A Unified Approach to Scrambling Filter Design

Chwan-Wen King and Ching-An Lin

Abstract— **Most speech scrambling systems are either linear periodic filters or can be modeled as such. It is well-known that from an input-output point of view, a periodic filter is equivalent to a multi-input multi-output linear time-invariant system and thus a rational transfer matrix. In this paper, we propose a framework, based on transfer matrices and frequency domain descriptions, for analysis and design of speech scrambling filters. We derive necessary and sufficient conditions for a scrambling system to be insensitive to frame synchronization error. We propose a procedure for the design of scrambling filters that are insensitive to synchronization error and have zero bandwidth expansion. An illustrative design example is given.**

I. Introduction

MANY schemes [6], [11], [17] have been proposed for analog scrambling of speech. Most existing analog speech scrambling filters [1], [4], [6], [10], even the so-called 2-D scrambler [13], are either linear periodic filters or can be modeled as such. From an input-output point of view, a linear periodic filter is equivalent to a multi-input multi-output linear time-invariant system and thus a rational transfer matrix [12]. It is shown in this paper that by modeling periodic filters as transfer matrices and using their frequency domain characteristics [12], many design requirements of scrambling and descrambling filters become transparent and thus design procedures become easy to develop.

The need for frame synchronization in scrambling systems complicates the implementation and makes the recovered speech sensitive to channel conditions. Lee [6], [7] and Del Re [13] proposed scrambling schemes in which the frequency bands of input speech are interchanged and the original speech is correctly recovered in spite of the existence of frame synchronization error. This capability improves reliability and feasibility of scrambling systems and reduces implementation complexity. However, general conditions under which a scrambling system is insensitive to synchronization error have not been investigated.

In this paper, we develop necessary and sufficient conditions for a scrambling system to be insensitive to frame synchronization error. Based on these conditions, we propose an algorithm for designing scrambling systems that do not require frame synchronization. The implementation of such scrambling systems is also simplified since it only requires realization of periodic filters. The realization of periodic filters represented as transfer matrices can be found in [8].

The contributions of this paper are as follows. We propose a framework, based on transfer matrices and frequency domain descriptions, for analysis and design of speech scrambling filters. We derive necessary and sufficient conditions for a scrambling system to be insensitive to frame synchronization error. We propose a procedure for the design of scrambling filters that are insensitive to synchronization error and have zero bandwidth expansion.

In Section II, we review the properties of the periodic system in both time and frequency domains. In Section III, we introduce a generic model for scrambling systems. Performance analysis is studied under this model and the performance of an FFT algorithm-based scrambler is qualitatively analyzed. In Section IV, we derive necessary and sufficient conditions for scrambling systems to be insensitive to frame synchronization error. A design algorithm is given in Section V. An illustrative design example with simulation and experimental results is then given in Section VI. A brief conclusion is given in the last section.

II. Properties of Periodic Filter

In this section, we review descriptions of periodic filters in time-domain and frequency-domain and discuss briefly their properties.

A. Time Domain Description

Consider a linear finite dimensional SISO causal N-periodic filter, described by

$$\begin{aligned} x(k+1) &= A(k)x(k) + b(k)u(k), \\ y(k) &= c(k)x(k) + d(k)u(k) \end{aligned} \tag{1}$$

where $A(k) \in \mathbb{R}^{n\times n}$, $b(k) \in \mathbb{R}^{n\times 1}$, $c(k) \in \mathbb{R}^{1\times n}$, and $d(k) \in \mathbb{R}$ are N-periodic, i.e., $A(k+N) = A(k)$, $b(k+N) = b(k)$, $c(k+N) = c(k)$, and $d(k+N) = d(k)$ for all k.

Let

$$\bar{x}(k) = x(kN),$$
$$\bar{u}(k) = [u(kN) \quad u(kN+1) \quad \cdots \quad u(kN+N-1)]^T$$

and

$$\bar{y}(k) = [y(kN) \quad y(kN+1) \quad \cdots \quad y(kN+N-1)]^T$$

then [12]

$$\begin{aligned} \bar{x}(k+1) &= \bar{A}\bar{x}(k) + \bar{B}\bar{u}(k), \\ \bar{y}(k) &= \bar{C}\bar{x}(k) + \bar{D}\bar{u}(k) \end{aligned} \tag{2}$$

Manuscript received October 9, 1992; revised February 10, 1995. This work was supported by the National Science Council of the Republic of China under Grant NSC-79-0404-E-009-07. The associate editor coordinating the review of this paper and approving it for publication was Prof. Sergio D. Cabrera.

C.-W. King is with the Institute of Electronics, National Chiao-Tung University, Hsinchu, Taiwan, R.O.C.

C.-A. Lin is with the Department of Control Engineering, National Chiao-Tung University, Hsinchu, Taiwan, R.O.C.

IEEE Log Number 9412627.

1053–587X/95$04.00 © 1995 IEEE

where

$$\bar{A} = A(N-1)A(N-2)\cdots A(1)A(0),$$
$$\bar{B} = [\bar{b}_0 \quad \bar{b}_1 \quad \cdots \quad \bar{b}_{N-1}]$$
$$\text{with} \quad \bar{b}_i = \begin{cases} A(N-1)\cdots A(i+1)b(i), & i = 0,\cdots,N-2 \\ b(N-1), & i = N-1 \end{cases},$$
$$\bar{C}^T = [\bar{c}_0^T \quad \bar{c}_1^T \quad \cdots \quad \bar{c}_{N-1}^T]$$
$$\text{with} \quad \bar{c}_i = \begin{cases} c(0), & i = 0 \\ c(i)A(i-1)\cdots A(0), & i = 1,\cdots,N-1 \end{cases},$$
$$\bar{D} = [\bar{d}_{i,j}]$$
$$\text{with } \bar{d}_{i,j} = \begin{cases} 0, & i < j \\ d(i), & i = j \\ c(i)A(i-1)\cdots A(j+1)b(j), & i > j \end{cases}. \qquad (3)$$

With zero initial conditions, the systems (1) and (2) describe the same input-output relation, except that in (2), the input and output are in blocks of size N. Since (2) is linear time-invariant, the transfer function $G(z) = \bar{C}(zI - \bar{A})^{-1}\bar{B} + \bar{D}$ is defined. We call $G(z)$ the *block transfer matrix* of the N-periodic filter (1). We note that linear periodic filters described in other forms, e.g., polyphase model [2], [16], coefficient varying model [2], and difference equation [12], all can be converted into the corresponding MIMO linear time-invariant system of the form (2). Clearly the periodic system (1) is stable if and only if the block system (2) is stable. Also, since (2) is linear time-invariant, it allows a frequency domain description of the periodic filter.

B. Frequency Domain Description

Let

$$Y(z) = \sum_{k=0}^{\infty} y(k)z^{-k} \quad \text{and} \quad U(z) = \sum_{k=0}^{\infty} u(k)z^{-k} \qquad (4)$$

be the z-transform of $y(k)$ and $u(k)$, respectively. And let

$$\bar{Y}(z) = \sum_{k=0}^{\infty} \bar{y}(k)z^{-k} =: \begin{bmatrix} \bar{Y}_0(z) \\ \vdots \\ \bar{Y}_{N-1}(z) \end{bmatrix}$$

and

$$\bar{U}(z) = \sum_{k=0}^{\infty} \bar{u}(k)z^{-k} =: \begin{bmatrix} \bar{U}_0(z) \\ \vdots \\ \bar{U}_{N-1}(z) \end{bmatrix} \qquad (5)$$

be the z-transform of $\bar{y}(k)$ and $\bar{u}(k)$, respectively. The z-transforms are related by [12]

$$Y(z) = \sum_{l=0}^{N-1} z^{-l}\bar{Y}_l(z^N) \qquad (6)$$

and

$$\bar{U}_i(z^N) = \frac{z^i}{N}\sum_{k=0}^{N-1} W_N^{ki}U(zW_N^k) \qquad (7)$$

where $W_N = e^{-j\frac{2\pi}{N}}$. Let $G(z) = [G_{l,k}(z)]_{l=0,k=0}^{N-1}$. Since $\bar{Y}(z) = G(z)\bar{U}(z)$, thus

$$\bar{Y}_l(z) = \sum_{i=0}^{N-1} G_{l,i}(z)\bar{U}_i(z). \qquad (8)$$

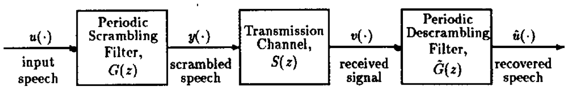

$u(\cdot)$
input speech
Periodic Scrambling Filter, $G(z)$
$y(\cdot)$
scrambled speech
Transmission Channel, $S(z)$
$v(\cdot)$
received signal
Periodic Descrambling Filter, $\tilde{G}(z)$
$\hat{u}(\cdot)$
recovered speech

Fig. 1. Generic model of a scrambling system.

From (6)–(8) we have

$$Y(z) = \frac{1}{N}\sum_{k=0}^{N-1}\left(\sum_{i=0}^{N-1} W_N^{ki}\left(\sum_{l=0}^{N-1} z^{i-l}G_{l,i}(z^N)\right)\right)U(zW_N^k)$$
$$=: \frac{1}{N}\sum_{k=0}^{N-1} H_k(z)U(zW_N^k) \qquad (9)$$

where $H_k(z) = \sum_{i=0}^{N-1} W_N^{ki}\sum_{l=0}^{N-1} z^{i-l}G_{l,i}(z^N)$, $k = 0,\cdots,N-1$, is called the transfer function of the kth shifted band, or simply called the *kth transfer function* of the periodic system (1). The relàtion (9) shows that the output spectrum is a sum of shifted and shaped versions of the input spectrum. We note that the relation (9) can also be obtained through filter bank representation of periodic filters [16, ch. 10].

