

# SPECTRAL AND CONVERGENCE ANALYSIS OF THE DISCRETE ALIF METHOD\*

ANTONIO CICONE<sup>†</sup>, CARLO GARONI<sup>‡</sup>, AND STEFANO SERRA-CAPIZZANO<sup>§</sup>

**Abstract.** The Adaptive Local Iterative Filtering (ALIF) method is a recently proposed iterative procedure to decompose a signal into a finite number of “simple components” called intrinsic mode functions. It is an alternative to the well-known and widely used empirical mode decomposition, a method proposed in 1998 by Huang and his research team at NASA. In this paper we consider the discrete version of the ALIF method and we perform a spectral analysis of the related iteration matrices, with particular focus on the eigenvalue clustering and the eigenvalue distribution. Based on the eigenvalue distribution, we formulate a necessary condition for the convergence of the Discrete ALIF method. We also provide a simple criterion to construct appropriate filters satisfying this condition. Our analysis, which represents a first step toward a complete mathematical understanding of the ALIF method, relies on recent matrix-theoretic results about sampling matrices and, above all, on the theory of generalized locally Toeplitz sequences, which we extend in this paper. Several numerical examples in support of the theoretical analysis are presented.

**Key words.** iterative filtering, adaptive local iterative filtering, empirical mode decomposition, convergence analysis, eigenvalue distribution, eigenvalue clustering, generalized locally Toeplitz sequences

**AMS subject classifications.** 94A12, 68W40, 15A18, 47B06, 15B05

**1. Introduction.** The Adaptive Local Iterative Filtering (ALIF) method was proposed in [12] as a generalization of the Iterative Filtering (IF) method due to Lin et al. [29]. The purpose of both these techniques is to iteratively decompose a given signal into a finite number of “simple components” called Intrinsic Mode Functions (IMFs). The idea of looking for such a decomposition dates back to the pioneering work by Huang et al. [23]. In particular, according to [23, section 4], an IMF is a signal possessing the following two properties:

- the number of extrema and the number of zero crossings must be either equal or differ at most by one;
- at any point, the mean value of the envelope connecting the local maxima and the envelope connecting the local minima must be zero.

For more details, we refer the reader to [23].

The first method able to decompose a signal into IMFs was proposed by Huang et al. [23] and goes under the name of Empirical Mode Decomposition (EMD). The EMD and its stable version, the Ensemble Empirical Mode Decomposition (EEMD) [42], have found extensive applications in many scientific fields: in Medicine, to automatically identify seizure-free electroencephalographic signals [35], study the gastroesophageal reflux disease [28], derive the respiratory sinus arrhythmia from the heartbeat time series [7], analyze the heart rate variability [17], denoise the elec-

---

\*Submitted to the editors DATE.

**Funding:** Antonio Cicone and Carlo Garoni are Marie-Curie fellows of the Italian INdAM (Istituto Nazionale di Alta Matematica) under grant agreement PCOFUND-GA-2012-600198.

<sup>†</sup>Department of Information Engineering, Computer Science and Mathematics, University of L’Aquila, L’Aquila, Italy ([antonio.cicone@univaq.it](mailto:antonio.cicone@univaq.it)).

<sup>‡</sup>Institute of Computational Science, University of Italian Switzerland, Lugano, Switzerland, and Department of Science and High Technology, Insubria University, Como, Italy ([carlo.garoni@usi.ch](mailto:carlo.garoni@usi.ch), [carlo.garoni@uninsubria.it](mailto:carlo.garoni@uninsubria.it)).

<sup>§</sup>Department of Science and High Technology, Insubria University, Como, Italy, and Division of Scientific Computing, Department of Information Technology, Uppsala University, Uppsala, Sweden ([stefano.serrac@uninsubria.it](mailto:stefano.serrac@uninsubria.it), [stefano.serra@it.uu.se](mailto:stefano.serra@it.uu.se)).

trocardiographic signal and correct the baseline wander [9], study the dengue virus spread [22], identify the coupling between prefrontal and visual cortex [21]; in Geophysics, to study the evolution of land surface air temperature trend [25], analyze the global mean sea-level rise [10], extract the solar cycle signal from the stratospheric data [16], identify near-fault ground-motion characteristics [30]; in Engineering, to diagnose faults in rotating machinery [27], separate two sources from a single-channel recording [31], control the wind response of a building with variable stiffness tuned mass damper [41]; in Information Technology, to analyze images [33] and texture [34]; in Economics, to analyze the price of the crude oil [44]. This seemingly long list of applications is actually far from being complete. In fact, the first paper ever published on the EMD algorithm [23] has received so far, based on the Scopus database, more than 9000 citations.

Despite the applicative success, the convergence analysis of EMD and EEMD is still an open problem. This is due to the fact that a key step of both these methods is the computation of the moving average of a given signal by means of cubic splines. The IF and ALIF algorithms have been recently proposed in [12, 29] as alternatives to EMD and EEMD. Their structure is the same as the structure of EMD and EEMD, but the moving average of a signal is now computed by integrating a properly weighted version of the signal itself. Such a computation by means of integrals instead of cubic splines opens the door to a mathematically rigorous convergence analysis of the IF and ALIF methods. Over the last two years, both these methods have gained a considerable attention from the international scientific community [2, 3, 4, 5, 6, 26, 32, 37, 40, 43]. However, while the convergence analysis of the IF method has already been carried out in the continuous and discrete settings [12, 15], so far nothing is known about the convergence of the ALIF method.

In this paper we focus on the discrete version of the ALIF method,<sup>1</sup> which will be introduced in section 2.3. Using recent results about sampling matrices [1] and the theory of Generalized Locally Toeplitz (GLT) sequences [18, 19, 38, 39], we analyze the asymptotic spectral properties of the ALIF iteration matrices, with particular attention to the eigenvalue clustering (Theorem 4.1) and the eigenvalue distribution (Theorems 4.4 and 4.5). In doing so, we also provide a noteworthy extension of the theory of GLT sequences (Theorem 4.6). Based on our spectral analysis, we formulate a necessary condition for the convergence of the Discrete ALIF method (condition (4.21)), and we provide in Remark 4.10 a simple criterion for constructing appropriate filters which ensure that (4.21) is satisfied. It is worth noting that a completely analogous criterion was obtained in [12, section 2.1], when analyzing the convergence of the simpler IF method. A few numerical examples are provided in section 5 to support the analysis of section 4. In section 6 we draw conclusions and suggest future lines of research.

## 2. The ALIF method.

**2.1. Terminology.** Throughout this paper, any real function  $g : \mathbb{R} \rightarrow \mathbb{R}$  is also referred to as a signal. Without loss of generality, we assume that the domain on which every signal  $g$  is studied is the reference interval  $[0, 1]$ . Outside the reference interval, the signal is usually not known and so, whenever necessary, we have to make assumptions, that is, we have to impose boundary conditions. The extrema of a signal  $g$  are the points belonging to  $(0, 1)$  where  $g$  attains its local maxima and minima. If  $\mathbf{g} = [g_0, \dots, g_{n-1}]$  is a vector in  $\mathbb{R}^n$ , the extrema of  $\mathbf{g}$  are the indices

---

<sup>1</sup>A Matlab implementation of the Discrete ALIF method is available at <http://www.cicone.com>.

belonging to  $\{1, \dots, n-2\}$  where  $\mathbf{g}$  attains its local maxima and minima, i.e., the indices  $j \in \{1, \dots, n-2\}$  such that  $\mathbf{g}_j > \max(\mathbf{g}_{j-1}, \mathbf{g}_{j+1})$  or  $\mathbf{g}_j < \min(\mathbf{g}_{j-1}, \mathbf{g}_{j+1})$ . A filter  $k$  is an even, nonnegative, bounded, measurable,<sup>2</sup> and compactly supported function from  $\mathbb{R}$  to  $\mathbb{R}$  satisfying the normalization condition  $\int_{\mathbb{R}} k(y)dy = 1$  and the following additional property: let  $\ell = \sup\{y > 0 : k(y) > 0\}$ , then  $k(y) = 0$  for  $|y| \geq \ell$ . We refer to  $\ell$  as the length of the filter  $k$ . Note that  $0 < \ell < \infty$  and the support of  $k$  is contained in  $[-\ell, \ell]$ .

---

**Algorithm 2.1 (ALIF Algorithm)** IMFs = ALIF( $g$ )

---

```

IMFs = {}
initialize the remaining signal  $r = g$ 
while the number of extrema of  $r$  is  $\geq 2$  do
  for each  $x \in [0, 1]$  compute the filter  $k_x$ , whose length  $\ell(x)$  changes from  $x$  to  $x$ 
  based on  $r$  itself
   $g_1 = r$ 
   $m = 1$ 
  while the stopping criterion is not satisfied do
    compute the moving average  $f_m$  of the signal  $g_m$  as
     $f_m(x) = \int_{\mathbb{R}} g_m(y)k_x(x-y)dy$ 
     $g_{m+1} = g_m - f_m$ 
     $m = m + 1$ 
  end while
  IMFs = IMFs  $\cup \{g_m\}$ 
   $r = r - g_m$ 
end while

```

---

**2.2. The ALIF method.** As mentioned in section 1, the ALIF method is an iterative procedure whose purpose is to decompose a signal  $g$  into a finite number of “simple components”, the so-called IMFs of  $g$ . Algorithm 2.1 shows the pseudocode of the ALIF method, in which the input is a signal  $g$  and the output is the set of the IMFs of  $g$ . The ALIF algorithm contains two loops. The inner loop captures a single IMF, while the outer loop produces all the IMFs embedded in  $g$ . Considering the first iteration of the ALIF outer loop in which  $g_1 = g$ , we see that the key idea to extract the first IMF consists in computing the moving average of  $g_m$  and subtract it from  $g_m$  itself so as to capture the fluctuation part  $\mathcal{S}_m(g_m) = g_m - f_m = g_{m+1}$ . This is repeated iteratively and, assuming convergence, the first IMF is obtained as  $\text{IMF}_1 = \lim_{m \rightarrow \infty} \mathcal{S}(g_m)$ . In practice, however, we cannot let  $m$  go to  $\infty$  and we have to use a stopping criterion, as indicated in Algorithm 2.1. Assuming convergence, one can stop the inner loop at the first index  $m$  such that the difference  $g_{m+1} - g_m$  is small in some norm (possibly, a norm for which the convergence is known). A safer stopping criterion also imposes a limit on the maximum number of iterations. Once the first IMF is obtained, to produce the second IMF we apply the previous process to the remaining signal  $r = g - \text{IMF}_1$ . We then iterate this procedure to obtain all the IMFs of  $g$ , and we stop as soon as the remaining signal becomes a trend signal, meaning that it possesses at most one extremum. Clearly, the sum of all the IMFs of  $g$  produced by the ALIF method with the final trend signal  $r$  is equal to  $g$ .

---

<sup>2</sup>Throughout this paper, the word “measurable” always means “Lebesgue measurable”.

REMARK 2.1. *In the case where  $\ell(x)$  is chosen at each iteration of the outer loop as a constant  $\ell$ , depending on the remaining signal  $r$  but not on  $x$ , the ALIF method reduces to the IF method, whose convergence has been studied in [12, 24].*

REMARK 2.2. *In practical applications of the ALIF method,  $k_x$  is chosen as<sup>3</sup>*

$$(2.1) \quad k_x(y) = \frac{k\left(\frac{y}{\ell(x)}\right)}{\ell(x)},$$

where  $k$  is an a priori fixed filter with length 1 (so that the length of  $k_x$  is  $\ell(x)$  as in Algorithm 2.1).

**2.3. The Discrete ALIF method.** In practice, we usually do not know a signal  $g$  on the whole reference interval  $[0, 1]$ . What we actually know are the samples of  $g$  over a fine grid in  $[0, 1]$ . We therefore need a discrete version of the ALIF algorithm, which is able to (approximately) capture the IMFs of  $g$  by exploiting this sole information. From now on, we make the following assumptions.

- For any signal  $g$ , no other information about  $g$  is available except for its samples at the  $n$  points  $x_i = \frac{i}{n-1}$ ,  $i = 0, \dots, n-1$ . Moreover,  $g = 0$  outside  $[0, 1]$  (so we are imposing homogeneous Dirichlet boundary conditions).
- The filter  $k_x$  is defined as in (2.1) in terms of an a priori fixed filter  $k$  with length 1.

