Autoregressive Modeling of Temporal/Spectral Envelopes With Finite-Length Discrete Trigonometric Transforms

Han-Wen Hsu and Chi-Min Liu

***Abstract*—The theory of autoregressive (AR) modeling, also known as linear prediction, has been established by the Fourier analysis of infinite discrete-time sequences or continuous-time signals. Nevertheless, for various finite-length discrete trigonometric transforms (DTTs), including the discrete cosine and sine transforms of different types, the theory is not well established. Several DTTs have been used in current audio coding, and the AR modeling method can be applied to reduce coding artifacts or exploit data redundancies. This paper systematically develops the AR modeling fundamentals of temporal and spectral envelopes for the sixteen members of the DTTs. This paper first considers the AR modeling in the generalized discrete Fourier transforms (GDFTs). Then, we derive the modeling to all the DTTs by introducing the analytic transforms which convert the real-valued vectors into complex-valued ones. Through the process, we build the compact matrix representations for the AR modeling of the DTTs in both time domain and DTT domain. These compact forms also illustrate that the AR modeling for the envelopes can be performed through the Hilbert envelope and the power envelope. These compact forms can be used to develop new coding technologies or examine the possible defects in the existing AR modeling methods for DTTs, We apply the forms to analyze the current temporal noise shaping (TNS) tool in MPEG-2/4 advanced audio coding (AAC).**

***Index Terms*—Autoregressive (AR) modeling, discrete cosine transform (DCT), discrete trigonometric transform (DTT), generalized discrete Fourier transform (GDFT), Hilbert envelope, linear prediction, frequency-domain linear prediction, linear prediction in spectral domain, temporal noise shaping (TNS).**

I. Introduction

AUTOREGRESSIVE (AR) modeling is widely used for power spectrum estimation [1], [2]. An order-P AR model, also referred to as linear prediction model, is defined by $x(n) = -\sum_{k=1}^{P} a_k x(n-k) + e(n)$, where $x(n)$ is the data sequence and $e(n)$ is a random sequence of variance $\boldsymbol{\sigma}_e^2$. The power spectral density, or power spectrum, of the order-P AR model is given by

$$P_X^{\mathrm{AR}}(f) = \boldsymbol{\sigma}_e^2 \cdot \left| 1 + \sum_{k=1}^{P} a_k \exp(-j2\pi fk) \right|^{-2} \quad \text{for } 0 \le f \le 1.$$

From the aspect of filtering, an AR model describes a data sequence generated from an all-pole filter fed by a white excitation sequence. As a source-filter model with low dimension parameters, the AR modeling method has been successfully utilized to model the acoustical system of speech production, which comprises the vocal tract and the glottal excitation, and has become an essential method in many speech applications, such as speech coding, speech recognition, pitch tracking, and formant estimation.

The AR modeling or linear prediction has received more and more applications in audio coding and processing. For example, the predictive error filter is used as an inverse filter for acoustic echo cancellation [3]. The spectral band replication (SBR) [4], which is adopted in MPEG-4 high efficiency advanced audio coding (HE-AAC) [5] for bandwidth extension, uses a second-order linear predictor for inverse filtering; and MPEG-4 audio lossless coding (ALS) [6] generates residuals with a smaller dynamic range by linear prediction. Some audio coding approaches are based on the linear prediction performed on a warped frequency scale [7], [8]. In addition to the applying of AR modeling in the time domain, according to the duality between the squared Hilbert envelope and the power spectrum [5], AR modeling can be applied to spectral sequences for temporal envelope estimation. Because of its ability to provide a smooth temporal representation of signals, AR modeling is very useful for feature extraction and classification [9], [10]. The temporal noise shaping (TNS) tool [12]–[15], which is adopted for reducing the pre-echo artifacts [16] in MPEG-2/4 AAC [5], applies an open-loop prediction [17] across frequency lines prior to quantization in encoder to shape the temporal envelope of the quantization noise.

Although AR modeling is always preformed on windowed signals in practice, its theoretical derivation was given through the Fourier theory of infinite discrete-time sequences or continuous signals in the literatures such as [1] and [2]. In [12]–[15], Herre and Johnston explained the concept of TNS through the duality between the squared Hilbert envelope and the power spectrum for continuous signals. Likewise, Kumaresan *et al.* [18]–[22] formulated the linear prediction in spectral domain equations for the AR modeling of temporal envelope in the indirect way. No exact derivation for finite sequences was developed until Athineos and Ellis [11] formulated via matrix operations the solution of the problem that finds an AR model of a discrete spectrum and relates it to the temporal envelope of a finite time-domain sequence. However, the derivation in [11] was limited in the scenario when the discrete spectrum is the odd type-I discrete cosine transform (DCT-I) coefficients.

Manuscript received August 26, 2009; accepted March 04, 2010. Date of publication April 01, 2010; date of current version June 16, 2010. This work was supported in part by the National Science Council under NSC98-2221-E-009-120. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Maciej Niedzwiecki.

The authors are with the Department of Computer Science, National Chiao Tung University, Hsinchu 300, Taiwan (e-mail: hwhsu@csie.nctu.edu.tw).

Digital Object Identifier 10.1109/TSP.2010.2047105

In this paper, we concern the temporal and spectral AR modeling of a finite sequence when one of the 16 members of the discrete trigonometric transform (DTT) is used in the temporal and spectral domains. Different from DCT-I, other DTTs have a 1/2-sample delay in the time domain and/or a 1/2-sample advance in the frequency domain. When considering a finite-length sequence as a discrete periodic signal obtained by sampling a continuous signal, we need to consider the aliasing effect in the dual domain. We systematically establish the AR modeling fundamentals for the DTTs by exploiting the relationship of the DTTs and the generalized discrete Fourier transforms (GDFTs) [27]. We address the AR problem with GDFTs by extending the well-known relationship between the autocorrelation and the power spectrum to the GDFT/Inverse GDFT domains. Then we define new finite-length analytic transforms based on GDFTs. Through the analytic transforms, we establish the AR modeling fundamentals for DTTs by relating the DTT spectra to the corresponding GDFT spectra with appropriate symmetric extension or zero padding operations. In addition to the temporal Hilbert envelope, we also concern the power envelope estimation for a real-valued sequence without introducing the Hilbert signal. The dual results can be derived with the consistent representation in the temporal and spectral domains, i.e., both are periodic and finite. Our formulation is expressed entirely in the discrete finite domain in matrix form. The compact expressions not only disclose that the AR modeling concept can hold in each DTT domain, not limited in DCT-I domain, but also can be used for more clearly examining the related methods based on AR modeling in DTTs.

This paper is organized as follows. Section II summarizes the definitions and properties on GDFTs and DTTs. Section III formulates the AR modeling of temporal and spectral envelopes in GDFTs/IGDFTs. Section IV derives the AR modeling for the DTTs/IDTTs. Section V applies the result to analyze the TNS in MPEG-2/4 AAC. Section VI concludes the paper.

II. Preliminaries

To derive the AR modeling with the DTTs, we summarize the essential results on the convolution-multiplication and periodicity properties for the GDFTs. The convolution–multiplication property sets the fundamental for modeling the temporal/spectral envelopes. The periodicity property is fundamental to discuss the effect of finite-length transforms when applying to periodic sequences. Then, the terms, transform formula, periodicity, and the relation with GDFTs are summarized for the 16 DTTs.

A. Notations

Throughout this paper, we use calligraphic capital letters to denote matrices (e.g., $\boldsymbol{A}, \boldsymbol{G}, \boldsymbol{T}$) and calligraphic lower case letters to denote column vectors (e.g., $\boldsymbol{a}, \boldsymbol{x}, \boldsymbol{y}$). For vectors and matrices, both row and column indices used are zero based. We use the notations $\boldsymbol{x}_k$ or $(\boldsymbol{x})_k$ to refer to the kth entry of $\boldsymbol{x}$. Some operation notations are described as follows. Superscripts $(T), (H)$, and $(^*)$ denote the transpose, Hermitian transpose, and conjugate operations, respectively. The notation $(\circ)$ denotes the Hadamard product (i.e., the entry-wise product of two vectors or matrices); $(\|\cdot\|)$ denotes 2-norm. In terms of linear algebra, we consider and represent a finite sequence $x(n)$, for $n_0 \le n \le n_1$, as a column vector $\boldsymbol{x}$ of length $(n_1 - n_0 + 1)$: $\boldsymbol{x} = [x(n_0), x(n_0+1), x(n_0+2), \ldots, x(n_1)]^T$.

B. Generalized Discrete Fourier Transform

The generalized DFT (GDFT) [25] of a finite sequence $x(n)$, $n = 0, 1, \ldots, N-1$ is defined as

$$y(k) = \sum_{n=0}^{N-1} x(n) \exp\{-j2\pi(k+a)(n+b)/N\} \quad \text{for } k = 0, 1, \ldots, N-1. \quad (1)$$

Four special forms of the GDFT arise when a and b take on the values 0 or 1/2. They are classified and named as follows [27]: 1) DFT (discrete Fourier transform): $a = 0$ and $b = 0$; 2) OTDFT (Odd-Time DFT): $a = 0$ and $b = 1/2$; 3) OFDFT (Odd-Frequency DFT): $a = 1/2$ and $b = 0$; 4) O^{2}DFT (Odd-Time Odd-Frequency DFT): $a = 1/2$ and $b = 1/2$. The last three transforms can be regarded as the modified version of the DFT with a 1/2-sample delay in the time domain and/or a 1/2-sample advance in the frequency domain. The GDFT matrix is defined by $[\boldsymbol{G}_{a,b}]_{k,n} = \exp\{-j2\pi(k+a)(n+b)/N\}$, where the row and column indices are $k, n = 0, 1, \ldots, N-1$. Since the inverse GDFT (IGDFT) matrix is the scaled Hermitian transpose of the forward GDFT matrix, the IGDFT matrix is related to the forward matrix as

$$\boldsymbol{G}_{a,b}^{-1} = \frac{1}{N}\boldsymbol{G}_{a,b}^{H} = \frac{1}{N}\boldsymbol{G}_{b,a}^{*}. \quad (2)$$

C. Generalized-Periodic Sequence, Periodic Convolution, and GDFT

Considering the generalized-periodic sequence (GPS) by extending a finite sequence into an infinite sequence in either strictly periodic or anti-periodic way, Martucci summarized the four different periodic relationships for the four special GDFTs/IGDFTs in [27, Table I]. A sequence is said to be anti-periodic with period N if $x(n) = -x(n+N)$ for all n. For a period- N GPS, we refer to the samples in the base period (for index $n = 0, 1, \ldots, N-1$) as the representative samples or vector.

The periodic convolution of two period-N GPSs, $\tilde{x}(n)$ and $\tilde{y}(n)$, of the same type (i.e., both of them are either strictly periodic or anti-periodic) is defined as $\tilde{x}(n) \otimes \tilde{y}(n) = \sum_{k=0}^{N-1} \tilde{x}(k)\tilde{y}(n-k) = \sum_{k=0}^{N-1} \tilde{x}(n-k)\tilde{y}(k)$. The periodic convolution is also a GPS with the same periodic type and period. On the other hand, the circular and skew-circular convolutions of two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ are defined by

$$(\boldsymbol{x} \textcircled{c} \boldsymbol{y})_n = \sum_{k=0}^{n} \boldsymbol{x}_k \boldsymbol{y}_{n-k} + \sum_{k=n+1}^{N-1} \boldsymbol{x}_k \boldsymbol{y}_{n-k+N}, \quad \text{for } n = 0, 1, \ldots, N-1. \quad (3)$$

$$(\boldsymbol{x} \textcircled{s} \boldsymbol{y})_n = \sum_{k=0}^{n} \boldsymbol{x}_k \boldsymbol{y}_{n-k} - \sum_{k=n+1}^{N-1} \boldsymbol{x}_k \boldsymbol{y}_{n-k+N}, \quad \text{for } n = 0, 1, \ldots, N-1. \quad (4)$$

The circular and skew-circular convolutions of two length-N sequences are, respectively, equivalent to the representative vectors of the periodic convolutions of the period-N strictly peri-

odic and anti-periodic sequences extended from the finite sequences.

D. Convolution-Multiplication Property of GDFT

The DFT has the convolution-multiplication property that the inverse transform after entry-wise multiplication gives the same result as the circular convolution of the original sequences. Vernet [26] and Martucci [27] derived such properties for other GDFTs. We summarize the results in matrix form as follows. Let $\boldsymbol{u} = \boldsymbol{x} \textcircled{c} \boldsymbol{y}$ and $\boldsymbol{w} = \boldsymbol{x} \textcircled{s} \boldsymbol{y}$, then the following properties hold:

$$\boldsymbol{u} = \boldsymbol{G}_{0,0}^{-1}[(\boldsymbol{G}_{0,0}\boldsymbol{x}) \circ (\boldsymbol{G}_{0,0}\boldsymbol{y})] \tag{5}$$

$$\boldsymbol{u} = \boldsymbol{G}_{0,\frac{1}{2}}^{-1}\left[\left(\boldsymbol{G}_{0,\frac{1}{2}}\boldsymbol{x}\right) \circ (\boldsymbol{G}_{0,0}\boldsymbol{y})\right] \tag{6}$$

$$\boldsymbol{w} = \boldsymbol{G}_{\frac{1}{2},0}^{-1}\left[\left(\boldsymbol{G}_{\frac{1}{2},0}\boldsymbol{x}\right) \circ \left(\boldsymbol{G}_{\frac{1}{2},0}\boldsymbol{y}\right)\right] \tag{7}$$

$$\boldsymbol{w} = \boldsymbol{G}_{\frac{1}{2},\frac{1}{2}}^{-1}\left[\left(\boldsymbol{G}_{\frac{1}{2},\frac{1}{2}}\boldsymbol{x}\right) \circ \left(\boldsymbol{G}_{\frac{1}{2},0}\boldsymbol{y}\right)\right]. \tag{8}$$

Notice that the implied periodicity of $\boldsymbol{u}_n$ is strictly periodic and that of $\boldsymbol{w}_n$ is anti-periodic due to the periodicity inherence of the periodic convolution of the original sequences represented by $\boldsymbol{x}$ and $\boldsymbol{y}$.

E. Discrete Trigonometric Transform

The family of DTTs comprises eight versions of the discrete cosine transform (DCT) and eight versions of the discrete sine transform (DST). Martucci formulated the DTTs through the convolution forms as defined in [27, Appendix]. The orthogonal-like relations between the inverse and forward DTTs are

$$\boldsymbol{T}_I^{-1} = \frac{1}{M}\boldsymbol{T}_I, \quad \boldsymbol{T}_{II}^{-1} = \frac{1}{M}\boldsymbol{T}_{III}, \quad \boldsymbol{T}_{III}^{-1} = \frac{1}{M}\boldsymbol{T}_{II}, \text{ and}$$
$$\boldsymbol{T}_{IV}^{-1} = \frac{1}{M}\boldsymbol{T}_{IV} \tag{9}$$

where the DTTs in both sides of each equality must be the same in the categories of cosine or sine and even or odd, and M is $2N$ and $2N-1$ for the even and odd cases, respectively.

