An effective criterion for a stable factorisation of strictly non-singular 2 × 2 matrix functions

Adukova, N. V.; Adukov, V. M.; Mishuris, G.

Published in:
Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences

DOI:
10.1098/rspa.2024.0116

Publication date:
2024

Citation for published version (APA):
Adukova, N. V., Adukov, V. M., & Mishuris, G. (2024). An effective criterion for a stable factorisation of strictly non-singular 2 × 2 matrix functions: Use of the ExactMPF package. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *480*(2299), Article 20240116. https://doi.org/10.1098/rspa.2024.0116

PROCEEDINGS A

royalsocietypublishing.org/journal/rspa

Research

Cite this article: Adukova NV, Adukov VM, Mishuris G. 2024 An effective criterion for a stable factorisation of strictly non-singular 2 × 2 matrix functions: use of the ExactMPF package. *Proc. R. Soc. A* **480**: 20240116.
https://doi.org/10.1098/rspa.2024.0116

Received: 11 February 2024
Accepted: 10 June 2024

Subject Category:
Mathematics

Subject Areas:
analysis

Keywords:
Wiener–Hopf factorisation, strictly non-singular matrix functions, canonical factorisation, stable factorisation, explicit conditions for stable factorisation, approximate factorisation

Author for correspondence:
N. V. Adukova
e-mail: naa53@aber.ac.uk

One contribution to a special feature "Mathematical theory and applications of multiple wave scattering" organised by guest editors Luke G. Bennetts, Michael H. Meylan, Malte A. Peter, Valerie J. Pinfield and Olga Umnova.

Electronic supplementary material is available online at https://doi.org/10.6084/m9.figshare.c.7318642.

THE ROYAL SOCIETY PUBLISHING

An effective criterion for a stable factorisation of strictly non-singular 2×2 matrix functions: use of the ExactMPF package

N. V. Adukova[1], V. M. Adukov[2] and G. Mishuris[1]

[1]Department of Mathematics, Aberystwyth University, Aberystwyth SY23 3BZ, UK
[2]Institute of Natural Sciences and Mathematics, South Ural State University, Chelyabinsk 454080, Russia

NVA, 0000-0002-6758-6101; VMA, 0000-0002-3594-3565; GM, 0000-0003-2565-1961

In this paper, we propose a method to factorise arbitrary strictly non-singular 2×2 matrix functions, allowing for stable factorisation. For this purpose, we use the ExactMPF package working within the Maple environment previously developed by the authors and perform an *exact* factorisation of a non-singular polynomial matrix function. A crucial point in the present analysis is the evaluation of a stability region of the canonical factorisation of the polynomial matrix functions. This, in turn, allows us to propose a sufficient condition for the given matrix function admitting stable factorisation.

1. Introduction and outline of the main result

Factorisation of matrix functions is a challenging task when it comes to its practical implementation. Indeed, any available numerical method is only justified under the condition that partial indices of the matrix function in question are stable [1]. On the other hand, there is no general explicit criterion allowing to determine those indices [2].

Factorisation of matrix functions plays the central role in various applications, e.g. integration of nonlinear differential equations by the inverse scattering method [3,4], in the theory of the Markushevich problem on the unit circle [5] and in scattering and diffraction of elastic waves in bodies with obstacles [6–10].

Most of the problems discussed in applications refer to vectorial Wiener–Hopf problems with 2×2 matrix functions

$$A(t) = \begin{pmatrix} a_{11}(t) & a_{12}(t) \\ a_{21}(t) & a_{22}(t) \end{pmatrix}. \tag{1.1}$$

It is known [11–14] that a continuous invertible matrix function $A(t)$ admits a *right Wiener–Hopf factorisation*

$$A(t) = A_-(t)D(t)A_+(t), \ t \in \mathbb{T}, \tag{1.2}$$

where $A_+(t)$ and $A_-(t)$ are continuous matrix functions on $\mathbb{T}$ that can be extended analytically to the domains $\mathcal{D}_+ = \{z \in \mathbb{C}\colon |z| < 1\}$ and $\mathcal{D}_- = \{z \in \mathbb{C} \cup \{\infty\}\colon |z| > 1\}$ and are invertible in the respective domain. Factor $D(t)$ is the diagonal matrix function $D(t) = \mathrm{diag}[t^{\varkappa_1}, t^{\varkappa_2}]$. Integers $\varkappa_1 \leq \varkappa_2$ are called the *right partial indices* of $A(t)$, and they are unique. By contrast, the factors are not unique. We discuss possible normalisation that guarantees the uniqueness in the following.

Factorisation (1.2) is said to be *canonical* if all partial indices are equal to zero ($D(t) = I_2$). In virtue of the Gohberg–Krein–Bojarskii criterion [1], a non-singular matrix function $A(t)$ admits the stable factorisation if and only if $\varkappa_2 - \varkappa_1 \leq 1$. For example, a matrix function with a positive definite real component admits the canonical factorisation [1], while a matrix function not containing zero in its numerical range has partial indices equal to each other [15]. Hence, both classes admit stable factorisations.

We call a matrix function *strictly* non-singular 2×2 on the unit circle $\mathbb{T}$ if the following conditions are satisfied:

$$a_{11}(t) \neq 0, \ \ \det A(t) \neq 0, \ t \in \mathbb{T}. \tag{1.3}$$

Note that matrix functions with positive definite real components and matrix functions not containing zero in their numerical ranges satisfy conditions (1.3) automatically ([11, ch. II, section 6] and [14, ch. II, section 1.3]), but not vice versa. Some results on the partial indices of strictly non-singular matrices were obtained in Adukov [16]. Recently, an effective algorithm to construct an approximate numerical factorisation of a matrix function being arbitrarily close to a given 2×2 strictly non-singular matrix function has been proposed in Ephremidze & Spitkovsky [17]. However, to the best of the authors' knowledge, there is still no rigorous proof available to attribute the constructed approximation to the given matrix function.

In this paper, we close this gap. Namely, we develop a method allowing us to determine and prove whether a *strictly* non-singular 2×2 matrix function possesses a stable set of partial indices and simultaneously performs its approximate factorisation. A crucial point of our analysis is the development of efficient conditions allowing us to determine the stability region (where the approximate matrix functions preserve their *stable* partial indices). As an approximation, we use polynomial matrix functions. This allows us to use the ExactMPF package working within the Maple Software environment [18] developed by Adukov *et al*. [19]. It is worth noting that the paper by Adukov *et al*. [19] was based on the method of essential polynomials [20]. This powerful tool allows for simultaneous left- and right-side factorisation of the given analytical matrix function. The method, however, is difficult for realisation for external users. For this reason, it was implemented in software as the ExactMPF package working in the Maple Software [18] environment where the factorisation process is fully automated. ExactMPF can be used as a tool for numerical experiments with matrix factorisation and in any applications requiring factorisation of polynomial matrix functions. The listing of the package can be found in the electronic supplementary material of Adukov *et al*. [19]. Below we outline the main idea of our approach.

Let $A(t)$ be an arbitrary invertible matrix function from the Wiener algebra $W^{2\times 2}$ and $\mathcal{EF} \subset W^{2\times 2}$ a class of matrix functions admitting an explicit solution of the factorisation problem. By the explicit solution of the factorisation problem, we understand a clearly defined algorithmic procedure that definitely terminates after a finite number of steps. Additionally, we note that if (i) the input data belongs to the Gaussian field $\mathbb{Q}(i)$ of complex rational numbers and (ii) all (finite) steps of the explicit algorithm can be performed in rational arithmetic, then we say that the problem can be solved exactly [19]. Examples of the class $\mathcal{EF}$ are a class of Laurent matrix polynomials [21,22], a class of meromorphic matrix functions [20], or a class of triangular 2×2 matrix functions defined in Adukov *et al*. [23].

Now, suppose we have approximated $A(t)$ by $\tilde{A}(t) \in \mathcal{EF}l$ such that $\|A - \tilde{A}\|_W < \delta$ and $\tilde{A}(t)$ admits the canonical factorisation $\tilde{A}(t) = \tilde{A}_-(t)\tilde{A}_+(t)$. Since the canonical factorisation is stable, there exists a neighbourhood of $\tilde{A}(t)$ consisting of matrix functions admitting the canonical factorisation. Assume we can obtain an explicit estimate of the radius R of such a neighbourhood as $R \le \tilde{r}_{\mathrm{cf}}$, where the formula fo $\tilde{r}_{\mathrm{cf}}$ will be discussed below. If $\delta < \tilde{r}_{\mathrm{cf}}$ then $A(t)$ also admits the canonical factorisation (figure 1). Similar reasoning can be applied to a stable factorisation ($\rho_2 - \rho_1 = 1$), where $\tilde{r}_{\mathrm{cf}}$ will have a different representation though. Note that both estimates depend on the normalisation of matrix functions during the approximation as well as the factorisation procedures.

We note that factorisation of an arbitrary strictly non-singular matrix function $A(t)$ from (1.1) can be explicitly reduced to factorisation of an auxiliary (invertible) matrix function in the form (see §2 for details):

$$a(t) = \begin{pmatrix} 1 & \beta_-(t) \\ \alpha_+(t) & t^{\theta} + \alpha_+(t)\beta_-(t) \end{pmatrix}, \tag{1.4}$$

where $\theta = \operatorname{ind}\det A(t) = \operatorname{ind}\det A(t) - 2\ \operatorname{ind} a_{11}(t)$.

The approach, highlighted above for matrix function $A(t)$ in its general form, can be applied to the matrix function $a(t)$. As a class $\mathscr{EF}$, we take Laurent matrix polynomials in the form:

$$a_N(t) = \begin{pmatrix} 1 & \beta_-^{(N)}(t) \\ \alpha_+^{(N)}(t) & t^{\theta} + \alpha_+^{(N)}(t)\beta_-^{(N)}(t) \end{pmatrix}, \tag{1.5}$$

where $\alpha_+^{(N)}(t)$ and $\beta_-^{(N)}(t)$ are the Laurent polynomials of degree N of the functions $\alpha_+(t)$, $\beta_-(t)$.

To obtain the explicit estimate of $\|a - a_N\|_W$, we assume that $a(t)$ is analytical in some annulus containing the unit circle $|t| = 1$. In this case, the matrix Fourier coefficients of $a(t)$ coincide with the Laurent coefficients of $a(t)$, and we can use the Cauchy inequalities [24] for the latter.

The factorisation of $a_N(t)$ can be explicitly constructed by the method of essential polynomials (refer to [20,21,25]). Moreover, if the Laurent coefficients of $\alpha_+^{(N)}(t)$ and $\beta_-^{(N)}(t)$ belong to the field $\mathbb{Q}(i)$, then the factorisation problem for $a_N(t)$ can be solved *exactly* [19] with the help of the package ExactMPF. If the coefficients do not belong to the field $\mathbb{Q}(i)$, we can find their rational approximations and apply ExactMPF. The use of the error-free calculations performed with the package is the key idea of the approach as it guarantees the exact computation of the partial indices.

