A Generalized Weighted Overlap-Add (WOLA) Filter Bank for Improved Subband System Identification

Mohit Sharma, Robbe Van Rompaey, Wouter Lanneer, and Marc Moonen, *Fellow, IEEE*

***Abstract*—This paper addresses the challenges in short-time Fourier transform (STFT) domain subband adaptive filtering, in particular, subband system identification. Previous studies in this area have primarily focused on setups with subband filtering at a downsampled rate, implemented using the weighted overlap-add (WOLA) filter bank, popular in audio and speech-processing for its reduced complexity. However, this traditional approach imposes constraints on the subband filters when transformed to their full-rate representation. This paper makes three key contributions. First, it introduces a generalized WOLA filter bank that repositions subband filters before the downsampling operation, eliminating the constraints on subband filters inherent in the conventional WOLA filter bank. Second, it investigates the mean square error (MSE) performance of the generalized WOLA filter bank for full-band system identification, establishing analytical ties between the order of subband filters, the full-band system impulse response length, the decimation factor, and the prototype filters. Third, to address the increased computational complexity of the generalized WOLA, the paper proposes a low-complexity implementation termed per-tone weighted overlap-add (PT-WOLA), which maintains computational complexity on par with conventional WOLA. Analytical and empirical evidence demonstrates that the proposed generalized WOLA filter bank significantly enhances the performance of subband system identification.**

***Index Terms*—Weighted overlap-add (WOLA), subband adaptive filtering, crossband filtering, system identification, filter banks, short-time Fourier transform (STFT), acoustic echo cancellation.**

Received 14 January 2025; revised 22 July 2025; accepted 23 September 2025. Date of publication 29 September 2025; date of current version 10 November 2025. This work was supported in part by ESAT Laboratory of KU Leuven, in the frame of FWO Research Project G0A9921N "Per-Tone Equalization and Per-Tone Precoding for Speech and Audio Signal Processing", in part by the Research Council KU Leuven Project C14-21-0075 "A holistic approach to the design of integrated and distributed digital signal processing algorithms for audio and speech communication devices", and in part by VLAIO Project HBC.2022.0428 "SPARTA- Signal Processing for next-generation opticAl netwoRks and remoTe Audio experiences". The associate editor coordinating the review of this article and approving it for publication was Behtash Babadi. *(Corresponding author: Mohit Sharma.)*

Mohit Sharma and Marc Moonen are with the Department of Electrical Engineering (ESAT), KU Leuven, 3000 Leuven, Belgium (e-mail: mohitsharma.nith@gmail.com).

Robbe Van Rompaey and Wouter Lanneer are with Nokia Bell Labs, 2018 Antwerp, Belgium.

Digital Object Identifier 10.1109/TSP.2025.3614849

I. Introduction

ADAPTIVE filtering problems appear in various signal processing fields. Particularly in system identification problems when dealing with systems with long impulse responses, such as in audio and speech signal processing applications, adaptive filtering is performed in the short-time Fourier transform (STFT)-domain to reduce computational complexity [1], [2]. STFT-domain subband adaptive filtering typically employs a uniform discrete Fourier transform (DFT) modulated filter bank for efficient implementation.

Subband adaptive filtering utilizing a uniform DFT modulated filter bank with specific analysis and synthesis filters along with constrained subband processing, can be represented equivalently as a full-band time-domain adaptive filtering operation. Such subband adaptive filtering is typically labeled as frequency-domain adaptive filtering (FDAF) [2], [3], [4].

General oversampled uniform DFT modulated filter bank-based subband adaptive filtering, with zero-order analysis and synthesis filter polyphase components, is commonly referred to as weighted overlap-add (WOLA) based subband adaptive filtering. The term 'zero-order' implies that the prototype filter length does not exceed the number of subbands N, resulting in scalar polyphase components [5]. Fig. 1 illustrates the state-of-the-art setup for WOLA based subband system identification, where $W(z)$ is the system to be identified, with input signal $U(z)$ and noise $\eta(z)$. Moreover, it incorporates analysis filter $H_k(z)$, synthesis filter $F_k(z)$ and subband filter $C_k(z)$ associated with k^{th} subband. When the subband filters are one-tap filters, the setup represents conventional WOLA based subband processing. With higher-order subband filters, the setup represents WOLA based subband processing with inter-frame filtering.

WOLA based subband adaptive filtering offers reduced computational complexity by closely approximating the long time-domain impulse response of the system with a lower-order subband filter in each subband, and by performing the subband adaptive filtering at a downsampled rate. Hence, it is widely popular in audio and speech signal processing applications [6]. However, this reduction in computational complexity comes at the cost of undesirable aliasing and distortion effects. Consequently, WOLA based subband adaptive filtering does not correspond to a full-band time-domain adaptive filtering operation, which limits its overall modeling accuracy.

1053-587X © 2025 IEEE. All rights reserved, including rights for text and data mining, and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: KU Leuven University Library. Downloaded on November 19,2025 at 14:23:26 UTC from IEEE Xplore. Restrictions apply.

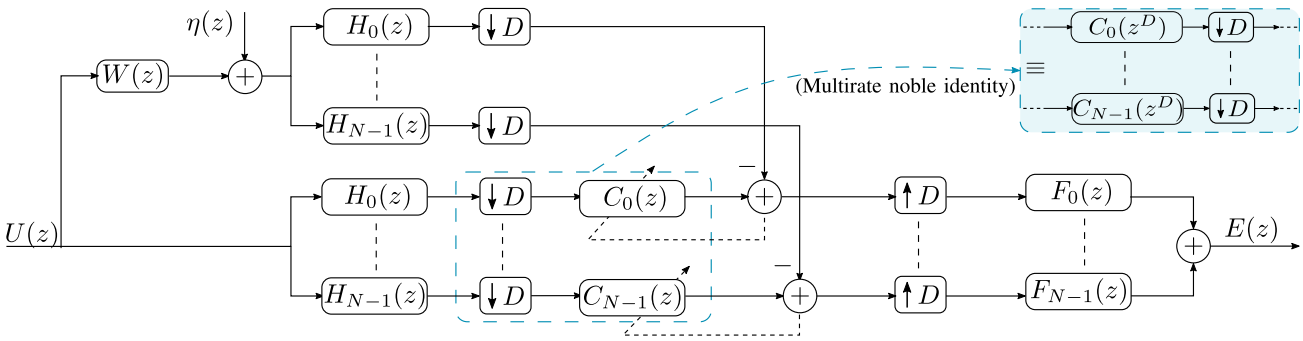


Fig. 1. Weighted overlap-add (WOLA) filter bank based subband system identification.

Various techniques have been proposed to improve the modeling accuracy of WOLA based subband adaptive filtering. These techniques can be broadly classified into two categories: those aimed at reducing aliasing effects originating from the downsampling and non-ideal prototype filters in the analysis filter bank, and those aimed at reducing cyclic distortion effects originating from STFT-domain processing of full-band filtering.

To reduce the impact of aliasing effects, adaptive cross-band filters have been proposed in [7] and analyzed in [8], [9]. In [10], a post-processing stage has been proposed that generates additional subband filter coefficients for time-frequency interpolated reference signals. Additionally, multiple prototype filter design techniques have been proposed in [11], [12], [13], [14], [15], [16], [17], [18], [19] to minimize the inter-subband aliasing, thereby reducing the aliasing effects.

To mitigate cyclic distortion effects, constrained subband processing, similar to FDAF, has been proposed in [2]. However, the constraints introduced by this method add complexity and limit "the degrees of freedom", restricting the means of improvement in filter bank design. In [20], [21], the relation between full-band and STFT-domain subband adaptive filtering has been analyzed, motivating the use of full-band errors for adaptation, similar to a closed-loop structure. However, this method introduces additional delay to the adaptation path, leading to slow and unreliable convergence, as noted in [4]. In [22], a novel scheme for WOLA based subband adaptive filtering has been proposed. The proposed adaptive filter structure in [22] employs an additional set of adaptive subband filters and leverages the relation between linear convolution and the root-Hann prototype analysis filter. However, this approach introduces additional computational complexity due to the additional set of subband filters and limits the filter bank design to the root-Hann prototype analysis filter.

To the best of our knowledge, no prior work has specifically addressed the impact of the prototype synthesis filter design on cyclic distortions. Additionally, there is a notable gap in research concerning the theoretical performance analysis of WOLA filter banks from a time-domain perspective. Previous studies on WOLA filter bank-based adaptive filtering have predominantly focused on subband filter adaptation at a downsampled rate, as illustrated in Fig. 1. By applying the so-called multirate noble identities [5] to move the subband filters before the downsampling operation, it is seen that, at full rate, the subband filters are constrained to the $C_k(z^D)$ structure. This structure imposes $D-1$ zeroes between subsequent filter coefficients, thus restricting the conventional WOLA filter bank to inter-frame filtering. The predominant focus on subband adaptive filtering at a downsampled rate in existing research can be attributed to the practical challenges associated with the computational complexity of performing subband adaptive filtering at full rate. Although it is theoretically feasible to position the subband filters before the downsampling operation, as shown in Fig. 2, the high computational demands have rendered such a system impractical for real-world applications.

Nevertheless, this paper aims to analyze a generalized WOLA filter bank with full rate subband adaptive filtering, as shown in Fig. 2, and analyze the impact of various factors on the overall system performance from a time-domain perspective. The paper also aims at seeking a low-complexity implementation of such a system. The paper makes three key contributions:

First, the paper introduces a generalized WOLA filter bank for subband adaptive filtering, which is motivated by the per-tone equalizer (PTEQ) structure in asymmetric digital subscriber line (ADSL) systems [23], [24] and fundamentally removes the primary limitation of conventional WOLA. As shown in Fig. 2, by repositioning the subband filters to operate at the full rate before the downsampling operation, the filters are no longer constrained to the sparse $C_k(z^D)$ form, but become a general, unconstrained $C_k(z)$ filter, providing the necessary degrees of freedom for the accurate system modeling that was previously not possible[1]. The paper further investigates the application of the generalized WOLA filter bank in subband system identification, with a particular focus on its time-domain representation, addressing the limitations of conventional WOLA filter banks in subband system identification.

Second, the paper examines the mean square error (MSE) performance of the generalized WOLA filter bank in full-band system identification and establishes analytical ties between various factors, such as the order of subband filters, the length of the full-band system impulse response, the decimation factor, and the design of prototype analysis and synthesis filters.

[1]The conventional WOLA with inter-frame filtering in Fig. 1 can be represented by the proposed generalized WOLA, by constraining its subband filters $C_k(z)$ in Fig. 2 to follow $C_k(z^D)$.

Authorized licensed use limited to: KU Leuven University Library. Downloaded on November 19,2025 at 14:23:26 UTC from IEEE Xplore. Restrictions apply.

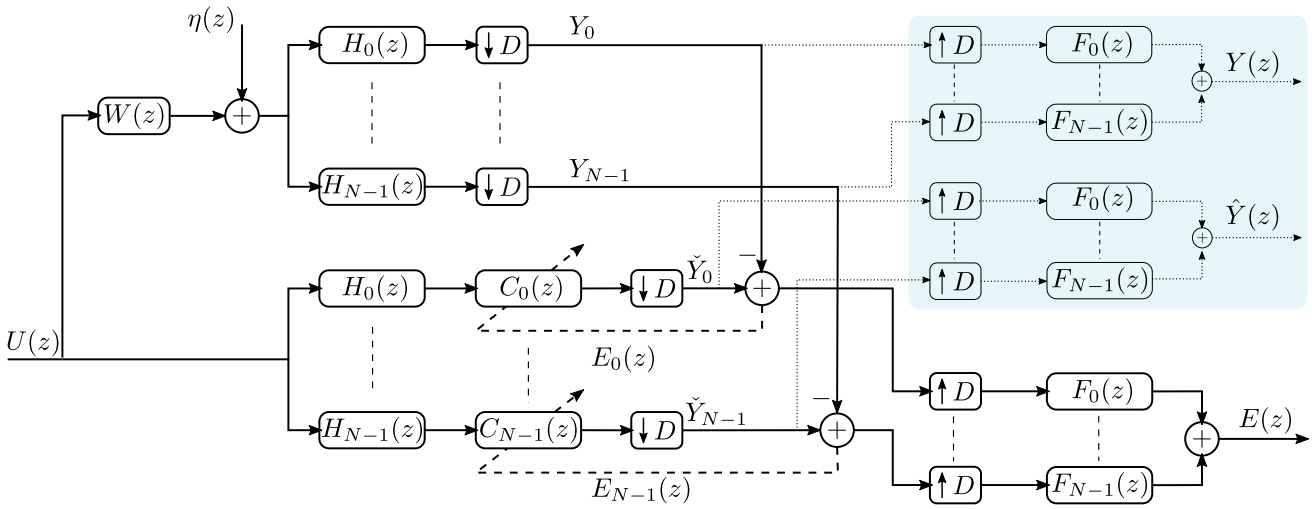


Fig. 2. Generalized weighted overlap-add (WOLA) filter bank based subband system identification.

