Asymptotic accuracy of Iterative Feedback Tuning *

R. Hildebrand†, A. Lecchini‡, G. Solari* and M. Gevers*

November 2, 2004

†Laboratoire de Modélisation et Calcul (LMC)
Université Joseph Fourier (UJF) Grenoble, France
roland.hildebrand@imag.fr

‡Department of Engineering, University of Cambridge, UK
al394@cam.ac.uk

*Center for Systems Engineering and Applied Mechanics (CESAME)
Université Catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium
{solari, gevers}@csam.ucl.ac.be

Abstract

Iterative Feedback Tuning (IFT) is a widely used procedure for controller tuning. It is a sequence of iteratively performed special experiments on the plant interlaced with periods of data collection under normal operating conditions. In this paper we derive the asymptotic convergence rate of IFT for disturbance rejection, which is one of the main fields of application.

*Paper supported by the the Belgian Programme on Interuniversity Attraction Poles, initiated by the Belgian Federal Science Policy Office and the European Research Network on System Identification (ERNSI) funded by the European Union. The scientific responsibility rests with its authors.

1 Introduction

Iterative Feedback Tuning (IFT) is a data-based method for the tuning of restricted complexity controllers. It has proved to be very effective in practice and is now widely used in process control, often for disturbance rejection. The reader is referred to [6] for a recent overview. The objective of IFT is to minimize a quadratic performance criterion. IFT is a stochastic gradient descent scheme in a finitely parameterized controller space. The gradient of the cost function at each step is estimated from data. These data are collected with the actual controller in the loop. Under suitable assumptions the algorithm converges to a local minimum of the performance criterion. For more details of the procedure see [7].

In this paper we provide an analytic expression for the asymptotic convergence rate of IFT for a disturbance rejection. The convergence rate depends on the covariance of the gradient estimates. Therefore, the calculation of this covariance is a part of our analysis.

The remainder of the paper is structured as follows. In the next section we summarize the details of the IFT algorithm for disturbance rejection. In Section 3 we derive an expression for the asymptotic convergence rate dependent on the covariance of the gradient estimates. In Section 4 the asymptotic expression of this covariance is calculated. Conclusions are given in Section 5. The Appendix contains all the technical proofs.

2 IFT for disturbance rejection

In this section we review the IFT method for the disturbance rejection problem with a classical LQ criterion. For a more general and detailed presentation of IFT the reader is referred to [7, 8].

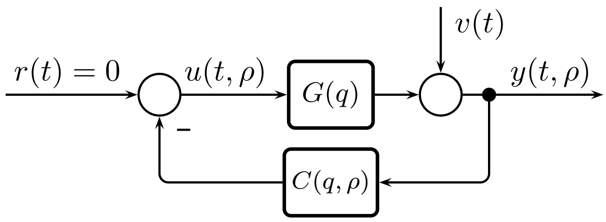

Figure 1: The control system under normal operating conditions.

Consider a SISO discrete time system described by

$$y(t) = G(q)u(t) + v(t)\,, \tag{1}$$

where $y(t)$ is the output, $u(t)$ is the input, $G(q)$ is a linear time-invariant transfer function, with q being the shift operator, and $v(t)$ is the process disturbance, assumed to be quasistationary with zero mean and spectral density $\Phi_v(\omega)$. The transfer function $G(q)$ and the disturbance spectrum $\Phi_v(\omega)$ are unknown.