F. DTT and GDFT

1) Symmetric-Extension Operator: Just as the special forms of the GDFT provide representations for GPSs, the symmetry and periodicity of the basis functions of the DTTs establish a one-to-one correspondence between the DTTs and the 16 symmetric-periodic sequences (SPSs) that are summarized in [27, Fig. 2]. Since these SPSs are also generalized-periodic, the SPS extended from a DTT of a finite sequence can be produced from the corresponding GDFT of that sequence after having been symmetrically extended to a base period. Therefore, each DTT can be directly constructed in terms of its corresponding GDFT by cascading an appropriate symmetric-extension operator as defined in Table I, where the subscripts are in terms of Martucci's naming rules, including whole-sample symmetry (WS), whole-sample anti-symmetry (WA), half-sample symmetry (HS), and half-sample anti-symmetry (HA). The notations $\boldsymbol{I}_N$ and $\boldsymbol{J}_N$ mean the identity matrix and reversal matrix of order N. A reversal matrix is a square matrix whose nonzero entries are ones on the antidiagonal. For example, we have

TABLE I
MATRIX FORMS FOR SYMMETRIC EXTENSION OPERATOR

$M=2N$		$M=2N-1$	
$\boldsymbol{E}_{WSWS}$	$\begin{bmatrix} \boldsymbol{I}_{N+1} \\ \hline \boldsymbol{0}_{N-1\times 1} \vdots \boldsymbol{J}_{N-1} \vdots \boldsymbol{0}_{N-1\times 1} \end{bmatrix}$	$\boldsymbol{E}_{WSHS}$	$\begin{bmatrix} \boldsymbol{I}_N \\ \hline \boldsymbol{0}_{N-1\times 1} \vdots \boldsymbol{J}_{N-1} \end{bmatrix}$
$\boldsymbol{E}_{WAWA}$	$\begin{bmatrix} \boldsymbol{0}_{1\times N-1} \\ \hline \boldsymbol{I}_{N-1} \\ \hline \boldsymbol{0}_{1\times N-1} \\ \hline -\boldsymbol{J}_{N-1} \end{bmatrix}$	$\boldsymbol{E}_{WAHA}$	$\begin{bmatrix} \boldsymbol{0}_{1\times N-1} \\ \hline \boldsymbol{I}_{N-1} \\ \hline -\boldsymbol{J}_{N-1} \end{bmatrix}$
$\boldsymbol{E}_{HSHS}$ $\boldsymbol{E}_{HAHS}$	$\begin{bmatrix} \boldsymbol{I}_N \\ \hline \boldsymbol{J}_N \end{bmatrix}$	$\boldsymbol{E}_{HSWS}$ $\boldsymbol{E}_{HAWS}$	$\begin{bmatrix} \boldsymbol{I}_N \\ \hline \boldsymbol{J}_{N-1} \vdots \boldsymbol{0}_{N-1\times 1} \end{bmatrix}$
$\boldsymbol{E}_{HAHA}$ $\boldsymbol{E}_{HSHA}$	$\begin{bmatrix} \boldsymbol{I}_N \\ \hline -\boldsymbol{J}_N \end{bmatrix}$	$\boldsymbol{E}_{HAWA}$ $\boldsymbol{E}_{HSWA}$	$\begin{bmatrix} \boldsymbol{I}_{N-1} \\ \hline \boldsymbol{0}_{1\times N-1} \\ \hline -\boldsymbol{J}_{N-1} \end{bmatrix}$
$\boldsymbol{E}_{WSWA}$	$\begin{bmatrix} \boldsymbol{I}_N \\ \hline \boldsymbol{0}_{1\times N} \\ \hline \boldsymbol{0}_{N-1\times 1} \vdots -\boldsymbol{J}_{N-1} \end{bmatrix}$	$\boldsymbol{E}_{WSHA}$	$\begin{bmatrix} \boldsymbol{I}_N \\ \hline \boldsymbol{0}_{N-1\times 1} \vdots -\boldsymbol{J}_{N-1} \end{bmatrix}$
$\boldsymbol{E}_{WAWS}$	$\begin{bmatrix} \boldsymbol{0}_{1\times N} \\ \hline \boldsymbol{I}_N \\ \hline \boldsymbol{J}_{N-1} \vdots \boldsymbol{0}_{N-1\times 1} \end{bmatrix}$	$\boldsymbol{E}_{WAHS}$	$\begin{bmatrix} \boldsymbol{0}_{1\times N-1} \\ \hline \boldsymbol{I}_{N-1} \\ \hline \boldsymbol{J}_{N-1} \end{bmatrix}$

$\boldsymbol{y} = \boldsymbol{G}_{0,(1/2)}\boldsymbol{E}_{\mathrm{HSHS}}\,\boldsymbol{x} = \boldsymbol{C}_{II}^{e}\boldsymbol{x}$, where $\boldsymbol{y}$ is $P \times 1, \boldsymbol{G}_{0,(1/2)}$ is $P \times 2N, \boldsymbol{E}_{\mathrm{HSHS}}$ is $2N \times N, \boldsymbol{C}_{II}^{e}$ is $P \times N, \boldsymbol{x}$ is $N \times 1$, and positive integer P determines how long the output SPS is captured.

2) Zero-Padding and Selective Matrices: For capturing the standard index range of the DTT output, the transposed zero-padding matrix is introduced. The $(r+p+q)\times p$ zero-padding matrix $\boldsymbol{Z}_{r,p,q}$ is defined as $[\boldsymbol{0}_{p\times r}\ \boldsymbol{I}_p\ \boldsymbol{0}_{p\times q}]^T$. The name "zero-padding" comes from the fact that left-multiplying a length-p column vector $\boldsymbol{x}$ by $\boldsymbol{Z}_{r,p,q}$ is equivalent to padding $\boldsymbol{x}$ to the up and down by r and q zeros, respectively. On the contrary, left-multiplying a length-$(r+p+q)$ column vector $\boldsymbol{x}$ by $\boldsymbol{Z}_{r,p,q}^T$ is equivalent to selecting $\boldsymbol{x}_n$ for $r \leqq n \leqq p+r-1$. For instance, $\boldsymbol{Z}_{3,4,2}^T \cdot [\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_8]^T = [\boldsymbol{x}_3, \boldsymbol{x}_4, \boldsymbol{x}_5, \boldsymbol{x}_6]^T$. Hence, we name the transposed zero-padding matrix as the "selective" matrix.

3) Relationship of DTT and GDFT: By using the selective matrix and symmetric-extension operator defined above, we can express the relation between a DTT matrix and its corresponding GDFT matrix as

$$\boldsymbol{T}_q = \boldsymbol{Z}_q^T \cdot \boldsymbol{F}_q \cdot \boldsymbol{E}_q \tag{10}$$

where $\boldsymbol{T}_q$ denotes the DTT matrix, $\boldsymbol{F}_q$ denotes the GDFT matrix, $\boldsymbol{Z}_q^T$ denotes the selective matrix, $\boldsymbol{E}_q$ denotes the symmetric- extension operator, and subscript q indicates the type of DTT, which takes on I, II, III, and IV. Alternatively, by left-multiplying the DTT matrix by a symmetric-extension operator to obtain another half of samples in the base period of the corresponding GDFT, we can define the correspondent symmetric-extension operator $\boldsymbol{E}'_q$ through the following relation:

$$\boldsymbol{E}'_q \cdot \boldsymbol{T}_q = \boldsymbol{F}_q \cdot \boldsymbol{E}_q. \tag{11}$$

TABLE II
DEFINITIONS OF RELATED MATRICES FOR DTT

	q	T_q	Input Index Range	Z_q	F_q	E_q	E'_q	A_q^+	$W_q^{'+}$	K_q	$r_{Freq.}$
$M=2N$	I	C_I^e	$0\to N$	$Z_{0,N+1,N-1}$	$G_{0,0}$	E_{WSWS}	E_{WSWS}	A_I^{e+}	W_{N+1}^{-II}	$G_{0,0}$	C
	II	C_{II}^e	$0\to N-1$	$Z_{0,N,N}$	$G_{0,\frac{1}{2}}$	E_{HSHS}	E_{WSWA}	A_{II}^{e+}	W_N^{-I}	$G_{0,\frac{1}{2}}$	S
	III	C_{III}^e	$0\to N-1$	$Z_{0,N,N}$	$G_{\frac{1}{2},0}$	E_{WSWA}	E_{HSHS}	A_{III}^{e+}	$2I_N$	$G_{0,0}$	C
	IV	C_{IV}^e	$0\to N-1$	$Z_{0,N,N}$	$G_{\frac{1}{2},\frac{1}{2}}$	E_{HSHA}	E_{HSHA}	A_{IV}^{e+}	$2I_N$	$G_{0,\frac{1}{2}}$	S
	I	S_I^e	$1\to N-1$	$Z_{1,N-1,N}$	$jG_{0,0}$	E_{WAWA}	E_{WAWA}	A_I^{e+}	$2I_{N-1}$	$G_{0,0}$	C
	II	S_{II}^e	$1\to N$	$Z_{0,N,N}$	$jG_{0,\frac{1}{2}}$	E_{HAHA}	E_{WAWS}	A_{II}^{e+}	W_N^{-I}	$G_{0,\frac{1}{2}}$	S
	III	S_{III}^e	$0\to N-1$	$Z_{1,N,N-1}$	$jG_{\frac{1}{2},0}$	E_{WAWS}	E_{HAHA}	A_{III}^{e+}	$2I_N$	$G_{0,0}$	C
	IV	S_{IV}^e	$0\to N-1$	$Z_{0,N,N}$	$jG_{\frac{1}{2},\frac{1}{2}}$	E_{HAHS}	E_{HAHS}	A_{IV}^{e+}	$2I_N$	$G_{0,\frac{1}{2}}$	S
$M=2N-1$	I	C_I^o	$0\to N-1$	$Z_{0,N,N-1}$	$G_{0,0}$	E_{WSHS}	E_{WSHS}	A_I^{o+}	W_N^{-I}	$G_{0,0}$	C
	II	C_{II}^o	$0\to N-1$	$Z_{0,N,N-1}$	$G_{0,\frac{1}{2}}$	E_{HSWS}	E_{WSHA}	A_{II}^{o+}	W_N^{-I}	$G_{0,\frac{1}{2}}$	S
	III	C_{III}^o	$0\to N-1$	$Z_{0,N,N-1}$	$G_{\frac{1}{2},0}$	E_{WSHA}	E_{HSWS}	A_{III}^{o+}	W_N^{-III}	$G_{0,0}$	C
	IV	C_{IV}^o	$0\to N-2$	$Z_{0,N-1,N}$	$G_{\frac{1}{2},\frac{1}{2}}$	E_{HSWA}	E_{HSWA}	A_{IV}^{o+}	$2I_{N-1}$	$G_{0,\frac{1}{2}}$	S
	I	S_I^o	$1\to N-1$	$Z_{1,N-1,N-1}$	$jG_{0,0}$	E_{WAHA}	E_{WAHA}	A_I^{o+}	$2I_{N-1}$	$G_{0,0}$	C
	II	S_{II}^o	$1\to N-1$	$Z_{0,N-1,N}$	$jG_{0,\frac{1}{2}}$	E_{HAWA}	E_{WAHS}	A_{II}^{o+}	W_{N-1}^{-I}	$G_{0,\frac{1}{2}}$	S
	III	S_{III}^o	$0\to N-2$	$Z_{1,N-1,N-1}$	$jG_{\frac{1}{2},0}$	E_{WAHS}	E_{HAWA}	A_{III}^{o+}	W_{N-1}^{-III}	$G_{0,0}$	C
	IV	S_{IV}^o	$0\to N-1$	$Z_{0,N,N-1}$	$jG_{\frac{1}{2},\frac{1}{2}}$	E_{HAWS}	E_{HAWS}	A_{IV}^{o+}	W_N^{-III}	$G_{0,\frac{1}{2}}$	S

4) Relationship of IDTT and IGDFT: The dual formula related to (11) is derived as follows:

$$\begin{aligned}\boldsymbol{E}_q\boldsymbol{T}_q^{-1} &= \boldsymbol{F}_q^{-1}(\boldsymbol{F}_q\boldsymbol{E}_q)\boldsymbol{T}_q^{-1} = \boldsymbol{F}_q^{-1}(\boldsymbol{E}'_q\boldsymbol{T}_q)\boldsymbol{T}_q^{-1}\\ &= \boldsymbol{F}_q^{-1}\boldsymbol{E}'_q. \end{aligned} \tag{12}$$

Note that $\boldsymbol{E}_q$ and $\boldsymbol{E}'_q$ are interchanged in (11) and (12). Taking conjugate of (12) and using the properties (2) and (9) lead to $\boldsymbol{E}_q\boldsymbol{T}_{\Phi(q)} = \boldsymbol{F}_{\Phi(q)}\boldsymbol{E}'_q$, and hence $\boldsymbol{E}_q = \boldsymbol{E}'_{\Phi(q)}$, where subscript $\Phi(q)$ indicates the pair type according to (9) (i.e., $\Phi(I) = I, \Phi(II) = III, \Phi(III) = II$, and $\Phi(IV) = IV$). On the other hand, by (2) and (9), the dual formula related to (10) is derived as follows:

$$\begin{aligned}\boldsymbol{T}_q^{-1} &= \frac{1}{M}\boldsymbol{T}_{\Phi(q)} = \frac{1}{M}\boldsymbol{Z}_{\Phi(q)}^T\boldsymbol{F}_{\Phi(q)}^*\boldsymbol{E}_{\Phi(q)}\\ &= \boldsymbol{Z}_{\Phi(q)}^T\boldsymbol{F}_q^{-1}\boldsymbol{E}_{\Phi(q)} = \boldsymbol{Z}_{\Phi(q)}^T\boldsymbol{F}_q^{-1}\boldsymbol{E}'_q. \end{aligned} \tag{13}$$

Since $\boldsymbol{T}_{\Phi(q)}$ is real-valued, the conjugate operation can be applied in the second equality. In the above generic formulas, the specific definitions of the related matrices are given in the first six columns of Tables II and III. As an example, according to (10) and Table II, the relation between the odd DST-II of length $N-1$ and the OTDFT of length M that equals $2N-1$ is given by $\boldsymbol{S}_{II}^o\boldsymbol{x} = \boldsymbol{Z}_{0,N-1,N}^T \cdot j\boldsymbol{G}_{0,(1/2)} \cdot \boldsymbol{E}_{\mathrm{HAWA}}\boldsymbol{x} = \boldsymbol{y}$, where $\boldsymbol{x} = [x(1), x(2), \ldots, x(N-1)]^T$ and $\boldsymbol{y} = [y(0), y(1), \ldots, y(N-2)]^T$. According to (11) and Table II, another relation is expressed by $\boldsymbol{E}_{\mathrm{WAHS}}\boldsymbol{S}_{II}^o\boldsymbol{x} = j\boldsymbol{G}_{0,(1/2)}\boldsymbol{E}_{\mathrm{HAWA}}\boldsymbol{x}$. Moreover, in [28], Britanak and Rao presented the sparse matrix factorizations of GDFTs to DCTs and DSTs. These factorizations are useful to visually understand how the relationship of DTTs and GDFTs hold.

III. AUTOREGRESSIVE MODELING AND GDFT

In this section, we first establish the time–frequency relation between the periodic autocorrelation and power spectrum in the GDFT frequency domain. Then we show that, like the traditional approach, the Yule–Walker equations consisting of periodic autocorrelations are derived in the least square error sense for evaluating the AR parameters in the finite length problems with GDFTs.

Before proceeding, we show a general property of two GPSs with period N, which will be heavily used in later derivation.

