm{tr}(M_j^{-1} - M^{-1}). \tag{154}$$

Further, from (18)

$$\begin{aligned} \mathbb{E}\,\mathrm{tr}(M_j^{-1} - M^{-1}) &= \mathbb{E}\,\mathrm{tr}\left\{ \Theta_j \Omega_j^{(q)} \right\} \\ &= \mathbb{E}\,\mathrm{tr}\left(\Theta_j \tilde{\Omega}_j \right) + \mathbb{E}\,\mathrm{tr}\left(\Theta_j \tilde{\Omega}_j (\tilde{\Omega}_j^{-1} - (\Omega_j^{(q)})^{-1}) \Omega_j^{(q)} \right) \end{aligned} \tag{155}$$

with

$$\Theta_j = \begin{pmatrix} \varepsilon_{(j)}' M_j^{-2} \varepsilon_{(j)} & \varepsilon_{(j)}' M_j^{-2} C_j D_j^{-1} \varepsilon_{(j)} \\ \varepsilon_{(j)}' D_j^{-1} C_j' M_j^{-2} \varepsilon_{(j)} & \varepsilon_{(j)}' D_j^{-1} C_j' M_j^{-2} C_j D_j^{-1} \varepsilon_{(j)} \end{pmatrix}.$$

Note that

$$\mathbb{E}\Theta_j = \frac{1}{T} \mathbb{E} \begin{pmatrix} \mathrm{tr}\left\{ M_j^{-2} \right\} & \mathrm{tr}\left\{ M_j^{-2} C_j D_j^{-1} \right\} \\ \mathrm{tr}\left\{ D_j^{-1} C_j' M_j^{-2} \right\} & \mathrm{tr}\left\{ D_j^{-1} C_j' M_j^{-2} C_j D_j^{-1} \right\} \end{pmatrix}$$

Now, using Lemma 3 and the definition of M_j, we obtain

$$\sup_{x_p \in [a,b]} \| \mathbb{E}\Theta_j \| \leq C T^{-1} \mathbb{E} \frac{\mu_{\max,j}}{\mu_{\min,j}^3} \left(p \underline{\epsilon}^{-2} + p F_{pj}\left([a', b'] \right) y_p^{-2} \right) \leq C. \tag{156}$$

This inequality and the fact that the entries of $\tilde{\Omega}_j$ are bounded (which is proved similarly to the boundedness of the entries of $\Omega_j^{(d)}$) imply that

$$\max_j \sup_{x_p \in [a,b]} \mathbb{E}\,\mathrm{tr}\left(\Theta_j \tilde{\Omega}_j \right) \leq C. \tag{157}$$

Further, for any $\delta \in (0,1)$ by Lemma OW8 and Lemma 24,

$$\begin{aligned} & \mathbb{E}\,\mathrm{tr}\left(\Theta_j \tilde{\Omega}_j (\tilde{\Omega}_j^{-1} - (\Omega_j^{(q)})^{-1}) \Omega_j^{(q)} \right) \\ \leq\ & C \left(\mathbb{E} \left(\|\Theta_j\|^{1+\delta} \left\| \tilde{\Omega}_j^{-1} - (\Omega_j^{(q)})^{-1} \right\|^{1+\delta} \right) \right)^{1/(1+\delta)} \left(\mathbb{E} \left\| \Omega_j^{(q)} \right\|^{(1+\delta)/\delta} \right)^{\delta/(1+\delta)} \\ \leq\ & C y_p^{-5} \left(\mathbb{E} \left(\|\Theta_j\|^{1+\delta} \left\| \tilde{\Omega}_j^{-1} - (\Omega_j^{(q)})^{-1} \right\|^{1+\delta} \right) \right)^{1/(1+\delta)} \\ \leq\ & C y_p^{-5} \left(\mathbb{E} \|\Theta_j\|^2 \right)^{1/2} \left(\mathbb{E} \left\| \tilde{\Omega}_j^{-1} - (\Omega_j^{(q)})^{-1} \right\|^{2(1+\delta)/(1-\delta)} \right)^{(1-\delta)/2(1+\delta)} \\ \leq\ & C y_p^{-6} p^{-1/2} \left(\mathbb{E} \|\Theta_j\|^2 \right)^{1/2}. \end{aligned}$$

From (143) and C_r-inequality, for any $\rho > 0$,

$$\mathbb{E}_{\mathcal{W}} \left|\xi' \mathcal{W} \xi\right|^{\rho} \leq CT^{-\rho} \left((\operatorname{tr} \mathcal{W}\mathcal{W}^{*})^{\rho/2} + |\operatorname{tr} \mathcal{W}|^{\rho} \right), \tag{158}$$

where $\mathbb{E}_{\mathcal{W}}$ denotes expectation conditional on $\mathcal{W}$, ξ is independent from $\mathcal{W}$ and distributed as $N(0, I_p/T)$. This implies that the second absolute moment of any of the elements of Θ_j is of order y_p^{-4}. Indeed, take for example the upper left element, $\varepsilon_{2j-1}' M_j^{-2} \varepsilon_{2j-1}$. Using (158), we obtain

$$\sup_{x_p \in [0,1]} \mathbb{E} \left|\varepsilon_{2j-1}' M_j^{-2} \varepsilon_{2j-1}\right|^2 \leq CT^{-2} \mathbb{E} \left(p\mu_{\min,j}^{-4} y_p^{-4} + p^2 \mu_{\min,j}^{-4} y_p^{-4} \right) \leq C y_p^{-4}. \tag{159}$$

