Separation of Superimposed Digital Communication and Speech Signals: Time-varying ARMA Approach

Serkalem Melese Adigeh

A Thesis
in
The Department
of
Electrical and Computer Engineering

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Applied Science at
Concordia University
Montreal, Quebec, Canada

May 1996

ISBN 0-612-18370-X

ABSTRACT

Separation of Superimposed Digital Communication
and Speech Signals Time-varying ARMA Approach

Serkalem Melese Adigeh

The problem of separating frequency overlapping signals such as a digital communication signal and a speech signal is considered. It is assumed that the signals share the same bandwidth and appear simultaneously, precluding the use of both FDM and TDM techniques

The proposed approach is based on structural signal processing, which learns from examples by constructing an input-output mapping for the problem at hand Structural signal processing is attractive for practical application by virtue of its ability to deal with nonlinearity, non stationarity, and non-Gaussianity. More over, it offers robustness with respect to parameters tuning, which are for a setting of optimal value parameters by non expert users In this research, a special linear time-varying model, which takes advantage of the intrinsic properties of a digital communication signal, is developed to separate signals whose spectra overlap The present approach uses an Almost Symmetric-Autoregressive Moving Average (AS-ARMA) model that is combined with the proposed nonlinear blanking algorithm used to nullify residual interferences that appear at the output of Moving Average model.

To motivate the development of the thesis, two cases of the problem are presented. The first problem is to detect the presence of a digital communication signal when the received signal is composed of a weak digital communication signal with a strong speech signal and channel noise. The second case assumes that the speech signal of interest is interfered by a strong quasi-periodic signal and channel noise The simulation results show that the performance of the signal separation improves the signal-to-interference ratio for the digital communication signal detector Thus the presence of a strong speech signal does not degrade the performance of signal detection. On the other hand, a reconstructed speech signal is of good quality, so that the difference between the original and the reconstructed speech signal is minimal.

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to Dr. E.I Plotkin for his guidance and support throughout the course of this work. He led me to acquire the necessary background in this area and gave me freedom to develop my own ideas. I deeply appreciate his encouragement and his efforts in motivating me, which allowed me to complete this work.

Table of contents

List of Figures

List of Tables

List of Symbols and Abbreviations

α ==> The symmetry factor of AS-ARMA model

ADC ==> Analog-to-Digital Converter

AR ==> Autoregressive

ARMA ==> Autoregressive Moving Average

AS-ARMA ==> Almost Symmetrical-Autoregressive Moving Average

ASK ==> Amplitude Shift Keying

ASTV-ARMA ==> Almost Symmetric Time-varying-Autoregressive Moving Average

AT&T ==> AT&T microelectronics

ATDM ==> Asynchronous Time Division Multiplexing

ATM ==> Asynchronous Transfer Mode

BER ==> Bit Error Rate

BPSK ==> Binary Phase Shift Keying

CATV ==> Community Antenna Television

CDMA ==> Code Division Multiplexer Access

CPM ==> Continuous Phase Modulation

$D^m \varphi_l(k) = \varphi_l(k-m)$ ==> A shift operator

DFT ==> Discrete Fourier Transform

DM ==> Delta Modulation

DPCM ==> Delta Pulse Code Modulation

DSVD ==> Digital Simultaneous Voice and Data

FBL ==> Feedback Linearization

FDM ==> Frequency Division Multiplexing

FIR ==> Finite Impulse Response

FM ==> Frequency Modulation

$\{\varphi_i(t)\}_{i=1}^{N} \quad t \in \Re$ ==> A system of any N functions

Υ^{k_0} ==> Nonlinear noise output at synchronous instant

$H_{AR}(q^{-1})$ ==> Anti-null operator transfer function

$H_{ARMA}(q^{-1})$ ==> Null-anti-null operator transfer function

$H_{MA}(q^{-1})$ ==> Null operator transfer function

ISDN ==> Integrated Services Digital Network

IDFT ==> Inverse Discrete Fourier Transform

IF ==> Intermediate Frequency

IIR ==> Infinite Impulse Response

J ==> The cost Function

LPC ==> Linear Predictive Coding

MA ==> Moving Average

MFSK ==> Minimum Frequency Shift Keying

$min\{E(s-\tilde{s})^2\}$ ==> Mean-square error of a speech signal

MSE ==> Mean Square Error

MSK ==> Minimum Shift Keying

μ ==> A small positive step size, which is a viable recursive solution to the cost function

$n_{II}^{\alpha}(k)$ ==> Leakage noise of AS-ARMA model

Ω ==> N-dimensional signal subspace

OPSK ==> Offset Phase Shift Keying

PCM ==> Pulse Code Modulation

PDF ==> Probability Density Function

$PDS(m)$ ==> Power Density Spectrum

PN ==> Pseudo-Noise

PSK ==> Phase Shift Keying

q^{-1} ==> A delay operator

$Q_1[\cdot]$ ==> Linear time-varying transparent operator

QPSK ==> Quadrature Phase Shift Keying

RF ==> Radio Frequency

SNR ==> Signal-to-Noise Ratio

SSR ==> Structural Signal Representation

STDM ==> Synchronous Time Division Multiplexing

TDM ==> Time Division Multiplexing

$U(\cdot)$ ==> A unit step function

$W_O^{N+1}(\cdot)$ ==> An augmented Wronskian homogeneous equation

CHAPTER 1

INTRODUCTION

In general, the function of any communication system is to convey from transmitter to receiver a sequence of messages which is selected from a large number of predetermined messages. A specific aspect of a signal (data or speech) is the information that it carries rather than an average power or energy.

With the increased reliance on digital communication, bandwidth constraints have become the principal concern. It has become important to consider the efficient usage of a given bandwidth.

The convergence of computer, television, and telephone technologies has been gaining momentum for many years, but perhaps nowhere with as much impact as it is beginning to have in the area of residential access to information [1]. This is information in the broad sense: anything that can be stored and delivered in the form of bits, ranging from personalized newspapers to digitally transmitted versions of classic movies. All this is moving rapidly into the public consciousness under the terms Information Highway, interactive TV, multimedia, and so on. Expectations that something significant is about to happen are being raised, although there are many different opinions as to what the tide of convergence will actually leave on the beach. It is clear, however, that communication standards are needed in order to make this information revolution effective. More than perhaps any other field, communication require standards, from the physical medium to the user's application. Without standards, communication cannot occur since both sides of the line will not understand each other. To date the standards work undertaken in this field has been very successful

The type of applications that are envisioned place a wide variety of requirements on the network that will deliver the data to the user. In the case of voice and video, real-time considerations are critical. the audio and video streams must be reconstructed as they are at the source. Delays in transmission must not exceed the buffering capacity of the destination. Most of the delay penalty is in digitization, compression, and buffering, leaving a very strident allowance for network delay The transmission of data and still images, while more variable yet, is generally more delay-tolerant

The other physical issue is noise ingress, which resides at base-band signals Analog filtering before the analog-to-digital converter (ADC) stage is ultimately related to the definition of bandlimiting. Where the definition of bandlimiting deals with the content of the signals that may be present, analog filtering before ADC represents a signal processing stage where certain frequencies can be attenuated It is important to know the signals that can be presented before filtering and the filter's transfer function With the knowledge of both of these, the true spectrum of the signal to be digitized can be determined Sampling the Nyquist rate presents a large and often impractical demand on the filter used before digitization (anti-aliasing filter). Ideally, an anti-aliasing filter placed before an ADC would pass all of the desired frequencies up to some cutoff frequency and provide infinite attenuation for frequencies above the cutoff frequency Thus sampling at the Nyquist rate would be two times the cutoff frequency and no spectrum overlap would occur Unfortunately, practically realizable filters cannot provide this type of ideal response The attenuation of real filters increases more gradually from the cutoff frequency to the stop band Therefore, for a given cutoff frequency on a real filter, sampling at two times this cutoff frequency will produce some spectrum overlap

In general, more complicated filters are required to achieve steeper transitions and higher attenuation in the stopband therefore, more complicated techniques are required to reduce the distortion in the sampled signal due to spectrum overlap for a given sampling rate. Oversampling, i.e., sampling at rates greater than the Nyquist sampling rate, eases requirements on anti-aliasing filters. One of the benefits of oversampling is that the copies of the spectrum of the analog signal $F(f)$ [2] that are present in the spectrum of the sampled signal $F_S(f)$ become increasingly separated as the sampling rate increases beyond the Nyquist rate. The trade-off, of course, is that increasingly faster ADCs are required to digitize relatively low frequency signals

In quadrature sampling, the signal to be digitized is split into two signals One of these signals is multiplied by a sinusoid to down-convert it to a zero center frequency and form the inphase component of the original signal The other signal is multiplied by a 90-degree phase-shifted sinusoid to down-convert it to a zero center frequency and form the quadrature-phase component of the original signal. Each of these components occupies only one-half of the bandwidth of the original signal and can be sampled at one-half the sampling rate required for the original signal. Therefore, quadrature sampling reduces the required sampling rate by a factor of two at the expense of using two phase-locked ADCs instead of one.

Sampling at rates lower than $2f_{max}$ can still allow for an exact reconstruction of the information content of the analog signal if the signal is a bandpass signal An ideal bandpass signal has no frequency components below a certain frequency, f_l, and above a certain frequency, f_h. For a bandpass signal, the minimum requirement on the sampling rate to allow for exact reconstruction is that the sampling rate be at least two times the bandwidth $f_h - f_l$ of the analog signal.

A bandpass signal can be used to down-convert a signal from a bandpass signal at an RF or IF to a bandpass signal at a lower IF. Since the bandpass signal is repeated at integer multiples of the sampling frequency, selecting the appropriate spectral replica of the original bandpass signal provides the down-conversion function. Theoretically, bandpass sampling allows sampling rates to be much lower than rates required for sampling at two or more times the highest frequency content of the bandpass signal. This means that ADCs with lower sampling rates may be used [3]

Until recently, teleconferencing required two phone lines, one for voice and one for data. Recently, three technologies have been developed for the purpose of using one line for carrying both voice and digital communication signal simultaneously: Digital Simultaneous Voice and Data (DSVD), Voice-span, and Voice-view.

The DSVD technology digitally compresses voice and multiplexes it with the data stream of a 28 8 kbps, V.34-compatible modem [4].

The Voice-span technology from AT&T Paradyne transmits analog voice signals by modulating the amplitude of the transmitted digital data while simultaneously detecting the voice-and-data modulation at the receiving end.

The Voice-view method, unlike the other two methods, is not a simultaneous voice-over-data technology However, voice and data can be carried simultaneously over a standard telephone line using this technique by toggling between voice and data.

The DSVD technique has three states: analog voice transmission, digital data signal transmission, and the DSVD state When it is in analog voice transmission state the modem simply switches the analog voice call to the down-line phone, completing a normal connection No data transmission occurs, and the call is ended by placing the down-line phone on the hook In digital data signal transmission state, the modem transfers only data using the V.34 protocol. To end data transfer, one modem sends a stop command to the other modem. When it is in its own (DSVD) state, the modem samples and digitally compresses the analog signal from the down-line phone for transmission The modem multiplexes the compressed voice with user data by using the V 42 link access procedure for modem data-link protocol The voice and data are transmitted to another DSVD-capable modem The data link connection identifier, a 6-bit address field, implements the logical channels The DSVD specification multiplexes audio onto the data stream, using a separate data link connection identifier for audio Once connected in analog state, a modem can make the transition to a DSVD state. The originator modem transmits a V 8 start-up sequence that the answering modem detects. Upon detection of the start-up, sequence, the handsets are muted for a few seconds. This silencing prevents the users from hearing the data transfer while negotiations between modems occur. After a short protocol-negotiation phase, voice-over-data operation begins, and digital compressed audio is routed to the down-line phone. When a V.34 modem is transmitting at 28.8 kbps, the modem can transfer data at 19.2 kbps. This transmission rate allows for a small 1 1 kbps overhead rate when the voice is compressed to 8 5 kbps. When the modem detects audio silence, the data-transmission bandwidth automatically expands to 28.8 kbps

The current thesis considers structural signal processing for the purpose of separating frequency overlapped signals, which are the result of multiple usage of the same bandwidth The usual method of passing a received signal through a filter, which keeps the interested signal relatively unchanged while suppressing the interference, is not necessarily appropriate.

The thesis takes advantage of the intrinsic properties of signals and their structural nature to provide specific modeling for the separation of signals whose spectra overlap A linear time-varying ARMA model is considered to achieve the preferred task [5] This technique allows the use of the telephone channel for multiple purposes,-- for instance, sending data while speech occupies the same bandwidth. The algorithm is very simple to implement

To evaluate the performance of this model, mean-square-error (MSE) is used to measure the integrity of the reconstructed speech, and a bit-error-rate (BER) assesses the digital communication system.

1.1 Statement of the Problem

In certain communication situations, a primary input is available consisting of a signal component with an additive undesired interferences. The problem of the reception is to gain information from the received input signal, which is the sum of multiple user signals transmitted using a single physical path Conventional techniques employ multiplexing mechanisms to combine different user signals in a single physical path. Based on structural signal processing, the present work uses a telephone channel for conversation and data communication simultaneously. In this thesis the work, which is to separate the interest signal from interferences, is done by exploiting the intrinsic properties of the digital communication signal. A null-operator associated with a blanking algorithm is modeled to suppress the digital communication signal totally from the received input signal. The remaining speech signal is passed through an anti-null operator to be reconstructed. The problem under consideration may be stated as the retrieving of a weak signal in the presence of a strong interference, whose spectra entirely overlap with the spectra of the useful signal.

To motivate the development of the thesis, two cases of the problem are presented. These problems are specially difficult to solve using the theories of measuring the power spectrum. Consider the signal model,

$$r(t) = \sum_{k=1}^{M} s_k(t) + x(t) + n_r(t) \tag{1.1}$$

where $s_k(t)$ is a non-Gaussian interfering signal that temporarily and spectrally overlaps $x(t)$, $x(t)$ is a communication signal of interest, and $n_r(t)$ is a broadband noise. All the signals and noise are considered statistically independent of each other. The power level of the noise is assumed unknown and is varying within the observation interval

The first problem is to detect the presence of $x(t)$ while a strong interference occupies the whole bandwidth. Methods of detecting $x(t)$ that are based on measuring the power spectrum are ineffective because the power level of noise and interference is unknown and time-varying; setting a threshold for energy detection is difficult

Consider a weak digital communication signal $x(t)$, that is superimposed with a strong speech signal. The BER is used to evaluate the system performance of detection.

Next assume that $s_k(t)$ is a speech signal of interest which is interfered by a strong quasi-periodic signal, and noise. In this case the concern is a strong digital communication signal that corrupts a speech over the same bandwidth.

Since both the signals occupy the same frequency bandwidth and appear at the input of the receiver simultaneously, the problem is to retrieve the weak signal, whose spectra are entirely overlapped by a strong interference.

1.2 Organization of the Thesis

Chapter 2 is devoted to the mathematical representation of data-communication and speech signals, and error analysis on signal detection and speech reconstruction as it relates to the thesis work

Chapter 3 covers structural signal modeling. It presents the concept of structural signal representation, the linear time-invariant structural signal representations, such as MA, AR and ARMA modeling, as well as time-varying and nonlinear structural signal representations.

Chapter 4 illustrates the concept and rules for the design of a linear time-varying null-anti-null operator. Basic properties of such a time-varying null-anti-null operator are then discussed. It shows that physically implementable anti-null operator cannot be perfectly symmetrical to a null-operator. The transparency property can be achieved by introducing a certain strategy to control the symmetrical nature of the transparent operator. This chapter explores the almost symmetrical ARMA model and its unique properties such as *annihilation* and *transparency.*

In Chapter 5 the *transparency* problem of the ARMA model is studied. It shows that a symmetrical anti-null-operator can provide an ideal *transparency*, but it is sensitive to undefined initial conditions The MA part of the AS-ARMA model creates spurious spectral components when a digital communication signal changes its phase. Using an extra blanking algorithm, which eliminates the null-operator residues, will lead the model to become an almost transparent model by choosing the symmetric factor near unity.

Chapter 6 presents the investigation of sources of noise, which affect the detection of a digital communication signal and the reconstruction of a speech signal. A signal detection problem is considered when a weak digital communication signal is imbedded in a speech signal and channel noise, and a signal estimation problem when a weak speech signal is imbedded in an interfering digital communication signal and channel noise. This nonlinear noise, which is seen at the output of the AS-ARMA system, will affect the detection of a digital communication signal as well as the reconstruction of speech signal. In order to verify the system performance, the bit-error-rate (BER) is measured and the quality of a speech signal reconstruction is evaluated.

Chapter 7 outlines a novel self-synchronized signal controlled AS-ARMA model. The basic approach is to define parameters that will minimize the mean-square separation error over a short segment of the input signal. The resulting parameters are then assumed to be the parameters of the system function, $H(z)$, in the model for signal separation.

Chapter 8 concludes with a summary of works which are presented in the thesis and suggestions for future work.

Chapter 2

Representation of Data-communication and Speech Signals

Most of the engineering processes can be represented by means of signals. These signals are represented by different mathematical models that are convenient for the analysis Representation of data-communication and speech signals can be based on time-domain techniques, frequency-domain techniques, or a combination of these techniques. In this section the focus is mainly on time-domain representations

2.1 Representation of Data-communication Signals and their Properties

In general, a signal can be classified as either analog or discrete. Analog signals can be represented by a series of discrete samples, each of which represents a value of $f(t$ at a particular sampling point. One of the simplest ways of discretizing an analog signal is by equally spaced sampling. Such a type of sampling procedure is usually called uniform sampling [6].

A signal whose Fourier spectrum vanishes outside an interval can be easily realized, it is called a bandlimited signal. A continuous time signal $f(t)$, band limited to $|w| < W$ can be reconstructed from its equally spaced samples if the sampling frequency is at least equal to the Nyquist frequency. A typical practical example is the voice signal in telephone systems where the highest frequency is approximately equal to 3 3 kHz, and the sampling frequency is chosen to be 8 kHz, which is above its Nyquist frequency If the minimum sampling rate, which is 2 W, is not met, the spectral components will overlap In this case, the original spectrum will appear with that of the other components and cannot be uniquely determined. This process is called aliasing.

In general any communication signal may be described as a modulated signal, having a carrier frequency f_c:

$$x(t) = R(t)\cos[2\pi f_c t + \phi(t)] \tag{2 1}$$

where $R(t)$ represents the real envelope of $x(t)$ and $\phi(t)$ represents the phase deviation.

Equation 2 1 can be written in the form.

$$x(t) = Re\{R(t)e^{j\phi(t)}e^{j2\pi tf_c}\} \tag{2.2}$$

or

$$x(t) = Re\{\tilde{x}(t)e^{j2\pi tf_c}\} \tag{2.3}$$

where the quantity $\tilde{x}(t)$ is called a complex envelope of a real signal $x(t)$ Clearly $\tilde{x} = R(t)e^{j\phi(t)}$ is a complex function of time that is independent of the carrier frequency f_c. It is important to note that a complex envelope involves signals that are usually slowly varying with respect to a carrier frequency. Since the bandwidth of a bandpass signal is significantly less than f_c, it takes a much lower sampling frequency to represent a complex envelope $\tilde{x}(t)$, than to represent a real-time signal $x(t)$.

The result is a smaller number of samples for a given time segment of $x(t)$. The complex envelope is usually expressed in rectangular form $\tilde{x}(t) = x_d(t) + jx_q(t)$ where $x_d(t)$ is the direct (or real) component of $\tilde{x}(t)$ and $x_q(t)$ is the quadrature (or imaginary) component of $\tilde{x}(t)$. Assuming that the carrier frequency is known, the complex envelope contains all of the information contained in the original signal $x(t)$.

- **Signal Generation**

Both deterministic and random signals exist in almost all communication systems. Models must be developed for each of these signal types that can be implemented in digital computer simulation. Deterministic signals are usually generated using the defining equation for the signal [7] Random signals are usually generated using either a linear congruential algorithm or a Pseudo-Noise PN sequence algorithm. Although the mathematical descriptions of these two algorithms are somewhat different, they are essentially equivalent. Since a digital computer is a finite-state machine, it is not possible to generate a truly random signal on a computer, and all computer-generated sequences are periodic. Within a period, the pseudorandom sequence approximates many of the properties of a random signal. One is therefore able to generate "noise-like" wavefoms for use in a simulation to represent both random signals and noise, hence the term Pseudo-Noise (PN) sequences.

- **Base-band Signal Modulation**

Base-band signals can be modulated onto a sinusoidal carrier by modulating one or more of its three basic parameters: amplitude, frequency, and phase. Correspondingly, there are three basic modulation schemes in digital communication. Amplitude Shift Keying (ASK), Frequency Shift Keying (FSK) and Phase Shift Keying (PSK).

The digital communication schemes can be classified into two large categories: constant envelope and non-constant envelope. The PSK schemes have constant envelope but discontinuous phase transitions from symbol to symbol The Continuous Phase Modulation (CPM) schemes have not only constant envelope, but also continuous phase transitions. Thus they have less side lobe energy in their spectra in comparison with the PSK schemes From Table 2.1, it is clear that a Binary Phase Shift Keying (BPSK) and Quadrature/ Offset Quadrature Phase Shift Keying (QPSK/OQPSK) have the same power efficiency, but QPSK/OQPSK have twice the bandwidth efficiency. MSK is just a special case of a large class of constant amplitude modulation schemes called Continuous Phase Modulation (CPM). This class of modulation is jointly power and bandwidth efficient. The CPM signal is defined by:

$$x(t) = \sqrt{2\frac{E}{T}} \cos\left(2\pi f_0 t + \Phi(t, a)\right) \tag{2.4}$$

where the transmitted M-ary symbols a_i appear in the phase is given by Eq. 2.5:

$$\Phi(t, a) = 2\pi h \sum_{i=-\infty}^{\infty} a_i q(t - iT) \tag{2.5}$$

with

$$q(t) = \int_{-\infty}^{t} g(\tau)\, d\tau$$

The function $g(t)$ has smooth pulse shape over finite time interval $0 \le t \le LT$, and is zero outside When $L >$, there is a partial-response pulse shape; when $L \le$, there is a full-response pulse shape.

Modulation schemes	Bandwidth Nyquist	Efficiency(bps/Hz) Null-to-null	$E_b/N_0\ (dB)$ $p_b = 10^{-5}$
BPSK	1.0	0.5	9.6
QPSK	2.0	1.0	9.6
OQPSK	2.0	1.0	9.6
MSK	n/a	2/3	9.6

Table 2 1 Comparison of power and bandwidth efficiencies for various signal modulations

The present thesis uses BPSK, FSK and MFSK modulation techniques to represent a digital communication signal. The modulated signal, which is represented by one of the three techniques mixed with a speech signal, will be transmitted using a single path to remote. In this thesis the task is to research a model that can be used to separate a speech signal from a digital communication signal without affecting the original quality of a speech and data detection. Conventional techniques exist which allows different users to occupy a single physical path for communication

- **Multiplexing**

Multiplexing combines and bundles together a number of communication channels and transmits them over one physical common broadband channel. At the receiving end, demultiplexing separates and recovers the original channels. The main conventional techniques of multiplexing are Frequency Division Multiplexing Access (FDMA), Time Division Multiplexing Access

(TDMA), and Code Division Multiplexing Access (CDMA)

- **Frequency Division Multiplexing Access (FDMA)**

In FDM systems, a portion of the frequency spectrum (that is, a frequency band) is allocated in each transmitter The transmitted signal spectral component must be confined to the allocated frequency band.