Under these hypotheses, what we may ask to a discrete version of the ALIF algorithm is to compute the (approximated) samples of the IMFs of  $g$  at the sampling points  $x_i$ ,  $i = 0, \dots, n-1$ . This is done by approximating the moving average at the points  $x_i$  through the rectangle formula or any other quadrature rule. Setting for convenience  $x_j = \frac{j}{n-1}$  for all  $j \in \mathbb{Z}$ , the rectangle formula yields the approximation

$$f_m(x_i) = \int_{\mathbb{R}} g_m(y) k_{x_i}(x_i - y) dy \approx \frac{1}{n-1} \sum_{j \in \mathbb{Z}} g_m(x_j) k_{x_i}(x_i - x_j), \quad i = 0, \dots, n-1,$$

where we note that the sum is finite because  $k_{x_i}$  is compactly supported. Assuming that, at each iteration of the ALIF inner loop, the signal  $g_m$  is set to zero outside the reference interval  $[0, 1]$ , the previous equation becomes

$$f_m(x_i) \approx \frac{1}{n-1} \sum_{j=0}^{n-1} g_m(x_j) k_{x_i}(x_i - x_j), \quad i = 0, \dots, n-1.$$

We then obtain

$$(2.2) \quad \begin{aligned} g_{m+1}(x_i) &= g_m(x_i) - f_m(x_i) \\ &\approx g_m(x_i) - \frac{1}{n-1} \sum_{j=0}^{n-1} g_m(x_j) k_{x_i}(x_i - x_j), \quad i = 0, \dots, n-1. \end{aligned}$$

Denoting by  $\mathbf{g} = [g(x_0), \dots, g(x_{n-1})]^T$  the vector containing the samples of the signal  $g$  at the sampling points  $x_i$ , we can rewrite (2.2) in matrix form as follows:

$$(2.3) \quad \mathbf{g}_{m+1} \approx (I_n - K_n) \mathbf{g}_m,$$

<sup>3</sup>Note that  $k_x$  in (2.1) is indeed a filter according to the terminology introduced in section 2.1.

where  $I_n$  is the  $n \times n$  identity matrix and

$$(2.4) \quad K_n = \left[ \frac{1}{n-1} k_{x_i}(x_i - x_j) \right]_{i,j=0}^{n-1} = \left[ \frac{k\left(\frac{x_i - x_j}{\ell(x_i)}\right)}{(n-1)\ell(x_i)} \right]_{i,j=0}^{n-1} = \left[ \frac{k\left(\frac{i-j}{(n-1)\ell(x_i)}\right)}{(n-1)\ell(x_i)} \right]_{i,j=0}^{n-1}.$$

---

**Algorithm 2.2 (Discrete ALIF Algorithm) IMF<sub>s</sub> = ALIF(**g**)**


---

**IMF<sub>s</sub>** = {}  
initialize the remaining signal **r** = **g**  
**while** the number of extrema of **r** is  $\geq 2$  **do**  
  for each  $x_i = x_0, \dots, x_{n-1}$  compute the filter  $k_{x_i}$ , whose length  $\ell(x_i)$  changes from  $x_i$  to  $x_i$  based on **r** itself  
  **g**<sub>1</sub> = **r**  
   $m = 1$   
  **while** the stopping criterion is not satisfied **do**  
    extend **g** <sub>$m$</sub>  to  $\mathbb{Z}$  by setting  $(\mathbf{g}_m)_j = 0$  for  $j \notin \{0, \dots, n-1\}$   
    compute the moving average **f** <sub>$m$</sub>  of **g** <sub>$m$</sub>  as  
     $(\mathbf{f}_m)_i = \frac{1}{n-1} \sum_{j \in \mathbb{Z}} (\mathbf{g}_m)_j k_{x_i}(x_i - x_j)$ ,  $i = 0, \dots, n-1$   
    **g** <sub>$m+1$</sub>  = **g** <sub>$m$</sub>  - **f** <sub>$m$</sub>   
     $m = m + 1$   
  **end while**  
  **IMF<sub>s</sub>** = **IMF<sub>s</sub>**  $\cup$  {**g** <sub>$m$</sub> }  
  **r** = **r** - **g** <sub>$m$</sub>   
**end while**

---

The pseudocode for the Discrete ALIF method<sup>4</sup> is reported in Algorithm 2.2. The input is a vector  $\mathbf{g} = [\mathbf{g}_0, \dots, \mathbf{g}_{n-1}]^T = [g(x_0), \dots, g(x_{n-1})]^T$  containing the samples of a signal  $g$  at the sampling points  $x_i = \frac{i}{n-1}$ ,  $i = 0, \dots, n-1$ , while the output is the set of vectors containing the (approximated) samples of the IMF<sub>s</sub> of  $g$  at the same points  $x_i$ . Note that the first four lines inside the inner loop of Algorithm 2.2 can be replaced by the sole equation  $\mathbf{g}_{m+1} = (I_n - K_n)\mathbf{g}_m$ , which is obtained from (2.3) by turning “ $\approx$ ” into “ $=$ ”. Assuming convergence, the vector **IMF**<sub>1</sub> containing the (approximated) samples of the first IMF is obtained as **IMF**<sub>1</sub> =  $(I_n - K_n)^m \mathbf{r}$  with  $\mathbf{r} = \mathbf{g}$  and  $m$  large enough so that the stopping criterion is met. Similarly, **IMF**<sub>2</sub> =  $(I_n - K_n)^m \mathbf{r}$  with  $\mathbf{r} = \mathbf{g} - \mathbf{IMF}_1$  and  $m$  large enough, **IMF**<sub>3</sub> =  $(I_n - K_n)^m \mathbf{r}$  with  $\mathbf{r} = \mathbf{g} - \mathbf{IMF}_1 - \mathbf{IMF}_2$  and  $m$  large enough, etc. Note that the matrix  $K_n$  used to compute **IMF** <sub>$i$</sub>  is different in general from the matrix  $K_n$  used to compute **IMF** <sub>$j$</sub>  if  $i \neq j$ . Indeed, the matrix  $K_n$  changes at every iteration of the outer loop because, although the filter  $k$  is fixed, the length  $\ell(x_i)$  depends on the remaining signal **r** and changes with it.

REMARK 2.3. A necessary condition for the convergence of the Discrete ALIF method is that

$$(2.5) \quad |1 - \lambda_i(K_n)| \leq 1, \quad i = 1, \dots, n.$$

Indeed, if (2.5) is violated then  $\rho(I_n - K_n) > 1$  and  $(I_n - K_n)^m \mathbf{r}$  diverges to  $\infty$  (with respect to any norm of  $\mathbb{R}^n$ ) for almost every vector  $\mathbf{r} \in \mathbb{R}^n$ .

---

<sup>4</sup>Available at [www.cicone.com](http://www.cicone.com).

**3. Spectral analysis tools.** We present in this section the tools for analyzing the asymptotic spectral properties of the ALIF iteration matrix  $K_n$ . Throughout this paper, a matrix-sequence is any sequence of the form  $\{A_n\}_n$ , where  $A_n$  is a square matrix of size  $n$ .

**Singular value and eigenvalue distribution of a matrix-sequence.** Let  $C_c(\mathbb{C})$  be the space of continuous complex-valued functions with bounded support defined on  $\mathbb{C}$  and let  $\mu_p$  be the Lebesgue measure in  $\mathbb{R}^p$ . If  $A$  is a square matrix of size  $n$ , the singular values and the eigenvalues of  $A$  are denoted by  $\sigma_1(A), \dots, \sigma_n(A)$  and  $\lambda_1(A), \dots, \lambda_n(A)$ , respectively.

DEFINITION 3.1. *Let  $\{A_n\}_n$  be a matrix-sequence and let  $f : D \subset \mathbb{R}^p \rightarrow \mathbb{C}$  be a measurable function defined on a set  $D$  with  $0 < \mu_p(D) < \infty$ .*

- *We say that  $\{A_n\}_n$  has a singular value distribution described by  $f$ , and we write  $\{A_n\}_n \sim_\sigma f$ , if for all  $F \in C_c(\mathbb{C})$  we have*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n F(\sigma_i(A_n)) = \frac{1}{\mu_p(D)} \int_D F(|f(y_1, \dots, y_p)|) dy_1 \dots dy_p.$$

- *We say that  $\{A_n\}_n$  has an eigenvalue distribution described by  $f$ , and we write  $\{A_n\}_n \sim_\lambda f$ , if for all  $F \in C_c(\mathbb{C})$  we have*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n F(\lambda_i(A_n)) = \frac{1}{\mu_p(D)} \int_D F(f(y_1, \dots, y_p)) dy_1 \dots dy_p.$$

If  $\{A_n\}_n$  has both a singular value and an eigenvalue distribution described by  $f$ , we write  $\{A_n\}_n \sim_{\sigma, \lambda} f$ .

**Informal meaning of the singular value and eigenvalue distribution.**

Assuming  $f$  is Riemann-integrable, the eigenvalue distribution  $\{A_n\}_n \sim_\lambda f$  has the following informal meaning [18, section 3.1]: all the eigenvalues of  $A_n$ , except possibly for  $o(n)$  outliers, are approximately equal to the samples of  $f$  over a uniform grid in  $D$  (for  $n$  large enough). For instance, if  $p = 1$  and  $D = [a, b]$ , then, assuming we have no outliers, the eigenvalues of  $A_n$  are approximately equal to

$$f\left(a + i \frac{b-a}{n}\right), \quad i = 1, \dots, n,$$

for  $n$  large enough. Similarly, if  $p = 2$ ,  $n = m^2$  and  $D = [a_1, b_1] \times [a_2, b_2]$ , then, assuming we have no outliers, the eigenvalues of  $A_n$  are approximately equal to

$$f\left(a_1 + i \frac{b_1 - a_1}{m}, a_2 + j \frac{b_2 - a_2}{m}\right), \quad i, j = 1, \dots, m,$$

for  $n$  large enough. A completely analogous meaning can also be given for the singular value distribution  $\{A_n\}_n \sim_\sigma f$ .

**Rearrangement.**

DEFINITION 3.2. *Let  $f : D \subset \mathbb{R}^p \rightarrow \mathbb{C}$  be a measurable function defined on a set  $D$  with  $0 < \mu_p(D) < \infty$ , and let  $\varphi : E \subset \mathbb{R}^q \rightarrow \mathbb{C}$  be a measurable function defined on a set  $E$  with  $0 < \mu_q(E) < \infty$ . We say that  $\varphi$  is a rearranged version of  $f$  if*

$$(3.1) \quad \frac{1}{\mu_q(E)} \int_E F(\varphi(t_1, \dots, t_q)) dt_1 \dots dt_q = \frac{1}{\mu_p(D)} \int_D F(f(y_1, \dots, y_p)) dy_1 \dots dy_p$$

for all  $F \in C_c(\mathbb{C})$ .

Since for every  $F$  in  $C_c(\mathbb{C})$  the function  $F(|\cdot|)$  belongs to  $C_c(\mathbb{C})$  as well, eq. (3.1) immediately implies that

$$(3.2) \quad \frac{1}{\mu_q(E)} \int_E F(|\varphi(t_1, \dots, t_q)|) dt_1 \dots dt_q = \frac{1}{\mu_p(D)} \int_D F(|f(y_1, \dots, y_p)|) dy_1 \dots dy_p$$

for all  $F \in C_c(\mathbb{C})$ . Thus, if  $\varphi$  is a rearranged version of  $f$  and  $\{A_n\}_n \sim_\lambda f$  (resp.,  $\{A_n\}_n \sim_\sigma f$ ) then  $\{A_n\}_n \sim_\lambda \varphi$  (resp.,  $\{A_n\}_n \sim_\sigma \varphi$ ). The following result can be found in [18, section 3.2].

**PROPOSITION 3.3.** *Let  $D = [a_1, b_1] \times \dots \times [a_p, b_p]$  be a hyperrectangle in  $\mathbb{R}^p$  and let  $f : D \rightarrow \mathbb{R}$  be a Riemann-integrable function. Compute, for each  $r \in \mathbb{N}$ , the uniform samples*

$$f\left(a_1 + i_1 \frac{b_1 - a_1}{r}, \dots, a_p + i_p \frac{b_p - a_p}{r}\right), \quad i_1, \dots, i_p = 1, \dots, r,$$

sort them in non-decreasing order and put them in a vector  $(s_1, s_2, \dots, s_{r^p})$ . Let  $\varphi_r : [0, 1] \rightarrow \mathbb{R}$  be the piecewise linear non-decreasing function that interpolates the samples  $(s_0 = s_1, s_1, s_2, \dots, s_{r^p})$  over the nodes  $(0, \frac{1}{r^p}, \frac{2}{r^p}, \dots, 1)$ , i.e.,

$$(3.3) \quad \begin{cases} \varphi_r\left(\frac{i}{r^p}\right) = s_i, & i = 0, \dots, r^p, \\ \varphi_r \text{ linear on } \left[\frac{i}{r^p}, \frac{i+1}{r^p}\right] \text{ for } i = 0, \dots, r^p - 1. \end{cases}$$

Suppose  $\varphi_r$  converges in measure over  $[0, 1]$  to some function  $\varphi$  as  $r \rightarrow \infty$  (this is always the case in real-world applications). Then  $\varphi$  is a rearranged version of  $f$ .

**Clustering.** If  $z \in \mathbb{C}$  and  $\varepsilon > 0$ , we denote by  $D(z, \varepsilon)$  the disk with center  $z$  and radius  $\varepsilon$ , i.e.,  $D(z, \varepsilon) = \{w \in \mathbb{C} : |w - z| < \varepsilon\}$ . If  $S \subseteq \mathbb{C}$  and  $\varepsilon > 0$ , we denote by  $D(S, \varepsilon)$  the  $\varepsilon$ -expansion of  $S$ , which is defined as  $D(S, \varepsilon) = \bigcup_{z \in S} D(z, \varepsilon)$ .

**DEFINITION 3.4.** *Let  $\{A_n\}_n$  be a matrix-sequence and let  $S \subseteq \mathbb{C}$ .*

- We say that  $\{A_n\}_n$  is strongly clustered at  $S$  in the sense of the eigenvalues if for every  $\varepsilon > 0$  the number of eigenvalues of  $A_n$  lying outside  $D(S, \varepsilon)$  is bounded by a constant  $C_\varepsilon$  independent of  $n$ ; that is, if for every  $\varepsilon > 0$

$$(3.4) \quad \#\{j \in \{1, \dots, n\} : \lambda_j(A_n) \notin D(S, \varepsilon)\} = O(1).$$

- We say that  $\{A_n\}$  is weakly clustered at  $S$  in the sense of the eigenvalues if for every  $\varepsilon > 0$

$$(3.5) \quad \#\{j \in \{1, \dots, n\} : \lambda_j(A_n) \notin D(S, \varepsilon)\} = o(n).$$

By replacing ‘‘eigenvalues’’ with ‘‘singular values’’ and  $\lambda_j(A_n)$  with  $\sigma_j(A_n)$  in (3.4)–(3.5), we obtain the definitions of a matrix-sequence strongly or weakly clustered at a subset of  $\mathbb{C}$  in the sense of the singular values.