Inequality (159) and Lemma 24 yield

$$\max_j \sup_{x_p \in [a,b]} \mathbb{E} \operatorname{tr} \left(\Theta_j \tilde{\Omega}_j (\tilde{\Omega}_j^{-1} - (\Omega_j^{(q)})^{-1}) \Omega_j^{(q)} \right) \leq C y_p^{-8} p^{-1/2} \leq C. \tag{160}$$

Using (157) and (160) in (155), we obtain

$$\max_j \sup_{x_p \in [a,b]} \left| \frac{1}{T} \mathbb{E} \operatorname{tr}(M_j^{-1} - M^{-1}) \right| \leq C p^{-1}$$

as required. Hence,

$$\max_j \sup_{x_p \in [a,b]} |\mathbb{E} v - \mathbb{E} v_j| \leq C p^{-1}. \tag{161}$$

One can similarly prove that

$$\max_j \sup_{x_p \in [a,b]} |\mathbb{E} u - \mathbb{E} u_j| \leq C p^{-1}, \tag{162}$$

$$\max_j \sup_{x_p \in [a,b]} |\mathbb{E} \tilde{v} - \mathbb{E} \tilde{v}_j| \leq C p^{-1}. \tag{163}$$

The above three displays yield the lemma.□

To finish the proof of Lemma 23, it remains to use Lemmas 24 and 25 in the decomposition (149).

4.4 Step 3: Convergence of $\mathbb{E} m - m_0$

4.4.1 Proof of Lemma OW15 (bounds on errors $\bar{e}_k$)

Two auxiliary lemmas (bounds on $\mathbb{E} \left\|\Omega_j^{(q)}\right\|^{\rho}$ and $\mathbb{E} \left(\left\|(\Omega_j^{(d)})^{-1} - (\Omega_j^{(q)})^{-1}\right\|^2 \right)$)

Lemma 26 *For any $\rho > 0$, there exists $C > 0$, s.t. $\max_j \sup_{x_p \in [a,b]} \mathbb{E} \left\|\Omega_j^{(q)}\right\|^{\rho} \leq C$.*

Proof: We have, for any $\tau > \rho$

$$\begin{aligned} \mathbb{E} \left\|\Omega_j^{(q)}\right\|^{\rho} &\leq C\mathbb{E} \left\|\Omega_j^{(q)} - \Omega_j^{(d)}\right\|^{\rho} + C \left\|\Omega_j^{(d)}\right\|^{\rho} \\ &= C\mathbb{E} \left\|\Omega_j^{(d)} \left((\Omega_j^{(d)})^{-1} - (\Omega_j^{(q)})^{-1} \right) \Omega_j^{(q)}\right\|^{\rho} + C \left\|\Omega_j^{(d)}\right\|^{\rho} \\ &\leq C \left\|\Omega_j^{(d)}\right\|^{\rho} \mathbb{E} \left(||(\Omega_j^{(d)})^{-1} - (\Omega_j^{(q)})^{-1}||^{\rho} \times ||\Omega_j^{(q)}||^{\rho} \right) + C \left\|\Omega_j^{(d)}\right\|^{\rho} \\ &\leq C \left\|\Omega_j^{(d)}\right\|^{\rho} \left(\mathbb{E} \left(||(\Omega_j^{(d)})^{-1} - (\Omega_j^{(q)})^{-1}||^{\tau} \right) \right)^{\rho/\tau} \left(\mathbb{E} \left(||\Omega_j^{(q)}||^{\tau\rho/(\tau-\rho)} \right) \right)^{(\tau-\rho)/\tau} + C \left\|\Omega_j^{(d)}\right\|^{\rho}. \end{aligned}$$

Using Lemmas OW13, OW8, and 23, we obtain

$$\max_j \sup_{x_p \in [a,b]} \mathbb{E} \left\|\Omega_j^{(q)}\right\|^{\rho} \leq C \max_j \sup_{x_p \in [a,b]} \left\|\Omega_j^{(d)}\right\|^{\rho} \left(p^{-\rho/2} y_p^{-6\rho} + 1 \right) \leq C. \square$$

Lemma 27 *There exists* $C > 0$, *s.t.* $\max_j \sup_{x_p \in [a,b]} \mathbb{E}\left(\left\|(\Omega_j^{(d)})^{-1} - (\Omega_j^{(q)})^{-1}\right\|^2\right) \le Cp^{-1}$.