Lemma 1: Given two GPSs, $\tilde{x}(n)$ and $\tilde{y}(n)$, which are either strictly periodic or anti-periodic with period N. Then any summation over successive N terms of their product is equal to the summation over the base period from 0 to $N-1$. That is, $\sum_{n=i}^{i+N-1}\tilde{x}(n)\tilde{y}(n) = \sum_{n=0}^{N-1}\tilde{x}(n)\tilde{y}(n)$, for any integer i.

Proof: Since $\tilde{x}(n+N)\tilde{y}(n+N) = \tilde{x}(n)\tilde{y}(n)$, the product $\tilde{x}(n)\tilde{y}(n)$ is a strictly periodic sequence with period N. Thus, any summation over successive N terms of their product has the same result. ■

A. Autocorrelation and Power Spectrum in GDFT

The periodic correlation of two period-N GPSs $\tilde{x}(n)$ and $\tilde{y}(n)$ of the same type is defined as $\tilde{x}(n) * \tilde{y}(n) = (1/N)\sum_{m=0}^{N-1}\tilde{x}^*(m)\tilde{y}(n+m)$. Note that the periodic correlation is also a period-N GPS that has the same periodic type as the input GPSs. Similarly, to distinguish the strictly

TABLE III
DEFINITIONS OF RELATED MATRICES FOR IDTT

	q	T_q^{-1}	Input Index Range	$Z_{\Phi(q)}$	F_q^{-1}	E_q	E_q'	A_q^-	$W_q'^-$	Γ_q^{-1}	r_{Time}
$M=2N$	I	$C_I^{e\,-1}$	$0\to N$	$Z_{0,N+1,N-1}$	$G_{0,0}^{-1}$	E_{WSWS}	E_{WSWS}	A_I^{e-}	W_{N+1}^{II}	$G_{0,0}^{-1}$	c
	II	$C_{II}^{e\,-1}$	$0\to N-1$	$Z_{0,N,N}$	$G_{0,\frac{1}{2}}^{-1}$	E_{HSHS}	E_{WSWA}	A_{II}^{e-}	$2I_N$	$G_{0,0}^{-1}$	c
	III	$C_{III}^{e\,-1}$	$0\to N-1$	$Z_{0,N,N}$	$G_{\frac{1}{2},0}^{-1}$	E_{WSWA}	E_{HSHS}	A_{III}^{e-}	W_N^I	$G_{\frac{1}{2},0}^{-1}$	s
	IV	$C_{IV}^{e\,-1}$	$0\to N-1$	$Z_{0,N,N}$	$G_{\frac{1}{2},\frac{1}{2}}^{-1}$	E_{HSHA}	E_{HSHA}	A_{IV}^{e-}	$2I_N$	$G_{\frac{1}{2},0}^{-1}$	s
	I	$S_I^{e\,-1}$	$1\to N-1$	$Z_{1,N-1,N}$	$-jG_{0,0}^{-1}$	E_{WAWA}	E_{WAWA}	A_I^{e-}	$2I_{N-1}$	$G_{0,0}^{-1}$	c
	II	$S_{II}^{e\,-1}$	$0\to N-1$	$Z_{1,N,N-1}$	$-jG_{0,\frac{1}{2}}^{-1}$	E_{HAHA}	E_{WAWS}	A_{II}^{e-}	$2I_N$	$G_{0,0}^{-1}$	c
	III	$S_{III}^{e\,-1}$	$1\to N$	$Z_{0,N,N}$	$-jG_{\frac{1}{2},0}^{-1}$	E_{WAWS}	E_{HAHA}	A_{III}^{e-}	W_N^I	$G_{\frac{1}{2},0}^{-1}$	s
	IV	$S_{IV}^{e\,-1}$	$0\to N-1$	$Z_{0,N,N}$	$-jG_{\frac{1}{2},\frac{1}{2}}^{-1}$	E_{HAHS}	E_{HAHS}	A_{IV}^{e-}	$2I_N$	$G_{\frac{1}{2},0}^{-1}$	s
$M=2N-1$	I	$C_I^{o\,-1}$	$0\to N-1$	$Z_{0,N,N-1}$	$G_{0,0}^{-1}$	E_{WSHS}	E_{WSHS}	A_I^{o-}	W_N^I	$G_{0,0}^{-1}$	c
	II	$C_{II}^{o\,-1}$	$0\to N-1$	$Z_{0,N,N-1}$	$G_{0,\frac{1}{2}}^{-1}$	E_{HSWS}	E_{WSHA}	A_{II}^{o-}	W_N^{III}	$G_{0,0}^{-1}$	c
	III	$C_{III}^{o\,-1}$	$0\to N-1$	$Z_{0,N,N-1}$	$G_{\frac{1}{2},0}^{-1}$	E_{WSHA}	E_{HSWS}	A_{III}^{o-}	W_N^I	$G_{\frac{1}{2},0}^{-1}$	s
	IV	$C_{IV}^{o\,-1}$	$0\to N-2$	$Z_{0,N-1,N}$	$G_{\frac{1}{2},\frac{1}{2}}^{-1}$	E_{HSWA}	E_{HSWA}	A_{IV}^{o-}	$2I_{N-1}$	$G_{\frac{1}{2},0}^{-1}$	s
	I	$S_I^{o\,-1}$	$1\to N-1$	$Z_{1,N-1,N-1}$	$-jG_{0,0}^{-1}$	E_{WAHA}	E_{WAHA}	A_I^{o-}	$2I_{N-1}$	$G_{0,0}^{-1}$	c
	II	$S_{II}^{o\,-1}$	$0\to N-2$	$Z_{1,N-1,N-1}$	$-jG_{0,\frac{1}{2}}^{-1}$	E_{HAWA}	E_{WAHS}	A_{II}^{o-}	W_{N-1}^{III}	$G_{0,0}^{-1}$	c
	III	$S_{III}^{o\,-1}$	$1\to N-1$	$Z_{0,N-1,N}$	$-jG_{\frac{1}{2},0}^{-1}$	E_{WAHS}	E_{HAWA}	A_{III}^{o-}	W_{N-1}^I	$G_{\frac{1}{2},0}^{-1}$	s
	IV	$S_{IV}^{o\,-1}$	$0\to N-1$	$Z_{0,N,N-1}$	$-jG_{\frac{1}{2},\frac{1}{2}}^{-1}$	E_{HAWS}	E_{HAWS}	A_{IV}^{o-}	W_N^{III}	$G_{\frac{1}{2},0}^{-1}$	s

periodic and anti-periodic cases, the circular and skew-circular correlations of two length-N vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ are defined as

$$(\boldsymbol{x}_C^* \boldsymbol{y})_n = \frac{1}{N}\left[\sum_{k=0}^{N-n-1} \boldsymbol{x}_k^* \cdot \boldsymbol{y}_{n+k} + \sum_{k=N-n}^{N-1} \boldsymbol{x}_k^* \cdot \boldsymbol{y}_{n+k-N}\right], \quad \text{for } n = 0, 1, \ldots, N-1. \tag{14}$$

$$(\boldsymbol{x}_S^* \boldsymbol{y})_n = \frac{1}{N}\left[\sum_{k=0}^{N-n-1} \boldsymbol{x}_k^* \cdot \boldsymbol{y}_{n+k} - \sum_{k=N-n}^{N-1} \boldsymbol{x}_k^* \cdot \boldsymbol{y}_{n+k-N}\right], \quad \text{for } n = 0, 1, \ldots, N-1. \tag{15}$$

The circular and skew-circular correlations of two length-N vectors are, respectively, equivalent to the representative vectors of the periodic correlations of the period-N strictly periodic and anti-periodic sequences extended from the finite vectors.

To express a periodic correlation in terms of a periodic convolution, the flip operation on a GPS is introduced and defined as $\mathrm{FLIP}\{\tilde{x}(n)\} = \tilde{x}(-n)$. The flip operation (also referred to as time or frequency reversal) can also preserve the periodicity of the input GPS. For finite sequences of length N, the strict-flip and anti-flip operations are defined in matrix form as

$$\boldsymbol{S}_{\mathrm{FLIP}} = \left[\begin{array}{c|c} 1 & \boldsymbol{0}_{1\times(N-1)} \\ \hline \boldsymbol{0}_{(N-1)\times 1} & \boldsymbol{J}_{N-1} \end{array}\right]$$

and

$$\boldsymbol{A}_{\mathrm{FLIP}} = \left[\begin{array}{c|c} 1 & \boldsymbol{0}_{1\times(N-1)} \\ \hline \boldsymbol{0}_{(N-1)\times 1} & -\boldsymbol{J}_{N-1} \end{array}\right]. \tag{16}$$

The strict-flip and anti-flip operations of a length-N vector are equivalent to the representative vectors of the flip operations of the period-N strictly periodic and anti-periodic sequences extended from the finite vector. For instance, let $\boldsymbol{x} = [1, 2, 3, 4]^T$, then $\boldsymbol{S}_{\mathrm{FLIP}}\boldsymbol{x} = [1, 4, 3, 2]^T$ and $\boldsymbol{A}_{\mathrm{FLIP}}\boldsymbol{x} = [1, -4, -3, -2]^T$.

Theorem 1: Given two period-N GPSs, $\tilde{x}(n)$ and $\tilde{y}(n)$, of the same periodic type. Then,

$$\tilde{x}(n) * \tilde{y}(n) = \frac{1}{N}\mathrm{FLIP}\{\tilde{x}^*(n)\} \otimes \tilde{y}(n).$$

Proof: Since both $\tilde{x}^*(m)$ and $\tilde{y}(n+m)$ for any fixed n are either strictly periodic or anti-periodic with period N, by Lemma 1, we have

$$\sum_{m=0}^{N-1} \tilde{x}^*(m)\tilde{y}(n+m) = \sum_{m=0}^{N-1} \tilde{x}^*(-m)\tilde{y}(n-m).$$

■

Corollary 1: For two column vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ of length N, the following properties hold:

$$\boldsymbol{x}_C^* \boldsymbol{y} = \frac{1}{N}(\boldsymbol{S}_{\mathrm{FLIP}}\boldsymbol{x}^*) \textcircled{c}\, \boldsymbol{y}, \text{ and}$$
$$\boldsymbol{x}_S^* \boldsymbol{y} = \frac{1}{N}(\boldsymbol{A}_{\mathrm{FLIP}}\boldsymbol{x}^*) \textcircled{s}\, \boldsymbol{y}. \tag{17}$$

■

Like the conjugate relation between the DFTs of a vector and its strict-flipped conjugate [29], we extend without proof such properties for other GDFTs in the next lemma.

Lemma 2: Consider a column vectors $\boldsymbol{x}$ of length N.

1) The DFT and OTDFT of the strict-flipped conjugated $\boldsymbol{x}$ can be evaluated by

$$(\boldsymbol{G}_{0,b}\boldsymbol{S}_{\mathrm{FLIP}}\boldsymbol{x}^*)_k = \exp(-j4\pi kb/N) \cdot (\boldsymbol{G}_{0,b}\boldsymbol{x})_k^*, \quad \text{for } b = 0 \text{ or } 1/2, \text{ and } k = 0, 1, \ldots, N-1. \tag{18}$$

2) The OFDFT and O^{2}DFT of the anti-flipped conjugated $\boldsymbol{x}$ can be evaluated by

$$\left(\boldsymbol{G}_{\frac{1}{2},b}\boldsymbol{A}_{\mathrm{FLIP}}\boldsymbol{x}^*\right)_k = \exp\left(-j4\pi\left(k+\frac{1}{2}\right)b\Big/N\right)\cdot\left(\boldsymbol{G}_{\frac{1}{2},b}\boldsymbol{x}\right)^*_k, \quad \text{for } b=0 \text{ or } 1/2, \text{ and } k=0,1,\ldots,N-1. \tag{19}$$

■

The periodic autocorrelation of a GPS $\tilde{x}(n)$ with period N is defined as

$$r_{\tilde{x}}(n) = \tilde{x}(n) * \tilde{x}(n) = \frac{1}{N}\sum_{m=0}^{N-1}\tilde{x}^*(m)\tilde{x}(n+m). \tag{20}$$

For a vector $\boldsymbol{x}$ of length N, the circular and skew-circular autocorrelations are defined as $\boldsymbol{r}^C_{\boldsymbol{x}} = \boldsymbol{x}^*_C\boldsymbol{x}$ and $\boldsymbol{r}^S_{\boldsymbol{x}} = \boldsymbol{x}^*_S\boldsymbol{x}$. It is well known that the DFT of the circular autocorrelation of a vector equals the DFT power spectrum of the vector. We extend such important relations for other GDFTs and GPSs.

Theorem 2: Consider a column vector $\boldsymbol{x}$ of length N.

1) The relation between the circular autocorrelation and DFT/OTDFT power spectra is given by

$$\begin{aligned}\boldsymbol{r}^C_{\boldsymbol{x}} &= \frac{1}{N}\boldsymbol{G}^{-1}_{0,0}[(\boldsymbol{G}_{0,0}\boldsymbol{x})\circ(\boldsymbol{G}_{0,0}\boldsymbol{x})^*] \\ &= \frac{1}{N}\boldsymbol{G}^{-1}_{0,0}\left[\left(\boldsymbol{G}_{0,\frac{1}{2}}\boldsymbol{x}\right)\circ\left(\boldsymbol{G}_{0,\frac{1}{2}}\boldsymbol{x}\right)^*\right].\end{aligned} \tag{21}$$

2) The relation between the skew-circular autocorrelation and OFDFT/O^{2}DFT power spectra is given by

$$\begin{aligned}\boldsymbol{r}^S_{\boldsymbol{x}} &= \frac{1}{N}\boldsymbol{G}^{-1}_{\frac{1}{2},0}\left[\left(\boldsymbol{G}_{\frac{1}{2},0}\boldsymbol{x}\right)\circ\left(\boldsymbol{G}_{\frac{1}{2},0}\boldsymbol{x}\right)^*\right] \\ &= \frac{1}{N}\boldsymbol{G}^{-1}_{\frac{1}{2},0}\left[\left(\boldsymbol{G}_{\frac{1}{2},\frac{1}{2}}\boldsymbol{x}\right)\circ\left(\boldsymbol{G}_{\frac{1}{2},\frac{1}{2}}\boldsymbol{x}\right)^*\right].\end{aligned} \tag{22}$$

Proof: We first consider the case of skew-circular autocorrelation and OFDFT in part 2). By Corollary 1, we have $\boldsymbol{r}^S_{\boldsymbol{x}} = ((1/N)\boldsymbol{A}_{\mathrm{FLIP}}\boldsymbol{x}^*)\ \textcircled{S}\ \boldsymbol{x}$. Thus, using (5) yields $\boldsymbol{r}^S_{\boldsymbol{x}} = (1/N)\boldsymbol{G}^{-1}_{(1/2),0}[(\boldsymbol{G}_{(1/2),0}\boldsymbol{A}_{\mathrm{FLIP}}\boldsymbol{x}^*)\circ(\boldsymbol{G}_{(1/2),0}\boldsymbol{x})]$. Then, by Lemma 2 2), we obtain $\boldsymbol{r}^S_x = (1/N)\boldsymbol{G}^{-1}_{(1/2),0}[(\boldsymbol{G}_{(1/2),0}\boldsymbol{x})^*\circ(\boldsymbol{G}_{(1/2),0}\boldsymbol{x})]$. Due to $|(\boldsymbol{G}_{(1/2),0}\boldsymbol{x})_k| = |(\boldsymbol{G}_{(1/2),(1/2)}\boldsymbol{x})_k|$, part 2) is proved completely. Similarly, part 1) can be proved by the same method and using (7). ■

Corollary 2: Consider a column vector $\boldsymbol{y}$ of length N.