This paper is organised as follows. In §2, we present important facts about the matrix Wiener algebra and the Toeplitz operators used in the following, discuss possible normalisation of the factorisation $A(t)$ and finally show the reduction of an arbitrary strictly non-singular matrix function equation (1.1) to (1.4). In §3, we give the approximation of the matrix function $a(t)$ and estimate its accuracy. Section 4 contains the main result of this article, namely a criterion of the

royalsocietypublishing.org/journal/rspa Proc. R. Soc. A 480: 20240116

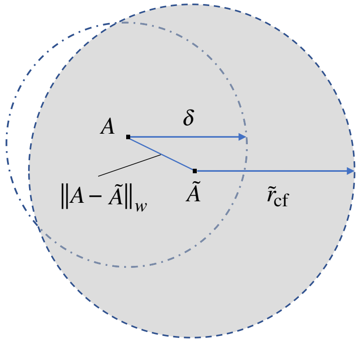

Figure 1. Neighbourhoods of A and $\widetilde{A}$ for which $A(t)$ is expected to admit the canonical factorisation. If $\widetilde{A}$ admits an *explicit* canonical factorisation, then two inequalities, $\left\|A-\widetilde{A}\right\|_W < \delta$ and $\delta < \widetilde{r}_{\rm cf}$, together guarantee that A also possesses the same property. The goal of this paper is to deliver a tool to effectively verify the conditions.

stable factorisation for $a(t)$. In §5, we obtain the explicit estimates for the absolute error in the approximate calculation of the factorisation factors. Finally, in §6, we present some numerical results highlighting three different cases: canonical factorisation, a factorisation with equal partial indices and a stable factorisation. The electronic supplementary material contains a full set of the data in the form of tables.

2. Preliminary considerations

(a) Main definitions and known facts

In this section, we provide some important definitions and follow the notation from Clancey & Gohberg, Litvinchuk & Spitkovskii, Gohberg & Feldman and Gohberg & Kaashoek [11–13,26].

Let $W^{p\times p}$ be the $p\times p$ matrix Wiener algebra consisting of matrix functions with entries from W. Thus any $A(t)\in W^{p\times p}$ expands into an absolutely convergent matrix Fourier series, $A(t)=\sum_{k=-\infty}^{\infty}A_k t^k$, such that $\sum_{k=-\infty}^{\infty}\|A_k\|<\infty$. Here, A_k belongs to the algebra $\mathbb{C}^{p\times p}$ complex matrices equipped with the norm represented by any matrix multiplicative norm on $\mathbb{C}^{p\times p}$, preserving unity. Algebra $W^{p\times p}$ becomes a Banach algebra if we endow it with a norm $\left\|A\right\|_W=\sum_{k=-\infty}^{\infty}\|A_k\|_1$, where $\|\cdot\|_1$ is the maximum column sum matrix norm. Also, we denote a group of invertible elements of the algebra $W^{p\times p}$ by $GW^{p\times p}$.

If $A_k=(a_{ij}^{(k)})_{i,j=1,\dots,p}$, where $a_{ij}^{(k)}$ are the Fourier coefficients of the element $a_{ij}(t)$ of the matrix function $A(t)$, then it is easily seen that $\left\|A_k\right\|_1\le\sum_{i,j}\left|a_{ij}^{(k)}\right|$. Then $\left\|A\right\|_W\le\sum_{i,j}\|a_{ij}\|_W$. Let us define

$$W_+^{p\times p}=\left\{A(t)\in W^{p\times p}: A(t)=\sum_{k=0}^{\infty}A_k t^k\right\},$$

$$W_-^{p\times p}=\left\{A(t)\in W^{p\times p}: A(t)=\sum_{k=-\infty}^{0}A_k t^k\right\},$$

$$(W_-^{p\times p})_0=\left\{A(t)\in W^{p\times p}: A(t)=\sum_{k=-\infty}^{-1}A_k t^k\right\}.$$

It is known that $W_\pm^{p\times p}$, $(W_-^{p\times p})_0$ are closed subalgebras of $W^{p\times p}$ and

$$W^{p\times p}=W_+^{p\times p}\oplus(W_-^{p\times p})_0.$$

Thus $W_{\pm}^{p\times p}$ is a decomposing algebra.

By $\mathcal{P}_+$ we denote a projector from $W^{p\times p}$ onto $W_+^{p\times p}$ along $(W_-^{p\times p})_0$ and $\mathcal{P}_- = \mathcal{I} - \mathcal{P}_+$. Here, $\mathcal{I}$ is the identity operator. Operator $\mathcal{P}_+$ acts according to the rule

$$\mathcal{P}_+ \sum_{k=-\infty}^{\infty} A_k t^k = \sum_{k=0}^{\infty} A_k t^k$$

and is a linear bounded operator in $W^{p\times p}$, while $\|\mathcal{P}_+\| = 1$.

Let $A(t) \in W^{p\times p}$. On the Banach space $W_+^{p\times p}$, we define the operator T_A acting according to the rule

$$T_A X(t) = \mathcal{P}_+ A(t) X(t)\,.$$

It is obvious that T_A is a linear bounded operator and $\|T_A\| \le \|A\|_W$. T_A is called *the Toeplitz operator with the matrix symbol* $A(t)$. Note that even though the Toeplitz operator is usually considered on the Banach space $W_+^{p\times 1}$, for us it is more convenient to operate on $W_+^{p\times p}$. It is straightforward to prove the *partial multiplicativity* property of the mapping $A \to T_A$, that is, for any $A_\pm \in {W_\pm}^{p\times p}$ and $B \in W^{p\times p}$, the relations

$$T_{BA_+} = T_B T_{A_+}, \quad T_{A_- B} = T_{A_-} T_B$$

are valid. As a result, all properties of the standard Toeplitz operator are preserved.

We will also need the following two statements.

If $A(t) = A_-(t)D(t)A_+(t)$ is the right Wiener–Hopf factorisation of an invertible matrix function $A(t)$, then the Toeplitz operator T_A is the right (left) invertible if and only if all right partial indices are non-positive (non-negative). In this case, $T_A^{(-1)} = T_{A_+^{-1}} T_{D^{-1}} T_{A_-^{-1}}$ is its one-sided inverse [13, ch. VIII, cor. 4.1].

If a linear bounded operator $\mathcal{A}$ is one-side invertible and $\mathcal{A}^{(-1)}$ is its one-side inverse, then any operator $\tilde{\mathcal{A}}$ satisfying the inequality

$$\|\mathcal{A} - \tilde{\mathcal{A}}\| < \frac{1}{\|\mathcal{A}^{(-1)}\|} \tag{2.1}$$

is also one-side invertible (from the same side as $\mathcal{A}$). Moreover, if $\mathcal{A}$ is one-side invertible but not invertible, then $\tilde{\mathcal{A}}$ is also not invertible [27, ch. II, th. 5.4]. Note that (2.1) is true in any elements of an abstract Banach algebra.

(b) Normalised factorisation of matrix functions

Normalisation of the factorisation plays a crucial role when performing factorisation numerically. This is specifically in cases when there is a need to compare the consequent approximations. Usually, a normalisation is chosen ad hoc in a line with chosen numerical/asymptotic procedure without any justification, and success depends on each specific case (e.g. [28–30]).

Unfortunately, being important, this issue has still not been fully resolved. The effective results are known for (i) 2×2 matrix functions [31]; for matrix functions of arbitrary dimension $p \times p$ $(p > 2)$, when (ii) the set of partial indices is stable [32] or (iii) all partial indexes are different [33].

Since the paper deals with 2×2 matrices, we present here the normalisation only for such cases. Two cases should be distinguished.

(1) If the matrix function admits a factorisation $A(t) = t^{\varkappa_1} A_-(t)A_+(t)$, $(\varkappa_1 = \varkappa_2)$, then it can be trivially normalised by the condition $A_-(\infty) = I_2$, where I_2 is the unit matrix, and such factorisation is unique.

(2) In the general case $(\varkappa_1 < \varkappa_2)$, there exists the so-called P-normalised factorisation of $A(t)$ that guarantees its uniqueness [31]. The type of normalisation is determined by a 2×2 permutation matrix P, $P = I_2$ or $P = J_2$: $= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

The latter can be written in an explicit form: if $(A_-(\infty))_{11} \neq 0$ (or $(A_-(\infty))_{21} \neq 0$), then $A(t)$ admits the I_2-normalised (or J_2-normalised) factorisation [32, th. 2.1].

In a particular case of the stable factorisation $(\rho_2 - \rho_1 = 1)$ when $(A_-(\infty))_{11} \neq 0$, the I_2-normalisation is carried out as follows (see more details in Adukov [31,32]). Let $A_-(t) = A_0 + A_1 t^{-1} + \ldots$, and $A_0 = L_0 U_0$ is LU-factorisation of the limiting matrix function $A_0 = A_-(\infty)$. Define a matrix polynomial $Q_-(t) = Q_0 + Q_1 t^{-1}$ such that

$$Q_0 = U_0^{-1}, \quad Q_1 = -\frac{1}{(A_0)_{11}} \begin{pmatrix} 0 & (A_1 U_0^{-1})_{12} \\ 0 & 0 \end{pmatrix}. \tag{2.2}$$

Let $C_-(t) = A_-(t)Q_-(t)$, $C_+(t) = D^{-1}(t)Q_-^{-1}(t)D(t)A_+(t)$. Then $A(t) = C_-(t)D(t)C_+(t)$ is the sought for I_2-normalised factorisation of $A(t)$. J_2-normalised factorisation constructed analogously by the JLU-factorisation of A_0.

Finally, P-normalisation is stable under a small perturbation $A(t)$, that is, for arbitrary sufficiently small $\delta > 0$, each matrix function $\tilde{A}(t)$ possessing the same set of the right partial indices ρ_1, ρ_2, as $A(t)$ and satisfying the inequality $\| A - \tilde{A} \|_W < \delta$ has the same type of P-normalisation as $A(t)$ [32, th. 4.1].

(c) Reduction of matrix function $A(t)$ to the form $a(t)$

In this subsection, we deliver a representation of the matrix function $A(t)$ given in the form (1.1) allowing us to consider a simpler matrix function, $a(t)$, in the form (1.4). Let

$$a_{11}(t) = a_{11}^-(t)t^{\varkappa}a_{11}^+(t), \quad \Delta(t) = \Delta_-(t)t^{\theta+2\varkappa}\Delta_+(t) \tag{2.3}$$

be the Wiener–Hopf factorisations of the functions $a_{11}(t)$ and $\Delta(t) = \det A(t)$, respectively. Note that the indices $\varkappa$ and $\theta + 2\varkappa$ in (2.3) do not coincide, generally speaking, with the pair $\varkappa_1$, $\varkappa_1 + \varkappa_2$ (compare with (1.2)). The only relationship between them is $\varkappa_1 + \varkappa_2 = \theta + 2\varkappa$.

The matrix function $A(t)$ can be equivalently written in the form:

$$A(t) = t^{\varkappa} \begin{pmatrix} a_{11}^-(t) & 0 \\ 0 & \dfrac{\Delta_-(t)}{a_{11}^-(t)} \end{pmatrix} \begin{pmatrix} 1 & \beta(t) \\ \alpha(t) & \dfrac{a_{22}a_{11}^-a_{11}^+}{t^{\varkappa}\Delta_-\Delta_+} \end{pmatrix} \begin{pmatrix} a_{11}^+(t) & 0 \\ 0 & \dfrac{\Delta_+(t)}{a_{11}^+(t)} \end{pmatrix}, \tag{2.4}$$

where

$$\alpha(t) = \frac{a_{11}^-(t)a_{21}(t)}{\Delta_-(t)t^{\varkappa}a_{11}^+(t)}, \quad \beta(t) = \frac{a_{11}^+(t)a_{12}(t)}{\Delta_+(t)t^{\varkappa}a_{11}^-(t)}. \tag{2.5}$$

Introducing functions $\alpha_\pm(t) = \mathcal{P}_\pm\alpha(t)$, $\beta_\pm(t) = \mathcal{P}_\pm\beta(t)$, matrix function $A(t)$ can be further transformed to the form

$$A(t) = t^{\varkappa}\begin{pmatrix} a_{11}^{-}(t) & 0 \\ \dfrac{\Delta_{-}(t)\alpha_{-}(t)}{a_{11}^{-}(t)} & \dfrac{\Delta_{-}(t)}{a_{11}^{-}(t)} \end{pmatrix} a(t) \begin{pmatrix} a_{11}^{+}(t) & \dfrac{\Delta_{+}(t)\beta_{+}(t)}{a_{11}^{+}(t)} \\ 0 & \dfrac{\Delta_{+}(t)}{a_{11}^{+}(t)} \end{pmatrix}, \tag{2.6}$$

where

$$a(t) = \begin{pmatrix} 1 & 0 \\ -\alpha_{-}(t) & 1 \end{pmatrix} \begin{pmatrix} 1 & \beta(t) \\ \alpha(t) & \dfrac{a_{22}a_{11}^{-}a_{11}^{+}}{t^{\varkappa}\Delta_{-}\Delta_{+}} \end{pmatrix} \begin{pmatrix} 1 & -\beta_{+}(t) \\ 0 & 1 \end{pmatrix}. \tag{2.7}$$

It remains to be observed that this matrix function has the form (1.4):

$$a(t) = \begin{pmatrix} 1 & \beta_{-}(t) \\ \alpha_{+}(t) & t^{\theta} + \alpha_{+}(t)\beta_{-}(t) \end{pmatrix}. \tag{2.8}$$

For $\theta = 0$, this reduction was carried out in Adukov [16], where explicit formulas for the partial indices of $a(t)$ were obtained under the additional condition that $\alpha_{+}(t)$ is a polynomial in t or $\beta_{-}(t)$ is a polynomial in t^{-1}.