III. Model and Analysis of Speech Scrambling Systems

Typically, most scrambling filters can be regarded as a periodic filters. Based on this fact, we develop the following generic model for a scrambling system.

A. Model Setup

The functional block diagram of a typical speech scrambling system, shown in Fig. 1, consists of a periodic scrambling filter, transmission channel, and a periodic descrambling filter. In our analysis, the periodic scrambling filter is an N-periodic filter whose block transfer matrix is $G(z)$; the transmission channel is characterized by a transfer function $S(z)$; and the periodic descrambling filter is an N-periodic filter whose block transfer matrix is $\tilde{G}(z)$.

Referring to Fig. 1, we have the following relations in frequency domain:

a) The scrambled signal $Y(z)$ is obtained by

$$Y(z) = \frac{1}{N}\sum_{k=0}^{N-1} H_k(z)U(zW_N^k) \qquad (10)$$

where $U(z)$ is the input signal

$$H_k(z) = \sum_{i=0}^{N-1} W_N^{ki}F_i(z) \qquad (11)$$

and

$$F_i(z) = \sum_{l=0}^{N-1} z^{i-l}G_{l,i}(z^N). \qquad (12)$$

Remarks: i) Obviously, $F_i(z) \in \mathbb{R}_p(z)$ and $H_k(z) \in C_p(z)$, where $\mathbb{R}_p(z)$ and $C_p(z)$ denote the set of proper rational functions in z with real and complex coefficients respectively.

ii) In the polyphase model, the N-periodical filter is implemented by connecting N time-invariant filters

$P_l(z)$, $l = 0, \cdots, N-1$ in parallel and the output signal is obtained by selecting the output signals from these filters periodically. It can be shown that the relation between $P_l(z)$ and $G(z)$ is given by $P_l(z) = \sum_{i=0}^{N-1} z^{i-l} G_{l,i}(z^N)$. Note that $G_{l,i}(z)$ for $0 \leq l \leq N-1$ is the polyphase component of the N-component polyphase representation of $z^{-i}F_i(z)$, while $G_{l,i}(z)$ for $0 \leq i \leq N-1$ is just the polyphase component of the second type polyphase decomposition of $z^{l-(N-1)}P_l(z)$.

b) The signal $V(z)$ received at the input of the descrambling filter is given by

$$V(z) = S(z)Y(z). \tag{13}$$

c) The recovered signal $\hat{U}(z)$ is obtained by

$$\hat{U}(z) = \frac{1}{N}\sum_{k=0}^{N-1} \tilde{H}_k(z)V(zW_N^k) \tag{14}$$

where

$$\tilde{H}_k(z) = \sum_{i=0}^{N-1} W_N^{ki}\tilde{F}_i(z) \tag{15}$$

$$\tilde{F}_i(z) = \sum_{l=0}^{N-1} z^{i-l}\tilde{G}_{l,i}(z^N) \tag{16}$$

and

$$\tilde{G}(z) = [\tilde{G}_{l,k}(z)]_{l=0,k=0}^{N-1}. \tag{17}$$

By (10), (13), and (14), the relation between the input signal and the recovered output signal in frequency domain is given by

$$\hat{U}(z) = \frac{1}{N^2}\sum_{l=0}^{N-1} \tilde{H}_l(z)S(zW_N^l)\sum_{k=0}^{N-1} H_k(zW_N^l)U(zW_N^{l+k}). \tag{18}$$

For simplicity, rewrite (18) as

$$\hat{U}(z) = \sum_{k=0}^{N-1} M_k(z)U(zW_N^k) \tag{19}$$

where

$$M_k(z) = \frac{1}{N^2}\sum_{l=0}^{N-1} \tilde{H}_l(z)S(zW_N^l)H_{[k-l]}(zW_N^l) \tag{20}$$

and $[k-l] := (k-l)$ modulo N, $0 \leq [k-l] \leq N-1$. We call $M_k(z)$ the *kth recovered gain* of the scrambling system shown in Fig. 1. We note that for perfect speech recovery we should have

$$M_0(z) = 1 \quad \text{and} \quad M_k(z) = 0 \quad \text{for } k = 1, \cdots, N-1. \tag{21}$$

From a design point of view, the goal is to find transfer matrices $G(z)$ and $\tilde{G}(z)$ so that the transfer functions $M_k(z)$ in (19) are as desired. Clearly G and $\tilde{G}$ must be real rational to be realizable in the form (2). The rational functions $G_{l,i}(z)$ for $0 \leq l \leq N-1$ can be uniquely obtained from $F_i(z)$ since they are simply the polyphase component of the N-component polyphase representation of $z^{-i}F_i(z)$ [5], [16]. In particular, G is real rational iff F_i, $i = 0, \cdots, N-1$, is real rational. The following result shows that G is real rational iff $H_k(z)$ have a certain symmetric property. The proof is straightforward and hence omitted.

Proposition 3.1: Suppose that $F_i(z)$ and $H_k(z)$ are related by (11). Then, the following holds:

i) For $i = 0, \cdots, N-1$, $F_i(z)$ is uniquely determined by

$$F_i(z) = \frac{1}{N}\sum_{k=0}^{N-1} W_N^{-ik}H_k(z). \tag{22}$$

ii) $F_i(z) \in \mathbb{R}_p(z)$ for each i if and only if

$$H_{[N-k]}(z) = H_k^*(z) \tag{23}$$

holds for $k = 0, \cdots, N-1$, where $H_k^*(z)$ is obtained from $H_k(z)$ by conjugating the coefficients.

Similar results hold for $\tilde{F}_i(z)$ and $\tilde{H}_k(z)$. With the constraint (23) on H_k and $\tilde{H}_k$, the design of scrambling system is equivalent to choosing the transfer functions $H_k(z)$ and $\tilde{H}_k(z)$.

B. Performance Analysis

Commonly used criteria for evaluating the performance of a scrambling system are stability, descrambling capability, security level (residual intelligibility and key space), distortion caused by channel characteristics, distortion caused by synchronization error, expansion of bandwidth, amplification of noise, operation delay, etc. The generic model proposed here is useful for evaluating the performance of scrambling systems based on the stated criteria.

To quantitatively evaluate performance we use a performance index called the deformation factor (DF) of the scrambling system, defined as

$$\mathrm{DF} = \frac{1}{2\pi}\int_0^{2\pi}\left((|M_0(e^{j\theta})| - 1)^2 + \sum_{k=1}^{N-1}|M_k(e^{j\theta})|^2\right)d\theta. \tag{24}$$

We note that for perfect recovery (21), DF is zero. Roughly, we expect that two scrambling systems with the same DF have similar quality of speech recovery.

With the generic model and DF, the performance of a scrambling system can be evaluated as follows:

1) Stability: The scrambling and descrambling filters are BIBO stable iff $G(z)$ and $\tilde{G}(z)$ have all their poles inside the unit disk.
2) Descrambling capability: The descrambling capability is a basic requirement for a scrambling system. It can be evaluated by choosing $S(z) = 1$ and calculating the associated DF. Small DF indicates good recovery quality.
3) Distortion caused by frame synchronization error: An m-sample synchronization error in the descrambling process can be characterized by $S(z) = z^{-m}$. With this specific form of $S(z)$, we obtain the recovered gains by (20) and evaluate the distortion due to the frame synchronization error through the index DF. In order to estimate the net effect of the frame synchronization

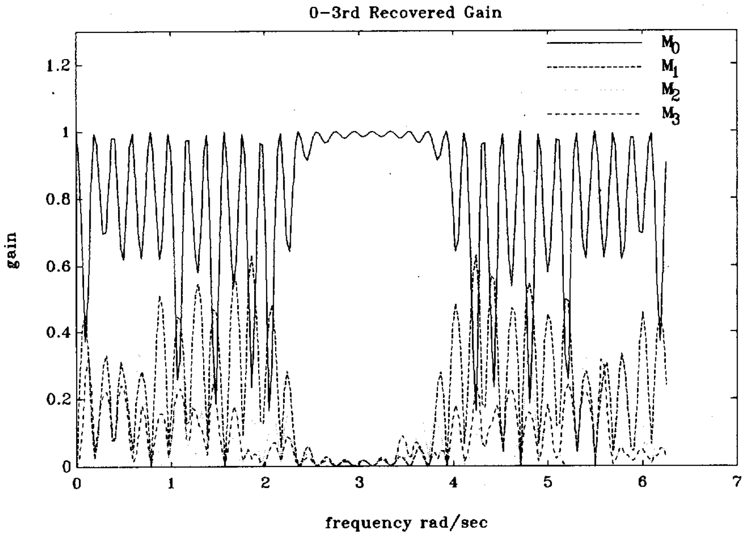


Fig. 2. First four recovered gains as synchronization error = 3 samples.

error, the difference between the resultant DF and the DF without synchronization error can be used as a new performance index. More precisely, the new performance index ΔDF = DF (calculated with $S(z) = z^{-m}$) − DF (calculated with $S(z) = 1$). For properly designed systems, $\Delta\text{DF} \geq 0$ in general. Large ΔDF indicates sensitivity to synchronization error.

