IMPROVING ENERGY COMPACTION OF ADAPTIVE FOURIER DECOMPOSITION

Adam Borowicz
Faculty of Computer Science, Department of Digital Media and Computer Graphics
Bialystok University of Technology
Wiejska str. 45A, 15-351 Bialystok, Poland
a.borowicz@pb.edu.pl

***Abstract*—Adaptive Fourier decomposition (AFD) provides an expansion of an analytic function into a sum of basic signals, called mono-components. Unlike the Fourier series decomposition, the AFD is based on an adaptive rational orthogonal system, hence it is better suited for analyzing non-stationary data. The most popular algorithm for the AFD decomposes any signal in such a way that the energy of the low-frequency components is maximized. Unfortunately, this results in poor energy compaction of high-frequency components. In this paper, we develop a novel algorithm for the AFD. The key idea is to maximize the energy of any components no matter how big or small the corresponding frequencies are. A comparative evaluation was conducted of the signal reconstruction efficiency of the proposed approach and several conventional algorithms by using speech recordings. The experimental results show that with the new algorithm, it is possible to get a better performance in terms of the reconstruction quality and energy compaction property.**

***Index Terms*—Fourier series, Takenaka-Malmquist system, mono-components, adaptive decomposition, energy compaction**

I. Introduction

Spectrum analysis is the core of many signal processing applications. In engineering, it is the process of decomposing a signal into simpler components. Fourier-based methods are probably the most popular tool used for this purpose. Unfortunately, conventional methods such as Fourier series expansion may be inefficient in the analysis of non-stationary data. Recently, the adaptive Fourier decomposition (AFD) has been proposed [1]. The AFD has a firm mathematical foundation in complex analysis and analytic function theories. It provides an expansion of an analytic function into a sum of basic components with a non-negative analytic phase derivative. Measurable functions with this property are called mono-components. The AFD is based on the rational orthogonal system in which parameters are selected adaptively for a given signal. Fourier series decomposition (FD) can be considered as a special case of the AFD when these parameters are all chosen to be zeros. In opposition to the Fourier series, mono-components decomposed by the AFD are with time-varying instantaneous frequencies. Hence it is better suited for analyzing non-stationary data.

In recent years, the AFD has found several applications including system identification, signal compression, and noise reduction [2]–[4]. However, two major issues limit wider use of the AFD: its high computational complexity and poor energy compaction of high-frequency components. Possible solutions for the first problem include using the FFT algorithm [5], numerical optimization methods [6] and GPU-based implementation [7]. In this paper, we deal with the second problem.

The most popular approach, known as core AFD algorithm [1], [8], forces to extract, at the consecutive steps, the maximal energy in low-frequency components, which is not natural. Such a strategy works well only when the spectral content of a decomposed signal is dominated by low frequencies. However, in practice, there are many signals containing components that fall off at high frequencies. Applying the core AFD to such signals results in a significant reconstruction error or a large number of non-zero decomposition coefficients.

In the literature [8], we can also find several modifications of the core AFD including unwinding AFD and cyclic AFD. It was reported in [8] that the unwinding AFD achieves a faster energy convergence rate for the high-frequency signals. Although the algorithm proved to be suitable for compression of the electrocardiogram signals [4], it is complicated and computationally expensive. Furthermore, the unwinding AFD has not a simple form as it involves, at each decomposition level, factorization of the remainder into inner and outer functions. The cyclic AFD [8], [9] directly minimizes the reconstruction error for a given number of decomposition levels. Unfortunately, attaining the minimum does not always mean that we also get optimal energy compaction of the AFD coefficients. Indeed, the cost function can have multiple local minima that correspond to various energy distributions across the AFD spectrum.

In this paper, we propose a novel algorithm for the AFD that extracts the components of large energy portions no matter how big or small the corresponding frequencies are. The new optimization strategy can be viewed as an extension of the core AFD. The proposed method is evaluated using real speech recordings. It achieves better energy compaction than the conventional approaches and turns out to be well suited to speech processing applications. Unfortunately, it is also more computationally demanding compared to the core AFD.

This work was supported by Bialystok University of Technology under the grant WZ/WI-III/4/2020.