Proof: As follows from Lemma 25, it is sufficient to prove that

$$\max_j \sup_{x_p \in [a,b]} \mathbb{E}\left(\left\|\tilde{\Omega}_j^{-1} - (\Omega_j^{(q)})^{-1}\right\|^2\right) \le Cp^{-1}.$$

As in the proof of Lemma 24, split $\tilde{\Omega}_j^{-1} - (\Omega_j^{(q)})^{-1}$ into two parts

$$\tilde{\Omega}_j^{-1} - (\Omega_j^{(q)})^{-1} = \left(\tilde{\Omega}_j^{-1} - \hat{\Omega}_j^{-1}\right) + \left(\hat{\Omega}_j^{-1} - (\Omega_j^{(q)})^{-1}\right).$$

Consider the upper left entry of $\tilde{\Omega}_j^{-1} - \hat{\Omega}_j^{-1}$, that is $\mathbb{E}v_j - v_j$. We have

$$\mathbb{E}\left|v_j - \mathbb{E}v_j\right|^2 = \mathbb{E}\left|\sum_{i:\ i\neq j} (\mathbb{E}_i - \mathbb{E}_{i-1})\, v_j\right|^2 = \mathbb{E}\left|\frac{1}{T}\sum_{i:\ i\neq j} (\mathbb{E}_i - \mathbb{E}_{i-1}) \operatorname{tr}\left(M_j^{-1} - M_{ji}^{-1}\right)\right|^2.$$

Since $(\mathbb{E}_i - \mathbb{E}_{i-1}) \operatorname{tr}\left(M_j^{-1} - M_{ji}^{-1}\right)$ is a martingale difference sequence, by Burkholder inequality (see Lemma 2.2 in BS98),

$$\mathbb{E}\left|\frac{1}{T}\sum_{i\neq j} (\mathbb{E}_i - \mathbb{E}_{i-1}) \operatorname{tr}\left(M_j^{-1} - M_{ji}^{-1}\right)\right|^2 \le CT^{-2}\mathbb{E}\sum_{i\neq j}\left|(\mathbb{E}_i - \mathbb{E}_{i-1}) \operatorname{tr}\left(M_j^{-1} - M_{ji}^{-1}\right)\right|^2.$$

Similarly to (18), we have

$$M_j^{-1} = M_{ji}^{-1} - M_{ji}^{-1}\alpha_{ji}\Omega_{ji}^{(q)}\alpha_{ji}' M_{ji}^{-1},$$

with α_{ji} and $\Omega_{ji}^{(q)}$ being obvious analogues of α_j and $\Omega_j^{(q)}$. Therefore,

$$\operatorname{tr}\left(M_j^{-1} - M_{ji}^{-1}\right) = \operatorname{tr}\left(\Theta_{ji}\Omega_{ji}^{(q)}\right) = \operatorname{tr}\left(\Theta_{ji}\tilde{\Omega}_{ji}\right) + \operatorname{tr}\left(\Theta_{ji}\tilde{\Omega}_{ji}(\tilde{\Omega}_{ji}^{-1} - (\Omega_{ji}^{(q)})^{-1})\Omega_{ji}^{(q)}\right) \tag{164}$$

with

$$\Theta_{ji} = \begin{pmatrix} \varepsilon_{(i)}' M_{ji}^{-2}\varepsilon_{(i)} & \varepsilon_{(i)}' M_{ji}^{-2} C_{ji} D_{ji}^{-1}\varepsilon_{(i)} \\ \varepsilon_{(i)}' D_{ji}^{-1} C_{ji}' M_{ji}^{-2}\varepsilon_{(i)} & \varepsilon_{(i)}' D_{ji}^{-1} C_{ji}' M_{ji}^{-2} C_{ji} D_{ji}^{-1}\varepsilon_{(i)} \end{pmatrix}.$$

Consider the second absolute moment of any of the entries of Θ_{ji}. Take, for example the upper left element, $\varepsilon_{2i-1}' M_{ji}^{-2}\varepsilon_{2i-1}$, we have by (158)

$$\begin{aligned} \sup_{x_p\in[a,b]} \mathbb{E}\left|\varepsilon_{2i-1}' M_{ji}^{-2}\varepsilon_{2i-1}\right|^2 &\le CT^{-2}\sup_{x_p\in[a,b]}\mathbb{E}\left(\operatorname{tr}\left(M_{ji}^{-2}M_{ji}^{*-2}\right) + \left|\operatorname{tr} M_{ji}^{-2}\right|^2\right) \\ \le\ & CT^{-2}\mathbb{E}\left(p\mu_{\min,ji}^{-4}\left(\underline{\epsilon}^{-4} + F_{p,ji}\left([a',b']\right)y_p^{-4}\right) + p^2\mu_{\min,ji}^{-4}\left(\underline{\epsilon}^{-2} + F_{p,ji}\left([a',b']\right)y_p^{-2}\right)^2\right) \end{aligned}$$

Using this, an analogue of equation (OW35), Hölder's inequality and Lemma 5, we get

$$\sup_{x_p\in[a,b]} \mathbb{E}\left|\varepsilon_{2i-1}' M_{ji}^{-2}\varepsilon_{2i-1}\right|^2 \le C + Cy_p^{-4}\mathbb{E}\left(\mu_{\min,ji}^{-4}F_{p,ji}^2\left([a',b']\right)\right).$$

Further, by the Cauchy-Schwarz inequality,

$$\mathbb{E}\left(\mu_{\min,ji}^{-4}F_{p,ji}^2\left([a',b']\right)\right) \le \left(\mathbb{E}\mu_{\min,ji}^{-8}\right)^{1/2}\left(\mathbb{E}F_{p,ji}^4\left([a',b']\right)\right)^{1/2} \le C\left(\mathbb{E}F_{p,ji}^2\left([a',b']\right)\right)^{1/2} \le Cy_p^5.$$

Therefore,

$$\sup_{x_p\in[a,b]} \mathbb{E}\left|\varepsilon_{2i-1}' M_{ji}^{-2}\varepsilon_{2i-1}\right|^2 \le C.$$

