Supplement to
Model Identification via Total Frobenius Norm of Multivariate Spectra

Tucker S. McElroy[*]and Anindya Roy[†]

Disclaimer: This report is released to inform interested parties of research and to encourage discussion. The views expressed on statistical issues are those of the authors and not those of the U.S. Census Bureau.

Appendix A Asymptotic Theory of Linear and Quadratic Functionals of the Periodogram

In the main paper it is asserted that $\widehat{Q}_{\varphi_1,\varphi_2}(I,I)$ and $\widehat{L}_{\varphi_3}(I)$ are consistent estimators of $Q_{\varphi_1,\varphi_2}(f,f)$ and $L_{\varphi_3}(f)$. (Actually, if we integrate over all frequencies, the estimators $Q_{\varphi_1,\varphi_2}(I,I)$ and $L_{\varphi_3}(I)$ are also consistent, but the variability in the former estimator will differ from what is given in Proposition A.2 below.) This Appendix provides asymptotic theory for $\widehat{L}_{\varphi_3}(I)$ and $\widehat{Q}_{\varphi_1,\varphi_2}(I,I)$, based on assumptions involving summability conditions on the cumulants of the $\{x_t\}$ process. Under Assumption 1 the asymptotic behavior of the first moments of the linear and quadratic functionals can be established.

Proposition A.1. *(Convergence of First Moments) Let φ_1, φ_2, φ_3 be continuous real matrix-valued functions of frequency λ, and let $\widehat{L}_{\varphi_3}$ and $\widehat{Q}_{\varphi_1,\varphi_2}$ be defined via (10). Assume that $\{x_t\}$ is strictly stationary with spectral density $\widetilde{f}$, and satisfies Assumption 1(k) for $k = 2,3,4$. Then as $T \to \infty$,*

$$\mathbb{E}\widehat{L}_{\varphi_3}(I) = L_{\varphi_3}(\widetilde{f}) + O(T^{-1})$$
$$\mathbb{E}\widehat{Q}_{\varphi_1,\varphi_2}(I,I) = Q_{\varphi_1,\varphi_2}(\widetilde{f},\widetilde{f}) + \langle [[\varphi_1 \widetilde{f}]]\, [[\varphi_2 \widetilde{f}]] \rangle_0 + O(T^{-1}),$$

where L and Q are defined via (9).

Remark A.1. Lemma 3.1.1 of Taniguchi and Kakizawa (2000) shows that

$$\widehat{L}_{\varphi_3}(I) \xrightarrow{P} L_{\varphi_3}(\widetilde{f}) \tag{A.1}$$

in the special case that $\varphi_3(-\lambda)' = \varphi_3(\lambda)$. Proposition A.1 relaxes this condition, establishing convergence of the first moments for linear functionals, and Proposition A.2 below shows convergence of second moments (for linear functionals) as well. Together, these propositions establish (A.1), thereby generalizing part (i)

[*]Research and Methodology Directorate, U.S. Census Bureau, 4600 Silver Hill Road, Washington, D.C. 20233-9100, tucker.s.mcelroy@census.gov

[†]University of Maryland, Baltimore County and U.S. Census Bureau

of Lemma 3.1.1 of Taniguchi and Kakizawa (2000). In addition, Propositions A.1 and A.2 establish first and second moment convergences for quadratic functionals, which are not considered in Lemma 3.1.1 of Taniguchi and Kakizawa (2000).

Next, we introduce the notation $\overline{L}_{\varphi_3}(I)$ and $\overline{Q}_{\varphi_1,\varphi_2}(I,I)$ for the estimated functionals centered by their respective asymptotic means, i.e.,

$$\begin{aligned}\overline{L}_{\varphi_3}(I) &= \widehat{L}_{\varphi_3}(I) - L_{\varphi_3}(\widetilde{f})\\ \overline{Q}_{\varphi_1,\varphi_2}(I,I) &= \widehat{Q}_{\varphi_1,\varphi_2}(I,I) - Q_{\varphi_1,\varphi_2}(\widetilde{f},\widetilde{f}) - \langle [[\varphi_1\,\widetilde{f}]]\,[[\varphi_2\,\widetilde{f}]]\rangle_0.\end{aligned}$$

In the next result we utilize the notation $[h]$ for $[[\varphi_h\,\widetilde{f}]]$, $[hi]$ for $[[\varphi_h\,\widetilde{f}\,\varphi_i\,\widetilde{f}]]$, $[hij]$ for $[[\varphi_h\,\widetilde{f}\,\varphi_i\,\widetilde{f}\,\varphi_j\,\widetilde{f}]]$, and $[hijk]$ for $[[\varphi_h\,\widetilde{f}\,\varphi_i\,\widetilde{f}\,\varphi_j\,\widetilde{f}\,\varphi_k\,\widetilde{f}]]$, where $h,i,j,k \in \{1,2,3,4\}$. It can happen that a function φ appears in such a term with its argument reflected and the matrix transposed, i.e, $\varphi(-\lambda)'$, in which case we place an underscore under the index, e.g., $[1\underline{2}3]$ denotes $[[\varphi_1(\lambda)\,\widetilde{f}(\lambda)\,\varphi_2(-\lambda)'\,\widetilde{f}(\lambda)\,\varphi_3(\lambda)\,\widetilde{f}(\lambda)]]$. The tri-spectral density is denoted by $\widetilde{f}$ with four sub-indices and three frequency arguments, and is defined via

$$\widetilde{f}_{\ell ksr}(\lambda,\theta,\omega) = \sum_{h_1,h_2,h_3\in\mathbb{Z}} \gamma_{\ell ksr}(h_1,h_2,h_3)\, e^{-i(\lambda h_1+\theta h_2+\omega h_3)}.$$

We also use the notation

$$[[A(\lambda)\,\widetilde{f}(\lambda,-\lambda,\omega)\,B(\omega)]] = \sum_{\ell,k,r,s} A_{k\ell}(\lambda)\,\widetilde{f}_{\ell ksr}(\lambda,-\lambda,\omega)\,B_{rs}(\omega),$$

which is a sort of double-trace of the 4-array. If integration is applied over both λ and ω, and the whole divided by $4\pi^2$, we use a $\langle\langle\cdot\rangle_0\rangle_0$ notation. Finally, a capital letter $\widetilde{F}$ for an m-dimensional 4-array $\widetilde{f}_{\ell ksr}$ will denote a m^2-dimensional matrix defined as follows: let C be the $m^2\times 2$ matrix of row and column indices corresponding to the vec of an $m\times m$ matrix. Then for $1\leq p,n\leq m^2$, let $(\ell,k)=C[p,]$ and $(s,r)=C[n,]$ and $\widetilde{F}_{pn}=\widetilde{f}_{\ell ksr}$.

Proposition A.2. *(Convergence of Second Moments) Let φ_1, φ_2, φ_3, φ_4 be continuous real matrix-valued functions of frequency λ, and let $\widehat{L}_{\varphi_3}$ and $\widehat{Q}_{\varphi_1,\varphi_2}$ be defined via (10). Assume that $\{x_t\}$ is strictly stationary with spectral density $\widetilde{f}$ and satisfies Assumption 1(k) for $2\leq k\leq 8$. Then as $T\to\infty$,*

$$\begin{aligned}Cov\left(\sqrt{T}\,\widehat{L}_{\varphi_1}(I),\sqrt{T}\,\widehat{L}_{\varphi_2}(I)\right) &\to \langle[12]\rangle_0 + \langle[1\underline{2}]\rangle_0\\ &\quad + \langle\langle[[\varphi_1(\lambda)\,\widetilde{f}(\lambda,-\lambda,\omega)\,\varphi_2(\omega)]]\rangle_0\rangle_0\end{aligned}$$

and

$$\begin{aligned}Cov\left(\sqrt{T}\,\widehat{Q}_{\varphi_1,\varphi_2}(I,I),\sqrt{T}\,\widehat{L}_{\varphi_3}(I)\right) &\to \langle[1]\,[23]\rangle_0 + \langle[132]\rangle_0 + \langle[123]\rangle_0 + \langle[2]\,[13]\rangle_0\\ &\quad + \langle[1]\,[2\underline{3}]\rangle_0 + \langle[1\underline{3}2]\rangle_0 + \langle[12\underline{3}]\rangle_0 + \langle[2]\,[1\underline{3}]\rangle_0\\ &\quad + \langle\langle g_{123}(\lambda,\omega)\rangle_0\rangle_0\end{aligned}$$

and

$$
\begin{aligned}
Cov\left(\sqrt{T}\,\widehat{Q}_{\varphi_1,\varphi_2}(I,I), \sqrt{T}\,\widehat{Q}_{\varphi_3,\varphi_4}(I,I)\right) &\to \langle\langle g_{1234}(\lambda,\omega)\rangle_0\rangle_0 \\
&+ \langle[1][4][32]\rangle_0 + \langle[1][243]\rangle_0 + \langle[1][234]\rangle_0 + \langle[1][3][24]\rangle_0 \\
&+ \langle[4][132]\rangle_0 + \langle[1342]\rangle_0 + \langle[3][142]\rangle_0 + \langle[1432]\rangle_0 \\
&+ \langle[2][4][31]\rangle_0 + \langle[31][42]\rangle_0 + \langle[1423]\rangle_0 + \langle[2][143]\rangle_0 \\
&+ \langle[4][123]\rangle_0 + \langle[1243]\rangle_0 + \langle[2][134]\rangle_0 + \langle[1324]\rangle_0 \\
&+ \langle[14][23]\rangle_0 + \langle[3][2][14]\rangle_0 + \langle[1234]\rangle_0 + \langle[3][124]\rangle_0 \\
&+ \langle[1][2\underline{43}]\rangle_0 + \langle[\underline{4}][1][2\underline{3}]\rangle_0 + \langle[1][\underline{3}][2\underline{4}]\rangle_0 + \langle[1][2\underline{34}]\rangle_0 \\
&+ \langle[1\underline{34}2]\rangle_0 + \langle[1\underline{3}2][\underline{4}]\rangle_0 + \langle[1\underline{4}2][\underline{3}]\rangle_0 + \langle[1\underline{43}2]\rangle_0 \\
&+ \langle[\underline{4}][2][1\underline{3}]\rangle_0 + \langle[1\underline{3}][2\underline{4}]\rangle_0 + \langle[12\underline{43}]\rangle_0 + \langle[12\underline{3}][\underline{4}]\rangle_0 \\
&+ \langle[1\underline{4}2\underline{3}]\rangle_0 + \langle[1\underline{43}][2]\rangle_0 + \langle[12\underline{4}][\underline{3}]\rangle_0 + \langle[12\underline{34}]\rangle_0 \\
&+ \langle[1\underline{34}][2]\rangle_0 + \langle[1\underline{3}2\underline{4}]\rangle_0 + \langle[1\underline{4}][2][\underline{3}]\rangle_0 + \langle[1\underline{4}][2\underline{3}]\rangle_0,
\end{aligned}
$$

where the functions $g_{123}(\lambda,\omega)$ and $g_{1234}(\lambda,\omega)$ depend upon the tri-spectral density, and are given by (C.5) and (C.4) of the proof.

We here introduce a notation for the limiting covariances: write $V_{\varphi_1|\varphi_2}$, $V_{\varphi_1,\varphi_2|\varphi_3}$, and $V_{\varphi_1,\varphi_2|\varphi_3,\varphi_4}$, for the cases of linear-linear, quadratic-linear, and quadratic-quadratic, respectively. These quantities depend on the matrix-valued functions φ_i, as well as the spectral density $\widetilde{f}(\lambda)$ and the tri-spectral density $\widetilde{f}(\lambda,-\lambda,\omega)$. Based on the moment convergence, we can formulate a central limit theorem for vectors of linear and quadratic functionals.

Theorem A.1. *(CLT) Let $\vartheta_1,\ldots,\vartheta_r$, $\varphi_1,\ldots,\varphi_s$, and $\psi_1,\ldots,\psi_s$ be continuous real matrix-valued functions of frequency λ. Let $\widehat{L}$ and $\widehat{Q}$ be defined via (10). Assume that $\{x_t\}$ is strictly stationary and satisfies Assumption 1(k) for $k \geq 2$. Then the vector of r linear and s quadratic functionals are jointly asymptotically normal:*

$$
\sqrt{T}\,\left[\overline{L}_{\vartheta_1}(I),\ldots,\overline{L}_{\vartheta_r}(I),\overline{Q}_{\varphi_1,\psi_1}(I,I),\ldots,\overline{Q}_{\varphi_s,\psi_s}(I,I)\right] \stackrel{\mathcal{L}}{\Longrightarrow} \mathcal{N}\left(0,V\right),
$$

$$
V = \begin{bmatrix}
V_{\vartheta_1|\vartheta_1} & \cdots & V_{\vartheta_1|\vartheta_r} & V_{\vartheta_1|\varphi_1,\psi_1} & \cdots & V_{\vartheta_1|\varphi_s,\psi_s} \\
\vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\
V_{\vartheta_r|\vartheta_1} & \ddots & V_{\vartheta_r|\vartheta_r} & V_{\vartheta_r|\varphi_1,\psi_1} & \ddots & V_{\vartheta_r|\varphi_s,\psi_s} \\
V_{\varphi_1,\psi_1|\vartheta_1} & \ddots & V_{\varphi_1,\psi_1|\vartheta_r} & V_{\varphi_1,\psi_1|\varphi_1,\psi_1} & \ddots & V_{\varphi_1,\psi_1|\varphi_s,\psi_s} \\
\vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\
V_{\varphi_s,\psi_s|\vartheta_1} & \ddots & V_{\varphi_s,\psi_s|\vartheta_r} & V_{\varphi_s,\psi_s|\varphi_1,\psi_1} & \ddots & V_{\varphi_s,\psi_s|\varphi_s,\psi_s}
\end{bmatrix}.
$$

This theory is somewhat more general than needed for our particular applications below, but nonetheless provides a complete framework for understanding how second and fourth cumulant functions impact the covariances of linear and quadratic functionals of the periodogram.

Appendix B Testing the Spectral Residual

Section 5 treats the goodness-of-fit problem when the model is separable, but here we consider a broader framework that allows for non-separable models or even nonparametric methods. The drawback is that there is no accounting for parameter estimation uncertainty. However, in some situations – for example, where the data process has been modeled with a tapered autocovariance estimator – the data may be whitened without an explicit parametric model, and the following techniques may be useful. So instead of analyzing $I f_{\widehat{\theta}}^{-1}$ directly, we consider the periodogram J of the residual process that results from some whitening procedure. This approach may be attractive even in scenarios where there is a parametric model, but it is not separable and involves transformations, fixed regressors, and stochastic effects – such that f_θ is only a partial description of the model.

Let J denote the periodogram of the estimated residuals, with $\widetilde{g}$ denoting the spectral density of the true residual process, and consider the testing problem $|\widetilde{g}|^2 - \|\langle \widetilde{g} \rangle_0\|^2 = 0$; this quantity is always non-negative, but equals zero if and only if $\widetilde{g}$ corresponds to white noise. Rewriting this functional as

$$|\widetilde{g}|^2 - \|\langle \widetilde{g} \rangle_0\|^2 = Q_{\mathrm{id,id}}(\widetilde{g}, \widetilde{g}) - [[\langle \widetilde{g} \rangle_0^2]],$$

it can be seen that substituting J for $\widetilde{g}$ yields a statistic that converges to $|\widetilde{g}|^2 + \langle [[\widetilde{g}]]^2 \rangle_0 - [[\langle \widetilde{g} \rangle_0^2]]$ (this is shown in the proof of Proposition B.1, below). In the case that $\widetilde{g} \equiv \Sigma$, corresponding to a white noise process, the limit reduces to $[[\Sigma]]^2$, which is nonzero except in trivial cases. In order to obtain a statistic that converges to zero (the null hypothesis value of $|\widetilde{g}|^2 - \|\langle \widetilde{g} \rangle_0\|^2$), we propose to subtract the quantity $[[\langle J \rangle_0]]^2$, which converges to $[[\Sigma]]^2$ under the null. These arguments suggest defining a model evaluation functional $\mathrm{Eval}(\widetilde{g})$ and estimator $\widehat{\mathrm{Eval}}(J)$ as follows:

$$\begin{aligned}
\mathrm{Eval}(\widetilde{g}) &= |\widetilde{g}|^2 - \|\langle \widetilde{g} \rangle_0\|^2 + \langle [[\widetilde{g}]]^2 \rangle_0 - [[\langle \widetilde{g} \rangle_0]]^2 = Q_{\mathrm{id,id}}(\widetilde{g}, \widetilde{g}) + \langle [[\widetilde{g}]]^2 \rangle_0 - [[\langle \widetilde{g} \rangle_0^2]] - [[\langle \widetilde{g} \rangle_0]]^2 \\
\widehat{\mathrm{Eval}}(J) &= |J|^2 - \|\langle J \rangle_0\|^2 - [[\langle J \rangle_0]]^2 = \widehat{Q}_{\mathrm{id,id}}(J, J) - [[\langle J \rangle_0^2]] - [[\langle J \rangle_0]]^2.
\end{aligned}$$

The estimator $\widehat{\mathrm{Eval}}(J)$ can be alternatively expressed with the final two terms, $[[\langle J \rangle_0^2]]$ and $[[\langle J \rangle_0]]^2$, having their integral replaced by a sum over Fourier frequencies; this can make calculation easier. However, because these two terms involve integrals/sums over a linear function of the periodogram, the asymptotic theory is the same, whether we use integrals or sums (cf. Chen and Deo (2000)). The evaluation estimator converges in probability to the correct quantity for the testing problem, as shown below.