II. Adaptive Fourier Decomposition

Suppose a real-valued function $s(e^{it})$ defined in $t \in \langle 0; 2\pi)$. By using the Gabor method [10] one can associate it with a complex-valued function of the form:

$$f(e^{it}) = s(e^{it}) + i\mathcal{H}s(e^{it}), \tag{1}$$

where $\mathcal{H}$ stands for the Hilbert transformation of the context. This function has only positive frequencies in the Fourier spectrum and it is referred to as an analytic signal. For a given $f(e^{it})$, the AFD is expressed as follows [1], [8]:

$$f(e^{it}) = \sum_{k=1}^{\infty} c_k B_k(e^{it}), \tag{2}$$

where c_k is kth decomposition coefficient defined using the inner product:

$$c_k = \langle f, B_k \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(e^{it}) \bar{B}_k(e^{it}) dt, \tag{3}$$

with $\bar{B}_k$ denoting the complex conjugate of B_k. The function:

$$B_k = B_{\{a_1,...,a_k\}}(e^{it}) = \frac{\sqrt{1-|a_k|^2}}{1-\bar{a}_k e^{it}} \prod_{j=1}^{k-1} \frac{e^{it}-a_j}{1-\bar{a}_j e^{it}}, \tag{4}$$

is generated by a sequence $(a_1, ..., a_k)$ of complex numbers of the open unit disc, i.e., $a_j \in \mathbb{D}$, where $1 \leq j \leq k$ and $\mathbb{D} := \{z \in \mathbb{C} : |z| < 1\}$. The set of functions $\{B_k\}_{k=1}^{\infty}$ forms orthonormal system with respect to the inner product (3). It is also known as the Takenaka-Malmquist (TM) system, and it is complete in the Hardy space if and only if the following condition is satisfied:

$$\sum_{k=1}^{\infty} (1 - |a_k|) = \infty. \tag{5}$$

In practice, a discrete model of (3) is used and it is assumed that the signal $f(e^{it})$ is sampled uniformly at points $t \in \{t_n = 2\pi(n-1)/N, n = 1, \dots N\}$. Furthermore, the expansion (2) is truncated to a finite number of terms, so that $f(e^{it})$ can be approximated by the Kth partial sum:

$$f_K(e^{it_n}) = \sum_{k=1}^{K} c_k B_k(e^{it_n}), \quad n = 1, \dots, N, \tag{6}$$

with the squared reconstruction error given by

$$\|f - f_K\|^2 = \|f\|^2 - \sum_{k=1}^{K} |c_k|^2. \tag{7}$$

where $\|.\|$ denotes Euclidean norm. Please note that, if all parameters of the sequence $(a_1, ..., a_K)$ are fixed to zero, the AFD reduces to the Fourier series decomposition. If one chooses $a_1 = 0$ and $|a_k| < 1$ for $k = 2, ..., K$, then the functions $\{B_k\}_{k=1}^{K}$ are all mono-components. Most frequently, these parameters are adaptively chosen by following a given optimization strategy. In the case of the core AFD algorithm, they are selected consecutively, one-by-one by applying the so-called maximal selection principle (MSP). Namely, in order to identify the parameter a_k, the following optimization problem is solved [1]:

$$a_k = \arg\max_{a \in \mathbb{D}} \{|\langle f, B_{\{a_1,...,a_{k-1},a\}} \rangle|^2\}, \tag{8}$$

for $a_1, ..., a_{k-1}$ being previously determined. This procedure is repeated until either the decomposition level K is reached or (7) drops below some reasonably small threshold.

III. The AFD algorithm based on energy compaction

An energy convergence of the AFD has been theoretically proved in [1]. In practice, for sufficiently large K we can get near-zero reconstruction error for any signal and sequence $(a_1, ..., a_K)$ as long as $a_j \in \mathbb{D}$ for all $j = 1, ..., K$. However, the resulting AFD spectrum may contain a large number of non-zero coefficients that are difficult to encode. On the other hand, the packing energy of the signal into the low-frequency components as in the core AFD may be ineffective. Therefore, instead to minimize K, we propose to reduce the number of non-zero coefficients by maximizing the energy of any AFD component no matter how big or small corresponding frequencies are.