Examples of FDM systems are the telephone network, cable television (that is, the Community Antenna Television CATV), voice frequency multiplexers, satellite systems, etc

- **Time Division Multiplexing Access (TDMA)**

In TDM systems, the entire frequency bandwidth of transmission media is allocated in each station but only for a limited portion of time, called a time "slot" There are two basic techniques for time division multiplexing. Synchronous Time Division Multiplexing (STDM) and Asynchronous Time Division Multiplexing (ATDM)

In STDM each source is repeatedly (hence, the word synchronous) assigned a portion of time of the capacity. TDM is understood to imply STDM. Circuit switched telephone networks use (synchronous) TDM.

On the other hand, in ATDM, each source is assigned a portion of the transmission capacity only as it is needed ATDM is used in statistical multiplexers, known also as St-Mux or concentrators, packet switches, and asynchronous transfer mode (ATM) switches proposed for broadband ISDN.

- **Code Division Multiplexing Access (CDMA)**

In CDMA, each user is provided with an individual and distinctive pseudorandom noise (PN) code. These codes are almost uncorrelated. To illustrate the principle of operation of CDMA, consider that at a given time, each of k users is transmitting data at the same carrier frequency f_0, using a direct sequence spread spectrum, and his particular code $g_i(t$

Then, each receiver is presented with the same input waveform,

$$v(t) = \sum_{i=1}^{k} \sqrt{2P_s}\, g_i(t)\, d_i(t) \cos(\omega_0 t + \theta_i) \tag{2.6}$$

where each signal is assumed to present the same power P_s to the receiver, each PN sequence $g_i(t)$ has the same chip rate f_c, and $d_i(t)$ is the data transmitted by user i. The data rate for each user is the same. If the receiver is required to receive each of the k users, it needs k correlators. The advantage of a CDMA system is that collisions are not destructive, i.e., each of the signals involved in a collision would be received with only a slight increase in error rate.

	Time-Domain	*Frequency Domain*	*Remark*
FDMA	Superimposed	Separated	Channel Banks
TDMA	Separated	Identical	Time Slot
CDMA	Superimposed	Overlapped	Orthogonal Code

Table 2.2 Comparison of Multiplexing Techniques

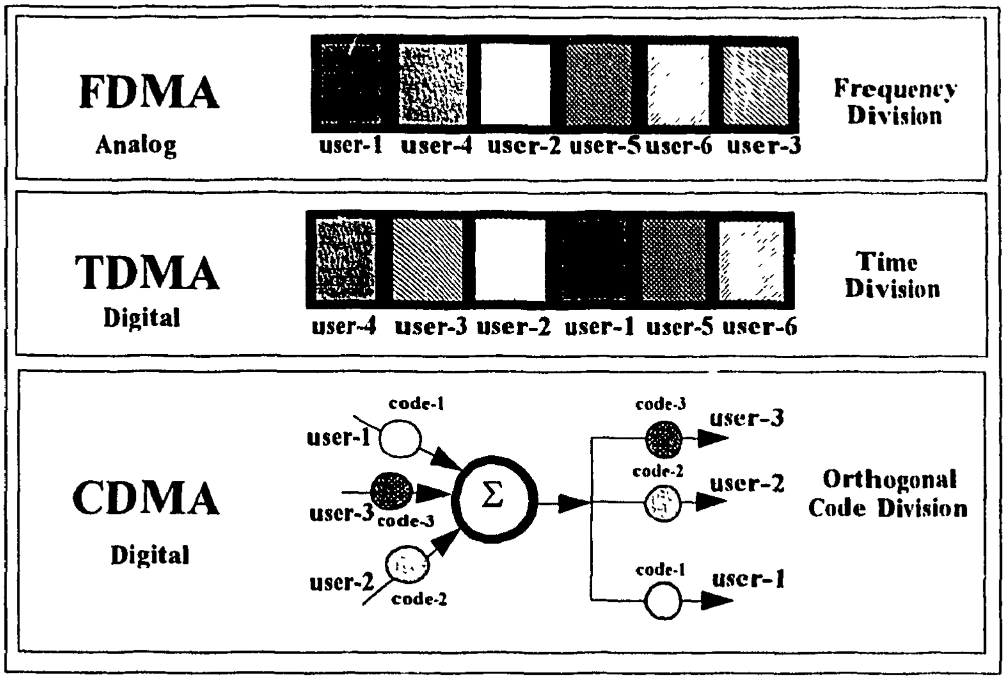


Figure 2.1 Conventional Multiplexing Techniques

A Structural Signal representation can be considered as one of the multiplexing techniques which drives different users into a single information path. In Structural Signal Multiplexing Access (SSMA) the intrinsic properties of each digital communication signal is used to model a signal separator. An Almost-symmetric ARMA model whose function is to separate an interest signal from the mixture will be introduced in later chapters. Overlaying CDMA with SSMA is a possible technique in ever-growing band width constraints. Each digital communication signal modulates the base-band signals, which are mentioned in the previous section, and it is transmitted to its destination in a single channel

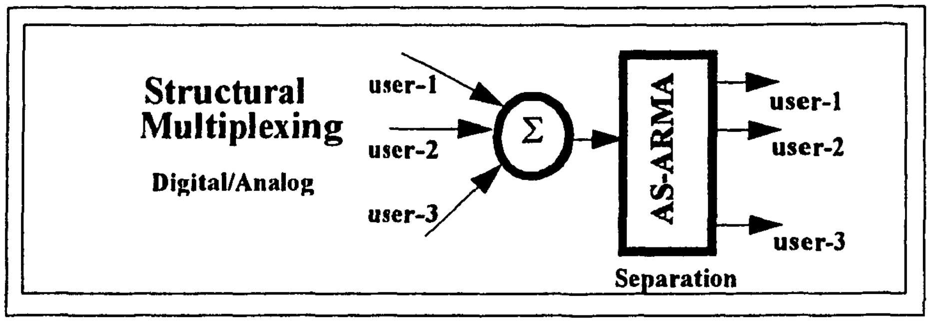


Figure 2.2 Structural Multiplexing Technique

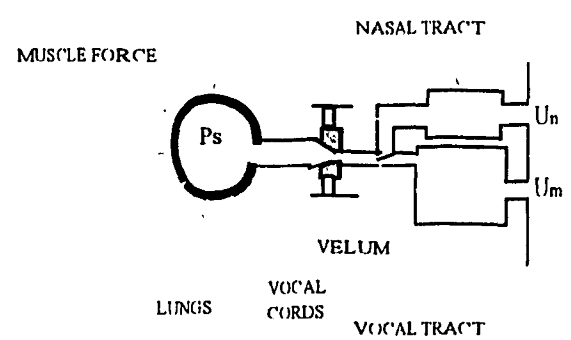
This multiplexing and demultiplexing technique uses the structural signal properties of the modulated signal. The transmitted signal, which is the sum of modulated signals mixed with speech, covers the entire telephone bandwidth. A speech signal is a broadband random signal. When it is transmitted through a telephone channel its bandwidth is limited by a lowpass filter.

2.2 Representation of a Speech Signal.

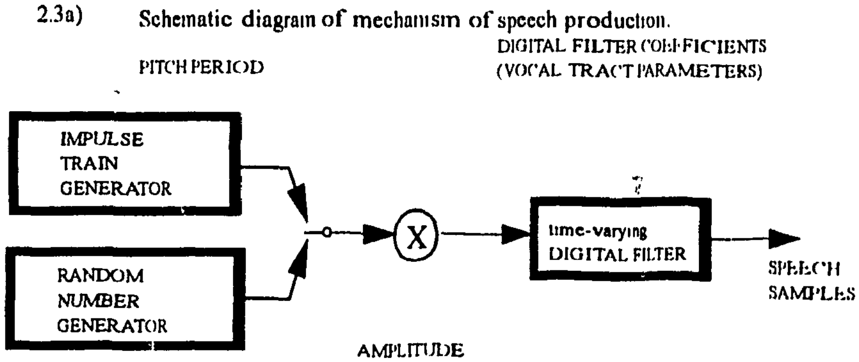
The notion of representation of a speech signal is central to almost every area of speech communication research. Often the form of representation of a speech signal is not singled out for special attention or concern, but yet is implicit in the formulation of the problem or in the design of the system. A good example of analog representation of a speech signal is in telephony, where speech is in fact represented by fluctuations in current for purposes of long distance transmission. In other situations, however, one must often pay strict attention to the choice and method of implementation of a speech signal. This is true, for example, in such diverse areas as speech transmission, computer storage of speech and computer voice response, speech synthesis, speech aids for the handicapped, speaker verification and identification, and speech recognition. In all of these area digital representations, i.e., representations based on a sequence of numbers, are becoming increasingly dominant. There are two basic reasons for this. First, through the use of small, general purpose digital computers, speech researchers have been able to apply a wide variety of digital signal processing techniques to communication problems [7]. These techniques cover a range of complexity and sophistication that is impossible for analog methods to match Second, the recent and predicted future developments in integrated circuit technology makes it possible to realize digital speech processing schemes economically as the hardware devices having the same sophistication and flexibility as a computer program implementation are available in these days.

• **A digital Model of a Speech Signal**

The vocal tract is an acoustic tube that is terminated at one end by the vocal cords and at the other end by the lips. An ancillary tube, the nasal tract, can be connected or disconnected by the movement of the velum. The shape of the vocal tract is determined by the position of the lips, jaw, tongue, and velum [8].



2.3a) Schematic diagram of mechanism of speech production.



2.3b) Digital processing model for production of speech signals [8]

Sound is generated in this system in three ways Voiced sounds are produced by exciting the vocal tract with quasi periodic pulses of air pressure, caused by vibration of the vocal cords Fractive sounds are produced by forming a constiuction somewhere in the vocal tract and forcing air through the construction, thereby creating turbulence which produces a source of noise to excite the vocal tract. Plosive sounds are created by completely closing off the vocal tract, building up a pressure, and then quickly releasing it All these sources create a wideband excitation of the vocal tract which in turn acts as a linear time-varying filter, which imposes its transmission properties on the frequency spectra of the sources The vocal tract can be characterized by its natural frequencies (or formants) which correspond to resonances in the sound transmission characteristics of the vocal tract Because both the sound sources and vocal tract shape are relatively independent, a reasonable approximation is to model them separately, as shown in Figure 2.3b.

In this digital model, samples of a speech waveform are assumed to be the output of a time-varying digital filter that approximates the transmission properties of the vocal tract and the spectral properties of the glottal pulse shape This model is the basis of a wide variety of representations of speech signals. These are conveniently classified as either waveform representations or parametric representations depending upon whether a speech waveform is represented directly or whether the representation is in terms of time-varying parameters of the basic speech model. The choice of the digital representation is governed by three major considerations processing complexity, information (bit) rate, and flexibility. By complexity, it means the amount of processing required to obtain the chosen representation. A low bit rate means that the digital representation of a speech signal can be transmitted over a low capacity channel, or stored efficiently in digital memory. Flexibility is a measure of how a speech can be manipulated or altered for applications other than transmission, e g., voice response, speech recognition, or speaker verification.

In general, greater complexity is the price paid to lower the bit rate and increase the flexibility. In transmission and voice response applications, the quality and intelligibility of the reconstituted speech are also prime considerations.

The following techniques are capable of producing good quality and highly intelligible speech·

• **Digital Waveform Coding**

Conceptually, the simplest digital representation of speech is concerned with direct representation of speech waveform. Such schemes as pulse code modulation (PCM), delta modulation (DM) and adaptive differential PCM (DPCM) are all based on Shannon's sampling theorem, which says that any bandlimited signal can be exactly reconstructed from samples taken periodically in time if the sampling rate is twice the highest frequency of the signal.

The objective of digital waveform coding is to represent a speech waveform as accurately as possible so that an acoustic signal can be reconstructed from the digital representation In many speech processing problems, however, one is not interested in constructing an acoustic signal but rather one is concerned with representing a speech signal in terms of a set of properties or parameters of the model discussed in the previous section Some rather simple, but useful, characterizations can be derived by simple measurements of the waveform itself, i.e., upon a PCM representation of the waveform

The key to these and indeed the key to all parametric representations, is the concept of short-time analysis. For a short period of time (10-to 30-ms duration) it is quite probable that the properties of the waveform remain roughly invariant over that interval

• **Homomorphic Speech Analysis**

Homomorphic filtering is a class of nonlinear signal processing techniques that is based on generalization of the principle of superposition that defines linear systems Such techniques have been applied in separating signals that have been combined by multiplication and convolution. The application of these techniques to speech processing is again based on the assumption that although speech production is a time-varying process, it can be viewed on a short-time basis as the convolution of an excitation function (either random noise or a quasi periodic pulse train) with the vocal tract impulse response These methods for separating the components of a convolution are of interest [9] The cepstrum is an excellent basis for estimating the fundamental period of voiced speech for determining whether a particular speech segment is voiced or unvoiced.

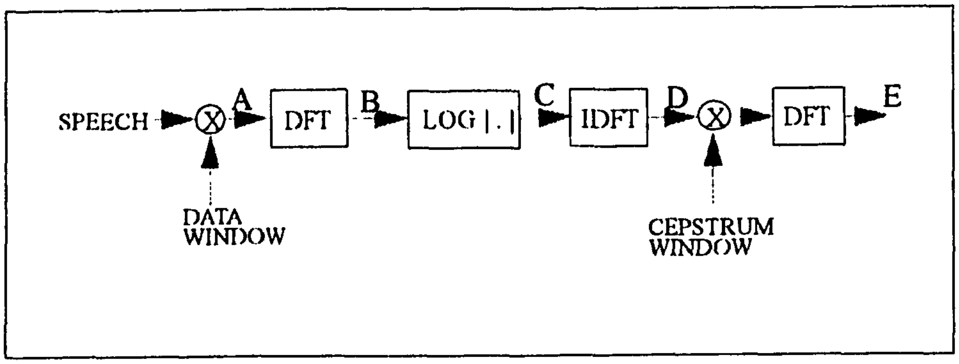


Figure 2.4 Homomorphic processing of speech

The signal at A, which is shown in Figure 2.4, is assumed the discrete convolution of the excitation and the vocal tract impulse response. Then the short-time Fourier transform (i.e., the spectrum of the windowed signal), is the product of the Fourier transforms of the excitation and the vocal tract impulse response. Taking the logarithm of the magnitude of the Fourier transform, one obtains at C the sum of the logarithms of the transforms of the excitation and vocal tract impulse response. Since the inverse discrete Fourier transform (IDFT) is a linear operation, the result at D (called the cepstrum of the input A) is an additive combination of the cepstra of the excitation and vocal tract components. Thus, the effect of operations, windowing, DFT, log magnitude, and IDFT is to approximately transform convolution into addition.

The vocal tract transfer function, often called the spectrum envelope, can be obtained by removing the rapidly varying components of the log magnitude spectrum by linear filtering. One approach to this filtering operation involves computing the IDFT of the log magnitude spectrum (to give the cepstrum), multiplying the cepstrum by an appropriate window that only passes the short-time components, and then computing the DFT of the resulting windowed cepstrum. This method corresponds to the fast convolution method-in this case being applied to filter a function of frequency rather than of time. The smoothed spectrum obtained by the above method is in many respects comparable to a short-time spectrum obtained by direct analysis using a short data window. The major difference, however, is that the cepstrum method is based upon the initial computation of narrow-band spectrum, which involves a wide time window, while the wideband spectrum is computed using a very narrow-time window. The smoothing is done upon the narrow-band log-magnitude spectrum rather than upon the short-time Fourier transform itself, as is the case for wideband analysis Thus, for speech segments in which the basic parameters such as pitch period and formant frequencies are not changing, one should expect the cepstrum method to produce superior results to direct spectrum analysis When the speech spectrum is changing rapidly, as in the case of voiced/unvoiced boundary, the direct method may produce a better representation than the cepstrum method due to its shorter averaging time. Specifically, voiced/unvoiced classification of the exciting is indicated by the presence or absence of a strong peak in the cepstrum [10]. The presence of a strong peak for voiced speech is independent upon there being many harmonics present in the spectrum. In cases where this is not true, such as voiced stops, zero crossing measurements are helpful in distinguishing voiced from unvoiced speech [11]. If a strong peak is present, its location is a good indicator of the pitch period.

- **Linear Predictive Analysis**

Among the most useful methods of speech analysis are those based upon the principle of linear prediction. These methods are important because of their accuracy and their speed of computation. The basic idea behind linear predictive coding (LPC) is that a sample of speech can be approximated as a linear combination of the past speech samples By minimizing the square difference between the actual speech samples and the linearly predicted ones, one can determine the predictor coefficients; i.e., the weighting coefficients of the linear combination.

Once the predictor coefficients have been obtained, they can be used in various ways to represent the properties of a speech signal.

The results of most of these techniques can be applied in a variety of speech processing applications including speech recognition, speech synthesis, and speech verification

2.3 Error Analysis on Signal Detection and Speech Reconstruction

A major issue which arises when trying to understand the results of this study is how to interpret the various error scores This is one problem for which one has no simple answer other than that it depends on the intended application. The level at which various types of errors become significant depends strongly on the application

In order to evaluate a system in analog or digital communication, one needs to measure errors; estimating or detecting an observed signal at the receiver and comparing it to the original. The error rate in signal detection is a function of the distance S of the signal from the decision threshold compared with the level of noise. If the noise exceeds S at the decision time, there is an error. This quantity is given in bit-error-rate (BER), which is the probability of the noise exceeding S

In speech enhancement, a system evaluation is based on many factors such as intelligibility, naturalness, cost of implementation, etc. These factors of speech quality are governed by the interference and background noise.

Intelligibility varies from fair to excellent. It is worst when the recovered voice is weaker and best when voices are strong. Naturalness of the recovered speech is striking. The sound is unmistakably that of a person, not a machine, and the voice is recognizable.

- **Error Analysis on Detection of communication Signal**

It can take a long time to accumulate enough errors to accurately assess a digital communication system. To accurately measure the error rate of a digital communication system, one must record a fairly large number of errors [1]. If n errors are counted, then the inaccuracy is about $\frac{1}{\sqrt{n}}$ For example, for 400 recorded errors, the inaccuracy is 5%. But if the error rate is very low, it may take hours or even days to accumulate 400 errors. Measurement time can be reduced if the goal is to determine an upper bound on the error rate rather than to pin down the rate itself. If the system is error free for a period T, then there's 95% confidence that the error rate is less than $\frac{3}{T}$. For example, if no errors occur for one hour, then 95% of the time the system will have less than three errors per hour. Digital communication system must have very low BER. Measurement of such low rates presents a dilemma: either the test will take a long time or results won't be accurate Suppose the communication system under test has a bit rate of f_b and the BER must be less than P_e This limit corresponds to $f_b \times P_e$ errors per second. In general the error rate r is the bit-error ratio times the bit rate.

$$r = BER \times f_b \tag{2 7}$$

The error rate is continually calculated according to the formula:

$$r = \frac{n}{T} \tag{2 8}$$

where n is the number of errors counted during T. For short measurement times, r varies wildly, but it settles down as time increases. A Poisson process presumes an actual or average error rate r that can be determined from the process itself. Our task is to get an estimate r' of this actual rate by measuring n errors in a period of T and dividing:

$$r' = \frac{n}{T} \tag{2.9}$$

If T is one hour and if one takes many one-hour measurements of n, one will get a range of answers with deviation σ of the measurements in the rms of the difference from this average:

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(n_i - \bar{n}\right)^2} \tag{2.10}$$

where N is the number of measurements, about *68%* of the measurements will lie within of the average $\bar{n}$.

The standard deviation of n is given in terms of n:

$$\sigma \approx \sqrt{n}$$

That is, n is within $\sqrt{n}$ of the expected count *68%* of the time. If one takes the inaccuracy of the measurement to be as a fraction of n, then:

$$inaccuracy_{0.68} = \frac{\sigma}{n} \approx \frac{1}{\sqrt{n}} \tag{2.11}$$

This relationship, or the curve for *68%* confidence, is easily plotted (Figure 2.5). As an example, suppose that the desired accuracy is *0.10* or *10%*. Then from the inaccuracy 68% equation, the test must continue until $n=100$. If the time required to collect *100* errors turns out to be $T=19$ hours, then $r'=100/19=5.26/hour$, and this is within *10%* of the actual error rate r. That is, r is inferred to lie between $5.26-0.526=4.634/hour$ and $5.26+0.526=5.634/hour$ (with a confidence of *68%*). Because of the measurements, statistical nature, n -*100* can be more than *10%* away from the expected measurement, but *68%* of the time it will be less than *10%* away.

The confidence level can be improved by using 2σ. The measured $n=100$ is within 2σ (or *20%* here) of the expected count *95%* of the time (Figure2.5). In the example of $r'=5.26/hour$, r is inferred to lie between $5.26-1.052=4.208/hour$ with a confidence of *95%*. There's greater confidence that the inaccuracy won't be exceeded, but the inaccuracy is twice as large. To maintain a confidence of *95%* (the higher figure in Figure2.5) and still have an inaccuracy of *10%*, one must count more errors. If 2σ is to *10%* of n, then σ is *5%* of n, or

$$\frac{\sigma}{n} = 0.05 \tag{2.12}$$

Using $\sigma \approx \sqrt{n}$, the inaccuracy can be calculated, $\frac{1}{\sqrt{n}}=0.05$, *or* $n=400$. The general expression with *95%* confidence is:

$$Inaccuracy_{0.95} = \frac{2\sigma}{n} \approx \frac{2}{\sqrt{n}} \tag{2.13}$$

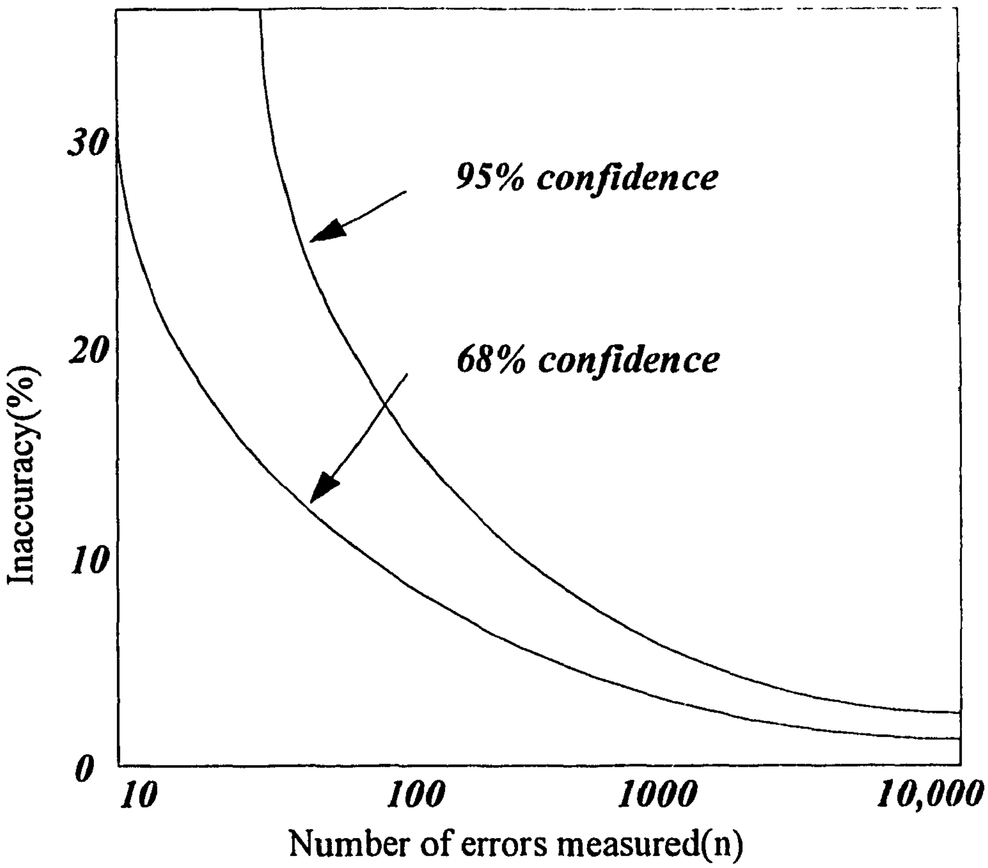


Figure 2.5 Inaccuracy of BER measurement.