Proposition B.1. *Assume that $\{x_t\}$ is strictly stationary with spectral density $\widetilde{f}$, and satisfies Assumption 1(k) for $k = 2, 3, 4$. Then as $T \to \infty$, $\widehat{Eval}(J) \xrightarrow{P} Eval(\widetilde{g})$ and*

$$\mathbb{E}[|J|^2] = O(T^{-1}) + |\widetilde{g}|^2 + \langle [[\widetilde{g}]]^2 \rangle_0. \tag{B.1}$$

In order to test the null hypothesis, we need a limit theory for the model evaluation estimator. Unlike the case of the Ljung-Box statistic, based upon a finite number of sample autocovariances, in this case the limit distribution is normal – this is essentially due to the inclusion of the bias-correction term $[[\langle J \rangle_0]]^2$.

Theorem B.1. *Assume that $\{x_t\}$ is strictly stationary with spectral density $\widetilde{g}$ and satisfies Assumption 1(k)* *for $k \geq 2$. Then as $T \to \infty$*

$$\sqrt{T}\left(\widehat{Eval}(J) - Eval(\widetilde{g})\right) \xLongrightarrow{\mathcal{L}} \mathcal{N}(0, v'\,W\,v)$$

$$W = \begin{bmatrix} V_{\langle\widetilde{g}\rangle_0|\langle\widetilde{g}\rangle_0} & V_{\langle\widetilde{g}\rangle_0|id} & V_{\langle\widetilde{g}\rangle_0|id,id} \\ V_{id|\langle\widetilde{g}\rangle_0} & V_{id|id} & V_{id|id,id} \\ V_{id,id|\langle\widetilde{g}\rangle_0} & V_{id,id|id} & V_{id,id|id,id} \end{bmatrix},$$

where $v' = \{-2, -2[[\langle\widetilde{g}\rangle_0]], 1\}$.

In general, the calculation of the entries of W is daunting due to the presence of the tri-spectrum, but under the null hypothesis of white noise there is a remarkable simplification to the limiting variance – all the dependence on the tri-spectrum vanishes.

Corollary B.1. *Assume that $\{x_t\}$ is strictly stationary with white noise spectral density $\widetilde{g} = \Sigma$, and satisfies Assumption 1(k) for $k \geq 2$. Then as $T \to \infty$*

$$\sqrt{T}\,\widehat{Eval}(J) \xLongrightarrow{\mathcal{L}} \mathcal{N}(0, 4\,[[\Sigma^4]] + 4\,[[\Sigma^2]]^2).$$

Corollary B.1 can be applied by substituting the sample variance matrix of the residual process, i.e., $\widehat{\Sigma} = \widehat{\Gamma}(0)$, in the expression for the asymptotic variance.

Appendix C Proofs

Proof of Proposition A.1. Without loss of generality, we replace the sample mean in the DFT by the true mean, because the error in doing so is of lower order. The result for linear functionals is already known in the special case that $\varphi_3(-\lambda)' = \varphi_3(\lambda)$ – see result (i) of Lemma 3.1.1 in Taniguchi and Kakizawa (2000), which can be applied because its condition (B) holds. However, we provide a general proof, beginning with the linear functional. Note that $L_{\varphi_3}(I) = T^{-1} \langle (d^* \varphi_3 d) \rangle_0$, which by Lemma P5.1 of Brillinger (2001) approximates the estimator $\widehat{L}_{\varphi_3}(I)$ up to $O_P(T^{-1})$. Based on such an approximation, one could work with the estimator $L_{\varphi_3}(I)$ instead of $\widehat{L}_{\varphi_3}(I)$, but as mentioned in the text it is simpler to focus upon the latter. To that end, we write

$$\widehat{L}_{\varphi_3}(I) = T^{-2} \sum_{j=1}^{T} (d^* \varphi_3 d)(\lambda_j).$$

Here, λ_j is a Fourier frequency defined as $\lambda_j = 2\pi j/T - \pi$. (The subtraction by π ensures that we stay in the interval $[-\pi, \pi]$.) Denote entry r, s of φ via $\varphi(r, s; \lambda)$; each of these components are a function of λ. Equation (4.3.15) of Brillinger (2001) – and utilizing Assumption 1(2) – yields

$$\mathrm{cum}(d_\ell^*(\lambda_j), d_k(\lambda_i)) = O(1) + \Delta^{(T)}(\lambda_j - \lambda_i)\, \widetilde{f}_{\ell k}^*(\lambda_j), \tag{C.1}$$

where $\Delta^{(T)}(\omega)$ equals T if $\omega = 0$, but equals zero otherwise. Applying (C.1),

$$\begin{aligned}
\mathbb{E}\widehat{L}_{\varphi_3}(I) &= T^{-2} \sum_{j=1}^{T} \sum_{\ell,k,} \varphi_3(\ell, k; \lambda_j)\, \mathbb{E}[d_\ell^*(\lambda_j)\, d_k(\lambda_j)] \\
&= T^{-1} \sum_{j=1}^{T} \sum_{\ell,k,} \varphi_3(\ell, k; \lambda_j)\, \widetilde{f}_{\ell k}^*(\lambda_j) + O(T^{-1}) \\
&\to \langle [[\varphi_3\, \widetilde{f}]] \rangle_0.
\end{aligned}$$

Next, we consider the quadratic case:

$$Q_{\varphi_1,\varphi_2}(I, I) = \langle [[\varphi_1\, I\, \varphi_2\, I]] \rangle_0 = T^{-2} \langle (d^* \varphi_1 d)\, (d^* \varphi_2 d) \rangle_0.$$

By Lemma P5.1 of Brillinger (2001) this approximates the estimator $\widehat{Q}_{\varphi_1,\varphi_2}(I, I)$ with error of order T^{-1}. (Again, we could work with the estimator $Q_{\varphi_1,\varphi_2}(I, I)$, but it is simpler to focus on $\widehat{Q}_{\varphi_1,\varphi_2}(I, I)$.) Next, we write

$$\widehat{Q}_{\varphi_1,\varphi_2}(I, I) = T^{-3} \sum_{j=1}^{T} (d^* \varphi_1 d)(\lambda_j)\, (d^* \varphi_2 d)(\lambda_j). \tag{C.2}$$

Then it follows that

$$\mathbb{E}\widehat{Q}_{\varphi_1,\varphi_2}(I, I) = T^{-3} \sum_{j=1}^{T} \sum_{\ell,k,r,s} \varphi_1(\ell, k; \lambda_j)\, \varphi_2(r, s; \lambda_j)\, \mathbb{E}[d_\ell^*(\lambda_j)\, d_k(\lambda_j)\, d_r^*(\lambda_j)\, d_s(\lambda_j)],$$

and we can apply (C.1). The inner expectation of the four DFTs is broken into a sum over all indecomposable partitions; because the mean of a DFT is zero, we only need to consider three partitions that each involve two pairs. For two of these partitions, we would obtain $\Delta^{(T)} = T$, but for the third partition we obtain $\Delta^{(T)} = 0$; writing the table as $\{\ell\, k\, r\, s\}$, the substantive partitions are $\{(\ell\, k)\, (r\, s)\}$ and $\{(\ell\, s)(k\, r)\}$. Therefore, by applying Assumption 1(3) and 1(4), we obtain

$$
\begin{aligned}
\mathbb{E}\widehat{Q}_{\varphi_1,\varphi_2}(I,I) = T^{-3} \sum_{j=1}^{T} \sum_{\ell,k,r,s} & \varphi_1(\ell,k;\lambda_j)\,\varphi_2(r,s;\lambda_j)\cdot \\
& \Big\{ \Big(O(1) + T\,\widetilde{f}^{*}_{\ell k}(\lambda_j) \Big) \Big(O(1) + T\,\widetilde{f}^{*}_{rs}(\lambda_j) \Big) \\
& + \Big(O(1) + T\,\widetilde{f}^{*}_{\ell s}(\lambda_j) \Big) \Big(O(1) + T\,\widetilde{f}_{kr}(\lambda_j) \Big) \Big\} \\
\to \langle [[\varphi_1\,\widetilde{f}]]\,[[\varphi_2\,\widetilde{f}]] \rangle_0 & + \langle [[\varphi_1\,\widetilde{f}\,\varphi_2\,\widetilde{f}]] \rangle_0
\end{aligned}
$$

as $T \to \infty$, where the final line uses the fact that $\widetilde{f}$ is Hermitian. □

Proof of Proposition A.2. We provide the proof for the hardest case (the third), noting that similar techniques yield the easier two cases. Applying (C.2) twice, we obtain

$$
\begin{aligned}
& \mathrm{Cov}\left(\sqrt{T}\,\widehat{Q}_{\varphi_1,\varphi_2}(I,I), \sqrt{T}\,\widehat{Q}_{\varphi_3,\varphi_4}(I,I) \right) \\
& = T^{-5} \sum_{j_1,j_2=1}^{T} \sum_{\ell_1,\ell_2} \sum_{k_1,k_2} \sum_{r_1,r_2} \sum_{s_1,s_2} \varphi_1(\ell_1,k_1;\lambda_{j_1})\,\varphi_2(r_1,s_1;\lambda_{j_1})\,\varphi_3(\ell_2,k_2;\lambda_{j_2})\,\varphi_4(r_2,s_2;\lambda_{j_2})\cdot \\
& \quad \mathrm{cum}\left(d^{*}_{\ell_1}(\lambda_{j_1})\,d_{k_1}(\lambda_{j_1})\,d^{*}_{r_1}(\lambda_{j_1})\,d_{s_1}(\lambda_{j_1}), d^{*}_{\ell_2}(\lambda_{j_2})\,d_{k_2}(\lambda_{j_2})\,d^{*}_{r_2}(\lambda_{j_2})\,d_{s_2}(\lambda_{j_2}) \right).
\end{aligned}
$$

To compute the cumulant we utilize Theorem 2.3.2 of Brillinger (2001), which indicates that we proceed by summing over all indecomposable partitions of the *table* with two rows and four columns, multiplying the cumulants for sets of random variables (DFTs) corresponding to each set of a given partition. Hence, any partitions involving a 1-element set contribute zero, because the cumulant of a single DFT is its mean, which is zero. Which partitions are relevant depends on whether the sum over frequencies collapses to a single summation: if $\lambda_{j_1} = \pm\lambda_{j_2}$, the sum over frequencies collapses to a single summation, and the only partitions we need consider are those involving four sets of size 2 (proved below); otherwise, if $\lambda_{j_1} \neq \pm\lambda_{j_2}$ there is a double summation and the relevant partitions involve one set of size 4 and two sets of size 2 (proved below). In determining which partitions are relevant, we apply Assumption 1(k) for $2 \leq k \leq 8$, so that we can focus on those indecomposable partitions of the table that yield the highest order in T, all other partitions of lesser order being asymptotically negligible.

Diagonal Case: First suppose that $\lambda_{j_1} = \pm\lambda_{j_2}$. Because no 1-element sets need be considered, the maximal number of sets in a partition is four (which must be four 2-element sets) – and we show that some of these partitions will yield a cumulant $O(T^4)$. Any other type of partition would have fewer than three sets, so that the cumulant would be at most $O(T^3)$, and thus can be ignored. First setting $\lambda_{j_1} = \lambda_{j_2}$, we write the table

$$
\begin{array}{cccc}
d^{*}_{\ell_1}(\lambda_{j_1}) & d_{k_1}(\lambda_{j_1}) & d^{*}_{r_1}(\lambda_{j_1}) & d_{s_1}(\lambda_{j_1}) \\
d^{*}_{\ell_2}(\lambda_{j_1}) & d_{k_2}(\lambda_{j_1}) & d^{*}_{r_2}(\lambda_{j_1}) & d_{s_2}(\lambda_{j_1}).
\end{array}
$$

A four 2-element set partition that is indecomposable must have at least one 2-element set with an element in both of the two rows. There are many of these, but only 20 of them are $O(T^4)$: using (C.1), we only need consider 2-element sets where the sum of the corresponding frequencies is zero, i.e., sets where one element corresponds to a DFT and the other element to a conjugate DFT. We denote these 20 partitions with the following notation: the symbols $\sharp$, $\flat$, $\natural$, and $\star$ will denote membership in a particular 2-element set:

$$\begin{bmatrix} \sharp & \sharp & \flat & \natural \\ \natural & \flat & \star & \star \end{bmatrix} \quad \begin{bmatrix} \sharp & \sharp & \flat & \natural \\ \star & \flat & \natural & \star \end{bmatrix} \quad \begin{bmatrix} \sharp & \sharp & \flat & \natural \\ \natural & \star & \star & \flat \end{bmatrix} \quad \begin{bmatrix} \sharp & \sharp & \flat & \natural \\ \star & \star & \natural & \flat \end{bmatrix}$$

$$\begin{bmatrix} \sharp & \flat & \natural & \sharp \\ \flat & \natural & \star & \star \end{bmatrix} \quad \begin{bmatrix} \sharp & \flat & \natural & \sharp \\ \flat & \star & \star & \natural \end{bmatrix} \quad \begin{bmatrix} \sharp & \flat & \natural & \sharp \\ \star & \star & \flat & \natural \end{bmatrix} \quad \begin{bmatrix} \sharp & \flat & \natural & \sharp \\ \star & \natural & \flat & \star \end{bmatrix}$$

$$\begin{bmatrix} \sharp & \flat & \natural & \natural \\ \flat & \sharp & \star & \star \end{bmatrix} \quad \begin{bmatrix} \sharp & \flat & \natural & \star \\ \flat & \sharp & \star & \natural \end{bmatrix} \quad \begin{bmatrix} \sharp & \flat & \natural & \star \\ \star & \sharp & \flat & \natural \end{bmatrix}$$

$$\begin{bmatrix} \sharp & \flat & \natural & \natural \\ \star & \sharp & \flat & \star \end{bmatrix} \quad \begin{bmatrix} \sharp & \flat & \flat & \natural \\ \natural & \sharp & \star & \star \end{bmatrix} \quad \begin{bmatrix} \sharp & \flat & \flat & \natural \\ \star & \sharp & \natural & \star \end{bmatrix}$$

$$\begin{bmatrix} \sharp & \flat & \natural & \natural \\ \flat & \star & \star & \sharp \end{bmatrix} \quad \begin{bmatrix} \sharp & \flat & \natural & \star \\ \flat & \natural & \star & \sharp \end{bmatrix} \quad \begin{bmatrix} \sharp & \flat & \natural & \star \\ \star & \natural & \flat & \sharp \end{bmatrix}$$

$$\begin{bmatrix} \sharp & \flat & \natural & \natural \\ \star & \star & \flat & \sharp \end{bmatrix} \quad \begin{bmatrix} \sharp & \flat & \flat & \natural \\ \natural & \star & \star & \sharp \end{bmatrix} \quad \begin{bmatrix} \sharp & \flat & \flat & \natural \\ \star & \star & \natural & \sharp \end{bmatrix}$$