Note that the matrix function $a(t)$ is invertible ($\det A(t) = t^{\theta}$) and thus can be factorised:

$$a(t) = a_{-}(t)d_{r}(t)a_{+}(t), \quad d_{r}(t) = \mathrm{diag}[t^{\rho_1}, t^{\rho_2}], \tag{2.9}$$

where $\rho_1 \le \rho_2$ and $\theta = \rho_1 + \rho_2$.

Comparing the latter with (2.9) and (2.6), we conclude that the partial indices of the matrices $A(t)$ and $a(t)$ are related in the following manner:

$$\varkappa_1 = \varkappa + \rho_1, \quad \varkappa_2 = \varkappa + \rho_2\,. \tag{2.10}$$

In the following, we will focus on the stable factorisation of the matrix function $a(t)$. Hence,

$$\begin{aligned} \rho_1 = \rho_2 = \nu \quad &\text{for even} \quad \theta = 2\nu, \\ \rho_1 = \nu, \rho_2 = \nu + 1 \quad &\text{for odd} \quad \theta = 2\nu{+}1\,. \end{aligned} \tag{2.11}$$

As we have already mentioned above, when the authors of Ephremidze & Spitkovsky [17] built their approximation, they used a completely different representation of the matrix function $A(t)$, in comparison with (2.6):

$$S(t) = \begin{pmatrix} s_{11}^{+}(t) & 0 \\ \dfrac{t^{-k_1}s_{21}(t)}{s_{11}^{-}(t)} & \dfrac{\Delta^{+}(t)}{s_{11}^{+}(t)} \end{pmatrix} \begin{pmatrix} t^{k_1} & 0 \\ 0 & t^{k-k_1} \end{pmatrix} \begin{pmatrix} s_{11}^{-}(t) & \dfrac{t^{-k_1}s_{12}(t)}{s_{11}^{+}(t)} \\ 0 & \dfrac{\Delta^{-}(t)}{s_{11}^{-}(t)} \end{pmatrix},$$

where $s_{11}(t) = s_{11}^{+}(t)t^{k_1}s_{11}^{-}$ and $\det S(t) = \Delta^{+}(t)t^{k}\Delta^{-}(t)$.

3. Approximation of the matrix function $a(t)$ by the Laurent matrix polynomial $a_N(t)$

In this section, we approximate $a(t)$ by the Laurent matrix polynomial $a_N(t)$ and estimate the norm of the difference $\|a - a_N\|_W$.

Let $\alpha_{+}(t) = \sum_{k=0}^{\infty}\alpha_k t^k$ and $\beta_{-}(t) = \sum_{k=-\infty}^{-1}\beta_k t^k$. We denote

$$\alpha_+^{(N)}(t) = \sum_{k=0}^{N} \alpha_k t^k, \ \ \beta_-^{(N)}(t) = \sum_{k=-N}^{-1} \beta_k t^k$$

and consider the Laurent matrix polynomial defined in (1.5).

Similarly to the matrix function $a(t)$, since $\det a_N(t) = t^\theta$, then $a_N(t)$ is an invertible element of $W^{2\times 2}$ for any N.

Now we can estimate the norm $\|a - a_N\|_W$ as

$$\|a - a_N\|_W \le \|\alpha_+ - \alpha_+^{(N)}\|_W + \|\beta_- - \beta_-^{(N)}\|_W + \|\alpha_+\beta_- - \alpha_+^{(N)}\beta_-^{(N)}\|_W .$$

Representing

$$\alpha_+(t)\beta_-(t) - \alpha_+^{(N)}(t)\beta_-^{(N)} = \alpha_+(t)\left(\beta_-(t) - \beta_-^{(N)}(t)\right) + \beta_-^{(N)}(t)\left(\alpha_+(t) - \alpha_+^{(N)}(t)\right)$$

and taking into account that $\|\beta_-^{(N)}\|_W \le \|\beta_-\|_W$, we get

$$\|a - a_N\|_W \le (1 + \|\alpha_+\|_W)\|\beta_- - \beta_-^{(N)}\|_W + (1 + \|\beta_-\|_W)\|\alpha_+ - \alpha_+^{(N)}\|_W . \tag{3.1}$$

The last inequality gives us a qualified estimate of the norm of the difference between the matrix functions $a_N(t)$ and $a(t)$ defined in (1.4) and (1.5), respectively.

To evaluate explicitly the convergence rate, we make additional assumptions on the rate of decay of the Fourier coefficients. Namely, we restrict ourselves to the case when the matrix function $a(t)$ is analytic in the annulus $r_1 < |t| < r_2$ containing the unit circle, $\mathbb{T}$, while r_1 and r_2 can be zero and infinity, respectively. This effectively means that the function $\alpha_+(t)$ is analytic in the domain $|t| < r_2$, while $\beta_-(t)$ is analytic in the domain $|t| > r_1$.

Obviously, the restriction of $a(t)$ on the unit circle $|t| = 1$ belongs to the Wiener algebra $W^{2\times 2}$, and its matrix Fourier coefficients a_k coincide with the coefficients of the Laurent series for $a(t)$ in the annulus $r_1 < |t| < r_2$.

The Laurent coefficients $|\alpha_k|$ of the function $\alpha_+(t)$ hold Cauchy's inequalities [24]:

$$|\alpha_k| \le \frac{M_+(\zeta)}{\zeta^k} \tag{3.2}$$

for any ζ, $0 < \zeta < r_2$. Here, $M_+(\zeta) = \max_{|t|=\zeta} |\alpha_+(t)|$.

Fixing a value $\zeta = \zeta_2$, $(1 < \zeta_2 < r_2)$, we can estimate the sums of series $\sum_{k=0}^{\infty} |\alpha_k|$, $\sum_{k=N+1}^{\infty} |\alpha_k|$ and obtain the following inequalities:

$$\|\alpha_+\|_W \le \frac{\zeta_2 M_+(\zeta_2)}{\zeta_2 - 1}, \ \|\alpha_+ - \alpha_+^{(N)}\|_W \le \frac{M_+(\zeta_2)}{\zeta_2^N(\zeta_2 - 1)}, \ M_+(\zeta_2) = \max_{|t| \le \zeta_2} |\alpha_+(t)| . \tag{3.3}$$

If $r_2 < \infty$ and $\alpha_+(t)$ is continuously extended on $|t| = r_2$, then the Cauchy inequalities hold for $\zeta = r_2$ and, in this case, we can take also $\zeta_2 = r_2$.

For the function $\beta_-(t)$, we can obtain similar estimates fixing a value $\zeta = \zeta_1$, $(r_1 < \zeta_1 < 1)$:

$$\|\beta_-\|_W \le \frac{\zeta_1 M_-(\zeta_1)}{1 - \zeta_1}, \ \|\beta_- - \beta_-^{(N)}\|_W \le \frac{\zeta_1^{N+1} M_-(\zeta_1)}{(1 - \zeta_1)}, \tag{3.4}$$

where $M_-(\zeta_1) = \max_{|t|=\zeta_1} |\beta_-(t)| = \max_{|t|\ge\zeta_1} |\beta_-(t)|$. If $r_1 > 0$ and $\beta_-(t)$ admits a continuous extension on the circle $|t| = r_1$, the value $\zeta_1 = r_1$ is also admissible.

This allows us to obtain an explicit estimate (3.1)

$$\|a - a_N\|_W \le \delta_N(\zeta_1, \zeta_2), \tag{3.5}$$

where $r_1 < \zeta_1 < 1$, $1 < \zeta_2 < r_2$ and

$$\delta_N(\zeta_1,\zeta_2)=\left(1+\frac{\zeta_2 M_+(\zeta_2)}{\zeta_2-1}\right)\frac{\zeta_1^{N+1}M_-(\zeta_1)}{1-\zeta_1}+\left(1+\frac{\zeta_1 M_-(\zeta_1)}{1-\zeta_1}\right)\frac{M_+(\zeta_2)}{\zeta_2^N(\zeta_2-1)}. \tag{3.6}$$

The values $\zeta_1=r_1$ and $\zeta_2=r_2$ can be also admissible.

It is obvious that $\delta_N(\zeta_1,\zeta_2)$ is monotonic decreasing as $N\to\infty$ and tends to zero for any fixed ζ_1,ζ_2. Although, if ζ_1 and ζ_2 are close enough to 1, then the convergence is slow. On the other hand, the function increases when ζ_1 or ζ_2 are close to the other ends of their intervals. Therefore, it is desirable to optimise the choice ζ_1, ζ_2 minimising the function $\delta_N(\zeta_1,\zeta_2)$ on the rectangle $(r_1,1)\times(1,r_2)$. Depending on whether the series converges on the closed or open domain, we can distinguish the following four respective cases:

(1) $r_1>0, r_2<\infty$, that is, the function $\alpha_+(t)$ is analytic into $|t|<r_2$, and $\beta\_(t)$ is analytic into $|t|>r_1$.

(2) $r_1>0, r_2=\infty$, that is, the function $\alpha_+(t)$ is an entire function, and $\beta\_(t)$ is analytic into $|t|>r_1$.

(3) $r_1=0, r_2<\infty$, that is, the function $\alpha_+(t)$ is analytic into $|t|<r_2$, and $\beta\_(t)$ is analytic into $\overline{\mathbb{C}}\setminus\{0\}$.

(4) $r_1=0, r_2=\infty$, that is, the function $\alpha_+(t)$ is an entire function, and $\beta\_(t)$ is analytic into $\overline{\mathbb{C}}\setminus\{0\}$.

In order to effectively use (3.6), we need to estimate $M_-(\zeta_1)$, $M_+(\zeta_2)$ in the respective domains with the best possible accuracy. Here, a difficulty stems from the fact that the minimisation is sought in the open set $(r_1,1)\times(1,r_2)$. In practice, we restrict ourselves to a closed rectangular K embedded into $(r_1,1)\times(1,r_2)$ that provides quite a reasonable approximation.