1) The relation between the skew-circular autocorrelation and IOTDFT/IO^2 DFT power spectra is given by

$$\begin{aligned}\boldsymbol{r}^S_{\boldsymbol{y}} &= \boldsymbol{G}_{0,\frac{1}{2}}\left[\left(\boldsymbol{G}^{-1}_{0,\frac{1}{2}}\boldsymbol{y}\right)\circ\left(\boldsymbol{G}^{-1}_{0,\frac{1}{2}}\boldsymbol{y}\right)^*\right] \\ &= \boldsymbol{G}_{0,\frac{1}{2}}\left[\left(\boldsymbol{G}^{-1}_{\frac{1}{2},\frac{1}{2}}\boldsymbol{y}\right)\circ\left(\boldsymbol{G}^{-1}_{\frac{1}{2},\frac{1}{2}}\boldsymbol{y}\right)^*\right].\end{aligned} \tag{23}$$

2) The relation between the circular autocorrelation and IDFT/IOFDFT power spectra is given by

$$\begin{aligned}\mathrm{r}^C_{\mathrm{y}} &= \mathrm{G}_{0,0}\left[\left(\mathrm{G}^{-1}_{0,0}\mathrm{y}\right)\circ\left(\mathrm{G}^{-1}_{0,0}\mathrm{y}\right)^*\right] \\ &= \mathrm{G}_{0,0}\left[\left(\mathrm{G}^{-1}_{\frac{1}{2},0}\mathrm{y}\right)\circ\left(\mathrm{G}^{-1}_{\frac{1}{2},0}\mathrm{y}\right)^*\right].\end{aligned} \tag{24}$$

Proof: We first represent (21) and (22) in Theorem 2 in generic form as

$$\mathrm{r}_{\mathrm{x}} = \frac{1}{N}\mathrm{G}^{-1}_{a,0}[(\mathrm{G}_{a,b}\mathrm{x})\circ(\mathrm{G}_{a,b}\mathrm{x})^*]. \tag{25}$$

By taking conjugate of both sides of (25) and using the property (2), we have

$$\begin{aligned}\boldsymbol{r}^*_{\boldsymbol{x}} &= \frac{1}{N^2}\boldsymbol{G}_{0,a}\left[\left(N\boldsymbol{G}^{-1}_{b,a}\boldsymbol{x}^*\right)\circ\left(N\boldsymbol{G}^{-1}_{b,a}\boldsymbol{x}^*\right)^*\right] \\ &= \boldsymbol{G}_{0,a}\left[\left(\boldsymbol{G}^{-1}_{b,a}\boldsymbol{x}^*\right)\circ\left(\boldsymbol{G}^{-1}_{b,a}\boldsymbol{x}^*\right)^*\right].\end{aligned} \tag{26}$$

Letting $\boldsymbol{x}$ be y^* in (26) yields $\boldsymbol{r}^*_{\boldsymbol{y}^*} = \boldsymbol{G}_{0,a}[(\boldsymbol{G}^{-1}_{b,a}\boldsymbol{y})\circ(\boldsymbol{G}^{-1}_{b,a}\boldsymbol{y})^*]$. Since $\boldsymbol{r}^*_{\boldsymbol{y}^*} = \boldsymbol{r}_{\boldsymbol{y}}$ in both circular and skew-circular autocorrelations, the proof is accomplished. ■

Thus far, the relation between the periodic autocorrelation and GDFT power spectra has been connected in Theorem 2 and Corollary 2. These results are the theoretical fundamental of AR modeling of GDFT spectra in later derivation.

B. AR Modeling With GDFT

An order-P AR model for a period-N GPS $\tilde{x}(n)$ with parameters $\alpha(1),\alpha(2),\ldots,\alpha(P)$ is defined as

$$\tilde{x}(n) = -\sum_{k=1}^{P}\alpha(k)\tilde{x}(n-k) + \tilde{e}(n) \tag{27}$$

where $\tilde{e}(n)$ is a generalized-periodic excitation sequence, which is regarded as residuals in linear prediction sense. The residual term can be written as

$$\tilde{e}(n) = \sum_{k=0}^{N-1}\alpha(k)\tilde{x}(n-k) = \tilde{x}(n)\otimes\tilde{\alpha}(n), \text{ where } \alpha(0)=1$$

and $\alpha(n)=0$ for $n=P{+}1,P{+}2,\ldots,N{-}1$. The periodic type of $\tilde{\alpha}(n)$ that is the GPS extended from $\alpha(n), n=0,1,\ldots,N-1$, must be the same as that of $\tilde{x}(n)$ to make the periodic convolution computable. Let $\mathrm{x},\alpha,\mathrm{e}$ denote the representative vectors for $\tilde{x}(n),\tilde{\alpha}(n)$, and $\tilde{e}(n)$, respectively. Then, by the convolution-multiplication properties (5)–(8), $\boldsymbol{e}$ can be transformed into the GDFT frequency domain as $\mathrm{G}_{a,b}\mathrm{e} = (\mathrm{G}_{a,b}\mathrm{x})\circ(\mathrm{G}_{a,0}\alpha)$, where a and b take on the values 0 or 1/2 depending on the periodicity type of $\tilde{x}(n)$ and the GDFT type used. Hence, by viewing $\tilde{e}(n)$ as a white noise, the power spectrum of the order-P AR model is given by

$$\begin{aligned}|(\boldsymbol{G}_{a,b}\boldsymbol{x})_k|^2 &= |(\boldsymbol{G}_{a,b}\boldsymbol{e})_k|^2|(\boldsymbol{G}_{a,0}\boldsymbol{\alpha})_k|^{-2} \\ &\approx \boldsymbol{\sigma}^2_e\cdot|(\boldsymbol{G}_{a,0}\boldsymbol{\alpha})_k|^{-2}, \\ &\qquad \text{for } k=0,1,\ldots,N-1.\end{aligned} \tag{28}$$

To evaluate the AR parameters, the Yule–Walker equations can be derived through the least square error (LSE) approach for the GDFT family. We present the result without proof.

Theorem 3: For a GPS $\tilde{x}(n)$ with period N, the parameters for the order-P AR modeling in the LSE criterion (i.e., minimize $\sum_{n=0}^{N-1}|\tilde{e}(n)|^2$) can be obtained by solving the Yule–Walker equations consisting of the periodic autocorrelations: $r_{\tilde{x}}(i) = -\sum_{n=1}^{P}r_{\tilde{x}}(i-n)\alpha(n)$, for $i=1,2,\ldots,P$. Furthermore, the LSE is given by $\sum_{n=0}^{N-1}|\tilde{e}(n)|^2 = N\cdot\sum_{n=0}^{P}r_{\tilde{x}}(-n)\alpha(n)$. ■

Remarkable, in both strictly periodic and anti-periodic cases, the Yule–Walker equations can be expressed in terms of a Toeplitz matrix, and hence the Levinson–Durbin algorithm [1]

can be used for efficiently computing AR parameters. Also, according to the next theorem, only $P+1$ periodic autocorrelation entries, $r_{\tilde{x}}(i), i = 0,1,\ldots,P$, are required to comprise the Yule–Walker equations.

Theorem 4: Let $\tilde{x}(n)$ be a GPS with period N. Then its periodic autocorrelation has the conjugate-symmetric property as $r_{\tilde{x}}(-n) = r^*_{\tilde{x}}(n)$.

Proof: Applying Lemma 1 to $\tilde{x}^*(m)$ and $\tilde{x}(-n+m)$ leads to

$$\begin{aligned} N \cdot r_{\tilde{x}}(-n) &= \sum_{m=0}^{N-1} \tilde{x}^*(m)\tilde{x}(-n+m) \\ &= \sum_{m=n}^{N+n-1} \tilde{x}^*(m)\tilde{x}(-n+m) \\ &= \sum_{m=0}^{N-1} \tilde{x}^*(m+n)\tilde{x}(m) = N \cdot r^*_{\tilde{x}}(n). \end{aligned}$$

■

The next theorem describes the well-known Parseval's Theorem [29] for GDFTs.

Theorem 5: Let $\boldsymbol{x}$ be a column vector of length N. Then $\|\boldsymbol{x}\|^2 = (1/N)\|\boldsymbol{G}_{a,b}\boldsymbol{x}\|^2 = N\|\boldsymbol{G}^{-1}_{a,b}\boldsymbol{x}\|^2$ for $a, b = 0, 1/2$. ■

To summarize, like the traditional AR modeling method, the Yule–Walker equations in Theorem 3 can be solved to yield AR parameters in finite length problems. Then $|(\boldsymbol{G}_{a,b}\boldsymbol{x})_k|^2$ can be estimated by $\boldsymbol{\sigma}_e^2|(\boldsymbol{G}_{a,0}\boldsymbol{\alpha})_k|^{-2}$. By Theorem 5, we have $\boldsymbol{\sigma}_e^2 \approx (1/N)\|\boldsymbol{G}_{a,b}\boldsymbol{e}\|^2 = \|\boldsymbol{e}\|^2$. Also, $|(\boldsymbol{G}_{a,0}\boldsymbol{x})_k| = |(\boldsymbol{G}_{a,(1/2)}\boldsymbol{x})_k|$. Thus, when viewing $\boldsymbol{x}$ in the strictly periodic sense, we can approximate its DFT and OTDFT power spectra by

$$\begin{aligned} |(\boldsymbol{G}_{0,0}\boldsymbol{x})_k|^2 &= \left|\left(\boldsymbol{G}_{0,\frac{1}{2}}\boldsymbol{x}\right)_k\right|^2 \approx \boldsymbol{\sigma}_e^2|(\boldsymbol{G}_{0,0}\boldsymbol{\alpha})_k|^{-2} \\ &= \boldsymbol{\sigma}_e^2\left|\left(\boldsymbol{G}_{0,\frac{1}{2}}\alpha\right)_k\right|^{-2}. \end{aligned}$$

Oppositely, when viewing $\boldsymbol{x}$ in the anti-periodic sense, we can approximate its OFDFT and O^2 DFT power spectra by

$$\begin{aligned} \left|\left(\boldsymbol{G}_{\frac{1}{2},0}\boldsymbol{x}\right)_k\right|^2 &= \left|\left(\boldsymbol{G}_{\frac{1}{2},\frac{1}{2}}\boldsymbol{x}\right)_k\right|^2 \approx \boldsymbol{\sigma}_e^2\left|\left(\boldsymbol{G}_{\frac{1}{2},0}\boldsymbol{\alpha}\right)_k\right|^{-2} \\ &= \boldsymbol{\sigma}_e^2\left|\left(\boldsymbol{G}_{\frac{1}{2},\frac{1}{2}}\boldsymbol{\alpha}\right)_k\right|^{-2}. \end{aligned}$$

Likewise, for the frequency-domain AR modeling, we can estimate the squared temporal envelope by

$$\begin{aligned} \left|\left(\boldsymbol{G}^{-1}_{0,b}\boldsymbol{y}\right)_n\right|^2 &= \left|\left(\boldsymbol{G}^{-1}_{\frac{1}{2},b}\boldsymbol{y}\right)_n\right|^2 \approx N^{-2}\boldsymbol{\sigma}_e^2\left|\left(\boldsymbol{G}^{-1}_{0,b}\boldsymbol{\alpha}\right)_n\right|^{-2} \\ &= N^{-2}\boldsymbol{\sigma}_e^2\left|\left(\boldsymbol{G}^{-1}_{\frac{1}{2},b}\boldsymbol{\alpha}\right)_n\right|^{-2} \end{aligned}$$

where $b = 0$ or 1/2 depending on the forward GDFT used, and $\boldsymbol{\sigma}_e^2 \approx (1/N)\|\boldsymbol{G}^{-1}_{a,b}\boldsymbol{e}\|^2 = (1/N^2)\|\boldsymbol{e}\|^2$. As an instance, Fig. 1 illustrates the spectral power envelopes of a speech segment of 2048 samples at 44.1 kHz. The power spectra obtained from its DFT and OFDFT are shown in Fig. 1(b) and (c), where the spectral power envelopes of order-24 AR modeling are obtained by solving the Yule–Walker equations consisting of the circular and skew-circular autocorrelations, respectively. The two spectral power envelopes are depicted together in Fig. 1(d) to compare their difference. In the low-frequency part, the two envelopes are almost identical, whereas the major deviation occurs in the high-frequency part and reveals the difference of the circular and skew-circular autocorrelations.

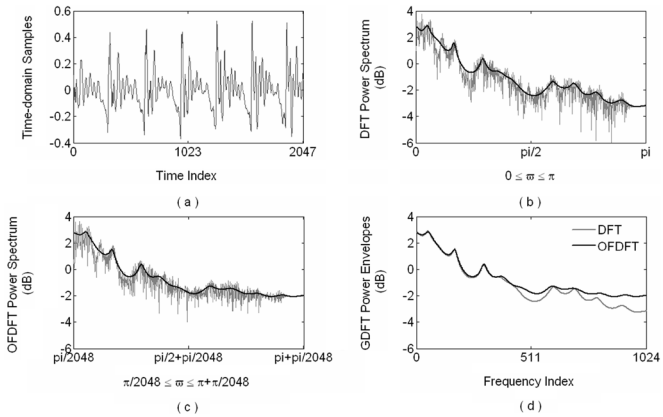

Fig. 1. Comparison of spectral power envelopes (i.e., squared envelope). (a) The time-domain speech segments of 2048 points at 44.1 kHz. (b) The DFT power spectrum and the spectral power envelope. (c) The OFDFT power spectrum and the spectral power envelope. Both the power envelopes are obtained by order-24 AR modeling, for which the Yule–Walker equations are comprised by the circular and skew-circular autocorrelations in (b) and in (c), respectively. (d) The comparison of the two power envelopes.

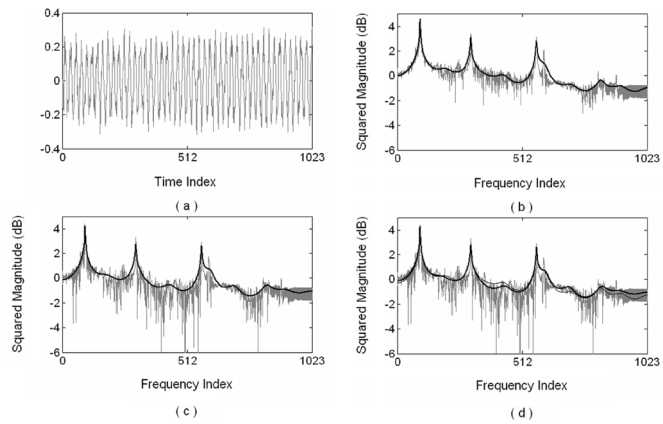

Fig. 2. Comparison of squared spectral envelopes. (a) The time-domain audio segment of 1024 samples at 44.1 kHz. (b) The squared analytic transform of the odd DST-IV spectrum and the squared Hilbert envelope through AR modeling. (c) The squared odd DST-IV spectrum and the power envelope. (d) The power.

IV. Autoregressive Modeling and DTT

In this section, we derive the theorems for the AR modeling in DTTs. First, associated with GDFTs, we derive the analytic transform matrices which convert real-valued vectors into analytic vectors. Then, based on the analytic vectors in DTT/IDTT, we derive the close form between DTT/IDTT with the GDFT/IGDFT. Combing the AR modeling in last section with the close form, we derive the AR modeling formulation with the DTTs in both temporal and spectral domains.