• **Test for Upper Limit**

The proposed method for determining an upper bound on an error rate is to require the system under test to be error-free for a measurement period *T*. The longer *T* is, the lower the error rate boundary. Suppose one wants to be sure that the actual error rate of a system is less than a specified error rate of *r=5.56/hour*, then one must choose *T* so that an error rate of *r=5.56/hour* or greater will have at least one error in the period *T*. Because of the statistical nature of the measurement, this cannot be absolute no matter how large *T* is.

If one cannot be *100%* certain that the actual error rate is less than the specified error rate *r*, the next best thing is to settle for an acceptable level of certainty, say *90%* of an actual rate of less than *r*. Choose *T* so that *r=5.56/hour* will fail the test *90%* of the time. In other words, choose T so that the probability of measuring r=*zero* errors is only *10%* when the error rate is at the limit r. The side-bar on Poisson errors gives the probability of measuring $n=0$ as:

$$p(0) = e^{-rT} \tag{2.14}$$

Then set $p(0)=0.10$ or 10% and solve for T:

$$T = \frac{[-ln(0.10)]}{r} = \frac{2.3}{r} \tag{2.15}$$

For r=*5.56/hour*, this gives T=*0.414 hours*. If the system is error-free for *0.414 hours* (*25* minutes), one can be *90%* confident that the error rate is less than r=*5.56/hour*. In general, for a confidence level C that the error rate is less than r, the error free period is given by:

$$T = -\frac{1}{r} ln(1-C) \tag{2.16}$$

For example, if C=*0.99* and r=*5.56/hour*, then T=*0.826 hours*. By doubling the test time, the confidence level is increased from *90%* to *99%*.

Where the objective is to measure the error rate accurately to determine an upper bound on the error rate, one can decrease the measurement time dramatically by stressing the system under test. The stress produces a higher error rate, and the higher error rate can be measured more quickly. Then, if the error rate as a function of stress is known, one can extrapolate to the error rate the system would have when it is not stressed.

- **Quality and Intelligibility of Speech Signals**

Getting a speech signal through a noisy, interference-ridden channel is a well-known problem in communication. Attempts to find a new speech enhancement technique have met with varying degrees of success. In order to evaluate speech processing algorithms, it would be useful to be able to identify the similarities and differences in perceived quality and subjectively measured intelligibility. The quality of speech addresses "how" a speaker conveys an utterance and may include such attributes as "naturalness," or speaker recognizability. In contrast, speech intelligibility is concerned with what the speaker has said· the meaning or information content behind the words.

The difficulty in separating the notions of quality and intelligibility is due in part to the difficulty in isolating and characterizing those acoustic-correlates of quality and intelligibility in speech. However, extensive research has been carried out in developing both subjective and objective tests to ascertain quality and intelligibility.

Subjective measures are based on the opinion of a listener or a group of listeners regarding the quality of an utterance. As suggested by Hecker and Williams (1966) [12], one means of classifying subjective quality measures is to group them as utilitarian or analytical. Utilitarian measures employ testing procedures that are both efficient and reliable and that produce a measure of speech quality on a unidimensional scale The main advantage is that a single number results, which can be used to directly compare speech processing systems.

In contrast, analytical methods seek to identify the underlying psychological components that determine perceived quality. These methods are oriented more toward characterizing speech perception than measuring perceived quality, and typically use more than one dimension for reporting results.

Maximum intelligibility is a function of effective signal-to-noise ratio. Most methods of speech enhancement have in common the assumption that the power spectrum of a signal corrupted by uncorrelated noise is equal to the sum of the signal spectrum and the noise spectrum. Thus, power subtraction is used to separate a speech signal from the noise observation. The suppression rules for the power subtraction, Wiener filtering, and maximum likelihood algorithms are used.

A useful criterion for estimating a speech signal is to minimize the mean-square error $min\ \{E(s-\tilde{s})^2\}$ There are many methods in research for speech estimation. The most known area of research is system identification [13] A least square analysis system identification technique is based on optimal estimation. The output speech is modelled as having been obtained via linear filtering of the input signal. For the least square method the model is unknown, and one makes an optimal estimate so that the estimated output and the desired speech have little difference [14].

The least mean-square adaptation algorithm is an iterative, minimum-seeking method for determining the least square solution. Assuming that $\underline{\hat{h}}_i$ is an estimate of $\underline{\hat{h}}$ at i^{th} iteration, the new estimate $\underline{\hat{h}}_{i+}$ is determined as:

$$\underline{\hat{h}}_{i+1} = \underline{\hat{h}} - u\nabla$$

(2.17)

where ∇ is the gradient of $\|\hat{e}(n)\|^2$ with respect to $\underline{\hat{h}}$ and u is a constant. Basically, ∇ determines the direction in which the correction is made for the $i+1$ iteration and u is a constant which controls the size of the step taken in that direction. Since $\|\hat{e}(n)\|^2$ is a quadratic function of $\underline{\hat{h}}$, a single minimum exists in the error surface and it can be shown that the algorithm converges to this minimum if the step size is not too large. The LMS adaptation algorithm uses the gradient of a single error:

$$\nabla \approx -2x(n-m)\,\hat{e}(n)$$

(2.18)

New estimates of $\underline{\hat{h}}$ are then computed on a sample-by-sample basis as data samples $x(n)$ and $y(n)$ become available. The new estimate of the m^{th} coefficient of $\underline{\hat{h}}$ is then computed as:

$$\hat{h}_{n+1}(m) = \hat{h}_n(m) + 2ux(n-m)\,\hat{e}(n)$$

(2.19)

Since the choice of u depends on the variance of $x(n)$, a self-normalizing form of the LMS adaptation algorithm,

$$\hat{h}_{n+1}(m) = \hat{h}_n(m) + Kx(n-m)\,\hat{e}(n)/\sigma_x^2$$

(2.20)

was used.

To measure the performance of these techniques, the Q measure is defined. The Q measure is basically the ratio of the norm of the coefficient error vector to the norm of the true coefficient vector. It is useful for characterizing how well the estimate $\hat{h}$ approximates the true h. The second measure, Q', is a frequency-weighted measure which is useful for characterizing the performance of system identification methods for nonwhite inputs. The Q measure has the form:

$$Q = 10 \cdot \log \left| \frac{\sum_{m=0}^{M-1} [h(m) - \hat{h}(m)]^2}{\sum_{m=0}^{M-1} h^2(m)} \right|$$

(2.21)

It can be shown that for a white input signal $x(n)$ and with uncorrelated white noise $e(n)$, the quantity Q is a simple function of the system parameters, namely N, M and the signal-to-noise ratio, $\frac{s}{n} = 10 \cdot \log \left[\sigma_v^2 / \sigma_e^2 \right]$, at the output of the system, and is of the form:

$$Q \Big|_{\substack{white \\ input}} \cong 10 \cdot \log\left(\frac{M}{N}\right) - \frac{s}{n} (dB)$$

(2.22)

where M is the estimated duration of the impulse response, and N is the signal duration.

The above equation predicts the performance of the least square analysis system identification method for white uncorrelated inputs. The quantity Q is directly dependent on the signal-to-noise ratio at the output of the system. It improves (decreases) by 3 dB per doubling of the block size, and it degrades (increases) with $\log M$. For the case of nonwhite inputs, it is not possible to express Q in the form of Eq. 2.16. In general, for nonwhite inputs, the value of Q will be larger than Eq.2.16 and in this sense Eq.2.16 represents a lower bound on the expected value of Q. That is, a white uncorrelated input signal is the best form of input signal to use in the system identification problem. The modified Q

measure, Q', applies frequency weighting which is equal to that of the frequency response of the filter which is used to create the nonwhite signal from the white signal $x(n)$ [15]. This weighting can conveniently be achieved by convolving $h(n)$ and $\hat{h}(n)$ by nonwhitening impulse response $g(n)$. This procedure serves to weight the performance measure by the frequency spectrum of the input signal.

$$Q' = 10 \cdot \log \left| \frac{\sum_n [(h(n) - \hat{h}(n)) \otimes g(n)]^2}{\sum_n [h(n) \otimes g(n)]^2} \right| \tag{2.23}$$

The properties of Q' as a function of *N*, *M*, and s/n are somewhat more complicated than those of the Q measure, since the "coloring unwhitening" filter, $g(n)$, affects the result. However, it can be shown that the properties of Q' and Q are quite similar.

In this work, the basis of the separation process is to use an Almost Symmetrical Time-Varying ARMA model to track the BPSK signal, which interferes with the speech signal. This algorithm is simple to implement and possesses a fast convergence property. The factors which affect the quality and the intelligibility of a speech signal are the presence of the residual BPSK signal and the background noise. The aim of the work is to check the influence of the speech-to-digital communication signal ratio on the quality and intelligibility.

The first part of the ARMA model is suppressing the BPSK signal without affecting the intelligibility of the speech. At the output of the MA part of the model, a residual BPSK signal was found at the synchronous point when a digital signal changes from one level to another. An extra blanking algorithm is used to eliminate this residual signal.

In the research a formal listening test is performed to assess the quality and intelligibility of the enhanced speech. In all aspects the intelligibility of enhanced speech was the same as the unprocessed signal. The speech-to-digital communication signal ratio affects the quality more than the intelligibility of a speech signal. The intelligibility is more dependent on the speaker, on context, and on the phonetic content.

Chapter 3

Structural Signal Modeling

Work on structural signal modeling has been escalating. Structural signal processing can approximate any continuous input-output mapping to any desired degree of approximation, given a sufficient number of hidden units. This property is also shared by classical methods based on the use of smooth functions such as algebraic or trigonometric polynomials. What is really important, therefore, is the rate of convergence with which the unknown function is approximated for a prescribed set of basis functions.

In contrast to conventional filtering, structural signal processing-based methods are attractive for practical applications by virtue of the ability to deal with nonlinearity, non-stationarity, and non Gaussianity. Moreover, these methods offer robustness with respect to parameter tuning and sample properties, which are important for a good setting of user-tunable parameters by nonexpert users.

To start with the design one would assume a complete knowledge of the signal properties and initial state of structure, and the exact time history of the input signal. Consider, for example, the simplest situation in which a structure is an ideal first order autoregressive model:

$$y(n) - a_1 y(n-1) = x(n) \tag{3.1}$$

where $x(n)$ represents the input signal.

If it is assumed that $x(n)$ on the right hand side of Eq. 3.1 is precisely known, then the output of the structure can be computed exactly. This is called a deterministic structural signal modeling. On the other hand, if the input signal of the above model is random, obviously the output becomes random. Examples of input signals which are essentially random are numerous in signal processing. Speech and digital communication signals fall into this general category.

The form and material of structural signal properties dictate their behavior, which in turn dictates the character of the analytical model. A structural signal model is linear if the relationship between input and output signal is represented by a linear operator If this operator is not linear, the structural signal model becomes a nonlinear one. A structural signal model is time-invariant or time-varying, depending on whether the parameters of the model are time-invariant or time-varying during signal processing. All these behavioral aspects of the structural signal modeling will have a significant influence on the nature of the analysis used in studying structural signal processing [16] Finally, the nature of the input signal, which is independent of the function of the model, will also influence the analysis. For example, a simple analytical model may suffice to reject a known frequency sinusoid interference other than that which would be required for enhancing speech signal which is interfered with by other speech signals It is important to recognize at the outset that the concepts that will be presented can be extended to the solutio. of many other classes of structural signal modeling problems.

3.1 The Concept of Structural Signal Representation (SSR); Signal-system Duality

The basic of the concept of structural signal representation has been explained as "the idea that a process under analysis is a particular realization of an ensemble of different transformations of the same signal" [16]. Although the characteristics of different realizations may vary from one to another, there are some intrinsic properties which are considered to belong to the ensemble, as well as to each individual realization.

Structural signal representation is the subdiscipline where different realizations of the same signal $s(i, \alpha)$ are combined into an ensemble by the application of a certain rule. This kind of description is well-suited to clarify many important issues related to the representation, prediction, and identification of signal processes For instance, the spectral density $S(j\omega)$ of signal $s(i, \alpha)$ may be considered as a particular realization of signal $s(i)$ given by Fourier transform; the autocorrelation function $R_{ss}(\tau)$ of signal $s(i)$ is yet another realization with a new independent variable-time-lag τ. Because the autocorrelation function is the expected value of a signal $s(i$ multiplied by a delayed version of itself, which is a quadratic transformation, it can be viewed as the expected value of a signal $y(i, \tau)$ that is obtained by nonlinear transforming $s(i)$, [17]

$$y(i, \tau) = s\left(i + \frac{\tau}{2}\right)s\left(i - \frac{\tau}{2}\right)$$

(3.2)

and the expected value of the signal is:

$$R_{ss}(\tau) = E[y(i, \tau)]$$

(3.3)

In particular, if the autocorrelation function is zero for all τ $R_{ss}(\tau) \equiv 0$, the operation becomes a null-operation, which characterizes a property of the ensemble $s_k(i)$. Similar to classical signal processing, structural signal representation SSR may be classified as either analog or discrete (based on the properties of independent variables).Depending on the properties of the operators, SSR may be categorized as either linear or nonlinear, time-invariant or time-varying. There exists, however, one essential property which belongs explicitly to SSR, which is parameter invariance (partial or complete).

To illustrate this property, consider a null-operator given by

$$z(t) = \sum_{i=0}^{p} a_i(t)s(t-i) = 0$$

(3.4)

where s_t is a continuous sinusoid of arbitrarily chosen parameters (frequency, amplitude, phase) and the model parameter value $a_i(t)$ depends on the signal frequencies [18]. A null-operator transforms the input signal to zero independently of such parameters as amplitudes and phases.

Another example is a null-operator which applies to any arbitrary signal with time-varying parameters and transfer to zero is presented [19]. A constant envelope FM signal is one example in this situation. When a strong FM signal, which the instantaneous phase *a priori* known, is burying a weak narrow band digital communication signal, one way of performing the removal of this strong FM signal is to apply a nonuniform sampling and use a second order time-invariant AS-ARMA model. Applying a nonuniform sampling to the input mixture will concentrate a wideband FM signal into a single spectrum line, and transform a digital communication signal into wideband spectrum. A second order AS-ARMA model will cancel the FM signal, while leaving unchanged the desired signal [18]. Another approach to this problem is to use a second order time-varying AS-ARMA model. An *a priori* knowledge of the instantaneous phase of the FM signal is used to compute the model parameters. A perfect separation can be achieved if the true phase of the FM signal is available. If the coefficients are computed by using an estimated phase, then the parameters of the model deviate from their true values, and an ideal separation becomes impossible.

3.2 Linear Time-invariant SSR: MA, AR, ARMA Modeling.

For decades, the time-invariant parametric model has been a subject of intensive study in the field of signal processing. The properties and applications of these time-invariant models are well known.

The first known example of a parametric model-based signal processing is the model for describing exponential by roots of exponential polynomials, first suggested by Baron de Prony. Yule introduced the idea of the regression equation and Walker combined such a representation with the LMS criterion for the analysis of damped sinusoidal functions corrupted by noise. The result of this analysis has been the so-called Yule-Walker equation which is very broadly used in regression analysis [20].

The Yule-Walker equation is the basis of some very efficient algorithms for high resolution spectral analysis. In recent years Moving Average (MA), Autoregressive (AR), and the more general Autoregressive Moving Average (ARMA) models have been intensively investigated.

If a signal can be represented by a linear combination of a complete set of independent functions, such a set of linear independent functions constitute a system of fundamental solutions (Wronskian-system) for a particular linear homogenous equation. Thus the signal may be considered as a solution of this linear homogenous equation. Such a homogenous equation is defined as the MA parametric model associated with this signal. If the coefficients of this homogenous equation are fixed values, the parametric model is a time-invariant one, if at least one of the coefficients of this model is time-varying, then the model is a time-varying one

If a signal can be represented by a certain W-system, then the output of the parametric model associated with such a signal is zero. Therefore such a parametric model may be treated as a null-operator with respect to a given signal

When an $AR(p)$ model is assumed for the N sample input signals $x_1, x_2, \ldots x_n$ the prediction error signal $e(p, n)$ is written by the following equation

$$e(p,n) = x(n) + \sum_{i=1}^{p} a_i x(n-i) \tag{3.5}$$

where a_i is the prediction coefficient This process is using N sample of the output signal to compute optimum coefficients, which are used in the model to estimate the current input signal There are a number of methods to determine the model coefficients. All of them target to minimize the prediction error

The theory of ARMA models has its roots in the early contribution [21,22] on AR models, and has found many applications in connection with time-series prediction and identification. These models are also called parametric models. The properties and applications of these time-invariant models are well known. In a certain sense, time-invariant signal modelling is equivalent to the problem of signal filtering based on knowledge of the Fourier spectrum of signal under analysis

A model for a linear system, suitable for implementation on a digital computer, is usually determined from the transfer function of the system $H(f)$, or the unit impulse response $h(t)$. If the transfer function $H(f)$, is for a lowpass type system, the computer model is easily determined directly from $H(f)$ using one of the standard digital filter synthesis techniques that map a transfer function into an equivalent digital filter Perhaps the most popular synthesis techniques are those that yield impulse-invariant, step-invariant, and bilinear z-transform filters. All of these synthesis techniques involve approximations; thus, it is important that the approximations be understood if the simulation user is to have confidence in the simulation result.

A complex envelope signal representation is generally used for bandpass signals If the system is a bandpass system, the unit-impulse response of the system will be a bandpass signal The complex envelope is usually expressed in rectangular form

$$\tilde{x}(t) = x_d + jx_q(t)$$

(3.6)

where $\tilde{x}_d(t)$ is the direct (or real) component of $\tilde{x}(t)$ and $\tilde{x}_q(t)$ is the quadrature (or imaginary) component of $\tilde{x}(t)$. As such, that unit-impulse response is usually represented by the complex envelope model of the bandpass system, defined by,

$$\tilde{h}(t) = h_d(t) + jh_q(t)$$

(3.7)

The complex envelope of the system output $\tilde{y}(t)$, is the convolution of the complex envelope of the input, represented by Eq. 3.6, and as given by Eq.3.7. This yields

$$\tilde{y}(t) = [x_d + jx_q(t)] \otimes [h_d(t) + jh_q(t)]$$

(3.8)

where $\otimes$ denotes convolution. The preceding expression can be written as:

$$\tilde{y}(t) = [x_d(t) \otimes h_d(t) - x_q(t) \otimes h_q(t)]$$

$$+ (j[x_d(t) \otimes h_q(t) + x_q(t) \otimes h_d(t)])$$

(3.9)

This yields the structure shown in Figure 3.1 Since the functions $h_d(t)$ and $h_q(t)$ represent lowpass signals, computer models for these signals can be realized using the same techniques described in the preceding paragraph. Two filters will be necessary, one for $h_d(t)$ and one for $h_q(t)$.

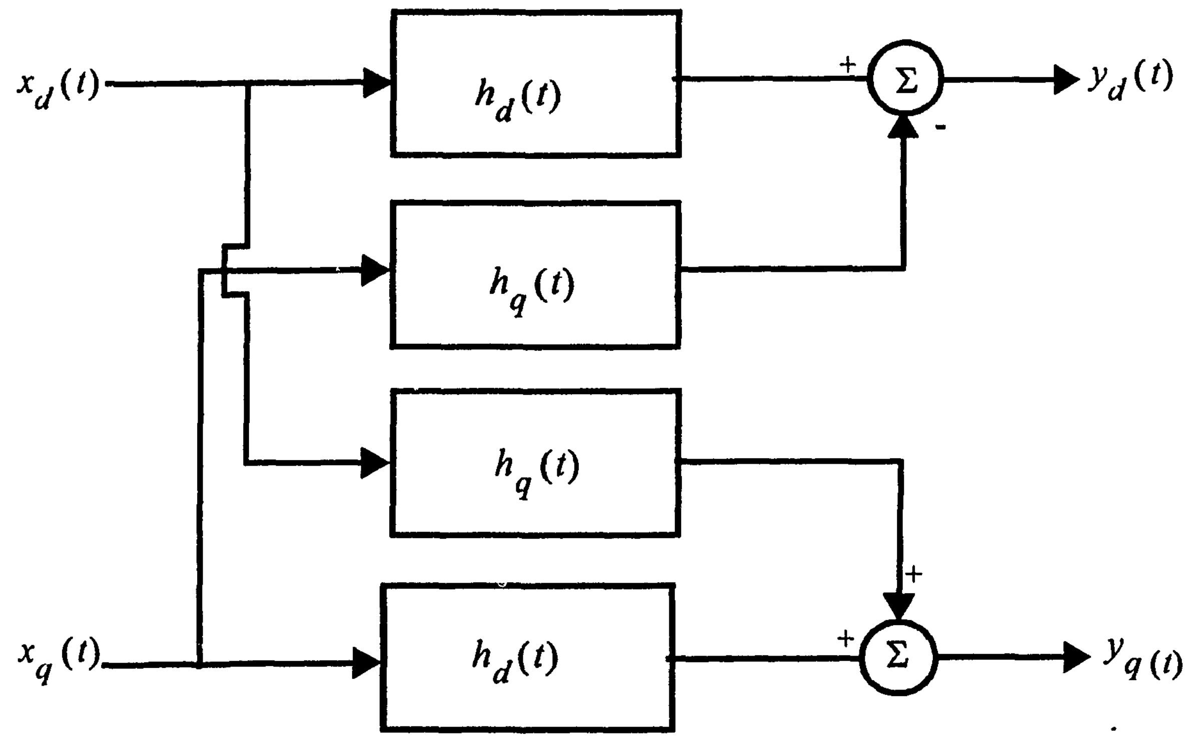


Figure 3.1 Complex envelope representation of bandpass linear system.

Many of the linear systems used in a communication system involve a filtering operation. Since filters have memory, past input or output samples are used in forming the current system output. Efficient filtering routines are therefore essential elements in a simulation program.

3.3 Linear Time-varying SSR

It is a well-known fact that an ideal separation of signals by time-invariant filtering may be achieved only if the Fourier spectra of the superimposed signals are located in non-overlapping frequency ranges. However, in most signal processing procedures, this is not the case. It is an established fact that a time-invariant model is optimal for the processing of signals with rational Fourier spectra. However, if the Fourier spectrum of a signal is not rational, only an infinite order time-invariant parametric model can be associated with such a signal. On the other hand, it is possible to find a finite order time-varying parametric model to represent a signal with nonrational Fourier spectrum

In the separation of signals for which the Fourier spectra are overlapping, using a conventional linear time-invariant filter becomes almost an impossible task In contrast, a linear time-varying structural signal representation is considered in order to provide an effective model for such signals separation. Since the process of constructing a time-varying parametric model is essentially independent of the knowledge of the Fourier spectrum of the input signal, such a model may be used for the separation of a certain class of signals with overlapping Fourier spectra.