The next proposition can be found in [18, section 3.3]. It provides an important relation between the notions of eigenvalue distribution and clustering. If  $C$  is a constant, we denote by  $C$  also the function which is identically equal to  $C$  and is defined on some subset  $D$  of some  $\mathbb{R}^k$  with  $0 < \mu_k(D) < \infty$ .

**PROPOSITION 3.5.** *Let  $\{A_n\}_n$  be a matrix-sequence and let  $C$  be a constant. Then, the following properties hold.*

- $\{A_n\}_n \sim_\lambda C$  if and only if  $\{A_n\}_n$  is weakly clustered at  $\{C\}$  in the sense of the eigenvalues.
- $\{A_n\}_n \sim_\sigma C$  if and only if  $\{A_n\}_n$  is weakly clustered at  $\{|C|\}$  in the sense of the singular values.

The final result about singular value and eigenvalue clustering that we shall need in this paper is reported in the next proposition [1, Theorem 3]. If  $A$  is an  $n \times n$  matrix and  $1 \leq p \leq \infty$ , we denote by  $\|A\|_p$  the Schatten  $p$ -norm of  $A$ , i.e., the  $p$ -norm of the vector  $(\sigma_1(A), \dots, \sigma_n(A))$  formed by the singular values of  $A$ . The Schatten  $\infty$ -norm  $\|A\|_\infty$  is the largest singular value of  $A$  and coincides with the spectral norm  $\|A\|$ . Note that  $\|A\| \leq \|A\|_p$  for all matrices  $A$  and all  $p \in [1, \infty]$ . The Schatten 1-norm  $\|A\|_1$  is the sum of the singular values of  $A$  and is often referred to as the trace-norm of  $A$ . The Schatten 2-norm  $\|A\|_2$  coincides with the Frobenius norm, i.e.,  $\|A\|_2 = (\sum_{i,j=1}^n |a_{ij}|^2)^{1/2}$ . For more on Schatten  $p$ -norms, see [8].

**PROPOSITION 3.6.** *Let  $\{A_n\}_n$  be a matrix-sequence and suppose there exist  $p \in [1, \infty)$  and a constant  $M$  such that  $\|A_n\|_p \leq M$  for all  $n$ . Then  $\{A_n\}_n$  is strongly clustered at  $\{0\}$  both in the sense of singular values and eigenvalues.*

**Zero-distributed sequences.** A matrix-sequence  $\{Z_n\}_n$  such that  $\{Z_n\}_n \sim_\sigma 0$  is referred to as a zero-distributed sequence. In other words,  $\{Z_n\}_n$  is zero-distributed if and only if  $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n F(\sigma_i(Z_n)) = F(0)$  for all  $F \in C_c(\mathbb{C})$ . Proposition 3.7 is proved in [18, section 3.4]. It provides an important characterization of zero-distributed sequences together with a useful sufficient condition for detecting such sequences. For convenience, throughout this paper we use the natural convention  $1/\infty = 0$ .

**PROPOSITION 3.7.** *Let  $\{Z_n\}_n$  be a matrix-sequence.*

- $\{Z_n\}_n$  is zero-distributed if and only if  $Z_n = R_n + N_n$  with

$$\lim_{n \rightarrow \infty} n^{-1} \text{rank}(R_n) = \lim_{n \rightarrow \infty} \|N_n\| = 0.$$

- $\{Z_n\}_n$  is zero-distributed if there is a  $p \in [1, \infty]$  such that

$$\lim_{n \rightarrow \infty} n^{-1/p} \|Z_n\|_p = 0.$$

**Sequences of diagonal sampling matrices.** If  $n \in \mathbb{N}$  and  $a : [0, 1] \rightarrow \mathbb{C}$ , the  $n$ th diagonal sampling matrix generated by  $a$  is the  $n \times n$  diagonal matrix given by

$$D_n(a) = \text{diag}_{i=1, \dots, n} a\left(\frac{i}{n}\right).$$

$\{D_n(a)\}_n$  is called the sequence of diagonal sampling matrices generated by  $a$ .

**Toeplitz sequences.** If  $n \in \mathbb{N}$  and  $f : [-\pi, \pi] \rightarrow \mathbb{C}$  is a function in  $L^1([-\pi, \pi])$ , the  $n$ th Toeplitz matrix generated by  $f$  is the  $n \times n$  matrix

$$T_n(f) = [f_{i-j}]_{i,j=1}^n = \begin{bmatrix} \hat{f}_0 & \hat{f}_{-1} & \hat{f}_{-2} & \cdots & \cdots & \hat{f}_{-(n-1)} \\ \hat{f}_1 & \ddots & \ddots & \ddots & & \vdots \\ \hat{f}_2 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \hat{f}_{-2} \\ \vdots & & \ddots & \ddots & \ddots & \hat{f}_{-1} \\ \hat{f}_{n-1} & \cdots & \cdots & \hat{f}_2 & \hat{f}_1 & \hat{f}_0 \end{bmatrix},$$



where the numbers  $\hat{f}_k$  are the Fourier coefficients of  $f$ ,

$$\hat{f}_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\theta) e^{-ik\theta} d\theta, \quad k \in \mathbb{Z}.$$

$\{T_n(f)\}_n$  is called the Toeplitz sequence generated by  $f$ .

**Approximating classes of sequences.** The notion of approximating classes of sequences (a.c.s.) is fundamental to the theory of GLT sequences and it is deeply studied in [18, chapter 5].

**DEFINITION 3.8.** *Let  $\{A_n\}_n$  be a matrix-sequence and let  $\{\{B_{n,m}\}_n\}_m$  be a sequence of matrix-sequences. We say that  $\{\{B_{n,m}\}_n\}_m$  is an approximating class of sequences (a.c.s.) for  $\{A_n\}_n$  if the following condition is met: for every  $m$  there exists  $n_m$  such that, for  $n \geq n_m$ ,*

$$A_n = B_{n,m} + R_{n,m} + N_{n,m}, \quad \text{rank}(R_{n,m}) \leq c(m)n, \quad \|N_{n,m}\| \leq \omega(m),$$

where  $n_m, c(m), \omega(m)$  depend only on  $m$ , and  $\lim_{m \rightarrow \infty} c(m) = \lim_{m \rightarrow \infty} \omega(m) = 0$ .

Roughly speaking,  $\{\{B_{n,m}\}_n\}_m$  is an a.c.s. for  $\{A_n\}_n$  if, for large  $m$ , the sequence  $\{B_{n,m}\}_n$  approximates  $\{A_n\}_n$  in the sense that  $A_n$  is eventually equal to  $B_{n,m}$  plus a small-rank matrix (with respect to the matrix size  $n$ ) plus a small-norm matrix. It turns out that the notion of a.c.s. is a notion of convergence in the space of matrix-sequences  $\mathcal{E} = \{\{A_n\}_n : \{A_n\}_n \text{ is a matrix-sequence}\}$ , i.e., there exists a topology  $\tau_{\text{a.c.s.}}$  on  $\mathcal{E}$  such that  $\{\{B_{n,m}\}_n\}_m$  is an a.c.s. for  $\{A_n\}_n$  if and only if  $\{\{B_{n,m}\}_n\}_m$  converges to  $\{A_n\}_n$  in  $(\mathcal{E}, \tau_{\text{a.c.s.}})$ ; see [18, section 5.2]. Motivated by this topological interpretation, we will use the convergence notation  $\{B_{n,m}\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n$  to indicate that  $\{\{B_{n,m}\}_n\}_m$  is an a.c.s. for  $\{A_n\}_n$ .

**GLT sequences.** A GLT sequence  $\{A_n\}_n$  is a special matrix-sequence equipped with a measurable function  $\kappa : [0, 1] \times [-\pi, \pi] \rightarrow \mathbb{C}$ , the so-called symbol (or kernel). We use the notation  $\{A_n\}_n \sim_{\text{GLT}} \kappa$  to indicate that  $\{A_n\}_n$  is a GLT sequence with symbol  $\kappa$ . The properties of GLT sequences that we shall need in this paper are listed below; the corresponding proofs can be found in [18].

**GLT 0.** If  $\{A_n\}_n \sim_{\text{GLT}} \kappa$  and  $\{A_n\}_n \sim_{\text{GLT}} \xi$  then  $\kappa = \xi$  almost everywhere (a.e.) in  $[0, 1] \times [-\pi, \pi]$ . Conversely, if  $\{A_n\}_n \sim_{\text{GLT}} \kappa$  and  $\kappa = \xi$  a.e. in  $[0, 1] \times [-\pi, \pi]$  then  $\{A_n\}_n \sim_{\text{GLT}} \xi$ .

**GLT 1.** If  $\{A_n\}_n \sim_{\text{GLT}} \kappa$  then  $\{A_n\}_n \sim_{\sigma} \kappa$ . If  $\{A_n\}_n \sim_{\text{GLT}} \kappa$  and the matrices  $A_n$  are Hermitian then  $\{A_n\}_n \sim_{\lambda} \kappa$ .

**GLT 2.** If  $\{A_n\}_n \sim_{\text{GLT}} \kappa$  and  $A_n = X_n + Y_n$ , where

- every  $X_n$  is Hermitian,
- $\|X_n\|, \|Y_n\| \leq C$  for some constant  $C$  independent of  $n$ ,
- $n^{-1}\|Y_n\|_1 \rightarrow 0$ ,

then  $\{A_n\}_n \sim_{\lambda} \kappa$ .

**GLT 3.** We have

- $\{T_n(f)\}_n \sim_{\text{GLT}} \kappa(x, \theta) = f(\theta)$  if  $f \in L^1([-\pi, \pi])$ ,
- $\{D_n(a)\}_n \sim_{\text{GLT}} \kappa(x, \theta) = a(x)$  if  $a : [0, 1] \rightarrow \mathbb{C}$  is Riemann-integrable,
- $\{Z_n\}_n \sim_{\text{GLT}} \kappa(x, \theta) = 0$  if and only if  $\{Z_n\}_n \sim_{\sigma} 0$ .

**GLT 4.** If  $\{A_n\}_n \sim_{\text{GLT}} \kappa$  and  $\{B_n\}_n \sim_{\text{GLT}} \xi$  then

- $\{A_n^*\}_n \sim_{\text{GLT}} \bar{\kappa}$ ,
- $\{\alpha A_n + \beta B_n\}_n \sim_{\text{GLT}} \alpha \kappa + \beta \xi$  for all  $\alpha, \beta \in \mathbb{C}$ ,
- $\{A_n B_n\}_n \sim_{\text{GLT}} \kappa \xi$ .

**GLT 5.**  $\{A_n\}_n \sim_{\text{GLT}} \kappa$  if and only if there exist GLT sequences  $\{B_{n,m}\}_n \sim_{\text{GLT}} \kappa_m$  such that  $\{B_{n,m}\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n$  and  $\kappa_m \rightarrow \kappa$  in measure over  $[0, 1] \times [-\pi, \pi]$ .

**4. Spectral and convergence analysis of the Discrete ALIF method.** We are now ready to perform our spectral and convergence analysis of the Discrete ALIF method described in section 2.3. We begin with the clustering analysis in section 4.1, we then investigate the singular value and eigenvalue distribution of the ALIF iteration matrices in section 4.2, and we finally focus on the convergence analysis in section 4.3. Our main results are Theorems 4.1, 4.4, 4.5, 4.6, the necessary condition (4.21) for the convergence of the Discrete ALIF method, and the simple criterion provided in Remark 4.10 to construct a filter  $k$  which ensures that (4.21) is satisfied.

**4.1. Clustering analysis.** The main result of this section is the following.

**THEOREM 4.1.** *Let  $K_n$  be the ALIF iteration matrix (2.4). Let  $k$  be any filter and suppose that  $\ell(x) \geq \ell_*$  for some constant  $\ell_* > 0$  and for all  $x \in [0, 1]$ . Then,  $\{K_n\}_n$  is strongly clustered at  $\{0\}$  both in the sense of eigenvalues and singular values. In particular,  $\{K_n\}_n \sim_{\sigma, \lambda} 0$ .*

To prove Theorem 4.1 we need the following lemma. The result of this lemma may somehow be derived from more general results on sampling matrices [1], but for the reader's convenience we include a short and direct proof.

**LEMMA 4.2.** *Let  $A_n = \frac{1}{n-1} [\phi(x_i, x_j)]_{i,j=0}^{n-1}$ , where  $\phi(x, y)$  is a bounded function over  $[0, 1]^2$  and  $0 \leq x_0 \leq \dots \leq x_{n-1} \leq 1$ . Then  $\|A_n\|_2 \leq \frac{n}{n-1} \|\phi\|_\infty$  for all  $n$ .*

*Proof.* It suffices to observe that

$$\|A_n\|_2^2 = \sum_{i,j=0}^{n-1} |(A_n)_{ij}|^2 = \frac{1}{(n-1)^2} \sum_{i,j=0}^{n-1} |\phi(x_i, x_j)|^2 \leq \frac{n^2}{(n-1)^2} \|\phi\|_\infty^2. \quad \square$$

Note that the estimate in Lemma 4.2 is sharp. Indeed, if we take  $\phi(x, y) = 1$  identically, then  $\|A_n\|_2 = \frac{n}{n-1} \|\phi\|_\infty$ .