As a result, the covariance of the quadratic forms has an asymptotic contribution from the diagonal case (with $\lambda_{j_1} = \lambda_{j_2}$) of

$$\begin{aligned} T^{-1} \sum_{j=1}^{T} \sum_{\ell_1,\ell_2} \sum_{k_1,k_2} \sum_{r_1,r_2} \sum_{s_1,s_2} & \varphi_1(\ell_1,k_1;\lambda_j)\, \varphi_2(r_1,s_1;\lambda_j)\, \varphi_3(\ell_2,k_2;\lambda_j)\, \varphi_4(r_2,s_2;\lambda_j) \cdot \\ & \Big\{ \widetilde{f}^*_{\ell_1 k_1}(\lambda_j)\, \widetilde{f}^*_{r_1 k_2}(\lambda_j)\, \widetilde{f}_{s_1 \ell_2}(\lambda_j)\, \widetilde{f}^*_{r_2 s_2}(\lambda_j) + \widetilde{f}^*_{\ell_1 k_1}(\lambda_j)\, \widetilde{f}^*_{r_1 k_2}(\lambda_j)\, \widetilde{f}_{s_1 r_2}(\lambda_j)\, \widetilde{f}^*_{\ell_2 s_2}(\lambda_j) \\ & + \widetilde{f}^*_{\ell_1 k_1}(\lambda_j)\, \widetilde{f}^*_{r_1 s_2}(\lambda_j)\, \widetilde{f}_{s_1 \ell_2}(\lambda_j)\, \widetilde{f}_{k_2 r_2}(\lambda_j) + \widetilde{f}^*_{\ell_1 k_1}(\lambda_j)\, \widetilde{f}^*_{r_1 s_2}(\lambda_j)\, \widetilde{f}_{s_1 r_2}(\lambda_j)\, \widetilde{f}^*_{\ell_2 k_2}(\lambda_j) \\ & + \widetilde{f}^*_{\ell_1 s_1}(\lambda_j)\, \widetilde{f}_{k_1 \ell_2}(\lambda_j)\, \widetilde{f}^*_{r_1 k_2}(\lambda_j)\, \widetilde{f}^*_{r_2 s_2}(\lambda_j) + \widetilde{f}^*_{\ell_1 s_1}(\lambda_j)\, \widetilde{f}_{k_1 \ell_2}(\lambda_j)\, \widetilde{f}^*_{r_1 s_2}(\lambda_j)\, \widetilde{f}_{k_2 r_2}(\lambda_j) \\ & + \widetilde{f}^*_{\ell_1 s_1}(\lambda_j)\, \widetilde{f}_{k_1 r_2}(\lambda_j)\, \widetilde{f}^*_{r_1 s_2}(\lambda_j)\, \widetilde{f}^*_{\ell_2 k_2}(\lambda_j) + \widetilde{f}^*_{\ell_1 s_1}(\lambda_j)\, \widetilde{f}_{k_1 r_2}(\lambda_j)\, \widetilde{f}^*_{r_1 k_2}(\lambda_j)\, \widetilde{f}^*_{\ell_2 s_2}(\lambda_j) \\ & + \widetilde{f}^*_{\ell_1 k_2}(\lambda_j)\, \widetilde{f}_{k_1 \ell_2}(\lambda_j)\, \widetilde{f}^*_{r_1 s_1}(\lambda_j)\, \widetilde{f}^*_{r_2 s_2}(\lambda_j) + \widetilde{f}^*_{\ell_1 k_2}(\lambda_j)\, \widetilde{f}_{k_1 \ell_2}(\lambda_j)\, \widetilde{f}^*_{r_1 s_2}(\lambda_j)\, \widetilde{f}_{s_1 r_2}(\lambda_j) \\ & + \widetilde{f}^*_{\ell_1 k_2}(\lambda_j)\, \widetilde{f}_{k_1 r_2}(\lambda_j)\, \widetilde{f}^*_{r_1 s_2}(\lambda_j)\, \widetilde{f}_{s_1 \ell_2}(\lambda_j) + \widetilde{f}^*_{\ell_1 k_2}(\lambda_j)\, \widetilde{f}_{k_1 r_2}(\lambda_j)\, \widetilde{f}^*_{r_1 s_1}(\lambda_j)\, \widetilde{f}^*_{\ell_2 s_2}(\lambda_j) \\ & + \widetilde{f}^*_{\ell_1 k_2}(\lambda_j)\, \widetilde{f}_{k_1 r_1}(\lambda_j)\, \widetilde{f}_{s_1 \ell_2}(\lambda_j)\, \widetilde{f}^*_{r_2 s_2}(\lambda_j) + \widetilde{f}^*_{\ell_1 k_2}(\lambda_j)\, \widetilde{f}_{k_1 r_1}(\lambda_j)\, \widetilde{f}_{s_1 r_2}(\lambda_j)\, \widetilde{f}^*_{\ell_2 s_2}(\lambda_j) \\ & + \widetilde{f}^*_{\ell_1 s_2}(\lambda_j)\, \widetilde{f}_{k_1 \ell_2}(\lambda_j)\, \widetilde{f}^*_{r_1 s_1}(\lambda_j)\, \widetilde{f}_{k_2 r_2}(\lambda_j) + \widetilde{f}^*_{\ell_1 s_2}(\lambda_j)\, \widetilde{f}_{k_1 \ell_2}(\lambda_j)\, \widetilde{f}^*_{r_1 k_2}(\lambda_j)\, \widetilde{f}_{s_1 r_2}(\lambda_j) \\ & + \widetilde{f}^*_{\ell_1 s_2}(\lambda_j)\, \widetilde{f}_{k_1 r_2}(\lambda_j)\, \widetilde{f}^*_{r_1 k_2}(\lambda_j)\, \widetilde{f}_{s_1 \ell_2}(\lambda_j) + \widetilde{f}^*_{\ell_1 s_2}(\lambda_j)\, \widetilde{f}_{k_1 r_2}(\lambda_j)\, \widetilde{f}^*_{r_1 s_1}(\lambda_j)\, \widetilde{f}^*_{\ell_2 k_2}(\lambda_j) \\ & + \widetilde{f}^*_{\ell_1 s_2}(\lambda_j)\, \widetilde{f}_{k_1 r_1}(\lambda_j)\, \widetilde{f}_{s_1 \ell_2}(\lambda_j)\, \widetilde{f}_{k_2 r_2}(\lambda_j) + \widetilde{f}^*_{\ell_1 s_2}(\lambda_j)\, \widetilde{f}_{k_1 r_1}(\lambda_j)\, \widetilde{f}_{s_1 r_2}(\lambda_j)\, \widetilde{f}^*_{\ell_2 k_2}(\lambda_j) \Big\}. \end{aligned}$$

This converges to

$$\begin{aligned} & \langle [1][4][32] \rangle_0 + \langle [1][243] \rangle_0 + \langle [1][234] \rangle_0 + \langle [1][3][24] \rangle_0 \\ & + \langle [4][132] \rangle_0 + \langle [1342] \rangle_0 + \langle [3][142] \rangle_0 + \langle [1432] \rangle_0 \\ & + \langle [2][4][31] \rangle_0 + \langle [31][42] \rangle_0 + \langle [1423] \rangle_0 + \langle [2][143] \rangle_0 \\ & + \langle [4][123] \rangle_0 + \langle [1243] \rangle_0 + \langle [2][134] \rangle_0 + \langle [1324] \rangle_0 \\ & + \langle [14][23] \rangle_0 + \langle [3][2][14] \rangle_0 + \langle [1234] \rangle_0 + \langle [3][124] \rangle_0 \end{aligned}$$

Next, setting $\lambda_{j_1} = -\lambda_{j_2}$ we write the table

$$
\begin{array}{cccc}
d^*_{\ell_1}(\lambda_{j_1}) & d_{k_1}(\lambda_{j_1}) & d^*_{r_1}(\lambda_{j_1}) & d_{s_1}(\lambda_{j_1}) \\
d_{\ell_2}(\lambda_{j_1}) & d^*_{k_2}(\lambda_{j_1}) & d_{r_2}(\lambda_{j_1}) & d^*_{s_2}(\lambda_{j_1}).
\end{array}
$$

Again, there are 20 relevant partitions:

$$
\begin{bmatrix} \sharp & \sharp & \flat & \star \\ \flat & \natural & \natural & \star \end{bmatrix}
\quad
\begin{bmatrix} \sharp & \sharp & \flat & \natural \\ \flat & \natural & \star & \star \end{bmatrix}
\quad
\begin{bmatrix} \sharp & \sharp & \flat & \natural \\ \star & \star & \flat & \natural \end{bmatrix}
\quad
\begin{bmatrix} \sharp & \sharp & \flat & \natural \\ \star & \natural & \flat & \star \end{bmatrix}
$$

$$
\begin{bmatrix} \sharp & \flat & \natural & \sharp \\ \star & \flat & \natural & \star \end{bmatrix}
\quad
\begin{bmatrix} \sharp & \flat & \natural & \sharp \\ \natural & \flat & \star & \star \end{bmatrix}
\quad
\begin{bmatrix} \sharp & \flat & \natural & \sharp \\ \star & \star & \natural & \flat \end{bmatrix}
\quad
\begin{bmatrix} \sharp & \flat & \natural & \sharp \\ \natural & \star & \star & \flat \end{bmatrix}
$$

$$
\begin{bmatrix} \sharp & \flat & \natural & \natural \\ \sharp & \flat & \star & \star \end{bmatrix}
\quad
\begin{bmatrix} \sharp & \flat & \natural & \star \\ \sharp & \flat & \natural & \star \end{bmatrix}
\quad
\begin{bmatrix} \sharp & \flat & \flat & \star \\ \sharp & \natural & \natural & \star \end{bmatrix}
$$

$$
\begin{bmatrix} \sharp & \flat & \flat & \natural \\ \sharp & \natural & \star & \star \end{bmatrix}
\quad
\begin{bmatrix} \sharp & \flat & \star & \natural \\ \sharp & \natural & \star & \flat \end{bmatrix}
\quad
\begin{bmatrix} \sharp & \flat & \star & \star \\ \sharp & \natural & \natural & \flat \end{bmatrix}
$$

$$
\begin{bmatrix} \sharp & \flat & \flat & \star \\ \natural & \natural & \sharp & \star \end{bmatrix}
\quad
\begin{bmatrix} \sharp & \flat & \flat & \star \\ \natural & \star & \sharp & \natural \end{bmatrix}
\quad
\begin{bmatrix} \sharp & \flat & \star & \star \\ \natural & \flat & \sharp & \natural \end{bmatrix}
$$

$$
\begin{bmatrix} \sharp & \flat & \natural & \star \\ \natural & \flat & \sharp & \star \end{bmatrix}
\quad
\begin{bmatrix} \sharp & \flat & \star & \star \\ \natural & \natural & \sharp & \flat \end{bmatrix}
\quad
\begin{bmatrix} \sharp & \flat & \natural & \star \\ \natural & \star & \sharp & \flat \end{bmatrix}
$$

Hence, the covariance of the quadratic forms has an asymptotic contribution from the diagonal case (with $-\lambda_{j_1} = \lambda_{j_2}$) of

$$
\begin{aligned}
T^{-1} \sum_{j=1}^{T} \sum_{\ell_1,\ell_2} \sum_{k_1,k_2} \sum_{r_1,r_2} \sum_{s_1,s_2} & \varphi_1(\ell_1, k_1; \lambda_j)\, \varphi_2(r_1, s_1; \lambda_j)\, \varphi_3(\ell_2, k_2; -\lambda_j)\, \varphi_4(r_2, s_2; -\lambda_j) \cdot \\
& \Big\{ \widetilde{f}^*_{\ell_1 k_1}(\lambda_j)\, \widetilde{f}^*_{r_1 \ell_2}(\lambda_j)\, \widetilde{f}^*_{k_2 r_2}(\lambda_j)\, \widetilde{f}_{s_1 s_2}(\lambda_j) + \widetilde{f}^*_{\ell_1 k_1}(\lambda_j)\, \widetilde{f}^*_{r_1 \ell_2}(\lambda_j)\, \widetilde{f}_{s_1 k_2}(\lambda_j)\, \widetilde{f}_{r_2 s_2}(\lambda_j) \\
& + \widetilde{f}^*_{\ell_1 k_1}(\lambda_j)\, \widetilde{f}^*_{r_1 r_2}(\lambda_j)\, \widetilde{f}_{s_1 s_2}(\lambda_j)\, \widetilde{f}_{\ell_2 k_2}(\lambda_j) + \widetilde{f}^*_{\ell_1 k_1}(\lambda_j)\, \widetilde{f}^*_{r_1 r_2}(\lambda_j)\, \widetilde{f}_{s_1 k_2}(\lambda_j)\, \widetilde{f}_{\ell_2 s_2}(\lambda_j) \\
& + \widetilde{f}^*_{\ell_1 s_1}(\lambda_j)\, \widetilde{f}_{k_1 k_2}(\lambda_j)\, \widetilde{f}^*_{r_1 r_2}(\lambda_j)\, \widetilde{f}_{\ell_2 s_2}(\lambda_j) + \widetilde{f}^*_{\ell_1 s_1}(\lambda_j)\, \widetilde{f}_{k_1 k_2}(\lambda_j)\, \widetilde{f}^*_{r_1 \ell_2}(\lambda_j)\, \widetilde{f}_{r_2 s_2}(\lambda_j) \\
& + \widetilde{f}^*_{\ell_1 s_1}(\lambda_j)\, \widetilde{f}_{k_1 s_2}(\lambda_j)\, \widetilde{f}^*_{r_1 r_2}(\lambda_j)\, \widetilde{f}_{\ell_2 k_2}(\lambda_j) + \widetilde{f}^*_{\ell_1 s_1}(\lambda_j)\, \widetilde{f}_{k_1 s_2}(\lambda_j)\, \widetilde{f}^*_{r_1 \ell_2}(\lambda_j)\, \widetilde{f}^*_{k_2 r_2}(\lambda_j) \\
& + \widetilde{f}^*_{\ell_1 \ell_2}(\lambda_j)\, \widetilde{f}_{k_1 k_2}(\lambda_j)\, \widetilde{f}^*_{r_1 s_1}(\lambda_j)\, \widetilde{f}_{r_2 s_2}(\lambda_j) + \widetilde{f}^*_{\ell_1 \ell_2}(\lambda_j)\, \widetilde{f}_{k_1 k_2}(\lambda_j)\, \widetilde{f}^*_{r_1 r_2}(\lambda_j)\, \widetilde{f}_{s_1 s_2}(\lambda_j) \\
& + \widetilde{f}^*_{\ell_1 \ell_2}(\lambda_j)\, \widetilde{f}_{k_1 r_1}(\lambda_j)\, \widetilde{f}^*_{k_2 r_2}(\lambda_j)\, \widetilde{f}_{s_1 s_2}(\lambda_j) + \widetilde{f}^*_{\ell_1 \ell_2}(\lambda_j)\, \widetilde{f}_{k_1 r_1}(\lambda_j)\, \widetilde{f}_{s_1 k_2}(\lambda_j)\, \widetilde{f}_{r_2 s_2}(\lambda_j) \\
& + \widetilde{f}^*_{\ell_1 \ell_2}(\lambda_j)\, \widetilde{f}_{k_1 s_2}(\lambda_j)\, \widetilde{f}_{s_1 k_2}(\lambda_j)\, \widetilde{f}^*_{r_1 r_2}(\lambda_j) + \widetilde{f}^*_{\ell_1 \ell_2}(\lambda_j)\, \widetilde{f}_{k_1 s_2}(\lambda_j)\, \widetilde{f}^*_{k_2 r_2}(\lambda_j)\, \widetilde{f}^*_{r_1 s_1}(\lambda_j) \\
& + \widetilde{f}^*_{\ell_1 r_2}(\lambda_j)\, \widetilde{f}_{k_1 r_1}(\lambda_j)\, \widetilde{f}_{\ell_2 k_2}(\lambda_j)\, \widetilde{f}_{s_1 s_2}(\lambda_j) + \widetilde{f}^*_{\ell_1 r_2}(\lambda_j)\, \widetilde{f}_{k_1 r_1}(\lambda_j)\, \widetilde{f}_{\ell_2 s_2}(\lambda_j)\, \widetilde{f}_{s_1 k_2}(\lambda_j) \\
& + \widetilde{f}^*_{\ell_1 r_2}(\lambda_j)\, \widetilde{f}_{k_1 k_2}(\lambda_j)\, \widetilde{f}_{\ell_2 s_2}(\lambda_j)\, \widetilde{f}^*_{r_1 s_1}(\lambda_j) + \widetilde{f}^*_{\ell_1 r_2}(\lambda_j)\, \widetilde{f}_{k_1 k_2}(\lambda_j)\, \widetilde{f}^*_{r_1 \ell_2}(\lambda_j)\, \widetilde{f}_{s_1 s_2}(\lambda_j) \\
& + \widetilde{f}^*_{\ell_1 r_2}(\lambda_j)\, \widetilde{f}_{k_1 s_2}(\lambda_j)\, \widetilde{f}_{\ell_2 k_2}(\lambda_j)\, \widetilde{f}^*_{r_1 s_1}(\lambda_j) + \widetilde{f}^*_{\ell_1 r_2}(\lambda_j)\, \widetilde{f}_{k_1 s_2}(\lambda_j)\, \widetilde{f}^*_{r_1 \ell_2}(\lambda_j)\, \widetilde{f}_{s_1 k_2}(\lambda_j) \Big\}.
\end{aligned}
$$

Noting that the argument of φ_3 and φ_4 is $-\lambda_j$, and using the underscore notation, the above converges to

$$
\begin{aligned}
&\langle [1][2\underline{43}] \rangle_0 + \langle [\underline{4}][1][2\underline{3}] \rangle_0 + \langle [1][\underline{3}][2\underline{4}] \rangle_0 + \langle [1][2\underline{34}] \rangle_0 \\
&+ \langle [1\underline{34}2] \rangle_0 + \langle [1\underline{3}2][\underline{4}] \rangle_0 + \langle [1\underline{4}2][\underline{3}] \rangle_0 + \langle [1\underline{43}2] \rangle_0 \\
&+ \langle [\underline{4}][2][1\underline{3}] \rangle_0 + \langle [1\underline{3}][2\underline{4}] \rangle_0 + \langle [12\underline{43}] \rangle_0 + \langle [12\underline{3}][\underline{4}] \rangle_0 \\
&+ \langle [1\underline{4}2\underline{3}] \rangle_0 + \langle [1\underline{43}][2] \rangle_0 + \langle [12\underline{4}][\underline{3}] \rangle_0 + \langle [12\underline{34}] \rangle_0 \\
&+ \langle [1\underline{34}][2] \rangle_0 + \langle [1\underline{3}2\underline{4}] \rangle_0 + \langle [1\underline{4}][2][\underline{3}] \rangle_0 + \langle [1\underline{4}][2\underline{3}] \rangle_0 .
\end{aligned}
$$

This accounts for the entire contribution from the diagonal case.