Third, since the proposed approach leads to a substantial increase in computational complexity, to address this, the paper presents an alternative low-complexity implementation of the generalized WOLA filter bank, referred to as per-tone weighted overlap-add (PT-WOLA), inspired by the term PTEQ from [23]. The PT-WOLA maintains a computational complexity comparable to that of conventional WOLA, while demonstrating performance that is comparable to, or slightly exceeds, that of the generalized WOLA filter bank for the same subband filter order.

The structure of the paper is as follows: Section II presents the system model for generalized WOLA filter bank-based subband adaptive filtering, also establishing a time-domain representation of the subband filtering. Section III presents the steady-state solution for the generalized WOLA filter bank based subband system identification, while Section IV analyzes its the steady-state MSE performance. Section V discusses the computational complexity of the generalized WOLA filter bank and proposes a computationally efficient implementation, termed PT-WOLA. Section VI presents simulation results for acoustic echo cancellation (AEC) as a use-case scenario. Finally, Section VII concludes the paper.

II. Generalized WOLA and Time-Domain Representation of Subband Filtering

The subband system identification problem comprises two primary components: subband filtering and the adaptation of subband filter coefficients that minimize the subband errors[2]. This section focuses on the subband filtering aspect and derives a relation between the subband filtering and the overall time-domain transfer function of the generalized WOLA filter bank, shown in Fig. 2, which will be utilized in subsequent sections.

[2]In literature, the so-called open-loop subband system identification is widely used due to its faster convergence, which minimizes the subband errors, in contrast to the closed-loop subband system identification, wherein the full-band error after the synthesis filter bank is minimized [4], [25].

The relation between the input signal $u(n)$ and output signal $y(n)$ for an unknown linear time-invariant (LTI) system $w(n)$, ignoring noise, is first given as

$$y(n) = u(n) \star w(n) \tag{1}$$

where $\star$ represents the convolution operation, which can equivalently be expressed in the z-domain as

$$\underbrace{\sum_{m=0}^{\infty} y(m)z^{-m}}_{Y(z)} = \underbrace{\sum_{p=0}^{\infty} u(p)z^{-p}}_{U(z)} \underbrace{\sum_{q=0}^{L-1} w(q)z^{-q}}_{W(z)} \tag{2}$$

where, $Y(z)$, $U(z)$ and $W(z)$ represent the z-transform of $y(n)$, $u(n)$ and $w(n)$, respectively, and L denotes the finite impulse response (FIR) length of the unknown LTI system. The unknown LTI system, in this case, acts on the full-band signal $U(z)$ as a whole. Therefore, if L is large, accurately approximating the full-band LTI system may necessitate a high filter order.

In the generalized WOLA filter bank in Fig. 2, however, $Y(z)$ represents the input signal $U(z)$ convolved with the unknown LTI system transfer function $W(z)$ and additionally passed through the WOLA filter bank, which is assumed to satisfy the perfect reconstruction property and hence only introduces additional delay. $\hat{Y}(z)$ represents the estimate of $Y(z)$ using a subband filter $C_k(z)$ in the k^{th} subband $(k = 0 \cdots N-1)$ and $E_k(z)$ represents the k^{th} subband error. Furthermore, the generalized WOLA filter bank in Fig. 2 consists of analysis filters $H_k(z)$, synthesis filters $F_k(z)$, and a D-fold decimation and interpolation $(D \leq N)$. The relationship between the input and output of the filter bank is characterized by the so-called overall transfer function of the filter bank, which includes a distortion function $T(z)$ and alias transfer functions $A_m(z)$, i.e.

$$Y(z) = \underbrace{\left\{ \frac{1}{D} \sum_{k=0}^{N-1} H_k(z)F_k(z) \right\}}_{T(z)} U(z)W(z)$$

Authorized licensed use limited to: KU Leuven University Library. Downloaded on November 19,2025 at 14:23:26 UTC from IEEE Xplore. Restrictions apply.

$$+\sum_{m=1}^{D-1}\left\{\underbrace{\frac{1}{D}\left(\sum_{k=0}^{N-1}H_k(z\alpha_D^{-m})F_k(z)\right)}_{A_m(z)}U(z\alpha_D^{-m})W(z\alpha_D^{-m})\right\} \tag{3}$$

where $\alpha_D \triangleq \exp^{j2\pi/D}$.

The analysis and synthesis filters are designed to ensure perfect reconstruction, i.e.

$$T(z) = z^{-\tau}\ , \tag{4a}$$
$$A_m(z) = 0,\ \ \forall m = 1\cdots D-1. \tag{4b}$$

Substituting (4) in (3), the $Y(z)$ in Fig. 2 is indeed given as

$$Y(z) = U(z)W(z)z^{-\tau}. \tag{5}$$

Hence, it can be observed that (5) is equivalent to (2), corrected with the (unavoidable) delay represented by τ, i.e.

$$y(n) = u(n-\tau)\star w(n). \tag{6}$$

Similarly, for $\hat{Y}(z)$ in Fig. 2, the input-output relation is given as

$$\hat{Y}(z) = \underbrace{\left\{\frac{1}{D}\sum_{k=0}^{N-1}H_k(z)F_k(z)C_k(z)\right\}}_{\hat{T}(z)}U(z) + \sum_{m=1}^{D-1}\underbrace{\left\{\frac{1}{D}\sum_{k=0}^{N-1}H_k(z\alpha_D^{-m})F_k(z)C_k(z\alpha_D^{-m})\right\}}_{\hat{A}_m(z)}U(z\alpha_D^{-m}) \tag{7}$$

where, $\hat{T}(z)$ and $\hat{A}_m(z)$ are the so-called distortion function and alias transfer functions of the filter bank with subband processing, respectively. Moreover, $C_k(z)$ represents the k^{th} subband FIR filter with length T, i.e.

$$C_k(z) = \sum_{q=0}^{T-1}c_k(q)z^{-q} \tag{8}$$

where $c_k(q)$ is the q^{th} filter coefficient of the k^{th} subband filter and $\mathbf{c}_k \triangleq \left[c_k(0),\cdots,c_k(T-1)\right]$. The time-domain equivalent of (7) is then given as

$$\hat{y}(n) = u(n)\star\hat{t}(n) + \sum_{m=1}^{D-1}\left(u(n)\star\hat{a}_m(n)\alpha_D^{nm}\right) \tag{9}$$

The following subsections provide a separate analysis for the distortion function and the aliasing transfer functions.

A. Distortion Function

In a uniform DFT modulated filter bank, the k^{th} analysis and synthesis filter are derived from a prototype analysis filter $\mathbf{h}_0 \triangleq \left[h_0(0),\cdots,h_0(N-1)\right]$ and a prototype synthesis filter $\mathbf{f}_0 \triangleq \left[f_0(0),\cdots,f_0(N-1)\right]$, respectively, i.e.[3]

$$h_k(n) = h_0(n)\alpha_N^{nk} \quad \text{and} \quad f_k(n) = f_0(n)\alpha_N^{nk}\alpha_N^{k}$$
$$\Leftrightarrow$$
$$H_k(z) = H_0(z\alpha_N^{-k}) \quad \text{and} \quad F_k(z) = F_0(z\alpha_N^{-k})\alpha_N^{k}. \tag{10}$$

Therefore the distortion function in (7) can be expressed as

$$\hat{T}(z) = \frac{1}{D}\sum_{k=0}^{N-1}H_0(z\alpha_N^{-k})F_0(z\alpha_N^{-k})\alpha_N^{k}C_k(z). \tag{11}$$

The corresponding time-domain filter coefficients $\hat{\mathbf{t}} \triangleq \left[\hat{t}(0),\ \cdots,\ \hat{t}(2N+T-3)\right]$ are then given as

$$\hat{t}(n) = \frac{1}{D}\sum_{k=0}^{N-1}\left[\left(h_0(n)\alpha_N^{nk}\right)\star\left(f_0(n)\alpha_N^{nk}\alpha_N^{k}\right)\star c_k(n)\right] \tag{12}$$

$$= \frac{1}{D}\sum_{k=0}^{N-1}\sum_{l=0}^{T-1}\sum_{m=0}^{N-1}h_0(n-m-l)\,\alpha_N^{k(n-m-l)}f_0(m)\alpha_N^{km}\alpha_N^{k}c_k(l)$$
$$= \sum_{k=0}^{N-1}\sum_{l=0}^{T-1}\underbrace{\left\{\frac{1}{D}\sum_{m=0}^{N-1}h_0(n-m-l)f_0(m)\right\}}_{g_0(n-l)}\alpha_N^{k(n-l+1)}c_k(l) \tag{13}$$

where $g_0(n)$ represents the scaled convolution of the prototype analysis filter $h_0(n)$ and the prototype synthesis filter $f_0(n)$, of length $2N-1$, i.e.

$$g_0(n) = \frac{1}{D}\sum_{m=0}^{N-1}h_0(n-m)f_0(m). \tag{14}$$

Alternatively, (13) can be written as

$$\hat{t}(n) = \sum_{l=0}^{T-1}g_0(n-l)\underbrace{\left\{\sum_{k=0}^{N-1}\alpha_N^{k(n-l+1)}c_k(l)\right\}}_{\tilde{c}^l(n-l)}$$
$$= \sum_{l=0}^{T-1}g_0(n-l)\tilde{c}^l(n-l) \tag{15}$$

where $\tilde{c}^l(n)$ are time-domain subband filter coefficients, obtained as the (circularly shifted) N-point inverse discrete Fourier transform (IDFT) of the collection of l^{th} filter coefficients for all subband filters, i.e.

$$\tilde{c}^l(n) \triangleq \sum_{k=0}^{N-1}c_k(l)\alpha_N^{k(n+1)} \tag{16}$$

which is periodic with period N: $\tilde{c}^l(n+N) = \tilde{c}^l(n) \quad \forall l$.

The result in (15) can also be written as

$$\hat{\mathbf{t}} = \sum_{l=0}^{T-1}\left\{\begin{bmatrix}\mathbf{0}_{1,l} & \mathbf{g}_0 & \mathbf{0}_{1,T-l-1}\end{bmatrix}\odot\begin{bmatrix}\mathbf{0}_{1,l} & \tilde{\mathbf{c}}^l & \mathbf{0}_{1,T-l-1}\end{bmatrix}\right\} \tag{17}$$

[3]The scaling factor α_N^k in (10) for the synthesis filters is introduced primarily for the sake of mathematical convenience [5].

Authorized licensed use limited to: KU Leuven University Library. Downloaded on November 19,2025 at 14:23:26 UTC from IEEE Xplore. Restrictions apply.

where, the symbol $\odot$ denotes the Hadamard (element-wise) product, $\mathbf{g}_0 \triangleq \left[g_0(0),\ g_0(1),\ \cdots,\ g_0(2N-2)\right]$ and $\tilde{\mathbf{c}}^l \triangleq \left[\tilde{c}^l(0),\ \cdots,\ \tilde{c}^l(N-2),\ \tilde{c}^l(0),\ \cdots,\ \tilde{c}^l(N-1)\right]$, are vectors of length $2N-1$. The two periodic images of $\tilde{c}^l(n)$ in $\tilde{\mathbf{c}}^l$, represent the circular convolution effect in the distortion function $\hat{\mathbf{t}}$. It is noted that conventional WOLA has $T=0$, and hence the summation in (17) is reduced to one term.