Consider the feedback loop around $G(q)$ depicted in Figure 1, where $C(q, \rho)$ is a one-degree-of-freedom controller belonging to a parameterized set of controllers with parameter $\rho \in \mathbf{R}^n$. The transfer function from $v(t)$ to $y(t, \rho)$ is named sensitivity function and is denoted by $S(q, \rho)$. We assume that in the control system of Figure 1 the reference signal $r(t)$ is set at zero under normal operating conditions. Our goal is to tune the controller $C(q, \rho)$ so that the variance of the noise-driven closed loop output is as small as possible subject to a penalty on the control effort. Thus we want to find a minimizer for the cost function

$$J(\rho) = \frac{1}{2}\,\mathbf{E}\left[y(t, \rho)^2 + \lambda u(t, \rho)^2\right]\,, \tag{2}$$

where $\lambda \geq 0$ is chosen by the user. The IFT method yields an approximate solution to the above problem. IFT is based on the possibility of obtaining an unbiased estimate of the gradient $\frac{\partial J}{\partial \rho}(\rho)$ of the cost function at $\rho = \rho_n$ from data collected from the closed-loop system with the controller $C(\rho_n)$ operating on the loop. The cost function $J(\rho)$ can then be minimized with an iterative stochastic gradient descent scheme of Robbins-Monro type [1]. In that scheme, a sequence of controllers $C(q, \rho_n)$ is computed and applied to the plant. In the n-th iteration step, data obtained from the system with the controller $C(\rho_n)$ operating on the loop are used to construct the next parameter vector ρ_{n+1}. The data-based iterative procedure is as follows.

IFT PROCEDURE

1. Collect a sequence $\{u^1(t, \rho_n),\, y^1(t, \rho_n)\}_{t=1,\dots,N}$ of N input-output data under normal operating conditions, i.e. without reference signal.

2. Collect a sequence $\{u^2(t, \rho_n),\, y^2(t, \rho_n)\}_{t=1,\dots,N}$ of N input-output data by performing a special experiment with reference signal

$$r_n^2(t) = -K_n(q) y^1(t, \rho_n)$$

where $K_n(q)$ is any stable minimum-phase prefilter.

3. Construct the estimates of the gradients of $u^1(t, \rho_n)$ and $y^1(t, \rho_n)$ as

$$\begin{aligned} est\left[\frac{\partial u^1}{\partial \rho}(t, \rho_n)\right] &= \frac{1}{K_n(q)} \frac{\partial C}{\partial \rho}(q, \rho_n)\, u^2(t, \rho_n)\ , \\ est\left[\frac{\partial y^1}{\partial \rho}(t, \rho_n)\right] &= \frac{1}{K_n(q)} \frac{\partial C}{\partial \rho}(q, \rho_n)\, y^2(t, \rho_n)\, . \end{aligned}$$

4. Form the estimate of the gradient of $J(\rho)$ at ρ_n as

$$est_N\left[\frac{\partial J}{\partial \rho}(\rho_n)\right] = \frac{1}{N}\sum_{t=1}^{N}\left[y^1(t, \rho_n) est\left[\frac{\partial y^1}{\partial \rho}(t, \rho_n)\right] + \lambda u^1(t, \rho_n) est\left[\frac{\partial u^1}{\partial \rho}(t, \rho_n)\right]\right].$$

5. Calculate the new parameter vector ρ_{n+1} according to

$$\rho_{n+1} = \rho_n - \gamma_n R_n^{-1}\, est_N \left[\frac{\partial J}{\partial \rho}(\rho_n)\right]$$

where γ_n is a positive step size and R_n is a symmetric positive definite matrix.

We recall that the estimate of the gradient calculated in step 4 is unbiased under the assumption that the disturbance realizations $v_n^1(t)$, in the first experiment, and $v_n^2(t)$, in the second experiment, are independent. This assumption can be considered fulfilled if the two experiments in the algorithm are sufficiently separated in time. In the procedure, the sequences γ_n and R_n are basically left to the choice of the user. The matrix R_n should be an approximation of the Hessian of the cost function in ρ_n. A biased estimate of the Hessian, obtained from data, has been proposed in [7]. The prefilter $K_n(q)$ is also a degree of freedom in the algorithm; it affects the signal to noise ratio in the second experiment. Two possible choices for prefilter $K_n(q)$, derived from the results presented in the present paper, are discussed in [4] and [3], respectively.