In SSR, the model coefficients, synchronization and being time-varying are derived from the input mixture. In some cases the computation of coefficients of such a model could be performed by applying the input mixture to a phase-locked loop These coefficients are used to design a specific filter based on a structural signal representation of a signal under analysis.

The straight forward time-varying structural signal representations are MA, AR and ARMA models. For instance, a nonstationary ARMA process $O(t)$ may be represented by the following difference equation,

$$O(t) = e(t) + \sum_{i=1}^{p} a_i(t) O(t-i) + \sum_{i=0}^{q} b_i(t) r(t-i) \tag{3.10}$$

where $a_i(t)$ and $b_i(t)$are time-varying parameters of AR and MA models respectively, and e_t is a null-operator residue. Figure 3 2 illustrates a time-varying ARMA model, which may be used in a variety of problem, such as signal synthesis, spectral analysis (spectral estimation), signal classification, and data compression [23] The above model becomes an AR model if all the coefficients of an MA model are zero, $b_i(t) = 0$ Assuming that the given signal can be represented as the output of the system of Eq.3 10, a standard problem is to estimate the parameters $a_i(t)$, $b_i(t)$ and the variance of e_t from a knowledge of N output samples. A reasonable method of estimating these quantities is to use the minimum mean-square error (MMSE) criterion

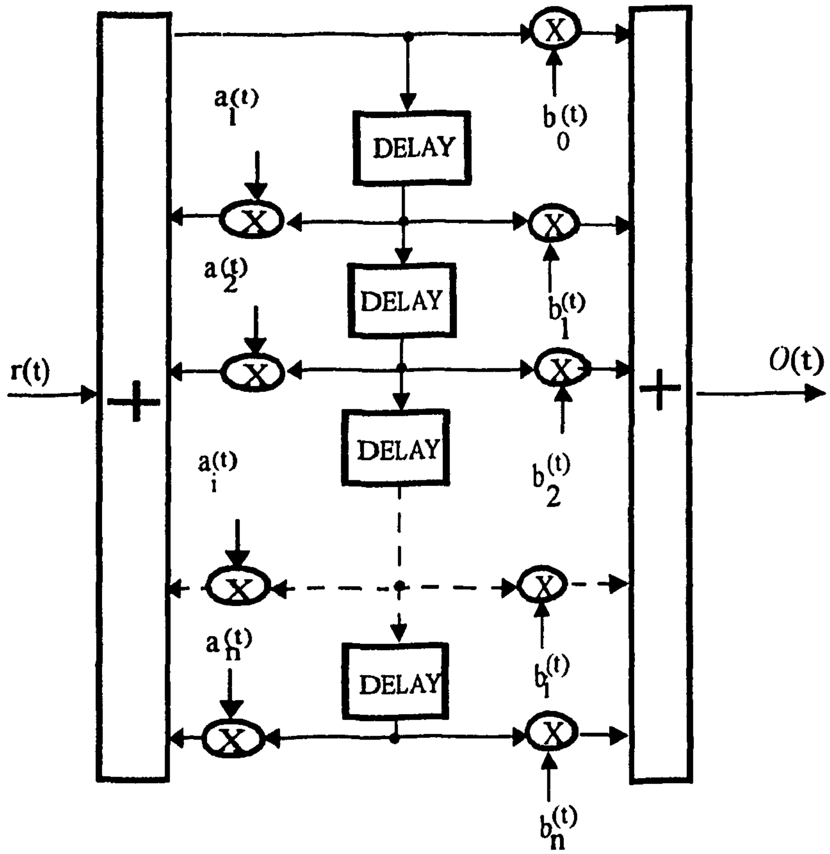


Figure 3.2 Time-varying ARMA model

3.4 Nonlinear Systems and Nonlinear SSR

Nonlinear system models are common in many areas of scientific inquiry, such as physics, biology, and engineering. A common practice is to determine an appropriate model for a particular system, and then to design and carry out experiments to determine the model's parameters.

A popular approach to the identification of nonlinear dynamical systems from input-output measurements is to model the system in terms of Volterra series, which is a generalization of the power-series (or polynomial) representation of a memoryless system to systems with memory, and then to identify one-by-one the Volterra kernels, each one of which characterizes one term in the series representation. The first kernel is the impulse response of the linear part of the system. The second kernel is a two-dimensional generalization of the impulse response of the quadratic part of the system, and so on. Common approaches to identifying the kernels are based on cross-correlation measurements between the unknown system output and specially designed nonlinear functions of the system input A well-known example is the Volterra series expansion. Unfortunately, this expansion is computationally expensive and therefore rarely used.

- **Models for Nonlinear Systems**

Nonlinear and time-varying systems present special difficulties when bandpass models for these systems are needed. While complex envelope models exist for linear, time-varying systems, there is no guarantee that a complex envelope model exists for systems that are both nonlinear and time-varying. One must rely on approximation methods to model these devices.

Little can be said about the most general class of time-varying, nonlinear systems. The only method that ensures that these systems can be accurately modeled is to translate the complex envelope back to a bandpass signal and pass it through an appropriate device model.

To develop more computationally efficient models, one must make assumptions about the device model. There are also a variety of models for nonlinear but time-invariant systems. There is a special class of nonlinear devices that have very short, or no memory. In a true memoryless device, such as a square-law device, the output is only a function of the current input. If a single sinusoid is placed into these devices, the output will have terms only at the harmonics of the input frequency, and if the bandpass filter follows the memoryless nonlinearity, all but the first harmonic term can be removed. Thus, a sinusoidal input produces a sinusoidal output, where the amplitude of the output may be a nonlinear function of the amplitude of input. This type of device lends itself well to the complex envelope representation.

Another class of interesting systems have "short" memory, i.e., the time constant of the nonlinearity is long with respect to the carrier frequency, but short with respect to the message waveform. These systems can be called complex envelope memoryless systems, or envelope nonlinearities. This is because the complex envelope of the output can be approximated by the memoryless, but nonlinear, function of the complex envelope of the input. As with the truly memoryless nonlinearity, the complex envelope of the input is decomposed into its amplitude and phase. The amplitude is both passed through a nonlinear device and used to alter the phase of the signal. If the input to the (assumed memoryless) nonlinearity is:

$$x(t) = A(t) \cos[2\pi f_c t + \phi(t)] \tag{3.11}$$

then the output is represented by,

$$y(t) = f[A(t)] \cos\{2\pi f_c t + g[A(t)] + \phi(t)\} \tag{3.12}$$

The function $f[A(t)]$ is known as the AM-to-AM conversion characteristic and $g[A(t)]$ is known as the AM-to-PM conversion characteristic. For a constant envelope of $x(t)$, $A(t)$ is a constant and thus $f[A(t)]$ and $g[A(t)]$ are constants. This explains the interest in constant envelope modulation techniques.

• **Nonlinear SSR**

Due to the severity of signal interference, structural models may respond nonlinearly when subjected to these signal hazards. In addition, the interference characteristics reveal nonlinear properties and consequently have to be modeled by a nonlinear structure. MA, AR, and ARMA models are successful models of linear systems because they allow a complicated system to be represented in terms of a small number of parameters [24]. This statement applies to nonlinear versions of these models. A system with an infinite number of nonzero infinite length Volterra kernels could possibly be represented with a small number of parameters using a nonlinear version of the AR model.

A nonlinear MA model is a linear MA (all zero) model with nonlinear feed-forward paths. Such a model has a finite number of non-zero Volterra kernels (high order impulse responses) and the kernel sequence is also finite, i.e., the model has a finite impulse response (FIR). Similarly, a nonlinear AR model is defined as an all pole linear AR model with nonlinear feedback. In this case, the model has an infinite number of non-zero Volterra Kernels and each of these kernel sequences is infinite in general. The nonlinear AR model has an infinite impulse response (IIR) A linear system with nonlinear feedback loops is called feedback linearization (FBL) because nonlinear feedback transformations can map such systems to linear systems by redefining the nonlinear feedback as input, so that the system appears to be linear.

The linear AR model is the most popular approach to time series modeling when only output data is available. This is because accurate estimates of AR parameters are found by solving a set of linear equations [25]. The nonlinear AR model also has the very desirable property that all parameters, even the nonlinear ones, are found via linear equations Applications of the nonlinear AR model results indicate that this data may be represented by a nonlinear model that contains fewer parameters (and less computations) than a linear model representing the same data.

Nonlinear ARMA models are an extension of ARMA models where the sampled response of a system is modelled as the weighted sum of the previous input and response values plus nonlinear combination of these input and response samples. The nonlinear differential equations are mapped to nonlinear ARMA models that relate the sampled input to the sampled response of a nonlinear system. Briefly, a nonlinear differential equation of the form [26]:

$$\sum_i \alpha_i \frac{d^i y}{dt_i} + \sum_i \beta_i y(t)^i = x(t) \tag{3.13}$$

is divided into its linear part and nonlinear elements are transferred to the right hand side and grouped with input.

Where $\sum_i \alpha_i \frac{d^i y}{dt_i}$ is the sum of linear difference functions, and $\sum_i \beta_i y(t)^i$ is the sum of nonlinear functions. An impulse invariant mapping of the linear part of the model yields an AR model. Nonlinear ARMA models can be obtained by cascading a nonlinear or linear MA into a nonlinear or linear AR model with at least one section nonlinear. Also, since nonlinear or linear cascades are not commutative, reversing the order (AR into MA) produces another set of nonlinear ARMA models.

A nonlinear system could be time-invariant or time-varying depending on whether the parameters are time-invariant or time-varying; the system is to be one with memory if the input at an earlier time will affect the output at a later time, or memoryless if the output is determined from the instantaneous value of the input signal.

CHAPTER 4

Structural ARMA Modeling and Separation of Frequency Overlapping Signals

Separating signals from other signals whose spectra overlap the entire bandwidth is of major interest in digital signal processing. In this section the concept of signal subspace, time-varying null-operator and its property is introduced. First, the concept of signal parametric modeling is established by exploring the relationship between a homogeneous parametric model and its fundamental solutions. If a signal can be represented by a linear combination of a complete set of independent functions, such a set of linear independent functions constitutes a system of fundamental solutions for a particular linear homogeneous equation. Thus, the signal may be considered as a solution of this homogeneous equation. In 1815 Hoene Wronski a Polish mathematician, published his celebrated memoir, "Philosophie de la Technique Algorithmique", where he gave the following fundamental result (as presented in [27]):

Let $\{\varphi_i(t)\}_{i=1}^{N}$ $t \in \Re$, be a system of any N functions which together with their first $N-1$) derivatives are continuous. The necessary and sufficient condition for such a system of functions $\{\varphi_1(t), \varphi_2(t), \ldots, \varphi_N(t)\}$ to be the linear independent solutions of the homogeneous equation is that a Wronskian determinant be nonvanishing at any instant of time.

$$M^{N}_{\frac{d}{dt}}\left(\varphi_1(t), \varphi_2(t), \ldots, \varphi_N(t)\right) \equiv \begin{bmatrix} \varphi_1(t) & \varphi_2(t) & \ldots & \varphi_N(t) \\ \varphi'_1(t) & \varphi'_2(t) & \ldots & \varphi'_N(t) \\ \ldots & \ldots & \ldots & \ldots \\ \varphi_1^{[N-1]}(t) & \varphi_2^{[N-1]}(t) & \ldots & \varphi_N^{[N-1]}(t) \end{bmatrix} \neq 0 \tag{4.1}$$

In 1880, *E. Casorati* [28] presented a discrete version of Wronskian for a system of discrete functions $\{\varphi_i(k)\}_{i=1}^{N}$ $k \in Z^{+}$.

$$M_D^N\left(\varphi_1(k), \varphi_2(k), \ldots, \varphi_N(k)\right) \equiv \begin{bmatrix} \varphi_1(k) & \varphi_2(k) & \ldots & \varphi_N(k) \\ D\varphi_1(t) & D\varphi_2(k) & \ldots & D\varphi_N(k) \\ \ldots & \ldots & \ldots & \ldots \\ D^{N-1}\varphi_1(k) & D^{N-1}\varphi_2(k) & \ldots & D^{N-1}\varphi_N(k) \end{bmatrix} \neq 0$$

(4.2)

where D denotes a shift operator, $D^m \varphi_l(k) = \varphi_l(k-m)$.

Later, Bortolotti extended his idea to a general functional setting. He stated if $\{\varphi_i(t)\}_{i=1}^{N}$ be a system of analytic functions having a common domain of convergence, and O be a one-to-one mapping functional operation, the necessary and sufficient condition for such a system of functions being associated with a homogeneous equation with coefficients with respect to O is

$$M_O^N\left(\varphi_1(k), \varphi_2(k), \ldots, \varphi_N(k)\right) \equiv \begin{bmatrix} \varphi_1(k) & \varphi_2(k) & \ldots & \varphi_N(k) \\ O\varphi_1(t) & O\varphi_2(k) & \ldots & O\varphi_N(k) \\ \ldots & \ldots & \ldots & \ldots \\ O^{N-1}\varphi_1(k) & O^{N-1}\varphi_2(k) & \ldots & O^{N-1}\varphi_N(k) \end{bmatrix} \neq 0$$

(4.3)

These three works provide a concrete base in the theory of linear equations, and a mathematical tool for linear structural signal processing. For a system of N linearly independent functions, there exists a unique monic homogeneous equation of order N given by:

$$Q_O(x) \equiv \frac{W_O^{N+1}\left(\varphi_1(t), \varphi_2(t), \ldots, \varphi_l(t), \varphi_{l+1}(t), \ldots, \varphi_N(t)\, x(t)\right)}{W_O^N\left(\varphi_1(t), \varphi_2(t), \ldots, \varphi_N(t)\right)} = 0$$

(4.4)

where $W_O^{N+1}(\ldots)$ is an augmented Wronskian. Such a homogeneous equation also defines a linear operator, which can be expressed as a linear homogeneous equation in the form:

$$\left[1 + a_1(t) O + a_2(t) O^2 + \ldots + a_N(t) O^N\right] x(t) = 0 \tag{4.5}$$

where,

$$a_l(t) = (-1)^{l+1} \frac{M_O^N(\varphi_1(t), \varphi_2(t), \ldots, \varphi_l(t), \varphi_{l+1}(t), \ldots, \varphi_N(t))}{W_O^N(\varphi_1(t), \varphi_2(t), \ldots, \varphi_N(t))} \tag{4.6}$$

and $M_O^N(\ldots)$ is the minor of W_O^{N+1} with respect to the Laplace expansion of $x(t$. Bortolotti's result is of special interest, since it provides a mathematical support to structural signal processing, where the operator O can be treated as an arbitrary chosen building block.

4.1 Signal Subspace and Time-varying Null-operator

Given a linear homogeneous equation, one can find a system of fundamental solutions such that the Wronskian of this system of functions constitutes a positive function. An N-dimensional W-system of function spans an N-dimensional signal subspace, Ω. A linear span of W-system of functions $\{\varphi_i(k)\}_{i=1}^{N}$ constitutes an N-dimensional signal space. Such a signal space is *a* null singularity of time-varying null-operator. An N^{th}-order time-varying null-operator is defined by the homogeneous equation.

Consider two signals that belong to two different subspaces Ω, and Ω_2, and which have the following properties:

$$X(k) \in \Omega_1 \qquad S(k) \in \Omega_2 \tag{4.7}$$

$$
\begin{array}{ll}
\text{I} & \Omega_1 \cup \Omega_2 = \Omega \\
\text{II} & \Omega_1 \cap \Omega_2 = \varnothing
\end{array}
\tag{4.8}
$$

Consider that the above two signals are transmitted simultaneously using a single path and are received at the destination. If the two signals' Fourier spectrum is not overlapped, conventional techniques can be used to filter the interest signal. However, in most cases, and in this application, the two signal Fourier spectra are overlapped. If the interest signal is $S(k)$, one must build a transparent operator Q_1 such as:

$$
\begin{array}{ll}
\text{I} & Q_1[X(k)] = 0 \\
II & Q_1[X(k) + S(k)] = S(k)
\end{array}
\tag{4.9}
$$

If a signal can be represented by a certain W-system, then the output of the parametric model associated with such a signal is zero. Therefore, such a parametric model may be treated as a null-operator with respect to the given signal. The relationship between a null-operator and its null singularity is a very important key to the parametric modeling of a signal. The non-harmonic system usually represents a linear time-varying model. The basic approach to carrying out separation of signals whose Fourier spectra are overlapped, is to build a special time-varying transparent operator. The time-varying transparent operator consists of a time-varying null-operator and time-varying anti-null-operator. The first objective of Eq. 4.9 can be accomplished by a time-varying null-operator. However, to fulfill the second objective one needs both operators. The suppression process distorts other signals which do not belong to the signal subspace Ω_1. In order to satisfy our second objective one needs to compensate the distortion in signal subspace Ω_2.

The *transparency* property of a linear operator Q_1 is defined by Eq. 4.10, which states that if a signal does not belong to a subspace, where a null-operator is modeled, it will pass through the model without change. The distortion, which is caused by the null-operator, is compensated by the anti-null-operator.

$$Q_1[S] = S, if \quad \forall S \quad \varepsilon \quad \bar{\Omega}_1, \; and, \; \Omega_1 \cup \bar{\Omega}_1 = I, \quad \Omega_1 \cap \bar{\Omega}_1 = \varnothing$$

4.10

In the case of a discrete time-varying null-operator and anti-null-operator, where this study's interests are focused, the transparent operator may be expressed by Moving Average (MA) and Autoregressive (AR) models. The moving-average model suppresses all the signals belonging to subspace Ω_1 and constitutes the so-called *annihilation* property, and the autoregressive model recovers those signals which do not belong to the subspace Ω_1 leading to the *transparency* property. Such a model that possesses the *annihilation* and *transparency* properties may be represented in the cascaded form, which can be written as:

$$H^{[1]}_{ARMA}(q^{-1}) = H^{[1]}_{MA}(q^{-1})H^{[1]}_{AR}(q^{-1})$$

(4.11)

where q^{-1} is a delay operator. For time-invariant cases, this model represents a conventional digital filter, where the zeros are created by the MA part and the poles are created by the AR part. The zeros are subtracting energy from the input signal while the poles are adding energy in a very narrowly defined range of frequencies to the output signal.

Annihilation	$H^{[1]}_{MA}(q^{-1})H^{[1]}_{AR}(q^{-1})X(q^{-1}) = 0$
Transparency	$H^{[1]}_{MA}(q^{-1})H^{[1]}_{AR}(q^{-1})S(q^{-1}) = S(q^{-1})$

(4.12)

It is seen that a time-varying null-operator fulfills the first requirement. But this achievement is shared with the expense of distortion to other signals which are not constituted as a null singularity. With that in mind, to ensure the second objective, an anti-null-operator

(AR model) is considered. An anti-null-operator associated with such an MA model is dedicated to compensate the distortions which are created by the use of a time-varying null-operator.

The necessary and sufficient condition for a time-varying operator Q_1 to be transparent to any signals other than its null singularity is that the time-varying anti-null-operator be perfectly symmetrical to the time-varying null-operator.

Thus,

$$H^{[1]}_{ARMA}(q^{-1})\,\{X(q^{-1}) + S(q^{-1})\} = S(q^{-1})$$

(4.13

The second order time-invariant null-anti-null-operator maybe presented as:

$$\boxed{\begin{aligned} &y(k) + a_1(k)y(k-1) + a_2(k)y(k-2) \\ &= x(k) + b_1(k)x(k-1) + b_2(k)x(k-2) \end{aligned}}$$

(4.14)

where, $a_1(k), a_2(k), b_1(k),\ and\ b_2(k)$ are parameters of the operator. If these parameters are constant for all k, the operator is a time-invariant; otherwise, it is a time-varying operator.

4.2 Input of the Model

Figure 4.1 illustrates the application of the processing being considered based on the null-anti-null-operator as applied to the problem of interest: separation of data-communication and speech signals. The digital communication signal is a quasi-periodic signal, while the speech signal is a broadband random process. It is assumed that the input signal occupies the whole bandwidth of the channel. It is true that conventional techniques, such as applying a time-invariant filter to perform this task, is fruitless.

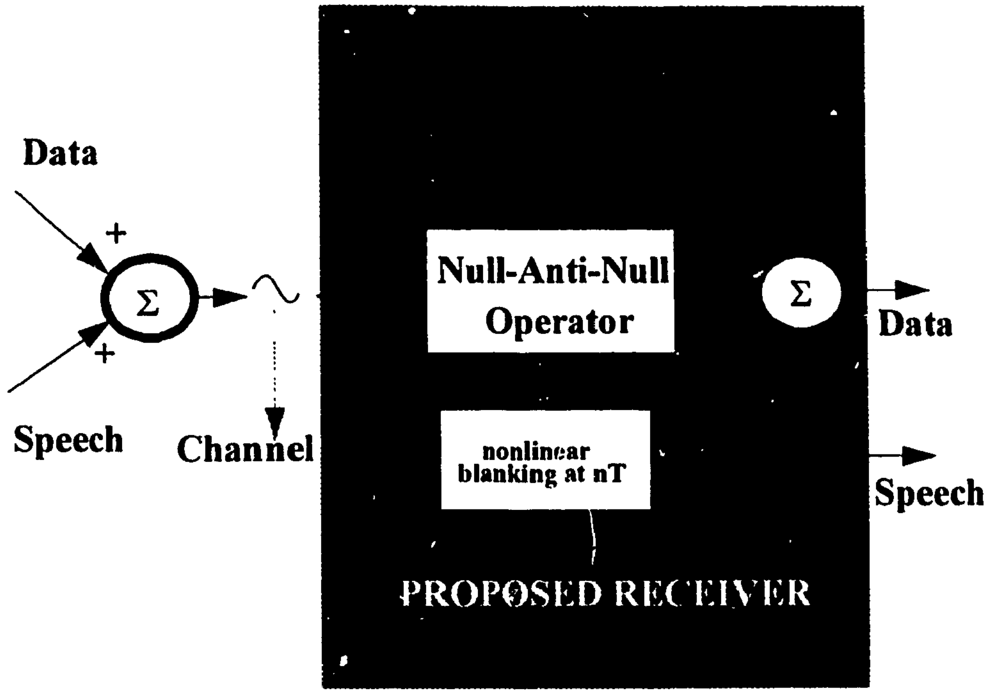


Figure 4.1 Proposed signal separation model

The present work is concerned with one specific example of this problem. When conversation is on in a particular telephone channel it is obvious that this specific channel is not available for other purposes, unless bandwidth or time allocation sharing techniques are applied. This is because the entire frequency band is occupied by the speech signal. If one wants to use the same channel simultaneously for other purposes, for example data-communication, one should find a way to retrieve the signal without losing any information. By exploiting the properties of structural communication signal, the present work shows a technique which can separate a digital communication signal that is mixed with a speech signal.

Let $r(t)$ denote the input to the model, which is the sum of a digital communication signal $x(t)$, a speech signal $s(t)$, and a background noise $n(t)$, which is expressed as the following:

$$r(t) = x(t) + s(t) + n(t) \tag{4.15}$$

where $x(t) = A \cdot \cos[\Psi(t) + \phi$.