A. Analytic Transform Based on GDFT and IGDFT

Marple has proposed a DFT-based method for computing the analytic signal corresponding to a finite real-valued sequence of an even length [30]. The $N \times N$ analytic transform matrix $\boldsymbol{A}$ converting a real-valued vector $\boldsymbol{x}$ into a complex-valued analytic vector $\boldsymbol{a}$ is decomposed in matrix product form

$$\boldsymbol{A} = \boldsymbol{F}^{-1}\boldsymbol{Z}\boldsymbol{W}\boldsymbol{Z}^T\boldsymbol{F} \tag{29}$$

TABLE IV
DEFINITIONS OF RELATED MATRICES FOR ANALYTIC TRANSFORM

	A_q^+	F_q	Z_q^+	W_q^+	A_q^-	F_q^{-1}	Z_q^-	W_q^-
$M=2N$	A_I^{e+}	$G_{0,0}$	$Z_{0,N+1,N-1}$	W_{N+1}^{II}	A_I^{e-}	$G_{0,0}^{-1}$	$Z_{0,N+1,N-1}$	W_{N+1}^{II}
	A_{II}^{e+}	$G_{0,\frac{1}{2}}$		W_{N+1}^{II}	A_{II}^{e-}	$G_{0,\frac{1}{2}}^{-1}$	$Z_{0,N,N}$	$2I_N$
	A_{III}^{e+}	$G_{\frac{1}{2},0}$	$Z_{0,N,N}$	$2I_N$	A_{III}^{e-}	$G_{\frac{1}{2},0}^{-1}$	$Z_{0,N+1,N-1}$	W_{N+1}^{II}
	A_{IV}^{e+}	$G_{\frac{1}{2},\frac{1}{2}}$		$2I_N$	A_{IV}^{e-}	$G_{\frac{1}{2},\frac{1}{2}}^{-1}$	$Z_{0,N,N}$	$2I_N$
$M=2N-1$	A_I^{o+}	$G_{0,0}$	$Z_{0,N,N-1}$	W_N^I	A_I^{o-}	$G_{0,0}^{-1}$	$Z_{0,N,N-1}$	W_N^I
	A_{II}^{o+}	$G_{0,\frac{1}{2}}$		W_N^I	A_{II}^{o-}	$G_{0,\frac{1}{2}}^{-1}$		W_N^{III}
	A_{III}^{o+}	$G_{\frac{1}{2},0}$		W_N^{III}	A_{III}^{o-}	$G_{\frac{1}{2},0}^{-1}$		W_N^I
	A_{IV}^{o+}	$G_{\frac{1}{2},\frac{1}{2}}$		W_N^{III}	A_{IV}^{o-}	$G_{\frac{1}{2},\frac{1}{2}}^{-1}$		W_N^{III}

where $\boldsymbol{F}$ denotes the DFT matrix, $\boldsymbol{Z}$ denotes the zero-padding matrix $\boldsymbol{Z}_{0,N/2+1,N/2-1}$, and $\boldsymbol{W}$ is the weighting matrix $\mathrm{diag}\{1,2,2,\ldots,2,1\}$ of order $N/2+1$. As can be seen, the analytic transform proposed by Marple discards the negative DFT frequencies. With the appropriate weighting by $\boldsymbol{W}$, the analytic vector can have two desired properties. First, the real part of $\boldsymbol{a}$ exactly equals the original vector

$$\mathrm{Re}(\boldsymbol{a}_n) = \boldsymbol{x}_n, \quad \text{for } n = 0,1,\ldots,N-1. \tag{30}$$

Second, the real and imaginary parts of $\boldsymbol{a}$ are orthogonal:

$$\sum_{n=0}^{N-1} \mathrm{Re}(\boldsymbol{a}_n)\cdot \mathrm{Im}(\boldsymbol{a}_n) = 0. \tag{31}$$

As the fundamental for establishing the analytic transform via the GDFT, we show without proof the generalized symmetry of the GDFT of a real-valued input.

Lemma 3: Given a column vector $\boldsymbol{x}$ of length M, and $\boldsymbol{y} = \boldsymbol{G}_{a,b}\boldsymbol{x}$. Then $\boldsymbol{x}$ is real-valued if and only if the following conjugate symmetric/anti-symmetric property of $\boldsymbol{y}$ holds.

1) For DFT and OTDFT, i.e., $a = 0$,

$$\boldsymbol{y}_{\lceil M/2\rceil+k} = s\cdot \boldsymbol{y}^*_{\lfloor M/2\rfloor-k}, \quad \text{for } 0\le k\le \lfloor \boldsymbol{M}/2\rfloor. \tag{32}$$

2) For OFDFT and O^{2}DFT, i.e., $a = 1/2$,

$$\boldsymbol{y}_{\lceil M/2\rceil+k-1} = s\cdot \boldsymbol{y}^*_{\lfloor M/2\rfloor-k}, \quad \text{for } 0\le k\le \lfloor \boldsymbol{M}/2\rfloor. \tag{33}$$

where $s = 1$ when $b = 0$ and $s = -1$ when $b = 1/2$. ■

Based on the generalized symmetry of GDFTs, we can construct other analytic transform matrices as shown in the next theorem.

Theorem 6: Via each GDFT, we can define the analytic transform matrix, which satisfies both the properties (30) and (31), in the generic form

$$\boldsymbol{A}_q^+ = \boldsymbol{F}_q^{-1}\boldsymbol{Z}_q^+\boldsymbol{W}_q^+(\boldsymbol{Z}_q^+)^T\boldsymbol{F}_q \tag{34}$$

where $\boldsymbol{A}_q^+$ is the $M\times M$ analytic transform matrix, $\boldsymbol{F}_q$ is the GDFT matrix, $\boldsymbol{Z}_q^+$ is the zero-padding matrix, and $\boldsymbol{W}_q^+$ is the weighting matrix. The specific matrices are tabulated in Table IV, where $\boldsymbol{W}_q^+$ belongs to one of the following diagonal matrices of order N or $N+1$ denoted as subscripts: $2\boldsymbol{I}_N = \mathrm{diag}\{2,2,\ldots,2\}$, $\boldsymbol{W}_N^I = \mathrm{diag}\{1,2,2,\ldots,2\}$, $\boldsymbol{W}_{N+1}^{II} = \mathrm{diag}\{1,2,2,\ldots,2,1\}$, and $\boldsymbol{W}_N^{III} = \mathrm{diag}\{2,2,\ldots,2,1\}$.

Proof: In the same approach in [30], we can construct the analytic vector by defining the conjugate symmetric and anti-symmetric functions. For example, we consider the construction of odd $\boldsymbol{A}_{III}^+$. Let $\boldsymbol{y} = \boldsymbol{G}_{(1/2),0}\boldsymbol{x}$ and $\boldsymbol{a} = \boldsymbol{a}^{\mathrm{Re}} + j\boldsymbol{a}^{\mathrm{Im}} = \boldsymbol{A}_{III}^{o+}\boldsymbol{x}$ for a real-valued vector $\boldsymbol{x}$ of length M, where $\boldsymbol{a}^{\mathrm{Re}}$ and $\boldsymbol{a}^{\mathrm{Im}}$ denote the real and imaginary parts of $\boldsymbol{a}$. Let $\boldsymbol{z}^{\mathrm{Re}}$ and $\boldsymbol{z}^{\mathrm{Im}}$ denote the OFDFTs of $\boldsymbol{a}^{\mathrm{Re}}$ and $\boldsymbol{a}^{\mathrm{Im}}$, respectively. For the condition (30), we must define $\boldsymbol{z}^{\mathrm{Re}} = \boldsymbol{y}$ to have $\boldsymbol{a}^{\mathrm{Re}} = \boldsymbol{x}$. On the other hand, we define

$$(\boldsymbol{z}^{\mathrm{Im}})_k = \begin{cases} -j\cdot \boldsymbol{y}_k & 0\le k\le (M-1)/2 \\ 0 & k = (M+1)/2 \\ j\cdot \boldsymbol{y}_k & (M+3)/2\le k\le M-1 \end{cases}. \tag{35}$$

Note that adding $\boldsymbol{z}^{\mathrm{Re}}$ to $j\boldsymbol{z}^{\mathrm{Im}}$ not only eliminates the negative spectrum but also leads to the definition of $\boldsymbol{W}_{III}^{o+}$ as $\mathrm{diag}\{2,2,\ldots,2,1\}$ of order $(M+1)/2$. The definition (35) indeed implies that $\boldsymbol{a}^{\mathrm{Im}} = \boldsymbol{G}_{(1/2),0}^{-1}\boldsymbol{z}^{\mathrm{Im}}$ is purely real-valued since the conjugate symmetric property (33) holds; thus, the condition (30) is satisfied. The orthogonal property (31) can be confirmed as follows:

$$\begin{aligned}(\boldsymbol{a}^{\mathrm{Re}})^T\boldsymbol{a}^{\mathrm{Im}} &= \frac{1}{M}\left(\boldsymbol{G}_{\frac{1}{2},0}\boldsymbol{a}^{\mathrm{Re}}\right)^H\left(\boldsymbol{G}_{\frac{1}{2},0}\boldsymbol{a}^{\mathrm{Im}}\right)\\ &= \frac{1}{M}(\boldsymbol{z}^{\mathrm{Re}})^H\boldsymbol{z}^{\mathrm{Im}} = 0.\end{aligned}$$

In the last step, the conjugate symmetric property of $\boldsymbol{y}$ is used. Other analytic transform matrices can be constructed in the same way. ■

In the next corollary, we show the dual formula of (34).

Corollary 3: Let $\boldsymbol{A}_q^- = (\boldsymbol{A}_{\Phi(q)}^+)^*$, $\boldsymbol{Z}_q^- = \boldsymbol{Z}_{\Phi(q)}^+$, and $\boldsymbol{W}_q^- = \boldsymbol{W}_{\Phi(q)}^+$. We can define the analytic transform matrix which converts a real-valued spectral vector into a spectral analytic vector as

$$\boldsymbol{A}_q^- = \boldsymbol{F}_q\boldsymbol{Z}_q^-\boldsymbol{W}_q^-(\boldsymbol{Z}_q^-)^T\boldsymbol{F}_q^{-1}. \tag{36}$$

Proof: We take conjugate of (34) and use the property (2) to have $(\boldsymbol{A}_q^+)^* = \boldsymbol{F}_{\Phi(q)}\boldsymbol{Z}_q^+\boldsymbol{W}_q^+(\boldsymbol{Z}_q^+)^T\boldsymbol{F}_{\Phi(q)}^{-1}$. Then replacing q by $\Phi(q)$ yields $(\boldsymbol{A}_{\Phi(q)}^+)^* = \boldsymbol{F}_q\boldsymbol{Z}_{\Phi(q)}^+\boldsymbol{W}_{\Phi(q)}^+(\boldsymbol{Z}_{\Phi(q)}^+)^T\boldsymbol{F}_q^{-1}$. ■

Take for example, let $\boldsymbol{x} = [1, -2, -3, 7, 11]^T$, we have $\mathrm{A}_{\mathrm{II}}^{\mathrm{o}-}\mathrm{x} = (\mathrm{A}_{\mathrm{III}}^{\mathrm{o}+}\mathrm{x})^*$ and

$$\begin{aligned}\boldsymbol{A}_{II}^{\boldsymbol{o}+}\boldsymbol{x} &= [1 + j6.5489, -2 + j1.8809, -3 \\ &\quad - j4.0867, 7 - j8.1816, 11 + j3.8385]^T \\ \boldsymbol{A}_{III}^{\boldsymbol{o}+}\boldsymbol{x} &= [1 - j6.1211, -2 - j0.1534, -3 \\ &\quad - j6.9929, 7 - j8.7628, 11 + j4.1978]^T.\end{aligned}$$

B. DTT and Analytic Transform

The next theorem illustrates the interpretation of DTT spectra as the GDFT spectra of analytic vectors.

Theorem 7: Given a temporal column vector $\boldsymbol{x}$ and $\boldsymbol{y} = \boldsymbol{T}_q\boldsymbol{x}$. Then the IGDFT of the zero-padded and scaled DTT equals the analytic transform of the symmetrized temporal vector, that is,

$$\boldsymbol{A}_q^+(\boldsymbol{E}_q\boldsymbol{x}) = \boldsymbol{F}_q^{-1}(\boldsymbol{Z}_q\boldsymbol{W'}_q^+\boldsymbol{y}) \tag{37}$$

where the related matrices are defined in Table II.

Proof: To relate the DTT and the analytic transform, we combine the generic formulas (10) and (34) into the form

$$\begin{aligned}\boldsymbol{A}_q^+(\boldsymbol{E}_q\boldsymbol{x}) &= \boldsymbol{F}_q^{-1}\boldsymbol{Z}_q^+\boldsymbol{W}_q^+(\boldsymbol{Z}_q^+)^T\boldsymbol{F}_q\boldsymbol{E}_q\boldsymbol{x} \\ &= \boldsymbol{F}_q^{-1}\boldsymbol{Z}_q^+\boldsymbol{W}_q^+(\boldsymbol{Z'}_q\boldsymbol{T}_q\boldsymbol{x}) = \boldsymbol{F}_q^{-1}(\boldsymbol{Z}_q\boldsymbol{W'}_q^+\boldsymbol{y}).\end{aligned}$$

In the second step, the purpose of zero-padding matrix Z'_q is to gather the inherent zero output at the boundary indices of $\boldsymbol{T}_q$ to make up the total output length of N or $N+1$. The matrix product $\boldsymbol{Z}_q^+\boldsymbol{W}_q^+\boldsymbol{Z}'_q$ can be rewritten as $\boldsymbol{Z}_q^+\boldsymbol{Z}'_q\boldsymbol{W}_q'^+$, where $\boldsymbol{W}_q'^+$ is obtained from $\boldsymbol{W}_q^+$ by removing the boundary diagonal terms and reducing the order if necessary. Also, $\boldsymbol{Z}_q^+\boldsymbol{Z}'_q$ can pad the DTT output to fill the lengths $2N$ and $2N-1$ for the even and odd cases, respectively; hence, $\boldsymbol{Z}_q^+\boldsymbol{Z}'_q$ equals $\boldsymbol{Z}_q$, and the last step is arrived. ∎

Remarkably, when combined with DST, the matrices $\boldsymbol{F}_q$ and $\boldsymbol{F}_q^{-1}$ associated with $\boldsymbol{A}_q^+$ should be scaled by j and $-j$, and the derivation above has no affection. In the dual manner, the next corollary illustrates the interpretation of temporal vectors as the IGDFT of DTT-domain analytic vectors.

Corollary 4: Given a spectral column vector $\boldsymbol{y}$ and $\boldsymbol{x} = \boldsymbol{T}_q^{-1}\boldsymbol{y}$. Then the GDFT of the zero-padded and scaled temporal vector equals the analytic transform of the symmetrized DTT vector. That is,

$$\boldsymbol{A}_q^-(\boldsymbol{E}'_q\boldsymbol{y}) = \boldsymbol{F}_q(\boldsymbol{Z}_{\Phi(q)}\boldsymbol{W}_q'^+\boldsymbol{x}) \tag{38}$$

where the related matrices are defined in Table III.

