SCIENCE CHINA
Mathematics

• ARTICLES •

December 2021 Vol. 64 No. 12: 2703–2718
https://doi.org/10.1007/s11425-020-1706-5

A class of iterative greedy algorithms related to Blaschke product

Tao Qian[1], Lihui Tan[2,*] & Jiecheng Chen[3]

[1]*Macao Institute of Systems Engineering, Macau University of Science and Technology, Macao, China;*
[2]*School of Applied Mathematics, Guangdong University of Technology, Guangzhou* 510006*, China;*
[3]*Department of Mathematics, Zhejiang Normal University, Jinhua* 321004*, China*

Email: tqian1958@gmail.com, lihuitan@ymail.com, jcchen@zju.edu.cn

Received March 2, 2020; accepted May 10, 2020; published online August 12, 2021

Abstract Möbius transforms, Blaschke products and starlike functions as typical conformal mappings of one complex variable give rise to nonlinear phases with non-negative phase derivatives with the latter being defined by instantaneous frequencies of signals they represent. The positive analytic phase derivative has been a widely interested subject among signal analysts (see Gabor (1946)). Research results of the positive analytic frequency and applications appears in the literature since the middle of the 20th century. Of the positive frequency study a directly related topic is positive frequency decomposition of signals. The mainly focused methods of such decompositions include the maximal selection method and the Blaschke product unwinding method, and joint use of the mentioned methods. In this paper, we propose a class of iterative greedy algorithms based on the Blaschke product and adaptive Fourier decomposition. It generalizes the Blaschke product unwinding method by subtracting constants other than the averages of the remaining functions, aiming at larger winding numbers, and subtracting n-Blaschke forms of the remaining functions, aiming at generating larger numbers of zero-crossings, to fast reduce energy of the remaining terms. Furthermore, we give a comprehensive and rigorous proof of the converging rate in terms of the zeros of the remainders. Finite Blaschke product methods are proposed to avoid the infinite phase derivative dilemma, and to avoid the computational difficulties.

Keywords complex Hardy space, Möbius transform, Blaschke product, rational orthogonal system, Takenaka-Malmquist system, mono-component, adaptive Fourier decomposition, unwinding Blaschke expansion

MSC(2020) 30B50, 30A10, 42C10, 65T99

Citation: Qian T, Tan L H, Chen J C. A class of iterative greedy algorithms related to Blaschke product. Sci China Math, 2021, 64: 2703–2718, https://doi.org/10.1007/s11425-020-1706-5

1 Introduction

1.1 Analytic signals with positive phase derivatives: Momo-components

The recently developed empirical mode decomposition (EMD) has provided a general method for examining the time frequency distribution, and has been successfully applied to various areas, e.g., medical studies, meteorology, geophysical studies and image analysis (see [19]). In spite of considerable success of EMD, all of the EMD algorithms are based on empirical and heuristic procedures that make them

* Corresponding author

hard to analyze mathematically, and EMD may suffer from mode mixing, detrend uncertainty, aliasing and end effect artefacts (see [23]). In order to develop the related mathematical theory of EMD, some researchers hope to decompose a complicated signal $f(t)$ into the following form:

$$f(t) = \sum_{k=1}^{n} a_k(t) \cos \theta_k(t) + r(t), \tag{1.1}$$

where $r(t)$ is the residual component, and $a_k(t) \cos \theta_k(t)$ in (1.1) should satisfy the equation

$$(H a_k(\cdot) \cos \theta_k(\cdot))(t) = a_k(t) \sin \theta_k(t), \quad a_k(t) \geqslant 0, \quad \theta'_k(t) \geqslant 0. \tag{1.2}$$

Here, H stands for the Hilbert transform in (1.2). The functions satisfying (1.2) are referred to as *mono-components* (see [7,9]) and well developed in the literature (see [27,31]), and they are analytic signals with the positive instantaneous frequency. Some algorithms are proposed for decomposing a signal into the model (1.1) based on the theories of Fourier analysis and wavelet analysis (see [16,20,40]). In this paper, we will make use of complex analysis methods to establish the rigorous mathematical analysis for (1.2) and (1.1).

Denote by

$$\tau_a(z) = \frac{z-a}{1-\bar{a}z}$$

the canonical Möbius transform at $a \in \boldsymbol{D}$, where $\boldsymbol{D}$ denotes the unit disc. τ_a maps $\boldsymbol{D}$ to $\boldsymbol{D}$, and its boundary $\partial\boldsymbol{D} = \{z = \mathrm{e}^{\mathrm{i}t}, 0 \leqslant t < 2\pi\}$ to itself. It follows that there exists a real-valued function $\theta_a(t)$ such that $\tau_a(\mathrm{e}^{\mathrm{i}t}) = \mathrm{e}^{\mathrm{i}\theta_a(t)}$. Since the mapping on the boundary is one to one and onto, and keeps the orientation, the real-valued function $\theta_a(t)$ is strictly increasing. It concludes that $\theta'_a(t) \geqslant 0$ for all t. Through a simple computation we have

$$\theta'_a(t) = \frac{1-|a|^2}{|\mathrm{e}^{\mathrm{i}t} - a|^2} > 0, \quad \forall\, t \in [0, 2\pi).$$

The right-hand side of the above equality is, in fact, the 2π-multiple of the Poisson kernel at a (see [18]). A finite Blaschke product (a Blaschke product with finite zeros) is the product of a finite number of Möbius transforms, and its boundary phase derivative, as a function, is the 2π-multiple of the sum of the corresponding Poisson kernels, and thus is a strictly positive function. Some early studies of the analytic phase derivative were found in [10,26]. Some recent studies on analytic signals with nonnegative instantaneous frequencies are contained in [9,27,28,31,34,43,45,47]. The circle context corresponds to periodic signals. The line context corresponds to signals defined in the whole time range that uses the Hardy space of the upper-half complex plane. The theories of the two contexts are parallel.

Let f be an analytic function in the unit disc $\boldsymbol{D}$. Then the amplitude-phase representation of f may be written in the polar coordinates as $f(r\mathrm{e}^{\mathrm{i}t}) = \rho_r(t)\mathrm{e}^{\mathrm{i}\theta_r(t)}$ with $\rho_r(t) \geqslant 0$, and $\theta_r(t)$ being real-valued, $t \in [0, 2\pi)$. Through a direct computation one has

$$\theta'_r(t) = \mathrm{Re}\left\{ z\frac{f'(z)}{f(z)} \right\}, \quad z = r\mathrm{e}^{\mathrm{i}t}, \quad 0 < r < 1.$$

We note that on the boundary a phase function may not be well defined but a phase derivative function may be defined. In [27], it is shown that for any inner function $f(r\mathrm{e}^{\mathrm{i}t}) = \rho_r(t)\mathrm{e}^{\mathrm{i}\theta_r(t)}$ the limits in (1.3), as a consequence of the Wolff-Julia-Carathéodory theorem, exist for all t but possibly a Lebesgue null set. In the inner function case, the boundary phase derivative $\theta'(t)$ is well defined and positive, as

$$\theta'(t) \triangleq \lim_{r\to 1-} \theta'_r(t) \geqslant 0, \quad \text{a.e.} \tag{1.3}$$

It is noted that for the outer functions, under mild conditions ensuring absolute continuity, the boundary phase derivatives defined through the same limit on the left-hand side of (1.3) satisfy

$$\int_0^{2\pi} \theta'(t)dt = 0 \tag{1.4}$$

(see [27]). This shows that unless $\theta'(t)$ is identical with the zero function, the phase derivative must be negative in a set of the positive Lebesgue measure. The phase derivative is closely related to the angular derivative that has been closely studied in a number of papers (see [1, 14, 15, 17, 27]). For the higher dimensions, the subject frequency analysis and its practice have not yet been fully developed. For the relevant references we refer to [30, 37, 46].

1.2 Momo-component decompositions of signals in the Hardy space

Closely related study is mono-component expansions of signals. The Fourier series expansion, and essentially the Fourier inversion formula as well, are positive- (or negative-) constant-frequency expansions. Mono-component expansions are generalizations, in both the concept and the methodology, of the Fourier type expansions. Mono-component decompositions can be divided into two categories, or rather strategies, since they are methodology-related, of which both are of a number of variations. They may be called the adaptive Fourier decomposition (abbreviated as AFD), as they are related to boundary limits of analytic functions with the polynomials and rational functions are particular cases, and adaptive. Of the two strategies, the first is based on the reproducing kernel property, also called the Szegö kernel, of the Hardy space. This kernel approach brings the interpolation property and creates zero-crossings. Through a maximal selection principle, the interpolating approximation series can fast get close in the energy sense to the given signal. The standard remainders generated through the approximation belong to the Beurling-Lax shift-invariant subspaces. This interpolation approach in some general spaces can be used to define Blaschke products. The Blaschke product as an operator has been extended to multivariate contexts in such a way (see [4, 5]).