B. Aliasing Transfer Functions

Similar to (12), each alias transfer function $\hat{A}_m(z)$, $(m=1\cdots D-1)$, in (7) corresponds to time-domain filter coefficients given as

$$\hat{a}_m(n) = \frac{1}{D}\sum_{k=0}^{N-1}\left(h_0(n)\alpha_N^{nk}\alpha_D^{nm}\right) \star \left(f_0(n)\alpha_N^{nk}\alpha_N^{k}\right) \star \left(c_k(n)\alpha_D^{nm}\right) \tag{18}$$

Similar to the derivation in Section II-A, it can be shown that the expression in (18) can be simplified to (19), shown at the bottom of the page, and

$$\hat{\mathbf{a}}_m = \sum_{l=0}^{T-1}\alpha_D^{lm}\left[\mathbf{0}_{1,l}\quad \boldsymbol{\psi}_m\quad \mathbf{0}_{1,T-l-1}\right]\odot\left[\mathbf{0}_{1,l}\quad \tilde{\mathbf{c}}^l\quad \mathbf{0}_{1,T-l-1}\right] \tag{20}$$

where, $\hat{\mathbf{a}}_m \triangleq \left[\hat{a}_m(0),\ \hat{a}_m(1),\ \cdots,\ \hat{a}_m(2N-2)\right]$ and $\psi_m(n)$ represents the scaled modulated convolution of the prototype analysis and synthesis filter, of length $2N-1$, i.e.

$$\psi_m(n) = \frac{1}{D}\sum_{p=0}^{N-1} h_0(p)f_0(n-p)\alpha_D^{pm}. \tag{21}$$

and finally $\boldsymbol{\psi}_m \triangleq \left[\psi_m(0),\ \psi_m(1),\ \cdots,\ \psi_m(2N-2)\right]$. It is noted again that conventional WOLA has $T=0$, and hence the summation in (20) is reduced to one term.

III. Steady-State Solution

This section derives the steady-state MSE solution for the generalized WOLA filter bank based subband system identification.

For each subband, the subband filter coefficients are estimated independently by minimizing the following MSE cost function

$$\underset{\mathbf{c}_k}{\text{minimize}}\quad \mathbb{E}\left\{\left|\check{Y}_k(q)-Y_k(q)\right|^2\right\} \tag{22}$$

It can be verified that for a white stationary input signal, the steady-state solution for (22) corresponds to the least-squares (LS) solution of

$$H_k(z)C_k(z) \overset{\text{LS}}{=} H_k(z)W(z) \tag{23}$$

For a uniform DFT modulated filter bank, the corresponding time-domain relation can be written as

$$\left(h_0(n)\alpha_N^{nk}\right)\star c_k(n) \overset{\text{LS}}{=} \left(h_0(n)\alpha_N^{nk}\right)\star w(n) \tag{24}$$

Defining, $\mathbf{w} \triangleq \left[w(0)\quad w(1)\quad \cdots\quad w(L-1)\quad \mathbf{0}_{N-L,1}\right]$ as the column vector of length N representing the impulse response of the unknown LTI system, (23) can be expressed in matrix form as (25), shown at the bottom of the page, where $\mathbf{H}_c$ and $\mathbf{H}_w$ are a Toeplitz matrix of size $(2N-1)\times T$ and $(2N-1)\times N$ respectively. The LS solution for (25) can then be written as

$$\operatorname{diag}\left(\mathbb{F}_N(k,0:T-1)\right)\mathbf{c}_k^{\mathsf{T}} = \underbrace{\left(\mathbf{H}_c\right)^{\ddagger}\mathbf{H}_w}_{\mathbf{Z}_0}\operatorname{diag}\left(\mathbb{F}_N(k,:)\right)\mathbf{w}^{\mathsf{T}} \tag{26}$$

where $\mathbb{F}_N$ is the $N\times N$ DFT-matrix and $(\mathbf{H}_c)^{\ddagger} \triangleq \left((\mathbf{H}_c)^{\mathsf{T}}\mathbf{H}_c\right)^{-1}(\mathbf{H}_c)^{\mathsf{T}}$ represents the left Moore–Penrose inverse (pseudoinverse) of $\mathbf{H}_c$. Therefore, $\mathbf{Z}_0$ is a matrix of size $T\times N$.

$$\begin{aligned}\hat{a}_m(n) &= \sum_{l=0}^{T-1}\left(\underbrace{\frac{1}{D}\sum_p h_0(p)f_0(n-l-p)\alpha_D^{(p+l)m}}_{\psi_m(n-l)\alpha_D^{lm}}\underbrace{\sum_{k=0}^{N-1}\alpha_N^{(n+1)k}c_k(l)\alpha_N^{-lk}}_{\tilde{c}^l(n-l)}\right)\\ &= \sum_{l=0}^{T-1}\alpha_D^{lm}\ \psi_m(n-l)\ \tilde{c}^l(n-l)\end{aligned} \tag{19}$$

$$\underbrace{\left[\begin{array}{cccc} h_0(0) & 0 & \cdots & 0\\ h_0(1) & h_0(0) & \cdots & 0\\ \vdots & \ddots & \ddots & \vdots\\ 0 & 0 & \cdots & h_0(N-1)\\ \hline & & \mathbf{0}_{N-T,T} & \end{array}\right]}_{\mathbf{H}_c}\operatorname{diag}\left(\mathbb{F}_N(k,0:T-1)\right)\mathbf{c}_k^{\mathsf{T}} \overset{\text{LS}}{=} \underbrace{\left[\begin{array}{cccc} h_0(0) & 0 & \cdots & 0\\ h_0(1) & h_0(0) & \cdots & 0\\ \vdots & \ddots & \ddots & \vdots\\ 0 & 0 & \cdots & h_0(N-1)\end{array}\right]}_{\mathbf{H}_w}\operatorname{diag}\left(\mathbb{F}_N(k,:)\right)\mathbf{w}^{\mathsf{T}} \tag{25}$$

Authorized licensed use limited to: KU Leuven University Library. Downloaded on November 19,2025 at 14:23:26 UTC from IEEE Xplore. Restrictions apply.

The l^{th} filter coefficient of the k^{th} subband filter ($\mathbf{c}_k$) in (26), is then given as

$$\mathbb{F}_N(k,l)c_k(l) = \mathbf{Z}_0(l,:)\text{diag}\left(\mathbb{F}_N(k,:)\right)\mathbf{w}^\intercal \tag{27}$$

Stacking the l^{th} filter coefficients of all the subband filters, then leads to

$$\text{diag}\left(\mathbb{F}_N(l,:)\right)\begin{bmatrix} c_0(l) \\ \vdots \\ c_{N-1}(l)\end{bmatrix} = \begin{bmatrix}\mathbf{Z}_0(l,:) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \mathbf{Z}_0(l,:)\end{bmatrix} \begin{bmatrix}\text{diag}\left(\mathbb{F}_N(0,:)\right) \\ \vdots \\ \text{diag}\left(\mathbb{F}_N(N-1,:)\right)\end{bmatrix}\mathbf{w}^\intercal \tag{28}$$

which can be simplified as

$$\text{diag}\left(\mathbb{F}_N(l,:)\right)\begin{bmatrix} c_0(l) \\ \vdots \\ c_{N-1}(l)\end{bmatrix} = \mathbb{F}_N\text{diag}\left(\mathbf{Z}_0(l,:)\right)\mathbf{w}^\intercal \tag{29}$$

From (16), it follows that $\begin{bmatrix}c_0(l) & \cdots & c_{N-1}(l)\end{bmatrix}^\intercal = \mathbb{F}_N\left(\tilde{\tilde{\mathbf{c}}}^l\right)^\intercal$ with $\tilde{\tilde{\mathbf{c}}}^l \triangleq \begin{bmatrix}\tilde{c}^l(0) & \cdots & \tilde{c}^l(N-1)\end{bmatrix}$, and hence (29) becomes

$$\text{diag}\left(\mathbb{F}_N(l,:)\right)\mathbb{F}_N\left(\tilde{\tilde{\mathbf{c}}}^l\right)^\intercal = \mathbb{F}_N\text{diag}\left(\mathbf{Z}_0(l,:)\right)\mathbf{w}^\intercal. \tag{30}$$

Finally, the time-domain representation of subband filters appearing in (17) and (20), is then given as $\tilde{\mathbf{c}}^l = [\tilde{\tilde{\mathbf{c}}}^l \mid [\tilde{\tilde{\mathbf{c}}}^l(0:N-2)]$ and

$$\left(\tilde{\tilde{\mathbf{c}}}^l\right)^\intercal = \underbrace{\left(\text{diag}\left(\mathbb{F}_N(l,:)\right)\mathbb{F}_N\right)^{-1}\mathbb{F}_N}_{\mathcal{I}_N^l}\text{diag}\left(\mathbf{Z}_0(l,:)\right)\mathbf{w}^\intercal \tag{31}$$

where $\mathcal{I}_N^l$ is the l-point circularly shifted identity matrix of size N, with the element in the i^{th} row and j^{th} column defined as

$$\mathcal{I}_N^l(i,j) = \delta_{((i+l) \bmod N),j} \tag{32}$$

and $\delta_{a,b}$ denotes the Kronecker delta, which equals 1 if $a=b$ and 0 otherwise.

The matrix $\mathbf{Z}_0$ depends only on the prototype analysis filter $\mathbf{h}_0$, presented in $\mathbf{H}_w$ and $\mathbf{H}_c$, as can be seen from (25) and (26). Moreover, from (25) it can be seen that $\mathbf{H}_c$ is a submatrix of $\mathbf{H}_w$, such that $\mathbf{H}_c = \mathbf{H}_w(:,0:T-1)$. Therefore, $\mathbf{H}_w$ can be written as

$$\mathbf{H}_w = \begin{bmatrix}\mathbf{H}_c & \mathbf{H}_w(:,T:N-1)\end{bmatrix} \tag{33}$$

Substituting (33) in $\mathbf{Z}_0$ as defined in (26) leads to

$$\mathbf{Z}_0 = \left(\left(\mathbf{H}_c\right)^\intercal\mathbf{H}_c\right)^{-1}\left(\mathbf{H}_c\right)^\intercal\begin{bmatrix}\mathbf{H}_c & \mathbf{H}_w(:,T:N-1)\end{bmatrix} \tag{34}$$

Therefore, it can be observed that $\mathbf{Z}_0$ has the following structure

$$\mathbf{Z}_0 = [\mathbf{I}_T \mid \boldsymbol{\xi}_0] \tag{35}$$

where, $\boldsymbol{\xi}_0 \triangleq \left(\left(\mathbf{H}_c\right)^\intercal\mathbf{H}_c\right)^{-1}\left(\mathbf{H}_c\right)^\intercal\mathbf{H}_w(:,T:N-1)$ is a (non-sparse) matrix of size $(T \times N-T)$, only dependent on the prototype analysis filter.

It is noted that the last $N-L$ components of $\mathbf{w}$ are zero, hence the last $N-L$ columns of $\mathbf{Z}_0$ do not play a role.

It can be observed from (35) that as the subband filter order T increases, the non-zero part of the $\mathbf{Z}_0$ matrix approaches an identity matrix. For $L=T$, $\tilde{\mathbf{c}}^l = \begin{bmatrix}w(l) & \mathbf{0}_{N-1,1}\end{bmatrix}$. Furthermore, from (31) and (35), the maximum length of $\tilde{\mathbf{c}}^l$ is $L-T+1$, as will also be exemplified with simulations in Section VI (Fig. 7).

IV. Steady-State MSE Performance of Subband System Identification

This section addresses the minimum mean square error (MMSE)-based subband system identification problem, deriving the overall steady-state MSE performance of the generalized WOLA filter bank. The analysis is framed within the context of echo return loss enhancement (ERLE), which is commonly employed in AEC applications, but also serves as a more broadly applicable metric, and examines the impact of the subband filter order, the FIR length of the unknown LTI system, and the prototype analysis and synthesis filters on the system's overall performance.

The normalized MSE in AEC is conventionally represented by the ERLE, which provides a measure of the achieved echo suppression and is inversely related to the normalized MSE. The ERLE for the considered system is given as

$$\text{ERLE} = \frac{\mathbb{E}\left\{|y(n)|^2\right\}}{\mathbb{E}\left\{|y(n)-\hat{y}(n)|^2\right\}} \propto \frac{1}{\text{MSE}} \tag{36}$$

where $\mathbb{E}\{\cdot\}$ represents the expected value operation. Substituting (6) and (9) into (36), and for simplification, assuming the input signal is a zero mean white signal with variance σ_u^2, the ERLE expression can be rewritten as a function of the impulse response of the unknown LTI system, the distortion function and aliasing transfer functions, i.e.

$$\text{ERLE} = \frac{\|\mathbf{w}_{ext}\|^2\sigma_u^2}{\left(\left\|\hat{\mathbf{t}}-\mathbf{w}_{ext}\right\|^2 + \sum_{m=1}^{D-1}\|\hat{\mathbf{a}}_m\|^2\right)\sigma_u^2} \tag{37}$$

The $\mathbf{w}_{ext}$ represents the impulse response of the unknown LTI system, padded with zeros, i.e.

$$\mathbf{w}_{ext} = \begin{bmatrix}\mathbf{0}_{1,N-1} & \mathbf{w} & \mathbf{0}_{1,T-1}\end{bmatrix} \tag{38}$$

with $\mathbf{w} \triangleq \begin{bmatrix}w(0) & w(1) & \cdots & w(L-1) & \mathbf{0}_{1,N-L}\end{bmatrix}$ the row vector of length N representing the impulse response of the unknown LTI system. The zeros appended at the front account for the delay introduced by the subband filtering in the considered filter bank[4], while the zeros at the end ensure the extended impulse response matches with the length of the distortion function.