Proof: By Theorem 7 and the properties $(\boldsymbol{A}_q^-)^* = \boldsymbol{A}_{\Phi(q)}^+$ and $\boldsymbol{E}_{\Phi(q)} = \boldsymbol{E}'_q$, we have $[\boldsymbol{A}_q^-(\boldsymbol{E}'_q\boldsymbol{y})]^* = \boldsymbol{A}_{\Phi(q)}^+(\boldsymbol{E}_{\Phi(q)}\boldsymbol{y}) = \boldsymbol{F}_{\Phi(q)}^{-1}(\boldsymbol{Z}_{\Phi(q)}\boldsymbol{W}_{\Phi(q)}'^+\boldsymbol{T}_{\Phi(q)}\boldsymbol{y})$. Then, by taking conjugate again and using the property $\boldsymbol{W}_q'^- = \boldsymbol{W}_{\Phi(q)}'^+$ due to $\boldsymbol{W}_q^- = \boldsymbol{W}_{\Phi(q)}^+$, the dual formula for IDTT is derived as $\boldsymbol{A}_q^-(\boldsymbol{E}'_q\boldsymbol{y}) = \boldsymbol{F}_q(\boldsymbol{Z}_{\Phi(q)}\boldsymbol{W}_{\Phi(q)}'^+\boldsymbol{T}_q^{-1}\boldsymbol{y}) = \boldsymbol{F}_q\,(\boldsymbol{Z}_{\Phi(q)}\boldsymbol{W}_q'^-\boldsymbol{x})$. ∎

The dual formulas (37) and (38) describe how a vector is related to its DTT and IDTT in the sense of analytic transform through appropriate symmetrization, zero-padding and weighting operations.

C. Autocorrelation and Squared Hilbert Envelope

Thanks to Corollary 2, we have linked the GDFT-domain periodic autocorrelation with the IGDFT-domain (temporal) envelope. Combining the corollary with Theorem 7, we can immediately obtain the time–frequency relation between the DTT-domain periodic autocorrelation and the Hilbert envelope (i.e., the magnitude envelope of the analytic signal) for a time-domain finite sequence.

Theorem 8: Given a temporal real-valued vector $\boldsymbol{x}$ and $\boldsymbol{y} = \boldsymbol{T}_q\boldsymbol{x}$. Let $\hat{\boldsymbol{y}} = \boldsymbol{Z}_q\boldsymbol{W}_q'^+\boldsymbol{y}$, then $\boldsymbol{r}_{\hat{\boldsymbol{y}}} = \boldsymbol{K}_q[(\boldsymbol{A}_q^+\boldsymbol{E}_q\boldsymbol{x}) \circ (\boldsymbol{A}_q^+\boldsymbol{E}_q\boldsymbol{x})^*]$,, where $\boldsymbol{r}_{\hat{\boldsymbol{y}}}$ is the circular or skew-circular autocorrelation of $\hat{\boldsymbol{y}}$ depending on the type of $\boldsymbol{K}_q$ that is a DFT or OTDFT matrix. (The specific types of transform and autocorrelation are defined in Table II, where notations (c) and (s) denote the circular and skew-circular autocorrelations, respectively.)

Proof: From Theorem 7, we have $\boldsymbol{F}_q^{-1}\hat{\boldsymbol{y}} = \boldsymbol{A}_q^+\boldsymbol{E}_q\boldsymbol{x}$; by Corollary 2, the proof is accomplished. ∎

The next theorem gives the dual formulation for estimating the spectral Hilbert envelope.

Theorem 9: Given a spectral real-valued vector $\boldsymbol{y}$ and $\boldsymbol{x} = \boldsymbol{T}_q^{-1}\boldsymbol{y}$. Let $\hat{\boldsymbol{x}} = \boldsymbol{Z}_{\Phi(q)}\boldsymbol{W}_{\Phi(q)}'^+\boldsymbol{x}$, then $\boldsymbol{r}_{\hat{\boldsymbol{x}}} = (1/M)\boldsymbol{\Gamma}_q^{-1}[(\boldsymbol{A}_q^-\boldsymbol{E'}_q\boldsymbol{y}) \circ (\boldsymbol{A}_q^-\boldsymbol{E}'_q\boldsymbol{y})^*]$, where $\boldsymbol{r}_{\hat{\mathrm{x}}}$ is the circular or skew-circular autocorrelation of $\hat{\boldsymbol{x}}$ depending on the type of $\boldsymbol{\Gamma}_q^{-1}$ that is a IDFT or IOFDFT matrix; and the scale factor M is the length of $\hat{\boldsymbol{x}}$. (The specific types of transform and autocorrelation are defined in Table III.)

Proof: From Corollary 4, we have $\boldsymbol{F}_q\hat{\boldsymbol{x}} = \boldsymbol{A}_q^-\boldsymbol{E}'_q\boldsymbol{y}$; by Theorem 2, the proof is accomplished. ∎

Theorems 8 and 9 permit to model the squared temporal or spectral Hilbert envelope of the symmetrized time-domain or DTT-domain vector by fitting an AR model to the zero-padded and weighted vector in the dual domain. The AR parameters obtained by the Yule–Walker equations should be zero-padded to length M and transformed by $\boldsymbol{F}_q$ and $\boldsymbol{F}_q^{-1}$, respectively, for the spectral and temporal envelope estimation.

D. Autocorrelation and DTT Power Envelope

Rather than the squared Hilbert envelope, modeling the squared DTT spectrum (i.e., evaluating the DTT power envelope) is another choice. This can be analyzed by applying the periodic autocorrelation analysis to the symmetrized data vector, instead of the zero-padded and scaled data vector.

Theorem 10: Given a spectral real-valued vector $\boldsymbol{y}$ and $\boldsymbol{x} = \boldsymbol{T}_q^{-1}\boldsymbol{y}$. Let $\ddot{\boldsymbol{x}} = \boldsymbol{E}_q\boldsymbol{x}$, then $\boldsymbol{r}_{\ddot{\boldsymbol{x}}} = (1/M)\boldsymbol{\Gamma}_q^{-1}[(\boldsymbol{E}'_q\boldsymbol{y}) \circ (\boldsymbol{E}'_q\boldsymbol{y})$, where $\mathbf{r}_{\ddot{\mathrm{x}}}$ is the circular or skew-circular autocorrelation of $\ddot{\boldsymbol{x}}$ depending on the type of Γ_q^{-1}; and M is the length of $\ddot{\mathrm{x}}$. (The specific types of transform and autocorrelation are defined in Table III.)

Proof: From (11), it implies $\boldsymbol{F}_q\boldsymbol{E}_q\boldsymbol{x} = \boldsymbol{E}'_q\boldsymbol{T}_q\boldsymbol{x} = \boldsymbol{E}'_q\boldsymbol{y}$; by Theorem 2, the proof is accomplished. ∎

Oppositely, the next theorem provides the fundamental for estimating the temporal power envelope.

Theorem 11: Given a temporal real-valued vector $\boldsymbol{x}$ and $\boldsymbol{y} = \boldsymbol{T}_q\boldsymbol{x}$. Let $\ddot{\boldsymbol{y}} = \boldsymbol{E}'_q\boldsymbol{y}$, then $\boldsymbol{r}_{\ddot{\boldsymbol{y}}} = \boldsymbol{K}_q[(\boldsymbol{E}_q\boldsymbol{x}) \circ (\boldsymbol{E}_q\boldsymbol{x})]$, where $\boldsymbol{r}_{\ddot{y}}$ is the circular or skew-circular autocorrelation of $\ddot{\boldsymbol{y}}$ depending on the type of $\boldsymbol{K}_q$. (The specific types of transform and autocorrelation are defined in Table II.)

Proof: From (12), it implies $\boldsymbol{F}_q^{-1}\boldsymbol{E}_q'\boldsymbol{y} = \boldsymbol{E}_q\boldsymbol{T}_q^{-1}\boldsymbol{y} = \boldsymbol{E}_q\boldsymbol{x}$; by Corollary 2, the proof is accomplished. ■

The dual formulas in Theorems 10 and 11 permit to estimate the temporal or spectral power envelope of the symmetrized time-domain or DTT-domain vector by fitting an AR model to the symmetrized vector in the dual domain.

E. Remarks and Examples

We notice that the periodic autocorrelation of the zero-padded samples is equivalent to the linear autocorrelation used in the autocorrelation method [2] for linear prediction. Therefore, the traditional autocorrelation method with scaling can be interpreted as the Hilbert envelope estimation for DTT/IDTT spectra as illustrated in Theorems 8 and 9. However, looking at the time-domain AR modeling in Theorems 9 and 10, the spectral squared Hilbert and power envelopes will be close when the temporal input signal is steady and the AR order is much smaller than the length of the input segment. The phenomenon can be interpreted by the approximate results of periodic autocorrelations $\boldsymbol{r}_{\ddot{\boldsymbol{x}}}$ and $\boldsymbol{r}_{\hat{\boldsymbol{x}}}$ in that condition. Likewise, when a windowing operation, such as sine windowing, is applied to the input samples, the effect of the aliasing part of periodic autocorrelation attenuates; hence, the resultant spectral Hilbert or power envelopes of all DTTs have no significant difference. Oppositely, the differences of the two kinds of envelopes are easier to be observed in the frequency-domain AR modeling because the DTT coefficients are usually unsteady and have large energy variation in the low-frequency part. Moreover, we can expect the temporal Hilbert envelop should be more smooth due to the imaginary part added by the analytic transform, while the power envelope can fit the temporal samples better in the LSE sense.

In the following, some examples are provided to illustrate the remarks. Fig. 2 compares spectral envelopes estimated by the two approaches on a time-domain audio segment of 1024 samples at 44.1 kHz. For the symmetrized odd DST-IV spectrum, in Fig. 2(b) the skew-circular autocorrelations of the zero-padded and scaled time-domain samples are used to evaluate its spectral Hilbert envelope, while in Fig. 2(c) the skew-circular autocorrelations of the symmetrized time-domain samples are used to evaluate its spectral power envelope. In the two cases, the AR order is 24, and the squared envelopes are computed by squaring the length-2048 O ^{2} DFT (or OFDFT) of the AR parameters. For comparison, Fig. 2(d) depicts the two squared envelopes together with the squared odd DST-IV spectrum. Since $\|\ddot{\boldsymbol{x}}\|^2 \approx 1/2 \cdot \|\hat{\boldsymbol{x}}\|^2$, the estimated Hilbert envelope should have energy alignment by 1/2 when compared with the squared DTT spectrum. As can be seen, the envelopes associated with (b) and (c) are highly close. The observation can be interpreted from the approximation of the two skew-circular autocorrelations, especially when the AR order is much smaller than the sample number.

Fig. 3(b) and (c) show the temporal squared Hilbert and power envelopes of a length-1024 even DCT-IV spectrum in the frequency-domain AR modeling of order 24. In addition to the two envelopes in Fig. 3, Fig. 4 depicts the two kinds of temporal envelope evaluated from the odd DCT-I spectrum of the same time-domain segment. Here, the estimated Hilbert envelopes

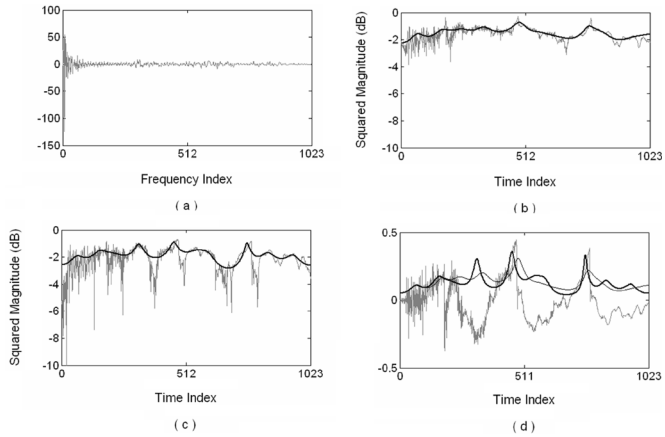

Fig. 3. Comparison of squared temporal envelopes. (a) The even DCT-IV coefficients of an audio segment of 1024 samples at 44.1 kHz. (b) The squared analytic transform of the symmetrized time-domain samples and the squared temporal Hilbert envelope through AR modeling. (c) The squared symmetrized time-domain samples and the temporal power envelope. (d) The time-domain samples and the two (non-squared) envelopes depicted in linear scale. The two squared temporal envelopes are obtained by order-24 AR modeling. The symmetrized parts are not depicted in (b)–(d).

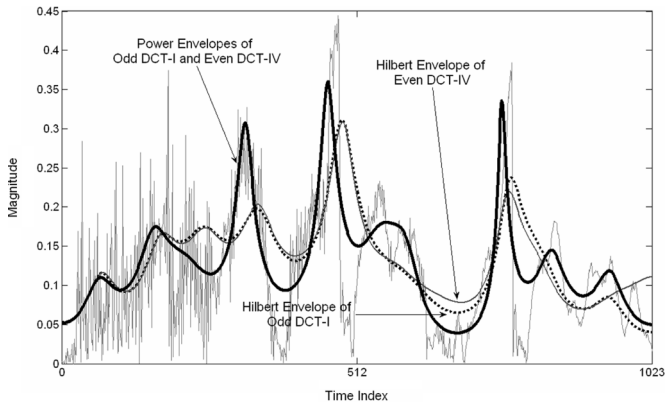

Fig. 4. Comparison of temporal envelopes evaluated from even DCT-IV and odd DCT-I coefficients, where the two (non-squared) envelopes and the magnitude of the time-domain samples in Fig. 3 are depicted in linear scale. Furthermore, the two envelopes evaluated from odd DCT-I coefficients are shown for comparison. Note that the two envelopes corresponding to the squared symmetrized time-domain samples are too close to distinguish.

have energy alignment by 1/2. The temporal envelopes are evaluated from the length-2048IO^2 DFT (or IOTDFT) of the AR parameters for even DCT-IV and from the length-2048 IDFT (or IOFDFT) of the AR parameters for odd DCT-I, respectively. As can be seen, the two Hilbert envelopes have minor deviation since the corresponding analytic signals are similar but different in magnitude, while the other two envelopes are too close to distinguish because they are corresponding to the squared time-domain samples which are symmetrized by $\boldsymbol{E}_{\mathrm{HSHA}}$ and $\boldsymbol{E}_{\mathrm{HAWS}}$, respectively, and only differ by one zero sample. We note that the temporal power envelope can fit well the temporal audio segment in the LSE sense and fit the valley better than the Hilbert envelope. Fig. 5 illustrates another instance, where an even DCT-IV spectrum of length 2048, which has strong low-frequency sinusoid component, is analyzed by the order-50 AR modeling. In Fig. 5(d), the Hilbert envelope without energy alignment fits the magnitude peaks of the sinusoid component, while the square-root power envelope leads to a sawtooth response. Since $\lim_{T\to\infty}(1/T)\int_0^T a^2\cos^2(\omega\cdot t)dt = (a^2/2)$,

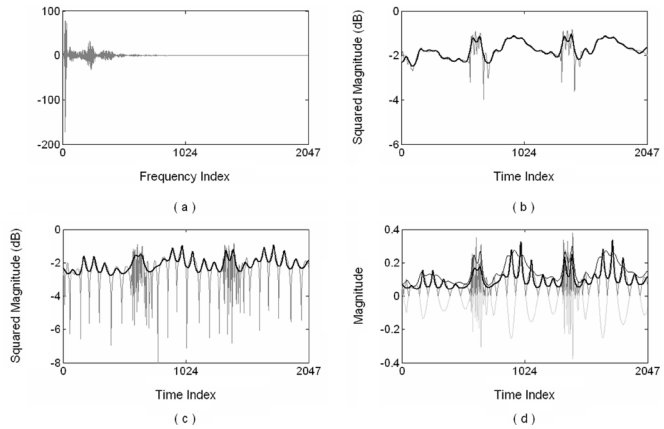

Fig. 5. Comparison of squared temporal envelopes. (a) The even DCT-IV coefficients of an audio segment of 2048 samples at 44.1 kHz. (b) The squared analytic transform of the symmetrized time-domain samples and the squared temporal Hilbert envelope through AR modeling. (c) The squared symmetrized time-domain samples and the temporal power envelope. (d) The time-domain samples, its magnitude, the square-root power envelopes (thick line), and the Hilbert envelope without energy alignment are depicted in linear scale. The two temporal envelopes are obtained by order-50 AR modeling.

the ratio of the energy and the squared amplitude of a sinusoid is 1/2; hence, the Hilbert envelope without energy alignment in this example can well fit the peaks of the sinusoid component.

