ISSN 1066-5307, Mathematical Methods of Statistics, 2015, Vol. 24, No. 2, pp. 135–146. © Allerton Press, Inc., 2015.

Baxter's Inequality for Triangular Arrays

M. Meyer[1*], T. McMurry[2], and D. Politis[3***]**

[1] TU Braunschweig, Inst. f. Math. Stochastik, Braunschweig, Germany
[2] Univ. Virginia, Public Health Sci., Charlottesville, USA
[3] Dept. of Math., Univ. California, San Diego, USA

Received May 5, 2014

Abstract—A central problem in time series analysis is prediction of a future observation. The theory of optimal linear prediction has been well understood since the seminal work of A. Kolmogorov and N. Wiener during World War II. A simplifying assumption is to assume that one-step-ahead prediction is carried out based on observing the infinite past of the time series. In practice, however, only a finite stretch of the recent past is observed. In this context, Baxter's inequality is a fundamental tool for understanding how the coefficients in the finite-past predictor relate to those based on the infinite past. We prove a generalization of Baxter's inequality for triangular arrays of stationary random variables under the condition that the spectral density functions associated with the different rows converge. The motivating examples are statistical time series settings where the autoregressive coefficients are re-estimated as new data are acquired, producing new fitted processes — and new predictors — for each n.

Keywords: linear prediction, time series, weak stationarity.

2000 Mathematics Subject Classification: 62M20, 62M10, 60G10.

DOI: 10.3103/S1066530715020040

1. INTRODUCTION

Baxter's inequality ([1], [2]) provides a fundamental tool for understanding the behavior of linear predictors based on a finite observed history. In particular, let $(X_t)_{t\in\mathbb{Z}}$ be a mean zero, weakly stationary time series. Under reasonably general conditions (see, for example, [9], [10], [11]), such a process admits an AR(∞) representation

$$X_t = \sum_{k=1}^{\infty} a_k X_{t-k} + \varepsilon_t,$$

where $(\varepsilon_t)_{t\in\mathbb{Z}}$ is a white noise, stationary innovation process.

In principle, the sequence of autoregressive (AR) coefficients $(a_k)_{k=1}^{\infty}$ describes the optimal coefficients for the prediction of X_t based on an infinite observed history $(X_{t-k})_{k=1}^{\infty}$. Of course, in practice only a finite history $X_0, X_1, \ldots, X_{t-1}$ is observed, and predictions are made based only on the p most recent observations — with p typically much smaller than t.

It is well known ([3, p.167]) that the optimal prediction coefficients for the one-step-ahead linear predictor based on the p most recent observations are the solutions to the Yule–Walker equations

$$\begin{pmatrix} \gamma(0) & \ldots & \gamma(p-1) \\ \vdots & \ddots & \vdots \\ \gamma(p-1) & \ldots & \gamma(0) \end{pmatrix} \begin{pmatrix} a_1(p) \\ \vdots \\ a_p(p) \end{pmatrix} = \begin{pmatrix} \gamma(1) \\ \vdots \\ \gamma(p) \end{pmatrix},$$

*E-mail: marco.meyer@tu-bs.de
**E-mail: tmcmurry@virginia.edu
***E-mail: dpolitis@ucsd.edu

where $\gamma(k) = E[X_t X_{t-k}]$ is the autocovariance function of $(X_t)_{t\in\mathbb{Z}}$ at lag k. Baxter's inequality provides a bound on the vector difference between the finite-past predictor sequence $(a_k(p))_{k=1}^p$ and the infinite-past predictor sequence $(a_k)_{k=1}^p$; the bound depends on the tail decay of $(a_k)_{k=p+1}^\infty$. The original result already contains weighting functions which ensure sufficiently fast rates of convergence, as $p \to \infty$, if certain conditions are met.

Since Baxter's early work the result has been generalized to different scenarios. Hannan and Deistler (see [7]) give a proof in the context of vector-valued processes which contains the same weighting functions as in the univariate setting. Cheng and Pourahmadi (see [4]) discuss L^1- and L^2-versions for vector-valued processes which are, however, unweighted. In this work, we will introduce another important generalization to situations which occur naturally in many applications.

In statistical problems, the coefficients $a_1(p), \ldots, a_p(p)$ must be estimated from the data, and in principle will be re-estimated if/when additional data accumulates. For example, when a time series is observed on a regular basis, e.g., daily or monthly, then the observed sample size n increases as time goes by. The result is a triangular array of predictor coefficients $\big(a_k^{(n)}(p)\big)_{k=1}^p$, where both the coefficients and the order p depend on n.

As another example, consider a practitioner that wants to resample a given time series of length n. There are many bootstrap methods available in the literature; see the review by [8]. In each of these methods, notably, the resampling mechanism is done via a probability measure that depends on n; hence, the bootstrap time series represents the nth row of a triangular array.

In the present paper we establish a version of Baxter's inequality appropriate for this setting. Our inequality concerns a triangular array of observations

$$X_{1,n}, \ldots, X_{n,n} \tag{1}$$

from a sequence of processes with autocovariance functions $\gamma_n(\cdot)$ and spectral densities $f_n(\cdot)$ converging to a limiting autocovariance function $\gamma(\cdot)$ and spectral density $f(\cdot)$.

The remainder of the paper is structured as follows: Section 2 describes our assumptions and contains the statements of our main results. We will also state an auxiliary result that might be of its own interest. It transfers convergence of the Fourier coefficients of functions (in our case spectral densities) to convergence of the optimal factors of these functions in the spectral decomposition. Section 3 contains our proofs.

2. MAIN RESULTS

Consider a triangular array of observations

$$X_{1,n}, \ldots, X_{n,n}, \quad n \geq 1, \tag{2}$$

where, for each $n \in \mathbb{N}$, the data $X_{t,n} = X_t^{(n)}$ are generated by a weakly stationary, real-valued process $(X_t^{(n)})_{t\in\mathbb{Z}}$ with autocovariance function $\gamma_n(\cdot)$ and spectral density $f_n(\cdot)$. We impose the following assumptions.