[4]The expression in (38) does not account for the delay of $\left\lfloor\frac{N}{D}-1\right\rfloor D$ samples associated with the buffer for block processing, represented by τ in (5). For a filter bank, this buffer-induced delay depends solely on the number of subbands N and the decimation (or interpolation) factor D, and is therefore not considered in this analysis. The delay in (38) pertains to the overall transfer function in (17), which corresponds to the suppression of the first periodic image of $\tilde{c}^l(n-l)$ to avoid circular convolution effects. Alternatively, the definition of $\mathbf{w}_{ext}$ can be adjusted to $\mathbf{w}_{ext} = \begin{bmatrix}\mathbf{w} & \mathbf{0}_{1,2N+T-L-2}\end{bmatrix}$, corresponding to the suppression of the second periodic image of $\tilde{c}^l(n-l)$. A detailed explanation has been provided by the authors in [26].

Authorized licensed use limited to: KU Leuven University Library. Downloaded on November 19,2025 at 14:23:26 UTC from IEEE Xplore. Restrictions apply.

Substituting (31) into (17) and (20), the expression for the overall distortion function and aliasing transfer functions can be expressed in terms of the unknown LTI impulse response, as given in (39), shown at the bottom of the page.

Finally, the analytical ERLE can be expressed as a function of the prototype filters and the system impulse response. Taking the ERLE formulation from (37), the matrix representations of $\hat{\mathbf{t}}$ and $\hat{\mathbf{a}}_m$ from (39a) and (39b), respectively, are substituted into the expression. The distortion error term $\left\|\hat{\mathbf{t}}-\mathbf{w}_{ext}\right\|^2$ becomes $\left\|\mathbf{G}_{\xi}^{\intercal}\mathbf{w}-\mathbf{w}_{ext}\right\|^2$, and the aliasing error $\sum_{m=1}^{D-1}\left\|\hat{\mathbf{a}}_m\right\|^2$ becomes $\sum_{m=1}^{D-1}\left\|\mathbf{\Psi}_{\boldsymbol{\xi}_m}^{\intercal}\mathbf{w}\right\|^2$. Inserting these into (37) yields the final expression for the analytical ERLE:

$$\mathrm{ERLE}=\frac{\left\|\mathbf{w}_{ext}\right\|^2\sigma_u^2}{\left(\left\|\mathbf{G}_{\xi}^{\intercal}\mathbf{w}-\mathbf{w}_{ext}\right\|^2+\sum_{m=1}^{D-1}\left\|\mathbf{\Psi}_{\boldsymbol{\xi}_m}^{\intercal}\mathbf{w}\right\|^2\right)\sigma_u^2}. \tag{40}$$

V. PT-WOLA: Computationally Efficient Generalized WOLA

The preceding sections have established the theoretical advantages and superior modeling potential of the generalized WOLA filter bank framework. However, despite its promise, the framework's high computational complexity limits its practical implementation, particularly as the subband filter order increases. This section addresses the critical trade-off between performance and computational efficiency by first introducing the PT-WOLA, a novel architecture that preserves the performance of the generalized WOLA while significantly reducing its computational complexity. A detailed analysis of computational complexity is then presented in Section V-A.

The large computational complexity of the generalized WOLA filter bank compared to the conventional WOLA filter bank, stems from two reasons. First, although the subband error for the subband adaptive filtering is computed at the downsampled rate (22), the subband filtering is performed at the full-rate, before the downsampling operation. In conventional WOLA, the subband filters are zeroth-order, and hence can be moved after the downsampling operation, allowing the subband filtering operation at the downsampled rate. This is not directly feasible with the higher-order subband filters in the generalized WOLA filter bank. Second, additional computational complexity stems from the computation of the input signals for the subband filters, as will be explained next.

In Fig. 2, the output of the k^{th} subband filter $C_k(z)$, after the downsampling operation, can be formulated as

$$\begin{aligned}\check{Y}_k(q)&=\left[u(n)\star\left(h_0(n)\alpha_N^{nk}\right)\star c_k(n)\right]_{\downarrow D}\\&=\sum_{l=0}^{T-1}c_k(l)\sum_{m=0}^{N-1}h_0(m)\alpha_N^{mk}u(qD-m-l)\end{aligned} \tag{41}$$

where $[\cdot]_{\downarrow D}$ represents the D-fold downsampling operation and q represents the time index at the downsampled rate (frame index). This can be expressed in matrix form as (42), shown at the bottom of the next page, where $\check{\mathbf{u}}_k[q]$ represents the input signal vector for the k^{th} subband filter.

Given that the vector $\left[\alpha_N^0,\cdots,\alpha_N^{(N-1)k}\right]$ represents the k^{th} row of the IDFT matrix, (42) can be reformulated as

$$\begin{aligned}\check{Y}_k(q)&=\check{\mathbf{u}}_k[q]\,\mathbf{c}_k^{\intercal}=\underbrace{\mathbb{F}_N^*(k,:)\mathrm{diag}\left(\mathbf{h}_0\right)}_{\triangleq\mathbb{F}_N^{win^*}(k,:)}\mathbf{U}[q]\,\mathbf{c}_k^{\intercal}\\&=\mathrm{row}_k\underbrace{\left(\mathbb{F}_N^{win^*}\mathbf{U}[q]\right)}_{T\text{-IDFT operations}}\mathbf{c}_k^{\intercal}\end{aligned} \tag{43}$$

where, $\mathbb{F}_N^{win^*}$ represents the windowed IDFT matrix $\mathbb{F}_N^{win^*}\triangleq\mathbb{F}_N^*\mathrm{diag}\left(\mathbf{h}_0\right)$. Therefore, from (43) it can be observed that as opposed to performing a single IDFT operation for each time frame (i.e. for each q) in conventional WOLA, $T-$IDFT operations seem to be necessary for each frame, i.e., an IDFT operation for every column in the matrix $\mathbf{U}[q]$.

$$\begin{aligned}\left(\hat{\mathbf{t}}\right)^{\intercal}&=\sum_{l=0}^{T-1}\left\{\mathrm{diag}\left(\begin{bmatrix}\mathbf{0}_{1\times l}&\mathbf{g}_0&\mathbf{0}_{1\times(T-l-1)}\end{bmatrix}\right)\begin{bmatrix}\mathbf{0}_{l\times 1}\\\hline\boldsymbol{\mathcal{I}}_N^l(0:N-2,:)\mathrm{diag}\left(\mathbf{Z}_0(l,:)\right)\mathbf{w}^{\intercal}\\\boldsymbol{\mathcal{I}}_N^l\mathrm{diag}\left(\mathbf{Z}_0(l,:)\right)\mathbf{w}^{\intercal}\\\hline\mathbf{0}_{(T-l-1)\times 1}\end{bmatrix}\right\}\\&=\underbrace{\sum_{l=0}^{T-1}\left\{\mathrm{diag}\left(\begin{bmatrix}\mathbf{0}_{1,l}&\mathbf{g}_0&\mathbf{0}_{1,T-l-1}\end{bmatrix}\right)\begin{bmatrix}\mathbf{0}_{l\times(2N-1)}\\\hline\mathbf{I}_{2N-1}\\\hline\mathbf{0}_{(T-l-1)\times(2N-1)}\end{bmatrix}\begin{bmatrix}\boldsymbol{\mathcal{I}}_N^l(0:N-2,:)\\\hline\boldsymbol{\mathcal{I}}_N^l\end{bmatrix}\mathrm{diag}(\mathbf{Z}_0(l,:))\right\}}_{\mathbf{G}_{\xi}}\mathbf{w}^{\intercal}\end{aligned} \tag{39a}$$

$$\begin{aligned}\left(\hat{\mathbf{a}}_m\right)^{\intercal}&=\sum_{l=0}^{T-1}\left\{\alpha_D^{lm}\mathrm{diag}\left(\begin{bmatrix}\mathbf{0}_{1\times l}&\boldsymbol{\psi}_m&\mathbf{0}_{1\times(T-l-1)}\end{bmatrix}\right)\begin{bmatrix}\mathbf{0}_{l\times 1}\\\hline\boldsymbol{\mathcal{I}}_N^l(0:N-2,:)\mathrm{diag}\left(\mathbf{Z}_0(l,:)\right)\mathbf{w}^{\intercal}\\\boldsymbol{\mathcal{I}}_N^l\mathrm{diag}\left(\mathbf{Z}_0(l,:)\right)\mathbf{w}^{\intercal}\\\hline\mathbf{0}_{(T-l-1)\times 1}\end{bmatrix}\right\}\\&=\underbrace{\sum_{l=0}^{T-1}\left\{\alpha_D^{lm}\mathrm{diag}\left(\begin{bmatrix}\mathbf{0}_{1\times l}&\boldsymbol{\psi}_m&\mathbf{0}_{1\times(T-l-1)}\end{bmatrix}\right)\begin{bmatrix}\mathbf{0}_{l\times(2N-1)}\\\hline\mathbf{I}_{2N-1}\\\hline\mathbf{0}_{(T-l-1)\times(2N-1)}\end{bmatrix}\begin{bmatrix}\boldsymbol{\mathcal{I}}_N^l(0:N-2,:)\\\hline\boldsymbol{\mathcal{I}}_N^l\end{bmatrix}\mathrm{diag}(\mathbf{Z}_0(l,:))\right\}}_{\mathbf{\Psi}_{\boldsymbol{\xi}_m}}\mathbf{w}^{\intercal}\end{aligned} \tag{39b}$$

Authorized licensed use limited to: KU Leuven University Library. Downloaded on November 19,2025 at 14:23:26 UTC from IEEE Xplore. Restrictions apply.

However, it is worth noting that the T-IDFT operations can be computed efficiently for any general cosine window used as a prototype analysis filter, including the rectangular window as a special case. This section primarily focuses on the extensively used (root-)raised cosine windows as the prototype analysis filter. Nevertheless, the analysis and findings discussed herein can be extended to other general cosine windows.

A prototype analysis filter, represented by a root-raised cosine window of length N, can be written as

$$h_0(n) = \left(\rho - \eta \cos\frac{2\pi n}{N}\right)^{0.5}, \forall n \in [0, N-1] \tag{44}$$

where, ρ and η are scalar parameters in the range $[0,1]$, with $\eta \leq \rho$ to ensure the expression remains real-valued for all n.