*Proof of Theorem 4.1.* As it is clear from (2.4), the ALIF iteration matrix  $K_n$  can be written in the form

$$K_n = \frac{1}{n-1} [\phi(x_i, x_j)]_{i,j=0}^{n-1},$$

where  $\phi(x, y) = k((x-y)/\ell(x))/\ell(x)$  and  $x_i = \frac{i}{n-1}$  for  $i = 0, \dots, n-1$ . Due to the assumption that  $\ell(x) \geq \ell_* > 0$  for all  $x \in [0, 1]$ , the function  $\phi(x, y)$  is bounded over  $[0, 1]^2$ . The hypotheses of Lemmas 4.2 are then satisfied and so  $\|A_n\|_2 \leq M$  for some constant  $M$  and for all  $n$ . It follows from Proposition 3.6 that  $\{K_n\}_n$  is strongly clustered at  $\{0\}$  both in the sense of eigenvalues and singular values. In particular,  $\{K_n\}_n \sim_{\sigma, \lambda} 0$  by Proposition 3.5.  $\square$

**4.2. Singular value and eigenvalue distribution.** By Theorem 4.1, the sequence  $\{K_n\}_n$  has an eigenvalue distribution described by the identically zero function. This sole information, however, is not particularly helpful in view of a convergence analysis, because it does not provide any indication on how to choose the filter  $k$  and the length  $\ell(x)$  so as to ensure the convergence of the Discrete ALIF method. Indications on how to make these choices will be given in section 4.3 on the basis of the results of the present section, where we investigate the eigenvalue distribution of  $\{K_n\}_n$  under the artificial (but useful) assumption that the function

$$(4.1) \quad L(x) = (n-1)\ell(x)$$

is independent of  $n$ . Assuming (4.1), the ALIF iteration matrix (2.4) takes the form

$$(4.2) \quad K_n = \left[ \frac{k\left(\frac{i-j}{L(x_i)}\right)}{L(x_i)} \right]_{i,j=0}^{n-1}.$$

REMARK 4.3. *Suppose  $L(x)$  is bounded over  $[0, 1]$ . Then, the matrix  $K_n$  is banded with bandwidth bounded by  $2\|L\|_\infty + 1$ . Indeed, if  $|i - j| > \|L\|_\infty$  then the point  $(i - j)/L(x_i)$  lies outside  $[-1, 1]$  and hence outside the support of the filter  $k(y)$ , so  $(K_n)_{ij} = 0$ . Suppose now that  $L(x)$  is bounded over  $[0, 1]$  and  $L(x) \geq L_*$  for some constant  $L_* > 0$  and for all  $x \in [0, 1]$ . Then  $\|K_n\| \leq (2\|L\|_\infty + 1)\|k\|_\infty/L_*$  for all  $n$  by virtue of the inequality*

$$(4.3) \quad \|A\| \leq \sqrt{\left( \max_{i=1, \dots, n} \sum_{j=1}^n |a_{ij}| \right) \left( \max_{j=1, \dots, n} \sum_{i=1}^n |a_{ij}| \right)},$$

which holds for all  $n \times n$  matrices  $A$ ; see, e.g., Corollary 2.3.2 of [20].

The main results of this section are the following two theorems. Throughout this paper, the characteristic (indicator) function of a set  $E$  is denoted by  $\chi_E$ .

THEOREM 4.4. *Suppose  $L(x) = (n - 1)\ell(x)$  is independent of  $n$ . Assume that  $L(x)$  is a step function of the form  $L(x) = \sum_{r=0}^{s-1} \alpha_r \chi_{I_r}(x)$ , where  $\alpha_0, \dots, \alpha_{s-1} > 0$  and  $\{I_0, \dots, I_{s-1}\}$  is a partition of  $[0, 1]$  such that  $I_r = [\xi_r, \xi_{r+1})$  for  $r = 0, \dots, s - 2$  and  $I_{s-1} = [\xi_{s-1}, \xi_s]$ , with  $0 = \xi_0 < \dots < \xi_s = 1$ . Then, for the sequence of ALIF iteration matrices  $\{K_n\}_n$  we have*

$$(4.4) \quad \{K_n\}_n \sim_{\text{GLT}} \kappa(x, \theta),$$

$$(4.5) \quad \{K_n\}_n \sim_{\sigma, \lambda} \kappa(x, \theta),$$

where

$$(4.6) \quad \kappa(x, \theta) = \frac{1}{L(x)} \sum_{j \in \mathbb{Z}} k\left(\frac{j}{L(x)}\right) e^{ij\theta} = \frac{1}{L(x)} \left( k(0) + 2 \sum_{j=1}^{\lceil L(x) \rceil - 1} k\left(\frac{j}{L(x)}\right) \cos(j\theta) \right).$$

Note that the second equality in (4.6) is simply due to the fact that the length of the filter  $k$  is 1.

THEOREM 4.5. *Suppose  $L(x) = (n - 1)\ell(x)$  is independent of  $n$ . Assume that the filter  $k$  is continuous on  $\mathbb{R}$  and that  $L(x)$  is continuous on  $[0, 1]$  with  $L(x) \geq L_*$  for some constant  $L_* > 0$  and for all  $x \in [0, 1]$ . Then, for the sequence of ALIF iteration matrices  $\{K_n\}_n$  the relations (4.4)–(4.5) hold with  $\kappa(x, \theta)$  defined as in (4.6).*

In order to prove Theorems 4.4–4.5, we first establish the following result, which is stated as a theorem because it is interesting also in itself. Indeed, it provides a noteworthy extension of the theory of GLT sequences and may therefore be regarded as a further main result of this paper. We recall that the direct sum of two matrices  $X_1 \in \mathbb{C}^{s_1 \times t_1}$  and  $X_2 \in \mathbb{C}^{s_2 \times t_2}$  is denoted by  $X_1 \oplus X_2$  and is defined as

$$X_1 \oplus X_2 = \begin{bmatrix} X_1 & O \\ O & X_2 \end{bmatrix} \in \mathbb{C}^{(s_1+s_2) \times (t_1+t_2)}.$$

THEOREM 4.6. *Let  $A_n = T_{n_1}(f_1) \oplus \dots \oplus T_{n_s}(f_s)$ , where*

- $f_1, \dots, f_s \in L^1([-\pi, \pi])$ ,
- $n_1, \dots, n_s$  are  $s$  integers such that  $\sum_{j=1}^s n_j = n$  and  $\lim_{n \rightarrow \infty} n_j/n = t_j$  for  $j = 1, \dots, s$ .

Then  $\{A_n\}_n \sim_{\text{GLT}} \kappa(x, \theta)$ , where  $\kappa(x, \theta) = \sum_{j=1}^s \chi_{(r_{j-1}, r_j)}(x) f_j(\theta)$  and  $r_j = \sum_{i=1}^j t_i$  for  $j = 0, \dots, s$  ( $r_0 = 0$ ).

*Proof.* For  $j = 1, \dots, s$ , let

$$T_{n,j} = O_{n_1} \oplus \dots \oplus O_{n_{j-1}} \oplus T_{n_j}(f_j) \oplus O_{n_{j+1}} \oplus \dots \oplus O_{n_s},$$

where  $O_m$  denotes the  $m \times m$  zero matrix. We show that, for every  $j = 1, \dots, s$ ,

$$(4.7) \quad T_{n,j} = D_n(\psi_j) T_n(f_j) D_n(\psi_j) + R_{n,j},$$

where  $\psi_j = \chi_{(r_{j-1}, r_j)}$  and

$$(4.8) \quad \lim_{n \rightarrow \infty} \frac{\text{rank}(R_{n,j})}{n} = 0.$$

Once this is proved, the desired GLT relation  $\{A_n\}_n \sim_{\text{GLT}} \kappa(x, \theta)$  follows from the equation

$$A_n = \sum_{j=1}^s T_{n,j} = \sum_{j=1}^s D_n(\psi_j) T_n(f_j) D_n(\psi_j) + \sum_{j=1}^s R_{n,j}$$

and from **GLT 3–GLT 4**, considering that each sequence  $\{R_{n,j}\}_n$  is zero-distributed by Proposition 3.7. To prove (4.7)–(4.8), we set  $\varepsilon_j = t_j - n_j/n$  for  $j = 1, \dots, s$  (note that  $\varepsilon_j \rightarrow 0$  as  $n \rightarrow \infty$  by assumption). Moreover, we define  $N_j = \sum_{i=1}^j n_i$  and  $q_j = n \sum_{i=1}^j \varepsilon_i$  for  $j = 0, \dots, s$  ( $N_0 = q_0 = 0$ ). For  $j = 1, \dots, s$ , a direct computation shows that, for every  $u, v = 1, \dots, n$ ,

$$(T_{n,j})_{uv} = \begin{cases} (\widehat{f_j})_{u-v}, & \text{if } N_{j-1} < u, v \leq N_j, \\ 0, & \text{otherwise,} \end{cases}$$

$$(D_n(\psi_j) T_n(f_j) D_n(\psi_j))_{uv} = \begin{cases} (\widehat{f_j})_{u-v}, & \text{if } N_{j-1} + q_{j-1} < u, v < N_j + q_j, \\ 0, & \text{otherwise.} \end{cases}$$

Therefore, setting  $R_{n,j} = T_{n,j} - D_n(\psi_j) T_n(f_j) D_n(\psi_j)$ , for every  $u, v = 1, \dots, n$  the component  $(R_{n,j})_{uv}$  is zero for sure in the following cases:

(a) both  $u$  and  $v$  belong to

$$(\max(N_{j-1}, N_{j-1} + q_{j-1}), \min(N_j, N_j + q_j));$$

(b) at least one between  $u$  and  $v$  does not belong to

$$[\min(N_{j-1}, N_{j-1} + q_{j-1}), \max(N_j, N_j + q_j)].$$

It follows that

$$\text{rank}(R_{n,j}) \leq 2(|q_{j-1}| + |q_j|) \leq 4n \sum_{i=1}^s |\varepsilon_i|,$$

and (4.7)–(4.8) are proved.  $\square$

We are now ready to prove Theorems 4.4–4.5.

*Proof of Theorem 4.4.* For each fixed  $n$ , denote by  $i_0, \dots, i_s$  the indices such that  $0 = i_0 < \dots < i_s = n$  and  $x_i = \frac{i}{n-1} \in I_r$  for all  $i = i_r, \dots, i_{r+1} - 1$  and all  $r = 0, \dots, s-1$ . Then, for every  $r = 0, \dots, s-1$ , every  $i = i_r, \dots, i_{r+1} - 1$  and every  $j = 0, \dots, n-1$ , we have

$$(K_n)_{ij} = \frac{k\left(\frac{i-j}{L(x_i)}\right)}{L(x_i)} = \frac{k\left(\frac{i-j}{\alpha_r}\right)}{\alpha_r}.$$

Hence,

$$(4.9) \quad K_n = \begin{bmatrix} \hline K_{n,0} \\ \hline K_{n,1} \\ \hline \vdots \\ \hline K_{n,s-1} \end{bmatrix},$$

where

$$K_{n,r} = \left[ \frac{k\left(\frac{i-j}{\alpha_r}\right)}{\alpha_r} \right]_{\substack{i=i_r, \dots, i_{r+1}-1 \\ j=0, \dots, n-1}}, \quad r = 0, \dots, s-1.$$

The matrix  $K_{n,r}$  can be decomposed as follows:

$$K_{n,r} = [ A_{n,r} \mid T_{n,r} \mid B_{n,r} ],$$

where

$$A_{n,r} = \left[ \frac{k\left(\frac{i-j}{\alpha_r}\right)}{\alpha_r} \right]_{\substack{i=i_r, \dots, i_{r+1}-1 \\ j=0, \dots, i_r-1}}, \quad B_{n,r} = \left[ \frac{k\left(\frac{i-j}{\alpha_r}\right)}{\alpha_r} \right]_{\substack{i=i_r, \dots, i_{r+1}-1 \\ j=i_{r+1}, \dots, n-1}},$$

and

$$T_{n,r} = \left[ \frac{k\left(\frac{i-j}{\alpha_r}\right)}{\alpha_r} \right]_{i,j=i_r}^{i_{r+1}-1} = \left[ \frac{k\left(\frac{i-j}{\alpha_r}\right)}{\alpha_r} \right]_{i,j=0}^{i_{r+1}-i_r-1} = T_{n_r}(f_r),$$

with  $n_r = i_{r+1} - i_r$  and

$$(4.10) \quad f_r(\theta) = \sum_{j \in \mathbb{Z}} \frac{k\left(\frac{j}{\alpha_r}\right)}{\alpha_r} e^{ij\theta} = \frac{1}{\alpha_r} \left( k(0) + 2 \sum_{j=1}^{\lceil \alpha_r \rceil - 1} k\left(\frac{j}{\alpha_r}\right) \cos(j\theta) \right);$$

note that in the last equality we used the evenness of  $k$  and the fact that  $k(y) = 0$  for  $|y| \geq 1$  (because the length of  $k$  is 1). Let  $\tilde{K}_{n,r}$  be the matrix obtained from  $K_{n,r}$  by setting to zero the blocks  $A_{n,r}$  and  $B_{n,r}$ , i.e.,

$$\tilde{K}_{n,r} = [ O \mid T_{n,r} \mid O ] = [ O \mid T_{n_r}(f_r) \mid O ].$$

Let  $\tilde{K}_n$  be the matrix obtained from  $K_n$  by replacing each  $K_{n,r}$  in (4.9) with  $\tilde{K}_{n,r}$ , i.e.,

$$\tilde{K}_n = \begin{bmatrix} \hline \tilde{K}_{n,0} \\ \hline \tilde{K}_{n,1} \\ \hline \vdots \\ \hline \tilde{K}_{n,s-1} \end{bmatrix} = T_{n_0}(f_0) \oplus T_{n_1}(f_1) \oplus \dots \oplus T_{n_{s-1}}(f_{s-1}).$$