Assumption 2.1. *The weakly stationary processes $(X_t^{(n)})$ have finite second moments and are purely non-deterministic for each $n \in \mathbb{N}$.*

Assumption 2.2. *For a nondecreasing weighting function $\nu\colon \mathbb{N}_0 \to [1, \infty)$ which fulfils the norm condition $\nu(k) \leq \nu(j) \cdot \nu(|k - j|)$ for all $j, k \geq 0$, as well as the so-called GRS-condition*

$$\lim_{n\to\infty} \nu(nx)^{1/n} = 1, \quad \forall\, x \in \mathbb{R},$$

cf. [6]*, the autocovariances fulfil*

$$\sum_{h=-\infty}^{\infty} \nu(|h|)\,(1 + |h|) \cdot |\gamma_n(h)| \leq C < \infty, \quad \forall\, n \in \mathbb{N}. \tag{3}$$

Furthermore, the autocovariance functions γ_n *converge towards a limiting autocovariance function* γ *in the sense that*

$$\sum_{h=-\infty}^{\infty} \nu(|h|)\,(1+|h|)\cdot|\gamma_n(h)-\gamma(h)| \longrightarrow 0 \quad \text{as } n\to\infty. \tag{4}$$

Assumption 2.3. *Assumption 2.2 immediately implies that the spectral densities* $f_n(\omega) = (2\pi)^{-1}\sum_{h=-\infty}^{\infty}\gamma_n(h)\,e^{-ih\omega}$, $\omega\in(-\pi,\pi]$, *converge uniformly to a limiting spectral density function* $f(\omega) = (2\pi)^{-1}\sum_{h=-\infty}^{\infty}\gamma(h)\,e^{-ih\omega}$. *Assume that this limit is bounded away from zero, i.e., there exists* $c>0$ *such that* $f(\omega)\geq c$ *for all* $\omega\in(-\pi,\pi]$.

The weighting function ν in Assumption 2.2 allows for more flexibility in our results. Typical examples for ν, fulfilling the conditions stated in Assumption 2.2, are $\nu(h)=1$ for all h, $\nu(h)=(1+|h|)^r$ for some $r>0$ and $\nu(h)=\rho^{a|h|^b}$, $\rho>1$, $a\geq 0$, $0\leq b<1$.

We introduce the following notation: For any fixed norm function ν which fulfils Assumption 2.2, we denote by $\mathcal{C}_\nu$ the space of all complex-valued integrable functions on $(-\pi,\pi]$, with Fourier coefficients $\widetilde{f}_k = (2\pi)^{-1}\int_{-\pi}^{\pi} f(\omega)\,e^{-ik\omega}\,d\omega$ such that

$$\|f\|_\nu := \sum_{k=-\infty}^{\infty}\nu(|k|)\,|\widetilde{f}_k| < \infty.$$

Clearly, due to the norm property $\nu(k)\leq\nu(j)\cdot\nu(|k-j|)$, $\|\cdot\|_\nu$ is a submultiplicative norm on the space $\mathcal{C}_\nu$. If $f\in\mathcal{C}_\nu$ and all Fourier coefficients with negative index vanish, i.e., $\widetilde{f}_k=0$ for all $k<0$, then we say $f\in\mathcal{C}_\nu^+$. Analogously, if $f\in\mathcal{C}_\nu$ and $\widetilde{f}_k=0$ for all $k>0$, then $f\in\mathcal{C}_\nu^-$. Furthermore, for any fixed norm ν, we define the modified norm function $\nu^*\colon \mathbb{N}_0\to[1,\infty)$ via

$$\nu^*(k) := \nu(k)\cdot(1+k), \tag{5}$$

which also fulfils the norm conditions stated in Assumption 2.2.

Since the processes $(X_t^{(n)})$ fulfill Assumptions 2.1–2.3, there exists $n_0\in\mathbb{N}$ such that the spectral densities f_n are bounded from above and away from zero for $n\geq n_0$. Therefore it is known, cf. among others [4], (2.1), and Theorem 6.5 in [10], that the processes $(X_t^{(n)})$ possess one-sided autoregressive representations

$$X_t^{(n)} = \sum_{k=1}^{\infty} a_k^{(n)}\,X_{t-k}^{(n)} + \varepsilon_t^{(n)}, \quad n\geq n_0, \tag{6}$$

where $(\varepsilon_t^{(n)})_{t\in\mathbb{Z}}$ are the innovation processes. It is also known that the summability condition on the autocovariances in Assumption 2.2 carries over to the autoregressive coefficients $(a_k^{(n)})$, i.e.,

$$\sum_{k=1}^{\infty}\nu(|k|)\,(1+|k|)\cdot|a_k^{(n)}| < \infty, \tag{7}$$

cf. [4], Theorem 1.1 and (2.1), and also [11], Theorem 5.5 (for the connection of the autoregressive coefficients to the optimal factor stated in [4]).

Since the autoregressive coefficients in (6) are not observable, one is usually interested in finite-order prediction coefficients. For an arbitrary stationary process $(Y_t)_{t\in\mathbb{Z}}$ with finite second moments and some $p\in\mathbb{N}$, the coefficients $a_1(p),\ldots,a_p(p)$ minimizing the expression

$$E\Big(Y_t - \sum_{k=1}^{p} a_k(p)\,Y_{t-k}\Big)^2$$

are called the prediction coefficients of order p for (Y_t). They are determined by the L^2-projection of Y_t onto its finite past $Y_{t-1},\ldots,Y_{t-p}$ which is given by the solution of the minimization problem above. In the present setting of observations from a triangular array as in (2) it is particularly interesting to look at predictors with p depending on n, i.e., $p = p(n) \leq n-1$ for $(X_t^{(n)})$ since there are only n observations of $(X_t^{(n)})$ at hand. Since we are interested in convergence of the finite-order prediction coefficients towards the infinitely many coefficients $(a_k^{(n)})$ from (6), we need $p(n) \to \infty$ as $n \to \infty$. Therefore, we assume the following:

Assumption 2.4. *Let $(p(n))_{n\in\mathbb{N}}$ be a sequence of positive integers with $p(n) \leq n-1$ for all n and $p(n) \to \infty$ as $n \to \infty$.*

In order to simplify the notation we will suppress the dependence of p on n in the following and denote the prediction coefficients of order p of $(X_t^{(n)})$ by $a_1^{(n)}(p), \ldots, a_p^{(n)}(p)$ which are the minimizers of

$$E\Big(X_t^{(n)} - \sum_{k=1}^{p} a_k^{(n)}(p)\, X_{t-k}^{(n)}\Big)^2.$$

It is well known that the solution of this minimization problem is given by the solution of the Yule–Walker equations, i.e., it holds in the present setting

$$\begin{pmatrix} \gamma_n(0) & \ldots & \gamma_n(p-1) \\ \vdots & \ddots & \vdots \\ \gamma_n(p-1) & \ldots & \gamma_n(0) \end{pmatrix} \begin{pmatrix} a_1^{(n)}(p) \\ \vdots \\ a_p^{(n)}(p) \end{pmatrix} = \begin{pmatrix} \gamma_n(1) \\ \vdots \\ \gamma_n(p) \end{pmatrix}. \tag{8}$$

Under the stated assumptions the finite-order predictors $a_1^{(n)}(p), \ldots, a_p^{(n)}(p)$ and the autoregressive coefficients $(a_k^{(n)})$ from (6) converge to a common limit at a rate that is determined by the growth of the weighting function ν. The precise result is given by the following theorem and is a generalization of Baxter's inequality, cf. [1], Theorem 2.2.