Using the binomial expansion $(a+x)^\beta = a^\beta + \beta a^{\beta-1}x + \frac{\beta(\beta-1)}{2}a^{\beta-2}x^2 + \cdots$, for $|x| < |a|$, followed by trigonometric power-reduction identities (e.g. $\cos^2(x) = \frac{1}{2}(1+\cos(2x))$, etc.), (44) can be expanded into the Fourier cosine series, i.e.,

$$\begin{aligned} h_0(n) &= \underbrace{\left(\rho^2 - \frac{\eta^2}{16}\right)}_{\gamma(0)} - \underbrace{\left(\rho\eta + \frac{3}{64}\rho\eta^2\right)}_{-\gamma(1)} \\ &\quad \cos\frac{2\pi n}{N} - \underbrace{\frac{\eta^2}{16}}_{-\gamma(2)} \cos\frac{4\pi n}{N} - \underbrace{\frac{1}{64}\rho\eta^2}_{-\gamma(3)} \cos\frac{6\pi n}{N} + \cdots \\ &= \sum_{r=0}^{\infty} \gamma(r) \cos\frac{2\pi r n}{N} \quad \text{with } |\gamma(r)| > |\gamma(r+1)| \end{aligned} \tag{45}$$

Therefore, the IDFT applied to the windowed t^{th} column of $\mathbf{U}[q]$ in (42), evaluated for subband k, can be expressed as

$$\begin{aligned} \check{U}_k(qD-t) &= \sum_{m=0}^{N-1}\sum_{r=0}^{\infty} \gamma(r)\cos\frac{2\pi r m}{N} u(qD-t-m)\alpha_N^{mk} \\ &= \sum_{r=0}^{\infty}\frac{\gamma(r)}{2}\left[\begin{array}{l} \overbrace{\sum_{m=0}^{N-1} u(qD-t-m)\alpha_N^{(k-r)m}}^{\mathbb{F}_N^*((k-r)\%N,:)\cdot\mathbf{U}[q](:,t)} \\ \underbrace{+\sum_{m=0}^{N-1} u(qD-t-m)\alpha_N^{(k+r)m}}_{\mathbb{F}_N^*((k+r)\%N,:)\cdot\mathbf{U}[q](:,t)} \end{array}\right] \\ &\approxeq \sum_{r=0}^{R}\frac{\gamma(r)}{2}\left[\mathbb{F}_N^*((k-r)\%N,:)\right. \\ &\quad \left. + \mathbb{F}_N^*((k+r)\%N,:)\right]\mathbf{U}[q](:,t) \end{aligned} \tag{46}$$

where, % represents the modulo operation and $[\mathbb{F}_N^*((k\pm r)\%N,:)\cdot\mathbf{U}[q](:,t)]$ corresponds to the IDFT of the non-windowed (or rectangular windowed with $h_0(n)=1, \forall n\in[0,N-1]$) t^{th} column of $\mathbf{U}[q]$ in (42), evaluated for subbands $(k\pm r)\%N$, and where the infinite sum has been truncated to R, leading to $2R$ so-called "cross-terms". The number of cross-terms required for the computation of the windowed IDFT, based on the non-windowed IDFT, is dependent on the prototype analysis filter $h_0(n)$. For instance, utilizing a cosine window as the prototype analysis filter necessitates two cross-terms ($R=1$), while employing a root-Hann window as the prototype analysis filter achieves an acceptable approximation with $R\cong 4$, corresponding to 8 cross-terms.

Comparing the expression for $\check{U}_k(qD-t)$ in (42) and (43) with (46), it can be observed that

$$\mathbb{F}_N^{win^*}(k,:) = \boldsymbol{\gamma}_R\,\mathbb{F}_N^*((k-R:k+R)\%N,:) \tag{47}$$

where, $\boldsymbol{\gamma}_R$ is a symmetric vector containing the $\gamma(r)$ coefficients given as

$$\boldsymbol{\gamma}_R \triangleq \left[\gamma(R),\ \gamma(R-1),\ \cdots,\ \gamma(0),\ \gamma(1),\ \cdots,\ \gamma(R)\right] \tag{48}$$

Consequently, the windowed IDFT of the t^{th} column of $\mathbf{U}[q]$, evaluated for subband k, can be expressed as a weighted linear combination of the non-windowed IDFT of the t^{th} column of $\mathbf{U}[q]$, evaluated for subbands $k\pm r$ ($r=0\ \cdots\ R$).

Substituting (47) in (43) then leads to

$$\check{Y}_k(q) = \boldsymbol{\gamma}_R\,\mathbb{F}_N^*((k-R:k+R)\%N,:)\ \mathbf{U}[q]\ \mathbf{c}_k^{\mathsf{T}} \tag{49}$$

For further simplification, (49) can be reformulated as (50), shown at the bottom of the next page, where matrix $\boldsymbol{\mathcal{F}}_k$ reflects the (computationally complex) T-non-windowed IDFT operations for the k^{th} subband.

Fortunately, the structure in the matrix $\boldsymbol{\mathcal{F}}_k$ can be exploited to reduce the required number of non-windowed IDFT operations to just one for each frame, by decomposing it as a product of a lower block triangular matrix and a sparse matrix, as shown in (51) at the bottom of the next page, where $\mathbb{I}_R$ is a diagonal matrix given as

$$\mathbb{I}_R = \begin{bmatrix} \alpha_N^{-R} & 0 & \cdots & 0 \\ 0 & \alpha_N^{-R+1} & \cdots & 0 \\ \vdots & & \ddots & 0 \\ 0 & \cdots & & \alpha_N^{R} \end{bmatrix} \tag{52}$$

$$\check{Y}_k(q) = \underbrace{\begin{bmatrix} \alpha_N^0 \\ \alpha_N^k \\ \vdots \\ \alpha_N^{(N-1)k} \end{bmatrix}^{\mathsf{T}} \begin{bmatrix} h_0(0) & 0 & \cdots & 0 \\ 0 & h_0(1) & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & h_0(N-1) \end{bmatrix} \underbrace{\begin{bmatrix} u(qD) & \cdots & u(qD-T+1) \\ u(qD-1) & \cdots & u(qD-T) \\ \vdots & \cdots & \vdots \\ u(qD-N+1) & \cdots & u(qD-N-T+2) \end{bmatrix}}_{\mathbf{U}[q]} \begin{bmatrix} c_k(0) \\ c_k(1) \\ \vdots \\ c_k(T-1) \end{bmatrix}}_{\mathbf{c}_k^{\mathsf{T}}}}_{\check{\mathbf{u}}_k[q] \triangleq \left[\check{U}_k(qD), \cdots, \check{U}_k(qD-T+1)\right]} \tag{42}$$

Authorized licensed use limited to: KU Leuven University Library. Downloaded on November 19,2025 at 14:23:26 UTC from IEEE Xplore. Restrictions apply.

With (51), expression (50) can be written as

$$\check{Y}_k(q) = \mathbf{c}_k \, \overline{\boldsymbol{\gamma}_R^T} \, \overline{\mathbb{L}_k} \, \overline{\mathbb{F}_k} \, \mathbf{u}\,[q] \tag{53}$$

where both the $\overline{\boldsymbol{\gamma}_R^T}$ and $\overline{\mathbb{L}_k}$ are independent of the input signal $u(n)$.

Finally, to further reduce the computational complexity, the matrix $\left(\overline{\boldsymbol{\gamma}_R^T} \; \overline{\mathbb{L}_k}\right)$ can be integrated into the subband adaptive filter coefficients $\mathbf{c}_k$, i.e. $\left(\mathbf{c}_k \; \overline{\boldsymbol{\gamma}_R^T} \; \overline{\mathbb{L}_k}\right)$ is replaced by $\mathbf{v}_k$, where $\mathbf{v}_k$ then represents the modified subband adaptive filter coefficients for subband k. This integration allows each subband k in PT-WOLA to have a filter input with one direct term and $2R$ cross-terms (represented by the first $2R+1$ rows of the matrix $\overline{\mathbb{F}_k}$), which are calculated from the output of a single IDFT operation, plus $T-1$ so-called difference terms [23] (represented by the last $T-1$ rows of the matrix $\overline{\mathbb{F}_k}$). The simplified expression can be formulated as (54) where, $\mathbf{v}_k \triangleq \begin{bmatrix} v_k(0) & \cdots & v_k(2R+T-1) \end{bmatrix}$ are the modified subband adaptive filter coefficients for PT-WOLA.

In conclusion, the T windowed IDFT operations for each q in (42), can be replaced by a single non-windowed IDFT operation, accompanied by $T-1$ difference terms along with $2R$ cross-terms, as illustrated in (54), shown at the bottom of the page, where $U_k^{rec}(qD)$ represents the non-windowed IDFT as defined in (46) i.e., $U_k^{rec}(qD) \triangleq \mathbb{F}_N^*\left((k)\%N,:\right)\mathbf{U}\,[q]\,(:,1)$ and $\Delta u(q-j) \triangleq u(qD-j) - u(qD-j-N)$ is a so-called difference term [23].

A comparison between (54) for PT-WOLA and (42) for the generalized WOLA filter bank reveals that the PT-WOLA implementation necessitates an additional $2R$ subband adaptive filter coefficients; i.e. each subband adaptive filter has $T+2R$ taps instead of T taps. Nonetheless, in the simulations presented in Section VI-E, the number of difference terms (or cross-terms) in PT-WOLA is intentionally reduced to ensure the same number of subband adaptive filter coefficients across both implementations. This adjustment enables a fair comparison. Under these conditions, it will be subsequently demonstrated that PT-WOLA performance remains comparable to that of the generalized WOLA filter bank for identical total subband adaptive filter lengths.

The structure of PT-WOLA, with a single non-windowed IDFT operation and $T-1$ difference terms along with $2R$ cross-terms, is depicted in Fig. 3.

A. Computational Complexity

This subsection presents an analysis of the computational complexity of the generalized WOLA filter bank and PT-WOLA. The computational complexity is evaluated in terms of the number of real multiplications and real additions required to process a single frame. The following standard conventions are adopted in the analysis. A complex-by-complex multiplication operation is counted as 4 real multiplications and 2 real additions, and a real-by-complex multiplication operation is counted as 2 real multiplications. Similarly, both complex-plus-

$$\check{Y}_k(q) = \mathbf{c}_k \underbrace{\begin{bmatrix} \boxed{\gamma_R} & \cdots & \mathbf{0} \\ \mathbf{0} & \boxed{\gamma_R} & \mathbf{0} \\ \vdots & & \ddots \\ \mathbf{0} & \cdots & \boxed{\gamma_R} \end{bmatrix}}_{\overline{\boldsymbol{\gamma}_R^T}} \underbrace{\begin{bmatrix} \boxed{\mathbb{F}_N^*\left((k-R:k+R)\%N,:\right)} & \cdots & 0 \\ 0 & \boxed{\mathbb{F}_N^*\left((k-R:k+R)\%N,:\right)} & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & & \boxed{\mathbb{F}_N^*\left((k-R:k+R)\%N,:\right)} \end{bmatrix}}_{\boldsymbol{\mathcal{F}}_k} \underbrace{\begin{bmatrix} u(qD) \\ u(qD-1) \\ \vdots \\ u(qD-N-T+2) \end{bmatrix}}_{\mathbf{u}\,[q]} \tag{50}$$

$$\boldsymbol{\mathcal{F}}_k = \underbrace{\left[\begin{array}{c|c|c|c|c} \left(\alpha_N^k \mathbb{I}_R\right)^{-0} & \mathbf{0}_{(2R+1)\times 1} & \mathbf{0}_{(2R+1)\times 1} & \cdots & \mathbf{0}_{(2R+1)\times 1} \\ \left(\alpha_N^k \mathbb{I}_R\right)^{-1} & -\mathrm{diag}\left(\left(\alpha_N^k \mathbb{I}_R\right)^{-1}\right) & \mathbf{0}_{(2R+1)\times 1} & \cdots & \mathbf{0}_{(2R+1)\times 1} \\ \vdots & & \ddots & & \vdots \\ \left(\alpha_N^k \mathbb{I}_R\right)^{-(T-1)} & -\mathrm{diag}\left(\left(\alpha_N^k \mathbb{I}_R\right)^{-(T-1)}\right) & -\mathrm{diag}\left(\left(\alpha_N^k \mathbb{I}_R\right)^{-(T-2)}\right) & \cdots & -\mathrm{diag}\left(\left(\alpha_N^k \mathbb{I}_R\right)^{-1}\right) \end{array}\right]}_{\left(\overline{\mathbb{L}_k}\right)_{T(2R+1)\times(2R+T)}} \underbrace{\left[\begin{array}{c|c|c} \left(\mathbb{F}_N^*\left((k-R:k+R)\bmod N,:\right)\right) & \mathbf{0}_{2R+1\times(T-1)} \\ \hline \mathbf{I}_{T-1} & \mathbf{0}_{(T-1)\times N-T} & -\mathbf{I}_{T-1} \end{array}\right]}_{\overline{\mathbb{F}_k}} \tag{51}$$

$$\check{Y}_k(q) = \mathbf{v}_k \left[\begin{array}{c|c|c} \left(\mathbb{F}_N^*\left((k-R:k+R)\%N,:\right)\right) & \mathbf{0}_{2R+1\times(T-1)} \\ \hline \mathbf{I}_{T-1} & \mathbf{0}_{(T-1)\times N-T} & -\mathbf{I}_{T-1} \end{array}\right] \begin{bmatrix} u(qD) \\ u(qD-1) \\ \vdots \\ u(qD-N-T+2) \end{bmatrix} = \mathbf{v}_k \begin{bmatrix} U^{rec}{}_{(k-R)\%N}(qD) \\ \vdots \\ U^{rec}{}_{(k+R)\%N}(qD) \\ \Delta u(q) \\ \vdots \\ \Delta u(q-T+2) \end{bmatrix} \tag{54}$$

Authorized licensed use limited to: KU Leuven University Library. Downloaded on November 19,2025 at 14:23:26 UTC from IEEE Xplore. Restrictions apply.