We assume that the decomposition parameters can be optimized consecutively, one-by-one. Similarly, as in the cyclic AFD [8], we also assume that K is given *a priori*, and that the parameters in the sequence $(a_1, ..., a_K)$ are already initialized to some arbitrarily chosen values. It is easy to see that the function B_j depends on a_j and some $j-1$ previously determined parameters. Thus the parameter a_k, for $k < j$, may have an effect not only on the coefficient c_k, but also on the coefficients $c_{k+1}, ..., c_j$. The key idea of the proposed method is to use an optimization strategy that takes into account the energies of all these coefficients. In particular, the jth AFD coefficient, provided that the kth parameter $a_k = a$, can be written as follows:

$$c_j^k(a) = \langle f, B_{\{a_1,...,a_{k-1},a,a_{k+1},...,a_j\}} \rangle, \tag{9}$$

where $j = k, ..., K$. Let us define the function of $a \in \mathbb{D}$ that evaluates the squared magnitude of the AFD coefficient of the largest energy:

$$\lambda_{\max}^k(a) = \max\{|c_k^k(a)|^2, |c_{k+1}^k(a)|^2, ..., |c_K^k(a)|^2\}. \tag{10}$$

We propose the following optimization strategy for determining the kth decomposition parameter:

$$a_k = \arg\max_{a \in \mathbb{D}} \lambda_{\max}^k(a). \tag{11}$$

Please note that, contrary to (8), while searching for optimal a_k, we have to evaluate not only $c_k^k(a_k = a) = \langle f, B_{\{a_1,...,a_{k-1},a\}} \rangle$, but also high-frequency coefficients assuming that the sequence $(a_{k+1}, ..., a_K)$ is approximately known. In fact, if we ignore in (10) the terms $c_j^k(a)$ for all $j > k$, then the algorithm becomes equivalent to the core AFD. Thus, the proposed optimization strategy can be viewed as an extension of the MSP.

The cost function (10) is non-linear and complicated; hence calculating its gradient as well as finding an analytic solution to (11) is rather a challenge. In this paper, we propose the exhaustive search method similar to that of the core AFD [1]. It is based on a predefined subset:

$$\mathbb{D}_{L\times M} := \{\hat{a}_j \subset \mathbb{D} : j = 1, 2, ..., LM\}, \tag{12}$$

which can be viewed as a grid mesh of size $L \times M$ nodes. In particular, the elements of $\mathbb{D}_{L\times M}$ can be uniformly distributed on $\mathbb{D}$ and expressed in polar coordinates:

$$\hat{a}_{(l-1)M+m} = r_{\max}\frac{l-1}{L}e^{i2\pi\frac{m-1}{M}}, \tag{13}$$

for $1 \le l \le L$ and $1 \le m \le M$, where L, M correspond to the radial and angular discretization of the unit disc, respectively. The value of parameter $r_{\max}$ is empirically set in the range $(0; 1\rangle$ in order to prevent from dividing by near-zero values in the denominator of (4). In practice, it may be a good idea to gradually increase this value as long as no further improvement is observed in a signal reconstruction quality. Alternatively, the coordinates of $\hat{a}_j$ can be randomly generated from a uniform distribution as suggested in [6].

The most computationally demanding part of the proposed method is an evaluation of the AFD coefficients (9). Since the subset $\mathbb{D}_{L\times M}$ is of size LM, at the kth decomposition level, we have to compute $LM \times K - k + 1$ coefficients. For convenience, they can be stored in the following matrix:

$$\begin{bmatrix} c_k^k(\hat{a}_1) & c_{k+1}^k(\hat{a}_1) & \cdots & c_K^k(\hat{a}_1) \\ c_k^k(\hat{a}_2) & c_{k+1}^k(\hat{a}_2) & \cdots & c_K^k(\hat{a}_2) \\ \vdots & \vdots & \ddots & \vdots \\ c_k^k(\hat{a}_{LM}) & c_{k+1}^k(\hat{a}_{LM}) & \cdots & c_K^k(\hat{a}_{LM}) \end{bmatrix}. \tag{14}$$