Similar inequalities holds for the other entries of Θ_{ji}. This and the boundedness of $\left\|\tilde{\Omega}_{ji}\right\|$ imply that

$$\mathbb{E}\left|\operatorname{tr}\left(\Theta_{ji}\tilde{\Omega}_{ji}\right)\right|^2 \leq C.$$

The absolute second moment of the second term on the right hand side of (164) converges to zero as $p \to \infty$. Indeed, note that by Lemmas 24 and 26, and by the boundedness of $\left\|\tilde{\Omega}_{ji}\right\|$, for any $\rho > 2$,

$$\max_{i,j} \sup_{x_p \in [a,b]} \mathbb{E}\left\|\tilde{\Omega}_{ji}(\tilde{\Omega}_{ji}^{-1} - (\Omega_{ji}^{(q)})^{-1})\Omega_{ji}^{(q)}\right\|^\rho \leq Cp^{-\rho/2}y_p^{-\rho}. \tag{165}$$

On the other hand, for any $\rho > 2$, the absolute moment of order ρ of any of the entries of Θ_{ji} is bounded by $Cy_p^{-2\rho}$ Indeed, take for example the upper left element of Θ_{ji}. We have by (158)

$$\begin{aligned}
\sup_{x_p \in [a,b]} \mathbb{E}\left|\varepsilon_{2i-1}' M_{ji}^{-2}\varepsilon_{2i-1}\right|^\rho &\leq CT^{-\rho} \sup_{x_p \in [a,b]} \mathbb{E}\left(\left(\operatorname{tr}\left(M_{ji}^{-2}M_{ji}^{*-2}\right)\right)^{\rho/2} + \left|\operatorname{tr} M_{ji}^{-2}\right|^\rho\right) \\
&\leq CT^{-\rho} \sup_{x_p \in [a,b]} \mathbb{E}\left(\mu_{\min,ji}^{-2\rho} y_p^{-2\rho}\left(p^{\rho/2} + p^\rho\right)\right) \leq Cy_p^{-2\rho}.
\end{aligned}$$

Similar inequalities hold for the other entries of Θ_{ji}. Therefore,

$$\sup_{x_p \in [a,b]} \mathbb{E}\left\|\Theta_{ji}\right\|^\rho \leq Cy_p^{-2\rho}. \tag{166}$$

Now, by Hölder's inequality

$$\begin{aligned}
&\mathbb{E}\left|\operatorname{tr}\left(\Theta_{ji}\tilde{\Omega}_{ji}(\tilde{\Omega}_{ji}^{-1} - (\Omega_{ji}^{(q)})^{-1})\Omega_{ji}^{(q)}\right)\right|^2 \leq C\mathbb{E}\left(\left\|\Theta_{ji}\right\|^2\left\|\tilde{\Omega}_{ji}\left(\tilde{\Omega}_{ji}^{-1} - (\Omega_{ji}^{(q)})^{-1}\right)\Omega_{ji}^{(q)}\right\|^2\right) \\
\leq\ & C\left(\mathbb{E}\left\|\Theta_{ji}\right\|^3\right)^{2/3}\left(\mathbb{E}\left\|\tilde{\Omega}_{ji}\left(\tilde{\Omega}_{ji}^{-1} - (\Omega_{ji}^{(q)})^{-1}\right)\Omega_{ji}^{(q)}\right\|^6\right)^{1/3}.
\end{aligned}$$

Using (165) and (166), we obtain

$$\sup_{x_p \in [a,b]} \mathbb{E}\left|\operatorname{tr}\left(\Theta_{ji}\tilde{\Omega}_{ji}(\tilde{\Omega}_{ji}^{-1} - (\Omega_{ji}^{(q)})^{-1})\Omega_{ji}^{(q)}\right)\right|^2 \leq Cp^{-1}y_p^{-6} \to 0.$$

To summarize, the absolute second moment of the right hand side of (164) is bounded, and hence,

$$\mathbb{E}\left|v_j - \mathbb{E}v_j\right|^2 \leq Cp^{-1}.$$

Similar inequalities hold for the other entries of $\tilde{\Omega}_j^{-1} - \hat{\Omega}_j^{-1}$, so that

$$\mathbb{E}\left\|\tilde{\Omega}_j^{-1} - \hat{\Omega}_j^{-1}\right\|^2 \leq Cp^{-1}. \tag{167}$$

Now, let us consider $\hat{\Omega}_j^{-1} - (\Omega_j^{(q)})^{-1}$. Take, for example its upper left entry $v_j - v_{j,11}^{(q)}$. By (143), for any $x_p \in [a,b]$,

$$\begin{aligned}
\mathbb{E}\left|v_j - v_{j,11}^{(q)}\right|^2 &= \mathbb{E}\mathbb{E}_{-j}\left|v_j - v_{j,11}^{(q)}\right|^2 \leq \mathbb{E}Cp^{-2}\operatorname{tr} M_j^{-1}M_j^{*-1} \\
&\leq Cp^{-2}\mathbb{E}\mu_{\min,j}^{-2}\left(p\underline{\epsilon}^{-2} + pF_{pj}\left([a',b']\right)y_p^{-2}\right) \leq Cp^{-1}.
\end{aligned}$$