The terminology AFD was first used with the maximal selection approach (see [35]), and later renamed as the *core AFD*, being the central algorithm building block of the *cyclic* and *n-best AFD*, and *unwinding AFD*. The maximal selection type AFD was lately extended to reproducing kernel Hilbert spaces possessing the so called *boundary vanishing property* (BVP) (see [29]). When in a Hilbert space there are no functions playing the role like inner functions in the sifting process, then the best bet would be a maximal selection or a weak-maximal selection principle. The second strategy is the *Blaschke unwinding expansion*, based on the inner-outer function factorization, first explored with views of applications in the PhD dissertation of Nahon [25] in 2000 at Yale university under the guidance of Coifman [12] (see [28]). The earlier work of Weiss and Weiss [44] hints this development.

The unwinding methodology was independently studied in [28, 32]. The two strategies are, in fact, closely related: Both of them are related to zeros of functions and result in Blaschke products. A detailed description is as follows: Denote by

$$e_a = \frac{\sqrt{1-|a|^2}}{1-\overline{a}z}, \quad a \in D$$

the reproducing kernel of the Hardy space. At each decomposition step, denote by $P_n f$ the projection of f onto $\operatorname{span}\{e_{a_1}, \ldots, e_{a_n}\}$. Then $f - P_n f$ has the zeros $a_1, \ldots, a_n$, due to the reproducing kernel property and the orthogonality, and the remainder has the factor $(z-a_1)\cdots(z-a_n)$ with the latter giving rise to the corresponding Blaschke product factor of the remainder function. From the intrinsic and uniqueness point of view, we want the convergence to be fast. This would be gained, at least partially, from a maximal selection of an n-tuple of parameters $a_1, \ldots, a_n$. There is a delicate treatment when the parameters in a maximal selection of parameter n-tuples have multiples.

We note that in higher dimensions with the non-scalar-valued function there does not exist an effective inner-outer function theory and factorization, but one can adopt the above mentioned process to get a Blaschke product related expansion with a great similarity with the core AFD resulting in a Takenaka-Malmquist-like system (see [36]). For scalar-valued Blaschke products defined in the polydisc we refer to [38]. For functions defined on the unit disc with matrix values we refer to [4]. For matrix-valued functions defined on the polydisc we refer the reader to [3]. In the present paper we work only on the unit disc context with scalar-valued functions, and will concentrate on the Blaschke unwinding methodology via factorization.

Fast convergence of Blaschke unwinding expansion is intimately related to the following result in digital signal processing (DSP) (see [14, 28]), phrased as the *energy front-loading property of outer functions* (minimum-phase physically realizable signals): $F = IG$, where $F, G \in H^2(\mathbf{D})$ and I is an inner function. Let $F(z) = \sum_{k=0}^{\infty} c_k z^k$ and $G(z) = \sum_{k=0}^{\infty} d_k z^k$. As a consequence of the Plancherel theorem,

$$\sum_{k=0}^{\infty} |d_k|^2 \geqslant \sum_{k=0}^{\infty} |c_k|^2.$$

One, however, can prove that for all integers $N > 0$, $\sum_{k=0}^{N} |d_k|^2 \geqslant \sum_{k=0}^{N} |c_k|^2$. This shows that the Fourier series of the unwound function G converges in a faster pace than that of the original F. The DSP result suggests that the more zero factors one factorizes out, the faster is the convergence rate of the Fourier series of the remaining function.

1.3 The unwinding expansion by subtracting the averages: The existing model

The following *unwinding Blaschke expansion* was first studied by Coifman and Nahon (see [25]). However, the study remained unpublished until 2016 (see Coifman and Steinerberger [12]). In 2010, the literature [28] introduced what was called *intrinsic mono-component decomposition* of analytic signals. [28, Remark 4.4], in particular, presents the same unwinding method of [12] together with a proof of its H^2-convergence. In 2013, the terminology *unwinding AFD* (UAFD) was used for this maximal selection combined unwinding method, coinciding with the name given in [25]. This method is observed to have a rapid convergence rate while giving rise to positive frequency constructive blocks.

This study was further expanded in [11, 13]. A number of applications of the unwinding expansion were explored with practical algorithms, including some in signal analysis, system identification and image processing (see [21, 24]).

In self-explanatory notations the Blaschke unwinding method is formulated as follows. Let $F = F_0 \in H^2(\boldsymbol{D})$. In each of the following iterations, we have

$$G_k(z) = G_k(z) - G_k(0) + G_k(0) = zB_{k+1}G_{k+1}(z) + G_k(0), \tag{1.5}$$

where B_{k+1} is the Blaschke product part of F_{k+1} and G_{k+1} is the product of the outer factor and the singular inner function factor parts of F_{k+1}. We leave out the z factor in each of the sifting steps, except the case where $k = 0$. It follows

$$\begin{aligned} F(z) = F_0(z) &= B_0(z)(G_0(z) - G_0(0) + G_0(0)) = c_0B_0(z) + zB_0(z)F_1(z) \\ &= c_0B_0(z) + c_1zB_0(z)B_1(z) + z^2B_0(z)B_1(z)B_2(z)G_2(z) \\ &= \sum_{k=0}^{\infty} c_k z^k \prod_{j=0}^{k} B_j(z), \end{aligned} \tag{1.6}$$

where $c_k = G_k(0)$ is the average of the function G_k. We note that the standard unwinding process given above produces an orthogonal system. In proving this the factors z^k in the decomposition components $z^k \prod_{j=0}^{k} B_j$ play a crucial role. The obtained series converges in the square norm sense to the function on the left-hand side (see [12, 28]). For polynomial functions $F(z)$ the decomposition process ends after a finite number of iterations. By noticing the similarity between the backward-shift process to obtain the Taylor series and the unwinding process (1.5), the above process will be phrased as *sifting* through the *backward-shift operation* (*through subtracting the averages*). For the concept of the *generalized backward-shift operation*, see [35].

To the unwinding Blaschke expansion process there would exist the following concerns:

(1) The calculation of the Hardy space function F factorization $F = BG$, where B is the Blaschke product part and the calculation of the phase derivatives of B both have computational difficulties. The task of obtaining the Blaschke product part of F is equivalent to finding out all the zeros of F, and, of course, by no means easy. It is obviously more difficult than finding the roots of a general polynomial

of degree greater than 4. If, instead, one chooses to first compute the function G, as the product of the outer function and the singular inner function parts, it involves computation of the Hilbert transform with the latter being a continuing research topic in numerical computation, due to singularity of the Hilbert transform. The calculation of an outer function is, in fact, harder than just computing the Hilbert transform, as not only the Hilbert kernel is singular, but also the integrand function $\ln|F(\mathrm{e}^{\mathrm{i}t})|$ is singular at all the points that make $F(\mathrm{e}^{\mathrm{i}t}) = 0$. Algorithms have been developed for both of these computational strategies: First computing B or first computing G. For computing B or the zeros, see, for example, [22]. The literature promotes a finite zero method that leads to convenience of the unwinding expansion and, at the same time, obtains the intrinsic frequencies of the signal as the sum of the involved Poisson kernels. For computations of the outer functions via the Hilbert transform, see [25, 41]. A recent result explores an optimization method to extract out the outer function part without computing the Hilbert transform (see [42]). A more recent study computes the Hilbert transform by using a mechanical quadrature method (see [41]). The strategy of first finding out outer functions does not lose anything theoretically in itself but would lay difficulties of computing out the boundary phase derivatives and may fall in the dilemma of getting a Blaschke product that has infinite boundary phase derivatives at all the points of the circle (see below).

(2) The literature [27] constructed an inner function whose phase derivatives are a.e. infinite. The literature [1] and [39, p.184] constructed interpolating Blaschke products whose cluster set E is identical with $\partial\boldsymbol{D}$ and they have the infinite phase derivative at all the points of $\partial\boldsymbol{D}$. Blaschke products of such a type arouse considerable interest due to its connections with compactness of the composition operator in the Hardy space (see [39]). Such a substantial infinite phase derivative phenomenon, however, lays an ill impact to frequency analysis of unwinding Blaschke expansion: The expansion may automatically generate Blaschke products that have substantial infinite phase derivatives. In order to carry on meaningful frequency analysis one would need to restrict oneself to finite Blaschke product decompositions. The frequency decomposition of a signal does not seem to be unique.