Thus, the solution to (11) can be found as $a_k = \hat{a}_j$, where j is a row number of the matrix (14) that contains the AFD coefficient of the largest energy. Theoretically, the elements of (14) can be computed from (3) by means of numerical integration. However, a direct evaluation of the function (4) can be numerically unstable and erroneous. Instead, a so-called reduced remainder is used:

$$g_1(e^{it}) = f(e^{it}), \tag{15}$$

$$g_k(e^{it}) = \left[g_{k-1}(e^{it}) - c_{k-1}e_{a_{k-1}}(e^{it})\right]/b_{a_{k-1}}(e^{it}), \tag{16}$$

for $k > 1$, where

$$e_a(z) = \frac{\sqrt{1-|a|^2}}{1-\bar{a}z}, \quad b_a(z) = \frac{z-a}{1-\bar{a}z}, \tag{17}$$

denotes the evaluator and Blaschke functions, respectively. It can be verified that $\langle g_k, e_{a_k}\rangle = \langle f, B_k\rangle = c_k$ because of the orthogonalization of $\{B_k\}_{k=1}^K$ [8]. Therefore, the elements in the first column of the matrix (14) can be approximated by

$$c_k^k(a) \approx \frac{1}{N}\sum_{n=1}^{N} g_k(e^{it_n})\bar{e}_a(e^{it_n}). \tag{18}$$

Please note that since the functions (17) do not depend on the signal $f(e^{it})$, they can be pre-computed for each element of the subset (12) at the points $t_1, ..., t_N$ and reused at each decomposition level, so that the computational load can be reduced significantly. Furthermore, we can decrease the number of multiplications by initializing the parameters $a_{k+1}, ..., a_K$ to zeros. Then, the remaining AFD coefficients can be computed at a reduced cost as follows:

$$c_j^k(a) \approx \frac{1}{N}\sum_{n=1}^{N} g_j^k(e^{it_n}), \quad j > k, \tag{19}$$

where

$$g_{k+1}^k(e^{it}) = \left[g_k(e^{it}) - c_k^k(a)e_a(e^{it})\right]/b_a(e^{it}), \tag{20}$$

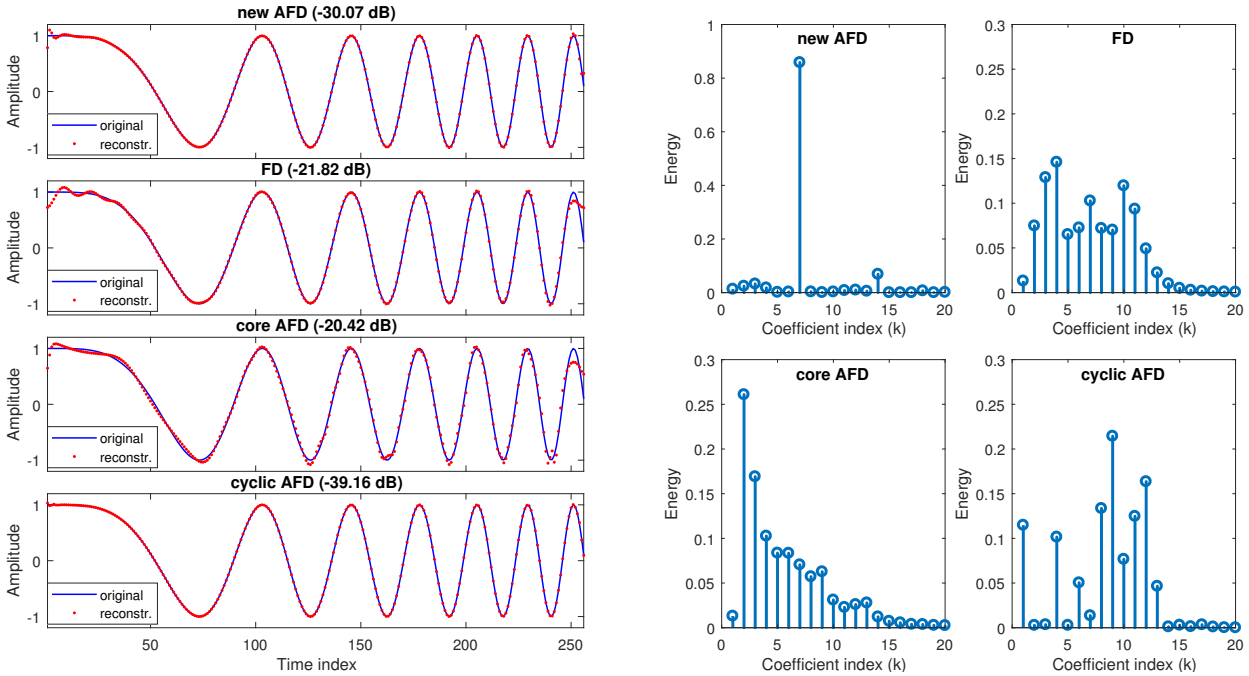
$$g_j^k(e^{it}) = \left[g_{j-1}^k(e^{it}) - c_{j-1}^k(a)\right]e^{-it}, \quad j > k+1. \tag{21}$$