4) Distortion caused by bandwidth expansion: Distortion caused by the bandwidth expansion can be evaluated as follows. Let $S(z)$ be the lowpass filter with bandwidth equal to the channel bandwidth. Calculate DF corresponding to this $S(z)$ but with a little modification so that the integral average is taken over the bandwidth of the input speech under consideration. We also can use ΔDF instead of DF to indicate the net distortion caused by the bandwidth expansion. Large ΔDF indicates severe distortion caused by expansion of bandwidth.
5) Noise amplification: Noises introduced during transmission are amplified by the descrambling matrix $\tilde{G}(z)$. If noise power distribution is known, the amplified noise power can be calculated. Typically, if the descrambling matrix $\tilde{G}(z)$ is designed to be orthogonal, i.e., $|\tilde{G}_{k,i}(z)\tilde{G}^*_{i,l}(z)| = \delta_{kl}$ for all z, then it will not enhance noise [17].
6) Operation delay: In traditional scrambler design, the operation delay usually equals the dimension of the blocked input vector, i.e., the period of the scrambler N. However, if the scrambler is represented by the block transfer matrix $G(z)$, then the minimal delay is proved to be the upper bandwidth of $G(\infty)$ [9], which is never greater than N.

C. Analyze an FFT Scrambler by Using the Generic Model

As an example, we analyze an FFT scrambler [15] with frame length of 32 samples at sampling frequency of 8 kHz. The frequency resolution is 250 Hz. Assume that the channel bandwidth is 3 kHz and the speech bandwidth is below 2500 Hz. Only 11 FFT coefficients corresponding to frequencies from 250 to 2750 Hz are rearranged according to a scrambling key; thus, the bandwidth does not expand. In the descrambling process, the 11 FFT coefficients are permuted in the reverse order. The scrambling matrix $G(z)$ and descrambling matrix $\tilde{G}(z)$ are simply given by $G(z) = z^{-1}Q^{-1}PQ$ and $\tilde{G}(z) = z^{-1}Q^{-1}P^{-1}Q$, respectively, where Q is the 32×32 DFT matrix and P is the permutation matrix. Thus we have perfect recovery and DF equals zero. To observe the effect of synchronization error, let $S(z) = z^{-m}$ and calculate the recovered gains. Fig. 2 shows the spectrum of the first four recovered gains with $m = 3$. The calculated DF is 0.38. For $m = 16$, DF increases to 0.58. To observe the effect of band limitation in the transmission channel, we calculate the recovered gains with $S(z)$ equals a lowpass filter whose bandwidth is 0.75π. Fig. 3 shows the first four recovered gains in this case. The corresponding DF is 0.49. In other words, the bandwidth of the scrambled signal does expand. The first four transfer functions $H_0(z) \sim H_3(z)$ are shown in Fig. 4. From Fig. 4, we can see how the input spectrum is shaped and split.

IV. Necessary and Sufficient Conditions for a Scrambling System to be Insensitive to Synchronization Error

With the generic model, it is simple to design a scrambling system that is insensitive to synchronization error.

A. Problem Formulation

It is argued in [6] that human ear is insensitive to the phase of a speech signal. A scrambling system satisfying

(S1) $|M_0(e^{j\theta})| = 1 \ \forall\theta \in [0, 2\pi)$,

(S2) $|M_i(e^{j\theta})| = 0$ for $i = 1, \cdots, N-1$, $\forall\theta \in [0, 2\pi)$

can be regarded as perfect so far as human listeners are concerned. Note that (S1) and (S2) imply that DF = 0.

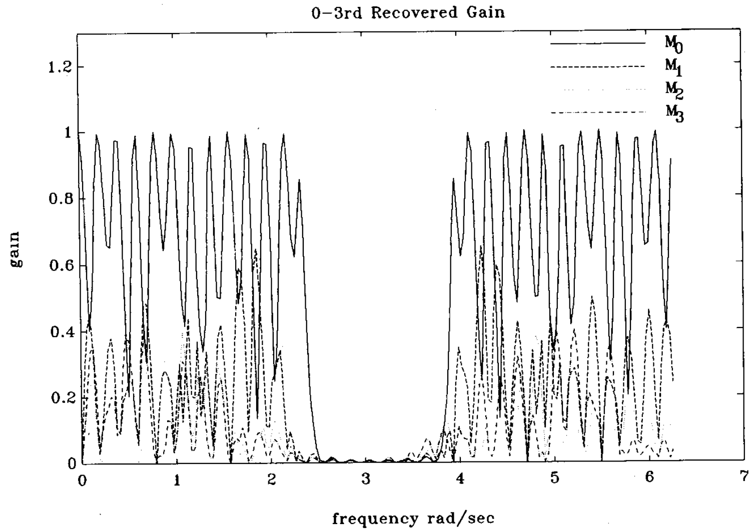


Fig. 3. First four recovered gains as channel bandwidth = $\frac{3}{4}\pi$.

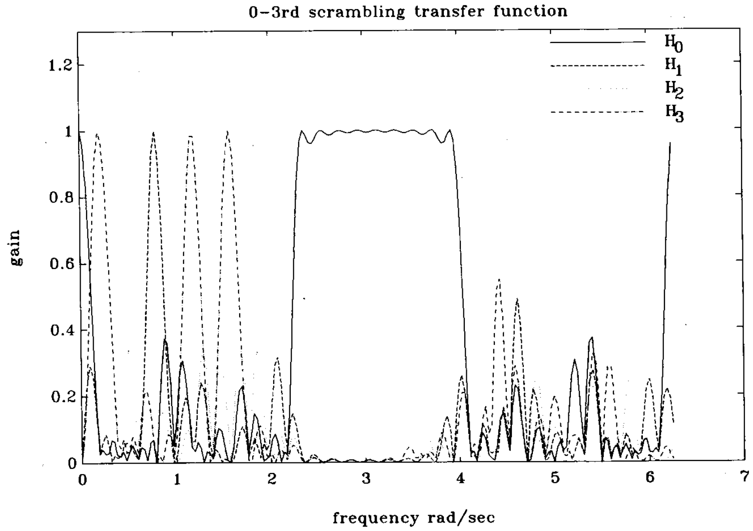


Fig. 4. Frequency response of the first four shifted transfer functions.

Since an m-sample synchronization error is modeled as $S(z) = z^{-m}$, a speech scrambling system that is insensitive to frame synchronization error should have the recovered gains obtained by (20) satisfying (S1) and (S2) for $S(z) = z^{-m}$, $m = 0, 1, \cdots, N-1$. More precisely, let

$$\hat{M}_{k,m}(e^{j\theta}) = \frac{e^{-jm\theta}}{N^2}\sum_{l=0}^{N-1} W_N^{lm}\tilde{H}_l(e^{j\theta})H_{[k-l]}(e^{j\theta}W_N^l) \quad (25)$$

which is the kth recovered gain computed at m-sample synchronization error, then a necessary and sufficient condition for the scrambling system to be insensitive to synchronization error can be stated as

(M1) $|\hat{M}_{0,m}(e^{j\theta})| = 1$ and

(M2) $|\hat{M}_{k,m}(e^{j\theta})| = 0$ for $k = 1, \cdots, N-1$,

for $\theta \in [0, 2\pi)$ and $m = 0, 1, \cdots, N-1$.

To find a solution $\{H_l(e^{j\theta})\}_{l=0}^{N-1}$ and $\{\tilde{H}_l(e^{j\theta})\}_{l=0}^{N-1}$ that satisfies (M1) and (M2), we formulate the problem as follows. Rewrite (25) as shown in (26) at the bottom of the next page, where

$$\hat{H}_{k,l}(e^{j\theta}) := \tilde{H}_l(e^{j\theta})H_{[k-l]}(e^{j\theta}W_N^l). \quad (27)$$

We note that the last matrix on the right-hand side of (26) is the N-point DFT matrix. By (26) we have

$$\hat{H}_{k,l}(e^{j\theta}) = N\sum_{m=0}^{N-1} e^{jm\theta}(W_N^*)^{ml}\hat{M}_{k,m}(e^{j\theta}),$$
$$0 \le k, l \le N-1. \quad (28)$$

From (26), (M1) is equivalent to

$$\left|\sum_{l=0}^{N-1} W_N^{lm}\hat{H}_{0,l}(e^{j\theta})\right| = N^2, \ \forall \ 0 \le m \le N-1, 0 \le \theta < 2\pi. \tag{29}$$

A necessary condition for (29) is that the union of the passbands of $\hat{H}_{0,l}$ for $l = 0, \cdots, N-1$ is the whole frequency region. More precisely, let $\Theta = [0, 2\pi)$ and let

$$I_{l,k} := \{\theta \mid |H_l(e^{j\theta}W_N^k)| \ne 0, \ \theta \in \Theta\}$$

and

$$J_{l,k} := \{\theta \mid |\tilde{H}_l(e^{j\theta}W_N^k)| \ne 0, \ \theta \in \Theta\}. \tag{30}$$

Then, (M1) implies that

(P1) $\bigcup_{l=0}^{N-1}(J_{l,0} \cap I_{[N-l],l}) = \Theta$.