Theorem 2.1. *Let ν be a norm function as defined in Assumption 2.2 and ν^* be the modified norm function as defined in (5). Let $(X_t^{(n)})$ be a sequence of stationary processes fulfilling Assumptions 2.1–2.3 and $(p(n))_{n\in\mathbb{N}}$ a sequence fulfilling Assumption 2.4 with the finite predictor coefficients $a_1^{(n)}(p), \ldots, a_p^{(n)}(p)$ given by (8) and the autoregressive coefficients given by (6). Then, there exist $n_0 \in \mathbb{N}$ and $C < \infty$, both not depending on n, such that*

$$\sum_{k=1}^{p} \nu^*(k)\, |a_k^{(n)}(p) - a_k^{(n)}| \leq C \cdot \sum_{k=p+1}^{\infty} \nu^*(k)\, |a_k^{(n)}|, \quad \forall\, n \geq n_0. \tag{9}$$

This bound remains true if one replaces ν^* with ν on both sides.

Note that the RHS of (9) is summable and therefore converges to zero as $n \to \infty$ due to Assumption 2.4 which yields convergence to zero for the LHS as well.

Remark 2.1. The bound (9) is the strong version of Baxter's inequality for triangular arrays. It is worth mentioning that the last statement of Theorem 2.1 ensures that the weaker version

$$\sum_{k=1}^{p} \nu(k)\, |a_k^{(n)}(p) - a_k^{(n)}| \leq C \cdot \sum_{k=p+1}^{\infty} \nu(k)\, |a_k^{(n)}|, \quad \forall\, n \geq n_0,$$

also holds. Since ν can be chosen to be $\nu(k) = 1$ for all k, this yields an unweighted version of the Baxter inequality.

The proof of Theorem 2.1 can be found in Section 3 and depends heavily on two auxiliary results. In order to state the first auxiliary result we first have to introduce the term *optimal factor*: Each function $g \in \mathcal{C}_{\nu^*}$ which is bounded away from zero can be decomposed as $g(\omega) = A(\omega)\,\overline{A(\omega)}$, where $A \in \mathcal{C}^{+}_{\nu^*}$, $\overline{A} \in \mathcal{C}^{-}_{\nu^*}$ and the complex-valued function A is called the optimal factor of g, cf. Theorem 1.1 in [4]. Here, $\overline{x}$ denotes conjugate transpose of x.

The following result transfers convergence of functions in the $\|\cdot\|_{\nu^*}$-sense to convergence of the respective logarithms and the optimal factors, and might be of its own interest.

Lemma 2.1. *Let ν be a norm function as defined in Assumption 2.2 and ν^* be the modified norm function as defined in (5). Let $f \in \mathcal{C}_{\nu^*}$ and $(f_n)_{n\in\mathbb{N}}$ be a sequence of functions in $\mathcal{C}_{\nu^*}$ with $\|f_n\|_{\nu^*} \le C$ for all $n \in \mathbb{N}$ and some $C < \infty$. Furthermore, let f, f_n be uniformly bounded away from zero, i.e., there exists $c > 0$ such that $f(\omega) \ge c, f_n(\omega) \ge c$ for all $\omega \in (-\pi, \pi]$ and all $n \in \mathbb{N}$. Then, f and f_n can be decomposed as*

$$f(\omega) = A(\omega)\,\overline{A(\omega)}, \quad f_n(\omega) = A_n(\omega)\,\overline{A_n(\omega)}, \quad n \in \mathbb{N}, \tag{10}$$

where A, A_n are the optimal factors defined above and it holds true:

(a) $\log f, \log f_n \in \mathcal{C}_{\nu^*}$ *and the sequence* $\|\log f_n\|_{\nu^*}$ *is bounded.*

Furthermore, if $\|f_n - f\|_{\nu^*} \to 0$, as $n \to \infty$, it follows

(b) $\|\log f_n - \log f\|_{\nu^*} \to 0$.

(c) $\|A_n - A\|_{\nu^*} \to 0$, $\|\overline{A_n} - \overline{A}\|_{\nu^*} \to 0$, $\|A_n^{-1} - A^{-1}\|_{\nu^*} \to 0$, $\left\|\overline{A_n}^{\,-1} - \overline{A}^{\,-1}\right\|_{\nu^*} \to 0$.

With this result we can prove the following lemma which is essential for our main Theorem 2.1 and provides a generalization of Theorem 1.1 of [2].

Lemma 2.2. *Let ν be a norm function as defined in Assumption 2.2 and ν^* be the modified norm function as defined in (5). Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of spectral densities of stationary processes $(X_t^{(n)})_{n\in\mathbb{N}}$ which fulfil Assumptions 2.1–2.3 and $(p(n))_{n\in\mathbb{N}}$ a sequence fulfilling Assumption 2.4, again suppressing the dependence of p on n for simplicity. Furthermore, let $h_1^{(n)}, \ldots, h_p^{(n)}$ be some real-valued coefficients for each $n \in \mathbb{N}$, and let $g_1^{(n)}, \ldots, g_p^{(n)}$ be defined by*

$$g_j^{(n)} := \int_{-\pi}^{\pi} \Big(\sum_{k=1}^{p} h_k^{(n)}\, e^{ik\omega}\Big) f_n(\omega)\, e^{-ij\omega}\, d\omega, \quad \forall\, j = 1, \ldots, p \tag{11}$$

for each $n \in \mathbb{N}$. Then, there exists $n_0 \in \mathbb{N}$ and $M < \infty$, both not depending on the coefficients $h_j^{(n)}, g_j^{(n)}$, and M not depending on n, such that

$$\sum_{k=1}^{p} \nu^*(k)\,|h_k^{(n)}| \le M \cdot \sum_{k=1}^{p} \nu^*(k)\,|g_k^{(n)}|, \quad \forall\, n \ge n_0. \tag{12}$$

The bound (12) remains true if one replaces ν^* with ν on both sides.