3 Analysis of the convergence rate of IFT

In this section we quantify the effect of the variability of the gradient estimate on the asymptotic convergence rate of the algorithm. The proposition below derives from a more general version of the same proposition for Robbins-Monro processes as can be found in [9, 11]. In the proposition we assume convergence of the sequence ρ_n. The reader is referred to [2, 5] for a detailed proof of convergence.

Proposition 3.1 *Assume that the sequence ρ_n converges to a local isolated minimum $\bar{\rho}$ of $J(\rho)$. Let $H(\bar{\rho})$ be the Hessian of $J(\rho)$ at $\rho = \bar{\rho}$. Suppose further that the following conditions hold.*

1. The sequence γ_n of step sizes is given by $\gamma_n = \frac{a}{n}$, where a is a positive constant. There exists an index $\bar{n}$ and a matrix R such that $R_n = R$ for all $n > \bar{n}$.

2. The matrix $A = \frac{1}{2}I - aR^{-1}H(\bar{\rho})$ is stable, i.e. the real parts of its eigenvalues are negative.

3. The covariance matrix $\mathbf{Cov}\left[est_N\left[\frac{\partial J}{\partial \rho}(\rho)\right]\right]$ at $\rho = \bar{\rho}$ is positive definite.

Then the sequence of random variables $\sqrt{n}(\rho_n - \bar{\rho})$ converges in distribution to a normally distributed zero mean random variable with covariance matrix

$$\Sigma = a^2 \int_0^\infty e^{At} R^{-1} \mathbf{Cov}\left[est_N\left[\frac{\partial J}{\partial \rho}(\bar{\rho})\right]\right] R^{-1} e^{A^T t} dt, \tag{3}$$

i.e. $\sqrt{n}(\rho_n - \bar{\rho}) \xrightarrow{D} \mathcal{N}(0, \Sigma)$. □

Proposition 3.1 shows that the asymptotic accuracy of the parameter estimate crucially depends on the covariance of the gradient estimate.

4 The covariance of the gradient estimate

In this section we compute an explicit expression for the covariance of $est_N\left[\frac{\partial J}{\partial \rho}(\rho_n)\right]$. We show that this covariance can be written as the sum of two terms. These two contributions originate in the variability of the noise realizations in the first and second experiment of iteration n, respectively.

It can be easily seen that the estimates of the gradients of $u^1(t, \rho_n)$ and $y^1(t, \rho_n)$ obtained in Step 3 of the IFT procedure are corrupted by the realization $v_n^2(t)$ of the noise in the second experiment as follows

$$\begin{aligned}
est\left[\frac{\partial u^1}{\partial \rho}(t, \rho_n)\right] &= \frac{\partial u^1}{\partial \rho}(t, \rho_n) - \frac{S(q, \rho_n)}{K_n(q)} C(q, \rho_n) \frac{\partial C}{\partial \rho}(q, \rho_n)\, v_n^2(t)\,, \\
est\left[\frac{\partial y^1}{\partial \rho}(t, \rho_n)\right] &= \frac{\partial y^1}{\partial \rho}(t, \rho_n) + \frac{S(q, \rho_n)}{K_n(q)} \frac{\partial C}{\partial \rho}(q, \rho_n)\, v_n^2(t)\,.
\end{aligned}$$

Therefore we can separate $est_N\left[\frac{\partial J}{\partial\rho}(\rho_n)\right]$ as