V. Formulation of Temporal Noise Shaping with Finite-Length DTTs

TNS aims to shape the temporal envelope of the quantization noise by incorporating an open-loop predictive filter [17] across frequency lines in audio encoders/decoders. In terms of z-transform, the principle of TNS can be explained as follows. As depicted in Fig. 6, $x\,(k)$ and $d(k)$ denote the input and predictive residual signals in frequency domain in the analysis part, whereas $x_r(k)$ and $d_r(k)$ denote the reconstructed signals related to $x(k)$ and $d(k)$, respectively, in the synthesis part. The relation between the reconstruction error $r(k)$, i.e., $x(k)-x_r(k)$, and the quantization noise $q(k)$, i.e., $d(k)-d_r(k)$, is expressed in z-transform as $R\,(z)=Q(z)/(1-H(z))$, where $R(z)$ and $Q(z)$ are the z-transforms of $r(k)$ and $q(k)$. If the magnitude response of the inverse filter $1/(1-H(z))$ can approximate the temporal envelope of the frequency-domain input signal $x(k)$, the quantization noise $\mathrm{Q}(\mathrm{e}^{-j\omega})$ (in time domain) can be amplified or attenuated with the temporal shape.

Herre and Johnston have proposed the TNS predictive filter by the duality between the squared temporal Hilbert envelope and the power spectrum for continuous-time signals [12]–[15]. Since there is no derivation for the finite sequences in DTT frequency domain in the literature, this section analyzes the fundamental of the TNS in DTTs based on the AR modeling of finite sequences in last section. Both the temporal Hilbert-envelope method used in standard and the power-envelope method will be analyzed.

A. Evaluation and Representation of the Whitening Filter

Let $\boldsymbol{x}$ denote the data vector and $\boldsymbol{y}=\boldsymbol{T}_q\boldsymbol{x}$. Depending on the kind of temporal envelope, we can have either the squared Hilbert envelope or power envelope for shaping the reconstruction noise. As defined in Theorems 8 and 11, $\mathbf{r}_{\hat{y}}$ and $\mathbf{r}_{\ddot{y}}$ consist of either the circular and skew-circular autocorrelation of $\hat{\mathrm{y}}$ and $\ddot{\mathrm{y}}$, respectively, where $\hat{\boldsymbol{y}}=[\hat{y}(0),\hat{y}(1),\ldots,\hat{y}(M-1)]^T=\boldsymbol{Z}_q\boldsymbol{W}_q'^{+}\boldsymbol{y}$ and $\ddot{\boldsymbol{y}}=[\ddot{y}(0),\ddot{y}(1),\ldots,\ddot{y}(M-1)]^T=\boldsymbol{E}'_q\boldsymbol{y}$. Subsequently, the parameters of the whitening filter are obtained by solving the Yule–Walker equations. Since the relations in Theorems 8 and 11 are based on length M instead of N, we assume that the whitening filter is applied to $\hat{\mathrm{y}}$ covering $\boldsymbol{y}$ in our derivation.

The whitening filter can be represented as a circulant or skew-circulant matrix [31] for the case of the circular or skew-circular convolution. By taking conjugate of both sides of (5)–(8) and assuming that the operands $\boldsymbol{x}$ and $\boldsymbol{y}$ are real-valued, we have the alternative relations

$$\begin{aligned}
\boldsymbol{u}&=N\cdot\boldsymbol{G}_{0,0}\left[\left(\boldsymbol{G}_{0,0}^{-1}\boldsymbol{x}\right)\circ\left(\boldsymbol{G}_{0,0}^{-1}\boldsymbol{y}\right)\right]\\
\boldsymbol{u}&=N\cdot\boldsymbol{G}_{\frac{1}{2}},0\left[\left(\boldsymbol{G}_{\frac{1}{2},0}^{-1}\boldsymbol{x}\right)\circ\left(\boldsymbol{G}_{0,0}^{-1}\boldsymbol{y}\right)\right]\\
\boldsymbol{w}&=N\cdot\boldsymbol{G}_{0,\frac{1}{2}}\left[\left(\boldsymbol{G}_{0,\frac{1}{2}}^{-1}\boldsymbol{x}\right)\circ\left(\boldsymbol{G}_{0,\frac{1}{2}}^{-1}\boldsymbol{y}\right)\right]\\
\boldsymbol{w}&=N\cdot\boldsymbol{G}_{\frac{1}{2},\frac{1}{2}}\left[\left(\boldsymbol{G}_{\frac{1}{2},\frac{1}{2}}^{-1}\boldsymbol{x}\right)\circ\left(\boldsymbol{G}_{0,\frac{1}{2}}^{-1}\boldsymbol{y}\right)\right].
\end{aligned}\tag{39}$$

According to (39), the matrix representation $\boldsymbol{H}$ for the whitening filter can be decomposed as

$$\boldsymbol{H}=\boldsymbol{F}_q\cdot\Phi\cdot\boldsymbol{F}_q^{-1}\tag{40}$$

where $\boldsymbol{\Phi}=N\cdot\mathrm{diag}\{(\mathbf{G}_{0,b}^{-1}\alpha)_n|n=0,1,\ldots,M-1\}$, where b is 0 or 1/2 depending on the convolution type.

In the MPEG standard, the TNS predictive error filter is performed through the linear convolution (filtering) in transform domain. In matrix form, the linear convolution $\boldsymbol{L}$ which is lower triangular is the same as the periodic convolution $\boldsymbol{H}$ except for the upper triangular entries. Thus, by padding the input data with suitable zeros, the periodic convolution equals the linear convolution. However, to reconstruct $\boldsymbol{y}$, all M residuals are necessary to be transmitted to the decoder to perform the periodic deconvolution $\boldsymbol{H}^{-1}$, while only the residuals corresponding to $\boldsymbol{y}$ are required for the linear deconvolution $\boldsymbol{L}^{-1}$ for it is still lower triangular. Interestingly, if $\boldsymbol{L}^{-1}\boldsymbol{u}=\boldsymbol{v}$ and $\boldsymbol{v}_n=0$ for $M-P\le n\le M-1$, then $\boldsymbol{H}^{-1}\boldsymbol{u}=\boldsymbol{H}^{-1}(\boldsymbol{L}\boldsymbol{v})=\boldsymbol{H}^{-1}(\boldsymbol{H}\boldsymbol{v})=\boldsymbol{v}$. Hence, $\boldsymbol{H}$ and $\boldsymbol{H}^{-1}$ are equivalent to $\boldsymbol{L}$ and $\boldsymbol{L}^{-1}$ on $\hat{\boldsymbol{y}}$ and the related residuals, respectively, and we can develop the TNS formulation on $\hat{\mathrm{y}}$ in the periodic convolution/deconvolution manner.

B. Formulation of Temporal Noise Shaping

The shaping effect of TNS can be formulated as follows. First, the dequantized residual $\boldsymbol{d}_r$ is given by

$$\boldsymbol{d}_r=\boldsymbol{d}+\boldsymbol{\varepsilon}=\boldsymbol{H}\hat{\boldsymbol{y}}+\boldsymbol{\varepsilon}\tag{41}$$

where $\boldsymbol{d}$ is the original residual, and $\boldsymbol{\varepsilon}$ is the additive quantization noise. After deconvolution, the reconstructed spectral vector $\hat{\mathrm{y}}_r$ is given by $\hat{\boldsymbol{y}}_r=\boldsymbol{H}^{-1}(\boldsymbol{H}\hat{\boldsymbol{y}}+\boldsymbol{\varepsilon})=\hat{\boldsymbol{y}}+\boldsymbol{H}^{-1}\boldsymbol{\varepsilon}$. In other words, the quantization noise $\boldsymbol{\varepsilon}$ can be shaped by the periodic deconvolution $\boldsymbol{H}^{-1}$ in transform domain. Notice that only the part of $\boldsymbol{d}$ corresponding to $\boldsymbol{y}$ is quantized and transmitted from the encoder to the decoder. We assume that the zero-padded part is

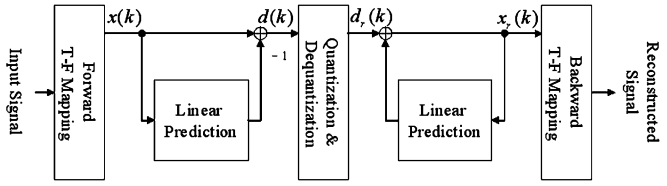

Fig. 6. Open-loop prediction coding scheme in TNS.

perfectly reconstructed and the reconstruction noise exists only for the non-zero-padded samples of $\hat{\boldsymbol{y}}$. Based on the assumption, we can confirm the equivalency of $\boldsymbol{H}^{-1}$ and $\boldsymbol{L}^{-1}$ on $\boldsymbol{\varepsilon}$ and have $\boldsymbol{H}^{-1}\boldsymbol{\varepsilon} = \boldsymbol{L}^{-1}\boldsymbol{\varepsilon} = \boldsymbol{Z_q n}$, where Z_q is the zero-padding matrix corresponding to $\boldsymbol{T}_q$, and $\boldsymbol{n}$ denotes the reconstruction noise related to $\boldsymbol{W}_q'^{+}\boldsymbol{y}$. This assumption implies that some "virtual" quantization noise should be imposed on the P samples after $\boldsymbol{y}$ to correct the noise propagation in the open-loop prediction. To check the temporal shaping effect, $\boldsymbol{T}_q^{-1}$ is applied to the part of $\hat{\boldsymbol{y}}_r$ related to $\boldsymbol{y}$, i.e., $(\boldsymbol{W}_q'^{+})^{-1}\boldsymbol{Z}_q^T\hat{\boldsymbol{y}}_r$, to yield the reconstructed temporal vector $\boldsymbol{x}_r$, where $(\boldsymbol{W}_q'^{+})^{-1}$ is multiplied for removing the weighting of $\boldsymbol{W}_q'^{+}$ on $\boldsymbol{y}$. Before formulating $\boldsymbol{x}_r$, we consider another relation between IDTT and IGDFT as follows. For an arbitrary vector $\boldsymbol{z}$, (37) can be rewritten as

$$\boldsymbol{A}_q^{+}\boldsymbol{E}_q\boldsymbol{T}_q^{-1}\boldsymbol{z} = \boldsymbol{F}_q^{-1}\boldsymbol{Z}_q\boldsymbol{W}_q'^{+}\boldsymbol{z}. \tag{42}$$

Thus, by the property that the real part of the analytic transform exactly equals the original sequence, for a real-valued vector $\boldsymbol{z}$, we have

$$\boldsymbol{E}_q\boldsymbol{T}_q^{-1}\boldsymbol{z} = \mathrm{Re}\left\{\boldsymbol{F}_q^{-1}\boldsymbol{Z}_q\boldsymbol{W}_q'^{+}\boldsymbol{z}\right\}. \tag{43}$$

Hence, by letting $\boldsymbol{z}$ be $(\boldsymbol{W}_q'^{+})^{-1}\boldsymbol{Z}_q^T\hat{\boldsymbol{y}}_r$ in (43), the reconstructed symmetrized temporal sequence is given by

$$\begin{aligned}\boldsymbol{E}_q\boldsymbol{x}_r &= \boldsymbol{E}_q\boldsymbol{T}_q^{-1}(\boldsymbol{W}_q'^{+})^{-1}\boldsymbol{Z}_q^T\hat{\boldsymbol{y}}_r \\ &= \mathrm{Re}\left\{\boldsymbol{F}_q^{-1}\boldsymbol{Z}_q\boldsymbol{W}_q'^{+}(\boldsymbol{W}_q'^{+})^{-1}\boldsymbol{Z}_q^T\hat{\boldsymbol{y}}_r\right\}.\end{aligned} \tag{44}$$

Substituting $\hat{\boldsymbol{y}}_r = \hat{\boldsymbol{y}} + \boldsymbol{H}^{-1}\boldsymbol{\varepsilon}$ into (44) leads to

$$\begin{aligned}\boldsymbol{E}_q\boldsymbol{x}_r &= \mathrm{Re}\left\{\boldsymbol{F}_q^{-1}\boldsymbol{Z}_q\boldsymbol{Z}_q^T(\hat{\boldsymbol{y}} + \boldsymbol{H}^{-1}\boldsymbol{\varepsilon})\right\} \\ &= \mathrm{Re}\left\{\boldsymbol{F}_q^{-1}\boldsymbol{Z}_q\boldsymbol{Z}_q^T(\boldsymbol{Z}_q\boldsymbol{W}_q'^{+}\boldsymbol{y} + \boldsymbol{H}^{-1}\boldsymbol{\varepsilon})\right\} \\ &= \mathrm{Re}\left\{\boldsymbol{F}_q^{-1}\boldsymbol{Z}_q\boldsymbol{W}_q'^{+}\boldsymbol{y} + \boldsymbol{F}_q^{-1}\boldsymbol{Z}_q\boldsymbol{Z}_q^T\boldsymbol{H}^{-1}\boldsymbol{\varepsilon}\right\} \\ &= \mathrm{Re}\left\{\boldsymbol{F}_q^{-1}\boldsymbol{Z}_q\boldsymbol{W}_q'^{+}\boldsymbol{y}\right\} + \mathrm{Re}\left\{\boldsymbol{F}_q^{-1}\boldsymbol{Z}_q\boldsymbol{Z}_q^T\boldsymbol{H}^{-1}\boldsymbol{\varepsilon}\right\} \\ &= \boldsymbol{E}_q\boldsymbol{T}_q^{-1}\boldsymbol{y} + \mathrm{Re}\left\{\boldsymbol{F}_q^{-1}\boldsymbol{H}^{-1}\boldsymbol{\varepsilon}\right\}.\end{aligned} \tag{45}$$

In the last step in (45), the property (43) is used, and the product $\boldsymbol{Z}_q\boldsymbol{Z}_q^T$ can be removed due to $\boldsymbol{H}^{-1}\boldsymbol{\varepsilon} = \boldsymbol{Z}_q\boldsymbol{n}$. Then substituting (40) to (45) yields

$$\begin{aligned}\boldsymbol{E}_q\boldsymbol{x}_r &= \boldsymbol{E}_q\boldsymbol{x} + \mathrm{Re}\left\{\boldsymbol{\Phi}^{-1}\boldsymbol{F}_q^{-1}\varepsilon\right\} \\ &= \boldsymbol{E}_q\boldsymbol{x} + \boldsymbol{\eta} \circ \mathrm{Re}\left\{\boldsymbol{\psi} \circ \boldsymbol{F}_q^{-1}\varepsilon\right\}\end{aligned} \tag{46}$$

where column vectors $\boldsymbol{\eta}$ and $\boldsymbol{\psi}$ are defined by $\boldsymbol{\eta}_i = |(\boldsymbol{\Phi}^{-1})_{ii}|$ and $\boldsymbol{\psi}_i = (\boldsymbol{\Phi}^{-1})_{ii} \cdot \boldsymbol{\eta}_i^{-1}$ for $i = 0, 1, \ldots, M-1$. Hence, $\boldsymbol{\eta}$ results in the temporal shaping effect. Furthermore, due to $\boldsymbol{\psi}$, the imaginary part of $\boldsymbol{F}_q^{-1}\varepsilon$ is also involved in the reconstruction noise.