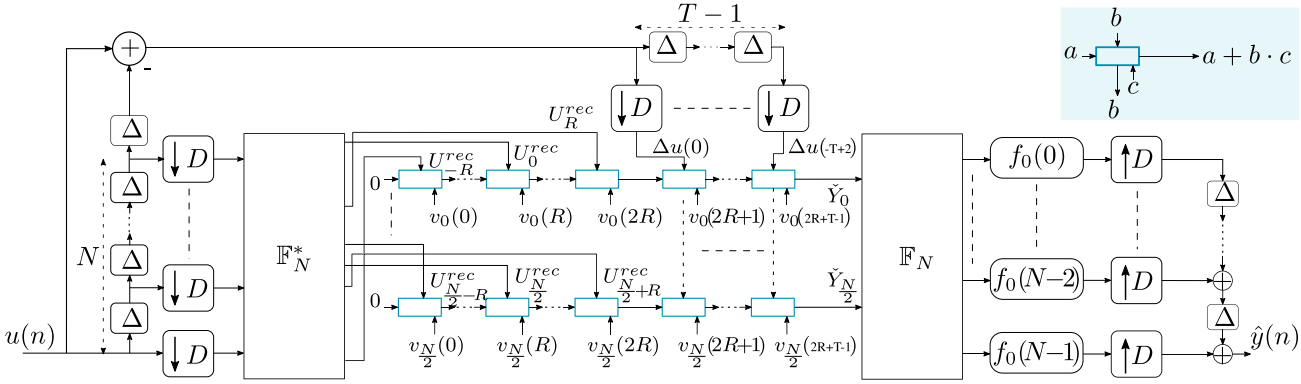


Fig. 3. Per-tone weighted overlap-add (PT-WOLA) filter bank.

complex and real-plus-complex addition operations are counted as 2 real additions, since the latter is typically implemented using a complex adder by treating the real input as having zero imaginary part. Finally, an N-point (I)DFT operation, based on the standard radix-2 algorithm, is assumed to require $2N\log_2(N)$ real multiplications and $3N\log_2(N)$ real additions.

1) Generalized WOLA filter bank: The first stage of the generalized WOLA filter bank based subband processing involves the computation of the term $\left(\mathbb{F}_N^{win^*}\mathbf{U}[q]\right)$ as shown in (43). This involves a windowing operation with N real multiplications followed by an N-point IDFT applied to each of the T columns of the input matrix $\mathbf{U}[q]$. Therefore, this transform stage requires $TN(2\log_2(N)+1)$ real multiplications and $3TN\log_2(N)$ real additions.

The second stage performs subband filtering through a complex-by-complex vector dot product between two T-element vectors, evaluated for $N/2$ unique subbands, followed by $T-1$ complex additions per subband. Hence the subband filtering stage requires $2NT$ real multiplications and $N(2T-1)$ real additions.

Overall, the computational complexity is dominated by the T separate N-point IDFT operations per frame. Therefore, the asymptotic complexity of the generalized WOLA filter bank is $\mathcal{O}(TN\log_2(N))$.

2) PT-WOLA: To allow a fair comparison with the generalized WOLA filter bank, the subsequent analysis assumes that the total number of filter taps in PT-WOLA is identical to that of generalized WOLA, with $2R+1$ cross-terms and $T-(2R+1)$ difference terms.

The first stage of PT-WOLA involves computing the input samples for the subband filtering stage. This consists of a single N-point IDFT applied to the first column of the input data matrix $U[q]$, along with the computation of $T-(2R+1)$ real difference terms. Therefore, this stage requires $2N\log_2(N)$ real multiplications and $3N\log_2(N)+(T-2R-1)$ real additions.

The second stage comprises subband filtering. It involves a complex-by-complex vector dot product between two $2R-1$-element vectors for the cross-terms, and a real-by-complex vector dot product between two $T-(2R+1)$-element vectors for the difference terms, both computed over $N/2$ unique subbands. The resulting T products are then summed, requiring $T-1$ complex additions. Hence, the subband filtering stage requires $N(2R+T+1)$ real multiplications and $N(2R+T)$ real additions.

In this case, the computational complexity is primarily dictated by the single N-point IDFT operation and the subband filtering stage. Therefore, the asymptotic complexity of PT-WOLA is $\mathcal{O}(N\log_2(N)+N(T+R))$. In typical scenarios where $N \gg T$, this simplifies to $\mathcal{O}(N\log_2(N))$, which is equivalent to the computational complexity of conventional WOLA.

The computational complexities for the generalized WOLA filter bank and the PT-WOLA are summarized in the Table I.

TABLE I
NUMBER OF FLOATING-POINT OPERATIONS REQUIRED BY GENERALIZED WOLA AND PT-WOLA BASED SUBBAND FILTERING

Metric	Generalized WOLA	PT-WOLA
Real Mult.	$TN(2\log_2(N)+3)$	$2N\log_2(N)+N(2R+T+1)$
Real Add.	$TN(3\log_2(N)+2)-N$	$3N\log_2(N)+(T-2R-1)+N(2R+T)$
Asymptotic	$\mathcal{O}(TN\log_2(N))$	$\mathcal{O}(N\log_2(N))$ (approx.)

VI. SIMULATION RESULTS

A. Acoustic Echo Cancellation (AEC)

For the evaluation of the subband system identification, single channel AEC is considered as a reference use-case scenario with the system distance and the ERLE as performance metrics. In this context, the room impulse response (RIR), represented by $w(n)$, describes the acoustic path, including all echoes and reverberation, from the loudspeaker to the microphone The system distance provides a measure of resemblance between the estimated distortion function $\hat{\mathbf{t}}$ and the actual RIR $\mathbf{w}_{ext}$, and is defined in decibels as $10\log_{10}\left(\frac{\|\mathbf{w}_{ext}-\hat{\mathbf{t}}\|^2}{\|\mathbf{w}_{ext}\|^2}\right)$. The ERLE provides a measure of the achieved echo suppression. For such AEC application in (1), $y(n)$ represents the microphone signal, $u(n)$ represents the far-end loudspeaker signal and $w(n)$ represents the RIR between the loudspeaker and the microphone. Moreover, $\mathbf{c}_k$ (or $\mathbf{v}_k$) represents the (converged) coefficients of the estimated low-order RIR in subband k.

Authorized licensed use limited to: KU Leuven University Library. Downloaded on November 19,2025 at 14:23:26 UTC from IEEE Xplore. Restrictions apply.

B. Simulation Scenario

A loudspeaker-enclosure-microphone (LEM) system is considered with a single microphone and a single loudspeaker. A non-stationary stochastic process with exponential decay, modeled after Polack's framework [27], is employed to generate the RIR of length $L = 512$ between the loudspeaker and microphone, i.e.

$$w(n) = \sqcup_L(n)\Omega(n)\exp^{-\kappa n} \tag{55}$$

where, $\sqcup_L$ is a rectangular function such that $\sqcup_L(n) = 1$ for $n = 0 \cdots L-1$, $\Omega(n)$ represents a zero-mean white Gaussian noise with unit variance and κ is the decay parameter of the exponential function, related to the reverberation time T_{60} by $\kappa = 3\ln(10)/T_{60}$. The far-end and near-end signals are both zero-mean white Gaussian noise signals, with an average echo-to-background ratio (EBR) of 20 dB. A WOLA filter bank is used with DFT size of $N = 1024$ and with 50% overlap (i.e., $D = 512$). Adaptation of subband coefficients $\mathbf{c}_k$ (or $\mathbf{v}_k$) is performed using the recursive least squares (RLS) algorithm [28]. For the simulations, the RLS algorithm was operated with a forgetting factor of $\lambda = 0.999$, and initialized with a covariance matrix $P(0) = 100 \cdot \mathbf{I}_T$ and all-zero filter coefficients. The simulation is performed at a sampling frequency of 16 kHz.

C. Prototype Filters and Time-Domain Subband Filters

For the performance evaluation, three windows were considered as prototype analysis filters; namely the rectangular window ($\sqcap_A$), the cosine window ($\cap_A$) and the root-Hann window ($\Cap_A$), as illustrated in Fig. 4. This selection ensures a broad evaluation across filters with distinct spectral characteristics. The rectangular and cosine window serve as representative extremes in terms of frequency selectivity, while the widely-used root-Hann window is included both for its relevance in practical applications and its effectiveness in highlighting the trade-offs inherent in the proposed PT-WOLA implementation (number of implemented cross-terms versus performance). The amplitude spectrum of the considered prototype analysis filters is depicted in Fig. 4(b). It can be observed that the cosine window exhibits the highest degree of frequency selectivity, while the rectangular window filter exhibits the lowest degree of frequency selectivity. For each prototype analysis filter, a corresponding perfect reconstruction prototype synthesis filter was designed, as shown in Fig. 5. The design was based on the (conventional) minimum-norm solution and on the recently proposed minimum-distortion solution presented in [26]. It can be observed that the rectangular window as prototype analysis filter and its corresponding minimum-distortion prototype synthesis filter, as depicted in Figs. 4(a) and 5(b), respectively, exemplify the prototype filters used in the overlap-save (OLS) method. Fig. 6 shows the (scaled) convolution $g_0(n)$ for the considered prototype analysis and synthesis filters.

Finally, Fig. 7 presents the values of $\sum_l \left|\tilde{c}^l(n)\right|$ for $n = 0 \cdots 2N-2$, derived from the converged subband filter coefficients as defined in (16). For brevity, results are only shown for the rectangular prototype analysis filter with a subband filter length of $T = 15$. However, without loss of generality, the same conclusions can be drawn for different prototype analysis filters and subband filter lengths. Moreover, to highlight the effect of the RIR length L, the values are presented for $L = 512$ and $L = 15$, as shown in Fig. 7(a) and 7(b), respectively. As can be observed from Fig. 7, a significant portion of the power of $\tilde{c}^l(n)$ is contained approximately between samples $n = 0 \cdots L-T$ (and between $n = N-1 \cdots N+L-T-1$ due to periodicity as derived in (16)). Therefore, the maximum length of $\tilde{\mathbf{c}}^l$ is $L-T+1$, which substantiates the theoretical derivations presented in (31) and (35).

D. Performance Evaluation: Generalized WOLA

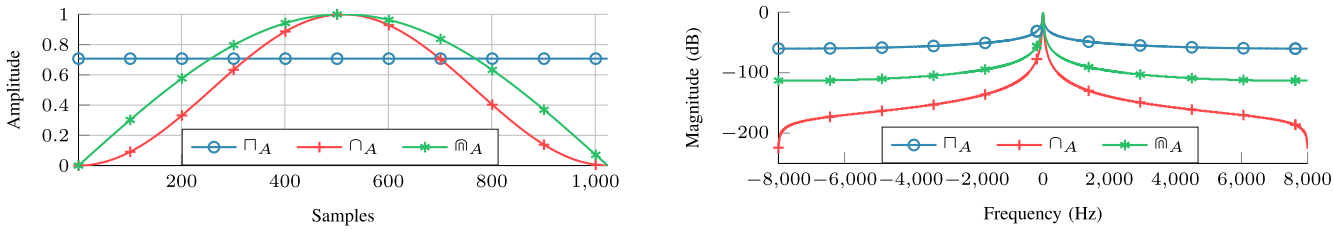
This subsection presents the experimental results that corroborate the theoretical conclusions drawn in Section IV for the considered AEC scenario. The system distance and ERLE curves are respectively shown in Fig. 8(a) and 8(b), for different prototype filters and subband filter lengths (T). It is worth noting that the subband filter length of $T = 1$, with different prototype analysis filters and their corresponding minimum-norm prototype synthesis filters, represents the conventional WOLA method.

It is observed that the system distance and the ERLE performance for all the prototype filters improve as the length of the subband filters (T) increases. This observation aligns with the findings derived in Section IV, as presented in Fig. 9. The figure illustrates the ERLE, calculated using (40), for the prototype analysis filters shown in Fig. 4, along with their respective minimum-norm (– – –) prototype synthesis filters.