Suppose at the outset that $\Psi(t)$ is known and A and ϕ are unknown. Applying the time-varying MA model that represents a null-operator associated with $x(t)$ to the received signal will reject a quasi-periodic signal $x(t)$, and obtain $\tilde{s}(t) + \tilde{n}(t)$ at its output (where $\tilde{s}(t)$ and $\tilde{n}(t)$ are distorted versions of $s(t)$ and $n(t)$ respectively). In the analysis of communication system, given a wave form $x(t)$, one shall be interested in normalized power, which is the quantity $\overline{x^2(t)}$, where the bar indicates the time-average value. In the case of periodic waveforms, the time averaging is done over one cycle.
For a sine wave of amplitude A, the power is simply calculated as

$$P_s = \frac{1}{2}A^2 \tag{4.16}$$

The power density spectrum $PDS(m)$, on other hand, can be calculated as:

$$PDS(m) = \frac{1}{N}|X(m)|^2 \tag{4.17}$$

where, $X(m)$ is the Discrete Fourier Transform (DFT) of the sequence of N samples of the signal $x(n)$:

$$X(m) = \sum_{n=0}^{N-1} x(n) e^{-j\frac{2\pi mn}{N}} \tag{4.18}$$

The BPSK signal, which represents a digital communication signal, has one fixed phase when the data is at one level, and the phase difference is π when the data is at other level. This change is instantaneous in time, which leads the spectrum of a digital communication signal to widen and overlap with the speech spectrum. The instantaneous phase of the BPSK signal is

$$\varphi(k) = \omega_o k + b_n \pi U(k - nT_b) \tag{4.19}$$

where ω_o is the carrier frequency, T_b is the symbol duration, and $U(\cdot)$ is a unit step function.

The phase change is considered to be at the zero crossing point so that the value of the BPSK sample is zero at this instant. To guarantee this condition the phase must satisfy:

$$mod_{\pi}[\omega_o n T_b] \equiv 0$$

(4.20)

If the present input sample is a null singularity of the null-operator, then the output will be zero. The sampling of a pure sinusoid is a good example. The parameter of the time-invariant null-operator can be determined by applying a pure sinusoid signal of frequency ω_0 in the following equation:

$$z(k) = x(k) + ax(k-1) + x(k-2) = 0$$

(4.21)

If $x(k) = A \cdot \cos(\omega_0 t + \varphi$, then the transfer function associated with this operator becomes:

$$1 + ae^{-j\omega_0} + e^{-2j\omega_0} = 0$$

(4.22)

From equation 4.22 the null-operator coefficient is extracted:

$$a = -\frac{1 + e^{-j2\omega_0}}{e^{-j\omega_0}} = -2\cos\omega_0$$

(4.23)

The input signal, whose spectra are shown in Figure4.1, represents a sum of a speech and a digital communication BPSK signal.

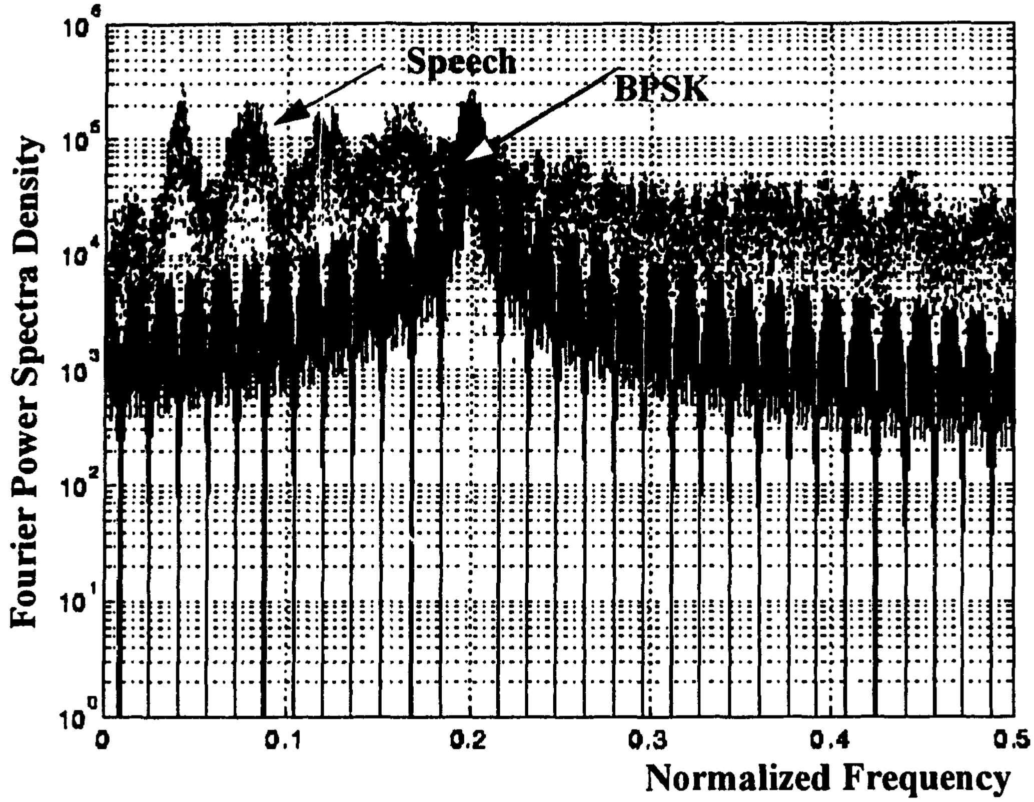


Figure 4.1 Input of the model: Spectra of speech mixed with BPSK signal

4.3 Transfer Function of the Null-anti-Null model.

In general, the input-output relationship of the second order time-varying null-anti-null-operator may be expressed as:

$$y_m = \left[\frac{1 + \sum_{n=1}^{2} b_n(m) z^{-n}}{1 + \sum_{k=1}^{2} a_k(m) z^{-k}} \right] x(m)$$

(4.24)

In order to satisfy the *transparency* property put $a_k = b_k$ for k=1,2

$$y(m) = \left[\frac{1 + \sum_{n=1}^{2} a_n(m) z^{-n}}{1 + \sum_{k=1}^{2} a_k(m) z^{-k}} \right] x(m)$$

(4 25)

Such a second order canonical form defines the linear time-varying transparent operator $Q_1[\cdot]$. It is shown that a physically implementable anti-null-operator cannot be perfectly symmetrical to a null-operator. The anti-null-operator in fact is a function generator, such that its impulse response is a causal function. The ideal *transparency* of the perfect symmetrical operator $Q_1[\cdot]$is based on the assumption that the system operates in the steady-state. Obviously, this is the ideal case: since any physical system has to be causal, once the input is applied to the system at a certain instant, the system contains previously unknown initial states which will produce an undesired residual output. In what follows, the causal transparent operator $Q_1[\cdot]$ cannot provide ideal *transparency* within a short-time. This can be attained by introducing a certain strategy to control the symmetrical nature of operator $Q_1[\cdot]$ The one-sided, almost symmetrical operator $Q^{\alpha}{}_1[\cdot]$, is defined as:

$$y_m \left[1 + \sum_{k=1}^{2} \alpha^k a_k(m) z^{-k} \right] = \left[1 + \sum_{k=1}^{2} a_k(m) z^{-k} \right] x_m$$

(4.26)

where real valued coefficient $\alpha < 1$ is called the symmetry factor. Thus the explicit expression of an AS-ARMA model may be written in a difference equation form

$$y(k) + \alpha a_1(m) y(k-1) + \alpha^2 a_2(m) y(k-2)$$
$$= x(k) + a_1(m) x(k-1) + a_2(m) x(k-2)$$

(4.27)

A one-sided, almost symmetrical transparent operator (AS-ARMA) is built with a null-operator, which is modelled on Ω_1, and an almost symmetric anti-null operator. The anti-null operator reconstructs those signals which are distorted by the null-operator and do not belong to Ω_1. When the input signal is a BPSK signal, the residual signal from the null-operator is found at synchronous time instants.The residues that appear at the input of the AR part are written as:

$$\Delta x\left(k_{t_0}\right) = -2\cos\left(\omega_0 k_{t_0}\right) \tag{4.28}$$

where k_{t_0} is the synchronous time instants.

The response of the AR part for this particular input signal sample is a damped sinusoid with frequency ω_0 has a form:

$$O(k) = -2\alpha^k \cos(\omega_0 k) \tag{4.29}$$

If the symmetry factor is a fixed value $\alpha < 1$, then the residual from the BPSK signal approaches zero $\Delta x(t) \rightarrow 0$ as observation time increases.

When the symmetry factor approaches unity $\alpha \rightarrow 1$, the output signal from the above residual stays for a longer time. The time duration (*L*), which takes this output to be zero, is of interest when system instability is considered. It is obvious that if $\alpha \ll 1$, the output signal damps over a short period. On the other hand, selecting the symmetry factor to be small does not provide the needed *transparency*. Such a *transparency* becomes almost ideal if the symmetry factor approaches unity, i.e. $\alpha \rightarrow 1^-$. A trade-off to this problem is made by choosing the symmetry factor near unity and using different algorithms to eliminate the nonlinear noise which is created at synchronous time instants.

CHAPTER 5

Transparency of the AS-ARMA System

In the previous chapter, the second order ARMA model that satisfies the null-singularity was introduced. It was shown that a symmetrical anti-null-operator can provide an ideal *transparency*; however, it is sensitive to undefined initial conditions. If the symmetry factor approaches unity, the model becomes almost transparent. However, the null-operator residues which happen at synchronous instants create problems for the signal reconstruction and detection models. To fulfill the almost ideal *transparency* property of the model, using a one-sided symmetry factor leading to an almost symmetrical anti-null-operator is introduced.

5.1 Ambiguity at Synchronous Instants

Consider a BPSK input signal sample when phase changes occur. This input signal sample does not satisfy a null-singularity of the MA model. The input sample at this particular instant is written as:

$$x(k) = e^{j(\omega_0 k + \pi)} = -e^{j\omega_0 k} \tag{5.1}$$

where $k = mT_b$ and $m = 1,2,..n$

Unless an extra blanking algorithm is used at this synchronous instant, the MA model creates an undesired impulse at its output, which will cause *transparency* imperfection and may be thought of as an "instability" of the AR system. The duration of this instability depends on the proximity of the symmetry factor value to unity. If the symmetry factor is unity $\alpha = 1$, these spikes which are generated at the output of a moving-average model at the synchronous instants will create additional sinusoid signal with frequency ω_0 at the AR output. In Figure 5.2 the autoregressive model response to these spikes is shown for two specific symmetry factor values when $\alpha < 1$.

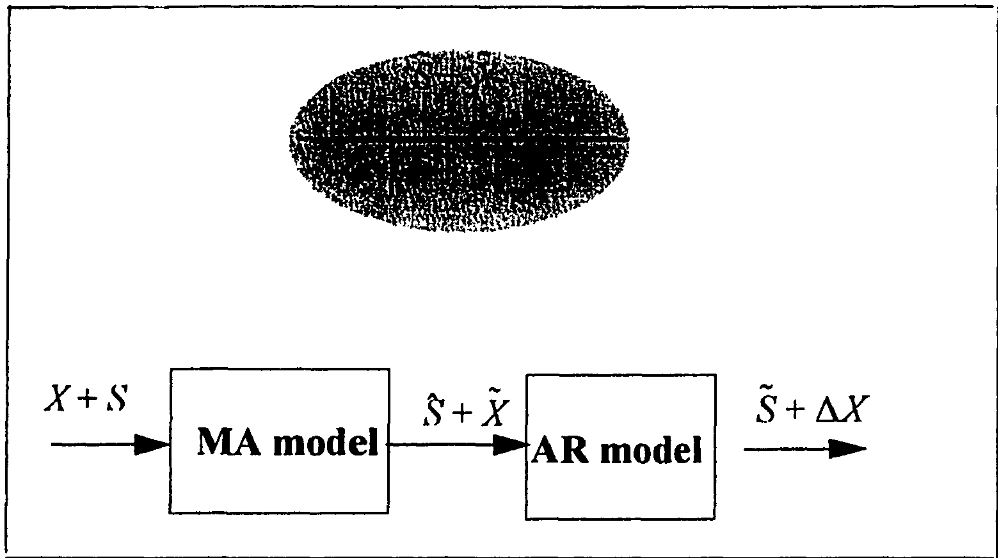


Figure 5.1 ARMA model

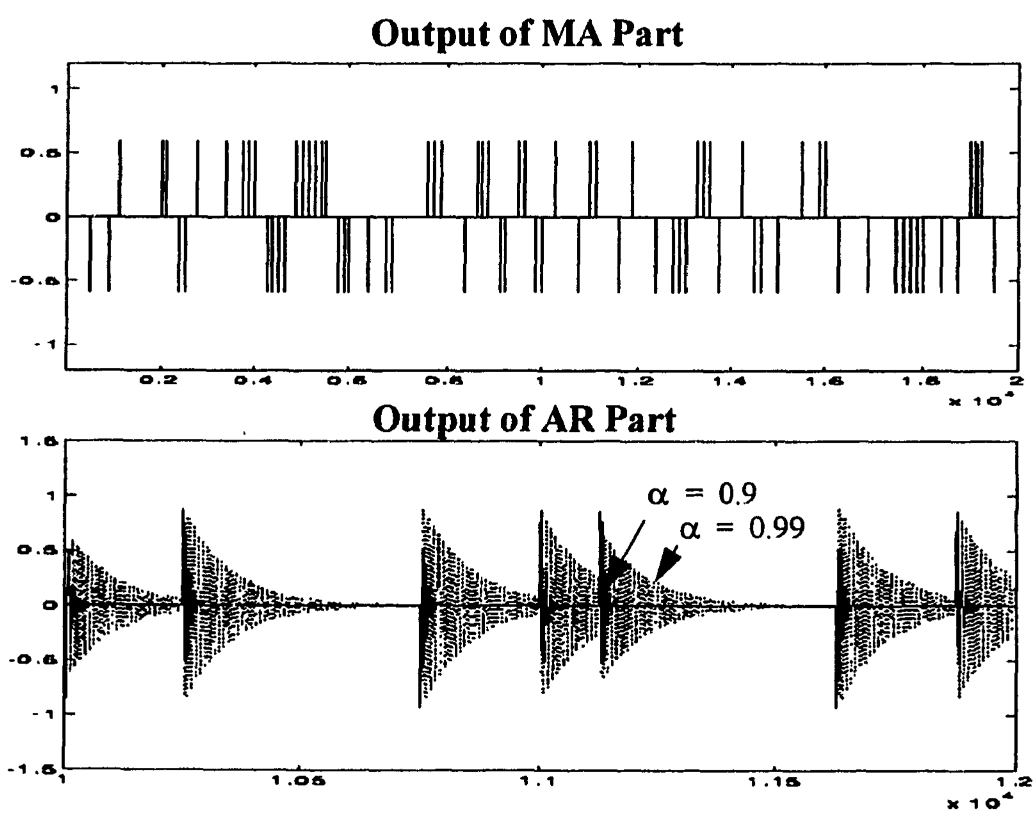


Figure 5.2 The model response to phase changes of BPSK signal

Table 5.1 shows the input output relation of BPSK signal phase keying at $k = mT_b$ instants.

n	k	INPUT (BPSK)	OUTPUT(SPIKE)
+	mT_b-1	$+e^{-j\omega_o}e^{jmT_b}$	0
+	T_b	$+e^{jmT_b}$	$-2\cos\omega_0$
− 1	T_b+	$-e^{-j\omega_o}e^{jmT_b}$	0

Table 5.1 Input-output relations at synchronous instant

5.2 Blanking Algorithm

For a given AS-ARMA model, the MA part is always stable. The stability problem arises only from the AR part. The MA part of the AS-ARMA model creates spurious spectral components when a digital communication signal changes its phase. Unless these spikes are not suppressed by an extra blanking algorithm, they will have an enormous effect on the signal detection and the speech reconstruction process. The first blanking algorithm approach is to replace every MA output sample at the synchronous point with zero. This blanking algorithm is called a deterministic blanking algorithm.

$$H_{MA}^{[1]}(q^{-1})\,[X(q^{-1})]\,\big|_k = \mathbf{0} \tag{5.2}$$

where $k = mT_b$ and $m = 1,2,..n$.

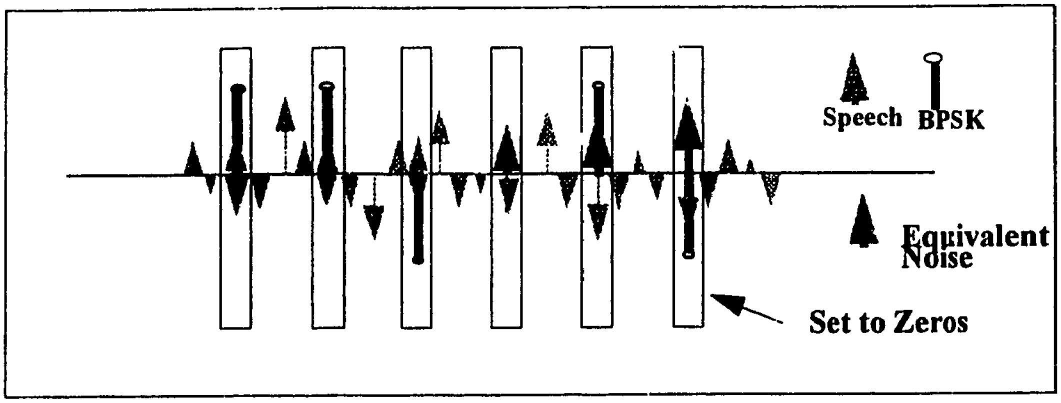


Figure 5.3 Nonlinear blanking operation

If blanking algorithms applied at each synchronous point, one succeeds in eliminating the BPSK residual where it exists. However, since the speech is superimposed on the BPSK signal, blanking operations at synchronous points will eliminate the speech components too, or equivalently, such a nonlinear operation will induce a new noise at the input of $H_{AR}^{[1]}(q^{-1})$ at synchronous points.

5.3 A Quasi-deterministic Periodic Blanking Algorithm

A close examination of this induced noise, which is generated at the output of the MA model, leads to two equivalent Time-Varying AS-ARMA models that cancel fully the BPSK signal. These models can be represented as the following:

$$
\begin{aligned}
& y(k) + \alpha a(m) y(k-1) + \zeta(k)\alpha^2 y(k-2) \\
& = x(k) + a(m) x(k-1) + \zeta(k) x(k-2)
\end{aligned}
$$

or

$$
\begin{aligned}
& y(k) + \zeta(k)\alpha a(m) y(k-1) + \alpha^2 y(k-2) \\
& = x(k) + \zeta(k) a(m) x(k-1) + x(k-2)
\end{aligned}
\tag{5.3}
$$

where

$$\zeta(k) = \begin{cases} \text{sgn}\{b_n b_{n-1}\} & k = nT_b \\ 1 & \textit{otherwise} \end{cases}$$

(5.4)

The above models transform the input BPSK signal into a pure sinusoid signal with a con stant frequency ω_0. The transformation is made by counteracting the phase change at synchronous instants which is caused by the data change from one level to the other. If the input is a pure sinusoid signal, a time-invariant MA model will cancel it perfectly Theoretically, such an ASTV-ARMA model can perform a perfect separation if the information $\{b_n\}$ is *a priori* known. Each system is quasi-deterministic and periodically time-varying. Therefore, if sign change is done to the ASTV-ARMA model parameters at synchronous instants, and the consecutive digital communication input signal $\{b_n\}$ does not change sign, the switch mechanism of the AS-ARMA model will introduce additional noise as the nonlinear blanking operation.

5.4 Trajectory of AS-ARMA Model Parameters

Consider the trajectory of the system coefficients where BPSK, MSK and FSK modulation are used. Figure 5.4 illustrates the model coefficients state within the bit duration The ASTV-ARMA model coefficients, which are used for suppression of the BPSK signal, are jumping between two points at synchronous instants When an ASTV-ARMA model is used to suppress MSK or FSK signals, the coefficients are constantly changing with time If the information source $\{b_n\}$ obeys a certain Markov model, then the optimal periodical AS-ARMA model should be a Markovian one.

If the model is time-invariant, the two coefficients are constant (Figure5.4a) The time-invariant MA model is capable of eliminating a pure sinusoid signal, whose frequency determines the model parameters. A perfect blanking of a BPSK signal is possible if the phase change of the modulated signal is already known This *a priori* knowledge changes the sign of the model parameters at synchronous instants, so that a BPSK signal will be transformed to a pure sinusoid one (Figure 5.4b). For an MSK input signal, the coefficients of the model move from one point to the other to track the input signal frequency. The input signal is a constant amplitude with a continuous frequen :y change, which determines the trajectory of the model coefficients. Figure 5.4c shows how the continuous phase change is translated into the model coefficient during the bit period. FSK modulation for binary data transmission uses two frequencies If the information bit is at one level it uses the first frequency, and the other when the information bit is changed. Figure5.4d shows the coefficients state. At a synchronous point the coefficients might jump from one point to the other if the transmitted data level changes.