Off-diagonal Case: Now we suppose that $\lambda_{j_1} \neq \lambda_{j_2}$, and hence the double sum does not collapse to a single summation. In this case, of the partitions with four 2-element sets they are all either decomposable or only contribute terms of order $O(T^2)$. This is because such a partition that is indecomposable must have two sets with an element in each row – see examples in the prior case. (The definition of indecomposable requires one set to have an element in each row, but as this would only leave three free slots in each row, we must have at least one other set with this property.) But because $\lambda_{j_1} \neq \lambda_{j_2}$, (C.1) ensures that the contribution to the cumulant from such sets is $O(1)$, indicating that the largest possible order from such a partition is $O(T^2)$. Such terms can be ignored, because there exist indecomposable partitions yielding $O(T^3)$ terms in the cumulants: these partitions involve one 4-element set and two 2-element sets. Moreover, no other partitions need be considered: the only other partitions (that don't involve 1-element sets) with three sets would have two 3-element sets and one 2-element set, but the cumulant of a 3-element set will never be $O(T)$, according to equation (4.3.15) of Brillinger (2001). This expression states that the m-fold cumulant of m DFTs can be $O(T)$ so long as the sum of the frequency arguments is zero; there is no way this can happen when the frequencies take the form $\lambda_{j_1}, -\lambda_{j_1}, \lambda_{j_2}, -\lambda_{j_2}$. The expression of equation (4.3.15) of Brillinger (2001) in the case $m = 4$ (using Assumption 1) is

$$
\mathrm{cum}(d_\ell(\lambda_{j_1}), d_k(\lambda_{j_2}), d_r(\lambda_{j_3}), d_s(\lambda_{j_4})) = O(1) + \Delta^{(T)} \left(\sum_{i=1}^{4} \lambda_{j_i} \right) \widetilde{f}_{\ell k r s}(\lambda_{j_1}, \lambda_{j_2}, \lambda_{j_3}). \tag{C.3}
$$

Using the table

$$
\begin{array}{cccc}
d^*_{\ell_1}(\lambda_{j_1}) & d_{k_1}(\lambda_{j_1}) & d^*_{r_1}(\lambda_{j_1}) & d_{s_1}(\lambda_{j_1}) \\
d^*_{\ell_2}(\lambda_{j_2}) & d_{k_2}(\lambda_{j_2}) & d^*_{r_2}(\lambda_{j_2}) & d_{s_2}(\lambda_{j_2}),
\end{array}
$$

we find that there are 16 indecomposable partitions consisting of one 4-element and two 2-element sets, such that the contribution is $O(T^3)$. The four elements of the 4-element set must be allocated with two elements in each row, because if three belong to a single row it is impossible for the sum of all frequencies to equal zero, as required in (C.3). (Also, if all four elements belonged to a single row, the partition would be decomposable.) Also, once the 4-element set is allocated with two members in each row, each of the remaining 2-element sets must be contained in a single row (otherwise the sum of frequencies cannot equal zero, and the contribution will be less than $O(T^3)$). There are 16 such partitions, which we list below:

$$\begin{bmatrix} \sharp & \sharp & \flat & \flat \\ \sharp & \sharp & \natural & \natural \end{bmatrix} \begin{bmatrix} \sharp & \sharp & \flat & \flat \\ \sharp & \natural & \natural & \sharp \end{bmatrix} \begin{bmatrix} \sharp & \sharp & \flat & \flat \\ \natural & \sharp & \sharp & \natural \end{bmatrix} \begin{bmatrix} \sharp & \sharp & \flat & \flat \\ \natural & \natural & \sharp & \sharp \end{bmatrix}$$
$$\begin{bmatrix} \sharp & \flat & \flat & \sharp \\ \sharp & \sharp & \natural & \natural \end{bmatrix} \begin{bmatrix} \sharp & \flat & \flat & \sharp \\ \sharp & \natural & \natural & \sharp \end{bmatrix} \begin{bmatrix} \sharp & \flat & \flat & \sharp \\ \natural & \sharp & \sharp & \natural \end{bmatrix} \begin{bmatrix} \sharp & \flat & \flat & \sharp \\ \natural & \natural & \sharp & \sharp \end{bmatrix}$$
$$\begin{bmatrix} \flat & \sharp & \sharp & \flat \\ \sharp & \sharp & \natural & \natural \end{bmatrix} \begin{bmatrix} \flat & \sharp & \sharp & \flat \\ \sharp & \natural & \natural & \sharp \end{bmatrix} \begin{bmatrix} \flat & \sharp & \sharp & \flat \\ \natural & \sharp & \sharp & \natural \end{bmatrix} \begin{bmatrix} \flat & \sharp & \sharp & \flat \\ \natural & \natural & \sharp & \sharp \end{bmatrix}$$
$$\begin{bmatrix} \flat & \flat & \sharp & \sharp \\ \sharp & \sharp & \natural & \natural \end{bmatrix} \begin{bmatrix} \flat & \flat & \sharp & \sharp \\ \sharp & \natural & \natural & \sharp \end{bmatrix} \begin{bmatrix} \flat & \flat & \sharp & \sharp \\ \natural & \sharp & \sharp & \natural \end{bmatrix} \begin{bmatrix} \flat & \flat & \sharp & \sharp \\ \natural & \natural & \sharp & \sharp \end{bmatrix}$$

The contribution to the covariance is therefore

$$\begin{aligned}
T^{-2} & \sum_{j_1 \neq j_2 = 1}^{T} \sum_{\ell_1,\ell_2} \sum_{k_1,k_2} \sum_{r_1,r_2} \sum_{s_1,s_2} \varphi_1(\ell_1,k_1;\lambda_{j_1})\,\varphi_2(r_1,s_1;\lambda_{j_1})\,\varphi_3(\ell_2,k_2;\lambda_{j_2})\,\varphi_4(r_2,s_2;\lambda_{j_2}) \cdot \\
& \Big\{ \widetilde{f}_{\ell_1 k_1 \ell_2 k_2}(-\lambda_{j_1},\lambda_{j_1},-\lambda_{j_2})\,\widetilde{f}^*_{r_1 s_1}(\lambda_{j_1})\,\widetilde{f}^*_{r_2 s_2}(\lambda_{j_2}) + \widetilde{f}_{\ell_1 k_1 \ell_2 s_2}(-\lambda_{j_1},\lambda_{j_1},-\lambda_{j_2})\,\widetilde{f}^*_{r_1 s_1}(\lambda_{j_1})\,\widetilde{f}^*_{k_2 r_2}(\lambda_{j_2}) \\
& + \widetilde{f}_{\ell_1 k_1 k_2 r_2}(-\lambda_{j_1},\lambda_{j_1},\lambda_{j_2})\,\widetilde{f}^*_{r_1 s_1}(\lambda_{j_1})\,\widetilde{f}^*_{\ell_2 s_2}(\lambda_{j_2}) + \widetilde{f}_{\ell_1 k_1 r_2 s_2}(-\lambda_{j_1},\lambda_{j_1},-\lambda_{j_2})\,\widetilde{f}^*_{r_1 s_1}(\lambda_{j_1})\,\widetilde{f}^*_{\ell_2 k_2}(\lambda_{j_2}) \\
& + \widetilde{f}_{\ell_1 s_1 \ell_2 k_2}(-\lambda_{j_1},\lambda_{j_1},-\lambda_{j_2})\,\widetilde{f}_{k_1 r_1}(\lambda_{j_1})\,\widetilde{f}^*_{r_2 s_2}(\lambda_{j_2}) + \widetilde{f}_{\ell_1 s_1 \ell_2 s_2}(-\lambda_{j_1},\lambda_{j_1},-\lambda_{j_2})\,\widetilde{f}_{k_1 r_1}(\lambda_{j_1})\,\widetilde{f}_{k_2 r_2}(\lambda_{j_2}) \\
& + \widetilde{f}_{\ell_1 s_1 k_2 r_2}(-\lambda_{j_1},\lambda_{j_1},\lambda_{j_2})\,\widetilde{f}_{k_1 r_1}(\lambda_{j_1})\,\widetilde{f}^*_{\ell_2 s_2}(\lambda_{j_2}) + \widetilde{f}_{\ell_1 s_1 r_2 s_2}(-\lambda_{j_1},\lambda_{j_1},-\lambda_{j_2})\,\widetilde{f}_{k_1 r_1}(\lambda_{j_1})\,\widetilde{f}^*_{\ell_2 k_2}(\lambda_{j_2}) \\
& + \widetilde{f}_{k_1 r_1 \ell_2 k_2}(\lambda_{j_1},-\lambda_{j_1},-\lambda_{j_2})\,\widetilde{f}^*_{\ell_1 s_1}(\lambda_{j_1})\,\widetilde{f}^*_{r_2 s_2}(\lambda_{j_2}) + \widetilde{f}_{k_1 r_1 \ell_2 s_2}(\lambda_{j_1},-\lambda_{j_1},-\lambda_{j_2})\,\widetilde{f}^*_{\ell_1 s_1}(\lambda_{j_1})\,\widetilde{f}_{k_2 r_2}(\lambda_{j_2}) \\
& + \widetilde{f}_{k_1 r_1 k_2 r_2}(\lambda_{j_1},-\lambda_{j_1},\lambda_{j_2})\,\widetilde{f}^*_{\ell_1 s_1}(\lambda_{j_1})\,\widetilde{f}^*_{\ell_2 s_2}(\lambda_{j_2}) + \widetilde{f}_{k_1 r_1 r_2 s_2}(\lambda_{j_1},-\lambda_{j_1},-\lambda_{j_2})\,\widetilde{f}^*_{\ell_1 s_1}(\lambda_{j_1})\,\widetilde{f}^*_{\ell_2 k_2}(\lambda_{j_2}) \\
& + \widetilde{f}_{r_1 s_1 \ell_2 k_2}(-\lambda_{j_1},\lambda_{j_1},-\lambda_{j_2})\,\widetilde{f}^*_{\ell_1 k_1}(\lambda_{j_1})\,\widetilde{f}^*_{r_2 s_2}(\lambda_{j_2}) + \widetilde{f}_{r_1 s_1 \ell_2 s_2}(-\lambda_{j_1},\lambda_{j_1},-\lambda_{j_2})\,\widetilde{f}^*_{\ell_1 k_1}(\lambda_{j_1})\,\widetilde{f}_{k_2 r_2}(\lambda_{j_2}) \\
& + \widetilde{f}_{r_1 s_1 k_2 r_2}(-\lambda_{j_1},\lambda_{j_1},\lambda_{j_2})\,\widetilde{f}^*_{\ell_1 k_1}(\lambda_{j_1})\,\widetilde{f}^*_{\ell_2 s_2}(\lambda_{j_2}) + \widetilde{f}_{r_1 s_1 r_2 s_2}(-\lambda_{j_1},\lambda_{j_1},-\lambda_{j_2})\,\widetilde{f}^*_{\ell_1 k_1}(\lambda_{j_1})\,\widetilde{f}^*_{\ell_2 k_2}(\lambda_{j_2}) \Big\}.
\end{aligned}$$

The tri-spectral density has the properties that $\widetilde{f}_{\ell k s r}(\lambda,-\lambda,\omega) = \widetilde{f}_{\ell k r s}(\lambda,-\lambda,-\omega)$, $\widetilde{f}_{\ell k s r}(\lambda,-\lambda,\omega) = \widetilde{f}_{k \ell s r}(-\lambda,\lambda,\omega)$, and $\widetilde{f}_{\ell k s r}(\lambda,-\lambda,\omega) = \widetilde{f}_{s r \ell k}(\omega,-\omega,\lambda)$. These are verified as follows:

$$\begin{aligned}
\widetilde{f}_{\ell k s r}(\lambda,-\lambda,\omega) &= \sum_{h_1,h_2,h_3} \gamma_{\ell k s r}(h_1,h_2,h_3)\, e^{-i\lambda(h_1-h_2)-i\omega h_3} \\
&= \sum_{h_1,h_2,h_3} \gamma_{\ell k r s}(h_1-h_3,h_2-h_3,-h_3)\, e^{-i\lambda(h_1-h_2)-i\omega h_3} \\
&= \sum_{k_1,k_2,k_3} \gamma_{\ell k r s}(k_1,k_2,k_3)\, e^{-i\lambda(k_1-k_2)+i\omega k_3} = \widetilde{f}_{\ell k r s}(\lambda,-\lambda,-\omega) \\
\widetilde{f}_{\ell k s r}(\lambda,-\lambda,\omega) &= \sum_{h_1,h_2,h_3} \gamma_{k \ell s r}(h_2,h_1,h_3)\, e^{-i\lambda(h_1-h_2)-i\omega h_3} \\
&= \sum_{k_1,k_2,k_3} \gamma_{k \ell s r}(k_1,k_2,k_3)\, e^{i\lambda(k_1-k_2)-i\omega k_3} = \widetilde{f}_{k \ell s r}(-\lambda,\lambda,\omega) \\
\widetilde{f}_{\ell k s r}(\lambda,-\lambda,\omega) &= \sum_{h_1,h_2,h_3} \gamma_{\ell k s r}(h_1,h_2,h_3)\, e^{-i\lambda(h_1-h_2)-i\omega h_3} \\
&= \sum_{h_1,h_2,h_3} \gamma_{s r \ell k}(h_3-h_2,-h_2,h_1-h_2)\, e^{-i\lambda(h_1-h_2)-i\omega h_3} \\
&= \sum_{k_1,k_2,k_3} \gamma_{s r \ell k}(k_1,k_2,k_3)\, e^{i\lambda k_3-i\omega(k_1-k_2)} = \widetilde{f}_{s r \ell k}(\omega,-\omega,\lambda).
\end{aligned}$$