Similar inequalities hold for the other entries of $\hat{\Omega}_j^{-1} - (\Omega_j^{(q)})^{-1}$. Hence,

$$\max_j \sup_{x_p \in [a,b]} \mathbb{E}\left\|\hat{\Omega}_j^{-1} - (\Omega_j^{(q)})^{-1}\right\|^2 \leq Cp^{-1}. \tag{168}$$

Combining (167) and (168) concludes our proof.□

The rest of the proof of Lemma OW15. Recall equation (OW45)

$$\Omega_j^{(d)} - \mathbb{E}\Omega_j^{(q)} = R_1 + R_2 + R_3 \tag{169}$$

where

$$\begin{aligned} R_1 &= \Omega_j^{(d)}\mathbb{E}((\Omega_j^{(q)})^{-1} - (\Omega_j^{(d)})^{-1})\Omega_j^{(d)}, \\ R_2 &= -\mathbb{E}\left(\left(\Omega_j^{(d)}((\Omega_j^{(q)})^{-1} - (\Omega_j^{(d)})^{-1})\right)^2 \Omega_j^{(d)}\right), \text{ and} \\ R_3 &= \mathbb{E}\left(\left(\Omega_j^{(d)}((\Omega_j^{(q)})^{-1} - (\Omega_j^{(d)})^{-1})\right)^3 \Omega_j^{(q)}\right). \end{aligned}$$

The decomposition (169) yields corresponding decompositions for $\bar{e}_k$. Specifically, we have

$$\begin{aligned} \bar{e}_1 &= \sum_{s=1}^{3} \bar{e}_{1s}(z) \equiv \sum_{s=1}^{3} \frac{1}{p} \sum_{j=1}^{T/2} \frac{1}{(1-z)^2} \operatorname{tr}\left(\left[I_2, r_j\nabla_j'\right] R_s \left[I_2, r_j\nabla_j'\right]'\right) \\ \bar{e}_2 &= \sum_{s=1}^{3} \bar{e}_{2s}(z) \equiv \sum_{s=1}^{3} \frac{1}{p} \sum_{j=1}^{T/2} \frac{1}{(1-z)^2} \operatorname{tr}\left(\left[I_2, r_j z\nabla_j'\right] R_s \left[I_2, z r_j\nabla_j'\right]'\right), \\ \bar{e}_3 &= \sum_{s=1}^{3} \bar{e}_{3s}(z) \equiv \sum_{s=1}^{3} \frac{1}{p} \sum_{j=1}^{T/2} \frac{1}{(1-z)^2} \operatorname{tr}\left(\left[I_2, r_j z\nabla_j'\right] R_s \left[I_2, r_j\nabla_j'\right]'\right), \\ \bar{e}_4 &= \sum_{s=1}^{3} \bar{e}_{4s}(z) \equiv -\sum_{s=1}^{3} \frac{1}{p} \sum_{j=1}^{T/2} \frac{1}{1-z} \operatorname{tr}\left(\left[0, I_2\right] R_s \left[I_2, r_j\nabla_j'\right]'\right). \end{aligned}$$

Lemmas OW13, 23 and 26 applied together with Hölder's inequality yield, for $k = 1, ..., 4$,

$$\sup_{x_p\in[a,b]} \left|\bar{e}_{k3}\left(z_p\right)\right| \le Cp^{-3/2}y_p^{-5} \le Cp^{-1}.$$

Clearly, we also have $\sup_{x_p\in[a,b]} \left|\bar{e}_{k3}\left(z_p\right)/z_p\right| \le Cp^{-3/2}y_p^{-6} \le Cp^{-1}$.

To establish similar bounds for $\bar{e}_{k1}\left(z_p\right)$, note that $\mathbb{E}(\Omega_j^{(q)})^{-1} = \tilde{\Omega}_j^{-1}$, and hence, $\mathbb{E}((\Omega_j^{(q)})^{-1} - (\Omega_j^{(d)})^{-1}) = \tilde{\Omega}_j^{-1} - (\Omega_j^{(d)})^{-1}$. Therefore, by Lemma 25,

$$\begin{aligned} \sup_{x_p\in[a,b]} \left|\bar{e}_{11}\left(z_p\right)\right| &\le Cp^{-1}s_1^2, \ \sup_{x_p\in[a,b]} \left|\bar{e}_{21}\left(z_p\right)\right| \le Cp^{-1}s_2^2, \\ \sup_{x_p\in[a,b]} \left|\bar{e}_{31}\left(z_p\right)\right| &\le Cp^{-1}s_1 s_2, \text{ and } \sup_{x_p\in[a,b]} \left|\bar{e}_{41}\left(z_p\right)\right| \le Cp^{-1}s_1. \end{aligned}$$

where

$$s_1 = \sup_{x_p\in[a,b]} \left\|\frac{1}{1-z}\left[I_2, r_j\nabla_j'\right]\Omega_j^{(d)}\right\| \text{ and } s_2 = \sup_{x_p\in[a,b]} \left\|\frac{1}{1-z}\left[I_2, r_j z\nabla_j'\right]\Omega_j^{(d)}\right\|.$$