4. A criterion of the stable factorisation for the matrix function $a(t)$

Having constructed a Laurent matrix polynomial (1.5) of degree N, which is an approximant of the matrix function (1.4) allowing for a stable factorisation as described in §3, we can compute the following measure:

$$q_N(\zeta_1,\zeta_2)=\sigma\,\delta_N(\zeta_1,\zeta_2)\,\|\left(a_+^{(N)}\right)^{-1}\|_W\,\|\left(a_-^{(N)}\right)^{-1}\|_W, \tag{4.1}$$

where

$$\sigma=\|d_r\|_W=\|d_r^{-1}\|_W=\begin{cases}1, & \text{if } \theta=2\nu,\\ 2, & \text{if } \theta=2\nu+1.\end{cases} \tag{4.2}$$

Theorem 4.1. *Let $a(t)$ admit a stable factorisation (2.9), and the factorisation is P-normalised. Then, there exists a natural N_0 such that, for $N>N_0$ and for all admissible pairs (ζ_1,ζ_2), the following conditions are fulfilled:*

1. $a_N(t)$aN(t) admits a stable P-normalised factorisation:

$$a_N(t)=a_-^{(N)}(t)d_r(t)a_+^{(N)}(t), \tag{4.3}$$

2. $q_N(\zeta_1,\zeta_2)<1$.

Proof. Since the sequence $a_N(t)$ converges to the matrix function $a(t)$ admitting a stable factorisation, it is always possible to find a natural N_0 such that, for all $N>N_0$, $a_N(t)$ admits the same factorisation (the same set of partial indices) [1]. However, to deliver an estimate for the stability domain, a more accurate analysis is required. This is the aim of this theorem.

Since $a_N(t)$ converges to the matrix function $a(t)$, there exists an integer N_0 such that

$$\|a(t)-a_N\|_W < \frac{1}{\sigma\|a_+^{-1}\|_W\|a_-^{-1}\|_W}, \quad \text{if } N \geq N_0. \tag{4.4}$$

From (4.4) and (4.2), it follows that

$$\|a-a_N\|_W < \frac{1}{\|a^{-1}\|_W}, \quad \text{if } N \geq N_0.$$

The latter means that $a_N(t)$ is an invertible element of $W^{2\times 2}$ and thus, it admits a factorisation $a_N(t) = a_-^{(N)}(t)d_r^{(N)}(t)a_+^{(N)}(t)$. The next step in the proof is to demonstrate that the latter inequality guarantees also that $a_N(t)$ belongs to the stability region (or equivalently, $d_r^{(N)}(t) = d_r(t)$).

By the assumption of the theorem, the matrix function $a(t)$ admits the stable factorisation $a(t) = a_-(t)d_r(t)a_+(t)$, that is, $d_r(t) = t^\nu I_2$ if $\theta = 2\nu$, and $d_r(t) = t^\nu\begin{pmatrix}1 & 0\\ 0 & t\end{pmatrix}$ for $\theta = 2\nu+1$. Moreover, since we consider the right partial indices of the matrix function $a(t)$, they are both non-negative for $\nu \geq 0$ and non-positive for $\nu < 0$. We will demonstrate now that $d_r^{(N)}(t) = d_r(t)$.

Let us consider Toeplitz operators T_a and T_{a_N} with the symbols $a(t)$ and $a_N(t)$, respectively. Since the partial indices have the same sign, then the operator T_a is one-sided invertible (invertible if $\nu = 0$) and $T_a^{(-1)} = T_{a_+^{-1}}T_{d_r^{-1}}T_{a_-^{-1}}$ is its one-sided inverse (inverse in the case $\nu = 0$), moreover (refer to [13])

$$\|T_a^{(-1)}\| \leq \|T_{a_+^{-1}}\|\,\|T_{d_r^{-1}}\|\,\|T_{a_-^{-1}}\| \leq \sigma\|a_+^{-1}\|_W\|a_-^{-1}\|_W.$$

Then, from (4.4), we have

$$\|T_a - T_{a_N}\| \leq \|a-a_N\|_W < \frac{1}{\sigma\|a_+^{-1}\|_W\|a_-^{-1}\|_W} \leq \frac{1}{\|T_a^{(-1)}\|}.$$

This in turn means that the operator T_{a_N} is also one-side invertible (or invertible if $\nu = 0$). Moreover, the operators T_a and T_{a_N} are invertible from the same side and, therefore, in the factorisation $a_N(t) = a_-^{(N)}(t)d_r^{(N)}(t)a_+^{(N)}(t)$, $d_r^{(N)}(t) = \mathrm{diag}[t^{\tilde\rho_1}, t^{\tilde\rho_2}]$, indices $\tilde\rho_1, \tilde\rho_2$ have the same signs as ρ_1, ρ_2.

The following three cases are possible:

(1) If $a(t)$ admits the canonical factorisation, i.e. $\nu = 0$, then T_{a_N} is invertible and, hence, $a_N(t)$ admits also the canonical factorisation.

(2) Let $a(t)$ admit a stable factorisation and $\theta = 2\nu$, that is, $d_r(t) = t^\nu I_2$. Consider auxiliary matrix functions $b(t) = t^{-\nu}a(t)$ and $c(t) = t^{-\nu}a_N(t)$. By the construction, $b(t)$ admits a canonical factorisation, $b(t) = a_-(t)a_+(t)$. On the other hand, since $\|t^\nu\|_W = \|t^{-\nu}\|_W = 1$ and, owing to inequality (4.4), we conclude

$$\|b-c\|_W < \frac{1}{\|b_+^{-1}\|_W\|b_-^{-1}\|_W}.$$

Hence (see the previous case), $c(t)$ also admits a canonical factorisation, $a_N(t) = t^\nu c(t)$ and $d_r^{(N)}(t) = d_r(t)$.

Finally, let $a(t)$ admit a stable factorisation and $\theta = 2\nu+1$, i.e. $d_r(t) = t^{\nu}\begin{pmatrix}1 & 0\\ 0 & t\end{pmatrix}$. Consider again the matrix functions $b(t) = t^{-\nu}a(t)$ and $c(t) = t^{-\nu}a_N(t)$. The auxiliary matrix function $b(t) = a_-(t)\begin{pmatrix}1 & 0\\ 0 & t\end{pmatrix}a_+(t)$ admits a stable factorisation, and

$$\| b - c \|_W < \frac{1}{2\| b_+^{-1}\|_W \| b_-^{-1}\|_W}.$$

This proves that the respective operators are invertible from the same side, and their right indices are non-negative. Thus, we conclude that $\tilde{\rho}_1 \geq \nu$ and $\tilde{\rho}_2 > \nu$, and, since $\tilde{\rho}_1 + \tilde{\rho}_2 = 2\nu+1$, this is possible only if $\tilde{\rho}_1 = \nu$, $\tilde{\rho}_2 = \nu+1$. This finishes the proof.

Summarising, we have shown that, under condition (4.4), the approximate matrix function $a_N(t)$ admits a stable factorisation with the same set of stable partial indices as the original matrix function $a(t)$.

It remains to prove the second statement of the theorem. Let $N \geq N_0$ and $a_N(t) = a_-^{(N)}(t)d_r(t)a_+^{(N)}(t)$ now be the stable P-normalised factorisation of $a_N(t)$. It follows from th. 5.1 of Adukov [31] that the factors $a_\pm^{(N)}(t)$ converge to normalised $a_\pm(t)$ as $N \to \infty$. Hence, the sequence $\|\left(a_+^{(N)}\right)^{-1}\|_W \cdot \|\left(a_-^{(N)}\right)^{-1}\|_W$ is bounded. As we noted in the previous section, the sequence $\delta_N(\zeta_1, \zeta_2)$ monotonically tends to zero for any fixed pair (ζ_1, ζ_2). Hence, starting from some N, the inequality $q_N(\zeta_1, \zeta_2) < 1$ is fulfilled. ∎

Theorem 4.2. *If conditions 1 and 2 are fulfilled for at least one N and one admissible pair* (ζ_1, ζ_2), *then* $a(t)$ *admits the stable factorisation with the same set of partial indices as* $a_N(t)$.

Proof. Since $\det a_N(t) = t^{\theta}$, $a_N(t)$ is invertible in the algebra $W^{2\times 2}$ for any N. Suppose that there exists N such that $a_N(t)$ admits a stable P-normalised factorisation $a_N(t) = a_-^{(N)}(t)d_r(t)a_+^{(N)}(t)$, and for some (ζ_1, ζ_2) it is fulfilled the inequality

$$\delta_N(\zeta_1, \zeta_2) < \frac{1}{\sigma \|\left(a_+^{(N)}\right)^{-1}\|_W \|\left(a_-^{(N)}\right)^{-1}\|_W}, \tag{4.5}$$

which coincides with $q_N < 1$ (compare with (4.1))

$$\| a - a_N \|_W \leq \delta_N(\zeta_1, \zeta_2) < \frac{1}{\sigma \|\left(a_+^{(N)}\right)^{-1}\|_W \|\left(a_-^{(N)}\right)^{-1}\|_W}. \tag{4.6}$$

Now we carry out the same reasoning as in the proof of theorem 4.1, only interchanging $a(t)$ by $a_N(t)$. Thus, in virtue of (4.6), we have

$$\| T_a - T_{a_N} \| \leq \| a - a_N \|_W < \frac{1}{\|\left(a_+^{(N)}\right)^{-1}\|_W \| d_r(t) \|_W \|\left(a_-^{(N)}\right)^{-1}\|_W}. \tag{4.7}$$

Since $a_N(t)$ admits the stable factorisation, the operator T_{a_N} is one-side invertible, and $T_{a_N}^{(-1)} = T_{(a_+^{(N)})^{-1}} T_{d_r^{-1}} T_{(a_-^{(N)})^{-1}}$ is its one-side inverse.

Now from (4.7) it follows that $\| T_a - T_{a_N} \| < \dfrac{1}{\| T_{a_N}^{(-1)} \|}$. This, in turn, means that the operator T_a is one-side invertible. The same arguments as mentioned above show that the matrix function $a(t)$ admits the stable factorisation. ∎

Remark 4.1. The normalisation of the factorisation is used only in the proof of the necessity of conditions 1 and 2. In fact, in theorem 4.1, we can omit the P-normalisation requirement and require that the sequence $\|\left(a_+^{(N)}\right)^{-1}\|_W \|\left(a_-^{(N)}\right)^{-1}\|_W$ is bounded from above. In theorem 4.2, we can use any factorisation. This allows us to avoid the computational difficulties associated with the procedure of P-normaliation. However, P-normalisation is a mandatory step in constructing an approximate factorisation with guaranteed accuracy.

Remark 4.2. Those theorems allow us to formulate a criterion for the case when the matrix function $a(t)$ does not admit a stable factorisation. However, this would require checking an infinite number of conditions (for each N) and, thus, is unlikely to be implemented in practice.

In fact, the results provided in §§3 and 4 allow us to address the main challenge of this paper, providing two explicit conditions $\|a - a_N\|_W < \delta_N(\zeta_1, \zeta_2)$ and $q_N < 1$ that (i) can be numerically verified and (ii) guarantee that matrix function $a(t)$ has the same set of stable partial indices as $a_N(t)$. To highlight things even further, we note that they represent those inequalities, $\|a - a_N\|_W < \delta$ and $\delta < \tilde{r}_{cf}$, discussed in figure 1. Indeed, the inequality $q_N < 1$ can be equivalently written in the form (4.5) where the right-hand side corresponds to the stability radius $\tilde{r}_{sf}$. The conditions are sufficient for stability in the numerical procedure used in the approximate factorisation, while continuity of the factors has been proven in Adukova [32, th. 5.1]. However, the quality of the approximation remains unknown and thus, if the procedure does not converge fast enough, this might be a serious obstacle for computations.

Having a solid theoretical estimate for the convergence rate is always a benefit. The next section is devoted to this task for a classic canonical case ($\varkappa_1 = \varkappa_2 = 0$) and a generalised canonical one ($\varkappa_1 = \varkappa_2$). Unfortunately, the remaining stable case ($\theta = 2\nu + 1$) requires more delicate treatment.