Considering that  $k\left(\frac{i-j}{\alpha_r}\right) = 0$  whenever  $|i-j| \geq \lceil \alpha_r \rceil$ , for every  $r = 0, \dots, s-1$  we have

$$\begin{aligned}
& \#\{i \in \{i_r, \dots, i_{r+1} - 1\} : \text{the } i\text{th row of } K_{n,r} - \tilde{K}_{n,r} \text{ is nonzero}\} \\
&= \#\{i \in \{i_r, \dots, i_{r+1} - 1\} : \\
&\quad \text{at least one between the } i\text{th row of } A_{n,r} \text{ and the } i\text{th row of } B_{n,r} \text{ is nonzero}\} \\
&= \#\left\{i \in \{i_r, \dots, i_{r+1} - 1\} : \right. \\
&\quad \left. k\left(\frac{i-j}{\alpha_r}\right) \neq 0 \text{ for some } j \in \{0, \dots, i_r - 1, i_{r+1}, \dots, n-1\}\right\} \\
&\leq \#\{i \in \{i_r, \dots, i_{r+1} - 1\} : \\
&\quad \{0, \dots, i_r - 1, i_{r+1}, \dots, n-1\} \cap \{i - \lceil \alpha_r \rceil + 1, \dots, i + \lceil \alpha_r \rceil - 1\} \neq \emptyset\} \\
&= \#\{i \in \{i_r, \dots, i_{r+1} - 1\} : i_r - 1 \geq i - \lceil \alpha_r \rceil + 1 \quad \vee \quad i_{r+1} \leq i + \lceil \alpha_r \rceil - 1\} \\
&\leq 2\lceil \alpha_r \rceil - 2.
\end{aligned}$$

Thus, setting  $E_n = K_n - \tilde{K}_n$ , we have

$$\begin{aligned}
\text{rank}(E_n) &\leq \#\{i \in \{0, \dots, n-1\} : \text{the } i\text{th row of } K_n - \tilde{K}_n \text{ is nonzero}\} \\
&\leq \sum_{r=0}^{s-1} (2\lceil \alpha_r \rceil - 2) \leq 2\|L\|_\infty s.
\end{aligned}$$

Considering that  $k$  is bounded and  $L(x) \geq L_* = \min(\alpha_0, \dots, \alpha_{s-1}) > 0$  for all  $x \in [0, 1]$ , the modulus of each entry of  $E_n$  is bounded by  $\|k\|_\infty / L_*$ . Moreover,  $E_n$  is banded with bandwidth bounded by  $2\|L\|_\infty + 1$ , as a consequence of the bandedness of  $K_n$  established in Remark 4.3. Thus, by (4.3) we obtain

$$\|E_n\| \leq (2\|L\|_\infty + 1) \frac{\|k\|_\infty}{L_*}.$$

We have therefore proved that the ALIF iteration matrix  $K_n$  can be written as

$$(4.11) \quad K_n = \tilde{K}_n + E_n = T_{n_0}(f_0) \oplus \dots \oplus T_{n_{s-1}}(f_{s-1}) + E_n,$$

where

- $n_0, \dots, n_{s-1}$  are  $s$  integers such that  $\sum_{j=1}^s n_j = n$  and

$$\lim_{n \rightarrow \infty} \frac{n_r}{n} = \lim_{n \rightarrow \infty} \frac{i_{r+1} - i_r}{n} = \xi_{r+1} - \xi_r, \quad r = 0, \dots, s-1,$$

- the functions  $f_0, \dots, f_{s-1}$  are defined by (4.10) for  $r = 0, \dots, s-1$ ,
- $E_n$  satisfies

$$(4.12) \quad \text{rank}(E_n) \leq 2\|L\|_\infty s,$$

$$(4.13) \quad \|E_n\| \leq (2\|L\|_\infty + 1) \frac{\|k\|_\infty}{L_*}, \quad L_* = \min(\alpha_0, \dots, \alpha_{s-1}).$$

By applying Theorem 4.6 and **GLT 0**, we obtain

$$\begin{aligned}
\{K_n\}_n &\underset{\text{GLT}}{\sim} \sum_{r=0}^{s-1} \chi_{I_r}(x) f_r(\theta) = \sum_{r=0}^{s-1} \chi_{I_r}(x) \sum_{j \in \mathbb{Z}} \frac{k\left(\frac{j}{\alpha_r}\right)}{\alpha_r} e^{ij\theta} \\
&= \sum_{r=0}^{s-1} \chi_{I_r}(x) \sum_{j \in \mathbb{Z}} \frac{k\left(\frac{j}{L(x)}\right)}{L(x)} e^{ij\theta} = \kappa(x, \theta),
\end{aligned}$$

which is (4.4). The singular value distribution in (4.5) follows immediately from (4.4) and **GLT 1**. The eigenvalue distribution in (4.5) follows from (4.4) and **GLT 2**, taking into account the decomposition (4.11), the symmetry of the matrix  $\tilde{K}_n$ , the bound  $\|K_n\| \leq (2\|L\|_\infty + 1)\|k\|_\infty/L_*$  (see Remark 4.3), the estimates (4.12)–(4.13), and the obvious inequality  $\|E_n\|_1 \leq \text{rank}(E_n)\|E_n\|$ .  $\square$

*Proof of Theorem 4.5.* Let  $\{L_m\}_m$  be a sequence of step functions as in Theorem 4.4 such that

$$(4.14) \quad L_* \leq L_m(x) \leq \|L\|_\infty$$

for all  $m$  and  $x \in [0, 1]$ , and  $L_m \rightarrow L$  uniformly over  $[0, 1]$ . Define  $\{K_{n,m}\}_n$  to be the sequence of ALIF iteration matrices associated with  $L_m$ , that is,

$$K_{n,m} = \left[ \frac{k\left(\frac{i-j}{L_m(x_i)}\right)}{L_m(x_i)} \right]_{i,j=0}^{n-1}.$$

By Theorem 4.4 we immediately have

$$(4.15) \quad \{K_{n,m}\}_n \underset{\text{GLT}}{\sim} \kappa_m(x, \theta) = \frac{1}{L_m(x)} \sum_{j \in \mathbb{Z}} k\left(\frac{j}{L_m(x)}\right) e^{ij\theta}.$$

Moreover, since  $L_m \rightarrow L$  uniformly over  $[0, 1]$ , it is clear that

$$(4.16) \quad \kappa_m(x, \theta) \rightarrow \kappa(x, \theta) \text{ a.e. (and hence also in measure) over } [0, 1] \times [-\pi, \pi].$$

We show that

$$(4.17) \quad \{K_{n,m}\}_n \xrightarrow{\text{a.c.s.}} \{K_n\}_n.$$

Once this is done, the application of **GLT 5** in combination with (4.15)–(4.17) yields (4.4). Since the support of the filter  $k$  is contained in  $[-1, 1]$ , if  $|i-j| > \|L\|_\infty$  then, by (4.14) and the definitions of  $K_n$  and  $K_{n,m}$ ,

$$(K_n)_{ij} = (K_{n,m})_{ij} = 0.$$

Hence,  $K_n$  and  $K_{n,m}$  are banded with bandwidth bounded by  $2\|L\|_\infty + 1$ . Let  $\omega_k$  be the modulus of continuity of  $k$ . For all  $i, j = 0, \dots, n-1$  such that  $|i-j| \leq \|L\|_\infty$  we have

$$\begin{aligned} |(K_n)_{ij} - (K_{n,m})_{ij}| &= \left| \frac{1}{L(x_i)} k\left(\frac{i-j}{L(x_i)}\right) - \frac{1}{L_m(x_i)} k\left(\frac{i-j}{L_m(x_i)}\right) \right| \\ &\leq k\left(\frac{i-j}{L(x_i)}\right) \left| \frac{1}{L(x_i)} - \frac{1}{L_m(x_i)} \right| + \frac{1}{L_m(x_i)} \left| k\left(\frac{i-j}{L(x_i)}\right) - k\left(\frac{i-j}{L_m(x_i)}\right) \right| \\ &\leq \frac{\|k\|_\infty}{(L_*)^2} \|L_m - L\|_\infty + \frac{1}{L_*} \omega_k\left(\frac{\|L\|_\infty}{(L_*)^2} \|L_m - L\|_\infty\right) = \varepsilon(m). \end{aligned}$$

By (4.3),

$$\|K_n - K_{n,m}\| \leq (2\|L\|_\infty + 1)\varepsilon(m) = \omega(m),$$

which tends to 0 as  $m \rightarrow \infty$ . Thus, the a.c.s. convergence (4.17) is proved and **GLT 5** yields (4.4). The singular value distribution in (4.5) follows immediately from (4.4) and **GLT 1**. To obtain the eigenvalue distribution in (4.5), we now define a symmetric

approximation  $\tilde{K}_n$  of the ALIF matrix  $K_n$  and we show that the assumptions in **GLT 2** are satisfied with  $X_n = \tilde{K}_n$  and  $Y_n = (K_n - \tilde{K}_n)$ . Let

$$\tilde{K}_n = \left[ \frac{k\left(\frac{i-j}{L(x_{i \wedge j})}\right)}{L(x_{i \wedge j})} \right]_{i,j=0}^{n-1},$$

where  $i \wedge j = \min(i, j)$ . The matrix  $\tilde{K}_n$  is clearly symmetric. As in Remark 4.3, one can show that  $(K_n)_{ij} = (\tilde{K}_n)_{ij} = 0$  whenever  $|i - j| > \|L\|_\infty$  (so  $K_n, \tilde{K}_n$  are banded with bandwidth bounded by  $2\|L\|_\infty + 1$ ) and  $\|K_n\|, \|\tilde{K}_n\| \leq C$  for some constant  $C$  independent of  $n$ . For all  $i, j = 0, \dots, n-1$  such that  $|i - j| \leq \|L\|_\infty$ , we have  $|x_i - x_{i \wedge j}| \leq \|L\|_\infty / (n-1)$  and

$$\begin{aligned} |(K_n)_{ij} - (\tilde{K}_n)_{ij}| &= \left| \frac{1}{L(x_i)} k\left(\frac{i-j}{L(x_i)}\right) - \frac{1}{L(x_{i \wedge j})} k\left(\frac{i-j}{L(x_{i \wedge j})}\right) \right| \\ &\leq k\left(\frac{i-j}{L(x_i)}\right) \left| \frac{1}{L(x_i)} - \frac{1}{L(x_{i \wedge j})} \right| + \frac{1}{L(x_{i \wedge j})} \left| k\left(\frac{i-j}{L(x_i)}\right) - k\left(\frac{i-j}{L(x_{i \wedge j})}\right) \right| \\ &\leq \frac{\|k\|_\infty}{(L_*)^2} |L(x_{i \wedge j}) - L(x_i)| + \frac{1}{L_*} \omega_k \left( \frac{\|L\|_\infty}{(L_*)^2} |L(x_{i \wedge j}) - L(x_i)| \right) \\ &\leq \frac{\|k\|_\infty}{(L_*)^2} \omega_L \left( \frac{\|L\|_\infty}{n-1} \right) + \frac{1}{L_*} \omega_k \left( \frac{\|L\|_\infty}{(L_*)^2} \omega_L \left( \frac{\|L\|_\infty}{n-1} \right) \right) = \delta(n). \end{aligned}$$

By (4.3),

$$\|K_n - \tilde{K}_n\| \leq (2\|L\|_\infty + 1)\delta(n),$$

which tends to 0 as  $n \rightarrow \infty$ . From the obvious inequality  $\|K_n - \tilde{K}_n\|_1 \leq n\|K_n - \tilde{K}_n\|$  we infer that  $\|K_n - \tilde{K}_n\|_1/n$  tends to 0 as well. In conclusion, all the assumptions of **GLT 2** are satisfied with  $X_n = \tilde{K}_n$  and  $Y_n = (K_n - \tilde{K}_n)$ , and the eigenvalue distribution in (4.5) is proved.  $\square$

At this point, we should say the following. Except for the essential assumption that  $L(x) = (n-1)\ell(x)$  is independent of  $n$ , the other hypotheses on  $L(x)$  and  $k$  appearing in the statements of Theorems 4.4–4.5 are a bit “artificial”, in the sense that they have just been formulated in view of the arguments used in the proofs. It is actually very likely that these hypotheses can be significantly relaxed according to the following conjecture.

**CONJECTURE 4.7.** *Suppose  $L(x) = (n-1)\ell(x)$  is independent of  $n$ . Assume that  $k(j/L(x))/L(x)$  is Riemann-integrable over  $[0, 1]$  for all  $j \in \mathbb{Z}$ . Then, for the sequence of ALIF iteration matrices  $\{K_n\}_n$  the relations (4.4)–(4.5) hold with  $\kappa(x, \theta)$  defined as in (4.6).*

Providing a clear motivation to Conjecture 4.7 may require several pages. We limit to say that this conjecture is suggested by the theory of GLT sequences [18, 19, 38, 39].