From (28), it follows that (M2) is equivalent to

$$\hat{H}_{k,l}(e^{j\theta}) = 0 \quad \text{for } k = 1, \cdots, N-1, \ l = 0, \cdots, N-1. \tag{31}$$

Equation (31) means that for $k = 1, \cdots, N-1, l = 0, \cdots, N-1$, the passbands of $\tilde{H}_l(e^{j\theta})$ and $H_{[k-l]}(e^{j\theta}W_N^l)$ have no intersection, or equivalently:

(P2) $J_{l,0} \cap I_{[k-l],l} = \emptyset$, for $k = 1, \cdots, N-1$, $l = 0, \cdots, N-1$.

Remarks:

1) If $H_l(e^{j\theta})$ and $\tilde{H}_l(e^{j\theta}), 0 \le l \le N-1$, satisfies (P1) and (P2), and in addition satisfies the magnitude condition (29), then it satisfies (M1) and (M2) also.
2) $I_{l,k}$ is the union of the frequency bands that are $k\frac{2\pi}{N}$ circular left-shifted from the passbands of H_l. Note that $I_{l,0}$ is the union of passbands of $H_l(e^{j\theta})$.
3) The set $J_{l,0} \cap I_{[k-l],l}$ equals the union of the passbands of $\hat{H}_{k,l}(e^{j\theta})$.
4) For practical applications, we only consider $I_{i,0}$, $J_{i,0}$ to be the union of a finite number of intervals with nonzero length.

B. Necessary and Sufficient Conditions

The following theorem gives necessary and sufficient conditions on H_i and $\tilde{H}_i$ so that (M1) and (M2) are satisfied.

Theorem 4.1: The set $\{H_l(e^{j\theta})\}_{l=0}^{N-1}$ and $\{\tilde{H}_l(e^{j\theta})\}_{l=0}^{N-1}$ is a solution of (M1) and (M2) if and only if

i) $$\bigcup_{l=0}^{N-1} I_{l,0} = \Theta, \quad I_{l,0} \cap I_{k,0} = \emptyset \quad \text{for } l \ne k, \tag{32}$$

ii) $$\bigcup_{l=0}^{N-1} J_{l,0} = \Theta, \quad J_{l,0} \cap J_{k,0} = \emptyset \quad \text{for } l \ne k, \tag{33}$$

iii) $$J_{l,0} = I_{[N-l],l} \quad \text{for } l = 0, \cdots, N-1, \tag{34}$$

iv) For $l = 0, \cdots, N-1$, $$|\tilde{H}_l(e^{j\theta})H_{[N-l]}(e^{j\theta}W_N^l)| = \begin{cases} N^2, & \text{for } \theta \in J_{l,0}, \\ 0, & \text{for } \theta \in \Theta, \ \theta \notin J_{l,0} \end{cases} \tag{35}$$

where $I_{i,j}$ and $J_{i,j}$ are defined in (30).

Hence, a scrambling system that is insensitive to frame synchronization error has the property that the passbands of $H_l(e^{j\theta})$ for $l = 0, \cdots, N-1$ fill up $[0, 2\pi)$ while none of the passbands intersect. Note that the passbands of the descrambling filter are uniquely determined from the passbands of the scrambling filter by (34). The following two lemmas, whose proofs are given in the Appendix, are used in the proof of Theorem 4.1.

Lemma 4.2: Let A, B, C, D_i and $E_i, i = 1, 2, \cdots, N$, be subsets of Ω. The following holds:

a) If $A \cap B = \emptyset$, then $(C \cap A) \subseteq (C - B)$;
b) $(A - (B \cup C)) \cap (B - (A \cup C)) = \emptyset$;
c) If $\bigcup_{i=1}^{N}(D_i - \bigcup_{j=1, j\ne i}^{N} D_j) = \Omega$, then $D_i \cap D_j = \emptyset$ for $i \ne j$;
d) If $\bigcup_{i=1}^{N}(E_i \cap D_i) = \Omega$ and $D_i \cap D_j = \emptyset$ for $i \ne j$, then $D_i \subseteq E_i, \forall i$.

Lemma 4.3: Let $\Theta = [0, 2\pi)$. Let S be the family of all subsets of Θ that are finite unions of disjoint intervals. For $i = 1, \cdots, N$, let $\theta_i \in \Theta$ and let $f_i : \Theta \to \Theta$ be the circular shift function defined by

$$f_i(x) = (x + \theta_i) \text{ modulo } 2\pi, 0 \le x < 2\pi. \tag{36}$$

Under these conditions, if

(L1) $A_i \in S$ and $\bigcup_{i=1}^{N} A_i = \Theta$; and **(L2)** $f_i(A_i) \cap f_j(A_j) = \emptyset$, for $i \ne j$, $1 \le i, j \le N$, then $\bigcup_{i=1}^{N} f_i(A_i) = \Theta$ and $A_i \cap A_j = \emptyset$, for $i \ne j$.

We note that $I_{l,m} = f_m(I_{l,0})$, where f_m is as defined in (36) with $\theta_m = -\frac{2\pi}{N}m$. Also, $I_{i,m} \circ I_{l,m} = f_m(I_{i,0}) \circ f_m(I_{l,0}) = f_m(I_{i,0} \circ I_{l,0})$, where $\circ$ is either union, intersection or difference operation of sets.

Proof of Theorem 4.1:

(Sufficiency): By (32), we have

$$I_{[N-l],l} \cap I_{[k-l],l} = \emptyset, \quad \text{for } k \ne N \text{ and } l = 0, \cdots, N-1. \tag{37}$$

$$N^2 \begin{bmatrix} \hat{M}_{0,0}(e^{j\theta}) & e^{j\theta}\hat{M}_{0,1}(e^{j\theta}) & \cdots & e^{j(N-1)\theta}\hat{M}_{0,N-1}(e^{j\theta}) \\ \hat{M}_{1,0}(e^{j\theta}) & e^{j\theta}\hat{M}_{1,1}(e^{j\theta}) & \cdots & e^{j(N-1)\theta}\hat{M}_{1,N-1}(e^{j\theta}) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{M}_{N-1,0}(e^{j\theta}) & e^{j\theta}\hat{M}_{N-1,1}(e^{j\theta}) & \cdots & e^{j(N-1)\theta}\hat{M}_{N-1,N-1}(e^{j\theta}) \end{bmatrix}$$
$$= \begin{bmatrix} \hat{H}_{0,0}(e^{j\theta}) & \hat{H}_{0,1}(e^{j\theta}) & \cdots & \hat{H}_{0,N-1}(e^{j\theta}) \\ \hat{H}_{1,0}(e^{j\theta}) & \hat{H}_{1,1}(e^{j\theta}) & \cdots & \hat{H}_{1,N-1}(e^{j\theta}) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{H}_{N-1,0}(e^{j\theta}) & \hat{H}_{N-1,1}(e^{j\theta}) & \cdots & \hat{H}_{N-1,N-1}(e^{j\theta}) \end{bmatrix} \times \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & W_N & \cdots & W_N^{N-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & W_N^{N-1} & \cdots & W_N^{(N-1)^2} \end{bmatrix} \tag{26}$$

It then follows from (34) that $J_{l,0} \cap I_{[k-l],l} = \emptyset$, or equivalently $\hat{H}_{k,l}(e^{j\theta}) = \tilde{H}_l(e^{j\theta}) H_{[k-l]}(e^{j\theta} W_N^l) \equiv 0$ for $k \neq N$ and $l = 0, \cdots, N-1$, and hence, by (25), (M2) is satisfied. By (34), it follows that the passbands of $\hat{H}_{0,l}$ equal the passbands of $\tilde{H}_l$. Since i) by (33) the intersection of passbands of $\hat{H}_{0,k}$ and $\hat{H}_{0,l}$ for $k \neq l$ is empty; ii) the union of passbands of each $\hat{H}_{0,l}$ is Θ; and iii) by (35) $|\hat{H}_{0,l}(e^{j\theta})| = N^2$ for θ belongs to the passbands of $\hat{H}_{0,l}$, for each $\theta \in \Theta$, we have that the summation in (25) reduces to only one term and thus (M1) is satisfied.