5. Accuracy estimate of the factors in approximation of a canonical factorisation

In §5a, we briefly recall the results of the work [34] on the continuity of factors $A_\pm(t)$ of the classic canonical factorisation (all partial indices are zero) and improve explicit estimates for the norm $\|A_\pm - A_\pm^{(N)}\|_W$ presented therein. Bearing in mind possible applications, we provide the analysis for matrix functions of an arbitrary order $p \geq 2$.

(a) Explicit estimates for the factors of canonical factorisation of the original matrix function $A(t)$

We will assume that the canonical factorisation $A(t) = A_-(t)A_+(t)$ is normalised by the condition $A_-(\infty) = A_0$, where A_0 is an arbitrary, pre-determined, invertible matrix. This normalisation guarantees the uniqueness of the canonical factorisation. Clearly, we can choose a unit matrix as the limiting value for the factor ($A_0 = I_p$); however, since the proof is not much dependent in this case of normalisation, we stay with a more general formulation.

We need the following elementary lemmas. Their proofs for $q = 1/2$ are given in Adukova & Dilman [34]. Even though the proofs do not differ much in the general case, the new results represent an improvement, making the delivery of the estimates more convenient for practical use.

Lemma 5.1. *Let $\mathcal{A}$ be an invertible bounded linear operator in some Banach space, and the operator $\tilde{\mathcal{A}}$ satisfies the inequality*

$$\|\mathcal{A} - \tilde{\mathcal{A}}\| \leq \frac{q}{\|\mathcal{A}^{-1}\|}, \quad 0 < q < 1. \tag{5.1}$$

Then $\tilde{\mathcal{A}}$ is also invertible, and, for solutions of the equations $\mathcal{A}x = b$, $\tilde{\mathcal{A}}\,\tilde{x} = \tilde{b}$, the following estimate holds true

$$\|x - \tilde{x}\| \le \frac{\|\mathcal{A}^{-1}\|}{1-q}\Big(\|b\|\,\|\mathcal{A}^{-1}\|\,\|\mathcal{A} - \tilde{\mathcal{A}}\| + \|b - \tilde{b}\|\Big).$$

Lemma 5.2. *Let the element $\mathcal{A}$ of a Banach algebra be invertible, and the element $\tilde{\mathcal{A}}$ satisfies the inequality (5.1). Then $\tilde{\mathcal{A}}$ is an invertible element and*

$$\|\mathcal{A}^{-1} - \tilde{\mathcal{A}}^{-1}\| \le \frac{1}{1-q}\|\tilde{\mathcal{A}}\|^2 \cdot \|\mathcal{A} - \tilde{\mathcal{A}}\|.$$

Theorem 5.1. *Let $A(t) \in GW^{p\times p}$ admit the canonical factorisation $A(t) = A_-(t)A_+(t)$, and this factorisation is normalised by the condition $A_-(\infty) = A_0$. If the matrix function $\tilde{A}(t) \in W^{p\times p}$ satisfies the inequality*

$$\|A - \tilde{A}\|_W \le \frac{q}{\|A_+^{-1}\|_W \|A_-^{-1}\|_W},\ 0 < q < 1, \tag{5.2}$$

then $\tilde{A}(t)$ is an invertible matrix function that admits the canonical factorisation $\tilde{A}(t) = \tilde{A}_-(t)\tilde{A}_+(t)$. For the unique factorisation normalised by the same condition $\tilde{A}_-(\infty) = A_0$, the following estimates are valid:

$$(i)\ \ \|A_+^{-1} - \tilde{A}_+^{-1}\|_W \le \frac{\|A_0\|}{1-q}\,\|A_+^{-1}\|_W^2\,\|A_-^{-1}\|_W^2 \|A - \tilde{A}\|_W, \tag{5.3}$$

$$(ii)\ \ \|A_- - \tilde{A}_-\|_W \le \Bigg(\|A_+^{-1}\|_W + \frac{\|A_0\|\ \ \|A\|_W}{1-q}\|A_+^{-1}\|_W^2\|A_-^{-1}\|_W^2 + \\ + \frac{q}{1-q}\|A_0\|\,\|A_+^{-1}\|_W\|A_-^{-1}\|_W\Bigg)\|A - \tilde{A}\|_W. \tag{5.4}$$

Moreover, if condition (5.2) is replaced by a stronger one:

$$\|A - \tilde{A}\|_W \le \frac{q}{\|A_+^{-1}\|_W\|A_-^{-1}\|_W} \cdot \frac{(1-q)}{\|A_0\|\,\|A_-^{-1}\|_W\|A_+\|_W\|A_+^{-1}\|_W}, \tag{5.5}$$

then

$$(iii)\ \ \|A_+ - \tilde{A}_+\|_W \le \frac{\|A_0\|}{(1-q)^2}\|A_+\|_W^2\|A_+^{-1}\|_W^2\|A_-^{-1}\|_W^2\|\,\|A - \tilde{A}\|_W. \tag{5.6}$$

Proof. To prove the first inequality (5.3), we consider operator equations $T_A X = A_0$ and $T_{\tilde{A}}\tilde{X} = A_0$, which both have unique solutions. These solutions are matrix functions $X = A_+^{-1}$ and $\tilde{X} = \tilde{A}_+^{-1}$, respectively. Indeed, $T_A A_+^{-1} = T_{A_-}T_{A_+}A_+^{-1} = \mathcal{P}_+A_-\mathcal{P}_+A_+A_+^{-1} = \mathcal{P}_+A_-I_p = \mathcal{P}_+A_- = A_-(\infty) = A_0$. Similar arguments apply to the second equation.

Using lemma 5.1 with $b = \tilde{b} = A_0$, we have

$$\|A_+^{-1} - \tilde{A}_+^{-1}\|_W = \|X - \tilde{X}\|_W \le \frac{\|A_0\|}{1-q}\,\|T_A^{-1}\|^2\,\|T_A - T_{\tilde{A}}\|.$$

Finally, since $\|T_A^{-1}\| \le \|A_+^{-1}\|_W\|A_-^{-1}\|_W$ and $\|T_A - T_{\tilde{A}}\| \le \|A - \tilde{A}\|_W$, we arrive at (5.3).

Now we can prove the continuity of the factor A_- with its explicit estimate (ii). Since both matrix functions $A(t)$ and $\tilde{A}(t)$ admit canonical factorisations, the following identity is valid:

$$A_- - \tilde{A}_- = \big(A - \tilde{A}\big)A_+^{-1} + \tilde{A}\big(A_+^{-1} - \tilde{A}_+^{-1}\big),$$

leading to the inequality

$$\|A_- - \tilde{A}_-\|_W \leq \|A_+^{-1}\|_W \|A - \tilde{A}\| + \|\tilde{A}\|_W \|A_+^{-1} - \tilde{A}_+^{-1}\|. \tag{5.7}$$

Using (5.2), we get an estimate of the first multiplayer of the second term, $\|\tilde{A}\|_W$, in the equality (5.7) in the form

$$\|\tilde{A}\|_W \leq \|A\|_W + \|A - \tilde{A}\|_W \leq \|A\|_W + \frac{q}{\|A_+^{-1}\|_W \|A_-^{-1}\|_W},$$

while the statement (i) of this theorem provides an estimate for the second multiplayer. Substituting those two inequalities into (5.7), we arrive at (5.4).

Finally, we can prove statement (iii), which represents an explicit stability condition for the factor $A_+(t)$ for a small perturbation of the matrix function $A(t)$.

Since condition (5.5) is stronger than (5.2), it also guarantees the existence of the canonical factorisation $\tilde{A}(t)$ fixed by its limiting value $A_-(\infty) = A_0$. Using lemma 5.2 to elements $\mathcal{A} = A_+^{-1}(t)$, $\tilde{\mathcal{A}} = \tilde{A}_+^{-1}(t)$ considered in the Banach algebra $W_+^{p\times p}$ and taking into account statement (i), which is the inequality (5.3) proven above, we arrive at (5.6). Condition (5.1) is fulfilled if we take a new radius of the vicinity of $A(t)$. ∎

To finalise this subsection, we note that the explicit estimates for the absolute errors $\|A_+ - \tilde{A}_+\|_W$, $\|A_- - \tilde{A}_-\|_W$ were obtained only for canonical factorisation of the original matrix function $A(t)$ of an arbitrary size $p \times p$.

(b) An explicit estimates for factors $a_\pm(t)$ for matrix function $a(t)$ from (1.4)

In this subsection, we specify the results for matrix function $a(t)$ in the form (1.4) for $\theta = 2\nu$ and assume that the partial indices are equal. Then we can always consider a new matrix function $b(t) = t^{-\nu} a(t)$ admitting classic canonical factorisation ($\theta = 0$). Below, we assume, without loss of generality, that $\theta = 0$ for both matrix functions $a(t)$ and $a_N(t)$. We will assume that the factorisation is normalised by the condition $a_-^{(N)}(\infty) = I_2$. Applying theorem 5.1 to the matrix function $a_N(t)$, we can formulate

Corollary 5.1. *Let for some N, ζ_1, ζ_2 we have $\|a - a_N\|_W < \delta_N(\zeta_1,\zeta_2)$, $a_N(t)$ admits the canonical factorizsation. Then*

$$\|a_+^{-1} - \left(a_+^{(N)}\right)^{-1}\|_W < \frac{\|\left(a_+^{(N)}\right)^{-1}\|_W^2 \, \|\left(a_-^{(N)}\right)^{-1}\|_W^2}{1 - q_N(\zeta_1,\zeta_2)} \delta_N(\zeta_1,\zeta_2), \tag{5.8}$$

if $q_N(\zeta_1,\zeta_2) := \delta_N(\zeta_1,\zeta_2) \|\left(a_+^{(N)}\right)^{-1}\|_W \, \|\left(a_-^{(N)}\right)^{-1}\|_W < 1$;

$$\|a_- - a_-^{(N)}\|_W < \delta_- = \left(\|\left(a_+^{(N)}\right)^{-1}\|_W + \frac{\|a_N\|_W \, \|\left(a_+^{(N)}\right)^{-1}\|_W^2 \|\left(a_-^{(N)}\right)^{-1}\|_W^2}{1 - q_N(\zeta_1,\zeta_2)} + \frac{q_N(\zeta_1,\zeta_2) \, \|\left(a_+^{(N)}\right)^{-1}\|_W \, \|\left(a_-^{(N)}\right)^{-1}\|_W}{1 - q_N(\zeta_1,\zeta_2)} \right) \delta_N(\zeta_1,\zeta_2), \tag{5.9}$$

if $q_N(\zeta_1,\zeta_2) < 1$;

$$\|a_+ - a_+^{(N)}\|_W < \delta_+ = \frac{\|a_+^{(N)}\|_W^2 \, \|\left(a_+^{(N)}\right)^{-1}\|_W^2 \, \|\left(a_-^{(N)}\right)^{-1}\|_W^2}{\left(1 - q_N^+(\zeta_1,\zeta_2)\right)^2} \delta_N(\zeta_1,\zeta_2), \tag{5.10}$$

if $\gamma_N(\zeta_1,\zeta_2) := 4\delta_N(\zeta_1,\zeta_2) \|a_+^{(N)}\|_W \|\left(a_+^{(N)}\right)^{-1}\|_W^2 \|\left(a_-^{(N)}\right)^{-1}\|_W^2 \leq 1$.

Here, $q_N^+(\zeta_1,\zeta_2) = \frac{1}{2}\left(1 - \sqrt{1 - \gamma_N(\zeta_1,\zeta_2)}\right)$.