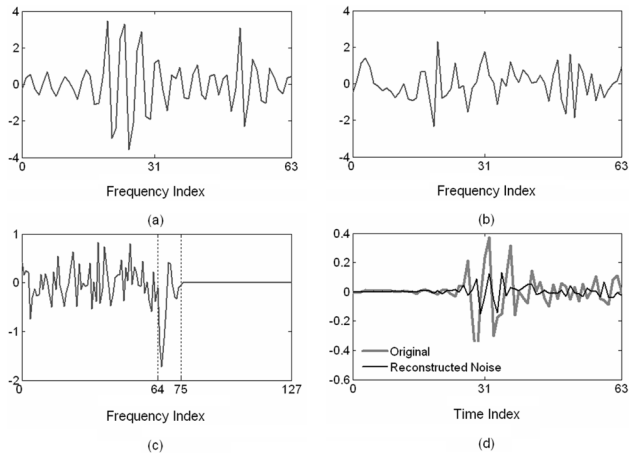

Fig. 7. TNS analysis. (a) The even DCT-IV coefficients of an audio segment of 64 samples at 8 kHz. (b) The predictive residuals by the order-12 whitening filter corresponding to temporal Hilbert envelope. (c) The quantization noise on the residuals indexed 0–63, and the virtual quantization noise indexed 64–75. (d) The original time-domain signal and the reconstruction time-domain noise.

C. Remarks and Examples

Fig. 7 illustrates an order-12 TNS analysis result on the even DCT-IV coefficients of 64 audio samples at 8 kHz. As shown in Fig. 7(c), although only 64 quantization noise samples are applied to the 64 residual samples transmitted, the 12 "virtual" quantization noise samples indexed from 64 to 75 occur when analyzed with the skew-circular convolution. In Fig. 7(d), the original time-domain signal and the reconstruction noise are depicted to show the shaping effect. Also notice that the TNS processing is applied to a data segment of length 64, but is analyzed in the O^{2}DFT domain of length 128. Because of symmetry, only one side is shown in this illustration. Fig. 8 illustrates the noise shaping effect for the Hilbert-envelope and power-envelope methods, where the two order-12 AR modeling methods are applied to a transient audio segment of 2048 samples at 44.1 kHz. The inverted magnitude responses of the two skew-circular predictors corresponding to the Hilbert and power envelopes are aligned in energy and depicted in Fig. 8(c). The quantization noise on the residual samples is simulated by a white random sequence shown in Fig. 8(d). The reconstructed temporal noises by the two predictors are shown in Fig. 8(e) and (f). As shown in Fig. 8(c), the magnitude response of the predictor corresponding to the power envelope is sharper than that corresponding to the Hilbert envelope at the silence segment. Therefore, the pre-echo artifact in Fig. 8(f) has higher attenuation than that in Fig. 8(e). The major difference of the two methods comes from the envelope estimation of the low-frequency lines due to the contribution of the low-frequency lines in the skew-circular autocorrelations. The Hilbert envelope can avoid the smoothing effect by removing the low-frequency lines in the calculation of the filter coefficients while applying the noise shaping to all the frequency lines to achieve similar as the power envelope method.

Another remark in above analysis is that the AAC employs the MDCT as the time-frequency mapping tool. The $N \times 2N$ MDCT matrix $\boldsymbol{M}$ can be factorized into the product of the "time-

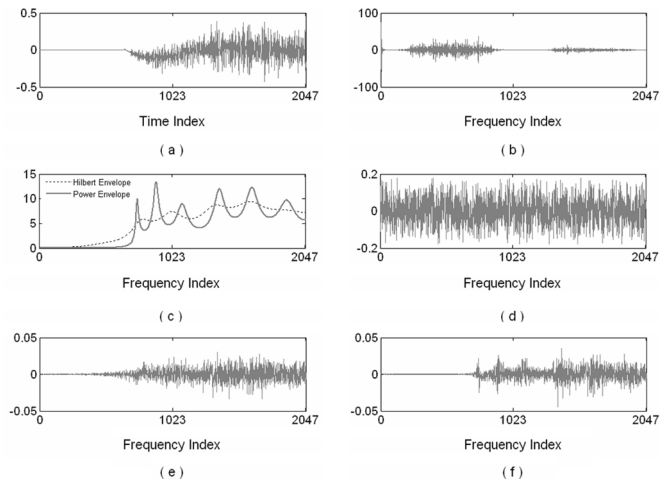

Fig. 8. Comparison of TNS effect by the order-12 predictors corresponding to the Hilbert and power envelopes. (a) A transient audio segment of 2048 samples at 44.1 kHz. (b) The even DCT-IV coefficients. (c) The energy-aligned inverted magnitude responses of the two skew-circular predictors corresponding to the Hilbert and power envelopes. (d) The simulated quantization noise. (e) The reconstruction temporal noises by the predictor corresponding to the Hilbert envelope. (f) The reconstruction temporal noise by the predictor corresponding to the power envelope.

domain aliasing" matrix and the $N \times N$ even DCT-IV matrix $\boldsymbol{C}_{IV}^{e}$ [24], [32]

$$\boldsymbol{M} = \boldsymbol{C}_{IV}^{e} \cdot \begin{bmatrix} \boldsymbol{0}_{N/2\times N/2} & \boldsymbol{0}_{N/2\times N/2} & -\boldsymbol{J}_{N/2} & -\boldsymbol{I}_{N/2} \\ \boldsymbol{I}_{N/2} & -\boldsymbol{J}_{N/2} & \boldsymbol{0}_{N/2\times N/2} & \boldsymbol{0}_{N/2\times N/2} \end{bmatrix}. \quad (47)$$

Accordingly, our formula associated with the even DCT-IV of length N should apply to the aliased time-domain segment. The impact on the aliased time-domain segment has led to the time-domain aliasing artifact [12], [16] which is another issue for the design of the whitening filter design, but the issue should not affect the comparison remarks on the Hilbert-envelope TNS and power-envelope TNS.

VI. Conclusion

This paper has established the compact forms for the AR modeling of both the temporal and spectral envelopes with finite-length DTTs. In the first part, we have developed the AR modeling forms with GDFTs which are highly related to DTTs with GPSs. In this part, we first represented the periodic correlation as the periodic convolution (see Theorem 1) to exploit the existing convolution theorems for GDFTs. Then the compact forms (see Theorem 2 and Corollary 2) between the periodic autocorrelation and power spectrum have been established. We have shown that the AR parameters can be obtained by solving the Yule–Walker equations consisting of periodic autocorrelations (see Theorems 3-5). The result is consistent with the traditional AR modeling method for DFTs. In the second part, we have considered the AR modeling with DTTs. The analytic transforms (see Theorem 6 and Corollary 3) based on GDFTs have been introduced for evaluating the analytic finite sequence corresponding to a real-valued finite sequence. Then the DTT spectra are represented as the GDFT spectra of the analytic vectors to exploit the AR modeling forms with GDFTs (see Theorem 7). The dual result for the IDTT spectra was also derived (see Corollary 4). Subsequently, the compact forms (see Theorems 8 and 9) between the periodic autocorrelation and the squared Hilbert envelope have been established. The compact forms (see Theorems 10 and 11) between the periodic autocorrelation and the power envelope have also been established. Based on these forms, the periodic autocorrelations of a zero-padded or symmetrized sequence can be used to comprise the Yule–Walker equations depending on the desired envelope estimation. The theorems have been applied to analyze the TNS tool in the MPEG-2/4 AAC audio coding scheme.

References

[1] S. M. Kay and S. L. Marple, Jr., "Spectrum analysis—A modern perspective," *Proc. IEEE*, vol. 69, no. 11, pp. 1380–1419, Nov. 1981.
[2] J. Makhoul, "Linear prediction: A tutorial review," *Proc. IEEE*, vol. 63, no. 4, pp. 561–580, Apr. 1975.
[3] T. van Waterschoot, G. Rombouts, P. Verhoeve, and M. Moonen, "Double-talk-robust prediction error identification algorithms for acoustic echo cancellation," *IEEE Trans. Signal Process.*, vol. 55, no. 3, pp. 846–858, Mar. 2007.
[4] P. Ekstrand, "Bandwidth extension of audio signals by spectral band replication," in *Proc. 1st IEEE Benelux Workshop Model Based Process. Coding Audio*, Leuven, Belgium, Nov. 2002, pp. 53–58.
[5] Bandwidth Extension,. Pattaya, Thailand, Mar. 2003, ISO/IEC JTC1/SC29/WG11/N5570, ISO/IEC, 14496-3:2001/FDAM1.
[6] Audio Lossless Coding (ALS), New Audio Profiles and BSAC Extensions, ISO/IEC, 14496-3:2005/Amd, vol. 3, 2006.
[7] A. Härmä and U. K. Laine, "A comparison of warped and conventional linear predictive coding," *IEEE Trans. Speech, Audio, Process.*, vol. 9, no. 5, pp. 579–588, Jul. 2001.
[8] M. Deriche and D. Ning, "A novel audio coding scheme using warped linear prediction model and the discrete wavelet transform," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 14, no. 6, pp. 2039–2048, Nov. 2006.
[9] M. Athineos and D. Ellis, "Frequency-domain linear prediction for temporal features," in *Proc. IEEE ASRU Workshop*, Dec. 2003, pp. 261–266.
[10] M. Athineos and D. Ellis, "Sound texture modeling with linear prediction in both time and frequency domains," in *Proc. IEEE ICASSP*, Apr. 2003, pp. 648–651.
[11] M. Athineos and D. Ellis, "Autoregressive modeling of temporal envelopes," *IEEE Trans. Signal Process.*, vol. 55, no. 11, pp. 5237–5245, Nov. 2007.
[12] J. Herre and J. D. Johnston, "Enhancing the performance of perceptual audio coders by using temporal noise shaping (TNS)," in *Proc. AES 101st Conv.*, Los Angeles, CA, Nov. 1996, preprint 4384.
[13] J. Herre and J. D. Johnston, "Continuously signal-adaptive filterbank for high-quality perceptual audio coding," in *Proc. IEEE ASSP Workshop*, Oct. 1997.
[14] J. Herre and J. D. Johnston, "Exploiting both time and frequency structure in a system that uses an analysis/synthesis filterbank with high frequency resolution," in *Proc. AES 103rd Conv.*, New York, Sep. 1997, preprint 5419.
[15] J. Herre, "Temporal noise shaping, quantization and coding methods in perceptual audio coding: A tutorial introduction," in *Proc. 17th AES Int. Conf.: High-Quality Audio Coding*, Sep. 1999, pp. 17–31.
[16] C. M. Liu, H. W. Hsu, and W. C. Lee, "Compression artifacts in perceptual audio coding," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 16, no. 4, pp. 681–695, May 2008.
[17] N. S. Jayant and P. Noll, *Digital Coding of Waveforms*. Englewood Cliffs, NJ: Prentice-Hall, 1984.
[18] R. Kumaresan, "An inverse signal approach to computing the envelope of a real valued signal," *IEEE Signal Process. Lett.*, vol. 5, no. 10, pp. 256–259, Oct. 1998.
[19] R. Kumaresan and A. Rao, "Model-based approach to envelope and positive instantaneous frequency estimation of signals with speech applications," *J. Acoust. Soc. Amer.*, vol. 105, no. 3, pp. 1912–1924, Mar. 1999.
[20] R. Kumaresan, "On minimum/maximum/all-pass decompositions in time and frequency domains," *IEEE Trans. Signal Process.*, vol. 48, no. 10, pp. 2973–2976, Oct. 2000.
[21] R. Kumaresan and Y. Wang, "On the relationship between line-spectral frequencies and zero-crossings of signals," *IEEE Trans. Speech Audio Process.*, vol. 9, no. 4, pp. 458–461, May 2001.

[22] R. Kumaresan and Y. Wang, "On representing signals using only timing information," *J. Acoust. Soc. Amer.*, vol. 110, no. 5, pp. 2421–2439, Nov. 2002.
[23] J. P. Princen, A. W. Johnson, and A. B. Bradley, "Subband/transform coding using filter bank designs based on time domain aliasing cancellation," in *Proc. IEEE ICASSP'87*, Dallas, TX, Apr. 1987, pp. 2161–2164.
[24] H. S. Malvar, *Signal Processing With Lapped Transforms*. Norwood, MA: Artech House, 1992.
[25] G. Bongiovanni, P. Corsini, and G. Frosini, "One-dimensional and two-dimensional generalized discrete Fourier transforms," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-24, no. 1, pp. 97–99, Feb. 1976.
[26] J. L. Vemet, "Real signals fast Fourier transform: Storage capacity and step number reduction by means of an odd discrete Fourier transform," *Proc. IEEE*, vol. 59, pp. 1531–1532, Oct. 1971.
[27] S. A. Martucci, "Symmetric convolution and the discrete sine and cosine transforms," *IEEE Trans. Signal Process.*, vol. 42, no. 5, pp. 1038–1051, May 1994.
[28] V. Britanak and K. R. Rao, "The fast generalized discrete Fourier transforms: A unified approach to the discrete sinusoidal transforms computation," *Signal Process.*, vol. 79, no. 12, pp. 135–150, Dec. 1999.
[29] A. V. Oppenheim, R. W. Schafer, and J. R. Buck, *Discrete-Time Signal Processing*, 2nd ed. Upper Saddle River, NJ: Prentice-Hall, 1999.
[30] L. S. Marple, Jr., "Computing the discrete-time 'analytic' signal via FFT," *IEEE Trans. Signal Process.*, vol. 47, no. 9, pp. 2600–2603, Sep. 1999.
[31] P. J. Davis, *Circulant Matrices*. New York: Wiley, 1979.
[32] M. H. Cheng and Y. H. Hsu, "Fast IMDCT and MDCT algorithms—A matrix approach," *IEEE Trans. Signal Process.*, vol. 51, no. 1, pp. 221–229, Jan. 2003.

Han-Wen Hsu was born in Tainan, Taiwan, in October 1977. He received the B.S. degree from the Division of Applied Mathematics, Department of Mathematics, National Tsing Hua University, Hsinchu, Taiwan, in 2000 and the M.S. and Ph.D. degrees from the Department of Computer Science and Information Engineering, National Chiao Tung University, Hsinchu, in 2004 and 2010, respectively.

His research interests are in audio coding and signal processing.

Chi-Min Liu received the M.S. and Ph.D. degrees in electronics from National Chiao Tung University (NCTU), Hsinchu, Taiwan, in 1987 and 1991, respectively.

He is currently a Professor in the Department of Computer Science, NCTU. His research interests include audio compression, audio effects, and video compression. He leads the Perceptual Signal Processing Laboratory at NCTU.