**4.3. Convergence analysis.** In real-world applications, the size  $n$  is fixed to a number  $N$ , which represents the cardinality of the set of points  $x_i$  where a given signal  $g$  has been sampled. Usually,  $N$  is quite large, but it cannot be increased further as it is given once and for all. Let

$$(4.18) \quad L(x) = (N-1)\ell(x).$$

From now on we assume that the sequence of ALIF iteration matrices  $\{K_n\}_n$ , with  $K_n$  defined as in (4.2) and  $L(x)$  defined as in (4.18), satisfies the thesis of Theorems 4.4–4.5



and Conjecture 4.7, i.e.,

$$\begin{aligned}\{K_n\}_n &\sim_{\text{GLT}} \kappa(x, \theta), \\ \{K_n\}_n &\sim_{\sigma, \lambda} \kappa(x, \theta),\end{aligned}$$

where

$$(4.19) \quad \kappa(x, \theta) = \frac{1}{L(x)} \sum_{j \in \mathbb{Z}} k\left(\frac{j}{L(x)}\right) e^{ij\theta} = \frac{1}{L(x)} \left( k(0) + 2 \sum_{j=1}^{\lceil L(x) \rceil - 1} k\left(\frac{j}{L(x)}\right) \cos(j\theta) \right).$$

According to the informal meaning of the eigenvalue distribution given in section 3, if  $n$  is large enough then the eigenvalues of  $K_n$ , except possibly for  $o(n)$  outliers, are approximately equal to the samples of  $\kappa(x, \theta)$  over a uniform grid in  $[0, 1] \times [-\pi, \pi]$ . Since  $N$  is large, we may certainly assume that the previous assertion applies to the matrix  $K_N$ , which is the actual ALIF iteration matrix we are given in practice. Thus, in order to ensure that the moduli of the eigenvalues of  $K_N$  belong to the closed disk with center 1 and radius 1, so that the necessary condition (2.5) for the convergence of the Discrete ALIF method is satisfied, it is necessary that

$$(4.20) \quad |1 - \kappa(x, \theta)| \leq 1, \quad (x, \theta) \in [0, 1] \times [-\pi, \pi].$$

Since  $\kappa(x, \theta)$  is symmetric in the so-called Fourier variable  $\theta$ , i.e.,  $\kappa(x, -\theta) = \kappa(x, \theta)$ , it suffices to require that (4.20) holds for  $(x, \theta) \in [0, 1] \times [0, \pi]$ . In conclusion, taking also into account that  $\kappa(x, \theta)$  is real, the necessary condition for the convergence of the Discrete ALIF method that arises from the analysis of this paper can be written as follows:

$$(4.21) \quad 0 \leq \kappa(x, \theta) \leq 2, \quad (x, \theta) \in [0, 1] \times [0, \pi],$$

where  $\kappa(x, \theta)$  is defined in (4.19) and  $L(x)$  is defined in (4.18).

REMARK 4.8. *It follows immediately from (4.19) that*

$$(4.22) \quad |\kappa(x, \theta)| \leq \frac{1}{L(x)} \sum_{j \in \mathbb{Z}} k\left(\frac{j}{L(x)}\right).$$

*For every  $x \in [0, 1]$ , the right-hand side of (4.22) is a Riemann sum for the integral  $\int_{\mathbb{R}} k(y) dy = 1$ , and it is therefore expected to be approximately equal to 1, also because  $N$  is normally large in practice and hence  $L(x)$  is large as well, due to its definition (4.18). Thus, the following approximate inequality is satisfied for any choice of the filter  $k$ :*

$$(4.23) \quad |\kappa(x, \theta)| \preccurlyeq 1, \quad (x, \theta) \in [0, 1] \times [0, \pi].$$

By Remark 4.8, the right inequality in (4.21) is always met in practice. It is then natural to ask in which cases the left inequality is met. As shown in Theorem 4.9, this happens if the filter  $k$  is obtained as the convolution of another filter  $w$  with itself, i.e., if<sup>5</sup>

$$(4.24) \quad k(y) = \int_{\mathbb{R}} w(y-z)w(z)dz, \quad y \in \mathbb{R}.$$

A completely analogous result was obtained in [12, section 2.1] when analyzing the convergence of the IF method.

<sup>5</sup>Note that  $k$  in (4.24) is indeed a filter according to the terminology introduced in section 2.1.

**THEOREM 4.9.** *Suppose that the filter  $k$  is obtained as the convolution of another filter  $w$  with itself according to (4.24). Then, the symbol  $\kappa$  in (4.19) satisfies the following properties: for every  $x \in [0, 1]$ ,*

$$(4.25) \quad \kappa(x, \theta) = \sum_{j \in \mathbb{Z}} |\widehat{w}_x(\theta + 2j\pi)|^2 \geq 0 \text{ a.e. in } [0, \pi],$$

where  $\widehat{w}_x(\theta) = \int_{\mathbb{R}} w_x(y) e^{-iy\theta} dy$  is the Fourier transform of the function

$$w_x(y) = \frac{w\left(\frac{y}{L(x)}\right)}{L(x)}, \quad y \in \mathbb{R}.$$

In particular,  $\kappa(x, \theta) \geq 0$  for all  $(x, \theta) \in [0, 1] \times [0, \pi]$ .

*Proof.* For every  $(x, \theta) \in [0, 1] \times [0, \pi]$ , by the evenness of  $w$  we have

$$\begin{aligned} \kappa(x, \theta) &= \frac{1}{L(x)} \sum_{j \in \mathbb{Z}} k\left(\frac{j}{L(x)}\right) e^{ij\theta} = \frac{1}{L(x)} \sum_{j \in \mathbb{Z}} \left( \int_{\mathbb{R}} w\left(\frac{j}{L(x)} - z\right) w(z) dz \right) e^{ij\theta} \\ &= \frac{1}{[L(x)]^2} \sum_{j \in \mathbb{Z}} \left( \int_{\mathbb{R}} w\left(\frac{j-y}{L(x)}\right) w\left(\frac{y}{L(x)}\right) dy \right) e^{ij\theta} \\ &= \sum_{j \in \mathbb{Z}} \left( \int_{\mathbb{R}} w_x(y-j) w_x(y) dy \right) e^{ij\theta} = \sum_{j \in \mathbb{Z}} W_x(-j) e^{ij\theta}, \end{aligned}$$

where  $W_x(t) = \int_{\mathbb{R}} w_x(y+t) w_x(y) dy$  is the so-called autocorrelation function of  $w_x$ . This immediately implies (4.25) for every  $x \in [0, 1]$ ; see, e.g., [11, p. 47]. Since for every  $x \in [0, 1]$  the function  $\theta \mapsto \kappa(x, \theta)$  is a trigonometric polynomial such that  $\kappa(x, \theta) \geq 0$  for almost every  $\theta \in [0, \pi]$ , we must actually have  $\kappa(x, \theta) \geq 0$  for all  $\theta \in [0, \pi]$ .  $\square$

**REMARK 4.10.** *In view of Remark 4.8 and Theorem 4.9, if the filter  $k$  is obtained as the convolution of another filter  $w$  with itself according to (4.24), then the necessary condition (4.21) for the convergence of the Discrete ALIF method is satisfied. Although this assertion is not completely rigorous from a mathematical viewpoint, it can be considered as a “theorem” in the world of practical applications.*

**5. Numerical examples.** This section presents a few numerical examples in support of the analysis carried out in section 4.

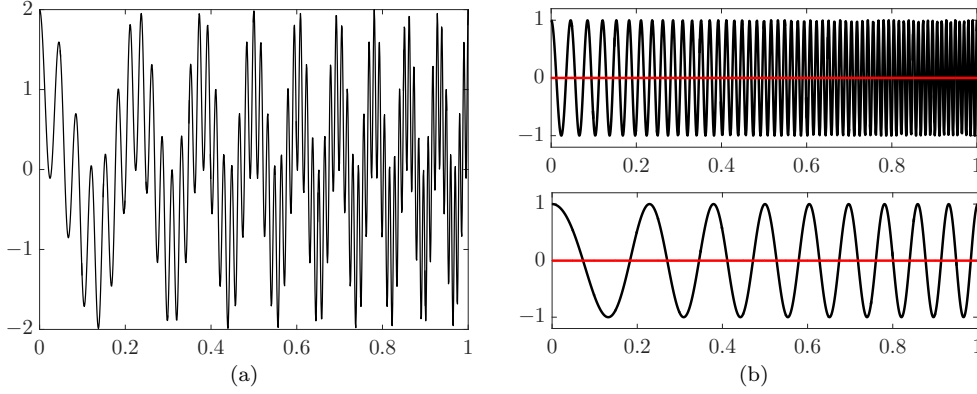
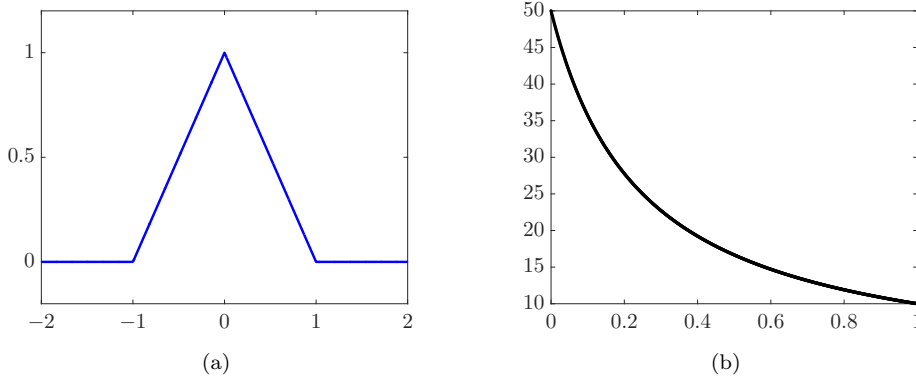
**EXAMPLE 5.1.** *Let  $g(x) = g_1(x) + g_2(x)$  be the signal obtained as the superposition of the two IMFs given by*

$$\begin{aligned} g_1(x) &= \cos(2\pi(40x^2 + 20x)), \\ g_2(x) &= \cos(2\pi(6x^2 + 3x)); \end{aligned}$$

see Fig. 5.1. Suppose  $g$  is sampled at the  $N = 1000$  points  $x_i = \frac{i}{N-1}$ ,  $i = 0, \dots, N-1$ , and  $g = 0$  outside the reference interval  $[0, 1]$ . Suppose further that the Discrete ALIF method (Algorithm 2.2) is applied with filter  $k_x$  as in (2.1) and with

$$(5.1) \quad k(y) = (y+1)\chi_{(-1,0]}(y) + (1-y)\chi_{(0,1)}(y),$$

$$(5.2) \quad \ell(x) = \frac{1}{20(4x+1)}.$$

FIG. 5.1. (a) Signal  $g$ ; (b) Components  $g_1$  and  $g_2$  used to create  $g$ .FIG. 5.2. (a) Double ramp filter; (b) Graph of the normalized filter length  $L(x) = (N - 1)\ell(x)$ .

The filter  $k$  is referred to as the double ramp filter; its graph is depicted in Fig. 5.2(a). The function  $\ell(x)$  is the reciprocal of the so-called instantaneous frequency of  $g_1(x)$ , which in turn is computed according to the definition proposed in [12]. The resulting ALIF iteration matrix  $K_N$  is defined by (4.2) for  $n = N$  and

$$(5.3) \quad L(x) = (N - 1)\ell(x) = \frac{999}{20(4x + 1)};$$

see Fig. 5.2(b).

By Theorem 4.5, we have  $\{K_n\}_n \sim_\lambda \kappa(x, \theta)$ , with  $\kappa(x, \theta)$  defined as in (4.19). Since  $\kappa(x, \theta)$  is symmetric in the Fourier variable  $\theta$ , the restriction  $\kappa|_{[0,1] \times [0,\pi]}$  is a rearranged version of  $\kappa$  according to Definition 3.2. Let  $\varphi$  be the rearranged version of  $\kappa$  obtained from Proposition 3.3 applied to  $f = \kappa|_{[0,1] \times [0,\pi]}$ . In Fig. 5.3(a) we plotted the graph of  $\varphi$  and the real parts of the eigenvalues of  $K_N$ . The imaginary parts of the eigenvalues are not shown because they are negligible (their maximum modulus is about  $9.4 \cdot 10^{-3}$ ). The graph of  $\varphi$  was obtained by plotting the graph of the function  $\varphi_r$  in (3.3) for a large value of  $r$  ( $r = 500$ ). The real parts of the eigenvalues of  $K_N$  have been arranged in non-decreasing order,  $\text{Re}(\lambda_1(K_N)) \leq \dots \leq \text{Re}(\lambda_N(K_N))$ , and are represented in the figure by the dashed line connecting the points

$$\left( \frac{i}{N}, \text{Re}(\lambda_i(K_N)) \right), \quad i = 1, \dots, N.$$

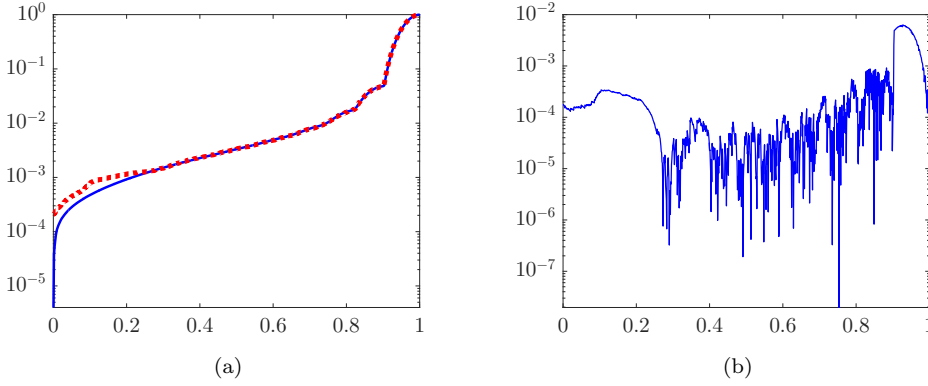


FIG. 5.3. (a) Comparison between  $\varphi$  (solid line) and the real parts of the eigenvalues of  $K_N$  (dashed line); (b) Differences, in absolute value, between the real parts of the eigenvalues of  $K_N$  and the corresponding values on the curve  $\varphi$ .

We see from the figure that, in accordance with the informal meaning of the eigenvalue distribution (see section 3), the eigenvalues of  $K_N$  are approximately equal to the uniform samples

$$\varphi\left(\frac{i}{N}\right), \quad i = 1, \dots, N.$$

This is also confirmed by Fig. 5.3(b), where we show the differences between the real parts of the eigenvalues of  $K_N$  and the corresponding values on the curve  $\varphi$ .