Proof. The proof is required only for estimate (5.10). Let us define the parameter $q_N^+(\zeta_1, \zeta_2)$ by the equation

$$q(1-q) = \tfrac{1}{4}\gamma_N(\zeta_1, \zeta_2)\,. \tag{5.11}$$

Since $\max_{[0,1]} q(1-q) = \frac{1}{4}$, the condition $\gamma_N(\zeta_1, \zeta_2) \le 1$ must be met. We have two roots of equation (5.11), but the root is such that $0 < q \le 1/2$, that is, $q_N^+(\zeta_1, \zeta_2) = \frac{1}{2}\left(1 - \sqrt{1-\gamma_N(\zeta_1, \zeta_2)}\right)$, gives the better estimate for $\|a_+ - a_+^{(N)}\|_W$. Now all conditions of theorem 5.1, item (iii), are fulfilled, and applying it, we arrive at (5.10). ∎

Note that, since the factorisation problem for a Laurent matrix polynomial $a_N(t)$ can be solved explicitly (refer to [20]), the estimates for $\|a_- - a_-^{(N)}\|_W$ and $\|a_+ - a_+^{(N)}\|_W$ can be obtained constructively in terms of some characteristics of the functions $\alpha_+(t)$, $\beta_-(t)$.

Suppose that the Laurent coefficients of $\alpha_+^{(N)}(t)$, $\beta_-^{(N)}(t)$ belong to the field $\mathbb{Q}(i)$. Since $\det a_N(t) = 1$, the factorisation problem for $a_N(t)$ can be solved exactly [19]. We can construct the exact factorisation with the help of the package ExactMPF [19]. If the Laurent coefficients α_k, β_k do not belong to the field $\mathbb{Q}(i)$, we can find the best rational approximations for them and include the respective corrections in $\delta_N(\zeta_1, \zeta_2)$.

6. Numerical experiments

(a) Matrix function admitting canonical factorisation ($\theta = 0$)

First, we consider matrix function $a(t)$:

$$a(t) = \begin{pmatrix} 1 & t^{-2}\sqrt{k_1^2 - t^2} \\ \sqrt{k_2^2 - t^2} & 1 + t^{-2}\sqrt{k_2^2 - t^2}\sqrt{k_1^2 - t^2} \end{pmatrix}, \quad k_1^{-1}, k_2 > 1. \tag{6.1}$$

Here, $\alpha_+(t)$ is the branch of $\sqrt{k_2^2 - t^2}$, $k_2 > 1$, which is analytic into the domain $\mathbb{C} \setminus ((-\infty, -k_2) \cup (k_2, +\infty))$, and has the positive value at $t = 0$.

For the function $\beta_-(t) = t^{-2}\sqrt{k_1^2 - t^2}$, $k_1 < 1$, we consider the branch of the square root $\sqrt{k_1^2 - t^2}$ with the cut $(-k_1, k_1)$ fixed by the value of $t^{-1}\sqrt{k_1^2 - t^2}$ at $t = \infty$ equal to i.

The function $\alpha_+(t)$ is expanded into the series

$$\alpha_+(t) = k_2\left(1 - \frac{t^2}{2k_2^2} - \sum_{n=2}^{\infty} \frac{(2n-3)!!}{(2n)!!}\frac{t^{2n}}{k_2^{2n}}\right),$$

that converges absolutely in the closet disc $|t| \le k_2$. So, $\alpha_+(t)$ is continuous on $|t| \le k_2$, analytic into $|t| < k_2$ and $\alpha_+(t)$, $|t| = 1$, belongs to the subalgebra W_+. We can easily prove that $M_+(\zeta_2) = \sqrt{k_2^2 + \zeta_2^2}$ for $1 < \zeta_2 \le k_2$, where function $M_+(\zeta_2)$ has been defined in equation (3.2).

Function $\beta_-(t)$ is expanded into the series

$$\beta_-(t) = \frac{i}{t}\left(1 - \frac{k_1^2}{2t^2} - \sum_{n=2}^{\infty} \frac{(2n-3)!!}{(2n)!!}\frac{k_1^{2n}}{t^{2n}}\right),$$

that converges absolutely in the close domain $|t| \ge k_1$. Moreover, $\beta_-(t)$ is continuous on $|t| \ge k_1$, analytic into $|t| > k_1$ and $\beta_-(t)$, $|t| = 1$, belongs to the subalgebra W_-^0. It is easy to prove that $M_-(\zeta_1)=\frac{\sqrt{k_1^2+\zeta_1^2}}{\zeta_1^2}$ for $k_1\le\zeta_1<1$. Thus, the admissible pairs (ζ_1,ζ_2) belong to the rectangular $[k_1,1)\times(1,k_2]$.

Now we have $\|a-a_N\|_W<\delta_N(\zeta_1,\zeta_2)$ for $k_1\le\zeta_1<1$, $1<\zeta_2\le k_2$, with the parameter $\delta_N(\zeta_1,\zeta_2)$ defined in (3.6)

$$\delta_N(\zeta_1,\zeta_2)=\left(1+\frac{\zeta_2\sqrt{k_2^2+\zeta_2^2}}{\zeta_2-1}\right)\frac{\zeta_1^{N-1}\sqrt{k_1^2+\zeta_1^2}}{1-\zeta_1}+\left(1+\frac{\sqrt{k_1^2+\zeta_1^2}}{\zeta_1(1-\zeta_1)}\right)\frac{\sqrt{k_2^2+\zeta_2^2}}{\zeta_2^N(\zeta_2-1)}. \tag{6.2}$$

Since the optimal choice of the parameters ζ_1,ζ_2 is desirable but not mandatory, we choose as optimal values ζ_1,ζ_2 the solution of the minimisation problem on a compact set $[k_1,1-\epsilon]\times[1+\epsilon,k_2]$ for sufficiently small $\epsilon>0$. For these purposes, we use the package Optimization of the system Maple.

Example 6.1. Now we present an approximate factorisation of the matrix function (6.1) with a guaranteed accuracy for fixed values of $k_1=1/5$ and $k_2=5$. With help procedure Minimize of the package Optimization, we verify that, for $\epsilon=0.01$ and for all $N\in[1,100]$, optimal values are $\zeta_1=1/5$, $\zeta_2=5$, thus

$$\delta_N(1/5,5)=\tfrac{15+2\sqrt{2}}{4}\cdot 5^{1-N}\approx 22.2855\cdot 5^{-N}. \tag{6.3}$$

For any fixed N_0 and $1\le N\le N_0$, the evident inequality

$$\|a_{N_0}-a_N\|_W\le\|a-a_N\|_W\le\delta_N. \tag{6.4}$$

We found the norms $\|a_{30}-a_N\|_W$, $\|a_{40}-a_N\|_W$, $\|a_{50}-a_N\|_W$ and compared them with the calculated values of δ_N for $N\in[1,30]$. The results are shown in figure 2 and the electronic supplementary material, table S1. All calculations were carried out with Digits = 10. The values for $\|a_{50}-a_N\|_W$ completely coincide with $\|a_{40}-a_N\|_W$. As a result, we conclude that $\|a_{40}-a_N\|_W=\|a-a_N\|_W$ for $1\le N\le 30$. Note that the inequality $\|a-a_N\|_W<\delta_N$ is far from being exact: its left-hand side is estimated as 6^{-N} (figure 2), while the right-hand side is of the order $22.3\cdot 5^{-N}$. From the first glance, since we know representations of the functions $a(t)$ and $a_N(t)$, we can compute also the norm $\|a-a_N\|_W$ itself:

$$\|a-a_N\|_W=\sum_{k=N}^{\infty}\|a_k\|.$$

Unfortunately, excluding the case when the series is computed exactly, we can have only an estimate of the norm from below, while the values of δ_N even are sometimes an overestimation that guarantees the inequality from the other side. Moreover, the value of δ_N is also involved in the computations of the second condition $q_N<1$ that is equally important for the theorems 4.1 and 4.2 to be valid.

Note that δ_N is also involved in the estimates of the factors (5.8) and (5.9). Since the Laurent coefficients of $\alpha_+^{(N)}(t)$, $\beta_-^{(N)}(t)$ belong to the field $\mathbb{Q}(i)$ and $\det a_N(t)=1$, the factorisation problem for $a_N(t)$ can be solved exactly. We apply the package ExactMPF for solving the factorisation problems for $a_N(t)$, $1\le N\le 30$.

It turns out that all $a_N(t)$ admits the canonical factorisations, which we normalised by the conditions $a_-^{(N)}(\infty)=I_2$. In the electronic supplementary material, table S1, we also show the values of $\|\left(a_\pm^{(N)}\right)^{-1}\|_W$ and q_N. For clarity, before being included in the tables, all exact rational numbers are converted into the floating format preserving their meaningful length.

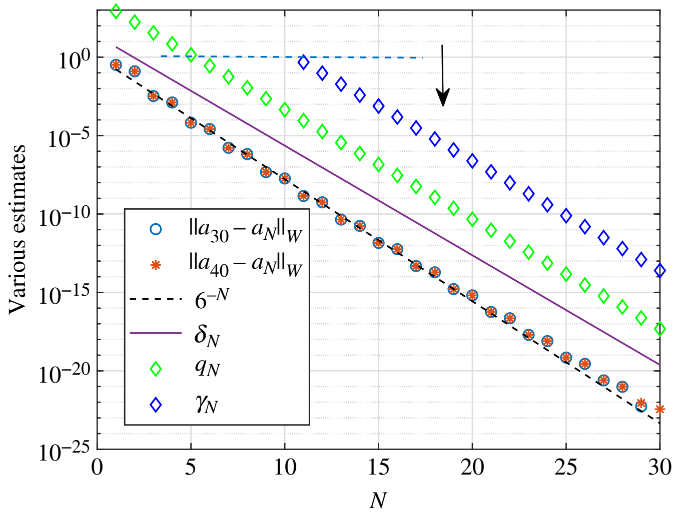

Figure 2. Verification of convergence of the approximation and estimate (6.2) for example 6.1.

Starting from $N = 6$, the conditions of theorem 4.1 are fulfilled. Therefore, the matrix functions $a(t)$ and its approximation, $a_N(t)$, both admit canonical factorisation.

Since the formulas for factors are cumbersome, we present them in the supplementary material.

In the electronic supplementary material, table S2, we also present the values of the right-hand sides of the inequalities (5.9) and (5.10) denoted by $\delta_-^{(N)}$ and $\delta_+^{(N)}$, respectively. Note that the equality (5.10) is valid if $\gamma_N \le 1$ (see corollary 5.1). From the electronic supplementary material, table S1, we see that this condition is fulfilled from $N = 11$. Hence, we calculate $\delta_+^{(N)}$ beginning from $N = 11$. They are estimations of the respective norms $\|a_\pm^{(N)} - a_\pm\|_W$. For example, $\|a_+ - a_+^{(15)}\|_W < 1.275 \cdot 10^{-3}$ and $\|a_- - a_-^{(15)}\|_W < 3.490 \cdot 10^{-4}$. In fact, the actual accuracy is much higher. Using the same line of reasoning as for the matrix functions mentioned above, we present in figure 3 the norms $\|a_\pm^{(40)} - a_\pm^{(N)}\|_W$ and their estimates $\delta_\pm^{(N)}$ for $1 \le N \le 30$. Again, theoretical estimates for the factors are rather crude in this case.

The respective data are given in the electronic supplementary material, table S2, where we also present the norms $\|a_\pm^{(N)} - a_\pm^{(N-1)}\|_W$.