(*Necessity*): Assume that $\{H_i(e^{j\theta})\}_{i=0}^{N-1}$ and $\{\tilde{H}_i(e^{j\theta})\}_{i=0}^{N-1}$ is a solution of (M1) and (M2), then (P1) and (P2) are satisfied. It follows from (P2) and Lemma 4.2(a) that

$$(I_{[N-l],l} \cap J_{l,0}) \subseteq \left(I_{[N-l],l} - \bigcup_{k=1}^{N-1} I_{[k-l],l} \right), \quad \text{for } l = 0, \cdots, N-1. \tag{38}$$

It then follows from (P1) and (38) that

$$\bigcup_{l=0}^{N-1} \left(I_{[N-l],l} - \bigcup_{k=1}^{N-1} I_{[k-l],l} \right) = \Theta. \tag{39}$$

By Lemma 4.2(b) we know that for $p \neq q$, $(I_{p,0} - \bigcup_{k=0,k\neq p}^{N-1} I_{k,0}) \cap (I_{q,0} - \bigcup_{k=0,k\neq q}^{N-1} I_{k,0}) = \emptyset$, hence with $p = [N-l]$ and $q = [N-m]$, we have

$$\left(I_{[N-l],0} - \bigcup_{k=1}^{N-1} I_{[k-l],0} \right) \cap \left(I_{[N-m],0} - \bigcup_{k=1}^{N-1} I_{[k-m],0} \right) = \emptyset, \quad \text{for } l \neq m. \tag{40}$$

By (39) and (40), it then follows from Lemma 4.3 that

$$\bigcup_{l=0}^{N-1} \left(I_{[N-l],0} - \bigcup_{k=1}^{N-1} I_{[k-l],0} \right) = \Theta \tag{41}$$

and

$$\left(I_{[N-l],l} - \bigcup_{k=1}^{N-1} I_{[k-l],l} \right) \cap \left(I_{[N-m],m} - \bigcup_{k=1}^{N-1} I_{[k-m],m} \right) = \emptyset, \quad \text{for } l \neq m. \tag{42}$$

Rewrite (41) as

$$\bigcup_{i=0}^{N-1} \left(I_{i,0} - \bigcup_{k=0,k\neq i}^{N-1} I_{k,0} \right) = \Theta \tag{43}$$

which implies that

$$\bigcup_{i=0}^{N-1} I_{i,0} = \Theta \tag{44}$$

and by Lemma 4.2(c), that

$$I_{i,0} \cap I_{l,0} = \emptyset, \quad \text{for } i \neq l. \tag{45}$$

Thus (32) holds. We now show that (34) holds. With (45), (42) reduces to

$$I_{[N-l],l} \cap I_{[N-m],m} = \emptyset, \quad \text{for } l \neq m. \tag{46}$$

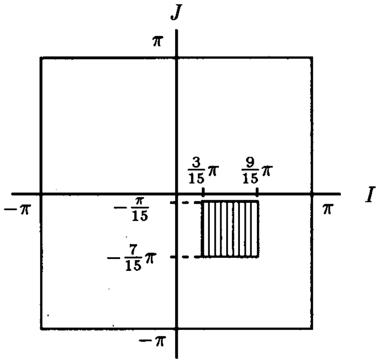


Fig. 5. Graphic representation of correspondence relation.

From (46), (P1) and Lemma 4.2(d), it follows that

$$I_{[N-l],l} \subseteq J_{l,0}. \tag{47}$$

However, by (P2) we have that $J_{l,0} \cap (\bigcup_{k=1}^{N-1} I_{[k-l],l}) = \emptyset$ and hence

$$J_{l,0} \subseteq \left(\Theta - \bigcup_{k=1}^{N-1} I_{[k-l],l} \right) = I_{[N-l],l} \tag{48}$$

where the last equality follows from (44) and (45). Thus (34) holds. And from (P1), (46), and (34), (33) follows.

It remains to show (35). From (32)–(34) it follows that for any $\theta \in \Theta$, the summation in (25) reduces to only one term, i.e.

$$\hat{M}_{0,m}(e^{j\theta}) = \frac{e^{-jm\theta}}{N^2} W_N^{lm} \tilde{H}_l(e^{j\theta}) H_{[N-l]}(e^{j\theta} W_N^l), \quad \text{for some } 0 \leq 1 < N-1. \tag{49}$$

Thus (35) follows from (M1). □

V. A New Design Algorithm

Theorem 4.1 gives precise constraints on the passbands of a set of analysis filters H_i and synthesis filters $\tilde{H}_i$ so that the scrambling system is insensitive to synchronization error. The conditions (32)–(35), although not very easy to use directly as a design tool, can be easily represented graphically. In the following, we introduce a graphical representation of (32)–(34) and the condition (23) called *correspondence map* and the associated operations that are useful for design purpose. Then, we propose a design algorithm.

A. Correspondence Map

To discuss the graphical representation of (32)–(34), we assume that the passbands of H_i and $\tilde{H}_j$ are all finite union of closed intervals in $[-\pi, \ \pi)$. We will call the axis for the passbands of H_i's the I-axis and draw it as a horizontal line segment. Similarly, we use J-axis for the passbands of $\tilde{H}_i$'s and draw it vertically. This is shown in Fig. 5. The condition (32) simply says that each point on I-axis should belong to one and only one passband of H_i and the passbands of H_i's can otherwise be arbitrarily assigned. Similar conditions for J-axis follow from (33).

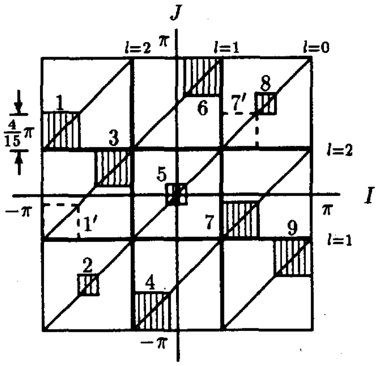


Fig. 6. Correspondence map for $N = 3$.

The divisions and assignments of passbands are constrained, however, since the passbands of $H_{[N-i]}(z)$ and $\tilde{H}_i(z)$ are related by (34). For example, if $N = 3$ and $[\frac{3}{15}\pi, \frac{9}{15}\pi]$ is an interval in I-axis corresponding to $H_2(z)$, then there must exist an interval $[-\frac{7}{15}\pi, -\frac{\pi}{15}]$ in J-axis corresponding to $\tilde{H}_1(z)$. Therefore, (34) defines the relation between the passbands of $H_{[N-i]}(z)$ and $\tilde{H}_i(z)$, the quantized shift relation between the passbands of the *correlated pair* $\{H_{[N-i]}, \tilde{H}_i(z)\}$, or equivalently, $\{H_i(z), \tilde{H}_{[N-i]}(z)\}$. Such a relation is represented graphically as a shaded square in the I-J plane as shown in Fig. 5. Such a square is called a *cell*. The I-axis and J-axis frequency regions, i.e., the projections in I-axis and J-axis, of the cell stand for the passbands of the correlated pair. Note that these cells are not allowed to be located arbitrarily since the quantized shift relation should hold. To see the shift relation, we draw N slant lines called the *mapping lines* in the I-J plane, as shown in Fig. 6. The mapping lines define the shift relations between I-axis and J-axis. If a cell lies on a mapping line, i.e., the diagonal of the cell is on a mapping line, then the quantized shift relation is satisfied. The cells lying in the ith mapping line represent the passbands of $H_i(z)$ and $\tilde{H}_{[N-i]}(z)$. Thus, cell 1 in Fig. 6 shows that $[-\pi, -\frac{11}{15}\pi]$ is a passband of $H_2(z)$ and $[\frac{5}{15}\pi, \frac{9}{15}\pi]$ is a passband of $\tilde{H}_1(z)$.

So far we have introduced the basic construction of the graphical representation. To locate these cells, both axes are divided into N regions, where N is the period. Thus there are totally N^2 square regions in the I-J plane. These square regions are called *blocks*. Fig. 6 shows a possible breakup of passbands of H_i and $\tilde{H}_i$ that satisfies (32)–(34). We call such a graphical representation a correspondence map.

In addition, we need to consider the condition (23). According to Proposition 3.1, we have that $H_{[N-k]}(z) = H_k^*(z)$. In other words, the I-axis frequency regions (i.e. the passbands) corresponding to $H_{[N-k]}(z)$ and $H_k(z)$ should be symmetric about the origin. Similarly, we also have $\tilde{H}_{[N-k]}(z) = \tilde{H}_k^*(z)$ or equivalently $\tilde{H}_k(z) = \tilde{H}_{[N-k]}^*(z)$, and hence the J-axis frequency regions corresponding to $\tilde{H}_k(z)$ and $\tilde{H}_{[N-k]}(z)$ are also symmetric about the origin. In the graphic representation, this means that the cells corresponding to the correlated pair $\{H_{[N-k]}(z), \tilde{H}_k(z)\}$ and the cells corresponding to the correlated pair $\{H_k(z), \tilde{H}_{[N-k]}(z)\}$ should be symmetric about both I- and J-axes.

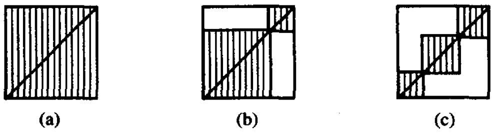


Fig. 7. Decompositions of cells. (a) Original cell; (b) and (c), decomposed cells.

Based on the discussions above, a correspondence map represents a realizable[1] scrambling system that is insensitive to frame synchronization error if and only if

(C1) all cells lie on the mapping lines;

(C2) the union of the I-axis (J-axis) frequency regions of these cells fills up $[-\pi, \ \pi)$;

(C3) the cells are pairwise disjoint;

(C4) all cells are symmetric about both I- and J-axes.

To make the graphical representation useful for design purpose, we introduce a set of graphic operations under which the constraints (C1)–(C4) are preserved. We begin with a few definitions. Two cells of the same size locate at the same relative position within their corresponding blocks are called *correlated cells*. Two cells symmetric with respect to the origin are called *symmetric cells*. Note that symmetric cells are of the same size. Any two correlated cells can be interchanged in horizontal (or vertical) position without violating the constraints (C1)–(C3). In Fig. 6, cells one and seven are correlated cells, so are cells one and four. Cell one and cell nine are symmetric cells. We can move cells one and seven into cell 1' and 7', respectively, while (C1)–(C3) still hold. A cell located at the center of a block is called a *center cell*. A cell that is not a center cell is called a *side cell*. In Fig. 6, cells two, five, and eight are center cells, the rest are side cells.