$$\begin{aligned}
est_N\left[\frac{\partial J}{\partial\rho}(\rho_n)\right] &= S_N(\rho_n)+E_N(\rho_n)\,,\ \text{with}\\
S_N(\rho_n) &= \frac{1}{N}\sum_{t=1}^{N}\left[y^1(t,\rho_n)\frac{\partial y^1}{\partial\rho}(t,\rho_n)+\lambda u^1(t,\rho_n)\frac{\partial u^1}{\partial\rho}(t,\rho_n)\right],\\
E_N(\rho_n) &= \frac{1}{N}\sum_{t=1}^{N}\left[y^1(t,\rho_n)\left[\frac{S(q,\rho_n)}{K_n(q)}\frac{\partial C}{\partial\rho}(q,\rho_n)\,v_n^2(t)\right]\right.\\
&+\left.\lambda u^1(t,\rho_n)\left[-\frac{C(q,\rho_n)S(q,\rho_n)}{K_n(q)}\frac{\partial C}{\partial\rho}(q,\rho_n)\,v_n^2(t)\right]\right].
\end{aligned}$$

The term $S_N(\rho_n)$ corresponds to the sampled estimate of the gradient of $J(\rho)$. This term is entirely dependent on the realization $v_n^1(t)$ of the noise in the first experiment. The second term $E_N(\rho_n)$ is an error due to the corruption of the estimates of the gradients of $u^1(t,\rho_n)$ and $y^1(t,\rho_n)$ by $v_n^2(t)$. The covariance of $est_N\left[\frac{\partial J}{\partial\rho}(\rho_n)\right]$ is described in the following proposition, which is the main result of this paper.

Proposition 4.1

1. *The following relation holds*

$$\mathbf{Cov}\left[est_N\left[\frac{\partial J}{\partial\rho}(\rho_n)\right]\right]=\mathbf{Cov}\left[S_N(\rho_n)\right]+\mathbf{Cov}\left[E_N(\rho_n)\right]\,.$$

2. *The following asymptotic frequency-domain expression of* $\mathbf{Cov}\left[E_N(\rho_n)\right]$ *holds*

$$\begin{aligned}
\lim_{N\to\infty} N\mathbf{Cov}\left[E_N(\rho_n)\right] &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{|K_n(e^{j\omega})|^2}|S(e^{j\omega},\rho_n)|^4\left[1+\lambda|C(e^{j\omega},\rho_n)|^2\right]^2\\
&\times\ \frac{\partial C}{\partial\rho}(e^{j\omega},\rho_n)\frac{\partial C^*}{\partial\rho}(e^{j\omega},\rho_n)\,\Phi_v^2(\omega)d\omega\,.
\end{aligned}$$

3. *Under the additional assumption that the 4th order cumulants of the noise* v *are zero (e.g the noise is normally distributed), the following asymptotic frequency-domain expression of*

Cov $[S_N(\rho_n)]$ *holds*

$$\begin{aligned}\lim_{N\to\infty} N\mathbf{Cov}\left[S_N(\rho_n)\right] &= 2\cdot\frac{1}{2\pi}\int_{-\pi}^{\pi}|S(e^{j\omega},\rho_n)|^4\Phi_v^2(\omega)\\ &\times \mathcal{R}e\left\{\left[G(e^{j\omega})-\lambda\bar{C}(e^{j\omega},\rho_n)\right]S(e^{j\omega},\rho_n)\frac{\partial C}{\partial\rho}(e^{j\omega},\rho_n)\right\}\\ &\times \mathcal{R}e\left\{\left[G(e^{j\omega})-\lambda\bar{C}(e^{j\omega},\rho_n)\right]S(e^{j\omega},\rho_n)\frac{\partial C}{\partial\rho}(e^{j\omega},\rho_n)\right\}^T d\omega\,.\end{aligned}$$

Proof: *see Appendix.* □

Proposition 4.1 shows that the covariance of the gradient estimate can be represented as the sum of the covariances of the separate contributions $S_N(\rho_n)$ and $E_N(\rho_n)$ (i.e. $S_N(\rho_n)$ and $E_N(\rho_n)$ are uncorrelated). Both **Cov** $[S_N(\rho_n)]$ and **Cov** $[E_N(\rho_n)]$ decay asymptotically as $1/N$ as the number of data tends to infinity. Their asymptotic frequency domain expressions as $N\to\infty$ have been given.