3. PROOFS

Proof of Theorem 2.1. The autoregressive coefficients from (6) are determined by the L^2-projection of $X_t^{(n)}$ onto its infinite past $\overline{sp}\{X_{t-k}^{(n)},\, k \in \mathbb{N}\}$. Therefore, the coefficients satisfy the system of equations

$$\begin{pmatrix} \gamma_n(0) & \gamma_n(1) & \gamma_n(2) & \dots \\ \gamma_n(1) & \gamma_n(0) & \gamma_n(1) & \dots \\ \gamma_n(2) & \gamma_n(1) & \gamma_n(0) & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} a_1^{(n)} \\ a_2^{(n)} \\ \vdots \end{pmatrix} = \begin{pmatrix} \gamma_n(1) \\ \gamma_n(2) \\ \vdots \end{pmatrix}, \tag{13}$$

which can be interpreted as the "limit" of the Yule–Walker equations (8), cf. Theorem 6.6 in [10]. Comparing the first p rows of this system with (8) it is obvious that the right-hand sides are identical which yields

$$\sum_{k=1}^{p} a_k^{(n)}(p)\, \gamma_n(k-j) = \sum_{k=1}^{\infty} a_k^{(n)}\, \gamma_n(k-j), \quad \forall\, j = 1, \dots, p,$$

because $\gamma_n(h) = \gamma_n(-h)$ for all n, h. Expressing the autocovariances in terms of the spectral densities via $\gamma_n(h) = \int_{-\pi}^{\pi} f_n(\omega)\, e^{ih\omega}\, d\omega$, these equations can be written as

$$\int_{-\pi}^{\pi} \Big(\sum_{k=1}^{p} a_k^{(n)}(p)\, e^{ik\omega} \Big) f_n(\omega)\, e^{-ij\omega}\, d\omega = \int_{-\pi}^{\pi} \Big(\sum_{k=1}^{\infty} a_k^{(n)}\, e^{ik\omega} \Big) f_n(\omega)\, e^{-ij\omega}\, d\omega, \quad \forall\, j = 1, \dots, p.$$

Splitting up the sum on the right-hand side and using linearity of the integral, this system is equivalent to

$$\int_{-\pi}^{\pi} \Big(\sum_{k=1}^{p} (a_k^{(n)}(p) - a_k^{(n)})\, e^{ik\omega} \Big) f_n(\omega)\, e^{-ij\omega}\, d\omega = g_j^{(n)}, \quad \forall\, j = 1, \dots, p, \tag{14}$$

where we used the abbreviation $g_j^{(n)} := \int_{-\pi}^{\pi} (\sum_{k=p+1}^{\infty} a_k^{(n)}\, e^{ik\omega}) f_n(\omega)\, e^{-ij\omega}\, d\omega$. We can now apply Lemma 2.2 to the system (14) which ensures that there exists $n_0 \in \mathbb{N}$ and $M < \infty$ such that

$$\sum_{k=1}^{p} \nu^*(k)\, |a_k^{(n)}(p) - a_k^{(n)}| \le M \cdot \sum_{k=1}^{p} \nu^*(k)\, |g_k^{(n)}|, \quad \forall\, n \ge n_0. \tag{15}$$

Inserting the definition of $g_k^{(n)}$ and again using $\gamma_n(h) = \int_{-\pi}^{\pi} f_n(\omega)\, e^{ih\omega}\, d\omega$, the right-hand side of (15) can be bounded by

$$M \cdot \sum_{k=1}^{p} \nu^*(k) \Big| \int_{-\pi}^{\pi} \Big(\sum_{j=p+1}^{\infty} a_j^{(n)}\, e^{ij\omega} \Big) f_n(\omega)\, e^{-ik\omega}\, d\omega \Big| \le M \cdot \sum_{k=1}^{p} \sum_{j=p+1}^{\infty} \nu^*(k)\, |a_j^{(n)}|\, |\gamma_n(j-k)|. \tag{16}$$

Using the norm condition $\nu^*(k) \le \nu^*(j) \cdot \nu^*(|j-k|)$ and (3), the right-hand side of (16) can further be bounded by

$$\begin{aligned} & M \cdot \sum_{j=p+1}^{\infty} \nu^*(j)\, |a_j^{(n)}| \sum_{k=1}^{p} \nu^*(|j-k|)\, |\gamma_n(j-k)| \\ & \quad \le M \cdot \sum_{k=-\infty}^{\infty} \nu^*(|k|)\, |\gamma_n(k)| \cdot \sum_{j=p+1}^{\infty} \nu^*(j)\, |a_j^{(n)}| \le MC \cdot \sum_{j=p+1}^{\infty} \nu^*(j)\, |a_j^{(n)}|. \end{aligned} \tag{17}$$

Altogether, with $C' := MC$ (not depending on n, cf. (3)), we get from (15)–(17)

$$\sum_{k=1}^{p} \nu^*(k)\, |a_k^{(n)}(p) - a_k^{(n)}| \le C' \cdot \sum_{k=p+1}^{\infty} \nu^*(k)\, |a_k^{(n)}|, \quad \forall\, n \ge n_0.$$

which gives (9). The entire proof remains true if one replaces ν^* with ν, even with the same constant C' because $\nu \leq \nu^*$. This completes the proof. $\square$

Proof of Lemma 2.1 (a) and (b). We will prove assertion (b) first and then get (a) as a by-product at the end of the proof. Per assumption, f and each f_n are equal to their absolutely convergent Fourier series, i.e.,

$$f(\omega) = \sum_{k=-\infty}^{\infty} b_k\, e^{ik\omega}, \qquad f_n(\omega) = \sum_{k=-\infty}^{\infty} b_k^{(n)}\, e^{ik\omega},$$

say. Since

$$\sum_{k=-\infty}^{\infty} |(ik)b_k\, e^{ik\omega}| \leq \sum_{k=-\infty}^{\infty} |k|\,|b_k| \leq \sum_{k=-\infty}^{\infty} \nu^*(|k|)\,|b_k| = \|f\|_{\nu^*} \leq C < \infty$$

is fulfilled (the same being true for $b_k^{(n)}$ instead of b_k), it is well known that f and f_n are continuously differentiable and the derivatives have absolutely convergent Fourier series

$$f'(\omega) = \sum_{k=-\infty}^{\infty} (ik)b_k\, e^{ik\omega}, \qquad f_n'(\omega) = \sum_{k=-\infty}^{\infty} (ik)b_k^{(n)}\, e^{ik\omega},$$

respectively, as can be obtained from termwise differentiation. Moreover, it is easy to see that $f_n' \in \mathcal{C}_\nu$ because $\|f_n'\|_\nu \leq \|f_n\|_{\nu^*} < \infty$. Also, $\|f_n\|_\nu \leq \|f_n\|_{\nu^*} \leq C < \infty$ and f_n is bounded away from zero. Therefore, the weighted version of Wiener's lemma, cf. Theorem 6.2 from [6] (note that ν and ν^* fulfil the GRS-condition required there), implies $(1/f_n) \in \mathcal{C}_\nu$. Lemma 1 in §2 of [5] shows that on the set of all functions $g \in \mathcal{C}_\nu$ with $(1/g) \in \mathcal{C}_\nu$, taking the inverse of any function is a continuous operation in the sense of the topology implied by $\|\cdot\|_\nu$. Therefore, since $f_n \to f$ in $\mathcal{C}_\nu$, we also get $(1/f_n) \to (1/f)$ in $\mathcal{C}_\nu$. Thus, we can choose $C' < \infty$ such that

$$\Big\| \frac{1}{f} \Big\|_\nu \leq C', \qquad \Big\| \frac{1}{f_n} \Big\|_\nu \leq C' \quad \forall\, n \in \mathbb{N}. \tag{18}$$