In the first calculation, this uses

$$\begin{aligned}\gamma_{\ell ksr}(h_1,h_2,h_3) &= \mathrm{cum}\{x_{h_1,\ell}, x_{h_2,k}, x_{h_3,s}, x_{0,r}\} \\ &= \mathrm{cum}\{x_{h_1-h_3,\ell}, x_{h_2-h_3,k}, x_{-h_3,r}, x_{0,s}\} = \gamma_{\ell krs}(h_1-h_3, h_2-h_3, -h_3)\end{aligned}$$

and the change of variable $k_1 = h_1 - h_3$, $k_2 = h_2 - h_3$, and $k_3 = -h_3$. The second calculation uses the fact that cumulant arguments can be permuted, and the change of variable $k_1 = h_2$, $k_2 = h_1$, and $k_3 = h_3$. The third calculation uses

$$\begin{aligned}\gamma_{\ell ksr}(h_1,h_2,h_3) &= \mathrm{cum}\{x_{h_1,\ell}, x_{h_2,k}, x_{h_3,s}, x_{0,r}\} \\ &= \mathrm{cum}\{x_{h_3-h_2,s}, x_{-h_2,r}, x_{h_1-h_2,\ell}, x_{0,k}\} = \gamma_{sr\ell k}(h_3-h_2, -h_2, h_1-h_2)\end{aligned}$$

and the change of variable $k_1 = h_3 - h_2$, $k_2 = -h_2$, and $k_3 = h_1 - h_2$. We can use these properties to express every occurence of the tri-spectral density in terms of the frequency arguments $(\lambda_{j_1}, -\lambda_{j_1}, \lambda_{j_2})$ by rearranging the subscript indices. Then it is possible to greatly simplify the summations, letting $T \to \infty$. (The double sum over frequencies will tend to a double integral, not-withstanding the omission of the diagonal portion, which has measure zero.) Then the limiting contribution to the covariance (where curly braces denote a matrix argument to the double bracket) simplifies to the double integral (weighted by $(2\pi)^{-2}$) over λ and ω of

$$\begin{aligned}g_{1234}(\lambda,\omega) = & [[\varphi_1(\lambda)\,\widetilde{f}(\lambda,-\lambda,\omega)\,\varphi_3(\omega)]]\,[[\varphi_2(\lambda)\,\widetilde{f}(\lambda)]]\,[[\varphi_4(\omega)\,\widetilde{f}(\omega)]] \\ & + [[\varphi_1(\lambda)\,\widetilde{f}(\lambda,-\lambda,\omega)\,\{\varphi_3(\omega)\,\widetilde{f}(\omega)\,\varphi_4(\omega)\}]]\,[[\varphi_2(\lambda)\,\widetilde{f}(\lambda)]] \\ & + [[\varphi_1(\lambda)\,\widetilde{f}(\lambda,-\lambda,\omega)\,\{\varphi_4(\omega)\,\widetilde{f}(\omega)\,\varphi_3(\omega)\}]]\,[[\varphi_2(\lambda)\,\widetilde{f}(\lambda)]] \\ & + [[\varphi_1(\lambda)\,\widetilde{f}(\lambda,-\lambda,\omega)\,\varphi_4(\omega)]]\,[[\varphi_2(\lambda)\,\widetilde{f}(\lambda)]]\,[[\varphi_3(\omega)\,\widetilde{f}(\omega)]] \\ & + [[\{\varphi_1(\lambda)\,\widetilde{f}(\lambda)\,\varphi_2(\lambda)\}\,\widetilde{f}(\lambda,-\lambda,\omega)\,\varphi_3(\omega)]]\,[[\varphi_4(\omega)\,\widetilde{f}(\omega)]] \\ & + [[\{\varphi_1(\lambda)\,\widetilde{f}(\lambda)\,\varphi_2(\lambda)\}\,\widetilde{f}(\lambda,-\lambda,\omega)\,\{\varphi_3(\omega)\,\widetilde{f}(\omega)\,\varphi_4(\omega)\}]] \\ & + [[\{\varphi_1(\lambda)\,\widetilde{f}(\lambda)\,\varphi_2(\lambda)\}\,\widetilde{f}(\lambda,-\lambda,\omega)\,\{\varphi_4(\omega)\,\widetilde{f}(\omega)\,\varphi_3(\omega)\}]] \\ & + [[\{\varphi_1(\lambda)\,\widetilde{f}(\lambda)\,\varphi_2(\lambda)\}\,\widetilde{f}(\lambda,-\lambda,\omega)\,\varphi_4(\omega)]]\,[[\varphi_3(\omega)\,\widetilde{f}(\omega)]] \\ & + [[\{\varphi_2(\lambda)\,\widetilde{f}(\lambda)\,\varphi_1(\lambda)\}\,\widetilde{f}(\lambda,-\lambda,\omega)\,\varphi_3(\omega)]]\,[[\varphi_4(\omega)\,\widetilde{f}(\omega)]] \\ & + [[\{\varphi_2(\lambda)\,\widetilde{f}(\lambda)\,\varphi_1(\lambda)\}\,\widetilde{f}(\lambda,-\lambda,\omega)\,\{\varphi_3(\omega)\,\widetilde{f}(\omega)\,\varphi_4(\omega)\}]] \\ & + [[\{\varphi_2(\lambda)\,\widetilde{f}(\lambda)\,\varphi_1(\lambda)\}\,\widetilde{f}(\lambda,-\lambda,\omega)\,\{\varphi_4(\omega)\,\widetilde{f}(\omega)\,\varphi_3(\omega)\}]] \\ & + [[\{\varphi_2(\lambda)\,\widetilde{f}(\lambda)\,\varphi_1(\lambda)\}\,\widetilde{f}(\lambda,-\lambda,\omega)\,\varphi_4(\omega)]]\,[[\varphi_3(\omega)\,\widetilde{f}(\omega)]] \\ & + [[\varphi_2(\lambda)\,\widetilde{f}(\lambda,-\lambda,\omega)\,\varphi_3(\omega)]]\,[[\varphi_1(\lambda)\,\widetilde{f}(\lambda)]]\,[[\varphi_4(\omega)\,\widetilde{f}(\omega)]] \\ & + [[\varphi_2(\lambda)\,\widetilde{f}(\lambda,-\lambda,\omega)\,\{\varphi_3(\omega)\,\widetilde{f}(\omega)\,\varphi_4(\omega)\}]]\,[[\varphi_1(\lambda)\,\widetilde{f}(\lambda)]] \\ & + [[\varphi_2(\lambda)\,\widetilde{f}(\lambda,-\lambda,\omega)\,\{\varphi_4(\omega)\,\widetilde{f}(\omega)\,\varphi_3(\omega)\}]]\,[[\varphi_1(\lambda)\,\widetilde{f}(\lambda)]] \\ & + [[\varphi_2(\lambda)\,\widetilde{f}(\lambda,-\lambda,\omega)\,\varphi_4(\omega)]]\,[[\varphi_1(\lambda)\,\widetilde{f}(\lambda)]]\,[[\varphi_3(\omega)\,\widetilde{f}(\omega)]].\end{aligned} \tag{C.4}$$

Similar calculations for the linear and mixed quadratic-linear cases yield terms involving the tri-spectrum as well, and

$$
\begin{aligned}
g_{123}(\lambda,\omega) = {} & [[\varphi_1(\lambda)\, \widetilde{f}(\lambda,-\lambda,\omega)\, \varphi_3(\omega)]]\, [[\varphi_2(\lambda)\, \widetilde{f}(\lambda)]] \\
& + [[\{\varphi_1(\lambda)\, \widetilde{f}(\lambda)\, \varphi_2(\lambda)\}\, \widetilde{f}(\lambda,-\lambda,\omega)\, \varphi_3(\omega)]] \\
& + [[\{\varphi_2(\lambda)\, \widetilde{f}(\lambda)\, \varphi_1(\lambda)\}\, \widetilde{f}(\lambda,-\lambda,\omega)\, \varphi_3(\omega)]] \\
& + [[\varphi_2(\lambda)\, \widetilde{f}(\lambda,-\lambda,\omega)\, \varphi_3(\omega)]]\, [[\varphi_1(\lambda)\, \widetilde{f}(\lambda)]].
\end{aligned}
\tag{C.5}
$$

□

Proof of Theorem A.1. We apply the method of cumulants to the functionals. The key result that is needed is that higher-order cumulants of the linear and quadratic functionals tend to zero. We demonstrate this through quadratic functionals only, the other cases being similar. Suppose we have an m-fold cumulant of normalized quadratic functionals, each of which takes the form $\sqrt{T}\, \widehat{Q}_{\varphi,\psi}(I,I)$. Generalizing the arguments of the proof of Proposition A.2, and using Assumption 1, we obtain a factor of order of T^{-2m} from the $2m$ periodograms, T^{-m} from the discretization of the m integrals, and $T^{m/2}$ from the normalizations, for an overall $T^{3m-m/2}$ in the denominator. For the cumulant, we must now consider a table with m rows and 4 columns. We claim that the largest possible order of the sums (over all frequencies) of such cumulants is $O(T^{2m+1})$. In the diagonal case, where the m sums really collapse to a single sum – in the manner discussed in the proof of Proposition A.2 – one obtains the highest order possible for the cumulant by taking an indecomposable partition with $2m$ sets of size 2. By (C.1), this would yield $O(T^{2m})$, which together with the single summation gives an overall $O(T^{2m+1})$. However, of the m summations we might allow some pairs to collapse to a single summation, and others may not.

Suppose we consider a pair of summations to be distinct, but all the others collapse to a single summation. Now, the partition involving $2m$ sets of size 2 will be of lower order, as any 2-element sets straddling distinct rows corresponding to the two frequencies of the paired summation will no longer satisfy $\lambda_{j_1} + \lambda_{j_2} = 0$; there is at least one such 2-element set (because the partition is indecomposable), so the cumulant order drops to T^{2m-1}. Moreover, by combining two 2-element sets into a 4-element set, we can obtain a factor of T if (C.3) is satisfied, although there will be at most $2m-1$ sets in such a partition – ultimately yielding $O(T^{2m-1})$. Because the total number of summations is two, we would obtain an overall $O(T^{2m+1})$ for this case.

Proceeding by the same argument, distinct summations over frequencies add an overall order of T but also limit the types of partitions that will yield cumulant terms of order T; we can always merge two 2-element sets into a 4-element set when determining the relevant partitions, in moving from a collapsed summation to a double summation – but this will decrease the size of the partition by one. This is compensated by having an additional summation – so the largest possible order is T^{2m+1}. Pairing this with the denominator $T^{3m-m/2}$, the m-fold cumulant is $O(T^{1-m/2})$, which tends to zero when $m > 2$.

Clearly, if we are talking about the m-fold cumulants of the same quadratic functional, this establishes that it is asymptotically normal. But because the discussion pertains to m-fold cumulants of any collection of quadratic functionals, joint asymptotic normality also follows from the Cramer-Wold device. Extending these arguments to joint relations with linear functionals completes the proof. □

Proof of Proposition 1. Because

$$
[[\Sigma\, \Gamma_\Phi(0)]] = \sigma'\, A(\phi)\, \sigma \quad \text{and} \quad [[\Sigma\, \Gamma_\Psi(0)]] = \sigma'\, B(\phi)\, \mathrm{vec}[\Omega],
$$

the objective function for the PTVs can be rewritten as

$$\sigma'\,A(\phi)\,\sigma - 2\,\sigma'\,B(\phi)\,\mathrm{vec}[\Omega].$$

Setting the gradient with respect to σ equal to zero and noting that $A(\phi)$ is a symmetric matrix yields the solution $A(\phi)^{-1}\,B(\phi)\,\mathrm{vec}[\Omega]$ for σ in terms of ϕ, where the symmetrization of $A(\phi)$ comes out of the vector calculus. Plugging this back in for σ now yields the concentrated objective function for ϕ; accounting for the minus sign, the minimization is turned to a maximization in (3). Once $\widetilde{\phi}$ is obtained, substitution into the formula for σ yields its PTV as well (4). □

Proof of Proposition 2. We can write the criterion function as

$$\begin{aligned}|I - f_\theta|^2 &= \sum_{i,k=1}^{K} G_{ik}[[\Theta_i\,\Theta_k']] - 2\sum_{i=1}^{K}[[\Theta_i\,\langle I g_i\rangle_0]] + \langle[[I^2]]\rangle_0 \\ &= [[\Theta\,(G\otimes 1_m)\,\Theta']] - 2[[\Theta\,\langle g\otimes I\rangle_0]] + \langle[[I^2]]\rangle_0.\end{aligned}$$

Computing the gradient, we now see that the stated formula (8) for the MOM estimator is a critical point, and a minimizer, of this criterion. The same proof, with $\widetilde{f}$ in place of I, shows that the formula for the PTV is obtained by replacing $\widetilde{f}$ for I in (8). □

Proof of Proposition 3. By our assumptions we can apply Propositions A.1 and A.2. So

$$\begin{aligned}\langle[[I^2]]\rangle_0 &= \left(\widehat{Q}_{\mathrm{id,id}}(I,I) - \mathbb{E}\widehat{Q}_{\mathrm{id,id}}(I,I)\right) + \mathbb{E}\widehat{Q}_{\mathrm{id,id}}(I,I) \\ &= \left(\widehat{Q}_{\mathrm{id,id}}(I,I) - \mathbb{E}\widehat{Q}_{\mathrm{id,id}}(I,I)\right) + Q_{\mathrm{id,id}}(\widetilde{f},\widetilde{f}) + \langle[[\widetilde{f}]]^2\rangle_0 + O_P(T^{-1}),\end{aligned}$$

where Proposition A.1 has been used in the second line. By Proposition A.2, $\mathrm{Var}\widehat{Q}_{\mathrm{id,id}}(I,I) = O(T^{-1})$, so that by the Chebyshev Inequality

$$\widehat{Q}_{\mathrm{id,id}}(I,I) - \mathbb{E}\widehat{Q}_{\mathrm{id,id}}(I,I) \xrightarrow{P} 0.$$

Therefore, $\langle[[I^2]]\rangle_0 \xrightarrow{P} \langle[[\widetilde{f}^2]]\rangle_0 + \langle[[\widetilde{f}]]^2\rangle_0$. Similarly, again by Proposition A.1

$$\langle[[I\,f_\theta]]\rangle_0 = \left(\widehat{L}_{f_\theta}(I) - \mathbb{E}\widehat{L}_{f_\theta}(I)\right) + L_{f_\theta}(\widetilde{f}) + O(T^{-1}).$$

Again by Proposition A.2, $\mathrm{Var}\widehat{L}_{f_\theta}(I) = O(T^{-1})$, so that by the Chebyshev Inequality

$$\widehat{L}_{f_\theta}(I) - \mathbb{E}\widehat{L}_{f_\theta}(I) \xrightarrow{P} 0.$$

Putting these results together, we obtain

$$|I - f_\theta|^2 = \langle[[I^2]]\rangle_0 - 2\,\langle[[I\,f_\theta]]\rangle_0 + |f_\theta|^2 \xrightarrow{P} \langle[[\widetilde{f}^2]]\rangle_0 + \langle[[\widetilde{f}]]^2\rangle_0 - 2\,\langle[[\widetilde{f}\,f_\theta]]\rangle_0 + |f_\theta|^2.$$

□

Proof of Theorem 1. Write $\widehat{F}(\theta)$ for $\mathrm{FD}(f_\theta, I)$, and $\widetilde{F}(\theta)$ for $\mathrm{FD}(f_\theta, \widetilde{f})$. First,

$$0 = \nabla\widehat{F}(\widehat{\theta}) = \nabla\widehat{F}(\widetilde{\theta}) + \nabla\nabla'\widehat{F}(\widetilde{\theta})\,(\widehat{\theta} - \widetilde{\theta}) + R_T$$

by a Taylor series expansion of $\nabla\widehat{F}(\theta)$ about $\theta=\widehat{\theta}$, where R_T depends on $\widehat{\theta}-\widetilde{\theta}$ quadratically. We proceed to compute the gradient and Hessian of $\widehat{F}(\theta)$. To make the notation less cumbersome, we abbreviate the partial derivative operator $\frac{\partial}{\partial\theta_j}$ by ∂_{θ_j}. Hence we find that

$$\begin{aligned}\partial_{\theta_j}\widehat{F}(\theta) &= -2\,[[\langle\partial_{\theta_j} f_\theta\,(I-f_\theta)\rangle_0]]\\ \partial_{\theta_j}\partial_{\theta_k}\widehat{F}(\theta) &= -2\,[[\langle\partial_{\theta_j}\partial_{\theta_k} f_\theta\,(I-f_\theta)\rangle_0]]+2\,[[\langle\partial_{\theta_j} f_\theta\,\partial_{\theta_k} f_\theta\rangle_0]],\end{aligned}$$

and the first term has asymptotic mean $-2\,[[\langle\partial_{\theta_j} f_\theta\,(\widetilde{f}-f_\theta)\rangle_0]]$. This we recognize as the derivative of $\widetilde{F}(\theta)$, which is zero at $\theta=\widetilde{\theta}$. Therefore

$$\nabla\widehat{F}(\widetilde{\theta})=-2\,[[\langle\nabla f_\theta\,(I-\widetilde{f})\rangle_0]],$$

where the trace operator does not act on the gradient. An application of Theorem A.1, in conjunction with our other assumptions yields

$$\sqrt{T}\,\nabla\widehat{F}(\widetilde{\theta})\overset{\mathcal{L}}{\Longrightarrow}\mathcal{N}(0,4\,V_{\widetilde{\theta}}). \tag{C.6}$$

Next, the Hessian of $\widehat{F}(\theta)$ converges in probability to $2\,M_\theta$, because the first term is actually $O_P(T^{-1/2})$ (by Proposition A.1). If this matrix is invertible at $\widetilde{\theta}$, we conclude that $\widehat{\theta}-\widetilde{\theta}=O_P(T^{-1/2})$ and further that

$$\widehat{\theta}-\widetilde{\theta}=\frac{1}{2}\,M_{\widetilde{\theta}}^{-1}\,\langle[[\nabla f_\theta\,(I-\widetilde{f})]]\rangle_0+o_P(T^{-1/2}).$$

So using (C.6), the theorem is proved. $\square$

To prove Theorem 2, we state a preliminary result.