5 Conclusions

In this contribution we have investigated the asymptotic accuracy of the IFT algorithm in the case of disturbance rejection. The result presented in this paper has been used to derive optimal choices for the prefilter $K_n(q)$ in two different situations. In [4], we consider the situation where the current controller is near the optimal controller, and we derive a prefilter which optimally increases the asymptotic accuracy of IFT under a constraint on the energy used during the special feedback experiment. In [3] we optimize the prefilter for accuracy of a single IFT step, under the same energy constraint. This second prefilter can be used when the current controller can be considered far from the optimal one (e.g. during the initial steps of the procedure).

References

[1] Julius R. Blum. Multidimensional stochastic approximation methods. *Annals of Mathematical Statistics*, 25:737–744, 1954.

[2] R. Hildebrand, A. Lecchini, G. Solari, and M. Gevers. A convergence theorem for Iterative Feedback Tuning. Technical Report 2003/17, CESAME - UCL, Louvain-la-Neuve, Belgium, 2003.

[3] R. Hildebrand, A. Lecchini, G. Solari, and M. Gevers. Prefiltering in Iterative Feedback Tuning: optimization of the prefilter for accuracy. *Submitted to IEEE Trans. on Automatic Control*, 2003.

[4] R. Hildebrand, A. Lecchini, G. Solari, and M. Gevers. Optimal prefiltering in Iterative Feedback Tuning. *Submitted to IEEE Trans. on Automatic Control*, 2004.

[5] H. Hjalmarsson. Performance analysis of iterative feedback tuning. Technical report, Dept of Signals, Sensors and Systems, Royal Institute of Technology, Stockholm, Sweden, 1998.

[6] H. Hjalmarsson. Iterative Feedback Tuning - an overview. *International Journal of Adaptive Control and Signal Processing*, 16(5):373–395, Special Issue: New Ideas in Recursive Estimation and Stochastic Adaptive Control 2002.

[7] H. Hjalmarsson, M. Gevers, S. Gunnarson, and O. Lequin. Iterative Feedback Tuning: theory and applications. *IEEE Control Systems Magazine*, 18(4):26–41, August 1998.

[8] H. Hjalmarsson, S. Gunnarson, and M. Gevers. A convergent iterative restricted complexity control design scheme. In *Proc. of the 33rd IEEE Conference on Decision and Control*, pages 1735–1740, Lake Buena Vista, FL, Dec. 1994.

[9] M. B. Nevelson and R. Z. Khasminskii. *Stochastic Approximation and Recursive Estimation*, volume 47 of *Translations of Mathematical Monographs*. American Mathematical Society, third printing 2000 edition, 1976.

[10] Murray Rosenblatt. *Stationary sequences and random fields*. Birkhaeuser, 1985.

[11] G. Yin. A stopped stochastic approximation algorithm. *Systems and Control Letters*, 11:107–115, 1988.

A Proof of Proposition 4.1

In order to prove the Proposition, we shall need the following technical results.

Lemma A.1 *Let e, f be two independent realizations of a zero mean white noise with variance σ^2. Let A, C be stable transfer functions and B, D be column vectors of stable transfer functions of equal length. Let $a = Ae, b = Bf, c = Ce, d = Df$ be signals obtained by filtering e and f through A, B, C, D. Then*

$$\lim_{N\to\infty} \mathbf{E}\left[\frac{1}{N}\sum_{t,s=1}^{N} a(t)b(t)c(s)d(s)^T\right] = \sigma^4 \frac{1}{2\pi}\int_{-\pi}^{\pi} \Phi_{ac}\bar{\Phi}_{bd^T}\, d\omega. \tag{4}$$

Here Φ_{gh} denotes the cross-spectrum of the signals g, h.