Submultiplicativity of $\|\cdot\|_\nu$ gives

$$\Big\| \frac{f_n'}{f_n} \Big\|_\nu \leq \| f_n' \|_\nu \cdot \Big\| \frac{1}{f_n} \Big\|_\nu < \infty,$$

as well as $(f'/f) \in \mathcal{C}_\nu$. The strategy is to infer $\|\log f_n - \log f\|_\nu \to 0$ from

$$\Big\| \frac{f_n'}{f_n} - \frac{f'}{f} \Big\|_\nu \longrightarrow 0, \quad n \to \infty. \tag{19}$$

Hence, we prove (19) first. We have

$$\Big\| \frac{f_n'}{f_n} - \frac{f'}{f} \Big\|_\nu = \Big\| \frac{f_n' f - f_n f'}{f_n f} \Big\|_\nu \leq \| f_n' f - f_n f' \|_\nu \cdot \Big\| \frac{1}{f} \Big\|_\nu \cdot \Big\| \frac{1}{f_n} \Big\|_\nu .$$

Thus, (18) implies (19), if we can show

$$\| f_n' f - f_n f' \|_\nu \longrightarrow 0, \quad n \to \infty. \tag{20}$$

The Fourier series of $f_n' f$ and $f_n f'$ can be obtained from straightforward multiplication of the Fourier series from above which yields

$$f_n'(\omega) f(\omega) = \sum_{k=-\infty}^{\infty} \Big(\sum_{j=-\infty}^{\infty} ij\, b_j^{(n)} b_{k-j} \Big) e^{ik\omega}, \qquad f_n(\omega) f'(\omega) = \sum_{k=-\infty}^{\infty} \Big(\sum_{j=-\infty}^{\infty} ij\, b_j b_{k-j}^{(n)} \Big) e^{ik\omega}.$$

Hence, using $\nu(|k|) \le \nu(|j|)\,\nu(\big||k|-|j|\big|) \le \nu(|j|)\,\nu(|k-j|)$ (recall that ν is nondecreasing), as well as $\nu(|j|)\,|j| \le \nu^*(|j|)$ and $\nu(|k-j|) \le \nu^*(|k-j|)$, we can derive

$$\begin{aligned}\|\, f_n' f - f_n f' \,\|_\nu &= \sum_{k=-\infty}^{\infty} \nu(|k|) \Big| \sum_{j=-\infty}^{\infty} ij\,(b_j^{(n)} b_{k-j} - b_j b_{k-j}^{(n)}) \Big| \\ &\le \sum_{k,j=-\infty}^{\infty} \nu(|k|)\,|j|\,\big|b_j(b_{k-j} - b_{k-j}^{(n)}) + (b_j^{(n)} - b_j) b_{k-j}\big| \\ &\le \sum_{k,j=-\infty}^{\infty} \nu^*(|j|)\,\nu^*(|k-j|)\,\big(|b_j|\,\big|b_{k-j} - b_{k-j}^{(n)}\big| + \big|b_j^{(n)} - b_j\big|\,|b_{k-j}|\big) \\ &= 2 \cdot \|f\|_{\nu^*} \cdot \|f_n - f\|_{\nu^*}\end{aligned}$$

which converges to zero as $n \to \infty$ according to the assumptions of Lemma 2.1 (b). This proves (20) and therefore also (19).

Now, let $\beta_k^{(n)}$ and β_k be the Fourier coefficients of f_n'/f_n and f'/f, respectively. In fact, it holds $\beta_0^{(n)} = 0$ since f_n'/f_n is the derivative of $\log f_n(\omega)$ and therefore, per definition,

$$\beta_0^{(n)} = (2\pi)^{-1}(\log f_n(\pi) - \log f_n(-\pi)) = 0.$$

Thus, the Fourier series are given by

$$\frac{f_n'(\omega)}{f_n(\omega)} = \sum_{k\in\mathbb{Z}\setminus\{0\}} \beta_k^{(n)}\, e^{ik\omega}, \qquad \frac{f'(\omega)}{f(\omega)} = \sum_{k\in\mathbb{Z}\setminus\{0\}} \beta_k\, e^{ik\omega}.$$

We have already shown that these series are absolutely convergent. Hence, it is well known that the Fourier series of their antiderivatives, $\log f_n(\omega)$ and $\log f(\omega)$, are given by

$$\log f_n(\omega) = \sum_{k\in\mathbb{Z}\setminus\{0\}} \frac{\beta_k^{(n)}}{ik}\, e^{ik\omega}, \qquad \log f(\omega) = \sum_{k\in\mathbb{Z}\setminus\{0\}} \frac{\beta_k}{ik}\, e^{ik\omega}, \tag{21}$$

and it follows

$$\|\log f_n - \log f\|_{\nu^*} = \sum_{k\in\mathbb{Z}\setminus\{0\}} \nu(|k|)\,\frac{1+|k|}{|k|}\big|\beta_k^{(n)} - \beta_k\big| \le 2 \cdot \Big\|\, \frac{f_n'}{f_n} - \frac{f'}{f} \,\Big\|_\nu,$$

which converges to zero due to (19). This completes the proof for Lemma 2.1](b). From this calculation it can also be seen that $\|\log f\|_{\nu^*} < \infty$ and $\|\log f_n\|_{\nu^*} \le \|f_n'\|_\nu \cdot \|\,(1/f_n)\,\|_\nu \le \widetilde{C}$ for some $\widetilde{C} < \infty$, even if $\|f_n - f\|_{\nu^*} \to 0$ is not fulfilled, which gives assertion (a) of Lemma 2.1. $\square$

Proof of Lemma 2.1 (*c*). We first show $\|A_n - A\|_{\nu^*} \to 0$ as $n \to \infty$. Lemma 2.1 (a) guarantees that $\log f,\, \log f_n \in \mathcal{C}_{\nu^*}$, i.e., the functions are equal to their absolutely convergent Fourier series given in (21). In the following we abbreviate the Fourier coefficients by $d_k^{(n)} := \beta_k^{(n)}/ik$ and $d_k := \beta_k/ik$. Also, we get from Lemma 2.1 (b) $\|\log f_n - \log f\|_{\nu^*} \to 0$ as $n \to \infty$, i.e.,

$$\sum_{k\in\mathbb{Z}\setminus\{0\}} \nu^*(|k|)\,|d_k^{(n)} - d_k| \longrightarrow 0 \quad \text{as } n \to \infty. \tag{22}$$