Lemma C.1. *Assume that $\{x_t\}$ is strictly stationary with spectral density $\widetilde{f}$ and satisfies Assumption 1(k) for $k\geq 2$. For a possibly non-square matrix A that is a function of λ,*

$$\sqrt{T}\,\langle A\,(vec(I-\widetilde{f}))\rangle_0\overset{\mathcal{L}}{\Longrightarrow}\mathcal{N}\left(0,2\,\langle A\,(\widetilde{f}'\otimes\widetilde{f})\,A'\rangle_0+\langle\langle A(\lambda)\widetilde{F}(\lambda,-\lambda,\omega)A(\omega)'\rangle_0\rangle_0\right).$$

Proof of Lemma C.1. To prove the Lemma, for each ℓ let α_ℓ be the matrix such that $\mathrm{vec}(\alpha_\ell')=\{e_\ell A\}'$. Then for any f

$$A\,\mathrm{vec}(f)=\left\{\begin{array}{c}e_1'A\,\mathrm{vec}(f)\\ e_2'A\,\mathrm{vec}(f)\\ \vdots\end{array}\right\}=\left\{\begin{array}{c}\mathrm{vec}(\alpha_1')'\mathrm{vec}(f)\\ \mathrm{vec}(\alpha_2')'\mathrm{vec}(f)\\ \vdots\end{array}\right\}=\left\{\begin{array}{c}[[\alpha_1 f]]\\ [[\alpha_2 f]]\\ \vdots\end{array}\right\}.$$

Now by Theorem A.1 and Proposition A.2 we have the joint CLT

$$\begin{aligned}\sqrt{T}\,\langle A\,(\mathrm{vec}(I-\widetilde{f}))\rangle_0 &= \sqrt{T}\left\{\begin{array}{c}\left[\left[\alpha_1(I-\widetilde{f})\right]\right]\\ \left[\left[\alpha_2(I-\widetilde{f})\right]\right]\\ \vdots\end{array}\right\}\\ &\overset{\mathcal{L}}{\Longrightarrow}\mathcal{N}\left(0,\left\{2\langle[[\alpha_j f_{\widetilde{\theta}}\alpha_k' f_{\widetilde{\theta}}]]\rangle_0+\langle\langle[[\alpha_j(\lambda)\widetilde{f}(\lambda,-\lambda,\omega)\alpha_k(\omega)]]\rangle_0\rangle_0\right\}_{j,k}\right).\end{aligned}$$

It can be shown using algebraic identities that

$$[[\alpha_j\widetilde{f}\alpha_k'\widetilde{f}]]=\mathrm{vec}(\alpha_j')'\,\mathrm{vec}(\widetilde{f}\alpha_k'\widetilde{f})=e_j'A\,\widetilde{f}'\otimes\widetilde{f}\,\mathrm{vec}(\alpha_k')=e_j'A\,\widetilde{f}'\otimes\widetilde{f}\,A'e_k,$$

and hence that $\langle [[\alpha_j \widetilde{f} \alpha_k' \widetilde{f}]] \rangle_0$ is the jkth entry of $A \langle \widetilde{f}' \otimes \widetilde{f} \rangle_0 A'$. Note that the transpose on the first appearance of $\widetilde{f}$ guarantees that this matrix is symmetric. Furthermore,

$$
\begin{aligned}
[[\alpha_j(\lambda) \widetilde{f}(\lambda, -\lambda, \omega) \alpha_k(\omega)]] &= \sum_{\ell,p,r,s} \alpha_j^{(\ell p)}(\lambda)\, \widetilde{f}_{p\ell sr}(\lambda, -\lambda, \omega)\, \alpha_k^{(rs)}(\omega) \\
&= \sum_{r,s} [[\alpha_j(\lambda)\, \widetilde{f}_{\cdot\cdot sr}(\lambda, -\lambda, \omega)]]\, \alpha_k^{(rs)}(\omega) \\
&= \sum_{r,s} \mathrm{vec}(\alpha_j'(\lambda))' \mathrm{vec}\left(\widetilde{f}_{\cdot\cdot sr}(\lambda, -\lambda, \omega)]]\, \alpha_k^{(rs)}(\omega) \right) \\
&= e_j' A(\lambda) \left\{ [[\widetilde{f}_{p\ell\cdot\cdot}(\lambda, -\lambda, \omega) \alpha_k(\omega)]] \right\}_{p\ell} \\
&= e_j' A(\lambda)\, \widetilde{F}(\lambda, -\lambda, \omega)\, A(\omega)' e_k,
\end{aligned}
$$

where the ordering over p, ℓ in the second to last equality corresponds to the vec operator, i.e., these are the indices of the C matrix given in the definition of $\widetilde{F}$. The final equality follows from another application of the trace identity, and we obtain the matrix $\widetilde{F}$ corresponding to the 4-tensor $\widetilde{f}_{p\ell sr}$. The result follows. □

Proof of Theorem 2. The result follows from Lemma C.1 upon writing

$$
\theta = \mathrm{vec}\, \Theta = \left\{ \begin{array}{c} \langle (g' G^{-1} e_1)\, \mathrm{vec}(f) \rangle_0 \\ \langle (g' G^{-1} e_2)\, \mathrm{vec}(f) \rangle_0 \\ \vdots \end{array} \right\} = \langle A\, \mathrm{vec}(f) \rangle_0,
$$

where $A = G^{-1} g \otimes 1_{m^2}$. □

Proof of Proposition 4. Noting that $|\widetilde{J}|^2 = [[\langle \widetilde{J}^2 \rangle_0]] = \widehat{Q}_{\underline{f}_{\widetilde{\theta}}^{-1}, \underline{f}_{\widetilde{\theta}}^{-1}}(I, I)$, it follows from Proposition A.1 that (13) holds; moreover, $|\widetilde{J}|^2 \xrightarrow{P} |\widetilde{g}|^2 + \langle [[\widetilde{g}]]^2 \rangle_0$ by Proposition A.2. Because $\widehat{\theta} \xrightarrow{P} \widetilde{\theta}$, it follows from Slutsky's Theorem and continuity of Ψ_θ that $\underline{f}_{\widehat{\theta}}^{-1}(\lambda) \xrightarrow{P} \underline{f}_{\widetilde{\theta}}^{-1}(\lambda)$ for each λ. Hence $\widehat{g}(\lambda) \xrightarrow{P} \widetilde{g}(\lambda)$ for all λ, and similarly it can be shown that $\widehat{J}(\lambda) - \widetilde{J}(\lambda) = o_P(1)$. Moreover, by Propositions A.1 and A.2 we obtain $[[\langle c\, (\widetilde{J} - \widetilde{g}) \rangle_0]] \xrightarrow{P} 0$ for a matrix function c. As a result, we obtain $[[\langle \widehat{J} \rangle_0]] \xrightarrow{P} [[\langle \widetilde{g} \rangle_0]]$. Next, write

$$
\begin{aligned}
|\widehat{J}|^2 &= |\widetilde{J}|^2 + [[\langle (\widehat{J} - \widetilde{J})(\widehat{J} + \widetilde{J}) \rangle_0]] \\
[[\langle \widehat{J} \rangle_0^{\,2}]] &= [[\langle \widetilde{g} \rangle_0^{\,2}]] + [[\langle \widehat{J} - \widetilde{g} \rangle_0 \langle \widehat{J} + \widetilde{g} \rangle_0]].
\end{aligned}
$$

Using the above results, we can then obtain

$$
\begin{aligned}
[[\langle \widehat{J} \rangle_0^{\,2}]] &\xrightarrow{P} [[\langle \widetilde{g} \rangle_0^{\,2}]] \\
|\widehat{J}|^2 &\xrightarrow{P} |\widetilde{g}|^2 + \langle [[\widetilde{g}]]^2 \rangle_0.
\end{aligned}
$$

as $T \to \infty$. Assembling these convergences yields the state result. □

Proof of Theorem 3. We first develop the CLT for the parameter estimates, which are the minimizers of the log determinant of the forecast error variance matrix $\widehat{\Sigma}(\theta)$. Equivalently, we can minimize $[[\widehat{\Sigma}(\theta)]]$, or $[[\langle \underline{f}_\theta^{-1} I \rangle_0]]$. Then by standard Taylor series arguments (see McElroy (2016) for the univariate case)

$$
\widehat{\theta} - \widetilde{\theta} = -H(\widetilde{\theta})^{-1} [[\langle \nabla \underline{f}_{\widetilde{\theta}}^{-1} (I - \widetilde{f}) \rangle_0]] + o_P(T^{-1/2}).
$$

The gradient operator applies to $\underline{f}_{\theta}^{-1}$, using the notational conventions discussed in the paper. Therefore, up to terms $o_P(T^{-1/2})$ the parameter error is

$$-H(\widetilde{\theta})^{-1}\left(L_{\nabla \underline{f}_{\widetilde{\theta}}^{-1}}(I) - L_{\nabla \underline{f}_{\widetilde{\theta}}^{-1}}(\widetilde{f})\right) = -H(\widetilde{\theta})^{-1}\,\overline{L}_{\nabla \underline{f}_{\widetilde{\theta}}^{-1}}(I).$$

We can use this result to analyze the difference between $\widehat{J}$ and $\widetilde{J}$:

$$\begin{aligned}
\widehat{J} - \widetilde{J} &= \left(\Psi_{\widehat{\theta}}(e^{-i\cdot})^{-1} - \Psi_{\widetilde{\theta}}(e^{-i\cdot})^{-1}\right) I \Psi_{\widehat{\theta}}(e^{i\cdot})^{-1\prime} + \Psi_{\widetilde{\theta}}(e^{-i\cdot})^{-1} I \left(\Psi_{\widehat{\theta}}(e^{i\cdot})^{-1\prime} - \Psi_{\widetilde{\theta}}(e^{i\cdot})^{-1\prime}\right) \\
&= o_P(T^{-1/2}) + (\widehat{\theta} - \widetilde{\theta})^{\prime} \nabla \Psi_{\widetilde{\theta}}(e^{-i\cdot})^{-1} \widetilde{f} \Psi_{\widetilde{\theta}}(e^{i\cdot})^{-1\prime} + \Psi_{\widetilde{\theta}}(e^{-i\cdot})^{-1} \widetilde{f} \nabla^{\prime} \Psi_{\widetilde{\theta}}(e^{i\cdot})^{-1\prime} (\widehat{\theta} - \widetilde{\theta}) \\
&= o_P(T^{-1/2}) - \overline{L}_{\nabla^{\prime} \underline{f}_{\widetilde{\theta}}^{-1}}(I)\, H(\widetilde{\theta})^{-1} \nabla \Psi_{\widetilde{\theta}}(e^{-i\cdot})^{-1} \widetilde{f} \Psi_{\widetilde{\theta}}(e^{i\cdot})^{-1\prime} \\
&\quad - \Psi_{\widetilde{\theta}}(e^{-i\cdot})^{-1} \widetilde{f} \nabla^{\prime} \Psi_{\widetilde{\theta}}(e^{i\cdot})^{-1\prime} H(\widetilde{\theta})^{-1}\, \overline{L}_{\nabla \underline{f}_{\widetilde{\theta}}^{-1}}(I).
\end{aligned}$$

Also, $\widehat{J} - \widetilde{g} = (\widehat{J} - \widetilde{J}) + (\widetilde{J} - \widetilde{g})$ and

$$\widetilde{J} - \widetilde{g} = \Psi_{\widetilde{\theta}}(e^{-i\cdot})^{-1} (I - \widetilde{f}) \Psi_{\widetilde{\theta}}(e^{i\cdot})^{-1\prime}.$$

With these results, we can analyze $\widehat{\text{Eval}}(\widehat{J}) - \text{Eval}(\widetilde{g})$ by breaking the difference into five terms:

$$\begin{aligned}
\widehat{\text{Eval}}(\widehat{J}) - \text{Eval}(\widetilde{g}) = &\left(|\widetilde{J}|^2 - \mathbb{E}[|\widetilde{J}|^2]\right) \\
&+ \left(\mathbb{E}[|\widetilde{J}|^2] - |\widetilde{g}|^2 - \langle [[\widetilde{g}]]^2 \rangle_0\right) \\
&- [[\langle \widehat{J} - \widetilde{g} \rangle_0 \langle \widehat{J} + \widetilde{g} \rangle_0]] \\
&- \langle [[\widehat{J} - \widetilde{g}]] \rangle_0 \langle [[\widehat{J} + \widetilde{g}]] \rangle_0 \\
&+ [[\langle (\widehat{J} - \widetilde{g})(\widehat{J} + \widetilde{g}) \rangle_0]].
\end{aligned}$$

The first term equals $\widehat{Q}_{\underline{f}_{\widetilde{\theta}}^{-1}, \underline{f}_{\widetilde{\theta}}^{-1}}(I, I)$ centered by its mean, and hence can be written as $\overline{Q}_{\underline{f}_{\widetilde{\theta}}^{-1}, \underline{f}_{\widetilde{\theta}}^{-1}}(I, I)$. The second term is $O(T^{-1})$ by (13). The third, fourth, and fifth terms, up to $o_P(T^{-1/2})$ summands, are

$$\begin{aligned}
&- 2[[\langle \widehat{J} - \widetilde{g} \rangle_0 \langle \widetilde{g} \rangle_0]] \\
&- 2\langle [[\langle \widehat{J} - \widetilde{g}]] \rangle_0 \langle [[\widetilde{g}]] \rangle_0 \\
&+ 2[[\langle (\widehat{J} - \widetilde{J}) \widetilde{g} \rangle_0]].
\end{aligned}$$

Terms three and four involve parameter estimation error, whereas the fifth term, up to $o_P(T^{-1/2})$, is

$$\begin{aligned}
&- 2\overline{L}_{\nabla^{\prime} \underline{f}_{\widetilde{\theta}}^{-1}}(I)\, H(\widetilde{\theta})^{-1} [[\langle \nabla \Psi_{\widetilde{\theta}}(e^{-i\cdot})^{-1} \widetilde{f} \Psi_{\widetilde{\theta}}(e^{i\cdot})^{-1\prime} \widetilde{g} \rangle_0]] \\
&- 2[[\langle \widetilde{g} \Psi_{\widetilde{\theta}}(e^{-i\cdot})^{-1} \widetilde{f} \nabla^{\prime} \Psi_{\widetilde{\theta}}(e^{i\cdot})^{-1\prime} \rangle_0]] H(\widetilde{\theta})^{-1}\, \overline{L}_{\nabla \underline{f}_{\widetilde{\theta}}^{-1}}(I).
\end{aligned}$$

Both summands are the same, because we can apply a transpose within the trace, and change $\lambda \mapsto -\lambda$, which amounts to applying a conjugate transpose. Term three can be expanded, up to terms $o_P(T^{-1/2})$, as

$$\begin{aligned}
&- 2[[\langle \widehat{J} - \widetilde{J} \rangle_0 \langle \widetilde{g} \rangle_0]] - 2[[\langle \widetilde{J} - \widetilde{g} \rangle_0 \langle \widetilde{g} \rangle_0]] \\
&= 2\overline{L}_{\nabla^{\prime} \underline{f}_{\widetilde{\theta}}^{-1}}(I)\, H(\widetilde{\theta})^{-1} [[\langle \nabla \Psi_{\widetilde{\theta}}(e^{-i\cdot})^{-1} \widetilde{f} \Psi_{\widetilde{\theta}}(e^{i\cdot})^{-1\prime} \rangle_0 \langle \widetilde{g} \rangle_0]] \\
&+ 2[[\langle \widetilde{g} \rangle_0 \langle \Psi_{\widetilde{\theta}}(e^{-i\cdot})^{-1} \widetilde{f} \nabla^{\prime} \Psi_{\widetilde{\theta}}(e^{i\cdot})^{-1\prime} \rangle_0]] H(\widetilde{\theta})^{-1}\, \overline{L}_{\nabla \underline{f}_{\widetilde{\theta}}^{-1}}(I) \\
&- 2\overline{L}_b(I),
\end{aligned}$$

which uses

$$[[\langle \widetilde{J}-\widetilde{g}\rangle_0\langle \widetilde{g}\rangle_0]] = [[\langle \Psi_{\widetilde{\theta}}(e^{-i\cdot})^{-1}(I-\widetilde{f})\Psi_{\widetilde{\theta}}(e^{i\cdot})^{-1\prime}\rangle_0\langle \widetilde{g}\rangle_0]] = [[\langle b\,(I-\widetilde{f})\rangle_0]].$$