The present study will generalize the existing unwinding processes (1.6) and (1.5). They at the same time are improvements of (1.6) and (1.5) with regard to the concerns just expressed.

1.4 Generalizations of the unwinding method: Main results of the paper

Our first main result, Theorem 2.4, deals with the cases where in the splitting

$$G_k(z) = G_k(z) - G_k(a_k) + G_k(a_k)$$

one may take $a_k \in \boldsymbol{D}$ other than $a_k = 0$. The a_k's may be selected to have the maximal winding (index) numbers. The technical difficulty is that in such a case the terms in the corresponding partial sum $\tilde{S}_n F$ for each n may not be orthogonal to each other. We overcome the difficulty by using a related interpolation property. With the convergence results we also prove convergence rates of the exponential type.

Our second main result Theorem 3.1 deals with the cases where instead of subtracting single values $G_k(a_k)$ one subtracts functions $\mathcal{G}(G_k)(z)$ generating prescribed or non-prescribed but induced zeros. In the induced zero case the functions $\mathcal{G}(G_k)(z)$ may be, for example, selected as Blaschke forms giving rise to maximal energy at each step.

2 Sifting process through subtracting function values other than the averages

The Blaschke unwinding method is based on subtracting the averages of the recursively generated functions G_k:

$$G_k(z) = G_k(z) - G_k(0) + G_k(0),$$

where

$$G_k(0) = \frac{1}{2\pi}\int_0^{2\pi} G_k(\mathrm{e}^{\mathrm{i}t})dt.$$

As a variation we can proceed with subtracting a function value $G_k(a_k)$, where a_k is not necessarily 0, to create more zeros in the remainders $G_k(z) - G_k(a_k), k = 0, 1, 2, \dots$ We may take, for example, for r being a fixed number in $(0,1)$, $a_k = \arg\max_{a\in r\boldsymbol{D}}\{\mathrm{Ind}_{G_k(\partial(r\boldsymbol{D}))}(a)\}$. There are maybe more than one a_k satisfying this relation. Then we have a different decomposition of the signal

$$\begin{aligned} F(z) &= B_0(z)[G_0(z) - G_0(a_0)] + B_0(z)G_0(z) \\ &= B_0(z)B_1(z)G_1(z) + B_0(z)G_0(a_0) \\ &= \cdots \\ &= G_0(a_0)B_0(z) + \sum_{k=1}^{n} G_k(a_k)\prod_{j=0}^{k} B_j(z) + G_{n+1}(z)\prod_{j=0}^{n+1} B_j(z), \end{aligned} \tag{2.1}$$

where $G_k(z) - G_k(a_k) = B_{k+1}(z)G_{k+1}(z)$, a_k is a point in $\boldsymbol{D}$, $B_0(z)$ is the Blaschke product defined by the zeros of $F(z)$, $B_{k+1}(z)$ is the Blaschke product defined by the zeros of $G_k(z) - G_k(a_k)$ (the zeros are counted according to their multiplicities) and $G_k(z)$ is a product of an outer function and an inner function. In order to design a computer algorithm easily, we restrict that each $B_k(z)$ in (2.1) is a finite Blaschke product. Let

$$E_{B_k}^{N_k} = \{b_j^{(k)}\}_{j=1}^{N_k}$$

be the set of all the zeros of $B_k(z)$ in the unit disc, where $b_1^{(k)} = a_{k-1}$ for $k \geqslant 1$. Then the sum of the first $n+1$ terms in the unwinding procedure (2.1) can be written as

$$\begin{aligned} (\widetilde{S}_nF)(z) &= G_0(a_0)B_0(z) + \sum_{k=1}^{n} G_k(a_k)\prod_{j=0}^{k} B_j(z) \\ &= G_0(a_0)\prod_{i=1}^{N_0}\frac{z - b_i^{(0)}}{1 - \overline{b_i^{(0)}}z} + \sum_{k=1}^{n} G_k(a_k)\prod_{j=0}^{k}\prod_{i=1}^{N_j}\frac{z - b_i^{(j)}}{1 - \overline{b_i^{(j)}}z}. \end{aligned}$$

When all a_k's are chosen to be 0, the components $\prod_{j=0}^{k} B_j$ in (2.1) are orthogonal to each other for z^k is a factor of $\prod_{j=0}^{k} B_j$. By virtue of the orthogonality, the result that $(\widetilde{S}_nF)(z)$ converges to $F(z)$ as $n \to \infty$ is proved in [28]. We cite this proof here to show how different it is between the two convergence proofs: One is for all $a_k = 0$, and the other is for all the selections of $a_k \in \boldsymbol{D}$.

Theorem 2.1. *If at each of the sifting steps of the unwinding process* (1.6) *an arbitrary sub-Blaschke product but including the z term is factorized out, then the resulted expansion* (1.6) *is orthogonal and converges to the originally given analytic signal.*

Proof. By (1.6), we know that

$$\begin{aligned} F(z) = F_0(z) &= B_0(z)[G_0(z) - G_0(0)] + B_0(z)G_0(0) \\ &= \left[G_0(0)B_0(z) + \sum_{k=1}^{n} G_k(0)z^k\prod_{j=0}^{k} B_j(z)\right] + \left[z^{n+1}\prod_{j=0}^{n+1} B_j(z)\right]G_{n+1}(z) \\ &=: T_n(z) + A_n(z), \end{aligned}$$

where $B_k(z)$ is the sub-Blaschke product of $G_{k-1}(z) - G_{k-1}(0)$ excluding the z factor. Now we show that T_n and A_n are orthogonal. In fact, for $k \leqslant n$, by Cauchy's integral theorem, we have

$$\begin{aligned} \left\langle A_n, \prod_{j=0}^{k} z^jB_j\right\rangle &= \frac{1}{2\pi}\int_0^{2\pi} \mathrm{e}^{\mathrm{i}(n+1-k)t}B_{k+1}(\mathrm{e}^{\mathrm{i}t})\cdots B_{n+1}(\mathrm{e}^{\mathrm{i}t})G_{n+1}(\mathrm{e}^{\mathrm{i}t})dt \\ &= \frac{1}{2\pi\mathrm{i}}\oint_{|z|=1} z^{n-k}B_k(z)\cdots B_{n+1}(z)G_{n+1}(z)dz = 0, \end{aligned}$$

as $n-k \geqslant 0$. Let $S_n(z)$ be the n-th partial sum of $F(z)$ collecting all the terms of the form $c_k z^k$ with $k \leqslant n$ in the Fourier expansion of $T_n(z)$, and hence

$$T_n(z) = \sum_{k=0}^{n} G_k(0) z^k \prod_{j=0}^{k} B_j(z) = S_n(z) + N_n(z),$$

where N_n collects all the terms of the form $a_k z^k$ with $k > n$ in the expansion of $T_n(z)$. The orthogonality between S_n and N_n is obvious. Due to the orthogonality, we have

$$\|F\|^2 = \|T_n\|^2 + \|A_n\|^2 = \|S_n\|^2 + \|N_n\|^2 + \|A_n\|^2.$$

The fact

$$\lim_{n\to\infty} S_n = F$$

forces

$$\lim_{n\to\infty} (\|N_n\|^2 + \|A_n\|^2) = 0.$$

Hence

$$\lim_{n\to\infty} \|A_n\|^2 = 0.$$

The last relation is equivalent to

$$\lim_{n\to\infty} \|F - T_n\|_2 = 0.$$

The proof is completed. □

The proof shows that the crucial point for the orthogonality and thus for the convergence as well is the presence of the factors z^k. In general cases where a_k's are not necessarily all chosen to be zero, the components in the decomposition (2.1) may not be orthogonal to each other. The above proof is thus not valid for general cases.