Now define

$$B(\omega) := \sum_{k=1}^{\infty} d_k\, e^{ik\omega}, \qquad B_n(\omega) = \sum_{k=1}^{\infty} d_k^{(n)}\, e^{ik\omega}.$$

The optimal factors A, A_n are then given by $A(\omega) = \exp(B(\omega))$, $A_n(\omega) = \exp(B_n(\omega))$, as can be seen from Chapter 2 in [11]. Therefore, $A(\omega), A_n(\omega) \neq 0$ for all ω and

$$A^{-1}(\omega) = \exp(-B(\omega)), \qquad A_n^{-1}(\omega) = \exp(-B_n(\omega)).$$

Since $A, A_n \in \mathcal{C}^+_{\nu^*}$, the weighted version of Wiener's lemma implies $A^{-1}, A_n^{-1} \in \mathcal{C}_{\nu^*}$, cf. Theorem 6.2 in [6]. In this case we also get $A^{-1}, A_n^{-1} \in \mathcal{C}^+_{\nu^*}$ because the Fourier series of A^{-1} can be derived using the fact that the exponential function is equal to its (absolutely convergent) power series expansion

$$A^{-1}(\omega) = \exp\left(-B(\omega)\right) = \sum_{j=0}^{\infty} \frac{1}{j!}\Big(\sum_{k=1}^{\infty}(-d_k)\, e^{ik\omega}\Big)^j.$$

Expanding the right-hand side gives the Fourier series of A^{-1} and it can easily be seen that it has the form $A^{-1}(\omega) = \sum_{k=0}^{\infty} c_k\, e^{ik\omega}$, for some c_k, i.e., the Fourier coefficients with negative index vanish. With the same arguments we get $A_n^{-1} \in \mathcal{C}^+_{\nu^*}$.

From (22) we immediately get

$$\|B_n - B\|_{\nu^*} \longrightarrow 0 \quad \text{as } n \to \infty. \tag{23}$$

We will now carry this result over to $\|A_n - A\|_{\nu^*}$. Using the power series expansion of the exponential function we get

$$\|A_n - A\|_{\nu^*} = \|\exp(B_n) - \exp(B)\|_{\nu^*} \leq \sum_{k=0}^{\infty} \frac{1}{k!} \left\| B_n^k - B^k \right\|_{\nu^*},$$

which converges to zero according to Lebesgue's dominated convergence theorem if we can show that

$$\| B_n^k - B^k \|_{\nu^*} \longrightarrow 0 \text{ as } n \to \infty \text{ for each } k \in \mathbb{N}, \tag{24}$$

$$\| B_n^k - B^k \|_{\nu^*} \leq C^k \text{ for some } C < \infty \text{ and } \forall\, n, k \in \mathbb{N}. \tag{25}$$

Of course, (24) follows from (23) inductively because

$$\begin{aligned}\| B_n^k - B^k \|_{\nu^*} &= \| (B_n - B)B^{k-1} + B_n(B_n^{k-1} - B^{k-1}) \|_{\nu^*} \\ &\leq \| B_n - B \|_{\nu^*} \cdot \| B \|_{\nu^*}^{k-1} + \| B_n \|_{\nu^*} \cdot \| B_n^{k-1} - B^{k-1} \|_{\nu^*},\end{aligned}$$

as $\| B_n \|_{\nu^*} \leq \| \log f_n \|_{\nu^*} \leq \widetilde{C}$ for some $\widetilde{C} < \infty$ and all n, cf. Lemma 2.1 (a). Furthermore,

$$\| B_n^k - B^k \|_{\nu^*} \leq \| B_n \|_{\nu^*}^k + \| B \|_{\nu^*}^k \leq (\widetilde{C} + \| B \|_{\nu^*})^k$$

yields (25) with $C = \widetilde{C} + \| B \|_{\nu^*} < \infty$. Therefore, it holds

$$\|A_n - A\|_{\nu^*} \longrightarrow 0 \quad \text{as } n \to \infty. \tag{26}$$

Since $\overline{A(\omega)} = \exp(\overline{B(\omega)})$, $\overline{A_n(\omega)} = \exp(\overline{B_n(\omega)})$ and

$$\overline{B(\omega)} := \sum_{k=1}^{\infty} d_k\, e^{-ik\omega}, \qquad \overline{B_n(\omega)} = \sum_{k=1}^{\infty} d_k^{(n)}\, e^{-ik\omega},$$

due to $d_k, d_n^{(n)} \in \mathbb{R}$ ($\log f, \log f_n$ are even functions), the same argument as above delivers $\overline{A}, \overline{A_n} \in \mathcal{C}^-_{\nu^*}$ and $\|\overline{A_n} - \overline{A}\|_{\nu^*} \to 0$. Also, exactly as for A^{-1}, A_n^{-1} above we get $\overline{A}^{\,-1}, \overline{A_n}^{\,-1} \in \mathcal{C}^-_{\nu^*}$ and

$$\|A_n^{-1} - A^{-1}\|_{\nu^*} \to 0, \qquad \left\| \overline{A_n}^{\,-1} - \overline{A}^{\,-1} \right\|_{\nu^*} \to 0 \quad \text{as } n \to \infty.$$

This completes the proof. □

Proof of Lemma 2.2. For each $n \in \mathbb{N}$ we define the complex-valued functions h_n, g_n to be the finite Fourier series

$$h_n(\omega) = \sum_{k=1}^{p} h_k^{(n)}\, e^{ik\omega}, \qquad g_n(\omega) = \sum_{k=1}^{p} g_k^{(n)}\, e^{ik\omega}. \tag{27}$$

Since f, f_n fulfil the assumptions of Lemma 2.1, the functions can be decomposed as in (11). We can observe that $g_1^{(n)}, \dots, g_p^{(n)}$ are exactly the first p Fourier coefficients with positive index of the function $2\pi\, h_n f_n$. To be more precise, the Fourier series of $2\pi\, h_n f_n$ is given by

$$2\pi\, h_n(\omega) f_n(\omega) = \sum_{k=-\infty}^{0} \widetilde{G}_k^{(n)}\, e^{ik\omega} + \sum_{k=1}^{p} g_k^{(n)}\, e^{ik\omega} + \sum_{k=p+1}^{\infty} \widetilde{G}_k^{(n)}\, e^{ik\omega}$$
$$=: G_1^{(n)}(\omega) + g_n(\omega) + G_2^{(n)}(\omega), \tag{28}$$

for some coefficients $\widetilde{G}_k^{(n)}$. Using (10), and suppressing dependence on ω, we can write

$$2\pi\, h_n A_n = G_1^{(n)} \cdot \overline{A_n}^{\,-1} + g_n \cdot \overline{A_n}^{\,-1} + G_2^{(n)} \cdot \overline{A_n}^{\,-1}. \tag{29}$$