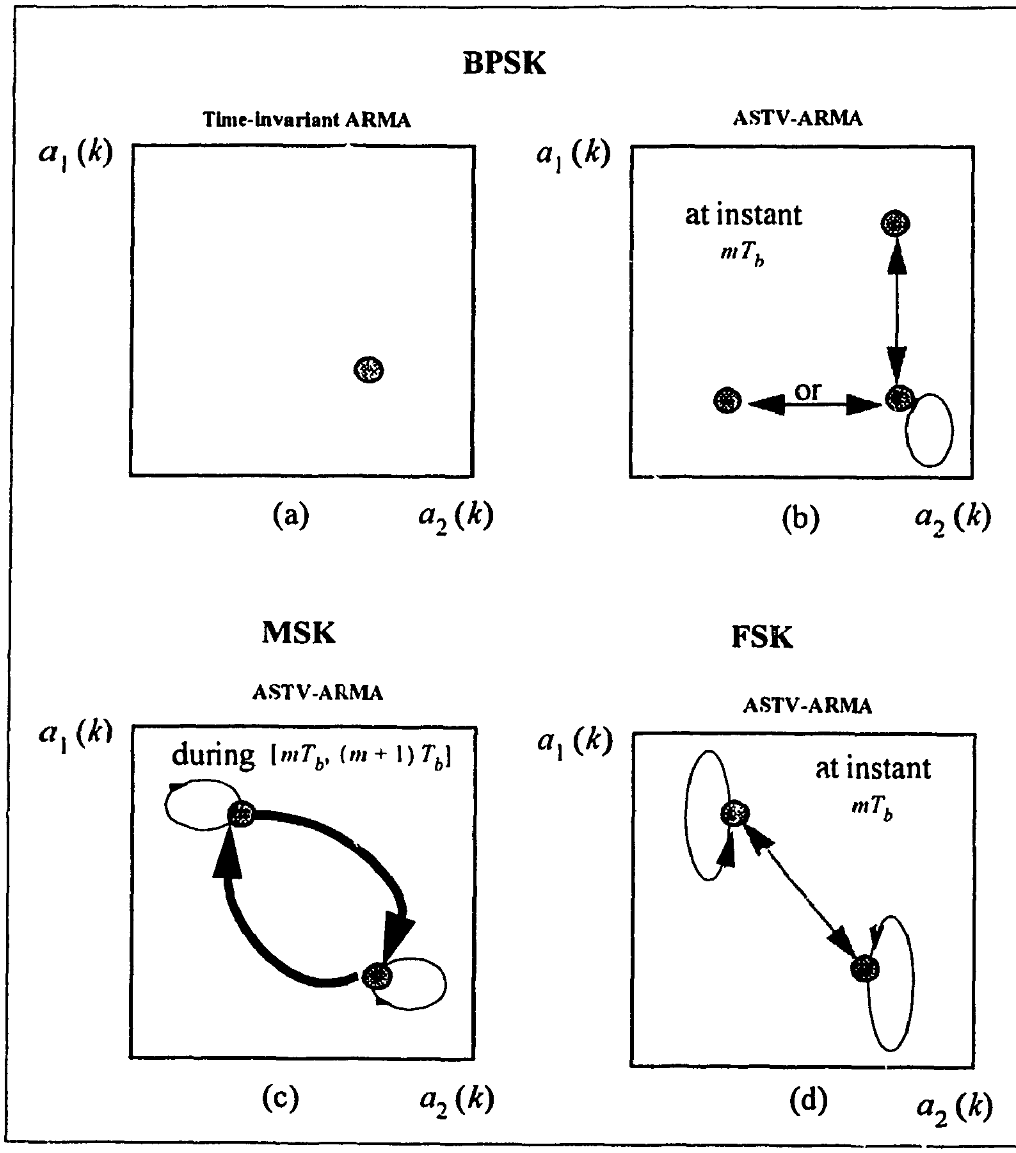


Figure 5.4 Trajectories of AS-ARMA model coefficients

CHAPTER 6

Detection and Estimation of Signals

In the previous chapters an AS-ARMA model is presented, which is used to extract a useful signal from a spectrum overlapped mixture. This technique exploits the property of structural signal representations, that is, a null-operator is modeled to suppress certain signals that belong to a subspace where a null-operator is modeled In the present chapter sources of noise which affect the detection of a digital communication signal and the reconstruction of a speech signal are investigated. There are two problems to consider one is the signal detection problem when a weak digital communication signal is imbedded in a speech signal and channel noise, and the other is the problem of signal estimation when a weak speech signal is imbedded in an interfering digital communication signal and channel noise. The process of separation rejects the digital communication signal from the received mix by using the null singularity property of the communication signal

6.1 Detection of a Weak BPSK Signal

To evaluate the effect of the system described in the previous chapters on the digital communication signal, the bit-error-rate of the digital communication signal detection must be computed. In simulation the received signal is assumed to be composed of many independent modulated sine waves, whose sampled amplitude has an almost Gaussian distribution. The received signal $r(m, k)$ may be expressed as:

$$r(m,k) = A(m,k)\, e^{j\omega(m,k)} + n_{IN}(k) \tag{6.1}$$

where, $A(m, k)$ is the required modulated digital communication signal amplitude, and $n_{IN}(k)$ is an additive Gaussian noise.

The modulated BPSK signal is completely cancelled by an MA null filter and extra blanking algorithm The nonlinear blanking zeros the output of the null-operator, which occurs at the synchronous instant to insure a complete rejection of a digital communication signal This process will introduce a nonlinear distortion to a system, when speech samples exist at synchronous instants The other distortion is caused by the value of the symmetry factor During the separation process, some input noise samples also leak through the system. Two questions are posed to this point (1) how does the separation process affects the overall SNR of the digital communication signal detector? and, (2) how is the speech quality degraded compared to the original speech signal? The overall AS-ARMA system output noise is written as

$$n_{OUT}(k) = n_I^{NL}(k) + n_{II}^{\alpha}(k) \tag{6.2}$$

where $n_{II}^{\alpha}(k)$ is the leakage of the input Gaussian noise (response of a linear system with transfer function $1 - H_{ARMA}(q^{-1})$ for $\alpha \neq 1$ to the input Gaussian noise $n_{IN}(k)$), and $n_I^{NL}(k)$ is the response to a nonlinear distortion, which is created during nonlinear blanking operation This nonlinear noise, which is seen at the output of the AS-ARMA system will affect the detection of a digital communication signal as well as the reconstruction of a speech signal In order to verify the system performance one has to measure the bit-error-rate (BER) and evaluate the quality of a speech signal. The accumulation of enough errors to accurately assess the above system can take a considerable amount of time. To accurately measure the error rate of a digital communication signal, one must record a fairly large number of errors. The proposed method for determining an upper bound on an error rate is to require the system under test to be error-free for a measurement period T. The longer the T is, the lower the error rate boundary In order to determine the upper bound of the error rate, the type of noise, which confuses the decision of the detector, is analyzed

6.2 Leakage of Additive Gaussian Noise

The power spectrum density of the leakage output $n_{II}^{\alpha}(k)$ is defined by the shaded area of the system transfer function $H_{ARMA}(z^{-1})$, which is shown in Figure 6.1 It is also well-known that the stationary Gaussian noise passes through a time-invariant linear system that remains Gaussian.

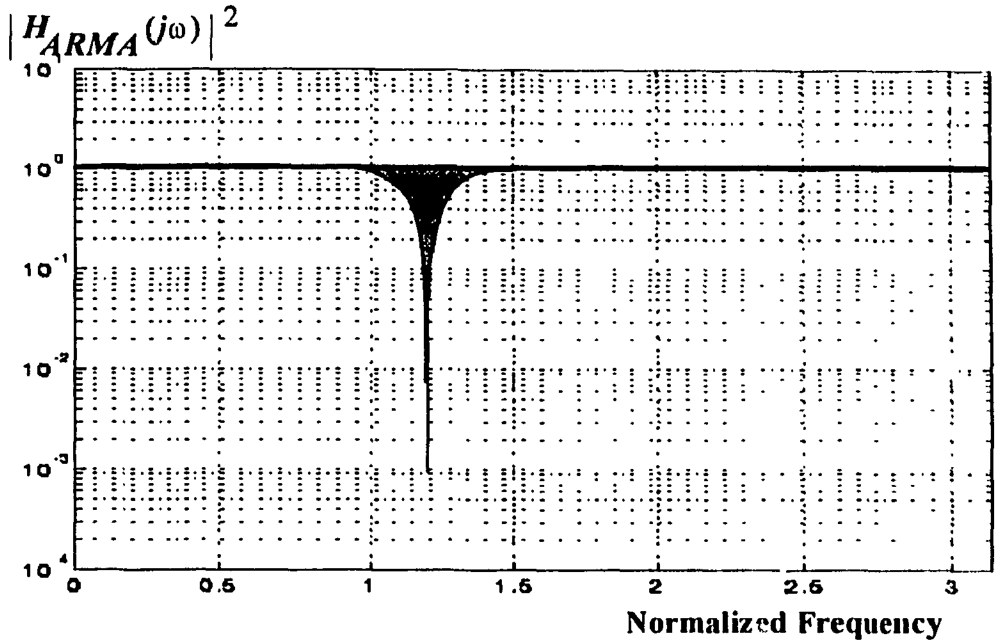


Figure 6.1 Leakage output of Gaussian noise

The variance σ^2_{II} of the leakage noise of AS-ARMA model $n^{\alpha}_{II}(k)$ is estimated by [23]

$$\Upsilon^{\alpha} = cov\{n^{\alpha}_{II}(k)\} = \int_0^{\pi}\left[1 - \left|\frac{1 + ae^{j\omega} + e^{j2\omega}}{1 + a\alpha e^{j\omega} + \alpha^2 e^{j2\omega}}\right|^2\right]\sigma^2\frac{d\omega}{\pi}$$

(6 3)

Eq. 6.3 can be further expressed as:

$$= \frac{\sigma^2}{\pi}\int_0^{\pi}\left[\frac{|1 + a\alpha e^{j\omega} + \alpha^2 e^{j2\omega}|^2 - |1 + ae^{j\omega} + e^{j2\omega}|^2}{|1 + a\alpha e^{j\omega} + \alpha^2 e^{j2\omega}|^2}\right]d\omega$$

(6.4)

$$\Upsilon^{\alpha} = \frac{\sigma^2}{\pi}\left[\frac{\alpha^2-1}{\alpha^2}\right]\int_0^{\pi}\left[\frac{\frac{1+a^2(\alpha^2-1)+\alpha^2(\alpha^2-2)}{4\alpha^2-4}+\frac{a(\alpha^3+\alpha-2)}{2\alpha^2-2}\cos\omega+\cos\omega^2}{\frac{1+\alpha^2(\alpha^2+a^2-1)\;\;+\alpha^2(\alpha^2-2)}{4\alpha^2}+\frac{a\alpha(1+\alpha^3)}{2\alpha^2}\cos\omega+\cos\omega^2}\right]d\omega$$

(6.5)

From this a closed form expression of Υ^{α} is obtained:

$$\Upsilon^{\alpha} = \frac{\sigma^2}{\pi}\left[\frac{\alpha^2-1}{\alpha^2}\right]\int_0^{\pi}\frac{(\cos\omega+A)^2+B}{(\cos\omega+C)^2+D}d\omega$$

(6.6)

where,

$$A = \frac{a[\alpha(1+\alpha^2)-2]}{4(\alpha^2-1)}$$

$$B = \frac{4\left[1+(\alpha^2-1)a^2+\alpha^2(\alpha^2-2)\right]-a^2(\alpha^2-2\alpha+2)^2(\alpha-1)}{16(\alpha^2-1)}$$

$$C = \frac{a(1+\alpha^3)}{4\alpha}$$

$$D = \frac{4\left[1+\alpha^2(\alpha^2+a^2-1)\right]-a^2(1+\alpha^3)^2}{16\alpha^2}$$

The PDF function of the output noise, which appears at the digital communication signal detector is written as

$$f_{ARMA}^{TOTAL}(z) = f_{AR}^{(\alpha)}(z) \otimes f_{MA}(z)$$

(6.7)

The PDF function $f_{ARMA}^{TOTAL}(z)$ is an even symmetry function, and the BER computed at the output AS-ARMA model is given by.

$$BER = \int_{1}^{\infty} f_{ARMA}^{TOTAL}(z)\, dz$$

(6.8)

In order to minimize the BER of the digital communication signal detector, one should minimize the variance of the model noise. The model noise depending on α is shown in Figure 6.1. There must be a unique symmetry factor $\alpha = \alpha^*$ such that the BER for the BPSK detection attains the minimum value, i.e.,

$$\min_{\alpha^*} \{BER\}$$

(6.9)

The calculation of the SNR usually requires that the waveform of interest (the test wave form) be compared to a "desired "or "ideal" waveform (local replica). The desired waveform is often chosen to be an amplitude-scaled and time-delayed version of the information-bearing waveform since amplitude scaling and time delay do not contribute to waveform distortion. The test waveform is then compared to the desired waveform and that portion of the test waveform that is orthogonal to the desired waveform is defined as noise. For this case the SNR estimate becomes [7]:

$$SNR = \frac{\rho^2}{1 - \rho^2}$$

(6.10)

where ρ represents correlation coefficients of model output and desired waveforms.

The gain of signal-to-noise ratio {SNR} in AS-ARMA model output is directly derived from the model output noise, and is illustrated in Figure 6.1:

$$G_{SNR} = \frac{SNR_{OUT}}{SNR_{IN}} \tag{6.11}$$

The SNR gain vs. the symmetric coefficient for $T_b = 10$ and $T_b = 100$ is shown. When α is near to zero the model acts as a low pass filter, and when α approaches the unity the AR part becomes unstable. The simulation presents the output SNR as a function of a symmetry factor. The optimum value of the symmetry factor is near (0.8-0.9).

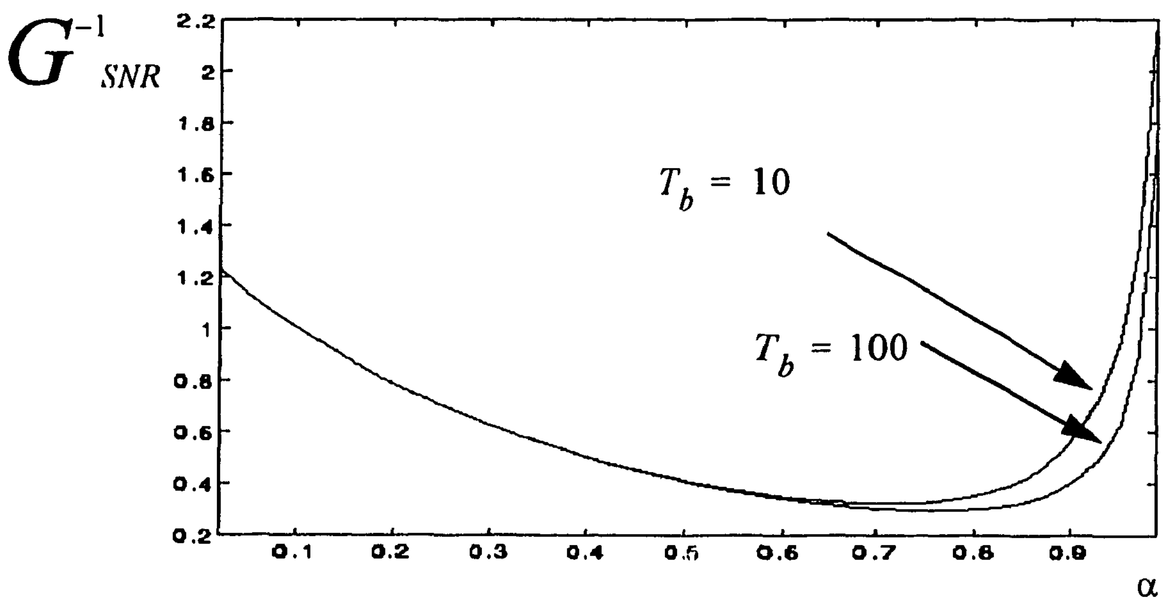


Figure 6.1 Effects of the symmetry factor to SNR

Figure 6.2 shows an upper bound on the bit-error-rate (BER) of a digital communication signal detection for three different schemes: (1) Signal detection with no signal separation process; a weak digital communication signal mixed with a strong interference is applied to the detector. (2) Signal detection with carrier suppression but blanking at synchronous instants is not applied: A null-operator is used to suppress the BPSK signal, but an extra blanking algorithm is not used at synchronous moments. The nonlinear noise created at synchronous instants degraded the detection performance (3) Signal detection with pre-processing of si‸nal separation: when a BPSK signal is canceled using a null-operator with sign change algorithm, the detection has the lowest upper bound of bit-error-rate.

Each signal processing model is simulated for different bit rate to estimate the BER. All three signal detection processes exhibit reduction on an upper bound when the bit rate decreases. It is obvious that a signal-to-noise ratio improvement is achieved by signal separation.

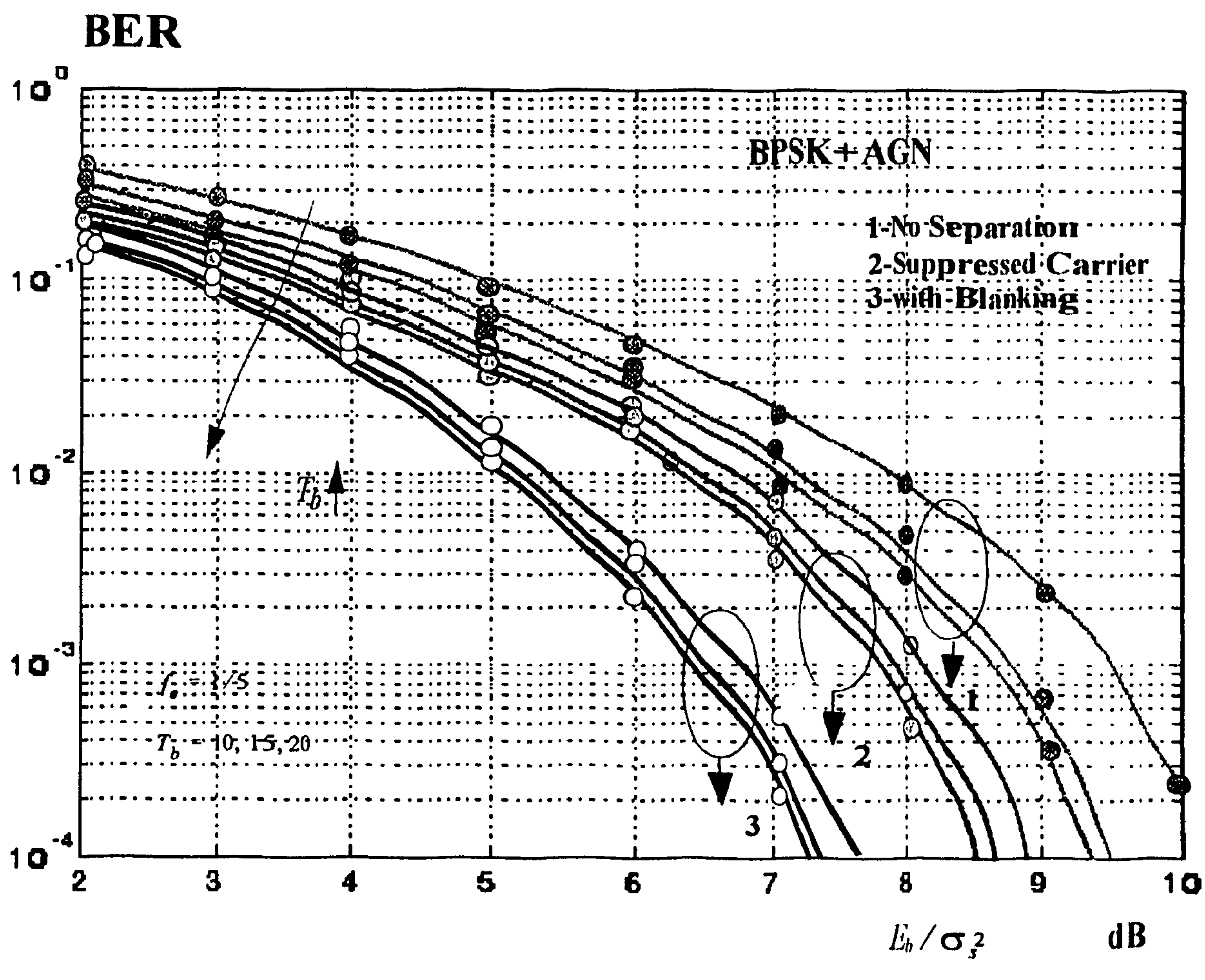


Figure 6.2 Bounds on the error probability and signal to interference ratio

The simulation is used to establish the test waveform for the system under study. As a simple example, if the complex envelope of the test waveform is:

$$\tilde{o} = A \cdot e^{j\theta} \tilde{x}(t - \tau) + n(t) \tag{6.12}$$

the SNR is $A^2 \frac{P_x}{P_n}$ where P_x and P_n are the signal and noise powers respectively. In most applications, the values of $A, \theta, \tau, P_x,$ *and* P_n must be estimated before the SNR can be determined. Simulation can assist in this undertaking.

In digital communication systems the probability of demodulations error P_e, is typically the prime performance measure. For simplicity binary communication systems will be considered and P_e will be referred to as BER. The techniques discussed here can typically be extended to include M-ary communication systems.

To minimize the BER, increasing the signal power is not always a solution. Figure 6.3 shows that for large values of SNR the residual signal from the MA model increases drastically so that the AS-ARMA model does not have any role in BER reduction and the BER remains constant.

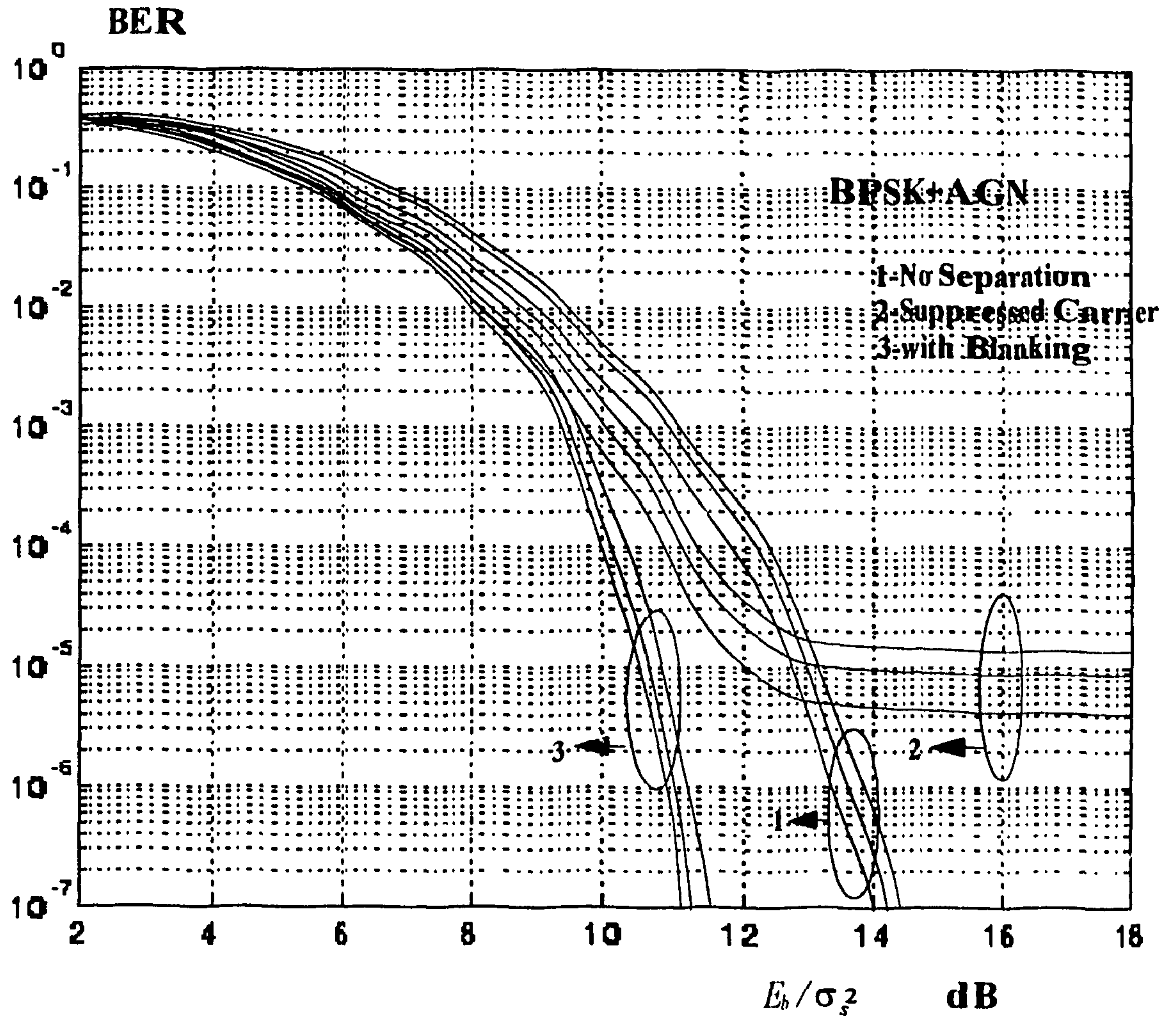


Figure 6.3 Increasing SNR beyond limit becomes no effect to BER

6.3 Reconstruction of a Weak Speech Signal

Processing of a speech which has been degraded due to the interference (digital communication signal) is the focus of this section. Most systems directed at processing speech in the presence of a background noise rely, at least to some extent, on a model of a speech waveform as the response of the vocal tract. This system is represented as a quasi-stationary linear system, to a pulse train excitation for voiced sounds or a noise-like excitations for unvoiced sounds.