To give a proof of the convergence for the general cases we first put all the zeros together: Let

$$\begin{aligned} E^{(n+1)}_{N^{(n+1)}} &= E^{N_0}_{B_0} \cup E^{N_1}_{B_1} \cup \cdots \cup E^{N_n}_{B_n} \cup E^{N_{n+1}}_{B_{n+1}} \\ &= \{b_1^{(0)}, \ldots, b_{N_0}^{(0)}, b_1^{(1)}, \ldots, b_{N_1}^{(1)}, \ldots, b_1^{(n+1)}, \ldots, b_{N_{n+1}}^{(n+1)}\} \\ &= \{b_0, b_1, \ldots, b_{N^{(n+1)}-1}\}. \end{aligned} \tag{2.2}$$

We found that the function

$$F(z) - (\widetilde{S}_n F)(z) = \left[\prod_{k=0}^{n+1} B_k(z)\right] G_{n+1}(z) = \left[\prod_{k=0}^{n+1}\prod_{j=1}^{N_k} \frac{z - b_j^{(k)}}{1 - \overline{b_j^{(k)}} z}\right] G_{n+1}(z)$$

satisfies the following interpolation conditions:

$$F^{(l(b_k)-1)}(b_k) = (\widetilde{S_n}F)^{(l(b_k)-1)}(b_k), \tag{2.3}$$

where $k = 0, 1, \ldots, N^{(n+1)} - 1$, and b_k is the $k+1$ term of the set $E^{(n+1)}_{N^{(n+1)}}$, $N^{(n+1)} = \sum_{k=0}^{n+1} N_k$ and $l(b_k)$ denotes the number of repeating times of b_k in the $(k+1)$-tuple sets $\{b_0, b_1, \ldots, b_k\}$. Let us define

$$\tilde{e}_{b_k}(z) = \begin{cases} \dfrac{1}{(1 - \overline{b_k} z)^{l(b_k)}}, & b_k \neq 0, \\ z^{(l(b_k)-1)}, & b_k = 0. \end{cases}$$

We say that $\tilde{e}_{b_k}(z)$ is *a higher order Szegö kernel* if $l(b_k) > 1$, and otherwise *a Szegö kernel*. The system $\{\tilde{e}_{b_0}(z), \tilde{e}_{b_1}(z), \ldots, \tilde{e}_{b_n}(z)\}$ is called the partial fraction system generated by $\{b_0, b_1, \ldots, b_n\}$. By the Gram-Schmidt (G-S) orthogonalization process applied to the *partial fractions* $\{\tilde{e}_{b_k}(z)\}_{k=0}^{n}$, a rational orthonormal system (referred as Takenaka-Malmquist (TM) system), denoted by $\{\mathcal{B}_k(z)\}_{k=0}^n$, can be obtained, where

$$\mathcal{B}_0(z) = \frac{\sqrt{1-|b_0|^2}}{1-\overline{b_0}z}, \quad \mathcal{B}_k(z) = \frac{\sqrt{1-|b_k|^2}}{1-\overline{b_k}z}\prod_{j=0}^{k-1}\frac{z-b_j}{1-\overline{b_j}z}. \tag{2.4}$$

Orthonormality of the TM system $\{\mathcal{B}_k(z)\}_{k=0}^\infty$ can be easily proved by using the Cauchy theorem: For $n<m$,

$$\begin{aligned}
\langle \mathcal{B}_m(\mathrm{e}^{\mathrm{i}t}), \mathcal{B}_n(\mathrm{e}^{\mathrm{i}t})\rangle &= \frac{1}{2\pi}\int_0^{2\pi}\mathcal{B}_m(\mathrm{e}^{\mathrm{i}t})\overline{\mathcal{B}_n(\mathrm{e}^{\mathrm{i}t})}dt \\
&= \frac{1}{2\pi\mathrm{i}}\oint_{|z|=1}\frac{\sqrt{1-|a_m|^2}}{1-\overline{a}_m z}\prod_{l=1}^{m-1}\frac{z-a_l}{1-\overline{a}_l z}\overline{\frac{\sqrt{1-|a_n|^2}}{1-\overline{a}_n z}\prod_{l=1}^{n-1}\frac{z-a_l}{1-\overline{a}_l z}}\frac{dz}{z} \\
&= \frac{1}{2\pi\mathrm{i}}\oint_{|z|=1}\frac{\sqrt{1-|a_m|^2}\sqrt{1-|a_n|^2}}{1-\overline{a}_m z}\prod_{l=n}^{m-1}\frac{z-a_l}{1-\overline{a}_l z}\frac{1}{z-a_n}dz \\
&= \frac{1}{2\pi\mathrm{i}}\oint_{|z|=1}\frac{\sqrt{1-|a_m|^2}}{1-\overline{a}_m z}\frac{\sqrt{1-|a_n|^2}}{1-\overline{a}_n z}\prod_{l=n+1}^{m-1}\frac{z-a_l}{1-\overline{a}_l z}dz \\
&= 0.
\end{aligned}$$

We need some technical preparation for proving the general convergence theorem. For the following two lemmas the reader can also be referred to [2, 8, 36].

Lemma 2.2. *Let $\phi_n(z) = \prod_{k=0}^n \frac{z-b_k}{1-\overline{b}_k z}$ be an n-Blaschke product for $n \geqslant 0$. Then we have*

$$\sum_{k=0}^n \overline{\mathcal{B}_k(\xi)}\mathcal{B}_k(z) = \frac{1-\overline{\phi_n(\xi)}\phi_n(z)}{1-\overline{\xi}z}, \quad \overline{\xi}z \neq 1, \quad |\xi| = 1 \tag{2.5}$$

and

$$\sup_{|z|=R}\frac{1}{|\phi_n(z)|} \leqslant \exp\left\{-\frac{R-1}{2R}\sum_{j=0}^n(1-|b_j|)\right\}, \quad R>1. \tag{2.6}$$

Proof. To make the paper self-contained we give outlines of the proofs. The detailed proofs can be found in [2]. The Christoffel-Darboux identity (2.5), from the Beurling direct sum decomposition point of view, can be easily obtained. In fact,

$$H^2(\boldsymbol{D}) = \mathrm{span}\{\mathcal{B}_k\}_{k=0}^n \oplus \phi_n H^2(\boldsymbol{D}).$$

The reproducing kernel of the shift-invariant subspace $\phi_n H^2(\boldsymbol{D})$ is $\frac{\overline{\phi_n(\xi)}\phi_n(z)}{1-\overline{\xi}z}$, and the kernel of the component in the span of $\{\mathcal{B}_k\}_{k=0}^n$ is, therefore,

$$\frac{1-\overline{\phi_n(\xi)}\phi_n(z)}{1-\overline{\xi}z}.$$

Hence the representation (2.5) is proved.

For proving the inequality (2.6), let $z = R\mathrm{e}^{\mathrm{i}\theta}$, $b_k = |b_k|\mathrm{e}^{\mathrm{i}\omega_k}$ and $\psi_k = \theta - \omega_k$. Then for $R>1$, a simple algebraic manipulation yields that

$$\begin{aligned}
\left|\frac{z-b_k}{1-\overline{b_k}z}\right|^2 &= \left|\frac{R\mathrm{e}^{\mathrm{i}\psi_k}-|b_k|}{1-R|b_k|\mathrm{e}^{\mathrm{i}\psi_k}}\right|^2 \\
&= 1-\frac{(1-|b_k|^2)(1-R^2)}{1+|b_k|^2R^2-2R|b_k|\cos\psi_k} \\
&\leqslant 1-(1-R)(1-|b_k|) \\
&\leqslant \exp\left[-(1-R)(1-|b_k|)\right],
\end{aligned}$$

where we used the elementary inequality $\mathrm{e}^{-x} \geqslant 1 - x$ for all x. Hence, we have

$$\sup_{|\xi|=R} \frac{1}{|\phi_n(\xi)|} = \sup_{|\xi|=\frac{1}{R}} |\phi_n(\xi)| \leqslant \exp\left\{ -\frac{R-1}{2R} \sum_{j=0}^{n} (1 - |b_j|) \right\}.$$

This completes the proof. □

Denote by $S_n F$ the partial sum

$$(S_n F)(z) = \sum_{k=0}^{n} \langle F, \mathcal{B}_k \rangle \mathcal{B}_k(z).$$

Based on (2.5) and hyperbolic non-separability condition, we have the following result (see [8]).

Lemma 2.3. *Let $\{\mathcal{B}_k(z)\}_{k=0}^{\infty}$ be the TM system determined by the set $E = \{b_0, b_1, \ldots\}$. If it holds that the hyperbolic non-separability condition $\sum_{i=0}^{\infty}(1 - |b_i|) = \infty$, then for any $F(z) \in H^2(\boldsymbol{D})$, we have*

$$\lim_{n\to\infty} \|(S_n F)(\mathrm{e}^{\mathrm{i}t}) - F(\mathrm{e}^{\mathrm{i}t})\|^2 = 0.$$

Moreover, $(S_n F)(z) = \sum_{k=0}^{n} \langle F(\mathrm{e}^{\mathrm{i}t}), \mathcal{B}_k(\mathrm{e}^{\mathrm{i}t}) \rangle \mathcal{B}_k(z)$ is the unique function in the rational space

$$\mathcal{R}_{n,n+1} = \left\{ \frac{P_n(z)}{\prod_{k=0}^{n} (1 - \overline{b_k} z)} \right\} \tag{2.7}$$

satisfying the interpolation condition $(S_n F)^{(l(b_k)-1)}(b_k) = F^{(l(b_k)-1)}(b_k)$ for $k = 0, 1, \ldots, n$, where $P_n(z)$ is an arbitrary polynomial with order not exceeding n.