Using the fact that

$$\Omega_j^{(d)} = \frac{1-z}{\delta_j^{(d)}} \begin{pmatrix} \frac{z}{1-z} r_j I_2 + z\mathbb{E}\tilde{v}I_2 & -\frac{1}{1-z} r_j\nabla_j' - \mathbb{E}uI_2 \\ -\frac{1}{1-z} r_j\nabla_j - \mathbb{E}uI_2 & \frac{1}{1-z} I_2 + \mathbb{E}vI_2 \end{pmatrix}$$

and the identity $\nabla_j'\nabla_j = r_j I_2$, we obtain

$$\begin{aligned} \frac{1}{1-z}\left[I_2, r_j\nabla_j'\right]\Omega_j^{(d)} &= \frac{1}{\delta_j^{(d)}}\left[I_2, r_j\nabla_j'\right]\begin{pmatrix} z\mathbb{E}\tilde{v}I_2 & -\mathbb{E}uI_2 \\ -\mathbb{E}uI_2 & \mathbb{E}vI_2 \end{pmatrix} - \frac{1}{\delta_j^{(d)}}\left[r_j I_2, 0\right], \text{ and} \\ \frac{1}{1-z}\left[I_2, r_j z\nabla_j'\right]\Omega_j^{(d)} &= \frac{1}{\delta_j^{(d)}}\left[I_2, r_j z\nabla_j'\right]\begin{pmatrix} z\mathbb{E}\tilde{v}I_2 & -\mathbb{E}uI_2 \\ -\mathbb{E}uI_2 & \mathbb{E}vI_2 \end{pmatrix} - \frac{1}{\delta_j^{(d)}}\left[0, r_j\nabla_j'\right]. \end{aligned}$$

On the other hand, as follows from the proof of Lemma OW13, $\min_j \inf_{x_p\in[a,b]} \delta_j^{(d)}$ is bounded away from zero for all sufficiently large p. Further, by inspection $\max_j \left\|r_j \nabla_j'\right\|$ is bounded. Finally, the boundedness of $\sup_{x_p\in[a,b]} |\mathbb{E}u|$, $\sup_{x_p\in[a,b]} |\mathbb{E}v|$, and $\sup_{x_p\in[a,b]} |\mathbb{E}\tilde{v}|$ follows, for example, from equations (130-132). Therefore, s_1 and s_2 are bounded, and

$$\sup_{x_p\in[a,b]} |\bar{e}_{k1}(z_p)| \leq Cp^{-1}$$

for $k = 1, ..., 4$.

We can slightly improve the latter inequality for $\bar{e}_{21}(z_p)$. Indeed, note that $\frac{1}{1-z}\left[I_2, r_j z\nabla_j'\right]\Omega_j^{(d)}$ can be represented in the form

$$\frac{1}{1-z}\left[I_2, r_j z\nabla_j'\right]\Omega_j^{(d)} = K_j \begin{pmatrix} zI_2 & 0 \\ 0 & I_2 \end{pmatrix},$$

where $K_j = \frac{1}{\delta_j^{(d)}}\left[\mathbb{E}\tilde{v}I_2 - r_j\mathbb{E}u\nabla_j', r_j\left(z\mathbb{E}v - 1\right)\nabla_j' - \mathbb{E}uI_2\right]$, so that $\max_j \sup_{x_p\in[a,b]} \|K_j\|$ is bounded. Therefore

$$\begin{aligned}
\bar{e}_{21}(z) &= \frac{1}{p}\sum_{j=1}^{T/2} \operatorname{tr}\left(K_j \begin{pmatrix} zI_2 & 0 \\ 0 & I_2 \end{pmatrix} \mathbb{E}((\Omega_j^{(q)})^{-1} - (\Omega_j^{(d)})^{-1}) \begin{pmatrix} zI_2 & 0 \\ 0 & I_2 \end{pmatrix} K_j'\right) \\
&= \frac{1}{p}\sum_{j=1}^{T/2} \operatorname{tr}\left(K_j \begin{pmatrix} zI_2 & 0 \\ 0 & I_2 \end{pmatrix} \begin{pmatrix} (\mathbb{E}v_j - \mathbb{E}v)I_2 & (\mathbb{E}u_j - \mathbb{E}u)I_2 \\ (\mathbb{E}u_j - \mathbb{E}u)I_2 & z(\mathbb{E}\tilde{v}_j - \mathbb{E}\tilde{v})I_2 \end{pmatrix} \begin{pmatrix} zI_2 & 0 \\ 0 & I_2 \end{pmatrix} K_j'\right) \\
&= z \times \frac{1}{p}\sum_{j=1}^{T/2} \operatorname{tr}\left(K_j \begin{pmatrix} z(\mathbb{E}v_j - \mathbb{E}v)I_2 & (\mathbb{E}u_j - \mathbb{E}u)I_2 \\ (\mathbb{E}u_j - \mathbb{E}u)I_2 & (\mathbb{E}\tilde{v}_j - \mathbb{E}\tilde{v})I_2 \end{pmatrix} K_j'\right).
\end{aligned}$$