Notably, the system distances achieved by the minimum-distortion prototype synthesis filters are smaller than those achieved by the minimum-norm synthesis filters for all prototype analysis filters and for all subband filter lengths, as shown in Fig. 8(a). In the open-loop subband system identification problem highlighted in (23), the converged filter coefficients are independent of the synthesis filter bank. Therefore, they remain unaltered under the same set of system conditions, such as prototype analysis filter, input signals, RIR, filter length, etc., regardless of the prototype synthesis filter used. Hence, the system distance in Fig. 8(a), illustrates that, for a given prototype analysis filter and corresponding converged subband filter coefficients, the minimum-distortion prototype synthesis filters transform the subband filter coefficients into a distortion function that models the actual RIR more accurately than the minimum-norm prototype synthesis filters. This improvement in system distance is partly due to the ability of minimum-distortion prototype synthesis filters to better suppress the circular convolution effect in the distortion function $\tilde{\mathbf{t}}$, which arises from the two periodic images in $\tilde{\mathbf{c}}^l$.

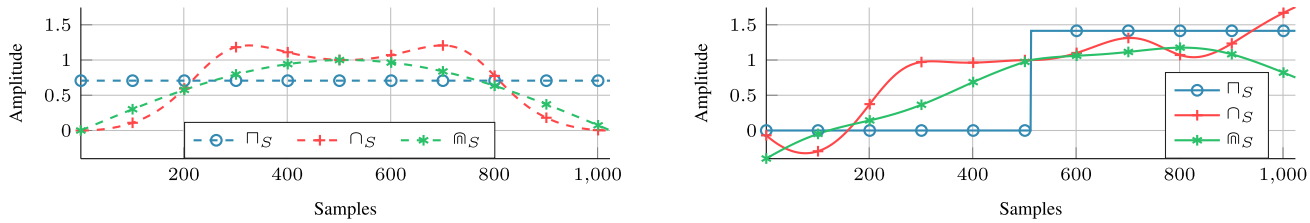
As shown in Fig. 6, the amplitude of the $g_0(n)$ samples corresponding to the (first) periodic image in $\tilde{\mathbf{c}}^l$ is significantly lower for minimum-distortion prototype synthesis filters than for minimum-norm prototype synthesis filters. Therefore, from (17), this leads a to better suppression of the circular convolution effect in the distortion function $\tilde{\mathbf{t}}$, consequently resulting in a lower system distance. This improvement in achieved system distance also improves the overall AEC performance of the

Authorized licensed use limited to: KU Leuven University Library. Downloaded on November 19,2025 at 14:23:26 UTC from IEEE Xplore. Restrictions apply.



(a) Prototype analysis filters (b) Amplitude spectrum of prototype analysis filters

Fig. 4. Prototype analysis filters and their corresponding amplitude spectrum (envelope).



(a) Minimum norm prototype synthesis filters (b) Minimum distortion prototype synthesis filters

Fig. 5. Minimum norm (– – –) and minimum-distortion (——) prototype synthesis filters of prototype analysis filters shown in Fig. 4.

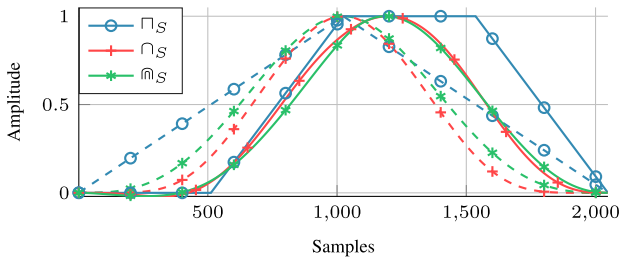


Fig. 6. Scaled convolution of the considered prototype analysis filters with their respective minimum-norm (– – –) and minimum-distortion (——) prototype synthesis filters.

system, as evident in Fig. 8(b), where the ERLE achieved by the minimum-distortion prototype synthesis filters consistently outperforms the ERLE achieved by the minimum-norm prototype synthesis filters for all prototype analysis filters and subband filter lengths.

Finally, the rectangular prototype analysis filter with the corresponding minimum-distortion prototype synthesis filter, despite representing the prototype filters for the OLS method, which facilitates perfect suppression of circular convolution effect, has its overall performance restricted by the poor frequency selectivity of the rectangular prototype analysis filter (Fig. 4b), leading to a poor convergence of subband filters and motivates the use of prototype filters with better frequency selectivity.

E. Performance Evaluation: PT-WOLA

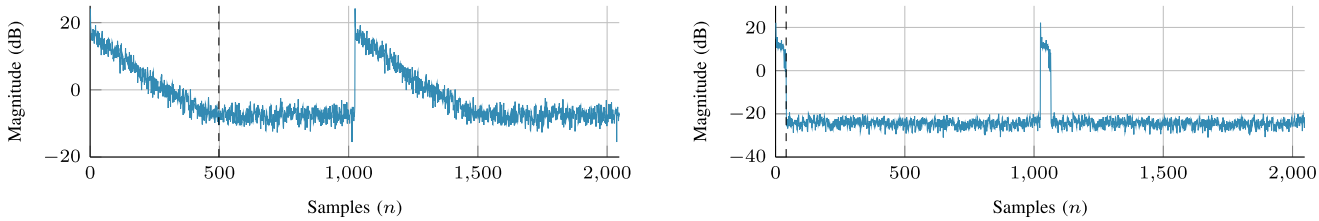
While generalized WOLA has demonstrated notable performance improvements, its high computational complexity, dominated by T windowed IDFT operations per time frame, limits its practical application. To address this issue, Section V proposes a computationally efficient PT-WOLA method, which reduces the T windowed IDFT operations per frame to a single non-windowed IDFT operation, $T-1$ difference terms, and $2R(+1)$ cross-terms ($+1$ direct term) per frame. The number of cross-terms depends on the prototype analysis filter (see (46)). This section compares the performance of PT-WOLA with generalized WOLA, with identical total subband adaptive filter lengths, for both minimum-norm and minimum-distortion prototype synthesis filters.

Fig. 10 presents a performance comparison between the PT-WOLA with $\mathbf{v}_k^{\mathsf{T}}$ and the generalized WOLA with $\mathbf{c}_k^{\mathsf{T}}$, considering the same total filter lengths and a consistent number of cross-terms $2R$, necessary to represent each prototype analysis filter. Consequently, given a total filter length of T, the generalized WOLA necessitates T windowed IDFT inputs, whereas the PT-WOLA implementation requires $2R$ $(+1)$ cross(+direct)-terms and $\acute{T} = T - (2R+1)$ difference terms, as highlighted in (54).

As demonstrated in Fig. 10, with a suitable number of cross-terms ($2R$), as defined in Section V, the performance exhibited by the PT-WOLA is comparable to the performance of the generalized WOLA for identical total filter lengths (T). Moreover, PT-WOLA delivers this comparable performance by utilizing a smaller or equal number of difference terms, (which represents the sliding IDFTs in generalized WOLA) with $\acute{T} = T - (2R+1) \leq T$. This is because, unlike the generalized WOLA, where any prototype analysis window implies having constant symmetric subband cross-term filter coefficients ($\gamma(-r) = \gamma(r)$ in (46)), PT-WOLA does not adhere to this symmetry constraint. This grants PT-WOLA extra degrees of freedom for identical total filter lengths. This has been meticulously observed in the converged subband filter coefficients for the cross-terms in PT-WOLA, which indeed in general are non-symmetric.

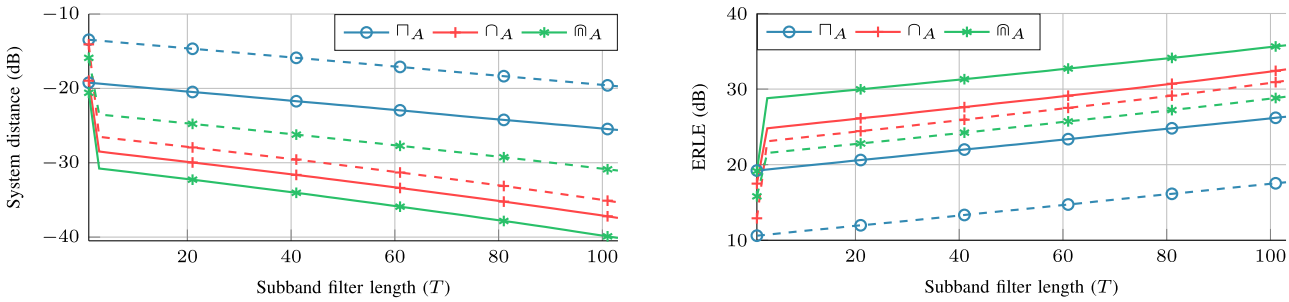
Finally, while the required number of single-sided cross-terms (R) to represent the windowing operation is finite for the

Authorized licensed use limited to: KU Leuven University Library. Downloaded on November 19,2025 at 14:23:26 UTC from IEEE Xplore. Restrictions apply.



(a) $\sum_l \left|\tilde{c}^l(n)\right|$ for $T = 15$ and $L = 512$ (b) $\sum_l \left|\tilde{c}^l(n)\right|$ for $T = 15$ and $L = 55$

Fig. 7. $\sum_l \left|\tilde{c}^l(n)\right|$ values in dB with rectangular prototype analysis filter and $T = 15$ for RIR length of $L = 512$ and $L = 55$. The vertical line (– – –) marks the sample point at $n = L$.



(a) System distance curves (b) ERLE curves

Fig. 8. System distance and ERLE achieved by subband processing with different subband filter lengths and considered prototype analysis filters with their respective minimum-norm (– – –) and minimum-distortion (——) prototype synthesis filters.

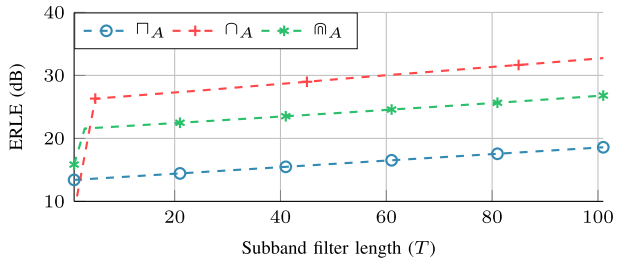


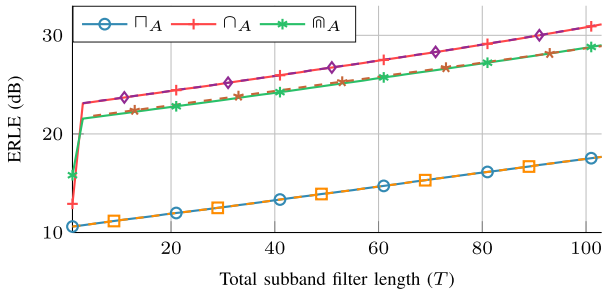
Fig. 9. ERLE calculated using (40) for the prototype analysis filters shown in Fig. 4 with their respective minimum-norm prototype synthesis filters.

rectangular ($R = 0$) and the cosine ($R = 1$) prototype analysis filters, root-Hann prototype filters demand an infinite number of cross-terms, as shown in (45). Consequently, the inter-subband aliasing effect diminishes as the number of cross-terms for the root-Hann prototype analysis filter is increased. However, for a fixed subband filter length (T), the increase in number of cross-terms reduces the available difference terms ($\acute{T} = T - (2R + 1)$), thereby amplifying the distortion error. Therefore, the necessary number of cross-terms (R) for PT-WOLA with the root-Hann prototype analysis filter to achieve comparable performance to generalized WOLA depends on whether the inter-subband aliasing or distortion error is more dominant. This is exemplified in Fig. 10(a), where PT-WOLA matches the performance of generalized WOLA with $R = 2$ for a minimum-norm prototype synthesis filter. However, as shown in Fig. 10(b), PT-WOLA requires a larger number of cross-terms, $R = 4$, to achieve the performance of generalized WOLA. This discrepancy arises because, for the root-Hann prototype analysis filter, the minimum-norm synthesis filter is subject to significant distortion error. Therefore, performance is limited by the distortion errors and PT-WOLA requires more difference terms to offset it and match the performance of the generalized WOLA. Meanwhile, with the minimum-distortion synthesis filter, the distortion errors are confined, and PT-WOLA performance is limited by the inter-subband aliasing, thereby requiring additional cross-terms to match the performance of the generalized WOLA filter bank.