(b) Matrix function with equal partial indices, ($\theta = 2\nu$)

Now we consider other examples.

$$a(t) = \begin{pmatrix} 1 & \dfrac{1}{\sqrt{k_1^2 + t^2}} \\ e^{k_2 t} & t^{2\nu} + \dfrac{e^{k_2 t}}{\sqrt{k_1^2 + t^2}} \end{pmatrix}, \quad 0 < k_1 < 1, \quad k_2 \in \mathbb{C}.$$

Here, $\alpha_+(t) = e^{k_2 t}$, $k_2 \in \mathbb{C}$. It is an entire function, represented by its Taylor series

$$\alpha_+(t) = \sum_{n=0}^{\infty} \frac{k_2^n t^n}{n!},$$

and the function M_+ defined in (3.2) is $M_+(\zeta_2) = e^{|k_2|\zeta_2}$, $1 < \zeta_2 < \infty$.

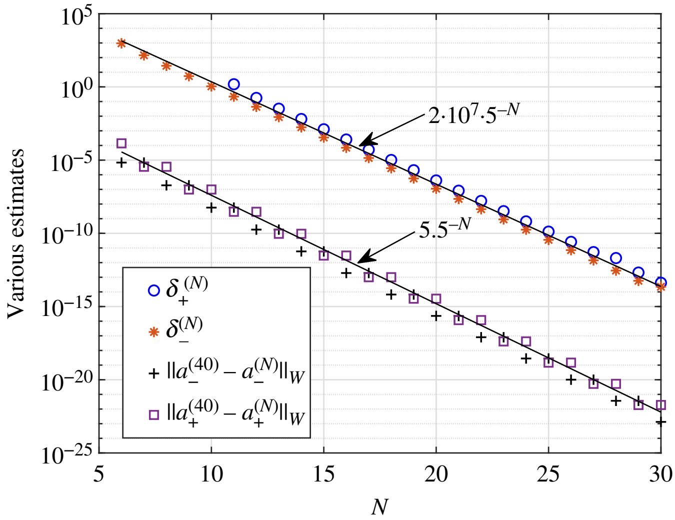

Figure 3. Estimates $\delta_{\pm}^{(N)}$ and the norms $\left\| a_{\pm}^{(40)} - a_{\pm}^{(N)} \right\|_W$ for example 6.1.

The other function, $\beta\_(t) = (k_1^2 + t^2)^{-1/2}$, $0 < k_1 < 1$, with the branch cut $(-k_1 i, k_1 i)$ on the imaginary axis, while the branch is fixed by the limiting value $t\beta\_(t) \to 1$ as $t = \infty$ is equal to 1. The function $\beta\_(t)$ is analytic into $|t| > k_1$ and $\beta\_(t)$ is expanded into the series

$$\beta\_(t) = \frac{1}{t}\left(1 + \sum_{n=1}^{\infty} (-1)^n \frac{(2n-1)!!}{(2n)!!}\frac{k_1^{2n}}{t^{2n}}\right),$$

that converges absolutely in the open domain $|t| > k_1$. Moreover, $\beta\_(t)$, for $|t| = 1$, belongs to the subalgebra W_-^0. It is easy to prove that $M\_(\zeta_1) = (\zeta_1^2 - k_1^2)^{-1/2}$ for $k_1 < \zeta_1 < 1$. From (3.5), we have for $k_1 < \zeta_1 < 1$, $1 < \zeta_2 < \infty$:

$$\| a - a_N \|_W \leq \delta_N(\zeta_1, \zeta_2) = \left(1 + \frac{\zeta_2 e^{|k_2|\zeta_2}}{\zeta_2 - 1}\right)\frac{\zeta_1^{N+1}}{(1-\zeta_1)\sqrt{\zeta_1^2 - k_1^2}} + \left(1 + \frac{\zeta_1}{(1-\zeta_1)\sqrt{\zeta_1^2 - k_1^2}}\right)\frac{e^{|k_2|\zeta_2}}{\zeta_2^N(\zeta_2 - 1)}.$$

Example 6.2. Let $k_1 = 1/5$, $k_2 = 1$, $\theta = 2\nu = 6$. Choosing an optimal choice of the parameters ζ_1, ζ_2, we observed that the pair has a small effect on the factorisation process. Thus, we took them based on the convenience of the numerical calculations. Numerical experiments lead us to the following values: $\zeta_1 = 1/4$, $\zeta_2 = 4$, and consequently

$$\delta_N = \frac{109e^4 + 60}{27}\frac{1}{4^N} \approx 222.6364 \cdot 4^{-N}. \tag{6.5}$$

Similarly to example 6.1, we computed the norms $\| a_{30} - a_N \|_W$, $\| a_{40} - a_N \|_W$ and $\| a_{50} - a_N \|_W$ and compared them with the calculated values of δ_N for $N \in [1, 30]$. The values for $\| a_{50} - a_N \|_W$ completely coincide with $\| a_{40} - a_N \|_W$ with the chosen accuracy (DIGID = 10). Thus we can assume that the lower bound $\| a_{40} - a_N \|_W \leq \| a - a_N \|_W$ is exact, and its left-hand side fully characterises the accuracy of the estimate (6.2). The results are shown in figure 4 and the electronic supplementary material, table S3.

As given above, the Laurent coefficients of $\alpha_+^{(N)}(t)$, $\beta_-^{(N)}(t)$ belong to the field $\mathbb{Q}(i)$ and $\det a_N(t) = 1$, hence the factorisation problem for $a_N(t)$ can be solved exactly with the package ExactMPF.

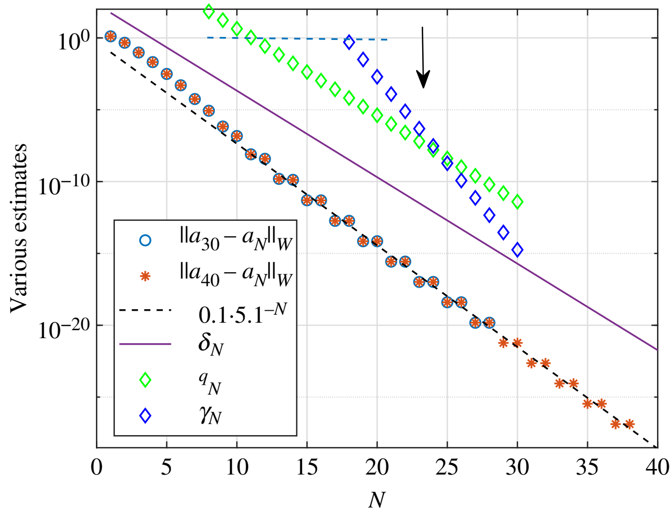

Figure 4. Verification of accuracy of the estimate (6.2) for example 6.2.

It turns out that all $a_N(t)$, $3 \le N \le 30$, admit the stable factorisations with the equal indices $\rho_1 = \rho_2 = \nu = 3$, in other words, $a_N(t) = t^3 a_-^{(N)}(t) a_+^{(N)}(t)$. We normalised the factorisations by the conditions $a_-^{(N)}(\infty) = I_2$. Obviously, estimates (5.9) and (5.10) hold true in this case.

In the electronic supplementary material, table S3, we show the values of $\|\left(a_\pm^{(N)}\right)^{-1}\|_W$ and $q_N(1/4, 4)$ (see (4.1)). From the table, we see that starting from $N = 12$, the stability criterion is fulfilled. Therefore, the matrix function $a(t)$ admits the stable factorisation with equal indices $\rho_1 = \rho_2 = \nu = 3$, and the stable factorisation of $a_N(t)$ gives an approximate stable factorisation of $a(t)$.

The calculations of the guaranteed accuracy give more or less satisfactory results for $N = 30$: $\|a_+ - a_+^{(30)}\|_W < 3.012 \cdot 10^{-3}$ and $\|a_- - a_-^{(30)}\|_W < 5.474 \cdot 10^{-7}$. The values of the norms $\|a_+^{(40)} - a_+^{(30)}\|_W = 3.779 \cdot 10^{-32}$, $\|a_-^{(40)} - a_-^{(30)}\|_W = 8.741 \cdot 10^{-24}$ show that the actual accuracy is much higher than predicted by the estimate. Comparison $\delta_\pm^{(N)}$ with these norms are presented in figure 5 for all values of N and in the electronic supplementary material, table S4.

(c) The stable factorisation of $a(t)$, $\theta = 2\nu + 1$

Now we consider the following matrix function

$$a(t) = \begin{pmatrix} 1 & \dfrac{e^{k_1/t}}{t} \\ e^{k_2 t} & t^\theta + \dfrac{e^{k_1/t + k_2 t}}{t} \end{pmatrix}, \quad k_1, k_2 \in \mathbb{C}, \quad \theta = 2\nu + 1.$$

Here, the function $\beta\_(t) = t^{-1} e^{k_1/t}$ is analytic on $\mathbb{C} \setminus \{0\}$, and $\beta\_(\infty) = 0$, while $\alpha_+(t) = e^{k_2 t}$ is an entire function. It is easy to check that $M_-(\zeta_1) = \frac{1}{\zeta_1} e^{|k_1|/\zeta_1}$ and $M_+(\zeta_2) = e^{|k_2|\zeta_2}$, and (3.5) gives for $0 < \zeta_1 < 1$, $1 < \zeta_2 < \infty$:

$$\delta_N(\zeta_1, \zeta_2) = \left(1 + \frac{\zeta_2 e^{|k_2|\zeta_2}}{\zeta_2 - 1}\right) \frac{\zeta_1^N e^{|k_1|/\zeta_1}}{1 - \zeta_1} + \left(1 + \frac{e^{|k_1|/\zeta_1}}{1 - \zeta_1}\right) \frac{e^{|k_2|\zeta_2}}{\zeta_2^N (\zeta_2 - 1)}.$$

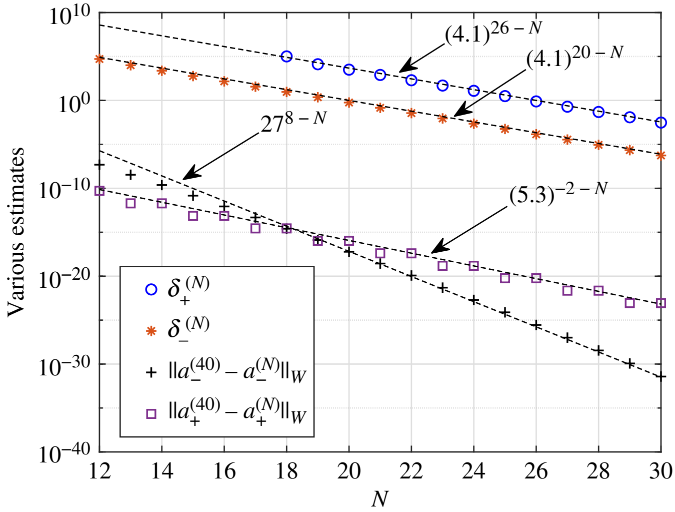

Figure 5. Comparison between $\delta_\pm^{(N)}$ with $\left\|a_\pm^{(40)} - a_\pm^{(N)}\right\|_W$ for example 6.2.

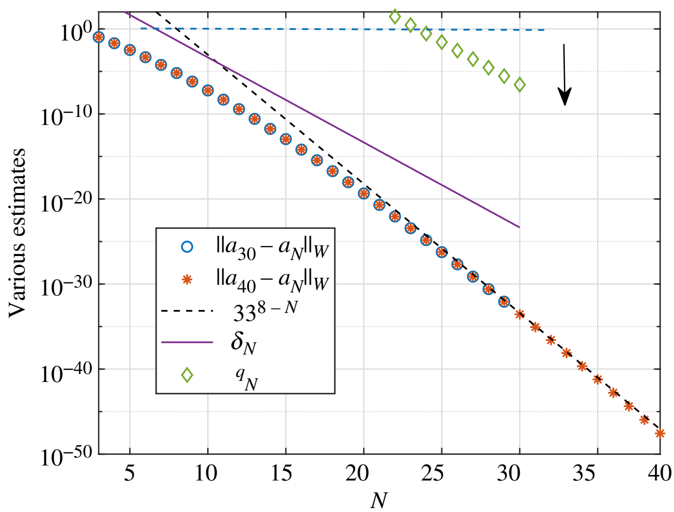

Figure 6. Accuracy of the estimate (6.2) for example 6.3.