Now we are ready to prove that $\widetilde{S}_n F$ converges to F as $n \to \infty$.

Theorem 2.4. *Let $F(z)$ be a nonzero function in $H^2(\boldsymbol{D})$ and $\{a_1, \ldots, a_n, \ldots\}$ be any sequence in $\boldsymbol{D}$ that induces the unwinding decomposition (2.1). Then*

$$(\widetilde{S}_n F)(z) = \frac{1}{2\pi \mathrm{i}} \oint_{|\xi|=1} \frac{F(\xi)[1 - \psi_n(z)\overline{\psi_n(\xi)}]}{(\xi - z)} d\xi + \frac{\psi_n(z)}{2\pi \mathrm{i}} \oint_{|\xi|=1} \frac{F(\xi)\overline{\psi_n(\xi)}}{(\xi - a_n)} d\xi. \tag{2.8}$$

If there exists r_0 such that $0 < r_0 < 1$ and $|a_k| \leqslant r_0$ for all k, then it holds that

$$\lim_{n\to\infty} \|F - \widetilde{S}_n F\|_2 = 0, \tag{2.9}$$

where $\psi_n(z) = \prod_{j=0}^{n} \prod_{i=1}^{N_j} \frac{z - b_i^{(j)}}{1 - \overline{b_i^{(j)}} z}$ and $b_1^{(j)} = a_{j-1}$ for $j \geqslant 1$.

Proof. By the unwinding procedure (2.1), $(\widetilde{S}_n F)(z)$ can be rewritten as

$$\begin{aligned}
(\widetilde{S_n F})(z) &= G_0(a_0) B_0(z) + G_1(a_1) B_0(z) B_1(z) + \cdots + G_n(a_n) \prod_{k=0}^{n} B_k(z) \\
&= G_0(a_0) \prod_{i=1}^{N_0} \frac{z - b_i^{(0)}}{1 - \overline{b_i^{(0)}} z} + \sum_{k=1}^{n} G_k(a_k) \prod_{j=0}^{k} \prod_{i=1}^{N_j} \frac{z - b_i^{(j)}}{1 - \overline{b_i^{(j)}} z} \\
&= \frac{\sum_{k=0}^{N^{(n)}} c_k z^k}{\prod_{j=0}^{n} \prod_{i=1}^{N_j} (1 - \overline{b_i^{(j)}} z)},
\end{aligned} \tag{2.10}$$

where $b_1^{(j)} = a_{j-1}$ for $j \geqslant 1$, $N^{(n)} = N_0 + N_1 + \cdots + N_n$ and $\{c_k\}_{k=0}^{N^{(n)}}$ are some complex constants in the complex plane. Let $E_{N^{(n+1)}}^{(n+1)} = \{b_0, b_1, \ldots, b_{N^{(n+1)}-1}\}$ be defined by (2.2). From the previous analysis, we know that $(\widetilde{S_n F})(z)$ satisfies the interpolation condition $(\widetilde{S}_n F)^{(l(b_k)-1)}(b_k) = F^{(l(b_k)-1)}(b_k)$ for all $b_k \in E_{N^{(n+1)}}^{(n+1)}$. Let

$$(S_n F)(z) = \sum_{k=0}^{N^{(n)}-1} \langle F, \mathcal{B}_k \rangle \mathcal{B}_k(z),$$

where $\mathcal{B}_k(z)$ is the TM system defined by the sequence in $E_{N^{(n)}}^{(n)}$ given in (2.2). Based on Theorem 2.3, we also know that $(S_nF)(z)$ is the unique function in the rational space $\mathcal{R}_{N^{(n)}-1,N^{(n)}}$ defined by (2.7) satisfying the interpolation condition $(\widetilde{S}_nF)^{(l(b_k)-1)}(b_k) = F^{(l(b_k)-1)}(b_k)$ for all $b_k \in E_{N^{(n)}}^{(n)}$. This implies that $(\widetilde{S}_nF)(z)-(S_nF)(z)$ will have zeros at $b_k \in E_{N^{(n)}}^{(n)}$ for $k = 0, 1, \dots, N^{(n)}-1$. It follows that $(\widetilde{S}_nF)(z)$ can be rewritten as

$$(\widetilde{S}_nF)(z) = (S_nF)(z) + \widetilde{C_n}\prod_{k=0}^{N^{(n)}-1}\frac{z-b_k}{1-\overline{b_k}z} = (S_nF)(z) + \widetilde{C_n}\prod_{j=0}^{n}\prod_{i=1}^{N_j}\frac{z-b_i^{(j)}}{1-\overline{b_i^{(j)}}z},$$

where $\widetilde{C_n}$ is an arbitrary constant. From the fact that $E_{B_{n+1}}^{N_{n+1}} = \{b_1^{(n+1)}, \dots, b_{N_{n+1}}^{(n+1)}\}$ at least has a point at a_n and the interpolation condition $(\widetilde{S}_nF)^{l(a_n)-1}(a_n) = F^{l(a_n)-1}(a_n)$, we obtain

$$\begin{aligned}[F(z)-(S_nF)(z)]^{(l(a_n)-1)}(a_n) &= \left[\frac{\psi_n(z)}{2\pi\mathrm{i}}\oint_{|\xi|=1}\frac{F(\xi)}{(\xi-z)\psi_n(\xi)}d\xi\right]_{z=a_n}^{(l(a_n)-1)}\\ &= \widetilde{C_n}[\psi_n(z)]_{z=a_n}^{(l(a_n)-1)},\end{aligned}$$

where $l(a_n)$ is the number of repeating times of a_n in the set $E_{N_n}^{(n)} \cup \{a_n\}$ and

$$\psi_n(z) = \prod_{k=0}^{N^{(n)}-1}\frac{z-b_j}{1-\overline{b_j}z} = \prod_{j=0}^{n}\prod_{i=1}^{N_j}\frac{z-b_i^{(j)}}{1-\overline{b_i^{(j)}}z}$$

has a zero at a_n with multiplicities $l(a_n)-1$. Hence, we have

$$\widetilde{C_n} = \frac{1}{2\pi\mathrm{i}}\oint_{|\xi|=1}\frac{F(\xi)}{(\xi-a_n)\psi_n(\xi)}d\xi.$$

This proves that the identity (2.8) holds. Next, we prove

$$\lim_{n\to\infty}\|F(z)-(\widetilde{S}_nF)(z)\|_2 = 0.$$

It is known that

$$\|F(z)-(\widetilde{S}_nF)(z)\|_2 = \|F(z)-(S_nF)(z)-\widetilde{C_n}\psi_n(z)\|_2 \leqslant \|F(z)-(S_nF)(z)\|_2 + |\widetilde{C_n}|.$$

If all $|a_k| < r$ for some $r < 1$, then we have the hyperbolic non-separability condition

$$\sum_{k=0}^{\infty}(1-|a_k|) = \infty.$$

Hence, by Theorem 2.3, we have

$$\lim_{n\to\infty}\|F(z)-(S_nF)(z)\|_2 = 0$$

and the $N^{(n)}$-th rational Fourier coefficients

$$c_{N^{(n)}} = \langle F, \mathcal{B}_{N^n}\rangle = \left\langle F, \frac{\sqrt{1-|a_n|^2}}{1-\overline{a_n}z}\prod_{k=0}^{N^{(n)}-1}\frac{z-b_n}{1-\overline{b_n}z}\right\rangle \to 0$$

as $n\to\infty$. From this we see that

$$\widetilde{C_n} = \frac{1}{\sqrt{1-|a_n|^2}}\langle F, \mathcal{B}_{N^n}\rangle \to 0$$

as $n\to\infty$ for $|a_n| < r$. The proof is completed. □

If all a_k's are chosen to be 0 and all B_k's are chosen to be z, then we have the Fourier expansion

$$F(z) = G_0(0) + G_1(0)z + G_2(0)z^2 + \cdots + G_n(0)z^n + \cdots .$$

By a simple computation, we know that $G_n(0) = F^{(n)}(0)/n!$. If all a_k's are chosen to be any constant in $\boldsymbol{D}$, all $B_k(z)$'s are chosen to be

$$B_k(z) = \frac{z - a_{k-1}}{1 - \overline{a_{k-1}}z}$$

for $k \geqslant 1$ and $B_0(z)$'s are chosen to be 1, then the single-zero unwinding procedure (2.1) generates

$$F(z) = G(a_0) + \cdots + G_n(a_n)\prod_{k=0}^{n-1}\frac{z - a_k}{1 - \overline{a_k}z} + \prod_{k=0}^{n}\frac{z - a_j}{1 - \overline{a_j}z}G_{n+1}(z). \tag{2.11}$$