We introduce the following notation: For an arbitrary function z, represented by its Fourier series $z(\omega) = \sum_{k=-\infty}^{\infty} \widetilde{z}_k\, e^{ik\omega}$, we define the following truncated Fourier series:

$$(z)^- := \sum_{k=-\infty}^{0} \widetilde{z}_k\, e^{ik\omega}, \qquad (z)^{p+} := \sum_{k=p+1}^{\infty} \widetilde{z}_k\, e^{ik\omega}, \qquad (z)^{p-} := \sum_{k=-\infty}^{-p-1} \widetilde{z}_k\, e^{ik\omega},$$

which will be used in the remainder of this proof. Since $A_n \in \mathcal{C}_{\nu^*}^+$, i.e., $A_n(\omega) = \sum_{k=0}^{\infty} \widetilde{A}_k^{(n)}\, e^{ik\omega}$ for some $\widetilde{A}_k^{(n)}$, the Fourier series of $2\pi\, h_n A_n$ can be obtained by straightforward multiplication of $\sum_{k=0}^{\infty} \widetilde{A}_k^{(n)}\, e^{ik\omega}$ with h_n as given in (27), which shows that $2\pi\, h_n A_n$ has only Fourier coefficients with strictly positive index, i.e., $(2\pi\, h_n A_n)^- = 0$. We now get from (29)

$$0 = (2\pi\, h_n A_n)^- = \big(G_1^{(n)} \cdot \overline{A_n}^{\,-1}\big)^- + \big(g_n \cdot \overline{A_n}^{\,-1}\big)^- + \big(G_2^{(n)} \cdot \overline{A_n}^{\,-1}\big)^-$$
$$= G_1^{(n)} \cdot \overline{A_n}^{\,-1} + \big(g_n \cdot \overline{A_n}^{\,-1}\big)^- + \big(G_2^{(n)} \cdot \overline{A_n}^{\,-1}\big)^-, \tag{30}$$

because multiplication of the series for $G_1^{(n)}$ and $\overline{A_n}^{\,-1} \in \mathcal{C}_{\nu^*}^-$ shows that all Fourier coefficients of $G_1^{(n)} \cdot \overline{A_n}^{\,-1}$ with positive index vanish. Now (30) yields

$$-G_1^{(n)}\overline{A_n}^{\,-1} = \big(g_n\overline{A_n}^{\,-1}\big)^- + \big(G_2^{(n)}\overline{A_n}^{\,-1}\big)^- = \big(g_n\overline{A_n}^{\,-1}\big)^- + \big(G_2^{(n)}(\overline{A_n}^{\,-1})^{p-}\big)^-. \tag{31}$$

The last equation holds because $\overline{A_n}^{\,-1} - (\overline{A_n}^{\,-1})^{p-} = \sum_{k=-p}^{0} \alpha_k^{(n)}\, e^{ik\omega}$ for some $\alpha_k^{(n)}$, and therefore its product with $G_2^{(n)}$ has only Fourier coefficients with strictly positive index. Applying the $\mathcal{C}_{\nu^*}$-norm to (31), and using the submultiplicativity and the obvious relation $\|(z)^-\|_{\nu^*} \le \|z\|_{\nu^*}$, we get

$$\|G_1^{(n)}\overline{A_n}^{\,-1}\|_{\nu^*} \le \|\overline{A_n}^{\,-1}\|_{\nu^*} \cdot \|g_n\|_{\nu^*} + \|G_2^{(n)} A_n^{-1}\|_{\nu^*} \cdot \|A_n(\overline{A_n}^{\,-1})^{p-}\|_{\nu^*}. \tag{32}$$

From Lemma 2.1 (c) we deduce $\|A_n\|_{\nu^*} \longrightarrow \|A\|_{\nu^*}$ and $\|\overline{A_n}^{\,-1}\|_{\nu^*} \longrightarrow \|\overline{A}^{\,-1}\|_{\nu^*}$ as $n \to \infty$. In particular, the sequences $\|A_n\|_{\nu^*}, \|\overline{A_n}^{\,-1}\|_{\nu^*}$ are bounded. Denoting the Fourier coefficients of $\overline{A}^{\,-1}$, $\overline{A_n}^{\,-1}$ by α_k, $\alpha_k^{(n)}$, respectively, it also holds

$$\|A_n(\overline{A_n}^{\,-1})^{p-}\|_{\nu^*} \le \|A_n\|_{\nu^*} \cdot \big(\|(\overline{A}^{\,-1})^{p-}\|_{\nu^*} + \|(\overline{A_n}^{\,-1} - \overline{A}^{\,-1})^{p-}\|_{\nu^*}\big)$$
$$= \|A_n\|_{\nu^*} \cdot \Big(\sum_{k=-\infty}^{-p-1} \nu^*(|k|)\, |\alpha_k| + \sum_{k=-\infty}^{-p-1} \nu^*(|k|)\, |\alpha_k^{(n)} - \alpha_k|\Big)$$
$$\longrightarrow 0 \quad \text{as } n \to \infty,$$

because of $p = p(n) \to \infty$, Lemma 2.1 (c) and the boundedness of $\|A_n\|_{\nu^*}$. Therefore, for arbitrary $\varepsilon \in (0, 1)$, we can obviously find $n_1 \in \mathbb{N}$ and $C < \infty$ such that $\|\overline{A_n}^{\,-1}\|_{\nu^*} \le C$ and $\|A_n(\overline{A_n}^{\,-1})^{p-}\|_{\nu^*} \le \varepsilon$ for all $n \ge n_1$. Then, (32) becomes

$$\|G_1^{(n)}\overline{A_n}^{\,-1}\|_{\nu^*} \le C \cdot \|g_n\|_{\nu^*} + \varepsilon \cdot \|G_2^{(n)} A_n^{-1}\|_{\nu^*}, \quad n \ge n_1. \tag{33}$$

In the next part of the proof we start again by using decomposition (10) in (28), but this time multiplying both sides with A_n^{-1} instead of $\overline{A_n}^{\,-1}$. Analogously to (29) this yields

$$2\pi\, h_n\overline{A_n} = G_1^{(n)} \cdot A_n^{-1} + g_n \cdot A_n^{-1} + G_2^{(n)} \cdot A_n^{-1}.$$