In the case of $\theta = 2\nu + 1$, the stability of the indices occurs at $\rho_1 = \nu$, $\rho_2 = \nu + 1$. Unfortunately, estimates of the accuracy of the calculation of $\|a_\pm - a_\pm^{(N)}\|_W$ are not available, since the issue of the stability of P-normalisation has not been completed yet.

Example 6.3. Let $k_1 = 1$, $k_2 = 1/2$, $\theta = -7$, i.e. $\nu = -4$. If we put $\zeta_1 = 1/10$, $\zeta_2 = 10$, then

$$\delta_N = \left(\tfrac{10}{9}e^{10} + \tfrac{110}{81}e^{15} + \tfrac{1}{9}e^{5}\right)\frac{1}{10^N} \approx 4.4639 \cdot 10^{6-N}.$$

These values of the parameters ζ_1, ζ_2 give sufficiently small values of δ_N beginning $N = 13$ (electronic supplementary material, table S5).

Calculations with the package ExactMPF show that for $3 \le N \le 30$, the matrix functions $a_N(t)$ admit the stable factorisations with the indices $\rho_1 = -4$, $\rho_2 = -3$. The factorisation constructed by ExactMPF admits I_2-normalisations starting from $N = 4$. We carry out these normalisations by formula (2.2). Respective estimates in graphical form are presented in figure 6. Data related to this example are given in the electronic supplementary material, table S5. Then we calculated $\|\left(a_\pm^{(N)}\right)^{-1}\|_W$, and the parameter q_N (electronic supplementary material, table S5). The criterion $q_N < 1$ is fulfilled beginning $N = 24$. Hence, the matrix function $a(t)$ admits the stable factorisation with the indices $\rho_1 = -4$, $\rho_2 = -3$.

7. Conclusions

We have developed a method allowing us to confidently identify whether a strictly non-singular 2×2 matrix function possesses a stable set of partial indices while simultaneously performing an approximate factorisation. For this purpose, we use the ExactMPF package working within the Maple Software environment [18] allowing for exact factorisation of an arbitrary non-singular polynomial matrix function [19]. We have developed the efficient condition determining the stability region.

Having this crucial information, we have constructed a sequence of polynomial matrix functions approximating the given matrix function in such a way that the distance between the initial matrix function and a consequent member of the approximation sequence decreases, being simultaneously smaller than the stability radius of the respective polynomial matrix function. This, in turn, proves the convergence of the sequence and the stability of the developed factorisation.

The developed approach has been complemented with three different numerical examples demonstrating the method's efficiency. We would like to underline, however, that a theoretical estimate for the factor approximation in the case where the partial indices are stable but different (example 6.3) has still to be developed since there are no effective estimates for the radius of the neighbourhood where A and $\widetilde{A}$ admit the same P-normalisation.

Data accessibility. Supplementary material is available online [35].

Declaration of AI use. We have not used AI-assisted technologies in creating this article.

Authors' contributions. N.V.A.: conceptualisation, data curation, formal analysis, investigation, methodology, software, validation, writing—original draft, writing—review and editing; V.M.A.: conceptualisation, data curation, formal analysis, investigation, methodology, software, supervision, validation, writing—original draft, writing—review and editing; G.M.: conceptualisation, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, supervision, validation, visualisation, writing—original draft, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interests. We declare we have no competing interests.

Funding. N.V.A. and G.M. gratefully acknowledge the support of the EU H2020 grant MSCA RISE-2020-101008140-EffectFact. G.M. wants to thank the Isaac Newton Institute for Mathematical Sciences (INI) for their support and hospitality during the programme 'Mathematical theory and applications of multiple wave scattering' (MWS), where this paper was undertaken and supported by EPSRC grant no. EP/R014604/1. G.M. is also grateful for the funding received from the Simons Foundation that supported their visit to INI during January–June 2023 and participation in the MWS programme.

References

1. Gohberg IC, Krein MG. 1958 Systems of integral equations on a half-line with kernels depending on the difference of arguments. *Usp. Mat. Nauk.* **13**, 3–72.
2. Rogosin S, Mishuris G. 2016 Constructive methods for factorisation of matrix-functions. *IMA J. Appl. Math.* **81**, 365–391. (doi:10.1093/imamat/hxv038)

3. Faddeev LD, Takhtajan LA. 1987 *Hamiltonian methods in the theory of solitons*. Berlin, Heidelberg, Germany: Springer Verlag. (doi:10.1007/978-3-540-69969-9)
4. Adukov VM, Mishuris G. 2023 Utilisation of the exactmpf package for solving a discrete analogue of the nonlinear SchrőInger equation by the inverse scattering transform method. *Proc. R. Soc. A* **479**, 20220144. (doi:10.1098/rspa.2022.0144)
5. Litvinchuk GS. 2000 *Solvability theory of boundary value problems and singular integral equations with shift*. Dordrecht, the Netherlands: Springer. (doi:10.1007/978-94-011-4363-9)
6. Kisil AV, Abrahams ID, Mishuris G, Rogosin SV. 2021 The Wiener–Hopf technique, its generalizations and applications: constructive and approximate methods. *Proc. R. Soc. A* **477**, 20210533. (doi:10.1098/rspa.2021.0533)
7. Abrahams ID. 1987 Scattering of sound by two parallel semi-Infinite screens. *Wave Motion* **9**, 289–300. (doi:10.1016/0165-2125(87)90002-3)
8. Peake N, Abrahams ID. 2020 Sound radiation from a semi-Infinite lined duct. *Wave Motion* **92**, 102407. (doi:10.1016/j.wavemoti.2019.102407)
9. Kisil AV. 2018 An iterative Wiener–Hopf method for triangular matrix functions with exponential factors. *SIAM J. Appl. Math.* **78**, 45–62. (doi:10.1137/17M1136304)
10. Priddin MJ, Kisil AV, Ayton LJ. 2020 Applying an Iterative method numerically to solve matrix Wiener–Hopf equations with exponential factors. *Phil. Trans. R. Soc. A* **378**, 20190241. (doi:10.1098/rsta.2019.0241)
11. Clancey KF, Gohberg I. 1981 Factorization of matrix functions and singular integral operators. In *Operator theory, advances and applications*, vol. **3**. Basel, Switzerland: Springer. (doi:10.1007/978-3-0348-5492-4)
12. Litvinchuk GS, Spitkovskii IM. 1987 *Factorization of measurable matrix functions*. Basel, Switzerland; Boston, MA: Birkhauser. (doi:10.1515/9783112471722)
13. Gohberg IC, Feldman IA. 1974 Convolution equations and projection method for their solution. *Amer. Math. Soc. Transl. Math. Monogr.* **41**.
14. Prössdorf S. 1978 Some classes of singular equations. In *Series North-Holland mathematical library*, vol. **17**. Amsterdam, the Netherlands; New York, NY: North-Holland Pub. Co.
15. Spitkovsky IM. 1974 Stability of partial indices of the Riemann boundary problem with a strictly nonsingular matrix. *Dokl. Akad. Nauk. SSSR.* **218**, 46–49.
16. Adukov V. 1995 On factorization indices of strictly nonsingular 2×2 matrix function. *Integr. equ. oper. theory* **21**, 1–11. (doi:10.1007/BF01262988)
17. Ephremidze L, Spitkovsky I. 2020 On explicit Wiener–Hopf factorization of matrices in a vicinity of a given matrix. *Proc. Math. Phys. Eng. Sci.* **476**, 20200027. (doi:10.1098/rspa.2020.0027)
18. Maplesoft*Technical Computing Software for Engineers, Mathematicians, Scientists, Instructors and Students*. See https://www.maplesoft.com/ (accessed 7 July 2024).
19. Adukov VM, Adukova NV, Mishuris G. 2022 An explicit Wiener-Hopf factorisation algorithm for matrix polynomials and its exact realizations within exactmpf package. *Proc. R. Soc. A* **478**, 20210941. (doi:10.1098/rspa.2021.0941)
20. Adukov VM. 1992 Wiener–hopf factorization of meromorphic matrix-valued functions. *St. Petersbg. Math. J.* **4**, 51–69. (doi:10.1070/SM2009v200n08ABEH004030)
21. Adukov VM. 1999 On a factorization of analytic matrix-valued functions. *Theor. Math. Phys.* **118**, 255–263. (doi:10.4213/tmf704)
22. Gohberg IC, Lerer L, Rodman L. 1978 Factorization indices for matrix polynomials. *Bull. Amer. Math. Soc.* **84**, 275–277. (doi:10.1090/S0002-9904-1978-14473-5)
23. Adukov VM, Mishuris G, Rogosin SV. 2020 Exact conditions for preservation of the partial indices of a perturbed triangular 2×2 matrix function. *Proc. R. Soc. A* **476**, 20200099. (doi:10.1098/rspa.2020.0099)
24. Ahlfors LV. 1979 *Complex analysis*, 2nd edn. New York, NY: McGraw-Hill.
25. Adukov VM. 1998 Generalized inversion of block Toeplitz matrices. *Linear Algebra Appl.* **274**, 85–124. (doi:10.1016/S0024-3795(97)00304-2)
26. Gohberg IC, Kaashoek MA, Spitkovsky IM. 2003 An overview of matrix factorization theory and operator applications. *Oper. Theor. Adv. Appl.* **141**, 1–102. (doi:10.1007/978-3-0348-8003-9_1)

27. Gohberg I, Krupnik N. 1992 One-dimensional linear singular integral equations. In *Operator theory: advances and applications*, vol. **53**. Verlag Basel Boston Berlin: Birkhiiuser. (doi:10.1007/978-94-011-4363-9)
28. Mishuris G, Rogosin S. 2014 An asymptotic method of factorization of a class of matrix functions. *Proc. R. Soc. A* **470**, 20140109. (doi:10.1098/rspa.2014.0109)
29. Mishuris G, Rogosin S. 2016 Factorization of a class of matrix-functions with stable partial indices. *Math. Methods Appl. Sci.* **39**, 3791–3807. (doi:10.1002/mma.3825)
30. Mishuris G, Rogosin S. 2018 Regular approximate factorization of a class of matrix-function with an unstable set of partial indices. *Proc. R. Soc. A* **474**, 20170279. (doi:10.1098/rspa.2017.0279)
31. Adukov VM. 2022 Normalization of Wiener–Hopf factorization for matrix functions and its application. *Ufimsk. Math. J.* **14**, 1–13. (doi:10.13108/2022-14-4-1)
32. Adukova NV. 2022 Stability of factorization factors of Wiener–Hopf factorization of matrix functions.. *Bull. South Ural Stat. Ser. Math. Mech. Phys.* **14**, 5–12. (doi:10.14529/mmph220101)
33. Adukov VM. 2022 Normalization of Wiener–Hopf factorisation for matrix functions with distinct partial indices. *Trans. A Razmadze Math. Inst.* **176**, 313–321.
34. Adukova NV, Dilman VL, South Ural State University (national research university). 2021 Stability of factorization factors of the canonical factorization of Wiener–Hopf matrix functions. *Bull. South Ural Stat. M.M.Ph.* **13**, 5–13. (doi:10.14529/mmph210101)
35. Adukova N, Adukov V, Mishuris G. 2024 Data from: An effective criterion for a stable factorisation of strictly non-singular 2 × 2 matrix functions: use of the exactmpf package. Figshare. (doi:10.6084/m9.figshare.c.7318642)

royalsocietypublishing.org/journal/rspa *Proc. R. Soc. A* **480:** 20240116