It can be verified that, for K decomposition levels, the time complexity of the proposed algorithm (in terms of complex multiplications) is of order $O(LMNK^2)$, while the complexity of the core AFD is of order $O(LMNK)$.

Figure 1a demonstrates the reconstruction of the chirp signal using core AFD, cyclic AFD, conventional Fourier series, and the proposed method. In Fig. 1b we also show the squared magnitudes (energies) of the decomposition coefficients obtained for each algorithm. The proposed method offers the best energy compaction, as the corresponding spectrum contains the smallest number of the 'non-zero' coefficients of the largest energy. Almost total energy was packed into a single component, while in the case of the other algorithms this energy leaks into the multiple components.

IV. Experiments

The simulations were performed in MATLAB environment using real-life signals. For comparative purposes, we also implemented core AFD, cyclic AFD, and conventional Fourier series decomposition (FD) as a baseline method. In the case of the AFD algorithms the size of the grid mesh (12) was 32×32 nodes and $r_{\max} = 0.9$. As an input data, a phonetically balanced set of 8 short sentences uttered by both male and female speakers was selected from the TIMIT database [11]. The sentences were originally recorded at a sampling rate of 16 kHz. For convenience, they were downsampled to 8 kHz, and merged together to form one sequence about 30 s long. This sequence was partitioned into frames of length $N = 128$ samples (16 ms) with no overlap, which gives as in total 1814 data frames. Since a real-valued signal was used, the frames were Hilbert transformed to obtain their analytic signal representations. For each frame, we computed decomposition coefficients and the original sequence was reconstructed in a twofold manner. In the first approach, we selected only $P < K$ coefficients of the largest energy, spread over the entire spectrum. This is equivalent to setting in (6) $K - P$ coefficients of the smallest magnitudes to zeros. The second approach was to reconstruct the frames by selecting the first P components no matter how big or small the corresponding magnitudes are. In both cases, the simulations were performed for P ranging from 1 to 32, and $K = 64$. The effectiveness of the reconstruction was measured using relative energy



(a) (b)

Fig. 1: Reconstruction of the chirp signal using various methods for $N = 256$ and $K = 20$. (a) Real parts of the original signal (blue line) and the partial sum (6) (red dots). The reconstruction errors are given in the titles. (b) Energies of the corresponding decomposition coefficients.

error defined as $E_K = \|f - f_K\|^2/\|f\|^2$. This quantity was computed for each speech frame and then averaged over the whole sequence. As can be seen in Fig. 2a, if we select coefficients with the largest magnitudes, the proposed algorithm gives the lowest reconstruction error for the same number of the coefficients among all methods. Whereas, if only the first P components are used for reconstruction (see Fig. 2b), the core AFD algorithm has the best performance, but the resulting reconstruction error is higher than that of the previous approach for all methods including core AFD itself. The effectiveness of the proposed method was also verified using spectrograms in Fig. 3. We see that compared to other AFD algorithms, the high frequencies are better represented in the reconstructed signal.

In some applications like speech coding, we need to encode not only the AFD coefficients but also the parameters of the TM system as a side data [12]. If the reconstruction is based on the first P coefficients, then P additional parameters must be encoded. However, when the coefficients are spread over the entire spectrum this requirement may increase up to K parameters. Fortunately, since we use the predefined set (12), they can be encoded more efficiently as the integer numbers (indices). Furthermore, as can be deduced from Fig. 2 (bottom charts), for the proposed method many of these parameters are simply zeros. Indeed, since the cost function (10) depends on the signal, there is no guarantee that the maximum always exists. Thus, if we initialize these parameters to zeros, many of them may stay unchanged. This observation can be used in the practical coding algorithm so that only non-zero parameters should be encoded. Despite the clear benefits offered by the proposed algorithm, the question arises whether this improvement is sufficient to build an efficient compression method? In fact, the development of such a method is not an easy task, so this question remains open and will be investigated in future work. On the other hand, there are many applications like noise reduction, harmonic analysis, where the primary requirement is good energy compaction, and the issues related to encoding side data are less important.