A commonly used and physically reasonable short-time model for a vocal tract is a linear system for which the transfer function is all-pole of the form [29]:

$$v(z) = \frac{1}{1 - \sum_{k=1}^{p} a_k z^{-k}}$$

(6.13)

Thus, on a short time basis, the speech waveform $s(n)$ is assumed to satisfy a difference equation of the form:

$$s(n) = \sum_{k=1}^{p} a_k s(n-k) + u(n) + e(n)$$

(6.14)

where $u(n)$ is the input excitation to the system and $e(n)$ represents the modeling error in considering a speech generation process as the output of an all-pole system excited by a simplified source. The basic problem is that of estimating the vocal-tract parameters a_k from a sequence of observations of $s(n)$.

The variance of speech reconstruction noise from the non-perfect symmetry of an AS-ARMA model is given as:

$$\Upsilon^{\alpha} = \int_{-\infty}^{\infty} \omega^2 f_{ARMA}^{(\alpha)}(\omega)\, d\omega$$

(6.15)

where $f_{ARMA}^{(\alpha)}(\omega)$ is the PDF of the ARMA model.

The second noise is generated when the AR input samples are put to ground at the synchronous points to remove the BPSK residues. Blanking at synchronous instants when speech sample is present creates a nonlinear noise Υ^{k_0} at the output.

The total MSE of a speech reconstruction noise at the output is:

$$\Upsilon^{\alpha} + \Upsilon^{k_0}$$

(6.16)

The simulation result showed that there exists a unique symmetry factor $\alpha = \alpha^*$ such that the MSE of speech reconstruction reaches its minimum value:

$$\min_{\alpha^*} \{\Upsilon^{\alpha} + \Upsilon^{k_0}\} \quad (6.17)$$

6.4 Simulation Results

In order to assess the subjective quality of the reconstructed speech signal, the speech separation and reconstruction systems have been simulated, which were described previously. The speech wave was lowpass filtered and sampled at frequency of 10kHz A speech signal mixed with a digital communication signal was applied to the system. The parameter a_k and the optimum symmetry factor value of the system were determined to reject a digital communication signal and to keep a system stable, respectively. From the simulation results it was concluded that the optimum symmetry factor, which provides an adequate representation of output speech, lies between 0.8 and 0.9.

The primary goal of a computer simulation of a communication link is to evaluate or predict the performance characteristics of the system. A number of performance estimates are now considered. In any speech segment, the amplitude of the n^{th} speech sample S_n can be decomposed into two parts: one part q_n contributed by the memory of the linear system carried out from the previous samples and the other part $\tilde{S}_n$ from the current sample. Thus,

$$S_n = q_n + \tilde{S}_n$$

(6.18)

Assume that $n=1$ is the first sample and $n=M$ the last sample of the current spee.·'' seg-ment. The first part q_n is gi en by

$$q_n = \sum_{k=1}^{2} \alpha^k a_k q_{n-k} \qquad 1 \le n \le M$$

(6.19)

where q_0 q_{-1} represents the memory of the predictor carried over from the previous speech samples.

Let P_s be the mean-square value :f the speech samples then P_s is given by,

$$P_s = \frac{1}{M} \sum_{n=1}^{M} \left(q_n + \tilde{S}_n \right)^2$$

(6.20)

The technique used here exploits the structural signal property to ıeject a digital communi-cation signal and passes the rest of the input signal with minor distortions. The two signals are added sample by sample and corrupted by additional random noise generated by the si. 'ator This sum represents the input of the desired system. Based on the structural sig-nal properties, first the nature of the communication signal is determ.ned, and then an AS-ARMA model is designed to minimize the BER of communication signal detector, and MSE between the reconstructed and the original speech signal.

Figure 6.4 shows the input of the AS-ARMA system, where the speech-to-digital commu-nication signal ratio is computed from the ·.ean-square value of the speech samples and the amplitude of digital communication signal.

$$10 \cdot \log\left(\frac{\frac{1}{M}\sum_{n=1}^{M} s_n^{\,2}}{0.5A^2}\right) = -13 \ dB$$

(6.21)

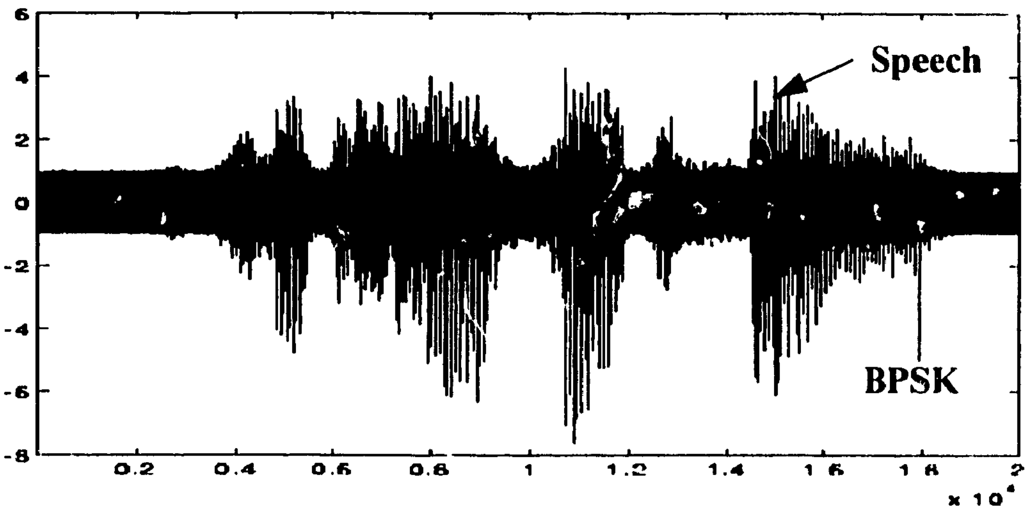


Figure 6.4 Input Speech signal superimposed with BPSK signal

The MA part of the AS-ARMA model suppresses the BPSK signal fully except at the synchronous points (Figure 6.5). The input signal energy is decreased by this process. Blanking at the synchronous instants will eliminate the BPSK residues completely. However, if a speech sample is present at that point, then this speech sample will be eliminated too. As a result, distortion is created.

OUTPUT OF MA MODEL

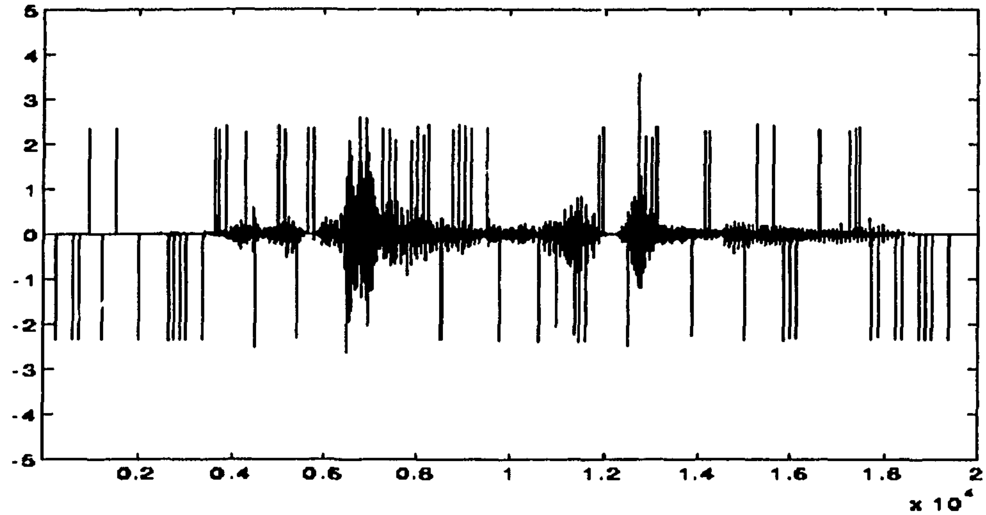


Figure 6.5 OUTPUT of $H_{MA}^{[1]}(q^{-1})$

After a blanking algorithm is applied to the output of the MA part, the residual passes into an anti-null-operator for enhancement. The AR part of the model reconstructs the speech signal, which is distorted by the null-operator. Contrary to the MA part, the AR part adds energy to its input signal. Figure 6.6 shows the reconstructed speech of the system:

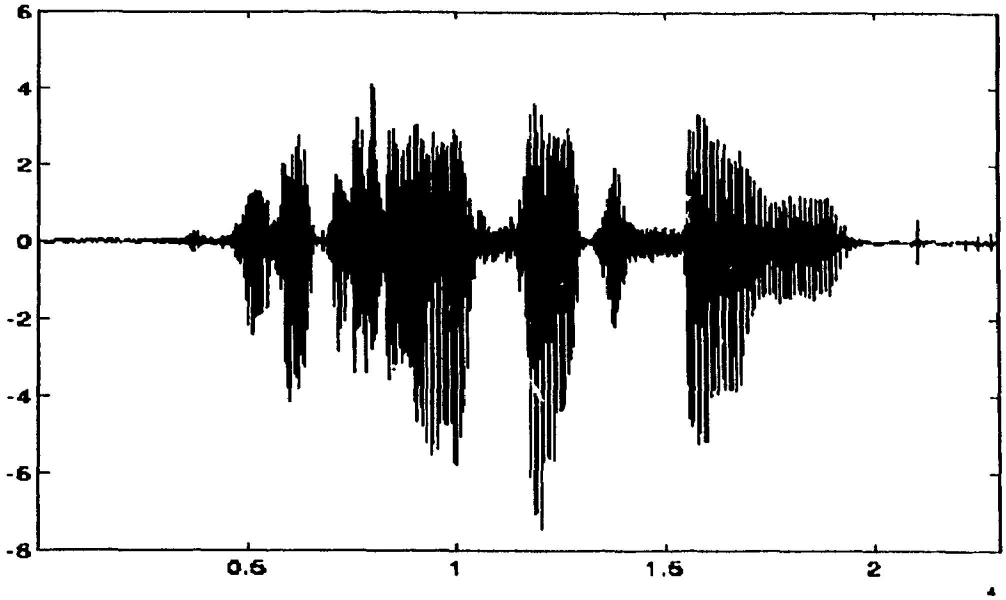


Figure 6.6 Output of $H_{AR}^{[1]}(q^{-1})$

Figure 6.7 shows the simulation result of the MSE function of the symmetry factor for various data rate. It is shown that when the symmetry factor approaches unity the MSE decreases significantly. However, the BER is affected when the symmetry factor nears unity. The optimal symmetry factor for the minimum MSE of speech reconstruction and the minimum value of BER of the BPSK detection are not the same. There exists an optimum symmetry factor for the minimum MSE of speech reconstruction and the minimum value of BER of the BPSK detection.

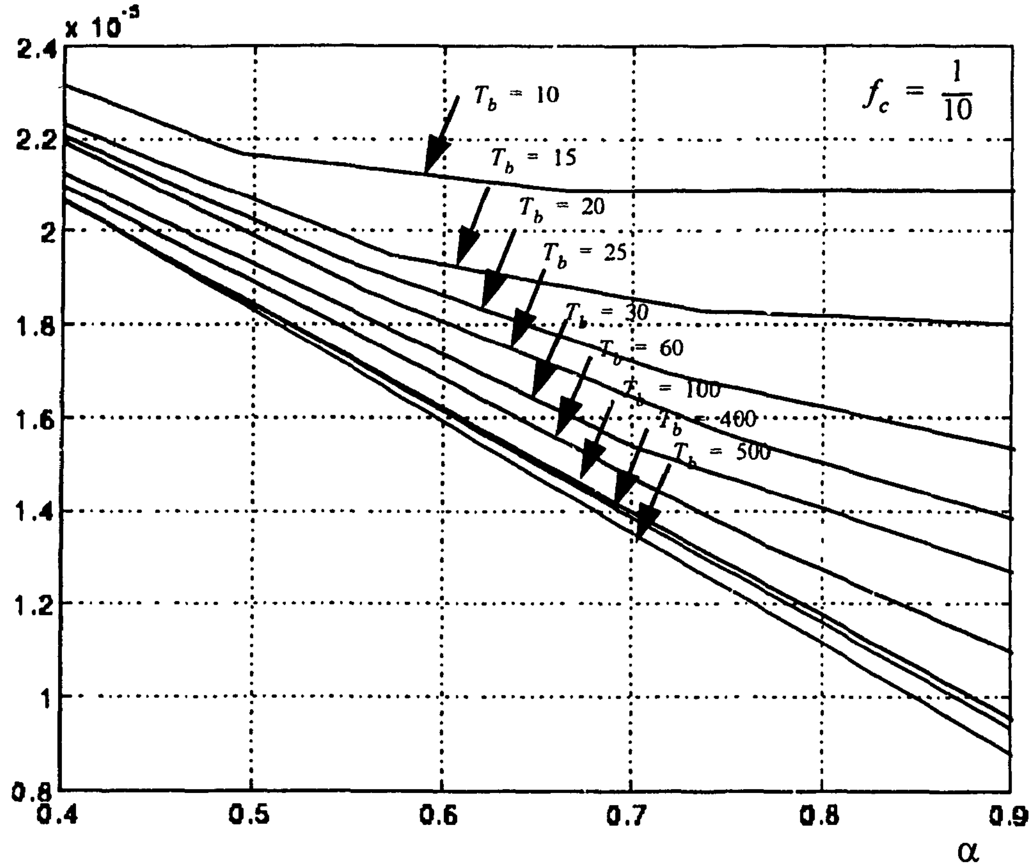


Figure 6.7 Normalized Speech Reconstruction MSE

CHAPTER 7

Dynamic Evaluation of Processing Parameters

7.1 Self-synchronized Signal Controlled AS-ARMA Model

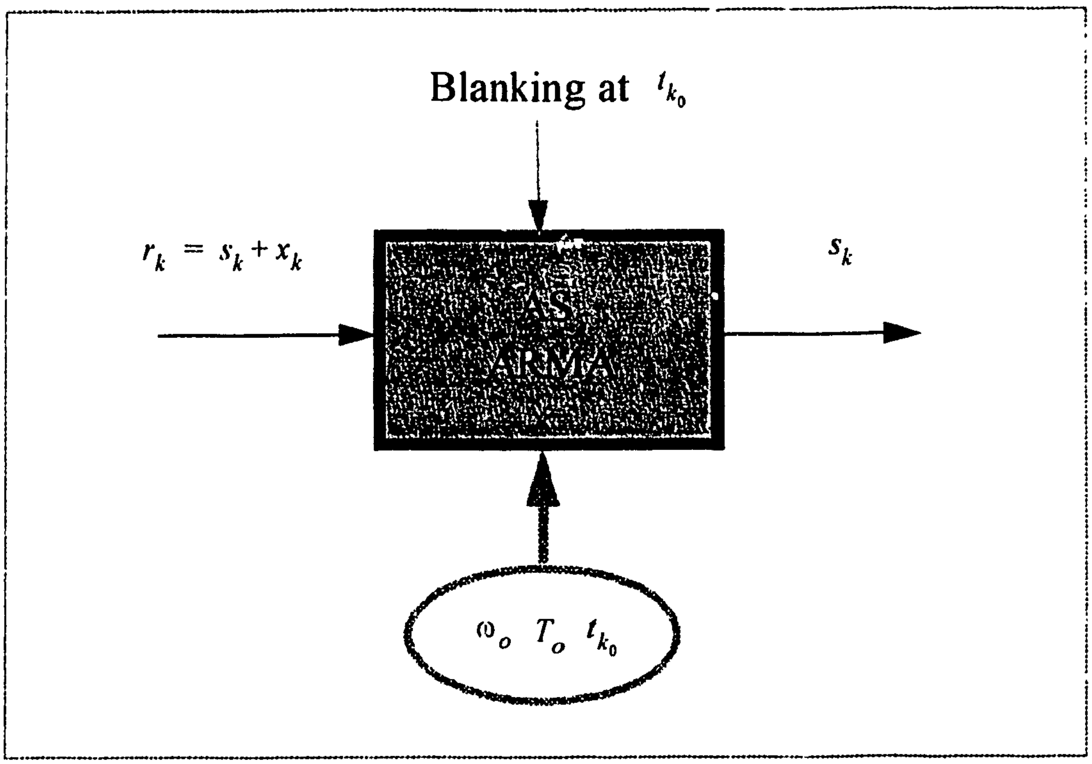
Retrieving a signal corrupted by a strong frequency overlapping interference is the problem under consideration [30]. The particular form of the model that is appropriate for signal separation analysis is shown in Figure 7.1. In this case a digital communication signal represented by a time-varying MA model whose steady-state system function is of the form:

$$H(z) = \frac{W(z)}{X(z)} = 1 + b_1(m)z^{-1} + b_2(m)z^{-2} \tag{7.1}$$

This system is excited by the sum of a speech signal and a BPSK signal The model will suppress the modulated signal while a speech signal will pass through the system with minor distortions. The basic problem is to determine model parameters from the input signal. The above model perfectly suppresses the modulated BPSK signal when the parameters of the signal are known and thus may be used to compute the model parameters If the parameters of the modulated signal are not known, the suppression process is not ideal and parameter estimation must be performed before the separation process is started The basic approach is to find parameters that will minimize the mean-square error in separation over a short segment of the input signal The resulting parameters are then assumed to be the parameters of the system function, $H(z)$, in the model for signal separation

7.2 Dynamic Evaluation of BPSK Signal Parameters

The three parameters which guarantee the perfect suppression of the received BPSK signal are the synchronous time instant t_{k_0}, the bit rate T_0, and the carrier frequency ω_0.



$s_k \rightarrow$ *Speech Signal*

$x_k \rightarrow$ *BPSK data*

$t_{k_0} \rightarrow$ *Synchronous Time Instant*

$T_0 \rightarrow$ *Symbol Width, Bit Rate*

$\omega_0 \rightarrow$ *Carrier Frequency*

Figure 7.2 AS-ARMA model

If one of the above parameters is not known, the suppression of the BPSK signal is not perfect. To assist the synchronization of the transmitter and receiver parameters, a dynamic control system is recommended for the first few samples of input signal. These input samples are used to train the system of updating the parameters, which later are implemented in the signal suppression process [31].

- **Evaluation of Carrier Frequency**

Figure 7.3 uses the basic adaptation formula that computes the new estimate as a sum of the old estimate and a correction term.

$$\begin{aligned} &\textit{Input:} && r_k = s_k + x_k \\ &\textit{Output:} && O_k = r_k - \tilde{x}_k \qquad \textit{residual signal} \end{aligned} \tag{7.2}$$

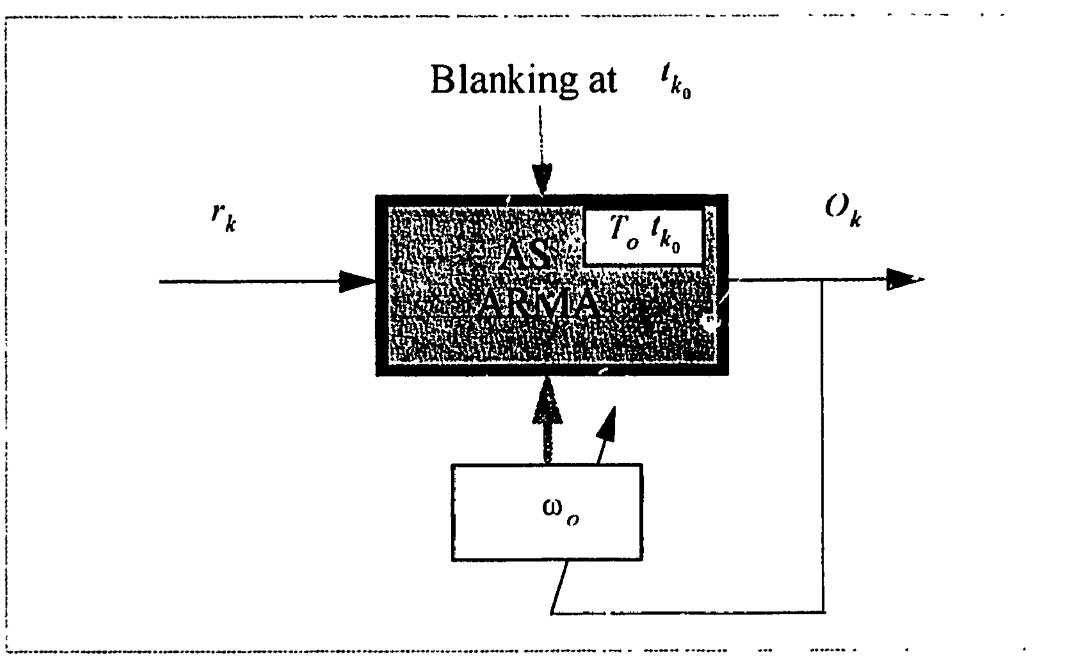


Figure 7.3 The case corresponding to carrier frequency unknown

The normalized MSE $E\{|x_k - \tilde{x}_k|^2$ between x_k and its series representation $\tilde{x}_k$ depends on the number of terms in the series and the (basis) functions used in the expression. The least-square method abandons ail need for stochastic information and treats the estimation problem as a deterministic optimization problem. Typically, the cost function is defined on the basis of MSE criterion, with the error signal itself being defined as the difference between a desired response and the actual output of the network produced in response to the corresponding input signal. The desired system learns from examples by constructing an input-output mapping for the problem at hand. The approximate gradient descent minimization of the average squared parameter evaluation is given as:

$$\omega^o_{k+1} = \omega^o_k + \mu\nabla^2 \boldsymbol{J} \tag{7.3}$$

where μ is a small positive step size, which is a viable recursive solution to the cost function.

$$\textit{Cost Function:} \quad \boldsymbol{J} = \min_{\{\omega_0\}} \boldsymbol{E}\{O^2_k\} \qquad \omega^o_k \in [0, \pi] \tag{7.4}$$

where,

$$O_k = r_k - 2\cos(\omega^o_k) r_{k-1} + r_{k-2} + 2\alpha\cos(\omega^o_k) O_{k-1} - \alpha^2 O_{k-2} \tag{7.5}$$

The LMS algorithm from Eq. 7.3 for evaluating the system parameter is given by the following expression:

$$\omega^o_{k+1} = \omega^o_k + 2\mu\sin(\omega^o_k) O_k (r_k - \alpha O_{k-1}) \tag{7.6}$$

The convergence properties of the algorithm are largely determined by the step size parameter μ and the power of the model output O_k. In general, making μ large speeds the convergence, while a smaller μ reduces the asymptotic cancellation error Figure 7.4 presents the convergence rate for different step size.