This implies that

$$\sup_{x_p\in[a,b]} |\bar{e}_{21}(z_p)/z_p| \leq Cp^{-1}.$$

For $\bar{e}_{k2}(z_p)$, we use Lemma 27 and the boundedness of s_1, s_2 and $\left\|\Omega_j^{(d)}\right\|$, to obtain inequalities

$$\sup_{x_p\in[a,b]} |\bar{e}_{k2}(z_p)| \leq Cp^{-1}.$$

Finally, for $\bar{e}_{22}(z_p)$, we have

$$\bar{e}_{22}(z) = -\frac{1}{p}\sum_{j=1}^{T/2} \operatorname{tr}\left(\mathbb{E}\left(K_j \begin{pmatrix} zI_2 & 0 \\ 0 & I_2 \end{pmatrix} ((\Omega_j^{(q)})^{-1} - (\Omega_j^{(d)})^{-1})\Omega_j^{(d)}((\Omega_j^{(q)})^{-1} - (\Omega_j^{(d)})^{-1}) \begin{pmatrix} zI_2 & 0 \\ 0 & I_2 \end{pmatrix} K_j'\right)\right).$$

On the other hand,

$$\begin{pmatrix} zI_2 & 0 \\ 0 & I_2 \end{pmatrix} ((\Omega_j^{(q)})^{-1} - (\Omega_j^{(d)})^{-1}) = \begin{pmatrix} z\left(v_j^{(q)} - \mathbb{E}vI_2\right) & u_j^{(q)\prime} - \mathbb{E}uI_2 \\ u_j^{(q)} - \mathbb{E}uI_2 & \tilde{v}_j^{(q)} - \mathbb{E}\tilde{v}I_2 \end{pmatrix} \begin{pmatrix} I_2 & 0 \\ 0 & zI_2 \end{pmatrix}$$

and

$$\begin{pmatrix} I_2 & 0 \\ 0 & zI_2 \end{pmatrix} \Omega_j^{(d)} = \bar{\Omega}_j^{(d)} \begin{pmatrix} zI_2 & 0 \\ 0 & I_2 \end{pmatrix}$$

with

$$\bar{\Omega}_j^{(d)} = \frac{1}{\delta_j^{(d)}} \begin{pmatrix} r_jI_2 + (1-z)\mathbb{E}\tilde{v}I_2 & -r_j\nabla_j' - (1-z)\mathbb{E}uI_2 \\ -r_j\nabla_j - (1-z)\mathbb{E}uI_2 & zI_2 + z(1-z)\mathbb{E}vI_2 \end{pmatrix}.$$

Hence,

$$\bar{e}_{22}(z) = -z \times \frac{1}{p}\sum_{j=1}^{T/2} \operatorname{tr}\left(\mathbb{E}\left(K_j \begin{pmatrix} z\left(v_j^{(q)} - \mathbb{E}vI_2\right) & u_j^{(q)\prime} - \mathbb{E}uI_2 \\ u_j^{(q)} - \mathbb{E}uI_2 & \tilde{v}_j^{(q)} - \mathbb{E}\tilde{v}I_2 \end{pmatrix} \bar{\Omega}_j^{(d)} \begin{pmatrix} z\left(v_j^{(q)} - \mathbb{E}vI_2\right) & u_j^{(q)\prime} - \mathbb{E}uI_2 \\ u_j^{(q)} - \mathbb{E}uI_2 & \tilde{v}_j^{(q)} - \mathbb{E}\tilde{v}I_2 \end{pmatrix} K_j'\right)\right).$$

Note that $\max_j \sup_{x_p\in[a,b]} \left\|\bar{\Omega}_j^{(d)}\right\|$ is bounded and

$$\max_j \sup_{x_p\in[a,b]} \mathbb{E}\left\|\begin{pmatrix} z\left(v_j^{(q)} - \mathbb{E}vI_2\right) & u_j^{(q)\prime} - \mathbb{E}uI_2 \\ u_j^{(q)} - \mathbb{E}uI_2 & \tilde{v}_j^{(q)} - \mathbb{E}\tilde{v}I_2 \end{pmatrix}\right\|^2 \leq Cp^{-1},$$

which can be established similarly to Lemma 27. So, finally,

$$\sup_{x_p\in[a,b]} |\bar{e}_{22}(z_p)/z_p| \leq Cp^{-1}.$$

4.4.2 Proof of Lemma OW16 (bounds on $\sup_{x_p\in[a,b]} \hat{e}_k(z_p)$)

Note that the system of equations (OW41-44) can be obtained from the system of equations (OW18-21) by replacing $m, v, u, \tilde{v}$ by their expected values and replacing $e_1, ..., e_4$ by $\bar{e}_1, ..., \bar{e}_4$. Therefore, the reduction of (OW41-44) to the simple system (OW47) parallels the reduction of (OW18-21) to (OW31).

In particular, proceeding as in Section 4.2.5, we obtain an equation analogous to (68)

$$\hat{e}_1 = (-\bar{e}_1 + \bar{e}_2/z + 2\bar{e}_4)\,\bar{\theta}^{-1},$$

where $\bar{\theta} = (2/cT)\sum_{j=1}^{T/2}\left(\delta_j^{(d)}\right)^{-1}$. Since, as follows from the proof of Lemma OW13, $\max_j \sup_{x_p\in[a,b]} \delta_j^{(d)}$ is bounded, we have

$$\sup_{x_p\in[a,b]} \left|\bar{\theta}^{-1}\right| \leq C. \tag{170}$$

Therefore, Lemma OW15 and equation (OW46) yield

$$\sup_{x_p\in[a,b]} |\hat{e}_1(z_p)| \leq Cp^{-1}.$$

Further, arguments that parallel those of Section 4.2.5 lead to equation

$$(2z\mathbb{E}v + \mathbb{E}u)\left((1 + \mathbb{E}v - z\mathbb{E}v)(1-c) - c\right) + 2c = \bar{\xi}_2,$$

where

$$\bar{\xi}_2 = -\frac{\mathbb{E}v}{\mathbb{E}u}\bar{\theta}^{-1}\left[(1-c)(-z\bar{e}_1 + \bar{e}_2 + 2z\bar{e}_4)(1 + \mathbb{E}v - z\mathbb{E}v) + (\mathbb{E}u\bar{e}_4/\mathbb{E}v + \bar{e}_3 - \bar{e}_2) - 2\bar{e}_4c(\mathbb{E}u + z\mathbb{E}v - 1)/\mathbb{E}v\right].$$