Term four is similar:

$$\begin{aligned}
&-2\langle[[\widehat{J}-\widetilde{J}]]\rangle_0\langle[[\widetilde{g}]]\rangle_0 - 2\langle[[\widetilde{J}-\widetilde{g}]]\rangle_0\langle[[\widetilde{g}]]\rangle_0\\
&= 2\overline{L}_{\nabla'\underline{f}_{\widetilde{\theta}}^{-1}}(I)\,H(\widetilde{\theta})^{-1}[[\langle \nabla\Psi_{\widetilde{\theta}}(e^{-i\cdot})^{-1}\widetilde{f}\Psi_{\widetilde{\theta}}(e^{i\cdot})^{-1\prime}\rangle_0]][[\langle\widetilde{g}\rangle_0]]\\
&+2[[\langle\widetilde{g}\rangle_0]][[\langle \Psi_{\widetilde{\theta}}(e^{-i\cdot})^{-1}\widetilde{f}\nabla'\Psi_{\widetilde{\theta}}(e^{i\cdot})^{-1\prime}\rangle_0]]H(\widetilde{\theta})^{-1}\,\overline{L}_{\nabla\underline{f}_{\widetilde{\theta}}^{-1}}(I)\\
&-2[[\langle\widetilde{g}\rangle_0]]\,\overline{L}_{\underline{f}_{\widetilde{\theta}}^{-1}}(I).
\end{aligned}$$

Combining and simplifying, we obtain

$$\begin{aligned}
\widehat{\mathrm{Eval}}(\widehat{J})-\mathrm{Eval}(\widetilde{g}) &= O(T^{-1})+o_P(T^{-1/2})+\overline{Q}_{\underline{f}_{\widetilde{\theta}}^{-1},\underline{f}_{\widetilde{\theta}}^{-1}}(I,I)\\
&+4[[\langle\widetilde{g}\rangle_0\langle \Psi_{\widetilde{\theta}}(e^{-i\cdot})^{-1}\widetilde{f}\nabla'\Psi_{\widetilde{\theta}}(e^{i\cdot})^{-1\prime}\rangle_0]]H(\widetilde{\theta})^{-1}\,\overline{L}_{\nabla\underline{f}_{\widetilde{\theta}}^{-1}}(I)-2\overline{L}_b(I)\\
&+4[[\langle\widetilde{g}\rangle_0]][[\langle \Psi_{\widetilde{\theta}}(e^{-i\cdot})^{-1}\widetilde{f}\nabla'\Psi_{\widetilde{\theta}}(e^{i\cdot})^{-1\prime}\rangle_0]]H(\widetilde{\theta})^{-1}\,\overline{L}_{\nabla\underline{f}_{\widetilde{\theta}}^{-1}}(I)-2[[\langle\widetilde{g}\rangle_0]]\,\overline{L}_{\underline{f}_{\widetilde{\theta}}^{-1}}(I)\\
&-4[[\langle\widetilde{g}\Psi_{\widetilde{\theta}}(e^{-i\cdot})^{-1}\widetilde{f}\nabla'\Psi_{\widetilde{\theta}}(e^{i\cdot})^{-1\prime}\rangle_0]]H(\widetilde{\theta})^{-1}\,\overline{L}_{\nabla\underline{f}_{\widetilde{\theta}}^{-1}}(I).
\end{aligned}$$

By regrouping the summands, we can re-express this (up to terms that are $O(T^{-1})$ and $o_P(T^{-1/2})$) as

$$v'\,[\overline{L}_b(I),\overline{L}_{\underline{f}_{\widetilde{\theta}}^{-1}}(I),\overline{Q}_{\underline{f}_{\widetilde{\theta}}^{-1},\underline{f}_{\widetilde{\theta}}^{-1}}(I,I),\overline{L}_{\nabla\underline{f}_{\widetilde{\theta}}^{-1}}(I)]'.$$

Now we apply Theorem A.1 to the four-component vector of functionals, multiplying by $T^{1/2}$; this concludes the proof. □

Proof of Corollary 1. Because the model is correctly specified, $\widetilde{f}(\lambda)=\Psi_{\widetilde{\theta}}(e^{-i\lambda})\Sigma\Psi_{\widetilde{\theta}}(e^{i\lambda})'$ and $\widetilde{g}=\Sigma$. Therefore, in the fourth component of the vector v the first two terms cancel out. The third term is

$$4[[\Sigma]][[\Sigma\langle\Psi_{\widetilde{\theta}}(e^{i\cdot})'\nabla'\Psi_{\widetilde{\theta}}(e^{i\cdot})^{-1\prime}\rangle_0]],$$

and this equals zero because $\Psi_{\widetilde{\theta}}(x)'\nabla'\Psi_{\widetilde{\theta}}(x)^{-1\prime}$ as a Laurent series in x only involves positive powers. (This uses the assumption that $\Psi_\theta(0)=1_m$ does not depend on θ.) Hence $v'=[-2,-2[[\Sigma]],1,0]$, and there is no contribution from parameter uncertainty! Next, taking $\widetilde{g}=\Sigma$ in the variance expressions of Theorem 3 yields

$$\begin{aligned}
V_{b|b} &= 2\,[[\Sigma^4]]+\frac{1}{(2\pi)^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}[[b(\lambda)\,\widetilde{f}(\lambda,-\lambda,\omega)\,b(\omega)]]\,d\lambda\,d\omega,\\
V_{\underline{f}_{\widetilde{\theta}}^{-1}|b} &= 2\,[[\Sigma^3]]+\frac{1}{(2\pi)^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}[[\underline{f}_{\widetilde{\theta}}^{-1}(\lambda)\,\widetilde{f}(\lambda,-\lambda,\omega)\,b(\omega)]]\,d\lambda\,d\omega,\\
V_{\underline{f}_{\widetilde{\theta}}^{-1}|\underline{f}_{\widetilde{\theta}}^{-1}} &= 2\,[[\Sigma^2]]+\frac{1}{(2\pi)^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}[[\underline{f}_{\widetilde{\theta}}^{-1}(\lambda)\widetilde{f}(\lambda,-\lambda,\omega)\underline{f}_{\widetilde{\theta}}^{-1}(\omega)]]\,d\lambda\,d\omega,
\end{aligned}$$

$$
\begin{aligned}
V_{\underline{f}_{\widetilde{\theta}}^{-1},\underline{f}_{\widetilde{\theta}}^{-1}|b} &= 4\,[[\Sigma]]\,[[\Sigma^3]] + 4\,[[\Sigma^4]] \\
&+ \frac{2}{(2\pi)^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\Big([[\Sigma]][[\underline{f}_{\widetilde{\theta}}^{-1}(\lambda)\widetilde{f}(\lambda,-\lambda,\omega)b(\omega)]] + [[b(\lambda)\widetilde{f}(\lambda,-\lambda,\omega)b(\omega)]]\Big)\,d\lambda\,d\omega, \\
V_{\underline{f}_{\widetilde{\theta}}^{-1},\underline{f}_{\widetilde{\theta}}^{-1}|\underline{f}_{\widetilde{\theta}}^{-1}} &= 4\,[[\Sigma]]\,[[\Sigma^2]] + 4\,[[\Sigma^3]] \\
&+ \frac{2}{(2\pi)^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\Big([[\Sigma]][[\underline{f}_{\widetilde{\theta}}^{-1}(\lambda)\widetilde{f}(\lambda,-\lambda,\omega)\underline{f}_{\widetilde{\theta}}^{-1}(\omega)]] + [[b(\lambda)\widetilde{f}(\lambda,-\lambda,\omega)\underline{f}_{\widetilde{\theta}}^{-1}(\omega)]]\Big)\,d\lambda\,d\omega, \\
V_{\underline{f}_{\widetilde{\theta}}^{-1},\underline{f}_{\widetilde{\theta}}^{-1}|\underline{f}_{\widetilde{\theta}}^{-1},\underline{f}_{\widetilde{\theta}}^{-1}} &= 8\,[[\Sigma]]^2\,[[\Sigma^2]] + 16\,[[\Sigma]]\,[[\Sigma^3]] + 12\,[[\Sigma^4]] + 4\,[[\Sigma^2]]^2 \\
&+ \frac{4}{(2\pi)^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\Big([[\underline{f}_{\widetilde{\theta}}^{-1}(\lambda)\widetilde{f}(\lambda,-\lambda,\omega)\underline{f}_{\widetilde{\theta}}^{-1}(\omega)]][[\Sigma]]^2 + [[\underline{f}_{\widetilde{\theta}}^{-1}(\lambda)\widetilde{f}(\lambda,-\lambda,\omega)b(\omega)]][[\Sigma]] \\
&+ [[b(\lambda)\widetilde{f}(\lambda,-\lambda,\omega)\underline{f}_{\widetilde{\theta}}^{-1}(\omega)]][[\Sigma]] + [[b(\lambda)\widetilde{f}(\lambda,-\lambda,\omega)b(\omega)]]\Big)\,d\lambda\,d\omega.
\end{aligned}
$$

Next, we claim that the double integral of $[[A(\lambda)\,\widetilde{f}(\lambda,-\lambda,\omega)\,B(\omega)]]$ is equal to that of $[[B(\omega)\,\widetilde{f}(\omega,-\omega,\lambda)\,A(\lambda)]]$, for any tri-spectrum $\widetilde{f}$ and matrices A and B. This is proved from the definition of the double bracket, and utilizing the third property of the tri-spectrum derived in the proof of Proposition A.2, namely that $f_{\ell ksr}(\lambda,-\lambda,\omega) = f_{sr\ell k}(\omega,-\omega,\lambda)$; two changes of variable in the sums allow us to swap pairs of indices, and thereby interchange the positions of A and B. Therefore it can be show that

$$
v'Wv = 4V_{b|b} + 4[[\Sigma]]^2 V_{\underline{f}_{\widetilde{\theta}}^{-1}|\underline{f}_{\widetilde{\theta}}^{-1}} + 8[[\Sigma]]V_{\underline{f}_{\widetilde{\theta}}^{-1}|b} + V_{\underline{f}_{\widetilde{\theta}}^{-1},\underline{f}_{\widetilde{\theta}}^{-1}|\underline{f}_{\widetilde{\theta}}^{-1},\underline{f}_{\widetilde{\theta}}^{-1}} - 4V_{\underline{f}_{\widetilde{\theta}}^{-1},\underline{f}_{\widetilde{\theta}}^{-1}|b} - 4[[\Sigma]]V_{\underline{f}_{\widetilde{\theta}}^{-1},\underline{f}_{\widetilde{\theta}}^{-1}|\underline{f}_{\widetilde{\theta}}^{-1}},
$$

after some cancellations, yields the stated limiting variance. □

Proof of Proposition B.1. First, $|J|^2 = [[\langle J^2\rangle_0]] = \widehat{Q}_{\mathrm{id},\mathrm{id}}(J,J)$, and so it follows from Proposition A.1 that (B.1) holds. Second, $|J|^2 \xrightarrow{P} |\widetilde{g}|^2 + \langle[[\widetilde{g}]]^2\rangle_0$ and $[[\langle J\rangle_0]] \xrightarrow{P} [[\langle\widetilde{g}\rangle_0]]$ by Proposition A.2. Next,

$$
[[{\langle J\rangle_0}^2]] = [[{\langle\widetilde{g}\rangle_0}^2]] + [[\langle J-\widetilde{g}\rangle_0\langle J+\widetilde{g}\rangle_0]],
$$

from which it follows again from Proposition A.2 that $[[{\langle J\rangle_0}^2]] \xrightarrow{P} [[{\langle\widetilde{g}\rangle_0}^2]]$ as $T\to\infty$. Assembling these convergences yields the state result. □

Proof of Theorem B.1. Along the lines of the proof of Theorem 3, we begin by expanding $\widehat{\mathrm{Eval}}(J) - \mathrm{Eval}(\widetilde{g})$ into four terms:

$$
\begin{aligned}
\widehat{\mathrm{Eval}}(J) - \mathrm{Eval}(\widetilde{g}) &= |J|^2 - |\widetilde{g}|^2 - \langle[[\widetilde{g}]]^2\rangle_0 \\
&= \Big(|J|^2 - \mathbb{E}[|J|^2]\Big) + \Big(\mathbb{E}[|J|^2] - |\widetilde{g}|^2 - \langle[[\widetilde{g}]]^2\rangle_0\Big) \\
&- [[\langle J-\widetilde{g}\rangle_0\,\langle J+\widetilde{g}\rangle_0]] - \langle[[J-\widetilde{g}]]\rangle_0\,\langle[[J+\widetilde{g}]]\rangle_0.
\end{aligned}
$$

The second term is $O(T^{-1})$ by (B.1). The third and fourth terms are analyzed along the lines given in the proof of Theorem 3, only there is no contribution from parameter estimation. So we obtain

$$
\widehat{\mathrm{Eval}}(J) - \mathrm{Eval}(\widetilde{g}) = \Big(|J|^2 - \mathbb{E}[|J|^2]\Big) - 2[[\langle\widetilde{g}\rangle_0\,\langle J-\mathbb{E}[J]\rangle_0]] - 2\,[[\langle\widetilde{g}\rangle_0]]\cdot[[\langle J-\mathbb{E}[J]\rangle_0]] + O(T^{-1}) + o_P(T^{-1/2}).
$$

In the notation of Theorem A.1, we are studying a linear combination of two linear and one quadratic functional, each centered by its expectation. Note that by passing $\langle\widetilde{g}\rangle_0$ into the inner integral, we obtain

$[[\langle\widetilde{g}\rangle_0 \langle J\rangle_0]] = L_{\langle\widetilde{g}\rangle_0}(J)$. Therefore

$$\begin{aligned}
&\sqrt{T}\,\left(\widehat{\mathrm{Eval}}(J) - \mathrm{Eval}(\widetilde{g})\right)\\
&= \sqrt{T}\,\left(-2\,\overline{L}_{\langle\widetilde{g}\rangle_0}(J) - 2[[\langle\widetilde{g}\rangle_0]]\,\overline{L}_{\mathrm{id}}(J) + \overline{Q}_{\mathrm{id,id}}(J,J)\right) + o_P(1)\\
&= \{-2, -2[[\langle\widetilde{g}\rangle_0]], 1\}\cdot\sqrt{T}\begin{bmatrix}\overline{L}_{\langle\widetilde{g}\rangle_0}(J)\\ \overline{L}_{\mathrm{id}}(J)\\ \overline{Q}_{\mathrm{id,id}}(J,J)\end{bmatrix} + o_P(1).
\end{aligned}$$