By introducing the generalized difference

$$F_B[a_0, a_1] = \frac{F(a_1) - F(a_0)}{\frac{a_1 - a_0}{1 - \overline{a_0}a_1}}$$

and inductively,

$$F_B[a_0, a_1, \ldots, a_n] = \frac{F_B[a_1, \ldots, a_n] - F_B[a_0, a_1, \ldots, a_{n-1}]}{\frac{a_n - a_0}{1 - \overline{a_0}a_n}},$$

we can show

$$G_n(a_n) = F_B[a_0, a_1, \ldots, a_n].$$

This means that the unwinding procedure proposed by (2.1) can be seen as a generalized version of the classical Newton interpolation formula in the complex plane. Hence, (2.11) may be written as

$$F(z) = F(a_0) + F_B[a_0, a_1]\frac{z - a_0}{1 - \overline{a_0}z} + \cdots + F_B[a_0, \ldots, a_n]\prod_{k=0}^{n-1}\frac{z - a_k}{1 - \overline{a_k}z} + \prod_{k=0}^{n}\frac{z - a_j}{1 - \overline{a_j}z}G_{n+1}(z).$$

The series part

$$\tilde{S}_n F(z) = F(a_0) + F_B[a_0, a_1]\frac{z - a_0}{1 - \overline{a_0}z} + \cdots + F_B[a_0, \ldots, a_n]\prod_{k=0}^{n-1}\frac{z - a_k}{1 - \overline{a_k}z}$$

may also be written in the rational Lagrange interpolation form, as

$$\tilde{S}_n F(z) = \sum_{k=0}^{n} l_k(z)F(a_k),$$

where

$$l_k(z) = \frac{\prod_{j=0, j\neq k}^{n}(z - a_j)\prod_{j=0}^{n-1}(1 - \overline{a_j}a_k)}{\prod_{j=0, j\neq k}^{n}(a_k - a_j)}\prod_{j=0}^{n-1}(1 - \overline{a_j}z).$$

For the multi-zero unwinding procedure we have the similar results.

Let $A(\boldsymbol{D}_R)$ denote the set of the functions that are holomorphic functions in $\boldsymbol{D}_R := \{z \mid |z| < R\}$ and continuous to the closure of $\boldsymbol{D}_R$. If $F(z) \in A(\boldsymbol{D}_R)$, $R > 1$, then it is easy to show that each fully factorized Blaschke product $B_j(z)$ in the unwinding procedure (2.1) is a finite Blaschke product.

It is well known that if F is analytic in the closure of $\boldsymbol{D}$ that is equivalent to $F \in A(\boldsymbol{D}_R)$ for some $R > 1$, then the Fourier series of F converges at an exponential rate. We will prove that the unwinding procedure given in (2.1) also gives rise to an exponential convergence rate.

Corollary 2.5. *Let nonzero functions $F(z) \in A(\boldsymbol{D}_R)$, $R > 1$, and*

$$(\widetilde{S}_n F)(z) = G_0(a_0)B_0(z) + \sum_{k=1}^{n} G_k(a_k)\prod_{j=0}^{k} B_j(z)$$

be produced by (2.1)*, where a_k is any sequence in $\boldsymbol{D}$. Then for $|z| \leqslant 1$, we have*

$$|F(z) - (\widetilde{S}_n F)(z)| \leqslant \frac{MR}{(R-1)} \exp\bigg\{ -\frac{R-1}{2R} \sum_{j=0}^{n} \sum_{i=1}^{N_j} (1 - |b_i^{(j)}|) \bigg\}, \tag{2.12}$$

where $B_j(z) = \prod_{i=1}^{N_j} \frac{z - b_i^{(j)}}{1 - \overline{b_i^{(j)}} z}$ and M is some constant.

Proof. For $|z| \leqslant 1$, by (2.5), we have

$$\begin{aligned} |F(z) - (\widetilde{S}_n F)(z)| &\leqslant |F(z) - (S_n F)(z)| + \bigg| \frac{\psi_n(z)}{2\pi \mathrm{i}} \oint_{|\xi|=1} \frac{F(\xi)}{(\xi - a_n)\psi_n(\xi)} d\xi \bigg| \\ &\leqslant \bigg| \frac{\psi_n(z)}{2\pi \mathrm{i}} \oint_{|\xi|=1} \frac{F(\xi)}{(\xi - z)\psi_n(\xi)} d\xi \bigg| + \bigg| \frac{\psi_n(z)}{2\pi \mathrm{i}} \oint_{|\xi|=1} \frac{F(\xi)}{(\xi - a_n)\psi_n(\xi)} d\xi \bigg|, \end{aligned}$$

where $\psi_n(z) = \prod_{j=0}^{n} \prod_{i=1}^{N_j} \frac{z - b_i^{(j)}}{1 - \overline{b_i^{(j)}} z}$. Since $F(z)$ is holomorphic in $\boldsymbol{D}_R$ and $|\psi_n(z)| \leqslant 1$ for $|z| \leqslant 1$, it follows that

$$\begin{aligned} |F(z) - (\widetilde{S}_n F)(z)| &\leqslant \frac{1}{2\pi} \oint_{|\xi|=R} \frac{|F(\xi)|}{|\xi - z||\psi_n(\xi)|} d\xi + \frac{1}{2\pi} \oint_{|\xi|=R} \frac{|F(\xi)|}{|\xi - a_n||\psi_n(\xi)|} d\xi \\ &\leqslant \frac{MR}{(R-1)} \sup_{|\xi|=R} \frac{1}{|\psi_n(\xi)|}, \end{aligned}$$

where M is a constant depending on the maximum value of F on the boundary of $\boldsymbol{D}_R$. Hence, by (2.6), we have

$$|F(z) - (\widetilde{S}_n F)(z)| \leqslant \frac{MR}{(R-1)} \exp\bigg\{ -\frac{R-1}{2R} \sum_{j=0}^{n} \sum_{i=1}^{N_j} (1 - |b_i^{(j)}|) \bigg\}.$$

This completes the proof. □

From (2.12), it seems that the rate of convergence of a Blaschke unwinding decomposition depends on two factors: One is how many zeros can be extracted in the unwinding steps, and the other is how close zeros $b_i^{(j)}$'s are to 0. It seems that the latter factor is more important. Figure 1 uses an example to show the relations among the subtracting various function values, their winding numbers and the related decompositions. The example function in use is

$$G(z) = 1 + 10(z - 0.95)\bigg(z - \frac{1}{3}\bigg)\bigg(z - \bigg[0.1 + \frac{\mathrm{i}}{1.01}\bigg]\bigg)\bigg(z - \bigg[0.05 + \frac{\mathrm{i}}{1.01}\bigg]\bigg).$$

The figure shows that with respect to the curve $\gamma(t) = G(\mathrm{e}^{\mathrm{i}t})$, $0 \leqslant t \leqslant 2\pi$, the point $G(0) = 4.12 - 0.16\mathrm{i}$, marked by a red color circle, has winding number 1. That point situates in only one of the loops formed by the curve γ. Also in the figure the point $1 = G(0.5) = G(\frac{1}{3}) = G(0.1 + \frac{\mathrm{i}}{1.01}) = G(0.05 + \frac{\mathrm{i}}{1.01})$, marked by a blue cross, has the winding number 4. It is one of the points that possess the greatest winding number 4. By subtracting $G(0.5) = 1$ one can factorize out a Blaschke product with 4 Möbius factors. Figure 2 is a magnifying enlargement of Figure 1 to present the local details. We colored all the loops for counting the winding numbers of the individual regions (connected open sets). Subtracting $G(0.5) = 1$ gives rise to a Blaschke unwinding decomposition, i.e.,

$$F(z) = 1 + \bigg(\prod_{k=1}^{4} \frac{z - a_k}{1 - \overline{a}_k z}\bigg) G_1(z) \tag{2.13}$$

with $G_1(z) = \prod_{k=1}^{4} (1 - \overline{a}_k z)$, which is different from the standard one cited in (1.6).