Suppose now that we apply the Discrete ALIF Algorithm 2.2<sup>6</sup> with the following options:

- One IMF — `ALIF.NIMFs = 1`
- Homogeneous Dirichlet boundary conditions — `ALIF.extensionType = z`
- Double ramp filter — `ALIF.FilterShape = DR`
- Stopping criterion — `ALIF.delta = 0.0001`

Then, the signal  $g$  is decomposed into one IMF and a remainder given in Fig. 5.4(a). The differences between the components produced by Algorithm 2.2 and the ground truth shown in Fig. 5.1(b) are plotted in Fig. 5.4(b). We clearly see that Algorithm 2.2 is convergent in this case. Hence, the necessary condition (4.21) for the convergence of the Discrete ALIF method should be satisfied. In fact, the filter  $k$  in (5.1) is the convolution of the unitary step filter  $\chi_{(-0.5, 0.5)}$  with itself, and so (4.21) is automatically satisfied by Remark 4.10. Actually, we could also infer that (4.21) is satisfied by looking at the graph of the rearranged version  $\varphi$  in Fig. 5.3(a), where we see that  $0 \leq \varphi(t) \leq C \approx 1$  over  $[0, 1]$  and, consequently,  $0 \leq \kappa(x, \theta) \leq C$  over  $[0, 1] \times [0, \pi]$ .

EXAMPLE 5.2. Let  $h(x) = h_1(x) + h_2(x)$  be the signal obtained as the superposition of the two IMFs given by

$$\begin{aligned} h_1(x) &= \cos\left(4\pi(-32(x-0.5)^2 - 20(x+0.5))\right), \\ h_2(x) &= \cos\left(4\pi(-32(x-0.5)^2 - 4(x+0.5))\right); \end{aligned}$$

see Fig. 5.5. Suppose  $h$  is sampled at the  $N = 1000$  points  $x_i = \frac{i}{N-1}$ ,  $i = 0, \dots, N-1$ , and  $h = 0$  outside  $[0, 1]$ . We apply the Discrete ALIF Algorithm 2.2 with the same

<sup>6</sup>Available at [www.cicone.com](http://www.cicone.com).

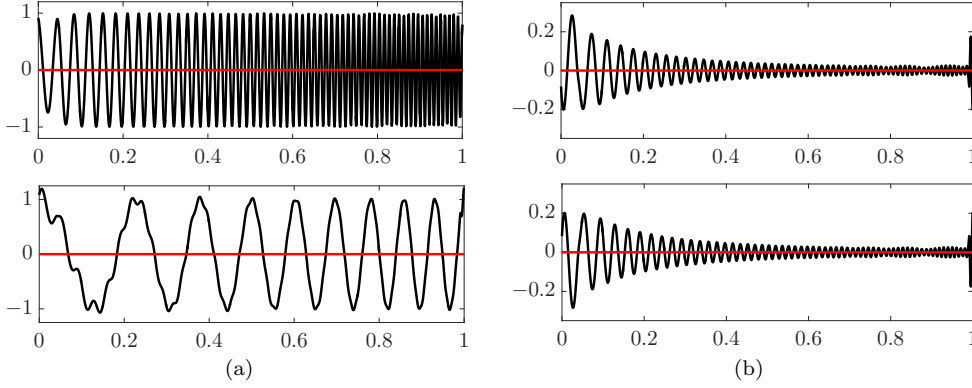


FIG. 5.4. (a) Decomposition of  $g$  produced by the Discrete ALIF Algorithm 2.2; (b) Differences between the components produced by the Discrete ALIF Algorithm 2.2 and the ground truth shown in Fig. 5.1(b).

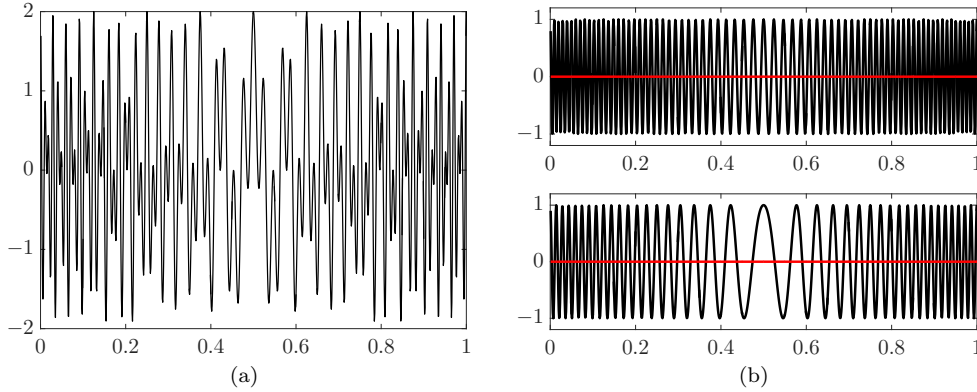


FIG. 5.5. (a) Signal  $h$ ; (b) Components  $h_1$  and  $h_2$  used to create  $h$ .

filter  $k_x$  as in (2.1), with  $k$  as in (5.1) and Fig. 5.2(a), and with

$$\ell(x) = \frac{1}{8(16|x - 0.5| + 5)}.$$

The function  $\ell(x)$  is obtained as the reciprocal of the instantaneous frequency of  $h_1(x)$ , according to the definition in [12]. The resulting matrix  $K_N$  is defined by (4.2) for  $n = N$  and

$$L(x) = (N - 1)\ell(x) = \frac{999}{8(16|x - 0.5| + 5)};$$

see Fig. 5.6. In this example, we repeat the same study as in Example 5.1. We will therefore skip several details for the sake of brevity.

In Fig. 5.7(a) we plot the graph of  $\varphi$  and the real parts of the eigenvalues of  $K_N$ . The imaginary parts of the eigenvalues are not shown because they are negligible (their maximum modulus is about  $6.2 \cdot 10^{-3}$ ). The graph of  $\varphi$  was obtained by plotting the graph of the function  $\varphi_r$  in (3.3) for  $r = 500$ . The real parts of the eigenvalues of  $K_N$  have been arranged in non-decreasing order and are represented by the dashed line connecting the points

$$\left(\frac{i}{N}, \text{Re}(\lambda_i(K_N))\right), \quad i = 1, \dots, N.$$

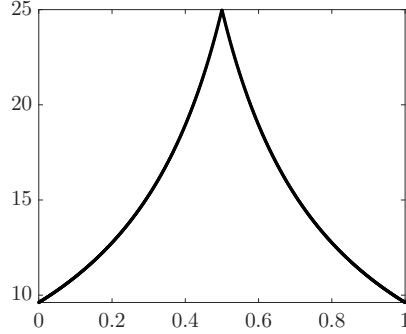


FIG. 5.6. Graph of the normalized filter length  $L(x) = (N - 1)\ell(x)$ .

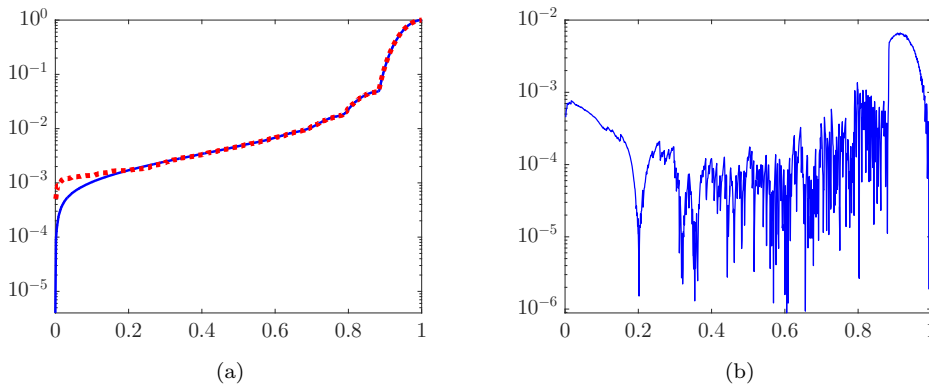


FIG. 5.7. (a) Comparison between  $\varphi$  (solid line) and the real parts of the eigenvalues of  $K_N$  (dashed line); (b) Differences, in absolute value, between the real parts of the eigenvalues of  $K_N$  and the corresponding values on the curve  $\varphi$ .

From the figure we see that the eigenvalues of  $K_N$  are approximately equal to the uniform samples

$$\varphi\left(\frac{i}{N}\right), \quad i = 1, \dots, N.$$

This is also confirmed by Fig. 5.7(b), where we plot the differences between the real parts of the eigenvalues of  $K_N$  and the corresponding values on the curve  $\varphi$ .

The Discrete ALIF Algorithm 2.2 with the same options as in Example 5.1 allows one to decompose  $h$  into one IMF and a remainder given in Fig. 5.8(a). The differences between the components produced by Algorithm 2.2 and the ground truth shown in Fig. 5.5(b) are shown in Fig. 5.8(b). We point out that Algorithm 2.2 is convergent even in this second example and hence condition (4.21) should be satisfied. This is clearly true for the same reason as given at the end of Example 5.1.

EXAMPLE 5.3. Let  $g$  be the signal of Example 5.1, which is sampled again at the  $N = 1000$  points  $x_i = \frac{i}{N-1}$ ,  $i = 0, \dots, N-1$ , and is assumed to be zero outside  $[0, 1]$ . We apply the Discrete ALIF Algorithm 2.2 with  $\ell(x)$  as in (5.2), with  $k_x$  as in (2.1), and with the unitary step filter

$$k(y) = \chi_{(-0.5, 0.5)}(y);$$

see Fig. 5.9. The resulting matrix  $K_N$  is defined by (4.2) for  $n = N$  and  $L(x)$  as

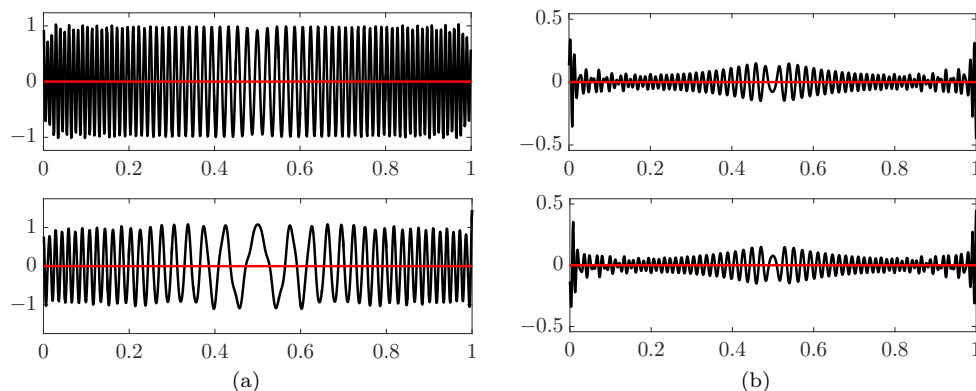


FIG. 5.8. (a) Decomposition of  $h$  produced by the Discrete ALIF Algorithm 2.2; (b) Differences between the components produced by the Discrete ALIF Algorithm 2.2 and the ground truth shown in Fig. 5.5(b).

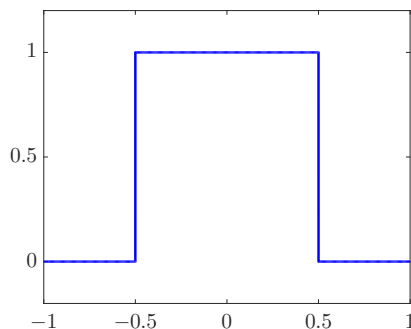


FIG. 5.9. Unitary step filter.

in (5.3) and Fig. 5.2(b). We are going to repeat once again the same study as in Examples 5.1 and 5.2. We will therefore skip many details for the sake of brevity.

In Fig. 5.10(a) we plot the graph of  $\varphi$  and the real parts of the eigenvalues of  $K_N$  (the imaginary parts are not shown as their maximum modulus is about  $2.4 \cdot 10^{-2}$  and hence negligible). In Fig. 5.10(b) we plot the differences between the real parts of the eigenvalues of  $K_N$  and the corresponding values on the curve  $\varphi$ . From Fig. 5.10 we clearly see that, even in this case, there is a very good agreement between the eigenvalues of  $K_N$  and the rearranged version  $\varphi$  of the symbol  $\kappa(x, \theta)$ .

We note that the necessary condition (4.21) for the convergence of the Discrete ALIF method is not satisfied in this case. Indeed, as it is clear from Fig. 5.10(a), the rearranged version  $\varphi$  (and hence also the symbol  $\kappa(x, \theta)$ ) takes negative values in the range  $[-0.2, 0)$ . Therefore, as expected and as it is clear from Fig. 5.10(a), the real parts of the eigenvalues of  $K_N$  take values in the same range, implying that the necessary condition (2.5) for the convergence of the Discrete ALIF method is not satisfied. If we apply the Discrete ALIF Algorithm 2.2, with the same options as in Example 5.1 and with the only exception that the filter  $k$  is now the unitary step filter, the algorithm does not converge anymore. If we stop the algorithm after 100 steps, we obtain the decomposition of  $g$  shown in Fig. 5.11, which clearly shows that the technique is now divergent. Note that the unitary step filter  $k$  cannot be generated as

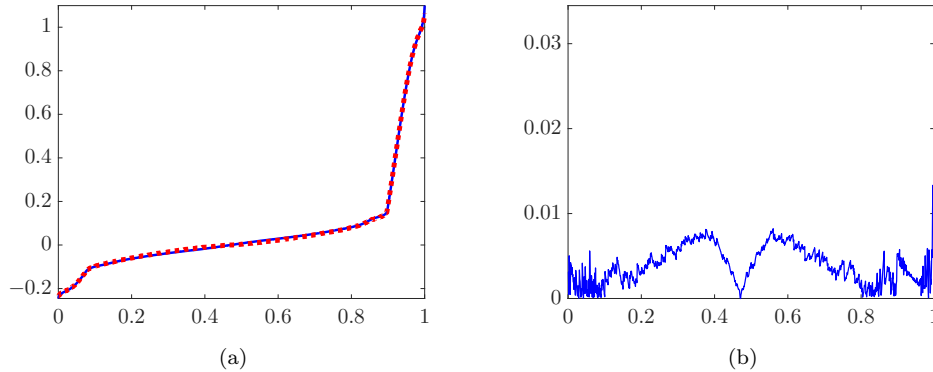


FIG. 5.10. (a) Comparison between  $\varphi$  (solid line) and the real parts of the eigenvalues of  $K_N$  (dashed line); (b) Differences, in absolute value, between the real parts of the eigenvalues of  $K_N$  and the corresponding values on the curve  $\varphi$ .