V. Conclusion

We developed a new algorithm aimed at improving the energy compaction of the AFD. As in the conventional optimization strategy, the decomposition parameters are selected consecutively one-by-one, but when maximizing inner product energy also higher frequency AFD coefficients are taken into account. Although the proposed method tends to be computationally demanding, it is more suitable for the processing of real-life data. The evaluation results clearly show that with the new method the speech signals can be encoded more efficiently by using a smaller number of coefficients. It was also shown that reconstructing the speech signals using the largest decomposition coefficients usually gives a smaller reconstruction error than with a typical AFD framework.

Future works include reducing the computational complexity of the proposed algorithm and developing practical applications like signal compression and denoising.

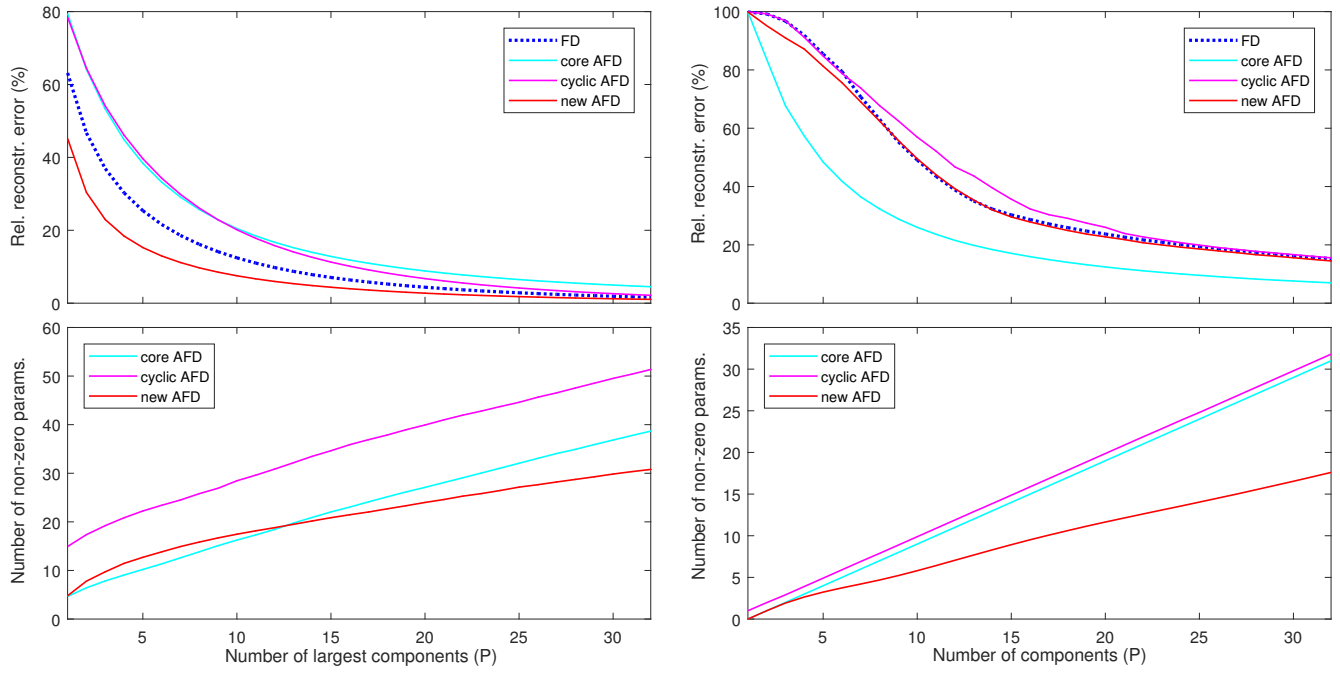


Fig. 2: Relative reconstruction error (top) and the number of non-zero decomposition parameters - side data - (bottom) averaged over all frames. (a) The scenario where the frames were reconstructed using P components of the largest energy. (b) The scenario where the frames were reconstructed using the first P components.