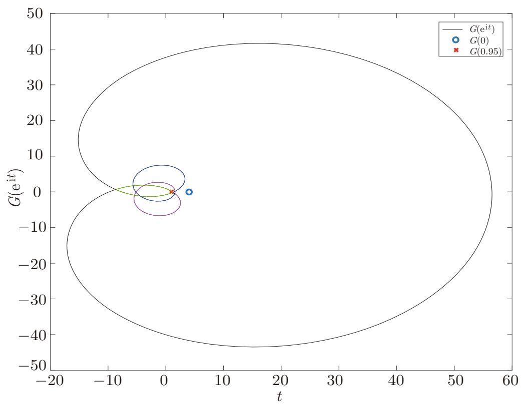

Figure 1 (Color online) The curve of $G(\mathrm{e}^{\mathrm{i}t})$

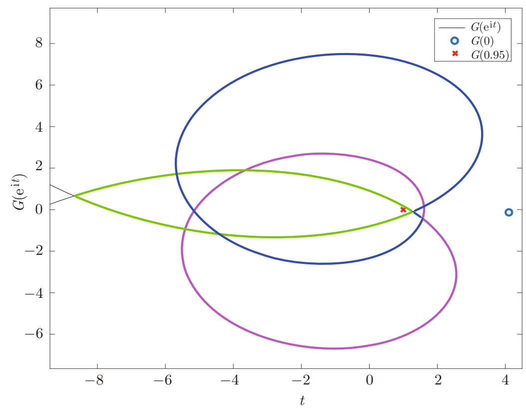

Figure 2 (Color online) Local details of $G(\mathrm{e}^{\mathrm{i}t})$

3 Sifting process through subtracting Blaschke forms

In AFD theory we have the following algebraic relation: For F being any function in the Hardy space $H^2(\boldsymbol{D})$ and $a_1, \ldots, a_n$ being any n points in the open disc $\boldsymbol{D}$, it holds that

$$G(z) = \sum_{k=1}^{n} \langle G, \mathcal{B}_k \rangle \mathcal{B}_k(z) + G_{n+1}(z) \prod_{k=1}^{n} \frac{z - a_k}{1 - \overline{a}_k z}, \tag{3.1}$$

where $\mathcal{B}_k$'s, called *weighted Blaschke products*, constitute the *rational orthonormal n-system* or the *Takenaka-Malmquist (TM) n-system*, determined by $a_1, \ldots, a_n$ with the expressions

$$\mathcal{B}_k(z) = \frac{\sqrt{1 - |a_k|^2}}{1 - \overline{a}_k z} \prod_{l=1}^{k-1} \frac{z - a_l}{1 - \overline{a}_l z},$$

and G_k's are the *reduced remainders*, obtained recursively through a *generalized backward shift transform*

$$G_{k+1}(z) = \frac{G_k(z) - \langle G_k, e_{a_k} \rangle e_{a_k}(z)}{\frac{z - a_k}{1 - \overline{a}_k z}}.$$

For a general $a \in \boldsymbol{D}$ the function

$$e_a(z) = \frac{\sqrt{1 - |a|^2}}{1 - \overline{a} z}$$

is the normalized reproducing kernel of the Hardy space with the parameter a. The orthogonality of the n-TM system implies the following useful relations: For each k,

$$\langle G_k, e_{a_k} \rangle = \langle G, \mathcal{B}_k \rangle = \langle H_k, \mathcal{B}_k \rangle, \tag{3.2}$$

where H_k is the *standard remainder*

$$H_k(z) = G(z) - \sum_{l=1}^{k-1} \langle G, \mathcal{B}_l \rangle \mathcal{B}_l(z).$$

By selecting all a_n's being equal to a real number or a complex number in the disc one obtains, respectively, the Laguerre and the two-parameter-Kautz systems. Both systems are orthonormal bases of the Hardy space and lead to convergence to the function G as $n \to \infty$.

By invoking the Cauchy-Schwarz inequality one has

$$\sup_{a \in \boldsymbol{D}} |\langle G_k, e_{a_k} \rangle| \leqslant \|G_k\| \leqslant \|G\| < \infty.$$

The essence of the AFD algorithm is that at each step the supremum is attainable (see [35]). For convenience we call

$$\mathcal{G}(G;\boldsymbol{a})(z)=\sum_{k=1}^{n}\langle G,\mathcal{B}_k\rangle\mathcal{B}_k(z)$$

as the *Blaschke form of* F *induced by* $\boldsymbol{a}=(a_1,\ldots,a_n)\in\boldsymbol{D}^n$. Note that (3.1) is an algebraic identity that achieves two aspects at the same time benefiting the approximation:

(1) the difference $G-\mathcal{G}(G;\boldsymbol{a})(z)$ has zeros $a_1,\ldots,a_n$ including multiples (see (3.1) or [35]);

(2) the energy $\|G-\mathcal{G}(G;\boldsymbol{a})\|$ can attain the minimum value over all the possible selections of n-tuple $\boldsymbol{a}$ (see [6,33]).

As an application of the results (1) and (2), the Blaschke forms $\mathcal{G}(G_k,\boldsymbol{a}^{(k)})$ can be designed to generate zeros in the unwinding process, where $\boldsymbol{a}^{(k)}$'s are arbitrary vectors in $\boldsymbol{D}^{n_k}$ of any prescribed length $n_k\geqslant 1$, or to maximally reduce energy as in [28,35]. In both cases the decompositions lead convergence to the originally given function.

Under this program the unwinding decomposition becomes

$$\begin{aligned}F(z)=F_0(z)&=B_0(z)(G_0(z)-\mathcal{G}(G_0;\boldsymbol{a}^{(0)})(z)+\mathcal{G}(G_0;\boldsymbol{a}^{(0)})(z))\\&=c_0(z)B_0(z)+B_0(z)B_{\boldsymbol{a}^{(0)}}(z)F_1(z)\\&=c_0(z)B_0(z)+c_1(z)B_0(z)B_{\boldsymbol{a}^{(0)}}(z)B_1(z)+B_0(z)B_{\boldsymbol{a}^{(0)}}(z)B_1(z)B_{\boldsymbol{a}^{(1)}}(z)F_2(z)\\&=c_0(z)B_0(z)+\sum_{k=1}^{n}c_k(z)B_0(z)\cdots B_k(z)B_{\boldsymbol{a}^{(0)}}(z)\cdots B_{\boldsymbol{a}^{(k-1)}}(z)\\&\quad+B_0(z)\cdots B_n(z)B_{\boldsymbol{a}^{(0)}}(z)\cdots B_{\boldsymbol{a}^{(n)}}(z)G_{n+1}(z),\end{aligned}$$

where $G_{k-1}(z)-\mathcal{G}(G_{k-1};\boldsymbol{a}^{(k-1)})(z)=B_{\boldsymbol{a}^{(k-1)}}(z)F_k(z)=B_{\boldsymbol{a}^{(k-1)}}(z)B_k(z)G_k(z)$ is the sifting process together with Nevanlinna factorization, $c_k(z)=\mathcal{G}(G_k;\boldsymbol{a}^{(k)})(z)$, and $B_{\boldsymbol{a}^{(k)}}(z)$ is the Blaschke product associated with $\boldsymbol{a}^{(k)}$. We note that with the Nevanlinna factorization $F_k(z)=B_k(z)G_k(z)$ the Blaschke product $B_k(z)$ does not have to exhaust all the zeros of $F_k(z)$ to make a proper sense of the instantaneous frequency and to reduce the computation complexity.

Theorem 3.1. *Under a similar condition regarding the hyperbolic non-separability of the zeros collected from the remainders as in Theorem* 2.4*, or using the maximal selections of* $\boldsymbol{a}^{(k)}$*'s as proved in* [28, 35]*, we can conclude*

$$F(z)=c_0(z)B_0(z)+\sum_{k=1}^{\infty}c_k(z)B_0(z)\cdots B_k(z)B_{\boldsymbol{a}^{(0)}}(z)\cdots B_{\boldsymbol{a}^{(k-1)}}(z).$$

If all $\boldsymbol{a}^{(k)}$'s have length 1, under the maximal selections of $\boldsymbol{a}^{(k)}$, the above model reduces to what is proceeded in [28].