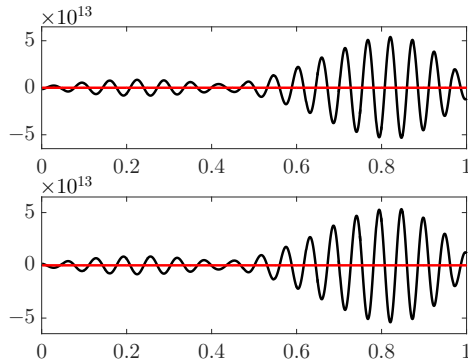


FIG. 5.11. Decomposition of  $g$  produced by the Discrete ALIF Algorithm 2.2.

the convolution of another filter  $w$  with itself, because any filter generated in this way is continuous on  $\mathbb{R}$ ; see, e.g., Lemma 5.9.4 of [36].

In all the numerical experiments we conducted, whenever condition (4.21) was satisfied the Discrete ALIF method was convergent. We therefore propose the following conjecture.

CONJECTURE 5.4. *The necessary condition (4.21) is also sufficient for the convergence of the Discrete ALIF method.*

**6. Conclusions and perspectives.** In this paper we have considered the discrete version of the ALIF method. We have performed a spectral analysis of the related iteration matrices, with particular focus on the eigenvalue clustering and the eigenvalue distribution; our main results in this direction are Theorems 4.1, 4.4 and 4.5. Based on the eigenvalue distribution, we have also formulated a necessary condition for the convergence of the Discrete ALIF method, that is, condition (4.21). Moreover, we have provided in Remark 4.10 a simple criterion to construct appropriate filters  $k$  for which (4.21) is satisfied. The analysis of this paper is based on recent results about sampling matrices [1] and, above all, on the theory of GLT sequences [18, 19, 38, 39], which we have extended in Theorem 4.6; in this sense, the latter theorem may be considered as a further main result of the paper. Several numerical examples in support



of the theoretical analysis have been presented in section 5.

The present paper is the first step toward a complete mathematical analysis of the ALIF method. Here we outline a few possible future lines of research.

1. Prove (or disprove) Conjectures 4.7 and 5.4.
2. This paper addresses the case of homogeneous Dirichlet boundary conditions. In particular, the signal  $\mathbf{g}_m$  is set to 0 outside  $\{0, \dots, n-1\}$  at every iteration of the Discrete ALIF inner loop (see Algorithm 2.2). However, other conditions are usually employed (periodic, reflective, anti-reflective, etc.). The analysis of the Discrete ALIF method in the case of boundary conditions other than Dirichlet's is certainly an interesting subject for future research.
3. Following what has been done for the IF method [13, 14], extend the ALIF algorithm and the analysis of this paper to higher dimensions.

#### REFERENCES

- [1] A. S. AL-FHAID, S. SERRA-CAPIZZANO, D. SESANA, AND M. Z. ULLAH, *Singular-value (and eigenvalue) distribution and Krylov preconditioning of sequences of sampling matrices approximating integral operators*, Numer. Linear Algebra Appl., 21 (2014), pp. 722–743.
- [2] X. AN, *Local rub-impact fault diagnosis of a rotor system based on adaptive local iterative filtering*, Trans. Inst. Meas. Control, 39 (2017), pp. 748–753.
- [3] X. AN, C. LI, AND F. ZHANG, *Application of adaptive local iterative filtering and approximate entropy to vibration signal denoising of hydropower unit*, J. Vibroeng., 18 (2016), pp. 4299–4311.
- [4] X. AN AND L. PAN, *Wind turbine bearing fault diagnosis based on adaptive local iterative filtering and approximate entropy*, Proc. Inst. Mech. Eng. C: J. Mech. Eng. Sci. (in press) <http://dx.doi.org/10.1177/0954406216642478>.
- [5] X. AN, W. YANG, AND X. AN, *Vibration signal analysis of a hydropower unit based on adaptive local iterative filtering*, Proc. Inst. Mech. Eng. C: J. Mech. Eng. Sci., 231 (2017), pp. 1339–1353.
- [6] X. AN, H. ZENG, AND C. LI, *Demodulation analysis based on adaptive local iterative filtering for bearing fault diagnosis*, Measurement, 94 (2016), pp. 554–560.
- [7] R. BALOCCHI, D. MENICUCCI, E. SANTARCANGELO, L. SEBASTIANI, A. GEMIGNANI, B. GHE-LARDUCCI, AND M. VARANINI, *Deriving the respiratory sinus arrhythmia from the heartbeat time series using empirical mode decomposition*, Chaos, Solitons & Fractals, 20 (2004), pp. 171–177.
- [8] R. BHATIA, *Matrix Analysis*, Springer, 1997.
- [9] M. BLANCO-VELASCO, B. WENG, AND K. E. BARNER, *ECG signal denoising and baseline wander correction based on the empirical mode decomposition*, Comput. Biol. Med., 38 (2008), pp. 1–13.
- [10] X. CHEN, X. ZHANG, J. A. CHURCH, C. S. WATSON, M. A. KING, D. MONSELESAN, B. LEGRESY, AND C. HARIG, *The increasing rate of global mean sea-level rise during 1993–2014*, Nat. Clim. Change, 7 (2017), pp. 492–495.
- [11] C. K. CHUI, *An Introduction to Wavelets*, Academic Press, 1992.
- [12] A. CICONE, J. LIU, AND H. ZHOU, *Adaptive local iterative filtering for signal decomposition and instantaneous frequency analysis*, Appl. Comput. Harmon. Anal., 41 (2016), pp. 384–411.
- [13] A. CICONE, J. LIU, AND H. ZHOU, *Hyperspectral chemical plume detection algorithms based on multidimensional iterative filtering decomposition*, Phil. Trans. R. Soc. A: Math. Phys. Eng. Sci., 374 (2016), pp. 20150196.
- [14] A. CICONE AND H. ZHOU, *Multidimensional iterative filtering method for the decomposition of high-dimensional non-stationary signals*, Numer. Math. Theory Methods Appl., 10 (2017), pp. 278–298.
- [15] A. CICONE AND H. ZHOU, *Numerical analysis of the iterative filtering algorithm for the decomposition of non-stationary and non-linear signals*, preprint.
- [16] K. T. COUGHLIN AND K. TUNG, *11-year solar cycle in the stratosphere extracted by the empirical mode decomposition method*, Adv. Space Res., 34 (2004), pp. 323–329.
- [17] J. C. ECHEVERRIA, J. A. CROWE, M. S. WOOLFSON, AND B. R. HAYES-GILL, *Application of empirical mode decomposition to heart rate variability analysis*, Med. Biol. Eng. Comput., 39 (2001), pp. 471–479.

- [18] C. GARONI AND S. SERRA-CAPIZZANO, *Generalized Locally Toeplitz Sequences: Theory and Applications*, volume I, Springer, 2017.
- [19] C. GARONI AND S. SERRA-CAPIZZANO, *Generalized Locally Toeplitz Sequences: Theory and Applications*, Techn. report 2017-002, Department of Information Technology, Uppsala University, 2017. Preliminary version of: C. GARONI AND S. SERRA-CAPIZZANO, *Generalized Locally Toeplitz Sequences: Theory and Applications*, volume II, in preparation.
- [20] G. H. GOLUB AND C. F. VAN LOAN, *Matrix Computations*, 4th ed., The Johns Hopkins University Press, 2013.
- [21] G. G. GREGORIOU, S. J. GOTTS, H. ZHOU, AND R. DESIMONE, *High-frequency, long-range coupling between prefrontal and visual cortex during attention*, *Science*, 324 (2009), pp. 1207–1210.
- [22] D. J. GUBLER, *Cities spawn epidemic dengue viruses*, *Nat. Med.*, 10 (2004), pp. 129–130.
- [23] N. E. HUANG, Z. SHEN, S. R. LONG, M. C. WU, H. H. SHIH, Q. ZHENG, N. C. YEN, C. C. TUNG, AND H. H. LIU, *The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis*, *Proc. R. Soc. A: Math. Phys. Eng. Sci.*, 454 (1998), pp. 903–995.
- [24] C. HUANG, L. YANG, AND Y. WANG, *Convergence of a convolution-filtering-based algorithm for empirical mode decomposition*, *Adv. Adapt. Data Anal.*, 1 (2009), pp. 561–571.
- [25] F. JI, Z. WU, J. HUANG, AND E. P. CHASSIGNET, *Evolution of land surface air temperature trend*, *Nat. Clim. Change*, 4 (2014), pp. 462.
- [26] S. J. KIM AND H. ZHOU, *A multiscale computation for highly oscillatory dynamical systems using Empirical Mode Decomposition (EMD)-type methods*, *Multiscale Model. Simul.*, 14 (2016), pp. 534–557.
- [27] Y. LEI, J. LIN, Z. HE, AND M. J. ZUO, *A review on empirical mode decomposition in fault diagnosis of rotating machinery*, *Mech. Syst. Signal Proc.*, 35 (2013), pp. 108–126.
- [28] H. LIANG, Q. LIN, AND J. D. Z. CHEN, *Application of the empirical mode decomposition to the analysis of esophageal manometric data in gastroesophageal reflux disease*, *IEEE Trans. Biomed. Eng.*, 52 (2005), pp. 1692–1701.
- [29] L. LIN, Y. WANG, AND H. ZHOU, *Iterative filtering as an alternative algorithm for empirical mode decomposition*, *Adv. Adapt. Data Anal.*, 1 (2009), pp. 543–560.
- [30] C. LOH, T. WU, AND N. E. HUANG, *Application of the empirical mode decomposition-Hilbert spectrum method to identify near-fault ground-motion characteristics and structural responses*, *Bull. Seismol. Soc. Am.*, 91 (2001), pp. 1339–1357.
- [31] B. MIJOVIC, M. DE VOS, I. GLIGORJEVIC, J. TAELEMAN, AND S. VAN HUFFEL, *Source separation from single-channel recordings by combining empirical mode decomposition and independent component analysis*, *IEEE Trans. Biomed. Eng.*, 57 (2010), pp. 2188–2196.
- [32] I. MITICHE, G. MORISON, A. NESBITT, M. HUGHES-NARBOROUGH, P. BOREHAM, AND B. G. STEWART, *An evaluation of total variation signal denoising methods for partial discharge signals*, *International Electrical Insulation Conference – INSUCON 2017*, pp. 1–5.
- [33] J. C. NUNES, Y. BOUAOUNE, E. DELECELLE, O. NIANG, AND P. BUNEL, *Image analysis by bidimensional empirical mode decomposition*, *Imag. Vision Comput.*, 21 (2003), pp. 1019–1026.
- [34] J. C. NUNES, S. GUYOT, AND E. DELECELLE, *Texture analysis based on local analysis of the bidimensional empirical mode decomposition*, *Mach. Vision Appl.*, 16 (2005), pp. 177–188.
- [35] R. B. PACHORI, *Discrimination between ictal and seizure-free EEG signals using empirical mode decomposition*, *Research Letters in Signal Processing*, 2008 (2008), 5 pages.
- [36] M. A. PICARDELLO, *Analisi Armonica: Aspetti Classici e Numerici*, Lecture Notes in Mathematics, University of Rome “Tor Vergata”, 02/02/2017. Available at <http://www.mat.uniroma2.it/~picard/SMC/didattica/materiali.did/Anal.Armon./LIBRO.pdf>.
- [37] M. PIERSANTI, M. MATERASSI, A. CICONE, L. SPOGLI, H. ZHOU, AND R. G. EZQUER, *Adaptive local iterative filtering: a promising technique for the analysis of nonstationary signals*, submitted.
- [38] S. SERRA-CAPIZZANO, *Generalized locally Toeplitz sequences: spectral analysis and applications to discretized partial differential equations*, *Linear Algebra Appl.*, 366 (2003), pp. 371–402.
- [39] S. SERRA-CAPIZZANO, *The GLT class as a generalized Fourier analysis and applications*, *Linear Algebra Appl.*, 419 (2006), pp. 180–233.
- [40] R. SHARMA, R. B. PACHORI, AND A. UPADHYAY, *Automatic sleep stages classification based on iterative filtering of electroencephalogram signals*, *Neural Comput. Appl.* (in press) <http://dx.doi.org/10.1007/s00521-017-2919-6>
- [41] N. VARADARAJAN AND S. NAGARAJAIAH, *Wind response control of building with variable stiffness tuned mass damper using empirical mode decomposition/Hilbert transform*, *J. Eng. Mech.*, 130 (2004), pp. 451–458.

- [42] Z. WU AND N. E. HUANG, *Ensemble empirical mode decomposition: a noise-assisted data analysis method*, *Adv. Adapt. Data Anal.*, 1 (2009), pp. 1–41.
- [43] D. YANG, B. WANG, G. CAI, AND J. WEN, *Oscillation mode analysis for power grids using adaptive local iterative filter decomposition*, *Int. J. Elec. Power Energy Syst.*, 92 (2017), pp. 25–33.
- [44] X. ZHANG, K. K. LAI, AND S. WANG, *A new approach for crude oil price analysis based on empirical mode decomposition*, *Energy Econ.*, 30 (2008), pp. 905–918.