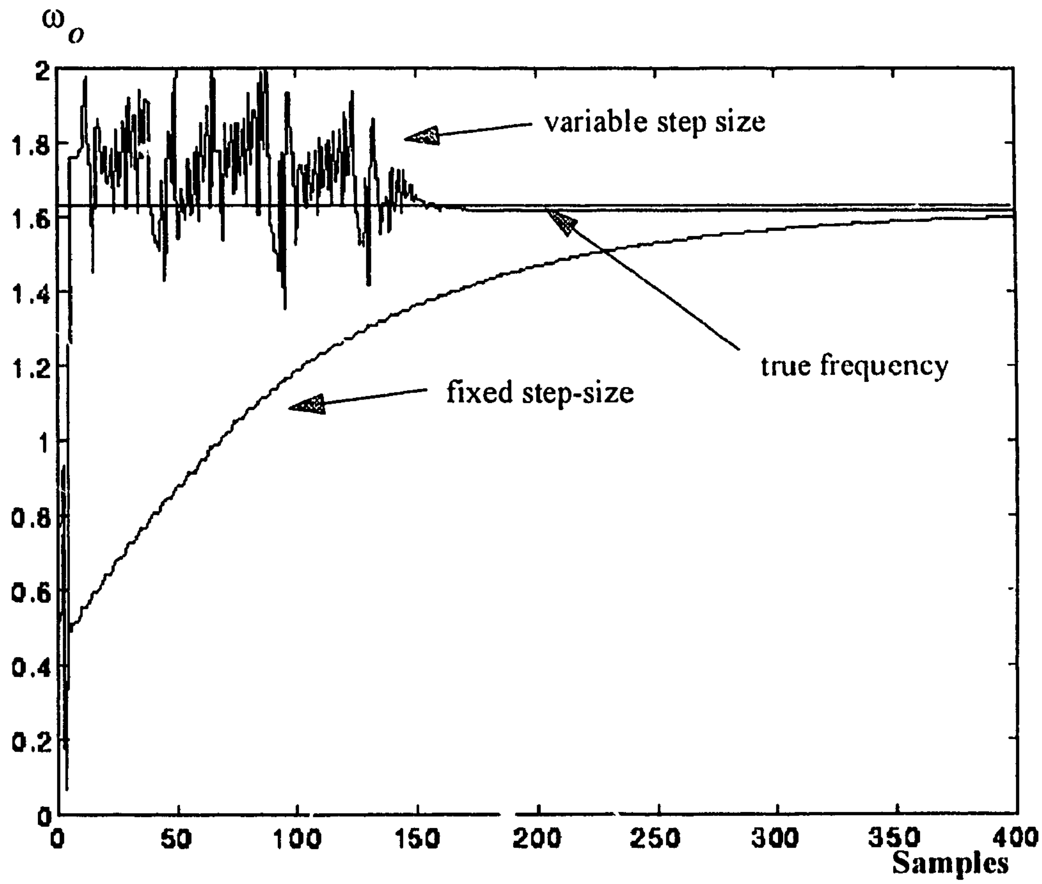


Figure 7.5 Dynamic evaluation of carrier frequency

- **Identifying Synchronous Instants**

A dynamic evaluation to identify synchronous instants is presented in Figure 7.6. The input of the system is a strong BPSK signal overlapped with speech, and the output of the system is a reconstructed speech interfered with a residual signal. This residual signal is the difference between the input BPSK signal and the estimated BPSK signal, which is evaluated by using the present synchronous point. This signal difference will be added to the reconstructed speech signal which increases the output signal power. The synchronous time instant together with the bit duration determine the length of signal processing. The process of searching for the optimum synchronous time instant is based on the LMS method. The process is started by choosing an arbitrary synchronous point. Using this synchronous time instant, the AS-ARMA system is modeled and the output signal is evaluated. A cost function is calculated as the cumulative output power for the period of an output signal sample damps, L

$$\text{Cost Function:} \quad J_k = \sum_{m=-L}^{0} O_{k+m}^2 \tag{7.7}$$

where L is evaluated to the nearest integer of the damping length, $L = \frac{2}{1-\alpha}$. This processing duration is determined by the symmetry factor α.

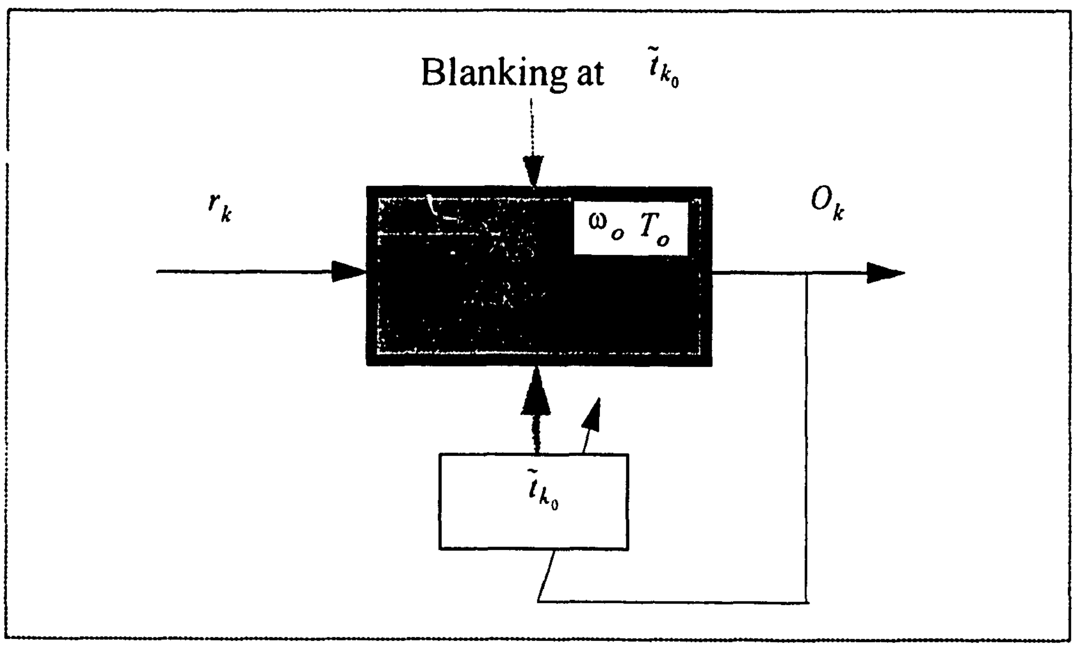


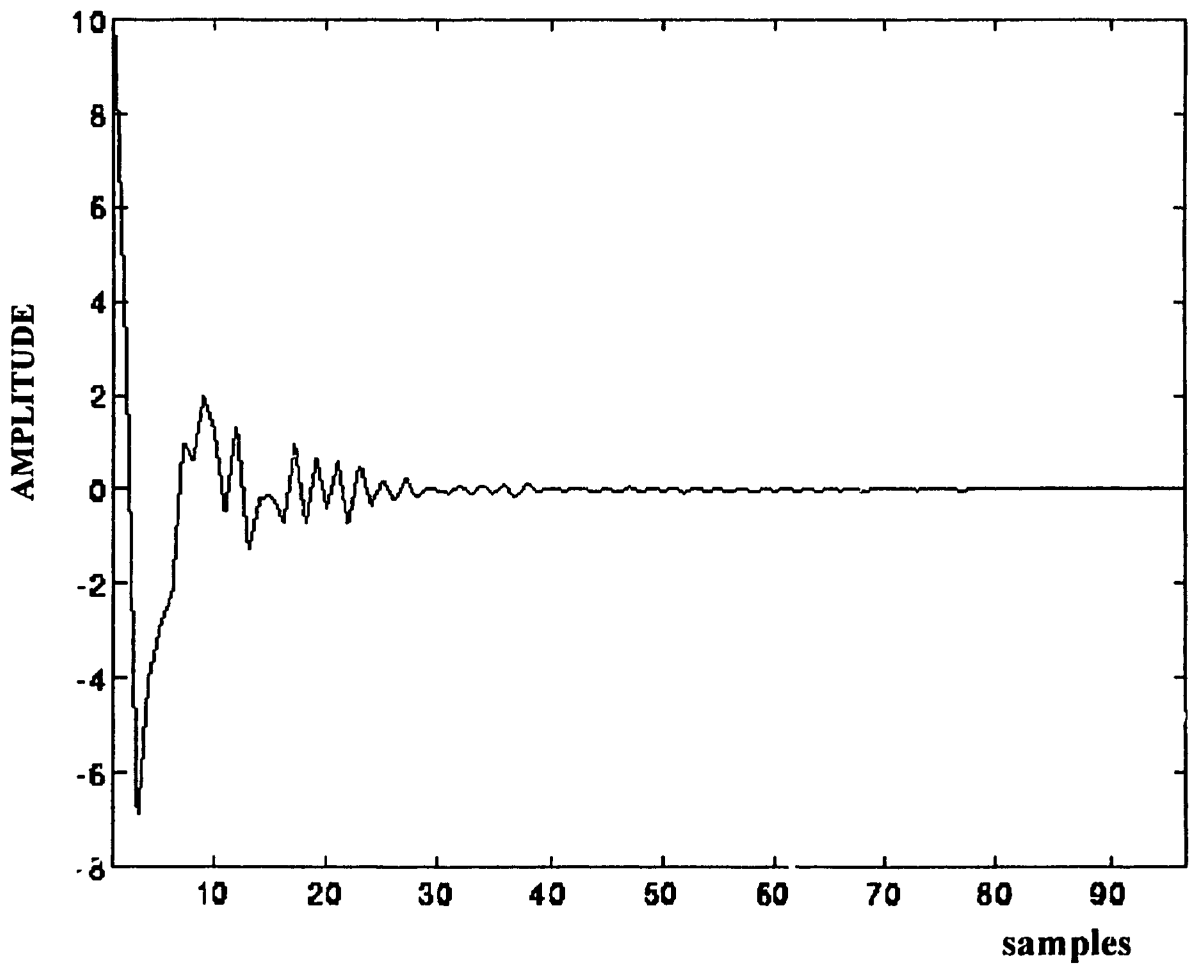
Figure 7.6 Identifying the synchronous instants

When the search approaches the real synchronous point the output signal energy becomes minimum, and Eq. 7.8 becomes lower than a predefined minimum threshold, $\gamma_k \leq \gamma_{k_0}$.

$$\gamma_k = J_k - \frac{1}{L} \sum_{m=-L}^{0} J_{k+m-\frac{L}{4}}$$

(7.8)

When the cost function reaches its minimum, the success of the control algorithm will be stopped and the signal separation process continues. Figure 7.7 shows the output signal of the AS-ARMA system. The input signal is a strong BPSK signal whose unknown synchronous time instant overlaps with weak speech.



7.2.2.2 System output (while evaluation of synchronous point)

- **Determining the Data Rate**

In communication, the data rate of the far end is unknown or varies from time to time. At the start of communication the receiver has to evaluate the bit length in order to synchronize with the received signal. Figure 7.8 shows a system which controls to determine the bit rate of the input BPSK signal. Clearly what is required is a practical measurement of the output signal power.

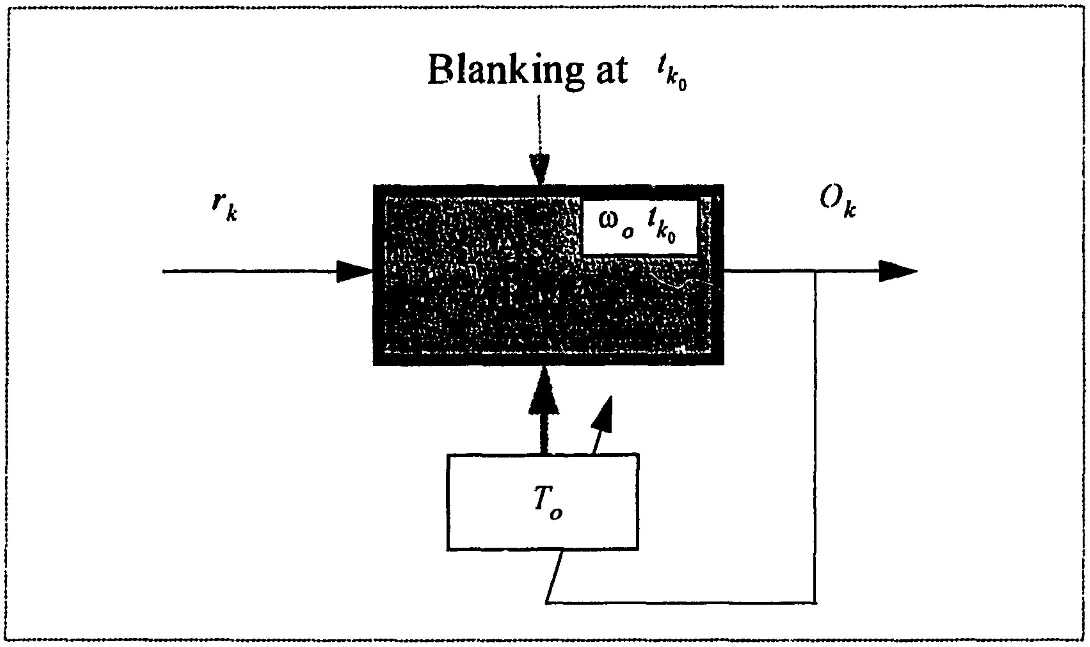


Figure 7.8 Determining the bit rate

The target is to minimize the output signal power. This quantity is computed for a predefined window.

$$\text{Cost Function:} \qquad J_k = \sum_{l=-nT_o^{(m)}}^{nT_o^{(m)}} O_{k+l}^2 \tag{7.9}$$

where $nT_0^{(m)}$, and $-nT_0^{(m)}$ are the upper and lower bound of the window respectively. If the cost function reaches its minimum and sustains for a number of windows, $\gamma_k < \gamma_{kT_0}$, the evaluation, registers the last bit rate as optimum.

$$\gamma_k = J_k - \frac{1}{N} \sum_{l=-\frac{N}{2}}^{\frac{N}{2}} J_{k+l} \tag{7.10}$$

where N is the number of windows, where the average cost function is computed. It has been shown that the convergence time constant is inversely proportional to the power of O_k, and that the algorithm will converge very slowly for low-power signals. To remedy that situation, the loop gain is usually normalized by an estimate of that power.

Chapter 8

Conclusions and Future Work

The thesis presented how separation of frequency overlapped signals, such as speech signal interfered by BPSK signal, can be achieved using a one-sided AS-ARMA model associated by a blanking algorithm, which nullify null-operator residues. The model is built from a symmetrical ARMA model, where its symmetric is controlled by a model parameter.

8.1 Conclusions

The research studied the problem of signal separation, where the received signal is the result of the simultaneous transmission of speech signal and digital communication signal operating within the same channel bandwidth. An almost symmetrical ARMA model, which is accompanied by a special design blanking procedure is constructed to solve a signal separation problem. The technique which is used to construct the model is based on the exploitation of the structural property of the digital communication signals Although some sophisticated algorithms such as Viterbi algorithm can be incorporated to challenge the described problem, a second order ASTV-ARMA or AS-ARMA model in conjunction with the nonlinear blanking operation showed that signal separation can be achieved with an acceptable quality of performance.

It is shown that a symmetrical anti-null operator can provide an ideal transparency, but it is sensitive to undefined initial conditions If the symmetry factor α approaches unity the model becomes almost transparent. However, the null-operator residues which happen at synchronous instants create instability for the system.

In this thesis an optimum symmetry factor that is found by simulation is used to design the almost symmetrical ARMA model. It was shown that for the small values of this factor $\alpha \ll 1$, the MA part output of the model vanishes over a short period. On the other hand, selecting the small value symmetry factor to be small does not provide the needed *transparency*. Such a *transparency* does become almost ideal if the symmetry factor approaches unity, i e $\alpha \rightarrow 1^{-}$ A trade-off associated with this problem requires to choose the symmetry factor close to unity $(\alpha = 0.9)$ and different algorithms to be considered to eliminate the nonlinear noise which is created at synchronous time instants

An additional blanking algorithm is introduced to prevent the residues of the null-operator from entering into the anti-null model, which will excite the model to be a frequency generator. The one sided null-anti-null operator and the blanking algorithm together permits separation of a speech signal, whose frequency overlap by a communication signal, with acceptable quality

The major contribution of this work is the development of two types of blanking algorithms that significantly improved AS-ARMA model transparency quality (see chapter 5). The main properties of this blanking procedures may be outlined as following:

(1) Periodical nonlinear-blanking AS-ARMA model

This model suppressed the BPSK signal totally at the output of a null-operator A perfect signal separation is achieved if speech samples do not exist at synchronous instants. Note that, speech samples that may appear at synchronous instants are canceled by the blanking algorithm, resulting in some speech signal distortion.

(2) Quasi-deterministic periodical time-varying AS-ARMA model

Instead of blanking at all synchronous instants, the algorithm detects the sign change of the BPSK signal to blank the null-operator residue at synchronous instant. This approach is more perfect than the previous blind blanking algorithm. Blanking is done when there is desired, so that speech samples are not deleted at all synchronous instants. This approach provides more reliable separation of the considered signals.

The performance of the model under analysis is assessed in two different cases. First a signal detection problem is considered, when a weak communication signal is mixed with a strong speech signal. The quality of the detector is expressed by measuring BER of signal detection. Second a signal reconstruction problem is considered, when a weak speech signal is interfered by a strong BPSK signal. The intelligibility of the reconstructed speech and the quality of the speech reconstruction are used to assess the model performance The intelligibility of the reconstructed speech is characterized by listening, and the mean square error is used to show the quality of speech signal reconstruction The result of the analysis are presented in chapter 6

A BPSK signal with unknown parameters is used in the last chapter to separate from the interfered speech signal A dynamic evaluation of AS-ARMA model is considered for this problem. The basic approach is to find parameters that will minimize the mean-square error in separation over a short period of input signal. The result shows that signal separation can be achieved with acceptable quality. It is necessary that more work has to be done in this direction.

8.2 Future Work

Among the future research topics that can be explained concerning the subjects studied in this thesis; it is necessary to mention directions of future development to be followed

(1) The problem of retrieving a signal corrupted by a strong impulsive interference can be addressed by using a time-varying AS-ARMA model. The advantages of AS-ARMA model that it offers easy control of bandwidth, an infinite null, and the capability of adoptively tracking the exact frequency of the interference are areas that need future work

(2) Possibility on increasing a communication system channel capacity by using a Structural Signal Multiplexing (SSM) based on the ability to separate superimposed frequency overlap signals. Overlying SSM with CDM is the area which future investigation should focus.

(3) Development of self-synchronized AS-ARMA modeling, which leads to an adaptive signal controlled modeling is a right direction to future work. This area of signal processing is very important in digital communication where system synchronization is driven from a received signal.

(4) System identification based on AS-ARMA modeling is another important area where future work should be considered. Further investigation of AS-ARMA model, where its coefficients define unknown system characteristics will lead to system identification. This area of work will address speech recognition, speech enhancement and speech synthesis.

References

[1] John W. Eng, James F. Mollenauer, "IEEE Project 802 14: Standards For Digital Convergence", IEEE communication, May 1995, vol 33 No.5, pp 20-23

[2] E. O. Brigham, "The Fast Fourier Transform and its Applications", Englewood Cliffs, NJ: Prentice Hall, Inc , 1988, pp. 83-86 and 320-337.

[3] Jeffery A. Wepman, "Analog-to-Digital Converters and their Applications in Radio Receivers". IEEE, communication May 1995 Vol. 33 No.5 pp. 39-45.

[4] John Gallant, "Transmitting Voice and Data", EDN, The Design Magazin of the Electronics Industry, September 1995, pp.61-68

[5] D. Wulich, E.I. Plotkin, M.N.S. Swamy and Wen Tong, "Stability of Two Second Order Time-varying Constrained Notch Filter with Externally Controlled Coefficients" Processing of Canadian Conference on Electrical and Computer Engineering, 1990, Ottawa, Canada

[6] A.V. Oppenheim and A.w. Schafer, "Digital Signal Processing", Prentice-Hall, 1965

[7] William H. Tranter and Kurt L. Kosbar "Simulation of Communication Systems", IE communication Magazine, July 1994 Vol 32 No.7

[8] L. R.Rabiner and R. W. Schafer, "Digital Processing of Speech Signals". Macmillan, 1993.

[9] A.V. Oppenheim and A.W. Schafer, and T.G. Stockham, Jr. "Nonlinear Filtering of Multiplied and Convolved Signals", Proc. IEEE, Vol 56,pp. 1264-1291,Augest1968

[10] A. M. Noll, "Cepstrum Pitch Determination," J. Acoust. Soc Amer. vol. 41 pp. 293-309, February 1966

[11] M. Pagano, "On Periodic and Multiple Autoregressive," Ann. Statist., vol.6,pp.1310-1316, 1978

[12] John R. Deller, JR., John G. Proakis John, H.L.Hansen "Discrete-Time Processing of Speech Signals", New York,1993

[13] B. Woodrow and S. D. Stearns, "Adaptive Signal Processing", Prentice-Hall, 1985.

[14] B.S. Atal, "Automatic Speaker Recognition Based on Pitch Contours" J. Acoust. Soc. Amer., vol. 52, pp. 1686-1696, December 1962

[15] L. R Rabiner, R. E Crochiere, and J. B Allen, "Comparison of System Identification Methods in the Presence of High Noise Levels and Bandlimited Inputs", Bell Laboratories, Murray Hill, New Jersy, 1978

[16] E I. Plotkin and M. N. S Swamy "Structural Signal Processing-Current Trends", ASME Diagnostics, Vehicle Dynamics and Special Topics, Book No H0508E,pp. 6-45, September 1989

[17] L. R. Rabiner and B. Gold, "Theory and Applications of Digital Signal Processing", Prentice-Hall, 1965

[18] William A. Garder "Cyclostatinarity in communications and Signal Processing" IEEE Press, New York 1994

[19] D Wulich, E.I. Plotkin, M.N S Swamy and Wen Tong, "Externally Controlled Time-Varying Constrained Notch Filter", ISCAS91, June 11-14, 1991, Singapore.

[21] D. Wulich and M. Bar, "The Convergence of PLL with Variable Parameters", Int. Circuit Theory Appl., vol 18, pp. 135-145, 1990

[22] D Wulich, E I. Plotkin, M.N.S. Swamy, "Synthesis of Discrete Time-varying Null Filters for Frequency Varying Signals using the Warping Technique", IEEE Trans., CAS-36, pp 966-990, August 1990

[22] J. S. Lim and A. V. Oppenheim, "Advance Topics in signal Processing", Prentice-Hall Signal Processing Series, 1988

[23] D. Wulich, E.I. Plotkin, M.N.S. Swamy, "Constrained Notch Filtering of Nonuniformly Spaced Sample for Enhancement of an Arbitrary Signal Corrupted by a Strong FM Interference", IEEE Trans. on Signal Processing (1991), Vol 39, No.10, pp. 2359-2363

[24] D. Wulich, E.I. Plotkin, M.N.S. Swamy and Wen Tong, "PLL based Discrete Time-varying Filter for Estimation of Parameters of a Sine Signal Corrupted by a Closely Spaced FM Interference", ISCA90, pp. 661-664, May 1-3, 1990, New Orleans, Louisiana

[25] R.A. Gable and R.A. Roberts, "Signal and Linear Systems", 2nd ed., John Wiley & Sons, New York, 1981

[26] R. D. De Groat, L. R. Linebarger, "Discrete-Time Nonlinear Systems Modeling", ICASSP, IEEE, pp. 2587-2590, 1990

[27] W. Hurewicz, "Lecture on Ordinary Differential Equations", John Wiley & Sons, Inc., New York, 1958

[28] T. Muir, "Theory of Determinants", Volume 1-4, Dover Publication, Inc., New York, 1906

[29] A. V. Vecchia, "Maximum Likelihood Estimation for Periodic Autoregressive Moving Average Models," Techometrics, Vol 26, pp. 365-384, 1985

[30] Wen Tong, "Almost-Symmetrical Time-Varying ARMA Model and Its Application for Separation of Signals with Overlapping Fourier Spectra", Ph.D. dissertation, Concordia University, Montreal, Quebec, Canada, 1993

[31] W. M. Brelsford, "Probability Predictions and Time Series with Periodic Structure", Ph.D. dissertation, Johns Hopkins University, Baltimore, Maryland, 1966