Let us introduce the graphic operations that preserve (C1)–(C3). The first operation is called *decomposition*. Every cell can be decomposed into smaller cells as shown in Fig. 7. The purpose of decomposition is for interchange. Any two symmetric center cells can be interchanged in horizontal (or vertical) position without violating the constraints (C1)–(C4), since they are correlated and symmetric. This type of interchange is called *symmetric change* and is illustrated in Fig. 8 for $N = 2$. Any two correlated side cells can be interchanged in horizontal (or vertical) position without violating (C1)–(C3), and to preserve (C4) their corresponding symmetric cells should also be interchanged accordingly. This type of interchange is called *asymmetric change* and is illustrated in Fig. 9.

B. Design Procedure

With the graphical representation and operations discussed, we are ready to propose a design procedure. The basic idea is very simple: to start with a simple correspondence map that satisfies (C1)–(C4) and use the three operations just defined to obtain a suitable correspondence map that would define the filters H_i and $\tilde{H}_j$. The simplest correspondence map is the *identity map* with all cells lying on the diagonal, as shown

[1] Here *realizable* means the existence of ideal bandpass filters.

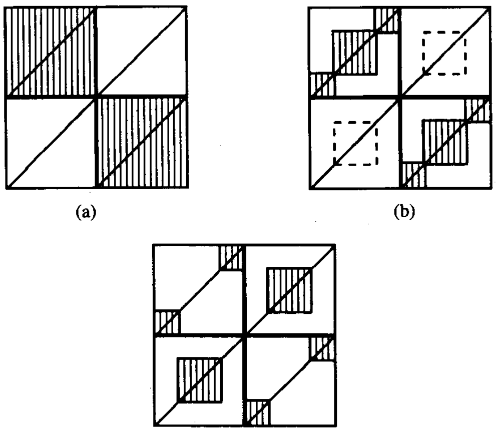


Fig. 8. Symmetric change. (a) Original cells; (b) decomposition; (c) interchange.

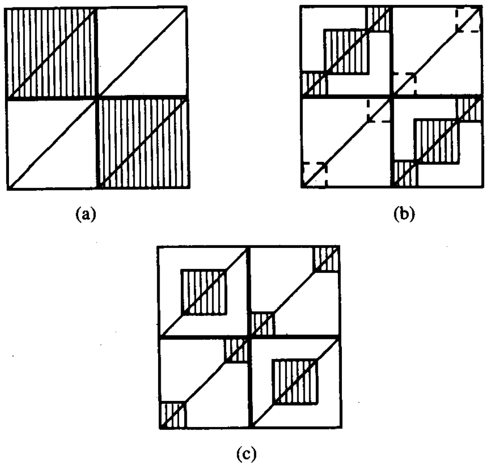


Fig. 9. Asymmetric change. (a) Original cells; (b) decomposition; (c) interchange.

in Fig. 10 for $N = 3$. Hence, the identity map corresponds to a system whose scrambled speech is just the input speech. Once the correspondence map satisfying (C1)–(C4), then the scrambling system is insensitive to synchronization error and realizable.

In addition to the conditions (C1)–(C4) on the correspondence map, the design still has to satisfy the magnitude constraint (35). A simple way to satisfy (35) is to have the ideal bandpass filters $H_i(z)$ and $\tilde{H}_i(z)$ to have magnitude N in the passbands. However, (35) will still be satisfied if the filters $H_i(z)$ and $\tilde{H}_i(z)$ are modified by any rational stable and minimum phase filter $A_i(z)$ to become $A_i(z)H_i(z)$ and $A_i^{-1}(z)\tilde{H}_{[N-i]}(z)$, respectively. Note that the condition (23) requires that $A_{[N-i]}(z) = A_i^*(z)$. The modification increases the flexibility in the design of scrambling systems.

To discuss the bandwidth expansion problem, let us consider (10) again. From (10), to keep the spectrum of the scrambled speech within the bandwidth of the original speech, say θ_B, the aliasing terms $H_k(e^{j\theta})U(e^{j\theta}W_N^k)$, $k \neq 0$ must vanish for $|\theta| > \theta_B$. A sufficient condition for avoiding bandwidth expansion is that the passbands of $H_k(z)$, $k \neq 0$ are all subsets of $[\theta_B, \theta_B]$; in other words, the cells corresponding to each $H_k(z)$, $k \neq 0$ should all lie inside the square centered at $(0, 0)$ with side length $2\theta_B$ as shown in Fig. 10, where $\theta_B = \frac{2}{3}\pi$. Note that $H_0(z)$ does not expand the output bandwidth, thus the cells outside the masked region correspond to $H_0(z)$ and remain in the diagonal.

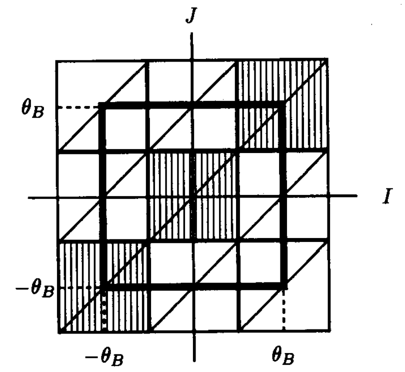


Fig. 10. Identity correspondence map for $N = 3$ and the masked region.

We now summarize the discussions so far into an algorithm for designing scrambling systems that are insensitive to frame synchronization error and free from bandwidth expansion.

Algorithm 5.1: **Data:** N, the period of the periodic filters and θ_B, the bandwidth of the scrambling system.

Step 1: Draw an identity correspondence map and make a square centered at $(0, 0)$ with side length $2\theta_B$ and call it the *changeable region.*

Step 2: Perform decompositions and symmetric and asymmetric changes within the changeable region to obtain the final correspondence map.

Step 3: With the passbands determined from the correspondence map, design the bandpass filters $H_i(z)$ and $\tilde{H}_i(z)$ with magnitude N in the passbands; modify the filters with stable minimum phase transfer functions $A_i(z)$ and $A_i^{-1}(z)$ if desired.

Step 4: Compute $F_i(z)$ and $\tilde{F}_i(z)$ by (22).

Step 5: By using the polyphase representation of $F_i(z)$ and $\tilde{F}_i(z)$, obtain $G(z)$ and $\tilde{G}(z)$, respectively.

Step 6: Realize $G(z)$ and $\tilde{G}(z)$ as periodic filters [8]; these are, respectively, the scrambling and descrambling filters.

Note that conditions (32)–(35) call for ideal bandpass filters, however, only their approximations can be implemented in practice. The quality of approximation directly relates to the DF of the scrambling system.

It is easy to see that the scrambling system proposed by Ishii [1] can be designed by using Algorithm 5.1. The cells have a fixed size of $\frac{2\pi}{N}$ and the filters are all ideal bandpass filters with a single passband. Algorithm 5.1 allows more general designs in that different cell sizes are allowed, that the bandpass filters may have multiple passbands, and that the filter characteristics could be modified by stable minimum phase transfer functions. It turns out that the schemes proposed by Lee *et al.* [6], [7]

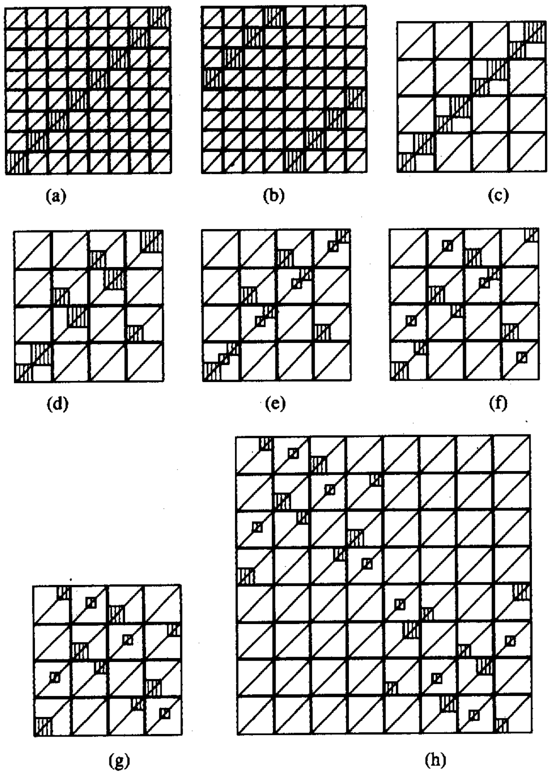


Fig. 11. Scrambling filter designed by Algorithm 5.1 with period =8.

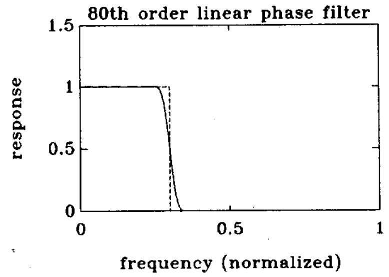


Fig. 12. Frequency response of the 80th order lowpass prototype filter.

and by Del Re [13] can also be designed by using Algorithm 5.1, in which the cell size is also fixed at $\frac{2\pi}{N}$.

High security level scrambling systems usually require small cell size. For example, in 8 kHz sampling of typical speech signal a frequency resolution of 50 Hz is required. This would require the period N to be at least 160 if the cell size were fixed at $\frac{2\pi}{N}$. In implementation, the period N is directly related to complexity and operation delay. We believe that with the flexibility provided by the proposed approach, larger key space is obtained and it is hence possible to achieve high level of security in scrambling systems with lower complexity.