Proof: The assertion is a direct consequence of the independence of e, f and Parseval's Theorem. □

Lemma A.2 *Let A, C be stable transfer functions and B, D be column vectors of stable transfer functions of equal length. Let v be a quasistationary zero mean stochastic process satisfying (6). Let $a = Av, b = Bv, c = Cv, d = Dv$ be signals obtained by filtering v through A, B, C, D and let $\alpha, \beta, \gamma, \delta$ be fixed delays. Then*

$$\bar{E}[a(t-\alpha)b(t-\beta)c(t-\gamma)d(t-\delta)^T] \;=\; R_{ab}(\beta-\alpha)R_{cd^T}(\delta-\gamma) + R_{ac}(\gamma-\alpha)R_{bd^T}(\delta-\beta)$$

$$+ \quad R_{bc}(\delta-\alpha)R_{ad^T}(\gamma-\beta), \tag{5}$$

where the time average is taken with respect to t *and* $R_{gh}(\tau)$ *denotes* $\bar{E}[g(t)h(t-\tau)]$.

Proof: The relation is easily verified by straightforward calculation using the fact that the autocorrelation coefficients of the signal v satisfy equation (6). □

Proof of Part 1 of Proposition 4.1

Since $E_N(\rho_n)$ has zero mean, we obtain

$$\begin{aligned}\mathbf{Cov}\left[S_N(\rho_n)+E_N(\rho_n)\right] &= \mathbf{Cov}\left[S_N(\rho_n)\right]+\mathbf{E}\left[E_N(\rho_n)\cdot S_N(\rho_n)^T\right]\\ &+ \mathbf{E}\left[E_N(\rho_n)\cdot S_N(\rho_n)^T\right]^T+\mathbf{Cov}\left[E_N(\rho_n)\right].\end{aligned}$$

Hence we have to show that $S_N(\rho_n)$ and $E_N(\rho_n)$ are uncorrelated. Note that $S_N(\rho_n)$ depends only on the noise realization $v_n^1(t)$. By independence of $v_n^1(t)$ and $v_n^2(t)$ we have

$$\begin{aligned}\mathbf{E}\left[E_N(\rho_n)\cdot S_N(\rho_n)^T\right] &= \frac{1}{N}\sum_{t=1}^{N}\mathbf{E}\left[y^1(t,\rho_n)S_N(\rho_n)\right]\mathbf{E}\left[\frac{S(q,\rho_n)}{K_n(q)}\frac{\partial C}{\partial\rho}(q,\rho_n)\,v_n^2(t)\right]\\ &+ \frac{\lambda}{N}\sum_{t=1}^{N}\mathbf{E}\left[u^1(t,\rho_n)S_N(\rho_n)\right]\mathbf{E}\left[-\frac{C(q,\rho_n)S(q,\rho_n)}{K_n(q)}\frac{\partial C}{\partial\rho}(q,\rho_n)\,v_n^2(t)\right].\end{aligned}$$

But $\mathbf{E}\left[\frac{S(q,\rho_n)}{K_n(q)}\frac{\partial C}{\partial\rho}(q,\rho_n)\,v_n^2(t)\right]=\mathbf{E}\left[-\frac{C(q,\rho_n)S(q,\rho_n)}{K_n(q)}\frac{\partial C}{\partial\rho}(q,\rho_n)\,v_n^2(t)\right]=0$ for all t, because $v_n^2(t)$ has zero mean. It follows that $\mathbf{E}\left[E_N(\rho_n)\cdot S_N(\rho_n)^T\right]=0$. □

Proof of Part 2 of Proposition 4.1

The claim follows from Lemma A.1 by writing $\mathbf{Cov}\left[E_N(\rho_n)\right]$ as a double sum over four separate terms and by inserting the cross-spectra of the corresponding transfer functions. □