This time around, it holds $(2\pi\, h_n\overline{A_n})^{p+} = 0$ and we get

$$0 = (2\pi\, h_n\overline{A_n})^{p+} = \big(G_1^{(n)} \cdot A_n^{-1}\big)^{p+} + \big(g_n \cdot A_n^{-1}\big)^{p+} + G_2^{(n)} \cdot A_n^{-1},$$

because all Fourier coefficients of $G_2^{(n)} \cdot A_n^{-1}$ vanish except for those with index $p+1$ or larger. Analogously as from (30) to (32), this yields

$$-G_2^{(n)}A_n^{-1} = \big(G_1^{(n)}A_n^{-1}\big)^{p+} + \big(g_nA_n^{-1}\big)^{p+} = \big(G_1^{(n)}(A_n^{-1})^{p+}\big)^{p+} + \big(g_nA_n^{-1}\big)^{p+}$$

and finally

$$\|G_2^{(n)}A_n^{-1}\|_{\nu^*} \le \|G_1^{(n)}\overline{A_n}^{\,-1}\|_{\nu^*} \cdot \|\overline{A_n}\,(A_n^{-1})^{p+}\|_{\nu^*} + \|A_n^{-1}\|_{\nu^*} \cdot \|g_n\|_{\nu^*}.$$

As above, we can show that $\|A_n^{-1}\|_{\nu^*}$ is bounded and $\|\overline{A_n}\,(A_n^{-1})^{p+}\|_{\nu^*} \longrightarrow 0$ as $n \to \infty$. Hence, for the same $\varepsilon \in (0,1)$ as in (33), we can find $C' < \infty$ and $n_2 \ge n_1$ such that

$$\|G_2^{(n)}A_n^{-1}\|_{\nu^*} \le \varepsilon \cdot \|G_1^{(n)}\overline{A_n}^{\,-1}\|_{\nu^*} + C' \cdot \|g_n\|_{\nu^*}, \quad n \ge n_2. \tag{34}$$

Note that for $n \ge n_2$ both inequalities (33) and (34) hold and we can insert the latter into the former to derive

$$\|G_1^{(n)}\overline{A_n}^{\,-1}\|_{\nu^*} \le C \cdot \|g_n\|_{\nu^*} + \varepsilon C' \cdot \|g_n\|_{\nu^*} + \varepsilon^2 \cdot \|G_1^{(n)}\overline{A_n}^{\,-1}\|_{\nu^*}, \quad n \ge n_2,$$

which yields

$$\|G_1^{(n)}\overline{A_n}^{\,-1}\|_{\nu^*} \le \frac{C + \varepsilon C'}{1-\varepsilon^2} \cdot \|g_n\|_{\nu^*}, \quad n \ge n_2. \tag{35}$$

The other way around, inserting (34) into (33) leads to

$$\|G_2^{(n)}A_n^{-1}\|_{\nu^*} \le \frac{C' + \varepsilon C}{1-\varepsilon^2} \cdot \|g_n\|_{\nu^*}, \quad n \ge n_2. \tag{36}$$

From (29) we get that $2\pi\,\|h_n\|_{\nu^*}$ can be bounded by

$$\|G_1^{(n)}\overline{A_n}^{\,-1}\|_{\nu^*}\|A_n^{-1}\|_{\nu^*} + \|g_n\|_{\nu^*}\|\overline{A_n}^{\,-1}\|_{\nu^*}\|A_n^{-1}\|_{\nu^*} + \|G_2^{(n)}A_n^{-1}\|_{\nu^*}\|\overline{A_n}^{\,-1}\|_{\nu^*}.$$

Since $\|A_n^{-1}\|_{\nu^*} \longrightarrow \|A^{-1}\|_{\nu^*}$, $\|\overline{A_n}^{\,-1}\|_{\nu^*} \longrightarrow \|\overline{A}^{\,-1}\|_{\nu^*}$, we can find $n_0 \ge n_2$ such that $\|A_n^{-1}\|_{\nu^*} \le 2\,\|A^{-1}\|_{\nu^*}$, $\|\overline{A_n}^{\,-1}\|_{\nu^*} \le 2\,\|\overline{A}^{\,-1}\|_{\nu^*}$ for all $n \ge n_0$. Thus, inserting (35) and (36) we get

$$\|h_n\|_{\nu^*} \le M_{\nu^*} \cdot \|g_n\|_{\nu^*}, \quad n \ge n_0,$$

where

$$M_{\nu^*} = \frac{1}{\pi}\Big(\|A^{-1}\|_{\nu^*} \cdot \frac{C+\varepsilon C'}{1-\varepsilon^2} + \|\overline{A}^{\,-1}\|_{\nu^*} \cdot \frac{C'+\varepsilon C}{1-\varepsilon^2} + 2\,\|A^{-1}\|_{\nu^*}\,\|\overline{A}^{\,-1}\|_{\nu^*}\Big)$$

does not depend on n. Since $\|g\|_\nu \le \|g\|_{\nu^*}$ for all functions g, the entire proof from (26) onwards remains valid if one replaces $\|\cdot\|_{\nu^*}$ with $\|\cdot\|_\nu$. This gives

$$\|h_n\|_\nu \le M_\nu \cdot \|g_n\|_\nu, \quad n \ge n_0,$$

where M_ν again does not depend on n and defining $M := \max\{M_{\nu^*}, M_\nu\}$ completes the proof. □

REFERENCES

1. G. Baxter, "An Asymptotic Result for the Finite Predictor", Math. Scand. **10**, 137–144 (1962).
2. G. Baxter, "A Norm Inequality for a Finite-Section Wiener–Hopf Equation", Illinois J. Math. **7**, 97–103 (1963).
3. P. Brockwell and R. Davis, *Time Series: Theory and Methods* (Springer, New York, 1991).
4. R. Cheng and M. Pourahmadi, M. "Baxter's Inequality and Convergence of Finite Predictors of Multivariate Stochastic Processes", Probab. Theory Rel. Fields **95**, 115–124 (1993).
5. I. Gel'fand, D. Raikov, and G. Shilov, *Commutative Normed Rings* (Chelsea Publishing Co., New York, 1964).
6. K. Gröchenig, "Weight Functions in Time-Frequency Analysis", in: *Fields Institute Comm.*, Vol. 52: *Pseudodifferential Operators: Partial Differential Equations and Time-Frequency Analysis*, Ed. by L. Rodino et al. (2007), pp. 343–366.
7. E. J. Hannan and M. Deistler, *The Statistical Theory of Linear Systems*, in *Wiley Series in Probab. and Math. Statist.* (1998).
8. J.-P. Kreiss and E. Paparoditis, "Bootstrap Methods for Dependent Data: A Review", J. Korean Statist. Soc. **40**, 357–378 (2011).
9. J. P. Kreiss, E. Paparoditis, and D. Politis, "On the Range of Validity of the Autoregressive Sieve Bootstrap", Ann. Statist. **39**, 2103–2130 (2011).
10. M. Pourahmadi, *Foundations of Time Series Analysis and Prediction Theory* (Wiley, New York, 2001).
11. N. Wiener and P. Masani, "The Prediction Theory of Multivariate Stochastic Processes, II", Acta Mathematica **99**, 93–137 (1958).