On the other hand, as follows from equations (130-132), there exists $C > 0$ s.t.

$$\begin{aligned} \inf_{x_p\in[a,b]} |\mathbb{E}u| &> C, \quad \inf_{x_p\in[a,b]} |\mathbb{E}v| > C, \text{ whereas} \\ \sup_{x_p\in[a,b]} |\mathbb{E}v| &< C, \quad \sup_{x_p\in[a,b]} |\mathbb{E}u| < C, \text{ and } \sup_{x_p\in[a,b]} |\mathbb{E}\tilde{v}| < C. \end{aligned} \tag{171}$$

Therefore,

$$\sup_{x_p\in[a,b]} \left|\bar{\xi}_2(z_p)\right| \leq Cp^{-1}.$$

Next, similarly to equation (80), we have

$$\mathbb{E}m = \left(c^{-1} - 1\right)\mathbb{E}v + \hat{e}_3,$$

where

$$\hat{e}_3 = \bar{\xi}_3 + 2\bar{\bar{\xi}}_3 + \frac{2c}{(1 + v - zv)(1-c) - c}\bar{\bar{\xi}}_3$$

with

$$\bar{\xi}_3 = \bar{e}_1 - 2\bar{e}_4 + \frac{cz\mathbb{E}v\bar{e}_1 - \bar{e}_3\mathbb{E}v + \bar{e}_2\mathbb{E}v(1-c) - \mathbb{E}u\bar{e}_4 + 2\bar{e}_4c(\mathbb{E}u - 1)}{(1 + \mathbb{E}v - z\mathbb{E}v)(1-c) - c}$$

and

$$\bar{\bar{\xi}}_3 = \frac{((1+\mathbb{E}v - z\mathbb{E}v)(1-c)-c)}{2c}\left(\mathbb{E}v\,(z\bar{e}_1 - \bar{e}_2 - 2z\bar{e}_4) - \frac{(\mathbb{E}u\bar{e}_4 + \mathbb{E}v\bar{e}_3 - \mathbb{E}v\bar{e}_2)}{(1+\mathbb{E}v - z\mathbb{E}v)}\right) + \frac{\mathbb{E}u\bar{\theta}\bar{\bar{\xi}}_2}{2c}$$

Since $\mathbb{E}v$ converges to v_0, $|(1+\mathbb{E}v - z\mathbb{E}v)(1-c)-c|$ and $|1+\mathbb{E}v - z\mathbb{E}v|$ are bounded away from zero, and we have

$$\sup_{x_p \in [a,b]} |\hat{e}_3(z_p)| \leq Cp^{-1}.$$

So continuing, now in parallel to Sections 4.2.5 and 4.2.5, we obtain

$$\sup_{x_p \in [a,b]} |\hat{e}_4(z_p)| \leq Cp^{-1} \text{ and } \sup_{x_p \in [a,b]} |\hat{e}_2(z_p)| \leq Cp^{-1}.$$

The details of such a derivation are tedious but straightforward and we omit them.

References

[1] Abramowitz, M. and I.A. Stegun, eds., (1970) *Handbook of Mathematical Functions,* Dover, New York.

[2] Bai, Z. D. and J. W. Silverstein, (1998) "No eigenvalues outside the support of the limiting spectral distribution of large-dimensional sample covariance matrices," *The Annals of Probability* 26, 316-345.

[3] Chen, Z. and J. J. Dongarra (2005) "Condition number of Gaussian random matrices," *SIAM J. Matrix. Anal. Appl.* 27, 603-620.

[4] Davidson, K.R. and S.J. Szarek (2001) Local operator theory, random matrices and Banach spaces, in: W.B. Johnson and J. Lindenstrauss (Eds.), Handbook of the Geometry of Banach Spaces, Vol. I, North-Holland, Amsterdam, 2001, 317–366.

[5] Golub, G.H. and C. F. Van Loan (1996) *Matrix Computations*, The John Hopkins University Press.

[6] Horn, R.A. and C.R. Johnson (1985) *Matrix Analysis*, Cambridge University press.

[7] Johansen, S. (1995) *Likelihood-based Inference in Cointegrated Vector Autoregressive Models*, Oxford University Press.

[8] Onatski, A. and Wang, C. (2017a) "Extreme canonical correlations and high-dimensional cointegration analysis," manuscript, University of Cambridge.

[9] Onatski, A. and Wang, C. (2017) "Alternative asymptotics for cointegration tests in large VARs," manuscript, University of Cambridge.

[10] Rudelson, M. and R. Vershynin (2010) "Non-asymptotic theory of random matrices: extreme singular values," *Proceedings of the International Congress of Mathematicians*, Hyderabad, India.