Remark 3.2. All the above mentioned unwinding expansions are in the following form:

$$\begin{aligned}F_0(z)&=B_0(z)\mathcal{G}(G_0)(z)+B_0(z)B_1(z)G_1(z)\\&=B_0(z)\mathcal{G}(G_0)(z)+B_0(z)B_1(z)\mathcal{G}(G_1)(z)+B_0(z)B_1(z)B_2(z)G_2(z)\\&=\cdots\\&=B_0\mathcal{G}(G_0)(z)+\mathcal{G}(G_n)(z)\sum_{k=1}^{n}\left[\prod_{j=0}^{k}B_j(z)\right]+\left[\prod_{k=0}^{n+1}B_k(z)\right]G_{n+1}(z),\end{aligned}$$

where $B_k(z)$ is a Blaschke product formed by a part or all of the zeros of $G_{k-1}(z)-\mathcal{G}(G_{k-1})(z)$ and $G_k(z):=\frac{G_{k-1}(z)-\mathcal{G}(G_{k-1})(z)}{B_k(z)}$. If one has targeted unwinding Blaschke factors, $\mathcal{G}(G_{k-1})(z)$ are specially designed to make $G_{k-1}(z)-\mathcal{G}(G_{k-1})(z)$ to have the corresponding zeros in the disc. In the above two sections, we, respectively, make $\mathcal{G}(G_{k-1})(z)=G_{k-1}(a_{k-1})$, where at a_{k-1} the function G_{k-1} has the maximal winding number, or make $\mathcal{G}(G_{k-1})(z)=\mathcal{G}(G_{k-1};\boldsymbol{a}^{(k-1)})(z)$ to reduce the maximal energy. The convergence can be verified through the non-separable hyperbolic property of the parameters of the TM system or through the maximal selection principle.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant Nos. 61471132 and 11671363) and the Science and Technology Development Fund, Macau Special Administration Region (Grant No. 0123/2018/A3). The first author and the second author thank Professors Ronald Coifman and Stefan Steinerberger for their beneficial discussions during their visits to Yale University in 2018 and their helpful suggestions for the improvement of this paper. The authors are grateful to the anonymous referees for their careful reading of the manuscript and many helpful suggestions.

References

1 Ahern P R, Clark D N. Radial nth derivatives of Blaschke products. Math Scand, 1971, 28: 189–201

2 Akcay H. Discrete-time system modelling in L_p with orthonormal basis functions. Systems Control Lett, 2000, 39: 365–376

3 Alpay D, Bolotnikov V, Rodman L. One-sided tangential interpolation for operator-valued Hardy functions on polydisks. Integral Equations Operator Theory, 1999, 35: 253–270

4 Alpay D, Colombo F, Qian T, et al. Adaptive orthonormal systems for matrix-valued functions. Proc Amer Math Soc, 2017, 145: 2089–2106

5 Alpay D, Colombo F, Qian T, et al. Adaptative decomposition: The case of the Drury-Arveson space. J Fourier Anal Appl, 2017, 23: 1426–1444

6 Baratchart L, Cardelli M, Olivi M. Identification and rational L^2 approximation: A gradient algorithm. Automatica, 1991, 27: 413–417

7 Boashash B. Estimating and interpreting the instantaneous frequency of a signal. I: Fundamental. In: Proceedings of the IEEE, vol. 80. Piscataway: IEEE, 1992, 520–538

8 Bultheel A, Gonzalez-Vera P, Hendriksen E, et al. Orthogonal Rational Functions. Cambridge: Cambridge University Press, 1999

9 Chen Q H, Li L Q, Qian T. Stability of frames generalized by nonlinear Fourier atoms. Int J Wavelets Multiresolut Inf Process, 2005, 3: 465–476

10 Cohen L. Time-Frequency Analysis: Theory and Applications. Upper Saddle River: Prentice Hall, 1995

11 Coifman R, Peyriére J. Phase unwinding, or invariant subspace decompositions of Hardy spaces. J Fourier Anal Appl, 2019, 25: 684–695

12 Coifman R, Steinerberger S. Nonlinear phase unwinding of functions. J Fourier Anal Appl, 2017, 23: 778–809

13 Coifman R, Steinerberger S, Wu H T. Carrier frequencies, holomorphy and unwinding. SIAM J Math Anal, 2017, 49: 4838–4864

14 Dang P, Qian T. Analytic phase derivatives, all-pass filters and signals of minimum phase. IEEE Trans Signal Process, 2011, 59: 4708–4718

15 Dang P, Qian T, You Z. Hardy-Sobolev spaces decomposition and applications in signal analysis. J Fourier Anal Appl, 2011, 17: 36–64

16 Daubechies I, Lu J F, Wu H T. Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool. Appl Comput Harmon Anal, 2011, 30: 243–261

17 Gallardo-Gutiérrez E, Gorkin P. Interpolating Blaschke products and angular derivatives. Trans Amer Math Soc, 2012, 364: 2319–2337

18 Garnett J. Bounded Analytic Functions. New York: Academic Press, 1981

19 Huang N E, Shen Z, Long S R, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc R Soc Lond Ser A Math Phys Eng Sci, 1998, 454: 903–995

20 Konstantin D, Dominique Z. Variational mode decomposition. IEEE Trans Signal Proces, 2014, 62: 531–544

21 Li Y D, Zhang L M, Qian T. 2D partial unwinding—A novel non-linear phase decomposition of images. IEEE Trans Image Process, 2019, 47: 516–525

22 Mai W X, Dang P, Zhang L M, et al. Consecutive minimum phase expansion of physically realizable signals with applications. Math Methods Appl Sci, 2016, 39: 62–72

23 Mandic D P, Rehman N, Wu Z, et al. Empirical mode decomposition-based time-frequency analysis of multivariate signals: The power of adaptive data analysis. IEEE Signal Process Mag, 2013, 30: 74–86

24 Mi W, Qian T. Frequency domain identification: An algorithm based on adaptive rational orthogonal system. Automatica, 2012, 48: 1154–1162

25 Nahon M. Phase evaluation and segmentation. PhD Thesis. New Haven: Yale University, 2000

26 Picinbono B. On instantaneous amplitude and phase of signals. IEEE Trans Signal Process, 1997, 45: 552–560

27 Qian T. Boundary derivatives of the phases of inner and outer functions and applications. Math Methods Appl Sci, 2009, 32: 253–263

28 Qian T. Intrinsic mono-component decomposition of functions: An advance of Fourier theory. Math Methods Appl Sci, 2010, 33: 880–891

29 Qian T. Two-dimensional adaptive Fourier decomposition. Math Methods Appl Sci, 2016, 39: 2431–2448

30 Qian T, Sproessig W, Wang J X. Adaptive Fourier decomposition of functions in quaternionic Hardy spaces. Math Methods Appl Sci, 2012, 35: 43–64

31 Qian T, Tan L H. Characterizations of mono-components: The Blaschke and starlike types. Complex Anal Oper Theory, 2018, 12: 1383–1399

32 Qian T, Tan L H, Wang Y B. Adaptive decomposition by weighted inner functions: A generalization of Fourier series. J Fourier Anal Appl, 2011, 17: 175–190

33 Qian T, Wang J Z, Mai W X. An enhancement algorithm for cyclic adaptive Fourier decomposition. Appl Comput Harmon Anal, 2019, 47: 516–525

34 Qian T, Wang R, Xu Y S, et al. Orthonormal bases with nonlinear phases. Adv Comput Math, 2010, 33: 75–95

35 Qian T, Wang Y B. Adaptive Fourier series—a variation of greedy algorithm. Adv Comput Math, 2011, 34: 279–293

36 Qian T, Wegert E. Optimal approximation by Blaschke forms. Complex Var Elliptic Equ, 2013, 58: 123–133

37 Qian T, Yang Y. Hilbert transforms on the sphere with the Clifford algebra setting. J Fourier Anal Appl, 2009, 15: 753–774

38 Rudin W. Function Theory in Polydiscs. New York: W A Benjamin, 1969

39 Shapiro J H, Taylor P D. Composition Operators and Classical Function Theory. Tracts in Mathematics. New York: Springer-Verlag, 1993

40 Singh P, Jodhi S D, Patney R K, et al. The Fourier decomposition method for nonlinear and non-stationary time series analysis. Proc R Soc Lond Ser A Math Phys Eng Sci, 2017, 473: 20160871

41 Sun X Y, Dang P. Numerical stability of circular Hilbert transform and its applications to signal decomposition. Appl Math Comput, 2019, 359: 357–373

42 Tan L H, Qian T. Extracting outer function part from Hardy space function. Sci China Math, 2017, 60: 2321–2336

43 Tan L H, Shen L X, Yang L H. Rational orthogonal bases satisfying the Bedrosian identity. Adv Comput Math, 2010, 33: 285–303

44 Weiss G, Weiss M. A derivation of the main results of the theory of H^p-spaces. Rev Un Mat Argentina, 1962, 20: 63–71

45 Xu Y, Yan D Y. The Bedrosian identity for the Hilbert transform of product functions. Proc Amer Math Soc, 2006, 134: 2719–2728

46 Yang Y, Qian T, Sommen F. Phase derivative of monogenic signals in higher dimensional spaces. Complex Anal Oper Theory, 2012, 6: 987–1010

47 Yu B, Zhang H Z. The Bedrosian identity and homogeneous semi-convolution equations. J Integral Equations Appl, 2008, 20: 527–568