VI. Design Example

In this section, we demonstrate the proposed algorithm with a design example and compare the design results with designs proposed in [3], [4], and [6].

TABLE I
The Associated ΔDF for Different Synchronization Errors

Sync. error	0	1	2	3	4	5	6	7
ΔDF	0.000	0.022	0.020	0.047	0.051	0.027	0.041	0.030

In the design we decide to set $N = 8$, and for simplicity, $\theta_B = \pi$. The design starts with the identity map in Fig. 11(a). By decomposition and interchanges we have Fig. 11(b), which is in fact the correspondence map of the frequency inversion scheme [4]. To keep the frequency inversion relation for low residual intelligibility [13], we perform interchanges within the second and fourth quadrants only. Since by (C4) the patterns in these two quadrants must be identical, we only have to perform interchanges for one quadrant and the other quadrant can be assigned accordingly. The fourth quadrant is selected for interchange, the process is shown in Fig. 11(c)–11(g). Fig. 11(h) shows the final correspondence map. We note that each block (of size $\frac{2\pi}{N} = \frac{\pi}{4}$) is divided into three cells with sizes $\frac{5\pi}{48}$, $\frac{3\pi}{48}$, and $\frac{4\pi}{48}$ respectively. The ideal generalized bandpass filters are approximated by 80th order linear phase FIR filters. The filters are all obtained from a lowpass prototype with appropriate frequency transformations. The bandpass filters are not modified, i.e., $A_i(z) = 1$. The frequency response of the lowpass linear phase filter prototype is shown in Fig. 12, where the passband is 0.3 and the dash curve is the ideal response. The maximum ripple in the passband is 0.00 29 and the transition band ranges from 0.27 34 (-0.5 dB) to 0.332 (-26 dB). We note that DF is zero if these filters are ideal. Such approximation with the nonideal filters used in our design increases DF from 0 to 0.12. For different frame synchronization error, the associated ΔDF is listed in Table I. The maximum ΔDF is 0.051 as synchronization error is half the period. The block transfer matrices $G(z)$ and $\tilde{G}(z)$ are computed. Since $G(\infty)$ and $\tilde{G}(\infty)$ are lower triangular and nonsingular, there is no operation delay.

To evaluate the design, an experiment is done as follows. A record of speech is sampled at 8 kHz, the sampled sequence is processed to yield the scrambled sequence, frame synchronization error is introduced, and the resulting signal is processed to give the recovered speech sequence. Digital-to-analog conversion is performed on both the scrambled sequence and the recovered sequence. All the signal processing and filter design are done with MATLAB under PC MS-DOS. The time and frequency domain waveforms of the original speech, the scrambled speech, and the recovered speech with three-sample synchronization error are shown in Figs. 13–15. Experimental results show that the quality of recovered speech is acceptable at this level of DF.

To evaluate the level of security, we also perform a residual intelligibility test. The procedure basically follows that of Jayant [3] and Lee [6]. The speech samples are four-digit numbers and there are 20 four-digit numbers in the test. Forty inexperienced listeners are divided into two groups to listen to the records and write down their best guess. Five scrambling designs are selected for the test. Scrambler 1 and 2 are designed by Algorithm 5.1 with correspondence maps

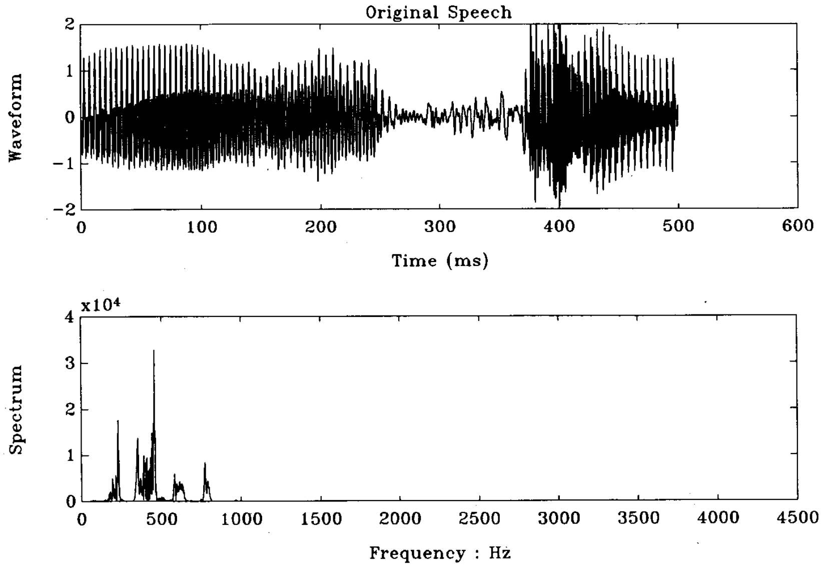


Fig. 13. Waveforms of original speech: time and frequency domain.

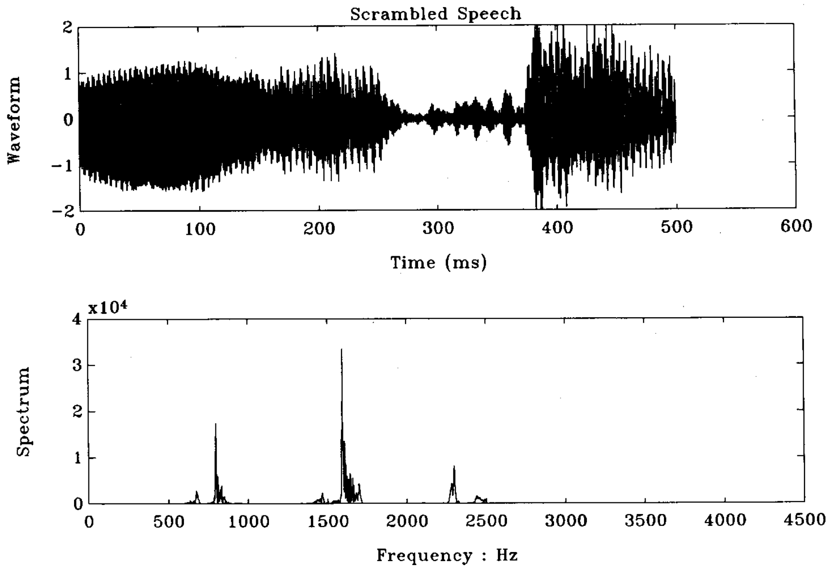


Fig. 14. Waveforms of scrambled speech: time and frequency domain.

shown in Figs. 11(h) and 16(a), respectively. Follow Lee's algorithm proposed in [6], we obtain scrambler 3. The period is eight and its equivalent correspondence map is given in Fig. 16(b). Scrambler 4 is obtained by frequency inversion. Scrambler 5 is obtained by block permutation combined with frequency inversion [4], in which the block length is two and the block number is four. The test result is shown in Table II. The results seem to indicate that acceptable level of security can be achieved by the proposed design algorithm with lower-than-usual period N. We hasten to add that the level of security (based on the RI test) depends heavily on the selection of key and that we do not claim the designs using Algorithm 5.1 are always better.

TABLE II
RESIDUAL INTELLIGIBILITY TEST RESULTS

Scheme	S1	S2	S3	S4	S5
R. I.	13.01	12.90	15.63	22.38	21.71

VII. CONCLUSION

The contributions of the paper are as follows. We propose a framework, based on transfer matrices and frequency domain descriptions, for analysis and design of speech scrambling filters. We derive necessary and sufficient conditions for a scrambling system to be insensitive to frame synchronization

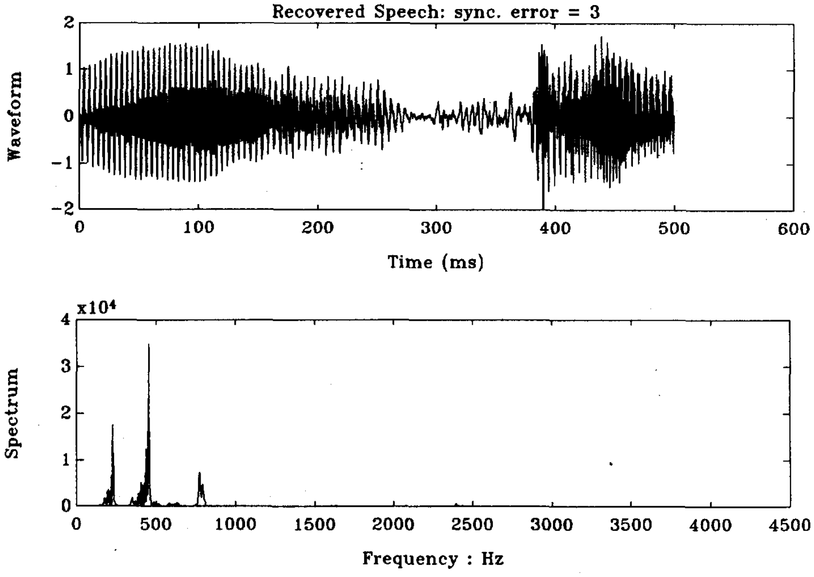


Fig. 15. Waveforms of recovered speech: time and frequency domain.