Proof of Part 3 of Proposition 4.1

The assumption that the 4th order cumulants of the noise v are zero means that the 4th order properties of v are related to its second order properties in the same way as for a Gaussian stochastic process. To be more precise, let us denote the autocorrelation function $\bar{E}[v(t)v(t-\tau)]$ of v by $R_v(\tau)$. Then this assumption can be rewritten as

$$\begin{aligned}\bar{E}[v(p+t)v(q+t)v(r+t)v(s+t)] &= R_v(p-r)R_v(q-s)+R_v(p-s)R_v(q-r)\\ &+ R_v(p-q)R_v(r-s) \qquad \forall\, p,q,r,s. \end{aligned} \tag{6}$$

Here the time average is taken with respect to t and the numbers p, q, r, s are assumed to be arbitrary, but fixed. Relation (6) is not very restrictive. It is satisfied e.g. for filtered zero mean i.i.d. white noise, if the probability density function of the white noise has zero *kurtosis* ("peakedness", see e.g. [10]). This is equivalent to the condition that the 2nd and 4th moments m_2, m_4 of this probability density function satisfy the relation $m_4 = 3m_2^2$. This relation holds e.g. for a normal distribution.

Consider now the notations and assumptions of Lemma A.2. For notational convenience, define a column vector Q_N by

$$Q_N = \frac{1}{N}\sum_{t=1}^{N}\left[a(t)b(t)+c(t)d(t)\right] . \tag{7}$$

and notice that S_N has the same structure as Q_N. We have

$$\mathbf{Cov}\,[Q_N] = \mathbf{E}[Q_N Q_N^T] - \mathbf{E}[Q_N]\mathbf{E}[Q_N]^T .$$

By (5), we obtain

$$\begin{aligned}\mathbf{E}[Q_N Q_N^T] &= \frac{1}{N^2}\sum_{t,s=1}^{N}[R_{ab}(0)R_{ab^T}(0)+R_{aa}(t-s)R_{bb^T}(t-s)+R_{ba}(t-s)R_{ab^T}(t-s)\\ &+ R_{ab}(0)R_{cd^T}(0)+R_{ac}(t-s)R_{bd^T}(t-s)+R_{bc}(t-s)R_{ad^T}(t-s)\\ &+ R_{cd}(0)R_{ab^T}(0)+R_{ca}(t-s)R_{db^T}(t-s)+R_{da}(t-s)R_{cb^T}(t-s)\\ &+ R_{cd}(0)R_{cd^T}(0)+R_{cc}(t-s)R_{dd^T}(t-s)+R_{dc}(t-s)R_{cd^T}(t-s)] .\end{aligned}$$

On the other hand, we have

$$\mathbf{E}[Q_N]\mathbf{E}[Q_N]^T = R_{ab}(0)R_{ab^T}(0) + R_{ab}(0)R_{cd^T}(0) + R_{cd}(0)R_{ab^T}(0) + R_{cd}(0)R_{cd^T}(0)\,.$$

Subtracting above equations and taking the limit $N \to \infty$ yields

$$\begin{aligned}
\lim_{N\to\infty} N\mathbf{Cov}\left[Q_N\right] &= \sum_{\tau=-\infty}^{\infty} [R_{aa}(\tau)R_{bb^T}(\tau) + R_{ba}(\tau)R_{ab^T}(\tau) + R_{ac}(\tau)R_{bd^T}(\tau) \\
&+ R_{bc}(\tau)R_{ad^T}(\tau) + R_{ca}(\tau)R_{db^T}(\tau) + R_{da}(\tau)R_{cb^T}(\tau) \\
&+ R_{cc}(\tau)R_{dd^T}(\tau) + R_{dc}(\tau)R_{cd^T}(\tau)]\,.
\end{aligned}$$

Applying the formula

$$\sum_{\tau=-\infty}^{+\infty} R_{ab}(\tau)R_{cd}(\tau) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \Phi_{ab}(\omega)\bar{\Phi}_{cd}(\omega)\,d\omega$$

componentwise and inserting the expressions for the cross-spectra finally furnishes the claim of Proposition 4.1 with the obvious substitutions applying. □