REFERENCES

[1] T. Qian, L. Zhang, and Z. Li, "Algorithm of adaptive Fourier decomposition," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 5899–5906, 2011.

[2] C. He, L. Zhang, X. He, and W. Jia, "A new image decomposition and reconstruction approach – adaptive Fourier decomposition," in *Int. Conf. on Multimedia Modeling*, 2015, pp. 227–236.

[3] Z. Wang, F. Wan, C. M. Wong, and L. Zhang, "Adaptive Fourier decomposition based ECG denoising," *Computers in Biology and Medicine*, vol. 77, pp. 195–205, 2016.

[4] C. Tan, L. Zhang, and H.-T. Wu, "A novel Blaschke unwinding adaptive Fourier decomposition based signal compression algorithm with application on ECG signals," *IEEE Journal of Biomedical and Health Informatics*, vol. PP, 2018.

[5] G. You, M. Ku, and T. Qian, "Fast algorithm of adaptive Fourier series," *Math. Method. Appl. Sci.*, pp. 2654–2663, 11 2017.

[6] Z. Wang, F. Wan, C. M. Wong, and T. Qian, "Fast basis search for adaptive Fourier decomposition," *EURASIP Journal on Advances in Signal Processing*, vol. 2018, no. 74, p. 14, 2018.

[7] A. Borowicz, "GPU implementation of adaptive Fourier decomposition," in *2019 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA)*, Sep. 2019, pp. 17–22.

[8] T. Qian, "Adaptive Fourier decompositions and rational approximations, part I: Theory," *Int. J. Wavelets Multi.*, vol. 12, no. 05, p. 1461008, 2014.

[9] T. Qian, J. Wang, and W. Mai, "An enhancement algorithm for cyclic adaptive Fourier decomposition," *Applied and Computational Harmonic Analysis*, vol. 47, no. 2, pp. 516–525, 2019.

[10] L. Cohen, *Time-frequency analysis*. Prentice Hall PTR Englewood Cliffs, N.J, 1995.

[11] J. Garofolo, L. Lamel, W. Fisher, J. Fiscus, D. Pallett, and N. Dahlgren, "DARPA TIMIT Acoustic-Phonetic Continuous Speech Corpus," National Institute of Standards and Technology (NIST), CD, 1993.

[12] K. Gopalan, "Speech coding using Fourier-Bessel expansion of speech signals," in *IECON'01. 27th Annual Conference of the IEEE Industrial Electronics Society*, vol. 3, Nov 2001, pp. 2199–2203.

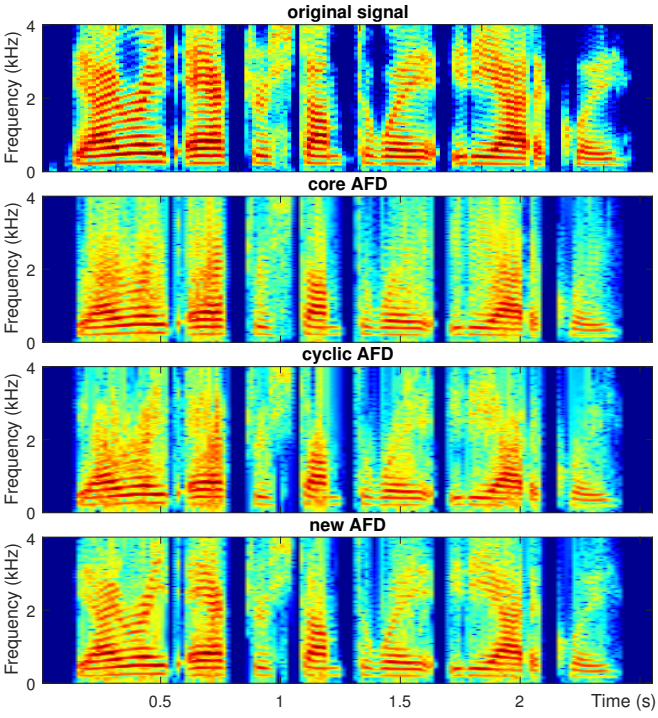


Fig. 3: Spectrograms of the original speech fragment (top) and the signals reconstructed by various AFD algorithms using $P = 16$ coefficients of the largest energy.