Applying Theorem A.1, the trivariate random vector is asymptotically normal with variance matrix W as given in the statement of the theorem, and with individual entries computed according to Proposition A.2. In particular, these are given by

$$\begin{aligned}
V_{\langle\widetilde{g}\rangle_0|\langle\widetilde{g}\rangle_0} &= 2\,\langle[[\langle\widetilde{g}\rangle_0\,\widetilde{g}\,\langle\widetilde{g}\rangle_0\,\widetilde{g}]]\rangle_0 + \frac{1}{(2\pi)^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}[[\langle\widetilde{g}\rangle_0\,\widetilde{g}(\lambda,-\lambda,\omega)\,\langle\widetilde{g}\rangle_0]]\,d\lambda\,d\omega,\\
V_{\mathrm{id}|\langle\widetilde{g}\rangle_0} &= 2\,\langle[[\widetilde{g}\,\langle\widetilde{g}\rangle_0\,\widetilde{g}]]\rangle_0 + \frac{1}{(2\pi)^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}[[\widetilde{g}(\lambda,-\lambda,\omega)\,\langle\widetilde{g}\rangle_0]]\,d\lambda\,d\omega,\\
V_{\mathrm{id}|\mathrm{id}} &= 2\,\langle[[\widetilde{g}^2]]\rangle_0 + \frac{1}{(2\pi)^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}[[\widetilde{g}(\lambda,-\lambda,\omega)]]\,d\lambda\,d\omega,\\
V_{\langle\widetilde{g}\rangle_0|\mathrm{id,id}} &= 4\,\langle[[\widetilde{g}]]\,[[\widetilde{g}\,\langle\widetilde{g}\rangle_0\,\widetilde{g}]]\rangle_0 + 4\,\langle[[\widetilde{g}^2\,\langle\widetilde{g}\rangle_0\,\widetilde{g}]]\rangle_0 + \frac{1}{(2\pi)^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} g_{123,\langle\widetilde{g}\rangle_0}(\lambda,\omega)\,d\lambda\,d\omega,
\end{aligned}$$

$$\begin{aligned}
g_{123,\langle\widetilde{g}\rangle_0}(\lambda,\omega) &= 2\,[[\widetilde{g}(\lambda,-\lambda,\omega)\,\langle\widetilde{g}\rangle_0]]\,[[\widetilde{g}(\lambda)]] + 2\,[[\widetilde{g}(\lambda)\,\widetilde{g}(\lambda,-\lambda,\omega)\,\langle\widetilde{g}\rangle_0]],\\
V_{\mathrm{id}|\mathrm{id,id}} &= 4\,\langle[[\widetilde{g}]]\,[[\widetilde{g}^2]]\rangle_0 + 4\,\langle[[\widetilde{g}^3]]\rangle_0 + \frac{1}{(2\pi)^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} g_{123,\mathrm{id}}(\lambda,\omega)\,d\lambda\,d\omega,\\
g_{123,\mathrm{id}}(\lambda,\omega) &= 2\,[[\widetilde{g}(\lambda,-\lambda,\omega)]]\,[[\widetilde{g}(\lambda)]] + 2\,[[\widetilde{g}(\lambda)\,\widetilde{g}(\lambda,-\lambda,\omega)]],\\
V_{\mathrm{id,id}|\mathrm{id,id}} &= 8\,\langle[[\widetilde{g}]]^2\,[[\widetilde{g}^2]]\rangle_0 + 16\,\langle[[\widetilde{g}]]\,[[\widetilde{g}^3]]\rangle_0 + 12\,\langle[[\widetilde{g}]]^4\rangle_0 + 4\,\langle[[\widetilde{g}^2]]^2\rangle_0\\
&\quad + \frac{1}{(2\pi)^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} g_{1234}(\lambda,\omega)\,d\lambda\,d\omega,\\
g_{1234}(\lambda,\omega) &= 4\,[[\widetilde{g}(\lambda,-\lambda,\omega)]]\,[[\widetilde{g}(\lambda)]\,[\widetilde{g}(\omega)]] + 4\,[[\widetilde{g}(\lambda,-\lambda,\omega)\,\widetilde{g}(\omega)]]\,[[\widetilde{g}(\lambda)]]\\
&\quad + 4\,[[\widetilde{g}(\lambda)\,\widetilde{g}(\lambda,-\lambda,\omega)]]\,[[\widetilde{g}(\omega)]] + 4\,[[\widetilde{g}(\lambda)\,\widetilde{g}(\lambda,-\lambda,\omega)\,\widetilde{g}(\omega)]].
\end{aligned}$$

The stated result now follows. $\square$

Proof of Corollary B.1. Taking $\widetilde{g} \equiv \Sigma = \langle \widetilde{g} \rangle_0$ in the variance expressions of Theorem B.1 yields

$$
\begin{aligned}
V_{\langle \widetilde{g} \rangle_0 | \langle \widetilde{g} \rangle_0} &= 2\,[[\Sigma^4]] + \frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} [[\Sigma\, \widetilde{g}(\lambda, -\lambda, \omega)\, \Sigma]]\, d\lambda\, d\omega, \\
V_{\mathrm{id} | \langle \widetilde{g} \rangle_0} &= 2\,[[\Sigma^3]] + \frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} [[\widetilde{g}(\lambda, -\lambda, \omega)\, \Sigma]]\, d\lambda\, d\omega, \\
V_{\mathrm{id} | \mathrm{id}} &= 2\,[[\Sigma^2]] + \frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} [[\widetilde{g}(\lambda, -\lambda, \omega)]]\, d\lambda\, d\omega, \\
V_{\langle \widetilde{g} \rangle_0 | \mathrm{id}, \mathrm{id}} &= 4\,[[\Sigma]]\,[[\Sigma^3]] + 4\,[[\Sigma^4]] + \frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} g_{123, \langle \widetilde{g} \rangle_0}(\lambda, \omega)\, d\lambda\, d\omega, \\
g_{123, \langle \widetilde{g} \rangle_0}(\lambda, \omega) &= 2\,[[\widetilde{g}(\lambda, -\lambda, \omega)\, \Sigma]]\,[[\Sigma]] + 2\,[[\Sigma\, \widetilde{g}(\lambda, -\lambda, \omega)\, \Sigma]], \\
V_{\mathrm{id} | \mathrm{id}, \mathrm{id}} &= 4\,[[\Sigma]]\,[[\Sigma^2]] + 4\,[[\Sigma^3]] + \frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} g_{123, \mathrm{id}}(\lambda, \omega)\, d\lambda\, d\omega, \\
g_{123, \mathrm{id}}(\lambda, \omega) &= 2\,[[\widetilde{g}(\lambda, -\lambda, \omega)]]\,[[\Sigma]] + 2\,[[\Sigma\, \widetilde{g}(\lambda, -\lambda, \omega)]], \\
V_{\mathrm{id}, \mathrm{id} | \mathrm{id}, \mathrm{id}} &= 8\,[[\Sigma]]^2\,[[\Sigma^2]] + 16\,[[\Sigma]]\,[[\Sigma^3]] + 12\,[[\Sigma^4]] + 4\,[[\Sigma^2]]^2 \\
&\quad + \frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} g_{1234}(\lambda, \omega)\, d\lambda\, d\omega, \\
g_{1234}(\lambda, \omega) &= 4\,[[\Sigma]]^2\,[[\widetilde{g}(\lambda, -\lambda, \omega)]] + 4\,[[\Sigma]]\,[[\widetilde{g}(\lambda, -\lambda, \omega)\, \Sigma]] \\
&\quad + 4\,[[\Sigma\, \widetilde{g}(\lambda, -\lambda, \omega)]]\,[[\Sigma]] + 4\,[[\Sigma\, \widetilde{g}(\lambda, -\lambda, \omega)\, \Sigma]].
\end{aligned}
$$

Now with $v' = \{-2, -2\,[[\Sigma]], 1\}$, we find the limiting variance is (after some cancellations)

$$
\begin{aligned}
& 4\, V_{\langle \widetilde{g} \rangle_0 | \langle \widetilde{g} \rangle_0} + 8\,[[\Sigma]]\, V_{\mathrm{id} | \langle \widetilde{g} \rangle_0} + 4\,[[\Sigma]]^2\, V_{\mathrm{id} | \mathrm{id}} - 4\, V_{\langle \widetilde{g} \rangle_0 | \mathrm{id}, \mathrm{id}} - 4\,[[\Sigma]]\, V_{\mathrm{id} | \mathrm{id}, \mathrm{id}} + V_{\mathrm{id}, \mathrm{id} | \mathrm{id}, \mathrm{id}} \\
&= 4\,[[\Sigma^4]] + 4[[\Sigma^2]]^2 + \frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} h_{1234}(\lambda, \omega)\, d\lambda\, d\omega, \\
h_{1234}(\lambda, \omega) &= 4\,[[\Sigma]]\,[[\widetilde{g}(\lambda, -\lambda, \omega)\, \Sigma]] - 4\,[[\Sigma]]\,[[\Sigma\, \widetilde{g}(\lambda, -\lambda, \omega)]].
\end{aligned}
$$

Using the fact proved in Corollary 1 that the double integral of $[[A\, f(\lambda, -\lambda, \omega)\, B]]$ is equal to that of $[[B\, f(\omega, -\omega, \lambda)\, A]]$, we find that the integral of h_{1234} is zero, and the result is proved. ☐

References

[1] Brillinger, D.R. (2001) *Time Series: Data Analysis and Theory*. Philadelphia: Siam.

[2] McElroy, T.S. (2016) Non-nested Model Comparisons for Time Series. *Biometrika* **103**, 905–914.

[3] Taniguchi, M., and Kakizawa, Y. (2000) *Asymptotic Theory of Statistical Inference for Time Series.* Springer-Verlag, New York.

Appendix D Supplementary Tables

D.1 Bivariate Inflation

We computed the Gaussian divergence (-2 times the log Gaussian likelihood) for both MOM estimates, unrestricted MLE (i.e., the MLE for the related trends model), and restricted MLE (i.e., the MLE for the common trends model), resulting in -1680.292, -1702.763, and -1708.075 respectively. Since lower values are better, we see that the MLE fits of both the related trends and common trends models are better than that of the MOM fit. (The common trends model is a restriction of the related trends model, but nonlinear optimization does not find the MLE to correspond to near-full correlation, and hence the common trends fit has a lower divergence.) Both the MOM and MLE fits of the related trends model provide an adequate whitening transformation (the portmanteau tests of the residuals fail to reject the null hypothesis of white noise). We can also directly compare the resulting covariance matrix estimates, expressed in units of millions in Table D.1.

	MOM		unrestricted MLE		restricted MLE	
Σ_μ	Core	Total	Core	Total	Core	Total
Core	15.660	34.396	4.871	6.483	4.912	5.393
Total	34.396	122.889	6.483	11.264	5.393	5.920
Σ_ι	Core	Total	Core	Total	Core	Total
Core	10.781	5.256	17.607	25.595	18.640	26.295
Total	5.256	81.805	25.595	161.374	26.295	170.666

Table D.1: Estimates of Σ_μ (upper rows) and Σ_ι (lower rows), the trend and irregular covariance matrix (units of millions), respectively, for bivariate inflation data, based on MOM, unrestricted MLE, and common trends MLE.

D.2 Four-variate Housing Starts

The model used for housing starts consists generalizes (7) slightly: the irregular component has the same definition, but the trend is second order and the seasonal $\{\xi_t\}$ is given as the sum of six atomic processes $\xi_t^{(j)}$ $(1 \leq j \leq 6)$:

$$\begin{aligned}
(1-B)^2 \mu_t &= \eta_t \\
\xi_t &= \xi_t^{(1)} + \xi_t^{(2)} + \xi_t^{(3)} + \xi_t^{(4)} + \xi_t^{(5)} + \xi_t^{(6)} \\
(1 - \sqrt{3}B + B^2)\xi_t^{(1)} &= \varepsilon_t^{(1)} \\
(1 - B + B^2)\xi_t^{(2)} &= \varepsilon_t^{(2)} \\
(1 + B^2)\xi_t^{(3)} &= \varepsilon_t^{(3)} \\
(1 + B + B^2)\xi_t^{(4)} &= \varepsilon_t^{(4)} \\
(1 + \sqrt{3}B + B^2)\xi_t^{(5)} &= \varepsilon_t^{(5)} \\
(1 + B)\xi_t^{(6)} &= \varepsilon_t^{(6)},
\end{aligned}$$

where the $\{\varepsilon_t^{(j)}\}$ are each independent vector white noise processes. For the reduced span of the last nine years, we fitted both MOM and MLE (unrestricted), with divergences 959.806 and 923.488 respectively. From the standpoint of likelihood, the MOM estimates are inferior to MLE, although both adequately whiten the data. As for the covariance estimates, there is a fairly close agreement between the MLE and MOM based on the nine-year span; the MOM covariances based on the full span are also quite close.

	9-year MLE				9-year MOM				49-year MOM			
Σ_μ	South	West	NE	MW	South	West	NE	MW	South	West	NE	MW
South	0.093	0.042	0.014	0.028	0.088	0.040	0.012	0.028	0.094	0.038	0.011	0.021
West	0.042	0.024	0.007	0.017	0.040	0.021	0.006	0.014	0.038	0.024	0.005	0.011
NE	0.014	0.007	0.004	0.006	0.012	0.006	0.003	0.005	0.011	0.005	0.003	0.004
MW	0.028	0.017	0.006	0.015	0.028	0.014	0.005	0.010	0.021	0.011	0.004	0.009
$\Sigma_{\xi^{(1)}}$	South	West	NE	MW	South	West	NE	MW	South	West	NE	MW
South	0.077	0.063	0.012	0.035	0.065	0.074	0.016	0.051	0.107	0.087	0.028	0.064
West	0.063	0.064	0.013	0.038	0.074	0.090	0.016	0.057	0.087	0.080	0.020	0.048
NE	0.012	0.013	0.006	0.011	0.016	0.016	0.005	0.013	0.028	0.020	0.009	0.017
MW	0.035	0.038	0.011	0.033	0.051	0.057	0.013	0.041	0.064	0.048	0.017	0.040
$\Sigma_{\xi^{(2)}}$	South	West	NE	MW	South	West	NE	MW	South	West	NE	MW
South	0.028	-0.002	0.004	0.020	0.015	-0.000	0.011	0.028	0.023	0.014	0.006	0.029
West	-0.002	0.011	0.004	0.006	-0.000	0.023	-0.006	-0.001	0.014	0.012	-0.001	0.015
NE	0.004	0.004	0.011	0.014	0.011	-0.006	0.012	0.020	0.006	-0.001	0.008	0.011
MW	0.020	0.006	0.014	0.043	0.028	-0.001	0.020	0.051	0.029	0.015	0.011	0.036
$\Sigma_{\xi^{(3)}}$	South	West	NE	MW	South	West	NE	MW	South	West	NE	MW
South	0.090	0.064	-0.015	0.027	0.051	0.065	-0.001	0.058	0.075	-0.013	0.023	0.041
West	0.064	0.055	-0.013	0.022	0.065	0.112	-0.015	0.059	-0.013	0.032	-0.004	-0.008
NE	-0.015	-0.013	0.010	-0.000	-0.001	-0.015	0.061	0.021	0.023	-0.004	0.020	0.011
MW	0.027	0.022	-0.000	0.020	0.058	0.059	0.021	0.079	0.041	-0.008	0.011	0.023
$\Sigma_{\xi^{(4)}}$	South	West	NE	MW	South	West	NE	MW	South	West	NE	MW
South	0.174	0.026	-0.035	-0.014	0.015	0.007	-0.014	0.014	0.042	-0.006	-0.002	-0.007
West	0.026	0.032	-0.004	0.004	0.007	0.068	-0.014	0.024	-0.006	0.002	-0.000	0.001
NE	-0.035	-0.004	0.011	0.003	-0.014	-0.014	0.014	-0.015	-0.002	-0.000	0.000	0.000
MW	-0.014	0.004	0.003	0.009	0.014	0.024	-0.015	0.018	-0.007	0.001	0.000	0.001
$\Sigma_{\xi^{(5)}}$	South	West	NE	MW	South	West	NE	MW	South	West	NE	MW
South	0.012	0.013	-0.000	-0.004	0.014	0.025	0.001	0.009	0.009	0.011	-0.002	0.004
West	0.013	0.024	0.000	0.008	0.025	0.044	0.002	0.014	0.011	0.013	-0.003	0.004
NE	-0.000	0.000	0.001	0.001	0.001	0.002	0.000	0.002	-0.002	-0.003	0.001	-0.003
MW	-0.004	0.008	0.001	0.021	0.009	0.014	0.002	0.026	0.004	0.004	-0.003	0.011
$\Sigma_{\xi^{(6)}}$	South	West	NE	MW	South	West	NE	MW	South	West	NE	MW
South	0.205	0.010	0.003	-0.002	0.111	0.003	0.009	0.043	0.030	-0.006	0.003	0.013
West	0.010	0.036	0.001	0.015	0.003	0.003	-0.001	0.000	-0.006	0.002	0.001	-0.003
NE	0.003	0.001	0.009	0.003	0.009	-0.001	0.001	0.004	0.003	0.001	0.002	0.001
MW	-0.002	0.015	0.003	0.042	0.043	0.000	0.004	0.017	0.013	-0.003	0.001	0.006
Σ_ι	South	West	NE	MW	South	West	NE	MW	South	West	NE	MW
South	3.783	-0.083	0.127	-0.537	10.419	-0.266	0.224	-3.556	7.862	0.075	0.155	-0.543
West	-0.083	0.901	0.171	0.456	-0.266	0.685	0.262	0.443	0.075	2.792	0.097	0.286
NE	0.127	0.171	0.347	0.160	0.224	0.262	0.512	-0.048	0.155	0.097	0.949	0.165
MW	-0.537	0.456	0.160	0.946	-3.556	0.443	-0.048	2.384	-0.543	0.286	0.165	1.824

Table D.2: For four-variate Starts data, estimates based on MLE (9-year span), MOM (9-year span), and MOM (49-year span) of covariance matrices for trend (Σ_μ), first atomic seasonal ($\Sigma_{\xi^{(1)}}$), second atomic seasonal ($\Sigma_{\xi^{(2)}}$), third atomic seasonal ($\Sigma_{\xi^{(3)}}$), fourth atomic seasonal ($\Sigma_{\xi^{(4)}}$), fifth atomic seasonal ($\Sigma_{\xi^{(5)}}$), sixth atomic seasonal ($\Sigma_{\xi^{(6)}}$), and irregular (Σ_ι).