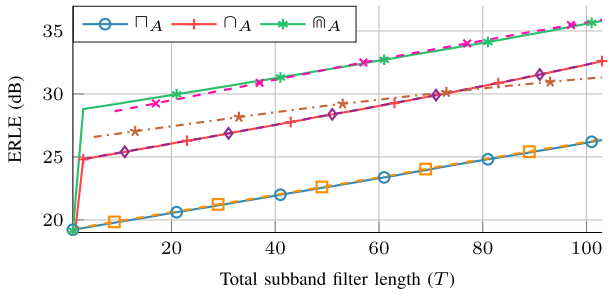
Fig. 11 presents a similar performance analysis for PT-WOLA with a fixed number of difference terms ($\acute{T}$). This implies that for a specific total filter length (T), the number of single-sided cross-terms is given by $R = \frac{T-\acute{T}-1}{2}$. For the sake of brevity and clarity, the performance analysis focuses on $\acute{T} = 5$ and $\acute{T} = 51$. However, it should be noted that these findings can be extrapolated to other values of $\acute{T}$, thus maintaining the generality of the observations.

For the cosine prototype analysis filter, minimal performance improvement is seen beyond $R = 1$, particularly when using the minimum-norm synthesis filter. This trend is attributed to the frequency selectivity of the cosine prototype analysis filter (Fig. 4b), where the magnitude spectrum for the cosine prototype analysis filter drops significantly beyond $R = 1$, limiting its impact on the inter-subband aliasing error. With the rectangular prototype analysis filter, which exhibits poor frequency selectivity (Fig. 4b), performance is largely dictated by the inter-subband aliasing error. Consequently, adding cross-terms improves performance, even outperforming the generalized

Authorized licensed use limited to: KU Leuven University Library. Downloaded on November 19,2025 at 14:23:26 UTC from IEEE Xplore. Restrictions apply.

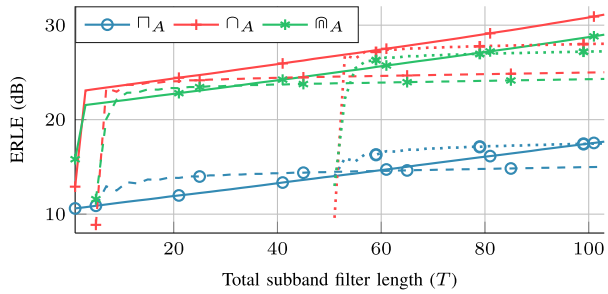


(a) ERLE achieved with minimum-norm prototype synthesis filters.

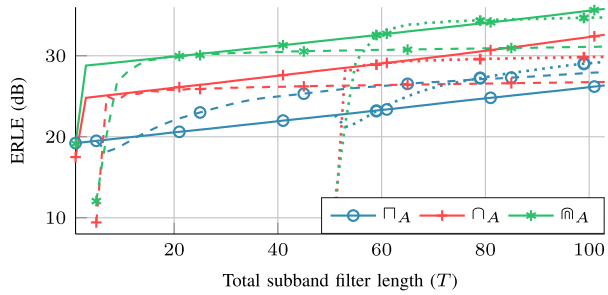


(b) ERLE achieved with minimum-distortion prototype synthesis filters.

Fig. 10. ERLE achieved by PT-WOLA (– – –) and generalized WOLA (——) for $L = 512$ with different total subband filter lengths and constant number of cross-terms (R) for prototype filters shown in Figs. 4 and 5. The cross-terms for different prototype filters are: $\sqcap_A$ $R = 0$ (- □ -), $\cap_A$ $R = 1$ (- ◇ -), $\Cap_A$ $R = 2$ (- ⋆ -), $\Cap_A$ $R = 4$ (- ⋆ -).



(a) ERLE achieved with minimum-norm prototype synthesis filters



(b) ERLE achieved with minimum-distortion prototype synthesis filters

Fig. 11. ERLE achieved by PT-WOLA ($\acute{T} = 5$ (– – –), $\acute{T} = 51$ (·····)) and generalized WOLA (——) for $L = 512$ with different total subband filter lengths and constant number of difference terms ($\acute{T}$) for prototype filters shown in Figs. 4 and 5.

WOLA for the same filter length. However, this performance improvement plateaus earlier for the minimum-norm synthesis filter (Fig. 11a) compared to the minimum-distortion synthesis filter (Fig. 11b). This is due to the higher distortion error for the rectangular prototype analysis filter with the minimum-norm synthesis filter compared to the minimum-distortion synthesis filter (Fig. 6). Hence, as the inter-subband aliasing error decreases with an increased number of cross-terms, the distortion error dominates the aliasing error much earlier for the minimum-norm synthesis filter than for the minimum-distortion synthesis filter. Lastly, a similar pattern can be observed for the root-Hann prototype analysis filter, where an increase in the number of cross-terms results in a lower inter-subband aliasing error. Nevertheless, this performance enhancement becomes minimal with an increasing number of cross-terms of R, as their contribution to inter-subband aliasing drops significantly, as shown in Fig. 4(b).

Furthermore, the analysis presented in Fig. 11 also serves as a direct comparison with state-of-the-art cross-band filtering methods from the literature [8], [9]. These methods are structurally equivalent to a special case of the PT-WOLA framework where the number of difference terms is minimal ($\acute{T} = 1$) and performance is improved by adding cross-terms ($R > 0$). Fig. 11 illustrates precisely this concept, showing how the ERLE evolves as the number of cross-terms is increased for a fixed number of difference terms.

VII. Conclusion

This paper has addressed the constraints inherent in conventional WOLA filter banks, and introduced a generalized WOLA filter bank for subband adaptive filtering and investigated its performance for subband system identification.

The MSE performance of the generalized WOLA filter bank has been analyzed, highlighting the critical factors for accurate full-band system identification, including the order of the subband filters, the full-band system impulse response length, decimation factor, and the design of the analysis and synthesis filter bank.

Both analytical and empirical evidence has confirmed that the proposed generalized WOLA filter bank substantially enhances performance in subband system identification compared to the conventional WOLA filter bank. To mitigate the increased computational complexity associated with the generalized WOLA filter bank, the PT-WOLA was proposed, which maintains computational complexity on par with conventional WOLA, while demonstrating performance that is comparable to, or slightly exceeds, that of the generalized WOLA filter bank for the same subband filter order.

For future research, a promising direction is to extend the application of the generalized WOLA and PT-WOLA frameworks from single-channel system identification to multi-channel spatial filtering. Conventional WOLA filter banks are widely used for tasks such as adaptive beamforming for noise reduction and

Authorized licensed use limited to: KU Leuven University Library. Downloaded on November 19,2025 at 14:23:26 UTC from IEEE Xplore. Restrictions apply.

Direction of Arrival (DOA) estimation. The improved modeling capabilities demonstrated in this paper could offer significant performance enhancements in these multi-channel scenarios.

ACKNOWLEDGMENT

The scientific responsibility is assumed by its authors.

DATA AVAILABILITY SECTION

The complete source code is publicly available at: https://github.com/mohit-nith/GeneralizedWOLA-SystemIdentification.

REFERENCES

[1] R. Crochiere, "A weighted overlap-add method of short-time Fourier analysis/synthesis," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 28, no. 1, pp. 99–102, Feb. 1980.

[2] J. Shynk, "Frequency-domain and multirate adaptive filtering," *IEEE Signal Process. Mag.*, vol. 9, no. 1, pp. 14–37, Jan. 1992.

[3] G. Clark, S. Parker, and S. Mitra, "A unified approach to time- and frequency-domain realization of FIR adaptive digital filters," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 31, no. 5, pp. 1073–1083, Oct. 1983.

[4] P. S. R. Diniz, *Subband Adaptive Filters.* Boston, MA, USA: Springer US, 2013, pp. 501–549.

[5] P. Vaidyanathan, *Multirate Systems and Filter Banks* (Prentice-Hall Signal Processing Series). Englewood Cliffs, NJ, USA: Prentice Hall, 1993.

[6] A. Gilloire and M. Vetterli, "Adaptive filtering in subbands," in *Proc. Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 1988, vol. 3, pp. 1572–1575.

[7] A. Gilloire and M. Vetterli, "Adaptive filtering in subbands with critical sampling: Analysis, experiments, and application to acoustic echo cancellation," *IEEE Trans. Signal Process.*, vol. 40, no. 8, pp. 1862–1875, Aug. 1992.

[8] Y. Avargel and I. Cohen, "Performance analysis of cross-band adaptation for subband acoustic echo cancellation," in *Proc. Int. Workshop Acoust. Echo Noise Control (IWAENC),* 2006, pp. 1–4.

[9] Y. Avargel and I. Cohen, "System identification in the short-time Fourier transform domain with crossband filtering," *IEEE Trans. Audio, Speech, Language Process.*, vol. 15, no. 4, pp. 1305–1319, May 2007.

[10] M. Krini and G. Schmidt, "Method for temporal interpolation of short-term spectra and its application to adaptive system identification," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP),* 2012, pp. 45–48.

[11] Z. Shpiro and D. Malah, "Design of filters for discrete short-time Fourier transform synthesis," in *Proc.IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, vol. 10, pp. 537–540. 1985.

[12] S. Bagheri and D. Giacobello, "Robust STFT-domain multi-channel acoustic echo cancellation with adaptive decorrelation of the reference signals," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP),* 2021, pp. 131–135.

[13] M. Guo and B. Kuenzle, "Obtaining narrow transition region in STFT-domain processing using subband filters," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP),* 2019, pp. 970–974.

[14] M. L. Valero and E. A. Habets, "On the spatial coherence of residual echoes after STFT-domain multi-microphone acoustic echo cancellation," in *Proc. 16th Int. Workshop Acoust. Signal Enhancement (IWAENC)*, 2018, pp. 141–145.

[15] T. Tanaka, "A direct design of oversampled perfect reconstruction FIR filter banks," *IEEE Trans. Signal Process.*, vol. 54, no. 8, pp. 3011–3022, Aug. 2006.

[16] K. Eneman and M. Moonen, "DFT modulated filter bank design for oversampled subband systems," *Signal Process.*, vol. 81, no. 9, pp. 1947–1973, 2001.

[17] M. F. Mansour, "On the optimization of oversampled DFT filter banks," *IEEE Signal Process. Lett.*, vol. 14, no. 6, pp. 389–392, Jun. 2007.

[18] M. Renfors, J. Yli-Kaakinen, and F. J. Harris, "Analysis and design of efficient and flexible fast-convolution based multirate filter banks," *IEEE Trans. Signal Process.*, vol. 62, no. 15, pp. 3768–3783, Aug. 2014.

[19] J. H. Gunther and G. Wilson, "Mean-squared error analysis of adaptive subband-based system identification," *IEEE Trans. Audio, Speech, Language Process.*, vol. 16, no. 8, pp. 1452–1465, Nov. 2008.

[20] K. Eneman and M. Moonen, "Hybrid subband/frequency-domain adaptive systems," *Signal Process.*, vol. 81, no. 1, pp. 117–136, 2001.

[21] R. Merched and A. Sayed, "An embedding approach to frequency-domain and subband adaptive filtering," *IEEE Trans. Signal Process.*, vol. 48, no. 9, pp. 2607–2619, Sep. 2000.

[22] S. Shimauchi and H. Ohmuro, "Accurate adaptive filtering in square-root Hann windowed short-time Fourier transform-domain," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP),* 2014, pp. 1305–1309.

[23] K. Van Acker, G. Leus, M. Moonen, O. van de Wiel, and T. Pollet, "Per tone equalization for DMT-based systems," *IEEE Trans. Commun.*, vol. 49, no. 1, pp. 109–119, Jan. 2001.

[24] M. Sharma, J. Verdyck, Y. Lefevre, P. Tsiaflakis, and M. Moonen, "MIMO per-tone equalizer design for long reach xDSL," *IEEE Open J. Commun. Soc.*, vol. 3, pp. 51–64, 2022.

[25] R. Alves, M. Petraglia, and P. Diniz, "Convergence analysis of an oversampled subband adaptive filtering structure using global error," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2000, vol. 1, pp. 468–471.

[26] M. Sharma and M. Moonen, "Prototype filter design for weighted overlap-add filter bank based sub-band adaptive filtering applications," in *Proc. 31st Eur. Signal Process. Conf. (EUSIPCO)*, 2023, pp. 366–370.

[27] J.-D. Polack, "La Transmission de L'energie Sonore Dans Les Salles," Ph.D. Thesis, Thèse de doctorat dirigée par Bruneau, Michel Physique Le Mans, 1988.

[28] S. Haykin, *Adaptive Filter Theory*, 5th ed., International Edition. Harlow, U.K.: Pearson Education, 2014.

Authorized licensed use limited to: KU Leuven University Library. Downloaded on November 19,2025 at 14:23:26 UTC from IEEE Xplore. Restrictions apply.