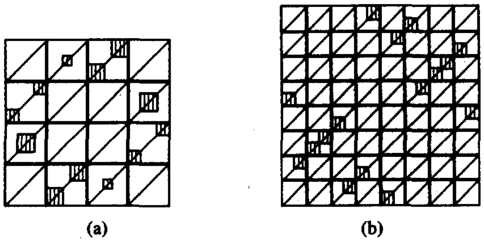


Fig. 16. Correspondence maps of scrambler 2 and 3.

error. We propose a procedure for the design of scrambling filter that is insensitive to synchronization error. The scrambling system also results in zero bandwidth expansion, low residual intelligibility, minimal operation delay, and simple hardware and software. An example is given to illustrate the design based on the proposed framework and algorithm.

APPENDIX

Proof of Lemma 4.2:

a) Suppose $x \in (C \cap A)$, then $x \in C$ and $x \in A$. Since $A \cap B = \emptyset$, $x \notin B$ and hence $x \in (C - B)$.

b) Suppose $x \in (A - (B \cup C))$, then $x \notin B$. Thus, $x \notin (B - (A \cup C))$.

c) Suppose that $x \in D_k$ for some k. Then $x \notin (D_l - D_k)$ for $l \neq k$. Since $(D_l - \bigcup_{m \neq l} D_m) \subseteq (D_l - D_k)$ for $k \neq l$, thus $x \notin (D_l - \bigcup_{m \neq l} D_m)$ for $l \neq k$. Now $\bigcup_{i=1}^{N}(D_i - \bigcup_{j \neq i} D_j) = \Omega$, hence $x \in (D_k - \bigcup_{m \neq k} D_m)$ and thus $x \notin D_m$ for $m \neq k$.

d) Assume that $x \in D_k$. Since $D_k \cap D_j = \emptyset$ for $j \neq k$, then $x \notin D_j$ for $j \neq k$. It implies that $x \notin (E_j \cap D_j)$ for $j \neq k$. Since $\bigcup_{i=1}^{N}(E_i \cap D_i) = \Omega$, $x \in (E_k \cap D_k)$ and hence $x \in E_k$. □

Proof of Lemma 4.3: Let $n(A)$ be the Lebesgue measure [14, p. 320] of A, where A is a subset of Θ. Condition (L1) implies that $n(\bigcup_{i=1}^{N} A_i) = 2\pi$, and hence $\sum_{i=1}^{N} n(A_i) \geq 2\pi$. Since $n(f_i(A_i)) = n(A_i)$, thus we also have

$$\sum_{i=1}^{N} n(f_i(A_i)) \geq 2\pi. \tag{A.1}$$

However, by (L2) we know that $n(\bigcup_{i=1}^{N} f_i(A_i)) = \sum_{i=1}^{N} n(f_i(A_i))$, thus (A.1) implies that $n(\bigcup_{i=1}^{N} f_i(A_i)) \geq 2\pi$. Since the measure of any set is bounded by 2π, it follows that

$$n\left(\bigcup_{i=1}^{N} f_i(A_i)\right) = 2\pi \tag{A.2}$$

i.e., the equality condition holds. Hence, we should have that $\sum_{i=1}^{N} n(A_i) = 2\pi = n(\bigcup_{i=1}^{N} A_i)$, or equivalently

$$n(A_i \cap A_j) = 0 \ \text{ for } i \neq j. \tag{A.3}$$

Since $A_i \in S$ and hence $f_i(A_i) \in S$, thus A_i and $f_i(A_i)$ only consist of finite number of disjoint intervals. These intervals, which may be open, closed, or half-open, have two boundary points. These boundary points can be partitioned into two categories: closed boundary point or open boundary point, according to whether they belong to the interval or not. Let n_i be the total number of open boundary points and m_i be the total number of closed boundary points of intervals of A_i. Define $N_o = \sum_{i=1}^{N} n_i$ and $N_c = \sum_{i=1}^{N} m_i$. Since $n(A_i \cap A_j) = 0$ for $i \neq j$, thus the intersection of A_i and A_j only can be empty set or isolated points. By (L1), then we should have

$$N_o \leq N_c. \tag{A.4}$$

Let $\bar{n}_i$ be the total number of open boundary points and $\bar{m}_i$ be the total number of closed boundary points of intervals of $f_i(A_i)$. Similarly, $\bar{N}_o = \sum_{i=1}^{N} \bar{n}_i$ and $\bar{N}_c = \sum_{i=1}^{N} \bar{m}_i$. By the fact that $n(\bigcup_{i=1}^{N} f_i(A_i)) = 2\pi$, we know that there is no nonzero length interval excluded from $\bigcup_{i=1}^{N} f_i(A_i)$. By (L2), then we should have $\bar{N}_c \leq \bar{N}_o$. However, the circular shift operation keeps the number of closed and open boundary points, or enhances both by one as the interval is just shifted across 2π. Thus, $\bar{N}_c \leq \bar{N}_o$ is equivalent to

$$N_c \leq N_o. \tag{A.5}$$

By (A.4) and (A.5), we conclude that $N_o = N_c$, which implies that $A_i \cap A_j = \emptyset$ for $i \neq j$ and $\bigcup_{i=1}^{N} f_i(A_i) = \Theta$. □

Acknowledgment

The authors thank Mr. G. T. Guh for conducting the residual intelligibility test.

References

[1] R. Ishii and M. Kakishita, "A design method for a periodically time-varying digital filter for spectrum scrambling," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, pp. 1219–1222, 1990.

[2] ——, "Analysis of periodically time-varying digital filters," in *Proc. 1986 ICASSP*, 1986, pp. 2607–2610.

[3] N. S. Jayant *et al.*, "A comparison of four methods for analog speech privacy," *IEEE Trans. Commun.*, vol. COM-29, pp. 18–23, 1981.

[4] S. C. Kak and N. S. Jayant, "On speech encryption using waveform scrambling," *Bell Syst. Tech. J.*, vol. 56, pp. 781–808, 1977.

[5] P. P. Khargonekar, K. Poolla, and A. Tannenbaum, "Robust control of linear time-invariant plants using periodic compensation," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 1088–1096, 1985.

[6] L. S. Lee, G. C. Chou, and C. S. Chang, "A new frequency domain speech scrambling system which does not require frame synchronization," *IEEE Trans. Commun.*, vol. COM-32, pp. 444–456, 1984.

[7] L. S. Lee and G. C. Chou, "A new time domain speech scrambling system which does not require frame synchronization," *IEEE J. Select. Areas Commun.*, vol. SAC-2, pp. 443–455, 1984.

[8] C. A. Lin and C. W. King, "Minimal periodic realizations of transfer matrices," *IEEE Trans. Automat. Contr.*, vol. 38, no. 3, pp. 462–466, Mar. 1993.

[9] ——, "Inverting periodic filters," *IEEE Trans. Signal Processing*, vol. 42, Jan. 1994.

[10] C. M. Loeffler and C. S. Burrus, "Optimal design of periodically time-varying and multirate digital filters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-32, pp. 991–997, Oct. 1984.

[11] A. Matsunaga, K. Koga, and M. Ohkawa, "An analog speech scrambling system using the FFT technique with high-level security," *IEEE J. Select. Areas Commun.*, vol. 7, pp. 540–547, 1989.

[12] R. A. Meyer and C. S. Burrus, "A unified analysis of multirate and periodically time-varying digital filters," *IEEE Trans. Circuits Syst. [Video Technol.]*, vol. CAS-22, pp. 162–168, 1975.

[13] E. Del Re, R. Fantacci, and D. Maffucci, "A new speech signal scrambling method for secure communications: Theory, implementation, and security evaluation," *IEEE J. Select. Areas Commun.*, vol. 7, pp. 474–480, 1989.

[14] W. Rudin, *Principles of Mathematical Analysis*. New York: McGraw-Hill, 1985.

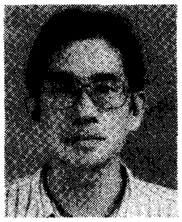
[15] K. Sakurai, K. Koga, and T. Muratani, "A speech scrambler using the fast Fourier transform technique," *IEEE J. Select. Areas Commun.*, vol. SAC-2, pp. 434–442, 1984.

[16] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*. Englewood Cliffs, NJ: Prentice-Hall, 1993, pp. 120–122.

[17] A. D. Wyner, "An analog scrambling scheme which does not expand bandwidth, part I: Discrete time," *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 261–274, 1979.

Chwan-Wen King was born in Taiwan, Republic of China, on June 13, 1957. He received the B.S. degree in electrical engineering from the National Taiwan University, Taiwan, Republic of China, in 1979, and the M.S. degree in electronics from the National Chiao-Tung University, Taiwan, Republic of China, in 1981. He is currently a Ph.D. candidate in control engineering at the same institution.

Since 1981, he has been an Assistant Scientist in the Chung-Shan Institute of Science and Technology, working on the area of development and evaluation of weapons systems. His primary research interests are the modeling and the design of periodic control systems and speech scrambling systems.



Ching-An Lin received the B.S. degree from the National Chiao-Tung University, Taiwan, in 1977, the M.S. degree from the University of New Mexico, Albuquerque, in 1980, and the Ph.D. degree from the University of California, Berkeley, in 1984, all in electrical engineering.

He was with the Chung-Shan Institute of Science and Technology from 1977 to 1979, and with Integrated Systems Inc. from 1984 to 1986. Since June 1986, he has been with the Department of Control Engineering, the National Chiao-Tung University, Taiwan, where he is a professor. His current research interests are in multivariable control and multirate signal processing.